

# **7PAM2000 Applied Data Science 1**

## **Assignment 1: Visualisation**

**Name:** Srijā Swarna

**Student ID:** 22010955

**Datasets:** <https://github.com/9908173113/ADS1.git>

## Contents

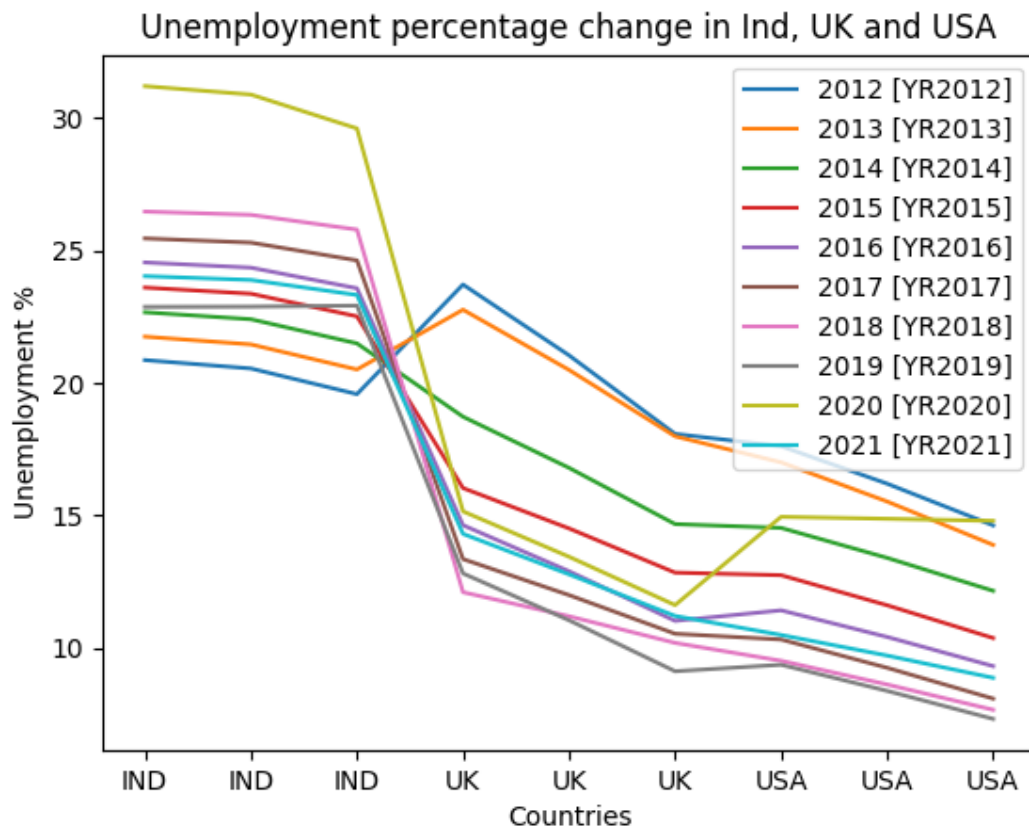
Overview of code and selected dataset .....	3
Visualisation 1: The line plot displays the unemployment rate of three countries.....	4
Visualisation 2: The bar chart displays the unemployment categories percentage of the three countries.....	5
Visualisation 3: The scatter plot displays the relationship between the overall unemployment percentage years and the three countries. ....	6

## **Overview of code and selected dataset**

The code implemented in the source file uses Python's pandas, seaborn, and matplotlib libraries to load and visualize the Unemployment dataset. The unemployment dataset contains unemployment data from the countries India, the UK and USA. The code first imports the necessary libraries and loads the unemployment dataset using the `pd.read_excel()` within which we load the excel file. The dataset is then converted to a pandas Data Frame after which it predefines the column names of the table.

The code defines three functions that create different types of plots. The first function creates a line plot that shows the unemployment rate of the three countries. The second function creates a bar chart that showcases the rank of three countries based on their unemployment rate. The third function creates a scatter plot that shows the relationship between the unemployment rate and the three countries, with each point coloured by its target class. Finally, the code calls each of the three plot functions to visualize the unemployment dataset in different ways. Overall, this code demonstrates how to load and visualize a dataset using pandas and matplotlib in Python. These functions generate three visualizations based on the unemployment dataset. The dataset consists of the unemployment rate of three countries from 2012 to 2021.

**Visualisation 1: The line plot displays the unemployment rate of three countries.**

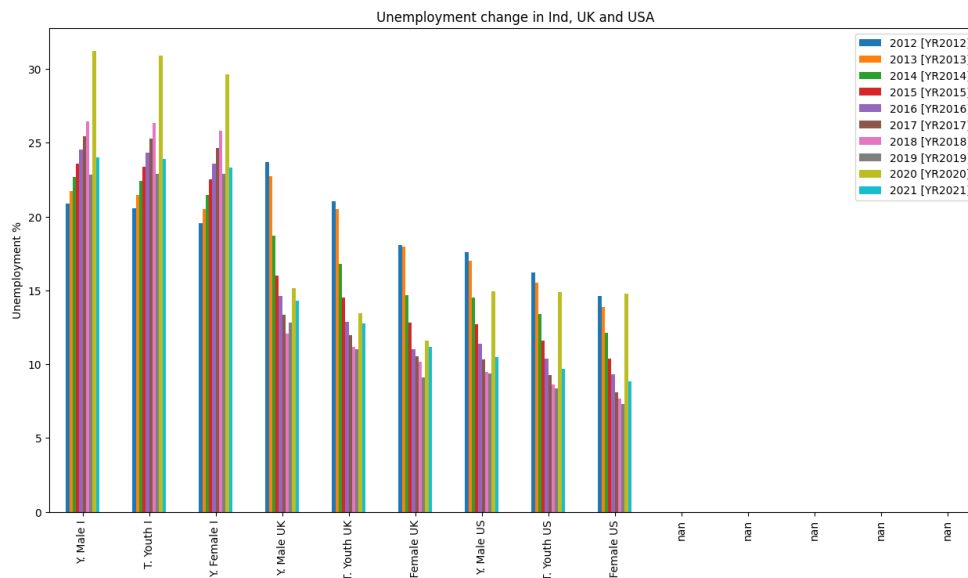


***Visualisation 1: Line Plot for Unemployment Change***

***(Source: Created by Author)***

The visualization of unemployment change in three countries provides insight into the variations in the unemployment percentage change in the three countries India, the UK and the US. It can be analysed from the figure that out of the three countries, India has the highest unemployment percentage in all three categories, Male Youth (1<sup>st</sup> in X-axis), Total Youth (2<sup>nd</sup> in X-axis) and Female Youth (3<sup>rd</sup> in X-axis). In terms of male unemployment, it lies between 20 to 30% same for both total youth and female youth unemployment. Concerning the UK, the male youth unemployment rate there generally lies around 15%, for total youth, it is between 10 to 20% and for female youth, it lies between 10 to 15%. In the USA, male youth unemployment generally lies between 10 to 20%, the same for total youth unemployment while female youth unemployment lies between 0 to 15%. This hints that out of the three countries, the USA has the lowest unemployment percentage in all three categories.

**Visualisation 2: The bar chart displays the unemployment categories percentage of the three countries.**

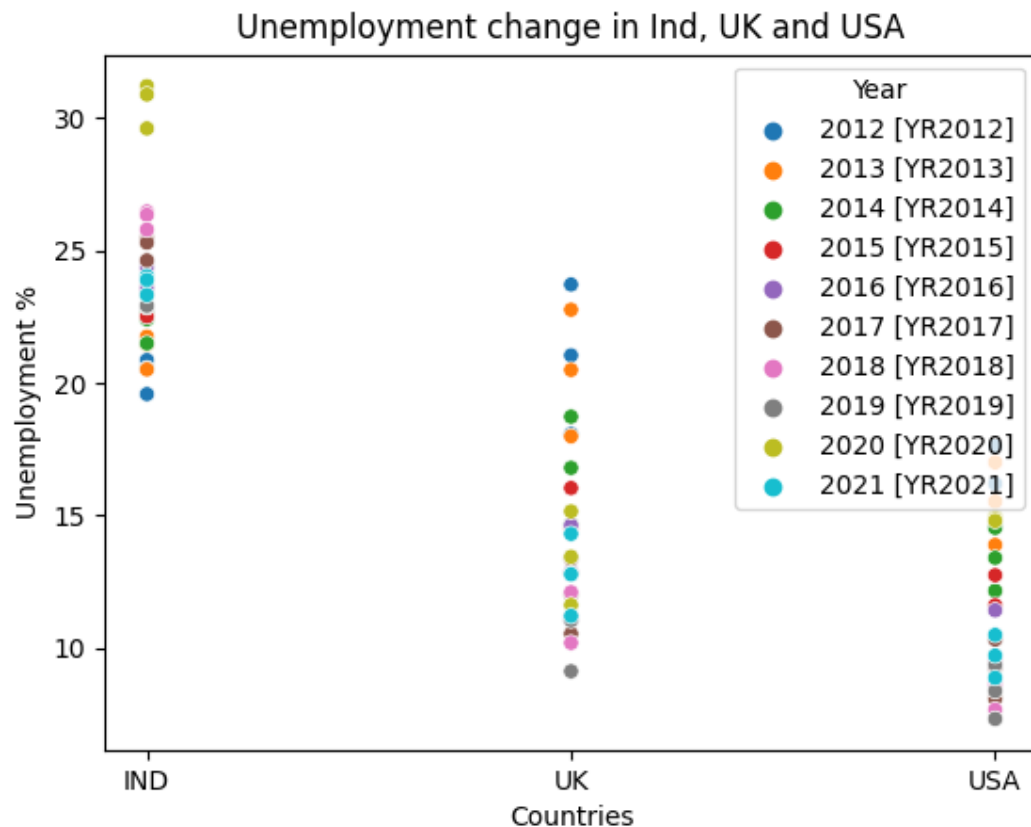


***Visualisation 2: Different Unemployment Categories of the three countries***

***(Source: Created by Author)***

The second visualization is a simple bar chart that provides a clear overview of the bar chart displays the unemployment categories percentage of the three countries. As the chart indicates, the chart presents that India out of all countries has the highest unemployment percentages in three categories. It can be visualised from the graph that the year 2020 was the time in which most youths were unemployed men (over 30%), female (around 30%) and total youth (over 30%). Concerning the UK, the year 2012 was the time in which most youths were unemployed men (around 25%), female (around 20%) and total youth (over 20%). In the USA, the year 2012 was the time in which most youths were unemployed men (between 15 to 20%), female (over 15%) and total youth (between 15 to 20%).

**Visualisation 3: The scatter plot displays the relationship between the overall unemployment percentage years and the three countries.**



***Visualisation 3: Overall Unemployment Percentage of the three countries***

***(Source: Created by Author)***

The third visualization is a scatter plot that is used to display the relationship between two variables, which in this case are sepal length and sepal width. The scatter plot represents each sample as a point and colours it based on its species. The plot shows that the year 2012 was the time in which India had the least overall unemployment percentage (below 20%). While for the UK the least unemployment year was 2019 with an overall unemployment percentage below 10%. And for the USA, the year 2019 faced the least overall unemployment rate of around 5%. It can be concluded from the figure that the UK and the USA, even with the highest overall unemployment, the countries are still below the average of India while the least unemployment in India is more than double the percentages of the two countries.

**Dataset reference:**

<https://fred.stlouisfed.org/series/UNRATE>

<https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS>