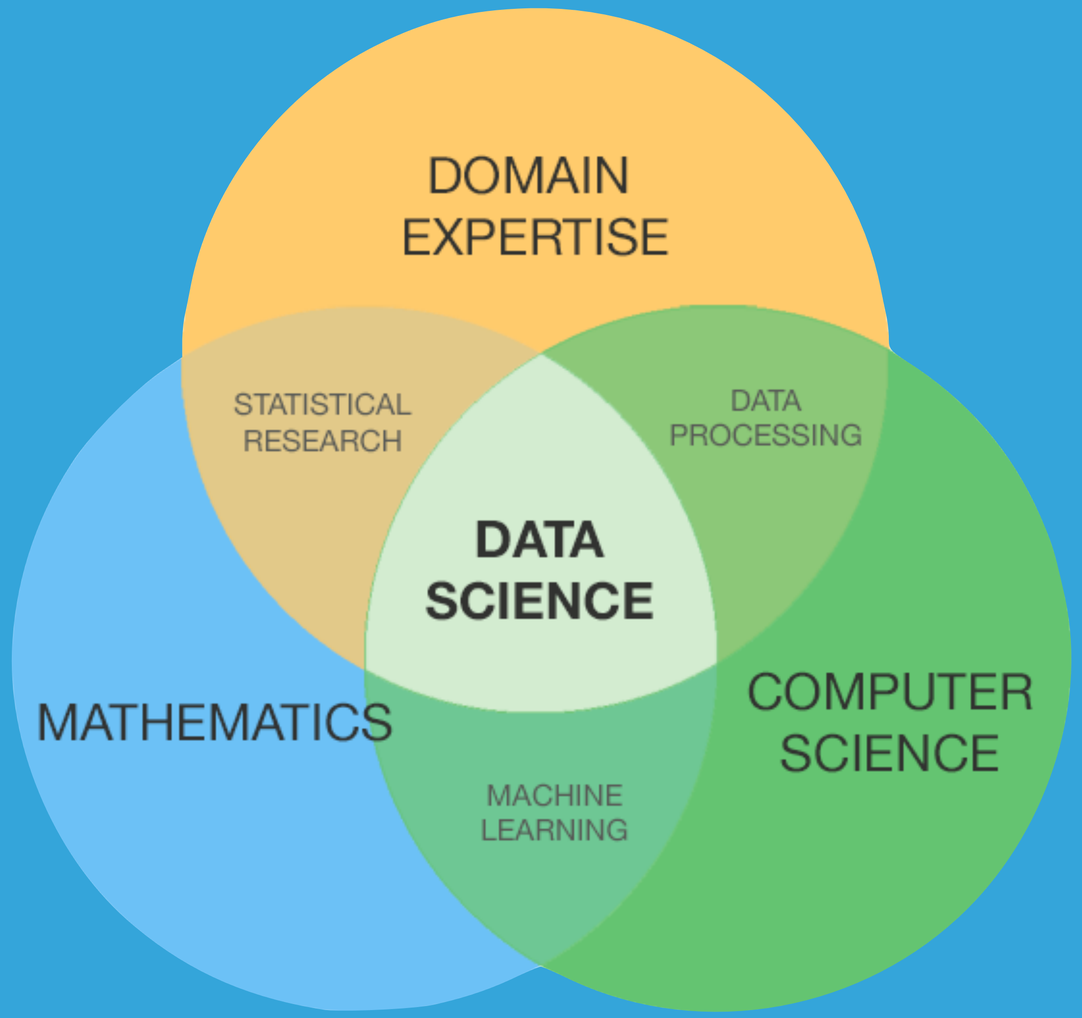
G136 - ENVIRONMENTAL DATA SCIENCE

# WEEK 1

# TODAY'S AGENDA

- Introductions
  - where are you? what are you into? what's keeping you busy these days?
- A quick overview of the class Syllabus & class Slack
- Background Info
- Intro to the Tools we are going to use (Github, JupyterHub)
- Explanation of how we hope the class will work
- Your assignment for the coming week
- Discussion

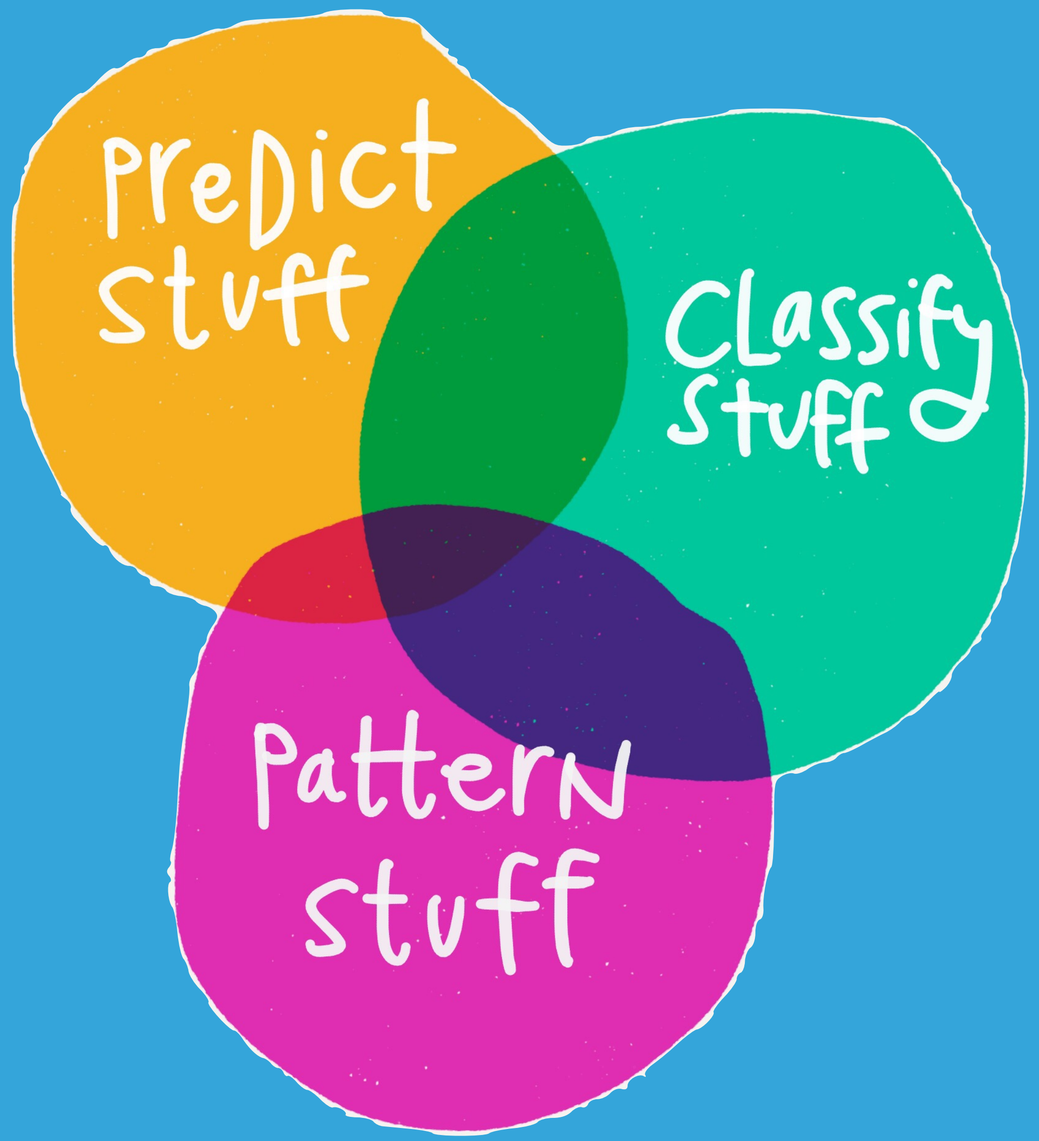
WHAT DATA SCIENCE IS:




"DATA SCIENTIST (N.): PERSON WHO IS BETTER AT STATISTICS THAN ANY SOFTWARE ENGINEER AND BETTER AT SOFTWARE ENGINEERING THAN ANY STATISTICIAN."

Josh Wills, Director of Data Engineering at Slack

WHAT DATA SCIENCE DOES:

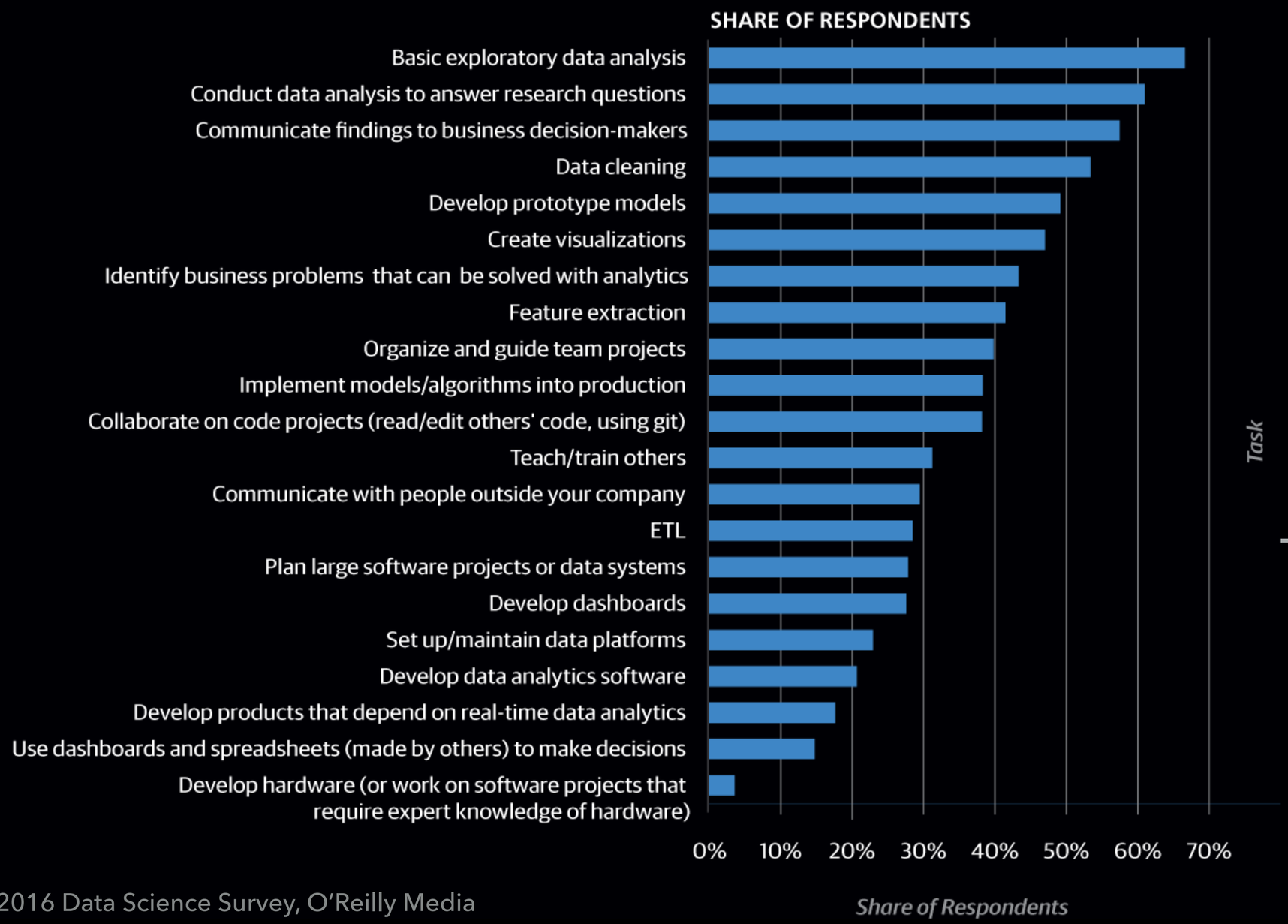



"ENVIRONMENTAL DATA SCIENTIST (N.): PERSON WHO IS BETTER AT COMPUTATIONS THAN ANY APPLIED ENVIRONMENTAL SCIENTIST AND BETTER AT ENVIRONMENTAL PROBLEM SOLVING THAN ANY COMPUTER SCIENTIST."

Me, Your Instructor

**TASKS** (RESPONDENTS COUNTED IF THEY SAID THEY HAVE "MAJOR INVOLVEMENT" IN THIS TASK)

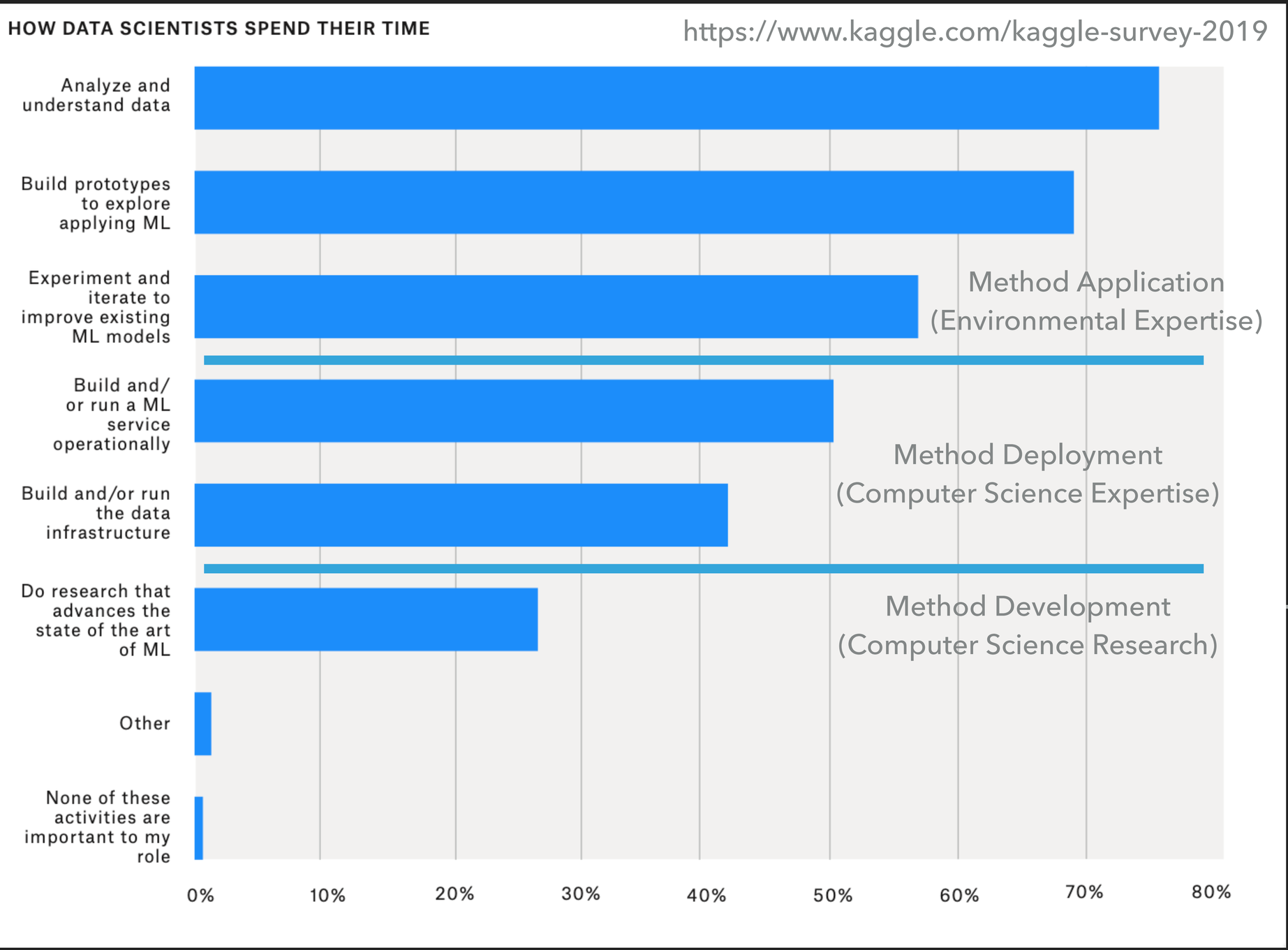
SHARE OF RESPONDENTS




# DATA SCIENCE

# WHAT IS IT?

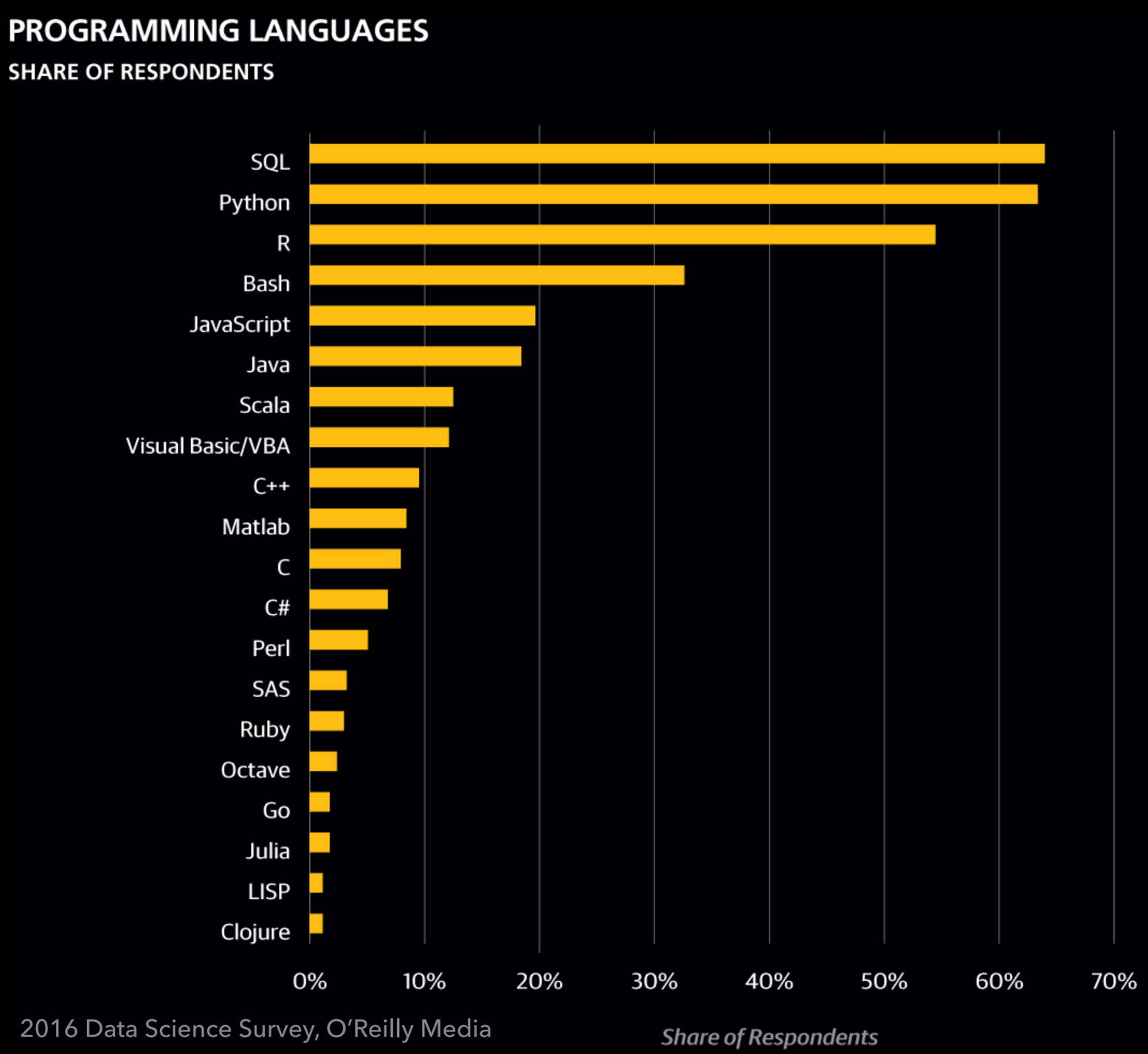
2016 Data Science Survey, O'Reilly Media




The “ground floor” of data science  is domain expertise. Environmental data science requires knowledge of the environment.

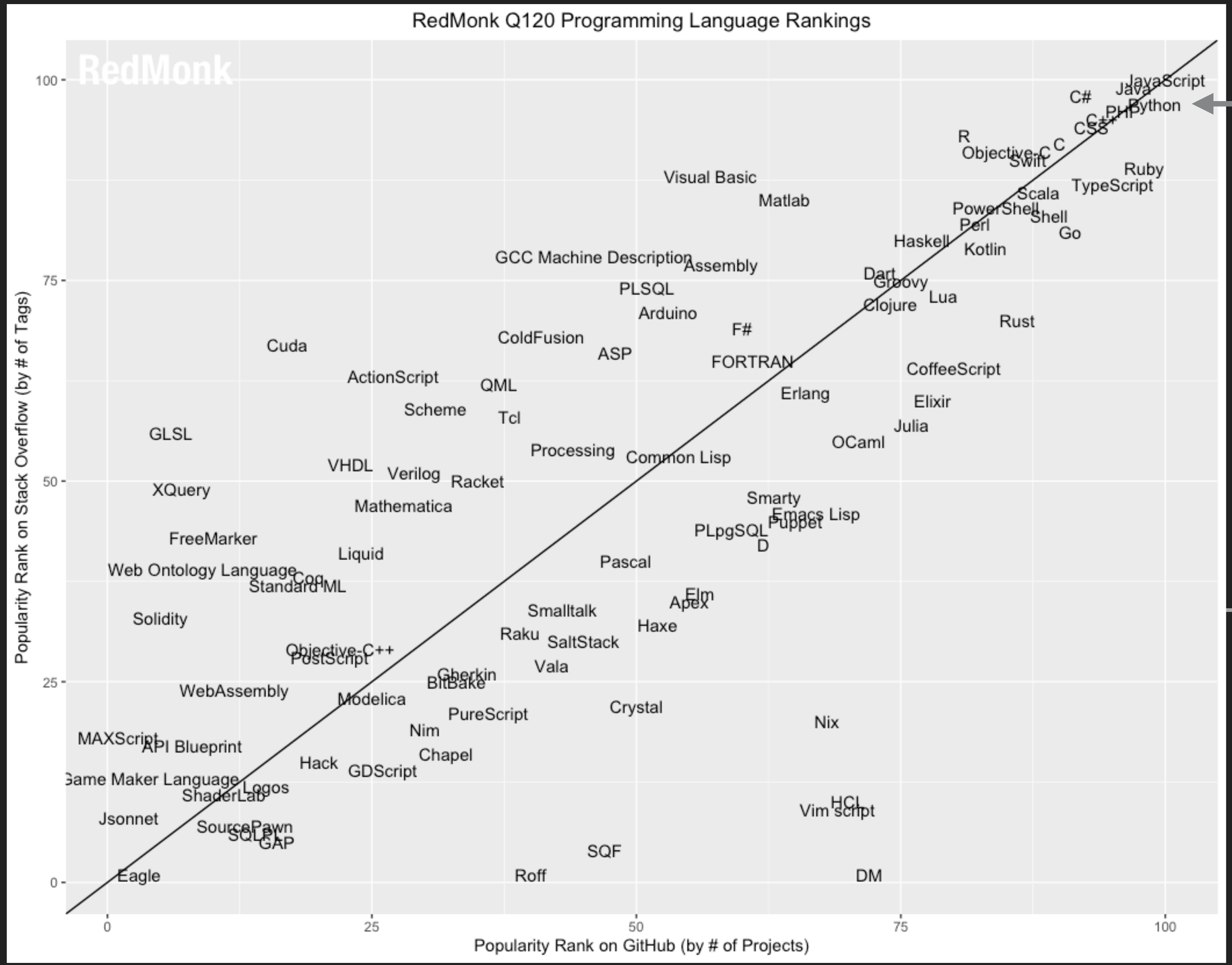
# DATA SCIENCE

# WHAT IS IT?




# WHY PYTHON?

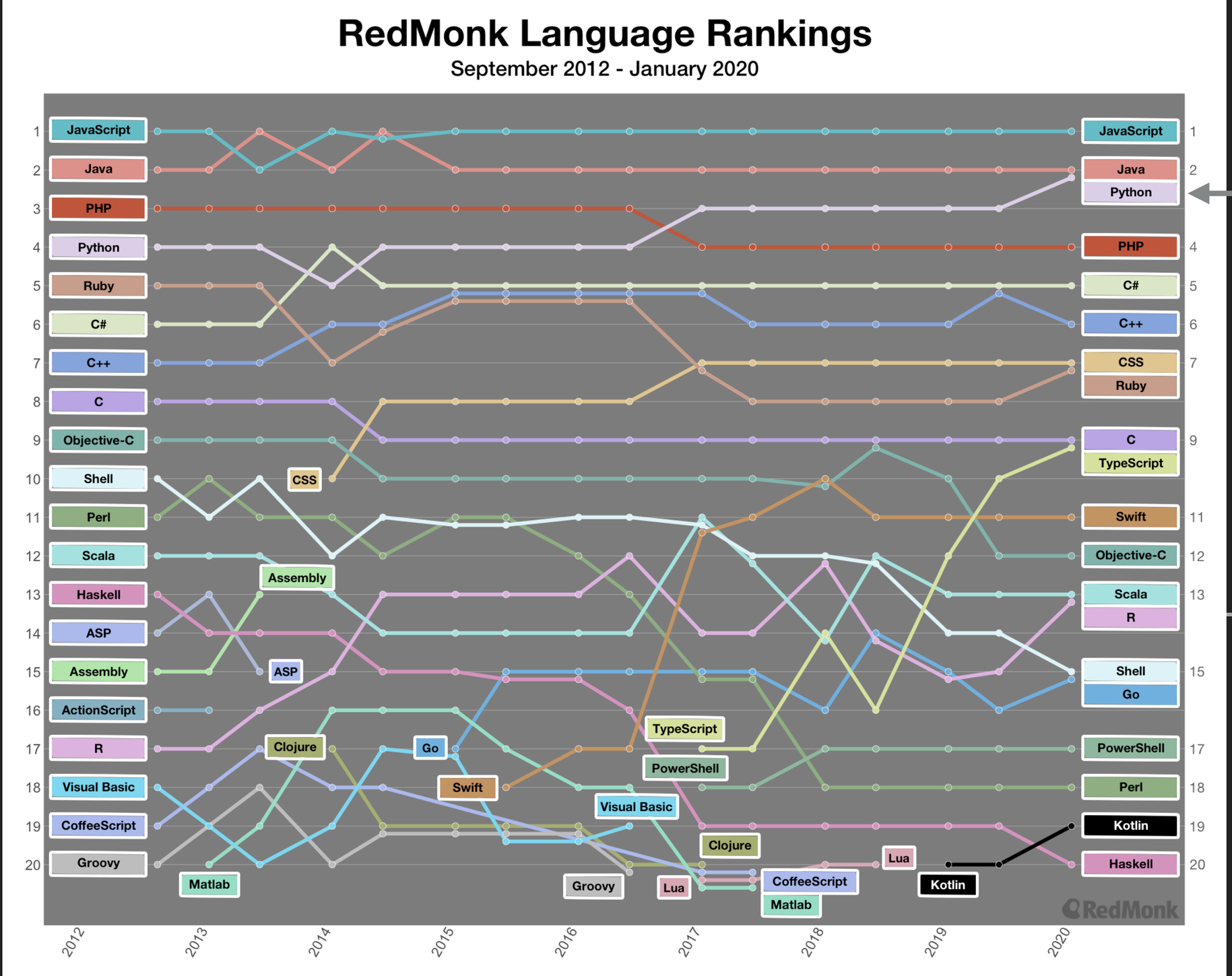
# WHAT TOOLS DO DATA SCIENTISTS USE?




Python

# WHY PYTHON?

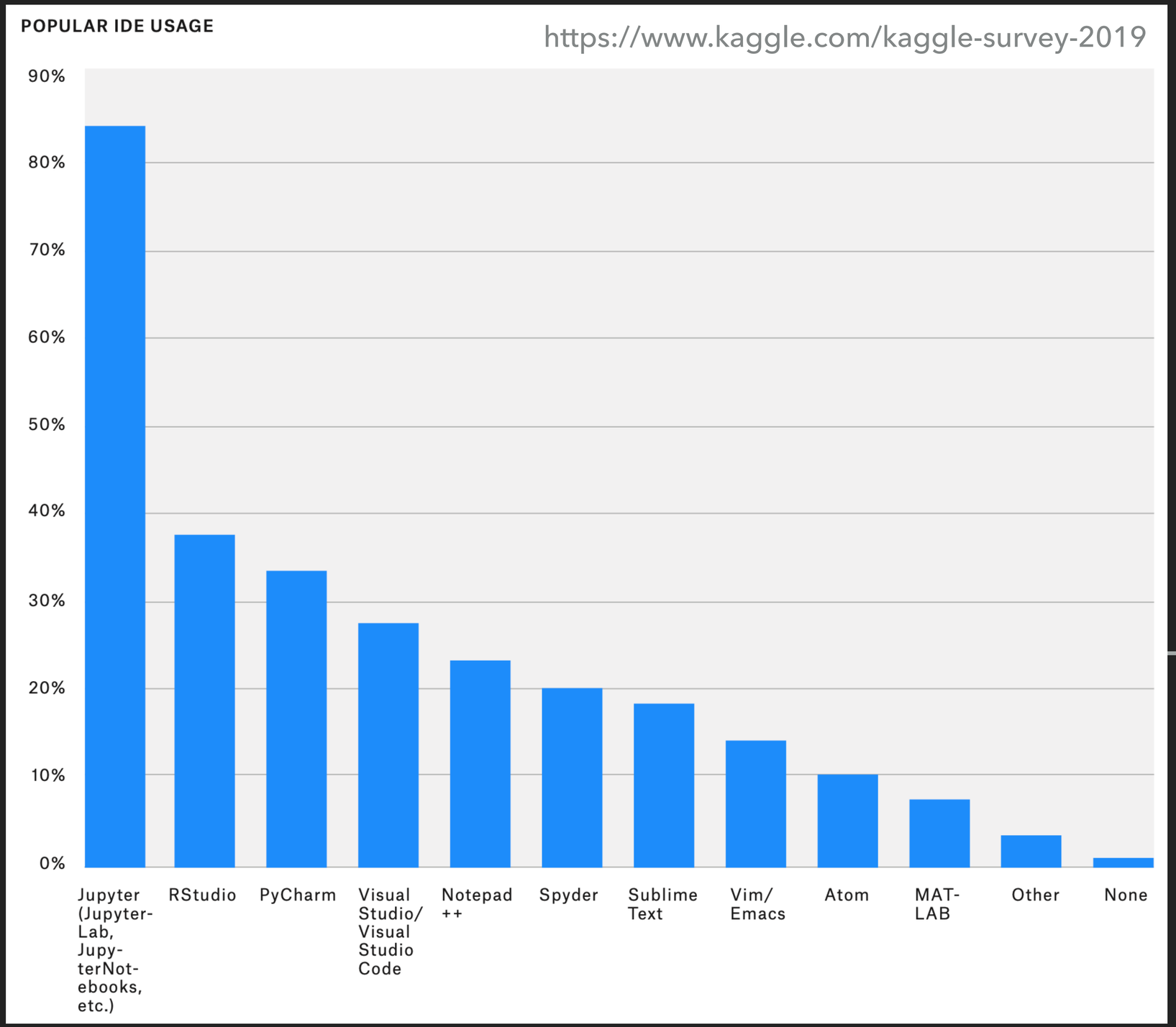
# WHAT LANGUAGES DO PROGRAMMERS USE?




Python (#2, tied w/Java)

# WHY PYTHON?

# WHAT LANGUAGES DO PROGRAMMERS USE?

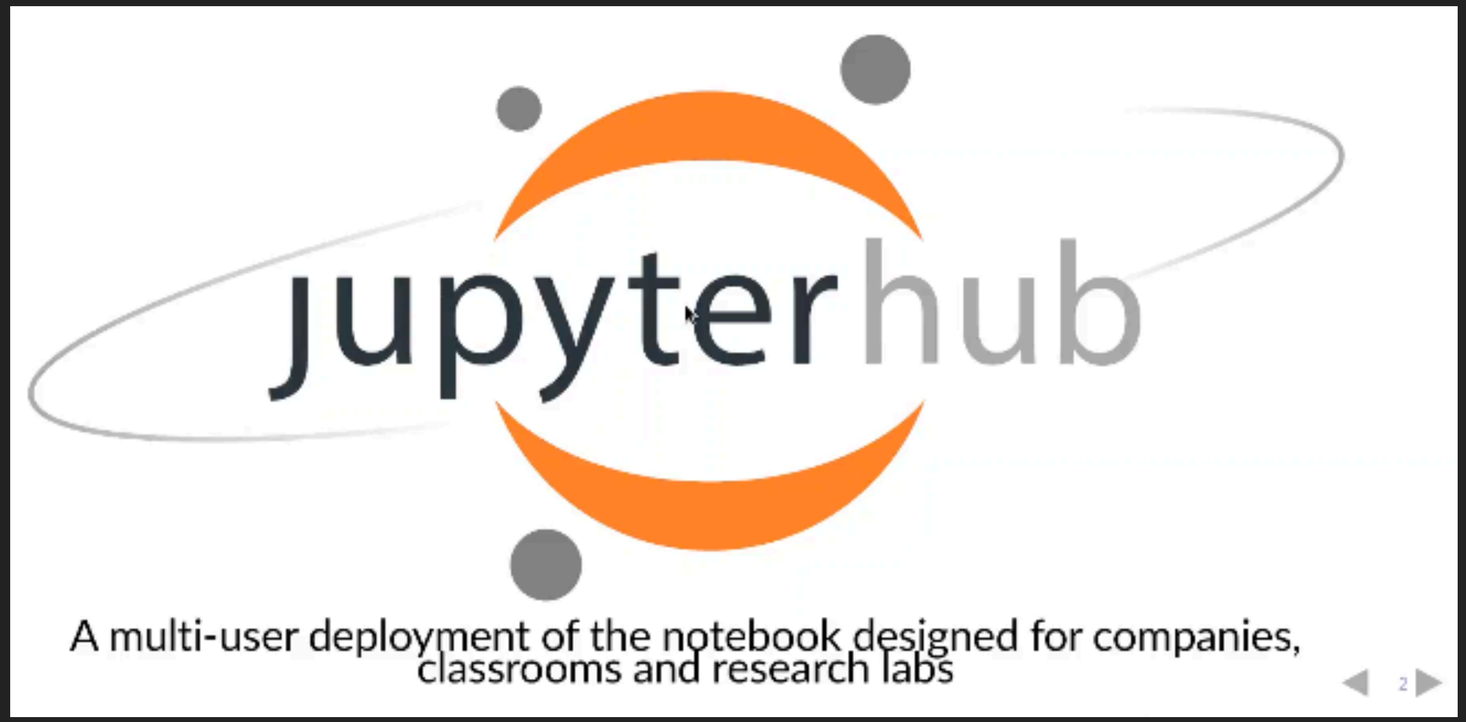



# WHY PYTHON?

# WHAT TOOLS DO DATA SCIENTISTS USE?

# INTRO TO THE TOOLS WE ARE GOING TO USE (GITHUB, JUPYTERHUB)

- https://g136.lsit.ucsb.edu/hub/login
- Getting our class repository:
  - https://github.com/environmental-data-science/envdatasci
- Adding class repository to your Jupyter Instance
  - `git clone`
  - `pip install`




## HOW PYTHON?

# JUPYTERHUB

Course JupyterHub Server: https://g136.lsit.ucsb.edu/

Github Organization: https://github.com/environmental-data-science

Class Repository: https://github.com/environmental-data-science/envdatasci

## COURSE MATERIALS

# GITHUB & GIT

Log into your individual JupyterLab instance

"Pull" Course Materials

Shared Class Repository

GitHub

"Push" Course Your Assignments

Your Personal Course Repository

GitHub

# HOW G136 WILL WORK

# YOUR LAPTOP/PC + GOOGLE CLOUD + GITHUB

Log into your individual JupyterLab instance

jupyterhub

A multi-user deployment of the notebook designed for companies, classrooms and research labs

"Pull" Course Materials

Shared Class Repository

GitHub

"Push" Course Your Assignments

Your Personal Course Repository

GitHub

SEND US YOUR GITHUB USERNAME!

# HOW G136 WILL WORK

# YOUR LAPTOP/PC + GOOGLE CLOUD + GITHUB