Group XL Capstone Project: USA Education and Economic Correlation

Zeeshan Pervaiz, Sarvani Dantuluri, Rob Dewan, Cheyenne Peterson

Introductions

Cheyenne

B.S in Anthropology and TESL.
Freelance artist on the side.

Zeeshan

B.A in Biology, M.A in Education.
Transitioning to Data career from Education.

Sarvani

B.S. in Chemistry and Minor in Business Management/Administration

Rob

B.S in Mathematics.
Likes Numbers and Music

Project Goals:

- Gather and analyze U.S Census data to determine if a correlation exists between Educational attainment level, Employment and Income for citizens in the United States.
- Utilize machine learning models to determine how well certain factors can be predicted based on the data gathered.

Initial Questions:

- What is the per capita income for the different levels of education?
- What is the average household income per state compared to the US as a whole?
- Does household income correlate to educational attainment?
- What is the unemployment rate for the different levels of education?
- How has educational attainment changed throughout the years in the US? income?

ETL

Sources:

- Current Population Survey
- National Center for Education Statistics
- Census Bureau

Cleaning Methods:

- Pandas Multi-Indexing
- Requests and Beautiful Soup
- Properties of Cyclic Groups

Dashboard Visualizations

- Analysis addressing our initial questions will be shown utilizing DASH.
- Our DASH dashboard includes multiple pages of visualizations that answer each exploratory question.

Machine Learning

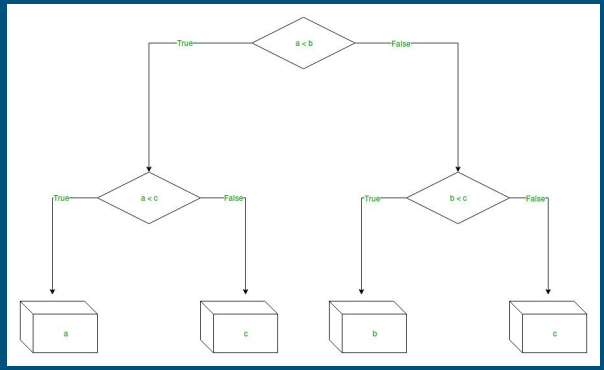
Utilized 4 ML algorithms to determine how well each is able to predict “unemployment rate” based on select features from the data.

Machine Learning Results

Model	Score
Linear Regression	0.54
Decision Tree Regression	0.78
K Nearest Neighbors	0.57
Ridge Regression	0.54

Based on the results obtained we can conclude that Decision Tree Regression was the most accurate in predicting unemployment rate.

- DT function differently then the other models.
- DT observes features and splits the data into a tree like model of decisions and consequences. Process repeats until a final prediction is made at a leaf node.



Findings and Analysis

- The average income has increased throughout the years.
- An individual with a degree typically has a higher income than someone who does not.
- Average state income has changed a lot from 1990 to 2019
- The working population total is increasing but so are the unemployment rates.
- Someone with an advanced degree is more likely to be employed that someone with just a GED.
- Average income has increased greatly for people with advanced degrees while income for people with no GED has remained relatively the same.

Application of Findings

Given the results of our analysis, Organizations can leverage the following recommendations:

- Targeted marketing based on educational attainment levels for each state.
- Adjust compensation to market standards based on location, education level, household size, etc.
- Utilize the unemployment data to offer positions in strategic locations.
- Offer traditional and/or alternative education based on education data for each state.
- Adjust educational requirements for positions based on the observed trends in educational attainment.

References

- *CPS Tables : U.S. Bureau of Labor Statistics*. (2023, January 25). https://www.bls.gov/cps/tables.htm
- *Median household income, by state: Selected years, 1990 through 2019*. (n.d.). https://nces.ed.gov/programs/digest/d21/tables/dt21_102.30.asp
- *Multi-Page Apps and URL Support | Dash for Python Documentation | Plotly*. (n.d.). https://dash.plotly.com/urls
- *Styling*. (n.d.). https://plotly.com/python/styling-plotly-express/
- US Census Bureau. (2022, July 4). *Educational Attainment*. Census.gov. https://www.census.gov/topics/education/educational-attainment.html
- Wan, M. (2021, December 14). *Beginner's Guide to Building a Multi-Page App using Dash, Plotly and Bootstrap*. Medium. https://towardsdatascience.com/beginners-guide-to-building-a-multi-page-dashboard-using-dash-5d06dbfc7599