Group project work: prediction task

- In the project, your group competes with us
  - A higher ranking usually leads to a better course grade
- Each group submits
  - Predictions for the competition
  - Two short Jupyter notebooks
    - They correspond to your two selected predictions in Kaggle
    - They contain only the necessary steps to produce your predictions
    - We may re-run your notebook to check whether the results are reproducible.
  - A project report
    - You can use a PDF or another Jupyter notebook
    - The report summarizes all attempts in your group, including exploratory data analysis, feature engineering, all models/algorithms no matter they are helpful or not
    - The report should include model interpretation

Course grade

- The course is letter-graded
  - We first calculate your points (0-100)
  - the points are converted to letters according to the ranges
    - A: 89-100
    - B: 77-88
    - C: 65-76
    - D: 53-64
    - E: 41-52
    - F: 0-40
- The points are calculated almost based on the project
- But you must pass the individual assignment first
  - Try to pass in the first chance (i.e. with the submission bef the deadline)
  - If you fail the first time,
    - we give you a second chance (resubmit by another given da
    - if your re-submission passes, you can continue to project, but with a deduction (-5) in your course points
  - If you fail both chances, you fail the course.

Deadlines are strict

- One minute late is late-submission ⇒ you will get deduction
- Submit early
  - Last minutes can be crowded
- DO NOT SUBMIT AFTER THE DEADLINES!
  - Because we grade your last attempt
  - Last attempt after the deadline = late-submission
- Most extension excuses are not acceptable
  - Start early

Course points

- Your course points are the sum of
  - base points of your group project
  - possible deductions
- Base points
  - $\propto$ the number of Virtual Teams (VTs) you defeat
  - max 100 (defeat all VTs) and min 41 (defeat 1 VT)
  - if you defeat 0 VT, you fail the course
  - VTs are prepared by the teachers and assistants
  - there will be 5 or more VTs
- Possible deductions
  - pass of individual assignment in the second chance (-5)
  - late submission ($\leq$ 3 days after the deadline) of the project (-30)
  - no exploratory data analysis (-3)
  - only one predictor is used (-3)
  - no feature engineering (-3)
  - no model interpretation (-3)
- More details will be announced later

Course points and letter grade (example)

- A student
  - passes the individual assignment in the first chance
  - submits the project in time
  - his/her team defeats 7 of 10 VTs
  - no model interpretation in the notebooks (-3)
- Then the student's course points
  - base points = $41 + \frac{100 - 41}{10 - 1} \times (7 - 1) \approx 80$
  - course points = $80 - 3 = 77$
  - rounded letter grade is B