Understanding and Analyzing Adult Education Levels

Pragati Pant
pantp@rpi.edu
Rensselaer Polytechnic Institute, Tetherless World Constellation, Troy, NY, United States
https://github.com/pantp/DataAnalyticsSpring2020/tree/master/project

Abstract

A nation's strength lies in its people. More specifically, it is important to analyze education as an indicator of a nation's growth. One area of education research is adult education. In this context, adult education refers to the concept that adults, approximately between the ages of 25 and 64, have finished some degree of schooling. There are three general levels of adult education: below secondary, upper secondary, and tertiary. The below secondary category includes adults who have attended some lower secondary schooling. The upper secondary category includes adults who have finished secondary schooling or high school. Finally, the tertiary category includes adults who have finished the highest degree of schooling, which entails some type of college degree. This research project aims to understand how adult education trends have changed within the United States and the world as a whole. Perhaps, this analysis will help in making predictions of where education is headed in the future.

For the purposes of analysis, the dataset used in this project is from the Organisation for Economic Co-operation and Development (OECD). This dataset has annual data points from various countries for three different levels of adult education. The three levels are defined as below secondary, upper secondary, and tertiary. All data points range from 1981 to 2018, but not every country is accounted for each year. Additionally, only data from eight countries is available for 1981. This changes as more years are included in the analysis. Therefore, in 2018, data for many more countries is available. With the addition of each new year, more and more data is available. This positive change in data availability represents the growth of the OECD in more countries.

This poster focuses on the various aspects of the project itself and draws on conclusions based on questions and hypotheses stated.

Questions and Hypotheses

The following are the two questions that served as motivation for this project.

1. **How have adult education levels changed within the United States and the world between the years of 1981 and 2018? Furthermore, is there a relationship between the different education levels and years passed?** I hypothesize that education level trends have changed and grown positively over time in both the United States as well as in the world.

2. **How has gender played a role in the change in trends of adult education levels between the years of 1981 and 2018?** I hypothesize that female education level trends have grown more positively than male education level trends in both the United States as well as in the world.

Analyzing Datasets and Deriving Conclusions

Workflow for Data Analysis

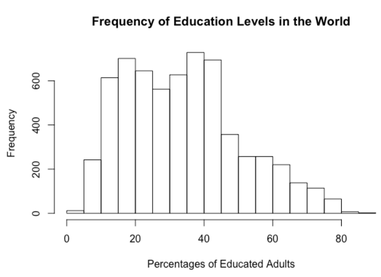
Hypothesize:
- Retrieved dataset from sources
- Formed questions/hypotheses
- Presented project proposal

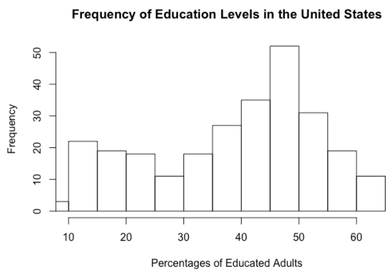
Analyze:
- Cleaned and filtered dataset
- Looked for initial data patterns
- Performed thorough data analysis

Visualize:
- Graphed initial data patterns
- Aggregated/graphed multiple regression models

Conclude:
- Analyzed trendlines and R values
- Found data analysis to be inconclusive based on models

Exploratory Data Analysis





Figures 1 and 2 (above) reflect the distributions of percentages of adults educated in the world and in the United States.

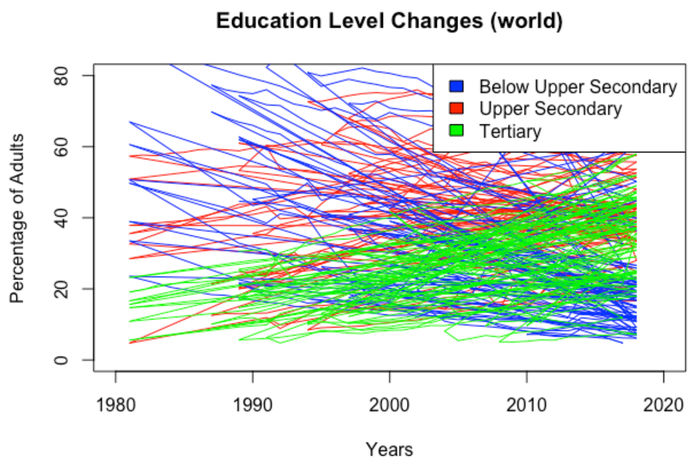


Figure 3 (above) graphs the initial adult education trends using the world data. It is evident that tertiary education percentages have grown positively while below upper secondary and upper secondary levels have grown negatively.

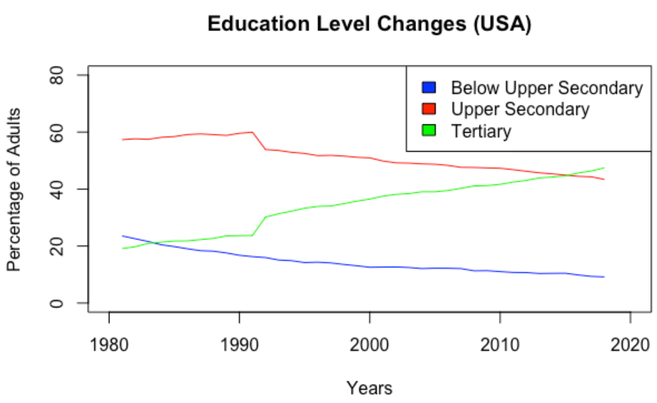


Figure 4 filters and plots the initial adult education trends using data for the United States.

Comprehensive Analysis

Question 1

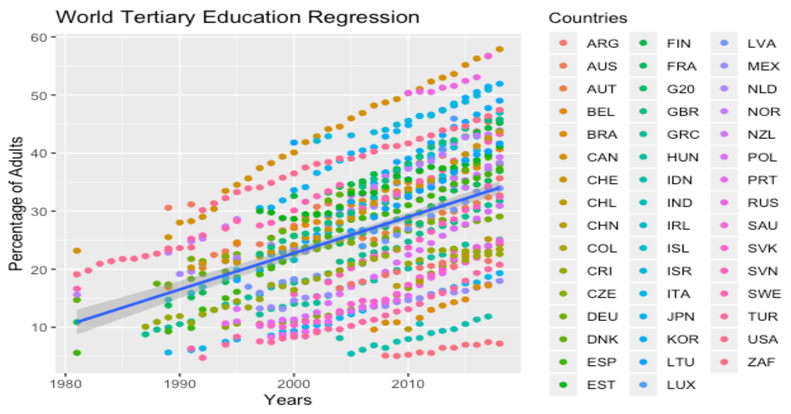
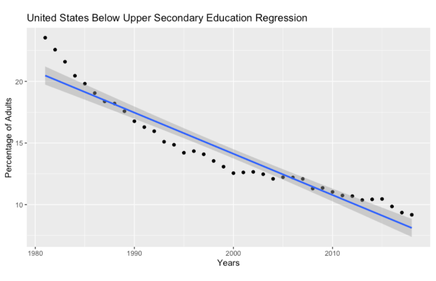


Figure 5 shows the world linear regression model for the tertiary educational level. The adjusted R-squared value is 0.2054, which shows a very weak linear relationship.



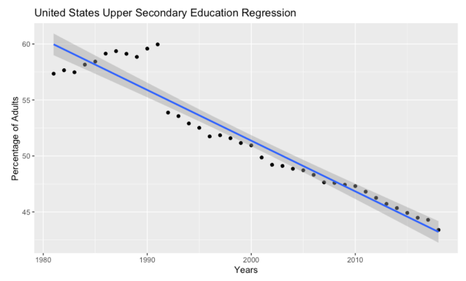
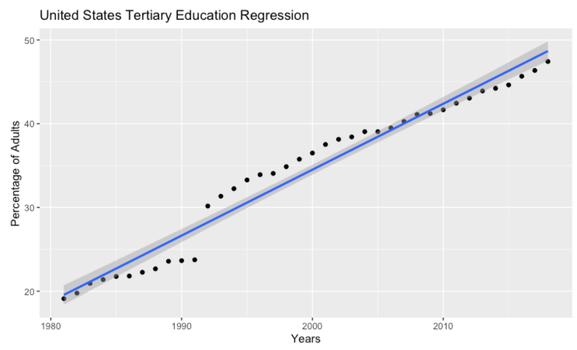
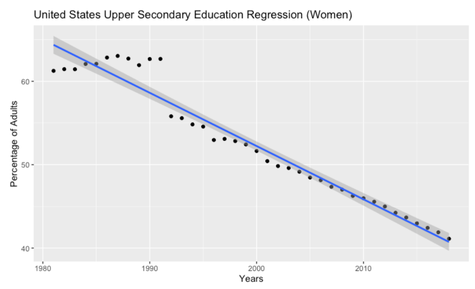


Figure 6 and 7 (above) demonstrate below upper secondary and upper secondary level plots for the United States. The adjusted R-squared values are 0.9148 and 0.9191 respectively.

Figure 8 (below) shows tertiary level for the United States. The adjusted R-squared value is 0.9602. This is a very strong linear correlation between United States adult education percentages and years past.



Question 2



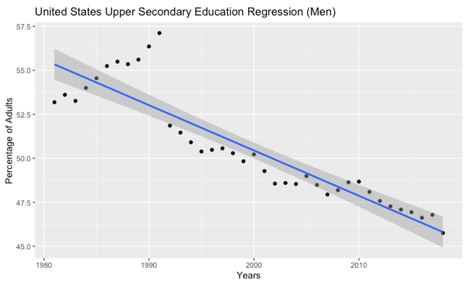
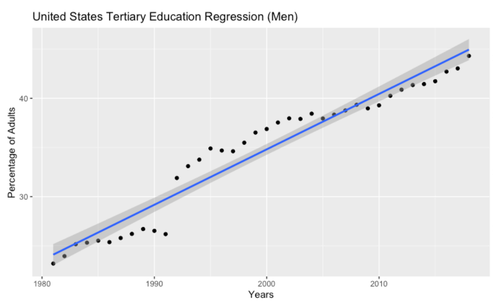


Figure 9 and 10 (above) demonstrate upper secondary level plots for the United States divided by gender. The adjusted R-squared values are 0.9491 and 0.8144 respectively. The gendered data shows strong linear correlation between United States adult education percentages and years past as well.



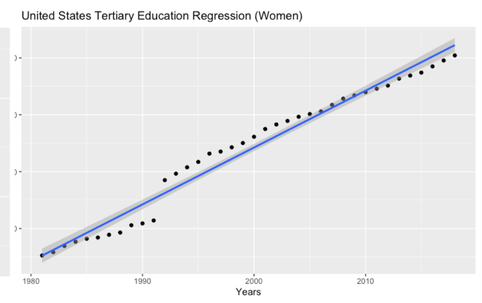


Figure 11 and 12 demonstrate tertiary level plots for the United States divided by gender. The adjusted R-squared values are 0.971 and 0.9343 respectively.

Conclusion

The data analysis was found to be inconclusive for the world data, but supportive for the United States data. There is a very weak relationship between world adult education levels' percentages and time. However, the United States data shows a strong linear correlation. The hypothesis is supported that more adults and more women are in the tertiary education level now than before.

Resources:

Dataset source: Adult education level.
Retrieved at https://data.oecd.org/eduatt/adult-education-level.htm
EDUCATION INDICATORS: An International Perspective.
Retrieved at https://nces.ed.gov/Pubs/eiip/eiip1s01.asp
ggplot2 package in R: https://ggplot2.tidyverse.org/