

# Final Project

Jahedur Rahman

2/9/2022

## Introduction

For many years people have been arguing over inequality in many different areas. Some things that are talked about are gender inequality or race inequality. The gender pay gap is something that is debated and argued every year for many years. It has been researched and studied that males have a higher income than females. However, other variables may or may not be affecting this difference in income. Employers look at a number of different information of a person before hiring them. The relationships between the different information can show how much does each variable affect the income.

## Research questions

Is there a relationship between age and income? Is there a relationship between education and income? Is there a relationship between experience and income? Is there a relationship between gender and education? Is there a relationship between gender and experience? Is there a relationship between race and education?

## Approach

The main approach I will be taking is finding the correlation between each relationship. In addition, I will calculate the covariance to find how strong the correlation is between the variables. Also, I want to plot the data so I can visually see the relationships between the variables.

## How your approach addresses (fully or partially) the problem.

The approaches address the problem because these approaches will find if there is a relationship between each variable. This is important because it will help in figuring out if any of these variables have a bigger impact on the wage gap than gender does.

## Data (Minimum of 3 Datasets - but no requirement on number of fields or rows)

(<https://www.kaggle.com/jonavery/incomes-by-career-and-gender>)  
(<https://www.kaggle.com/nilimajauhari/glassdoor-analyze-gender-pay-gap>)  
(<https://www.kaggle.com/uciml/adult-census-income>)

## Required Packages

readxl  
ggplot2  
dplyr  
ggm

## Plots and Table Needs

age and income  
education and income  
experience and income  
gender and education  
gender and experience  
race and education

## Questions for future steps

Can the data used represent all countries or just specific countries? Do certain years have substantial changes in numbers due to certain events, like Covid-19 in years 2020-2021?