



Assignment 6

Create Visualizations with Matplotlib

Wiley Winters

MSDS 670 Data Visualization

2024-FEB-25

Research Question



Does the country a person works in determine his/her annual salary and do factors such as company size and work setting also contribute to a person's annual pay.

Dataset Source and Cleaning

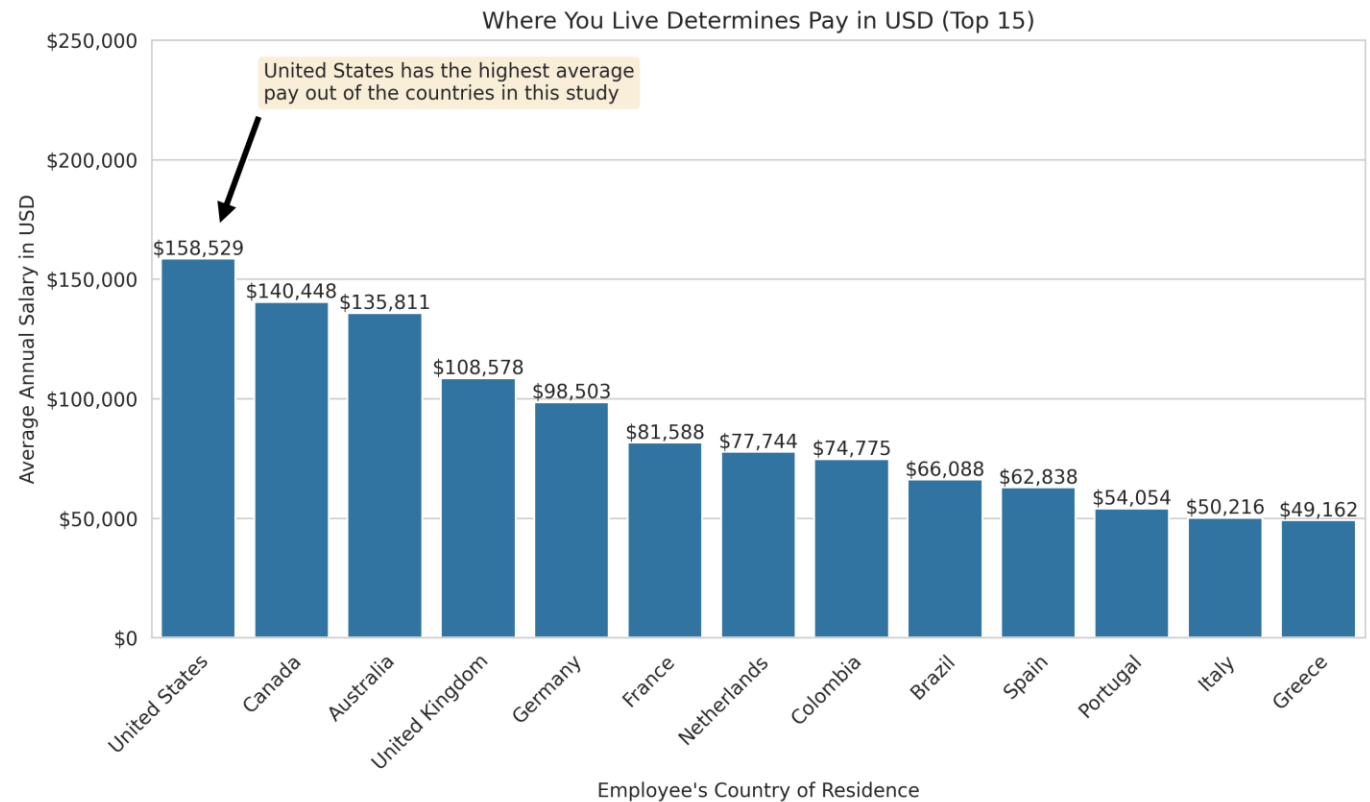


- The dataset used in this study was obtained from [Jobs and Salaries in Data Science](#)
- Research indicated the dataset was obtained from [ai-jobs.net](#) and modified to include a job_category column
- Dataset contained 12 columns and 9,355 rows
- 4014 duplicates were found and removed
- Analysis was conducted on the remaining data
- There are some unbalanced classes that were not handled

Average Salary by Country



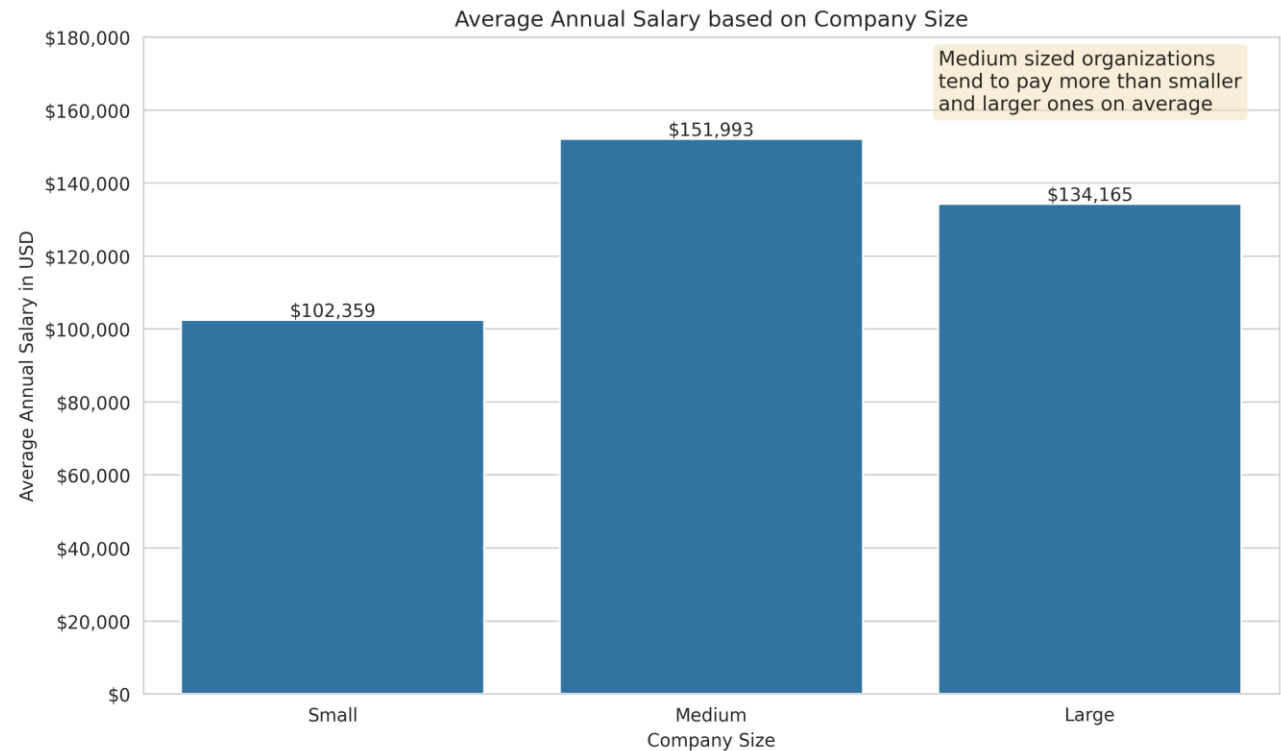
- The dataset contained many different countries, but since it was an informal survey, some countries did not have enough records to properly analyze
- Duplicates were identified and removed from the dataset before analysis was conducted.
- This visualization was conducted on the 15 countries that had the most records in the dataset
- Although it is a simple addition and the bars are arranged from largest to smallest, The text annotation really adds to the chart



Average Salary by Company Size



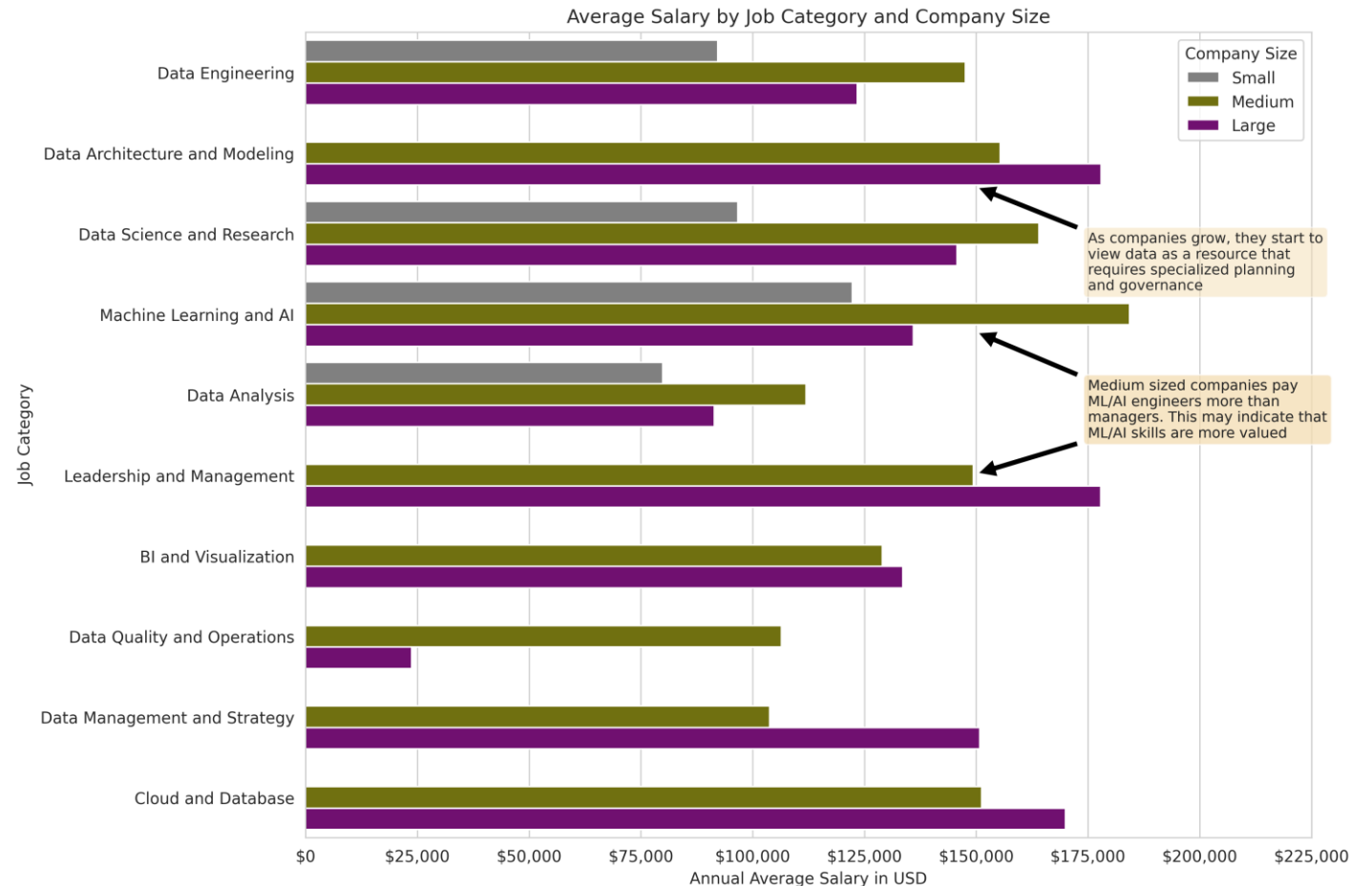
- Sorted the columns by company size instead of average income.
- Medium sized companies have a larger annual salary in the data science and related fields
- Since this data was taken from an informal survey, I am not sure of the criteria used to classified organizations by size.
- I wanted to add an annotation but couldn't decide if I should include an arrow pointer or not



Average Salary by Job Category and Company Size



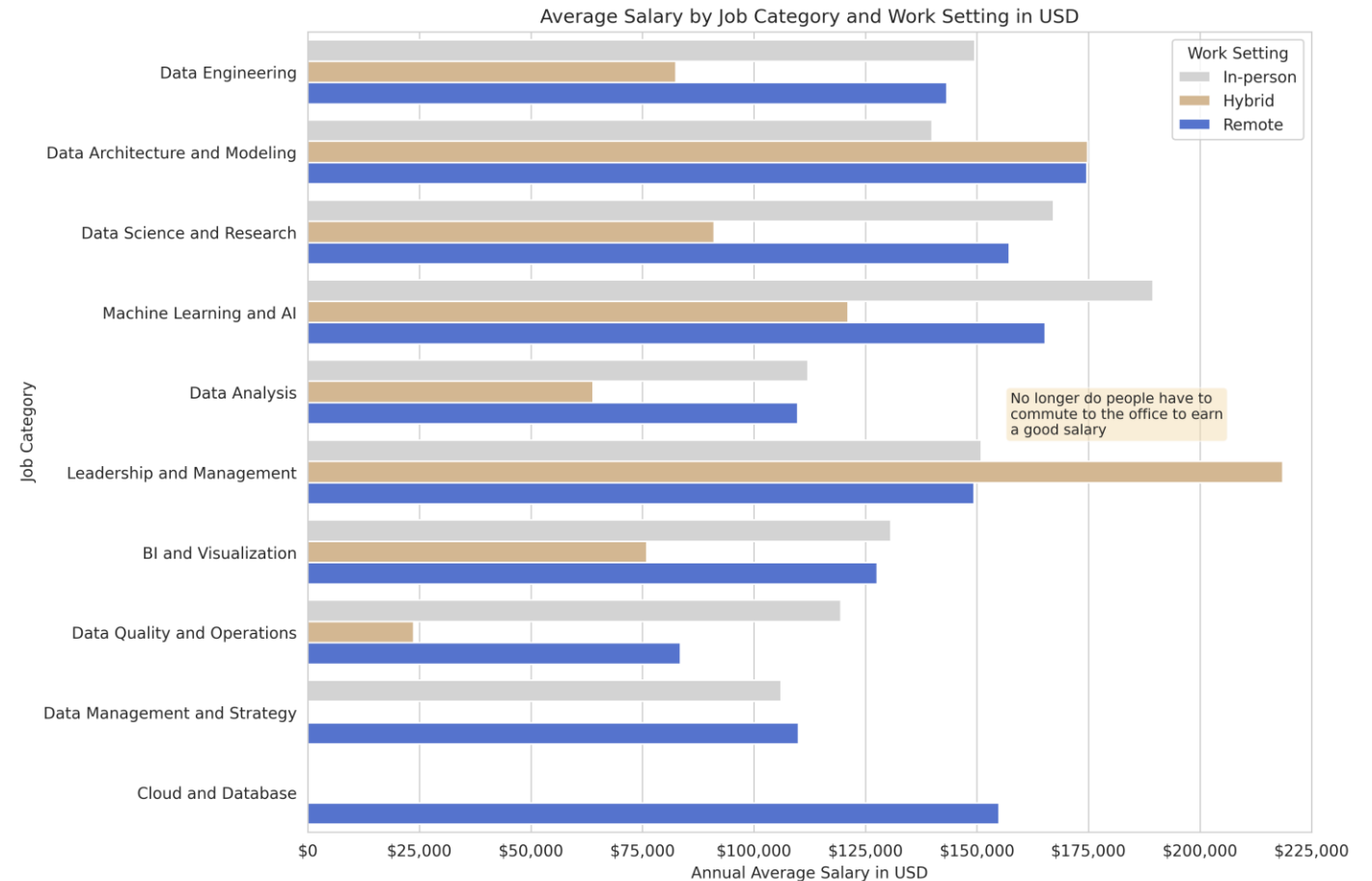
- The grouped bar chart allows the audience to compare different job categories, salary, and company size
- Pointed out some items on the chart that are interesting, like medium sized companies paying more for ML/AI engineers than management
- Small firms often do not have the resources to hire specialized employees, and this is reflected in the chart



Average Salary by Job Category and Work Setting



- This chart illustrates that going into the office everyday does not always pay the most.
- Leadership and management having the highest average salary for hybrid work indicates that people can work effectively in a hybrid setting
- Cloud and database having a high salary for remote work indicates this category of work can be done anywhere



Conclusions and Findings



- In this dataset the United States has the highest average annual salary of the countries being compared
- Company size does make a difference in annual salary with small and large companies lagging medium sized ones
- Small companies do not hire as many specialized employees as medium and larger ones
- People working in remote and hybrid settings can achieve salaries that rival people who go into the office everyday

Questions?

