QTM 350 Final Project

Bao Truong, Daolong Yang, Xipu Wang

November 29, 2024

Introduction

This is the final project for QTM 350 class. Our group chooses to work on the topic of economic development. In this project, we will try to identify the relationship between GDP(adjusted), GDP Growth, and Employment Ratio. We will analyze the trend to draw a meaningful insights to apply for other countries. Our research question is: "What is the correlation between GDP(adjusted), Employment Ratio, and GDP Growth. Does having higher GDP(adjusted) and higher Employment Ratio leads to a faster increase in GDP".
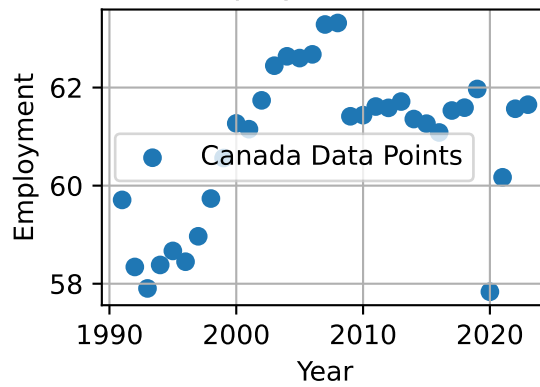
Data description

This project leverages World Bank data to explore and analyze the economic development trends in the North America region. The analysis focuses on three key economic indicators:
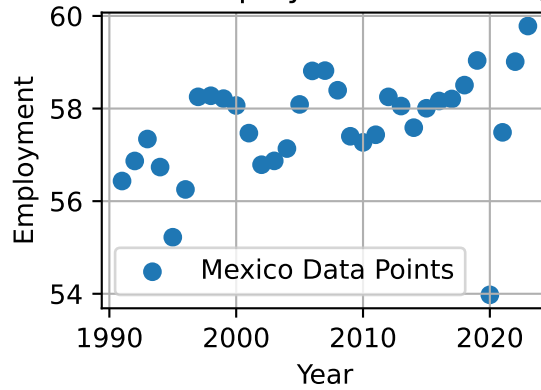
1. PPP: Measures the average economic output per person in the region adjusted for inflation.

2. GDP Growth: Tracks the annual percentage growth in the region's economic output.

3. Employment Ratio: Reflects the proportion of the working-age population that is employed. These variables provide a comprehensive overview of the region's economic health and development trajectory.

We clean the data using SQL. We try to filter out every country that's not USA, Canada, or Mexico. In addition, we drop the row in every dataset that we filter that has a null value in variable columns. Then, we also transform the column 'ppp' in the dataset PPP to 'log_ppp'. We then merge Employment Ratio and GDP Growth to visualize the residual to check for the condition of our model. We do similarly with PPP and GDP Growth.
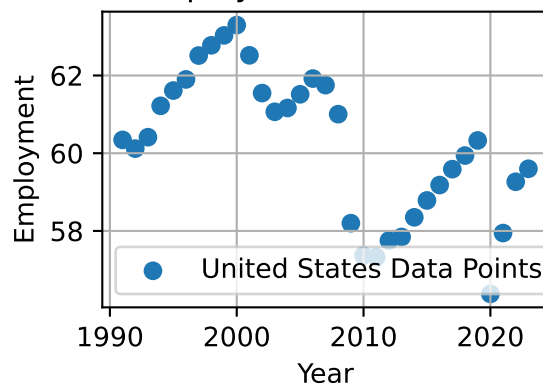
Scatter Plot: Employment vs. Year (Canada)

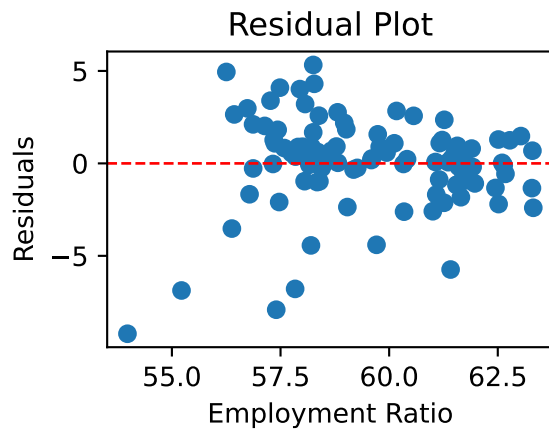## Scatter Plot: Employment vs. Year (Mexico)

Scatter Plot: Employment vs. Year (United States)

Data analysis

We create a residual plot to check for the requirements of linear regression model.

Residual Plot

[Residual plot showing Residuals vs Employment Ratio]

From both of the plots above, we see that the requirement for the linear regression is met. Therefore, we will conduct a linear regression to find out the correlation of the variables.

Quarto

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.