

WHY ARE LESS WOMEN IN COMPUTER SCIENCE?

GA - Data Science Final Project

By,

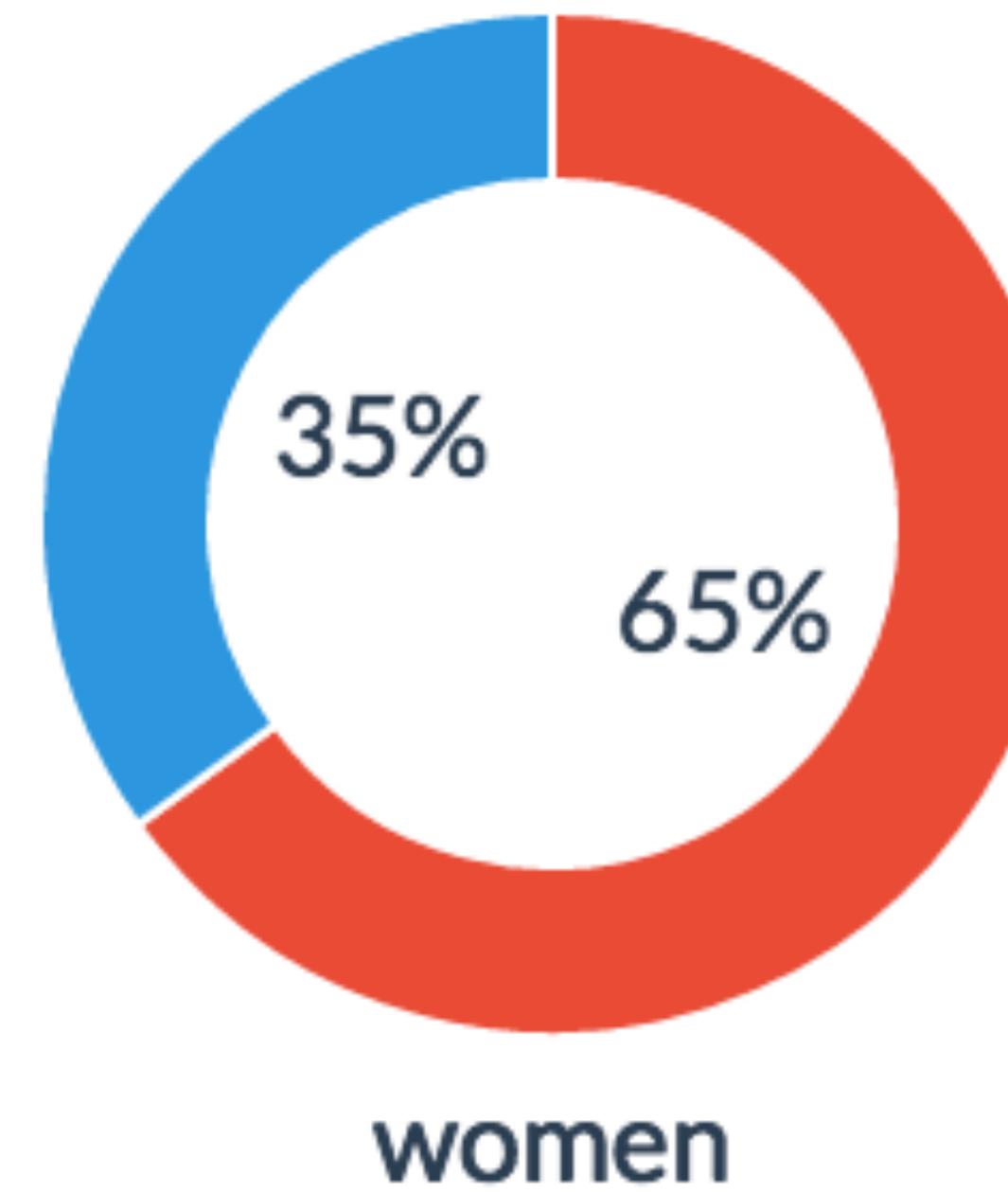
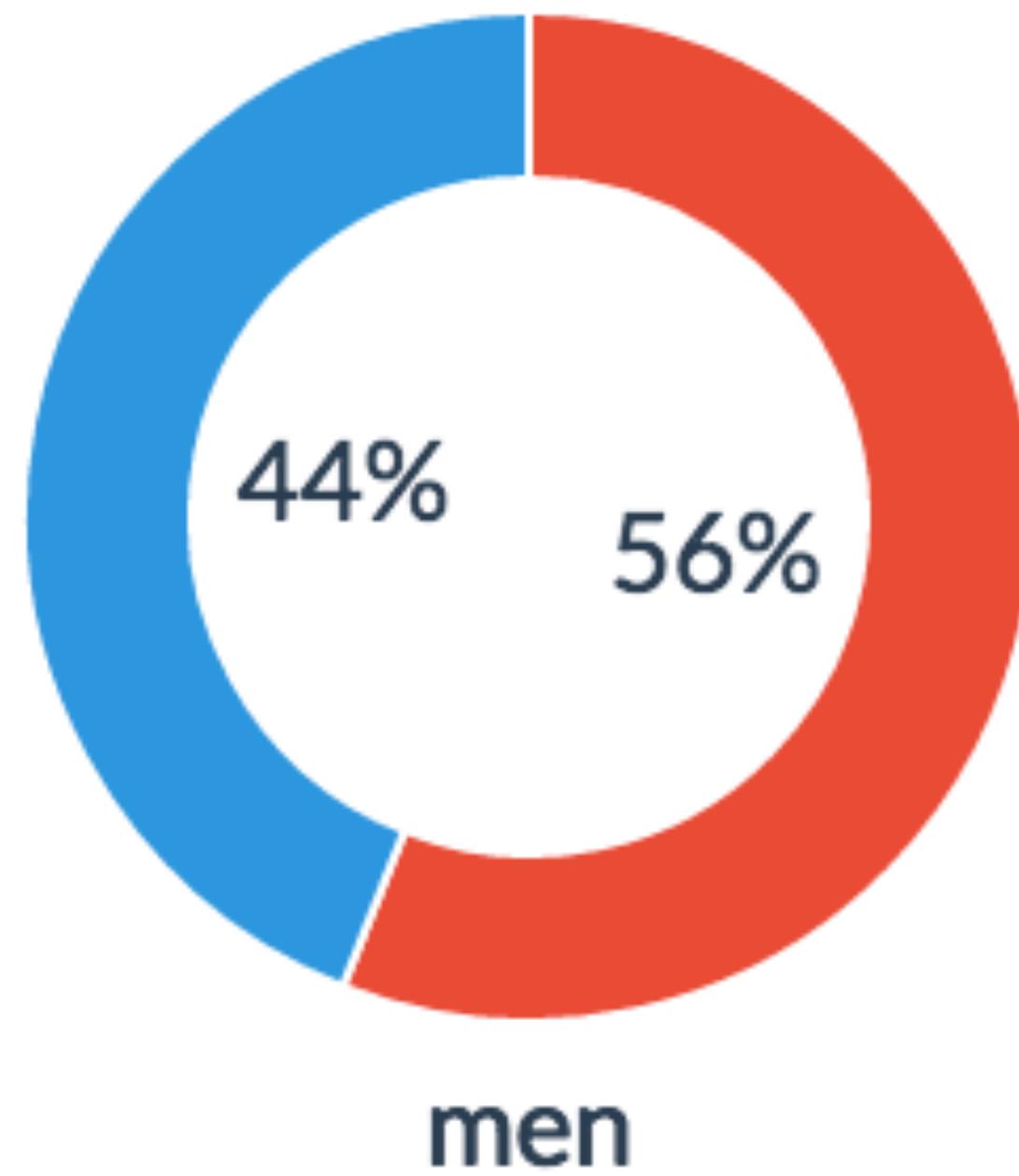
Sana

WHY THIS TOPIC??

- Girls outperformed boys in more countries in a science test given to 15-year-old students in 65 countries – but in the United States, boys led the girls.
- What is most startling about that is it does not represent progress. In 1985, women earned **37%** of computer-science undergraduate degrees but the number slowly started to decrease.
- Three decades later, STEM careers has become a much more vital gateway to high-paying jobs and chance to influence the software-driven future of society. Yet mostly more men than women are stepping through it.

LITTLE STATS:

- Women earn just **18%** of undergraduate degrees awarded for computer science.
- 30,000 students took the Advanced Placement Computer science exam in high school last year - 2014. Less than 6,000 of them were women.



accepted
rejected

A legend consisting of two colored squares: blue for 'accepted' and red for 'rejected'.

ASK RESEARCH QUESTIONS

This is what made me think about it

There were three main data files:

- all-ages.csv
- Recent Grad.csv (<28yrs) - Contains detailed breakdown by gender and type of job
- grad-students.csv (ages 25 +) - Contains basic earnings, labor force information, unemployment rate and details on graduate attendees.
- women-stem.csv - Contains total of Women enrolled in all different majors.
- Extracted from Census data (<http://www.census.gov/programs-surveys/acs/technical-documentation/pums.html>)

Header	Description
Rank	Rank by median earnings
Major_code	Major code, FO1DP in ACS PUMS
Major	Major description
Major_category	Category of major from Carnevale et al
Total	Total number of people with major
Sample_size	Sample size (unweighted) of full-time, year-round ONLY (used for earnings)
Men	Male graduates
Women	Female graduates
ShareWomen	Women as share of total
Employed	Number employed (ESR == 1 or 2)
Full_time	Employed 35 hours or more
Part_time	Employed less than 35 hours
Full_time_year_round	Employed at least 50 weeks (WKW == 1) and at least 35 hours (WKHP >= 35)
Unemployed	Number unemployed (ESR == 3)
Unemployment_rate	Unemployed / (Unemployed + Employed)
Median	Median earnings of full-time, year-round workers
P25th	25th percentile of earnings
P75th	75th percentile of earnings



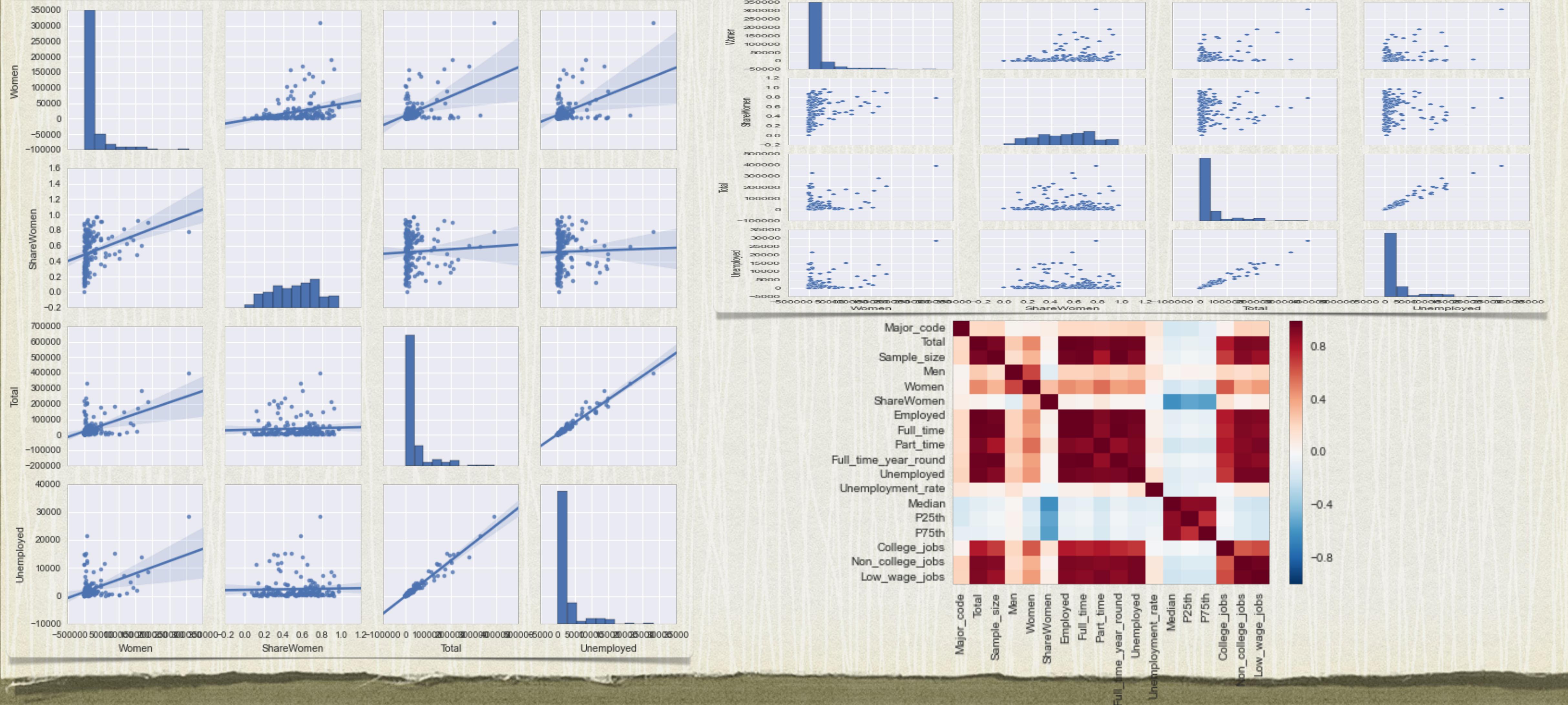
think

PHASES



- **Phase 1:** Data Collection- Exploring the data, cleaning and analyzing what could be done with the data
- **Phase 2:** Modeling Plan
 - Linear Regression
- **Phase 3:** Building my *Test Cases*.
 - Why are there less Women who sign up for Engineering major?
 - Is there a correlation between unemployment rate and Women count?





VISUALIZING LINEAR RELATIONSHIP

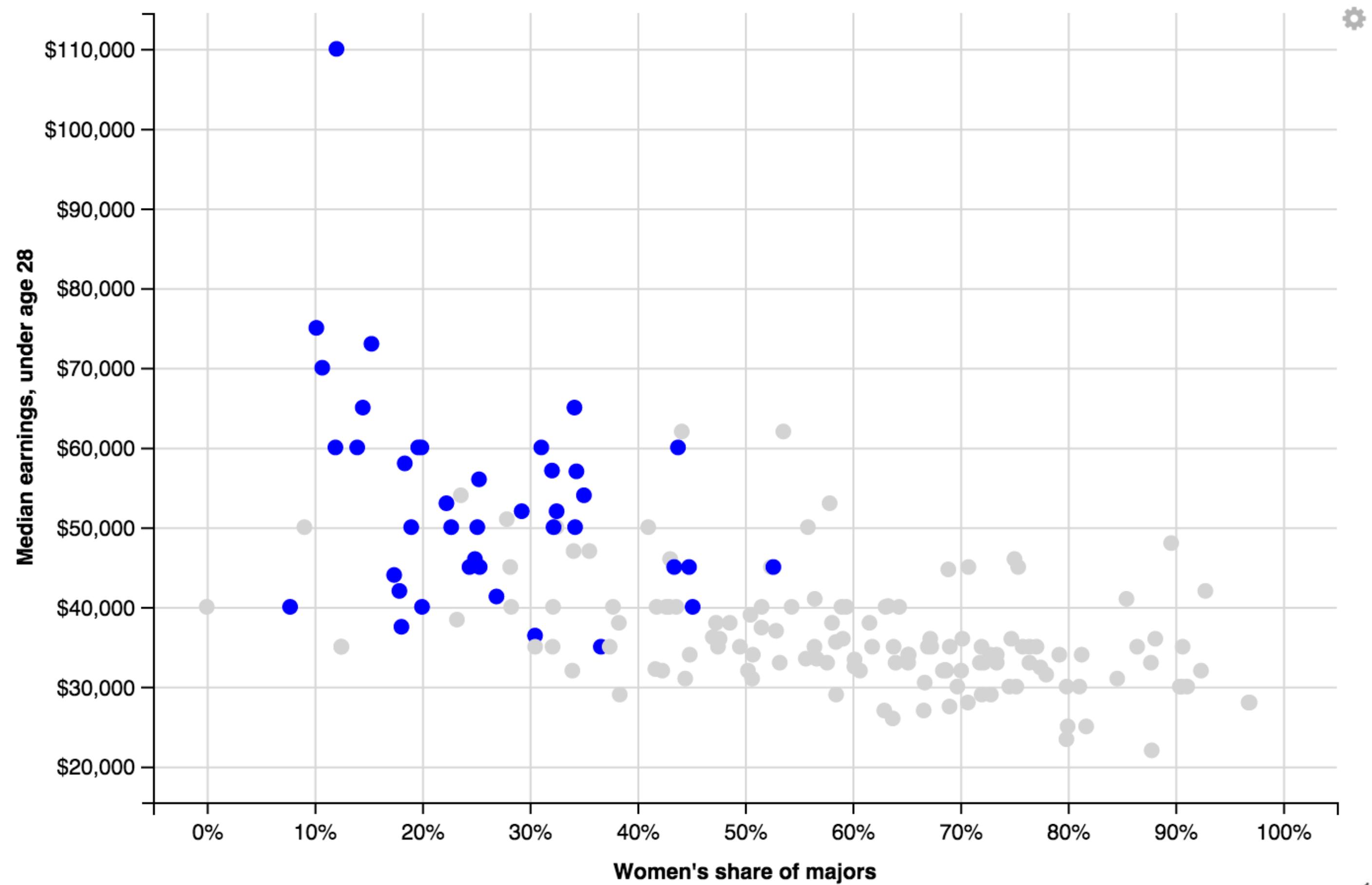
Using Seaborn

Select a category of major:

- All
- Arts, Social Sciences & Humanities
- Business, Law & Communications
- Engineering & Technology
- Health, Education & Social Work
- Natural Sciences
- Other

Women Opt For Lower-Paying Majors

What Majors Pay Vs. Women's Share Of Students



LESSONS LEARNT

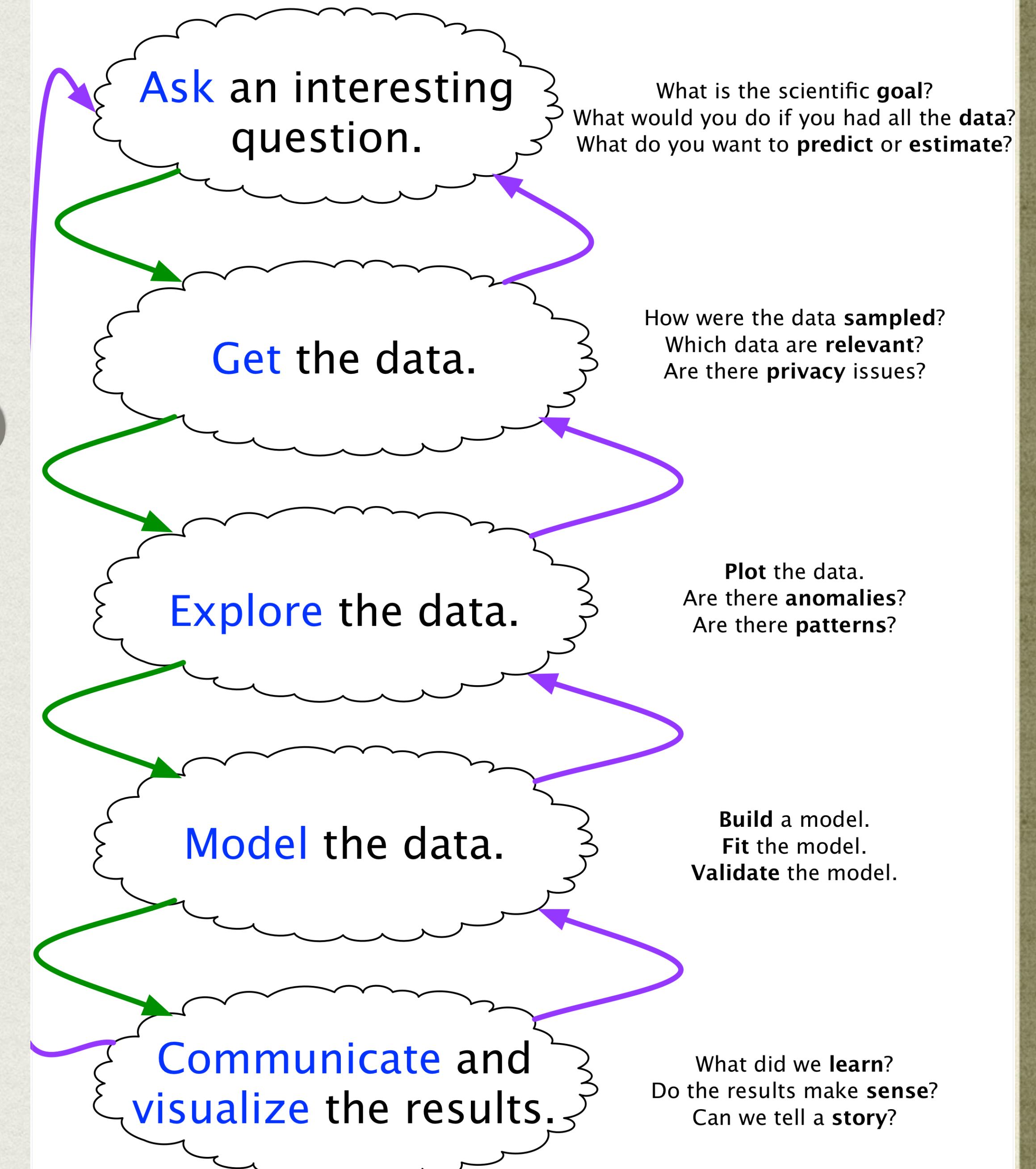
- Take a lot of time analyzing and understanding the data before jumping into code.
- Think about the goal and problem you are trying to solve.
- Explore the data well.
- Remember to keep it within scope - Keep the focus on the problem narrow with two or three questions at most.



The Data Science Process

BE PROCESS ORIENTED

Implement a procedural approach to any problem....



WHATS NEXT??

- Experimenting with other different modeling techniques and finding which one is better.
- Building a dashboard using Flask and predicting how many women will sign up for “Engineering” next year.
- Taking more time to think about the problem before starting with the code.
- Remember it never works the first time, might not work the second or the third - but just don't give up!

THANK YOU 😊

By,

Sana Nasar

