Group 3 - Reese W.N, Vinh Tran, Kelly Zhang, Swikara Dhikari (The Trendsetters)
DS 311
10/20/24
“Salary Dataset Summary”

Summary Report

The Salary dataset is composed of job data across various data-related roles. The Key Performance Indicator in this dataset is Paid Wage Per Year, which will be used as a basis for our analysis. Exploring this dataset provides insight into the various characteristics associated with different levels of paid wages. For our exploration, we’ve been breaking down various data-roles in a Jupyter notebook using Python . An important thing to note is how we will be breaking down associated job titles. The salary dataset has an exorbitant amount of unique job titles, therefore we will be breaking down the job titles by their corresponding Job title subgroups, which is only eight unique values. This enables us to conduct more efficient analysis across different job groups while still maintaining the integrity of our data analysis.

Data extrapolation paired with visualization helps the group identify patterns or trends when working with the data. As a start, we have begun cross-examining graphics across variables such as: Regions, Subgroups, Education level, Experience levels, College major, and the KPI Paid Wage Per Year. Some interesting facts we have came across are:

- *Fun Fact #1*: When observing the average paid wage per education level, going from a high school diploma to an associates degree can increase your salary potential from $62,000 to $90,000. This is the largest jump between any education level. The next step in the educational process of pursuing a bachelor’s only increases the salary chances by about $8,000 until capping out at a masters degree for an additional $2,000, reaching $100,000. The average pay then drops when going for a doctorate, going down to $85,000. From this, I would say if you are looking to only focus on experience and not necessarily post-graduate education, perhaps only going for a bachelors is ultimately worth it, considering that is a $37,000 salary increase in a matter of 4 years if not longer.

- Fun Fact #2: Another fun pattern that emerges looking at the various job subgroup categories is the concentration of software engineers. This subgroup accounts for nearly 60% of all entries in the dataset. It would be interesting to see how the current job market compares since the release of this data, especially since the rise of data-related programs and certifications for data science and analysis.

- Fun Fact #3:
- Fun Fact #4:
- Fun Fact #5:

Facts such as these provide a clearer picture of the salary data. At the same time, it raises questions digging deeper into why patterns or trends exist in the first place.