<https://www.kaggle.com/datasets/asaniczka/wages-by-education-in-the-usa-1973-2022?select=wages_by_education.csv>

Years 1973-2022

Race – White, Black, Hispanic

Gender – Male, Female

Level of Education – Less than HS, HS, Some College, Bachelors, Advanced

Questions:

1. On average over time does a female with an advanced degree earn as much as a male? April

Are things trending towards females earning as much as males.

1. For all races, does the increase in education affect the increase in salary for both men and women equally? Renee
2. If everyone has an advanced degree is there a difference between one of the groups based on race for last 5 years? Randy

Clone the git hub repository to each of our local repositories.

Anything from this point on commit to a branch separate for each of us.

Each finish code in Jupyter file by end of weekend.

Work on presentation slides and write up next week.

Project Proposal

Project Title: US Education vs Salary over time

Team Members: April Johnson, Randy Silvey, Renee Pleasnick

Project Description/Outline:

* Using the Wages by Education in the USA (1973-2022) data set found on Kaggle, we would like to determine if there are wage differences over time based on the criteria of gender and race.

Research Questions to Answer:

* Are things trending towards females earning as much as males for those with an advanced degree?
* Does the increase in education affect the increase in salary for both men and women equally?
* For those with an advanced degree, is there a difference between one of the groups based on race for the last 5 years?

Dataset to be used:

* Wages by Education in the USA (1973-2022) data set from Kaggle.com:
  + <https://www.kaggle.com/datasets/asaniczka/wages-by-education-in-the-usa-1973-2022?select=wages_by_education.csv>

Breakdown of tasks:

* Create a project proposal
* Create a repository for the project
* Read CSV file from Kaggle
* Team members each answer their assigned question
  + Manipulate data frames and wrangle data
  + Utilize APIs where appropriate
  + Create a minimum of two visuals for each question
  + Save those visuals as PNG files into our output data folder
  + Utilize statistical summaries to draw further comparisons
* Write a formal analysis in a .md file
* Prepare a slideshow presentation
  + Questions we found interesting and what motivates us to answer them
  + Where and how we found our data to answer our questions
  + The data exploration and clean up process (with our Jupyter notebook)
  + The analysis process (with our jupyter notebook)
  + Our conclusions, with a numerical summary and visualizations summary
  + Implications and limitations of our findings
* Practice presentation