

Case Studies

Fallstudien II

Prof. Dr. Matei Demetrescu, Dr. Paul Navas

Why Case Studies?

This is where you learn to tackle realistic problems using statistics:

- From the problem...
- ...to the data...
- ...to the report(s).

(Not to mention group work...)

Some details

Intended audience

- Master Econometrics (ME3), Master Data Science (MD4)
- Master Statistik – beware of the English baseline

Prerequisites

- All requirements obtained in case of conditional admission (any of modules ME/MD Req1 to ME Req7)
- All conditions must be fulfilled before the beginning of the course!
- **Highly recommended:** Successful pass of module ME1/MD2 (Statistical Theory)

Overall task

You will provide forecasts for the number of cases and deaths in the COVID pandemic for the US.

- Understand idiosyncracies of the situation
- Provide out-of-sample forecasts using linear and ML methods
- Consider forecast MSE as performance measure, use other forecast evaluation tools
- Think about explanations/interpretation why some methods work better than others.

COVID Forecasts

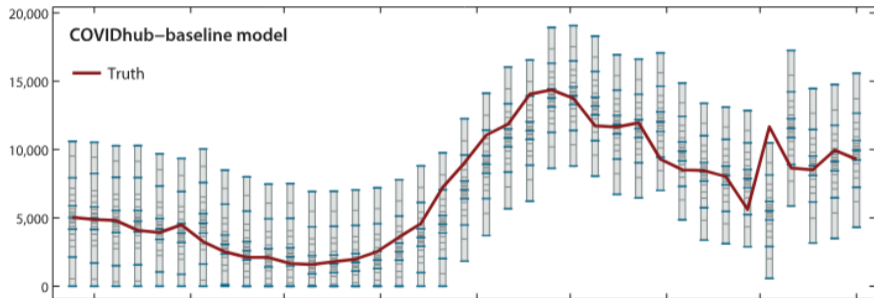


Abbildung: COVID Forecast (Quelle: Dimitriadis et al)

More specifically

Methods and findings are discussed together with the other participants and presented in formal reports:

- 1 Describe the data and use appropriate transformations.
- 2 Choose appropriate statistical methods and adapt them to the problem at hand.
- 3 Carry out a comprehensive analysis of the data, concentrating on forecast evaluation
- 4 Use the results to provide an answer to the research question.

Three Reports

Overall you will write two short reports and a longer one:

- Short report about basic forecasting with OLS predictive regressions (no more than 15 pages)
- Short report about a comparison of different statistical learning methods (no more than 15 pages)
- Longer report about an exhaustive analysis with comparisons of different methods, focusing on evaluation (no more than 30 pages)

Guidelines will be provided. You need to pass **all** three reports, but you have one chance to redo one of them.

Programming Tasks

- For each report there are different tasks in which you have to analyze data. For this you are supposed to use R but usage of Python is also allowed!
- Also, you will have to hand in your own code.
(The code might have to be looked at to confirm your results in the reports!)

Schedule

- Wednesdays 2pm to 6pm – attendance is compulsory! (some lenience in hardship cases)
- Room CDI 121
- There is a moodle room (passwd P4ndemic12!)
- First meeting: April 15th

Final details

Contact

- `mdeme@statistik.tu-dortmund.de`,
`navas@statistik.tu-dortmund.de`

Office hours

- by appointment; but before resorting to such desperate measures do not hesitate to ask simpler questions per email.

Q & A

Any specific questions?