

Kyiv School of Economics
Data analysis and visualization in R
Instructor: Maksym Obrizan

Time & Location: TBA

Office Hours: TBA. Please drop me an e-mail if I am not online.

E-mail: mobrizan@kse.org.ua or mobrizan@gmail.com

Textbooks

Rui Miguel Forte. 2017. "Mastering Predictive Analytics with R – Second Edition". *Packet publishing*.

Viswa Viswanathan, Shanthi Viswanathan, Atmajitsinh Gohil, Yu-Wei, Chiu (David Chiu). 2016. "R: Recipes for Analysis, Visualization and Machine Learning." *Packet publishing*.

Joshua N. Milligan. 2020. "Learning Tableau 2020 - Fourth Edition". *Packet publishing*.

Course description

This course is designed to get hands-on experience in working with data in business, economics, and finance. In addition to lectures, students will have numerous business cases to be solved in groups. The course will cover basic tools of data analytics and visualization with more than 10 practical cases analyzed in R and Tableau. Students will review [kaggle.com](https://www.kaggle.com) competitions, learn how to quickly find answers to programming questions at stackoverflow.com, and discover other great resources to deepen their knowledge. Prerequisites for this course are "Econometrics I" and "Econometrics II" covering descriptive statistics, linear regression, and basics of R programming.

Students are expected to participate in classes with a notebook with the most recent versions of R, RStudio and Tableau installed.

Learning Outcomes for the Course

By the end of this course, students will:

- Get familiar with business, finance and economic applications of big data
- Learn to apply R and RStudio software to data analytics problems
- Study business intelligence and data visualization tools
- Develop own big data project
- Formulate solutions to practical examples and exercises

List of topics and cases by week

Week 1

Data revolution in business, economics and finance

McAfee, Andrew and Erik Brynjolfsson (2012), "Big Data: The Management Revolution", Harvard Business Review, October 2012, pp. 60-68.

Davenport, Thomas and D.J. Patil (2012), "Data Scientist: The Sexiest Job of the 21st Century", Harvard Business Review, October 2012, pp. 70-76.

Big data web resources: [kaggle.com](https://www.kaggle.com), stackoverflow.com, [quora.com](https://www.quora.com), packtpub.com etc

Introduction to Data Visualization

Week 2

Review of R and Rstudio

Getting started with coding in R

Exploratory data analysis

Application: "Domain auction prices"

Week 3

Analysing text data

Application: "Airbnb property descriptions"

Week 4

Types of plots in R/RStudio

Building interactive plots with googleVis package

Creating dashboards with Shiny

Time series analysis

Application: "WDI Shiny dashboard"

Week 5

Recommender systems, association mining, measuring similarity and relationships

Application: "Customer analytics at Bigbasket - product recommendations. Indian Institute of Management, IMB Case 573."

Week 6

Overview of Business Intelligence. Types of visualizations and best practices of use

Visualizing data and digital storytelling with Tableau

Application: "Sample superstore"

Week 7

Road ahead: From Data Analytics to Artificial Intelligence

Presentations of student projects

Grading

Letter grades will be distributed according to the standard KSE grading scale based on attendance and practical assignments (30%), course project (30%) and final exam (40%). The final exam will be computer-based, focusing on examples studied in class. The score for the worst practical assignment will be dropped if **all assignments are submitted**. If assignment is not submitted it will count as zero and will not be dropped.

Policy on Student Attendance

It is expected that you attend classes regularly. In the case of an emergency when you must skip a class, it is your **sole responsibility to copy lecture notes** from a classmate. At least part of the exam material will be based on lectures and not on the textbook.