ProjectWork:
Some coaching

Raffaele Zenti (raffaele.zenti@virtualb.it)

Co-Founder, Chief Data Scientist, Virtual B SpA

Your goal

A working model:

- Solving a business problem
- Inspired by the data I have left you
- Somewhat documented and explained (presentation)

It is what happens in a company specializing in Data Analytics when starting a new product development or tackling a project from a customer

Structure

- Short presentation

- Code (very readable)

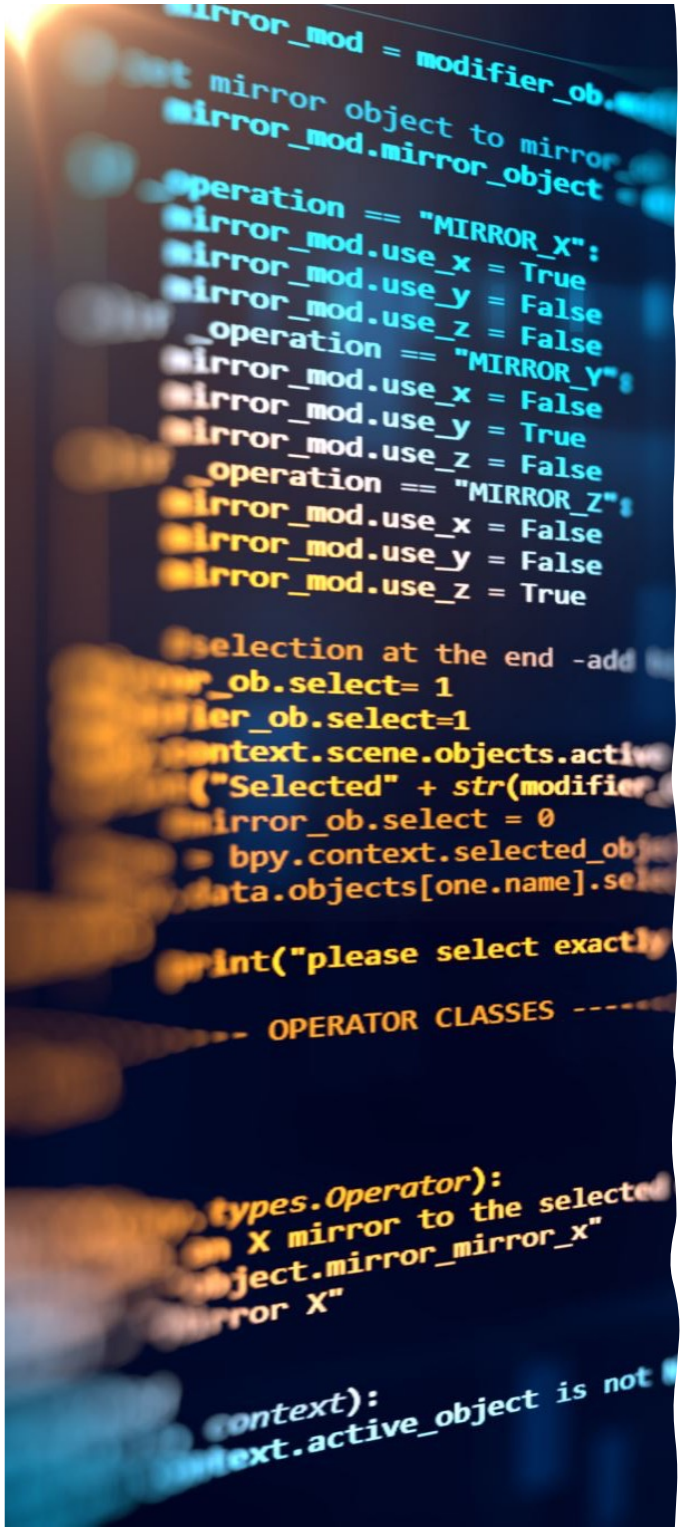
About the presentation

Short: 12-20 slides

You will not pitch, you will just send it to me

Straight to the point:

- **WHAT**: What have you done
- **WHY**: What problem solves
- **HOW** - How you did it
- **KEY ASPECTS (if any):** Any key aspects to underline, both of business and method
- **KEY RESULTS**: What results you got (they don't have to be exceptional, in Data Science you get both good and bad results, the important thing is the process)
- **OPTIONAL:**
  - Key references
  - Appendix



Code

- Live Editor Matlab /Jupyter Python, or both (you choose)
- Well commented, very readable
- Representing the ML pipeline (Data Retrieval/Data Exploration/Modeling/etc etc)
- Clean:
  - Only the key steps, not the various abortive attempts, skilful use of graphics (with legends, labels, etc) ... must be easily understood
  - Also, make it an html (can be done both in Matlab and Python), in addition to the original code

About the pipeline at this stage (i.e. model research)

Data ingestion

Data preparation: cleaning, exploration, transformation and feature selection/engineering

Data Segregation: split subsets of data in train validation/cross-validation, and test

Training and validation: assess the performance of the model(s) using the proper subset of data, pick the right metrics, and understand how accurate the prediction is

Deliverable: a tested, working model ready for deployment (which is another story)

Recap

Deliverables

- Presentation: pdf format
- Code: Matlab or/and Python
- HTML copy of the code

- Please **send everything to**:

  raffaele.zenti@virtualb.it

- (CC Prof. Marazzina, and all the team members)

- **When**: about 10 days before the oral examination