Project Report: Data Analysis of unemployment rate ages 24 - 65, from 2009 to 2021, Ireland.

GitHub URL

https://github.com/girardelligiuliana/UCDPA_giulianagirardelli

Abstract

The project is about the analysis of the Unemployment Rate of People between 25 to 64 years in Ireland from the years 2009 to 2021.

For the analysis of the data, I used Pandas, Numpy, Matplotlib and Seaborn for reading and visualizing the data.

Also I used different tools to present parts of the data such as info, shape, head and tail, columns, and describe.

After that, I used tools to verify that the data was not incomplete or duplicated, for this part I used duplicate, isnull and drop.

Once I had checked that the data was complete and without error I sorted the data by education level, sex and year for a better analysis.

Finally, I illustrated the data with a bar, histogram and line charts.

Introduction

I have chosen Unemployment rates in Ireland ages 25 to 64, in order to analyze the data and understand the current situation.

In this project, we can observe unemployment rates from 2009 to 2021, sorted by gender and level of education and analyze the progress since then.

After all the analysis, the results show that since 2009 unemployment levels have decreased notoriously, illustrating that males have the highest unemployment rates and that people with a primary level of education are the ones with the highest levels of unemployment. This data also shows how important it is to reach higher levels of education, as people with bachelor degrees, third level degrees and postgraduate degrees have the lowest rate of unemployment.

Dataset

For this project I have analyzed data, from the Central Statistics Office in ireland, CSO, where I have imported data about unemployment rates from 2009 to 2021, for people ages between 25 to 64 in Ireland, based on their level of education and sex.

I have chosen this dataset to show how unemployment rates varies by level of education, time, and sex, and the importance of reaching higher levels of education.

After analyzing this data, we can observe the progress and growth of the country through the time.

Implementation Process

I started the project importing *pandas*, which is a library used to read data, in this case it was used to read a CSV file. Second, I imported *NumPy*, which is a Python library used for working with arrays, this tool will help me along the project to illustrate the data. After that I imported two different visualization tools that are *matplotlib* and *seaborn*.

Once I had all the necessary libraries I started importing the file CSV, called Unemployment Rate of Persons 25-64 years.

From here on, I started to analyze the data with different tools:

First I used *info* to understand how the data was composed, that in this case it has 351 entries and 6 columns (Statistics, quarter, sex, education level, unit and value), that I also used *shape* for summarised that.

Secondly, I used *head* and *tail* to view a small sample of a series or the DataFrame object, that means the view of the first 5 rows and the last 5 rows of the data. And to that, I added the *describe* tool for a summarized description of the *std,max, min, mean* and count and *columns* attribute to return the column labels of the given Dataframe.

Once I had a sample of the data, I began to analyze if the data was complete, without blank spaces and without the information being duplicated. To analyze that, I used *df.duplicated* and *isnull* tools. And I added *dropna,* which is a tool that drops all rows with missing values and overwrites the dataset to double check.

After I was sure that the data was complete, I started sorting the data by:

- Education Level: the data has 9 levels of education, Primary, Lower Secondary, Upper Secondary, Post leaving cert, Third level, Higher Certificate, Ordinary Bachelor degree/professional qualification or both, Honorous bachelor degree/professional qualification or both and Postgraduate diploma/degree or Doctorate. This classification shows us the

percentage of unemployment based on the education level per year in a very simple table, easy to read.
- Sex: I sorted the data showing us the level of unemployement by sex, showing female unemployment rate, male and both genders.
- Year: data sorted by quarter, showing the percentage of unemployment for quarter in the year, with this information we can analyze the quantity of unemployment through time.
- Female: data sorted by female only, showing the percentage of unemployment through the years for only females.
- Male: data sorted by male only, showing the percentage of unemployment through the years for only males.

Consecutively, I used *groupby*, to see the mean, max, min and sum by sex.

After I gathered and analyzed all the information, I illustrated the data using bar charts, line charts and histogram charts. With these charts I illustrated the following information:

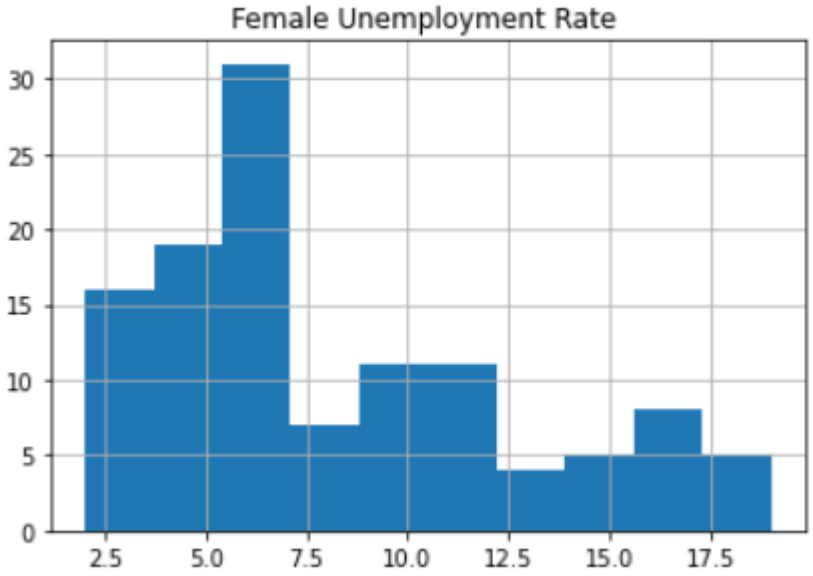
1. Unemployment rates by sex
2. Female and male unemployment rate
3. Unemployment rate through time
4. Mean of unemployment rate sorted by level of education

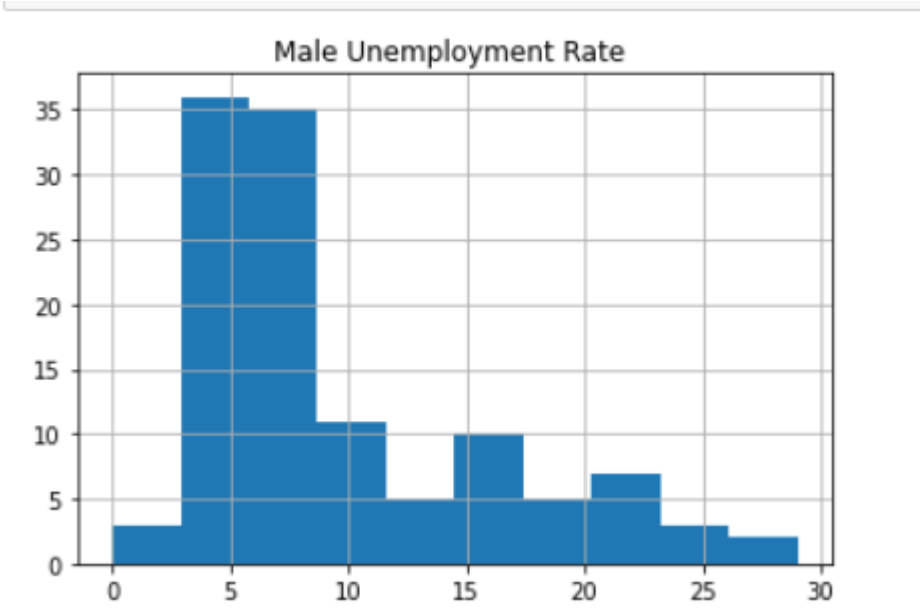
Results

This first chart is a histogram that shows the comparison rate of unemployment between females and male.

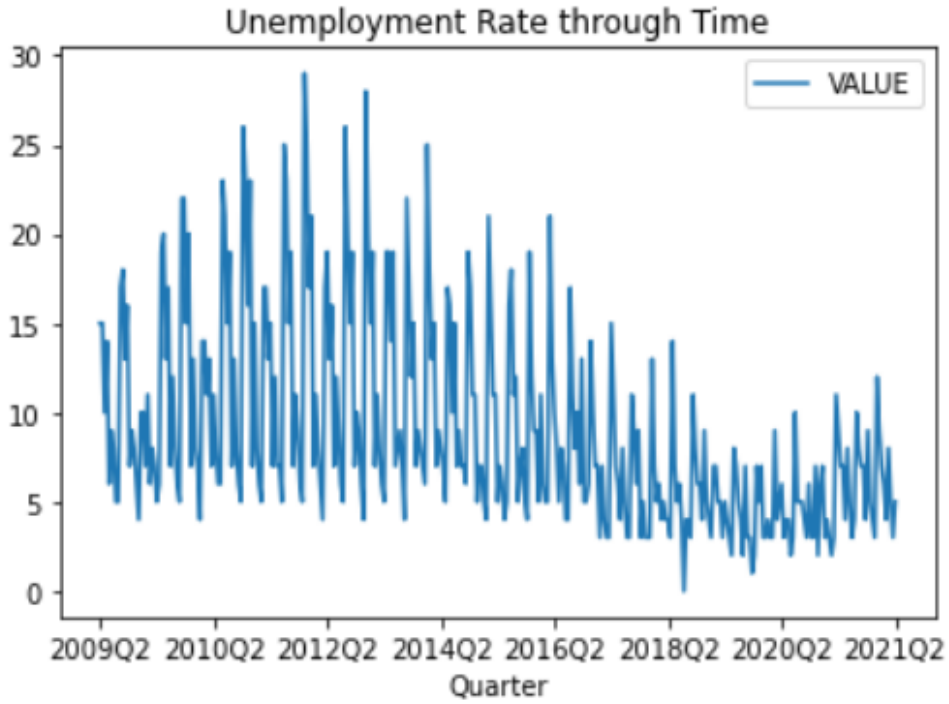
As we can observe, the male unemployment rates is significantly higher than the females rate.

We can observe the unemployment rate for female and male separately, for better examination.

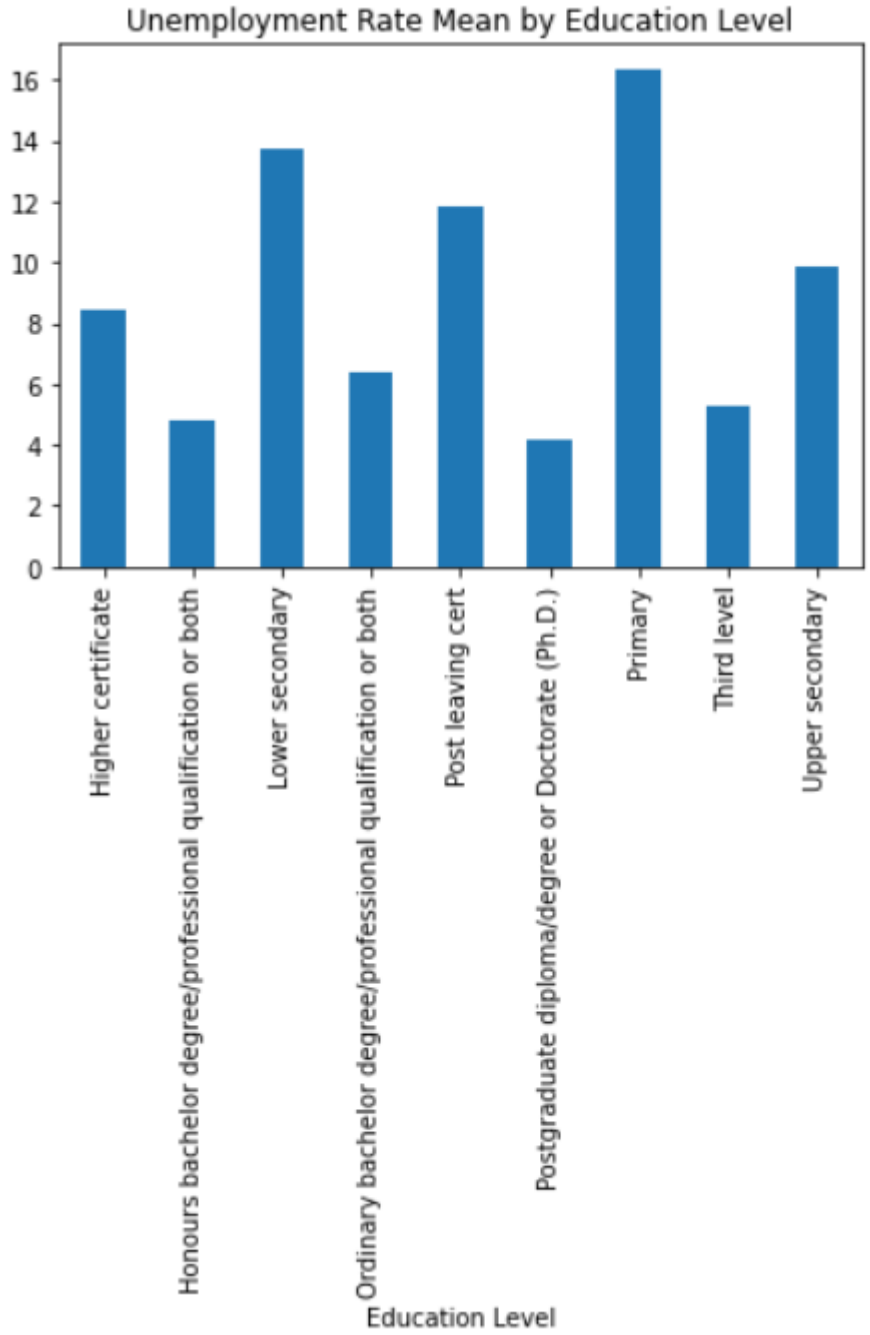




This Line chart illustrates the unemployment rate through time. This chart shows clearly that the unemployment rate has decreased on time since 2009, having the lowest rate in 2018/2020 and the highest in 2012.



This last chart is a bar chart that shows the unemployment rate sorted by education level. The data shows us that people with only a primary degree and with Lower Secondary degree are the people with the highest unemployment rate. While people with a third, bachelor and master degree have the lowest rate of unemployment.This proves the importance of reaching higher levels of education in Ireland to be employed.


Unemployment Rate Mean by Education Level
16
14
12
10
8
6
4
2
0
Higher certificate
Honours bachelor degree/professional qualification or both
Lower secondary
Ordinary bachelor degree/professional qualification or both
Post leaving cert
Postgraduate diploma/degree or Doctorate (Ph.D.)
Primary
Third level
Upper secondary
Education Level

MACHINE LEARNING

Having a high-quality data set is the first pillar to ensure success in machine learning models. The other two pillars are high quality models and an optimum combination of hyper-parameters. We can use Machine learning to predict the future based on this data.

I would use Machine learning to predict the rate of unemployment in the future, as we can predict that data based on the variable level of education. As we analyzed in the project, we can confirm that the higher level of education, lower is the unemployment rate, and the opposite, the lower is the level of education, higher is the unemployment rate.
With machine learning we can predict the future unemployment rate.
For this purpose we can use regression analysis, as it helps predicting continuous quantities, and the data would be

CONCLUSION

The conclusion after the analysis of this project is that higher education increases employability rate.

Unemployment rates for those aged 25-64 years old generally decreased as the level of education attained increased. In Q2 2020, males aged 25-64 years old with a primary education or below were more than twice as likely to be unemployed (10%) compared to males with a third level qualification (4%).
Employment rates for those aged 25-64 years old increased as the level of education attained increased and this relationship is stronger for females than for males. In Q2 2020, females aged 25-64 years old with a third level qualification were over three times more likely to be employed (80%) compared to females with a primary education

From 2008 to 2013 Ireland went through a financial crisis where many jobs were terminated, and we can see in the line chart how this period had the highest unemployment rates.
Since early 2014, the Irish Economy has exhibited particularly strong
growth in employment. As a result, the unemployment rate, which reached
16 per cent in 2012Q1, has fallen to 5.0 per cent according to the latest
data for the first quarter of 2019.
After that we can observe how the unemployment rates went down , having the lowest in 2020.

Insights

1. Show unemployment rates for people ages 25 to 64, between the years 2009-2021, in Ireland.
2. Show the unemployment rates by education level.
3. Show the unemployment rates by sex.
4. Show the unemployment rates by time.

5. Visualize the data using bar, lines and histogram charts.
6. Evaluate some descriptive statistics.
7. Utilize pandas, numpy, matplotlib and seaborn libraries for data analysis and visualization

References

The Data was obtained from: Data on Ireland sourced from the CSO's Labor Force Survey [https://data.cso.ie/table/QLF01]. European data sourced from Eurostat [https://appsso.eurostat.ec.europa.eu/nui/].