10122601-10 Big Data in Economics
— Online course —
Summer term 2021

Instructors: Lars Winkelmann, Elias Wolf.

Course website: all material and instructions will be posted on blackboard. You can register for this course via Campus Management.

Course description
We are living in a data revolution. We see a boom in data collection and analysis. Information on human behavior, environmental changes, economic activity, political institutions is not only being collected on a new scale, but data is becoming more accessible than ever before. Besides directly affecting real-world life in raising productivity and living standards, Big Data has important implications for the field of statistics and econometrics.

This course focuses on statistical methods and their application. To explore Big Data in Economics you need profound knowledge about statistical methods—their limitations and pitfalls. We will see that your standard stats (10120401) and econometric (10121201) methods fail in data-rich environments. We focus on: *i*) Time series, *ii*) multiple testing and *iii*) shrinkage and dimension reduction techniques.

Organizational issues and grading
The course is split into two parts. In the first part you learn basic statistical methods and concepts. We provide book chapters, videos and tutorials. All material will be available from April 19th on. You can read, watch and learn anytime you want. Until June 4th you need to submit your solutions of two mandatory homework assignments. The homework contributes 20% to your final grade. In the second part of the course you will write a paper review. We provide text and videos that tell you "how to write a paper review" and "how to use the text processing language LaTeX". You pick a Big Data paper from a list available on blackboard or propose your own Big Data paper. The submission deadline for your paper review is July 2rd. The review counts 80% of your final grade.

We will support and supervise your work over the whole semester via webex/phone/ regular office hours.

Prerequisites
Students are required to have taken the math (10120301) and stats courses (10120401). Having taken econometrics (10121201) or taking econometrics concurrently is strongly recommended. Students are assumed to be familiar with basic concepts in linear algebra, analysis, probability theory and statistical inference. Basic programming abilities in R are required to be able to actively follow the tutorials and to solve the homework. Students need to be willing to write their paper review in the processing language LaTeX.

Deadlines

- Due date for homework submissions: **June 4th, 2021, 24:00**, via mail to eliaswolf@zedat.fu-berlin.de. *No late homework assignments will be accepted.*

- Tell us which Big Data paper you will review. Write a mail with your name, student ID and title of the paper to eliaswolf@zedat.fu-berlin.de. Deadline until **June 4th, 2021, 24:00**.

- Submission deadline for the paper review: **July 2rd, 2021, 24:00**. Send your review in pdf to eliaswolf@zedat.fu-berlin.de. *There will be absolutely no extensions on the paper review, so please plan well in advance.*

Academic Policy

Discussing ideas and work-in-progress with fellow students and instructors is very important. The homework and paper review you ultimately turn in must be your own original writing, based on your own thoughts. The report must use appropriate citation practices (link). If you are caught plagiarizing in this class, you will fail the course. For information about the stance on academic honesty at FU Berlin, consult the Code of Hounor.

Appetizer

There are lots of starting points to get exited about Big Data in Economics. You may start your journey with an NBER Working Paper, the UN Global Pulse website, a nytimes article and/or a talk on open data.

Please be aware that the first part of the course focuses on abstract statistical methods—we will not talk about fancy data projects. Your paper review, however, can relate to a specific Big Data project.