



ΤΕΙ ΗΠΕΙΡΟΥ

ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ

Μεταπτυχιακό Πρόγραμμα Σπουδών

Λογιστική – Χρηματοοικονομική και

Διοικητική Επιστήμη

Κατεύθυνση: Διοικητική Επιστήμη και Οργανισμοί

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ:

«Ανάλυση χρηματοοικονομικών δεδομένων υψηλού όγκου και ταχύτητας. Μελέτη περίπτωσης: Χαρτοφυλάκιο Ελληνικών Ναυτιλιακών Εταιριών»

Τσίντζα Παναγιώτα Α.Μ. 109

Επιβλέπων Καθηγητής: Γκόγκος Χρήστος

Β' Επιβλέπων Καθηγητής: Καραμάνης Κωνσταντίνος

Πρέβεζα, Μάιος 2018



ΤΕΙ ΗΠΕΙΡΟΥ

ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ

Μεταπτυχιακό Πρόγραμμα Σπουδών

Λογιστική – Χρηματοοικονομική και

Διοικητική Επιστήμη

Κατεύθυνση: Διοικητική Επιστήμη και Οργανισμοί

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ:

«Ανάλυση χρηματοοικονομικών δεδομένων υψηλού όγκου και ταχύτητας. Μελέτη περίπτωσης: Χαρτοφυλάκιο Ελληνικών Ναυτιλιακών Εταιριών»

Τσίντζα Παναγιώτα Α.Μ. 109

Επιβλέπων Καθηγητής: Γκόγκος Χρήστος

Β' Επιβλέπων Καθηγητής: Καραμάνης Κωνσταντίνος

Πρέβεζα, Μάιος 2018

**ANALYSIS OF HIGH VOLUME AND SPEED FINANCIAL DATA.
CASE STUDY: PORTFOLIO OF GREEK SHIPPING COMPANIES**

Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Πρέβεζα, Ιούνιος 2018

ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ

1. Επιβλέπων καθηγητής
Χρήστος Γκόγκος,
Αναπληρωτής Καθηγητής

2. Μέλος επιτροπής
Κωνσταντίνος Καραμάνης,
Αναπληρωτής Καθηγητής

3. Μέλος επιτροπής
Χαρίλαος Ναζάκης,
Καθηγητής

© Τσίτζα, Παναγιώτα, 2018.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Δήλωση μη λογοκλοπής

Δηλώνω υπεύθυνα και γνωρίζοντας τις διατάξεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα μεταπτυχιακή εργασία είναι εξ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Τσίντζα Παναγιώτα

Υπογραφή

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία πραγματεύεται το ζήτημα των χρηματοοικονομικών προβλέψεων και εισάγει μια πρακτική εφαρμογή τους με τη χρήση των νέων τεχνολογιών και πιο συγκεκριμένα με την γλώσσα προγραμματισμού Python πάνω σε μεγάλα δεδομένα. Το ζήτημα των προβλέψεων είναι ένα ενδιαφέρον αλλά και πολύπλοκο θέμα μελέτης εφόσον εμπλέκει διάφορους τομείς όπως αυτός της Στατιστικής, των Χρηματοοικονομικών και της τεχνολογίας.

Η εργασία δομείται πάνω σε 3 ενότητες. Στην πρώτη αναλύονται οι χρηματοοικονομικές αγορές ως προς τη γραμμικότητα ή μη αλλά και οι διάφοροι τύποι κινδύνου που εμπεριέχουν. Στη δεύτερη ενότητα παρουσιάζονται οι διάφορες μέθοδοι προβλέψεων των χρηματιστηριακών αγορών μετά από μια σύντομη εισαγωγή στις χρονοσειρές και προβλέψεις. Περιλαμβάνει επίσης τα διάφορα πιθανά σφάλματα κατά την διαδικασία των προβλέψεων. Η τελευταία ενότητα πραγματεύεται μια ανασκόπηση των τεχνολογικών επιτευγμάτων πάνω στην θεωρία των προβλέψεων. Περιλαμβάνει επίσης την υλοποίηση ενός προβλεπτικού συστήματος καθώς και την περιγραφή των τεχνολογιών που χρησιμοποιήθηκαν.

Λέξεις-κλειδιά: Μεγάλα Δεδομένα, Python, Ανάλυση δεδομένων.

ABSTRACT

This Thesis discusses the issue of financial forecasts and introduces a practical application using new technologies and more specifically Python programming language on large data. The issue of forecasts is an interesting but complex subject of study since it involves a number of areas such as Statistics, Finance and Technology.

The work is structured on 3 different modules. In the first one, the financial markets are analyzed in terms of linearity or not, and the different types of stock market risks are introduced. The second section presents the various stock market forecasting methods after a brief introduction to time series and forecasts. It also includes the various possible errors in the forecasting process. The last section discusses the state of art of technological achievements on predictive theory. It also includes the implementation of a predictive system and a description of the technologies used.

Keywords: Big Data, Python, Data Analysis.

Περιεχόμενα

ΕΙΣΑΓΩΓΗ	1
1. ΓΡΑΜΜΙΚΗ ΚΑΙ ΜΗ-ΓΡΑΜΜΙΚΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΔΥΝΑΜΙΚΗ	3
1.1 Χρηματοοικονομικοί τίτλοι και παράγωγα χρηματοοικονομικά προϊόντα.....	3
1.1.1 Χρηματοοικονομικοί τίτλοι.....	4
1.1.2 Χρηματοοικονομικά παράγωγα	6
1.2 Η Υπόθεση Αποτελεσματικής Αγοράς	9
1.2.1 Ορισμός της Αποτελεσματικής Αγοράς.....	9
1.2.2 Η Υπόθεση των Αποτελεσματικών Αγορών και τα είδη τους	11
1.3 Αποτελεσματικότητα Αγοράς και ανωμαλίες.....	14
1.4 Μη-Γραμμικές δυναμικές στις αγορές Χρήματος και Κεφαλαίου.....	16
1.5 Προσαρμοστική Αγορά.....	18
1.6 Χρηματοοικονομικοί Κίνδυνοι	20
1.6.1. Κίνδυνοι αγοράς.....	22
1.6.2. Πιστωτικός κίνδυνος.....	23
1.6.3. Κίνδυνος ρευστότητας	24
1.6.4. Επιτοκιακοί κίνδυνοι.....	25
1.6.5. Λειτουργικοί Κίνδυνοι.....	25
1.6.6. Νομικοί κίνδυνοι.....	26
2. ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ.....	27
2.1 Οι τρεις τεχνικές προβλέψεων	28
2.1.1 Ποσοτικές Μέθοδοι.....	28
2.1.2 Ποιοτικές Μέθοδοι.....	29
2.1.3 Τεχνολογικές Μέθοδοι.....	30
2.2 Εισαγωγή και βασικές έννοιες Χρονοσειρών και Προβλέψεων	30
2.2.1 Προσθετικό Υπόδειγμα και συνιστώσες.....	32
2.2.2 Στοχαστικά Υποδείγματα χρονολογικών σειρών.....	37
2.3 Έλεγχος Μη-Γραμμικότητας σε χρονοσειρές.....	39
2.4 Προβλέψεις χρονοσειρών.....	40
2.4.1 Προβλέψεις με την χρήση Μεθόδων Εξομάλυνσης χρονοσειρών.....	41
2.4.1 Προβλέψεις χρονοσειρών με προβολή τάσης στο μέλλον	43
2.4.2 Προβλέψεις χρονοσειρών με εποχιακά στοιχεία.....	43
2.4.3 Προβλέψεις μέσω Μοντέλων Παλινδρόμησης	44
2.4.4 Προβλέψεις μέσω Παλίνδρομων Μοντέλων με ανεξάρτητα σφάλματα.....	45
2.4.5 Προβλέψεις μέσω Παλίνδρομων Μοντέλων με εξαρτώμενα σφάλματα	45
2.5 Γραμμικά Μοντέλα πρόβλεψης χρονοσειρών.....	46
2.5.1 Μοντέλα Προβλέψεων AR, ARMA και ARIMA	48
2.6 Μη-Γραμμικά Μοντέλα πρόβλεψης χρονοσειρών.....	52

2.7. Αστάθεια χρονοσειράς	58
3. Η ΧΡΗΣΗ ΤΗΣ ΓΛΩΣΣΑΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ Python ΣΤΙΣ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΕΣ ΣΥΝΑΛΛΑΓΕΣ.....	60
3.1 Λογισμικό Ανοικτού Κώδικα.....	62
3.1.1 Ορισμός ΕΛ/ΛΑΚ και άδειες χρήσης	62
3.1.2 Πλεονεκτήματα χρήσης ΕΛ/ΛΑΚ από τις επιχειρήσεις	64
3.2 Η γλώσσα προγραμματισμού Python.....	67
3.2.1 Δομές Δεδομένων.....	70
3.2.2 Επιλογή και επανάληψη	71
3.2.2 Βιβλιοθήκες.....	73
3.3 Data Science: Python και Big Data.....	75
3.3.1 Δουλεύοντας με Μεγάλα Δεδομένα: MapReduce.....	79
3.3.2 MapReduce και Hadoop.....	80
3.3.2 MapReduce και Spark	81
3.4 Εφαρμογές χρηματοοικονομικών συναλλαγών σε Python	82
3.4.1 VSTOXX.....	82
4. ΤΕΧΝΟΛΟΓΙΕΣ ΜΕΘΟΔΟΛΟΓΙΑΣ	85
4.1 Διαδικτυακές Αποθήκες Οικονομικών Δεδομένων	85
4.2 Υπολογιστική Νέφος (Cloud Computing)	97
4.2.1 Anaconda Cloud.....	99
4.2.2 Anaconda Jupyter Editor	100
4.3 Βιβλιοθήκες.....	102
4.3.1 Βιβλιοθήκη Numpy	103
4.3.2 Βιβλιοθήκη Pandas.....	104
4.3.3 Βιβλιοθήκη Matplotlib	107
4.3.4 Βιβλιοθήκες ggplot και ggplot2	108
5. ΥΛΟΠΟΙΗΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ	115
5.1 Μεταφορά, διαμόρφωση και οπτικοποίηση δεδομένων	117
5.2 Υλοποίηση και διαμόρφωση διαγραμμάτων των κινητών μέσων	122
5.3 Υλοποίηση χαρτοφυλακίου.....	125
5.4 Προσομοίωση χαρτοφυλακίου.....	136
ΕΠΙΛΟΓΟΣ	143
ΒΙΒΛΙΟΓΡΑΦΙΑ	145

ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ

Εικόνα 2.1: Γραφική απεικόνιση γραμμικής ανοδικής τάσης της S&P για τα έτη 1995-2000 (Πηγή StockCharts.com)	33
Εικόνα 2.2: Λογαριθμική τάση και πρόβλεψη του Bitcoin για τα έτη 2007-2030 (Πηγή StockCharts.com)	34
Εικόνα 2.3: Εκθετική τάση και πρόβλεψη της τιμής της μετοχής S&P 500 (Πηγή StockCharts.com)	34
Εικόνα 2.4: Συνοπτικό διάγραμμα τάσεων	35
Εικόνα 2.5: Ανακατασκευή χώρου καταστάσεων Πηγή http://users.auth.gr/dkugiui/Teach/Econophysics/Chp4.pdf	55
Εικόνα 3.1: Python Shell.....	69
Εικόνα 3.2: Το IDLE της Python	70
Εικόνα 3.3: Το Notebook του Anaconda.....	70
Εικόνα 3.4: VSTOXX Derivates.....	83
Εικόνα 4.1: Αρχική σελίδα του Yahoo Finance	86
Εικόνα 4.2: Ισοτιμία χρυσού	86
Εικόνα 4.3: Σχεδίαση γραφήματος χρονολογικής σειράς του FTSE 100 με κινητούς μέσους	87
Εικόνα 4.4: Αρχική σελίδα του IndexMundi	88
Εικόνα 4.5: Αρχική σελίδα του Energy Information.....	89
Εικόνα 4.6: Αρχική σελίδα του LME.....	89
Εικόνα 4.7: Εκκίνηση του Jupyter και ο κατάλογος εγγράφων	101
Εικόνα 4.8: Περιβάλλον συγγραφής και εκτέλεσης του Jupyter (cell)	102
Εικόνα 5.1: Κατασκευή πίνακα δεδομένων για την εταιρία AEGEAN	117
Εικόνα 5.2: Παράθυρο διαγράμματος των προσαρμοσμένων τιμών κλεισίματος.....	118
Εικόνα 5.3: Συγκριτικό διάγραμμα γραμμικού διαγράμματος και candlestick.....	119
Εικόνα 5.4: Συγκριτικός πίνακας και γραμμικό διάγραμμα των δύο ναυτιλιακών.....	120
Εικόνα 5.5: Διάγραμμα σύγκρισης λογαριθμικών διαφορών	121
Εικόνα 5.6: Διάγραμμα κινητού μέσου ανά μήνα του αποθέματος της AEGEAN για το διάστημα Ιανουάριος-Απρίλιος 2018.....	123
Εικόνα 5.7: Διάγραμμα κινητού μέσου ανά μήνα, δίμηνο και εξάμηνο του αποθέματος της AEGEAN για το διάστημα Ιανουάριος-Απρίλιος 2018	125
Εικόνα 5.8: Πίνακας διαφορών κινητών μέσων	128
Εικόνα 5.9: Εναλλαγές σε αργό και ταχύ κινητό μέσο σε πρόσφατα και σε ιστορικά δεδομένα .	129
Εικόνα 5.10: Πλήθος εναλλαγών αργού και ταχύ κινητού μέσου.....	130
Εικόνα 5.11: Έξοδος σημάτων πώλησης/αγοράς σύμφωνα με τις εναλλαγές	131
Εικόνα 5.12: Ετικέτες αγοράς και πώλησης μετοχών	132
Εικόνα 5.13: Υπολογισμός κέρδους/ζημίας του χαρτοφυλακίου για τέσσερις μήνες	133
Εικόνα 5.14: Υπολογισμός κέρδους/ζημίας σε ιστορικά δεδομένα	134
Εικόνα 5.15: Υπολογισμός κέρδους/ζημίας πάνω σε προσαρμοσμένα ιστορικά δεδομένα	135
Εικόνα 5.16: Λήψη χαμηλών τιμών για τις ημέρες συναλλαγών πάνω σε ιστορικά δεδομένα ...	136
Εικόνα 5.17: Προσομοίωση χαρτοφυλακίου	138
Εικόνα 5.18: Προσομοίωση σημάτων πώλησης/αγοράς των μετοχών του χαρτοφυλακίου	139
Εικόνα 5.19: Προσομοίωση απολαβών του χαρτοφυλακίου	140
Εικόνα 5.20: Τιμές SPY αρχής και τέλους περιόδου	141
Εικόνα 5.21: Αξία χαρτοφυλακίου SPY	141
Εικόνα 5.22: Γραφική συσχέτιση χαρτοφυλακίων.....	142

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

Πίνακας 4-1: Λειτουργίες σειρών	106
Πίνακας 4-2: Λειτουργίες DataFrames.....	106
Πίνακας 4-3: Βασικές συναρτήσεις της ggplot2	109
Πίνακας 4-4: Τιμές τροποποίησης προεπιλογής επικαλυπτόμενων αντικειμένων	111
Πίνακας 4-5: Στατιστικές συναρτήσεις της ggplot2	112
Πίνακας 4-6: Τιμές τροποποίησης προεπιλογής επικαλυπτόμενων αντικειμένων	113
Πίνακας 4-7: Συναρτήσεις αισθητικής	113
Πίνακας 4-8: Ετικέτες και λεζάντες	114
Πίνακας 4-9: Συναρτήσεις τροποποίησης καρτεσιανού επιπέδου	114

ΕΙΣΑΓΩΓΗ

Στις σύγχρονες οικονομίες, τόσο σε επίπεδο χωρών όσο και επιχειρήσεων, το σημαντικότερο ρόλο τον παίζουν οι αγορές χρήματος και κεφαλαίου. Ακόμη πιο σημαντική για τη σύγχρονη εποχή είναι ο όρος της αβεβαιότητας τόσο σε ατομικό ή κοινωνικό επίπεδο όσο και στο οικονομικό. Το ερώτημα αν η αξία του σημερινού χρήματος που διαθέτει θα είναι ίδια ή θα έχει αυξομειώσεις στο μέλλον είναι ένα καθημερινό πρόβλημα όλων των οντοτήτων, είτε αυτά είναι μεμονωμένα άτομα είτε επιχειρήσεις. Οι εμπλεκόμενες οντότητες λαμβάνουν καθημερινά αποφάσεις πάνω στην καλύτερη δυνατή στρατηγική μιας επένδυσης. Οι επενδύσεις θα πρέπει να πραγματοποιούνται με τέτοιο τρόπο έτσι ώστε στην περίπτωση που δεν επιφέρουν κέρδος να οδηγήσουν στην μικρότερη απώλεια.

Το 1920 ο Von Neumann φέρεται να δηλώνει ότι: *“Το Χρηματιστήριο είναι ένα Καζίνο όπου οι πιθανότητες επιτυχίας κλίνουν προς το μέρος του παίκτη”*, υποστηρίζοντας έτσι ότι στην περίπτωση του καζίνο οι πιθανότητες είναι υπέρ της επιχείρησης και όχι του πελάτη. Οι οικονομικές κρίσεις, που ακολούθησαν μετά αυτή την δήλωση¹ οδηγούν στο ερώτημα αν όντως στο Χρηματιστήριο οι πιθανότητες είναι υπέρ του συναλλαζόμενου, εφόσον κατά την διάρκεια τους οι τελευταίοι οδηγήθηκαν σε απώλεια τρισεκατομμυρίων δολαρίων.

Γίνεται λοιπόν αντιληπτό ότι η αβεβαιότητα αλλά και η τυχαία πορεία της αγοράς οδηγούν σε αναπόφευκτες απώλειες. Και αν αυτή η δήλωση ισχύει, προς αποφυγή ή ελαχιστοποίηση των απωλειών, αυτή η αβεβαιότητα μπορεί να ποσοτικοποιηθεί και αν ναι με ποιο τρόπο; Πως μπορούν οι γνωστές θεωρήσεις της γραμμικότητας και της μη-γραμμικότητας των αγορών να οδηγήσουν σε συμπεράσματα για την συμπεριφορά των αγορών και για την μελλοντική πρόβλεψη των τιμών;

Η συγκεκριμένη διατριβή λοιπόν πραγματεύεται το πρόβλημα της μελλοντικής πρόβλεψης των τιμών στις Χρηματιστηριακές αγορές.

¹ Η Μεγάλη Οικονομική Κρίση του 1929, η Μεγάλη Ύφεση του 1937, η Μεγάλη Πετρελαϊκή Ύφεση του 1974 και του 1981, η Μαύρη Δευτέρα του 1989 και η πιο πρόσφατη κρίση του 2007 με τις γνωστές συνέπειες για την ελληνική οικονομία.

Στο πρώτο κεφάλαιο παρουσιάζονται οι θεωρήσεις των αποτελεσματικών και προσαρμοστικών αγορών, θεωρίες που προσπαθούν να εξηγήσουν την συμπεριφορά των χρηματιστηρίων. Σε αυτό επίσης πραγματεύονται οι ανωμαλίες και τα ρίσκα των αγορών.

Το δεύτερο κεφάλαιο αποτελεί μια εισαγωγή στις χρονολογικές σειρές και τα γραμμικά κι μη-γραμμικά μοντέλα πρόβλεψης πάνω σε χρονικά δεδομένα.

Στο τρίτο κεφάλαιο πραγματοποιείται μια ιστορική αναδρομή στα τεχνολογικά συστήματα πρόβλεψης στις χρηματιστηριακές συναλλαγές που υλοποιούνται με χρήση ανοικτού κώδικα και πιο συγκεκριμένα με την γλώσσα προγραμματισμού Python.

Το τέταρτο και το πέμπτο κεφάλαιο παρουσιάζουν τις τεχνολογίες που εμπλέκονται στην υλοποίηση της εφαρμογής πρόβλεψης και παρουσιάζονται τα αποτελέσματα της τελευταίας.

1. ΓΡΑΜΜΙΚΗ ΚΑΙ ΜΗ-ΓΡΑΜΜΙΚΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΔΥΝΑΜΙΚΗ

Σε αυτό το εισαγωγικό κεφάλαιο αναλύονται οι βασικές έννοιες των οικονομικών επενδύσεων, τόσο μέσα από την παρουσίαση των χρηματοοικονομικών παραγώγων. Περαιτέρω παρουσιάζεται η έννοια της αποτελεσματικής αγοράς αλλά και οι ανωμαλίες που επηρεάζουν την τελευταία.

1.1 Χρηματοοικονομικοί τίτλοι και παράγωγα χρηματοοικονομικά προϊόντα

Στον χρηματοοικονομικό κλάδο η χρηματοδότηση των επιχειρήσεων πραγματοποιείται τόσο στα χρηματιστήρια αξιών, μέσω της αγοροπωλησίας χρηματοοικονομικών τίτλων όσο και στα χρηματιστήρια παραγώγων στα οποία συναλλάσσονται διάφορα παράγωγα χρηματοοικονομικά προϊόντα.

Οι πρώτες μορφές διαπραγμάτευσης εισάγονται στην Αρχαία Ελλάδα όπου σύμφωνα με τον (Αριστοτέλης), ο Θαλής ο Μιλήσιος προβλέπει την αυξημένη σοδειά ελιάς του επόμενου έτους και βάση αυτής της πρόβλεψης αποφασίζει την αγορά του δικαιώματος χρήσης των ελαιοτριβείων.

Στην σύγχρονη εποχή η πρώτη οργανωμένη διαπραγμάτευση παραγώγων προϊόντων ανάγεται στο 1688 με τόπο το Χρηματιστήριο του Άμστερνταμ (Amsterdam Bourse), έτος κατά το οποίο ξεκίνησε η διαπραγμάτευση των πρώτων συμβολαίων δικαιωμάτων προαίρεσης (Option Contracts) με αντικείμενο τον βολβό της τουλίπας. Το πρώτο οργανωμένο Χρηματιστήριο παραγώγων δημιουργείται το 1973 στο Σικάγο από το Chicago Board of Trades και το Chicago Mercantile Exchange. Σαν επακόλουθο δημιουργούνται, μεταγενέστερα, τα χρηματιστήρια της Νέας Υόρκης, του Μόντρεαλ, του Τόκιο κ.ά, με το Χρηματιστήριο Παραγώγων Αθηνών να δημιουργείται το έτος 1999. Τα τελευταία 20 χρόνια οι αγορές παραγώγων σε ολόκληρο το κόσμο γιγαντώθηκαν.

Στις επόμενες υποενότητες παρουσιάζονται οι βασικοί χρηματοοικονομικοί τίτλοι καθώς και τα παράγωγα αυτών (ΠΟΥΦΙΝΑΣ, και συν., Δεκέμβριος 2014).

1.1.1 Χρηματοοικονομικοί τίτλοι

Τα *χρεόγραφα* είναι επενδυτικά διαπραγματεύσιμα προϊόντα τα οποία μπορούν να εκδοθούν από κυβερνήσεις, εταιρίες ή άλλους οργανισμούς και αντιπροσωπεύουν μια οικονομική αξία που αποτελεί αποδεικτικό χρέους ή δικαίωμα σε διανεμόμενα κέρδη.

Το νομικό πρόσωπο το οποίο εκδίδει τα χρεόγραφα ονομάζεται *εκδότης* (*issuer*) και σε αυτά περιλαμβάνονται τα *ομόλογα*, τα *έντοκα γραμμάτια Δημοσίου*, τα *μερίδια αμοιβαίων κεφαλαίων*, τα *προθεσμιακά συμβόλαια (forwards)*, τα *συμβόλαια μελλοντικής εκπλήρωσης (futures)*, τα *συμβόλαια δικαιωμάτων προαίρεσης*, τα *παραστατικά απόκτησης μετοχών (warrants)*, οι *μετοχές* άλλα και όσα προϊόντα μπορούν να διαπραγματεύονται στη χρηματοπιστωτική αγορά.

1. **Μετοχές** οι οποίες ορίζονται ως ένα μερίδιο κεφαλαίου μιας ανώνυμης εταιρίας (Α.Ε.). Λόγω της φύσης τους, αποτελούν τίτλο ιδιοκτησίας του κατόχου στην εταιρία που τις έχει εκδώσει. Ειδικότερα, οι μετοχές δημιουργούνται λόγω της ανάγκης κεφαλαίου για την ίδρυση (ή την επέκταση αλλά και την χρηματοδότηση) μιας ανώνυμης εταιρίας. Πιο συγκεκριμένα το αρχικό κεφάλαιο χρηματοδότησης διαιρείται σε υπο-μονάδες² οι οποίες έρχονται στην κατοχή των μετόχων.

Οι κυριότερες μορφές τους³ είναι οι:

1.1 Κοινές μετοχές: Πέρα από την ιδιοκτησία μεριδίου σε μία επιχείρηση δίνουν επίσης δικαίωμα μίας ψήφου στον κάτοχό της. Σε ορισμένες περιπτώσεις ενδέχεται να δίνουν και δικαίωμα μερίσματος αν και η μη καταβολή του τελευταίου δεν συνιστά αθέτηση από μέρους της εταιρίας. Στη διαβάθμιση δικαιωμάτων που δίνονται στους κάτοχους μετοχών, αυτοί των κοινών μετοχών είναι οι τελευταίοι δικαιούχοι πάνω στα περιουσιακά στοιχεία ή το εισόδημα της επιχείρησης (*residual claim*) στην περίπτωση που τα περιουσιακά στοιχεία της

² Μικρότερα και ίσα μερίδια που ονομάζονται μετοχές

³ Η ταξινόμηση γίνεται με βάση τα δικαιώματα που προσφέρονται στους κατόχους

τελευταίας ρευστοποιηθούν. Επιπλέον, σε περίπτωση πτώχευσης η απώλεια για τον κάτοχο της κοινής μετοχής περιορίζεται στην αρχική του επένδυση (limited liability) εφόσον δεν υπάρχει επιπλέον απαίτηση επί των προσωπικών του περιουσιακών στοιχείων.

1.2 Προνομιούχες μετοχές: Οι προνομιούχες μετοχές, σε αντίθεση με τις κοινές, προσφέρουν στους κατόχους τους ένα ετήσιο μερίσμα. Και στην κατηγορία των προνομιούχων μετοχών η μη καταβολή του μερίσματος δεν συνιστά αθέτηση από μέρους της εταιρίας. Ωστόσο, οι κάτοχοι προνομιούχων μετοχών έχουν προτεραιότητα πληρωμής του μερίσματος σε σχέση με τους κατόχους κοινών μετοχών. Όσον αφορά το δικαίωμα ψήφου υπάρχουν προνομιούχες μετοχές με ή χωρίς αυτό το δικαίωμα.

2. **Δανειακά ομόλογα** (bonds) είναι ένας διαφορετικός τρόπος χρηματοδότησης τόσο των ανώνυμων εταιριών όσο του Δημοσίου και άλλων οργανισμών. Τα ομόλογα που εκδίδονται (όπως και τα λοιπά χρεόγραφα) αντιπροσωπεύουν μια οικονομική αξία και αποτελούν αποδεικτικό χρέους ή και δικαίωμα σε διανεμόμενα κέρδη με συγκεκριμένη ημερομηνία λήξης. Στην περίπτωση των ομολόγων ο κάτοχος τους είναι, ουσιαστικά, δανειστής του εκδότη και ο τελευταίος είναι υποχρεωμένος με το πέρας του χρόνου διαρκείας τους να καταλάβει στον κάτοχο τόσο το αρχικό ποσό όσο και τον τόκο του τίτλου.

Ανάλογα με το είδος επιτοκίου τα ομόλογα χωρίζονται στις κατηγορίες:

- Σταθερού επιτοκίου
- Κυμαινόμενου επιτοκίου
- Zero coupons bonds: πωλούνται με σημαντική έκπτωση έναντι της ονομαστικής τους αξίας⁴ αλλά σε αντίθεση με τις δύο προηγούμενες κατηγορίες οι κάτοχοι τους δεν έχουν δικαίωμα περιοδικών αποπληρωμών των επιτοκίων.

Η χρηματοδότηση των επιχειρήσεων, που πραγματοποιείται μέσω της διάθεσης των μετοχών και των ομολόγων τους στο επενδυτικό κοινό, πραγματοποιείται στο

⁴ **Ονομαστική αξία** ενός ομολόγου είναι η αξία που αναγράφεται στο ομόλογο όταν εκδίδεται και συνήθως είναι η τιμή εξόφλησης, δηλαδή το χρηματικό ποσό που θα αποδοθεί στον επενδυτή στη λήξη του ομολόγου.

Χρηματιστήριο. Το τελευταίο είναι ένας οργανισμός ο οποίος επιτηρεί τις συναλλαγές που λαμβάνουν χώρα πάνω σε κινητές αξίες, σε παράγωγα προϊόντα, με τελικό σκοπό την διασφάλιση της εύρυθμης λειτουργίας της αγοράς και τη προστασία του επενδυτικού κοινού. Στο Χρηματιστήριο, πέρα από τις μετοχές αλλά και τα ομόλογα (τραπεζικά και Ελληνικού Δημοσίου), συναλλάσσονται ομολογίες ανωνύμων εταιρειών καθώς και συμβόλαια μελλοντικής εκπλήρωσης αλλά και τα δικαιώματα προαίρεσης ή προτίμησης. Η άντληση των κεφαλαίων εκ μέρους της επιχείρησης πραγματοποιείται είτε με την εισαγωγή τους στο Χρηματιστήριο, είτε μέσω αυξήσεων του μετοχικού τους κεφαλαίου.

Από την πλευρά του επενδυτικού κοινού, απώτερος σκοπός μιας επένδυσης, είναι το κέρδος μιας ικανοποιητικής απόδοσης (μέσω μερισμάτων), υψηλότερης από αυτήν που προσφέρουν επενδύσεις όπως οι τραπεζικές καταθέσεις και τα κρατικά ομόλογα. Η διαδικασία περιλαμβάνει αγοροπωλησίες επενδύσεων με σκοπό την μεγιστοποίηση του κέρδους. Κατ' αυτόν τον τρόπο διαμορφώνεται μια τιμή των διαπραγματευόμενων προϊόντων όπου η προσφορά και η ζήτηση ισορροπούν κάθε δεδομένη χρονική στιγμή.

1.1.2 Χρηματοοικονομικά παράγωγα

Στον χρηματοοικονομικό κλάδο ως *παράγωγο προϊόν* ονομάζεται μια διμερής σύμβαση, η αξία της οποίας καθορίζεται από την αξία κάποιου άλλου, βασικότερου προϊόντος ή διαφορετικά *υποκείμενου προϊόντος (underlying asset)*. Κατ' αυτή την έννοια ένα παράγωγο προϊόν είναι ένα αξιόγραφο, η τιμή του οποίου καθορίζεται με άμεσο τρόπο από την τιμή του υποκείμενου τίτλου. Όπως σε όλα τα είδη συμβολαίων έτσι και εδώ εμπλέκονται δύο αντισυμβαλλόμενοι από τους οποίους ο ένας καταλαμβάνει την θέση του αγοραστή ενώ ο άλλος έχει τη θέση του πωλητή.

Τα παράγωγα προϊόντα προέρχονται και συνδέονται άμεσα με τα αντίστοιχα υποκείμενα με την έννοια ότι τα τελευταία μπορεί να είναι είτε προϊόντα που τίθενται υπό διαπραγμάτευση σε μία οργανωμένη δευτερογενή αγορά, (π.χ. ένα Χρηματιστήριο), είτε προϊόντα που εισάγονται σε μη οργανωμένες αγορές. Έτσι, γενικότερα, τα υποκείμενα προϊόντα μπορεί να είναι σχεδόν οτιδήποτε παρουσιάζει μια εμπορεύσιμη μορφή (από μετοχές και ομόλογα μέχρι αγροτικά προϊόντα, μέταλλα, ενέργεια κ.τ.λ) αλλά και τα παράγωγα αυτών.

Τα πιο γνωστά παράγωγα προϊόντα είναι:

1. **Προθεσμιακά συμβόλαια (ΠΣ)** που είναι και η πιο απλή μορφή παραγώγου. Οι διμερής συμφωνίες συνήθως πραγματοποιούνται μεταξύ δύο χρηματοοικονομικών ιδρυμάτων και όχι στην χρηματιστηριακή αγορά.

Τα προθεσμιακά συμβόλαια λαμβάνουν χώρα μεταξύ δύο συμβαλλομένων, με το ένα μέλος να υπόσχεται να αγοράσει και το άλλο να πουλήσει μία συγκεκριμένη ποσότητα ενός αγαθού, με προκαθορισμένη τιμή συναλλαγής και σε μία καθορισμένη ημερομηνία στο μέλλον.

Ο υπολογισμός του κέρδους ενός αγοραστή από ένα προθεσμιακό συμβόλαιο δίνεται ως:

$$S_T - K,$$

όπου S_T είναι η τιμή του υποκείμενου αγαθού κατά τον χρόνο της λήξης του συμβολαίου και K είναι η προκαθορισμένη τιμή συναλλαγής.

Το κέρδος του πωλητή, από την άλλη, υπολογίζεται ως:

$$K - S_T .$$

Η προκαθορισμένη τιμή συναλλαγής (K) ορίζεται την ημέρα σύναψης της διμερούς συμφωνίας ενώ η τιμή του υποκείμενου αγαθού κατά τον χρόνο λήξης του συμβολαίου (S_T) είναι τυχαία μεταβλητή, εφόσον δεν είναι γνωστή εξ αρχής. Η τιμή συναλλαγής K ορίζεται με τέτοιο τρόπο ώστε η αξία του προθεσμιακού συμβολαίου αρχικά να είναι μηδέν.

2. **Συμβόλαιο Μελλοντικής Εκπλήρωσης (ΣΜΕ)**. Και σε αυτή την περίπτωση πραγματοποιείται μία συμφωνία μεταξύ δύο συμβαλλομένων, ο ένας εκ των οποίων υπόσχεται να αγοράσει και ο άλλος να πουλήσει, μία συγκεκριμένη ποσότητα ενός αγαθού, σε μία καθορισμένη ημερομηνία στο μέλλον, σε μία προκαθορισμένη τιμή συναλλαγής.

Έτσι αυτός που αγοράζει προσβλέπει στην άνοδο της τιμής του αγαθού ενώ αντίθετα αυτός που πουλά στην πτώση της τιμής του αγαθού. Το υποκείμενο αγαθό υπό διαπραγμάτευση μπορεί να είναι οποιοδήποτε εμπόρευμα ή χρηματοοικονομικό αγαθό.

Τα ΣΜΕ συναλλάσσονται στο Χρηματιστήριο Παραγώγων (αυτή είναι και η διαφορά τους με τα ΠΣ) το οποίο παίρνει το μέρος του εγγυητή-μεσολαβητή

ανάμεσα στα συμβαλλόμενα μέλη μέσω της χρήσης λογαριασμών περιθωρίων⁵. Λόγω της παρουσίας τους στο Χρηματιστήριο Παραγώγων η αξία συναλλαγής των ΣΜΕ μεταβάλλεται ανάλογα με τις μεταβολές της τιμής του υποκείμενου αγαθού και τις προσδοκίες των επενδυτών. Μέσω της διαδικασίας του ημερήσιου διακανονισμού (marking to market) αγοραστές και πωλητές μπορούν να κλείσουν τη θέση τους οποιαδήποτε στιγμή (ακόμα και πριν τη λήξη του συμβολαίου) κάνοντας την αντίστροφη κίνηση (δλδ. οι αγοραστές πωλούν ενώ οι πωλητές αγοράζουν). Με αυτόν τον τρόπο το κέρδος (ή η ζημία) λαμβάνεται κατά τμήματα και πιστώνεται (ή χρεώνεται ανάλογα) κάθε χρηματιστηριακή ημέρα. Το περιθώριο ασφάλισης (margin) αποδεσμεύεται όταν ο επενδυτής κλείσει την ανοικτή του θέση.

3. Προϊόντα Δανεισμού Τίτλων (ΔΤ). Η διαδικασία λαμβάνει χώρα με δύο διαφορετικές μορφές:

- Οι μετοχές μεταβιβάζονται προσωρινά και ως “δάνειο” (stock lending – Repo) από κάποιον επενδυτή στο Χρηματιστήριο Παραγώγων. Κατ’ αυτόν τον τρόπο ο επενδυτής εισπράττει ένα μηνιαίο έσοδο αλλά επίσης δικαιούται να λάβει ένα τεχνητό μέρισμα από τις μετοχές του (repos).
- Οι μετοχές αποκτούνται προσωρινά ως “δάνειο” (stock borrowing - Reverse Repo) από κάποιον επενδυτή μέσω του Χρηματιστηρίου Παραγώγων έναντι ημερήσιου κόστους. Παρόλο που το έσοδο αυτής της επένδυσης είναι πάντα θετικό δεν είναι γνωστό εκ των προτέρων.

4. Δικαιώματα προαίρεσης (ΔΠ ή options). Η διμερής συμφωνία πραγματοποιείται μεταξύ του αγοραστή και τον πωλητή, με αντικείμενο διαπραγμάτευσης ενός δικαιώματος, με μεσολάβηση του Χρηματιστηρίου Παραγώγων. Στην συγκεκριμένη περίπτωση, ο αγοραστής έχει το δικαίωμα (άλλα όχι απαραίτητα την υποχρέωση) να αγοράσει (ή να πωλήσει) από τον πωλητή του δικαιώματος ένα αγαθό A (χρεόγραφο, συνάλλαγμα ή εμπόρευμα) σε μία προκαθορισμένη τιμή K, κατά τη διάρκεια μίας χρονικής περιόδου [0, T]. Η διαφορά του δικαιώματος προαίρεσης με τα ΣΜΕ και τα ΠΣ έγκειται στο γεγονός ότι ο αγοραστής του δικαιώματος μπορεί να εξασκήσει το δικαίωμά του μετά από δική του απόφαση και ο πωλητής του δικαιώματος είναι υποχρεωμένος να

⁵ Ο λογαριασμός περιθωρίων δημιουργείται με κατάθεση ενός ποσού υπό μορφή εγγύησης (ένα ποσοστό της τιμής συναλλαγής K) τόσο εκ μέρους του αγοραστή όσο και του πωλητή

ακολουθήσει την απόφαση του αγοραστή. Μέσω αυτής της διαδικασίας ο αγοραστής λαμβάνει ένα πλεονέκτημα και γι' αυτό το λόγο πρέπει να καταβάλει ένα αντίτιμο (ονομάζεται ασφάλιστρο ή τιμή δικαιώματος - Option price, option premium) στον πωλητή έτσι ώστε να αποκτήσει το δικαίωμα.

Επιμέρους κατηγοριοποίηση των δικαιωμάτων προαίρεσης γίνεται βάση του είδους του δικαιώματος (δικαίωμα αγοράς ή πώλησης), του υποκείμενου τίτλου ή εμπορεύματος, της ημερομηνίας λήξης (Αμερικάνικου τύπου ή Ευρωπαϊκού τύπου⁶), της προκαθορισμένης τιμής (strike price) Κ αγοράς (ή πώλησης) του δικαιώματος και του ύψους του ασφάλιστρου.

1.2 Η Υπόθεση Αποτελεσματικής Αγοράς

Ως χρηματοοικονομική ορίζεται ο κλάδος της επιστήμης των οικονομικών που ασχολείται με την μελέτη αναλυτικών μεθόδων απαραίτητων για τη λήψη επιτυχημένων χρηματοοικονομικών αποφάσεων κάτω από (πολύπλοκες) συνθήκες αβεβαιότητας που επικρατούν σε ένα ανταγωνιστικό και διεθνοποιημένο επιχειρηματικό περιβάλλον. Περιλαμβάνει μεθόδους λήψης αποφάσεων πάνω στην κατανομή και την ανταλλαγή χρηματοοικονομικών πόρων διαφόρων ομάδων επενδυτών όπως τα νοικοκυριά, οι επιχειρήσεις, οι πιστωτές, οι επενδυτές και οι οικονομίες ως σύνολο. Ο κλάδος των χρηματοοικονομικών εστιάζει, επίσης, στην έρευνα θεμάτων όπως η συμπεριφορά των χρηματοοικονομικών αγορών υπό ακραίες συνθήκες όπως αυτό μιας οικονομικής κρίσης.

1.2.1 Ορισμός της Αποτελεσματικής Αγοράς

Όπως αναφέρεται στο (De Moor, et al., 2013), η έννοια της αποτελεσματικής αγοράς είναι βασική στην ανάλυση αγορών χρήματος και κεφαλαίου. Η εισαγωγή της έγινε στα τέλη του 19^{ου} αιώνα από τον G. Gibson. Η αποτελεσματικότητα της αγοράς αφορά πολλά επίπεδα της αγοράς, αλλά από τους ερευνητές δίνεται ιδιαίτερη έμφαση σε τρία από αυτά τα οποία και είναι:

⁶Αμερικανικού τύπου (American option): το δικαίωμα προαίρεσης μπορεί να εξασκηθεί οποιαδήποτε στιγμή. Ευρωπαϊκού τύπου (European option): το δικαίωμα προαίρεσης μπορεί να εξασκηθεί μόνο κατά την ημερομηνία λήξης.

- το επίπεδο της κατανομής των πόρων. Εδώ μελετάται η κατανομή των τελευταίων με τον πιο παραγωγικό τρόπο (allocational efficiency),
- το επίπεδο της διεξαγωγής των συναλλαγών σε ανταγωνιστικό κόστος για τους συμμετέχοντες στην αγορά (operational efficiency) και
- το επίπεδο του βαθμού διαθεσιμότητας της πληροφορίας που αντανακλάται στις τιμές των χρεογράφων (informational efficiency) και εξαρτάται κυρίως από τα δυο προηγούμενα επίπεδα.

Το επίπεδο του βαθμού διαθεσιμότητας της πληροφορίας λαμβάνεται ως το πιο σημαντικό από τα τρία προαναφερθέντα επίπεδα λόγω της άποψης από μέρους των μελετητών ότι οι χρηματιστηριακές αγορές μπορούν να είναι αποτελεσματικές σε επίπεδο πληροφορίας ανεξάρτητα από τα άλλα δύο. Ποιοί είναι οι λόγοι που καθιστούν τον βαθμό διαθεσιμότητας της πληροφορίας ως τόσο σημαντική και χρηστική στην αποτελεσματικότητα της αγοράς; Η απάντηση είναι ότι οι χρηματοοικονομικές αγορές, στην ουσία, αποτελούν ένα κανάλι μεταφοράς κεφαλαίων συνδέοντας τους αποταμιευτές (που είναι οι μέτοχοι – επενδυτές) με τους δανειολήπτες (που είναι οι εταιρίες). Αυτή η διασύνδεση είναι που επηρεάζει την αποτελεσματικότητα της οικονομίας διαμέσου της τοποθέτησης κεφαλαίων σε παραγωγικές επενδύσεις. Κατ' επέκταση η αποτελεσματική λειτουργία των αγορών είναι αυτή που επηρεάζει άμεσα και με θετικό τρόπο την πορεία της οικονομίας. Στην αντίθετη περίπτωση, η μη αποτελεσματική λειτουργία των αγορών έχει σαν αποτέλεσμα την αποσταθεροποίηση της οικονομίας με ότι και αν αυτό συνεπάγεται.

Ένας επιπλέον λόγος που καθιστά την αποτελεσματικότητα της αγοράς τόσο σημαντικό παράγοντα και επομένως πεδίο έρευνας, είναι το γεγονός ότι μια αναποτελεσματική αγορά μεγιστοποιεί το επίπεδο του κινδύνου απόδοσης των επενδύσεων και με αυτόν τον τρόπο επηρεάζει και τα δύο διαπραγματεύσιμα μέλη της επένδυσης.

Λόγω τόσο της σημαντικότητας της αλλά και της πολυπλοκότητας της (αποτέλεσμα των διάφορων παραγόντων που εμπλέκονται στον ορισμό της αποτελεσματικότητας), μέχρι και σήμερα δεν έχουν οριστεί τα σαφή πλαίσια για το πότε μια αγορά χρήματος και κεφαλαίου είναι αποτελεσματική ή όχι. Σε γενικές γραμμές γίνεται δεκτό όμως ότι μια αγορά μπορεί να χαρακτηριστεί σαν αποτελεσματική στην περίπτωση στην οποία οι τιμές της αντανακλούν την θεμελιώδη πληροφορία για τις εταιρίες.

Βιβλιογραφικά έχουν καταγραφεί διάφορα χαρακτηριστικά που θα πρέπει να διέπουν μια αγορά χρήματος και κεφαλαίου έτσι ώστε αυτή να θεωρηθεί σαν αποτελεσματική όπως αυτό που περιγράφεται στο (Cuthbertson, et al., 2004) όπου οι συγγραφείς θεωρούν πως η αποτελεσματικότητα παρουσιάζεται σε μια αγορά που αποτελείται από ορθολογικούς επενδυτές. Συμπληρωματικά έρχεται η θεώρηση του Damodaran στο (Damodaran, 2006) στην οποία μια αγορά θεωρείται ως αποτελεσματική στην περίπτωση που ο επενδυτής μπορεί να οδηγηθεί σε μια αντικειμενική εκτίμηση της πραγματικής αξίας των αξιογράφων υπό διαπραγμάτευση, συνθήκη η οποία μπορεί να πραγματοποιηθεί μόνο στην περίπτωση που τα τελευταία περιέχουν όλη τη σχετική πληροφόρηση δημοσιευμένη ή όχι. Διαφορετική είναι η θεώρηση στο (Brealey, et al., 2017) όπου σύμφωνα με τους συγγραφείς ο όρος αποτελεσματικότητα σχετίζεται με την απόδοση της αγοράς και στο (Mishkin, et al., 2012) οι οποίοι ορίζεται ότι οι τιμές των χρεογράφων θα πρέπει να αντικατοπτρίζουν όλη τη διαθέσιμη πληροφορία.

Λόγω των δύο τελευταίων θεωρήσεων, σε μια αποτελεσματική αγορά οι πληροφορίες πάνω στα χρεόγραφα έχουν ήδη προεξοφληθεί και αντικατοπτρίζονται ήδη στην τιμή τους. Επιπρόσθετα σε ένα τέτοιο είδος αγοράς οι μετοχές δεν μπορούν να είναι ούτε υποτιμημένες αλλά ούτε υπερτιμημένες και κατ' επέκταση οι επενδυτές μπορούν να επιτύχουν μόνο κανονικές αποδόσεις⁷.

Συμπερασματικά, μπορεί να υποστηριχθεί ότι ο ορισμός της αποτελεσματικότητας αγοράς αφορά δύο διαφορετικά χαρακτηριστικά: πρώτο την διαφάνεια και την διαθεσιμότητα πληροφορίας πάνω στα αξιόγραφα υπό διαπραγμάτευση και δεύτερο, το γεγονός ότι κανένας επενδυτής δεν μπορεί να χρησιμοποιήσει πληροφορίες για να επιτύχει αποδόσεις πέρα του κανονικού.

1.2.2 Η Υπόθεση των Αποτελεσματικών Αγορών και τα είδη τους

Όπως προαναφέρθηκε, στο προηγούμενο υποκεφάλαιο, η *Υπόθεση* (ή *Θεωρία*) της *Αποτελεσματικής Αγοράς* (*ΥΑΑ* ή *ΘΑΑ*)⁸ θεωρείται κεντρική στα χρηματοοικονομικά και αποτελεί τον θεμέλιο λίθο της σύγχρονης χρηματοοικονομικής τόσο σε θεωρητικό όσο και

⁷Ως κανονική θεωρείται η απόδοση μετοχής η οποία είναι συνάρτηση του επενδυτικού κινδύνου τον οποίο έχει αναλάβει ο επενδυτής

⁸ Efficient Market Theory ή Efficient Market Hypothesis – (EMT ή EMH)

σε πρακτικό επίπεδο. Σύμφωνα με το θεμελιωτή της Υπόθεσης, τον E. Fama στο (Fama, 1965) αλλά και σε ένα πλήθος δημοσιεύσεων που ακολούθησαν, μια αγορά θεωρείται αποτελεσματική όταν “υπό συνθήκες **ισορροπίας** οι παρούσες τιμές των αξιογράφων αντικατοπτρίζουν πλήρως κάθε σχετική και διαθέσιμη πληροφορία κατά τρόπο άμεσο, γρήγορο και ακριβή και κατ’ επέκταση οι τιμές στην αγορά αντικατοπτρίζουν την πραγματική αξία του αξιογράφου”. Στις πληροφορίες πάνω στα αξιόγραφα συμπεριλαμβάνονται οι παρελθούσες τιμές των μετοχών, οι σημερινές τιμές θεμελιωδών μεταβλητών της οικονομίας όπως και εταιρικά νέα τα οποία μπορεί να επηρεάσουν στην απόδοσή τους. Κατ’ αυτό τον τρόπο οι τιμές μιας μετοχής θα πρέπει να διαφοροποιούνται μόνο στην περίπτωση στην οποία στις αγορές φτάνει μια νέα πληροφορία πάνω σε αυτές.

Στις χρηματοοικονομικές αγορές η πλήρης και άμεση διαθέσιμη πληροφορία μπορεί να είναι ένας σημαντικός παράγοντας πρόβλεψης διακύμανσης των τιμών δεν αρκεί όμως και δεν καλύπτει (ή δεν μπορεί να εξηγήσει) φαινόμενα ακραίων καταστάσεων όπως αυτό μιας οικονομικής κρίσης. Ο κύριος λόγος είναι ότι οι τελευταίες εισάγουν στην αγορά φαινόμενα έντονης μεταβλητότητας, με αποτέλεσμα να αυξάνεται η πολυπλοκότητα πρόβλεψης λόγω των απότομων και έντονων διακυμάνσεων στους δείκτες.

Στο (Fama, 1970) περιγράφονται οι τρεις βασικές υποθέσεις που οδηγούν σε μια αποτελεσματική αγορά είναι οι παρακάτω:

1. Μια αγορά για να θεωρηθεί αποτελεσματική θα πρέπει αντανακλά πλήρως και άμεσα τις πληροφορίες στις τιμές των μετοχών με τέτοιο τρόπο ώστε οι τιμές των μετοχών να αντικατοπτρίζουν πλήρως την αξία τους. Έτσι οι μεταβολές στις τιμές των αξιογράφων (και κατά συνέπεια και οι αποφάσεις των επενδυτών), θα πρέπει να μεταβάλλονται μόνο έπειτα από την εισροή στην αγορά νέων πραγματικών πληροφοριών για την κατάσταση και τις προοπτικές μιας εταιρίας
2. Επιπρόσθετα, η αγορά θα πρέπει να αποτελείται από ένα μεγάλο αριθμό ορθολογικών επενδυτών. Ως ορθολογικοί θεωρούνται εκείνοι οι επενδυτές που λειτουργούν προς όφελος τους και επομένως ανταγωνίζονται μεταξύ τους στην προσπάθεια μεγιστοποίησης των κερδών τους. Επιπλέον, οι επενδυτές θα πρέπει να παρουσιάζουν μια άμεση και ακριβής αντίδραση κατά την εισαγωγή στην αγορά της νέας πληροφορίας, αναπροσαρμόζοντας τις τιμές των αξιογράφων ανάλογα με την τελευταία.

3. Η πληροφορία πάνω στις μετοχές θα πρέπει να είναι καθολική, ισότιμη και ελεύθερη σε όλους τους επενδυτές.
4. Οι παρελθούσες τιμές που έχουν διαμορφωθεί πριν την εισροή νέας πληροφορίας, έχουν προσαρμοστεί σύμφωνα με την ίδια διαδικασία και επομένως έχουν ενσωματώσει όλες τις παρελθούσες πληροφορίες.

Το γεγονός ότι η ροή των πληροφοριών είναι συνεχής αλλά τυχαία καθιστά και τις μεταβολές πάνω στις τιμές των μετοχών μια τυχαία και διαρκή διαδικασία (random walk) και έχει σαν αποτέλεσμα η κατάσταση ισορροπίας της αγοράς να μεταβάλλεται συνεχώς. Ως εκ τούτου, αν όντως οι τιμές των μετοχών ακολουθούν ένα τυχαίο περίπατο γεγονός που σημαίνει ότι η παρούσα τιμή ενός αξιόγραφου δεν επηρεάζεται από τις παρελθούσες τιμές του, είναι αδύνατη η υιοθέτηση μιας στρατηγικής που θα αποφέρει κέρδος μεγαλύτερο από αυτό της αγοράς.

Τόσο το είδος όσο και ο όγκος της άμεσα διαθέσιμης πληροφορίας οδήγησαν στην κατηγοριοποίηση της αποτελεσματικότητας της αγοράς σε τρεις βασικές μορφές:

- Την *ασθενή μορφή* όπου η τιμή της μετοχής περιέχει (και επομένως αντικατοπτρίζει) μόνο ιστορική πληροφορία και σε αυτή περιλαμβάνονται οι προηγούμενες τιμές της μετοχής, όγκο συναλλαγών της στην αγορά κτλ. Λαμβάνοντας υπόψη μόνο την συγκεκριμένη πληροφορία οι τιμές των αξιογράφων είναι ανεξάρτητες μεταξύ τους και ακολουθούν μια κανονική κατανομή πιθανοτήτων γεγονός που οδηγεί στο συμπέρασμα ότι συγκεκριμένη μορφή αποτελεσματικότητας δεν μπορεί να οδηγήσει σε υπερβολικό κέρδος εκ μέρος των συμμετεχόντων.
- Την *ημι-σχυρή μορφή* όπου η τιμή της μετοχής πέρα από την ιστορική πληροφορία αντικατοπτρίζει επίσης την τρέχουσα διαθέσιμη πληροφορία, όπως για παράδειγμα τα επίπεδα επιτοκίων, τον πληθωρισμό κτλ.
- Την *ισχυρή μορφή*, όπου η τιμή της μετοχής περιλαμβάνει και αντικατοπτρίζει όλη την πιθανή πληροφορία είτε αυτή είναι δημόσια είτε ιδιωτική.

Η Υπόθεση Αποτελεσματικότητας της Αγοράς έχει αμφισβητηθεί κατά καιρούς από διάφορους επιστήμονες μέσω δημοσιευμένων εργασιών. Οι εργασίες αυτές καταλήγουν σε συμπεράσματα μη αποτελεσματικότητας των αγορών, συμπεράσματα που προκύπτουν μέσω επισημάνσεων συστηματικών μοτίβων και ιστορικών ανωμαλιών στις κινήσεις των

τιμών των μετοχών. Στο βιβλίο (Lo, et al., 1999) οι συγγραφείς υποστηρίζουν ότι η υπόθεση τυχαίου περιπάτου στην ΥΑΑ είναι εσφαλμένη, και ισχύει μόνο στην περίπτωση της ουδετερότητας του κινδύνου. Άλλοι επιστήμονες υποστηρίζουν ότι αποτελεσματικές αγορές μπορούν να υπάρχουν θέτοντας όμως σαν όρο την μη ύπαρξη κόστους απόκτησης της πληροφορίας.

Συμπερασματικά, η ΥΑΑ από την μία πλευρά θεωρείται σημαντική καθώς εισάγει την έννοια της διάχυσης της πληροφορίας από τη άλλη, όμως, από μόνη της δεν είναι σε θέση να εξηγήσει την εμφάνιση προτύπων στις τιμές των μετοχών.

1.3 Αποτελεσματικότητα Αγοράς και ανωμαλίες

Σε αυτό το υποκεφάλαιο θα εξεταστούν οι διάφορες ανωμαλίες που εμφανίζουν οι αγορές αξιογράφων αλλά και οι συνθήκες που τις εισάγουν. Ως ανωμαλίες της αγοράς (market anomalies) ορίζονται τα γεγονότα που δεν είναι τα αναμενόμενα (αυτά που ορίζονται δηλαδή από την Αποτελεσματικότητα της Αγοράς) σε μια ανταγωνιστική αγορά και γι' αυτό και θεωρούνται εξαιρέσεις στην αποτελεσματικότητα της αγοράς⁹.

Οι ανωμαλίες της αγοράς είναι αντίθετες με ορισμένες από τις θεμελιώδεις αρχές της σύγχρονης χρηματοοικονομικής, όπως είναι η υπόθεση περί αποτελεσματικότητας (Market Efficiency Hypothesis, MEH) των αγορών αλλά και το μοντέλο αποτίμησης περιουσιακών στοιχείων (Capital Asset Pricing Model, CAPM) που αναπτύχθηκαν από τους Sharpe, Lintner και Mossin στα (Sharpe, 1964), (Lintner, 1965) και (Mossin, 1966) αντίστοιχα.

Το μοντέλο αποτίμησης περιουσιακών στοιχείων έχει εισαχθεί ως θεωρία τιμολόγησης ενός περιουσιακού στοιχείου για τον υπολογισμό της απόδοσης μιας επένδυσης. Η τελευταία, σύμφωνα με τη θεωρία υπολογίζεται ως το επιτόκιο μηδενικού κινδύνου πλέον ενός premium (που αντιστοιχεί στον κίνδυνο που αναλαμβάνουν οι επενδυτές για το περιουσιακό στοιχείο) πολλαπλασιασμένο με τον συστηματικό κίνδυνο του περιουσιακού στοιχείου. Έτσι, το CAPM είναι ένα υπόδειγμα που περιγράφει τη σχέση που προκύπτει, στην περίπτωση χρεογράφων που εμπεριέχουν ρίσκο, μεταξύ του κινδύνου και της

⁹ Ειδικότερα για την ημι-ισχυρή της μορφή

αναμενόμενης απόδοσης ενός κεφαλαιακού περιουσιακού στοιχείου και δίνεται από τον τύπο:

$$\text{Αναμενόμενη Απόδοση} = \text{Ακίνδυνο Επιτόκιο} + \beta * \left(\text{Απόδοση Αγοράς} - \text{Ακίνδυνο Επιτόκιο} \right)$$

Έτσι, η ΥΑΑ σε συνδυασμό με το CAPM βεβαιώνουν ότι η αγοραία τιμή ενός αξιογράφου απεικονίζει την καλύτερη εκτίμηση της βασικής αξίας τους. Οι διαφορές στην εκτιμώμενη τιμολόγηση με την τιμή που όντως παρουσιάζεται στην αγορά αποτελεί μια ανωμαλία.

Οι σημαντικότερες από τις ανωμαλίες που παρουσιάζονται στις αγορές είναι:

Ανωμαλίες της θεμελιώδους αξίας

- Το φαινόμενο της Κεφαλαιοποίησης ή το μέγεθος της επιχείρησης (The size effect): η παρατήρηση ότι, με βάση τον προσαρμοσμένο κίνδυνο, κατά μέσο όρο οι μικρές εταιρείες έχουν μεγαλύτερα κέρδη από τις μεγάλες.
- Η επίδραση της μεταβλητής P/E στις αποδόσεις (The price to earnings effect P/E): η παρατήρηση ότι, μετοχές με χαμηλό δείκτη P/E αποδίδουν υπερκέρδη σε σχέση με την αγορά ενώ μετοχές με υψηλό δείκτη P/E να υπο-αποδίδουν.
- Η επίδραση των παρελθόντων αποδόσεων (Past returns effect Momentum and Contrarian Strategies): η παρατήρηση ότι οι παρούσες αποδόσεις των μετοχών σχετίζονται με τις παρελθούσες τους επιδόσεις.
- Η επίδραση της μεταβλητής BE/ME στις αποδόσεις (Book value equity/market value equity BV/MV effect).
- Η επίδραση της μερισματικής απόδοσης (The dividend-yield DY effect): η παρατήρηση ότι, μετοχές με υψηλή μερισματική απόδοση αποδίδουν υπερκέρδη σε σχέση με την αγορά.
- Ο λόγος των πωλήσεων ανά μετοχή προς την τιμή ανά μετοχή (The sales to price effect).
- Ο λόγος της λογιστικής προς την χρηματιστηριακή αξία (The book to market effect): η παρατήρηση ότι, μετοχές με χαμηλό δείκτη λογιστικής προς χρηματιστηριακής αξίας αποδίδουν μεγαλύτερα κέρδη από μετοχές με υψηλό δείκτη.
- Ο λόγος των ταμειακών Ροών προς την τιμή ανά μετοχή (The Cash Flow to Price Effect).

Ημερολογιακές ή εποχικές ανωμαλίες

- Το φαινόμενο του Ιανουαρίου (The January effect): η τάση των τιμών των μετοχών μικρών εταιριών να αποδίδουν μεγαλύτερα κέρδη από κατηγορίες άλλων μετοχών τις πρώτες τρεις εβδομάδες του Γενάρη.
- Το φαινόμενο της Δευτέρας (The Monday effect): η τάση των τιμών των μετοχών να μειώνονται τη Δευτέρα σε σχέση με το κλείσιμο της τιμής την προηγούμενη Παρασκευή.
- Το φαινόμενο της αλλαγής μήνα: η τάση των τιμών των μετοχών να αυξάνονται την τελευταία μέρα ενός μήνα και τις τρεις πρώτες μέρες του επόμενου.

Άλλες επιδράσεις που μπορεί να οδηγήσουν σε ανωμαλίες της αγοράς είναι αυτές της εσωτερικής πληροφόρησης, ορμής, της υπέρ-αντίδρασης και της υπό-αντίδρασης.

1.4 Μη-Γραμμικές δυναμικές στις αγορές Χρήματος και Κεφαλαίου

Πλήθος επιστημονικών άρθρων στην βιβλιογραφία υποστηρίζουν ότι στο σύγχρονο χρηματιστηριακό περιβάλλον οι αγορές που όντως μπορούν να χαρακτηριστούν ως αποτελεσματικές είναι ελάχιστες και μάλλον η εξαίρεση στον κανόνα. Υποστηρίζεται ότι, αντίθετα, τόσο στις αναπτυσσόμενες όσο και στις αναπτυγμένες αγορές μακροί περίοδοι αποτελεσματικότητας εναλλάσσονται με σύντομες περιόδους μη-γραμμικότητας. Να σημειωθεί ότι η μη-γραμμικότητα στις χρονοσειρές των αγορών είναι αποτέλεσμα της πολυπλοκότητας των αγορών και βασίζεται στο γεγονός ότι περισσότεροι παράγοντες έχουν επίδραση σε αυτές.

Οι αιτίες που προκαλούν σύντομες περιόδους μη-γραμμικότητας στις αγορές είναι πολλές. Ανάμεσα τους μπορούν να αναφερθούν:

- η ασυμμετρία και η έλλειψη αξιόπιστης πληροφόρησης,
- οι επενδυτές που υπό συνθήκες πίεσης δεν αντιδρούν ορθολογικά,
- οι περίοδοι οικονομικής κρίσης που μπορούν να επιφέρουν αλλαγές στην δομή και στην ρύθμιση των αγορών αλλά και διαφορές στις τιμές προσφοράς – ζήτησης
- δομικές αλλαγές,

- τεχνολογική καινοτομία,
- η χαμηλή ρευστότητα, και συχνότητα των συναλλαγών κ.α.

Όσον αφορά τα χαρακτηριστικά της μικροδομής των αγορών (όπως η ταχύτητα που πραγματοποιούνται οι συναλλαγές, τα κόστη τους, η σχέση μεταξύ όγκου συναλλαγών των αξιογράφων και της τιμής τους, κ.α.) επηρεάζουν άμεσα την γραμμικότητα των χρονοσειρών των αγορών, αφού οδηγούν την αγορά σε μια πολυπλοκότητα που μπορεί να επιφέρει σύγχυση αλλά και απροθυμία στους επενδυτές.

Όπως ήδη αναφέρθηκε, ένας άλλος λόγος μη-γραμμικότητας των αγορών είναι το γεγονός ότι οι πληροφορίες δεν είναι καθολικά διαθέσιμες στο σύνολο των επενδυτών καθώς αυτές διατίθενται στην αγορά με ένα επιπλέον κόστος συλλογής τους. Συνέπεια των παραπάνω είναι ότι οι επενδυτές διαθέτουν μια μη-συμμετρική πληροφόρηση η οποία προκαλεί μια αστάθεια και διαφορετικές προσδοκίες στους επενδυτές ακόμα και στην περίπτωση που αυτοί αντιδρούν ορθολογικά. Όσον αφορά την πληροφορία, το γεγονός ότι αυτή δεν είναι καθολική αλλά ούτε και άμεση έχει σαν αποτέλεσμα την δημιουργία δύο ομάδων επενδυτών τους καλά πληροφορημένους και τους απληροφόρητους, με τους τελευταίους να αναμένουν αλλά και να αναπαράγουν τη συμπεριφορά των ενημερωμένων. Κατ' αυτόν τον τρόπο παραβιάζεται η συνθήκη της αποτελεσματικότητας της αγοράς που θέλει η πληροφορία να ενσωματώνεται και να αντικατοπτρίζεται στην τιμή των μετοχών και έχει σαν αποτέλεσμα την συμπεριφορά μη-γραμμικότητας της αγοράς.

Μια επιπλέον αιτία μη-γραμμικότητας των αγορών αναφέρεται η ορθολογική ανάδραση των επενδυτών στις κινήσεις των μετοχών. Αυτή η ορθολογική ανάδραση οδηγεί τους πληροφορημένους επενδυτές να προσπαθήσουν να οδηγήσουν την τιμή μιας μετοχής στην τιμή ισορροπίας σε περίπτωση που η τελευταία είναι υπερεκτιμημένη (ή υποτιμημένη) και γενικά δεν αντανακλά την υπάρχουσα πληροφορία. Από την άλλη οι μη-πληροφορημένοι επενδυτές θα προσπαθήσουν να κρατήσουν την τρέχουσα τιμή της μετοχής. Αυτό έχει σαν αποτέλεσμα η νέα τιμή της μετοχής να μην είναι ίδια με την προηγούμενη αλλά και να παρουσιάζει μια απόκλιση αφού δεν αντικατοπτρίζει πλήρως την πληροφορία.

Επίσης σε ακραίες καταστάσεις, όπως αυτές μιας οικονομικής κρίσης, οι επενδυτές τείνουν να μην συμπεριφέρονται ορθολογικά. Η ορθολογικότητα μπορεί να παρουσιαστεί μόνο όταν οι επενδυτές επιδεικνύουν μια μη ανεκτικότητα στον κίνδυνο αλλά και αμερόληπτες προβλέψεις και ακαριαίες αντιδράσεις στην πληροφορία και έχει σαν

αποτέλεσμα οι τιμές των μετοχών να ανταποκρίνονται γραμμικά σε κάθε καινούργια πληροφορία. Η μη ορθολογικότητα από την άλλη έχει σαν αποτέλεσμα σημαντικές αλλά και διαρκείς αποκλίσεις τιμών από την πραγματική αξία τους και κατ' επέκταση επιφέρει μη-γραμμικότητα στις αγορές.

1.5 Προσαρμοστική Αγορά

Σύμφωνα με όσα παρουσιάστηκαν προηγουμένως τελικά οι αγορές δεν είναι τόσο αποτελεσματικές όσο αρχικά θεωρήθηκε με χαρακτηριστικό παράδειγμα την πτώση κατά 23% του Αμερικάνικου Χρηματιστηρίου σε μια μόλις ημέρα τον Οκτώβριο του 1987 χωρίς να είναι εμφανές γιατί οι επενδυτές άλλαξαν τόσο γρήγορα την παραδοχή τους πάνω στην αξία των μετοχών κατά το συγκεκριμένο γεγονός. Παρόλα αυτά η έννοια της Υπόθεσης της Αποτελεσματικής Αγοράς παραμένει σημαντική αφού στην ουσία αποτελεί μία πλήρη θεωρία. Οι αγορές περνούν ανάμεσα από μεγάλες περιόδους αποτελεσματικότητας σε, πιο βραχυπρόθεσμες, περιόδους μη-γραμμικότητας και η ΥΑΑ μας βοηθά να αναγνωρίσουμε τους λόγους των ανωμαλιών και κατ' επέκταση τον τρόπο με τον οποίο θα τις επαναφέρουμε σε μια κατάσταση αποτελεσματικότητας.

Αυτά τα συμπεράσματα οδήγησαν στην εισαγωγή νέων εννοιών όπως αυτόν της *Υπόθεσης Προσαρμοστικής Αγοράς* (ΥΠΑ) από τον Andrew Lo στα (Lo, 2004) και (Lo, 2005). Η ΥΠΑ συνδυάζει την θεωρία της ΥΑΑ και τη Συμπεριφοριστική Χρηματοοικονομική και θέτει το ανθρώπινο ένστικτο στο επίκεντρο της συμπεριφοράς της αγοράς.

Όπως ο ίδιος ο εισηγητής αναφέρει σε συνέντευξη του στο (Lo, 2012) *“Η βασική αρχή της υπόθεσης της “προσαρμοστικής αγοράς” είναι ο συνδυασμός των εννοιών της εξέλιξης και της προσαρμογής που θεωρώ ότι αποτελούν ουσιαστικά την κινητήρια δύναμη των αγορών. Οι αγορές είναι ανθρώπινοι θεσμοί που διακρίνονται για τη ρευστότητα και τη ροπή τους προς τις μεταβολές. Όπως και οι υπόλοιποι ανθρώπινοι θεσμοί, τείνουν να προσαρμόζονται σε μια ποικιλία δυνάμεων. Μία σημαντική αρχή είναι ότι τα risk premium που όλοι θεωρούμε δεδομένα δεν αποτελούν παγκόσμια σταθερά, αλλά μεταβάλλονται στο πέρασμα του χρόνου σε σχέση με έναν συγκεκριμένο πληθυσμό επενδυτών. Οι επενδυτές από την πλευρά τους συμπεριφέρονται κατά τρόπο αθροιστικά προβλέψιμο, κάτι που υποδεικνύει ότι περιοδικά πρέπει να περιμένουμε περιόδους μεγάλης ανάπτυξης, αλλά και αναταραχής.*

*Βλέποντας αυτά υπό το πρίσμα της εξελικτικής βιολογίας αλλάζει η άποψή μας αναφορικά με τις επενδύσεις όσο και σε σχέση με τις κυβερνητικές παρεμβάσεις και ρυθμίσεις.....
“Έτσι η αποτελεσματικότητα δεν είναι μια κατάσταση που μπορεί να θεωρηθεί καθολική αλλά κυρίως ένα δυναμικό χαρακτηριστικό που μεταβάλλεται τόσο στον χρόνο αλλά και από αγορά σε αγορά.”*

Η θεωρία βασίζεται σε διάφορους τομείς της επιστήμης όπως η ψυχολογία, η βιολογία, κ.α. και υποστηρίζει ότι μέσω αυτών των διεπιστημονικών προσεγγίσεων μπορεί να γίνουν εφικτές πιο ακριβείς σε επίπεδο πρόγνωσης, θεωρίες για την εξέλιξη των αγορών.

Σημαντικότερα μεταξύ αυτών των δυνάμεων θεωρούνται τα risk premium τα οποία δεν αποτελούν μια παγκόσμια σταθερά, αλλά μεταβάλλονται στη χρονοσειρά των αγορών ανάλογα με τον αριθμό των επενδυτών.

Οι βασικές αρχές που διέπουν την Υπόθεση της Προσαρμοστικής Αγοράς είναι (Lo, 2005) :

- Οι επενδυτές δεν είναι ούτε πάντα ορθολογικοί ούτε μη-ορθολογικοί, αλλά βιολογικές οντότητες των οποίων τα χαρακτηριστικά και οι συμπεριφορές διαμορφώνονται από τις δυνάμεις της εξέλιξης.
- Όπως όλες οι οντότητες, παρουσιάζουν συμπεριφοριστικές προκαταλήψεις και κατ' επέκταση αντίστοιχες αποφάσεις, αλλά οι επενδυτές μπορούν να μάθουν από την εμπειρία του παρελθόντος και να αναθεωρήσουν τις αποφάσεις τους.
- Οι επενδυτές έχουν την ικανότητα για αφηρημένη σκέψη και πιο συγκεκριμένα: για μελλοντική ανάλυση, προβλέψεις για το μέλλον βάσει προηγούμενης εμπειρίας και προετοιμασία για αλλαγές στο περιβάλλον τους. Αυτή είναι εξέλιξη στην ταχύτητα της σκέψης.
- Η δυναμική της χρηματοπιστωτικής αγοράς δημιουργείται από τις αλληλεπιδράσεις των επενδυτών, καθώς αυτοί συμπεριφέρονται, μαθαίνουν και προσαρμόζονται στο κοινωνικό, πολιτιστικό, πολιτικό, οικονομικό και φυσικό περιβάλλον στο οποίο δρουν.
- Η επιβίωση είναι η πιο σημαντική δύναμη που οδηγεί στον ανταγωνισμό, στην καινοτομία και στην προσαρμογή.

Η έννοια της προσαρμοστικότητας βασίζεται στο γεγονός ότι σε αντίθεση με την ΥΠΑ οι επενδυτές μπορούν να κάνουν λάθη στις αποφάσεις τους και να μάθουν από αυτά και να προσαρμόζουν την συμπεριφορά τους σε μελλοντικές συναλλαγές.

Στην θεωρία ΥΠΑ οι μεταβαλλόμενες καταστάσεις της αγοράς είναι παράγοντες πρόβλεψης των τιμών παρότι οι τελευταίες αντανακλούν πλέον μια πολύπλοκη πληροφορία που προκύπτει από το συνδυασμό του πλήθους και του τύπου των επενδυτών.

Σύμφωνα με τον Lo, ο οποίος και εισήγαγε την έννοια της προσαρμοστικότητας, η ΥΠΑ έχει διάφορες πρακτικές επιπτώσεις με πιο σημαντική ότι η σχέση ρίσκου / κέρδους δεν είναι σταθερή και ότι η μεταβολή της εξαρτάται από το πλήθος και το είδος των συμμετεχόντων στην αγορά. Επιπλέον, η αποτελεσματικότητα της αγοράς είναι μεταβλητή και εξαρτάται από το πλήθος των ορθολογικών και μη-ορθολογικών επενδυτών και την αναλογία τους. Αυτή η μεταβλητότητα της αποτελεσματικότητας πρέπει να λαμβάνεται υπόψη από τους επενδυτές οι οποίοι πρέπει να έχουν και την κατάλληλη αντίδραση.

Καταλήγουμε λοιπόν, ότι η αγορά και η αποτελεσματικότητα της παρουσιάζουν μια άμεση εξάρτηση από τις παραμέτρους του συστήματος και ότι οι αγορές είναι προσαρμόσιμες και εναλλάσσονται κατά καιρούς μεταξύ αποτελεσματικότητας και μη-αποτελεσματικότητας.

1.6 Χρηματοοικονομικοί Κίνδυνοι

Σύμφωνα με τον Holton ο όρος *κίνδυνος* εκφράζει την αβεβαιότητα και μεταβλητότητα στις αγορές και πιο συγκεκριμένα την αβεβαιότητα πάνω στα αναμενόμενα χρηματοοικονομικά αποτελέσματα (Holton, 2014). Ο κίνδυνος γίνεται αντιληπτός μέσω της διακύμανσης των αναμενόμενων αποδόσεων και μετράται με το βαθμό απόκλισης της αναμενόμενης απόδοσης από την μέση αναμενόμενη απόδοση της επένδυσης (Φίλιππας, 2005). Μια γενικότερη κατηγοριοποίηση των κινδύνων είναι ανάμεσα στους *επιχειρηματικούς* και *μη επιχειρηματικούς* κινδύνους.

Ως *επιχειρηματικοί* κίνδυνοι (*Business risks*) ορίζονται εκείνοι οι κίνδυνοι επενδυτικών αποφάσεων που μια επιχείρηση αναλαμβάνει εις γνώση της προκειμένου να αποκτήσουν ένα ανταγωνιστικό πλεονέκτημα σε σχέση με τους ανταγωνιστές της. Το αποτέλεσμα του ανταγωνιστικού πλεονεκτήματος είναι η επίτευξη πρόσθετης αξίας για τους μετόχους της.

Αφορούν το περιβάλλον μέσα στο οποίο μια επιχείρηση δραστηριοποιείται και τις αποφάσεις που αυτή λαμβάνει. Κάθε επιχείρηση αντιμετωπίζει συνεχώς μακροοικονομικούς κινδύνους που οφείλονται σε νομισματικές πολιτικές και σε διακυμάνσεις στα έσοδα.

Οι *μη επιχειρηματικοί* κίνδυνοι (*Non Business risks*), από την άλλη, είναι κίνδυνοι που η επιχείρηση δεν έχει αναλάβει και τους οποίους δεν μπορεί να ελέγξει. Στους μη επιχειρηματικούς κινδύνους συμπεριλαμβάνονται οι *στρατηγικοί* και οι *χρηματοοικονομικοί* κίνδυνοι.

Οι *χρηματοοικονομικοί* κίνδυνοι (*financial risks*) αφορούν τις επιχειρήσεις οφείλονται στο ύψος και το είδος του δανεισμού και στην μη εξυπηρέτησή του (*default*) και σχετίζονται με τις δυνητικές απώλειες στις αγορές χρήματος και κεφαλαίου. Οι κίνδυνοι αυτού του είδους οφείλονται σε λάθος προβλέψεις εσόδων, σε διακυμάνσεις στα επιτόκια, και διακυμάνσεις στο συνάλλαγμα. Στους κινδύνους επίσης περιλαμβάνονται απώλειες από αγορές αξιογράφων (Finard, et al., 1996).

Ως *στρατηγικοί* κίνδυνοι (*strategic risks*) ορίζονται εκείνοι που είναι αποτέλεσμα εφαρμογής διαφόρων στρατηγικών αποφάσεων της διοίκησης. Οι μη επιτυχημένες στρατηγικές έχουν σαν αποτέλεσμα την αδυναμία προσαρμογής μιας επιχείρησης σε αλλαγές στο οικονομικό και πολιτικό περιβάλλον.

Όπως ήδη αναφέρθηκε οι χρηματοοικονομικοί κίνδυνοι οφείλονται στις απώλειες στις αγορές χρήματος και κεφαλαίου. Διαφορετικά μπορεί να οριστεί ως η πιθανότητα ένα μη αναμενόμενο γεγονός να επηρεάσει αρνητικά την χρηματοοικονομική επίδοση ενός οργανισμού με αντίστοιχη επίδραση στα κέρδη. Οι χρηματοοικονομικοί κίνδυνοι μπορούν να κατηγοριοποιηθούν ως (Mark, 1996):

- Κίνδυνοι της αγοράς
- Πιστωτικοί κίνδυνοι
- Κίνδυνοι ρευστότητας
- Επιτοκιακοί κίνδυνοι
- Λειτουργικοί κίνδυνοι
- Νομικοί κίνδυνοι

1.6.1. Κίνδυνοι αγοράς

Σε αυτή την κατηγορία συμπεριλαμβάνονται οι κίνδυνοι που οφείλονται σε πιθανές απώλειες λόγω μεταβολών των τιμών ή των δεικτών της αγοράς. Ως μεταβολές των τιμών ορίζονται οι μεταβολές των επιτοκίων, των μετοχών και των συναλλαγματικών ισοτιμιών.

Ένας πιο επίσημος ορισμός είναι αυτός που δόθηκε από την Επιτροπή της Βασιλείας¹⁰ το 2005 όπου ως κίνδυνος αγοράς θεωρείται η απώλεια θέσεων εντός και εκτός του ισολογισμού σαν αποτέλεσμα κινήσεων των τιμών της αγοράς. Η επιτροπή της Βασιλείας κατηγοριοποιεί τους κινδύνους αγοράς σε γενικούς κινδύνους (general market risk) και στους κινδύνους που μόνο κάποιο συγκεκριμένο στοιχείο μεταβάλλεται σε σχέση με την αγορά (specific risk).

Σημαντικό ρόλο στην διαμόρφωση του κινδύνου αγοράς έχουν:

- Ο συναλλαγματικός κίνδυνος, όπου η υποτίμηση ενός νομίσματος μειώνει την αξία των τίτλων που συναλλάσσονται στο συγκεκριμένο νόμισμα,
- Ο κίνδυνος μετοχών,
- Ο κίνδυνος επιτοκίων, όπου μια πιθανή αύξηση επιτοκίων οδηγεί στη μείωση των τιμών των ομολόγων,
- Ο κίνδυνος εμπορευμάτων.

Ένας επιπλέον διαχωρισμός του κινδύνου αγοράς είναι αυτός σε *συστηματικό* και *μη συστηματικό* κίνδυνο. Ως *συστηματικός επενδυτικός κίνδυνος* ορίζεται ο κίνδυνος που παράγεται λόγω μη αναμενόμενων εξελίξεων στο μακροοικονομικό περιβάλλον της επένδυσης. Ως *μη αναμενόμενες εξελίξεις* θεωρούνται αυτές που είναι δύσκολο να προβλεφθούν, καθώς αφορούν διακυμάνσεις στις τιμές των τίτλων στην αγορά λόγω ενός εξωτερικού παράγοντα. Αντίθετα ένας *μη συστηματικός κίνδυνος* μπορεί να προβλεφθεί και να ποσοτικοποιηθεί με την βοήθεια τεχνικών διαχείρισης κινδύνου (*risk management*).

¹⁰ Η Επιτροπή της Βασιλείας για την Εποπτεία του Τραπεζικού Συστήματος (Basel Committee on Banking Supervision) ιδρύθηκε το 1974 με στόχο τη διαμόρφωση προτύπων εποπτείας και κατευθυντηρίων οδηγιών για τη λειτουργία του τραπεζικού συστήματος. Η Επιτροπή, αρχικά εισήγαγε ένα σύστημα κεφαλαιακής μέτρησης με την ονομασία Basel Capital Accord. Το έτος 1998 εξέδωσε το πλαίσιο εποπτείας του διεθνούς τραπεζικού συστήματος με την ονομασία Βασιλεία I με κύριο σκοπό την αντιμετώπιση του πιστωτικού κινδύνου μέσω της θέσπισης ελάχιστων κεφαλαιακών απαιτήσεων. Μεταγενέστερα εισήγαγε ένα νέο πλαίσιο (Βασιλεία II), με σκοπό την πληρέστερη απεικόνιση των αναλαμβανόμενων κινδύνων από τα πιστωτικά ιδρύματα και στη σύνδεση των κεφαλαιακών απαιτήσεων με τους κινδύνους αυτούς. Τέλος, στο πλαίσιο Βασιλεία III, παρουσιάζονται κανονιστικά πρότυπα που αφορούν την κεφαλαιακή επάρκεια και ρευστότητα των τραπεζών.

1.6.2. Πιστωτικός κίνδυνος

Ως πιστωτικός κίνδυνος (*credit risk*) ορίζονται οι περιπτώσεις πιθανών απωλειών ως αποτέλεσμα της αδυναμίας των αντισυμβαλλόμενων μερών, να εκπληρώσουν υποχρεώσεις και στόχους, που προκύπτουν από κάποια οικονομική σύμβαση. Ο πιστωτικός κίνδυνος είναι άμεσα συνδεδεμένος με τις υπηρεσίες εμπορικής τραπεζικής και με την πιστοληπτική ικανότητα των αντισυμβαλλομένων του πιστωτικού ιδρύματος, αποτελώντας το σημαντικότερο, ποιοτικά και ποσοτικά χρηματοοικονομικής φύσης κίνδυνο.

Ο πιστωτικός κίνδυνος έχει διάφορες επιπτώσεις στην αγοραία αξία μιας χρηματοοικονομικής θέσης, κυρίως λόγω του κόστους αναπλήρωσης των χρηματοροών λόγω αφερεγγυότητας του άλλου μέρους, αλλά και λόγω συγκέντρωσης μεγάλων ανοιγμάτων σε συγκεκριμένες επιχειρήσεις. Ο πιστωτικός κίνδυνος μπορεί να αντιμετωπιστεί μέσω της δημιουργία ανώτατων και κατώτατων ορίων σε παρούσες και μελλοντικές θέσεις.

Ο πιστωτικός κίνδυνος μπορεί να οφείλεται σε ένα από τους τέσσερεις επιμέρους κινδύνους:

- Ο κίνδυνος *αφερεγγυότητας ή αθέτησης (default risk)*: λαμβάνει χώρα στην περίπτωση την οποία η εξόφληση μιας υποχρέωσης δεν πραγματοποιείται στο συμφωνημένο χρονικό διάστημα,
- Ο κίνδυνος *έκθεσης ή ανοίγματος (exposure risk)*: έχει να κάνει με το συνολικό ποσό που βρίσκεται σε πιστωτικό κίνδυνο. Λαμβάνει χώρα σε περιπτώσεις όπου δεν υπάρχει προκαθορισμένος τρόπος εξόφλησης της υποχρέωσης και αφορά σε έκθεση που μπορεί να εμπλέκει δάνεια, παράγωγα, εγγυητικές επιστολές αλλά και ομολογίες,
- Ο κίνδυνος *ανάκτησης (recovery risk)*: έχει να κάνει με το συνολικό ποσό που βρίσκεται εκτεθειμένο σε κίνδυνο σε περίπτωση πτώχευσης του συμβαλλόμενου που έχει αναλάβει μια υποχρέωση και την αξία των εξασφαλίσεων της τράπεζας. Ουσιαστικά, ο συγκεκριμένος κίνδυνος έχει να κάνει με το ποσοστό του ποσού που οφείλεται κατά τη στιγμή της αθέτησης και που ο δανειστής είναι σε θέση να ανακτήσει (Recovery Rate),

- Ο κίνδυνος περιθωρίου (*credit spread risk*): έχει να κάνει με τις αλλαγές που προκύπτουν λόγω των πιστωτικών περιθωρίων (*spreads*), με την τεχνική υπολογισμού του κινδύνου των επιτοκίων και με τη μέθοδο που βασίζεται στη λήξη των χρεογράφων.

1.6.3. Κίνδυνος ρευστότητας

Ο κίνδυνος ρευστότητας (*liquidity risk*) είναι αποτέλεσμα του πιστωτικού κινδύνου καθώς μια αύξηση του τελευταίου μοιραία θα οδηγήσει σε αύξηση και του κινδύνου ρευστότητας. Ο κίνδυνος ρευστότητας αφορά στην αδυναμία εύρεσης επαρκών ρευστών διαθέσιμων για την κάλυψη των ληξιπρόθεσμων πληρωμών του οφειλομένου. Με άλλα λόγια, ο κίνδυνος ρευστότητας είναι η πιθανότητα αποτυχίας του οφειλομένου να ανταπεξέλθει στις υποχρεώσεις του απέναντι στους καταθέτες και τους δανειστές του. Ο κίνδυνος ρευστότητας επομένως συνδέεται άμεσα με οικονομικούς παράγοντες και για το λόγο αυτό αποτελεί ένα είδος συστημικού κινδύνου.

Εφόσον ο κίνδυνος ρευστότητας είναι αποτέλεσμα οικονομικών παραγόντων είναι και άμεσα συνδεδεμένος με τον ορίζοντα διακράτησης χαρτοφυλακίου (*portfolio holding period*). Αυτό συμβαίνει γιατί όταν οι συνθήκες στην αγορά δεν επιτρέπουν τη ρευστοποίηση μιας επένδυσης χαρτοφυλακίου, ένας χρηματοοικονομικός οργανισμός θα πρέπει να περιμένει οι τιμές να ανακάμψουν σε ικανοποιητικά επίπεδα. Η μη ικανότητα ρευστοποίησης έτσι ώστε ο οργανισμός να μπορέσει να ανταπεξέλθει στις υποχρεώσεις του μπορεί να επιφέρει ακόμη και την κατάρρευση.

Ο Jorion στο (Jorion, 2003) διακρίνει τον κίνδυνο ρευστότητας σε δύο είδη κινδύνου:

- τον κίνδυνο ρευστότητας επένδυσης ή αγοράς (*Asset Liquidity Risk or Market Liquidity Risk*). Στην ουσία πρόκειται για τον κίνδυνο μια συναλλαγή να λάβει χώρα όταν οι τιμές αγοράς δεν είναι ευνοϊκές. Ο κίνδυνος ρευστότητας αγοράς αντιμετωπίζεται θέτοντας χαμηλά όρια διαπραγμάτευσης σε αγορές χαμηλής ρευστότητας ή επιδιώκοντας την διαφοροποίηση (*diversification*) των θέσεων.
- τον κίνδυνο ρευστότητας χρηματοδότησης ή χρηματικών ροών (*Funding Liquidity Risk* ή *Cash-Flow Risk*). Είναι η αδυναμία ενός χρηματοοικονομικού ιδρύματος να εκπληρώσει τακτικές ή έκτακτες πληρωμές. Αυτή η αδυναμία εξαναγκάζει σε

πρόωρη ρευστοποίηση τίτλων. Λαμβάνει χώρα συνήθως σε περιπτώσεις χαρτοφυλακίων με μόχλευση (*leverage*) και υπόκεινται σε απαιτήσεις πρόσθετου περιθωρίου από τους δανειστές. Ο κίνδυνος αντιμετωπίζεται μέσω κατάλληλου προγραμματισμού εσόδων/ εξόδων προσπαθώντας να εξασφαλιστεί η επάρκεια ρευστών διαθεσίμων.

1.6.4. Επιτοκιακοί κίνδυνοι

Ο επιτοκιακός κίνδυνος (*interest rate risk*) έχει να κάνει με την δυσμενή μεταβολή των επιτοκίων, μεταβολή η οποία μπορεί να πλήξει μια επιτοκιακή θέση, όπως καταθέσεις, ομόλογα ή δάνεια (π.χ. μετά από λήψη δανείου με κυμαινόμενο επιτόκιο η αυξητική διακύμανση των επιτοκίων).

Ως γνωστό η διακύμανση των επιτοκίων πλήττει άμεσα την τιμή διαπραγμάτευσης ορισμένων χρηματοπιστωτικών μέσων (όπως ομόλογα ή παράγωγα). Αυτό γιατί, μια μη αναμενόμενη μεταβολή των επιτοκίων επηρεάζει με άμεσο τρόπο την αξία μιας επενδυτικής θέσης και κατά συνέπεια και το εισόδημα από την επενδυτική θέση αυτή.

Η μεταβολή της αξίας μιας επενδυτικής σχέσης, από την άλλη, επηρεάζει άμεσα την αξία ενός χαρτοφυλακίου ή την αξία ενός πιστωτικού ιδρύματος.

1.6.5. Λειτουργικοί Κίνδυνοι

Τα τρία προηγούμενα είδη κινδύνων μπορούν να θεωρηθούν ως ποσοτικά, από την άποψη ότι μπορούν να μετρηθούν μέσω τεχνικών και μοντέλων μέτρησης. Ο λειτουργικός κίνδυνος από την άλλη είναι είδος κινδύνου που δεν μπορεί να υπολογιστεί και να κατηγοριοποιηθεί και αφορά τις απώλειες τόσο ως αποτέλεσμα αποτυχημένων εσωτερικών διαδικασιών των οργανισμών, είτε ως αποτέλεσμα δυσχερειών της εταιρικής διοίκησης, αλλά και άλλων εξωγενών παραγόντων.

Με άλλα λόγια ο *λειτουργικός κίνδυνος* (*operational risk*), ορίζεται ως ο κίνδυνος άμεσης απώλειας που επέρχεται μέσω διαφόρων παραγόντων όπως η ανεπάρκεια εσωτερικών

διαδικασιών, το ανθρώπινο δυναμικό και τα συστήματα ή εξωγενών παραγόντων (Angelopoulos, et al., 2001).

Στα σύγχρονα πιστωτικά ιδρύματα, μεγάλο μέρος των ζημιών οφείλονται, συνήθως, σε κάποια πηγή λειτουργικού κινδύνου και στην μη σωστή διαχείριση της.

1.6.6. Νομικοί κίνδυνοι

Ο νομικός κίνδυνος (*legal risk*), αφορά στην αδυναμία διεξαγωγής μιας συναλλαγής, μέσω κολλημάτων που εισάγονται από τον ισχύον νομοθετικό πλαίσιο. Αυτό το είδος κινδύνου και έχει άμεση σχέση με τον πιστωτικό κίνδυνο. Συχνότερα ο νομικός κίνδυνος αφορά διεθνείς συναλλαγές και αυτό για το λόγο ότι σε κάθε χώρα το νομικό πλαίσιο, που ρυθμίζει τις τραπεζικές δραστηριότητες, μπορεί να διαφέρει. Ο νομικός κίνδυνος, όπως όλα τα είδη κινδύνου, μπορεί να οδηγήσει σε υποτίμηση των στοιχείων του ενεργητικού μιας εταιρείας.

Ο νομικός κίνδυνος εμπλέκει παράνομες ενέργειες, όπως αυτή της εσωτερικής πληροφόρησης (*inside trading*) και της χειραγώγησης τιμών (*price manipulation*), και για αυτό, για την ελαχιστοποίηση και αντιμετώπιση τους, εποπτεύονται και ελέγχονται από την νομική υπηρεσία της κάθε εταιρίας μαζί με την Ανώτατη Διοίκηση.

2. ΤΕΧΝΙΚΕΣ ΠΡΟΒΛΕΨΕΩΝ

Λόγω της φύσης των τιμών των χρηματιστηριακών αξιών, της μεταβλητότητας αυτών στο χρόνο αλλά και της δυσκολίας πρόβλεψης τους γίνεται χρήση διάφορων τεχνικών προβλέψεων¹¹. Οι προβλέψεις αυτές είναι απαραίτητες για ένα μεγάλο αριθμό αποφάσεων σχεδιασμού και προγραμματισμού.

Η σχετικά νέα επιστήμη των προβλέψεων αποδεικνύεται χρήσιμη και ίσως αναγκαία λόγω της ανάγκης του ανθρώπου για μοντελοποίηση μελλοντικών γεγονότων έτσι ώστε να γνωρίζει εκ των πρότερων μια πιθανή έκβαση των δραστηριοτήτων του με απώτερο σκοπό, φυσικά, την εκμετάλλευση επιτυχιών αλλά και την διαχείριση των αποτυχιών.

Οι μέθοδοι προβλέψεων που προτάθηκαν ανά καιρούς παρουσιάζουν βασικές διαφορές αφού κάποιες από αυτές βασίζονται μόνο σε θεωρητικό υπόβαθρο, ενώ άλλες, έχουν ανάγκη και εμπλέκουν τεχνολογία με μεγάλη υπολογιστική ισχύ.

Ειδικότερα, οι τεχνικές προβλέψεων κατηγοριοποιούνται αλλά και επιλέγονται βάση διάφορων παραμέτρων οι οποίοι είναι:

- η μορφή της πρόβλεψης.
- ο χρονικός ορίζοντας της πρόβλεψης (βραχυπρόθεσμος, μεσοπρόθεσμος ή μακροπρόθεσμος),
- η αξιοπιστία της μεθόδου και η ακρίβεια πρόβλεψης,
- τα ιστορικά δεδομένα και το πρότυπο συμπεριφοράς τους,
- η φιλικότητα προς τον χρήστη της επιλεγμένης μεθόδου πρόβλεψης και τέλος,
- το κόστος παραγωγής των προβλέψεων.

Όσον αφορά τον χρονικό ορίζοντα των προβλέψεων οι μέθοδοι διακρίνονται σε τρεις κατηγορίες:

- Βραχυπρόθεσμες,
- Μεσοπρόθεσμες, με χρονικό ορίζοντα 6-12 μηνών,
- Μακροπρόθεσμες, με χρονικό ορίζοντα ετών.

¹¹ Πρόβλεψη είναι η διαδικασία εκτίμησης μελλοντικών καταστάσεων βασισμένη, συνήθως, σε ιστορικά στοιχεία και χρησιμοποιείται όταν είναι επιτακτικές αποφάσεις που αφορούν “μελλοντικούς σχεδιασμούς”

Από την άλλη, τα βασικές κατηγορίες τεχνικών προβλέψεων, που θα αναλυθούν παρακάτω είναι τρεις:

- Ποσοτικές μέθοδοι,
- Ποιοτικές μέθοδοι και
- Τεχνολογικές μέθοδοι.

2.1 Οι τρεις τεχνικές προβλέψεων

2.1.1 Ποσοτικές Μέθοδοι

Η ονομασία των ποσοτικών μεθόδων (στη βιβλιογραφία αναφέρονται και ως αντικειμενικοί) οφείλεται στο γεγονός ότι βασικός σκοπός τους είναι η ποσοτικοποίηση σε αριθμούς ενός μεγάλου όγκου πληροφοριών. Οι ποσοτικές και αντικειμενικές μέθοδοι βασίζονται σε κάποιο μαθηματικό ή στατιστικό υπόδειγμα και σε ποσοτικά δεδομένα (model based forecasts). Τα υποδείγματα αυτά χωρίζονται σε δύο μεγάλες υποκατηγορίες, **τα αιτιατά (causal) και τα μη αιτιατά (non-causal)**.

Ο όγκος των δεδομένων θα πρέπει να έχει τέτοιο μέγεθος ώστε να θεωρηθεί ότι το πρότυπο συμπεριφοράς που θα παραχθεί θα παραμείνει αμετάβλητο και στο μέλλον. Οι ποσοτικές μέθοδοι χωρίζονται σε δύο επιμέρους κατηγορίες: το μοντέλο χρονοσειρών (time series model) και το αιτιοκρατικό ή επεξηγηματικό μοντέλο (causal relationship or explanatory model). Ειδικότερα:

Το *μοντέλο Χρονοσειρών*, το οποίο θα παρουσιαστεί αναλυτικότερα σε επόμενο υποκεφάλαιο, αποτελεί μέρος της στατιστικής επιστήμης με βασική χρήση για την πρόβλεψη ενός διαχρονικού μεγέθους με βάση το χρονικό δείγμα των τιμών του μεγέθους. Η μελέτη των χρονικών σειρών έχουν σαν στόχο την “αναγνώριση” ενός προτύπου συμπεριφοράς (ή επανάληψης) της χρονοσειράς υπό μελέτη έτσι ώστε να υπολογιστεί και η μελλοντική μεταβολή των τιμών¹². Στο μοντέλο χρονοσειρών συμπεριλαμβάνονται σαν

¹² Η πρόβλεψη πραγματοποιείται θεωρώντας ότι η μελλοντική μεταβολή της τιμής του μεγέθους, θα ακολουθήσει το ίδιο πρότυπο μεταβολής με τα παρελθοντικά στοιχεία. Με άλλα λόγια ότι η ιστορία επαναλαμβάνεται.

σύνθετοι μέθοδοι πρόβλεψης μεταξύ άλλων ο απλός κινητός μέσος, η εκθετική εξομάλυνση και η προβολή τάσης.

Το *αιτιοκρατικό* ή *επεξηγηματικό μοντέλο* όπου βασικός στόχος είναι η ανίχνευση μιας συσχέτισης ανάμεσα στο μέγεθος υπό εξέταση και σε άλλα μεγέθη που συνυπάρχουν στο ίδιο περιβάλλον έτσι ώστε να προβλέπεται η μελλοντική τιμή κάποιου μεγέθους, για διαφορετικούς συνδυασμούς των μεταβλητών εισόδου. Με την συγκεκριμένη μέθοδο τα διαφορετικά σενάρια που παράγονται από την μελέτη διαφορετικών εισόδων βοηθά στην υιοθέτηση μιας βέλτιστης επιλογής. Οι μέθοδοι παλινδρόμησης (απλή γραμμική παλινδρόμηση, πολλαπλή γραμμική παλινδρόμηση) αλλά και οι οικονομετρικές μέθοδοι ανήκουν στην μέθοδο αυτή. Βασικός περιορισμός της αιτιοκρατικής τεχνικής πρόβλεψης είναι η ανάγκη για ένα μεγάλο όγκο δεδομένων τόσο για το μέγεθος υπό μελέτη (μεταβλητή πρόβλεψης) όσο και για κάθε άλλη μεταβλητή που την επηρεάζει (ανεξάρτητες μεταβλητές).

2.1.2 Ποιοτικές Μέθοδοι

Οι ποιοτικές μέθοδοι (ή μέθοδοι κρίσεων) περιλαμβάνουν διάφορες υποκειμενικές μεθόδους πρόβλεψης οι οποίες είναι περισσότερο εμπειρικές και βασίζονται στην κρίση εμπειρών επιστημονικών αναλυτών. Λόγω του γεγονότος ότι οι τεχνικές πρόβλεψης θα πρέπει να είναι επιτυχής με μικρή ανοχή σε σφάλματα, οι ποιοτικές μέθοδοι χρησιμοποιούνται σε μεσοπρόθεσμες προβλέψεις σε συνδυασμό όμως με μια ποσοτική μέθοδο.

Η χρήση αυτών των μεθόδων γίνεται επιτακτική στην περίπτωση αδυναμίας συλλογής ιστορικών δεδομένων ή όταν αυτά τα τελευταία, δεν αρκούν για την πρόβλεψη μελλοντικών συνθηκών, οπότε οι ποσοτικές μέθοδοι δεν μπορούν να χρησιμοποιηθούν.

Η πιο ευρέως, ίσως, χρησιμοποιημένη ποιοτική μέθοδος προβλέψεων είναι η προσέγγιση Delphi. Στη συγκεκριμένη τεχνική, μια ομάδα επιστημόνων - άγνωστοι και σε απόσταση μεταξύ τους - απαντούν σε μια σειρά ερωτηματολογίων όπου οι απαντήσεις του καθενός χρησιμοποιούνται για τη σύσταση του επόμενου. Το τελευταίο ερωτηματολόγιο περιέχει πληροφορίες και απόψεις ολόκληρης της ομάδας. Η διαδικασία συνεχίζεται ζητώντας από κάθε επιστήμονα να ξανασκεφτεί και ίσως να αναθεωρήσει τις απαντήσεις του έχοντας

πλέον υπόψη τις απαντήσεις των υπόλοιπων επιστημόνων της ομάδας. Τα ερωτηματολόγια συνεχίζουν να συντάσσονται και να διανέμονται έως ότου ο υπεύθυνος της έρευνας να αποφασίσει ότι έχει επιτευχθεί ένας σημαντικός βαθμός ομοφωνίας μεταξύ των μελών της ομάδας. Στόχος της συγκεκριμένης μεθόδου δεν είναι η επίτευξη μιας μοναδικής απάντησης αλλά η οριοθέτηση των απόψεων μέσα σε ένα σχετικά μικρό πλάτος.

2.1.3 Τεχνολογικές Μέθοδοι

Τόσο η σημαντικότητα αλλά και η πολυπλοκότητα των προβλέψεων οδήγησαν σε μια τεχνολογική ανάπτυξη διάφορων συστημάτων πρόβλεψης μεταξύ των οποίων τα συστήματα SAP και τα Mangistics (Jain, 2003).

Η χρήση τους επικεντρώνονται στις μακροπρόθεσμες προβλέψεις και ειδικότερα σε εκείνες τις περιπτώσεις στις οποίες τα δεδομένα δεν ενδείκνυνται στην παραγωγή πρότυπων ή σχέσεων με το παρελθόν.

Η ταξινόμηση των συγκεκριμένων μεθόδων προβλέψεων γίνεται βάση του χρόνου των δεδομένων και εμπίπτουν σε δύο γενικές κατηγορίες:

- *Διερευνητικές*: τεχνικές οι οποίες βασίζονται σε μια προβολή των δεδομένων του παρελθόντος στο μέλλον κάνοντας χρήση δεδομένων και συνθηκών του παρόντος (S- καμπύλες, Delphi, αναλογίες και μορφολογική έρευνα)
- *Δεοντολογικές*: τεχνικές οι οποίες προβάλουν μια επιθυμητή (ή πιθανή) κατάσταση γεγονότων στο μέλλον. Στην συγκεκριμένη τεχνική ορίζεται μια σειρά βημάτων τα οποία (πιθανώς) θα οδηγήσουν στην επιθυμητή κατάσταση και υπολογίζεται η πιθανότητα πραγματοποίησής τους (δέντρα συσχετίσεων, μήτρες αποφάσεων, ανάλυση συστημάτων).

2.2 Εισαγωγή και βασικές έννοιες Χρονοσειρών και Προβλέψεων

Βασικός σκοπός της στοχαστικής ανάλυσης των χρονολογικών σειρών είναι η μελέτη της στοχαστικής διαδικασίας διαμόρφωσης των δεδομένων (data generating process) και η

διενέργεια προβλέψεων, δηλ βασικός στόχος των χρονοσειρών είναι η μελέτη διαφόρων μεγεθών η τιμή των οποίων δεν παραμένει σταθερή αλλά είναι δυναμική και μεταβάλλεται χρονικά. Πιο συγκεκριμένα, ο τομέας της στατιστικής που ασχολείται με την μελέτη των χρονοσειρών αναλύει και μελετά πως ένα μέγεθος, οι τιμές του οποίου μεταβάλλονται στον χρόνο, επηρεάζουν τις μελλοντικές τιμές της ίδιας μεταβλητής. Διαφορετικά, οι χρονοσειρές αναλύουν τη δυναμική εξέλιξη μιας διαδικασίας ή συστήματος που παράγει το μέγεθος που παρατηρούμε.

Ως προς την εξέταση οικονομικών φαινομένων γίνεται χρήση διάφορων τύπων δεδομένων μεταξύ των οποίων:

- *Διαστρωματικά Δεδομένα* (cross-section data): ο συγκεκριμένος τύπος δεδομένων έχει να κάνει με μετρήσεις μεταβλητών για διαφορετικές οικονομικές μονάδες (π.χ. άτομα, χώρες, επιχειρήσεις) σε κάποια χρονική περίοδο.
- *Χρονολογικές Σειρές* (time series): ο συγκεκριμένος τύπος δεδομένων χρησιμοποιείται για την μελέτη εξέλιξης και πρόβλεψης στο χρόνο συγκεκριμένων μεταβλητών.
- *Δεδομένα Panel* (panel data): τύπος ο οποίος χρησιμοποιείται στην μελέτη της χρονικής εξέλιξης μεταβλητών που αφορούν διαφορετικές οικονομικές μονάδες.

Γενικότερα τα δεδομένα που λαμβάνονται υπόψη είναι μία σειρά από παρατηρήσεις που συλλέγονται σε τακτικά χρονικά διαστήματα και μπορούν να αφορούν στα οικονομικά, χρηματοοικονομικά, την ιατρική και άλλους τομείς της επιστήμης.

Το χρονικό βήμα παρατήρησης και καταγραφής των δεδομένων είναι σταθερό και ονομάζεται **χρόνος δειγματοληψίας** (sampling time). Από την άλλη, το σύνολο τέτοιων παρατηρήσεων ονομάζεται **χρονική σειρά** ή **χρονοσειρά** (time series). Όσον αφορά την δειγματοληψία σε τιμές του χρηματιστηριακού τομέα σε διαδοχικές ημέρες ο χρόνος δειγματοληψίας δεν είναι σταθερός εφόσον δεν περιλαμβάνει τις χρονικές στιγμές που τα χρηματιστήρια δεν λειτουργούν, όπως σαββατοκύριακα και αργίες και ως εκ τούτου ο συγκεκριμένος τύπος χρονοσειράς άπτεται διαφορετικής επεξεργασίας.

Χρονοσειρά (ή time series) είναι ένα σύνολο παρατηρήσεων y_1, y_2, \dots, y_t όπου ο δείκτης παριστάνει ισαπέχοντα χρονικά σημεία ή διαστήματα (έτος, μήνας, ημέρα, εβδομάδα, ώρα, κ.α). Οι παρατηρήσεις y_1, y_2, \dots, y_t είναι συγκεκριμένες τιμές τυχαίων μεταβλητών Y_1, Y_2, \dots, Y_T και αποτελούν υποσύνολο μιας άπειρης ακολουθίας τυχαίων μεταβλητών. Έτσι,

στατιστικά, οι παρατηρήσεις λαμβάνουν την έννοια του δείγματος ενώ οι τυχαίες μεταβλητές την έννοια του πληθυσμού.

Βασικός στόχος του συγκεκριμένου κλάδου της στατιστικής είναι η μελέτη της χρονικής αλληλεξάρτησης μεταξύ των παρατηρήσεων. Η μελέτη περιλαμβάνει την ανάλυση των ιδιοτήτων της σειράς με σκοπό τον προσδιορισμό των χαρακτηριστικών που διέπουν τη συμπεριφορά της. Η μελέτη περιλαμβάνει επίσης τον προσδιορισμό υποδειγμάτων χρονολογικών σειρών (*time series models*) με σκοπό τη δημιουργία προβλέψεων και επομένως τη μείωση της αβεβαιότητας και τη καλύτερη εκτίμηση γεγονότων. Στόχος σε αυτή την περίπτωση είναι η πρόβλεψη του μέσου όρου της ζήτησης σε κάποια μελλοντική περίοδο. Για το σκοπό αυτό μπορούν να χρησιμοποιηθούν διάφορες μέθοδοι και τεχνικές όπως:

- *Απλός Μέσος Όρος (Simple Averages)*
- *Κινητός Μέσος Όρος (Simple Moving Averages)*
- *Σταθμικός Μέσος Όρος (Weighted Moving Averages)*
- *Εκθετική Μέθοδος (Exponential Smoothing)*
- *Προσαρμοσμένη Εκθετική Μέθοδος (Adjusted Exponential Smoothing)*
- *Εποχική Μέθοδος (Trend & Seasonal Methods)*

Στην μελέτη υποδειγμάτων χρονολογικών σειρών γίνεται χρήση κυρίως δύο βασικών κατηγοριών μεθόδων που διαφέρουν μεταξύ τους ως προς την πολυπλοκότητα, το κόστος υπολογισμού, την ταχύτητα πρόβλεψης και την διαθεσιμότητα των απαραίτητων δεδομένων.

2.2.1 Προσθετικό Υπόδειγμα και συνιστώσες

Η κλασική μέθοδος ανάλυσης μιας χρονολογικής σειράς ακολουθεί το διαχωρισμό της σειράς στις επιμέρους συνιστώσες της που είναι:

Μακρογρόνια Τάση (long-run trend - T_t)

Ως τάση ορίζεται το μαθηματικό μέγεθος που αντιπροσωπεύει τη μια μακροπρόθεσμη μεταβολή του μέσου επιπέδου τιμών της χρονοσειράς. Η τάση δίνει μια γενική εικόνα της χρονοσειράς, για το αν η πορεία του δείγματος είναι ανοδική, πτωτική ή σταθερή και

παριστάνεται γραφικά μέσω μιας ευθείας γραμμής ή μιας εκθετικής καμπύλης¹³. Σημαντικός παράγοντας στην σωστή εκτίμηση της τάσης είναι η συλλογή ενός ικανού αριθμού παρατηρήσεων σε ένα αντιπροσωπευτικό μήκος περιόδου.

Ειδικότερα τα υποδείγματα μέτρησης της τάσης μίας σειράς έχουν ως εξαρτημένη μεταβλητή τις τιμές των παρατηρήσεων της χρονολογικής σειράς και ως ανεξάρτητη μεταβλητή την χρονική περίοδο λήψης των παρατηρήσεων.

Οι πιο συνήθεις εξειδικεύσεις είναι οι ακόλουθες:

- *Γραμμική τάση*: υπολογίζεται μέσω του τύπου γραμμικής παλινδρόμησης

$$Y_t = \beta_0 + \beta_1 t + e_t$$

όπου e_t είναι ο όρος που αντικατοπτρίζει τις αποκλίσεις (τα λάθη) του ντετερμινιστικού μοντέλου.

Η γραμμική εξίσωση της τάσης είναι η απλούστερη και εκφράζεται με απόλυτους αριθμούς $Y_t =$ τάση της χρονολογικής σειράς για την περίοδο $t = (1, 2, \dots, N)$ και β_1 μετράει τη μέση μεταβολή της χρονοσειράς ανά μονάδα χρόνου.

Στην εικόνα 2.1 που ακολουθεί παρουσιάζεται γραφικά η περίπτωση τάσης του δείκτη S&P¹⁴:



Εικόνα 2.1: Γραφική απεικόνιση γραμμικής ανοδικής τάσης του S&P για τα έτη 1995-2000 (Πηγή StockCharts.com)

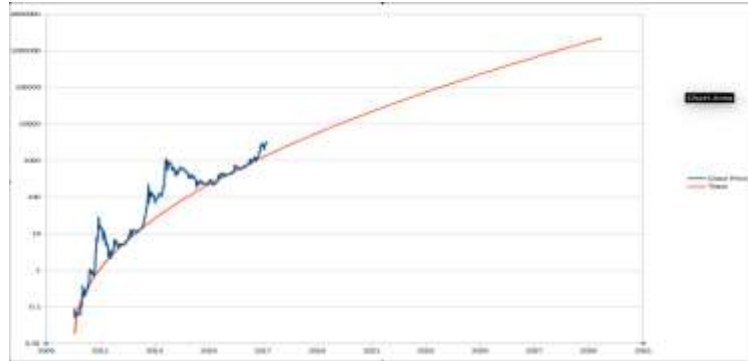
¹³ Αυτό δεν σημαίνει ότι η τάση δεν μπορεί να αναπαρασταθεί από κάποιο άλλο είδος καμπύλης

¹⁴ Ο S&P είναι ένας αμερικανικός δείκτης της χρηματιστηριακής αγοράς και βασίζεται στις κεφαλαιοποιήσεις της αγοράς των 500 μεγάλων εταιρειών που είναι εισηγμένες στο NYSE ή στο NASDAQ.

- *Λογαριθμική Τάση*: υπολογίζεται μέσω του τύπου

$$Y_t = \ln \beta_0 + \ln \beta_1 t + e_t$$

Στην εικόνα 2.2 που ακολουθεί παρουσιάζεται η λογαριθμική τάση και πρόβλεψη του Bitcoin για τα έτη 2007-2030



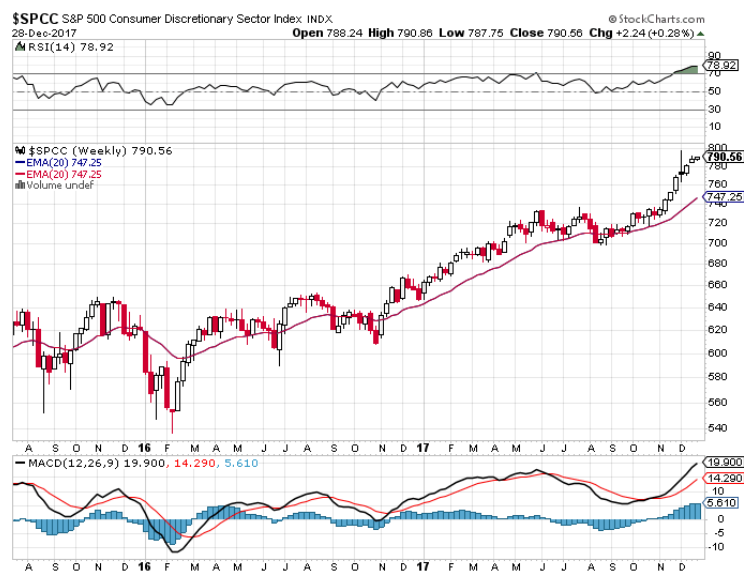
Εικόνα 2.2: Λογαριθμική τάση και πρόβλεψη του Bitcoin για τα έτη 2007-2030 (Πηγή StockCharts.com)

- *Εκθετική Τάση*

Η εκθετική τάση ενός τίτλου υπολογίζεται μέσω του τύπου

$$Y_t = \beta_0 * e^{\beta_1 t} * e_t$$

Ακολουθεί εικόνα εκθετικής τάσης και πρόβλεψης για την μετοχή της Intel Corp. του Nasdaq:

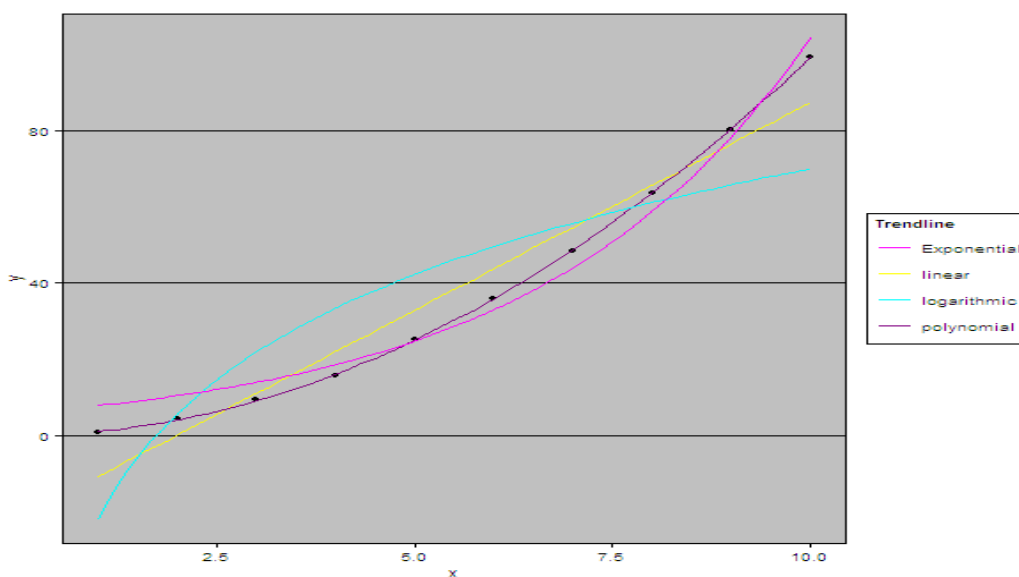


Εικόνα 2.3: Εκθετική τάση και πρόβλεψη της τιμής του δείκτη S&P 500 (Πηγή StockCharts.com)

- Πολυωνυμική Τάση που υπολογίζεται μέσω του πολυωνύμου:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + e_t$$

Το διάγραμμα της εικόνας 2.4 παρουσιάζει συνοπτικά τις καμπύλες των διαφόρων τάσεων που παρουσιάστηκαν.



Εικόνα 2.4: Συνοπτικό διάγραμμα τάσεων

Κυκλική συνιστώσα (CYCLICAL COMPONENT - C_t)

Η κυκλική συνιστώσα μιας χρονολογικής σειράς είναι μέγεθος μέτρησης αποκλίσεων των τιμών γύρω από τη μακροχρόνια τάση της και η διάρκεια της είναι μεγαλύτερη του έτους. Ο εντοπισμός της και η μέτρησή της πραγματοποιείται με τη μέθοδο του *ποσοστού της τάσης* (percentage of trend) ως εξής:

- Αρχικά υπολογίζεται το υπόδειγμα της χρονολογικής σειράς στη γραμμική του μορφή εκτιμάται με την μέθοδο ελαχίστων τετραγώνων (OLS).
- Χρησιμοποιώντας τη γραμμή τάσης, υπολογίζονται για κάθε χρονική περίοδο οι αντίστοιχες τιμές της.
- Υπολογίζεται ο λόγος των τιμών της χρονολογικής σειράς με τις αντίστοιχες τιμές της τάσης. Οι τυχόν αποκλίσεις του αποτελέσματος γύρω από τη μονάδα που

μπορεί και να παρουσιαστούν οφείλονται στις εποχιακές και τυχαίες επιδράσεις στη χρονολογική σειρά.

Εποχιακή Συνιστώσα και Εποχιακοί Δείκτες (SEASONAL COMPONENT SEASONAL INDICES - S_t)

Η εποχιακή συνιστώσα, με βασική εφαρμογή στις εμπορικές και οικονομικές χρονολογικές σειρές, έχει κοινά χαρακτηριστικά με την κυκλική συνιστώσα με βασική διαφορά στην διάρκεια (μπορεί πλέον να είναι μιας εβδομάδας, ενός μήνα ή και ενός τριμήνου). Η εποχιακή συνιστώσα μετράται με τους εποχιακούς δείκτες (seasonal indices) της χρονολογικής σειράς οι οποίοι κατασκευάζονται ακολουθώντας τα παρακάτω βήματα:

- Αρχικά πραγματοποιείται μια εξομάλυνση της χρονοσειράς μέσω του υπολογισμού των κινητών μέσων (moving averages) απομονώνοντας τις εποχιακές και τυχαίες επιδράσεις από τη χρονολογική σειρά.
- Στην συνέχεια υπολογίζεται ο λόγος των τιμών της χρονολογικής σειράς και των κινητών μέσων. Κατ' αυτό τον τρόπο ο κινητός μέσος μας δίνει την κυκλική συνιστώσα και τη συνιστώσα της τάσης (trend-cyclical component).
- Τέλος, για κάθε χρονική περίοδο υπολογίζονται οι αριθμητικοί μέσοι των λόγων έτσι ώστε να απομακρυνθεί η χρονολογική σειρά από τυχαίες αποκλίσεις.

Τυχαία Συνιστώσα (RESIDUAL VALUATION R_t)

Η τυχαία συνιστώσα (απόκλιση) μιας χρονολογικής σειράς δεν επηρεάζεται από τις άλλες συνιστώσες και δεν επαναλαμβάνεται σε τακτά χρονικά διαστήματα.

Λαμβάνοντας υπόψη τις τέσσερις συνιστώσες που παρουσιάστηκαν παραπάνω το προσθετικό υπόδειγμα μιας χρονολογικής σειράς υπολογίζεται ως το άθροισμα τους, δηλ:

$$Y_t = T_t + C_t + S_t + R_t$$

2.2.2 Στοχαστικά Υποδείγματα χρονολογικών σειρών

Η βασική ιδέα πίσω από τα στοχαστικά υποδείγματα (stochastic models) βασίζεται στην παρατήρηση ότι εάν οι διαδοχικές τιμές μιας χρονολογικής σειράς παρουσιάζουν μια ισχυρή συσχέτιση, τότε αυτές (πιθανότατα) έχουν παραχθεί από μια στοχαστική διαδικασία (stochastic process). Βασικό πρόβλημα που παρουσιάζεται στις χρονολογικές σειρές είναι το γεγονός ότι δεν μπορούμε να έχουμε περισσότερες από μια παρατηρήσεις για κάθε μια μεταβλητή σε συγκεκριμένο χρόνο. Το πρόβλημα παρακάμπτεται θεωρώντας τις παρατηρούμενες τιμές της χρονολογικής σειράς ως ένα δείγμα από ένα άπειρο πληθυσμό τέτοιων δειγμάτων, τα οποία θα μπορούσαν να είχαν παραχθεί από την ίδια στοχαστική διαδικασία.

Άλλο ένα πρόβλημα στην ανάλυση χρονολογικών σειρών είναι η επιλογή του κατάλληλου υποδείγματος προς εφαρμογή. Για την επιλογή του κατάλληλου υποδείγματος θα πρέπει να εξεταστεί η δομή των ιστορικών δεδομένων της χρονολογικής σειράς με συγκεκριμένα στατιστικά μέτρα τα οποία και θα αναλυθούν στη συνέχεια.

ΣΤΑΣΙΜΟΤΗΤΑ

Ως στάσιμη ορίζεται η διαδικασία της οποίας οι στατιστικές της ιδιότητες της σειράς δεν εξαρτώνται και δεν επηρεάζονται από κάποια μεταβολή στην αρχή του χρόνου της χρονοσειράς. Με άλλα λόγια, οι στατιστικές ιδιότητες (μέσος και διακύμανση) N παρατηρήσεων με αρχή την χρονική περίοδο t ($y_t, y_{t+1}, \dots, y_{t+N-1}$) παραμένουν αμετάβλητες και για τις N παρατηρήσεις με αρχή οποιαδήποτε άλλη περίοδο $t+k$ ($y_{t+k}, y_{t+k+1}, \dots, y_{t+k+N-1}$). Έτσι σε μια στάσιμη χρονολογική σειρά οι στατιστικές ιδιότητες θα είναι σταθερές για κάθε χρονική στιγμή t και άρα:

1. $E(y_t) = y_\mu$
2. $\text{Var}(y_t) = E[y_t - E(y_t)]^2 = \sigma_y^2$
3. $\text{Cov}(y_t, y_{t+k}) = \text{Cov}(y_{t+m}, y_{t+m+k})$ (δηλαδή η συνδιακύμανση μεταξύ δύο οποιονδήποτε τιμών y_t που απέχουν k περιόδους είναι συνάρτηση μόνο του k)

Ειδικότερα στις εφαρμογές των χρονοσειρών στην οικονομία οι περισσότερες από αυτές είναι μη στάσιμες (περιέχουν τάση, εποχικότητα και κυκλικές κυμάνσεις) και λόγω της

πολυπλοκότητας τους πρέπει να μετατρέπονται σε στάσιμες με την χρήση κατάλληλων τεχνικών. Για την μετατροπή μη στάσιμων χρονοσειρών σε στάσιμες μπορεί να αφαιρεθεί η τάση ή να θεωρηθούν διαδοχικές πρώτες διαφορές στο δείγμα.

ΑΥΤΟΣΥΝΔΙΑΚΥΜΑΝΣΗ & ΑΥΤΟΣΥΣΧΕΤΙΣΗ

Η αυτοσυνδιακύμανση είναι ένα μέτρο της συνδιακύμανσης μεταξύ δύο παρατηρήσεων της ίδιας χρονοσειράς οι οποίες παρατηρήσεις δεν ταυτίζονται αλλά βρίσκονται σε απόσταση μεταξύ τους. Η αυτοσυνδιακύμανση μεταξύ δύο παρατηρήσεων y_t και y_{t+k} που απέχουν k χρονικές περιόδους δίνεται από τον τύπο

$$\gamma_k = Cov(y_t, y_{t+k})$$

Σε αυτή την περίπτωση ορίζεται και ο συντελεστής αυτοσυσχέτισης ως:

$$\rho_k = \frac{Cov(y_t, y_{t+k})}{\sqrt{Var(y_t)} \sqrt{Var(y_{t+k})}}$$

που στην περίπτωση στασιμότητας της χρονοσειράς μετατρέπεται σε:

$$\rho_k = \frac{Cov(y_t, y_{t+k})}{Var(y_t)} = \frac{\gamma_k}{\gamma_0}$$

Ο συντελεστής αυτοσυσχέτισης ρ_k παίρνει τιμές στο διάστημα $[-1, 1]$ και στην περίπτωση που δύο παρατηρήσεις που απέχουν k χρονικές περιόδους έχουν μεγάλη σχέση μεταξύ τους τότε η τιμή του συντελεστή είναι κοντά στη μονάδα.

Το διάγραμμα συνάρτησης αυτοσυσχέτισης (correlogram) παράγεται από την διαγραμματική απεικόνιση των τιμών της ρ_k . Μέσω του διαγράμματος αυτοσυσχέτισης μπορεί να διακριθεί μια στάσιμη χρονοσειρά στην οποία οι συντελεστές αυτοσυσχέτισης θα φθίνουν γρήγορα προς το μηδέν καθώς μεγαλώνει ο αριθμός των υστερήσεων ενώ αντίθετα δεν θα συμβαίνει το ίδιο στις μη στάσιμες χρονολογικές σειρές.

Οι σχέσεις εξάρτησης των χρονοσειρών μπορούν να είναι γραμμικές, μη-γραμμικές ή συνδυασμός τους.

2.3 Έλεγχος Μη-Γραμμικότητας σε χρονοσειρές

Όπως αναλύθηκε στο πρώτο κεφάλαιο, υπάρχει διάσταση απόψεων ανάμεσα στους ερευνητές ως προς την αποτελεσματικότητα ή μη των χρηματοοικονομικών αγορών. Όπως και να έχει οι αγορές μπορούν να παρουσιαστούν μέσω χρονοσειρών και στην περίπτωση που αυτές είναι αποτελεσματικές θα παρουσιάζει διαχρονικά μια τυχαία συμπεριφορά στα δείγματα των τιμών ενώ στην αντίθετη περίπτωση θα παρουσιάζει αυτοσυσχέτιση μεταξύ τρεχουσών και ιστορικών τιμών (παράδειγμα σχέσεις εξάρτησης γραμμικές, μη-γραμμικές ή συνδυασμό τους).

Στον τομέα των χρηματοοικονομικών αγορών, λόγω της δυναμικής φύσης της ΥΑΑ αλλά και της ισορροπίας ρίσκου/απόδοσης ο δεσμευμένος μέσος είναι η παράμετρος που μελετάται με μεγαλύτερο ενδιαφέρον. Στην πραγματικότητα ο υπολογισμός του δεσμευμένου μέσου μιας χρονοσειράς και των δυναμικών (γραμμικών ή μη-γραμμικών) που τον χαρακτηρίζουν, είναι ιδιαίτερα σημαντικός εφόσον λάθος αποτελέσματα μπορούν να οδηγήσουν σε παραπλανητικά συμπεράσματα πάνω στις υποθέσεις που μελετώνται¹⁵. Η σημαντικότητα στην μη-γραμμικότητα στο μέσο οφείλεται κυρίως στο γεγονός ότι αυτή είναι κύριο χαρακτηριστικό στις χρονοσειρές των αποδόσεων των χρεογράφων, τα οποία δημιουργούν και είναι μέρος ενός δυναμικού περιβάλλοντος.

Έτσι, το πρώτο, κρίσιμο, βήμα στη μοντελοποίηση των χρονοσειρών¹⁶, είναι ο προσδιορισμός γραμμικότητας τους στο μέσο. Η γραμμικότητα στο μέσο σημαίνει ότι ο δεσμευμένος μέσος είναι γραμμικός συνδυασμός παραγόντων (μεταβλητών) που παρέχονται από τις πληροφορίες. Σαν δεύτερο βήμα της μοντελοποίησης πραγματοποιούνται έλεγχοι σχετικά με την διακύμανση.

Το τρίτο βήμα της μοντελοποίησης είναι η επιλογή του καταλληλότερου μη-γραμμικού μοντέλου. Η ορθή επιλογή του μη-γραμμικού μοντέλου οδηγεί σε πιο επιτυχημένες προβλέψεις σε σχέση με την χρήση ενός μοντέλου που είναι γραμμικό.

Τα είδη των ελέγχων που χρησιμοποιούνται στην απόδειξη μη-γραμμικότητας χωρίζονται στους γενικούς ελέγχους μέσω των οποίων ελέγχονται τα είδη των μη-γραμμικών δομών

¹⁵ Διαφορετικά, ο δεσμευμένος μέσος θα πρέπει να μοντελοποιηθεί με τέτοιο τρόπο ώστε να αποφευχθούν οικονομετρικές αστοχίες στη μελέτη της δεσμευμένης διακύμανσης.

¹⁶ Μπορεί να θεωρηθεί και σαν έλεγχος καταλληλότητας του επιλεγμένου μοντέλου προσομοίωσης των δεδομένων

στα δεδομένα¹⁷ και πιο ειδικοί που ελέγχουν την ύπαρξη συγκεκριμένων μη-γραμμικών δομών. Και τα δύο είδη ελέγχων μπορούν να εμπλέκουν το σύνολο των δεδομένων αλλά και υποσύνολα τους, αρκεί να μην ταυτίζονται και να ανταποκρίνονται στις τιμές διαφορετικών χρονικών περιόδων. Οι ερευνητές δεν χρησιμοποιούν έναν έλεγχο αλλά ένα σύνολο ελέγχων καθένας από τους οποίους ελέγχει ένα διαφορετικό τύπο μη-γραμμικότητας. Ονομαστικά, διάφοροι έλεγχοι που χρησιμοποιούνται στην έρευνα της χρηματιστηριακής αγοράς (και από τον παλαιότερο στο νεώτερο) είναι: ο έλεγχος bispectrum που παρουσιάζεται στο (Hinich, 1982), ο έλεγχος των McLeod and Li (McLeod, et al., 1983), και άλλοι. Πιο πρόσφατοι έλεγχοι είναι αυτός των νευρωνικών δικτύων (Lee, et al., 1993), ο έλεγχος Brock – Dechert – Scheinkman (BDS) (Brock, et al., 1996) και ο biconrelation (Hinich, 1996).

2.4 Προβλέψεις χρονοσειρών

Η διαδικασία πρόβλεψης μελλοντικών καταστάσεων είναι σημαντική σε πολλούς και διάφορους τομείς της επιστήμης και ειδικότερα της χρηματοοικονομικής, εφόσον μια ακριβής πρόβλεψη είναι θεμέλιο σχεδιασμού αλλά και πολιτικής. Οποιοσδήποτε συναλλαζόμενος μπορεί να κάνει σχέδια και να αναλάβει δράσεις, κυρίως, μέσω των προβλέψεων εφόσον οι τελευταίες έχουν σαν αποτέλεσμα την δημιουργία μιας προβολής του μελλοντικού περιβάλλοντος στο οποίο θα λάβουν χώρα οι δραστηριότητες του.

Αρχικό βήμα στην διαδικασία προβλέψεων είναι η συλλογή και ανάλυση ιστορικών στοιχείων και πληροφοριών και ακολουθεί η αξιολόγηση των παραγόντων που επηρεάζουν την πρόβλεψη.

Η διαδικασία προβλέψεων περιλαμβάνει την δημιουργία υποθέσεων πάνω στις μελλοντικές συνθήκες των αγορών, την δημιουργία σχεδίων και την πρόβλεψη των αποτελεσμάτων που θα έχει η εφαρμογή των τελευταίων.

Η όλη διαδικασία δεν είναι γραμμική αλλά κυκλική εφόσον σαν τελικό βήμα περιλαμβάνει την αξιολόγηση σε σχέση με την πραγματική έκβαση των γεγονότων. Με αυτό τον τρόπο βελτιώνεται η αξιοπιστία της πρόβλεψης μέσω μιας διαδικασίας κατανόησης των αιτιών που οδήγησαν σε λανθασμένες εκτιμήσεις.

¹⁷ Δεν γίνεται αναζήτηση συγκεκριμένων μη-γραμμικών μοντέλων

Είναι λοιπόν φανερό το πόσο σημαντική είναι η πρόβλεψη των μελλοντικών τιμών των αγορών. Στα επόμενα υποκεφάλαια παρουσιάζονται συνοπτικά οι διάφοροι μέθοδοι πρόβλεψης που αναφέρονται στην βιβλιογραφία.

2.4.1 Προβλέψεις με την χρήση Μεθόδων Εξομάλυνσης χρονοσειρών

Κύριος στόχος των μεθόδων εξομάλυνσης μιας χρονοσειράς, όπως αναφέρεται από τους συγγραφείς (Makridakis, et al., 1983), είναι αυτός της εξάλειψης της τυχαίας συνιστώσας και έχουν σαν αποτέλεσμα την υποβάθμιση της τάσης αλλά και της εποχιακής και κυκλικής συνιστώσας. Διαφορετικά, η εξομάλυνση έχει σαν αποτέλεσμα την στρογγυλοποίηση των μετρήσεων έτσι ώστε να παραχθεί ένα εξομαλυμένο στοιχείο το οποίο και αντικατοπτρίζει τη συστηματική κίνηση των σειρών.

Οι βασικές μέθοδοι εξομάλυνσης χρονολογικών σειρών, οι οποίες παρουσιάζονται στη συνέχεια, είναι δύο: η **μέθοδος κινητών μέσων** (simple moving average (MA) method) και η **μέθοδος εκθετικής εξομάλυνσης** (exponential smoothing method).

ΜΕΘΟΔΟΣ ΚΙΝΗΤΩΝ ΜΕΣΩΝ

Το όνομα της μεθόδου βασίζεται στο γεγονός πως κάθε νέα διαθέσιμη παρατήρηση αντικαθιστά την παλαιότερη και επομένως κάθε φορά υπολογίζεται ένας νέος μέσος όρος με αποτέλεσμα ο μέσος όρος να αλλάξει ή να μετακινηθεί όσο νέες παρατηρήσεις γίνονται διαθέσιμες. Η μέθοδος των κινητών μέσων είναι μια διαδικασία τριών βημάτων. Στο 1^ο βήμα γίνεται ο διαχωρισμός των παρατηρήσεων σε κινητά σύνολα. Βασική, αλλά και κρίσιμη, προϋπόθεση ο καθορισμός του αριθμού των επιθυμητών περιόδων. Το 2^ο υπολογίζει το μέσο όρο των κινητών συνόλων. Στην περίπτωση που τα κινητά σύνολα περιέχουν περιττό αριθμό παρατηρήσεων δεν υπολογίζονται κινητοί μέσοι για τις πρώτες και τελευταίες $\frac{k-1}{2}$ χρονικές περιόδους (όπου k ο αριθμός παρατηρήσεων για κάθε κινητό σύνολο). Στο 3^ο βήμα υπολογίζονται οι αποκλίσεις της χρονοσειράς γύρω από τους κινητούς μέσους σαν διαφορά ανάμεσα στις τιμές τους.

Βασικό μειονέκτημα της μεθόδου κινητών μέσων είναι η απώλεια πληροφοριών στην περίπτωση που τα κινητά σύνολα περιλαμβάνουν περιττό αριθμό παρατηρήσεων και οι μέσοι όροι των ακραίων συνόλων δεν υπολογίζονται.

ΣΤΑΘΜΙΣΜΕΝΟΣ ΚΙΝΗΤΟΣ ΜΕΣΟΣ ΟΡΟΣ

Η μέθοδος του σταθμισμένου κινητού μέσου όρου είναι μια διαφοροποίηση της βασικής μεθόδου κινητού μέσου, όπου κάθε παρατήρηση του δείγματος έχει διαφορετικό βάρος και αυτό λαμβάνεται στον υπολογισμό. Εδώ κάθε τιμή μέσου υπολογίζεται από τον τύπο:

$$\hat{\sigma}_t = \sum_{i=t-\tau}^{t-1} \lambda \sigma_i$$

Όπου σ_i η εξομαλυμένη τιμή (η απόκλιση από το μέσο κινούμενο όρο) της σειράς την χρονική στιγμή i και λ το βάρος.

ΜΕΘΟΔΟΣ ΕΚΘΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ

Η μέθοδος εκθετικής εξομάλυνσης είναι μια μέθοδος που χρησιμοποιείται στις προβλέψεις οι οποίες χρησιμοποιούν τον σταθμισμένο μέσο όρο παρελθόντων δειγμάτων χρονοσειρών για την πρόβλεψη της τιμής χρονοσειρών επόμενης περιόδου. Η κύρια ιδέα είναι ότι, οι πρόσφατες τιμές της σειράς έχουν μεγαλύτερη βαρύτητα από τις προηγούμενες (παρελθοντικές) τιμές. Η τιμή εξομάλυνσης δίνεται από τον τύπο:

$$\sigma_t = \lambda Y_t + (1 - \lambda)\sigma_{t-1}$$

όπου λ το βάρος, σ_t η εξομαλυμένη τιμή της σειράς και Y_t η τιμή της χρονοσειράς την χρονική περίοδο t .

Ο βαθμός εξομάλυνσης της σειράς εξαρτάται από την τιμή του βάρους λ .

2.4.1 Προβλέψεις χρονοσειρών με προβολή τάσης στο μέλλον

Η μέθοδος πρόβλεψης με προβολής τάσης εφαρμόζεται στην πρόβλεψη μελλοντικών τιμών χρονοσειρών στην περίπτωση που αυτές όμως παρουσιάζουν μια *μακροχρόνια γραμμική τάση*, και ανήκει στα περιγραφικά μοντέλα στα οποία δεν υπολογίζονται κατάλοιπα με πιθανοτικούς όρους.

Βασική λειτουργία τη μεθόδου είναι ο καθορισμός μιας γραμμικής σχέσης μεταξύ των παραμέτρων της αγοράς και του χρόνου με τέτοιο τρόπο που να προσεγγίζει (ιδανικά να ταυτίζεται) με την τάση.

2.4.2 Προβλέψεις χρονοσειρών με εποχιακά στοιχεία

Σημαντικό μέρος των τιμών αξιογράφων παρουσιάζεται μέσω χρονικών σειρών που παρουσιάζει τάση αλλά και εποχιακά στοιχεία. Οι εποχικοί παράγοντες υπολογίζονται ως ποσοστό της τάσης.

Οι χρονοσειρές αυτού του τύπου αντιμετωπίζονται με από-εποχοποίηση, με την απομάκρυνση δηλ των εποχιακών παραγόντων και των εποχιακών στοιχείων από την χρονοσειρά. Αποτέλεσμα της προηγούμενης διαδικασίας είναι η παραγωγή μιας χρονοσειράς με τάση η οποία μας οδηγεί στην υιοθέτηση της μεθόδου πρόβλεψης με προβολή τάσης στο μέλλον.

Σε συνέχεια της διαδικασίας πρόβλεψης, στα αποτελέσματα της προβολής τάσης ενσωματώνονται οι εποχικοί παράγοντες και ακολουθεί προσαρμογή της τάσης.

Στην πραγματικότητα, στις χρονοσειρές πέρα από την τάση και τους εποχιακούς παράγοντες συμπεριλαμβάνονται και τυχαίοι παράγοντες¹⁸ αλλά και κυκλικές μεταβλητές¹⁹. Λαμβάνοντας υπόψη όλους τους παράγοντες ο μαθηματικός τύπος που ακολουθεί δίνει την πραγματική τιμή της χρονοσειράς:

¹⁸Οι τυχαίοι παράγοντες είναι τυχαία φαινόμενα (μεταβλητές) στις χρονοσειρές οι οποίοι δεν εμπίπτουν στην τάση ή την εποχικότητα

¹⁹Η κυκλική μεταβλητή είναι μια εποχική μεταβλητή (και επομένως ποσοστό της τάσης) που αναφέρεται σε μεγάλο χρονικό διάστημα. Το μέγεθος του χρονικού διαστήματος, το οποίο μπορεί και να ποικίλει, καθιστούν δύσκολη την διαδικασία συλλογής δεδομένων.

$$Y_t = T_t \times C_t \times S_t \times I_t$$

Όπου:

- Y_t η τιμή της χρονοσειράς τη χρονική περίοδο t ,
- T_t η τάση της χρονοσειράς,
- C_t η κυκλική μεταβλητή,
- S_t οι εποχικοί παράγοντες και
- I_t οι τυχαίοι παράγοντες.

2.4.3 Προβλέψεις μέσω Μοντέλων Παλινδρόμησης

Τα μοντέλα παλινδρόμησης εισήχθησαν ως εργαλεία πρόβλεψης για τις περιπτώσεις στις οποίες η τιμή μιας χρονικής σειράς θεωρείται ως εξαρτημένη μεταβλητή και επιτρέπουν στατιστικά πορίσματα και προβλέψεις. Σκοπός της μεθόδου είναι κατ' αρχάς ο καθορισμός των ανεξάρτητων προβλεπτικών μεταβλητών και στη συνέχεια η διατύπωση μιας παλινδρομης εξίσωσης για την επίτευξη πρόβλεψης των χρονοσειρών.

Η βασική δυσκολία της μεθόδου έγκειται στο πρώτο βήμα καθώς η δυναμική μορφή αλλά και η πολυπλοκότητα των περισσότερων περιπτώσεων υπό πρόβλεψη που πραγματεύεται αυτή η εργασία εμπλέκουν ένα μεγάλο αριθμό ανεξάρτητων μεταβλητών.

Στο υπό εξέταση μοντέλο πρόβλεψης συνήθως γίνεται χρήση της πολλαπλής παλινδρόμησης η οποία εμπλέκει τουλάχιστον πέντε οικονομικές αλλά και χρονικές ανεξάρτητες μεταβλητές. Η ακρίβεια της πρόβλεψης, και γι' αυτό το πιο κρίσιμο σημείο της, εξαρτάται από την φάση επιλογής των ανεξάρτητων μεταβλητών (και τον βαθμό συσχέτισης με την τιμή της χρονοσειράς) και την συλλογή δεδομένων για αυτές.

Ένας δεύτερος τύπος παλινδρομου μοντέλου είναι το αυτοπαλίνδρομο, όπου η εξαρτημένη τιμή είναι πάλι η τιμή της χρονοσειράς και οι ανεξάρτητες οι ιστορικές στιγμές της ίδιας.

Ο συνδυασμός των δύο προηγούμενων παλινδρομων μοντέλων που εμπλέκει τόσο ανεξάρτητες οικονομικές και χρονικές μεταβλητές όσο και ιστορικές τιμές της ίδιας χρονοσειράς ορίζουν ένα τρίτο τύπο μοντέλου παλινδρόμησης.

Στην περίπτωση αυτή τα μοντέλα διακρίνονται επίσης σε *αιτιοκρατικά* και *στοχαστικά*. Στα αιτιοκρατικά μοντέλα αν καταφέρουμε να προσδιορίσουμε την τιμή της ανεξάρτητης μεταβλητής x τότε η τιμή της εξαρτημένης μεταβλητής y δύναται να προσδιοριστεί ακριβώς. Στην αντίθετη περίπτωση στην οποία η τιμή της εξαρτημένης μεταβλητής y δε μπορεί να προσδιοριστεί ακριβώς, το μοντέλο συσχέτισης των x και y ονομάζεται *στοχαστικό*. Το στοχαστικό μοντέλο χρησιμοποιείται στον υπολογισμό της κατανομής πιθανοτήτων πάνω σε ενδεχόμενα αποτελέσματα εισάγοντας έτσι μια ποικιλομορφία πάνω στα δεδομένα.

2.4.4 Προβλέψεις μέσω Παλίνδρομων Μοντέλων με ανεξάρτητα σφάλματα

Όσον αφορά τα προηγούμενα μοντέλα δεν λαμβάνονται υπόψη τα τυχόν τυχαία και ανεξάρτητα σφάλματα που είναι ανεξάρτητα από το χρόνο. Τα ανεξάρτητα σφάλματα παρόλα αυτά είναι μια κρίσιμη μεταβλητή που θα πρέπει να λαμβάνεται υπόψη σε τομείς της οικονομίας. Αυτό γιατί στους τελευταίους παρουσιάζονται πολλές περιπτώσεις που η εξαρτημένη μεταβλητή παραμένει η χρονοσειρά ενώ οι ανεξάρτητες μεταβλητές είναι οι επιχειρηματικοί δείκτες ή άλλες σειρές πρόβλεψης.

Για το λόγο αυτό εισήχθησαν τα παλινδρομικά μοντέλα για δεδομένα χρονοσειρών με ανεξάρτητα σφάλματα σαν υποπεριπτώσεις μοντέλων πολλαπλής παλινδρόμησης.

Τα παλίνδρομα μοντέλα που κατασκευάζονται για αυτές τις περιπτώσεις περιέχουν ανεξάρτητες μεταβλητές που είναι υστερημένες με βασικό πλεονέκτημα ότι οι δεν είναι απαραίτητο να προβλεφθούν οι ανεξάρτητες μεταβλητές ώστε να γίνει μια επιτυχής πρόβλεψη της εξαρτημένης μεταβλητής.

2.4.5 Προβλέψεις μέσω Παλίνδρομων Μοντέλων με εξαρτώμενα σφάλματα

Επαναλαμβάνεται συχνά το γεγονός, κατά την πρώτη φάση της δημιουργίας παλίνδρομων μοντέλων, κάποια (-ες) μεταβλητή (-ες) να μην θεωρηθούν και να παραλειφθούν. Επίσης είναι συχνό φαινόμενο οι κύριες τυχαίες επιρροές στις χρονοσειρές να τείνουν να

διατηρούνται για αρκετές περιόδους. Και στις δύο περιπτώσεις το αποτέλεσμα είναι η εισαγωγή στη θεώρηση τυχαίων σφαλμάτων που σχετίζονται χρονικά.

Στη μέθοδο προβλέψεων μέσω παλίνδρομων μοντέλων με εξαρτώμενα σφάλματα τα τελευταία σχετίζονται χρονικά από χρονική περίοδο σε χρονική περίοδο. Σε αυτές τις περιπτώσεις τα σφάλματα ονομάζονται αυτοσυσχετιζόμενα ή συσχετιζόμενα σειριακά.

Το πιο κοινά χρησιμοποιημένο μοντέλο στην περίπτωση των σφαλμάτων που σχετίζονται από χρονική περίοδο σε χρονική είναι το καθώς ονομάζεται το *πρώτης τάξης αυτοπαλίνδρομο υπόδειγμα λάθους*, το οποίο δίνεται από τον τύπο:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

όπου:

- ε_t είναι η τιμή του υποδείγματος λάθους την χρονική στιγμή t ,
- ρ είναι η παράμετρος αυτοσυσχέτισης και λαμβάνει τιμές στο διάστημα $(-1, 1)$ και
- u είναι τα ανεξάρτητα σφάλματα

Το μοντέλο παλινδρόμησης με εξαρτώμενα σφάλματα υποθέτει πως το σφάλμα ε_t για την μελλοντική χρονική περίοδο t περιλαμβάνει το σφάλμα ε_{t-1} της παρούσας χρονικής στιγμής αυξημένο κατά ένα τυχαίο και ανεξάρτητο χρονικά διαταρακτικός όρος u_t .

2.5 Γραμμικά Μοντέλα πρόβλεψης χρονοσειρών

Όπως ήδη έχει αναφερθεί τα γραμμικά μοντέλα είναι αυτά που χρησιμοποιούνται σε στατιστικές εφαρμογές όπου το πρόβλημα είναι αυτό της συσχέτισης δύο ή περισσότερων τυχαίων μεταβλητών. Η μελέτη των γραμμικών μοντέλων επικεντρώνεται στον καθορισμό ενός μαθηματικού μοντέλου ως προς τον προσδιορισμό της σχέσης των υπό μελέτη μεταβλητών. Το ουσιώδες ζήτημα στην κατασκευή του μοντέλου είναι η προϋπόθεση, εκ των προτέρων, μιας κάποια γνώσης πάνω στην φύση της σχέσης μεταξύ των μεταβλητών. Έστω, για παράδειγμα, ένας επενδυτής στην χρηματοπιστωτική αγορά. Για την πραγματοποίηση της αγοράς ενός μεγάλου αριθμού ομολόγων ο ενδιαφερόμενος επιθυμεί να προβλέψει την τιμή των μετοχών σε βάθος χρόνου που αυτός επιθυμεί. Θα πρέπει να

λάβει λοιπόν υπόψη του τους παράγοντες που μπορεί να επηρεάσουν την τιμή αυτή όπως τα επιτόκια, ο πληθωρισμός, η ζήτηση, η οικονομική κατάσταση κτλ. Έστω τώρα ότι η τιμή των ομολόγων εξαρτάται γραμμικά από την ζήτηση ομολόγων. Σύμφωνα με το γραμμικό μοντέλο η σχέση αυτή περιγράφεται από τον τύπο ντετερμινιστικών μοντέλων:

$$Y = \alpha + \beta X$$

όπου α και β είναι σταθερές, Y η εξαρτημένη μεταβλητή (η μελλοντική τιμή των ομολόγων) και X η ανεξάρτητη μεταβλητή (η ζήτηση). Κατ' αυτό τον τρόπο το πρόβλημά συνίσταται στον προσδιορισμό μιας ευθείας που ονομάζεται *ευθεία παλινδρόμησης* (regression line).

Οι περιπτώσεις πραγματικής εφαρμογής, σπάνια παρουσιάζουν περιπτώσεις δύο μεγεθών με μια τέλεια γραμμική σχέση, εφόσον συνήθως το επηρεαζόμενο μέγεθος καθορίζεται από περισσότερες από μια μεταβλητές, κάποιες από τις οποίες μπορεί να μην είναι καν μετρήσιμες. Για το λόγο αυτό εισάγονται μοντέλα που περικλείουν στοιχεία μιας κάποιας τυχαιότητας (randomness) και που ονομάζονται *μοντέλα πιθανότητας* (probabilistic models). Τα μοντέλα πιθανότητας κατασκευάζονται μέσω ενός ντετερμινιστικού μοντέλου που προσεγγίζει τη σχέση υπό μελέτη, στο οποίο προστίθεται ο τυχαίος όρος των αποκλίσεων (ή λάθους) του ντετερμινιστικού όρου. Με την προσθήκη του τυχαίου όρου προσεγγίζονται οι τιμές των μεταβλητών που επηρεάζουν την πραγματική τιμή της εξαρτημένης μεταβλητής, οι οποίες όμως δεν είναι μέρος του μοντέλου.

Μοντέλα πιθανότητας είναι τα:

1. $Y = \alpha + \beta X + \varepsilon$ (απλή γραμμική παλινδρόμηση, simple linear regression).
2. $Y = \alpha + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon$ (πολλαπλή παλινδρόμηση, multiple regression).
3. $Y = \alpha + \beta_1 X + \beta_2 X^2 + \dots + \beta_k X^k + \varepsilon$ (πολυωνομική παλινδρόμηση, multinomial regression).

Στη συνέχεια, για την εφαρμογή των γραμμικών μοντέλων πρόβλεψης στις χρηματοοικονομικές αγορές, θεωρούμε σαν δειγματικό χώρο τη χρονοσειρά των μεταβολών του χρηματοοικονομικού δείκτη $\{x_i\}_{i=1}^N$, θεωρώντας ότι η χρονοσειρά δίνεται από τις πρώτες διαφορές ή από τις αποδόσεις του δείκτη. Η χρονοσειρά θεωρείται στάσιμη και αποτελεί την πραγματοποίηση μιας στοχαστικής διαδικασίας, τον ορισμό δηλαδή κάποιου συστήματος που οδηγείται από θόρυβο ή τυχαία βήματα.

2.5.1 Μοντέλα Προβλέψεων AR, ARMA και ARIMA

Όπως ήδη αναφέρθηκε στην εισαγωγή του υποκεφαλαίου πάνω στα γραμμικά μοντέλα προβλέψεων, τα μοντέλα παλινδρόμησης (regression models) έχουν ως βασικό στόχο τον ορισμό μιας εξαρτημένης μεταβλητής ως συνάρτηση κάποιων άλλων ανεξάρτητων μεταβλητών. Ειδικότερα, στην περίπτωση των γραμμικών μοντέλων παλινδρόμησης η συνάρτηση συσχέτισης είναι γραμμική και έτσι η υπό εξέταση εξαρτημένη μεταβλητή δίνεται ως γραμμικός συνδυασμός των ανεξάρτητων μεταβλητών.

Μοντέλο AR

Το πρώτο είδος μοντέλου παλινδρόμησης που εισάγεται είναι τα αυτοπαλινδρομούμενα μοντέλα (ή *AutoRegressive models*, ή AR). Τα μοντέλα AR είναι μοντέλα γραμμικής παλινδρόμησης, στα οποία η εξαρτημένη μεταβλητή είναι μια τυχαία μεταβλητή της χρονοσειράς σε μια χρονική στιγμή t (x_t), ενώ οι ανεξάρτητες μεταβλητές είναι τυχαίες τιμές της χρονοσειράς σε προηγούμενους χρόνους ($x_{t-1}, x_{t-2}, \dots, x_{t-p}$). Σε αυτή την περίπτωση, ο αριθμός των υστερήσεων (p) καλείται η τάξη (order) του αυτοπαλινδρομούμενου μοντέλου. Από την άλλη ένα αυτοπαλινδρομούμενο μοντέλο τάξης p συμβολίζεται με AR(p) και ορίζεται από τον τύπο:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p} + z_t$$

όπου φ_i είναι οι συντελεστές του μοντέλου και $\{z_t\} \sim iid$ ²⁰ με μέση τιμή 0 και διασπορά (του λευκού θορύβου) σ_z^2 . Για τον προσδιορισμό του μοντέλου AR είναι απαραίτητη η διασπορά του λευκού θορύβου καθώς και οι συντελεστές του μοντέλου τα οποία και υπολογίζονται ως μια εκτίμηση μέσω της χρονοσειράς. Από τον προηγούμενο τύπο υπολογισμού της εξαρτημένης μεταβλητής του μοντέλου AR(p) προκύπτει ότι η μεταβλητή της χρονοσειράς τη χρονική στιγμή t εξαρτάται από το γραμμικό συνδυασμό των p τελευταίων τιμών της χρονοσειράς και από ένα δεύτερο μέρος που καλείται *στοχαστικό*. Ο στοχαστικός παράγοντας (z_t), με την σειρά του, επηρεάζεται από εξωγενείς επιδράσεις κατά την χρονική στιγμή t ²¹.

Η κατασκευή του μοντέλου AR(p) το φέρνει σε άμεση σχέση με την αυτοσυσχέτιση. Αλλά μια μικρή αυτοσυσχέτιση βραχείας κλίμακας οδηγεί σε αδυναμία πρόβλεψης με το

²⁰ iid (independent and identically distributed) είναι η διασπορά του λευκού θορύβου

²¹ Στα χρηματοοικονομικά ο στοχαστικός παράγοντας z_t αναφέρεται και ως σοκ (shock) της χρονοσειράς

συγκεκριμένο μοντέλο. Γι' αυτό το λόγο τα AR μοντέλα δεν παρουσιάζουν προβλεπτική ικανότητα ειδικά σε χρονοσειρές όπου οι μεταβολές των χρηματοοικονομικών δεικτών δεν έχουν σημαντικές συσχετίσεις.

Η διαδικασία πρόβλεψης μέσω του μοντέλου AR περιλαμβάνει την διαδικασία προσαρμογής του μοντέλου σε στάσιμη χρονοσειρά. Για την προσαρμογή αυτή θα πρέπει να γίνει υπολογισμός των παραγόντων που επηρεάζουν το μοντέλο και γι' αυτό θα πρέπει να:

- επιλεγεί η τάξη p του μοντέλου,
- γίνει η εκτίμηση των παραμέτρων φ_i και σ_z^2 του μοντέλου και
- γίνει διαγνωστικός έλεγχος καταλληλότητας του μοντέλου.

Η επιλογή της τάξης p του μοντέλου AR και η εκτίμηση των παραμέτρων, πραγματοποιούνται μέσω του κριτηρίου της μερικής αυτοσυσχέτισης (partial autocorrelation). Η τελευταία ορίζεται για κάθε υστέρηση p . Για μια χρονοσειρά μήκους $p+1$, $\{x_i\}_{i=t}^{t-p}$, η μερική αυτοσυσχέτιση μεταξύ των x_t και x_{t-p} εξουδετερώνει την συσχέτιση των ενδιάμεσων μεταβλητών. Ο υπολογισμός της μερικής αυτοσυσχέτισης γίνεται από την εκτίμηση των παραμέτρων του μοντέλου με αυξανόμενη τάξη. Στην, γενική, αυξημένη τάξη i το μοντέλο θα είναι:

$$x_t = \varphi_{1,i}x_{t-1} + \varphi_{2,i}x_{t-2} + \dots + \varphi_{i,i}x_{t-i} + z_t$$

και οι συντελεστές του μοντέλου είναι οι $\varphi_{i,i}$ και x_{t-i} .

Στις χρονοσειρές που προέρχονται από ένα μοντέλο AR(p), ο συντελεστής $\varphi_{i,i}$ ($i < p$) δεν είναι μηδενικός, ενώ για τις τάξεις μοντέλου $i > p$ οι συντελεστές είναι (στατιστικά) μηδενικοί. Βάσει αυτού του κριτηρίου και δημιουργώντας συνεχόμενες αυξημένες τάξεις είμαστε σε θέση να προσδιορίσουμε την τάξη του AR.

Τον υπολογισμό της τάξης και των συντελεστών του μοντέλου ακολουθεί η αξιολόγηση επίδοσης του μοντέλου. Ο έλεγχος αυτός πραγματοποιείται μέσω του υπολογισμού των σφαλμάτων προσαρμογής, και τουτέστιν βασίζεται στην εκτίμηση των διαφορών πραγματικών και εκτιμωμένων τιμών που αποτελούν και εκτίμηση του z_t . Αν η διασπορά που εκτιμάτε δεν είναι λευκού θορύβου (iid) τότε υπάρχουν συσχετίσεις στη χρονοσειρά (συσχετίσεις που δεν έχουν συμπεριληφθεί) και επομένως το μοντέλο θεωρείται ως επαρκές.

Μοντέλο ARMA

Στο δεύτερο γραμμικό μοντέλο που παρουσιάζεται λαμβάνεται υπόψη η υπόθεση ότι οι εξωγενείς παράγοντες προηγούμενων χρόνων (μέρος κινούμενου μέσου ή *moving average*) δύναται να επηρεάσουν τη μεταβλητή της χρονοσειράς τη χρονική στιγμή t . Το μοντέλο που συμπεριλαμβάνει το μέρος κινούμενου μέσου όρου είναι το αυτοπαλινδρομούμενο μοντέλο κινούμενου μέσου (*AutoRegressive Moving Average* ή *ARMA*). Το συγκεκριμένο μοντέλο δίνεται μέσω του τύπου:

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p} + z_t - \theta_1 z_{t-1} - \dots - \theta_q z_{t-q}$$

Σύμφωνα με τον παραπάνω τύπο το αυτοπαλίνδρομο μοντέλο χωρίζεται σε ένα παλίνδρομο (AR) μέρος τάξης p και σε ένα κινούμενου μέσου (MA) μέρους τάξης q . Το αυτοπαλίνδρομο μοντέλο συμβολίζεται ως $ARMA(p,q)$.

Τα μοντέλα τύπου $ARMA(p,q)$ (που συμπεριλαμβάνουν τα AR και MA) είναι κατάλληλα προς εφαρμογή προβλέψεων στάσιμων χρονοσειρών.

Μοντέλο ARIMA

Στην περίπτωση που η χρονοσειρά μεταβολών του δείκτη $\{x_t\}_{t=1}^N$ (ανεξάρτητη μεταβλητή) προσδιορίζεται από ένα μοντέλο $ARMA(p,q)$ τότε η χρονοσειρά μεταβολών του δείκτη $\{y_t\}_{t=0}^N$ (εξαρτημένη μεταβλητή) προσδιορίζεται μέσω ενός ολοκληρωμένου αυτοπαλινδρομούμενου μοντέλου κινούμενου μέσου ή ολοκληρωμένο μικτό μοντέλο (*autoregressive integrated moving average model*, ή *ARIMA*). Το μοντέλο συμβολίζεται $ARIMA(p,1,q)$. Στο συμβολισμό αυτό ο δείκτης 1 δείχνει ότι οι διαφορές σε πρώτο χρόνο της χρονοσειράς $\{y_t\}_{t=0}^N$, δηλαδή οι μεταβολές του $\{x_t\}_{t=1}^N$, προσδιορίζονται μέσω του μοντέλου $ARMA(p,q)$. Συμπερασματικά τα μοντέλα $ARIMA$ ορίζονται μέσω μιας επανάληψης του μετασχηματισμού των πρώτων d διαφορών. Η επανάληψη συνεχίζεται μέχρις ότου η παραγόμενη χρονοσειρά να είναι στάσιμη. Ένα μοντέλο $ARIMA$ που παράγεται μέσα από d επαναλήψεις συμβολίζεται ως $ARIMA(p,d,q)$.

Στην περίπτωση που, η υπό εξέταση χρονοσειρά, δεν είναι στάσιμη θα πρέπει να γίνει μετατροπή της τελευταίας σε στάσιμη και στην συνέχεια να εφαρμοστούν μοντέλα πρόβλεψης τύπου $ARMA$. Η στασιμότητα της χρονοσειράς επιτυγχάνεται μέσω της

απαλοιφής της τάσης μιας χρονοσειράς παίρνοντας τις πρώτες διαφορές. Στην περίπτωση των χρηματοοικονομικών η απαλοιφή αυτή πραγματοποιείται μέσω των αποδόσεων. Στην περίπτωση που η προκύπτουσα χρονοσειρά δεν είναι στάσιμη, επαναλαμβάνουμε η διαδικασία έως ότου αυτή να μετατραπεί σε στάσιμη. Όπως ήδη αναφέρθηκε, η διαδικασία d επαναλήψεων των πρώτων διαφορών για την επίτευξη μιας στάσιμης χρονοσειράς οδηγεί σε ένα μοντέλο ARIMA(p,d,q) ή διαφορετικά μοντέλα Box-Jenkins.

Η όλη διαδικασία γραμμικής ανάλυσης χρονοσειρών στα πλαίσια πρόβλεψης ακολουθεί τα παρακάτω βήματα:

1. Αρχικά δημιουργείται το γράφημα αναπαράστασης της χρονοσειράς και της αυτοσυσχέτισης.
2. Στην περίπτωση στην οποία η υπολογιζόμενη αυτοσυσχέτιση για κάθε υστέρηση είναι στατιστικά ασήμαντη, τότε δεν ακολουθείται η προσαρμογή του γραμμικού μοντέλου. Εδώ στα δεδομένα προσαρμόζεται ένα μη γραμμικό μοντέλο. Αντίθετα, στην περίπτωση που η αυτοσυσχέτιση αποδεικνύεται ισχυρή και παρατηρείται μια φθίνουσα πορεία στις τιμές της σε σχέση με την υστέρηση τότε η χρονοσειρά δεν είναι στάσιμη και ξεκινάει το βήμα μετατροπής της χρονοσειράς σε στάσιμη παίρνοντας τις πρώτες διαφορές.
3. Υπολογίζονται η τάξη και οι συντελεστές του μοντέλου χρησιμοποιώντας την μερική αυτοσυσχέτιση.
4. Στη χρονοσειρά που προκύπτει από τον υπολογισμό των πρώτων διαφορών προσαρμόζεται μοντέλο ανάλογα το κατάλληλο μοντέλο AR ή ARMA τάξης που έχει επιλεχτεί.
5. Ακολουθεί η προσαρμογή του μοντέλου στη χρονοσειρά και ο έλεγχος καταλληλότητας του (ελέγχοντας αν η χρονοσειρά των υπολοίπων είναι iid).
6. Χρησιμοποιείται το επιλεγμένο μοντέλο ως προς την πρόβλεψη και τον μετασχηματισμό των προβλέψεων στην αρχική χρονοσειρά.

Για την πραγματοποίηση του 6^{ου} βήματος εργαζόμαστε ως εξής: Από την αρχική χρονοσειρά $\{y_0, y_1, \dots, y_n\}$ προκύπτει η χρονοσειρά των πρώτων διαφορών $\{x_1, x_2, \dots, x_n\}$, όπου $x_t = y_t - y_{t-1}$. Στην χρονοσειρά διαφορών που προκύπτει εφαρμόζεται το μοντέλο ARMA(p,q) έχοντας έτσι σαν αποτέλεσμα την πρόβλεψη για ένα χρονικό βήμα $x_n(1)$. Αυτή η πρόβλεψη μεταφέρεται στην αρχική σειρά ως:

$$y_n(1) = y_n + x_n(1).$$

Γενικότερα, η πρόβλεψη για ένα χρονικό διάστημα T βημάτων δίνεται από την σχέση

$$y_n(T) = y_n(T-1) + x_n(T)$$

Στη σχέση αυτή ο όρος $x_n(T)$ είναι η πρόβλεψη x_{n+T} υπολογιζόμενη μέσω του μοντέλου ARMA (p,q) και ο όρος $y_n(T-1)$ έχει υπολογιστεί σαν πρόβλεψη του προηγούμενου βήματος.

2.6 Μη-Γραμμικά Μοντέλα πρόβλεψης χρονοσειρών

Όσον αφορά την πρόβλεψη χρηματοοικονομικών χρονοσειρών (δεικτών) οι τελευταίες δεν μπορούν να παρουσιαστούν μέσω ενός γραμμικού μοντέλου. Αυτό έχει σαν αποτέλεσμα οι χρηματιστηριακές και οικονομικές χρονοσειρές να μην μπορούν να χαρακτηριστούν μέσω των βασικών στατιστικών μετρήσεων (μέσος όρος, διασπορά και αυτοσυσχέτιση) και επομένως να μην μπορούν παρασταθούν μέσω μιας κανονικής κατανομής (Gaussian) αλλά και μέσω γραμμικών μοντέλων όπως το ARMA και το ARIMA. Σε αυτές τις περιπτώσεις, λοιπόν, εφαρμόζονται προσεγγίσεις *μη-γραμμικής ανάλυσης χρονοσειρών (nonlinear time series analysis)*. Αυτές οι προσεγγίσεις εμπλέκουν τη διερεύνηση και εκτίμηση χαρακτηριστικών σύνθετων (μη-γραμμικών) δομών και έχουν ως στόχο την παραγωγή περισσότερης πληροφορίας για το υπό μελέτη μη- γραμμικό σύστημα. Πιο συγκεκριμένα, οι μέθοδοι αυτοί βασίζονται στην *θεωρία των μη-γραμμικών δυναμικών συστημάτων* και του *χάους*.

Η πραγματική *τυχειότητα* διακρίνεται σε ένα δυναμικό σύστημα μέσω μορφής θορύβου. Ο θόρυβος είναι διαταραχές που οφείλονται σε κάποιο τυχαίο εξωτερικό παράγοντα και όχι στο ίδιο το σύστημα. Σαν *χάος* ορίζεται μια φαινομενικά τυχαία συμπεριφορά που παρατηρείται στα διάφορα συστήματα, πίσω από την οποία όμως κρύβεται μια προσδιοριστική (ντετερμινιστική) δυναμική. Εάν η παρουσία του θορύβου είναι μικρή, η χαοτική δομή δε χάνεται. Η δυναμική αυτή εκφράζεται από μικρό αριθμό ανεξάρτητων μεταβλητών ενώ η δυναμική εξέλιξη στο χρόνο είναι προϋπόθεση ως προς τον έλεγχο της χαοτικής συμπεριφοράς (Συριόπουλος, και συν., 2000).

Σημαντική διαφορά των μη-γραμμικών δυναμικών συστημάτων με τα γραμμικά συστήματα είναι το γεγονός ότι τα πρώτα δεν κάνουν χρήση του θορύβου t_z για την δημιουργία των διαφόρων καταστάσεων (καταστάσεις τις οποίες τα γραμμικά συστήματα δεν είναι σε θέση να αναπαραστήσουν).

Τα μη- γραμμικά χαρακτηριστικά που εκτιμούνται μέσω της μεθοδολογίας της μη- γραμμικής ανάλυσης είναι η *μορφοκλασματική διάσταση (fractal dimension)* και οι *εκθέτες Lyapunov (Lyapunov exponents)*. Στα διάφορα μη-γραμμικά μοντέλα περιλαμβάνονται τα *νευρωνικά δίκτυα (neural networks)*, τα *μη- γραμμικά προσθετικά μοντέλα (nonlinear additive models)* καθώς και τα *τοπικά γραμμικά μοντέλα (local linear models)*.

Χρονοσειρές δυναμικών συστημάτων

Οι χρονοσειρές δυναμικών συστημάτων ορίζονται με 2 διαφορετικούς τρόπους σε σχέση με το χρόνο δειγματοληψίας:

1. Αν ο χρόνος δειγματοληψίας είναι διακριτός η χρονοσειρά ορίζεται από εξισώσεις διαφορών,
2. Αν ο χρόνος δειγματοληψίας είναι συνεχής η χρονοσειρά ορίζεται μέσω διαφορικών εξισώσεων με αποτέλεσμα η δειγματοληψία να είναι πολύ πυκνή.

Στις περιπτώσεις οικονομετρικών εφαρμογών ο χρόνος θεωρείται διακριτός και επομένως η χρονοσειρά ορίζεται μέσω διαφορών. Πιο συγκεκριμένα, στην περίπτωση του διακριτού χρόνου το δυναμικό σύστημα ορίζεται μέσω του τύπου (Κουγιουμτζής, 2005):

$$s_t = f^t(s_0)$$

όπου:

- s_t : είναι το διάνυσμα θέσης (κατάστασης) του συστήματος τη χρονική στιγμή t .
- $s_t \in \mathbb{R}^d$ όπου d είναι η διάσταση του Ευκλείδειου χώρου καταστάσεων του συστήματος.
- s_0 : είναι η αρχική συνθήκη δηλαδή το διάνυσμα θέσης για την χρονική στιγμή 0.
- $f^t: \mathbb{R}^d \rightarrow \mathbb{R}^d$ η συνάρτηση του συστήματος που απεικονίζει το s_0 στο s_t .
- t : είναι ο χρόνος που μπορεί να είναι συνεχής ή διακριτός χρόνος.

Η τροχιά του συστήματος $\{s_0, s_1, \dots\}$ ορίζεται ξεκινώντας από ένα αρχικό σημείο (ή αρχική συνθήκη) s_0 και εφαρμόζοντας επαναληπτικά την παραπάνω εξίσωση. Η τροχιά περιορίζεται στο διάστημα R^d και καταλήγει ασυμπτωτικά σε ένα σύνολο σημείων που ονομάζεται *ελκυστής* (*attractor*). Λόγω του τρόπου κατασκευής του ο ελκυστής μπορεί να είναι:

- ένα ευσταθές σημείο ισορροπίας του συστήματος (*stable equilibrium point*),
- ένα πεπερασμένο σύνολο σημείων (παρουσιάζεται σε τροχιές που παρουσιάζουν περιοδικότητα),
- κάποιο άλλο μη πεπερασμένο σύνολο σημείων.

Στους ελκυστές της τρίτης κατηγορίας, που αποτελούνται από ένα μη πεπερασμένο σύνολο σημείων, περιλαμβάνεται ο *παράξενος ελκυστής* (*strange attractor*), που παρουσιάζει την ιδιότητα της *αυτο-ομοιότητας* (*self-similarity*) σε διαφορετικές κλίμακες του χώρου, είναι δηλαδή ένα *μορφοκλασματικό σύνολο* (*fractal*).

Τα *μορφοκλασματικά σύνολα* είναι σύνολα τα οποία χαρακτηρίζονται από τη *μορφοκλασματική διάσταση* (*fractal dimension*), δηλαδή έναν μη-ακέραιο αριθμό (μικρότερο της Ευκλείδειας διάστασης του ελκυστή) που αντιπροσωπεύει το βαθμό αυτό-ομοιότητας του συνόλου.

Από την άλλη, οι παράξενοι ελκυστές βασίζονται στις τροχιές δυναμικών συστημάτων. Αυτά τα συστήματα παρουσιάζουν *ευαίσθητη εξάρτηση στις αρχικές συνθήκες* (*sensitive dependence on initial conditions*) και ονομάζονται *χαοτικά δυναμικά συστήματα* (*chaotic dynamical systems*). Η ευαισθησία των χαοτικών δυναμικών συστημάτων στις αρχικές συνθήκες, έχει σαν αποτέλεσμα κοντινές τροχιές τους να αποκλίνουν γρήγορα και έτσι να παρουσιάζουν μια στοχαστική συμπεριφορά.

Μια χρονοσειρά $\{x_t\}_{t=0}^N$, προέρχεται από ένα δυναμικό σύστημα όταν για κάθε στοιχείο της χρονοσειράς ισχύει η σχέση:

$$x_t = h(s_t).$$

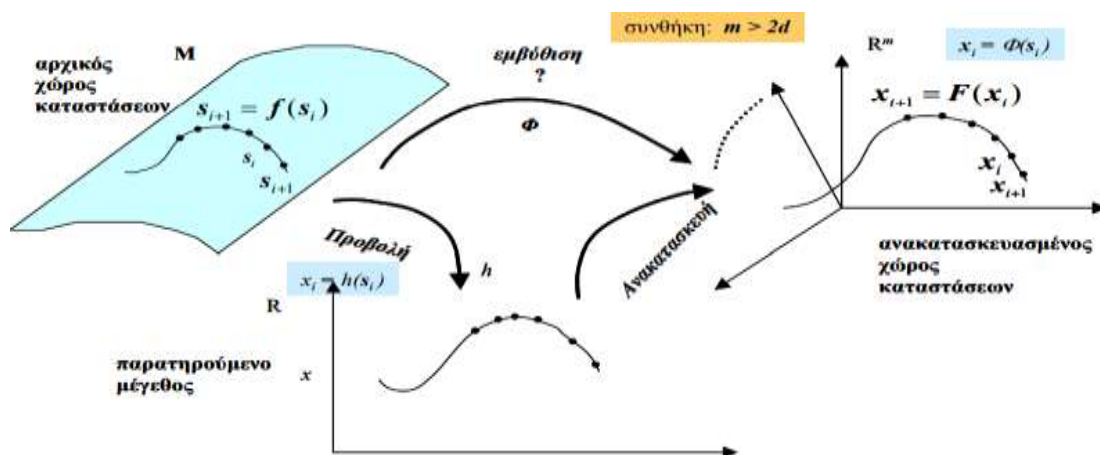
Στην παραπάνω σχέση η συνάρτηση $\square: R^d \rightarrow R$ ονομάζεται *συνάρτηση παρατήρησης* (*observation function*) και ορίζει μια προβολή για κάθε s_t της τροχιάς πάνω σε έναν από τους άξονες R^d .

Η διαχείριση των χαοτικών χρονοσειρών περιλαμβάνει δύο βήματα τα οποία και αναλύονται στην συνέχεια.

Ανακατασκευή χώρου καταστάσεων

Η συνάρτηση παρατήρησης παράγει μεν κάποια αποτελέσματα τα οποία όμως δεν επαρκούν για την δυναμική ανάλυση χρονοσειρών, για τον καθορισμό δηλαδή τόσο της συνάρτησης του συστήματος f όσο της διάστασης του χώρου καταστάσεων d στο οποίο βρίσκονται οι τροχιές που παράγει το σύστημα αλλά και τη συνάρτηση παρατήρησης h . Εισάγεται λοιπόν η ανάγκη σχηματισμού ειδώλου (ισοδύναμου) του αρχικού ελκυστή. Η δημιουργία του ειδώλου πραγματοποιείται μέσω του θεωρήματος του Takens (Takens theorem). Το συγκεκριμένο θεώρημα ουσιαστικά επιτρέπει (υπό συνθήκες) την κατασκευή ενός νέου χώρου καταστάσεων διαφορετικής διάστασης m στον οποίο χώρο οι ανακατασκευασμένες τροχιές $\{x_t\} \in \mathbb{R}^m$ από τη χρονοσειρά $\{s_t\}$, διατηρούν τις ίδιες χαρακτηριστικές. Αυτό σημαίνει πως στα δύο συστήματα παρουσιάζεται η ιδιότητα εμφάνισης (embedding) όπου τόσο το αρχικό όσο και το ανακατασκευασμένο παρουσιάζουν τα ίδια δυναμικά χαρακτηριστικά.

Βασική συνθήκη του θεωρήματος του Takens πάνω στις διαστάσεις είναι ότι πρέπει να ισχύει ότι $m \geq 2D + 1$, όπου D είναι η μορφοκλασματική διάσταση του ελκυστή. Τα αποτελέσματα της ανακατασκευής είναι προσεγγιστικά για το λόγο ότι το θεώρημα αφορά χρονοσειρές με άπειρο μήκος και χωρίς θόρυβο, συνθήκη η οποία πρακτικά δεν συναντάται.



Εικόνα 2.5: Ανακατασκευή χώρου καταστάσεων Πηγή <http://users.auth.gr/dkugiu/Teach/Econophys/Chp4.pdf>

Όπως παρουσιάζεται και στην Εικόνα 5 η ανακατασκευή του χώρου καταστάσεων πραγματοποιείται μέσω της δημιουργίας μιας σειράς t σημείων στο χώρο \mathbb{R}^m μέσω της προβολής των αρχικών παρατηρήσεων x_n .

Μεταξύ των διαφόρων μεθόδων ανακατασκευής του χώρου καταστάσεων είναι η μέθοδος των υστερήσεων (*method of delays*) χαρακτηριστική για την απλότητα της. Σε αυτή τα σημεία x_t σχηματίζονται ως

$$x_t = [x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}]'$$

Όπου:

- m είναι η διάσταση της εμβάθυνσης που ορίζει τον αριθμό των παρατηρήσεων που γίνονται συνιστώσες του ανακατασκευασμένου διανύσματος.
- τ είναι η υστέρηση (delay) που ορίζει με ποια χρονική διαφορά επιλέγονται οι m παρατηρήσεις για χρόνους μικρότερους της χρονικής στιγμής t .

Στον υπολογισμό το χρονικό παράθυρο ανακατασκευής $(m - 1)\tau$ δεν θα πρέπει να είναι πολύ μικρό, γιατί έτσι δε θα έχει χρησιμοποιηθεί αρκετή πληροφορία, αλλά ούτε πολύ μεγάλο, γιατί θα υπάρχει περιττή.

Όσον αφορά την επιλογή των τιμών m και τ συνήθως τίθεται $\tau = 1$ ενώ για την παράμετρο m δεν υπάρχει κάποια μέθοδος που να εκτιμά πάντα κατάλληλες τιμές.

Τοπικά μοντέλα πρόβλεψης

Όπως ήδη αναφέρθηκε, στην ανακατασκευή του χώρου των καταστάσεων η ανακατασκευασμένη χρονοσειρά είναι η προβολή κάποιου δυναμικού συστήματος, προβολή που ορίζεται μέσω της συνάρτησης $F: \mathbb{R}^m \rightarrow \mathbb{R}^m$. Όσον αφορά τα προβλήματα πρόβλεψης της χρονοσειράς δε είναι απαραίτητη η εκτίμηση της συνάρτησης F^T αλλά αρκεί ο υπολογισμός της συνιστώσας της $F^T: \mathbb{R}^m \rightarrow \mathbb{R}$ που ορίζει το στοιχείο $x_{i+\tau}$.

Στα τοπικά μοντέλα πρόβλεψης (*local prediction models* ή *local state space models*) η χρονοσειρά αντιμετωπίζεται τμηματικά ή σε παράθυρα μεγέθους m που μοιάζουν με το τελευταίο τμήμα μεγέθους m που έχουμε στη διάθεση μας. Στην περίπτωση που η

χρονοσειρά έχει δημιουργηθεί βασιζόμενοι σε κάποιο αιτιοκρατικό δυναμικό σύστημα τότε τα όμοια τμήματα θα έχουν και όμοιες εικόνες, δηλαδή οι παρατηρήσεις των δύο τμημάτων θα είναι πολύ κοντά (πχ το x_{j+1} θα είναι πολύ κοντά στο x_{n+1}). Κατ' αυτό τον τρόπο γίνεται και η πρόβλεψη του x_{n+1} .

Στον ανακατασκευασμένο χώρο καταστάσεων \mathbb{R}^m τα τμήματα της χρονοσειράς αντιστοιχούν σε γειτονικά σημεία του $x_n = [x_n, x_{n-1}, \dots, x_{n-m+1}]'$, δηλαδή του τελευταίου ανακατασκευασμένου σημείου από τη χρονοσειρά $\{x_1, \dots, x_n\}$ που αντιστοιχεί στη χρονική στιγμή n . Έστω $x_n(1)$ το κοντινότερο σημείο στο x_n που βρέθηκε για χρόνους μικρότερους του n . Έτσι η πρόβλεψη μηδενικής τάξης του x_{n+1} θα είναι είναι:

$$\hat{x}_{n+1} \equiv x_n(1) = x_{n(1)+1}$$

Η πρόβλεψη συνεχίζεται πραγματοποιώντας το μέσο όρο απεικονίσεων (*local average mapping, LAM*) σε K γειτονικά σημεία δηλαδή

$$x_n(1) = \frac{1}{K} \sum_{j=1}^K x_{n(j)+1}$$

Ο μέσος όρος πρόβλεψης μπορεί να είναι σταθμισμένος μέσο όρο, θεωρώντας ως βάρος την απόσταση του κάθε γειτονικού σημείου από το x_n .

Ένα πιο πολύπλοκο τοπικό μοντέλο πρόβλεψης από το προηγούμενο του μέσου όρου απεικονίσεων είναι το τοπικό γραμμικό μοντέλο (*local linear model, LLM*). Σε αυτό η εικόνα του x_{n+1} ως γραμμικό συνδυασμό των τελευταίων m στοιχείων x_{n-m+1}, \dots, x_n δλδ:

$$x_{n+1} = \varphi_0^{(n)} + \varphi_1^{(n)} x_n + \dots + \varphi_m^{(n)} x_{n-m+1}$$

Στον τύπο αυτό οι παράμετροι $\varphi_i^{(n)}$ δεν είναι σταθεροί αλλά εκτιμήσεις για κάθε χρονική στιγμή n . Κατ' αυτό τον τρόπο το τοπικό γραμμικό μοντέλο πρόβλεψης προσομοιώνει την συμπεριφορά ενός γραμμικού μοντέλου AR(m) μοντέλου με βασική διαφορά τόσο την έλλειψη του θορύβου όσο και το γεγονός της χρήσης K γειτονικών δεδομένων του αρχικού δείγματος στον ανακατασκευασμένο χώρο καταστάσεων \mathbb{R}^m .

Στην περίπτωση ενός τοπικού γραμμικού μοντέλου η εκτίμηση των παραμέτρων γίνεται μέσω της μεθόδου των ελαχίστων τετραγώνων ενώ η πρόβλεψη του x_{n+1} δίνεται από το γραμμικό μοντέλο με τις εκτιμήσεις των παραμέτρων $\varphi_0^{(n)}, \varphi_1^{(n)}, \dots, \varphi_m^{(n)}$. Ανάλογες είναι

και οι πιο μακροχρόνιες προβλέψεις T χρονικών βημάτων στις οποίες λειτουργούμε με 2 τρόπους: είτε αλλάζοντας στα μοντέλα πρόβλεψης τις εικόνες των γειτονικών σημείων για ένα χρονικό βήμα, $x_{n(j)+1}$, με τις εικόνες για T χρονικά βήματα, $x_{n(j)+T}$, είτε κάνοντας την πρόβλεψη επαναληπτικά προβλέποντας κάθε φορά ένα χρονικό βήμα μπροστά.

Το βήμα της αξιολόγησης της απόδοσης των μοντέλων πρόβλεψης ακολουθεί την ίδια διαδικασία των γραμμικών μοντέλων. Το λάθος πρόβλεψης μετριέται από το $nrmse$ και αν η χρονοσειρά $x\{t\}$ δεν είναι στάσιμη, γίνεται πρώτα στάσιμη μέσω των ίδιων διαδικασιών που χρησιμοποιούνται στο μοντέλο AR (τις πρώτες διαφορές ή τις αποδόσεις) διαδικασίες. Η πρόβλεψη σε αυτήν την περίπτωση γίνεται όπως στη διαδικασία Box-Jenkins. Έτσι για την πρόβλεψη της μη-στάσιμης $y\{t\}$ γίνεται πρώτα η πρόβλεψη στη στάσιμη χρονοσειρά $x\{t\}$ και η πρόβλεψη για τη χρονική στιγμή n και ένα βήμα μπροστά σε αυτή τη χρονοσειρά, $x_n(1)$, προστίθεται στην τιμή της αρχικής χρονοσειράς για χρόνο n , y_n , για να πάρουμε την πρόβλεψη $y_n(1) = y_n + x_n(1)$.

Η επιλογή της παράμετρο της διάστασης εμβάθυνσης m επηρεάζει άμεσα την πρόβλεψη τόσο με τα τοπικά γραμμικά μοντέλα όσο και με τα μη-γραμμικά μοντέλα, σε σημαντικότερο βαθμό μάλιστα από την επίδραση της τάξης p στα μοντέλα AR, όπου για μεγαλύτερη τάξη οι προβλέψεις δεν αλλάζουν σημαντικά. Επίσης σημαντικό ρόλο παίζει η επιλογή του αριθμού των γειτονικών σημείων K εφόσον στην τοπική πρόβλεψη μέσου όρου και την πρόβλεψη με τοπικά γραμμικά μοντέλα επηρεάζει άμεσα την απόδοση του μοντέλου. Ανεξάρτητα από την καταλληλότητα του μοντέλου και την επιλογή των παραμέτρων η πρόβλεψη χαοτικών χρονοσειρών έχει νόημα για μικρό χρονικό ορίζοντα που ορίζεται από τον μέγιστο εκθέτη Lyapunov.

2.7. Αστάθεια χρονοσειράς

Συχνό φαινόμενο στις χρονοσειρές είναι η αυξομείωση των τιμών της γεγονός που ονομάζεται ως *μεταβλητότητα* ή *αστάθεια* (*volatility*). Η αστάθεια ορίζεται ως η διασπορά της απόδοσης της παρατηρούμενης μεταβλητής σε κάθε χρονική στιγμή t , σ_t^2 (ή αντίστοιχα η τυπική απόκλιση σ_t). Στις στάσιμες χρονοσειρές η διασπορά είναι σταθερή και για τον λόγο αυτό δεν μιλάμε για αστάθεια. Στις πραγματικές εφαρμογές όμως, όπως στις αποδόσεις χρηματο-οικονομικών δεικτών, η διασπορά είναι μεταβαλλόμενη, παρ' όλο

που ασυμπτωτικά (σε μεγάλο χρονικό ορίζοντα) συγκλίνει σε κάποια τιμή (δεν αυξάνει με τη χρονική κλίμακα).

Υποθέτοντας πως η μέση τιμή της χρονοσειράς απόδοσης ενός δείκτη $\{x\}_{t=1}^N$ είναι 0, η αστάθεια σε κάθε χρόνο t μπορεί να εκτιμηθεί στιγμιαία από το τετράγωνο της απόδοσης ως $s_t^2 = x_t^2$ ή με αναφορά στην τυπική απόκλιση από την απόλυτη τιμή της ως $s_t = |x_t|$. Η στιγμιαία αστάθεια της χρονοσειράς μπορεί να μην εμφανίζεται τυχαία. Οι αποδόσεις χρηματιστηριακών δεικτών μπορεί να μην έχουν συσχετίσεις αλλά τα τετράγωνα τους να έχουν.

Η αστάθεια συνήθως δεν εκτιμάται από μια χρονική στιγμή, αλλά σε ένα παράθυρο χρόνου. Έτσι αν το δείγμα για την χρονοσειρά ενός δείκτη μετριέται ανά λεπτό, η αστάθεια σε κάθε λεπτό μπορεί να εκτιμάται από τη διασπορά την τελευταία ώρα, ενώ αν ο δείκτης είναι ημερήσιος, η αστάθεια για κάθε μέρα μπορεί να ορίζεται με τη διασπορά την τελευταία βδομάδα (5 εργάσιμες μέρες) ή τον τελευταίο μήνα (προσεγγιστικά τις τελευταίες 25 μέρες). Έτσι σχηματίζεται η χρονοσειρά της αστάθειας του δείκτη πάνω στην οποία γίνεται και η πρόβλεψη για τις επόμενες χρονικές στιγμές.

Με τα τοπικά μοντέλα, με την κατάλληλη επιλογή του πλήθους των γειτονικών σημείων K μπορούμε να πετύχουμε καλύτερες προβλέψεις της αστάθειας του δείκτη από τα μοντέλα AR.

3. Η ΧΡΗΣΗ ΤΗΣ ΓΛΩΣΣΑΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ Python ΣΤΙΣ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΕΣ ΣΥΝΑΛΛΑΓΕΣ

Η ραγδαία ανάπτυξη της τεχνολογίας έχει επηρεάσει έντονα τα τελευταία χρόνια τους περισσότερους οικονομικούς τομείς και επιστήμες. Η εισαγωγή και χρήση εξειδικευμένων τεχνολογικών εφαρμογών έχει ως αποτέλεσμα τόσο την βελτίωση της υλικής, κοινωνικής, πολιτιστικής και πνευματικής ζωής του ανθρώπου, όσο και την ανάπτυξη και ευημερία διαφόρων τομών της οικονομίας.

Στον χρηματιστηριακό και τραπεζικό τομέα η ανάπτυξη της τεχνολογίας αποτελεί ένα σημαντικό ανταγωνιστικό πλεονέκτημα σε σχεδόν όλα τα χρηματοπιστωτικά ιδρύματα. Το γεγονός αυτό είναι και ο λόγος που τα περισσότερα χρηματοπιστωτικά ιδρύματα επενδύουν στην τεχνολογία μεγάλα κεφάλαια και μάλιστα σε ετήσια βάση (Hilpisch). Σύμφωνα με δήλωση του Crosman του 2013, *“Οι τράπεζες θα ξοδέψουν 4,2% περισσότερα χρήματα για τεχνολογία το 2014 από ό, τι το 2013, σύμφωνα με αναλυτές της IDC²². Οι συνολικές δαπάνες πάνω σε τεχνολογίες της πληροφορικής στις χρηματοπιστωτικές υπηρεσίες παγκοσμίως θα υπερβούν τα 430 δισ. \$ το 2014 και θα ξεπερνούν τα 500 δισεκατομμύρια δολάρια μέχρι το 2020, αναφέρουν οι αναλυτές.”*

Από την άλλη, σύμφωνα με ανακοίνωση της Deutsche Börse Group του 2008 *“Οι τεχνολογικές καινοτομίες συνέβαλαν στη βελτίωση της αποτελεσματικότητας της ανταγωνιστικής αγοράς. Μέσα από καινοτομίες στην τεχνολογία των συναλλαγών, οι συναλλαγές στο Eurex εκτελούνται σήμερα πολύ ταχύτερα από ό, τι πριν από δέκα χρόνια, παρά την έντονη αύξηση του όγκου συναλλαγών και τον αριθμό των προσφορών ... αυτές οι σημαντικές βελτιώσεις οφείλονται στις συνεχείς και υψηλές επενδύσεις σε τεχνολογίες πληροφορικής από γραφεία ανταλλαγών παραγώγων και εκκαθαριστικών.”*

Αυτή η ραγδαία αύξηση του όγκου των συναλλαγών έχει επιφέρει και αρκετά αρνητικά, εφόσον από μόνη της αυξάνει τον κίνδυνο και καθιστά την διαχείριση του ολοένα και πιο

²² Η International Data Corporation (IDC) που ιδρύθηκε το 1964 είναι ένας πάροχος πληροφοριών αγοράς, συμβουλευτικών υπηρεσιών και εκδηλώσεων για τις αγορές της τεχνολογίας των πληροφοριών, των τηλεπικοινωνιών και των καταναλωτικών τεχνολογιών. Είναι θυγατρική της International Data Group (IDG), η κορυφαία εταιρεία παροχής υπηρεσιών πληροφορικής και μάρκετινγκ παγκοσμίως.

περίπλοκη. Χαρακτηριστικό παράδειγμα είναι το λεγόμενο flash crash²³ του Μαΐου 2010, όπου η αυτοματοποιημένη πώληση οδήγησε σε μεγάλες ενδοημερήσιες πτώσεις σε ορισμένα χρηματιστήρια και δείκτες μετοχών.

Ένα επιπλέον πλεονέκτημα, λοιπόν, της χρήσης νέων τεχνολογιών και πληροφορικής από την χρηματοπιστωτική βιομηχανία είναι η ταχύτητα και η συχνότητα με τις οποίες οι συναλλαγές αποφασίζονται και εκτελούνται. Στο (Lewis, 2014) ο Lewis περιγράφει το λεγόμενο *flash trading*, το είδος συναλλαγών δηλαδή με τις υψηλότερες δυνατές ταχύτητες σε πραγματικό χρόνο. Το flash trading έχει σαν αποτέλεσμα την *αύξηση των διαθέσιμων δεδομένων* σε ολόένα και μικρότερες κλίμακες και κατ' επέκταση καθιστά αναγκαία την *αντίδραση σε πραγματικό χρόνο*. Έτσι η αύξηση της ταχύτητας και της συχνότητας των συναλλαγών οδηγεί αναπόφευκτα και στην αύξηση του όγκου δεδομένων αλλά και στην μείωση της μέσης χρονική κλίμακας των χρηματοπιστωτικών συναλλαγών. Αρκεί να σκεφτούμε ότι για 30 χρόνια λειτουργίας ενός χρηματιστηρίου τα δεδομένα μιας μετοχής, πάνω στα οποία στηρίζεται και το μεγαλύτερο μέρος της θεωρίας της οικονομίας²⁴, αγγίζουν περίπου τους 7.500 δείκτες.

Με βάση τα προηγούμενα λοιπόν το ζητούμενο από τις νέες τεχνολογίες είναι η *επεξεργασία των δεδομένων* κατά προτίμηση σε πραγματικό χρόνο και η *ταχύτητα ανάλυσης* ενός τεράστιου όγκου των τελευταίων. Όσον αφορά την χρήση οικονομικών θεωριών και προβλέψεων πολλά θα πρέπει να γίνουν ακόμη εφόσον οι θεωρίες που χρησιμοποιούνται είναι μεν εύρωστες αλλά και ελλιπείς. Η ανάλυση των οικονομικών στοιχείων και δεδομένων είναι ένας συνδυασμός λογισμικού και τεχνολογίας μαζί με αλγόριθμους και μεθόδους με κύριο σκοπό την συλλογή, επεξεργασία και ανάλυση δεδομένων προκειμένου να αποκτηθούν γνώσεις, να ληφθούν αποφάσεις και να εκπληρωθούν τυχόν κανονιστικές απαιτήσεις.

Όσον αφορά την επεξεργασία και ανάλυση δεδομένων δύο είναι τα βασικά προβλήματα που πρέπει να αντιμετωπιστούν: ο μεγάλος όγκος δεδομένων (ή μεγάλα δεδομένα – Big Data) και η ανάλυση τους σε πραγματικό χρόνο.

²³ Ένα flash crash είναι μια πολύ γρήγορη, βαθιά και πτητική πτώση σε εξαιρετικά σύντομο χρονικό διάστημα. Συχνά προέρχεται από συναλλαγές με τεχνική black-box, σε συνδυασμό με συναλλαγές υψηλών ταχυτήτων των οποίων η ταχύτητα και η διασύνδεση μπορούν να έχουν ως αποτέλεσμα την απώλεια εκατομμυρίων δολαρίων σε λεπτά ή και δευτερόλεπτα.

²⁴ Για παράδειγμα η σύγχρονη θεωρία χαρτοφυλακίου (MPT), το μοντέλο τιμολόγησης ενεργητικού κεφαλαίου (CAPM) και η αξία σε κίνδυνο (VaR) έχουν όλα τα θεμέλιά τους στα καθημερινά στοιχεία των τιμών των μετοχών

Στην συνέχεια του κεφαλαίου αναλύονται διεξοδικότερα τόσο οι όροι των μεγάλων δεδομένων όσο και της επεξεργασίας αυτών σε συνδυασμό με διάφορες τεχνολογικές εφαρμογές.

3.1 Λογισμικό Ανοικτού Κώδικα

Το λογισμικό ανοιχτού κώδικα (Ελεύθερο Λογισμικό / Λογισμικό Ανοικτού Κώδικα - ΕΛ/ΛΑΚ) που επινοήθηκε ως όρος το 1998 αποτελεί την νέα σύγχρονη τάση στις νέες τεχνολογίες και την πληροφορική. Ο όρος αναφέρεται σε κάτι το οποίο δέχεται τροποποίησης και διανομής λόγω του ότι ο σχεδιασμός του είναι ανοικτός στο κοινό. Επίσης προέρχεται από μια συγκεκριμένη προσέγγιση για τη δημιουργία προγραμμάτων ηλεκτρονικών υπολογιστών. Γενικότερα ο όρος ανοικτού κώδικα υποδηλώνει ένα ευρύτερο σύνολο αξιών - αυτό που αποκαλείται “τρόπος ανοιχτού κώδικα”. Οι εργασίες αλλά και τα προϊόντα του υιοθετούν μια λογική ανοικτής ανταλλαγής, ομαδικής συμμετοχής, ταχείας παραγωγής πρωτοτύπων, διαφάνειας, αξιοκρατίας και ανάπτυξης με γνώμονα την κοινότητα.

Παραδείγματα λογισμικού ανοιχτού κώδικα είναι το MediaWiki (MediaWiki, 2016) βάσει του οποίου λειτουργεί η Wikipedia, το OpenStreetMap (OpenStreetMap, 2004), το SUGAR (SugarCRM, 2004), το Moodle (moodle, 1999), το UBUNTU/LINUX (Linux, 1983), το OLPC XO (OLPC, 2008), το Arduino (Arduino, 2006), το Raspberry Pi (Raspberry, 2012) και το REPRAP (RepRap, 2005). Η πιο ενδιαφέρουσα σύγχρονη εφαρμογή ανοικτού λογισμικού είναι το επονομαζόμενο (ηλεκτρονικό) υπολογιστικό νέφος το οποίο και αναπτύσσεται με ταχείς ρυθμούς.

3.1.1 Ορισμός ΕΛ/ΛΑΚ και άδειες χρήσης

Σύμφωνα με την επίσημη σελίδα του Open Source Initiative (Ope18), μιας πρωτοβουλίας που διανύει τον εικοστό της χρόνο, το λογισμικό ανοικτού κώδικα (σε αντίθεση με το κλασσικό λογισμικό κλειστού κώδικα όπως το Microsoft Office και το Adobe Reader) δίνει στο χρήστη την δυνατότητα να έχει πρόσβαση στον πηγαίο κώδικα του λογισμικού

και να παρέμβει (τροποποιήσει και διανέμει) πάνω σε αυτόν προσθέτοντας του έτσι νέες λειτουργίες ή αντιμετωπίζοντας κάποιες από τις προβληματικές που μπορεί να παρουσιάζει. Τόσο το LibreOffice και το πρόγραμμα χειρισμού εικόνων GNU αποτελούν παραδείγματα λογισμικού ανοιχτού κώδικα.

Σύμφωνα με το μανιφέστο του 1986 των κατασκευαστών του ελεύθερου λειτουργικού συστήματος GNU (1992), το ελεύθερο λογισμικό σημαίνει ότι στους χρήστες ενός προγράμματος παρέχονται τέσσερις θεμελιώδεις ελευθερίες:

- Η ελευθερία εκτέλεσης του προγράμματος από μέρους των χρηστών, για οποιονδήποτε σκοπό (ελευθερία 0).
- Η ελευθερία τόσο της μελέτης του προγράμματος όσο και της προσαρμογής του στις ιδιαίτερες ανάγκες του κάθε χρήστη (ελευθερία 1). Η πρόσβαση στον πηγαίο κώδικα είναι απαραίτητη προϋπόθεση γι' αυτό.
- Η ελευθερία αναδιανομής αντιγράφων με σκοπό την διευκόλυνση άλλων χρηστών (ελευθερία 2).
- Η ελευθερία τροποποίησης και βελτίωσης προγραμμάτων με την σύγχρονη αναδημοσίευση των αλλαγών στο ευρύ κοινό, ώστε να επωφελείται ολόκληρη η κοινότητα (ελευθερία 3). Η πρόσβαση στον πηγαίο κώδικα είναι απαραίτητη προϋπόθεση γι' αυτό.

Τα λογισμικά open source διαθέτουν όρους αδειάς και χρήσης, με διάφορα επίπεδα δικαιωμάτων, οι οποίοι όροι όμως διαφέρουν από τους όρους ιδιοκτησίας που περιλαμβάνουν τα λογισμικά κλειστού κώδικα. Πιο συγκεκριμένα, οι άδειες των open source (Creative Commons και GNU GPL-Compatible Free Software Licenses) καθορίζουν τον τρόπο με τον οποίο οι χρήστες τους μπορούν να χρησιμοποιήσουν, να τροποποιήσουν και να αναδιανέμουν το λογισμικό. Αυτές οι άδειες δεν επηρεάζουν και δεν επιβάλλουν τον τρόπο και σκοπό χρήσης του λογισμικού.

Ένα εκ των ειδών αδειάς είναι οι επονομαζόμενες άδειες GPL ή “copyleft” σύμφωνα με τις οποίες ο χρήστης που αναδιανέμει ένα τροποποιημένο πρόγραμμα ανοιχτού κώδικα είναι υποχρεωμένος να θέσει ελεύθερο προς χρήση και τροποποίηση τον πηγαίο κώδικα.

Γενικότερα και σύμφωνα με Οδηγίες για το Ελεύθερο Λογισμικό του Debian (DFSG) στο (Organization, 2004) και το FSF (Free Software Foundation) (Fre18), οι άδειες διανομής του λογισμικού ανοιχτού κώδικα θα πρέπει να πληρούν διάφορα κριτήρια όπως:

1. Ελεύθερη Αναδιανομή,
2. Πρόσβαση στον πηγαίο κώδικα,
3. Άδεια τροποποίησης και παράγωγα έργα,
4. Ακεραιότητα του πηγαίου κώδικα,
5. Δεν υπάρχουν διακρίσεις εις βάρος προσώπων ή ομάδων,
6. Δεν υπάρχουν διακρίσεις εις βάρος των τομέων χρήσης,
7. Διανομή άδειας,
8. Η άδεια χρήσης δεν πρέπει να είναι συγκεκριμένη για ένα προϊόν
9. Η άδεια χρήσης δεν πρέπει να περιορίζει άλλο λογισμικό
10. Η άδεια χρήσης πρέπει να είναι τεχνολογία-ουδέτερη

3.1.2 Πλεονεκτήματα χρήσης ΕΛ/ΛΑΚ από τις επιχειρήσεις

Η χρήση των ΕΛ/ΛΑΚ επιφέρει έναν σύνολο πλεονεκτημάτων στους φορείς που τα εφαρμόζουν. Ανάμεσα σε αυτά τα πλεονεκτήματα μπορούν να αναφερθούν τα παρακάτω:

- Το ΕΛ/ΛΣΚ είναι ένα απόλυτα νόμιμο λογισμικό η διάθεση του οποίου πραγματοποιείται χωρίς χρηματική επιβάρυνση (ή με ένα ελάχιστο κόστος) και έχει πολύ μικρότερο κόστος συντήρησης σε σχέση με τα λογισμικά κλειστού κώδικα. Αυτό συμβαίνει γιατί ο χρήστης δεν εξαρτάται από εταιρίες ανάπτυξης λογισμικού. Αντίθετα τόσο η ανάπτυξη όσο η διανομή αλλά και η διόρθωση σφαλμάτων πραγματοποιείται από μια οποιαδήποτε καταρτισμένη ομάδα με αποτέλεσμα να δημιουργείται ένα περιβάλλον ανταγωνισμού και κατ' επέκταση στην παροχή υψηλών υπηρεσιών υποστήριξης σε χαμηλές.
- Παρέχει τη δυνατότητα μετατροπής του και προσαρμογής του στις ιδιαίτερες ανάγκες του κάθε χρήστη
- Είναι εξ' ορισμού ένα λογισμικό που ενσωματώνει πολύ γρήγορα τις πιο σύγχρονες τεχνολογικές εξελίξεις εφόσον επιτρέπει συχνές ενημερώσεις και νέες δυνατότητες
- Λόγο της δυνατότητας μετατροπής του παρέχεται στους χρήστες ένας μεγάλος αριθμός πρόσθετων προγραμμάτων με εύκολη και δωρεάν εγκατάσταση

- Παρέχει σταθερότητα και ασφάλεια εφόσον είναι ασφαλέστερο αλλά και πιο αξιόπιστο σε σχέση με το λογισμικό κλειστού κώδικα που παρέχεται στο διαδίκτυο.
- Τόσο στην χρήση όσο και στην τροποποίηση του υποστηρίζεται άμεσα από μια μεγάλη κοινότητα ανάπτυξης και υποστήριξης του λογισμικού

Το ΕΛ/ΛΑΚ περιλαμβάνει μια μεγάλη κοινότητα χρηστών και προγραμματιστών, η συνεργασία των οποίων οδηγεί σε μια συνεχή αλλά και ταχεία κυκλοφορία βελτιωμένων εκδόσεων του λογισμικού. Αυτό επίσης οδηγεί σε μια αμεσότητα επέμβασης από μέρους των προγραμματιστών σε περιπτώσεις εμφάνισης προβλημάτων ή κενών ασφαλείας.

Επίσης το λογισμικό είναι εύκολα προσβάσιμο εφόσον διατίθεται μέσω Διαδικτύου, γεγονός που βοήθησε και στην γρήγορη διάδοση του. Σημαντικό να σημειωθεί ότι η ανάπτυξη και η λειτουργία της νέας γενιάς Διαδικτύου βασίζεται, κατά κύριο λόγο, σε ΕΛ/ΛΑΚ.

Εφαρμογές Λογισμικό Ανοιχτού Κώδικα, με αξιόπιστα και σταθερά εργαλεία αλλά και με ελάχιστο κόστος εγκατάστασης και αναβάθμισης, χρησιμοποιούνται σε διάφορους τομείς όπως η εκπαίδευση (π.χ. η εφαρμογή Open E-Class), η δημόσια διοίκηση αλλά και οι επιχειρήσεις. Ειδικά όσον αφορά τις σύγχρονες επιχειρήσεις, τα τελευταία χρόνια έχουν συμπεριλάβει στις λειτουργίες τους εφαρμογές που αφορούν τόσο λειτουργικά συστήματα (υπολογιστών και εξυπηρετητών), ενδιάμεσο λογισμικό (middleware), υπολογιστικό νέφος (cloud), αλλά και τηλεφωνία και δικτυακούς τόπους. Τα πλεονεκτήματα που οι εφαρμογές επιφέρουν μέσω της χρήσης τους σε μια επιχείρηση είναι:

- **Ευελιξία λύσεων**

Σε περίπτωση προβλημάτων, ο ανοιχτός κώδικας, λόγω της φύσης του, επιτρέπει μια μεγαλύτερη ευελιξία λύσεων. Οι λύσεις αυτές έχουν το χαρακτηριστικό ότι υλοποιούνται πολύ πιο γρήγορα αλλά και σε μια μεγαλύτερη γκάμα σε σχέση με το λογισμικό που υλοποιείται από εταιρίες. Στις ευέλικτες αυτές λύσεις περιλαμβάνονται εφαρμογές cloud, η εγκατάσταση Red Hat Enterprise Linux ή η βάση δεδομένων MongoDB.

➤ **Ταχύτητα εφαρμογής λύσεων**

Όπως προαναφέρθηκε, οι λύσεις που παρέχονται από λογισμικό ελεύθερου κώδικα χαρακτηρίζονται από την ταχύτητα υλοποίησης και εφαρμογής τους λόγω της ύπαρξης κοινοτήτων. Στις επιχειρήσεις παρέχεται η δυνατότητα άμεσης εγκατάστασης του λογισμικού που έχει ως αποτέλεσμα την άμεση προσθήκη αξίας στην επιχείρηση.

➤ **Εφαρμογή σε όλα τα επίπεδα ανάπτυξης μιας επιχείρησης**

Η εφαρμογή μιας λύσης ανοιχτού κώδικα μπορεί να εφαρμοστεί ανεξάρτητα από το επίπεδο ανάπτυξης της επιχείρησης, ακόμα και αν αυτό είναι πολύ χαμηλό. Μεταγενέστερα και με την ανάπτυξη της επιχείρησης αυτή μπορεί να υιοθετήσει λύσεις πιο εμπορικές ή να συνεχίσει με το λογισμικό ανοικτού κώδικα.

➤ **Χαμηλό κόστος/υψηλές υπηρεσίες**

Βασικό πλεονέκτημα του ελεύθερου λογισμικού είναι το μηδαμινό ή ελάχιστο κόστος σε σχέση με αυτό των αντίστοιχων εμπορικών λογισμικών. Έτσι νεοσυσταθείσες ή μικρές επιχειρήσεις έχουν την δυνατότητα να αναζητήσουν και να υιοθετήσουν λύσεις ανοιχτού κώδικα, που ικανοποιούν πλήρως τις ανάγκες τους.

➤ **Ασφάλεια πληροφοριών**

Ένα επιπλέον πλεονέκτημα που παρέχει η υιοθέτηση της λύσης του ελεύθερου λογισμικού είναι ότι η κοινότητα ανοιχτού κώδικα μπορεί άμεσα να εντοπίσει και να επιλύσει προβλήματα ασφαλείας, γεγονός που δεν είναι δυνατό με την χρήση εμπορικού λογισμικού.

➤ **Ελάχιστο κόστος συντήρησης**

Η κοινότητα του ελεύθερου λογισμικού πραγματοποιεί μια συνεχή συντήρηση και βελτίωση του υπάρχοντος κώδικα. Τα κόστη συντήρησης και ανάπτυξης των εφαρμογών μοιράζεται με τη κοινότητα.

3.2 Η γλώσσα προγραμματισμού Python

Η γλώσσα προγραμματισμού Python, η οποία χρησιμοποιείται στην παρούσα πτυχιακή για την υλοποίηση μιας εφαρμογής χρηματοοικονομικού ενδιαφέροντος, είναι μια γλώσσα ανοικτού κώδικα, υψηλού επιπέδου²⁵ και με ένα πολύ μεγάλο αριθμό εφαρμογών σε διάφορους τομείς. Στην επίσημη ιστοσελίδα (2018) δίνεται, για την γλώσσα, η παρακάτω περιγραφή:

“Η Python είναι μια εύκολη στην εκμάθηση, ισχυρή γλώσσα προγραμματισμού. Έχει αποδοτικές δομές δεδομένων υψηλού επιπέδου και μια απλή αλλά αποτελεσματική προσέγγιση στον αντικειμενοστραφή προγραμματισμό. Η κομψή σύνταξη της Python και οι δυναμικοί τύποι της, μαζί με τη λειτουργία της ως διερμηνευόμενη (αντί μεταγλωττιζόμενης) γλώσσα, την καθιστούν την ιδανική γλώσσα για δημιουργία σεναρίων εντολών και για ταχεία ανάπτυξη εφαρμογών σε πολλούς τομείς και στις περισσότερες πλατφόρμες.”

Τα τελευταία χρόνια η γλώσσα έχει βρει μεγάλη ανταπόκριση και έχει υιοθετηθεί στην υλοποίηση πολλών εφαρμογών σε πλατφόρμες όπως η Google, το YouTube, το DropBox, το Facebook, το Pinterest, το Instagram, το Spotify μέχρι την NASA. Σύμφωνα με το (Aby18) η επιτυχία της γλώσσας οφείλεται στα βασικά χαρακτηριστικά της τα οποία και είναι:

➤ Απλότητα και ευκολία εκμάθησης

Η σύνταξη της Python παρουσιάζει μια μεγάλη ομοιότητα με εντολές ψευδοκώδικα, γεγονός που την καθιστά πραγματικά απλή και εύκολη στην εκμάθηση. Σκοπός είναι η επικέντρωση στην λύση ενός προβλήματος αντί στην ίδια τη γλώσσα.

➤ Λογισμικό ΕΛ/ΛΑΚ

Η Python είναι ένα ΕΛ/ΛΑΚ και ως εκ τούτου παρουσιάζει όλα τα χαρακτηριστικά των τελευταίων. Έτσι τα αντίγραφα της διανέμονται ελεύθερα, με δυνατότητα παρεμβάσεων στον πηγαίο κώδικα. Επίσης διαθέτει την δική της κοινότητα που εργάζεται πάνω στην γλώσσα και τις μετατροπές της.

²⁵ Γλώσσα υψηλού επιπέδου: είναι γλώσσες προγραμματισμού απλές και ανεξάρτητες από συγκεκριμένο τύπο υπολογιστή

➤ **Γλώσσα προγραμματισμού υψηλού επιπέδου (high-level programming language)**

Οι γλώσσες προγραμματισμού υψηλού επιπέδου είναι γλώσσες απλές και ανεξάρτητες από κάποιο συγκεκριμένο τύπο υπολογιστή. Τα προγράμματα και οι εφαρμογές που υλοποιούνται σε γλώσσες υψηλού επιπέδου μπορούν να εκτελεστούν και σε διαφορετικούς υπολογιστές από αυτούς στους οποίους δημιουργήθηκαν, χαρακτηριστικό που ονομάζεται **μεταφερσιμότητα**. Λόγω του ότι ο προγραμματισμός δεν γίνεται σε γλώσσα μηχανής αλλά με την εκτέλεση των προγραμμάτων μέσω μεταφραστών ή μεταγλωττιστών, ο δημιουργός δεν προβληματίζεται με λεπτομέρειες χαμηλού επιπέδου όπως ο χώρος και η θέση αποθήκευσης μιας μεταβλητής κτλ.

➤ **Γλώσσα φορητή**

Η Python έχει υλοποιηθεί με τέτοιο τρόπο ώστε να λειτουργεί στις περισσότερες πλατφόρμες και λειτουργικά συστήματα. Έτσι το ίδιο πρόγραμμα σε Python μπορεί να εκτελεστεί σωστά τόσο στο Linux, στα Windows, στα Macintosh και σε ένα μεγάλο αριθμό λειτουργικών συστημάτων!

➤ **Γλώσσα διερμηνευόμενη**

Η φάση εκτέλεσης κάθε πρόγραμμα, υλοποιημένο πηγαία γλώσσα σε γλώσσα μηχανής (δυναδικός κώδικας δηλαδή ακολουθίες 0 και 1). Η παραπάνω ενδιάμεση διαδικασία ονομάζεται μεταγλώττιση και πραγματοποιείται από προγράμματα όπως ο μεταγλωττιστής ή ο διερμηνέας με διάφορες σημαίες και επιλογές. Στην περίπτωση της Python αυτό δεν είναι απαραίτητο. Αυτό συμβαίνει γιατί η ίδια η γλώσσα μετατρέπει τον πηγαίο κώδικα σε μια ενδιάμεση μορφή (bytecode), στη συνέχεια το μεταφράζει στη γλώσσα μηχανής και μετά το εκτελεί. Λόγο της έλλειψης χρήσης μεταγλωττιστή, η γλώσσα αποδεικνύεται ακόμα πιο εύκολη στη χρήση της εφόσον ο προγραμματιστής δε χρειάζεται να ασχοληθεί με την μεταγλώττιση του προγράμματος, τη σύνδεση με τις κατάλληλες βιβλιοθήκες, κ.λπ. Η δυνατότητα ενδιάμεσης μετατροπής που παρέχει η γλώσσα καθιστά τα προγράμματα εξαιρετικά φορητά, αφού αυτά θα δουλεύουν το ίδιο σε οποιοδήποτε υπολογιστή.

➤ Γλώσσα Αντικειμενοστραφής

Η Python είναι γλώσσα που μπορεί να χρησιμοποιηθεί τόσο για διαδικασιοστραφή προγραμματισμό (procedure-oriented) όσο και για αντικειμενοστραφή προγραμματισμό (object-oriented). Η διαφορά μεταξύ των δύο ομάδων έγκειται στον τρόπο με τον οποίο αντιμετωπίζεται ο τρόπος προγραμματισμού. Στο *διαδικασιοστραφή προγραμματισμό*, στο πρόγραμμα υλοποιούνται διαδικασίες και συναρτήσεις (επαναχρησιμοποιήσιμα κομμάτια από προγράμματα) ενώ ο *αντικειμενοστραφής* προγραμματισμός, είναι προσανατολισμένος στην δημιουργία αντικειμένων και κλάσεων που μοιράζονται κοινά χαρακτηριστικά και λειτουργικότητα.

➤ Γλώσσα επεκτάσιμη και ενσωματώσιμη

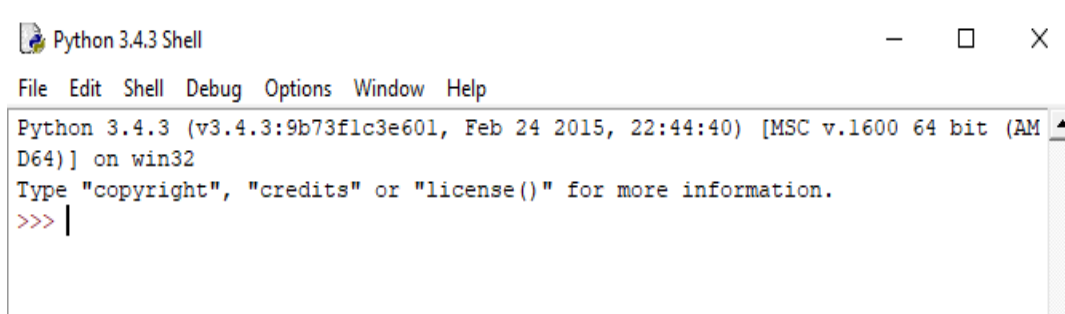
Στα προγράμματα Python μπορούν εύκολα να προσαρμοστούν κομμάτια υλοποιημένα σε μια άλλη γλώσσα προγραμματισμού (C ή C++).

➤ Εκτεταμένες βιβλιοθήκες

Η Πρότυπη βιβλιοθήκη της Python περιέχει ένα τεράστιο αριθμό υλοποιήσεων που αφορά στο μεγαλύτερο μέρος των εφαρμογών. Η βιβλιοθήκη περιλαμβάνει αρχεία σχετικά με κανονικές εκφράσεις, δημιουργία τεκμηρίωσης, δοκιμές μονάδων, νημάτωση, βάσεις δεδομένων, περιηγητές ιστού, CGI, FTP, email, XML, XML-RPC, HTML, αρχεία WAV, κρυπτογράφηση, γραφικές διεπαφές χρήστη (GUI -graphical user interfaces).

Από την γλώσσα προσφέρονται διάφορες επιλογές ως προς το ποιο περιβάλλον προγραμματισμού θα χρησιμοποιηθεί:

➤ Terminal / Shell



```
Python 3.4.3 Shell
File Edit Shell Debug Options Window Help
Python 3.4.3 (v3.4.3:9b73f1c3e601, Feb 24 2015, 22:44:40) [MSC v.1600 64 bit (AMD64)] on win32
Type "copyright", "credits" or "license()" for more information.
>>> |
```

Εικόνα 3.1: Python Shell

- IDLE (προεπιλεγμένο περιβάλλον της Python)



```
L1=[]; L2=[]; max=0
for i in range(3):
    x=input("Δώσε ονοματεπώνυμο")
    L1.append(x)
print("Τα ονοματεπώνυμα είναι",L1)
for i in range(3):
    misthos=float(input("Δώσε μισθό"))
    if misthos>0 and misthos<2500:
        L2.append(misthos)
print("Ο μισθός είναι",L2)
for i in range(3):
    if L2[i]>max:
        max=L2[i]
print(L1[i])
```

Εικόνα 3.2: Το IDLE της Python

- Notebook iPython



Εικόνα 3.3: Το Notebook του Anaconda

3.2.1 Δομές Δεδομένων

Όπως σε όλες τις γλώσσες υψηλού επιπέδου έτσι και στη Python, συμπεριλαμβάνονται διάφορες *δομές δεδομένων* όπως οι λίστες, τα αλφαριθμητικά, οι πλειάδες και τα λεξικά.

Πιο συγκεκριμένα:

- Οι *λίστες* είναι οι πιο δημοφιλείς δομές της Python και αυτό λόγω της ευελιξίας τους. Μια λίστα μπορεί να οριστεί γράφοντας, μέσα σε αγκύλες, μια λίστα τιμών

που χωρίζονται με κόμμα. Δεν παρουσιάζεται κανένας περιορισμός στον τύπο των δεδομένων της λίστας, γεγονός που σημαίνει ότι σε μια λίστα μπορούν να περιέχονται τόσο αριθμητικές τιμές όσο και αλφαριθμητικά, αν και συνήθως τα στοιχεία έχουν όλα τον ίδιο τύπο. Τέλος, οι λίστες Python είναι μεταβλητές και επομένως τα μεμονωμένα στοιχεία μιας λίστας μπορούν να αλλάξουν.

- Οι *συμβολοσειρές* (strings) ορίζονται κάνοντας χρήση μονών ή και διπλών εισαγωγικών (‘ και “ ”), π.χ. ‘Hello Word!’. Ότι περικλείεται μέσα σε εισαγωγικά αντιλαμβάνονται σαν αλφαριθμητικά, δεν είναι δυνατό να μετατραπούν και μπορούν να επεκταθούν σε περισσότερες από μια γραμμές. Στα αλφαριθμητικά μπορεί να γίνει χρήση του ειδικού χαρακτήρα διαφυγής \.
- Οι *πλειάδες* (tuples) ορίζονται από μια ακολουθία τιμών που χωρίζονται μεταξύ τους με κόμμα. Οι πλειάδες είναι σταθερές (αν και μπορούν να διατηρούν μεταβλητά δεδομένα αν χρειαστεί) και στην έξοδο περιβάλλονται από παρενθέσεις, έτσι ώστε οι ένθετες πλειάδες να επεξεργάζονται σωστά. Λόγο του γεγονότος ότι οι πλειάδες είναι σταθερές επεξεργάζονται ταχύτερα από τις λίστες που είναι μεταβλητές.
- Τα *λεξικά* (dictionaries) είναι ένα μη ταξινομημένο σύνολο ζευγών με την σύνταξη κλειδί: αξία μέσα σε άγκιστρα (π.χ. extension = {‘Anna’:1234, ‘Nick’:12345, ‘Jim’:968}). Βασική προϋπόθεση είναι ότι τα κλειδιά είναι μοναδικά μέσα στο ίδιο λεξικό. Ένα ζευγάρι άγκιστρα δημιουργεί ένα κενό λεξικό: {}.

3.2.2 Επιλογή και επανάληψη

Στην Python γίνεται, επίσης, χρήση μεθόδους επιλογής και επανάληψης.

- Η επανάληψη *for* που είναι η πιο κοινή μέθοδος επανάληψης. Η σύνταξη της έχει ως εξής:

```
for i in [list]:  
    __statement__
```

Στην Python δεν χρησιμοποιούνται βρόγχοι ({}) οι οποίοι έχουν αντικατασταθεί από τις εσοχές.

- Η επανάληψη *while* με σύνταξη:

```
while condition:
    __statement__
```

- Η μέθοδος επιλογής *if ...elif ... else* με σύνταξη:

```
if condition1:
    __statement1__
elif condition2:
    __statement2__
elif condition3:
    __statement3__
else condition4:
    __statement4__
```

Όπως έχει ήδη αναφερθεί προηγούμενα η γλώσσα Python υποστηρίζεται από ένα μεγάλο αριθμό βιβλιοθηκών. Η ενσωμάτωση μιας βιβλιοθήκης (π.χ. της βιβλιοθήκης *math* των μαθηματικών συναρτήσεων) μπορεί να γίνει με 2 διαφορετικούς τρόπους:

```
import math as m ή from math import *
```

Στην πρώτη περίπτωση ορίζεται ένα *alias* (ψευδώνυμο) *m* στη βιβλιοθήκη μαθηματικών. Στην συνέχεια μπορεί να γίνει χρήση διάφορων λειτουργιών από τη βιβλιοθήκη μαθηματικών (π.χ. *factorial*) με την αναφορά της χρησιμοποιώντας το *m.factorial()*.

Στην δεύτερη περίπτωση γίνεται εισαγωγή ολόκληρου του χώρου ονομάτων στην βιβλιοθήκη. Έτσι μπορεί να γίνει απευθείας χρήση της συνάρτησης *factorial()* χωρίς να αναφέρετε η βιβλιοθήκη *math* από την οποία μεταφέρεται.

3.2.2 Βιβλιοθήκες

Ανάμεσα σε ένα μεγάλο αριθμό βιβλιοθηκών που συμπεριλαμβάνονται στη γλώσσα θα αναφερθούν διάφορες βιβλιοθήκες που είναι χρήσιμες στην ανάλυση δεδομένων.

- Η βιβλιοθήκη *NumPy* η οποία παίρνει το όνομα της ως ακρώνυμο του *Numerical (Αριθμητική) Python*. Βασικό χαρακτηριστικό της βιβλιοθήκης NumPy είναι οι πίνακες n -διαστάσεων. Στη βιβλιοθήκη περιέχονται επίσης βασικές λειτουργίες και συναρτήσεις της γραμμικής άλγεβρας, οι μετασχηματισμοί Fourier, διάφορες προηγμένες δυνατότητες τυχαίων αριθμών (random numbers) και εργαλεία ενσωμάτωσης με άλλες γλώσσες χαμηλού επιπέδου όπως Fortran, C και C ++.
- Η βιβλιοθήκη *SciPy* παίρνει το όνομα της ως ακρώνυμο του *Scientific (Επιστημονική) Python* και είναι βασισμένη πάνω στην NumPy. Πρόκειται για μία από τις πιο χρήσιμες βιβλιοθήκες με εφαρμογή σε ποικίλες ενότητες της επιστήμης και της μηχανικής υψηλού επιπέδου και συμπεριλαμβάνει λειτουργίες του διακριτού μετασχηματισμού Fourier, της Γραμμικής Άλγεβρας, της Βελτιστοποίησης και των Αραιών Μητρών.
- Εξίσου χρήσιμη αποδεικνύεται και η βιβλιοθήκη *Matplotlib* ως προς το σχεδιασμό μιας τεράστιας ποικιλίας γραφικών παραστάσεων, από τα πιο απλά ιστογράμματα ως τα πιο σύνθετα γραμμικά γραφήματα και γραφικές παραστάσεις.
- Βασική βιβλιοθήκη πάνω στις λειτουργίες δομημένων δεδομένων και στον χειρισμό τους είναι η *Pandas* η οποία χρησιμοποιείται εκτενώς για την αποθήκευση και την προετοιμασία δεδομένων. Οι Pandas προστέθηκαν σχετικά πρόσφατα στην Python και έχουν συμβάλει στην ενίσχυση της χρήσης της Python στην κοινότητα ανάλυσης επιστημονικών δεδομένων.
- Η βιβλιοθήκη *Agate* είναι η πιο νέα βιβλιοθήκη στην επίλυση προβλημάτων ανάλυσης δεδομένων. Αναπτύχθηκε με γνώμονα τη δημοσιογραφία και έχει πολλά εξαιρετικά χαρακτηριστικά για την ανάλυση δεδομένων. Μπορεί να χρησιμοποιηθεί στην ανάλυση και σύγκριση υπολογιστικών φύλλων, στη συλλογή στατιστικών στοιχείων μιας βάσης δεδομένων κτλ.
- Η βιβλιοθήκη *Scikit Learn* η οποία χρησιμοποιείται σε εφαρμογές μηχανικής μάθησης (Machine Learning) και βασίζεται πάνω στις βιβλιοθήκες NumPy, SciPy και matplotlib. Η συγκεκριμένη βιβλιοθήκη περιέχει πολλά χρήσιμα εργαλεία για

τη μηχανική μάθηση και τη στατιστική μοντελοποίηση, συμπεριλαμβανομένης της κατηγοριοποίησης, της παλινδρόμησης, της συσσώρευσης και της μείωσης των διαστάσεων.

- Η βιβλιοθήκη *Statsmodels* η οποία χρησιμοποιείται στη στατιστική μοντελοποίηση και αποτελεί μια ενότητα της Python που επιτρέπει στους χρήστες να διερευνήσουν δεδομένα, να εκτιμήσουν στατιστικά μοντέλα και να εκτελέσουν στατιστικές δοκιμές. Με την χρήση της βιβλιοθήκης είναι διαθέσιμος εκτενής κατάλογος περιγραφικών στατιστικών, στατιστικών δοκιμών, λειτουργιών σχεδίασης και στατιστικών αποτελεσμάτων για διάφορους τύπους δεδομένων και για κάθε εκτιμητή.
- Η βιβλιοθήκη *Seaborn* είναι μια βιβλιοθήκη για τη δημιουργία ελκυστικών και ενημερωτικών στατιστικών γραφικών στην Python και είναι βασισμένη στην βιβλιοθήκη *matplotlib* και στοχεύει να μετατρέψει την απεικόνιση σε κεντρικό μέρος της διερεύνησης και κατανόησης των δεδομένων.
- Η βιβλιοθήκη *Bokeh* χρησιμοποιείται στην δημιουργία διαλογικών διαγραμμάτων, πινάκων ελέγχου και εφαρμογών δεδομένων σε σύγχρονα προγράμματα περιήγησης ιστού. Εξουσιοδοτεί τον χρήστη να παράγει κομψά και συνοπτικά γραφικά στο στυλ του D3.js. Επιπλέον, έχει την ικανότητα διαδραστικότητας υψηλής απόδοσης σε πολύ μεγάλες σειρές δεδομένων ή streaming.
- Η βιβλιοθήκη *Blaze* είναι μια επέκταση της δυνατότητας των βιβλιοθηκών *Numpy* και *Pandas* σε κατανεμημένα και συνεχόμενα σύνολα δεδομένων. Μπορεί να χρησιμοποιηθεί για πρόσβαση σε δεδομένα από πολλές πηγές, όπως *Bigz*, *MongoDB*, *SQLAlchemy*, *Apache Spark*, *PyTables* κλπ. Μαζί με την βιβλιοθήκη *Bokeh*, η *Blaze* μπορεί να λειτουργήσει ως ένα πολύ ισχυρό εργαλείο για τη δημιουργία αποτελεσματικών απεικονίσεων και dashboards σε τεράστια κομμάτια δεδομένων.
- Η βιβλιοθήκη *Scrapy* με χρήση την ανίχνευση στον ιστό. Είναι ένα πολύ χρήσιμο εργαλείο για τη λήψη συγκεκριμένων προτύπων δεδομένων. Έχει την δυνατότητα να ξεκινήσει σε μια ιστοσελίδα με συγκεκριμένο url και στη συνέχεια να εξορύξει πληροφορίες μέσα από τις ιστοσελίδες που αναφέρονται μέσα στην ιστοσελίδα.

- Η βιβλιοθήκη *SymPy* χρησιμοποιείται σε συμβολικούς υπολογισμούς. Έχει ευρείες δυνατότητες από τη βασική συμβολική αριθμητική μέχρι τον υπολογισμό, την άλγεβρα, τα διακριτά μαθηματικά και την κβαντική φυσική. Ένα άλλο χρήσιμο χαρακτηριστικό είναι η δυνατότητα μορφοποίησης του αποτελέσματος των υπολογισμών ως κώδικας LaTeX.
- Η βιβλιοθήκη *Requests* η οποία χρησιμοποιείται για πρόσβαση στο διαδίκτυο. Λειτουργεί παρόμοια με την τυπική βιβλιοθήκη της Python *urllib2*, αλλά είναι πολύ πιο εύκολο να κωδικοποιηθεί.

3.3 Data Science: Python και Big Data

Η επιστήμη των δεδομένων (*Data Science*) είναι ένα μεγάλο πεδίο που καλύπτει τη συλλογή δεδομένων, τον καθαρισμό, την τυποποίηση, την ανάλυση, την απεικόνιση και την αναφορά τους. Ανάλογα μεγάλο είναι και το πεδίο εφαρμογών της επιστήμης των δεδομένων αφού μπορούν να χρησιμοποιηθούν στην ανάλυση μιας γλώσσα, στην εξατομικευμένη πρόταση προβολής ενός βίντεο ή στον προσδιορισμό νέων προϊόντων από δεδομένα πελατών ή μάρκετινγκ. Έτσι τόσο η επιστήμη των δεδομένων όσο και η ανάλυση τους χρησιμοποιούνται ευρέως στην επίλυση προβλημάτων σε τομείς έρευνας, στις επιχειρήσεις και τις εταιρίες.

Ο όρος των *μεγάλων δεδομένων* (*Big Data*) στην επιστήμη των δεδομένων, οφείλεται στην δεδομένων μεγάλης κλίμακας. Το μέγεθος της κλίμακας, φυσικά, εξαρτάται από το πεδίο εφαρμογής. Έτσι τα *Big Data* είναι επίσης δεδομένα αλλά με ένα τεράστιο μέγεθος. Είναι ένας όρος που χρησιμοποιείται για να περιγράψει τη συλλογή ενός τεράστιου αριθμού δεδομένων, αριθμός που αυξάνεται εκθετικά με το χρόνο. Αυτά τα δεδομένα είναι τόσο μεγάλα και πολύπλοκα που κανένα από τα παραδοσιακά εργαλεία διαχείρισης δεδομένων δεν είναι σε θέση να το αποθηκεύσει ή να το επεξεργαστεί αποτελεσματικά. Για παράδειγμα το χρηματιστήριο της Νέας Υόρκης παράγει δεδομένα της τάξης του ενός *terabyte* στην διάρκεια μιας μέρας, το Facebook αποθηκεύει νέα δεδομένα της τάξης των *500+ terabytes* σε μια ημέρα, ενώ μια αεροπορική πτήση παράγει πάνω από *10 terabyte* δεδομένα μόνο σε μισή ώρα.

Μερικοί από τους τομείς που παράγουν Big Data είναι:

- *Δεδομένα μαύρου κουτιού*: δεδομένα που παράγονται από την αντίστοιχη συσκευή ελικοπτέρων, αεροπλάνων και αεριοθούμενων κλπ.
- *Δεδομένα κοινωνικών μέσων*: Τα κοινωνικά μέσα όπως το Facebook και το Twitter κατέχουν πληροφορίες και τις απόψεις που δημοσιεύονται από εκατομμύρια ανθρώπους σε όλο τον κόσμο.
- *Δεδομένα Χρηματιστηρίου*: Τα στοιχεία του χρηματιστηρίου συγκεντρώνουν πληροφορίες σχετικά με τις αποφάσεις αγοράς και πώλησης των μετοχές των διαφόρων εταιρειών που πραγματοποιούν οι πελάτες.
- *Δεδομένα ηλεκτρικής ενέργειας (power grid)*: Τα δεδομένα ηλεκτρονικής ενέργειας διατηρούν πληροφορίες που συγκεντρώνονται από έναν συγκεκριμένο κόμβο σε έναν σταθμό βάσης.
- *Δεδομένα μεταφοράς*: Τα δεδομένα μεταφοράς περιλαμβάνουν το μοντέλο, τη χωρητικότητα, την απόσταση και τη διαθεσιμότητα ενός οχήματος.

Τα μεγάλα δεδομένα μπορούν να ταξινομηθούν σε τρεις κατηγορίες:

- **Δομημένα**: δεδομένα με σταθερό μορφότυπο (γνωστό εκ των προτέρων) που μπορούν να αποθηκευτούν, είναι προσβάσιμα και μπορούν να υποστούν επεξεργασία. Ένας πίνακας μιας σχεσιακής βάσης δεδομένων είναι παράδειγμα δομημένων δεδομένων. Έχουν αναπτυχθεί πολλές δημοφιλείς τεχνικές για την συλλογή αλλά και επεξεργασία των δομημένων δεδομένων. Σήμερα παρουσιάζονται διάφορα ζητήματα πάνω στο μέγεθος τους (μέγεθος που αυξάνεται με τεράστιο ρυθμό) που αγγίζει την τάξη πολλών zettabyte.
- **Μη-δομημένα**: δεδομένα με άγνωστη μορφή ή δομή. Σε αυτή την κατηγορία η δυσκολία δεν άπτεται μόνο στο μέγεθος των δεδομένων που είναι τεράστιο, αλλά και στην επεξεργασία τους έτσι ώστε να αποκομιστεί αξία από αυτά. Παραδείγματα μη δομημένων δεδομένων είναι αρχεία κειμένου, εικόνες, βίντεο, τα αποτελέσματα μιας αναζήτησης σε κάποια μηχανή αναζήτησης κλπ. Ο τομέας παρουσιάζει μεγάλο επιστημονικό ενδιαφέρον εφόσον το μεγαλύτερο μέρος δεδομένων είναι αδόμητο.
- **Ημιδομημένα**: δεδομένα που μπορούν να περιέχουν και τους δύο προηγούμενους τύπους δεδομένων. Τα συγκεκριμένα δεδομένα έχουν δομημένη μορφή, αλλά στην πραγματικότητα δεν καθορίζονται π.χ. μέσω ενός σχεσιακού πίνακα ΣΔΒΔ.

Παράδειγμα ημι-δομημένων δεδομένων είναι δεδομένα που περιέχονται σε ένα αρχείο XML.

Για την ακριβέστερη ανάλυση των μεγάλων δεδομένων, η οποία θα οδηγήσει σε μια πιο επιτυχή λήψη αποφάσεων και επομένως σε μεγαλύτερη λειτουργική αποδοτικότητα, μείωση του κόστους και μειωμένους κινδύνους για μια επιχείρηση, έχουν αναπτυχθεί διάφορες νέες τεχνολογίες. Για την αξιοποίηση των μεγάλων δεδομένων, είναι απαραίτητη μια υποδομή που να μπορεί να διαχειρίζεται και να επεξεργάζεται τεράστιους όγκους δομημένων και αδόμητων δεδομένων σε πραγματικό χρόνο και να προστατεύει την ιδιωτικότητα και την ασφάλεια των δεδομένων αυτών (tutorialspoint, 2018).

Στην σημερινή αγορά παρουσιάζονται διάφορες τεχνολογίες χειρισμού μεγάλων δεδομένων που έχουν εισαχθεί από επιχειρήσεις όπως το Amazon, η IBM, η Microsoft κ.λπ. Οι δύο βασικότερες τεχνολογίες χειρισμού και επεξεργασίας μεγάλων δεδομένων είναι οι ακόλουθες:

- **Operational Big Data:** η κατηγορία περιλαμβάνει συστήματα όπως το MongoDB το οποίο παρέχει επιχειρησιακές δυνατότητες εντοπισμού και αποθήκευσης δεδομένων σε πραγματικό χρόνο και διαδραστικά. Περιλαμβάνουν επίσης συστήματα δεδομένων NoSQL²⁶ Big Data εφόσον οι νέες αρχιτεκτονικές υπολογιστών, επιτρέπουν την εκτεταμένη και αποτελεσματική εκτέλεση μαζικών υπολογισμών με αποτέλεσμα τα λειτουργικά φορτία δεδομένων να διαχειρίζονται πιο εύκολα, φθηνότερα και να μπορούν να υλοποιηθούν ταχύτερα. Επιπλέον κάποια από τα συστήματα NoSQL μπορούν να παρέχουν πληροφορίες για τα πρότυπα και τις τάσεις βάσεις δεδομένων σε πραγματικό χρόνο και μάλιστα με ελάχιστη κωδικοποίηση.
- **Analytical Big Data:** η κατηγορία περιλαμβάνει συστήματα όπως το Massively Parallel Processing (MPP) και το MapReduce που παρέχουν εκτεταμένες δυνατότητες για αναδρομική και σύνθετη ανάλυση που αφορούν τα περισσότερα ή όλα τα δεδομένα. Το MapReduce παρέχει μια νέα μέθοδο ανάλυσης δεδομένων που συμπληρώνει τις δυνατότητες που παρέχονται από την SQL και θα αναλυθεί εκτενέστερα στη συνέχεια.

²⁶ NoSql: μη σχεσιακές βάσεις δεδομένων

Αυτές οι δύο κατηγορίες τεχνολογίας είναι συμπληρωματικές και συχνά αναπτύσσονται μαζί.

Ένα μεγάλο μέρος των εφαρμογών δεν απαιτούν μεγάλα δεδομένα, γεγονός που σημαίνει ότι το σύνολο δεδομένων είναι αρκετά μικρό για να αποθηκευτεί και να αναλυθεί με την χρήση ενός προσωπικού υπολογιστή. Τα περισσότερα μεγάλα προβλήματα δεδομένων προκύπτουν από δεδομένα που δεν μπορούν να διατηρηθούν σε έναν υπολογιστή. Αν η κλίμακα των δεδομένων υπό εξέταση είναι μεγάλη, γεγονός που απαιτεί την αποθήκευσή τους σε πολλούς υπολογιστές, τότε μπορεί να γίνει χρήση μεγάλων βιβλιοθηκών και αναλυτικών στοιχείων ανάλυσης δεδομένων.

Όπως ήδη αναφέρθηκε, τόσο οι τράπεζες όσο και άλλα χρηματοπιστωτικά ιδρύματα βρέθηκαν αντιμέτωπες με τεράστιες ποσότητες δεδομένων πριν ακόμη ο όρος “μεγάλα δεδομένα” να επινοηθεί. Ο όγκος των χρηματοοικονομικών δεδομένων υπό ανάλυση έχει αυξηθεί σημαντικά με την πάροδο του χρόνου, απαιτώντας έτσι τόσο μια αυξημένη υπολογιστική ισχύ όσο και μεγαλύτερες αποθηκευτικές ικανότητες.

Αν και η πλειονότητα των σημερινών αναλύσεων οικονομικής και εταιρικής φύσης έχεις να κάνει με μεγάλα δεδομένα (της τάξης των peta), οι απλές εργασίες αναλύσεων χρησιμοποιούν ένα υπο-σύνολο δεδομένων της τάξης των μεσαίων δεδομένων. Μια πρόσφατη μελέτη κατέληξε στο συμπέρασμα:

“Οι μετρήσεις μας καθώς και άλλες πρόσφατες εργασίες δείχνουν ότι η πλειοψηφία των αναλυτικών εργασιών πραγματικού κόσμου επεξεργάζεται λιγότερο από 100 GB εισόδου, αλλά δημοφιλείς υποδομές όπως το Hadoop / MapReduce σχεδιάστηκαν αρχικά για επεξεργασία δεδομένων της τάξης των peta.” (Appuswamy, et al., 2013)

Στα πλαίσια μιας οικονομικής ανάλυσης και σε γενικές γραμμές, τα δεδομένα υπό επεξεργασία δεν ξεπερνάει σε μέγεθος τα δύο gigabytes (GB). Το συγκεκριμένο μέγεθος δεδομένων προς επεξεργασία πραγματοποιείται με εύκολο τρόπο από την Python με την βοήθεια ενσωματωμένων βιβλιοθηκών, όπως οι NumPy, pandas και PyTables. Η ανάλυση των δεδομένων από την άλλη πραγματοποιείται απευθείας στην μνήμη, οδηγώντας σε γενικά υψηλές ταχύτητες με την χρήση των σύγχρονων CPUs και GPUs.

Η Python έχει αναδειχθεί τα τελευταία χρόνια ως ηγέτης στον προγραμματισμό της επιστήμης των δεδομένων. Άλλες γλώσσες και εφαρμογές που χρησιμοποιούνται είναι η

R, το SPSS, η Julia κ.α. Η επιλογή της Python έχει να κάνει τόσο με την χρήση των βιβλιοθηκών που αναφέρθηκαν όσο και με το πλεονέκτημα απόδοσης που έχει αυτή σε σχέση με άλλες γλώσσες προγραμματισμού, συμπεριλαμβανομένων των *Matlab* και *R*, τόσο από την άποψη της ταχύτητας υπολογισμού όσο και της χρήσης της μνήμης. Ακόμη πιο σημαντικό πλεονέκτημα στην χρήση της Python είναι ότι αυτή περιέχει νέα εργαλεία που καθιστούν δυνατή την συγγραφή κώδικα, ο οποίος εκτελείται με σχεδόν την ταχύτητα του κώδικα που είναι γραμμένος σε C ή FORTRAN - εντολές γρηγορότερες από αυτές της R.

3.3.1 Δουλεύοντας με Μεγάλα Δεδομένα: MapReduce

Η χρήση του MapReduce πάνω σε μεγάλα σύνολα δεδομένων αρχικά εισήχθη το 2004 από την Google. Το MapReduce (Dean, et al., 2004) είναι μια μέθοδος η οποία χρησιμοποιείται στην επεξεργασία μεγάλων δεδομένων και επιτρέπει σε πρώτη φάση την *αντιστοίχιση των δεδομένων*²⁷ (map) και σε δεύτερη την *μείωση* τους με την χρήση ενός μηχανισμού μετασχηματισμού ή ενσωμάτωσης (reduce). Για παράδειγμα, σε μια συλλογή από γραμματόσημα, θα μπορούσε να γίνει αντιστοίχιση σε ομάδες με βάση τη χώρα προέλευσης και στη συνέχεια μείωση των ομάδων αθροίζοντας. Στο τέλος της διαδικασίας MapReduce, θα παραχθεί μια λίστα με όλες τις χώρες προέλευσης των γραμματοσήμων.

Συγκεκριμένα, μέσω της μεθόδου οι χρήστες καθορίζουν μια συνάρτηση αντιστοίχισης που επεξεργάζεται ζεύγη *κλειδιών / τιμών*. Η μέθοδος στην συνέχεια δημιουργεί ένα σύνολο ζευγών ενδιάμεσου κλειδιού / τιμής και μια λειτουργία μείωσης που συγχωνεύει όλες τις ενδιάμεσες τιμές που σχετίζονται με το ίδιο ενδιάμεσο κλειδί.

Το MapReduce είναι μια εφαρμογή εξαιρετικά επεκτάσιμη: ένας τυπικός υπολογισμός MapReduce επεξεργάζεται δεδομένα της τάξης των terabyte σε χιλιάδες μηχανές. Σχεδόν κάθε βιβλιοθήκη επιστημών δεδομένων έχει ενσωματωμένη κάποια λειτουργικότητα MapReduce εφόσον έχουν υλοποιηθεί εκατοντάδες προγράμματα μέσω της μεθόδου.

Υπάρχουν πολλές μεγάλες βιβλιοθήκες που μπορούν να χρησιμοποιηθούν για τη διαχείριση των δεδομένων με την βοήθεια του MapReduce και η Python μπορεί να

²⁷ Διαδικασία εξόρυξης δεδομένων κάνοντας χρήση ενός συγκεκριμένου χαρακτηριστικού, φίλτρου ή ομαδοποίησης

επικοινωνήσει με αυτές και να εξαγάγει αποτελέσματα για περαιτέρω αναφορές, οπτικοποίηση ή ειδοποίηση.

3.3.2 MapReduce και Hadoop

Το Hadoop είναι ένα από τα πιο δημοφιλή εργαλεία για το MapReduce. Αποτελεί ένα εργαλείο ανοικτού κώδικα της Apache υλοποιημένο σε Java και επιτρέπει την κατανεμημένη επεξεργασία μεγάλων συνόλων δεδομένων σε συστάδες υπολογιστών χρησιμοποιώντας απλά μοντέλα προγραμματισμού. Με άλλα λόγια η εφαρμογή Hadoop λειτουργεί σε ένα περιβάλλον που παρέχει κατανεμημένη αποθήκευση και υπολογισμό και έχει σχεδιαστεί για να λειτουργεί τόσο σε ένα μόνο διακομιστή όσο και σε χιλιάδες μηχανές, κάθε μια από τις οποίες προσφέρει τοπικούς υπολογισμούς και αποθήκευση.

Η αρχιτεκτονική του Hadoop περιλαμβάνει τέσσερις μονάδες:

- **Κοινό Hadoop:** σε αυτό το σύστημα περιλαμβάνονται οι βιβλιοθήκες Java και διάφορα βοηθητικά προγράμματα που απαιτούνται από τις άλλες μονάδες Hadoop. Αυτές οι βιβλιοθήκες παρέχουν γενικεύσεις τόσο στο σύστημα αρχείων όσο και στο λειτουργικό σύστημα και περιέχουν τα απαραίτητα αρχεία και τις δέσμες ενεργειών Java που απαιτούνται για την εκκίνηση του Hadoop.
- **Hadoop YARN:** η συγκεκριμένη μονάδα αφορά τον προγραμματισμό των εργασιών και τη διαχείριση των πόρων των συμπλεγμάτων.
- **Hadoop Distributed File System (HDFS TM):** μονάδα που περιλαμβάνει ένα κατανεμημένο σύστημα αρχείων και παρέχει πρόσβαση υψηλής απόδοσης στα δεδομένα της εφαρμογής.
- **Hadoop MapReduce:** Πρόκειται για σύστημα που βασίζεται στο Hadoop YARN για παράλληλη επεξεργασία μεγάλων συνόλων δεδομένων.

Όσον αφορά το Hadoop MapReduce τόσο η είσοδος όσο και η έξοδος αποθηκεύονται σε ένα σύστημα αρχείων. Το υπολογιστικό μοντέλο φροντίζει για τις εργασίες προγραμματισμού, την παρακολούθηση και την εκ νέου εκτέλεση των αποτυχημένων εργασιών.

Το Hadoop έχει την δυνατότητα λειτουργίας με οποιοδήποτε κατακεμημένο σύστημα αρχείων όπως το τοπικό σύστημα αρχείων, το HFTP FS, το S3 FS και άλλα, αλλά το πιο συνηθισμένο σύστημα αρχείων που χρησιμοποιεί η Hadoop είναι το Hadoop Distributed File System (HDFS).

Το σύστημα HDFS βασίζεται στο σύστημα αρχείων Google (GFS) και παρέχει ένα κατακεμημένο σύστημα αρχείων που έχει σχεδιαστεί για να λειτουργεί σε μεγάλες συστοιχίες (χιλιάδες υπολογιστές) μικρών υπολογιστών με αξιόπιστο και με τρόπο ανθεκτικό σε σφάλματα.

Το HDFS χρησιμοποιεί μια master / slave αρχιτεκτονική, όπου το master αποτελείται από ένα και μόνο nameNode που διαχειρίζεται τα μεταδεδομένα του συστήματος αρχείων και ένα ή περισσότερα slave dataNodes που αποθηκεύουν τα πραγματικά δεδομένα.

Το Hadoop παρουσιάζει μια σειρά από πλεονεκτήματα. Καταρχάς επιτρέπει στο χρήστη να υλοποιεί και να δοκιμάζει κατακεμημένα συστήματα με μεγάλη ταχύτητα. Επίσης είναι αποτελεσματικό, διανέμει αυτόματα τα δεδομένα και δουλεύει σε όλες τις μηχανές και χρησιμοποιεί τον παραλληλισμό των πυρήνων της CPU. Το Hadoop βασίζεται στην δική του βιβλιοθήκη²⁸ για να παρέχει αντοχή στα σφάλματα και υψηλή διαθεσιμότητα και όχι στο υλικό. Επιπλέον οι διακομιστές μπορούν να προστεθούν ή να αφαιρεθούν από το σύμπλεγμα δυναμικά και ο Hadoop συνεχίζει να λειτουργεί χωρίς διακοπή. Τέλος το Hadoop εκτός από το ότι είναι ανοιχτού κώδικα, είναι συμβατό σε όλες τις πλατφόρμες δεδομένου ότι βασίζεται στην Java.

Υπάρχουν πολλές βιβλιοθήκες της Python που μπορούν να χρησιμοποιηθούν για αποστολή δεδομένων στον Hadoop.

3.3.2 MapReduce και Spark

Το Spark είναι και αυτό ένα εξαιρετικό εργαλείο της Apache που χρησιμοποιείται στην περίπτωση που τα μεγάλα δεδομένα λειτουργούν καλύτερα σε ροή δεδομένων (δεδομένα σε πραγματικό χρόνο, δεδομένα καταγραφής, δεδομένα API). Το PySpark, το Python Spark API, επιτρέπουν μια γρήγορη εκκίνηση της αντιστοίχισης και της μείωσης του

²⁸Η βιβλιοθήκη έχει σχεδιαστεί για να εντοπίζει και να χειρίζεται τις αποτυχίες στο επίπεδο εφαρμογής

συνόλου δεδομένων αντίστοιχη με το MapReduce. Είναι επίσης εξαιρετικά δημοφιλές στα προβλήματα μηχανικής μάθησης, καθώς έχει ορισμένους ενσωματωμένους αλγορίθμους.

Υπάρχουν πολλά άλλα μεγάλης κλίμακας δεδομένα και βιβλιοθήκες εργασίας που μπορούν να χρησιμοποιηθούν με την Python.

3.4 Εφαρμογές χρηματοοικονομικών συναλλαγών σε Python

Από το 2005 και μετά, με την ωρίμανση βιβλιοθηκών (όπως NumPy και pandas) και τη διαθεσιμότητα σε προγραμματιστές της Python, η χρήση της γλώσσας στη χρηματοπιστωτική βιομηχανία αυξάνεται ραγδαία. Βασικός λόγος της ανάπτυξης είναι ότι η Python είναι κατάλληλη τόσο ως περιβάλλον διαδραστικής ανάλυσης όσο και για την γρήγορη ανάπτυξη ισχυρών συστημάτων. Η Python επίσης μπορεί να εκμεταλλευτεί το πλεονέκτημα της επεκτασιμότητας της και να ενσωματώσει εύκολα παλαιότερες βιβλιοθήκες που είναι ενσωματωμένες στη C ή C++.

Στην συνέχεια του κεφαλαίου αναφέρονται διάφορες εφαρμογές χρηματοοικονομικού ενδιαφέροντος που έχουν αναπτυχθεί στη γλώσσα προγραμματισμού Python.

3.4.1 VSTOXX

Το VSTOXX (επίσημα Euro Stoxx 50 Volatility Index) είναι μια διαδικτυακή εφαρμογή που παρέχεται από το Eurex Group στο (EurexGroup, 1999) και αποτελεί μια από τις πιο σημαντικές εφαρμογές πάνω στην μεταβλητότητα των ευρωπαϊκών δεικτών. Στην πραγματικότητα το VSTOXX καταμετράει την μεταβλητότητα των βραχυπρόθεσμων επιλογών του EuroStoxx 50, οι οποίες διαπραγματεύονται στο χρηματιστήριο Eurex. Στο VSTOXX® προσφέρονται παράγωγα έτσι ώστε να ληφθεί υπόψη η τεκμαρτή μεταβλητότητα των μετοχών.



Εικόνα 3.4: VSTOXX Derivates

Όπως αναφέρεται και στην περιγραφή της εφαρμογής:

Το *EURO STOXX 50®* δίνει στους επενδυτές την δυνατότητα να λάβουν γνώση της ευρωπαϊκής αστάθειας.

Η EURO STOXX παρέχει τα *EURO STOXX 50® Futures (EVAR)*:

Συμβόλαια που χρησιμοποιούνται ως συμπλήρωμα ανταλλαγής στο εξω-χρηματιστηριακό μέσο που αναπαράγουν το προφίλ απολαβών των συμβάσεων ανταλλαγής λόγω εξωχρηματιστηριακών διακυμάνσεων.

Η εφαρμογή βοηθά:

- Στην διαφοροποίηση του χαρτοφυλακίου μέσω της δυνατότητας προσθήκης νέων κατηγοριών περιουσιακών στοιχείων.
- Στην εξασφάλιση στην έκθεση του χαρτοφυλακίου σε κινδύνους.
- Στον καθορισμό στρατηγικών συναλλαγών

- Στη δυνατότητα εξερεύνησης των διαφορών μεταξύ ευρωπαϊκών και μη ευρωπαϊκών δεικτών

Τα επιπλέον βασικά πλεονεκτήματα που προσφέρει στους επενδυτές είναι:

- Έκθεση στη διαφορά της έμμεσης έναντι της πραγματοποιηθείσας μεταβλητότητας
- Δημιουργία δυνατότητας μετακίνησης στο μέλλον
- Επιτρέπεται στους χρήστες να επεκτείνουν τις στρατηγικές αντιστάθμισης του χαρτοφυλακίου των μετοχών τους
- Δίνει μεγαλύτερη διαφάνεια στις τιμές μέσω ενός κεντρικού βιβλίου παραγγελιών
- Μείωση του κινδύνου, καθώς το Eurex Clearing είναι ο κεντρικός αντισυμβαλλόμενος
- Διευρυμένη λειτουργικότητα και ελάχιστη καθυστέρηση μέσω της κορυφαία αρχιτεκτονική συναλλαγών T7

4. ΤΕΧΝΟΛΟΓΙΕΣ ΜΕΘΟΔΟΛΟΓΙΑΣ

Σε αυτό το κεφάλαιο πρόκειται να γίνει μια καταγραφή των τεχνολογιών που επιλέχθηκαν ως προς την υλοποίηση που πραγματοποιήθηκε στην συγκεκριμένη πτυχιακή εργασία. Θα γίνει τόσο αναφορά στις υπάρχουσες αποθήκες οικονομικών δεδομένων απαραίτητες για την αποθήκευση και ανάλυση των οικονομικών / χρηματιστηριακών δεδομένων, όσο και στις τεχνολογίες cloud και τις βιβλιοθήκες που χρησιμοποιήθηκαν.

4.1 Διαδικτυακές Αποθήκες Οικονομικών Δεδομένων

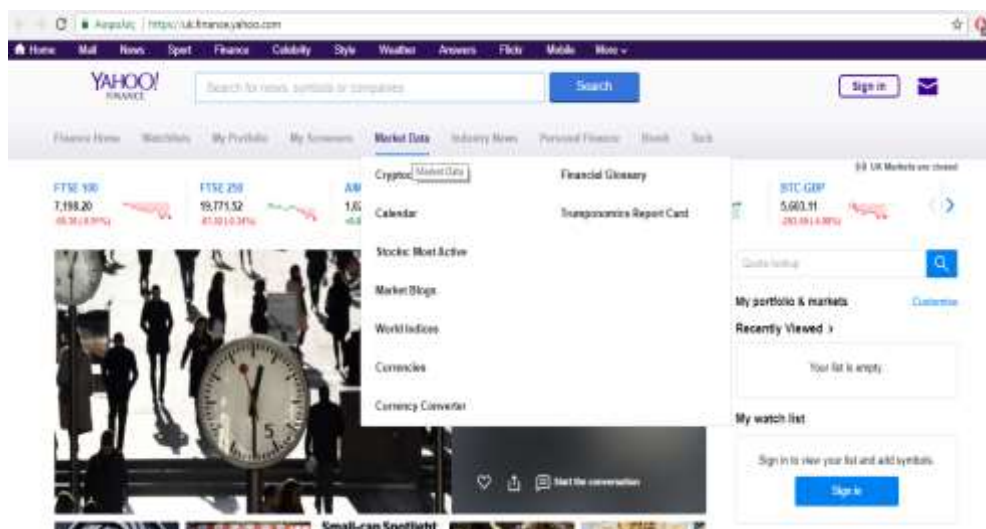
Τα οικονομικά / χρηματιστηριακά δεδομένα ανήκουν στην κατηγορία των Big Data και είναι αδύνατο να αποθηκευτούν, λόγω περιορισμών της μνήμης, σε τοπικούς και προσωπικούς υπολογιστές. Αυτό έχει οδηγήσει στην αποθήκευση των παραπάνω δεδομένων σε διαδικτυακούς τόπους και στο cloud. Οι αποθήκες και οι εφαρμογές ανάλυσης αυτών των δεδομένων είναι πολλές και αυτές με ελεύθερη πρόσβαση στα δεδομένα ξεπερνούν τις 50. Κάποιες από αυτές ταξινομημένες ανάλογα με το είδος δεδομένων που παρέχουν είναι:

Αποθήκες και διαδικτυακές εφαρμογές που παρέχουν δωρεάν δεδομένα αγορών είναι:

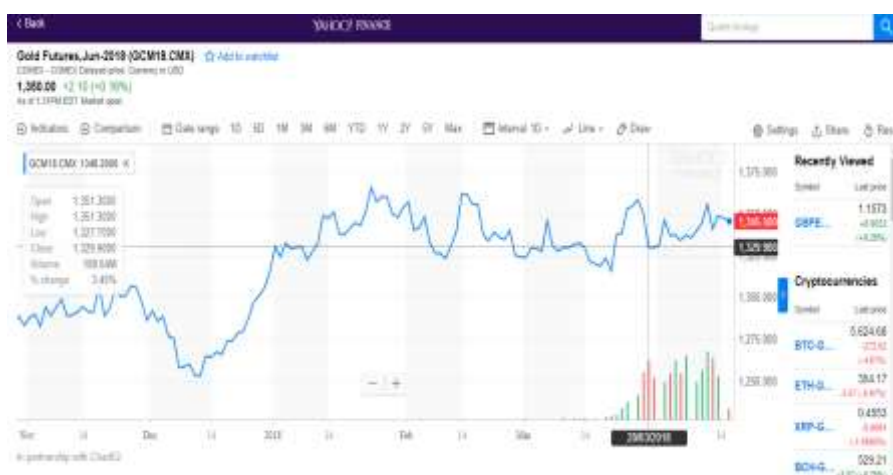
- **Yahoo Finance:** στον ιστότοπο συμπεριλαμβάνονται δεδομένα της αγοράς, ειδήσεις της βιομηχανίας και της προσωπικής χρηματοδότησης του Yahoo. Στην ενότητα δεδομένων αγοράς ο χρήστης έχει πρόσβαση σε πληροφορίες των αγορών²⁹ ενώ συμπεριλαμβάνει ένα blog στην αγορά. Στο χρήστη παρέχεται η δυνατότητα μέσω ενός κλικ πάνω σε ένα συγκεκριμένο απόθεμα ή μια αγορά, να έχει μια περίληψη σε πραγματικό χρόνο όπως και ένα μικρό γράφημα. Επιπλέον επιλέγοντας την δυνατότητα της “Πλήρης οθόνης”, το γράφημα μετατρέπεται σε πλήρες και παρουσιάζει τις τιμές του αποθέματος ή της αγοράς για οποιαδήποτε ημέρα. Με την επιλογή Σύγκριση ο χρήστης μπορεί να έχει πρόσβαση σε

²⁹ Στις πληροφορίες συμπεριλαμβάνονται τα νομίσματα, οι πιο ενεργές μετοχές και οι διεθνείς χρηματιστηριακούς δείκτες

περισσότερες τιμές μετοχών ή αγορών στην ίδια οθόνη για τις ίδιες ημερομηνίες. Ο ιστότοπος παρέχει επίσης την δυνατότητα δημιουργίας προσωπικού χαρτοφυλακίου.



Εικόνα 4.1: Αρχική σελίδα του Yahoo Finance



Εικόνα 4.2: Ισοτιμία χρυσού

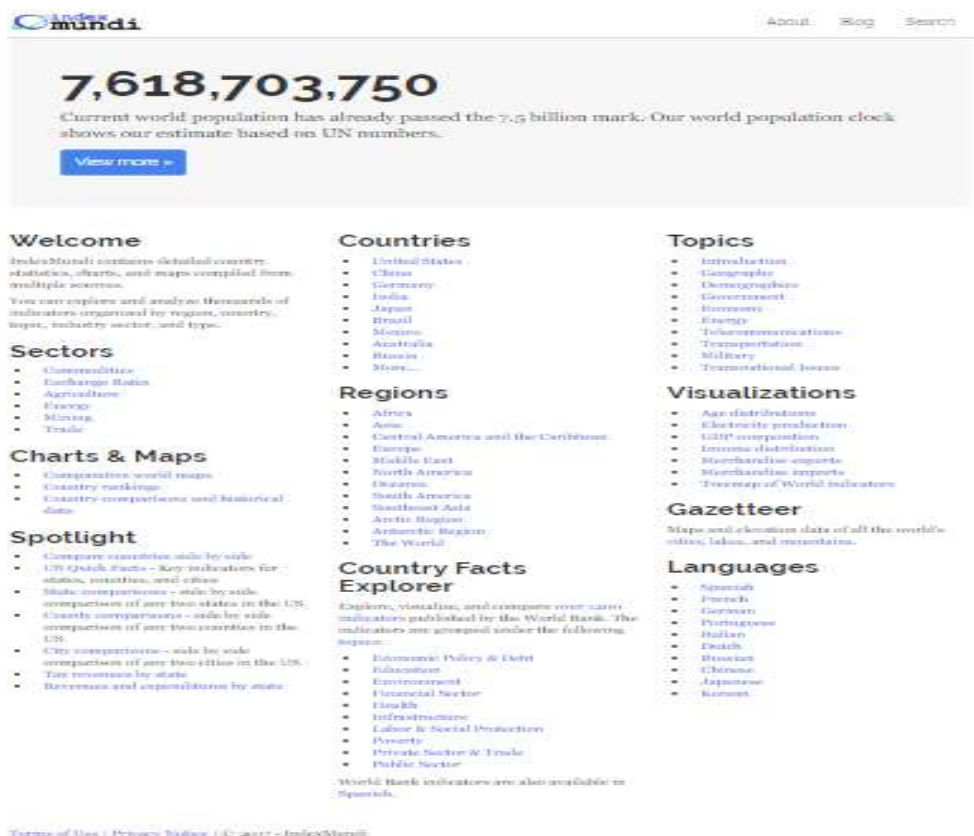
- **Telegraph Shares & Markets:** Πρόκειται για έναν καλά σχεδιασμένο και εύκολα προσπελάσιμο ιστότοπο για την εξεύρεση τιμών μετοχών και βασικών εμπορευμάτων. Ο ιστότοπος παρέχει στον χρήστη την δυνατότητα να έχει πρόσβαση σε δεδομένα για καθέναν από τους κύριους δείκτες χρηματιστηριακών αγορών και όλες τις μεμονωμένες μετοχές του συγκεκριμένου δείκτη. Άλλη μια δυνατότητα που παρέχεται είναι ο σχεδιασμός γραφημάτων χρονολογικών σειρών

οποιαδήποτε μετοχής για περιόδους από 1 έως 10 ετών. Τα γραφήματα μπορούν να είναι σε διαφορετικές μορφές, όπως π.χ. γραφήματα γραμμής, ενώ δίνεται η δυνατότητα προσθήκης των κινούμενων μέσων όρων και περαιτέρω προσθήκη γραφημάτων στο κάτω μέρος της οθόνης για μια σειρά άλλων δεικτών σε αυτό το μερίδιο (π.χ. ο όγκος και ο κινητός μέσος όρος, η κινητή μέση σύγκλιση / απόκλιση, η γρήγορη στοχαστική και ρυθμός μεταβολής). Επίσης, δίνεται η δυνατότητα δημιουργίας χαρτοφυλακίου.



Εικόνα 4.3: Σχεδίαση γραφήματος χρονολογικής σειράς του FTSE 100 με κινητούς μέσους

- **IndexMundi commodities:** Ο ιστότοπος IndexMundi διαθέτει ένα εκτεταμένο σύνολο στοιχείων για τις τιμές των βασικών προϊόντων που προέρχονται από το Διεθνές Νομισματικό Ταμείο. Τα προϊόντα ταξινομούνται σε χρονικές περιόδους 6 μηνών, 1, 5, 10, 15 ή 25 ετών και εμφανίζονται σε έναν πίνακα ο οποίος μπορεί να μεταφερθεί στο Excel. Ο ιστότοπος περιέχει επίσης στοιχεία χωρών.



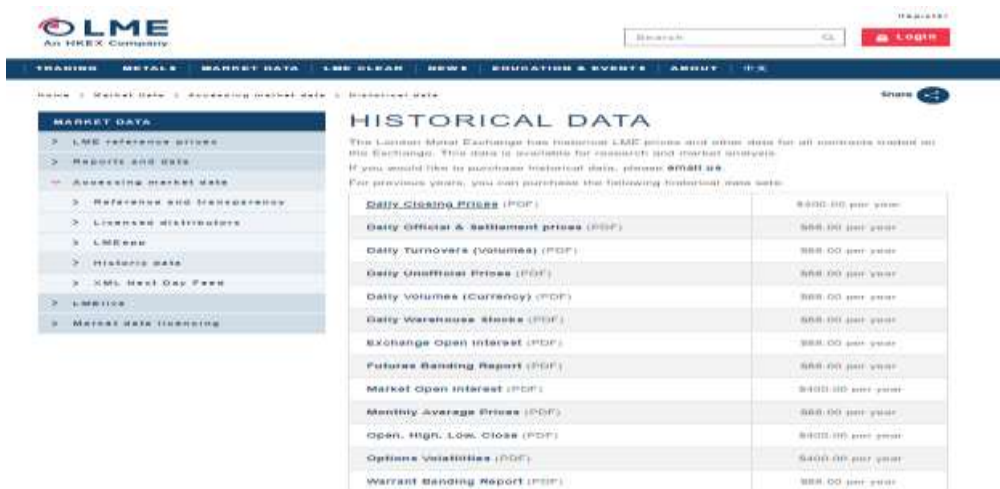
Εικόνα 4.4: Αρχική σελίδα του IndexMundi

- **Energy Information Administration:** Πρόκειται για ένα αμερικανικό κυβερνητικό ιστότοπο, που παρέχει διεθνή και αμερικανικά δεδομένα για την ηλεκτρική ενέργεια, τον άνθρακα, το φυσικό αέριο, την πυρηνική ενέργεια και το πετρέλαιο. Μέσω του ιστότοπου παρέχεται στον χρήστη η δυνατότητα να κατεβάσει αρχεία Excel με ημερήσιες, εβδομαδιαίες, μηνιαίες και ετήσιες τιμές πάνω σε διάφορα καύσιμα.



Εικόνα 4.5: Αρχική σελίδα του Energy Information

- **The London Metal Exchange:** ιστότοπος που παρέχει ιστορικά δεδομένα πάνω στις τιμές των διάφορων μετάλλων. Παρέχει την δυνατότητα κατεβάσματος τόσο υπολογιστικών φύλλων Excel και PDF για κάθε μήνα, του τρέχοντος έτους για κάθε μέταλλο.



Εικόνα 4.6: Αρχική σελίδα του LME

- **The UK government's Debt Management Office:** ιστότοπος της βρετανικής κυβέρνησης που παρέχει, μεταξύ άλλων, τα δεδομένα για τις αγορές χρυσού, και θησαυροφυλακίων.
- **BBC Market Data:** Ο ιστότοπος ειδήσεων του BBC διαθέτει τμήμα με δεδομένα αγοράς στο οποίο δίνονται οι τρέχουσες τιμές και τα διαγράμματα των διακυμάνσεων των τιμών κατά τη διάρκεια ενός, τριών ή δώδεκα μηνών χρηματιστηριακών αγορών, μετοχών, νομισμάτων και βασικών προϊόντων.
- **UK house prices:** Υπάρχουν πέντε τοποθεσίες για τα στοιχεία των τιμών κατοικιών του Ηνωμένου Βασιλείου.

UK House Price Indices (χρησιμοποιώντας τα στοιχεία του Κτηματολογίου): ιστότοπος που οδηγεί σε αναφορές σχετικά με την αγορά κατοικίας του Ηνωμένου Βασιλείου και τις τιμές κατοικιών. Παρέχει εκθέσεις με συνοπτικά δεδομένα και γραφήματα. Περιέχει επίσης συνδέσμους προς τα υποκείμενα σύνολα δεδομένων. Τα σύνολα δεδομένων προσαρμόζονται ανά θέμα, έτος κ.λπ. και η λήψη τους γίνονται με αρχεία Excel.

Nationwide house prices site είναι ιστότοπος που παρέχει τις τριμηνιαίες κινήσεις στις εθνικές τιμές κατοικιών από το 1952. Λεπτομερέστερες ακολουθίες που διακρίνουν τον τύπο ιδιοκτησίας, τον τύπο αγοραστή και την περιοχή δίδονται για τα πιο πρόσφατα έτη (1973, 1983 ή 1991).

Δείκτης τιμών Halifax House (Lloyds Banking Group) που παρουσιάζει τις κινήσεις στις τιμές των κατοικιών του Ηνωμένου Βασιλείου τους τελευταίους 12 μήνες.

Zoopla ιστότοπος που παρέχει εκτιμήσεις τιμών κατοικιών σε πραγματικό χρόνο με βάση τον ταχυδρομικό κώδικα, καθώς και τις τάσεις των τιμών κατοικιών ανά περιοχή.

Προβλέψεις για τις τιμές κατοικίας από το Γραφείο Ευθύνης Προϋπολογισμού: Το OBR προβλέπει τις μέσες τιμές κατοικιών, όπως μετράται από τον δείκτη τιμών των κατοικιών στο ONS, και τον αριθμό των συναλλαγών που πραγματοποιούνται, όπως αναφέρει η HMRC.

- **International property prices:** Υπάρχουν πέντε χρήσιμοι ιστότοποι για τις διεθνείς τιμές των ακινήτων.

Η Τράπεζα Διεθνών Διακανονισμών (Bank for International Settlements - BIS). Ο ιστότοπος του BIS περιλαμβάνει ένα τμήμα για τις τιμές των κατοικιών

και των εμπορικών ακινήτων σε ένα σύνολο χωρών που μπορούν να αποθηκευτούν ως αρχεία Excel.

Οδηγός παγκόσμιας ιδιοκτησίας (Global Property Guide). Ιστότοπος ο οποίος είναι σχεδιασμένος για επενδυτές ακινήτων σε άλλες χώρες αφού παρακολουθεί επίσημους και ημиеπίσημους δείκτες τιμών κατοικιών σε 57 χώρες. Διαθέτει επίσης στοιχεία σχετικά με τις ακαθάριστες αποδόσεις ενοικίασης, τις δομές της αγοράς, το κόστος μεταβίβασης, τον νόμο ιδιοκτητών και μισθωτών και τους πραγματικούς φορολογικούς συντελεστές για κατοικίες σε περισσότερες από 100 χώρες.

IMF (Διεθνές Νομισματικό Ταμείο). Η Global Housing Watch του ΔΝΤ παρακολουθεί τις εξελίξεις στις αγορές κατοικιών ανά τον κόσμο σε τριμηνιαία βάση. Παρέχει τρέχοντα στοιχεία για τις τιμές των κατοικιών καθώς και μετρήσεις που χρησιμοποιούνται για την αξιολόγηση της αποτίμησης στις αγορές κατοικιών, όπως οι τιμές κατοικίας προς ενοικίαση και οι αναλογίες κατοικίας-τιμής προς έσοδα.

The economist. Παρέχονται πληροφορίες, για μέχρι και 27 χώρες, για τους δείκτες τιμών κατοικίας, τις τιμές κατοικιών σε πραγματικούς όρους, τις τιμές κατοικιών έναντι των μέσων εισοδημάτων, τις τιμές κατοικιών έναντι ενοικίων και την ποσοστιαία μεταβολή των τιμών των κατοικιών.

Knight Frank. Ιστότοπος που δημοσιεύει διάφορες εκθέσεις για την αγορά κατοικίας, συμπεριλαμβανομένης μιας τριμηνιαίας παγκόσμιας έκθεσης. Περιέχει μια σειρά δεδομένων, όπως οι τελευταίες ετήσιες μεταβολές των τιμών των κατοικιών σε 56 χώρες και οι ετήσιες εκατοστιαίες μεταβολές του παγκόσμιου δείκτη τιμών για τα 21 τελευταία τρίμηνα

- **UK labour market data from Adzuna:** ιστότοπος αναζήτησης εργασίας, που ξεκίνησε το 2011. Διαθέτει μια μεγάλη βάση δεδομένων για θέσεις εργασίας για το Ηνωμένο Βασίλειο και για άλλες 10 χώρες. Περιλαμβάνει στατιστικά δεδομένα μηνιαίας μισθοδοσίας που αναφέρονται στο προηγούμενο έτος. Περιλαμβάνει, επίσης, εθνικά στοιχεία σχετικά με τους μισθούς και τις κενές θέσεις εργασίας και σύγκριση μισθών με τον εθνικό μέσο όρο για συγκεκριμένες θέσεις εργασίας.
- **worldwide-tax.com.** Ιστότοπος που περιέχει βασικά στοιχεία για τους παγκόσμιους φόρους, συμπεριλαμβανομένων των συντελεστών φορολογίας

εισοδήματος, πίνακα σύγκρισης των φορολογικών συντελεστών, τμήμα για τις επιχειρήσεις και τη χρηματοδότηση παγκοσμίως, κατάλογο παγκόσμιων πρεσβειών, προξενείων και διπλωματικών αποστολών, συνδέσμους με γενικές και οικονομικές περιοχές ειδήσεων, φορολογικά νέα, συνδέσεις με τις παγκόσμιες φορολογικές και φοροτεχνικές υπηρεσίες, κατάλογο διεθνών χρηματιστηρίων και συνδέσμους με τοποθεσίες επιχειρηματικών ευκαιριών που διοργανώνονται ανά χώρα.

- **Mergers and Acquisitions Statistics from the IMAA.** Το τμήμα πόρων του Ινστιτούτου συγχωνεύσεων, εξαγορών και συμμαχιών (IMAA) παρέχει εκτεταμένα και τακτικά ενημερωμένα στατιστικά στοιχεία σχετικά με τις συγχωνεύσεις και τις εξαγορές παγκοσμίως. Μεγάλο μέρος των δεδομένων έχουν δημόσια πρόσβαση και έχουν την μορφή: Περιφέρεια, Τύπος συναλλαγής, Βιομηχανία και Χώρα.

Στους άλλους διεθνείς ιστότοπους περιλαμβάνονται:

- **European Central Bank statistics:** στατιστικά της Ευρωπαϊκής κεντρικής τράπεζας. Στα στοιχεία περιλαμβάνονται το οικονομικό δελτίο και η ετήσια έκθεση.
- **European Economy:** δικτυακός τόπος Euroρα με δεδομένα διαθέσιμα και σε έντυπη μορφή.
- **OECD Statistics Portal:** ιστότοπος που παρέχει στατιστικά για καθεμία από τις 35 χώρες του ΟΟΣΑ και τη ζώνη του ευρώ. Οι στατιστικές ταξινομούνται ανά ομάδα θεμάτων, συμπεριλαμβανομένων των εθνικών λογαριασμών, των οικονομικών, της γεωργίας, της ανάπτυξης, του διεθνούς εμπορίου, της εργασίας, των τιμών, της δημόσιας διαχείρισης και των βραχυπρόθεσμων οικονομικών στατιστικών.
- **World Bank data sets:** ιστότοπος της Παγκόσμιας Τράπεζας. Περιέχει μια τεράστια βάση δεδομένων για οικονομικές, κοινωνικές και άλλες στατιστικές ανάπτυξης για όλες τις χώρες του κόσμου. Στα δεδομένα με ανοιχτή πρόσβαση, η αναζήτηση γίνεται ανά χώρα, θέμα και δείκτη.
- **IMF data sets:** περιέχει αναφορές σε οικονομικά πάνω σε όλες τις χώρες του κόσμου

- **WTP international trade statistics:** Στο τμήμα στατιστικών στοιχείων του ιστοτόπου είναι δυνατή η πρόσβαση σε διάφορες βάσεις δεδομένων και δημοσιεύσεις στα δεδομένα εμπορίου και τιμολογίων.
- **UNCTAD statistics:** ο ιστότοπος του United Nations Conference on Trade and Development (UNCTAD) διαθέτει μια ηλεκτρονική στατιστική βάση δεδομένων, την UNCTADstat η οποία περιλαμβάνει στοιχεία για το διεθνές εμπόριο, τις ξένες άμεσες επενδύσεις, τις τιμές των βασικών εμπορευμάτων, τις οικονομικές τάσεις, τον πληθυσμό και την εργασία, τους εξωτερικούς οικονομικούς πόρους και τις θαλάσσιες μεταφορές.
- **Bank for international settlements:** Η Τράπεζα Διεθνών Διακανονισμών είναι “ένας διεθνής οργανισμός που προωθεί τη διεθνή νομισματική και οικονομική συνεργασία και χρησιμεύει ως τράπεζα για τις κεντρικές τράπεζες”. Από τον ιστότοπο της παρέχεται πρόσβαση στις ιστοσελίδες καθεμιάς από τις κεντρικές τράπεζες του κόσμου.
- **Principal Global Indicators:** ιστότοπος που παρέχει στοιχεία για τις χώρες του G-20 και για επιπλέον δώδεκα χώρες της ευρωζώνης αλλά και το Χονγκ Κονγκ. Ο ιστότοπος φιλοξενείται από το IMF. Παρέχει μακροοικονομικά και οικονομικά στοιχεία σε μορφή πίνακα για τις 34 οικονομίες και η προσπέλαση των δεδομένων γίνεται ανά δείκτη ή ανά χώρα.
- **Nationmaster:** ιστότοπος που παρέχει ένα σύνολο περισσότερων από πέντε χιλιάδων δεδομένων, από πηγές όπως το CIA World Factbook, τον ΟΗΕ και τον ΟΟΣΑ. Εκτός από τα προφίλ των επιμέρους χωρών, συμπεριλαμβανομένων των χαρτών και των σημαιών τους, μέσω της σελίδας υπάρχει δυνατότητα πρόσβασης στα δεδομένα ανά θέμα. Υπάρχουν δεδομένα στην εκπαίδευσης, στην οικονομία, στα εγκλήματα, στη θνησιμότητα και στην υγεία. Ο ιστότοπος προσφέρει μια δυνατότητα δημιουργίας αναφορών συσχετισμού και διασκορπισμού κατά τη διάρκεια της εκμάθησης.
- **Gapminder:** επιτρέπει τον σχεδιασμό (με βάση τον χρόνο) τριών δεικτών για τις περισσότερες χώρες και ταυτόχρονα. Τα δεδομένα περιέχουν ιστορικά στοιχεία δεκαετών ή και αιώνων για ορισμένους δείκτες. Βασικός στόχος της εφαρμογής είναι η εύκολη οπτικοποίηση των διαφορών μεταξύ των χωρών.
- **Cia world factbook:** ιστότοπος της αμερικάνικης υπηρεσίας πληροφοριών που παρέχει μια σειρά από οικονομικά και άλλα στοιχεία, ανά χώρα.

- **FRED: Federal Reserve Economic Data.** Ιστότοπος που παρέχει οικονομικά δεδομένα της Αμερικής. Τα δεδομένα ανανεώνονται καθημερινά από Federal Reserve Bank του St. Louis.
- **EconStats from EconomyWatch**
- **Penn World Table:** Το Penn World Table v9.0 των Robert C. Feenstra, Robert Inklaar και Marcel P. Timmer παρέχεται από το Κέντρο Ανάπτυξης του Πανεπιστημίου του Groningen. Πρόκειται για μια βάση δεδομένων με πληροφορίες σχετικά με τα επίπεδα εισοδήματος, παραγωγής, εισροών και παραγωγικότητας, 182 χωρών από το 1950 και μεταγενέστερα. Τα αρχεία δεδομένων είναι συμβατά με Excel και Stata.
- **DataMarket:** Ο ιστότοπος παρέχει ένα μέσο αναζήτησης, σύγκρισης, οπτικοποίησης και λήψης ποσοτικών δεδομένων από μια μεγάλη ποικιλία διεθνών πηγών. Τα δεδομένα εντοπίζονται με αναζήτηση βάση χώρας ή βιομηχανίας. Τα δεδομένα λαμβάνονται μέσω του ΔΝΤ, του ΟΗΕ, της Μονάδας Πληροφοριών Οικονομίας, της Eurostat και της Παγκόσμιας Τράπεζας. Ο ιστότοπος περιλαμβάνει περίπου 125 εκατομμύρια χρονολογικές σειρές από περίπου 16.000 σύνολα δεδομένων ελεύθερης πρόσβασης. Οι μορφές αρχείων που υποστηρίζονται είναι Excel, PDF, SVG, κλπ.
- **Knoema:** ο ιστότοπος συλλέγει διεθνή δεδομένα από διάφορες πηγές και τα παρουσιάζει σε ένα κοινό χώρο, απεικονίζοντας στοιχεία, εφαρμόζοντας αναλυτικές λειτουργίες, δημιουργώντας μια σειρά από πίνακες ελέγχου και παρουσιάζοντας το αποτέλεσμα.
- **The Conference Board datasets:** ανεξάρτητος μη κερδοσκοπικός παγκόσμιος οργανισμός με σκοπό του είναι “να βοηθήσει τους ηγέτες να περιηγηθούν στα μεγαλύτερα ζητήματα που επηρεάζουν τις επιχειρήσεις και να υπηρετούν καλύτερα την κοινωνία”. Μεταξύ άλλων δραστηριοτήτων, παρέχει μια σειρά διεθνών δεδομένων όπως οικονομικά στοιχεία και ανάλυση, δείκτες παγκόσμιου επιχειρηματικού κύκλου, ετήσια στοιχεία για το ΑΕΠ, τον πληθυσμό, την απασχόληση, τις ώρες, την ποιότητα της εργασίας, τις κεφαλαιακές υπηρεσίες, την παραγωγικότητα της εργασίας και την παραγωγικότητα του συνολικού παράγοντα για 123 χώρες στον κόσμο.
- **Vizala:** ιστότοπος που περιέχει δεδομένα σε επίπεδο χώρας για 155 χώρες. Τα δεδομένα είναι οργανωμένα σε τρεις μεγάλες κατηγορίες: Οικονομία, Έρευνα

Αγοράς και Κοινωνία, καθένα από τα οποία στη συνέχεια χωρίζεται σε επτά υποκατηγορίες. Για παράδειγμα, η οικονομία χωρίζεται σε τράπεζες και δανειοδοτήσεις, καταναλωτές, χρέος / δημοσιονομική πολιτική, ΑΕΠ / ΑΕΕ, εργατικό δυναμικό / εργατικό δυναμικό, μηνιαία επισκόπηση και νομισματική / πληθωριστική πολιτική.

Σε καθεμία από τις υποκατηγορίες μπορούν να επιλεγούν μια σειρά ομάδων δεικτών. Για παράδειγμα, στην υποκατηγορία ΑΕΠ / ΑΕΕ μπορούν να επιλεγούν δείκτες στις ακόλουθες ομάδες: Επισκόπηση, βασικά ποσοστά ανάπτυξης, ΑΕΠ ανά Δαπάνη, ΑΕΠ ανά Δαπάνη Ανάπτυξης, Δαπάνες ως ποσοστό του ΑΕΠ, Προστιθέμενη αξία ανά Βιομηχανία, Προστιθέμενη αξία από τα ποσοστά ανάπτυξης της βιομηχανίας και τη βιομηχανία ως% του ΑΕΠ. Τα αποτελέσματα μπορούν να φιλτραριστούν ανά έτος, ομάδα χωρών, μέγεθος πληθυσμού και κατά κεφαλήν ΑΕΠ.

- **Asian Development Bank: Data and Research:** Η Ασιατική Αναπτυξιακή Τράπεζα δημοσιεύει μια σειρά οικονομικών, οικονομικών, κοινωνικών, περιβαλλοντικών και βιώσιμων αναπτυξιακών στόχων (SDG) στα 48 κράτη μέλη της. Στο τμήμα στατιστικών στοιχείων, μπορείτε να αναζητηθούν ανά χώρα οι κατηγορίες: Οικονομικοί δείκτες, δείκτες φτώχειας και στατιστικές εκθέσεις. Επιπλέον παρέχει τρεις χρήσιμες ετήσιες στατιστικές δημοσιεύσεις: Σειρά βασικών στατιστικών, Ασιατική Προοπτική Ανάπτυξης (ADO), Βασικοί δείκτες (KI) για την Ασία και τη Σειρά Ειρηνικού. Τα δεδομένα του ιστότοπου είναι διαθέσιμα για λήψη σε αρχεία Excel.

Τέλος, υπάρχουν αποθήκες και εφαρμογές δεδομένων που λαμβάνουν στοιχεία από άλλους διαδικτυακούς τόπους. Ορισμένες από αυτές είναι:

- **UK Data Service:** η Υπηρεσία δεδομένων του Ηνωμένου Βασιλείου που παρέχει πρόσβαση σε μια σειρά δεδομένων από εθνικούς και διεθνείς οργανισμούς. Εκτός από αυτά τα σύνολα δεδομένων, η υπηρεσία δεδομένων του Ηνωμένου Βασιλείου περιλαμβάνει δεδομένα από το Census.ac.uk, την Υπηρεσία Ασφαλούς Δεδομένων (SDS) και την Τράπεζα Ερωτήσεων Ερευνών (SQB). Η σύνδεση μόνο με λογαριασμό πανεπιστημίου / κολέγια του Ηνωμένου Βασιλείου. Δύο τμήματα του ιστότοπου με ιδιαίτερο ενδιαφέρον είναι: οι Διεθνείς

μακροσκοπήσεις που παρέχει πρόσβαση σε ένα ευρύ φάσμα διεθνών μακροοικονομικών στοιχείων από την Παγκόσμια Τράπεζα, τη Eurostat, το ΔΝΤ, τον Διεθνή Οργανισμό Ενέργειας (IEA), τη ΔΟΕ, τον ΟΟΣΑ και την UNIDO και οι Διεθνείς μακροσκοπήσεις που παρέχει πρόσβαση σε τράπεζες (π.χ. Διεθνές Χρηματοοικονομικό Στατιστικό (IFS) του ΔΝΤ).

- **Quandl:** ίσως η πιο δημοφιλής αποθήκη οικονομικών δεδομένων το Quandl περιέχει πάνω από 5 εκατομμύρια σύνολα δεδομένων χρονοσειρών από περισσότερες από 400 πηγές. Τα περισσότερα στοιχεία αφορούν τα οικονομικά στοιχεία και τις τιμές περιουσιακών στοιχείων, αλλά υπάρχουν και διάφορα οικονομικά σύνολα δεδομένων. Πρόκειται για μια συνδρομητική υπηρεσία που παρέχει και ελεύθερη πρόσβαση σε κάποια σύνολα δεδομένων. Παρέχεται επίσης η δυνατότητα απεικόνισης, αποθήκευσης, πιστοποίησης, επικύρωσης, ανεβάσματος και μετατροπής δεδομένων.
- **RFE Data Links:** μέρος του AEAWeb (πληροφορίες για οικονομολόγους) που παρέχει η Αμερικανική Οικονομική Ένωση. Από τη σελίδα “Δεδομένα” του τμήματος πόρων για τις οικονομολόγους (RFE) του ιστότοπου, υπάρχουν σύνδεσμοι σε τις πύλες δεδομένων σε απευθείας σύνδεση, οι οποίες ταξινομούνται ανά περιοχή. Οι κατηγορίες περιλαμβάνουν τα στοιχεία μακροεντολών και περιφερειακών δεδομένων των ΗΠΑ, τα δεδομένα παγκόσμιας και μη ΗΠΑ, τις χρηματοπιστωτικές και χρηματοπιστωτικές αγορές, τα αρχεία δεδομένων και τα αρχεία προγράμματος.
- **Development Gateway Foundation:** ηλεκτρονική πύλη πόρων για την ανάπτυξη πληροφοριών και την ανταλλαγή γνώσεων σε όλο τον κόσμο. Η Gateway Development είναι “ένας διεθνής μη κερδοσκοπικός οργανισμός με αποστολή να μειώσει τη φτώχεια και να επιτρέψει την αλλαγή στις αναπτυσσόμενες χώρες μέσω της τεχνολογίας των πληροφοριών”. Παρέχει πρόσβαση σε δεδομένα σχετικά με τη διεθνή βοήθεια ανά τύπο βοήθειας, δότη, αποδέκτη, σκοπό ή δραστηριότητα.
- **EcEdWeb:** ηλεκτρονική πύλη πύλης των Η.Π.Α. για τους πόρους της οικονομικής εκπαίδευσης. Στο τμήμα οικονομικών πληροφοριών του, διαθέτει ένα μεγάλο σύνολο συνδέσμων σε ιστότοπους οικονομικών δεδομένων και σε άλλους ιστότοπους σχετικούς με τη μελέτη των οικονομικών.
- **Zanran:** ιστότοπος αναζήτησης δεδομένων. Σύμφωνα με τον ιστότοπο: “*Το Zanran βοηθά στην εύρεση ημι-δομημένων δεδομένων στον ιστό. Αυτά είναι*

αριθμητικά δεδομένα που παρουσιάζονται ως γραφήματα και πίνακες. Για παράδειγμα, τα δεδομένα θα μπορούσαν να είναι ένα γράφημα σε μια αναφορά PDF ή ένας πίνακας ένα υπολογιστικό φύλλο Excel ή ένα barchart που εμφανίζεται ως εικόνα σε μια σελίδα HTML. Αυτή η τεράστια ποσότητα πληροφοριών μπορεί να είναι δύσκολο να βρεθεί χρησιμοποιώντας συμβατικές μηχανές αναζήτησης, οι οποίες επικεντρώνονται κατά κύριο λόγο στην εύρεση κειμένου παρά σε γραφήματα, πίνακες και γραφήματα.”

4.2 Υπολογιστική Νέφος (Cloud Computing)

Ένα παράδειγμα εφαρμογής ανοιχτού κώδικα (ΕΛ/ΛΑΚ), που αναπτύχθηκε εκτενέστερα στο 3^ο κεφάλαιο, είναι η υπολογιστική νέφος. Το υπολογιστικό νέφος είναι η παροχή υπολογιστικών πόρων, όπως δίκτυο, εξυπηρετητές, εφαρμογές και υπηρεσίες. Η έννοια αναφέρεται στη χρήση υπολογιστικής ισχύος που χωροταξικά είναι εγκατεστημένη σε ένα “σύννεφο” απόμακρων δικτύων. Σύμφωνα με το US National Institute for Standards and Technology το νέφος ορίζεται ως: “Ένα μοντέλο που δίνει τη δυνατότητα της συνεχούς, εύκολης και υψηλών απαιτήσεων πρόσβασης σε μια κοινόχρηστη συλλογή ρυθμιζόμενων υπολογιστικών πόρων, οι οποίοι τροφοδοτούνται και απελευθερώνονται με ελάχιστη προσπάθεια διαχείρισης και αλληλεπίδρασης παροχής υπηρεσιών”.

Σύμφωνα με τα παραπάνω η τεχνολογία υπολογιστικού νέφους αποτελεί οποιοδήποτε λογισμικό ή υλικό χρησιμοποιεί ο χρήστης, το οποίο όμως δεν βρίσκεται στον υπολογιστή του, αλλά σε κάποιον διακομιστή του Διαδικτύου. Γενικότερα θεωρείτε ως μια προσέγγιση στον χώρο των κατανεμημένων συστημάτων, με κάποιες υφιστάμενες τεχνολογίες.

Το υπολογιστικό νέφος μπορεί να χρησιμοποιηθεί τόσο στην αποθήκευση, την επεξεργασία και τη χρήση δεδομένων, αλλά και στην διαδικτυακή χρήση λογισμικού και υπηρεσιών μέσω απομακρυσμένων υπολογιστών. Επιπλέον υπηρεσίες που παρέχονται στο νέφος είναι η κατ' αίτηση παροχή εικονικών μηχανών, το διαδικτυακό ηλεκτρονικό ταχυδρομείο ή τα κοινωνικά δίκτυα.

Τα είδη εφαρμογών που παρέχονται στο υπολογιστικό νέφος μπορούν να χωριστούν σε 4 γενικές κατηγορίες:

- *εμπορικές εφαρμογές (business applications)*
- *εφαρμογές κοινωνικής δικτύωσης (social media applications)*
- *υπηρεσίες πληροφορικής (cloud IT service)*
- *εφαρμογές ενίσχυσης παραγωγικότητας*

Επίσης, το υπολογιστικό νέφος, ανάλογα με τις υπηρεσίες που παρέχει, μπορεί να κατηγοριοποιηθεί σε 3 διαφορετικά μοντέλα. Η συγκεκριμένη δομή αποτελεί το ονομαζόμενο μοντέλο SPI και περιλαμβάνει τις υποδομές:

- **SaaS (Software as a Service)** ή Υπηρεσίες Λογισμικού: στο συγκεκριμένο μοντέλο ο πελάτης (ή χρήστης) έχει πρόσβαση και μπορεί δυνατότητα χρήσης κάποιας εφαρμογής λογισμικού. Η συγκεκριμένη εφαρμογή δεν βρίσκεται πλέον εγκατεστημένη, όπως κλασσικά συμβαίνει, στον υπολογιστή του τελικού χρήστη αλλά φιλοξενείται, αναπτύσσεται και διαχειρίζεται, από τον πάροχο. Το μόνο λογισμικό που χρειάζεται να είναι εγκατεστημένο στον υπολογιστή του τελικού χρήστη είναι ένας Web browser (φυλλομετρητής). Οι χρήστες της εν λόγω υπηρεσίας περιορίζονται ως προς το πώς μπορούν να χρησιμοποιήσουν και να αλληλεπιδράσουν με την εφαρμογή, περιορισμοί που τίθενται από τον πάροχο. Στις υπηρεσίες που παρέχονται από το μοντέλο SaaS περιλαμβάνονται υπηρεσίες περιεχομένου, ηλεκτρονικού ταχυδρομείου, αλλά και επιχειρηματικές εφαρμογές, όπως για παράδειγμα οι CRM (ή εφαρμογές διαχείρισης σχέσεων με τους πελάτες)
- **PaaS (Platform as a Service)** ή Υπηρεσίες Πλατφόρμας. Το μοντέλο επιτρέπει στον χρήστη να δημιουργήσει και να διαχειριστεί τις εφαρμογές του νέφους. Αυτό δεν σημαίνει ότι ο τελικός χρήστης έχει δυνατότητα ελέγχου της υποδομής του νέφους. Υπηρεσίες που ανήκουν στο μοντέλο είναι το Google app Engine, το Microsoft Windows Azure, το Mozilla, το Bespin κ.α
- **IaaS (Infrastructure as a Service)** ή Υποδομή ως Υπηρεσία. Το συγκεκριμένο μοντέλο υπηρεσιών δίνει στους χρήστες την δυνατότητα να εκμισθώνουν κέντρα δεδομένων που βρίσκονται στους διακομιστές του πάροχου της υπηρεσίας. Η υπηρεσία δίνει στον χρήστη τη δυνατότητα να επιλέξει και να ρυθμίσει τους πόρους που αυτός χρειάζεται. Παράδειγμα μιας τέτοιας υποδομής υπηρεσίας είναι αυτές της Amazon.

Το μοντέλο υπηρεσιών IBM περιλαμβάνει ένα επιπλέον επίπεδο σε σχέση με το μοντέλο SPI. Το επίπεδο αυτό ονομάζεται Business Process as a Service/Υπηρεσία Επιχειρηματικής Διαδικασίας (BPaaS) και επιτρέπει στον χρήστη, να σχεδιάζει, να διαχειρίζεται και να ενσωματώνει συνεργατικές δραστηριότητες. Οι συνεργατικές δραστηριότητες συνδέονται άμεσα με υπηρεσίες SaaS. Υπηρεσίες που παρέχονται από το συγκεκριμένο επίπεδο είναι οι διαδικασίες δοκιμής λογισμικού και διαχείρισης των επιδομάτων των εργαζομένων.

4.2.1 Anaconda Cloud

Το Anaconda Cloud είναι μια υπηρεσία υπολογιστικού νέφους διανομής και περιβάλλοντος εργασίας της Python, που εμπίπτει στην κατηγορία του PaaS και, αποτελεί ίσως την πιο δημοφιλή πλατφόρμα νέφους για την υλοποίηση εφαρμογών τόσο της Python όσο και της γλώσσας R (ANACONDA CLOUD, 2018). Βασικό πλεονέκτημα της υπηρεσίας είναι η παροχή ισχυρής συνεργασίας και διαχείρισης πακέτων τόσο για έργα ανοιχτού κώδικα όσο και ιδιωτικά.

Όπως αναφέρεται και στην επίσημη ιστοσελίδα της υπηρεσίας, πρόκειται για μια υπηρεσία ανοιχτού κώδικα με πάνω από 6 εκατομμύρια χρήστες και με κύριο στόχο την διευκόλυνση της υλοποίησης προγραμμάτων σε σχέση με Data Science και εκμάθησης μηχανών σε Python αλλά και την συνεργασία σε πραγματικό χρόνο. Η διευκόλυνση πραγματοποιείται μέσω της δυνατότητας χρήσης περισσότερων από 250 πακέτων δημοφιλών δεδομένων, του πακέτου conda και του εικονικό περιβάλλοντος διαχείρισης για διάφορα λειτουργικά συστήματα (Windows, Linux και MacOS).

Ειδικότερα το πακέτο Conda καθιστά εύκολη την εγκατάσταση, εκτέλεση και αναβάθμιση πολύπλοκων επιστημονικών δεδομένων παρέχοντας επιπλέον περιβάλλοντα μηχανικής μάθησης όπως το Scikit-learn, το TensorFlow και το SciPy. Το εργαλείο conda του Anaconda απλοποιεί την εγκατάσταση πακέτων Python που χρησιμοποιούνται στην διαχείριση στατιστικών στοιχείων και στην ανάλυση δεδομένων, επειδή πολλά από αυτά τα πακέτα έχουν πολύπλοκες δυαδικές εξαρτήσεις.

Η υπηρεσία περιλαμβάνει επίσης μια μεγάλη κοινότητα (την Conda-Forge) με κύριο στόχο την κατασκευή πακέτων conda. Περίπου 3.200 πακέτα από την Conda-Forge είναι διαθέσιμα με πιο πρόσφατα ενημερωμένα τα παρακάτω:

cassandra-driver, μονάδα Python για εργασία με το Apache Cassandra και το δυαδικό πρωτόκολλο πρόσβασης δεδομένων.

pyinstaller, για την μεταμόρφωση μιας εφαρμογής Python ως ανεξάρτητου εκτελέσιμου αρχείου.

plotly, μια διαδραστική βιβλιοθήκη γραφικών.

openblas, μια βιβλιοθήκη για βασικά διανύσματα και μήτρες.

Το αποθετήριο της Anaconda περιέχει και διανέμει δωρεάν πάνω από 1000 πακέτα Python και R. Τα πακέτα όπως και κάθε υλοποίηση συντάσσονται και εκτελούνται σε ένα ασφαλές περιβάλλον παράγοντας βελτιστοποιημένα δυαδικά αρχεία που λειτουργούν με πανομοιότυπο τρόπο σε οποιοδήποτε λειτουργικό σύστημα.

Η υπηρεσία παρέχει επίσης και ένα εμπορικό προϊόν, το Anaconda Enterprise το οποίο είναι ένα λογισμικό που απευθύνεται σε οργανισμούς επιχειρήσεων διαχείρισης δεδομένων με Python και R. Με τη χρήση του λογισμικού δίνεται η δυνατότητα ασφαλούς συνεργασίας στους χρήστες αλλά και η αποθήκευση των παραγόμενων πακέτων σε ιδιωτικό χώρο αποθήκευσης. Μέσω του λογισμικού μπορεί να γίνει ταχεία ανάπτυξη μοντέλων ML και πινάκων ελέγχου για την παραγωγή συμπλεγμάτων υπολογιστών με Docker / Kubernetes, Hadoop και Spark τα οποία και “τρέχουν” σε τοπικά κέντρα δεδομένων αλλά και στις πλατφόρμες AWS, Azure και Google Cloud Platform.

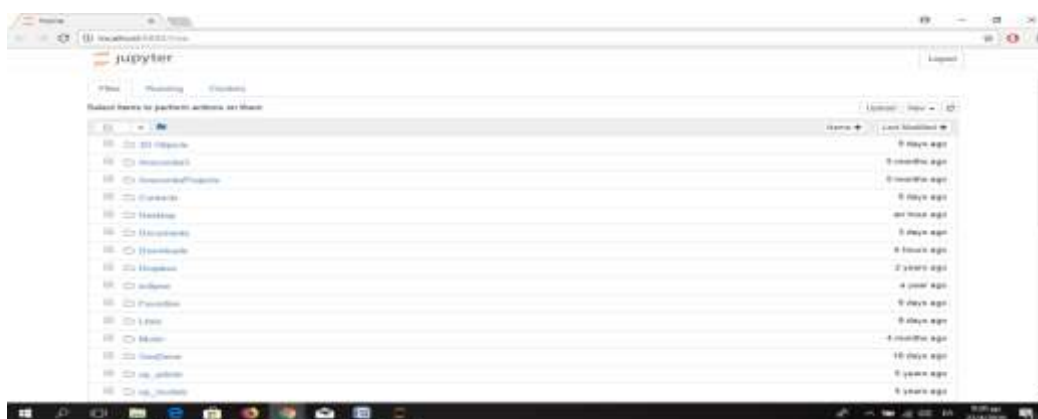
Η τρέχουσα έκδοση της υπηρεσίας Anaconda είναι η 5.1 και διατίθεται δωρεάν στο (ANACONDA DOWNLOADS, 2018).

4.2.2 Anaconda Jupyter Editor

Η εφαρμογή Jupyter Notebook είναι μια on-line εφαρμογή (διακομιστή-πελάτη) σημειωματάριου που επιτρέπει τόσο την επεξεργασία όσο και την εκτέλεση εγγράφων μέσω ενός φυλλομετρητή και χρησιμοποιείται στο Anaconda Cloud για την σύνταξη

προγραμμάτων (Jupyter Notebook, 2015). Η εφαρμογή Jupyter Notebook μπορεί να εκτελεστεί τόσο σε έναν τοπικό υπολογιστή, χωρίς την ανάγκη πρόσβασης στο Διαδίκτυο, όσο σε έναν απομακρυσμένο διακομιστή (με απαραίτητη την πρόσβαση στο Διαδίκτυο).

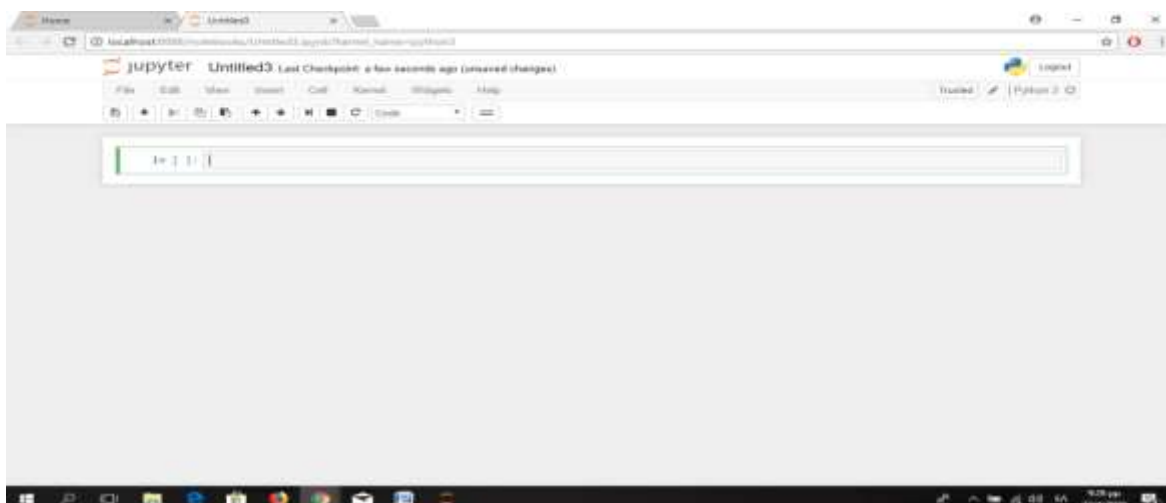
Ειδικότερα, η εφαρμογή, επιτρέπει την εμφάνιση, την επεξεργασία και γενικότερα τις λειτουργίες σημειωματάρων ενώ περεταίρω, διαθέτει έναν πίνακα ελέγχου (κατάλογο αρχείων) που περιέχει τα τοπικά αρχεία και επιτρέπει την πρόσβαση σε έγγραφα Notebook ή τον τερματισμό των πυρήνων τους.



Εικόνα 4.7: Εκκίνηση του Jupyter και ο κατάλογος εγγράφων

Ο πυρήνας αυτών των αρχείων από την άλλη είναι ο υπεύθυνος “μηχανισμός” για την εκτέλεση του κώδικα που περιέχεται σε ένα έγγραφο. Στην εφαρμογή διατίθενται επίσημοι πυρήνες για πολλές γλώσσες προγραμματισμού ενώ για την Python διατίθεται ο πυρήνας Ipython. Πιο συγκεκριμένα, κατά την εκκίνηση ενός εγγράφου εκκινείτε αυτόματα και ο συνδεδεμένος πυρήνας. Κατά την φάση εκτέλεσης ενός αρχείου³⁰, ο πυρήνας εκτελεί τον υπολογισμό και παράγει τα αποτελέσματα. Ανάλογα με τον τύπο των υπολογισμών, ο πυρήνας μπορεί να καταναλώνει σημαντική CPU και RAM.

³⁰ Η εκτέλεση μπορεί να γίνει με δύο τρόπους, είτε cell-by-cell είτε με το μενού Cell -> Run All



Εικόνα 4.8: Περιβάλλον συγγραφής και εκτέλεσης του Jupyter (cell)

4.3 Βιβλιοθήκες

Μια βιβλιοθήκη λογισμικού είναι αρχείο κώδικα που χρησιμοποιείται για την ανάπτυξη προγραμμάτων και εφαρμογών λογισμικού και έχει σχεδιαστεί για να βοηθήσει τόσο τον προγραμματιστή όσο και τον μεταγλωττιστή της γλώσσας.

Μια βιβλιοθήκη λογισμικού αποτελείται γενικά από προ-συντεταγμένο κώδικα, κλάσεις, διαδικασίες, σενάρια, δεδομένα διαμόρφωσης και άλλα. Η βιβλιοθήκη προστίθεται με μη αυτόματο τρόπο σε ένα πρόγραμμα, με κύριο σκοπό τη μεγαλύτερη λειτουργικότητα και την αυτοματοποίηση μιας διαδικασίας χωρίς να χρειάζεται η σύνταξη κώδικα. Για παράδειγμα, κατά την ανάπτυξη ενός μαθηματικού προγράμματος ή μιας εφαρμογής, ένας προγραμματιστής μπορεί να προσθέσει μια βιβλιοθήκη μαθηματικών στο πρόγραμμα για να εξαλείψει την ανάγκη σύνταξης σύνθετων λειτουργιών. Όλες οι διαθέσιμες λειτουργίες μέσα σε μια βιβλιοθήκη λογισμικού μπορούν απλά να καλούνται και στην συνέχεια να χρησιμοποιούνται μέσα στο σώμα του προγράμματος χωρίς την ανάγκη επαναπρογραμματισμού τους. Ομοίως, ένας μεταγλωττιστής μπορεί να προσθέσει αυτόματα μια σχετική βιβλιοθήκη λογισμικού σε ένα πρόγραμμα κατά τον χρόνο εκτέλεσης.

Στην συνέχεια αναφέρονται διεξοδικά οι βιβλιοθήκες της γλώσσας Python που χρησιμοποιήθηκαν στην υλοποίηση του επόμενου κεφαλαίου.

4.3.1 Βιβλιοθήκη Numpy

Η βιβλιοθήκη NumPy (Numeric Python) είναι ένα πακέτο Python. Πρόκειται για μια βιβλιοθήκη που αποτελείται από πολυδιάστατα αντικείμενα και μια συλλογή ρουτινών για την επεξεργασία μονοδιάστατων πινάκων (array ή συστοιχία).

Μέσω της χρήσης της βιβλιοθήκης NumPy παρέχονται διάφορες λειτουργίες όπως:

- Μαθηματικές και λογικές πράξεις σε συστοιχίες.
- Μετασχηματισμοί Fourier και ρουτίνες για χειρισμό σχημάτων.
- Λειτουργίες σχετικές με τη γραμμική άλγεβρα. Η NumPy διαθέτει ενσωματωμένες λειτουργίες για τη γραμμική άλγεβρα και την παραγωγή τυχαίων αριθμών.

Η NumPy χρησιμοποιείται συνήθως μαζί με πακέτα όπως το SciPy (Scientific Python) και το Matplotlib (βιβλιοθήκη σχεδίασης) που αναφέρεται στην συνέχεια. Αυτός ο συνδυασμός θεωρείται σήμερα ως η πιο σύγχρονη και πλήρης γλώσσα προγραμματισμού και χρησιμοποιείται ευρέως ως αντικαταστάτης του MatLab, μιας δημοφιλούς πλατφόρμας για τεχνικούς υπολογισμούς.

Η ενσωμάτωση της βιβλιοθήκης γίνεται μέσω της εντολής:

```
import numpy as np
```

Το πιο σημαντικό αντικείμενο της NumPy είναι ένας N-διάστατος πίνακας που ονομάζεται Narray και αποτελεί μια συλλογή αντικειμένων του ίδιου τύπου. Η πρόσβαση στα στοιχεία του αντικειμένου, όπως και σε όλα τα array, πραγματοποιείται χρησιμοποιώντας μηδενικό δείκτη. Κάθε στοιχείο του Narray καταλαμβάνει στη μνήμη μπλοκ του ίδιου μεγέθους και είναι ένα αντικείμενο του ίδιου τύπου δεδομένων (που ονομάζεται dtype). Ένα αντικείμενο τύπου δεδομένων εξαρτάται από τα παρακάτω:

- Τύπο δεδομένων (ακέραιος, κινητής υποδιαστολής ή αντικείμενο Python)

- Μέγεθος δεδομένων
- Τάξη των byte (little-endian ή big-endian)
- Σε περίπτωση δομημένων τύπων, εξαρτάται από τα ονόματα των πεδίων, τον τύπο δεδομένων κάθε πεδίου και από το μέρος του μπλοκ μνήμης που χρησιμοποιείται από κάθε πεδίο.
- Εάν ο τύπος δεδομένων είναι ένα υπο-array, εξαρτάται από τον τύπο δεδομένων

Η NumPy περιέχει ένα μεγάλο αριθμό διαφορετικών μαθηματικών συναρτήσεων όπως τριγωνομετρικές (`arcsin()`, `arcsin()` και `arctan()`), συναρτήσεις για αριθμητικές πράξεις (`power()`, `mod()`, `real()`, `imag()`, `conj()`, `angle()`), χειρισμό σύνθετων αριθμών (`around()`, `floor()`, `ceil()`) κλπ.

Η βιβλιοθήκη περιέχει επίσης στατιστικές συναρτήσεις για την εύρεση του μέγιστου (`amax()`), του ελάχιστου (`amin()`), την ποσοστιαία τυπική απόκλιση, τη διακύμανση, του διάμεσου (`median()`), του μέσου όρου (`mean()`), του σταθμισμένου μέσου όρου (`average()`) και άλλες.

4.3.2 Βιβλιοθήκη Pandas

Η Pandas (Panel Data) είναι μια βιβλιοθήκη ανοιχτού κώδικα της Python. Οι δομές που περιλαμβάνονται στην βιβλιοθήκη παρέχουν δυνατότητες επεξεργασίας και ανάλυσης δεδομένων με υψηλές αποδόσεις. Συνήθως χρησιμοποιείται σε συνεργασία με την μαθηματική βιβλιοθήκη NumPy.

Πριν την εισαγωγή της βιβλιοθήκης Pandas, η Python χρησιμοποιήθηκε σε μεγάλο βαθμό για την επεξεργασία και την προετοιμασία δεδομένων αλλά όχι και στην ανάλυση τους. Με την χρήση της βιβλιοθήκης επιτυγχάνονται τα βασικά βήματα επεξεργασίας και ανάλυσης δεδομένων (φόρτωση, προετοιμασία, χειρισμό, μοντελοποίηση και ανάλυση) ανεξάρτητα από την προέλευση τους.

Η Python με την Pandas, λόγω της φύσης της βιβλιοθήκης που επιτρέπει την επεξεργασία των δεδομένων, χρησιμοποιείται σε ακαδημαϊκούς αλλά και εμπορικούς τομείς, όπως χρηματοοικονομικά, οικονομικά, στατιστικές κ.λπ.

Η βιβλιοθήκη Pandas μεταξύ των άλλων προσφέρει:

- Γρήγορο και αποδοτικό DataFrame με προεπιλεγμένο αλλά και εξατομικευμένο ευρετήριο/κατάλογο.
- Εργαλεία για τη φόρτωση δεδομένων εντός μνήμης από διαφορετικές μορφές αρχείων.
- Ευθυγράμμιση δεδομένων και ολοκληρωμένο χειρισμό για δεδομένα που λείπουν.
- Επαναδιαμόρφωση και περιστροφή των data sets.
- Επεξεργασία και κατηγοριοποίηση με ετικέτες, ευρετηρίαση και καταμερισμός σε υποσύνολα μεγάλων συνόλων δεδομένων.
- Οι στήλες από μια δομή δεδομένων μπορούν να διαγραφούν ή να εισαχθούν.
- Ομαδοποίηση με βάση στοιχεία για ενσωμάτωση και μετασχηματισμούς.
- Υψηλής απόδοσης συγχώνευση και σύνδεση δεδομένων.
- Λειτουργίες πάνω σε Χρονικές Σειρές

Η βιβλιοθήκη, επίσης, λειτουργεί πάνω σε τρεις δομές δεδομένων: σε σειρές (μονοδιάστατα array χωρίς δυνατότητα προσθήκης/αφαίρεσης στοιχείων), DataFrames (δισδιάστατοι πίνακες με δυνατότητα προσθήκης/αφαίρεσης στοιχείων και με πιθανώς ετερογενείς στοιχεία στις στήλες) και Panel (τριδιάστατοι πίνακες με δυνατότητα προσθήκης/αφαίρεσης στοιχείων). Η δομή DataFrame είναι η πιο διαδεδομένη λόγω τρόπου ορισμού της.

Η λειτουργίες που παρέχονται από την βιβλιοθήκη πάνω στις σειρές είναι:

Χαρακτηριστικό ή μέθοδος	Περιγραφή
axes	Επιστρέφει μια λίστα με όλες τις ετικέτες του άξονα
dtype	Επιστρέφει τον τύπο δεδομένων ενός αντικειμένου
empty	Επιστρέφει την τιμή True if series is empty αν η σειρά είναι κενή
ndim	Επιστρέφει τον αριθμό των διαστάσεων των υποκείμενων δεδομένων
size	Επιστρέφει τον αριθμό των υποκείμενων δεδομένων

values	Επιστρέφει την σειρά σε μορφή μονοδιάστατου array
head (n)	Επιστρέφει τις πρώτες n γραμμές
tail (n)	Επιστρέφει τις τελευταίες n γραμμές

Πίνακας 4-1: Λειτουργίες σειρών

Από την άλλη, οι λειτουργίες που παρέχονται πάνω στα DataFrames είναι:

Χαρακτηριστικό ή μέθοδος	Περιγραφή
T	Μεταφορά και ανταλλαγή γραμμών και στηλών μεταξύ τους
axes	Επιστρέφει μια λίστα με τις ετικέτες του άξονα γραμμής και τις ετικέτες του άξονα στήλης
dtypes	Επιστρέφει τους τύπους των δεδομένων
empty	Επιστρέφει σαν τιμή True αν το NDFrame είναι κενό
ndim	Επιστρέφει τον αριθμό των αξόνων δηλ την διάσταση του πίνακα
shape	Επιστρέφει μια τριάδα που αντιπροσωπεύει την διάσταση του DataFrame.
size	Επιστρέφει τον Αριθμό στοιχείων του NDFrame.
Values	Επιστρέφει κάθε γραμμή σαν μονοδιάστατο array
head (n)	Επιστρέφει τις πρώτες n γραμμές
tail (n)	Επιστρέφει τις τελευταίες n γραμμές

Πίνακας 4-2: Λειτουργίες DataFrames

Η βιβλιοθήκη παρέχει μια σειρά μεθόδων για την υλοποίηση στατιστικών πάνω σε δεδομένα όπως οι: `sum()`- άθροισμα στηλών, `mean()`-μέσος όρος, `median()`- διάμεσος, `std()`- τυπική απόκλιση, `count()`-πλήθος και άλλες.

Για την κλήση/εισαγωγή της βιβλιοθήκης στο πρόγραμμα χρησιμοποιείται η εντολή:

```
import pandas as pd
```

4.3.3 Βιβλιοθήκη Matplotlib

Η Matplotlib (matplotlib, 2002-2012) είναι μια βιβλιοθήκη σχεδίασης γραφημάτων 2D της Python. Η σχεδίαση γίνεται μέσω της συνάρτησης `plot()`. Η βιβλιοθήκη παράγει ποιοτικά στοιχεία δημοσίευσης σε διάφορες μορφές (έντυπου αλλά και διαδραστικά περιβάλλοντα σε διάφορες πλατφόρμες). Η Matplotlib μπορεί να χρησιμοποιηθεί σε script της Python, στα κέλυφη Python και IPython, στο σημειωματάριο Jupyter που ήδη αναφέρθηκε και χρησιμοποιείται στην εφαρμογή, στους διακομιστές εφαρμογών ιστού και σε εργαλεία υλοποίησης γραφικών διεπαφών με τον χρήστη.

Με την χρήση της βιβλιοθήκης δίνεται η δυνατότητα δημιουργίας διαγραμμάτων, ιστογραμμάτων, γραμμικών γραφημάτων, διαγραμμάτων σφάλματος, scatterplots κ.λπ., με μερικές μόνο γραμμές κώδικα.

Η Matplotlib περιλαμβάνει μια σειρά προεπιλεγμένων ρυθμίσεων που επιτρέπουν την προσαρμογή όλων των ιδιοτήτων: μέγεθος σχήματος και dpi, πλάτος γραμμής, χρώμα και στυλ, άξονες, ιδιότητες άξονα και πλέγματος, ιδιότητες κειμένου και γραμματοσειράς κ.ο.κ.

Η βιβλιοθήκη περιλαμβάνει επίσης το `pyplot` που παρέχει μια διεπαφή τύπου MATLAB, ιδιαίτερα όταν συνδυάζεται με την IPython. Παρέχει επιπλέον την δυνατότητα πλήρη ελέγχου των γραμμικών στυλ, των γραμματοσειρών, των ιδιοτήτων των αξόνων. Η συγκεκριμένη δυνατότητα δίνεται μέσω ενός αντικειμενοστραφούς διεπαφής ή μέσω ενός συνόλου λειτουργιών που είναι γνωστές σε χρήστες του MATLAB.

Η τρέχουσα έκδοση της βιβλιοθήκης είναι η v2.2.2, ενώ η ενσωμάτωση της γίνεται μέσω της εντολής:

```
import matplotlib.pyplot as plt
```

Η χρήση της βιβλιοθήκης μπορεί να εξελιχθεί σε μια δύσκολη διαδικασία για πολλούς λόγους. Η βιβλιοθήκη είναι τεράστια, περίπου 70.000 συνολικά γραμμές κώδικα, ενώ φιλοξενεί αρκετές διαφορετικές διεπαφές και τρόπους κατασκευής μιας εικόνας και μπορεί να αλληλεπιδράσει με διαφορετικά backends. Ενώ περιέχει μια μεγάλη και επαρκή τεκμηρίωση κάποια κομμάτια της είναι παλιά.

4.3.4 Βιβλιοθήκες ggplot και ggplot2

Η ggplot2 είναι ένα πακέτο γραφικών της R και της Python για την παραγωγή γραφημάτων πάνω σε στατιστικά ή δεδομένα. Η βιβλιοθήκη, σε αντίθεση με άλλα πακέτα γραφικών, είναι υλοποιημένη πάνω στην Grammar of Graphics (Wilkinson, 2005) και αποτελείται από ένα σύνολο ανεξάρτητων στοιχείων που μπορούν να συνδυαστούν με ποικίλους τρόπους. Τα πεδία (plots) μπορούν να δημιουργηθούν επαναληπτικά και να επεξεργαστούν ακόμη και μετά την δημιουργία τους. Έχοντας επιλέξει προσεκτικά το σύνολο δεδομένων μπορεί να δημιουργηθεί ένα γραφικό υψηλής ποιότητας σε ελάχιστο χρόνο.

Η ggplot2 (Wickham, 2009) έχει σχεδιαστεί για να λειτουργεί με πολυεπίπεδο τρόπο, ξεκινώντας με ένα στρώμα που εμφανίζει τα ακατέργαστα δεδομένα και στη συνέχεια τελειοποιούνται προσθέτοντας στρώματα σχολιασμού και στατιστικές περιλήψεις. Η βιβλιοθήκη παρέχει επίσης την δυνατότητα δημιουργίας πολυεπίπεδων γραφικών αλλά και 3D γραφικών.

Η βιβλιοθήκη επιτρέπει την δημιουργία, με απλό τρόπο, γραφημάτων τόσο πάνω σε απλά δεδομένα όσο και πάνω σε πολύπλοκα μεταβλητά αριθμητικά και κατηγορηματικά δεδομένα. Οπτικά η ομαδοποίηση δεδομένων μπορεί να οπτικοποιείται μέσω χρωμάτων, συμβόλων και αχνότητας. Επίσης η δημιουργία ενός plot με δεδομένα υπό συνθήκες είναι σχετικά απλή.

Τα χαρακτηριστικά των γραφικών της ggplot2 μπορούν να τροποποιηθούν χρησιμοποιώντας τη λειτουργία `theme()` και προσθέτοντας γραφικές παραμέτρους μέσα

στη συνάρτηση `qplot()`. Η συνάρτηση `ggplot()` παρέχει περισσότερες επιλογές ελέγχου. Οι λειτουργίες `ggplot2` μπορούν να ακολουθούν η μια την άλλη (με μορφή αλυσίδας), για την δημιουργία της τελικής γραφικής παράστασης, χρησιμοποιώντας το σύμβολο “+”.

Όλα τα `plot` που παράγονται από την `ggplot2` με κλήση της συνάρτησης `ggplot()`, παρέχοντας προεπιλεγμένα δεδομένα και αισθητικές αντιστοιχίσεις, καθορίζονται από την συνάρτηση `aes()`. Στη συνέχεια προστίθενται στρώματα, κλίμακες, συντεταγμένες και πτυχές με την χρήση του συμβόλου “+”. Για την αποθήκευση ενός γραφήματος στο δίσκο, χρησιμοποιείται η συνάρτηση `ggsave()`.

Έτσι, οι βασικές συναρτήσεις της `ggplot2` είναι οι εξής:

Συνάρτηση	Περιγραφή
<code>ggplot</code>	Δημιουργεί ένα νέο πεδίο
<code>Aes</code>	Κατασκευή αισθητικών χαρτογραφημάτων
<code>+ .gg</code>	Προσθέτει συνιστώσες σε ένα γράφημα
<code>ggsave</code>	Αποθήκευση του γραφήματος στο δίσκο
<code>qplot</code> <code>quickplot</code>	Γρήγορο <code>plot</code>

Πίνακας 4-3: Βασικές συναρτήσεις της `ggplot2`

Ένα γεωμετρικό επίπεδο, που καθορίζει ένα μέρος του `plot`, συνδυάζει δεδομένα, αισθητική αντιστοίχιση, `geom` (γεωμετρικό αντικείμενο), `stat` (στατιστικό μετασχηματισμό) και προσαρμογή θέσης. Συνήθως, δημιουργούνται περισσότερα επίπεδα χρησιμοποιώντας μια γεωμετρική συνάρτηση, υπερκαλύπτοντας την προεπιλεγμένη θέση και τον στατικό μετασχηματισμό αν χρειαστεί. Η υλοποίηση ενός γεωμετρικού πεδίου `plot` μπορεί να εμπλέκει μια ή και περισσότερες από τις παρακάτω γεωμετρικές συναρτήσεις:

Συνάρτηση	Περιγραφή
<code>geom_abline geom_hline geom_vline</code>	Γραμμές αναφοράς: οριζόντια, κάθετη και διαγώνια αντίστοιχα
<code>geom_bar geom_col stat_count</code>	Γραφήματα στήλης
<code>geom_bin2d stat_bin_2d</code>	Αντιστοίχιση δισδιάστατων αποθηκευμένων δεδομένων
<code>geom_blank</code>	Σχεδίαση κενού πεδίου
<code>geom_boxplot stat_boxplot</code>	Σχεδιασμός του πεδίου με περίγραμμα (σε στυλ Tukey)
<code>geom_contour stat_contour</code>	Δισδιάστατο περίγραμμα σε τρισδιάστατη επιφάνεια
<code>geom_count stat_sum</code>	Καταμέτρηση επικαλυπτόμενων σημείων
<code>geom_density_2d stat_density_2d</code>	Περίγραμμα δισδιάστατης εκτίμησης πυκνότητας
<code>geom_density stat_density</code>	Ρυθμίσεις εκτίμησης πυκνότητας
<code>geom_dotplot</code>	Σχεδιασμός γραφήματος με κουκίδες
<code>geom_errorbarh</code>	Οριζόντιες γραμμές σφαλμάτων
<code>geom_hex stat_bin_hex</code>	Εξάγωνη χαρτογράφηση δισδιάστατων δεδομένων
<code>geom_freqpoly geom_histogram stat_bin</code>	Ιστογράμματα και πολύγωνα συχνότητας
<code>geom_jitter</code>	Σχεδίαση γραφημάτων με διακύμανση
<code>geom_crossbar geom_errorbargeom_linerange geom_pointrange</code>	Κάθετα διαστήματα: γραμμές, τομές και γραμμές σφαλμάτων

<code>geom_map</code>	Πολύγωνα κατασκευασμένα από ένα χάρτη αναφοράς
<code>geom_path</code> <code>geom_line</code> <code>geom_step</code>	Σύνδεση παρατηρήσεων
<code>geom_point</code>	Σημεία
<code>geom_polygon</code>	Πολύγωνα
<code>geom_qq</code> <code>stat_qq</code>	Σχεδιασμός πεδίου ποσοτικής-ποιοτικής συσχέτισης
<code>geom_quantile</code> <code>stat_quantile</code>	Ποσοτική παλινδρόμηση
<code>geom_ribbon</code> <code>geom_area</code>	Ribbons και περιοχή πεδίου
<code>geom_rug</code>	Υπόστρωμα περιθωρίων στα πεδία
<code>geom_segment</code> <code>geom_curve</code>	Γραμμές, διανύσματα και καμπύλες
<code>geom_smooth</code> <code>stat_smooth</code>	Εξομάλυνση υπό όρους
<code>geom_spoke</code>	Τα διανύσματα γραμμής παραμετροποιούνται ανάλογα με την τοποθεσία, την κατεύθυνση και την απόσταση
<code>geom_label</code> <code>geom_text</code>	Κείμενο
<code>geom_raster</code> <code>geom_rect</code> <code>geom_tile</code>	Παραλληλόγραμμα
<code>geom_violin</code> <code>stat_ydensity</code>	Violin πεδία
<code>stat_sf</code> <code>geom_sf</code> <code>coord_sf</code>	Οπτικοποίηση sf αντικειμένων

Πίνακας 4-4: Τιμές τροποποίησης προεπιλογής επικαλυπτόμενων αντικειμένων

Τα πεδία στατιστικών αποτελεσμάτων και τα επίπεδα τους υλοποιούνται πιο εύκολα με την χρήση ορισμένων στατιστικών συναρτήσεων, οι οποίες δίνουν έμφαση στον

στατιστικό μετασχηματισμό και όχι στην οπτική εμφάνιση. Οι στατιστικές συναρτήσεις τις βιβλιοθήκης είναι οι:

Συνάρτηση	Περιγραφή
<code>stat_ecdf</code>	Υπολογισμός της εμπειρικής σωρευτικής κατανομής
<code>stat_ellipse</code>	Υπολογισμός των κανονικών ελλείψεων εμπιστοσύνης
<code>stat_function</code>	Υπολογισμός της τιμής της συνάρτησης για κάθε τιμή του x
<code>stat_identity</code>	Αφήνει τα δεδομένα αμετάβλητα
<code>stat_summary_2d</code> <code>stat_summary_hex</code>	Τα δεδομένα συνοψίζονται σε δύο διαστάσεις με μορφή ορθογωνίου ή εξαγώνου
<code>stat_summary_bin</code> <code>stat_summary</code>	Συγκεντρώνει τις τιμές y της συνάρτησης για μοναδικά / συνδεδεμένα στοιχεία x
<code>stat_unique</code>	Αφαίρεση των διπλότυπων

Πίνακας 4-5: Στατιστικές συναρτήσεις της `ggplot2`

Όλα τα επίπεδα διαθέτουν μια ρύθμιση θέσης (προεπιλογή) που επιλύει τις περιπτώσεις των επικαλυπτόμενων γεωμετρικών αντικειμένων. Η κατάργηση της προεπιλογής πραγματοποιείται χρησιμοποιώντας το αντίστοιχο όρισμα θέσης στη γεωμετρική ή στατιστική συνάρτηση. Σε αυτές τις συναρτήσεις το όρισμα θέσης μπορεί να πάρει τις παρακάτω τιμές:

Τιμή θέσης	Περιγραφή
<code>position_dodge</code>	Αποφυγή επικάλυψης αντικειμένων που βρίσκονται δίπλα-δίπλα
<code>position_identity</code>	Αφήνει την θέση του αντικειμένου αμετάβλητη

<code>position_jitter</code>	Διακύμανση των σημείων προς αποφυγή της υπερκάλυψης τους
<code>position_jitterdodge</code>	Ταυτόχρονη αποφυγή και διακύμανση
<code>position_nudge</code>	Η μετατόπιση των σημείων είναι σε μια σταθερή απόσταση
<code>position_stack position_fill</code>	Στοιβάζει τα επικαλυπτόμενα αντικείμενα το ένα πάνω στο άλλο

Πίνακας 4-6: Τιμές τροποποίησης προεπιλογής επικαλυπτόμενων αντικειμένων

Από άποψη αισθητικής αυτή μπορεί να χρησιμοποιηθεί με μια από τις παρακάτω τιμές:

Συνάρτηση αισθητικής	Περιγραφή
<code>aes_colour_fill_alpha</code>	Χρώμα και γέμισμα
<code>aes_group_order</code>	Ομαδοποίηση
<code>aes_linetype_size_shape</code>	Αισθητική που σχετίζεται με τη διαφοροποίηση: τύπος γραμμής, μέγεθος, σχήμα
<code>aes_position</code>	Αισθητική που σχετίζεται με την θέση: x, y, xmin, xmax, ymin, ymax, xend, yend

Πίνακας 4-7: Συναρτήσεις αισθητικής

Οι οδηγοί (άξονες και λεζάντες) βοηθούν τους αναγνώστες να ερμηνεύουν τα πεδία γραφημάτων. Οι οδηγοί ελέγχονται ως επί το πλείστον μέσω μιας κλίμακας (π.χ. με τα όρια του πεδίου). Στη βιβλιοθήκη περιέχεται ένα σύνολο συναρτήσεων για την εισαγωγή ετικετών και λεζάντων:

Συνάρτηση	Περιγραφή
<code>Labeler</code>	Κατασκευή προδιαγραφών της ετικέτας

<pre>label_value label_both label_contextlabel_parsed label_wrap_gen</pre>	Χρήσιμες συναρτήσεις ετικετών
<pre>label_bquote</pre>	Ετικέτα με μαθηματικές εκφράσεις

Πίνακας 4-8: Ετικέτες και λεζάντες

Τέλος, το σύστημα συντεταγμένων καθορίζει τον τρόπο με τον οποίο συνδυάζεται η παρουσίαση x και y με τα στοιχεία θέσης στο πεδίο. Το προεπιλεγμένο σύστημα συντεταγμένων είναι καρτεσιανό (`coord_cartesian()`), το οποίο μπορεί να τροποποιηθεί με την χρήση των παρακάτω συναρτήσεων:

Συνάρτηση	Περιγραφή
<code>coord_cartesian</code>	Καρτεσιανές συντεταγμένες
<code>coord_fixed</code>	Καρτεσιανές συντεταγμένες με σταθερό λόγο διαστάσεων
<code>coord_flip</code>	Καρτεσιανές συντεταγμένες με τα x και y αντεστραμμένα
<code>coord_map</code> <code>coord_quickmap</code>	Προβολές χάρτη
<code>coord_polar</code>	Πολικές συντεταγμένες
<code>coord_trans</code>	Μετασχηματισμένο σύστημα καρτεσιανών συντεταγμένων

Πίνακας 4-9: Συναρτήσεις τροποποίησης καρτεσιανού επιπέδου

Η βιβλιοθήκη περιλαμβάνει περαιτέρω συναρτήσεις που δεν άπτονται της παρούσας πτυχιακής.

5. ΥΛΟΠΟΙΗΣΗ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Η παρούσα πτυχιακή επεξεργάζεται την μελέτη περίπτωσης ναυτιλιακών εταιριών εισηγμένων σε ξένα χρηματιστήρια (Θεοτοκάς, και συν., 2007). Πιο συγκεκριμένα λαμβάνονται υπόψη η εταιρία AEGEAN MARINE PETROL NETWORK INC, η εταιρία COSTAMARE INC και η Euroseas εισηγμένες στο χρηματιστήριο της Νέας Υόρκης (NYSE).

Η AEGEAN MARINE PETROL NETWORK INC (AEGEAN, 2018) είναι μια ελληνική ναυτιλιακή εταιρία συμφερόντων Δημητρίου Μελισσανίδη. Η εισαγωγή της στο χρηματιστήριο της Ν. Υόρκης πραγματοποιήθηκε τον Δεκέμβριο του 2006. Στη μετοχική σύνθεσή της συμμετέχουν οι κ.κ. Peter Γεωργιόπουλος και Γιάννης Ταβλάριος. Θεωρείται ως ένας από τους κορυφαίους ανεξάρτητους φυσικούς προμηθευτές καυσίμων πλοίων και λιπαντικών πλοίων παγκοσμίως. Η εταιρία κατέχει το μεγαλύτερο και πιο σύγχρονο στόλο δεξαμενόπλοιων. Ξεκίνησε το παγκόσμιο εμπόριο των καυσίμων το 2015 και επεκτείνεται σε 27 παγκόσμιες αγορές, λειτουργώντας παράλληλα με το υπάρχον δίκτυο φυσικού εφοδιασμού του ομίλου, προσφέροντας στους πελάτες της μια ολοκληρωμένη παγκόσμια λύση στις ανάγκες τους για ανεφοδιασμό. Δραστηριοποιείται σε 51 λιμάνια σε όλο τον κόσμο, ενώ ο χώρος αποθήκευσης των δεξαμενόπλοιων της ξεπερνάει τα 1,2 εκατομμύρια κυβικά μέτρα. Διαθέτει 61 δεξαμενόπλοια ανεφοδιασμού και η μέση ηλικία του στόλου της είναι τα 10 έτη. Η μετοχή της κυμαίνεται σήμερα στα \$2,85 (18 Μαΐου 2018) ενώ το τρίτο τρίμηνο του 2017 παρουσίασε μεικτά κέρδη της τάξης των \$67,9 εκατομμυρίων. Το ίδιο χρονικό διάστημα δημιούργησε λειτουργικό εισόδημα ύψους \$6,7 εκατ. και καταγεγραμμένη καθαρή ζημία που αναλογεί στους μετόχους της Aegean, ύψους 3,8 εκατ. δολαρίων ή 0,10 δολάρια ζημία ανά μετοχή. Τέλος το τρίτο τρίμηνο του 2017 η εταιρία δημιούργησε EBITDA³¹ ύψους 15,2 εκατομμυρίων δολαρίων.

Η Costamare Inc. (Costamare, 2018) συμφερόντων της οικογένειας Β. Κωνσταντακόπουλου είναι ένας κορυφαίος διεθνής ιδιοκτήτης πλοίων τύπου container,

³¹ Ο EBITDA (Earnings Before Interest, Tax, Depreciation, and Amortization) είναι ένας δείκτης, βάσει του οποίου εμφανίζονται τα κέρδη μιας οντότητας, πριν από την αφαίρεση των τόκων, των φόρων (εισοδήματος), και των αποσβέσεων (ενσώματων και άυλων παγίων).

που εισήχθη στο χρηματιστήριο της Ν. Υόρκης τον Νοέμβριο του 2010 με σημερινή (18 Μαΐου 2018) αξία ανά μετοχή στα \$7,42. Η εισαγωγή της εταιρίας στο NYSE της απέφερε το ποσό των \$160 εκατομμυρίων αφού πούλησε 13,3 εκατομμύρια μετοχές στην τιμή των \$12 έκαστη. Μέσω των θυγατρικών της η Costamare Inc. κατέχει ένα στόλο 71 πλοίων, συμπεριλαμβανομένων 10 νέων containerships κατά παραγγελία. Η στρατηγική της Costamare είναι να χρονομετρηθεί ο στόλος Costamare σε μια γεωγραφικά διαφοροποιημένη, οικονομικά ισχυρή και πιστή ομάδα κορυφαίων εταιρειών τακτικών γραμμών. Ο στόλος Costamare έχει εξαιρετική φήμη για ομαλή, χωρίς διακοπή εξυπηρέτηση, πλησιάζοντας επίπεδα χρήσης 100% (εξαιρουμένης της προγραμματισμένης συντήρησης).

Η Euroseas είναι μια ελληνική ναυτιλιακή εταιρία συμφερόντων του εφοπλιστή Αριστείδη Πίττα εισηγμένη στο NASDAQ. Η εταιρία δημιουργήθηκε τον Μάιο του 2005 και εισήχθη στην κεφαλαιαγορά του NASDAQ υπό τον τίτλο ESEA τον Ιανουαρίου 2007. Δραστηριοποιείται στις μεταφορές ξηρού φορτίου και εμπορευματοκιβωτίων. Η Euroseas είναι θυγατρική της Eurobulk Ltd. Η εταιρεία διαθέτει στόλο 17 σκαφών. Η σημερινή αξία της μετοχής (18 Μαΐου 2018) ανέρχεται στα \$2,45.

Η υλοποίηση που πραγματοποιήθηκε βασίζεται στην μελέτη και το συνδυασμό 2 διαφορετικών υποδειγμάτων υλοποίησης οικονομικών εφαρμογών για την πραγματοποίηση και τον έλεγχο απόδοσης ενός χαρτοφυλακίου που παρουσιάζονται στα (CURTIS MILLER, 2016), (Efstathopoulos, 2017).

Στόχος της υλοποίησης είναι η συλλογή δεδομένων (μέσω της αποθήκης δεδομένων της Yahoo), η επεξεργασία και η οπτικοποίηση τους μέσω της γλώσσας προγραμματισμού Python και των βιβλιοθηκών που αυτή παρέχει.

Η ανάλυση των δεδομένων (και η οπτικοποίησή τους) περιλαμβάνει την δημιουργία:

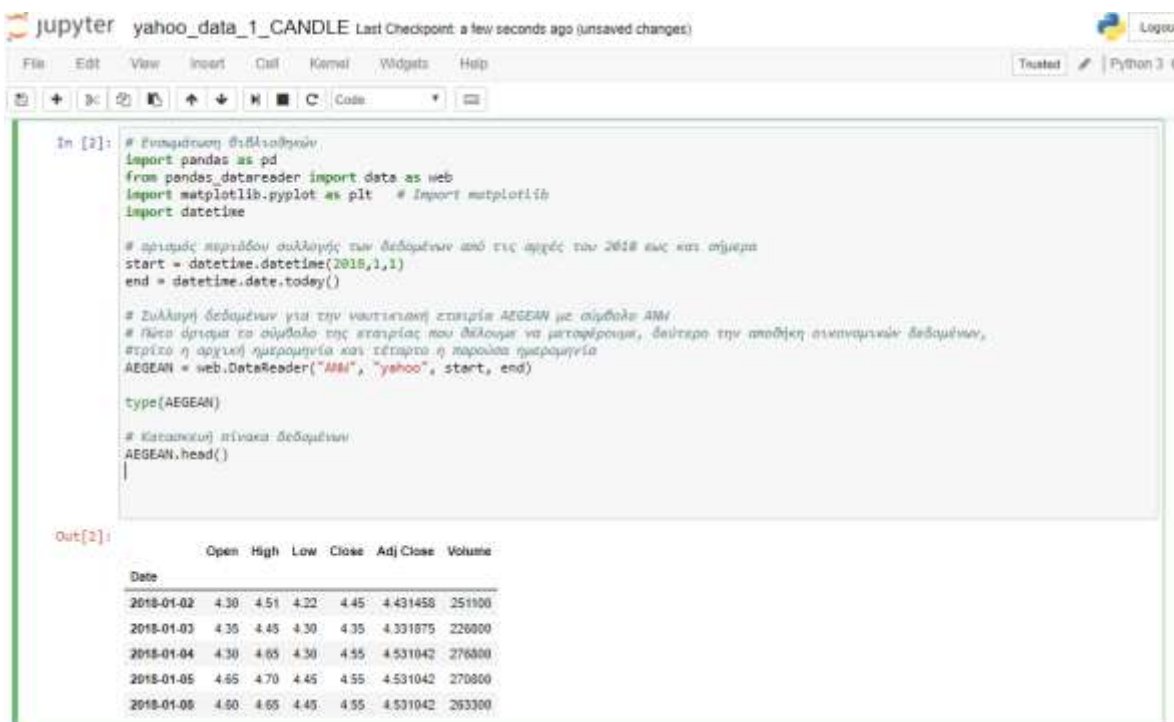
- Candlestick chart για την οπτικοποίηση των ημερών που η τιμή κλεισίματος της μετοχής μεταβάλλεται από κέρδος σε ζημία και αντίθετα
- Συγκριτικό διάγραμμα κλεισίματος των δύο διαφορετικών μετοχών.
- Διάγραμμα λογαριθμικών διαφορών ανά μέρα
- Διάγραμμα κινούμενων μέσων σε σύγκριση με το διάγραμμα candlestick και για διαφορετικά χρονικά διαστήματα

Στην συνέχεια του κεφαλαίου αναλύονται ένα προς ένα τα βήματα υλοποίησης του χαρτοφυλακίου χρήστη.

5.1 Μεταφορά, διαμόρφωση και οπτικοποίηση δεδομένων

Πριν τη χρήση δεδομένων αποθεμάτων, αυτά θα πρέπει πρώτα να συλλεχτούν και στην συνέχεια να μορφοποιηθούν έτσι ώστε να είναι δυνατή η επεξεργασία τους μέσω της βιβλιοθήκης pandas. Όπως ήδη αναφέρθηκε στο προηγούμενο κεφάλαιο, τα δεδομένα αποθεμάτων μπορούν να ληφθούν από το Yahoo! Finance, Google Finance ή μια σειρά από άλλες πηγές. Το πακέτο pandas παρέχει εύκολη πρόσβαση στο Yahoo! Finance και στοιχεία χρηματοδότησης της Google, μαζί με άλλες πηγές. Στην υλοποίηση που πραγματοποιήθηκε επιλέχθηκε σαν αποθήκη οικονομικών δεδομένων η Yahoo! Finance και σε πρώτη φάση γίνεται μεταφορά δεδομένων για την εταιρία AEGEAN.

Στον κώδικα μετά την ενσωμάτωση των απαραίτητων βιβλιοθηκών (pandas, pandas_datareader, matplotlib.pyplot και datetime), ορίζεται η χρονική περίοδος για την οποία θα γίνει η μεταφορά των δεδομένων.



```
In [2]: # Εγκατάσταση βιβλιοθηκών
import pandas as pd
from pandas_datareader import data as web
import matplotlib.pyplot as plt # Import matplotlib
import datetime

# ορισμός περιόδου συλλογής των δεδομένων από τις αρχές του 2018 ως και σήμερα
start = datetime.datetime(2018,1,1)
end = datetime.date.today()

# Συλλογή δεδομένων για την ναυτιλιακή εταιρία AEGEAN με σύμβολο AMW
# Πάντα όριζουμε το σύμβολο της εταιρίας που θέλουμε να μεταφέρουμε, δεύτερο την αποθήκη οικονομικών δεδομένων,
στρίξο η αρχική ημερομηνία και τρίτο η παρούσα ημερομηνία
AEGEAN = web.DataReader("AMW", "yahoo", start, end)

type(AEGEAN)

# Κατασκευή πίνακα δεδομένων
AEGEAN.head()
```

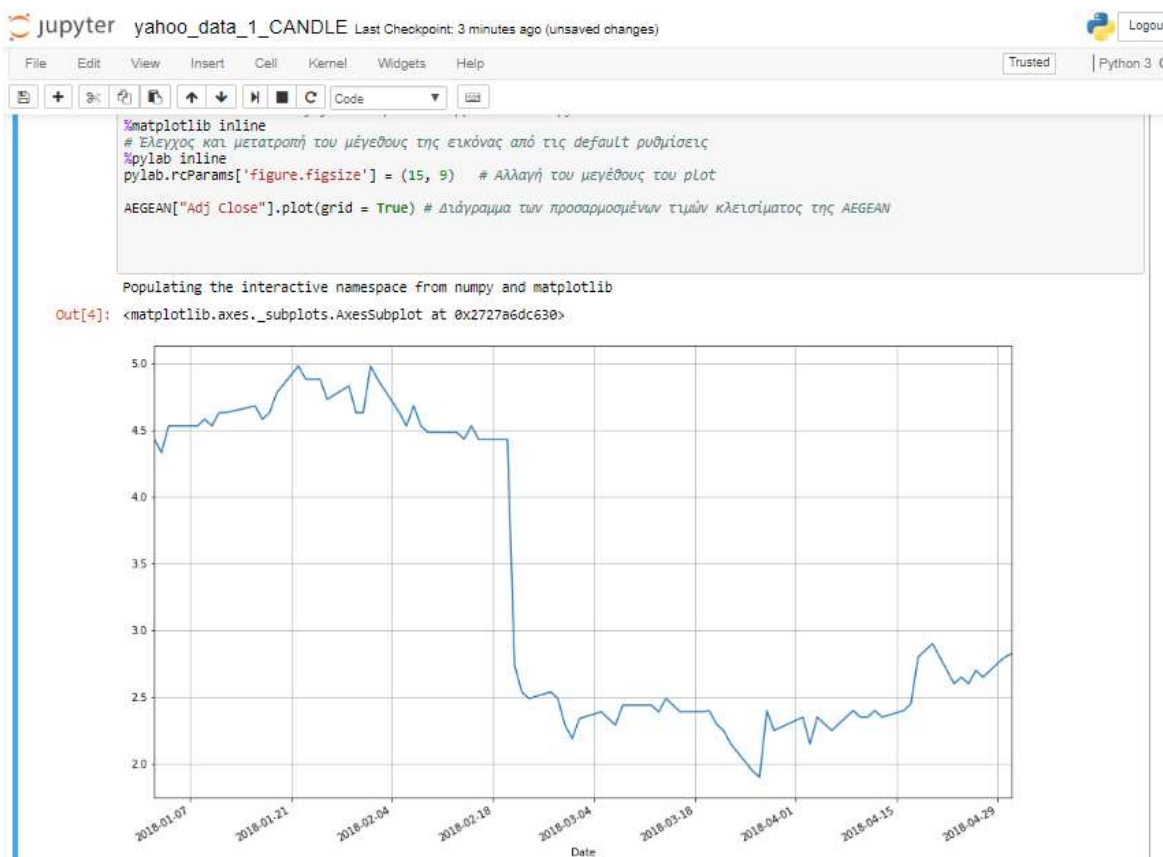
Out[2]:

	Open	High	Low	Close	Adj Close	Volume
Date						
2018-01-02	4.30	4.51	4.22	4.45	4.431458	251100
2018-01-03	4.35	4.45	4.30	4.35	4.331875	226800
2018-01-04	4.30	4.65	4.30	4.55	4.531042	276800
2018-01-05	4.65	4.70	4.45	4.55	4.531042	270800
2018-01-06	4.60	4.65	4.45	4.55	4.531042	263300

Εικόνα 5.1: Κατασκευή πίνακα δεδομένων για την εταιρία AEGEAN

Όπως φαίνεται και στην εικόνα 5.1 στην έξοδο παράγεται ο πίνακας τιμών της μετοχής ανά ημέρα. Στον πίνακα περιλαμβάνονται οι τιμές ανοίγματος όπως επίσης και η υψηλότερη και η χαμηλότερη τιμή της συνεδρίας. Στον πίνακα δεδομένων περιλαμβάνεται και ο όγκος τιμή που δείχνει τον αριθμό των μετοχών που διαπραγματεύθηκαν. Η τελευταία στήλη περιλαμβάνει την τιμή κλεισίματος των αποθεμάτων της μετοχής προσαρμοσμένη στην τιμή της μετοχής για όλες τις εταιρικές πράξεις. Αυτό γιατί ενώ οι τιμές των μετοχών καθορίζονται κυρίως από τους εμπόρους, τα αποθέματα μετοχών και τα μερίσματα (αποπληρωμή κερδών της εταιρείας ανά μετοχή) επηρεάζουν επίσης την τιμή ενός αποθέματος και πρέπει να συνυπολογίζονται.

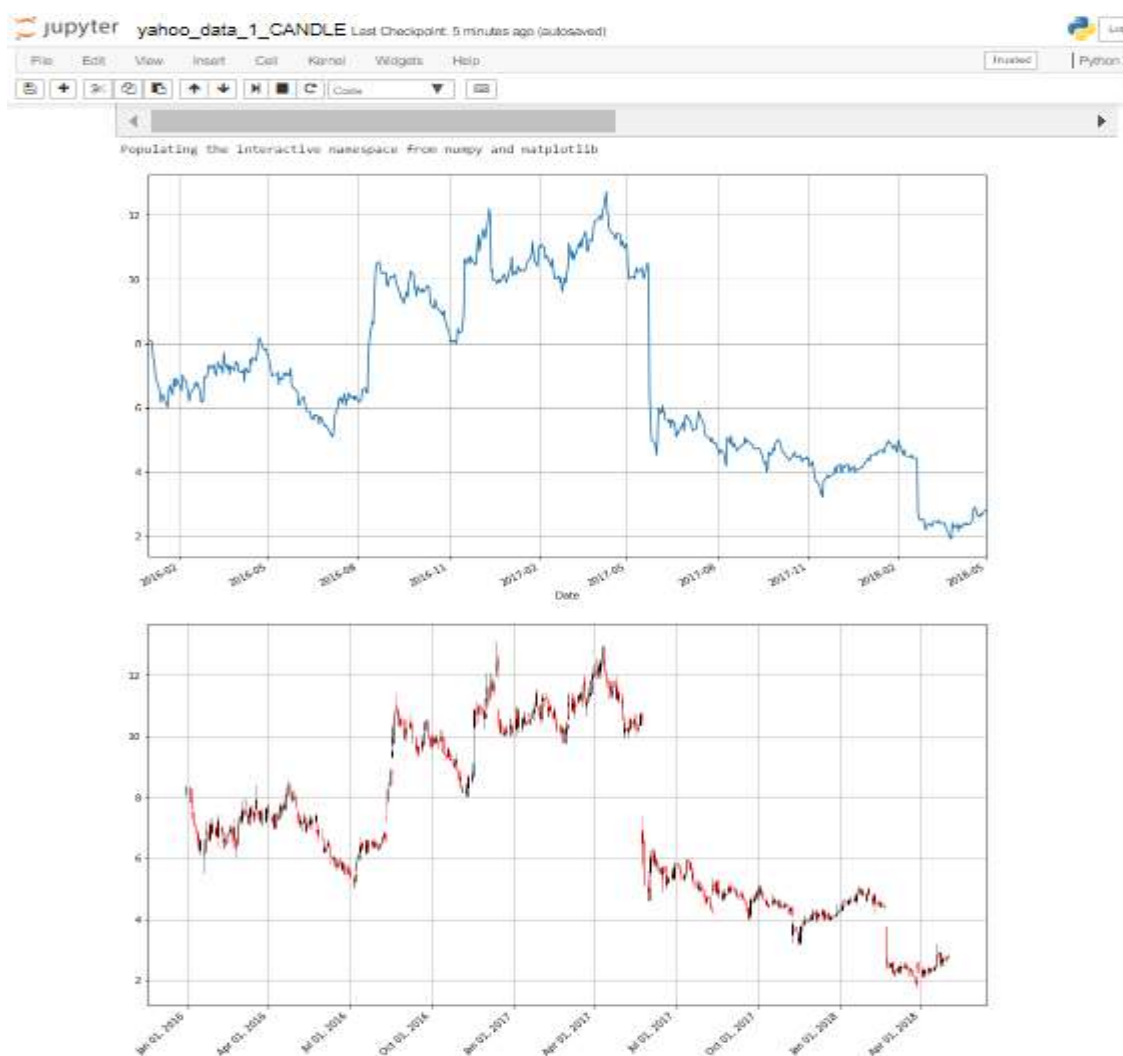
Μετά την μεταφορά των δεδομένων και την δημιουργία του πίνακα με τις απαραίτητες τιμές μπορεί να γίνει ο σχεδιασμός των προσαρμοσμένων τιμών κλεισίματος.



Εικόνα 5.2: Παράθυρο διαγράμματος των προσαρμοσμένων τιμών κλεισίματος

Το διάγραμμα της εξόδου παρουσιάζει γραφικά τις αυξομειώσεις των προσαρμοσμένων τιμών κλεισίματος του αποθέματος της μετοχής.

Η υλοποίηση ενός γραμμικού διαγράμματος πάνω σε μια μεταβλητή παρέχει περιορισμένες δυνατότητες οπτικοποίησης και χρήσης εφόσον θα πρέπει να λαμβάνονται υπόψη και οι τέσσερις μεταβλητές για κάθε ημερομηνία (ανοίγματος, υψηλή, χαμηλή και προσαρμοσμένη τιμή κλεισίματος). Συνήθως, η οπτικοποίηση οικονομικών στοιχείων πάνω σε περισσότερες μεταβλητές πραγματοποιείται μέσω των ιαπωνικών διαγραμμάτων candlestick. Τα συγκεκριμένα διαγράμματα προσφέρουν μια οπτικοποίηση, σε στήλες διαφορετικού χρώματος, ανάλογα με το αποτέλεσμα της σύγκρισης των τιμών ανοίγματος και κλεισίματος των μετοχών.



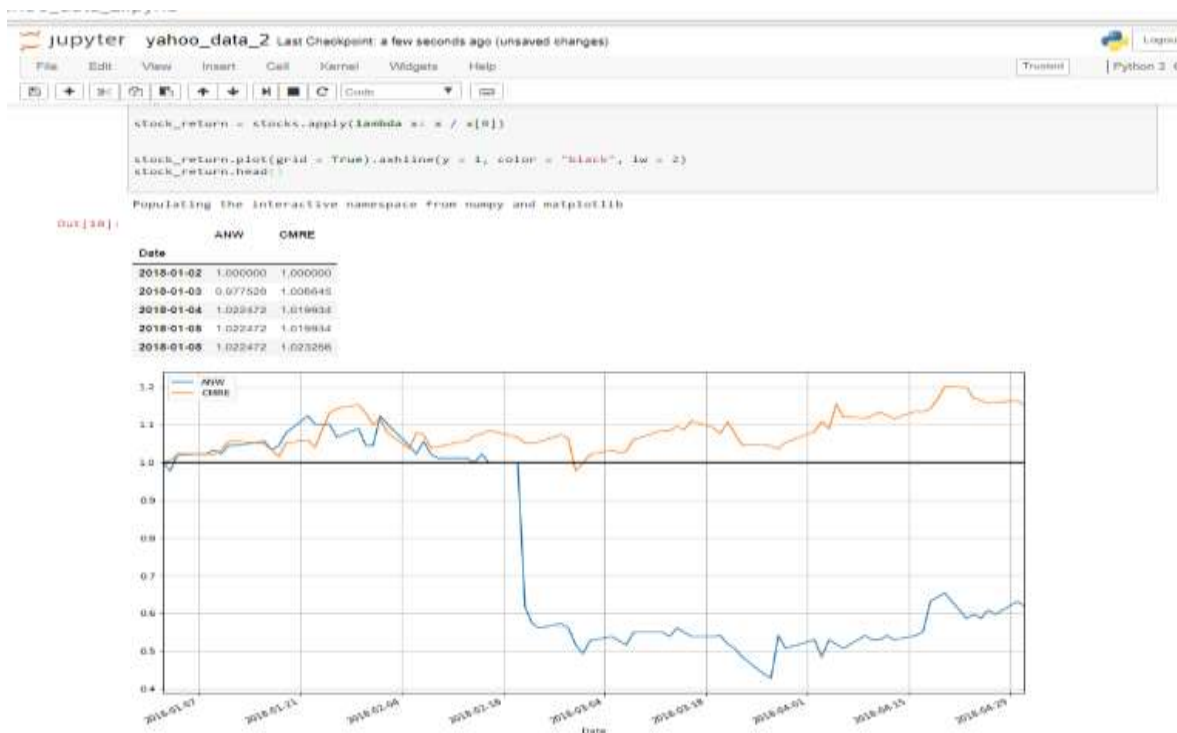
Εικόνα 5.3: Συγκριτικό διάγραμμα γραμμικού διαγράμματος και candlestick

Η έξοδος που παράχθηκε αφορά τα οικονομικά στοιχεία από το έτος 2016 μέχρι και την σημερινή ημερομηνία. Στο candlestick διάγραμμα που υλοποιήθηκε, μια *μαύρη στήλη* δηλώνει μια μέρα όπου η τιμή κλεισίματος ήταν υψηλότερη από την τιμή ανοίγματος (και άρα κέρδος), ενώ μια *κόκκινη στήλη* δηλώνει το αντίθετο (μια ημέρα όπου το άνοιγμα ήταν υψηλότερο από το κλείσιμο). Η αρχή και το τέλος των στηλών δηλώνουν το υψηλό και το χαμηλό ανάλογα με το χρώμα του σώματος.

Για την περίπτωση που υπάρχει ανάγκη σύγκρισης περισσότερων αποθεμάτων υλοποιήθηκαν πίνακες και διαγράμματα γραμμικής σύγκρισης. Η σύγκριση έγινε πάνω στις δύο ναυτιλιακές εταιρίες την AEGEAN (ANW) και COSTAMARE (CMRE).

Αν και η υλοποίηση διαγραμμάτων πάνω στις τιμές κλεισίματος των μετοχών είναι χρήσιμη είναι επίσης και προβληματική ειδικά στην περίπτωση που η σύγκριση που επιθυμούμε είναι ανάμεσα σε μετοχές με μεγάλη διαφορά στις τιμές. Για αυτό το λόγο η υλοποίηση πραγματοποιήθηκε πάνω στις αποδόσεις των αποθεμάτων και όχι πάνω στις απόλυτες τιμές τους. Έτσι αρχικά πραγματοποιήθηκε ο μετασχηματισμός των δεδομένων ως η απόδοση του αποθέματος από την αρχή της περιόδου ενδιαφέροντος ως:

$$return_{t,0} = \frac{price_t}{price_0}$$



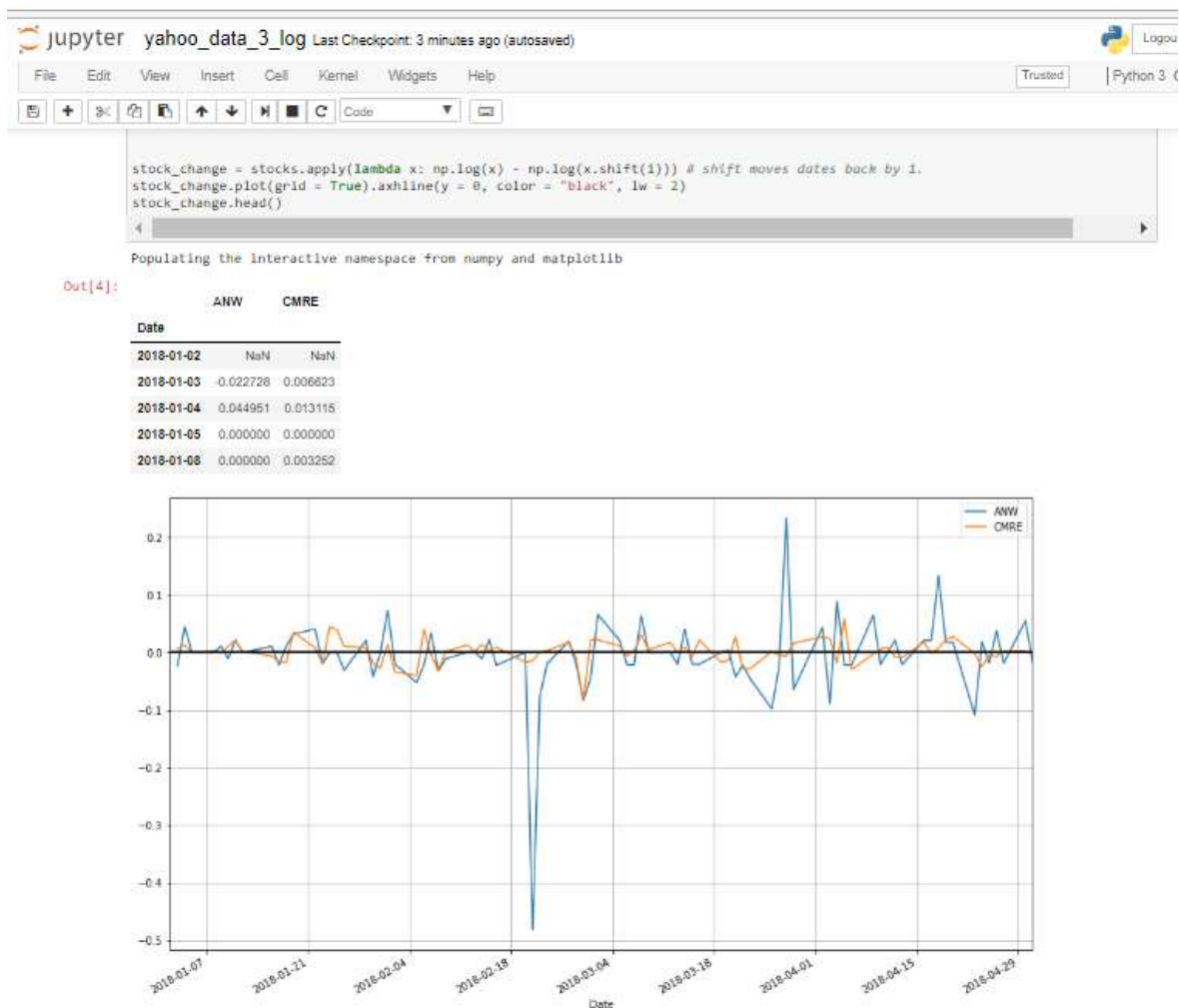
Εικόνα 5.4: Συγκριτικός πίνακας και γραμμικό διάγραμμα των δύο ναυτιλιακών

Στην έξοδο παρουσιάζεται ξεκάθαρα η μεγάλη πτώση της εταιρίας AEGEAN αλλά και η ταυτόχρονη αύξηση των τιμών των μετοχών της COSTAMARE κατά τους πρώτους πέντε μήνες του 2018.

Σε ένα δεύτερο τρόπο υλοποίησης των συγκριτικών διαγραμμάτων λήφθηκε υπόψη η λογαριθμική διαφορά των τιμών ανά ημέρα και όχι η απόδοση των αποθεμάτων. Έτσι η συνάρτηση υπολογισμού διαμορφώθηκε ως:

$$change_t = \ln(price_t) - \ln(price_{t-1})$$

Το πλεονέκτημα είναι ότι η λογαριθμική διαφορά πάνω στις τιμές μπορεί να ερμηνευτεί ως η ποσοστιαία μεταβολή σε ένα απόθεμα αλλά δεν εξαρτάται από τον παρονομαστή του κλάσματος. Η υλοποίηση απέδωσε τα αποτελέσματα της εικόνας 5.5.



Εικόνα 5.5: Διάγραμμα σύγκρισης λογαριθμικών διαφορών

5.2 Υλοποίηση και διαμόρφωση διαγραμμάτων των κινητών μέσων

Στην περίπτωση των χρηματοπιστωτικών αγορών τα διαγράμματα είναι ιδιαίτερα χρήσιμα εφόσον εκτός της οπτικοποίησης των αποτελεσμάτων δίνουν και την δυνατότητα τεχνικής ανάλυσης τους για εύρεση μοτίβων και προτύπων στην συμπεριφορά της αγοράς.

Μια από τις τεχνικές εύρεσης τάσης στα αποθέματα, όπως αναλύθηκε διεξοδικότερα στο κεφάλαιο 2, είναι αυτή των κινούμενων μέσων όρων. Υπενθυμίζεται ότι ένας ημερήσιος κινητός μέσος όρος q για μια σειρά δεδομένων x_t και ένα χρονικό σημείο t ορίζεται ως ο μέσος όρος των προηγούμενων q ημερών. Έτσι αν ο όρος MA_t^q δηλώνει μια διαδικασία κινητού μέσου τότε:

$$MA_t^q = \frac{1}{q} \sum_{i=0}^{q-1} x_{t-i}$$

Η χρήση τους γίνεται επωφελής για το λόγο ότι οι κινητοί μέσοι όροι εξομαλύνουν μια χρονοσειρά σειρά, εξαλείφοντας το “θόρυβο” και βοηθούν στον εντοπισμό των τάσεων. Όσο μεγαλύτερη είναι η τιμή του q (το διάστημα των ημερών δηλαδή που λαμβάνεται υπόψη στον υπολογισμό – *αργός μέσος*), τόσο λιγότερο ευαίσθητη είναι η μεταβολή του μέσου όρου στις βραχυπρόθεσμες διακυμάνσεις της σειράς x_t . Εδώ πρέπει να σημειωθεί ότι, οι γρήγορα μεταβαλλόμενοι μέσοι όροι έχουν μικρότερο q και παρακολουθούν στενότερα το απόθεμα, ενώ οι αργοί κινητοί μέσοι όροι έχουν μεγαλύτερο q , με αποτέλεσμα να ανταποκρίνονται λιγότερο στις διακυμάνσεις του αποθέματος και να είναι πιο σταθεροί.

Η βιβλιοθήκη `pandas` παρέχει τα εργαλεία (την συνάρτηση `mean()`) για τον εύκολο υπολογισμό των κινητών μέσων όρων. Στην μελέτη περίπτωσης που αναπτύσσεται δημιουργήθηκε ένας κινητός μέσος όρος 20 ημερών (ένα μήνα) για τα δεδομένα της AEGEAN. Ο σχεδιασμός του κινητού μέσου όρου γίνεται παράλληλα με το σχεδιασμό των προσαρμοσμένων τιμών του αποθέματος.

Η υλοποίηση πραγματοποιείται με την βοήθεια του κώδικα:

```
AEGEAN['20d'] = np.round(AEGEAN['Close'].rolling(window = 20, center =
False).mean(), 2)
pandas_candlestick_ohlc(AEGEAN.loc['2018-01-04':'2018-04-29',:],
otherseries = '20d')
```



Εικόνα 5.6: Διάγραμμα κινητού μέσου ανά μήνα του αποθέματος της AEGEAN για το διάστημα Ιανουάριος-Απρίλιος 2018

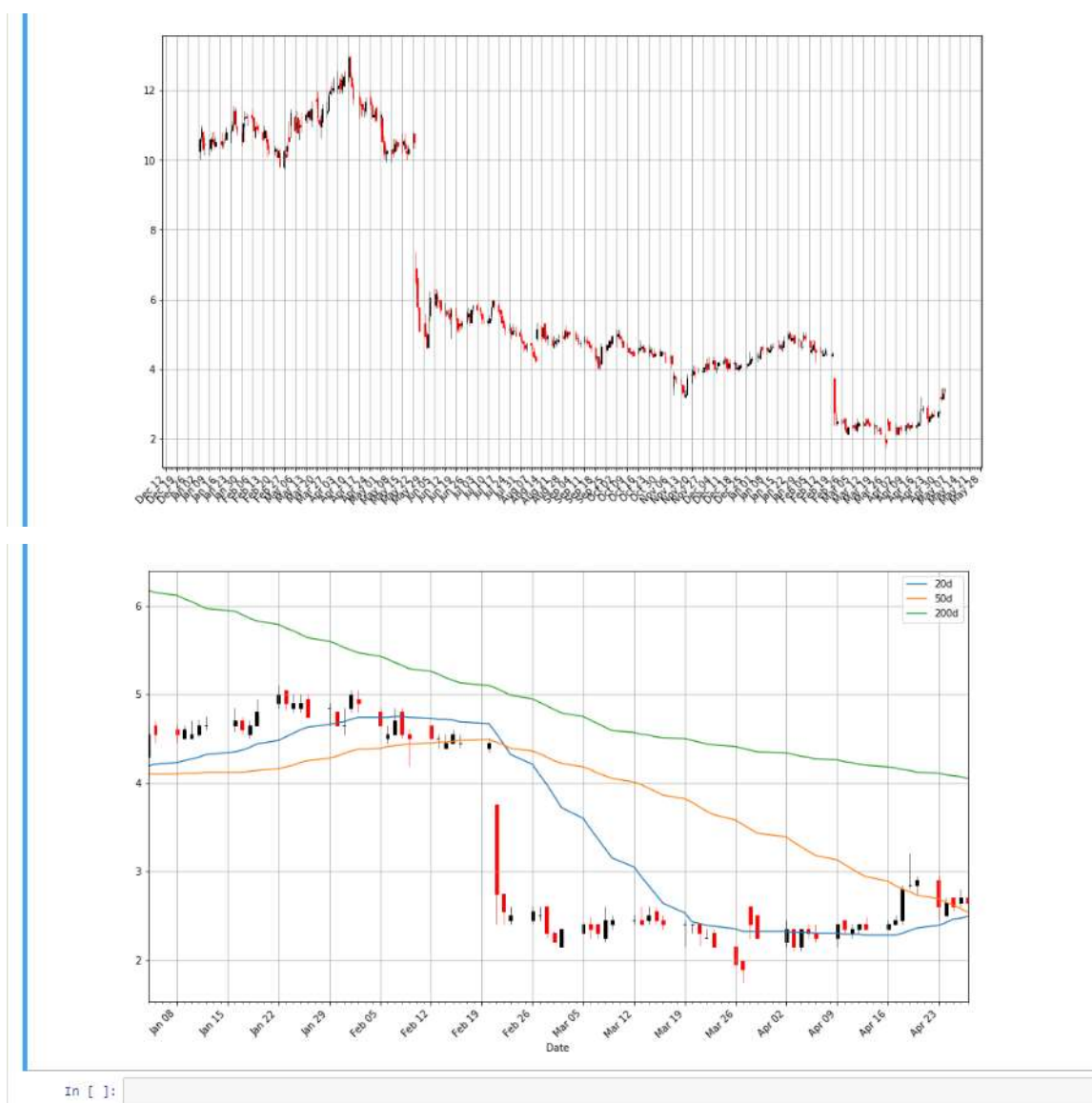
Από την εικόνα 5.6 γίνεται αντιληπτό ότι ο σχεδιασμός του κινητού μέσου ξεκινά με καθυστέρηση (μετά τις 29 Ιανουαρίου) εφόσον για τον υπολογισμό του κινητού μέσου είναι απαραίτητα παρελθοντικά δεδομένα πάνω στις προηγούμενες 20 ημέρες.

Για το λόγο αυτό επεκτείνονται τόσο τα ιστορικά στοιχεία (από τον Ιανουάριο του 2017) όσο ο υπολογισμός του κινητού μέσου στις 50 (2 μήνες) και στις 200 ημέρες (6 μήνες). Για την υλοποίηση ο προηγούμενος κώδικας μετατράπηκε ως:

```
AEGEAN['20d'] = np.round(AEGEAN['Close'].rolling(window = 20, center = False).mean(), 2)
AEGEAN['50d'] = np.round(AEGEAN['Close'].rolling(window = 50, center = False).mean(), 2)
AEGEAN['200d'] = np.round(AEGEAN['Close'].rolling(window = 200, center = False).mean(), 2)
pandas_candlestick_ohlc(AEGEAN.loc['2018-01-04':'2018-04-29',:],
otherseries = ['20d', '50d', '200d'])
```

Τα αποτελέσματα παρουσιάζονται στην εικόνα 5.7 που ακολουθεί.





Εικόνα 5.7: Διάγραμμα κινητού μέσου ανά μήνα, δίμηνο και εξάμηνο του αποθέματος της AEGEAN για το διάστημα Ιανουάριος-Απρίλιος 2018

5.3 Υλοποίηση χαρτοφυλακίου

Για την διαμόρφωση και την υλοποίηση χαρτοφυλακίου χρήστη πρέπει πρώτα να υλοποιηθούν οι στρατηγική συναλλαγών που θα ακολουθηθεί.

Μια συναλλαγή αποθεμάτων ονομάζεται *θέση αγοράς* (long position) στην περίπτωση που η συναλλαγή θα περατωθεί σε κάποια μελλοντική στιγμή αλλά υπό προϋποθέσεις και είναι

μία τεχνική που χρησιμοποιείται από τους επενδυτές με σκοπό να κερδίσουν από την πτώση των τιμών περιουσιακών στοιχείων.

Στις στρατηγικές *θέσης πώλησης* (short position) οι επενδυτές δανείζονται περιουσιακά στοιχεία από άλλους επενδυτές, τα διαπραγματεύονται και μετά από ένα προσυμφωνηθέν χρονικό διάστημα τα επιστρέφουν στο δανειστή με μια μικρή απόδοση.

Από την άλλη μια στρατηγική μπορεί να πραγματοποιηθεί ως εξής: η απόφαση αγοράς ενός αποθέματος (με την προσδοκία ότι η αξία του θα αυξηθεί), πραγματοποιείται ταυτόχρονα με ένα σχέδιο πώλησης του αποθέματος σε υψηλότερη τιμή. Σε αυτή την περίπτωση η στρατηγική καλείται *μακράς θέσης* και ο επενδυτής αποκομίζει κέρδος στην περίπτωση που το χρηματοπιστωτικό μέσο παρουσιάζει αυξήσεις αξίας, επωφελούμενος την διαφορά στις τιμές μεταξύ ημερομηνίας αγοράς-πώλησης. Στην στρατηγική μακράς θέσης το δυνητικό κέρδος είναι απεριόριστο ενώ οι πιθανές απώλειες περιορίζονται στην τιμή αγοράς του αποθέματος.

Από την άλλη πλευρά, εάν ο επενδυτής θεωρήσει ότι η αξία του αποθέματος πρόκειται να μειωθεί, μπορεί να πουλήσει το απόθεμα σε μια μεσιτική εταιρεία, με την προσδοκία να το αγοράσει αργότερα σε χαμηλότερη τιμή, αποκτώντας έτσι ένα κέρδος. Σε αυτή την περίπτωση μιλάμε για στρατηγική *θέσης πώλησης*. Το πιθανό κέρδος εξαρτάται από την τιμή του μετοχικού κεφαλαίου, ενώ οι απώλειες είναι απεριόριστες.

Κάθε επενδυτής ακολουθεί ένα σύνολο προσωπικών κανόνων που καθορίζουν πόσα από τα χρήματά του είναι πρόθυμος να επενδύσει, πως τι ρίσκο επιθυμεί να αναλάβει και ποιες είναι οι προϋποθέσεις εξόδου του από την θέση. Οι τελευταίες καθορίζουν πότε θα βγει από τη θέση, είτε λόγω κέρδους είτε λόγω απωλειών προκαθορίζοντας το ελάχιστο κέρδος ή την μέγιστη απώλεια (εντολής stop-loss, μια εντολή που ενεργοποιείται για να αποφευχθούν περαιτέρω απώλειες).

Στην συγκεκριμένη μελέτη περίπτωσης υλοποιείται μια στρατηγική που περιλαμβάνει σήματα για την εκκίνηση των συναλλαγών, έναν κανόνα για το πότε ένα χαρτοφυλάκιο βρίσκεται σε κίνδυνο (για οποιαδήποτε στρατηγική και να ακολουθείται), καθώς και μια ολοκληρωμένη στρατηγική εξόδου από την θέση.

Στην στρατηγική που θα ακολουθηθεί το χρηματικό ποσό του χαρτοφυλακίου που εμπλέκεται σε κάθε διαπραγμάτευση θα είναι ένα σταθερό ποσοστό (το 10% του

χαρτοφυλακίου). Από την άλλη, ως θέση εξόδου θα υιοθετηθεί η υπόθεση απώλειας του 20% της αξίας του αποθέματος.

Όσον αφορά την στρατηγική εισαγωγής και εξόδου από μια θέση με κέρδος θα χρησιμοποιηθούν δύο κινητοί μέσοι όροι, ένας γρήγορος και ένας αργός ως εξής:

- Πραγματοποιείται συναλλαγή του περιουσιακού στοιχείου όταν ο ταχέως μεταβαλλόμενος μέσος όρος διασχίζει τον αργό κινητό μέσο.
- Πραγματοποιείται έξοδος από την συναλλαγή όταν ο ταχέως μεταβαλλόμενος μέσος όρος περνά ξανά από τον αργό κινητό μέσο όρο.
- Μια συναλλαγή μακράς θέσης (long position) θα πραγματοποιηθεί όταν ο ταχέως μεταβαλλόμενος μέσος όρος θα διασχίσει από κάτω τον αργά κινητό μέσο όρο και η συναλλαγή θα διακόπτεται όταν ο ταχέως μεταβαλλόμενος μέσος όρος θα περάσει, σε μεταγενέστερο χρόνο, κάτω από τον αργό κινούμενο μέσο όρο.
- Μια θέση πώλησης (short position) θα πραγματοποιηθεί όταν ο ταχέως μεταβαλλόμενος μέσος όρος θα περάσει κάτω από τον αργό κινούμενο μέσο όρο και η έξοδος από την θέση θα πραγματοποιηθεί όταν ο ταχέως μεταβαλλόμενος μέσος όρος θα περάσει, σε μεταγενέστερο χρόνο, πάνω από τον αργό κινητό μέσο όρο.

Πριν την υλοποίηση και χρήση της στρατηγικής πρέπει να αξιολογηθεί η ποιότητα της. Η αξιολόγηση θα πραγματοποιηθεί μέσω ενός backtesting, το οποίο εξετάζει πόσο κερδοφόρα είναι η στρατηγική πάνω σε ιστορικά δεδομένα. Έτσι θα εξεταστεί η απόδοση του χαρτοφυλακίου για μετοχές της AEGEAN, λαμβάνοντας ως ταχύ κινητό μέσο όρο αυτό των 20 ημερών και ως αργό των 50 ημερών.

Αρχικά προσδιορίζουμε πότε ο κινητός μέσος όρος των 20 ημερών είναι κάτω από το μέσο όρο των 50 ημερών και αντίστροφα. Αυτός ο έλεγχος υλοποιείται μέσω του κώδικα:

```
AEGEAN['20d-50d'] = AEGEAN['20d'] - AEGEAN['50d']  
AEGEAN.tail() #ΠΙΝΑΚΑΣ ΤΩΝ ΤΕΛΕΥΤΑΙΩΝ ΤΙΜΩΝ
```

Το αποτέλεσμα παρουσιάζεται στην εικόνα 5.8:

```

In [2]: import pandas as pd
        from pandas_datareader import data as web
        import matplotlib.pyplot as plt # Import matplotlib

        import datetime
        import numpy as np
        # We will look at stock prices over last months, starting at January 1, 2018
        start = datetime.datetime(2018,1,1)
        end = datetime.date.today()

        # Let's get AEGEAN stock data; AEGEAN's ticker symbol is ANW
        # First argument is the series we want, second is the source ("yahoo" for Yahoo! Finance), third is the start date, fourth is the
        AEGEAN = web.DataReader("ANW", "yahoo", start, end)

        type(AEGEAN)

        # value matrix
        AEGEAN.head()

        AEGEAN["20d"] = np.round(AEGEAN["Close"].rolling(window = 20, center = False).mean(), 2)
        AEGEAN["50d"] = np.round(AEGEAN["Close"].rolling(window = 50, center = False).mean(), 2)
        AEGEAN["20d-50d"] = AEGEAN["20d"] - AEGEAN["50d"]
        AEGEAN.tail()
    
```

Date	Open	High	Low	Close	Adj Close	Volume	20d	50d	20d-50d
2018-04-30	2.05	2.20	2.03	2.00	2.00	253500	2.52	2.51	0.01
2018-05-01	2.75	2.85	2.75	2.75	2.75	108500	2.55	2.47	0.08
2018-05-02	3.20	3.45	3.10	3.15	3.15	3188400	2.60	2.46	0.14
2018-05-03	3.15	3.45	3.15	3.30	3.30	1239500	2.64	2.48	0.18
2018-05-04	3.35	3.45	3.30	3.45	3.45	1521100	2.70	2.48	0.22

Εικόνα 5.8: Πίνακας διαφορών κινητών μέσων

Μέσο αυτής της διαφοράς θα πραγματοποιηθεί ο έλεγχος και επομένως και επιλογή των βημάτων που θα ακολουθηθούν. Ο έλεγχος παρουσιάζεται στην συνέχεια:

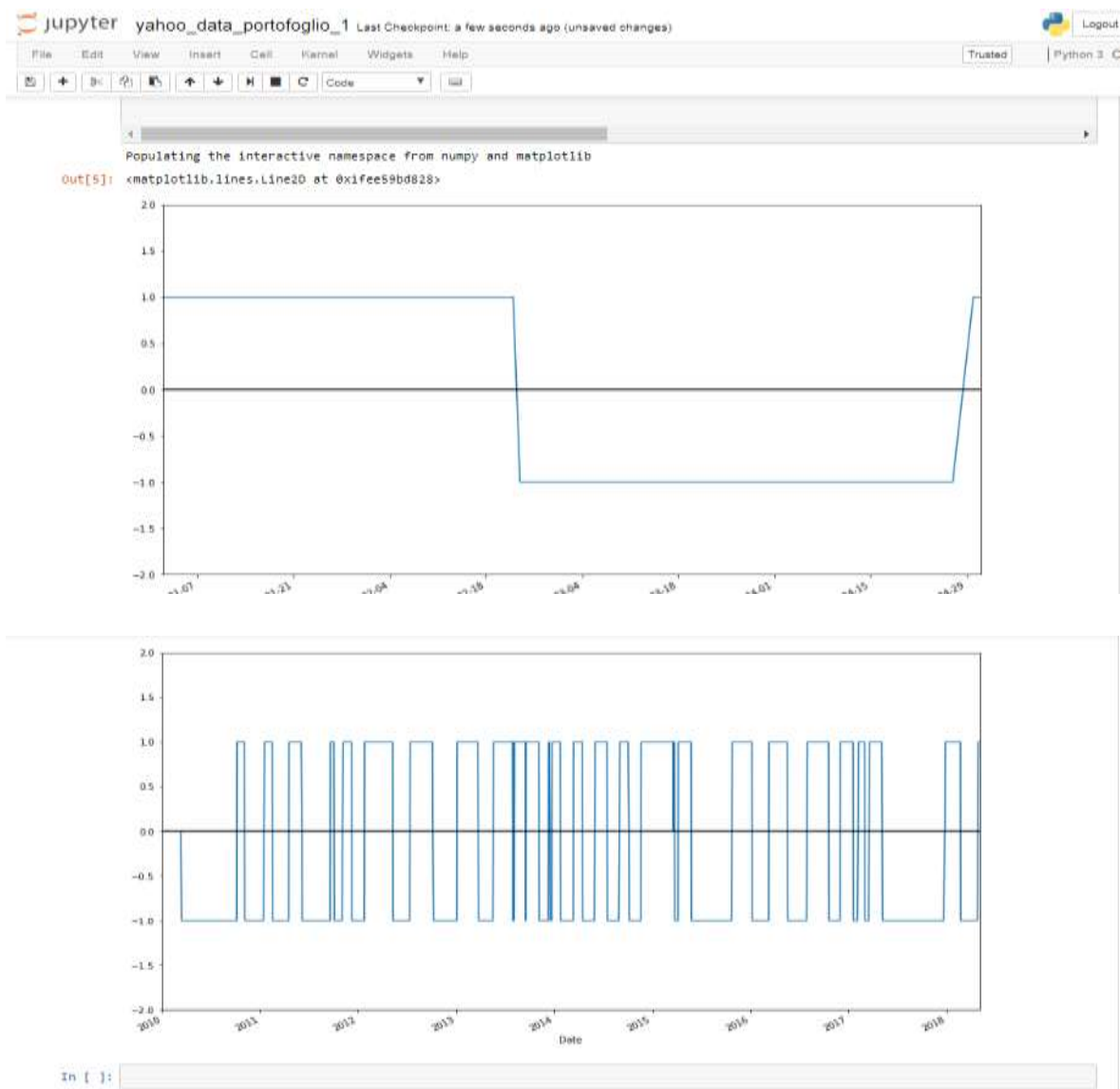
np.where() είναι μια διανυσματική συνάρτηση if-else, όπου μια συνθήκη ελέγχεται για κάθε στοιχείο ενός διανύσματος (το πρώτο όρισμα της συνάρτησης). Όταν η συνθήκη είναι αληθής δίνεται σαν τιμή επιστροφής το δεύτερο όρισμα, διαφορετικά το τρίτο

```

AEGEAN['Regime'] = np.where(AEGEAN['20d-50d'] > 0, 1, 0)
# η τιμή 1 ορίζει ότι ο ταχύς μέσος είναι πάνω από τον βραδύ και το 0 το αντίθετο
AEGEAN['Regime'] = np.where(AEGEAN['20d-50d'] < 0, -1, AEGEAN['Regime'])
# η τιμή -1 ορίζει ότι ο ταχύς μέσος είναι κάτω από τον βραδύ και το 0 το αντίθετο
    
```

```

AEGEAN.loc['2016-01-01':'2016-08-07', 'Regime'].plot(ylim = (-2,2)).axhline(y = 0, color = 'black', lw = 2)
    
```



Εικόνα 5.9: Εναλλαγές σε αργό και ταχύ κινητό μέσο σε πρόσφατα και σε ιστορικά δεδομένα

Για τον υπολογισμό του πλήθους των εναλλαγών για το χρονικό διάστημα από τον Ιανουάριο του 2018 ως και σήμερα χρησιμοποιήθηκε ο κώδικας:

```
AEGEAN['Regime'].value_counts()
```

και έδωσε σαν αποτέλεσμα την επόμενη εικόνα

```

AEGEAN["20d"] = np.round(AEGEAN["Close"].rolling(window = 20, center = False).mean(), 2)
AEGEAN["50d"] = np.round(AEGEAN["Close"].rolling(window = 50, center = False).mean(), 2)
AEGEAN["200d"] = np.round(AEGEAN["Close"].rolling(window = 200, center = False).mean(), 2)

AEGEAN['20d-50d'] = AEGEAN['20d'] - AEGEAN['50d']
AEGEAN.tail()

# np.where() is a vectorized if-else function, where a condition is checked for each component of a vector, and the first argument
AEGEAN["Regime"] = np.where(AEGEAN["20d-50d"] > 0, 1, 0)
# We have 1's for bullish regimes and 0's for everything else. Below I replace bearish regimes's values with -1, and to maintain

AEGEAN["Regime"] = np.where(AEGEAN["20d-50d"] < 0, -1, AEGEAN["Regime"])

#PLOT FOR A TIME INTERVAL
#AEGEAN.Loc['2018-01-01': '2018-05-01', "Regime"].plot(yLim = (-2,2)).axhline(y = 0, color = "black", Lw = 2)

#PLOT FOR HISTORICAL DATA
#AEGEAN["Regime"].plot(yLim = (-2,2)).axhline(y = 0, color = "black", Lw = 2)

AEGEAN["Regime"].value_counts()
    
```

Populating the interactive namespace from numpy and matplotlib

```

Out[10]: 0    49
        -1   32
         1    5
        Name: Regime, dtype: int64
    
```

In []:

Εικόνα 5.10: Πλήθος εναλλαγών αργού και ταχύ κινητού μέσου

Από την εικόνα 5.10 συμπεραίνεται ότι ο αργός μέσος πέρασε πάνω από τον ταχύ 32 φορές ενώ 5 φορές συνέβη το αντίθετο και ο ταχύς πέρασε τον αργό. Επίσης 49 φορές η κατάσταση παρέμεινε αμετάβλητη.

Τα έναυσμα (το σινιάλο) για μια συναλλαγή εμφανίζεται σε περιπτώσεις αλλαγής καθεστώτος. Όταν ξεκινά ένα διογκωτικό καθεστώς, ενεργοποιείται ένα σήμα αγοράς και όταν τελειώσει, ενεργοποιείται ένα σήμα πώλησης. Ομοίως, όταν ξεκινά ένα καθεστώς συρρίκνωσης, ένα σήμα πώλησης ενεργοποιείται και όταν λήξει το καθεστώς, ενεργοποιείται ένα σήμα αγορών.

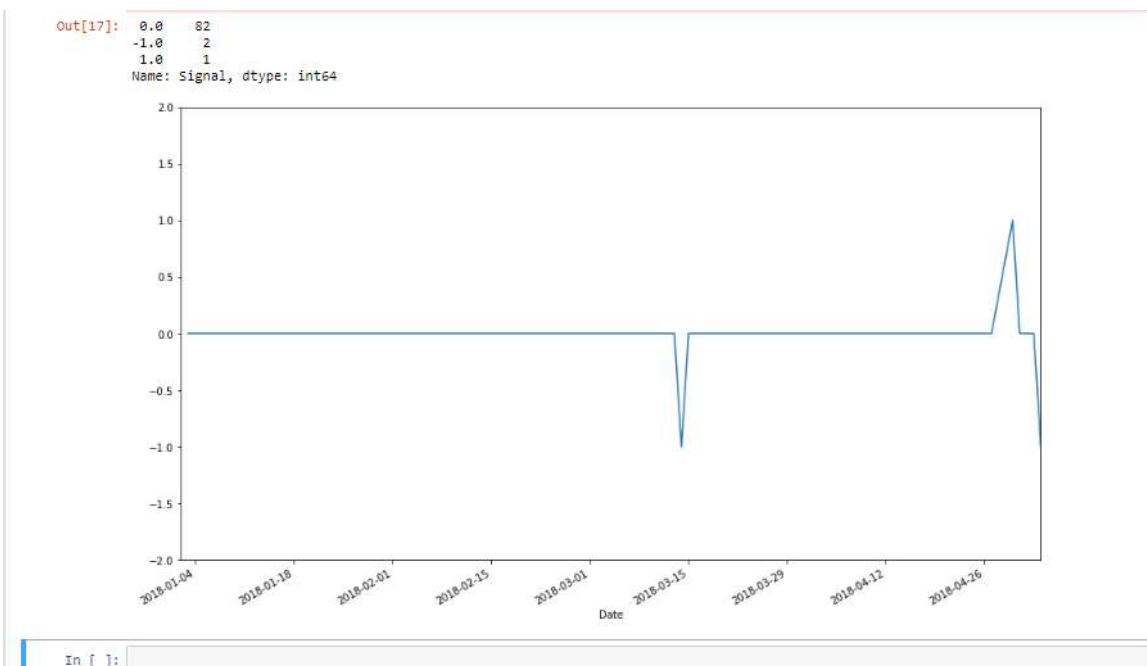
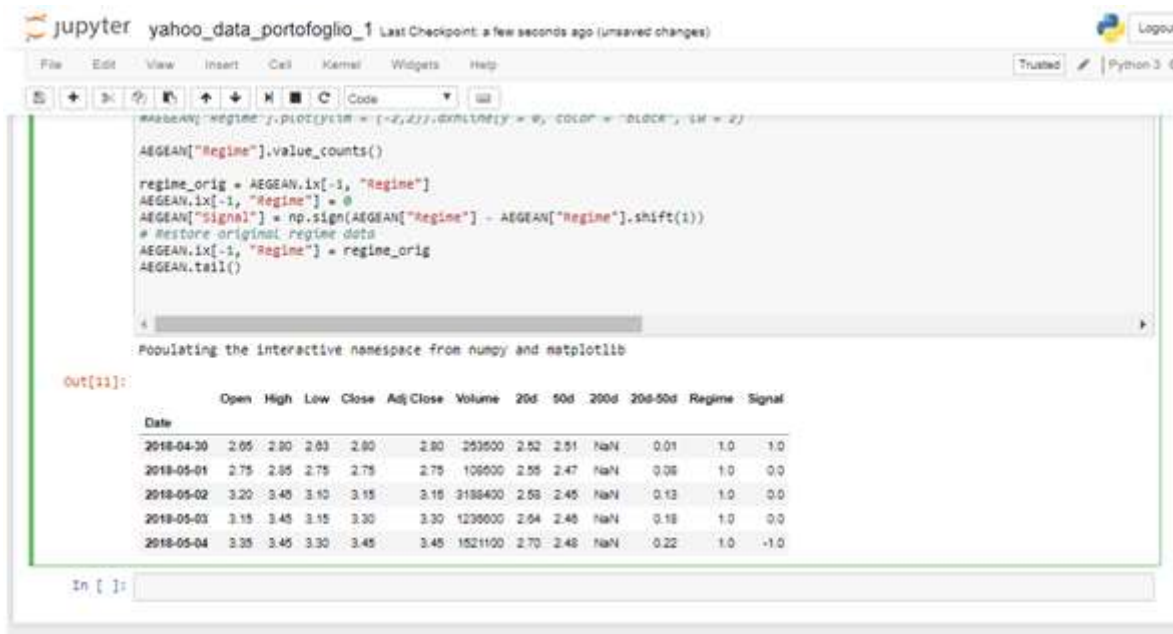
Έστω r_t το καθεστώς στο χρόνο t , και s_t το σήμα στο χρόνο t , τότε:

$s_t = \text{sign}(r_t - r_{t-1})$ όπου $s_t \in \{-1, 0, 1\}$ και -1 σημαίνει πώληση, 1 αγορά και 0 να μην πραγματοποιηθεί καμία κίνηση. Για την πραγματοποίηση του σήματος χρησιμοποιήθηκε ο κώδικας:

```

regime_orig = AEGEAN.ix[-1, 'Regime']
AEGEAN.ix[-1, 'Regime'] = 0
AEGEAN['Signal'] = np.sign(AEGEAN['Regime'] - AEGEAN['Regime'].shift(1))
# Restore original regime data
    
```

```
AEGEAN.ix[-1, 'Regime'] = regime_orig
AEGEAN.tail()
```



Εικόνα 5.11: Έξοδος σημάτων πώλησης/αγοράς σύμφωνα με τις εναλλαγές

Έτσι στο χρονικό διάστημα της μελέτης δίνεται σήμα για 2 πωλήσεις, 1 αγορά και 82 σήματα μη λήψης καμίας ενέργειας.

Στην συνέχεια γίνεται προσδιορισμός των τιμών σε κάθε αγοροπωλησία έτσι ώστε να υπολογιστεί το κέρδος ή η ζημία. Οι τιμές αγοροπωλησίας λαμβάνονται από τον:

```
AEGEAN.loc[AEGEAN['Signal'] == 1, 'Close']

AEGEAN.loc[AEGEAN['Signal'] == -1, 'Close']

# Δημιουργία ενός DataFrame των συναλλαγές, συμπεριλαμβανομένης της τιμής
# συναλλαγής και του καθεστώτος στο οποίο πραγματοποιείται.
AEGEAN_signals = pd.concat([

pd.DataFrame({'Price': AEGEAN.loc[AEGEAN['Signal'] == 1, 'Close'],
              'Regime': AEGEAN.loc[AEGEAN['Signal'] == 1, 'Regime'],
              'Signal': 'Buy'}),

pd.DataFrame({'Price': AEGEAN.loc[AEGEAN['Signal'] == -1, 'Close'],
              'Regime': AEGEAN.loc[AEGEAN['Signal'] == -1, 'Regime'],
              'Signal': 'Sell'})])

AEGEAN_signals.sort_index(inplace = True)

AEGEAN_signals
```

Παράγεται η ακόλουθη έξοδος:

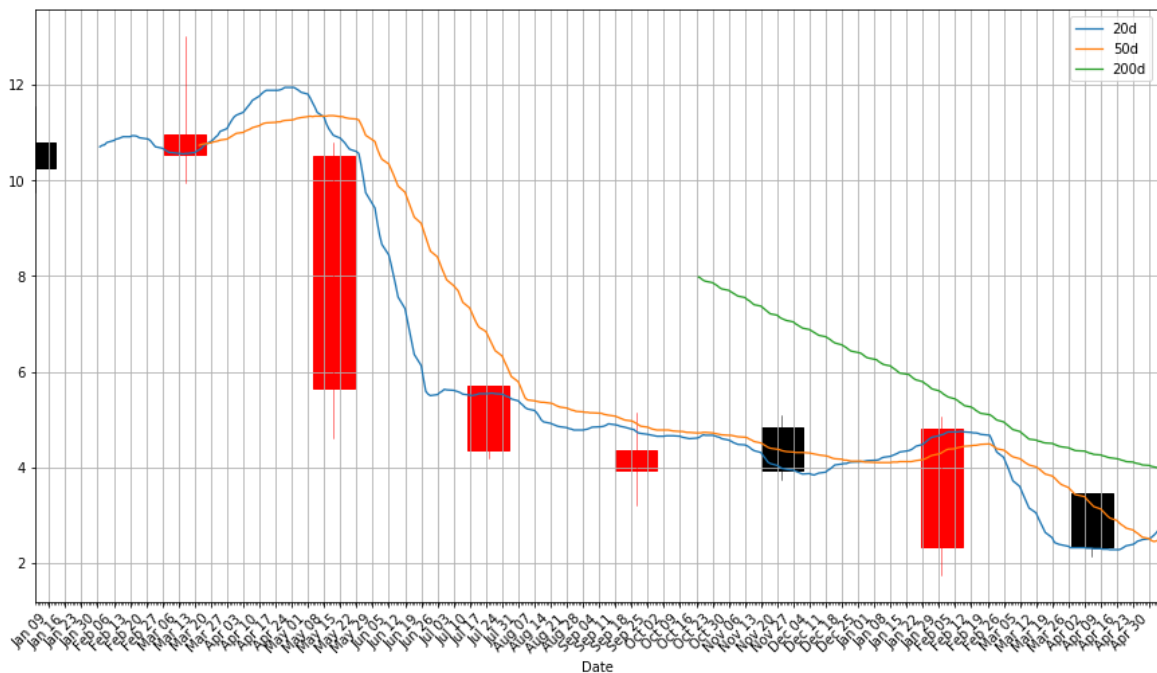
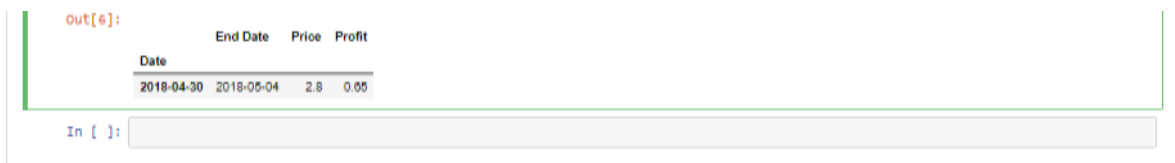


Εικόνα 5.12: Ετικέτες αγοράς και πώλησης μετοχών

Ο υπολογισμός του μακροχρόνιου κέρδους ή ζημιάς γίνεται μέσω του της δημιουργίας ενός νέου DataFrame πάνω στο προηγούμενο:

```
AEGEAN_long_profits =  
pd.DataFrame({'Price':AEGEAN_signals.loc[(AEGEAN_signals['Signal'] ==  
'Buy') &AEGEAN_signals['Regime'] == 1, 'Price'],  
  
'Profit':pd.Series(AEGEAN_signals['Price']AEGEAN_signals['Price'].shift(1  
(AEGEAN_signals['Regime'].shift(1) == 1)].index].tolist(),  
  
'End Date':  
AEGEAN_signals['Price'].loc[AEGEAN_signals.loc[(AEGEAN_signals['Signal'].  
shift(1) == 'Buy') & (AEGEAN_signals['Regime'].shift(1) ==  
1)].index].index  
  
)})  
  
AEGEAN_long_profits
```

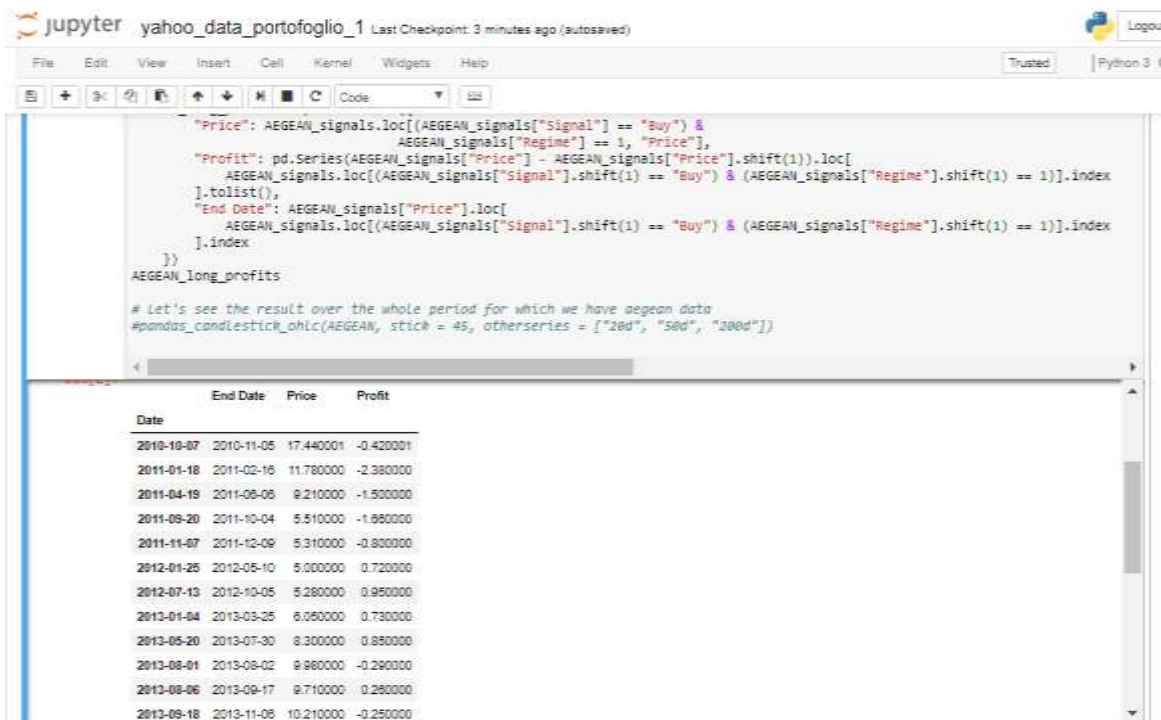
Και παράγει σαν έξοδο:



Εικόνα 5.13: Υπολογισμός κέρδους/ζημιάς του χαρτοφυλακίου για τέσσερις μήνες

Τα αποτελέσματα δεν είναι χαρακτηριστικά εφόσον κατά το τελευταίο πεντάμηνο η μετοχή επέφερε ένα ελάχιστο κέρδος μια και μόνο φορά.

Επιλέγοντας τα ιστορικά στοιχεία από την άλλη παράγονται περισσότερα αποτελέσματα και ένα πιο χαρακτηριστικό διάγραμμα:



Εικόνα 5.14: Υπολογισμός κέρδους/ζημίας σε ιστορικά δεδομένα

Η μέχρι τώρα υλοποίηση δεν καλύπτει όλες τις περιπτώσεις και δεν έχει τα αναμενόμενα αποτελέσματα σε περιπτώσεις διαίρεσης μετοχών ή πληρωμών μερισμάτων.

Μια λύση σε αυτή την προβληματική είναι η προσαρμογή των τιμών ώστε να λαμβάνονται υπόψη οι διαφορές αποθεμάτων και οι πληρωμές μερισμάτων. Η αποθήκη οικονομικών δεδομένων της Yahoo δεν παρέχει αυτές τις προσαρμοσμένες τιμές οι οποίες όμως μπορούν να υπολογιστούν μέσω της σχέσης:

$$price_t^{adj} = price_t * m_t$$

όπου m_t είναι ο συντελεστής που χρησιμοποιείται για την προσαρμογή.

Η έξοδος που παράγεται πάνω στα προσαρμοσμένα ιστορικά στοιχεία είναι η παρακάτω:



Εικόνα 5.15: Υπολογισμός κέρδους/ζητίας πάνω σε προσαρμοσμένα ιστορικά δεδομένα

5.4 Προσομοίωση χαρτοφυλακίου

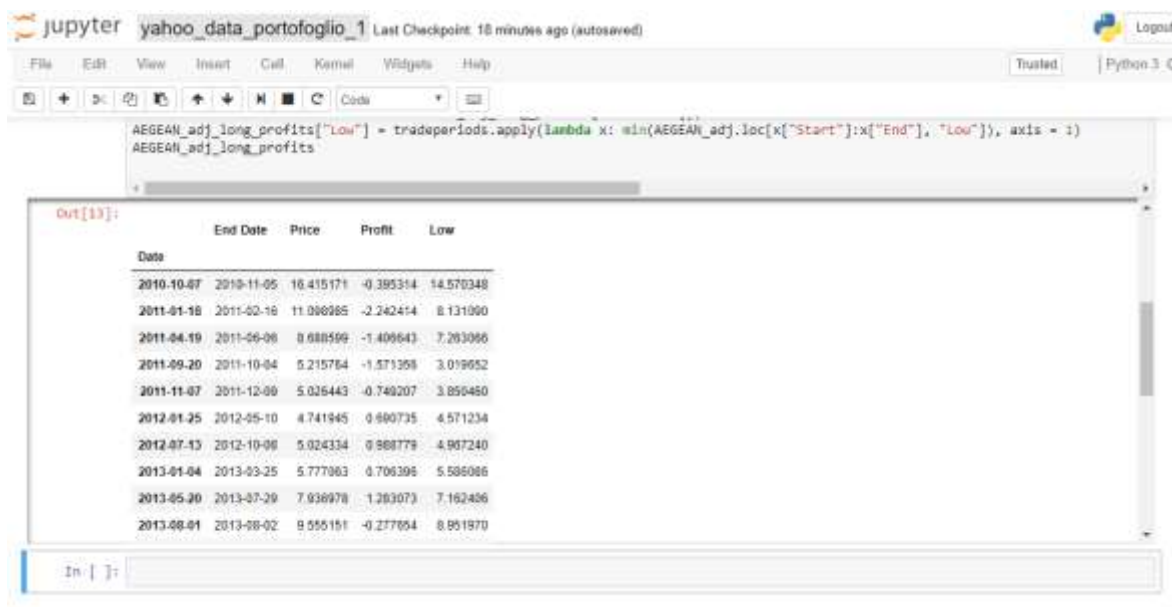
Έχοντας πλέον θεσπίσει τους κανόνες συναλλαγής δημιουργήθηκε η προσομοίωση ενός χαρτοφυλακίου \$1.000.000. Υπενθυμίζεται ότι οι κανόνες ορίζουν συναλλαγή του 10% του ποσού του χαρτοφυλακίου και έξοδο από την θέση αν οι απώλειες υπερβαίνουν το 20% της αξίας της συναλλαγής. Όσο αφορά τον κανόνα εξόδου πρέπει να πραγματοποιηθεί ένας έλεγχος για το πότε οι χαμηλά τιμές φθάνουν αρκετά χαμηλά για να προκαλέσουν τη διακοπή της απώλειας. Επιπλέον, κατά την προσομοίωση οι συναλλαγές πραγματοποιούνται σε πακέτα των 100 μετοχών.

Αρχικά πραγματοποιείται η λήψη των χαμηλών τιμών κάθε συναλλαγής και υλοποιείται μέσω του:

```
tradeperiods = pd.DataFrame({'Start': AEGEAN_adj_long_profits.index,
                             'End': AEGEAN_adj_long_profits['End Date']})
AEGEAN_adj_long_profits['Low'] = tradeperiods.apply(lambda x:
min(AEGEAN_adj.loc[x['Start']:x['End'], 'Low']), axis = 1)
```

AEGEAN_adj_long_profits

και παράγει σαν αποτέλεσμα:



Εικόνα 5.16: Λήψη χαμηλών τιμών για τις ημέρες συναλλαγών πάνω σε ιστορικά δεδομένα

Η προσομοίωση πρέπει να παράγει κατ' αρχάς έναν πίνακα δεδομένων μεταξύ των οποίων θα περιλαμβάνει στοιχεία για την τιμή μετοχής, την τιμή εισόδου και εξόδου από την θέση, την απολαβή (κέρδος ή ζημιά) και την αξία συναλλαγής.

```
cash = 1000000
AEGEAN_backtest = pd.DataFrame({'Start Port. Value': [],
                                'End Port. Value': [],
                                'End Date': [],
                                'Shares': [],
                                'Share Price': [],
                                'Trade Value': [],
                                'Profit per Share': [],
                                'Total Profit': [],
                                'Stop-Loss Triggered': []})

port_value = .1 # Ποσοστό του χαρτοφυλακίου προς επένδυση 10%
batch = 100     # Πακέτια μετοχών των 100
stoploss = .2   # % της απώλειας για έξοδο από την θέση

for index, row in AEGEAN_adj_long_profits.iterrows():
    batches = np.floor(cash * port_value) // np.ceil(batch *
    row['Price']) # Μέγιστος αριθμός πακέτων προς συναλλαγή
    trade_val = batches * batch * row['Price'] # Ποσό που επενδύεται σε
    κάθε συναλλαγή
    if row['Low'] < (1 - stoploss) * row['Price']: #μέτρηση της
    στρατηγικής διακοπής απώλειας
        share_profit = np.round((1 - stoploss) * row['Price'], 2)
        stop_trig = True
    else:
        share_profit = row['Profit']
        stop_trig = False
    profit = share_profit * batches * batch # Υπολογισμός απολαβών
    # Υλοποίηση προσομοίωσης
    AEGEAN_backtest = AEGEAN_backtest.append(pd.DataFrame({'
        'Start Port. Value': cash,
        'End Port. Value': cash + profit,
        'End Date': row['End Date'],
        'Shares': batch * batches,
        'Share Price': row['Price'],
        'Trade Value': trade_val,
        'Profit per Share': share_profit,
        'Total Profit': profit,
        'Stop-Loss Triggered': stop_trig
    }, index = [index]))
    cash = max(0, cash + profit)

AEGEAN_backtest
AEGEAN_backtest['End Port. Value'].plot()
```

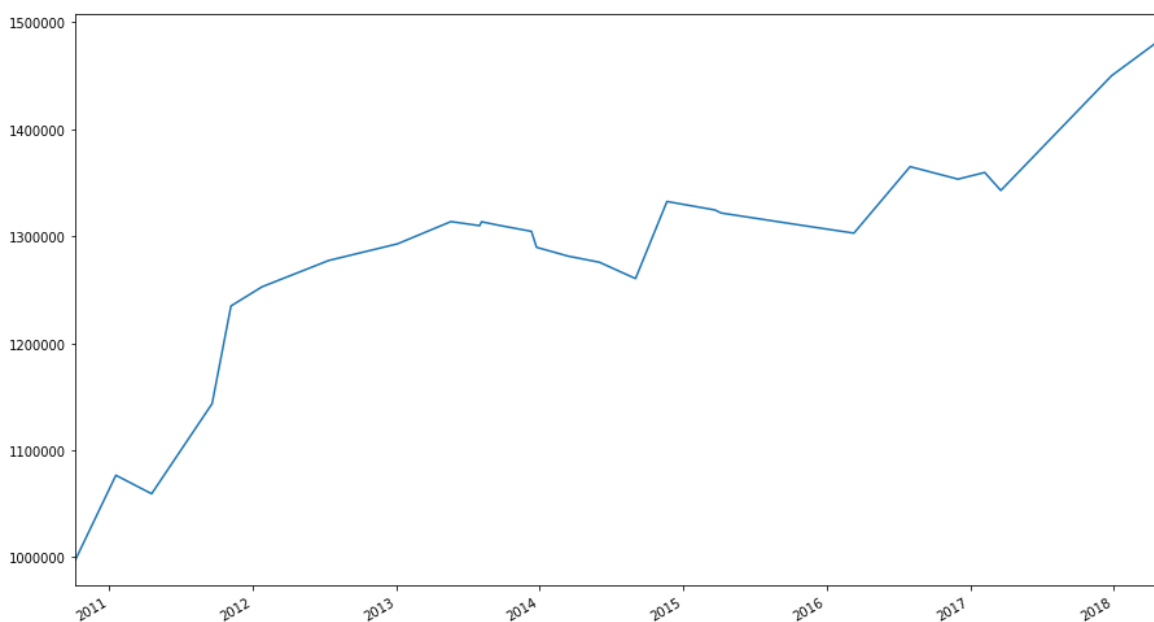
Η έξοδος παράχθηκε πάνω σε ιστορικά δεδομένα και έχει την μορφή της εικόνας 5.17:

jupyter yahoo_data_portofoglio_1 Last Checkpoint: 18 minutes ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

	End Date	End Port. Value	Profit per Share	Share Price	Shares	Start Port. Value	Stop-Loss Triggered	Total Profit	Trade Value
2010-10-07	2010-11-05	9.976281e+05	-0.395314	16.415171	6000.0	1.000000e+06	0.0	-2371.8840	98491.0280
2011-01-18	2011-02-16	1.076660e+06	8.880000	11.098085	8800.0	9.976281e+05	1.0	79032.0000	98780.9665
2011-04-19	2011-06-06	1.059358e+06	-1.405643	8.688599	12300.0	1.076660e+06	0.0	-17301.7089	106869.7677
2011-09-20	2011-10-04	1.143992e+06	4.170000	5.215764	20200.0	1.059358e+06	1.0	84234.0000	106358.4328
2011-11-07	2011-12-09	1.234846e+06	4.030000	5.025443	22700.0	1.143992e+06	1.0	91254.0000	114100.2561
2012-01-25	2012-05-10	1.252736e+06	0.690735	4.741945	25900.0	1.234846e+06	0.0	17699.0365	122816.3755
2012-07-13	2012-10-08	1.277357e+06	0.986779	5.024334	24800.0	1.252736e+06	0.0	24620.5971	125105.5180
2013-01-04	2013-03-25	1.292899e+06	0.706396	5.777083	22000.0	1.277357e+06	0.0	15546.7120	127095.3880
2013-05-20	2013-07-29	1.313684e+06	1.261073	7.839978	16200.0	1.292899e+06	0.0	20795.7828	128579.0436
2013-08-01	2013-08-02	1.309880e+06	-0.277054	9.555151	13700.0	1.313684e+06	0.0	-3803.8588	130905.5687
2013-08-08	2013-09-17	1.313514e+06	0.259823	9.296643	14000.0	1.309880e+06	0.0	3634.7220	130153.0020
2013-09-18	2013-11-06	1.310103e+06	-0.736828	9.708208	13400.0	1.313514e+06	0.0	-3211.0167	131138.5899

In []:



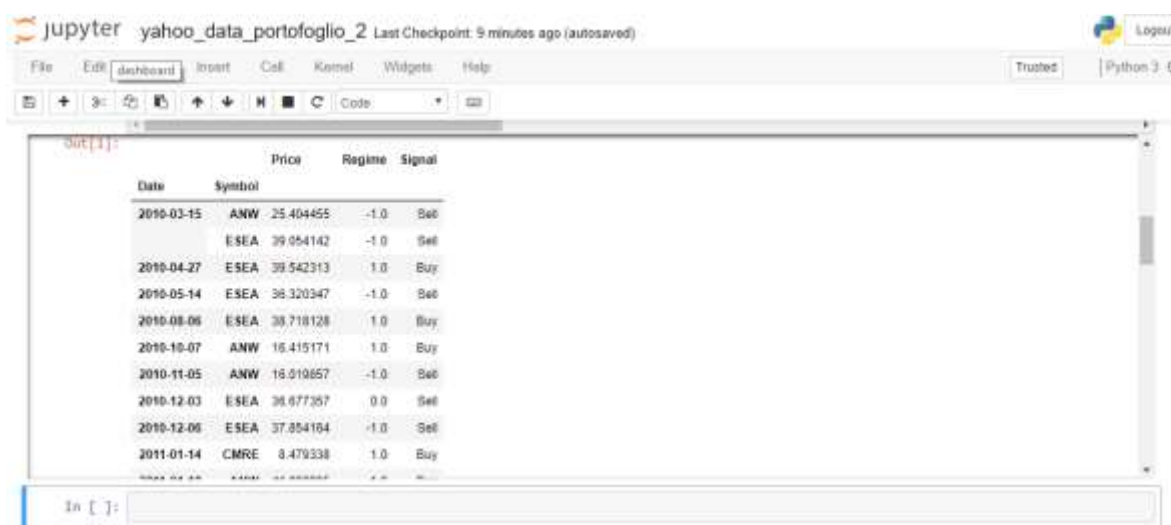
Εικόνα 5.17: Προσομοίωση χαρτοφυλακίου

Σύμφωνα με την προσομοίωση η αξία του χαρτοφυλακίου από το 2010 έως σήμερα αυξήθηκε κατά 4,9% σε περίπου 8 χρόνια. Δεδομένου ότι σε κάθε συναλλαγή επενδύθηκε μόνο το 10% της αξίας του χαρτοφυλακίου η απόδοση που παράχθηκε είναι πολύ καλή.

Μια πιο ρεαλιστική προσομοίωση χαρτοφυλακίου δεν πρέπει να εμπλέκει την συναλλαγή ενός αποθέματος αλλά περισσότερων. Σε αυτή την περίπτωση θα πρέπει να λαμβάνονται υπόψη όλες οι αγορές και πωλήσεις για όλες τις μετοχές του χαρτοφυλακίου. Για αυτό το λόγο στο χαρτοφυλάκιο προστέθηκαν οι μετοχές της EAGEAN, της COSTAMARE και της Euroseas.

Η συνάρτηση που υλοποιήθηκε λαμβάνει σαν όρισμα μια λίστα αποθεμάτων και καθορίζει πότε θα αγοράζεται ή θα πωλείται κάθε απόθεμα ανάλογα με τη στρατηγική crosswise cross-traffic, επιστρέφοντας ένα πλαίσιο δεδομένων με πληροφορίες σχετικά με το πότε αγοράζονται ή πωλούνται τα αποθέματα του χαρτοφυλακίου σύμφωνα με τη στρατηγική.

Η υλοποίηση των σημάτων πώλησης/αγοράς ανάλογα με την στρατηγική έδωσε σαν αποτέλεσμα την έξοδο της εικόνας 5. 18



The screenshot shows a Jupyter Notebook interface with a table of stock transactions. The table has the following data:

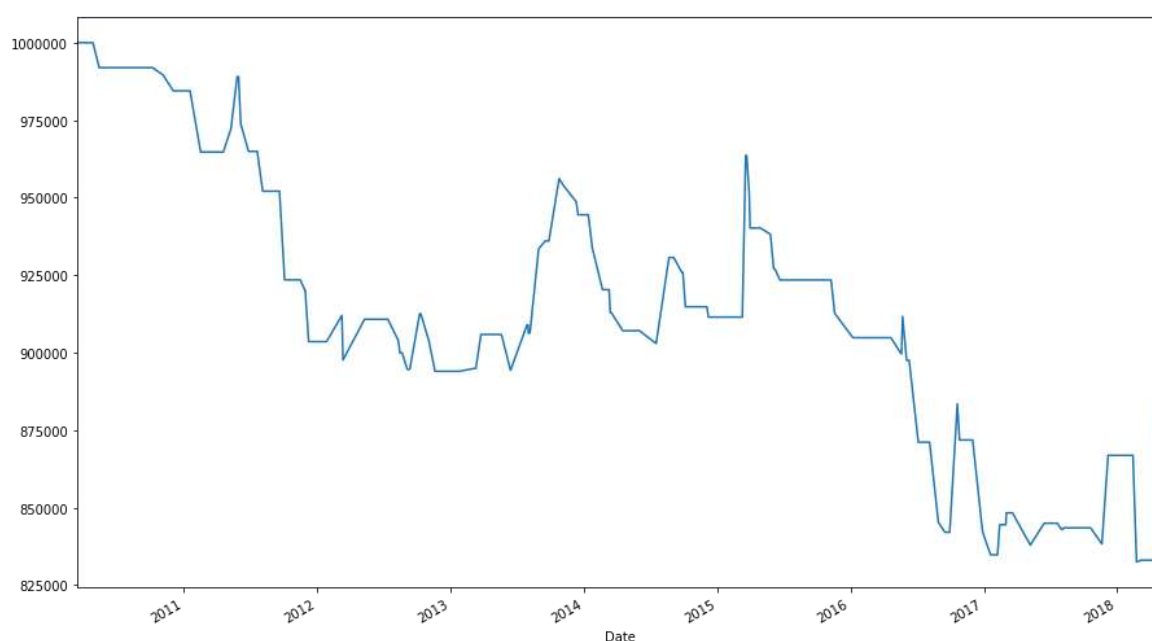
Date	Symbol	Price	Regime	Signal
2010-03-15	ANW	25.404455	-1.0	Sell
	ESEA	39.054142	-1.0	Sell
2010-04-27	ESEA	39.542313	1.0	Buy
2010-05-14	ESEA	36.320347	-1.0	Sell
2010-08-06	ESEA	39.718128	1.0	Buy
2010-10-07	ANW	16.415171	1.0	Buy
2010-11-05	ANW	16.010657	-1.0	Sell
2010-12-03	ESEA	36.677357	0.0	Sell
2010-12-06	ESEA	37.854104	-1.0	Sell
2011-01-14	CMRE	8.479338	1.0	Buy

Εικόνα 5.18: Προσομοίωση σημάτων πώλησης/αγοράς των μετοχών του χαρτοφυλακίου

Η υλοποίηση προσομοίωσης απολαβών του χαρτοφυλακίου περιλαμβάνει την δημιουργία ενός πίνακα με στήλες για την τελική αξία του χαρτοφυλακίου μετά από συναλλαγή, την απολαβή ανά συναλλαγή, την τιμή μετοχής, τον αριθμό μετοχών, την αξία συναλλαγής και τον τύπο συναλλαγής (αγορά / πώληση). Το διάγραμμα που δόθηκε σαν έξοδος δείχνει μια έντονη απώλεια στην αξία του χαρτοφυλακίου με περιόδους ανάκαμψης στα μέσα του 2011, στο τέλος του 2013 και στις αρχές του 2015 και έντονες τάσεις ανάκαμψης τις τελευταίες ημέρες (τέλη Απριλίου 2018). Όπως και να έχει η αξία του χαρτοφυλακίου βρίσκεται χαμηλότερα από την αρχική αξία με συνολική απώλεια της τάξης του 11%.

```
bk = backtest(signals, .1000000)
bk
```

Date	Symbol	End Cash	Portfolio Value	Profit per Share	Share Price	Shares	Start Cash	Total Profit	Trade Value	Type
2010-03-15	ANW	1000000.0000	1000000.0000	0.000000	25.404455	0.0	1000000.0000	0.0	0.0000	Sell
	ESEA	1000000.0000	1000000.0000	0.000000	39.054142	0.0	1000000.0000	0.0	0.0000	Sell
2010-04-27	ESEA	801144.2175	1000000.0000	0.000000	39.542313	2500.0	1000000.0000	0.0	98855.7825	Buy
2010-05-14	ESEA	901945.0850	991945.0850	-3.221968	36.320347	0.0	901144.2175	-0.0	90800.8675	Sell
2010-08-06	ESEA	895149.7650	991945.0850	0.000000	38.718128	2500.0	991945.0850	0.0	96795.3200	Buy
2010-10-07	ANW	795058.7300	991945.0850	0.000000	16.415171	6000.0	895149.7650	0.0	98491.0200	Buy
2010-11-05	ANW	892777.8810	869573.2010	-0.395314	16.019857	0.0	795058.7300	-0.0	96119.1420	Sell
2010-12-03	ESEA	984471.2735	984471.2735	-2.940771	36.677357	0.0	892777.8810	-0.0	91693.3925	Sell
2010-12-06	ESEA	984471.2735	984471.2735	-0.863964	37.854164	0.0	984471.2735	-0.0	0.0000	Sell
2011-01-14	CMRE	886110.9527	984471.2735	0.000000	8.479338	11600.0	984471.2735	0.0	98360.3208	Buy
2011-01-18	ANW	783439.8847	984471.2735	0.000000	11.098985	8800.0	886110.9527	0.0	97671.0660	Buy



Εικόνα 5.19: Προσομίωση απολαβών του χαρτοφυλακίου

Συμπερασματικά η στρατηγική που χρησιμοποιήθηκε μέχρι στιγμής δεν επέφερε κέρδη στον επενδυτή.

Αλλάζοντας στρατηγική και μειώνοντας το ποσοστό εξόδου από την θέση από 20% σε 10%, δεν παρατηρήθηκε καμιά διαφορά στην αξία του χαρτοφυλακίου.

Το Backtesting που πραγματοποιήθηκε αποτελεί μόνο μέρος της αξιολόγησης της αποτελεσματικότητας μιας στρατηγικής διαπραγμάτευσης. Ιδανικά η επιλεγμένη στρατηγική συγκρίνεται με άλλες γνωστές στρατηγικές, προκειμένου να καθορίσουμε την αποτελεσματικότητά της.

Η υπόθεση αποτελεσματικής αγοράς υποστηρίζει ότι είναι αδύνατο για μια οντότητα να αποκομίσει κέρδος πάνω από αυτό της αγοράς. Για το λόγο αυτό θα πρέπει να συγκρίνουμε την αξία του χαρτοφυλακίου, που αποτελείται από μετοχές εισηγμένες στα αμερικάνικα χρηματιστήρια, με μια τιμή που αντιπροσωπεύει την αξία των μετοχών του χρηματιστηριακού δείκτη S & P 500. Για το λόγο αυτό επιλέχθηκε ο SPY, ένα αμοιβαίο κεφάλαιο διαπραγματεύσιμο στο χρηματιστήριο σαν μετοχικό κεφάλαιο, η αξία του οποίου αντιπροσωπεύει την αξία των μετοχών του χρηματιστηριακού δείκτη S & P 500. Επειδή η σύγκριση πρέπει να γίνεται με το δείκτη του χρηματιστηρίου θα πραγματοποιηθεί μια απλή προσομοίωση αγοράς του SPY αλλά όχι προσομοίωση πώλησης του. Στην συνέχεια πραγματοποιείται μια παράλληλη οπτικοποίηση του χαρτοφυλακίου με την επίδοση του SPY.

Τα στοιχεία για τον SPY που παράχθηκαν για την αρχή και το τέλος της περιόδου που λήφθηκε υπόψη είναι τα ακόλουθα:



Date	Open	High	Low	Close	Adj Close	Volume
2009-12-31	112.759997	112.800003	111.389999	111.440002	94.356773	90637900
2018-05-04	251.519999	256.790009	251.149994	256.019989	256.019989	81780400

Εικόνα 5.20: Τιμές SPY αρχής και τέλους περιόδου

και υλοποιήθηκαν μέσω του:

```
spyder = web.DataReader('SPY', 'yahoo', start, end)
spyder.iloc[[0,-1],:]
```

Η τελική τιμή ενός χαρτοφυλακίου αξίας \$1.000.000 που αφορά μόνο τον SPY χωρίς στρατηγική πώλησης παράχθηκε όπως στην εικόνα 5:21

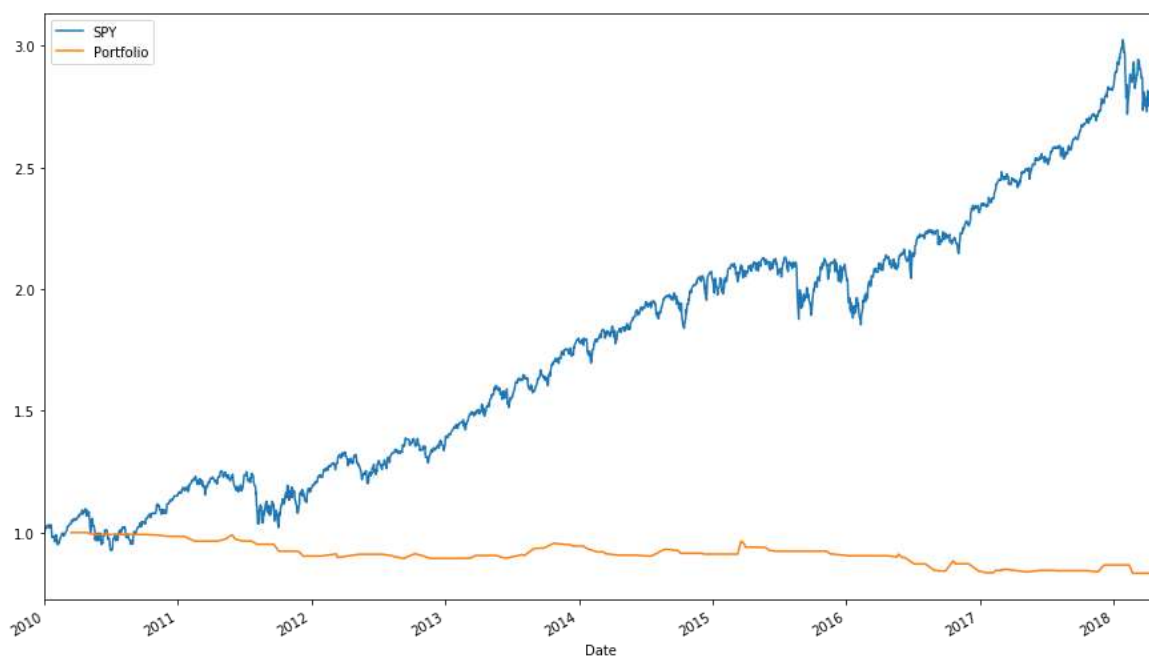


```
Out[1]: 2882463.7588888882
```

Εικόνα 5.21: Αξία χαρτοφυλακίου SPY

Η τελική αξία αποδεικνύει ότι στην περίοδο χρόνου που λήφθηκε υπόψη ο γενικός δείκτης είχε μια αύξηση σε αντίθεση με την απόδοση του χαρτοφυλακίου που δημιουργήθηκε υιοθετώντας την στρατηγική των κινούμενων μέσων.

Διαγραμματικά η συσχέτιση του χαρτοφυλακίου σε σχέση με την αξία του χαρτοφυλακίου μας που περιέχει μόνο την μετοχή SPY (εικόνα 5.22) οδηγεί στο ίδιο συμπέρασμα.



Εικόνα 5.22: Γραφική συσχέτιση χαρτοφυλακίων

ΕΠΙΛΟΓΟΣ

Η παρούσα πτυχιακή πραγματεύθηκε την ανάλυση χρηματοοικονομικών δεδομένων υψηλού όγκου και ταχύτητας όπως αυτά παρέχονται από την αποθήκη οικονομικών δεδομένων την Yahoo! Finance.

Στα πρώτα κεφάλαια της εργασίας πραγματοποιήθηκε μια αναφορά σε θεωρητικές έννοιες που εμπλέκονται τόσο στον χρηματοοικονομικό τομέα όσο και στον προγραμματισμό.

Ειδικότερα αναπτύχθηκαν έννοιες της γραμμικής και της μη γραμμικής χρηματοοικονομικής δυναμικής ενώ δόθηκε ιδιαίτερη έμφαση στις θεωρίες αγορών με κύρια την Υπόθεση Αποτελεσματικής αγοράς και τα προβλήματα που αυτές παρουσιάζουν. Σε αυτό το πλαίσιο αναφέρθηκαν επίσης τα είδη χρηματοπιστωτικών κινδύνων που επιφέρει μια χρηματιστηριακή συναλλαγή.

Αργότερα αναπτύχθηκαν οι σύγχρονες τεχνικές προβλέψεων πάνω στις χρονοσειρές και προβλέψεις. Σε αυτό το κομμάτι αναφέρθηκε οι τεχνικές του κινητού μέσου όρου και της λογαριθμικής διαφοράς που υιοθετήθηκε αργότερα στην υλοποίηση.

Την ανάπτυξη των χρηματοοικονομικών εννοιών ακολούθησε η ανάπτυξη εννοιών του προγραμματισμού. Αναφέρθηκαν , μεταξύ άλλων, το ελεύθερο λογισμικό και οι άδειες του καθώς επίσης και τα εργαλεία διαχείρισης των μεγάλων δεδομένων μεταξύ των οποίων το MapReduce και το Hadoop.

Σκοπός της υλοποίησης που πραγματοποιήθηκε ήταν η προσομοίωση ενός χαρτοφυλακίου πάνω σε μετοχές ελληνικών ναυτιλιακών μετοχών εισηγμένων στο χρηματιστήριο των Η.Π.Α. Για το σκοπό αυτό χρησιμοποιήθηκαν 3 διαφορετικές ναυτιλιακές εταιρίες και η στρατηγική που χρησιμοποιήθηκε –βασισμένη στους κινητούς μέσους - επέφερε μια ζημία της τάξης του 11% στην αξία του χαρτοφυλακίου.

Μπορούν να πραγματοποιηθούν διάφορες ενέργειες ως προς την βελτίωση απόδοσης του συστήματος. Μεταξύ αυτών αναφέρουμε τη διαφοροποίηση του χαρτοφυλακίου. Όλες οι μετοχές που θεωρήσαμε ήταν ελληνικών ναυτιλιακών εταιριών γεγονός που σημαίνει ότι αν ο κλάδος δεν πηγαίνει καλά αυτό έχει αντίκτυπο και στο χαρτοφυλάκιο μας. Θα μπορούσε επίσης να αναπτυχθεί ένα σύστημα το οποίο διαθέτει μικρά αποθέματα ώστε το χαρτοφυλάκιο να μπορέσει να επωφεληθεί από την κίνηση προς οποιαδήποτε κατεύθυνση.

Επίσης το σύστημα θα μπορούσε να περιλαμβάνει μια μελλοντική απόδοση και στρατηγικές για την υλοποίηση της.

Τέλος να σημειωθεί ότι η τεχνολογία Anaconda σε συνδυασμό με το Notebook της Jupyter αποδείχτηκε ιδιαίτερα ασταθής και δημιούργησε έναν αριθμό προβλημάτων κατά την υλοποίηση.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΞΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- Angelopoulos P. and Mourdoukoutas P.** «Banking Risk Management in a Globalising economy [Book]. - London : Westport, Conn. USA: Quorum Books, 2001.
- Brealey Richard, Myers Stewart and Allen Franklin** Principles of Corporate Finance, 12th Edition [Book]. - New York : McGraw-Hill/Irwin, 2017.
- Brock W.A. [et al.]** A test for independence based on the correlation dimension. [Article] // *Econometric Reviews*. - 1996. - Vol. 15. - pp. 197–235.
- Cuthbertson K. and Nitzsche D.** Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange, 2nd Edition [Book]. - [s.l.] : John Wiley&Sons LTD, 2004.
- Damodaran A.** Damodaran on Valuation: Security Analysis for Investment and Corporate Finance, 2nd Edition [Book]. - [s.l.] : USA: John Wiley & Sons, Inc, 2006.
- De Moor L, Van den Bossche F and Verheyden T** A tale of market efficiency. [Article] // HUB Research Paper Series, Vol. 5.. - 2013.
- Dean Jeffrey and Ghemawat Sanjay** MapReduce: simplified data processing on large clusters [Conference] // OSDI'04 Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation - Volume 6. - San Francisco : USENIX Association Berkeley, CA, USA, 2004. - p. 10.
- Fama E** The Behavior of Stock Market Prices [Book]. - [s.l.] : Ph.D Assertion, 1965.
- Fama E.** Efficient Capital Markets: A Review of Theory and Empirical Work [Article] // *Journal of Finance*. - 1970. - Vol. 25. - pp. 383–417.
- Finard J. and Stocks M.** A Framework for Corporate Financial Risk Management [Article] // included in Klein R. & Lederman J., *Derivatives Risk and Responsibility*, Irwin. - 1996.
- Hilpisch Yves** Python for Finance: Analyze Big Financial Data [Book]. - Sebastopol, CA 95472 : O' Reilly Media, Inc.. - 1st.
- Hinich M. J.** Testing for Gaussianity and linearity of a stationary time series. [Article] // *Journal of Time Series Analysis*. - 1982. - 3 : Vol. 3.
- Hinich M.J.** Testing for dependence in the input to a linear time series model [Article] // *Journal of Non-parametric Statistics*. - 1996. - 6. - pp. 205-221.
- Holton G. A.** Value-at-Risk: Theory and Practice [Book]. - [s.l.] : Elsevier Academic Press, 2014. - 2. - e-book at <http://value-at-risk.net>.
- Jain C. L.** Business Forecasting in the 21st Century [Article] // *The Journal of Business Forecasting* . - 2003. - pp. 2-6 .

Jorion P Financial Risk Manager Handbook [Book]. - New Jersey : John Wiley and Sons, 2003. - 2nd edition.

Lee T. H., White H. and Granger C. W. J. Testing for Neglected Nonlinearities in Time Series Models. [Article] // Journal of Econometrics. - 1993. - Vol. 56. - pp. 269-290.

Lewis Michael Flash Boys. [Book]. - New York : W. W. Norton & Company, 2014.

Lintner John The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets [Article] // The Review of Economics and Statistics. - Feb. 1965. - 1 : Vol. 47. - pp. 13-37.

Lo A Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis [Article] // Journal of Investment Consulting. - 2005. - 2 : Vol. 7. - pp. 21-44.

Lo A The adaptive markets hypothesis: market efficiency from an evolutionary perspective [Article] // Journal of Portfolio Management. - 2004. - Vol. 30. - pp. 15-29.

Lo A. Νέος Κόσμος: The Complete Hellenic Perspective [Ηλεκτρονικό]. - 7 Αυγούστος 2012. - <http://neoskosmos.com/news/en/node/24438>.

Lo A. and MacKinlay C. A Non-Random Walk Down Wall Street [Book]. - Princeton : NJ:Princeton University Press, 1999.

Lo A. Milken Institute Review [Online]: Adaptive markets: financial evolution at the speed of thought. - May 2005. - Ιανουάριος 5, 2018. - <http://www.milkenreview.org/articles/adaptive-markets-financial-evolution-at-the-speed-of-thought>.

Lo A. Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis [Article] // Journal of Investment Consulting. - 2 : Vol. 7. - pp. 21-44.

Makridakis S., Wheelwright SC. and McGee V.E. Forecasting: Methods and Applications [Book]. - New York : John Wiley and Sons, 1983. - 2.

Mark R. Innovative Strategies and Techniques for Pricing Contingent Credit Risk [Article] // included in Klein R. & Lederman J., Derivatives Risk and Responsibility, Irwin. - 1996.

McLeod A. I. and Li W. K. Diagnostic checking ARMA time series models using squared-residual autocorrelations. [Article] // Journal of Time Series Analysis. - 1983. - 4. - Vol. 4. - pp. 269-273.

Mishkin F. S. and Eakins S. G. Financial Markets and Institutions, 8th Edition [Book]. - Boston : Prentice Hall, 2012.

Mossin Jan Equilibrium in a Capital Asset Market [Article] // Econometrica. - Oct. 1966. - 4 : Vol. 34. - pp. 768-783.

Sharpe William F. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk [Article] // The Journal of Finance. - Sep. 1964. - 3 : Vol. 19. - pp. 425-442.

Wickham Hadley ggplot2, Use R [Book]. - New York : Springer Science+Business Media, LLC, 2009.

Wilkinson Leland Grammar of Graphics (Statistics and Computing) [Book]. - New York : Springer Science+Business Media LLC, 2005. - 2nd Edition.

ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Αριστοτέλης Πολιτικά [Βιβλίο] / μεταφρ. Δημήτρης Παπάδης Ι.. - Μετάφραση-Επιμέλεια Δημήτρης, Παπάδης Ι. : Εκδόσεις ΖΗΤΡΟΣ.

Θεοτοκάς Γ και Χαρλαύτη Τζ Έλληνες εφοπλιστές και ναυτιλιακές εταιρίες [Βιβλίο]. - Αθήνα : Εκδόσεις Αλεξάνδρεια, 2007.

Κουγιουμτζής Δημήτριος Προσομοίωση χασοτικών χρονοσειρών Μέρος Β - Μη-γραμμική ανάλυση χρονοσειρών [Βιβλίο]. - [s.l.] : Πανεπιστημιακές σημειώσεις, 2005.

ΠΟΥΦΙΝΑΣ ΘΩΜΑΣ και ΦΛΩΡΟΣ ΧΡΗΣΤΟΣ Β. ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΑ ΠΑΡΑΓΩΓΑ [Βιβλίο]. - ISBN13 9789609495400 : ΕΚΔΟΣΕΙΣ ΔΙΣΙΓΜΑ, Δεκέμβριος 2014.

Συριόπουλος Κώστας και Λεοντίσης Αλέξανδρος Χάος, Ανάλυση και πρόβλεψη χρονοσειρών [Βιβλίο]. - Θεσσαλονίκη : ΑΝΙΚΟΥΛΑ, 2000.

Φίλιππας Ν. Επενδύσεις [Βιβλίο]. - Αθήνα : Πανεπιστημιακό, 2005.

ΔΙΑΔΙΚΤΥΑΚΕΣ ΠΗΓΕΣ

A byte of Python [Ηλεκτρονικό] // [e-book] GitBook. - 17 Μάρτιος 2018. - <http://www.swaroopch.com/notes/Python>.

AEGEAN AEGEAN MARINE PETROL NETWORK INC [Ηλεκτρονικό] // AEGEAN MARINE PETROL NETWORK INC. - 2018. - 30 Απρίλιος 2018. - <http://www.ampni.com/>.

ANACONDA CLOUD ::ANACONDA CLOUD [Ηλεκτρονικό] // ANACONDA CLOUD. - 2018. - 22 Απρίλιος 2018. - <https://anaconda.org/>.

ANACONDA DOWNLOADS ANACONDA CLOUD [Ηλεκτρονικό] // DOWNLOADS. - ANACONDA CLOUD, 15 Φεβρουάριος 2018. - 22 Απρίλιος 2018. - <https://www.anaconda.com/download/>.

Appuswamy Raja [και συν.] Nobody Ever Got Fired for Buying a Cluster [Ηλεκτρονικό]. - Microsoft Research, Cambridge, England, 2 January 2013. - 13 4 2018. - <https://www.microsoft.com/en-us/research/publication/nobody-ever-got-fired-for-buying-a-cluster/>.

Arduino Arduino [Ηλεκτρονικό] // Arduino. - 2006. - <https://www.arduino.cc/>.

Costamare Costamare Inc. [Ηλεκτρονικό] // Costamare Inc.. - 2018. - 30 Απρίλιος 2018. - <http://www.costamare.com/>.

Curtis Miller An Introduction to Stock Market Data Analysis with Python (Part 1) [Ηλεκτρονικό] // WordPress.com. - 2016. - 20 Ιούνιος 2017. - <https://ntguardian.wordpress.com/2016/09/19/introduction-stock-market-data-python-1/>.

Efstathopoulos Georgios Python for Finance, Part I: Yahoo & Google Finance API, pandas, and matplotlib [Ηλεκτρονικό] // LearnDataSci. - 2017. - 20 Ιούνιος 2017. - <https://www.learn datasci.com/tutorials/python-finance-part-yahoo-finance-api-pandas-matplotlib/#>.

EurexGroup Eurex Exchange - Volatility Derivatives [Ηλεκτρονικό] // Eurex Group. - Eurex Group, 1999. - 15 Απρίλιος 2018.

Free Software Foundation [Ηλεκτρονικό]. - 8 Μάρτιος 2018. - <https://www.fsf.org/>.

GNU Operating System [Ηλεκτρονικό]. - 1992. - 8 Μάρτιος 2018. - <https://www.gnu.org/home.en.html>.

Jupyter Notebook Jupyter Notebook [Ηλεκτρονικό] // Jupyter Notebook. - Jupyter Notebook, 2015. - 23 Απρίλιος 2018. - <https://jupyter-notebook.readthedocs.io/en/stable/>.

Linux Linux.org [Ηλεκτρονικό] // Linux. - 1983. - <https://www.linux.org/>.

matplotlib Matplotlib Python [Ηλεκτρονικό] // matplotlib. - matplotlib, 2002-2012. - 26 Απρίλιος 2018. - <https://matplotlib.org/>.

MediaWiki MediaWiki [Ηλεκτρονικό] // MediaWiki.org. - 4 Δεκέμβριος 2016. - <https://www.mediawiki.org/wiki/MediaWiki>.

moodle moodle [Ηλεκτρονικό] // moodle. - 1999. - <https://moodle.org/>.

OLPC OLPC XO [Ηλεκτρονικό] // OLP XO. - 2008. - <http://laptop.org/en/laptop/start/activities.shtml>.

Open Source Initiative [Ηλεκτρονικό]. - March 2018. - [Opensource.org](https://opensource.org).

OpenStreetMap OpenStreetMap [Ηλεκτρονικό] // OpenStreetMap. - 2004. - <https://www.openstreetmap.org/#map=6/38.359/23.810>.

Organization Debian Debian Social Contract [Ηλεκτρονικό]. - 26 April 2004. - 8 Μάρτιος 2018. - https://www.debian.org/social_contract#guidelines.

Python [Ηλεκτρονικό]. - 17 Μάρτιος 2018. - <https://www.python.org>.

Raspberry Raspberry [Ηλεκτρονικό] // Raspberry. - 2012. - <https://www.raspberrypi.org/>.

RepRap RepRap [Ηλεκτρονικό] // RepRap. - 2005. - <http://reprap.org/>.

SugarCRM SugarCRM [Ηλεκτρονικό] // SugarCRM. - 2004. - <http://www.sugarcrm.com/download>.

tutorialspoint tutorialspoint [Ηλεκτρονικό] // `hadoop LEARN HADOOP big data analysis framework`. - 2018. - 15 Απρίλιος 2018. -
https://www.tutorialspoint.com/hadoop/hadoop_big_data_overview.htm.