

Prof. Anton Ovchinnikov

Prof. Spyros Zoumpoulis

DSB Classes 13-14, February 19, 2019

- **Project Presentations; Wrap-up**

Structure of the course

- SESSIONS 1-2 (AO): Data analytics process; from Excel to R
 - Tutorial 1: Getting comfortable with R
- SESSIONS 3-4 (AO): Time Series Models
- SESSIONS 5-6 (AO): Intro to classification, logistic regression and machine learning
- SESSIONS 7-8 (SZ): Advanced Classification; From .R to Notebooks
 - Tutorial 2: Midterm R help / classification
- SESSIONS 9-10 (SZ): Dimensionality Reduction; Clustering and Segmentation
 - Tutorial 3: Q&A on R for three main modules
- SESSIONS 11-12 (SZ): The Data Science Process; Guest speaker
 - Hands-on help with projects
- **SESSIONS 13-14 (AO+SZ): Project presentations**

Plan for the day

Learning objectives

- Final project presentations
 - Feedback: on the form + Q&A
- Discussion on learnings from the project and the course

The Process

- A group presents
- The group that presented right before is responsible for reacting, e.g.
 - what would you do differently, or/and
 - what did you like and why?
- When reacting, think of relevance, impact, feasibility, scalability, but also modeling or technical difficulty, cool implementation
- Everybody fills in the feedback forms for all (other) groups
- Overall, each group receives (privately) feedback from
 - Anton & Spyros
 - All of your classmates (anonymously)
- Last 20 minutes: Discussion. Think of two sets of learnings:
 - 3 learnings from your project
 - 3 learnings from the course

Running list of challenges and issues

- Ran out of time.
- Communication and coordination within team
 - Version control
- Team project management/decision making
 - Unsure what to do on drop/no drop project decisions
 - Computation time for a piece of the execution too long... and managing resulting second thoughts: is this a good path/not good path?
- Finding data is easy. Finding good quality/amenable to insights data is hard.
- Data is all over the place, in numerous different sources/files. Need to consolidate.
- Is the available training set meaningful/useful for my business problem?
- Technical issues / coding roadblocks

What is next after DSB?

- At INSEAD
 - Pricing Analytics (P3/P5)
 - Decision Models (P5)
- Outside INSEAD
 - Numerous good online courses on machine learning and coding on Datacamp, Coursera, ...
 - Rich online data science community: Kaggle, Stack Overflow,...
- Stay thirsty

The background of the slide is a green-tinted collage. At the top, there's a large crowd of people, possibly at a graduation ceremony. Below that, on the left, is a classroom scene with students sitting at desks. On the right, there's a modern building with large glass windows, identified by the INSEAD logo on its facade. At the bottom left, there are silhouettes of people in dynamic, athletic poses.

INSEAD

The Business School
for the World®

Europe

| Asia

| Middle East