

# Statistics for spatio-temporal data: Final evaluation

Isadora Antoniano-Villalobos

Academic year 2019/2020

A 50% of the final grade is given by a written project, including a data analysis, of which:

- 50% for the quality and originality of the analysis and methodology
- 15% for the readability and appearance of the work
- 10% for the clarity of the R code
- 25% for the project structure and consistency, from introduction and motivation to the conclusions

The remaining 50% will consist of an oral examination, including a 15 minute presentation of the project, of which:

- 40% Oral presentation
- 10% Appearance and clarity of the slides
- 50% Oral exam (response to questions regarding the project and non-included subjects)

All written outlets of the projects must be submitted via email to:

[isadora.antoniano@unive.it](mailto:isadora.antoniano@unive.it)

## Project guidelines:

Choose a subject of interest to you, if possible, related to the topic of your future or current dissertation, but not something that you have already done. You may derive inspiration from the literature but keep in mind that originality will be evaluated!

You can work individually or in pairs.

- The extent and quality of the written project must reflect the number of participants.
- If working in pairs, the same grade on the written project will be assigned to both members of the team. No complaints accepted: you choose your team mate, you accept the consequences!
- Oral presentation and evaluation will be individual.

Submission is organized in 3 stages:

1. **Project proposal** (roughly 1 page). To allow evaluation of the viability, interest and compatibility with the course, the proposal should include:
  - Introduction and/or motivation for the problem
  - A description of data (source, sample size, format, variables, etc)
  - Statistical methods that you intend to use

**No final projects will be accepted without a previously approved project proposal!**

2. **Project draft.** Not mandatory but strongly recommended. This will be the only opportunity to receive feedback on your project before the final submission. It should resemble the final project as closely as possible to allow for quality feedback.
3. **Final submission.** The report must be submitted in pdf format, with a 12pt font on a regular A4 layout (10 page maximum, including tables and figures, for individual projects; 15 for teams of 2 people).

Please use a scientific style, concise but complete. **The report should contain:**

- i. Introduction/ motivation
- ii. Methodology, detailing any assumptions and the statistical methods applied used. Reading this Section, it should be clear that you fully master these methods.
- iii. Data description and results
- iv. Discussion/conclusions and future work, summarizing the main results and outlining possible limitations new research questions emerging from the analysis.

**A file with the R code must be attached:** it must contain all necessary commands to reproduce the results. The code should be clear, complete and properly commented. Any unclear commands are subject to interrogation and evaluation will be affected by lack of justification

## Specific requirements

For the analysis of temporal data:

- Application of filtering or smoothing methods to the time series in order to identify possible trends and seasonality
- Verification the stationarity and the randomness of the time series
- Evaluate the data support for different ARIMA processes
- Forecast and forecast assessment

For the analysis of spatial data

- Appropriate exploratory analysis, including graphs and numerical summaries
- Trend removal and analysis of sample semivariograms to identify relevant features (i.e. lack of stationary or anisotropy)
- theoretical semivariogram fitting
- Prediction and goodness of fit assessment

## Possible sources of data

Possible sources of interesting data sets include, but are not limited to:

- [Climate Data Guide](#)
- [Data.gov](#)
- [eBird](#)
- [Google Flu and Dengue Trends](#)
- [Intergovernmental Panel on Climate Change Data Distribution Centre](#)
- [National Cancer Institute GDC Data Portal](#)
- [National Oceanic and Atmospheric Administration](#)
- [United States Environmental Protection Agency Air Quality Data](#)
- [U.S. Energy Information Administration](#)