Al Boot Camp Project 1

Team Members:

MARWA BENSALEM

SAMI CHOWDHURY

ROBERT FERRARI

MILEN KING

100 Thousand Dollar Professions Across the United States

Project Purpose / Description

- Our project highlights career paths across the U.S. with \$100K+ salaries, offering people a clear understanding of high-paying roles across different fields.
- This resource is crafted for anyone looking to increase their earning potential or find a stable career path, providing practical guidance to make informed career decisions based on qualifications and opportunities.

Goals/Questions to be addressed

- Which States had the highest earners?
- What education levels are amongst the highest earners?
- What profession or professions are amongst the highest earners?

Overview of data collection, cleanup and exploration process

• Data Collection:

- Census Data: Includes demographic, education, and income details.
- o BLS Data: Contains job titles, employment statistics, and wage information.

• Data Cleanup:

- o Removed duplicates, NaN values, and outliers.
- Standardized column names

• Data Transformation:

- Mapped occupation codes between datasets to enable merging.
- Separated national and state-level data for focused analysis.

• Data Exploration:

- Analyzed relationships between education, occupation, and income.
- Visualized state-level and national trends in wages and employment.

Approach taken to achieve goals

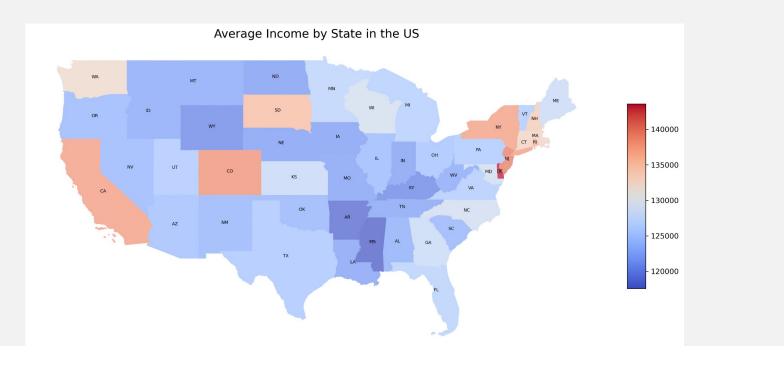
- Utilize GeoPandas to graph heatmap of United States to visualize wage gradients
- Bar, pie, box-and-whisker plots to create distributions of variables
- Analyze correlations between variables to determine relationships

Result/Conclusion 1

- The average income per state shows concentrations in certain geographic areas.

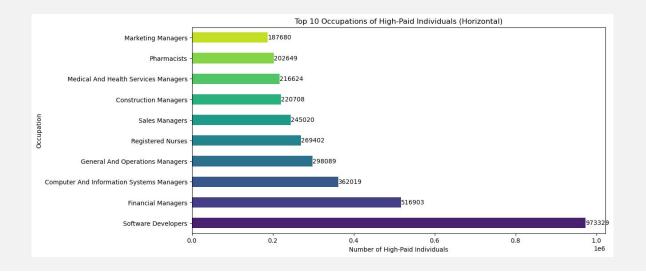
 O North East down to the Atlantic.

 - East Central to Mid North.
 - CA, CO, SD, and WA in the West



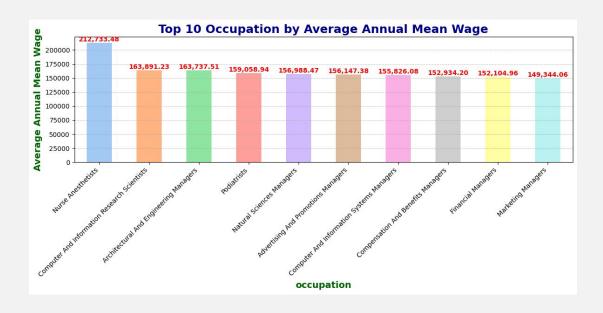
Result/Conclusion 2

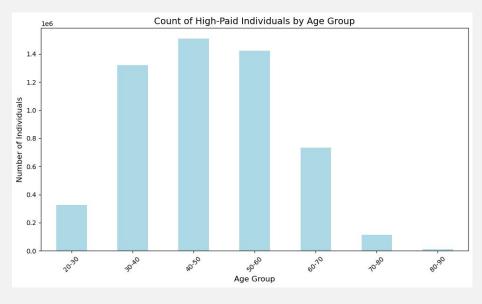
- The highest number of individuals earning +\$100K generally have managerial roles. The industries are mostly in Technology, Operations, and Healthcare.
 - Software Developers and CIS Managers
 - Financial Managers and Operations Managers
 - Registered Nurses and Pharmacists



Result/Conclusion 3

- The highest average income roles are still generally managers
 - Science and Healthcare professions
- The age distribution is normal with the majority of individuals between 40 to 50 years old





Summary

- 1. **Geographic and Industry Trends:** Delaware, California, and Texas top the list for high-paid individuals due to strong industries like tech and finance.
- Top Occupations and Education: Software developers and financial managers are the most common high-paid roles.
- 3. **Demographics and Correlation:** Ages 35–54 dominate high-paid demographics, with males leading in count but females earning competitive wages.

Problems Encountered

- Differentiating national and state-level data in BLS required filtering to focus on state-specific insights.
- Mapping education as a factor was challenging due to incompatible or missing data fields.
- Aligning education data with geographic maps needed extra processing for accuracy.
- Handling data inconsistencies and ensuring proper merges across datasets required careful validation.

Future Considerations



Summarize any additional questions that surfaced, what your group might research next if more time was available, or share a plan for future development.

New Questions:

- What drives gender income disparities despite similar education levels?
- How do urban vs. rural incomes differ within states?

Future Research:

- Explore industry-specific wage trends.
- Analyze income shifts over time.

Next Steps:

- Use machine learning for income prediction and pattern discovery.
- Build interactive dashboards for better data exploration.