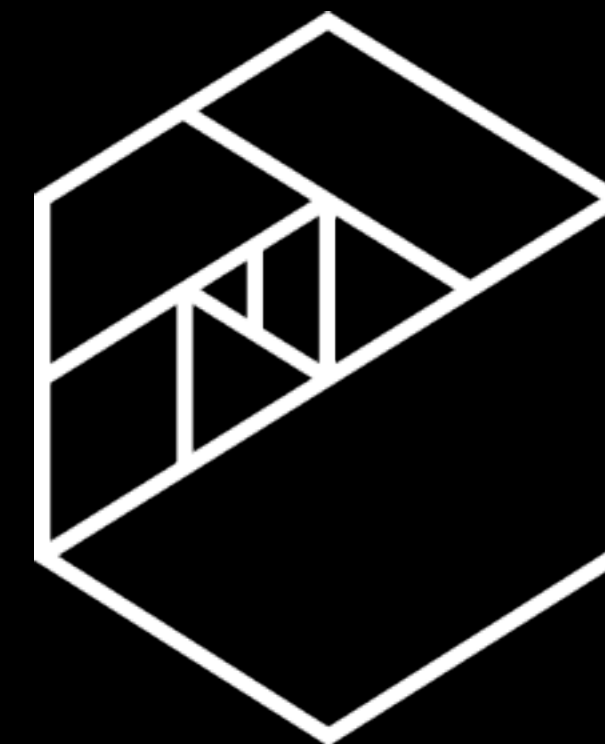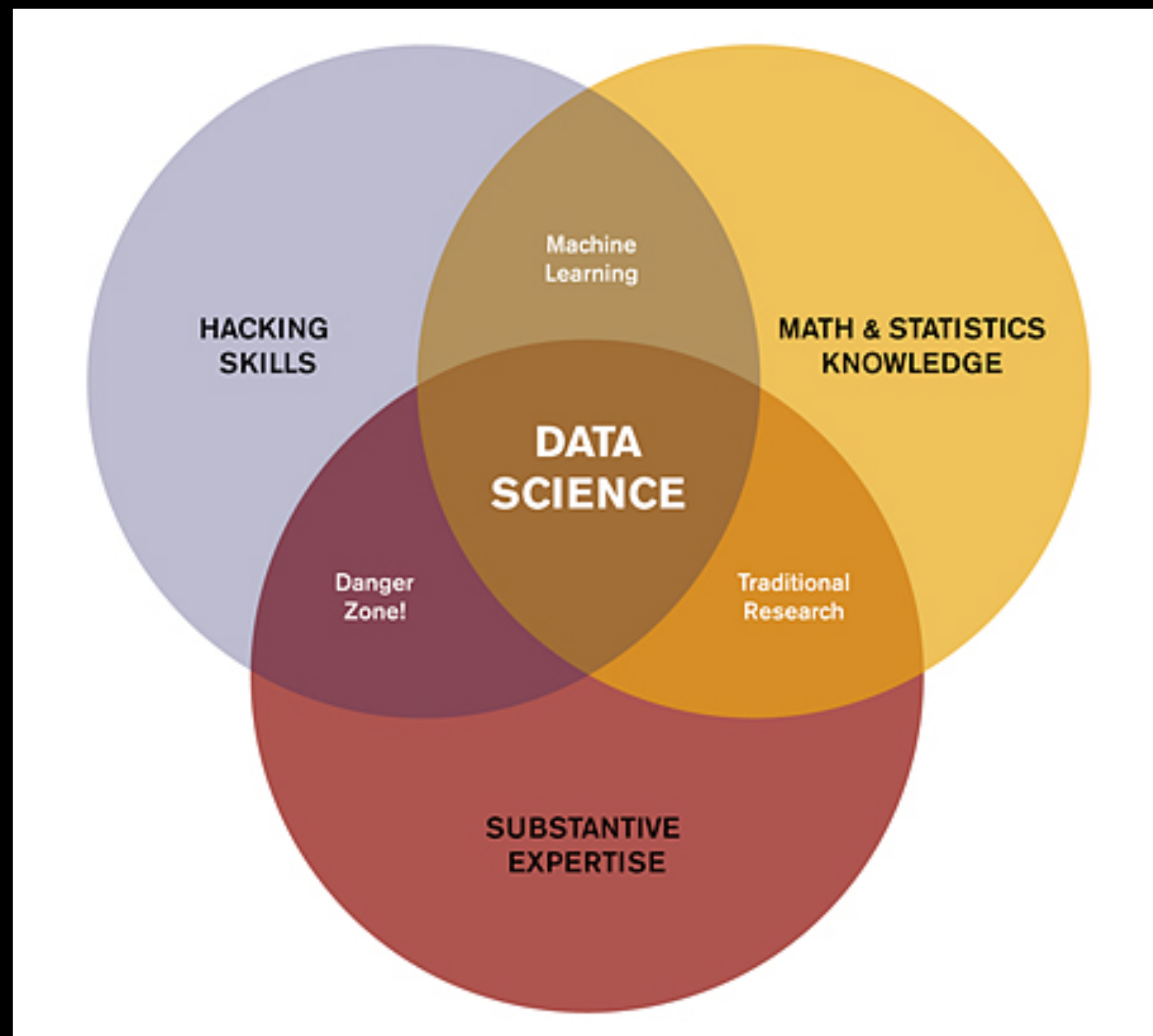# What is a Data Scientist

And how will you become one?
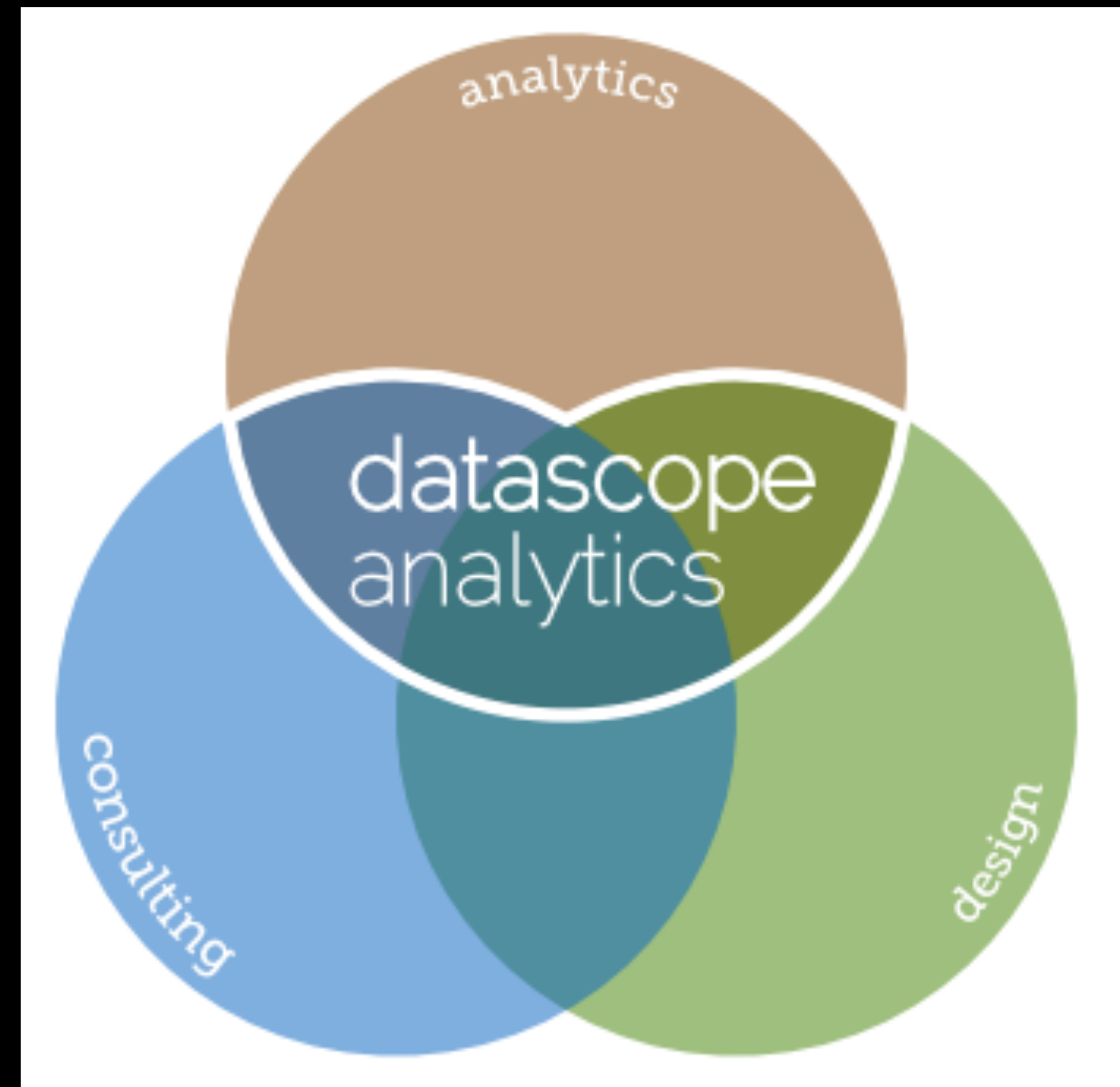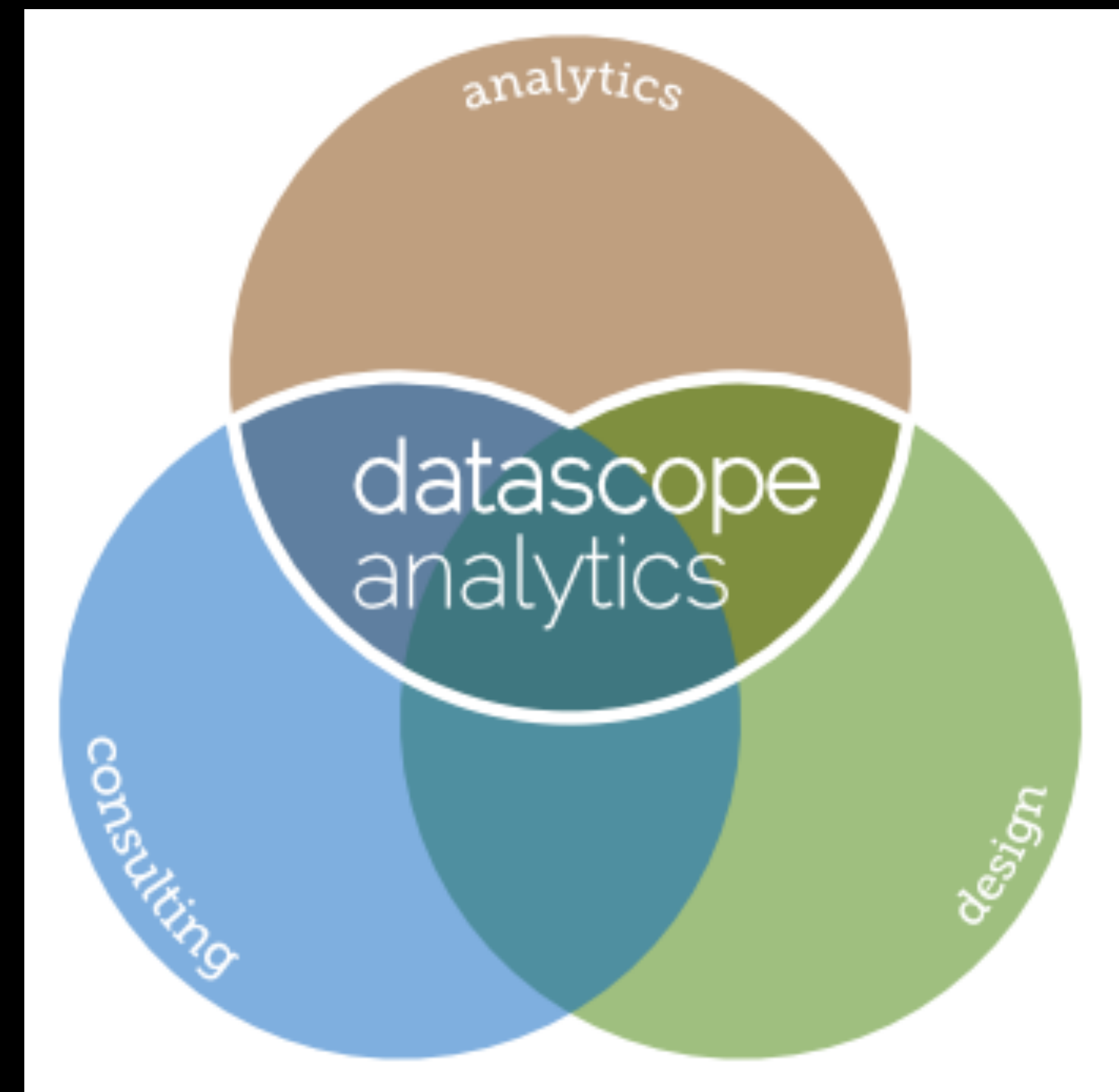
METIS

So... what is a data scientist?

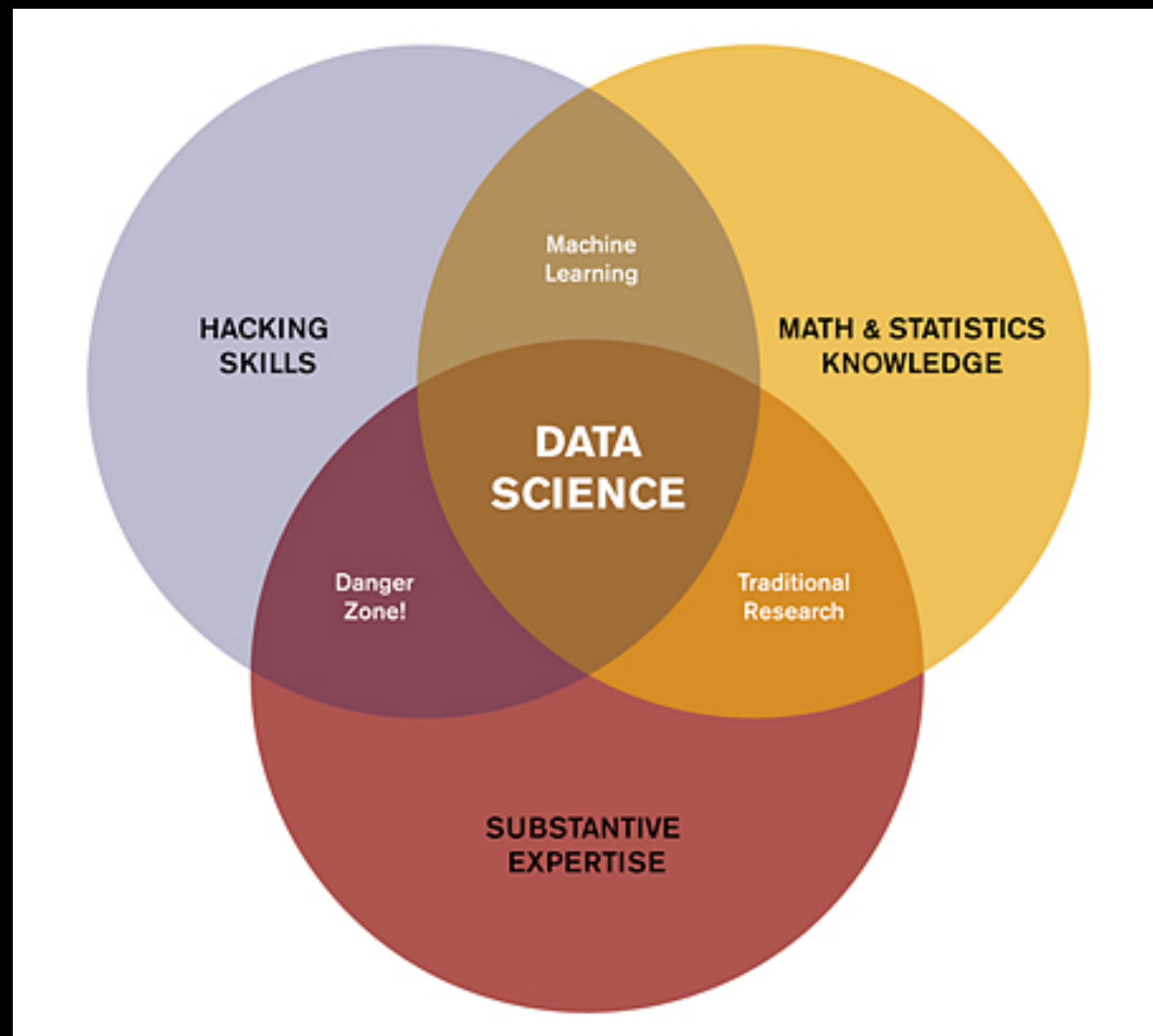# So... what is a data scientist?

So... what is a data scientist?

Skill domains of a data scientist

Statistics / Machine Learning
Programming
Communication
Design

Who can be a data scientist?

Personality Traits

Who can be a data scientist?

Personality Traits

Curiosity
Creativity
Grit

Who can be a data scientist?

Personality Traits

Continuous drive to learn
Creativity
Grit

Who can be a data scientist?

Personality Traits

Continuous drive to learn
Tons of different angles & ideas
Grit

Who can be a data scientist?

Personality Traits

Continuous drive to learn

Tons of different angles & ideas

Ability to push through walls

Who can be a data scientist?

Personality Traits

Continuous drive to learn
Tons of different angles & ideas
Ability to push through walls
Integrity

Who can be a data scientist?

Personality Traits

Continuous drive to learn
Tons of different angles & ideas
Ability to push through walls
Integrity
Humility

For the next three months, you will have...

Focused time

Peer and community support

Placement managers, hiring partners

Experienced practitioner-instructors

Bootcamp model

Training
Repetition
Culture/community

Training

Skills

O
S
E
M
N

Training

Skills

Obtain

Scrub

Explore

Model

iNterpret

Training

Skills

Obtain
Scrub
Explore
Model
iNterpret
Communicate

Training

Skills → tools

Data acquisition
Data exploration
Machine learning & statistics
Computer science
Web
Visualization
Domain awareness

Training

Tool selection

Comfort & facility with a core toolbox

Awareness of options

Some curveballs

Training

Tools

Python when possible

Training

Tools

Python, ipython notebook, git/github

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron
(flat files), PostgreSQL, mongoDB

Training

Tools

Python, ipython notebook, git/github

BeautifulSoup, selenium, cron

(flat files), PostgreSQL, mongoDB

numpy, scipy, matplotlib, pandas

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron
(flat files), PostgreSQL, mongoDB
numpy, scipy, matplotlib, pandas
statsmodels, scikit-learn

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron
(flat files), PostgreSQL, mongoDB
numpy, scipy, matplotlib, pandas
statsmodels, scikit-learn
HTML, CSS, JS, d3.js

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron
(flat files), PostgreSQL, mongoDB
numpy, scipy, matplotlib, pandas
statsmodels, scikit-learn
HTML, CSS, JS, d3.js
Amazon Web Services, Flask

Training

Tools

Python, ipython notebook, git/github
BeautifulSoup, selenium, cron
(flat files), PostgreSQL, mongoDB
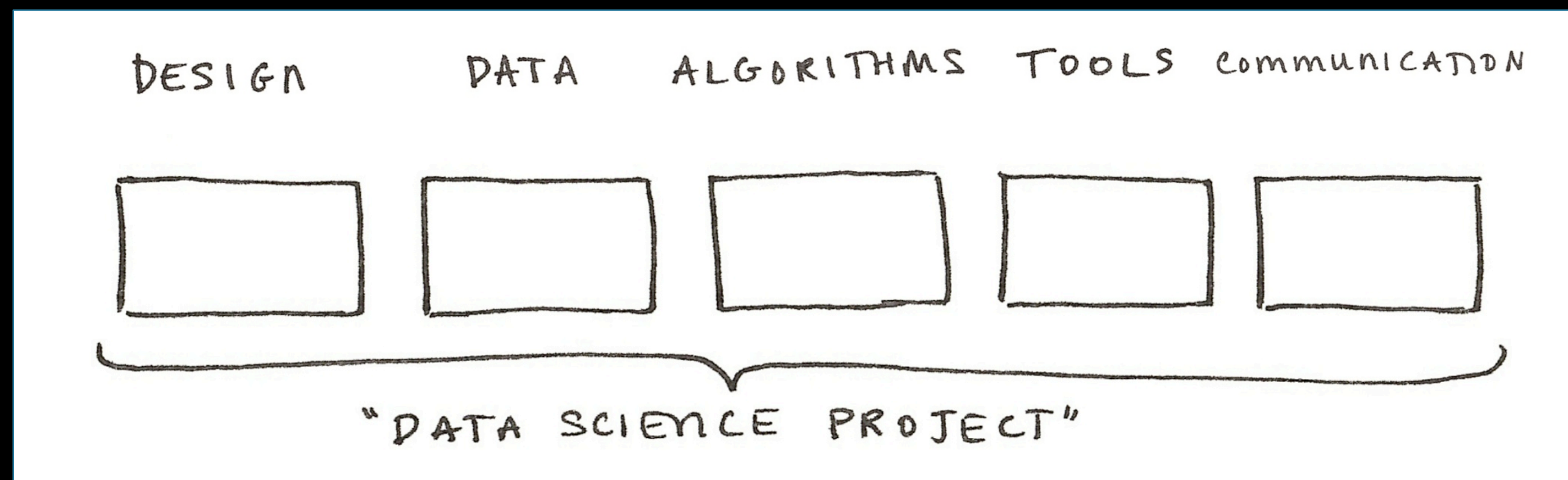numpy, scipy, matplotlib, pandas
statsmodels, scikit-learn
HTML, CSS, JS, d3.js
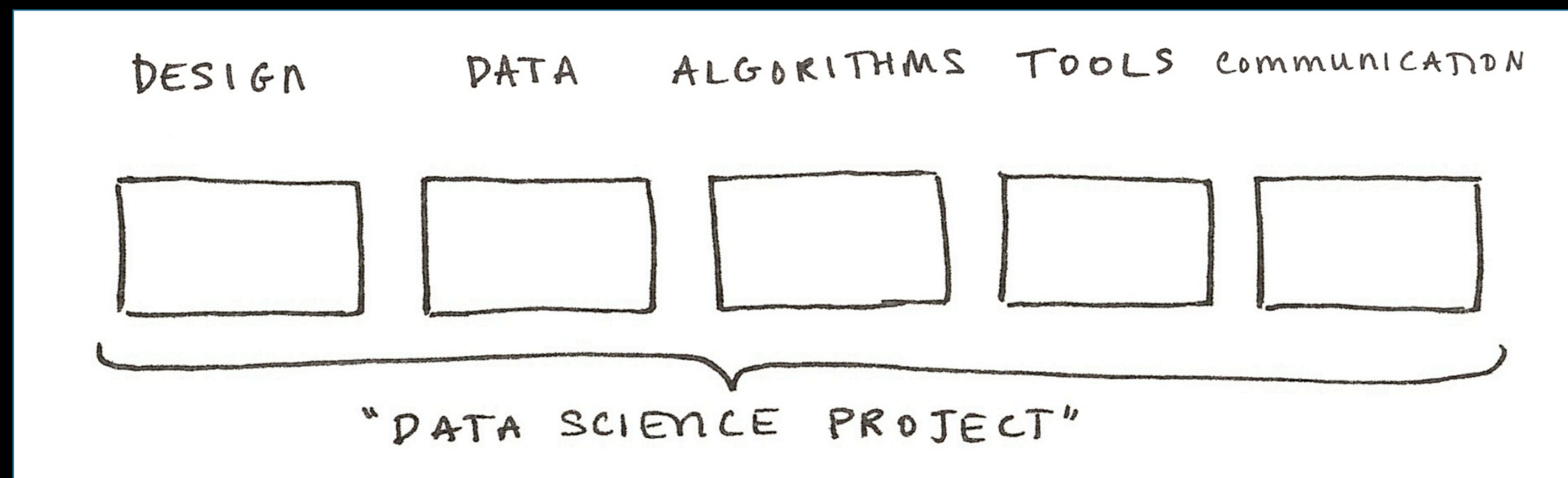Amazon Web Services, Flask
Google

Repetition

Project-based model
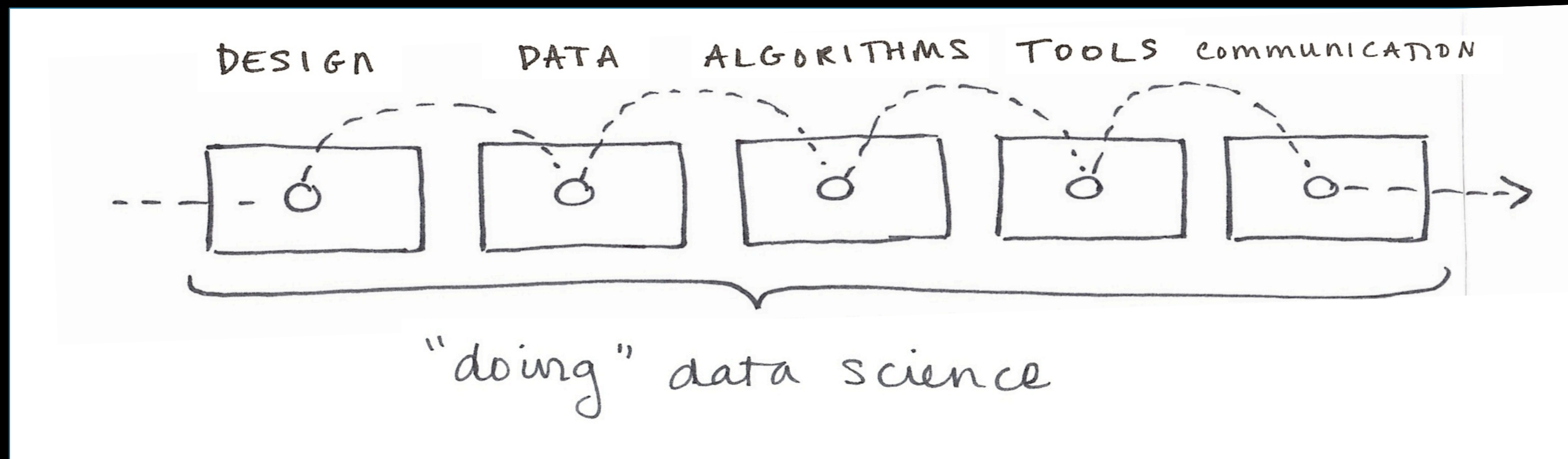
Repetition

Project-based model

Unit of repetition should be *the practice of doing data science*

DESIGN   DATA   ALGORITHMS   TOOLS   COMMUNICATION

"DATA SCIENCE PROJECT"
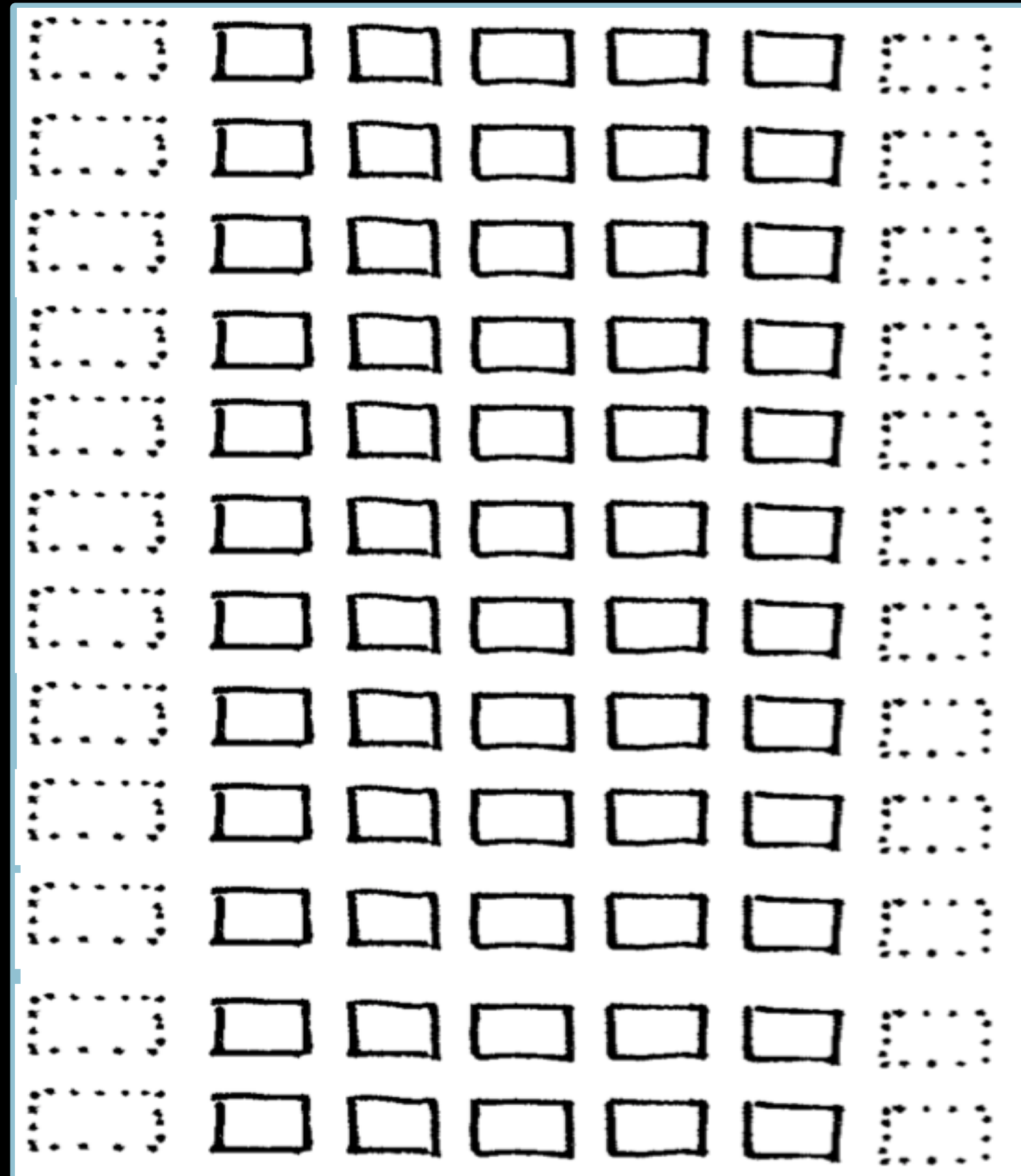
Repetition

Project-based model

Unit of repetition should be *the practice of doing data science*
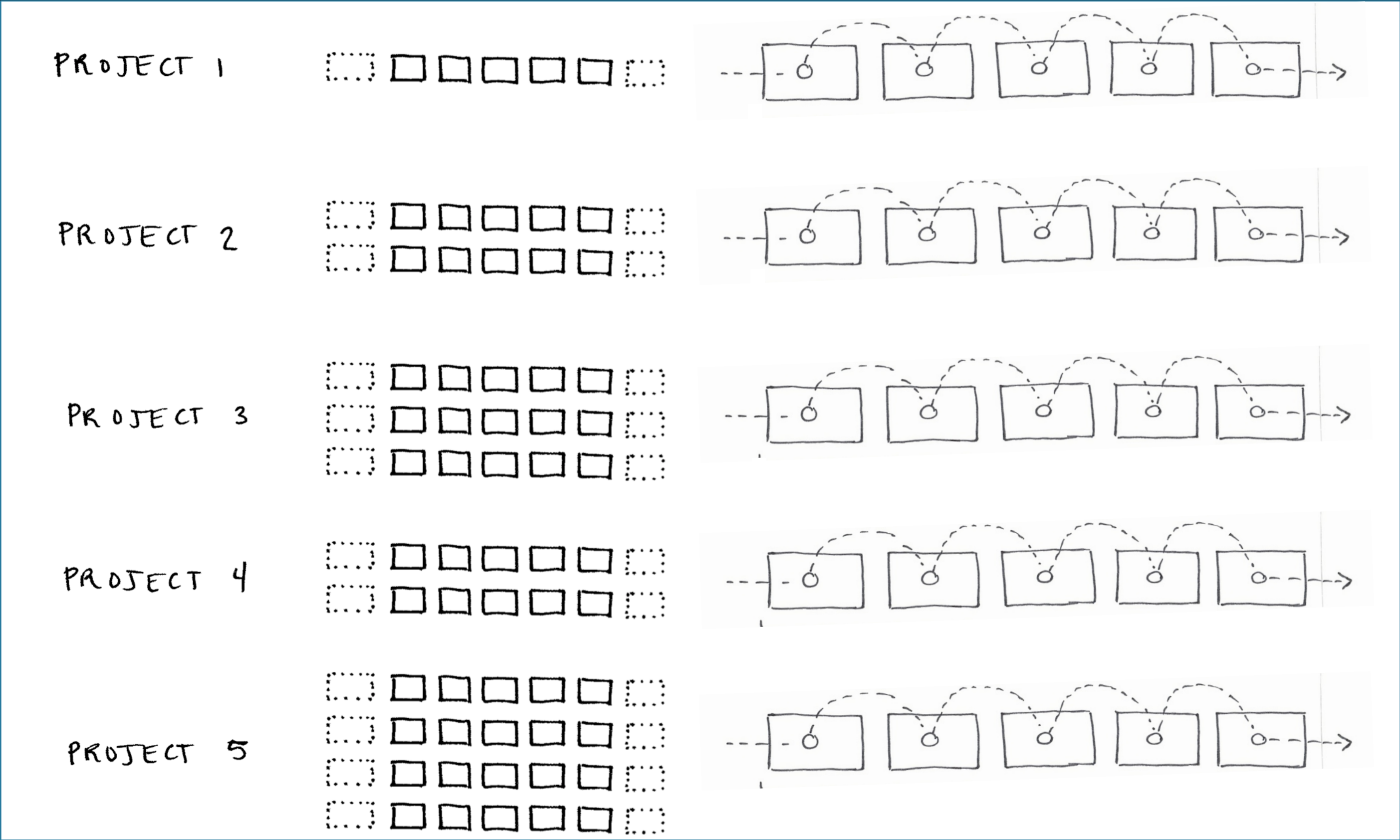
Repetition

Partitioning 12 weeks
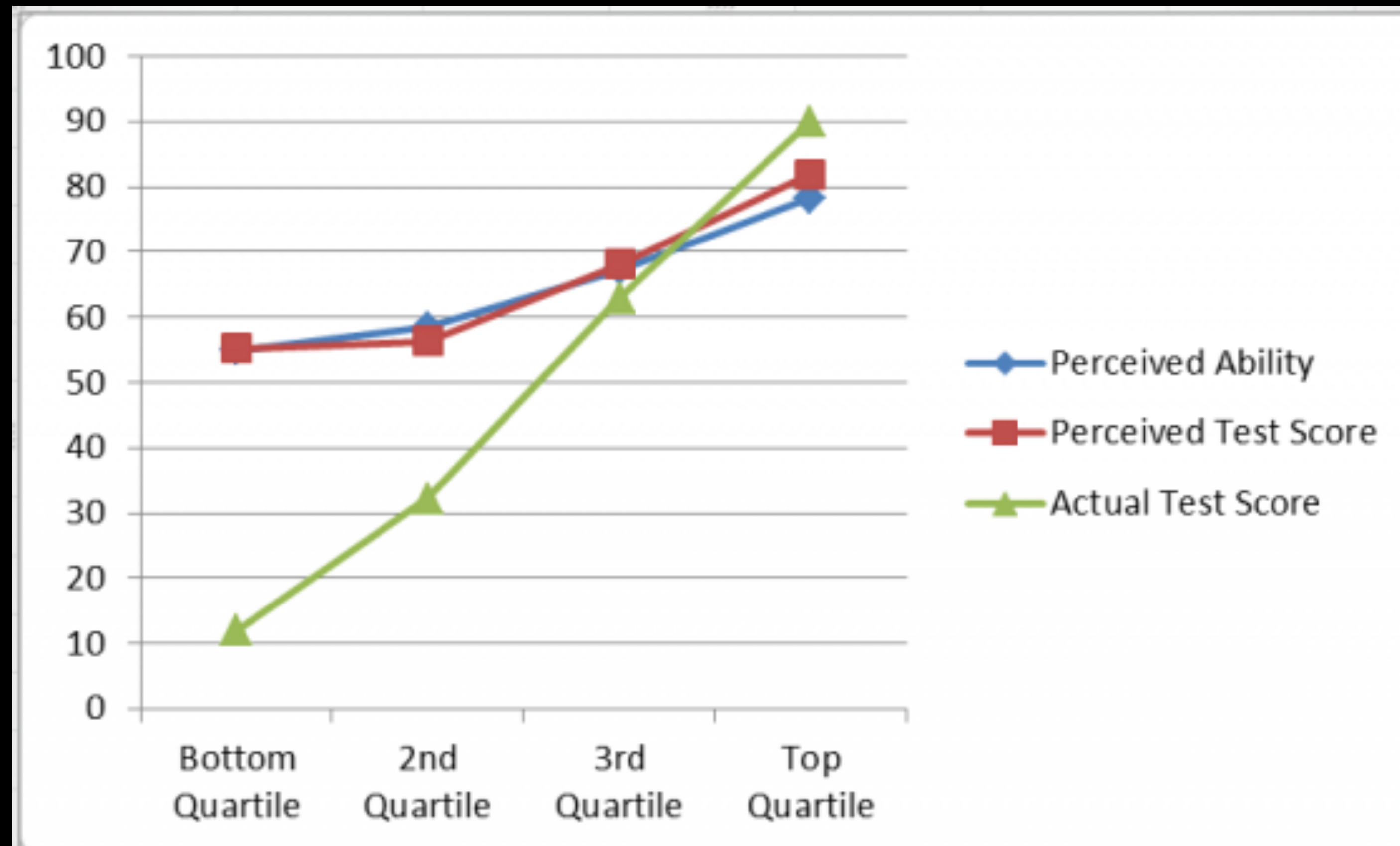
Repetition

Partitioning 12 weeks

Culture & Community

Must overcome:

Impostor syndrome

Perfectionism

Fighting the Dunning-Kruger effect

Culture & Community

Must overcome:
Impostor syndrome
group work
communication

Culture & Community

Must overcome:

Perfectionism
   unfairly short deadlines
   jumping into the unfamiliar
   presenting projects

Culture & Community

Becoming data scientists

Mixture of group and individual work

Guest speakers and meetups

Hiring partners

Career Support

- **One-On-One Meetings**

- **Exposure to industry experts via our Speakers Series**

- **Mock Interviews**

- **Career Day**

- **Post-graduation job support**

@missmegana22

Speaker Series

Some of our previous speakers...

Cassie Kozyrkov
Statistician, Google

Christine Hung
Executive Director, The New York Times

Claudia Perlich
Chief Data Scientist, Dstillery

Dan Valente
Data Scientist, Chartbeat

Dan Liebermann
Associate, Booz Allen Hamilton

Where our graduates are working

Fall bootcamp: September 21, 2015 - December 15, 2015

Early Application Deadline: Wednesday, August 5, 2015
Final Application Deadline: Monday, August 24, 2015

Winter bootcamp: January 11, 2016 – April 1, 2016

Early Application Deadline: Monday, November 16, 2015
Final Application Deadline: Monday, December 7, 2015