

Spring 2021 Syllabus

Social Data Science

Prof. David Garcia
Chair of Systems Design, ETH Zurich

Course dates: Monday 15.02.2021 to Friday 19.02.2021

Location: Online

Course time: 9:00-12:00 and 13:00-16:00

Course moodle: <https://moodle-app2.let.ethz.ch/course/view.php?id=14192>

All exercises are based on the R Statistical Language. Lecture handouts are available online before all sessions at <https://dgarcia-eu.github.io/SocialDataScience/>. Markdown files with exercise handouts will be available at <https://github.com/dgarcia-eu/SocialDataScience> and exercise solutions will be available on Moodle 24 hours after each practical session. Each day will have two introductory group sessions (at 9:00 and at 13:00) and two Q&A group sessions (at 11:00 and at 15:00). Videos and text handouts will be available to guide students at their own pace. Morning sessions will focus on foundations and methods. Afternoon sessions focus on tutorials and exercises.

The course assessment is based on final course project done in groups (2-4 students). The deadline to deliver the final project will be at the end of the Easter break, approximately 2 months after the course. Project grading will be based on three parts:

1. a short project description that has to be posted on the Moodle forum by 5.03.2021
2. the final text of the report (max. 6 pages)
3. the R code to generate the results of the project

Students receive input through the forum on their short project descriptions during the two weeks after the course. Optional short quizzes at the end of each day can increase the grade (up to 4% each for a total of at most 20%). Moodle forum participation in course discussions can increase the grade for up to 10%.

*Day 1. Introduction to Social Data Science**15.02.2021*

.....
Objectives: Students will learn the definition of Social Data Science and its relation to other disciplines. Students will learn the concept of long-term orientation and how to study it through search data. Students will become familiar with R and its packages through a test of a social science theory with search data.

- What is Social Data Science?
- Measuring temporal orientation with Google Trends
- R crash course
- Google Trends and World Development Indicators in R
- Exercise: Future orientation and economic development

*Day 2. Social Dynamics**16.02.2021*

.....
Objectives: Students will learn theories about social impact and how they are related to information spreading. Students will learn how to retrieve Twitter user data and to test the division of impact hypothesis.

- Social Impact Theory
- The Simmel Effect
- Linear regression and bootstrapping
- Data wrangling with dplyr
- The Twitter API in R
- Exercise: Division of impact on Twitter

Day 3. Computational Affective Science***17.02.2021***

.....
Objectives: Students will learn the theory and measurement of emotions in social media. Students will develop a critical understanding of sentiment analysis and how to apply it.

- Measuring emotions
- Supervised and unsupervised sentiment analysis
- Running sentiment analysis in R
- Exercise: Evaluating sentiment analysis methods
- Exercise: Twitter sentiment and retweeting

Day 4. Social Network Analysis***18.02.2021***

.....
Objectives: Students will learn how to construct and analyze social networks. Students will learn how to gather their own Twitter network data, how to analyze it, and how to visualize it.

- Introduction to social networks
- The Friendship paradox
- Centrality
- Analyzing networks in R
- Exercise: The friendship paradox on Twitter

Day 5. Social Network Phenomena***19.02.2021***

.....
Objectives: Students will become familiar with complex social network phenomena related to communities, assortativity, and distance. Students will learn how to calculate complex network metrics from Twitter data.

- Social resilience
- Structural holes and communities
- Assortativity
- Community detection
- Exercise: Assortativity among Swiss politicians on Twitter