Demographical Analysis of Poverty Rates and Salary from 2017-2021

The Data Vikings

Kendra Johnson, Same Moe, Jake Uhl

4/24/2023

**Introduction**

The two main concepts that our group will be investigating is employment and poverty. Our plan is to look at data from 2017 – 2021. We plan to look at data at the national level and state level, potentially drilling down to county level for a particular state (Stretch Goal)

This study began with asking:

1. How have salaries changed nationally? (2017 - 2021)
2. How have salaries changed from state to state? (2017-2021)
3. How have salaries changed in different industries?
4. How have poverty rates changed? (2017-2021)
5. How are changes in salary and changes in poverty correlated from state to state?
6. How does an increase in salary correlate to a decrease in poverty?
7. What occupation has the largest numbers of employees in 2021?
8. Which occupations are increasing from 2017-2021?
9. What effect, if any, did the covid 19 pandemic have on income and poverty in the US.

**Collaboration**

Using Trello, the Data Vikings will break the project down into smaller tasks that are assigned to particular team members.

* Tasks to competition
  + Load Github Repo / Clone to individual computers
  + Create data pipeline
  + Create team datablob
  + EDA
  + ETL
  + Create SQL DB
  + Create Kafka Producer/Consumer
    - Consume csv and produce copy into the datablob-storage
  + ML
  + Dashboard using PowerBi
    - Napkin Drawing

**Data Sources**

* Primary → Occupational Employment and Wage Statistics → <https://www.bls.gov/oes/tables.htm> → Planning to append tables since they are separated out by year
* Secondary → Small Area Income and Poverty Estimates (SAIPE) → <https://www.census.gov/data/developers/data-sets/Poverty-Statistics.html> → API call

**ETL**

This is included in the Jupyter Notebook.

**Conclusion**