

Στατιστική II

Γιώργος Τσιρογιάννης

Τμήμα Διοίκησης Επιχειρήσεων Αγροτικών
Προϊόντων και Τροφίμων,
Πανεπιστήμιο Πατρών

Βασικά προτεινόμενα συγγράμματα (ή μπορείτε να επιλέξετε οποιοδήποτε άλλο από την λίστα στον Εύδοξο)



Βασικά προτεινόμενα συγγράμματα



- Εξαιρετικά αναλυτική και πλήρης παρουσίαση της θεωρίας
- Πολύ μεγάλη συνάφεια των παραδειγμάτων με τον αγροτικό τομέα
- Λογική συνέχεια του 1^{ου} εξαμήνου με το 2^ο ως προς τους συμβολισμούς
- Το έχετε ήδη στην κατοχή σας οι περισσότεροι από το προηγούμενο εξάμηνο
- Πλήρες handbook που θα σας συνοδεύει στην μετέπειτα πορείας (και μετά την αποφοίτηση)

Βασικά προτεινόμενα συγγράμματα



- Πολύ γνωστό σύγγραμμα που τώρα είναι διαθέσιμο και στα ελληνικά
- Εμπλουτισμένο από τους Έλληνες συν-συγγραφείς με παραδείγματα και scripts στην στατιστική γλώσσα R
- Πολύτιμο εργαλείο για πρακτικές εφαρμογές
- Καλύπτει κάποια επιπλέον θέματα όπως παλινδρόμηση

Τι θα δούμε στο εξάμηνο

- Αναπαράσταση σχέσεων, συσχέτιση
- Παλινδρόμηση
- Σημειακή εκτίμηση και διαστήματα εμπιστοσύνης
- Στατιστικός έλεγχος υποθέσεων
- Ανάλυση διακύμανσης
- Έλεγχοι χ^2
-

Online υλικό

Εισαγωγικό:



<https://www.youtube.com/c/joshstarmmer/videos>

Online υλικό

Προχωρημένο:

Statistics for Applications

COURSE HOME <

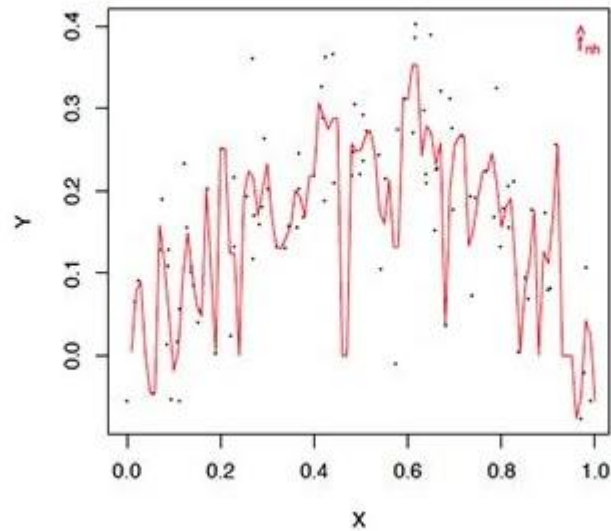
SYLLABUS

LECTURE SLIDES

LECTURE VIDEOS

ASSIGNMENTS

DOWNLOAD COURSE
MATERIALS



Nonparametric regression. (Image by Prof. Philippe Rigollet.)

Instructor(s)

Prof. Philippe Rigollet

MIT Course Number

18.650 / 18.6501

As Taught In

Fall 2016

Level

Undergraduate / Graduate

CITE THIS COURSE

<https://ocw.mit.edu/courses/mathematics/18-650-statistics-for-applications-fall-2016/lecture-videos/>

Πρακτικές εφαρμογές με την χρήση λογισμικού ανοικτού κώδικα

coursera

Explore ▾

What do you want to learn?



[Browse](#) > [Data Science](#) > [Probability and Statistics](#)

Statistics with Python Specialization

Practical and Modern Statistical Thinking For All. Use Python for statistical visualization, inference, and modeling

★★★★★ 4.6 2,974 ratings



Brenda Gunderson [+2 more instructors](#)

Enroll for Free

Starts Feb 22

Financial aid available

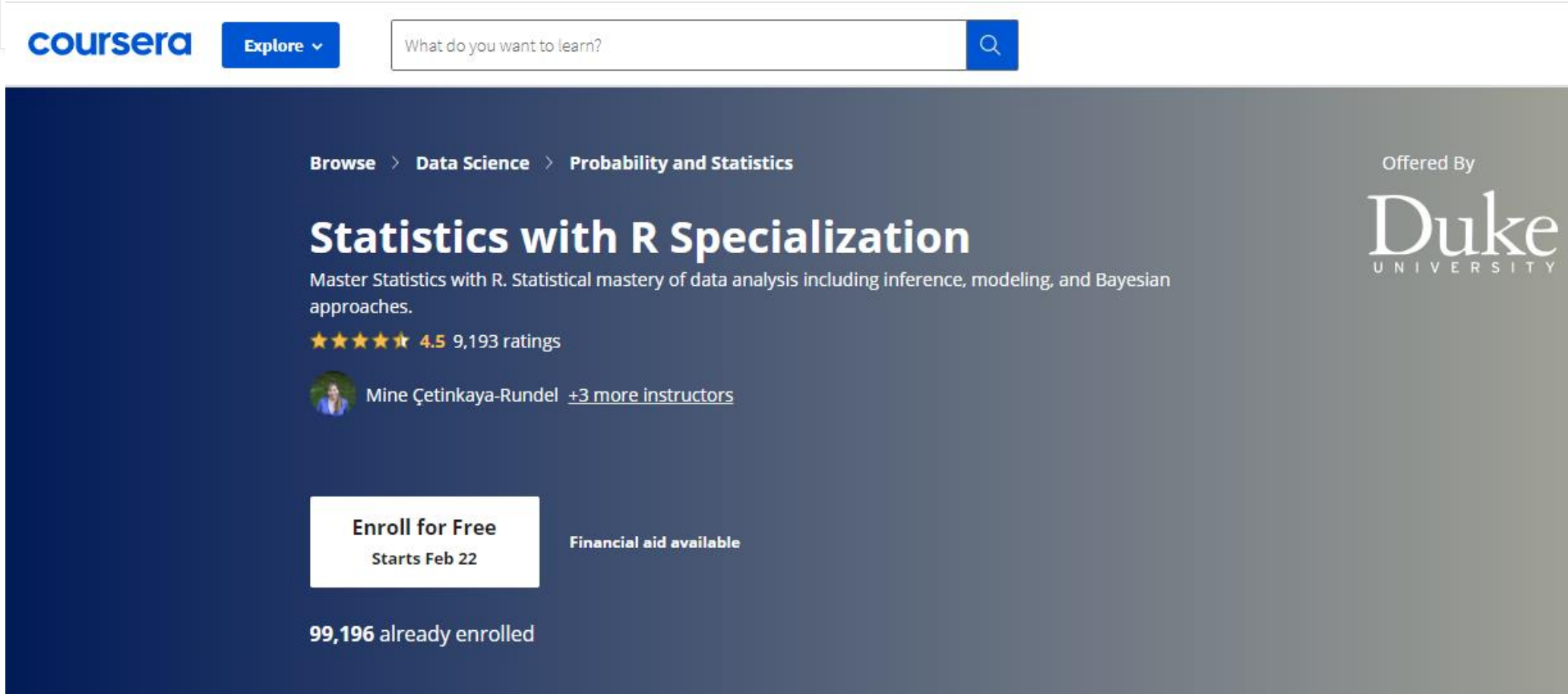
44,141 already enrolled

Offered By



<https://www.coursera.org/specializations/statistics-with-python>

Πρακτικές εφαρμογές με την χρήση λογισμικού ανοικτού κώδικα



The screenshot shows the Coursera website interface. At the top left is the Coursera logo. To its right is a blue 'Explore' button with a dropdown arrow. Further right is a search bar with the placeholder text 'What do you want to learn?' and a magnifying glass icon. Below the search bar, the breadcrumb navigation reads 'Browse > Data Science > Probability and Statistics'. The main title of the course is 'Statistics with R Specialization'. Below the title is a description: 'Master Statistics with R. Statistical mastery of data analysis including inference, modeling, and Bayesian approaches.' The course has a rating of 4.5 stars from 9,193 ratings. The instructor is Mine Çetinkaya-Rundel, with a link to '+3 more instructors'. A white button says 'Enroll for Free' with 'Starts Feb 22' below it. To the right of the button, it says 'Financial aid available'. At the bottom left, it says '99,196 already enrolled'. On the right side of the page, it says 'Offered By' followed by the Duke University logo.

<https://www.coursera.org/specializations/statistics>

Ελεύθερο στατιστικό λογισμικό που συναγωνίζεται επάξια τα εμπορικά πακέτα



[\[Home\]](#)

Download

[CRAN](#)

R Project

[About R](#)

[Logo](#)

[Contributors](#)

[What's New?](#)

[Reporting Bugs](#)

[Conferences](#)

[Search](#)

[Get Involved: Mailing Lists](#)

[Developer Pages](#)

[R Blog](#)

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News

- [R version 4.0.4 \(Lost Library Book\)](#) has been released on 2021-02-15.
- Thanks to the organisers of useR! 2020 for a successful online conference. Recorded tutorials and talks from the conference are available on the [R Consortium YouTube channel](#).
- [R version 3.6.3 \(Holding the Windsock\)](#) was released on 2020-02-29.
- You can support the R Foundation with a renewable subscription as a [supporting member](#)

News via Twitter

<https://www.r-project.org/>

Εγκατάσταση της R

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

<https://cran.biotoools.fr/>

Συνοδεύεται από ένα εξαιρετικό περιβάλλον ανάπτυξης



Products ▾ Solutions ▾ Customers Resources ▾ About ▾ Pricing

DOWNLOAD SUPPORT DOCS COMMUNITY 🔍



RStudio

Take control of your R code

RStudio is an integrated development environment (IDE) for R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. [Click here to see more RStudio features.](#)

RStudio is available in **open source** and **commercial** editions and runs on the desktop (Windows, Mac, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian/Ubuntu, Red Hat/CentOS, and SUSE Linux).

<https://rstudio.com/products/rstudio/>

Συνοδεύεται από ένα εξαιρετικό περιβάλλον ανάπτυξης

	RStudio Desktop Open Source License Free	RStudio Desktop Pro Commercial License \$995 /year	RStudio Server Open Source License Free	RStudio Server Pro Commercial License \$4,975 /year (5 Named Users)
	DOWNLOAD Learn more	BUY Learn more	DOWNLOAD Learn more	BUY Evaluation Learn more
Integrated Tools for R	✓	✓	✓	✓
Priority Support		✓		✓
Access via Web Browser			✓	✓
RStudio Professional Drivers		✓		✓
Connect to RStudio Server Pro remotely		✓		
Enterprise Security				✓
Project Sharing				✓
Manage Multiple R Sessions & Versions				✓

<https://rstudio.com/products/rstudio/download/>

RStudio

The image shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main editor window displays a script titled 'Kritiki_book_commands_Chapter_2.R' with the following R code:

```
1 library(readxl)
2 FinalData <- read_excel("C:/Users/user/Desktop/kritiki Book_diorthosei:
3 attach(FinalData)
4 View(FinalData)
5 ## Αέυάñáíá όό-ííòþòùí##
6 ## xñþόç cbind ##
7 cbind( Freq=table(score), Cumul=cumsum(table(score)), relative=prop.tal
8 ## Sturges##
9 factorx <- factor(cut(score, breaks=nclass.sturges(score)))
10 xout <- as.data.frame(table(factorx))
11 xout <- transform(xout, cumFreq = cumsum(Freq), relative = prop.table(i
12 print(xout)
13 #-----
14 library(descr)
15 freq(score, plot = FALSE)
16
17 # όόíÛñόçόç plot#.
18 plot(score)
19 plot(introyear)
20 plot(sex)
21 plot(introyear,score)
22
23
24 # όόíÛñόçόç hist#.
25 hist(introyear)
26 hist(introyear, nclass=14)
27 hist(score, breaks=c (5000,10000,15000,25000))
28
```

The console window at the bottom shows the R startup message:

```
R version 4.0.2 (2020-06-22) -- "Taking off Again"
Copyright (c) 2020 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

The Environment pane on the right shows 'Global Environment' and 'Environment is empty'. The bottom status bar includes 'Files', 'Plots', 'Packages', 'Help', and 'Viewer'.

Ελευθέρα βιβλία Στατιστικής (ξενόγλωσσα)

Introduction to Statistics

[Online Edition](#)

Primary author and editor:

David M. Lane¹

Other authors:

David Scott¹, Mikki Hebl¹, Rudy Guerra¹, Dan Osherson¹, and Heidi Zimmer²

https://onlinestatbook.com/Online_Statistics_Education.pdf

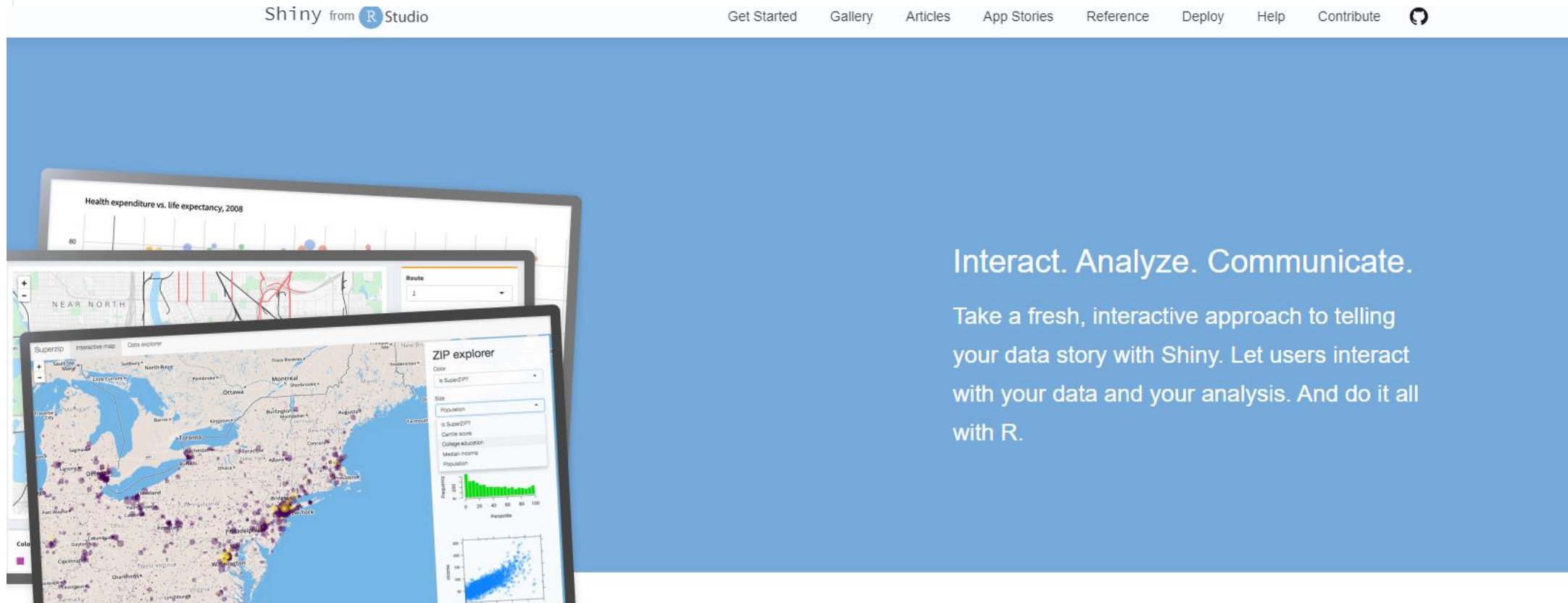
Ελευθέρη βιβλία Στατιστικής (ξενόγλωσσα)

Learning statistics with R:
A tutorial for psychology students and other beginners
(Version 0.6)

Danielle Navarro
University of New South Wales
d.navarro@unsw.edu.au

<http://compcogscisydney.org/learning-statistics-with-r>

Εργαλεία για συστήματα «από την ιδέα και τα δεδομένα στην δια-δραστική παρουσίασή»



The screenshot displays the Shiny website interface. At the top, the navigation bar includes the text "Shiny from R Studio" and a series of links: "Get Started", "Gallery", "Articles", "App Stories", "Reference", "Deploy", "Help", and "Contribute". Below the navigation bar, there are three overlapping panels showcasing interactive data visualizations. The top panel, titled "Health expenditure vs. life expectancy, 2008", shows a scatter plot with data points colored by country. The middle panel, titled "NEAR NORTH", features a map with a red route highlighted and a "Route" dropdown menu set to "2". The bottom panel, titled "ZIP explorer", shows a map of the United States with data points colored and sized by various variables. To the right of the map, there are several interactive controls: a "Color" dropdown menu set to "is SuperZIP?", a "Size" dropdown menu set to "Population", and a "ZIP explorer" section with a histogram and a scatter plot. The histogram shows the frequency distribution of ZIP codes, and the scatter plot shows the relationship between two variables.

Interact. Analyze. Communicate.

Take a fresh, interactive approach to telling your data story with Shiny. Let users interact with your data and your analysis. And do it all with R.

<https://shiny.rstudio.com/>

Εργαλεία για συστήματα «από την ιδέα και τα δεδομένα στην δια-δραστική παρουσίασή»

Explore the Fuel Economy of Cars in the U.S. Market

This interactive display features the [fuel economy data provided by the U.S. Environmental Protection Agency](#). It allows you to explore the fuel economy of cars in the U.S. market across time and other dimensions.

It is based on the 'ExPanD' display provided by the 'ExPanDaR' package. [Click here](#) to explore your own data with 'ExPanD'.

Otherwise: Scroll down and start exploring!

Subset factor

Full Sample

<https://shiny.rstudio.com/gallery/explore-panel-data.html>

Indicate whether you want to study the full sample or subset to a specific factor.

By default, this display uses all data from car makes with more than 100 cars in the 'fuelconomy.gov' database. Above, you can limit the analysis to cars from a certain make, class, country, fuel type or other factor present in the data.

Bar Chart

Select factor to display

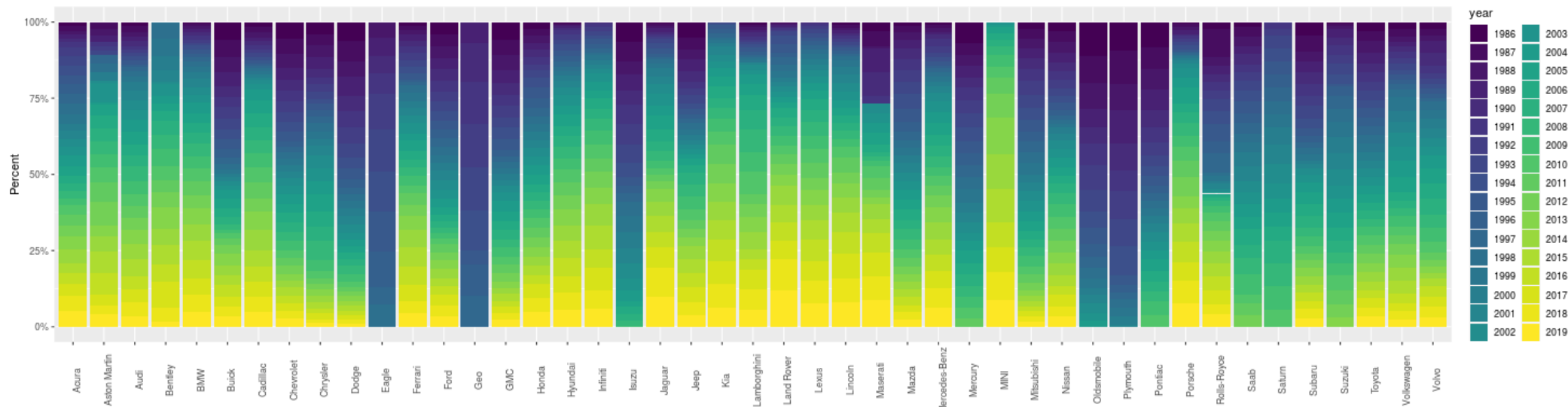
make

Select additional factor to display

year

Relative display

Check if you want to see the additional factor relative to the first factor.



Υλικό πέρα από την “κλασσική” Στατιστική

- Και στις δυο προσεγγίσεις θεωρούμε μοντέλα βάσει των οποίων δημιουργούνται τα δεδομένα
- Frequentist “κλασσική” στατιστική:
 - Μια παράμετρος (πχ μέση τιμή/διασπορά κανονικής κατανομής) είναι «σταθερή» αλλά άγνωστη σε εμάς
 - Η συμπερασματολογία πραγματοποιείται κάνοντας χρήση υποθετικών δειγματοληψιών των δεδομένων
 - Υπολογισμοί βάσει της αρχής: τα δεδομένα είναι τυχαία, αλλά οι παράμετροι σταθεροί
- Bayesian στατιστική:
 - Μια παράμετρος είναι τυχαία μεταβλητή που ακολουθεί μια κατανομή
 - Η συμπερασματολογία πραγματοποιείται λαμβάνοντας υπόψη τις παρατηρήσεις
 - Υπολογισμοί βάσει της αρχής: τα δεδομένα είναι σταθερά, αλλά οι παράμετροι τυχαίοι

Υλικό πέρα από την “κλασσική” Στατιστική

