APPLIED DATA SCIENCE I

Course Introduction

Kyle Scot Shank, '14
FA-21

01 INTRODUCTIONS

Let's all say hello to one another!

02 SYLLABUS REVIEW

Let's see what's in store for the next ten weeks/

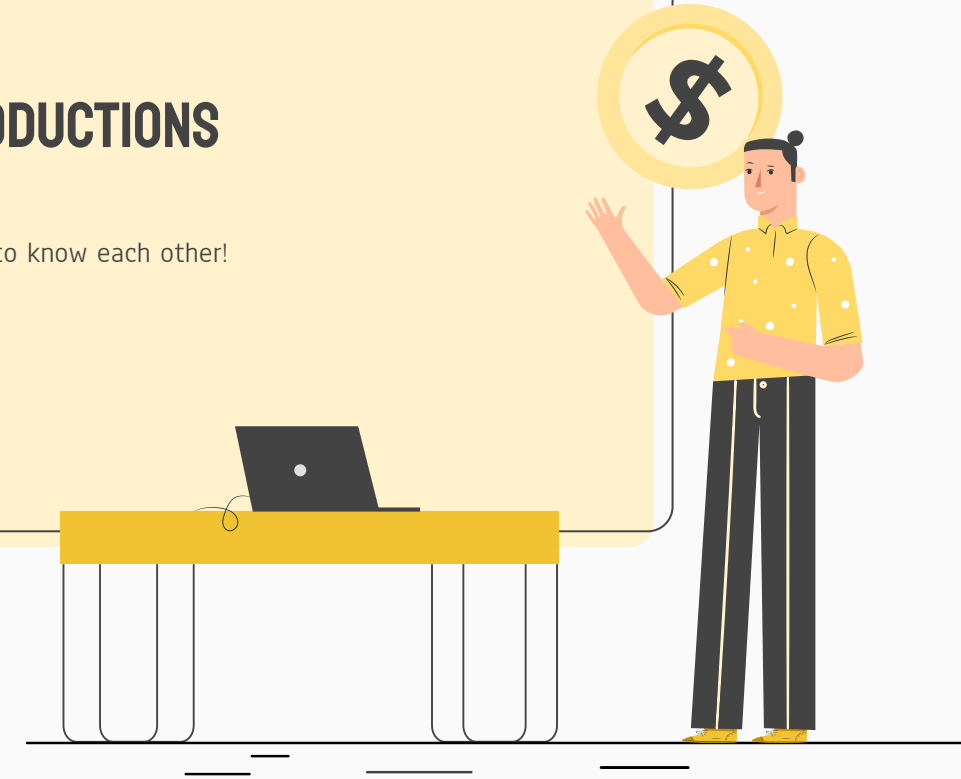
03 WHAT IS DATA SCIENCE?

A brief attempt at answering a very complicated question.

04 Q&A

INTRODUCTIONS

Let's all get to know each other!

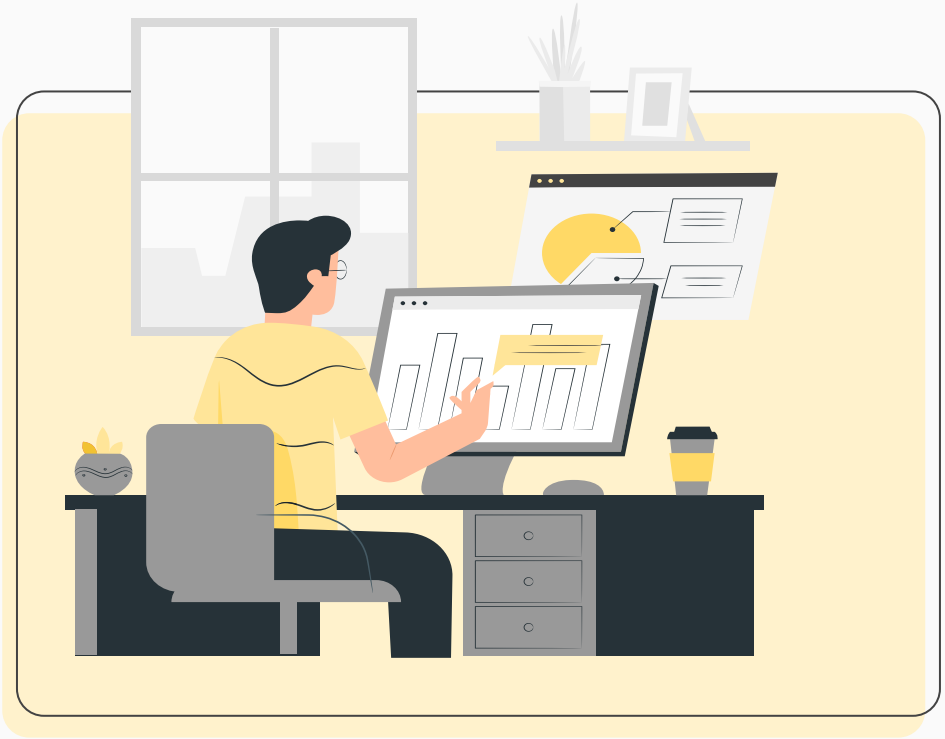


ABOUT KYLE...

- Visiting Faculty Member (this is my first time teaching @ COA!)

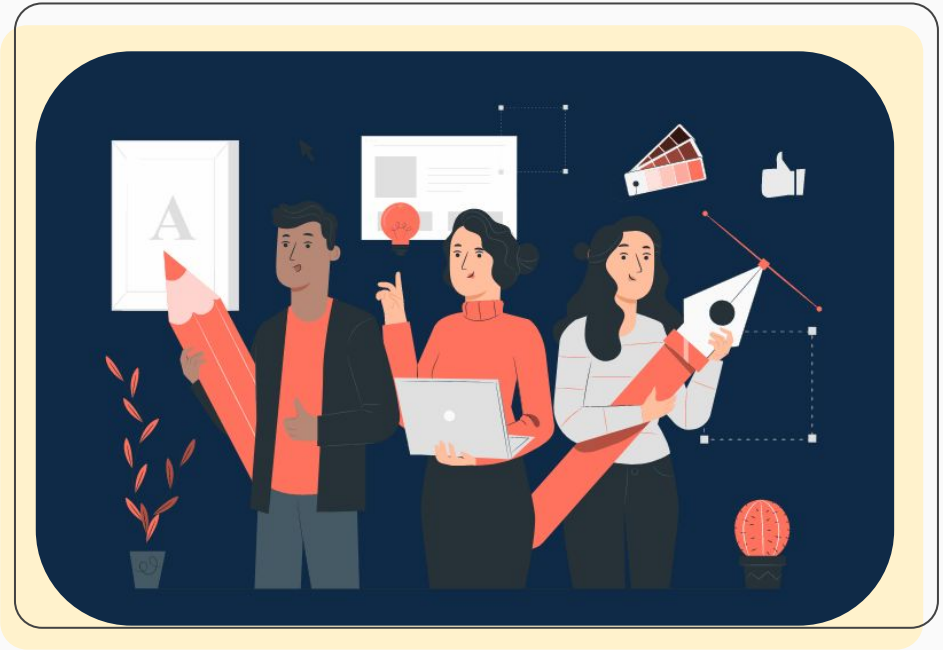
- Currently a quantitative researcher @ FACEBOOK, previously worked in Data Science-y jobs @ The Hershey Corporation, MDI Biological Laboratory, and energy efficiency consulting.

...also bartended in town!



THINGS I RESEARCH...

- How do individuals respond to various different forms of advertising and how can we make those ads more effective?
- How can we quantify and validate metrics related to creativity?
- How can we make housing more equitable and affordable on MDI?
- How can we use offline data from individual, small municipalities to identify broad, state-wide trends?

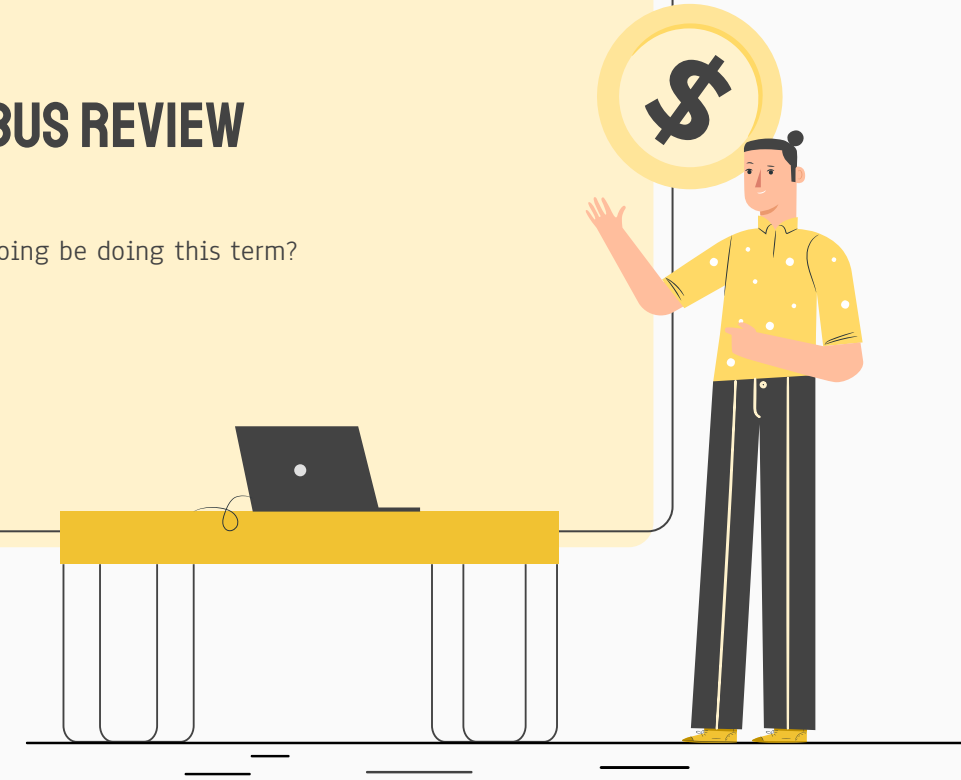


ABOUT YOU!

- What is your name?
- What year are you @ COA?
- What is one excellent thing that happened to you this summer?
- What do you want to learn in this course?

SYLLABUS REVIEW

What are we going be doing this term?



COURSE OBJECTIVES

01
DATA WRANGLING

Gain proficiency in loading, examining, and cleaning data sets in R - both at the command line and via scripting

02
EXPLORATORY DATA ANALYSIS

Plan, implement, and present analysis of data utilizing standard exploratory and statistical mechanics

03
DATA VISUALIZATION

Learn how to clearly and intuitively present data and insights to various audiences

04
DATA ETHICS

Practice in understanding the context of data, as well as identifying various ethical concerns relevant to data science

PROGRAM OBJECTIVES

You are here!

APPLIED DATA SCIENCE I

Main Focuses: Data Manipulation, Aggregation, and Visualization

APPLIED DATA SCIENCE II

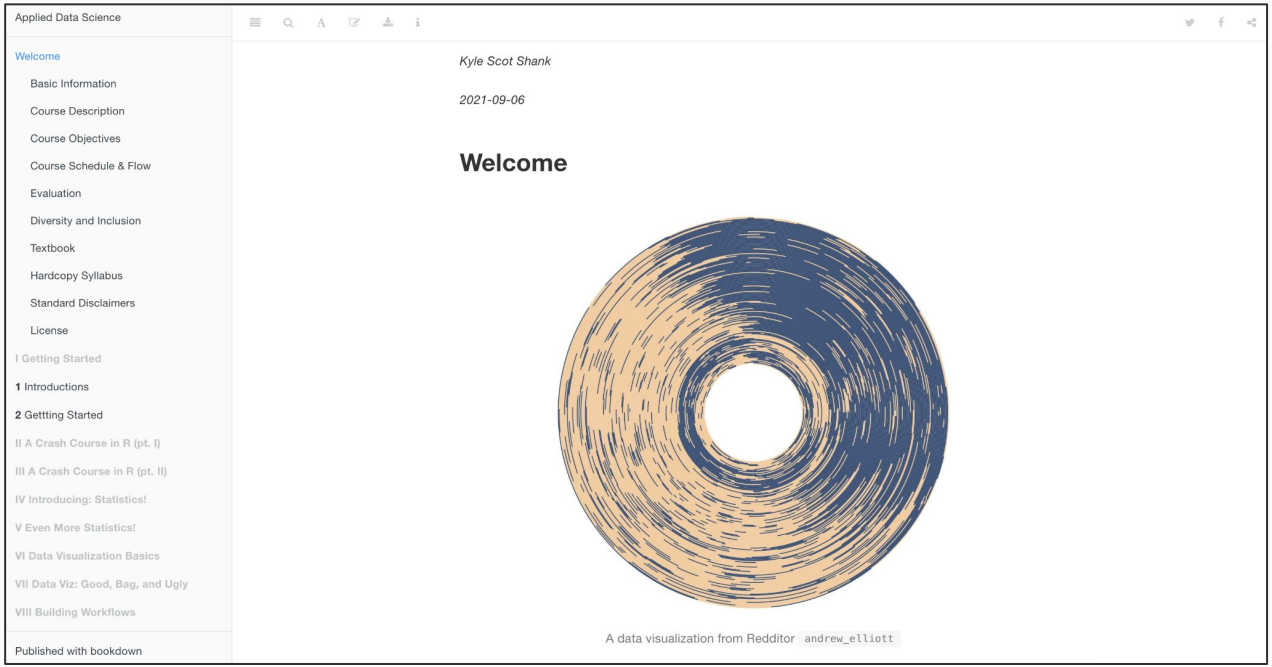
Main Focuses: Data Modeling, Inference, and Prediction

APPLIED DATA SCIENCE III

Main Focuses: Collaboration, Communication

COURSE WEBSITE

https://ads-coa.netlify.app/



Bookmark this site!

This is the syllabus, as well as a list of the readings for a given week, a copy of the lecture slides (if applicable), as well as copies of any files we may create and data sets that we get used.

I will (generally) push the content for the entire week over the preceding week - so you'll be able to browse content early if you like.

ASSIGNMENTS AND POLICIES

CLASS PARTICIPATION

We want to focus on *presence*, *attention*, and *preparation* - not necessarily just actively talking during class.

20% of total course grade

WRITTEN ASSIGNMENTS

We will produce 3 analyses of different data sets over the term. These can be done in a group but submitted individually.

40% of total course grade

FINAL PROJECT

An in-depth analysis of a data set of your choice - either done in a group or individually.

40% of total course grade

ASSIGNMENTS AND POLICIES (CONT'D)

PLAGIARISM

Don't copy other people's work or use other peopl's' work without attribution.

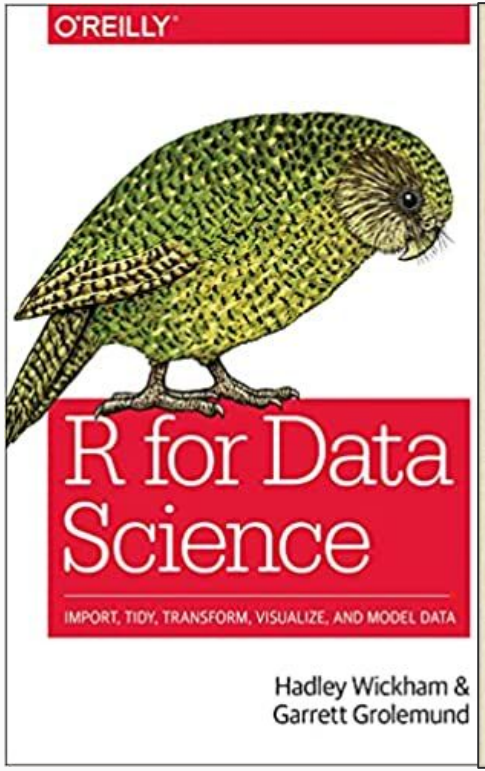
THE INTERNET IS YOUR FRIEND

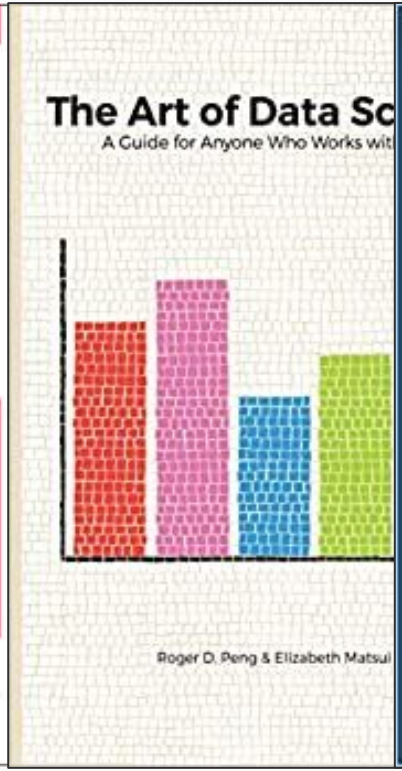
Modern Software Engineering + Data Science *explicitly* uses open source methods to get work done - so don't feel bad about using Google / StackOverflow (but cite your sources!)

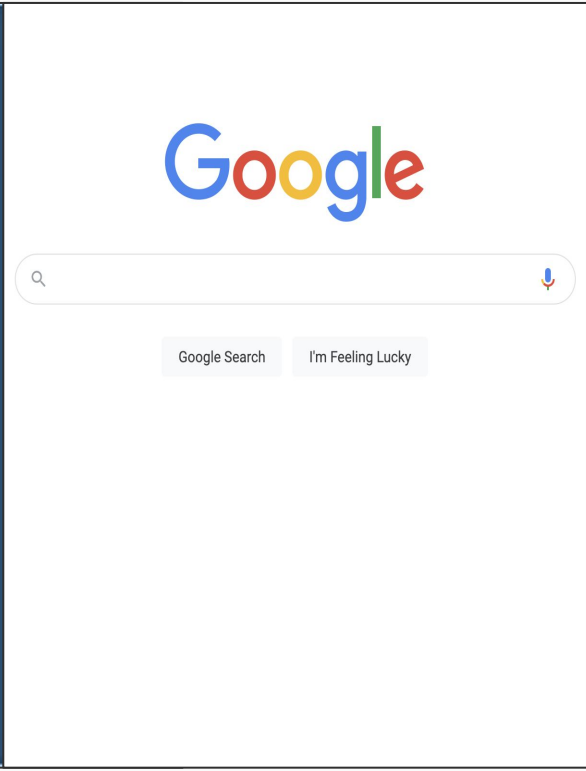
IF YOU'RE FEELING LOST OR STRESSED - REACH OUT

This class might contain a *lot* of new skills being learned all at once - it's okay to feel stressed! Please reach out early with any difficulties and I'll help.

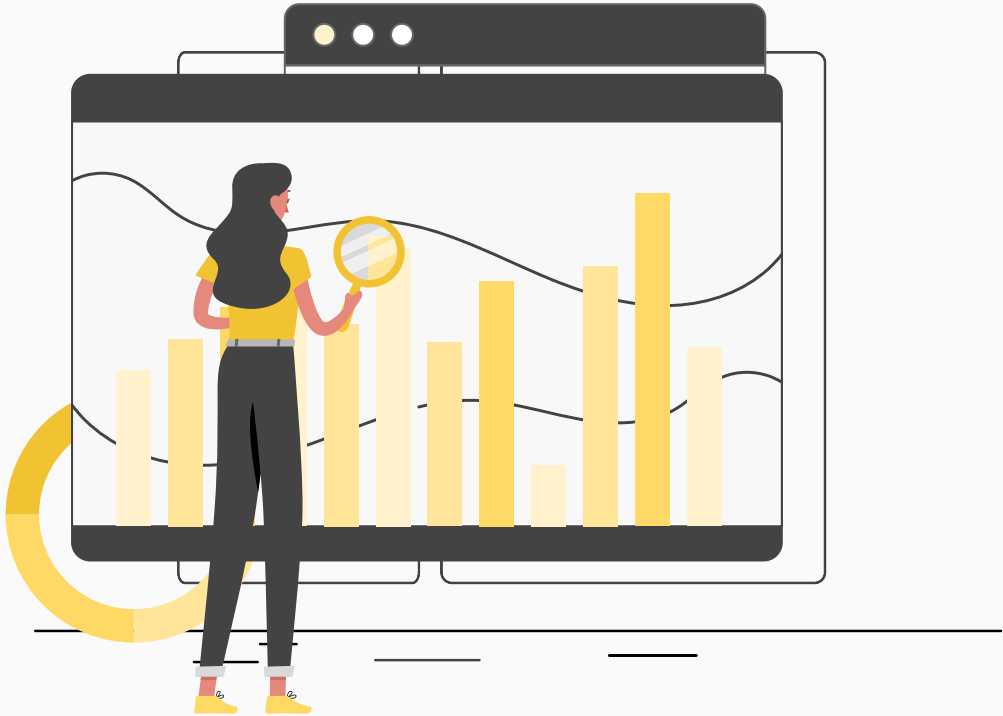
TEXTS + RESOURCES



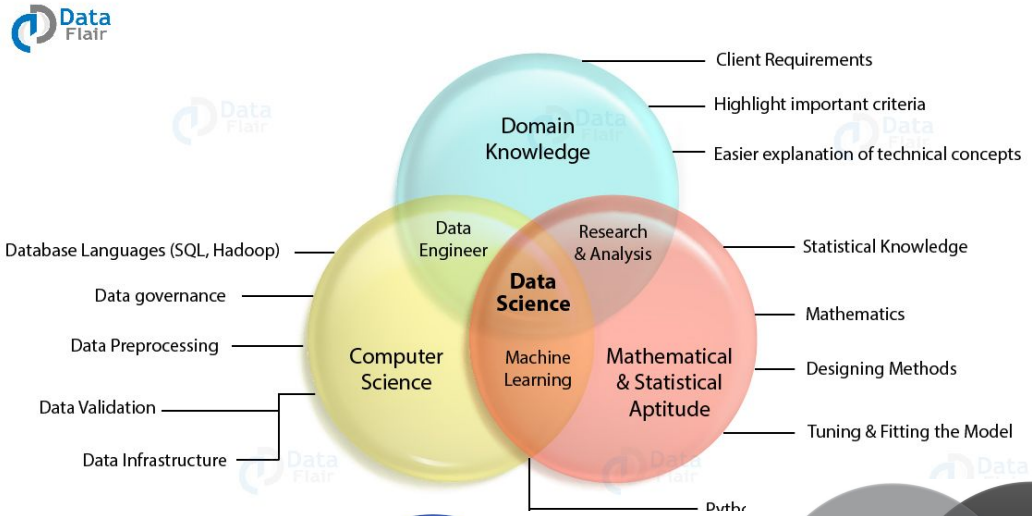


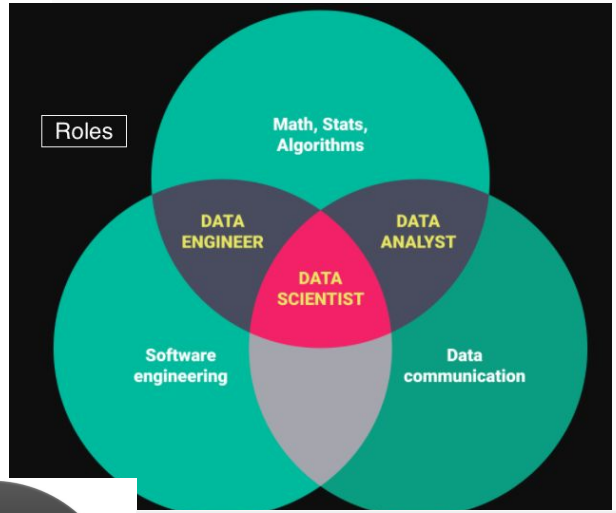


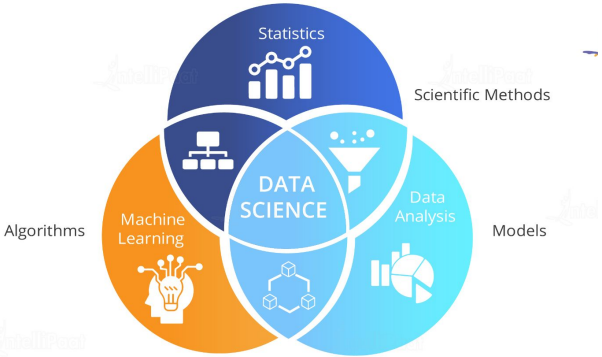
WHAT IS DATA SCIENCE?

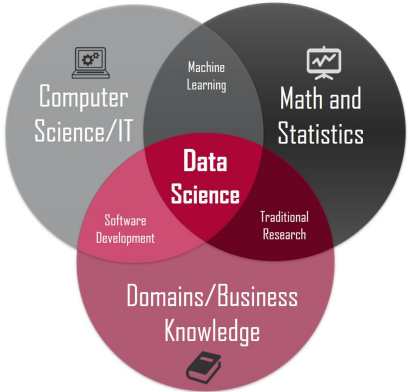


WHAT IS DATA SCIENCE?







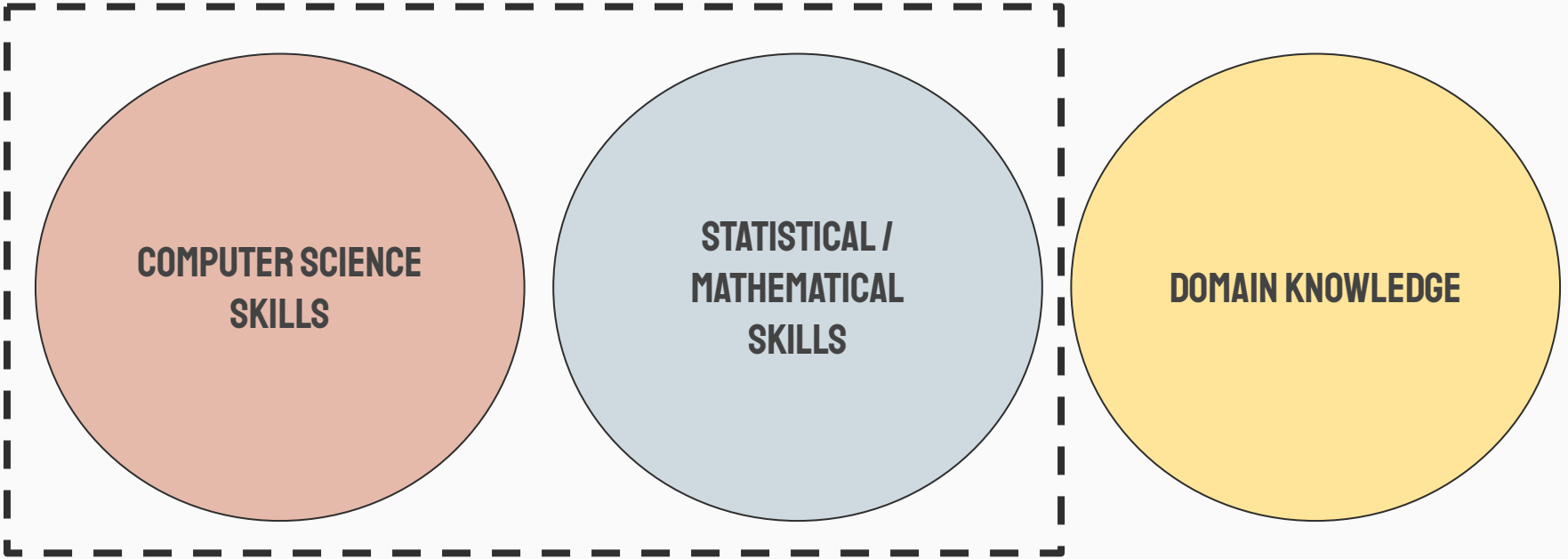


What is Data Science?

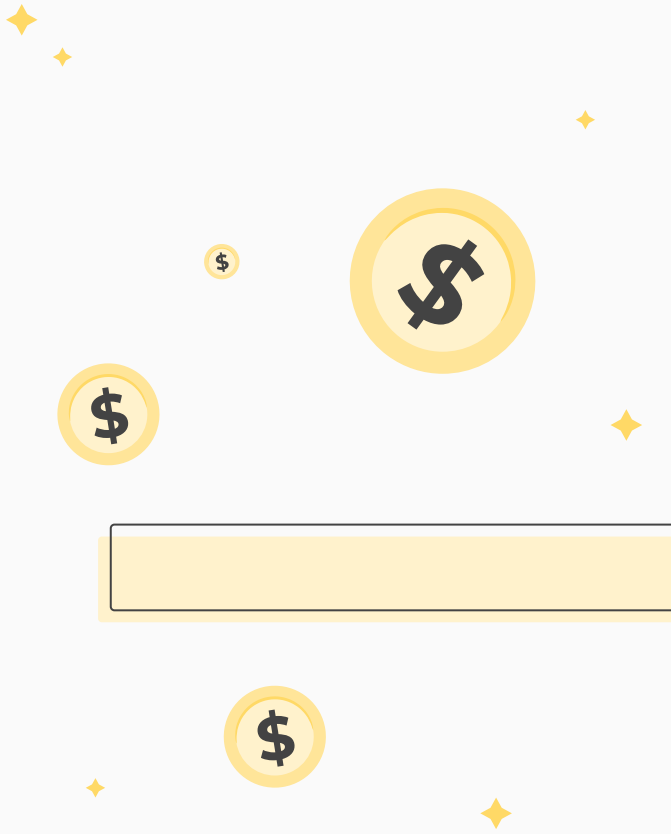
Data Science & its scope in IT industry

What is the difference between Big Data and Data Science?

WHAT IS DATA SCIENCE?



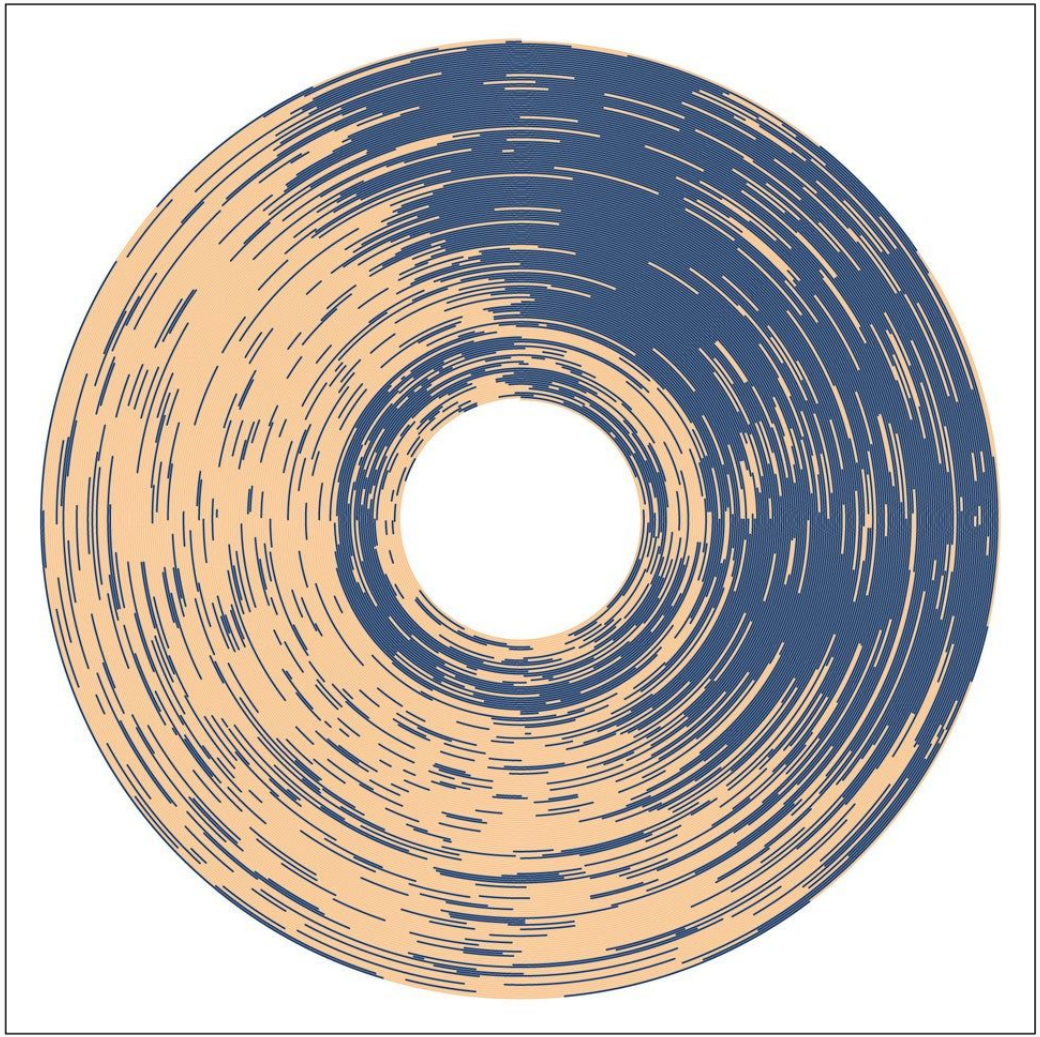
YOU CAN REALLY DO A LOT WITH JUST THIS PART!



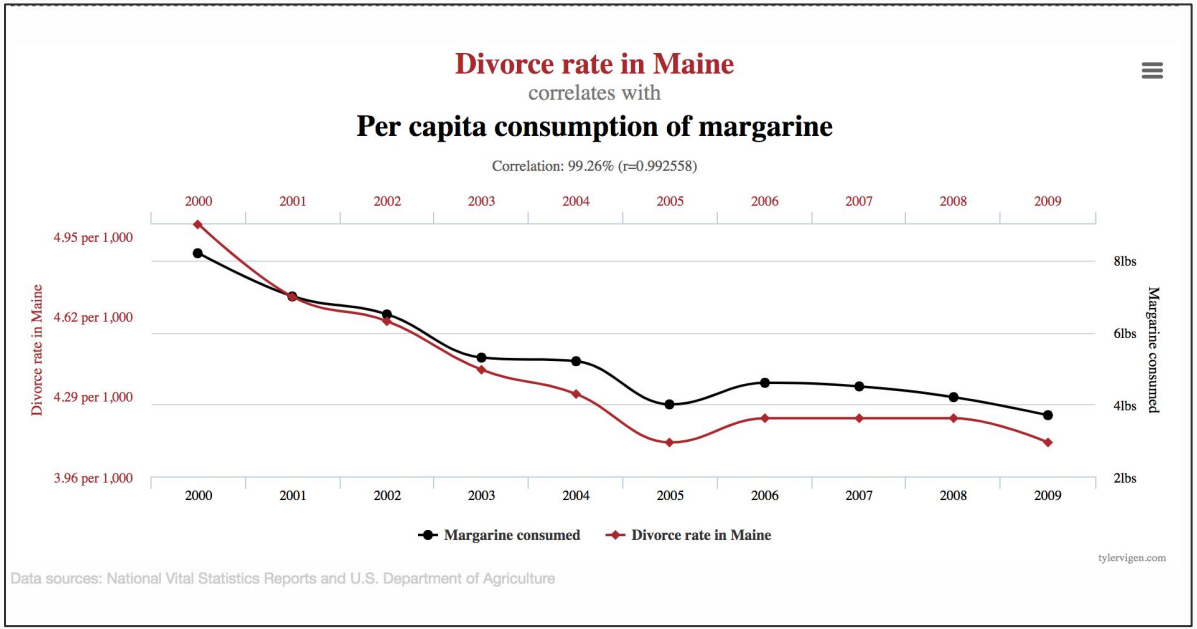
"I like to think of data as the new soil;
get in and get your hands dirty."

—DAVID MCCANDLESS

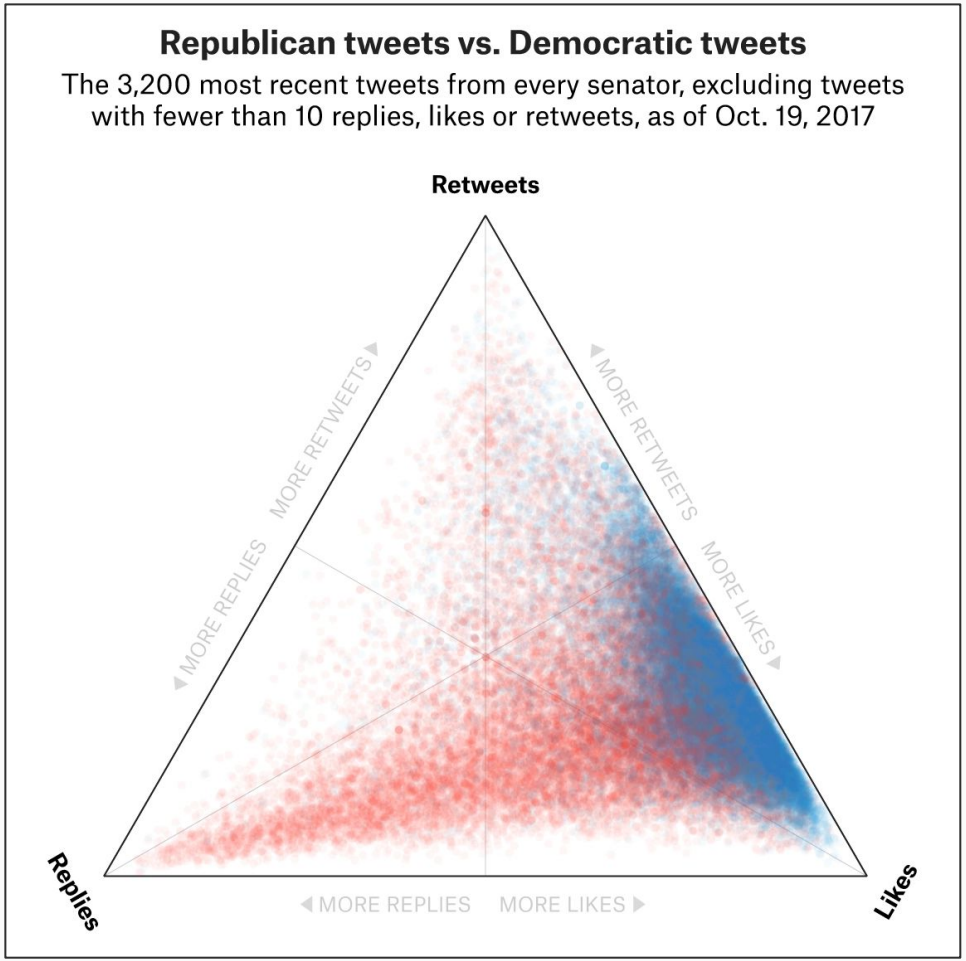
WHAT DO YOU THINK THIS IS VISUALIZING?



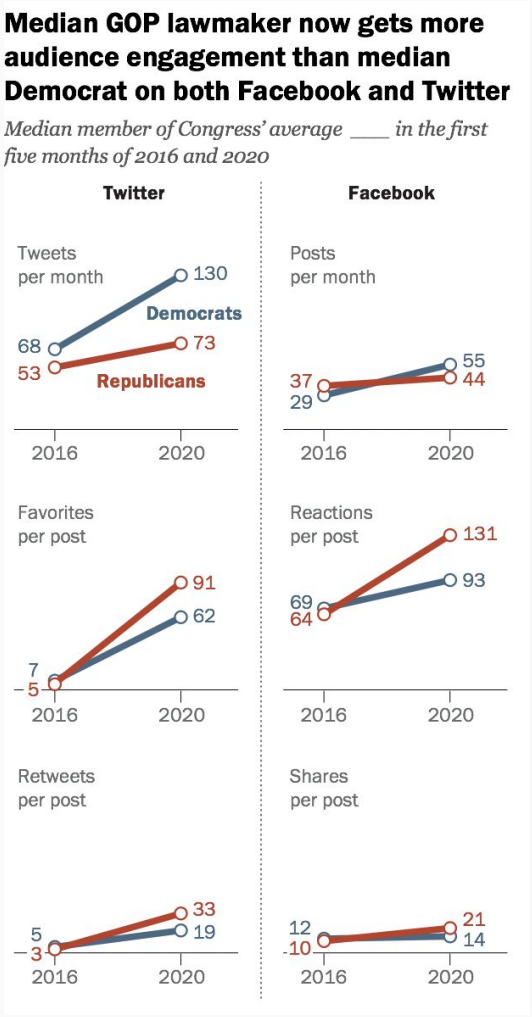
IS THIS INSIGHT ACTUALLY, UH, MEANINGFUL?



HOW CAN THE WAY WE AGGREGATE DATA FRAME THE INSIGHT?



HOW CAN THE WAY WE AGGREGATE DATA FRAME THE INSIGHT?



Q&A