

DSI - Basic Info



Edoardo Venturini
DSI Instructor

My name is **Edo**ardo

Structure

5 units

4 projects

1 capstone project

Units

Unit	Title	Topics Covered	Recommended Length of Time
Unit 1	Data Science Foundations	Fundamental Statistics, Databases, Python, & other Data Science Tools.	Week 1
Unit 2	Exploratory Data Analysis	SQL, Data cleaning, Transformation, & Visualization	Weeks 2-3
Unit 3	Classical Statistical Models	Regression, Classification, & KNN	Weeks 4-6
Unit 4	Machine Learning Models	Unsupervised Learning, CARTS, Naive Bayes, ARIMA, & NLP	Weeks 7-9
Unit 5	Advanced Topics	Bayesian Modeling, Split-Testing, & Deep Learning	Weeks 10-12

Projects

Projects	Deadline	Description	Requirements
Project 1	4th June 2018, Monday, 6am	Python + Pokemon Data	Students will demonstrate command of basic Python data structures, implement Python syntax for control log functions, and calculate descriptive statistics.
Project 2	15th June 2018, Friday, 6am	EDA + Double Datasets	Students will analyze SAT and drug use datasets & explicitly state their assumptions. They will form hypotheses and justify them with solid statistical testing in NumPy, while visualizing and interpreting their plots using Matplotlib and Seaborn.
Project 3	6th July 2018, Friday, 6am	Regression + Housing Data	Students will review the Ames housing dataset & explicitly state their assumptions. Then they will clean, impute, and explore the dataset, justifying their modeling choices and any feature engineering. They will need to plot, visualize, and interpret their data while clearly outlining their modeling strategy in response to the prompted questions. They will also need to explain their results to a technical audience.
Project 4	27th July 2018, Friday, 6am	Web Scraping + Career Data	Students will scrape and prepare their own data, then create and compare at least two models for each section: an ensemble model or decision tree, and a classifier or regression model. They will document their data and models in a notebook for technical peers, while also creating a written summary for non-technical stakeholders.

Capstone Project

Part	Description	Deadline
1	Pitch	22nd June 2018, Friday
2	Dataset	13th July 2018, Friday
3	EDA	20th July 2018, Friday
4	Findings	10th Aug 2018, Friday
5	Presentation	17th Aug 2018, Friday

Deadlines

Projects	Deadlines
Unit Project 1	4th June 2018, Monday, 6am
Unit Project 2	15th June 2018, Friday, 6am
CAPSTONE Project Stage 1 (Pitch/Datasets)	22nd June 2018, Friday
Unit Project 3	6th July 2018, Friday, 6am
CAPSTONE Project Stage 2 (Dataset)	13th July 2018, Friday
CAPSTONE Project Stage 3 (EDA)	20th July 2018, Friday
Unit Project 4	27th July 2018, Friday, 6am
CAPSTONE Project Stage 4 (Findings)	10th Aug 2018, Friday
CAPSTONE Project Stage 5 (Presentation)	17th Aug 2018, Friday

Communication

Slack channel

Slack private message

email

edoardo.venturini@generalasseb.ly

hafsa.laeque@generalassemb.ly

luther.lim@generalassemb.ly

Style

A lot of coding/labs

Not much math

Interactive lessons

Group work

**What I want
from you**

A lot of questions!

WHAT IF I TOLD YOU...



THERE ARE NO STUPID QUESTIONS.

100% focus during the lessons

Actively participate to the lessons!

(less boring + you learn more)

A few points

Yes, my english is really bad!



I'm not a university professor :)

Lessons start at 9:00

We're here to help you!

...but only for the next 12 weeks.

Google is your friend

Learning how to solve problems
is crucial!

80% of coding is

~~solving problems~~

googling + getting frustrated

Programmer



What my friends think I do



What my mom thinks I do



What society thinks I do



What my boss thinks I do



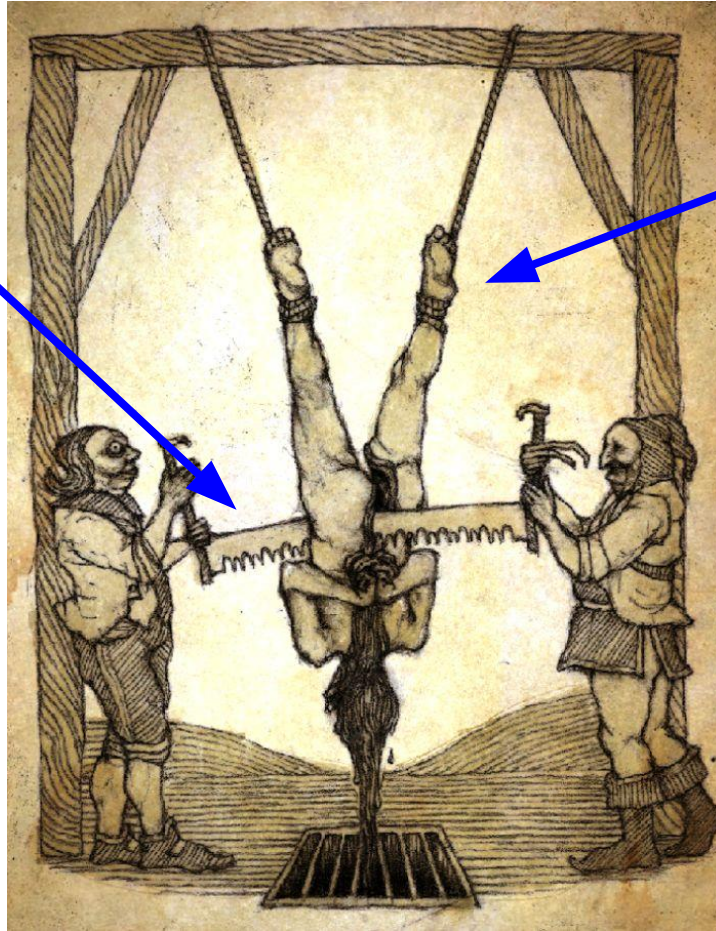
What I think I do



What I actually do

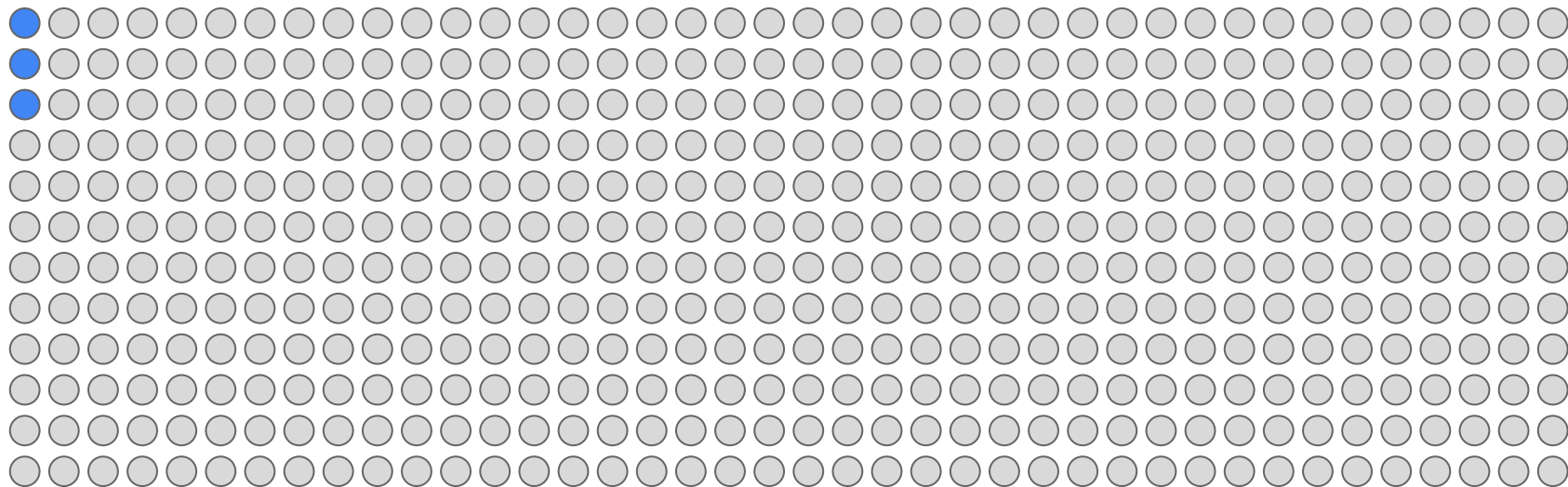
Python

You

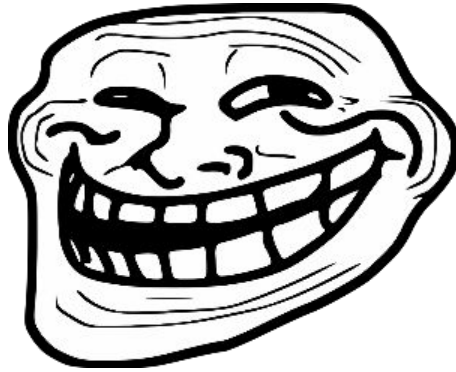


“They should understand they’ll have no private life in the next 12 weeks.”

Former DSI Instructor



The previous graph is partially false



Last but not least

You are a team!

Sharing + Helping

One of your classmates can be
your boss in a few years

Questions?

Prepare a 90-sec intro about yourself

1. Name
2. Background
3. Why you are here
4. What you expect from the course
5. What you want to do next (industry,etc)
6. Optional: fun fact about you