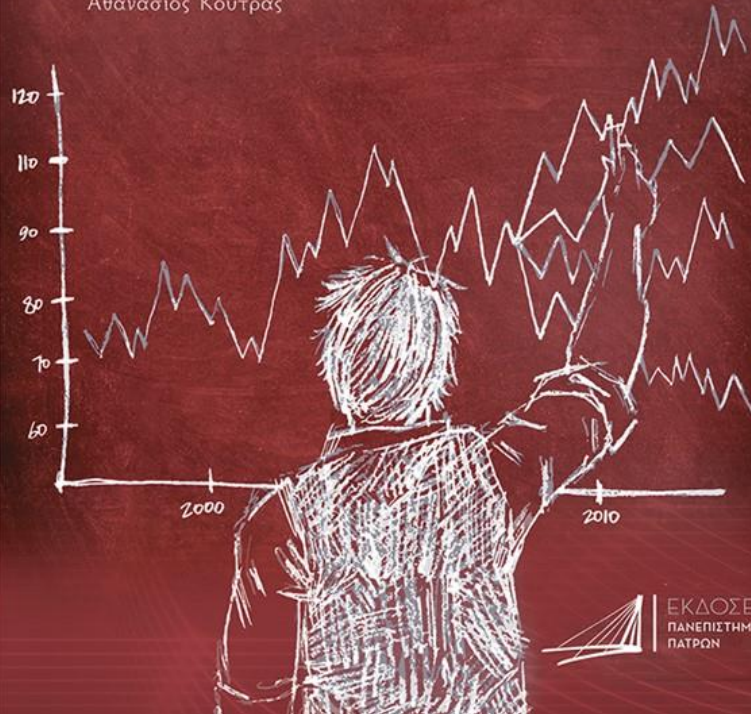


Rob J Hyndman
George Athanasopoulos

ΠΡΟΒΛΕΨΗ

ΑΡΧΕΣ ΚΑΙ ΠΡΑΚΤΙΚΗ

Επιστημονική Επιμέλεια - Απόδοση
Ιωάννης Α. Νίκας
Αθανάσιος Κούτρας



ΕΚΔΟΣΕΙΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ
ΠΑΤΡΩΝ

Πρόβλεψη: Αρχές και Πρακτική

ΣΥΓΓΡΑΦΕΙΣ:

Rob J. Hyndman, George Athanasopoulos

ΕΠΙΣΤΗΜΟΝΙΚΗ ΕΠΙΜΕΛΕΙΑ - ΑΠΟΔΟΣΗ:

Ιωάννης Νίκας, Αθανάσιος Κούτρας



Οι συντελεστές του βιβλίου

Rob Hyndman

Ο **Rob J Hyndman** είναι Καθηγητής Στατιστικής του Τμήματος Econometrics and Business Statistics στο Monash University, στην Αυστραλία. Από το 2005-2018 είχε διατελέσει Editor-in-Chief στο περιοδικό *International Journal of Forecasting*.

Είναι συγγραφέας 6 βιβλίων και περισσότερων από 200 ερευνητικών δημοσιεύσεων, και εκλεγμένος Fellow of the Academy for the Social Sciences στην Αυστραλία.

Για περισσότερα από 35 χρόνια, ο Hyndman διατηρεί θέση συμβούλου σε εκατοντάδες εταιρείες και οργανισμούς στο θεματικό πεδίο της πρόβλεψης. Του έχουν απονεμηθεί βραβεία για την έρευνα, τη διδασκαλία, τη συμβουλευτική και την επίβλεψη φοιτητών.



robjhyndman.com

George Athanasopoulos

Ο **George Athanasopoulos** είναι Καθηγητής, Διευθυντής Εκπαίδευσης και Πρόεδρος του Τμήματος Econometrics and Business Statistics στο Monash University, της Αυστραλίας.

Είναι, επίσης, Πρόεδρος του International Institute of Forecasters. Του έχει απονεμηθεί μεγάλος αριθμός βραβείων και διακρίσεων για την έρευνά του και τη διδασκαλία.

Είναι μέλος των Editorial Boards των περιοδικών *Journal of Travel Research* και *International Journal of Forecasting*



Ιωάννης Νίκας

Ο Ιωάννης Νίκας είναι Επίκουρος Καθηγητής στο Τμήμα Διοίκησης Τουρισμού στο Πανεπιστήμιο Πατρών και μέλος του εργαστηρίου Πληροφοριακών Συστημάτων & Προβλέψεων στον Τουρισμό.

Τα τρέχοντα ερευνητικά του ενδιαφέροντα περιλαμβάνουν την ανάπτυξη κι εφαρμογή τεχνικών ανάλυσης δεδομένων, την ανάπτυξη στοχαστικών και υπολογιστικών μεθόδων για την πρόβλεψη χρονοσειρών, οπτικοποίηση και χωρική ανάλυση δεδομένων με χρήση Γεωγραφικών Πληροφοριακών Συστημάτων (GIS) σε τουριστικά και αθλητικά δεδομένα, καθώς και την ανάπτυξη μεθόδων αριθμητικής ανάλυσης για την εύρεση αξιόπιστων λύσεων σε προβλήματα μη γραμμικών συστημάτων και βελτιστοποίησης.

Αθανάσιος Κούτρας

Ο **Αθανάσιος Κούτρας** είναι Αναπληρωτής Καθηγητής του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Παν. Πελοποννήσου. Είναι διπλωματούχος Ηλεκτρολόγος Μηχανικός και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών και διδάκτωρ του ιδίου Τμήματος.

Στα ερευνητικά του αντικείμενα περιλαμβάνονται αυτά της ψηφιακής επεξεργασίας σήματος και εικόνας, ανάλυσης βιοϊατρικού σήματος και εικόνας καθώς και πρόβλεψης χρονοσειρών.

Είναι μέλος του Τεχνικού Επιμελητηρίου Ελλάδας, της IEEE, της EMBS, του IEEE Signal Processing Society της European Neural Network Society (ENNS) και της International Speech Communication Association (ISCA), ενώ τα τελευταία 3 χρόνια συμμετέχει ως μέντορας στην IEEE.

*Αναγγείλατε τα επερχόμενα επ’
εσχάτου και γνωσόμεθα ότι θεοί
εστέ.*

*(πείτε μας τι επιφυλάσσει το μέλλον, για να γνωρίζουμε, εάν
πράγματι είστε Θεοί)*

Ησαΐας 41:23, ~700 π.Χ.

Διάσημες προβλέψεις...

- ✓ Πιστεύω πως υπάρχει μια παγκόσμια αγορά για το **πολύ πέντε υπολογιστές**.

Πρόεδρος της IBM, 1943

- ✓ Οι υπολογιστές στο μέλλον δεν θα ζυγίζουν περισσότερο από **1,5 τόνους**.

Popular Mechanics, 1949

- ✓ Δεν υπάρχει κανένας λόγος να θέλει κάποιος έναν υπολογιστή στο σπίτι του.

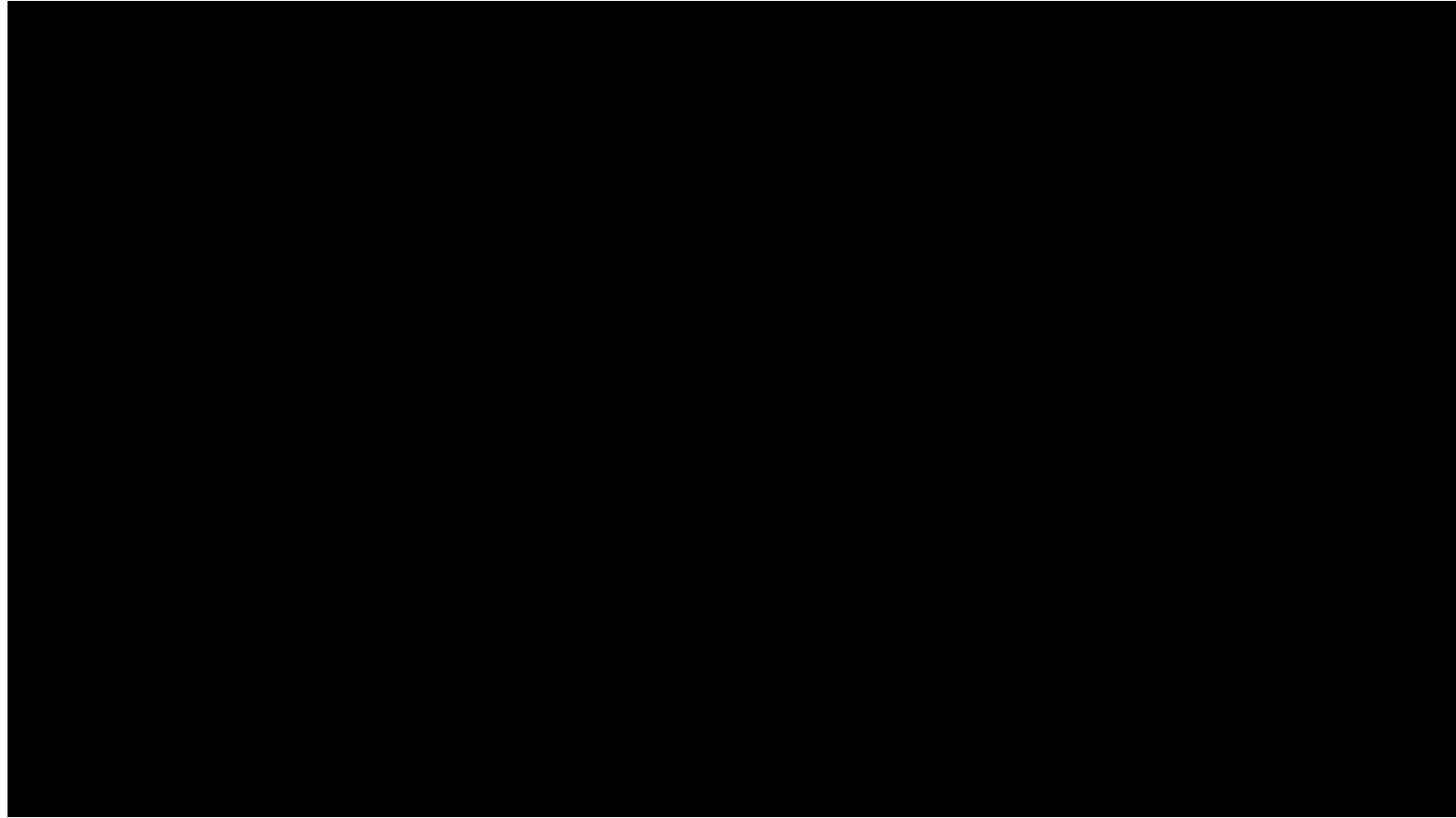
Πρόεδρος, DEC, 1977





Παρουσίαση του βιβλίου

Οι συγγραφείς
για το βιβλίο



Σκοπός του βιβλίου

Το βιβλίο έχει **σκοπό**

- να παρέχει μια εμπειριστατωμένη εισαγωγή στις μεθόδους πρόβλεψης
- να παρουσιάσει στους αναγνώστες του αρκετές πληροφορίες για κάθε μέθοδο με σκοπό την εύλογη χρήση τους.

Το βιβλίο **δεν έχει σκοπό**

- να παρουσιάσει διεξοδικά και αναλυτικά όλες τις θεωρητικές λεπτομέρειες κάθε μεθόδου

Παραθέτει **πλήθος βιβλιογραφικών αναφορών** στο τέλος κάθε κεφαλαίου

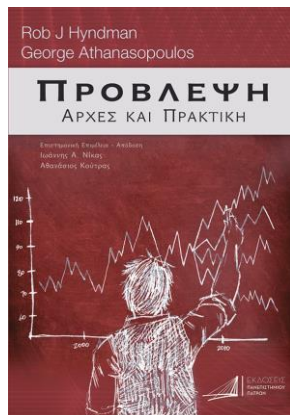
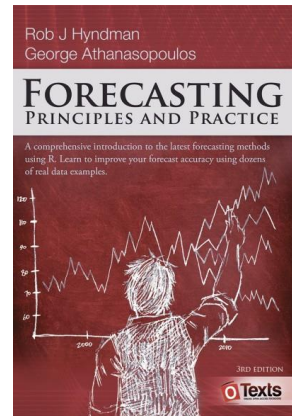
Σε ποιους απευθύνεται

Το βιβλίο είναι γραμμένο για **τρία είδη** ακροατηρίων:

- για **προπτυχιακούς φοιτητές** που σπουδάζουν διοίκηση επιχειρήσεων
- για **μεταπτυχιακούς φοιτητές** στη διοίκηση επιχειρήσεων (MBA) που μελετούν το αντικείμενο των προβλέψεων κατ' επιλογήν.
- για εκείνους που πραγματοποιούν προβλέψεις στις **επιχειρήσεις**, όταν μπορεί να μην είχαν οποιαδήποτε σχετική τυπική εκπαίδευση.

Σε ποιους απευθύνεται

Οι συγγραφείς χρησιμοποιούν το βιβλίο σε ένα μάθημα **τρίτου έτους** για φοιτητές που θα οδηγηθούν σε **Bachelor of Commerce** ή σε **Bachelor of Business** στο Πανεπιστήμιο Monash της Αυστραλίας.



Στην Ελλάδα χρησιμοποιείται σε μάθημα **τρίτου έτους** για φοιτητές που θα οδηγηθούν σε πτυχίο **Διοίκησης Τουρισμού** στο Πανεπιστήμιο της Πάτρας.

Γνωσιακές απαιτήσεις

Για τις περισσότερες ενότητες, οι αναγνώστες χρειάζεται να είναι εξοικειωμένοι μόνο με τις εισαγωγικές έννοιες της **στατιστικής και της άλγεβρας** επιπέδου Γυμνασίου - Λυκείου.

Μερικές (λίγες) ενότητες απαιτούν **γνώση πινάκων**, αλλά αυτές επισημαίνονται.

Μικρή εξοικείωση των αναγνωστών με **βασικές αρχές υπολογιστών** για την εγκατάσταση και εκτέλεση των προγραμμάτων που περιλαμβάνονται στο σύγγραμμα.

Πρακτική εφαρμογή

Σε όλο το βιβλίο χρησιμοποιείται η γλώσσα προγραμματισμού **R**.

www.r-project.org

Ο στόχος είναι να μάθουν οι φοιτητές πώς να κάνουν προβλέψεις με την γλώσσα αυτή.

Η R είναι **δωρεάν** και διαθέσιμη σχεδόν σε **κάθε λειτουργικό σύστημα**. Είναι ένα θαυμάσιο εργαλείο για όλες τις στατιστικές αναλύσεις, όχι μόνο για τις προβλέψεις.

Στο Παράρτημα του βιβλίου αναφέρονται οι οδηγίες εγκατάστασης της γλώσσας αυτής, καθώς και των πακέτων και βιβλιοθηκών που περιλαμβάνουν τις συναρτήσεις που υλοποιούν όλες τις μεθόδους που περιλαμβάνονται στο βιβλίο.

Εγκατάσταση της γλώσσας R

(χωρίς να είσαι προγραμματιστής)

Παράρτημα: Χρησιμοποιώντας την R

Το βιβλίο αυτό χρησιμοποιεί την γλώσσα R και είναι σχεδιασμένο για να διαβάζεται μαζί με την γλώσσα R. Η R είναι μια δωρεάν διαθέσιμη γλώσσα για όλα τα λειτουργικά συστήματα για την οποία υπάρχουν χιλιάδες επιπρόσθετα πακέτα με τα οποία μπορείτε να κάνετε οτιδήποτε θελήσετε. Προτείνουμε να χρησιμοποιείτε την R μαζί με το RStudio.

Εγκατάσταση R και RStudio

1. Μεταφόρτωση και εγκατάσταση R.¹
2. Μεταφόρτωση και εγκατάσταση RStudio.²
3. Εκτέλεση RStudio. Στην καρτέλα «Packages», επιλέξτε «Install» και εγκαταστήστε το πακέτο **fpp3** (Προσοχή. Θα πρέπει να έχετε επιλέξει την επιλογή «install dependencies»).

Αυτό είναι! Τώρα είστε σε θέση να ξεκινήσετε.

¹ <https://cran.r-project.org/>

² <https://bit.ly/rstudiodownload>

Μαθαίνοντας την R

(χωρίς να είσαι προγραμματιστής)

Ξεκινώντας με την R

Αν δεν έχετε χρησιμοποιήσει ξανά την R στο παρελθόν, προτείνουμε να ξεκινήσετε από την πρώτη ενότητα (Κεφάλαια 1-8) του «R for Data Science»³ των Garrett Golemund και Hadley Wickham. Αν και το βιβλίο αυτό δεν καλύπτει το αντικείμενο της ανάλυσης χρονοσειρών ή της πρόβλεψης, θα σας εισάγει στα βασικά της γλώσσας R, και στα πακέτα tidyverse. Επιπλέον προτείνεται το διαδικτυακό μάθημα (MOOC) «Coursera R Programming».⁴

Χρησιμοποιώντας τις ασκήσεις του βιβλίου αυτού, θα μάθετε πως να χρησιμοποιείτε την R για την πρόβλεψη χρονοσειρών.

Δυνατά σημεία του βιβλίου

Ενσωματώνει δεκάδες **πραγματικά παραδείγματα** δεδομένων από τη συμβουλευτική πρακτική των συγγραφέων.

Οι συγγραφείς έχουν συνεργαστεί με εκατοντάδες επιχειρήσεις και οργανισμούς που τους βοήθησαν σε ζητήματα πρόβλεψης.

Η εμπειρία έχει συμβάλει άμεσα σε πολλά από τα παραδείγματα που δίνονται, καθώς επίσης και στη διαμόρφωση της γενικότερης φιλοσοφίας του βιβλίου για τις προβλέψεις.

Δίνει μεγαλύτερη έμφαση σε **γραφικές μεθόδους** σε σχέση με τους περισσότερους προγνώστες.

Δυνατά σημεία του βιβλίου

Χρησιμοποιούνται **γραφήματα για τη διερεύνηση των δεδομένων**, την ανάλυση της εγκυρότητας των μοντέλων που εφαρμόστηκαν και της αναλυτικής παρουσίασης των αποτελεσμάτων των προβλέψεων.

Χρησιμοποιεί την γλώσσα R, η οποία είναι δωρεάν, ανοιχτού κώδικα και ένα εξαιρετικά ισχυρό λογισμικό για την πρακτική εφαρμογή των μεθόδων σε πραγματικά παραδείγματα.



Περιεχόμενα βιβλίου



Περιεχόμενα

Περιεχόμενα	v
Πρόλογος	ix
Πρόλογος για την ελληνική έκδοση	xiii
1 Εισαγωγή	1
1.1 Τι μπορεί να προβλεφθεί;	2
1.2 Πρόβλεψη, στόχοι και σχεδιασμός	4
1.3 Καθορίζοντας τι θα προβλεφθεί	5
1.4 Μέθοδοι και δεδομένα πρόβλεψης	6
1.5 Μερικές μελέτες περίπτωσης	10
1.6 Τα βασικά βήματα σε ένα πρόβλημα πρόβλεψης	12
1.7 Η προοπτική της στατιστικής πρόβλεψης	13
1.8 Ασκήσεις	16
1.9 Περαιτέρω μελέτη	16



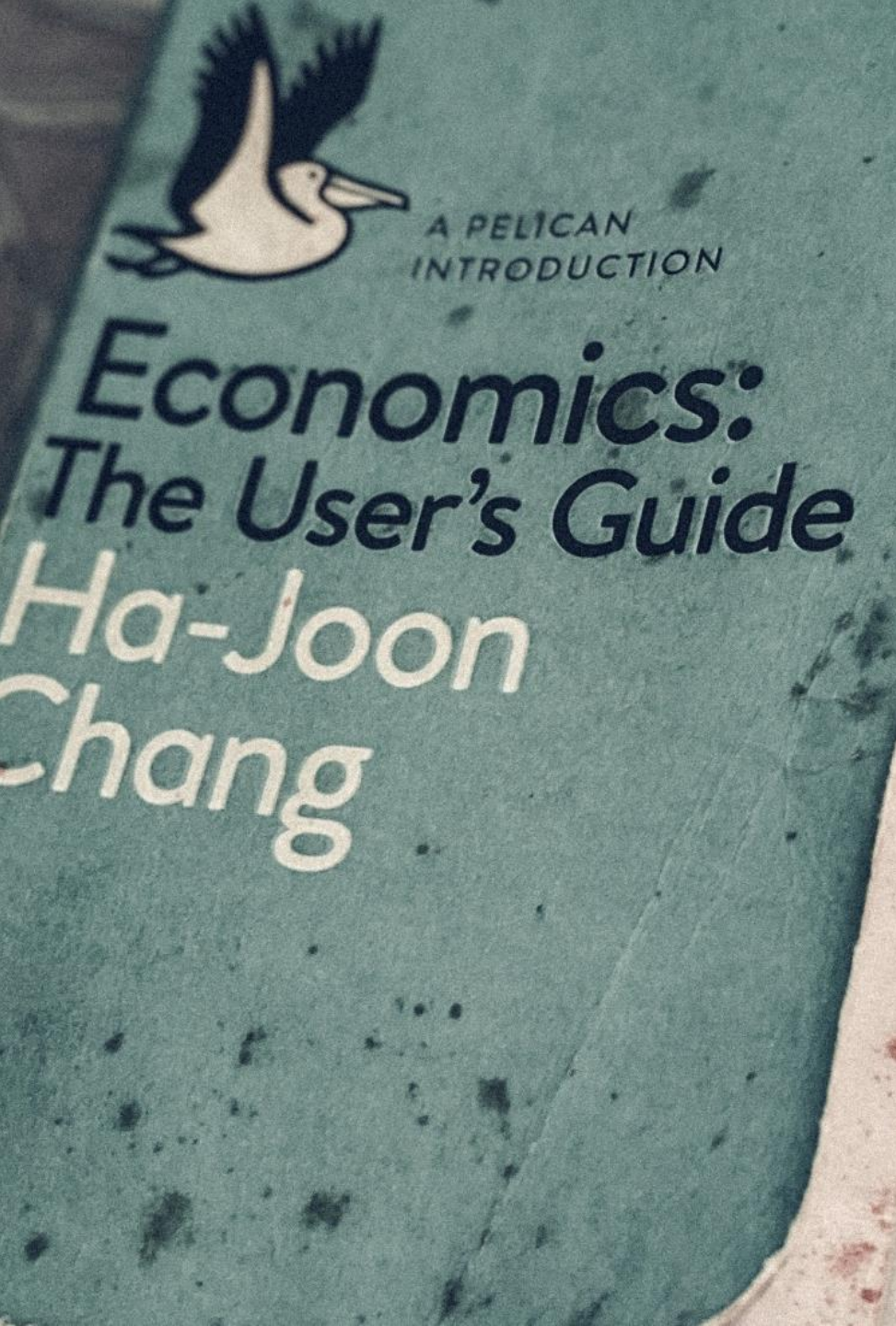
2	Απεικόνιση Χρονοσειρών	17
2.1	Τα αντικείμενα tsibble	17
2.2	Χρονοδιαγράμματα	25
2.3	Πρότυπα χρονοσειρών	27
2.4	Διαγράμματα εποχικότητας	29
2.5	Διαγράμματα εποχικότητας υποσειρών	32
2.6	Διαγράμματα διασποράς	35
2.7	Διαγράμματα υστέρησης	41
2.8	Αυτοσυσχέτιση	42
2.9	Λευκός θόρυβος	44
2.10	Ασκήσεις	45
2.11	Περαιτέρω μελέτη	48
3	Ανάλυση Χρονοσειρών	49
3.1	Μετασχηματισμοί και προσαρμογές	49
3.2	Συνιστώσες χρονοσειράς	55
3.3	Κινητοί μέσοι	59
3.4	Κλασική μέθοδος ανάλυσης	66
3.5	Μέθοδοι που χρησιμοποιούνται από επίσημους στατιστικούς οργανισμούς	68
3.6	Ανάλυση STL	72
3.7	Ασκήσεις	75
3.8	Περαιτέρω μελέτη	77



4	Χαρακτηριστικά Χρονοσειρών	79
4.1	Μερικά απλά στατιστικά μεγέθη	79
4.2	Χαρακτηριστικά αυτοσυσχέτισης (ACF)	81
4.3	Χαρακτηριστικά STL	82
4.4	Άλλα χαρακτηριστικά	85
4.5	Μελέτη τουριστικών δεδομένων στην Αυστραλία	87
4.6	Ασκήσεις	92
4.7	Περαιτέρω μελέτη	92
5	Η εργαλειοθήκη του προγνώστη	93
5.1	Μια τακτοποιημένη ροή εργασιών πρόβλεψης	93
5.2	Μερικές απλές μέθοδοι πρόβλεψης	98
5.3	Προσαρμοσμένες τιμές και υπόλοιπα	103
5.4	Διαγνωστικά υπολοίπων	105
5.5	Κατανομημένες προβλέψεις και διαστήματα πρόβλεψης	111
5.6	Πρόβλεψη χρησιμοποιώντας μετασχηματισμούς	119
5.7	Πρόβλεψη με ανάλυση	121
5.8	Αξιολόγηση ακρίβειας σημειακών προβλέψεων	124
5.9	Αξιολόγηση ακρίβειας κατανομημένων προβλέψεων	131
5.10	Διασταυρούμενη επικύρωση χρονοσειρών	135
5.11	Ασκήσεις	139
5.12	Περαιτέρω μελέτη	142



6	Πρόβλεψη με κρίση	143
6.1	Προσοχή στους περιορισμούς	144
6.2	Βασικές αρχές	145
6.3	Η μέθοδος Delphi	150
6.4	Προβλέψεις κατ' αναλογία	153
6.5	Πρόβλεψη σεναρίου	154
6.6	Πρόβλεψη νέου προϊόντος	155
6.7	Προσαρμογές στην κρίση	157
6.8	Περαιτέρω Μελέτη	160
7	Μοντέλα παλινδρόμησης χρονοσειρών	161
7.1	Το γραμμικό μοντέλο	161
7.2	Εκτίμηση ελάχιστων τετραγώνων	167
7.3	Αξιολόγηση του μοντέλου παλινδρόμησης	172
7.4	Μερικοί χρήσιμοι παράγοντες πρόβλεψης	179
7.5	Επιλογή παραγόντων πρόβλεψης	187
7.6	Πρόβλεψη με παλινδρόμηση	192
7.7	Μη γραμμική παλινδρόμηση	198
7.8	Συσχέτιση, αιτιότητα και πρόβλεψη	202
7.9	Διατύπωση με πίνακες	205
7.10	Ασκήσεις	207
7.11	Περαιτέρω μελέτη	210



8	Εκθετική εξομάλυνση	213
8.1	Απλή εκθετική εξομάλυνση	214
8.2	Μέθοδοι με τάση	220
8.3	Μέθοδοι με εποχικότητα	226
8.4	Μια ταξινόμηση των μεθόδων εκθετικής εξομάλυνσης	232
8.5	Καινοτόμος χώρος καταστάσεων μοντέλων για εκθετική εξομάλυνση	234
8.6	Εκτίμηση και επιλογή μοντέλου	238
8.7	Πρόβλεψη με μοντέλα ETS	241
8.8	Ασκήσεις	243
8.9	Περαιτέρω μελέτη	247
9	Μοντέλα ARIMA	249
9.1	Στασιμότητα και διαφοροποίηση	249
9.2	Τελεστής υστέρησης	258
9.3	Αυτοπαλίνδρομα μοντέλα	259
9.4	Μοντέλα κινητών μέσων	260
9.5	Μη εποχιακά μοντέλα ARIMA	262
9.6	Εκτίμηση και επιλογή τάξης	268
9.7	Μοντελοποίηση ARIMA στο fable	270
9.8	Πρόβλεψη	278
9.9	Εποχιακά μοντέλα ARIMA	281
9.10	Σύγκριση ARIMA και ETS	291
9.11	Ασκήσεις	297
9.12	Περαιτέρω μελέτη	302



10 Μοντέλα Δυναμικής Παλινδρόμησης	303
10.1 Εκτίμηση	304
10.2 Παλινδρόμηση με σφάλματα ARIMA χρησιμοποιώντας το fable	305
10.3 Πρόβλεψη	308
10.4 Στοχαστικές και αιτιοκρατικές τάσεις	313
10.5 Δυναμική αρμονική παλινδρόμηση	316
10.6 Παράγοντες πρόβλεψης υστέρησης	319
10.7 Ασκήσεις	322
10.8 Περαιτέρω μελέτη	325
 11 Πρόβλεψη ιεραρχικών και ομαδοποιημένων χρονοσειρών	 327
11.1 Ιεραρχικές και ομαδοποιημένες χρονοσειρές	328
11.2 Προσεγγίσεις ενός επιπέδου	338
11.3 Συμβιβαστική πρόβλεψη (Forecast reconciliation)	344
11.4 Πρόβλεψη του εγχώριου τουρισμού της Αυστραλίας	349
11.5 Συμβιβαστικές κατανομημένες προβλέψεις	353
11.6 Πρόβλεψη πληθυσμού των φυλακών της Αυστραλίας	353
11.7 Ασκήσεις	358
11.8 Περαιτέρω μελέτη	359



12 Προχωρημένες Μέθοδοι Πρόβλεψης	361
12.1 Σύνθετη εποχικότητα	361
12.2 Το μοντέλο Prophet	369
12.3 Διανυσματικές αυτοπαλινδρομήσεις	373
12.4 Μοντέλα νευρωνικών δικτύων	378
12.5 Δειγματοθέτηση και ενθυλάκωση	383
12.6 Ασκήσεις	387
12.7 Περαιτέρω μελέτη	388
13 Μερικά πρακτικά ζητήματα στην πρόβλεψη	389
13.1 Εβδομαδιαία, ημερήσια και υπο-ημερήσια δεδομένα	389
13.2 Χρονοσειρές με μετρήσεις	392
13.3 Διασφάλιση ότι οι προβλέψεις παραμένουν εντός ορίων	395
13.4 Συνδυαστική πρόβλεψη	397
13.5 Διαστήματα εμπιστοσύνης για συναθροίσεις	401
13.6 Προς τα πίσω πρόβλεψη	403
13.7 Πολύ μεγάλες και πολύ μικρές χρονοσειρές	404
13.8 Πρόβλεψη σε σύνολα εκπαίδευσης και αξιολόγησης	406
13.9 Αντιμετώπιση ακραίων και μη διαθέσιμων τιμών	408
13.10 Περαιτέρω Μελέτη	413
Παράρτημα: Χρησιμοποιώντας την R	415
Βιβλιογραφία	417



Δομή των Κεφαλαίων

Δομή κεφαλαίων βιβλίου

Θεωρητική προσέγγιση θέματος

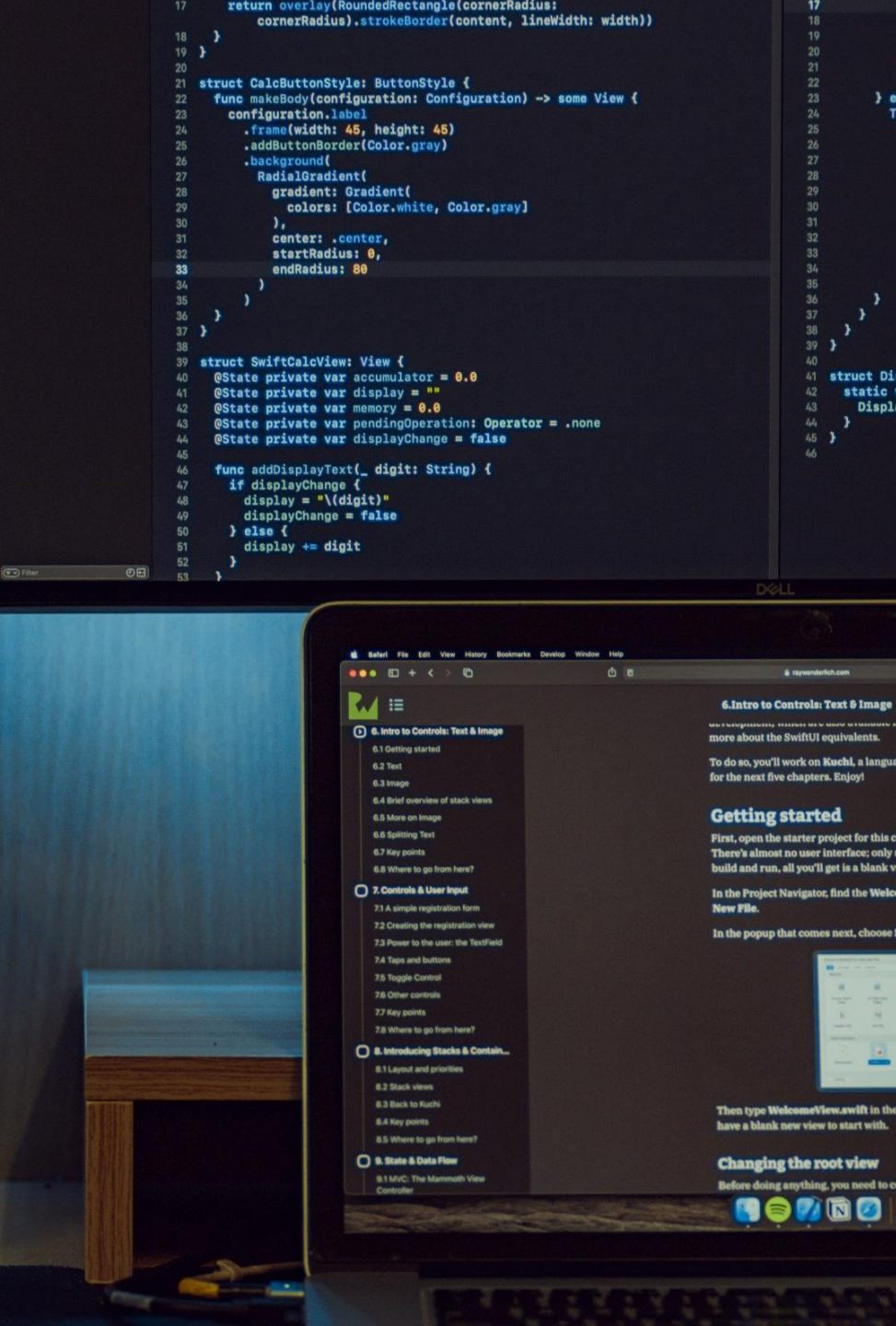
Πρακτική υλοποίηση μεθόδων στην γλώσσα R

Παρουσίαση παραδειγμάτων - γραφημάτων

Ασκήσεις κατανόησης

Περαιτέρω μελέτη

Βιβλιογραφία

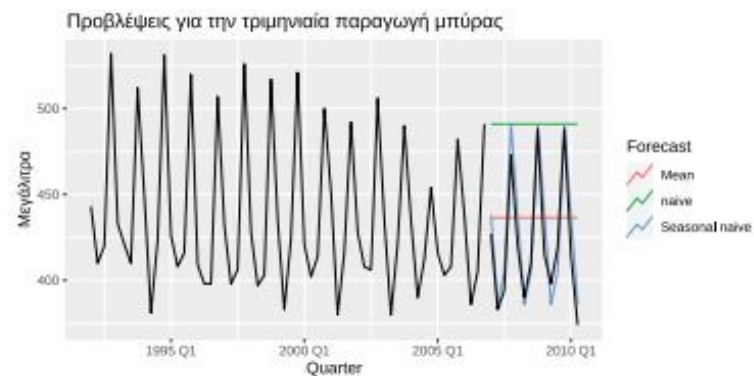


Παρουσίαση παραδειγμάτων - γραφημάτων

Παράδειγμα: Η τριμηνιαία παραγωγή μπύρας στην Αυστραλία

Το Σχήμα 5.7 δείχνει τις τρεις πρώτες μεθόδους όπως αυτές εφαρμόζονται στα δεδομένα της τριμηνιαίας παραγωγής μπύρας στην Αυστραλία από το 1992 έως το 2006, με τις προβλέψεις να συγκρίνονται με τις πραγματικές τιμές των επόμενων 3.5 χρόνων.

```
# Ορισμός δεδομένων εκπαίδευσης από το 1992 έως το 2006
train <- aus_production %>%
  filter_index("1992 Q1" ~ "2006 Q4")
# Προσαρμόζουμε τα μοντέλα
beer_fit <- train %>%
  model(
    Mean = MEAN(Beer),
    `naïve` = NAIVE(Beer),
    `Seasonal naïve` = SNAIVE(Beer)
  )
# Δημιουργία προβλέψεων για τα επόμενα 14 τρίμηνα
beer_fc <- beer_fit %>% forecast(h = 14)
# Σχεδίαση προβλέψεων μαζί με τις πραγματικές τιμές
beer_fc %>%
  autoplot(train, level = NULL) +
  autolayer(filter_index(aus_production, "2007 Q1" ~ .),
    color = "black") +
  labs(y = "Μεγάλιτρα",
    title = "Προβλέψεις για τριμηνιαία παραγωγή μπύρας" ) +
  guides(colour = guide_legend(title = "Forecast"))
```



Σχήμα 5.7: Προβλέψεις της τριμηνιαίας παραγωγής μπύρας στην Αυστραλία.



Ασκήσεις κατανόησης

5.11 Ασκήσεις

- Υπολογίστε τις προβλέψεις για τις επόμενες χρονοσειρές χρησιμοποιώντας οποιαδήποτε από τις μεθόδους $\text{NAIVE}(y)$, $\text{SNAIVE}(y)$ ή $\text{RW}(y \sim \text{drift}())$ είναι η καταλληλότερη σε κάθε περίπτωση:
 - Πληθυσμός Αυστραλίας (`global_economy`)
 - Τούβλα (`aus_production`)
 - Αμνοί (New South Wales) (`aus_livestock`)
 - Συνθήκες Διαβίωσης Νοικοκυριών (`hh_budget`).
 - Κύκλος εργασιών στο χώρο των τροφίμων της Αυστραλίας (`aus_retail`).
- Χρησιμοποιείτε τις τιμές της μετοχής της Facebook (data set `gafa_stock`) για να απαντήσετε στα ακόλουθα:
 - Σχεδιάστε την χρονοσειρά των δεδομένων.
 - Υπολογίστε τις προβλέψεις της τιμής χρησιμοποιώντας την μέθοδο περιπλάνησης και σχεδιάστε τα αποτελέσματα.
 - Δείξτε ότι οι προβλέψεις είναι ίδιες με αυτές που προκύπτουν αν επεκτείνετε την γραμμή μεταξύ της πρώτης και της τελευταίας παρατήρησης.
 - Προσπαθήστε να δοκιμάσετε μερικές από τις υπόλοιπες συναρτήσεις αναφοράς για να υπολογίσετε τις προβλέψεις στο ίδιο σύνολο δεδομένων. Ποια από αυτές πιστεύετε ότι είναι η καλύτερη; Για ποιον λόγο;

5.12 Περαιτέρω μελέτη

- Οι Ord, Fildes και Kourentzes (2017) παρέχουν περαιτέρω ανάλυση σε απλές μεθόδους πρόβλεψης αναφοράς.
- Μια ανασκόπηση των μεθόδων αξιολόγησης προβλέψεων δίνεται στην εργασία των Hyndman και Koehler (2006), όπου εξετάζονται τα πλεονεκτήματα και οι αδυναμίες των διαφορετικών προσεγγίσεων. Αυτή είναι η εργασία που εισήγαγε το MASE ως μέτρο ακρίβειας πρόβλεψης γενικού σκοπού.
- Για μια ανάλυση σχετικά με τις προβλέψεις που προκύπτουν με χρήση της STL, δείτε την εργασία του Theodosiou (2011).
- Μια εξαιρετική ανάλυση για την αξιολόγηση της ακρίβειας των προβλέψεων κατανομής παρέχεται από τους Gneiting και Katzfuss (2014).



Βιβλιογραφία

Οι αριθμοί στο δεξί περιθώριο υποδηλώνουν τις σελίδες αναφοράς κάθε βιβλιογραφικής πηγής.

- Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, **27**(1), 17–21. 38
- Armstrong, J. S. (1978). *Long-range forecasting: from crystal ball to computer*. John Wiley & Sons. 127
- Armstrong, J. S. (Επιμ.). (2001). *Principles of forecasting: a handbook for researchers and practitioners*. Kluwer Academic Publishers. 16
- Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting*, **25**, 146–166. 342, 359
- Athanasopoulos, G., & Hyndman, R. J. (2008). Modelling and forecasting Australian domestic tourism. *Tourism Management*, **29**(1), 19–31. 159
- Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & Petropoulos, F. (2017). Forecasting with temporal hierarchies. *European Journal of Operational Research*, **262**(1), 60–74. 359
- Athanasopoulos, G., Poskitt, D. S., & Vahid, F. (2012). Two canonical VARMA forms: scalar component models vis-à-vis the echelon form. *Econometric Reviews*, **31**(1), 60–83. 374



Διαθεσιμότητα του βιβλίου

Διαθεσιμότητα

Εκδόσεις Πανεπιστημίου Πατρών

<https://publications.upatras.gr>



ΕΚΔΟΣΕΙΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ
ΠΑΤΡΩΝ



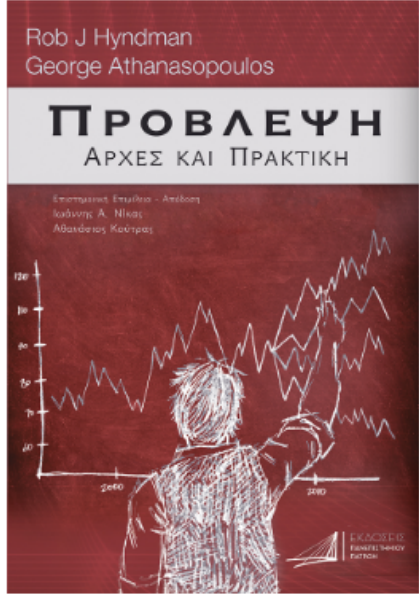
Ταυτότητα

Κατηγορίες ▾

Συγγραφείς

Νέα





Rob J Hyndman
George Athanasopoulos

ΠΡΟΒΛΕΨΗ
ΑΡΧΕΣ ΚΑΙ ΠΡΑΚΤΙΚΗ

Επιστημονική Επιμέλεια - Απόδοση
Ιωάννης Α. Νίκας
Αθανάσιος Κούτρας



George Athanasopoulos - Rob J Hyndman

Πρόβλεψη: Αρχές και Πρακτική

Μετάφραση από την τρίτη Αγγλική έκδοση
(1η Έκδοση)

Μετάφραση: Ιωάννης Α. Νίκας & Αθανάσιος Κούτρας

Αναλυτική Περιγραφή

Περίληψη περιεχομένου

Η πρόβλεψη είναι απαραίτητη σε πολλές περιπτώσεις. Η απόφαση κατασκευής μιας μονάδας παραγωγής ενέργειας μέσα στα επόμενα πέντε ...

Περισσότερα

Κωδικός: EPP-BK0113

Διαθεσιμότητα: Άμεσα διαθέσιμο

Τιμή: 28,50 €

Στην τιμή συμπεριλαμβάνεται ο Φ.Π.Α.

Χρόνος Παράδοσης: 3 έως 5 εργάσιμες μέρες

Ποσότητα

Προσθήκη στο καλάθι

Διαθεσιμότητα

ΕΥΔΟΞΟΣ

ΗΛΕΚΤΡΟΝΙΚΗ ΥΠΗΡΕΣΙΑ ΟΛΟΚΛΗΡΩΜΕΝΗΣ
ΔΙΑΧΕΙΡΙΣΗΣ ΣΥΓΓΡΑΜΜΑΤΩΝ

Κωδικός Βιβλίου: 112705742

Πρόβλεψη: Αρχές και Πρακτική

Κωδικός Βιβλίου στον Εύδοξο: 112705742

Συγγραφείς: Rob J Hyndman και George Athanasopoulos

Αριθμός Έκδοσης 1η (3η Αγγλική)

Έτος Τρέχ. Έκδοσης 2022

Λέξεις κλειδιά

Θεματικές Ενότητες

ISBN 9789605301835

Εκδόσεις ΕΤΑΙΡΕΙΑ ΑΞΙΟΠΟΙΗΣΗΣ-ΔΙΑΧΕΙΡΙΣΗΣ ΠΕΡΙΟΥΣΙΑΣ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΠΑΤΡΩΝ (ΕΚΔΟΣΕΙΣ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΠΑΤΡΩΝ)

Δέσιμο Μαλακό Εξώφυλλο

Διαστάσεις 17x24

Αριθμός Σελίδων 460

Διαθέτης (Εκδότης) Εταιρεία Αξιοποίησης και Διαχείρισης Περιουσίας Πανεπιστημίου Πατρών

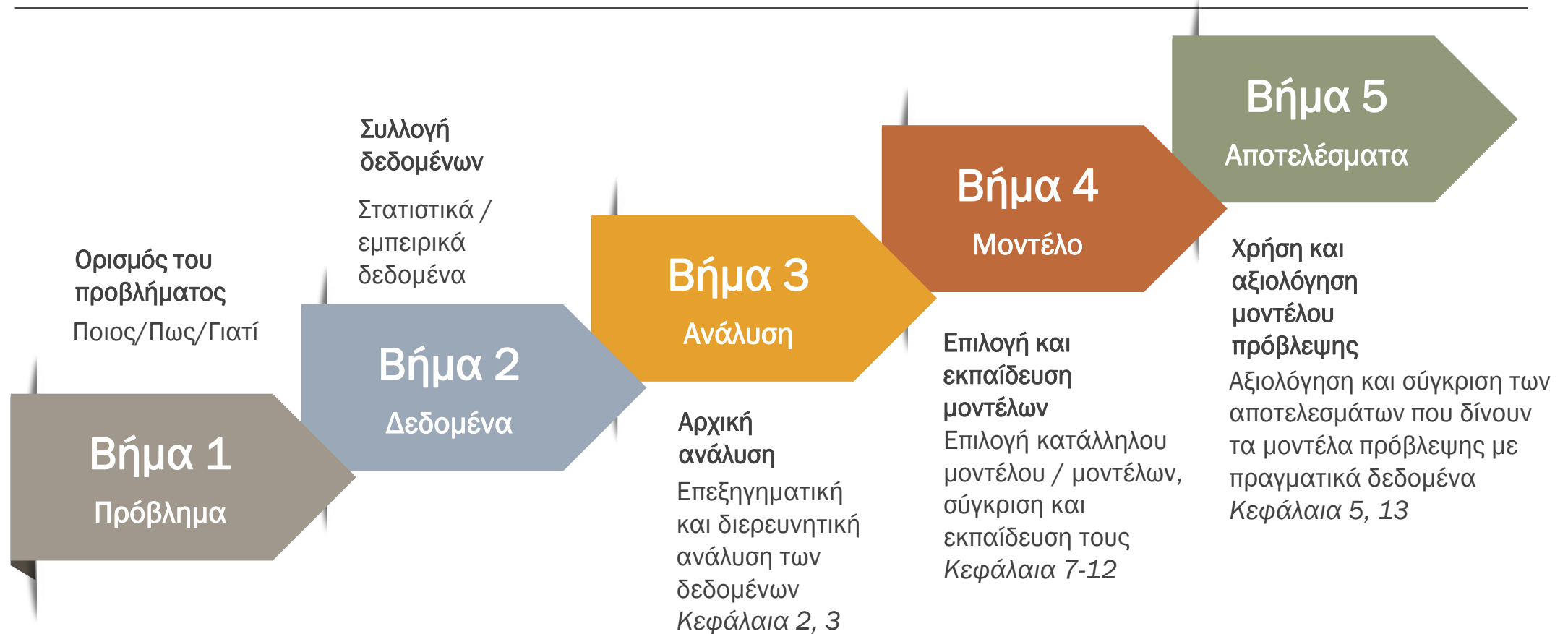
Τύπος Σύγγραμμα

Ιστοσελίδα Βιβλίου [Εδώ](#)



Ποή εργασίας

Βασικά βήματα πρόβλεψης



Απεικόνιση Χρονοσειρών

Τουρισμός

Κεφάλαιο 2ο __

Σειρά: ansett

Περιγραφή: ΑΕΡΟΠΟΡΙΚΗ
ΕΠΙΒΑΤΙΚΗ
ΚΙΝΗΣΗ

2.2 Χρονοδιαγράμματα

Για να απεικονίσουμε δεδομένα χρονοσειρών, το προφανές γράφημα με το οποίο μπορούμε να ξεκινήσουμε είναι ένα χρονοδιάγραμμα. Δηλαδή, οι παρατηρήσεις σχεδιάζονται ως συνάρτηση του χρόνου στον οποίο εμφανίζονται, και οι διαδοχικές παρατηρήσεις που ενώνονται μεταξύ τους με ευθείες γραμμές. Το ακόλουθο Σχήμα 2.1 δείχνει το πλήθος των επιβατών που ταξιδεύουν ανά εβδομάδα μεταξύ των δύο μεγαλύτερων πόλεων της Αυστραλίας με την αεροπορική εταιρεία Ansett Airlines.

Θα χρησιμοποιούμε την εντολή `autoplot()` αρκετά συχνά. Η εντολή αυτή δημιουργεί αυτόματα μια κατάλληλη γραφική παράσταση για οτιδήποτε δοθεί στο πρώτο της όρισμα. Σε αυτή την περίπτωση η εντολή αναγνωρίζει το `melsyd_economy` ως μια χρονοσειρά και δημιουργεί το αντίστοιχο χρονοδιάγραμμα.

```
melsyd_economy <- ansett %>%  
  filter(Airports == "MEL-SYD", Class == "Economy")  
autoplot(melsyd_economy, Passengers/1000) +  
  labs(title = "Επιβάτες οικονομικής θέσης της Ansett",  
        subtitle = "Melbourne-Sydney", y = "Επιβάτες (σε χιλιάδες)")
```

Απεικόνιση Χρονοσειρών

Τουρισμός

Κεφάλαιο 2ο _____

Σειρά: ansett

Περιγραφή: ΑΕΡΟΠΟΡΙΚΗ
ΕΠΙΒΑΤΙΚΗ
ΚΙΝΗΣΗ

Η γραφική παράσταση της χρονοσειράς μας, παρουσιάζει άμεσα μερικά πολύ ενδιαφέροντα χαρακτηριστικά:

- Υπήρξε μια περίοδος το 1989 όπου δεν μεταφέρθηκαν καθόλου επιβάτες – αυτό οφείλεται σε μια βιομηχανική διαμάχη που έλαβε χώρα την περίοδο αυτή.
- Υπήρξε μια περίοδος μειωμένου αριθμού επιβατών το 1992. Αυτό οφείλονταν σε μια δοκιμή κατά την οποία ορισμένες οικονομικές θέσεις αντικαταστάθηκαν από επιχειρηματικές.
- Μια μεγάλη αύξηση του αριθμού των επιβατών σημειώθηκε κατά το δεύτερο εξάμηνο του 1991.
- Υπάρχουν μερικές μεγάλες μειώσεις του αριθμού των επιβατών στην αρχή κάθε έτους. Αυτές οφείλονται στο φαινόμενο των διακοπών (καλοκαίρι στην Αυστραλία).
- Υπάρχει μια μακροπρόθεσμη διακύμανση των τιμών της χρονοσειράς, οι οποίες αυξάνονται κατά τη διάρκεια του 1987, μειώνονται το 1989 και ξανά αυξάνονται το 1990 και το 1991.

Οποιοδήποτε μοντέλο και να χρησιμοποιήσουμε, θα πρέπει να λάβει υπόψη όλα αυτά τα χαρακτηριστικά προκειμένου να προβλέψει αποτελεσματικά το μελλοντικό αριθμό των επιβατών.

Ανάλυση Χρονοσειρών

Απασχόληση

Κεφάλαιο 3ο __

Σειρά: us_employment

Περιγραφή: ΑΠΑΣΧΟΛΗΣΗ ΣΤΙΣ ΗΠΑ

Παράδειγμα: Απασχόληση στον τομέα του λιανικού εμπορίου των ΗΠΑ

Θα εξετάσουμε αρκετές μεθόδους για τη λήψη των συνιστωσών S_t , T_t και R_t αργότερα σε αυτό το κεφάλαιο, ωστόσο, είναι χρήσιμο πρώτα να δούμε ένα παράδειγμα. Θα αναλύσουμε τον αριθμό των ατόμων που απασχολούνται στο λιανικό εμπόριο, όπως φαίνεται στο Σχήμα 3.5. Τα δεδομένα δείχνουν το συνολικό μηνιαίο αριθμό ατόμων σε χιλιάδες που απασχολούνται στον τομέα του λιανικού εμπορίου των ΗΠΑ από το 1990.

```
us_retail_employment <- us_employment %>%  
  filter(year(Month) >= 1990,  
         Title == "Retail Trade") %>%  
  select(-Series_ID)  
  
autoplot(us_retail_employment, Employed) +  
  labs(y = "Απασχολούμενοι (χιλιάδες)",  
       title = "Συνολική απασχόληση στο λιανικό εμπόριο των ΗΠΑ")
```

Ανάλυση Χρονοσειρών

Απασχόληση

Κεφάλαιο 3ο _____

Σειρά: us_employment

Περιγραφή: ΑΠΑΣΧΟΛΗΣΗ ΣΤΙΣ ΗΠΑ

Για να επεξηγήσουμε τις ιδέες, θα χρησιμοποιήσουμε τη μέθοδο ανάλυσης STL, η οποία συζητείται στην Ενότητα 3.6.

```
dcmp <-  
  us_retail_employment %>%  
  model(stl = STL(Employed))  
  
components(dcmp)  
#> # A dable: 357 x 7 [1M]  
#> # Key:      .model [1]  
#> # :      Employed = trend + season_year + remainder  
#>      .model      Month Employed trend season_year remainder  
#>      <chr>      <mth>      <dbl> <dbl>      <dbl>      <dbl>  
#> 1 stl      1990 Jan      13256. 13288.      -33.0      0.836  
#> 2 stl      1990 Feb      12966. 13269.      -258.      -44.6  
#> 3 stl      1990 Mar      12938. 13250.      -290.      -22.1  
#> 4 stl      1990 Apr      13012. 13231.      -220.       1.05  
#> 5 stl      1990 May      13108. 13211.      -114.      11.3  
#> 6 stl      1990 Jun      13183. 13192.      -24.3      15.5  
#> 7 stl      1990 Jul      13170. 13172.      -23.2      21.6  
#> 8 stl      1990 Aug      13160. 13151.       -9.52     17.8  
#> 9 stl      1990 Sep      13113. 13131.      -39.5      22.0  
#> 10 stl     1990 Oct      13185. 13110.       61.6     13.2  
#> # ... with 347 more rows, and 1 more variable:  
#> #      season_adjust <dbl>
```

```
components(dcmp) %>% autoplot()
```

Πρόβλεψη με Ανάλυση Χρονοσειρών

Απασχόληση

Κεφάλαιο 5ο __

Σειρά: us_employment

Περιγραφή: ΑΠΑΣΧΟΛΗΣΗ ΣΤΙΣ ΗΠΑ

Παράδειγμα: Απασχόληση στον τομέα λιανικού εμπορίου των ΗΠΑ

Το Σχήμα 5.18 δείχνει naïve προβλέψεις για τα εποχικά προσαρμοσμένα δεδομένα της απασχόλησης στον τομέα λιανικής των ΗΠΑ. Αυτές στη συνέχεια «επαναπροσδιορίζουν την εποχικότητα» προσθέτοντας τις εποχιακές naïve προβλέψεις της εποχιακής συνιστώσας.

```
us_retail_employment <- us_employment %>%  
  filter(year(Month) >= 1990, Title == "Retail Trade")  
dcmp <- us_retail_employment %>%  
  model(STL(Employed ~ trend(window = 7), robust = TRUE)) %>%  
  components() %>% select(-.model)  
dcmp %>%  
  model(NAIVE(season_adjust)) %>%  
  forecast() %>% autoplot(dcmp) +  
  labs(y = "Αριθμός εργαζομένων",  
       title = "Απασχόληση στο λιανικό εμπόριο των ΗΠΑ")
```

Πρόβλεψη με Ανάλυση Χρονοσειρών

Απασχόληση

Κεφάλαιο 5ο _____

Σειρά: us_employment

Περιγραφή: ΑΠΑΣΧΟΛΗΣΗ ΣΤΙΣ ΗΠΑ

καθεμιά από τις αναλυόμενες συνιστώσες. Οι εποχιακές συνιστώσες του μοντέλου θα προβλεφθούν αυτόματα χρησιμοποιώντας τη συνάρτηση `SNAIVE()`, εάν δεν έχει καθοριστεί διαφορετικό μοντέλο. Η συνάρτηση θα κάνει, επίσης, τον επαναπροσδιορισμό της εποχικότητας, διασφαλίζοντας ότι λαμβάνονται οι προκύπτουσες προβλέψεις των αρχικών δεδομένων. Αυτά φαίνονται στο Σχήμα 5.19.

```
fit_dcmp <- us_retail_employment %>%  
  model(stlf = decomposition_model(  
    STL(Employed ~ trend(window = 7), robust = TRUE),  
    NAIVE(season_adjust)  
  ))  
fit_dcmp %>%  
  forecast() %>%  
  autoplot(us_retail_employment) +  
  labs(y = "Αριθμός εργαζομένων",  
       title = "Απασχόληση στο λιανικό εμπόριο των ΗΠΑ")
```


Αξιολόγηση Ακρίβειας Σημειακών Προβλέψεων

Παραγωγή μπίρας

Κεφάλαιο 5ο _____

Σειρά: aus_production

Περιγραφή: ΔΕΙΚΤΕΣ
ΠΑΡΑΓΩΓΗΣ ΣΤΗΝ
ΑΥΣΤΡΑΛΙΑ

Παραδείγματα

```
recent_production <- aus_production %>% filter(year(Quarter) >= 1992)
beer_train <- recent_production %>% filter(year(Quarter) <= 2007)
```

```
beer_fit <- beer_train %>%
  model(Mean = MEAN(Beer),
        `Naïve` = NAIVE(Beer),
        `Seasonal naïve` = SNAIVE(Beer),
```

```
        Drift = RW(Beer ~ drift()) )
```

```
beer_fc <- beer_fit %>% forecast(h = 10)
beer_fc %>%
  autoplot(aus_production %>% filter(year(Quarter) >= 1992),
           level = NULL) +
  labs(y = "Μεγάλιτρα",
       title = "Προβλέψεις τριμηνιαίας παραγωγής μπίρας") +
  guides(colour = guide_legend(title = "Πρόβλεψη"))
```

Αξιολόγηση Ακρίβειας Σημειακών Προβλέψεων

Παραγωγή μπύρας

Κεφάλαιο 5ο _____

Σειρά: **aus_production**

Περιγραφή: **ΔΕΙΚΤΕΣ
ΠΑΡΑΓΩΓΗΣ ΣΤΗΝ
ΑΥΣΤΡΑΛΙΑ**

Το Σχήμα 5.21 παρουσιάζει τέσσερις μεθόδους πρόβλεψης που εφαρμόζονται στην τριμηνιαία παραγωγή μπύρας στην Αυστραλία χρησιμοποιώντας δεδομένα μόνο έως το τέλος του 2007. Παρουσιάζονται, επίσης, οι πραγματικές τιμές για την περίοδο 2008-2010. Υπολογίζουμε τα μέτρα ακρίβειας των προβλέψεων για αυτήν την περίοδο.

```
accuracy(beer_fc, recent_production)
```

Μέθοδος	RMSE	MAE	MAPE	MASE
Μέθοδος Περιπλάνησης	64.90	58.88	14.58	4.12
Μέθοδος Μέσου	38.45	34.83	8.28	2.44
Μέθοδος Naïve	62.69	57.40	14.18	4.01
Μέθοδος Εποχιακή Naïve	14.31	13.40	3.17	0.94

Η συνάρτηση `accuracy()` θα εξαγάγει αυτόματα τις σχετικές περιόδους από τα δεδομένα (`recent_production` σε αυτό το παράδειγμα) για να ταιριάξει τις προβλέψεις κατά τον υπολογισμό των διαφόρων μέτρων ακρίβειας.

Σύγκριση μεθόδων πρόβλεψης

Πληθυσμός Αυστραλίας

Κεφάλαιο 9ο __

Σειρά: `global_economy`

Περιγραφή: ΠΑΓΚΟΣΜΙΑ
ΟΙΚΟΝΟΜΙΑ

Σύγκριση `ARIMA()` και `ETS()` σε μη εποχιακά δεδομένα

Μπορούμε να χρησιμοποιήσουμε διασταυρούμενη επικύρωση χρονοσειρών για να συγκρίνουμε ένα μοντέλο `ARIMA` και ένα μοντέλο `ETS`. Ας εξετάσουμε τον πληθυσμό της Αυστραλίας από το σύνολο δεδομένων `global_economy`, όπως παρουσιάστηκε στην Ενότητα 8.2.

```
aus_economy <- global_economy %>%  
  filter(Code == "AUS") %>%  
  mutate(Population = Population/1e6)
```

```
aus_economy %>%  
  slice(-n()) %>%  
  stretch_tsibble(.init = 10) %>%  
  model(  
    ETS(Population),  
    ARIMA(Population)  
  ) %>%  
  forecast(h = 1) %>%  
  accuracy(aus_economy) %>%  
  select(.model, RMSE:MAPE)
```

```
aus_economy %>%  
  model(ETS(Population)) %>%  
  forecast(h = "5 years") %>%  
  autoplot(aus_economy %>% filter(Year >= 2000)) +  
  labs(title = "Πληθυσμός Αυστραλίας", y = "Άνθρωποι (σε εκατομμύρια)")
```

Σύγκριση μεθόδων πρόβλεψης

Παραγωγή Τσιμέντου

Κεφάλαιο 9ο _____

Σειρά: aus_production

Περιγραφή: ΔΕΙΚΤΕΣ
ΠΑΡΑΓΩΓΗΣ ΣΤΗΝ
ΑΥΣΤΡΑΛΙΑ

Σύγκριση ARIMA() και ETS() σε εποχιακά δεδομένα

Σε αυτήν την περίπτωση θέλουμε να συγκρίνουμε τα εποχιακά μοντέλα ARIMA και ETS που εφαρμόζονται στα τριμηνιαία δεδομένα παραγωγής τσιμέντου (από το aus_production). Επειδή η σειρά είναι σχετικά μεγάλη, μπορούμε να χρησιμοποιήσουμε ένα εκπαιδευτικό σύνολο και ένα σύνολο αξιολόγησης αντί για διασταυρούμενη επικύρωση χρονοσειρών. Το πλεονέκτημα είναι ότι αυτό είναι πολύ πιο γρήγορο. Δημιουργούμε ένα σύνολο εκπαίδευσης από τις αρχές του 1988 έως το τέλος του 2007 και επιλέγουμε ένα ARIMA και ένα ETS μοντέλο χρησιμοποιώντας τις συναρτήσεις ARIMA() και ETS().

```
cement <- aus_production %>%  
  select(Cement) %>% filter_index("1988 Q1" ~ .)  
train <- cement %>% filter_index(. ~ "2007 Q4")
```

Το αποτέλεσμα παρακάτω δείχνει το μοντέλο ARIMA που έχει επιλεγεί και εκτιμηθεί από τη συνάρτηση ARIMA(). Το μοντέλο ARIMA τα καταφέρνει καλά στην καταγραφή όλων των δυναμικών στα δεδομένα, καθώς τα υπόλοιπα φαίνεται να είναι λευκός θόρυβος.

```
fit_arma <- train %>% model(ARIMA(Cement))  
report(fit_arma)
```

```
augment(fit_arma) %>%  
  features(.innov, ljung_box, lag = 16, dof = 6)
```


Σύγκριση μεθόδων πρόβλεψης

Παραγωγή Τσιμέντου

Κεφάλαιο 9ο _____

Σειρά: aus_production

Περιγραφή: ΔΕΙΚΤΕΣ
ΠΑΡΑΓΩΓΗΣ ΣΤΗΝ
ΑΥΣΤΡΑΛΙΑ

Το ακόλουθο αποτέλεσμα δείχνει, επίσης, το μοντέλο ETS που έχει επιλεγεί και εκτιμηθεί από τη συνάρτηση ETS(). Αυτό το μοντέλο τα καταφέρνει, επίσης, καλά στην καταγραφή όλων των δυναμικών στα δεδομένα, καθώς τα υπόλοιπα μοιάζουν να είναι λευκός θόρυβος.

```
fit_ets <- train %>% model(ETS(Cement))  
report(fit_ets)
```

```
fit_ets %>%  
  gg_tsresiduals(lag_max = 16)
```

```
augment(fit_ets) %>%  
  features(.innov, ljung_box, lag = 16, dof = 6)
```

Σύγκριση μεθόδων πρόβλεψης

Παραγωγή Τσιμέντου

Κεφάλαιο 9ο _____

Σειρά: aus_production

Περιγραφή: ΔΕΙΚΤΕΣ
ΠΑΡΑΓΩΓΗΣ ΣΤΗΝ
ΑΥΣΤΡΑΛΙΑ

```
# Δημιουργία προβλέψεων και σύγκριση ακρίβειας στο σύνολο αξιολόγησης
bind_rows(
  fit_arima %>% accuracy(),
  fit_ets %>% accuracy(),
  fit_arima %>% forecast(h = 10) %>% accuracy(cement),
  fit_ets %>% forecast(h = 10) %>% accuracy(cement)
) %>%
  select(-ME, -MPE, -ACF1)
```

Παρακάτω δημιουργούμε και σχεδιάζουμε προβλέψεις από το μοντέλο ARIMA για τα επόμενα 3 χρόνια.

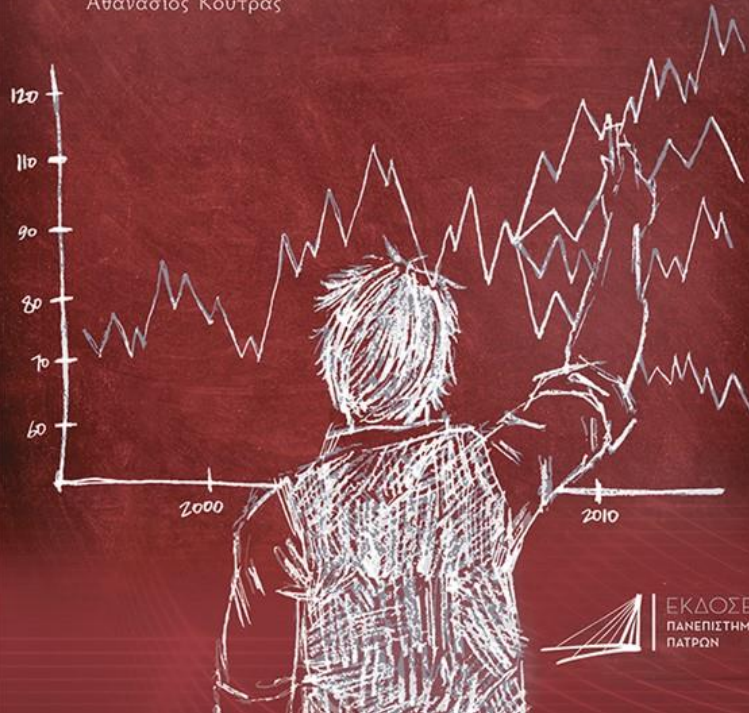
```
cement %>%
  model(ARIMA(Cement)) %>%
  forecast(h="3 years") %>%
  autoplot(cement) +
  labs(title = "Παραγωγή τσιμέντου στην Αυστραλία",
       y = "Τόνοι (σε χιλιάδες)")
```

Rob J Hyndman
George Athanasopoulos

ΠΡΟΒΛΕΨΗ

ΑΡΧΕΣ ΚΑΙ ΠΡΑΚΤΙΚΗ

Επιστημονική Επιμέλεια - Απόδοση
Ιωάννης Α. Νίκας
Αθανάσιος Κούτρας



ΕΚΔΟΣΕΙΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ
ΠΑΤΡΩΝ

Πρόβλεψη: Αρχές και Πρακτική

Σας ευχαριστούμε
για την προσοχή σας!