Coursework Assessment 6 PDF

by Sagari Muraliegaran

Submission date: 22-Nov-2024 09:47AM (UTC+0000)
Submission ID: 243413168
File name: 12762_Sagari_Muraliegaran_Coursework_Assessment_6_PDF_259784_1936741023.pdf (467.68K)
Word count: 1697
Character count: 8999

Analysis of Female Unemployment Rates: Insights from UK and USA

Introduction

In the United Kingdom (UK), female unemployment rates have fluctuated throughout time due to various factors, including recession, technological advances, and worldwide events like the COVID-19 pandemic. Unemployment measures people without a job who have been actively seeking work within the last four weeks and are available to start work within the next two weeks. This measure allows us to understand the trends and women's work opportunities.

In this report, there's statistical analysis and hypothesis testing which investigates unemployment patterns in the UK. The first hypothesis test uses statistics from "Office of National Statistics" *(Female Unemployment Rate (Aged 16 and over, Seasonally Adjusted): % - Office for National Statistics n.d.)*, which is a reliable source. The test examines if there was a significant difference in female unemployment rates in the UK between 2019 and 2020, as that was the period the COVID-19 pandemic (*Francis-Devine 2021)*.

A second hypothesis test compares average female unemployment rates in the UK and the United States of America (USA) (*Unemployment Rate Women U.S. 2023* n.d.) over a 20-year period (2000-2020). This test strives to identify differences and similarities between the two countries, taking into account their respective economic structures, cultural settings and labour market policy methods *(Women Bear Brunt of Coronavirus Economic Shutdown in UK and US | University of Cambridge 2020; Freeman and Wise 1982)*.

The aim of this study is to detect and evaluate statistical trends while taking into account the complex interaction of social, economic and policy factors influencing female unemployment. The findings are presented in clear visuals, summary data and comparisons, providing a detailed yet understandable examination of female unemployment trends in the UK and USA.

The Trend of Female Unemployment Rates in the USA and UK

In this section, there is a time series line graph that shows the comparison of the female unemployment rates in the USA and the UK. The R code and output used for the analysis are given in Appendix A.1.

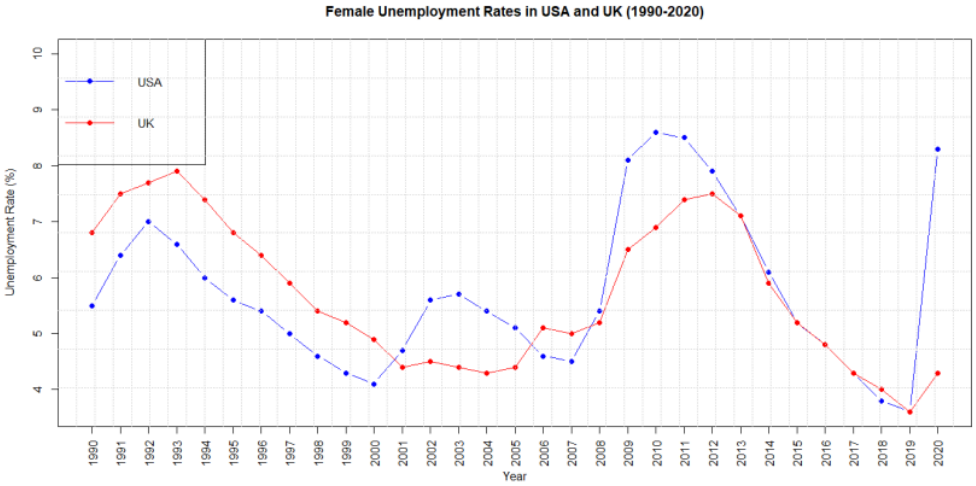


Figure 1 shows the trend of female unemployment in the USA and the UK over the periods between 1990-2020.

During the early 1990s recession in both the UK and the USA (Kochhar 2019; Industry employment and the 1990-91 recession n.d.), they faced economic downturns due to restrictive monetary policies like higher interest rates to control inflation and changes following the Cold and Gulf War, such as decreased defence spending. There was a decline in manufacturing and other sectors that led to job losses where female unemployment reached 7.0% in the USA and 7.9% in the UK.

Soon after in the mid-1990s, both economies began to recover due to various reforms, fiscal changes and a general tech boom, especially in the USA. As of this job growth improved, and female unemployment rates decreased. In the USA, the unemployment rate for women dropped to 4.1% by 2000, and in the UK it fell to around 5.9%.

By the late 2010s economic stability had resulted in record-low female unemployment rates: 3.6% in the USA (2019) and 4.3% in the UK, due to job creation and legislation. However, the COVID-19 pandemic in 2020 caused unemployment to skyrocket internationally, with the rate climbing to 8.3% in the USA and similarly in the UK, with women in service and part-time jobs being the most impacted.

Hypothesis Test 1: Comparing UK Female Unemployment Rates (2019-2020)

In this section, we examine the descriptive statistics and graphics for Hypothesis Test 1 and make an assumption on what type of test will be best. The values in Table 1 were calculated in R Studio where the code and output are presented in Appendix A.2.

Table 1: Descriptive Statistics for UK Female Unemployment Rates (2019-2020)

	Mean	Median	SD	IQR
2019 to 2020	3.95	3.95	0.4949747	0.35

Hypothesis Test 1 Hypotheses:

- Null Hypothesis (H0): There is no significant difference in female unemployment rate between 2019 and 2020.

  $H0: \mu 2019 = \mu 2020$

- Alternative Hypothesis (H1): There is a significant difference in female unemployment rates between 2019 and 2020.

  $H1: \mu 2019 \neq \mu 2020$

The descriptive statistics show the mean and median in Table 1 are the same (3.95), indicating a symmetric distribution of data. The low standard deviation and interquartile range predict that female unemployment rates will be stable in 2019 and 2020. The data's absence of skewness supports the assumptions of normality, making it appropriate for hypothesis testing. If the hypothesis test showed no significant difference ($p > \alpha$), it would support the null hypothesis of no significant change in female unemployment rates between 2019 and 2020. If the findings are significant ($p < \alpha$), it will support the alternative hypothesis.

Hypothesis Test 2: Comparing USA and UK Female Unemployment Rates (2000-2020)

In this section, we examine the descriptive statistics and graphics for Hypothesis Test 2 and make an assumption on what type of test will be best. The values in Table 2 and the histograms (figure 2 and figure 3) have the R code and outputs in Appendix A.3, Appendix A.4 and Appendix A.5.

Table 2: Comparative Statistics of Female Unemployment Rates in the USA and UK (2000-2020)

	Mean	Median	Standard Deviation	IQR
USA	5.780952	5.4	1.633897	2.5
UK	5.22381	4.9	1.181907	1.5

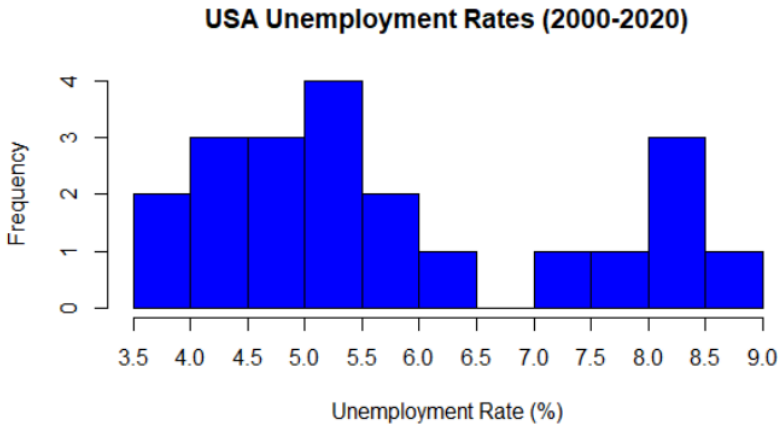


Figure 2 shows a bimodal distribution of USA unemployment rates from 2000 to 2020.

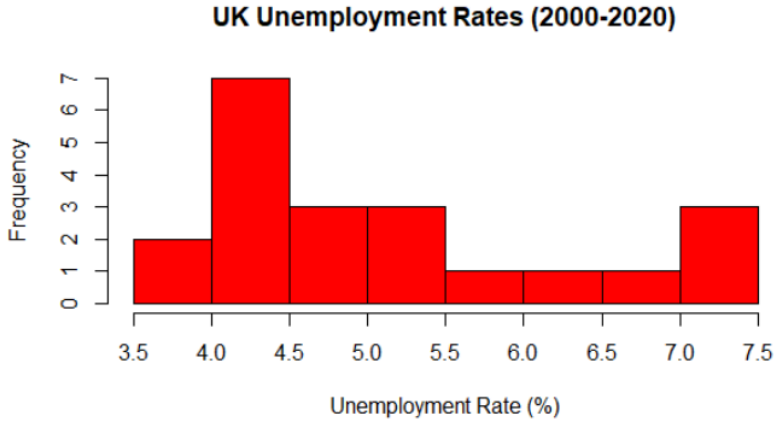


Figure 3 shows a slightly positively skewed distribution of UK unemployment rates from 2000 to 2020.

Hypothesis Test 2 Hypotheses:

- Null Hypothesis (H0): The mean unemployment rates in the USA and UK are equal.

  $H0: \mu_{USA} = \mu_{UK}$
- Alternative Hypothesis (H1): The mean unemployment rates in the USA and UK differ.

  $H1: \mu_{USA} \neq \mu_{UK}$

The descriptive statistics in Table 2 show that the USA has a higher mean for female unemployment rates than the UK, with greater variability and interquartile range. The USA's bimodal distribution shown in Figure 2, and the UK's slightly positive skewed histogram, shown in Figure 3, imply that normality assumptions may not apply, particularly for the USA. If normality fails, a hypothesis test may be a better option. Otherwise, a two-sample t-test

can test the null hypothesis, and the significant finding ($p < \alpha$) indicates differences in the mean female unemployment rate between the two countries.

Results

In this section, the results of the two hypotheses tests were conducted and presented in Table 3 and Table 4. The R code for conducting the hypothesis tests is given in Appendix A.6 and Appendix A.7

Table 3: Hypothesis Test Results for Comparing UK Female Unemployment Rates (2019-2020)

Results	Value
Chi-squared Value (X^2)	10,468
Degree of Freedom (df)	1
p-value	< 2.2e-16
Confidence Interval (95%)	[-0.007134168, -0.006865832]
Sample Proportion (2019)	0.036
Sample Proportion (2020)	0.043
Conclusion	Reject H0 ($p < 0.05$, a significant difference exists)

Table 4: Hypothesis Test Results for Comparing USA and UK Female Unemployment Rates (2000-2020)

Result	Value
t-Statistic (t)	1.2661
Degrees of Freedom (df)	36.431
p-value	0.2135
Confidence Interval (95%)	[-0.334953, 1.449239]
Mean (USA)	5.780952
Mean (UK)	5.223810
Conclusion	Fail to reject H0 ($p > 0.05$, no significant difference)

Conclusion

In conclusion, the two hypothesis tests generated conflicting results. Hypothesis Test 1, showed a significant difference ($p < 0.05$), thereby rejecting the null hypothesis. This aligns with the low variability observed and the likely economic shifts during this period. In contrast, Hypothesis Test 2, showed no significant difference ($p > 0.05$), consistent with the close means and overlapping distributions which led to the failure to reject the null hypothesis. However, the USA's bimodal distribution and the UK's skewed data may have affected normality assumptions.

To improve the data collection, a larger and more thorough dataset with geographical or economic factors would assist in eliminating bias. More research may look at long-term trends, regional disparities or how other variables impact unemployment.

References

Female Unemployment Rate (Aged 16 and over, Seasonally Adjusted): % - Office for National Statistics [Online]. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsz/lms [Accessed: 16 November 2024].

Francis-Devine, B. (2021). How has the coronavirus pandemic affected women in work? [Online]. Available at: https://commonslibrary.parliament.uk/how-has-the-coronavirus-pandemic-affected-women-in-work/ [Accessed: 21 November 2024].

Freeman, R.B. and Wise, D.A. eds. (1982). *The Youth Labor Market Problem: Its Nature, Causes, and Consequences*. A National Bureau of Economic Research conference report. Chicago: University of Chicago Press.

Industry employment and the 1990-91 recession [Online]. Available at: https://www.bls.gov/opub/mlr/1993/07/art2full.pdf [Accessed: 16 November 2024].

Kochhar, J.B. and R. (2019). Two Recessions, Two Recoveries. *Pew Research Center* [Online]. Available at: https://www.pewresearch.org/social-trends/2019/12/13/two-recessions-two-recoveries/ [Accessed: 16 November 2024].

Unemployment Rate Women U.S. 2023 [Online]. Available at: https://www.statista.com/statistics/193944/unemployment-rate-of-women-in-the-us-since-1990/ [Accessed: 16 November 2024].

Women Bear Brunt of Coronavirus Economic Shutdown in UK and US | University of Cambridge (2020). [Online]. Available at: https://www.cam.ac.uk/research/news/women-bear-brunt-of-coronavirus-economic-shutdown-in-uk-and-us [Accessed: 17 November 2024].

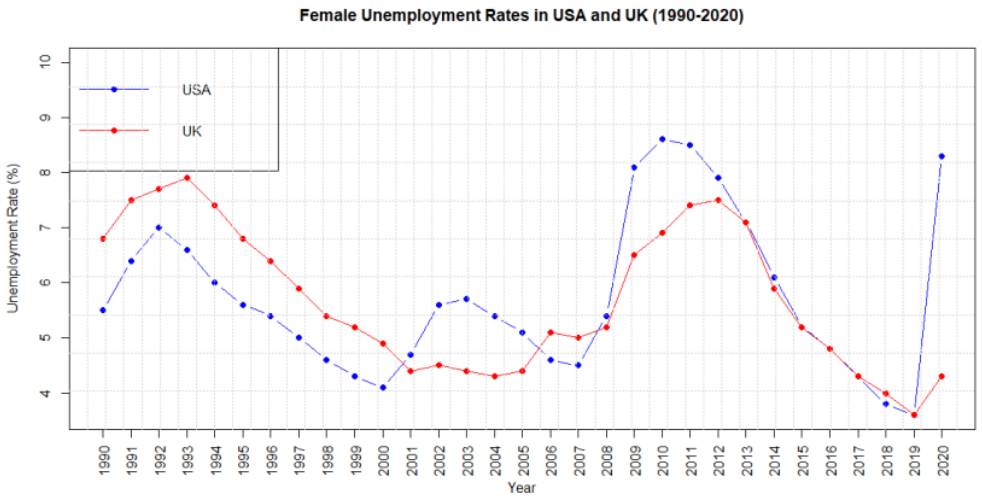
Appendix A.1

R code for Time series line graph – trends of female unemployment rates of USA and UK

```
> Female_Unemployment_Rates_in_USA_and_UK_1990_2020 <- read_excel("Female Unemployment Rates in USA and UK 1990-2020.xlsx")
>
> #Plot the USA female unemployment rate
> plot(Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,Female_Unemployment_Rates_in_USA_and_UK_1990_2020$USA,
+      type="b", #Plot both points and lines
+      col="blue",#Set the colour to blue for USA
+      ylim=c(min(c(Female_Unemployment_Rates_in_USA_and_UK_1990_2020$USA,Female_Unemployment_Rates_in_USA_and_UK_1990_2020$UK)),1
0),#Define y-axis limits
+      xlab="Year",#Label for x-axis
+      ylab="Unemployment Rate (%)",#Label for y-axis
+      main="Female Unemployment Rates in USA and UK (1990-2020)", # Title of the plot
+      xaxt="n") #Suppress default x-axis labels
>
> #Customize the x-axis with the years
> axis(1,
+      at=Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,#Position of labels
+      labels=Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,#Labels (years)
+      las=2) #Rotate labels for better readability
>
> #Add the UK female unemployment rate line
> lines(Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,
+       Female_Unemployment_Rates_in_USA_and_UK_1990_2020$UK,
+       col="red") #Set the colour to red for UK
>
> #Add points for USA
> points(Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,
+        Female_Unemployment_Rates_in_USA_and_UK_1990_2020$USA,
+        col="blue",
+        pch=19) #Use solid circle for points
```

```
> #Add points for UK
> points(Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year,
+        Female_Unemployment_Rates_in_USA_and_UK_1990_2020$UK,
+        col="red",
+        pch=19) #Use solid circles for points
>
> #Add a legend to the plot
> legend("topleft",#Position at the top left
+        legend=c("USA","UK"),#Label for the legend
+        col=c("blue","red"),#Colors corresponding to the data
+        lty=1,#line type
+        pch=19) #point type
>
> #Add a grid for better readability
> grid(nx=50,
+      ny=10,
+      col="lightgray",
+      lty="dotted") #Use dotted lines for the grid
```

The code produces the figure:



Appendix A.2

The R code and output for the values in Table 1: Descriptive Statistics for UK Female Unemployment Rates (2019-2020)

```
> uk_rates <-c(3.6,4.3) #UK Female Unemployment rates (2019,2020)
> mean(uk_rates) #Mean (Female Unemployment Rates (2019,2020))
[1] 3.95
> median(uk_rates) #Median (Female Unemployment Rates (2019,2020))
[1] 3.95
> sd(uk_rates) #SD (Female Unemployment Rates (2019,2020))
[1] 0.4949747
> IQR(uk_rates) #IQR (Female Unemployment Rates (2019,2020))
[1] 0.35
>
```

Appendix A.3

The R code and output for the values presented in Table 2: Comparative Statistics of Female Unemployment Rates in the USA and UK (2000-2020)

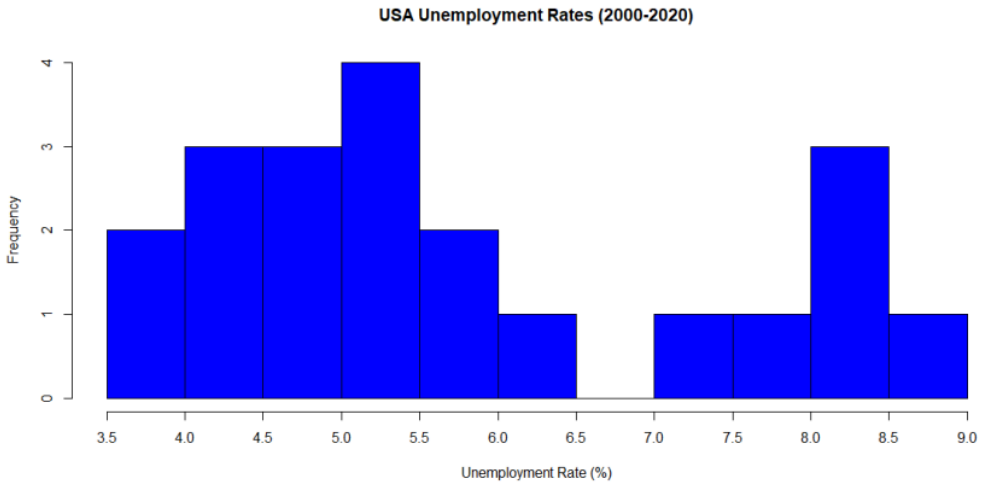
```
> #Filtered the data to show just 2000-2020 Female Unemployment Rate in USA and UK
> filtered_data <- Female_Unemployment_Rates_in_USA_and_UK_1990_2020[
+   Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year >= 2000 &
+     Female_Unemployment_Rates_in_USA_and_UK_1990_2020$Year <= 2020,]
>
> mean(filtered_data$USA) # Mean for USA (2000-2020)
[1] 5.780952
> median(filtered_data$USA) # Median for USA (2000-2020)
[1] 5.4
> sd(filtered_data$USA) # SD for USA (2000-2020)
[1] 1.633897
> IQR(filtered_data$USA) # IQR for USA (2000-2020)
[1] 2.5
>
> mean(filtered_data$UK) # Mean for UK (2000-2020)
[1] 5.22381
> median(filtered_data$UK) # Median for UK (2000-2020)
[1] 4.9
> sd(filtered_data$UK) # SD for UK (2000-2020)
[1] 1.181907
> IQR(filtered_data$UK) # IQR for UK (2000-2020)
[1] 1.5
>
```

Appendix A.4

The R code and output for Figure 2 USA Female Unemployment rates from 2000 to 2020

```
> #Create a histogram for USA Unemployment Rates (2000-2020)
> hist(filtered_data$USA, #Data to plot (USA unemploymentrates)
+      main = "USA Unemployment Rates (2000-2020)", # Main title of the histogram
+      xlab = "Unemployment Rate (%)", # Label for the x-axis
+      ylab = "Frequency", # Label for the y-axis
+      col = "blue", #Fill color for the bars
+      border = "black", #Board color for the bars
+      breaks = seq(3.5,9, by=0.5), # Bin intervals from 3.5 to 9 with 0.5 step
+      xlim = c(3.5,9), # Set limits for the x-axis
+      xaxt="n") #Suppress default x-axis
>
> #Add custom x=axis with tick marks
> axis(1, # Position of the axis (1=bottom)
+      at= seq(3.5,9, by=0.5)) # Custom tick marks from the 3.5 to 9 with 0.5 steps
>
```

The R code produces a figure:

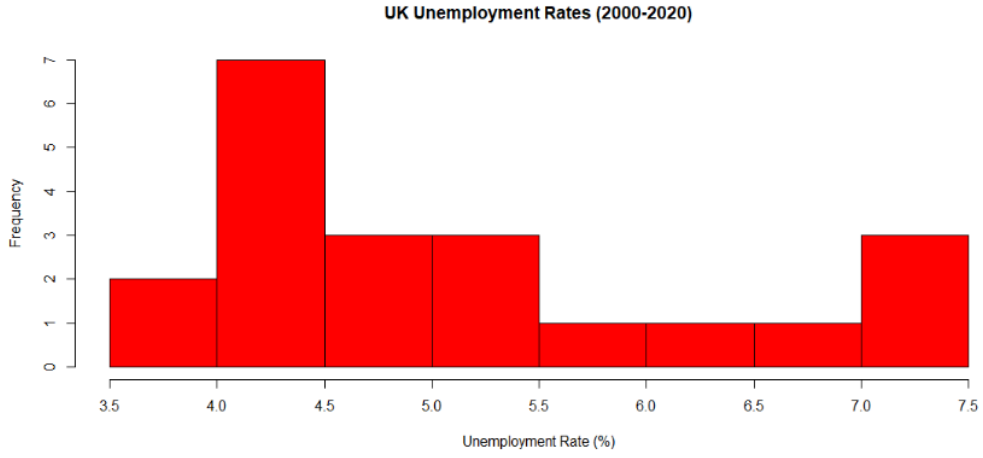


Appendix A.5

The R code and output of Figure 3 UK Female Unemployment Rates (2000-2020)

```
> #Add custom x=axis with tick marks
> axis(1, # Position of the axis (1=bottom)
+      at= seq(3.5,9, by=0.5)) # Custom tick marks from the 3.5 to 9 with 0.5 steps
> #Create a histogram for UK Unemployment Rates (2000-2020)
> hist(filtered_data$UK, #Data to plot (UK Unemployment rates)
+      main = "UK Unemployment Rates (2000-2020)", # Main title of the histogram
+      xlab = "Unemployment Rate (%)", # Label for the x-axis
+      ylab = "Frequency", # Label for the y-axis
+      col = "red", # Fill color for the bars
+      border = "black", # Board color for the bars
+      breaks = seq(3.5, 7.5, by=0.5), # Bins intervals from 3.5 to 7.5 with 0.5 steps
+      xlim = c(3.5,7.5), # Set limits for the x-axis
+      xaxt = "n") # Suppress default x-axis
>
> #Add custom x-axis with tick marks
> axis(1, #Position of the axis (1=bottom)
+      at= seq(3.5,7.5,by=0.5)) # Custom tick marks from the 3.5 to 7.5 with 0.5 steps
```

The R code produces a figure:



Appendix A.6

The R code and output for Table 3: Hypothesis Test Results for Comparing UK Female Unemployment Rates (2019-2020)

```
> # Define the number of unemployed individuals for 2019 and 2020
> unemployed <- c(585000,695000) # Unemployed counts for 2019 and 2020
> # Define the unemployed rates for 2019 and 2020 (convert percentages to proportions)
> rate <- c(3.6,4.3)/100 # Rates are given as percentages, dividing by 100 converts to proportions
> # Calucale the number of eligible individuals in the workforce
> eligible <- unemployed/rate # Workforce size is calculates as unemployment divided by the rate
> # Print the eligible population for 2019 and 2020
> print(eligible) # Displays the calculated workforce sizes
[1] 16250000 16162791
>
> #Perform a two-sample test for proportions
> prop.test(x=unemployed, # Observed counts of unemployed individuals
+           n=eligible, # Corresponding eligible population sizes
+           alternative="two.sided") # Specifies a two-sided test (default)

        2-sample test for equality of proportions with continuity correction

data:  unemployed out of eligible
X-squared = 10468, df = 1, p-value < 2.2e-16
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.007134168 -0.006865832
sample estimates:
prop 1 prop 2
 0.036  0.043
```

Appendix A.7

The R code and output for Table 4: Hypothesis Test Results for Comparing USA and UK Female Unemployment Rates (2000-2020)

```
> #Add custom x-axis with tick marks
> axis(1, #Position of the axis (1=bottom)
+      at= seq(3.5,7.5,by=0.5)) # Custom tick marks from the 3.5 to 7.5 with 0.5 steps
> #Perform a two sample t-test comparing UKA and UK Female unemplyment rates
> t_test_result <- t.test(
+   filtered_data$USA, # First sample: USA Female Unemployment Rates (2000-2020)
+   filtered_data$UK, # Second sample: UK Female Unemployment Rates (2000-2020)
+   alternative="two.sided", # Specifies a two-sided test (default), testing for any difference
+   var.equal = FALSE) # Assumes unequal variances between the two samples (Welch's t-test)
>
> # Print the t-test results to the console
> print(t_test_result) # Outputs the test statistics, p-value, confidence interval and means

        Welch Two Sample t-test

data:  filtered_data$USA and filtered_data$UK
t = 1.2661, df = 36.431, p-value = 0.2135
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.334953  1.449239
sample estimates:
mean of x mean of y
 5.780952  5.223810
```

Coursework Assessment 6 PDF

GRADEMARK REPORT

FINAL GRADE

75/100

GENERAL COMMENTS

Introduction: 15/15 nice explanation of the data and the general context - justified the choice of additional data - gave the structure of the report - explained and referenced where the data comes from - is the same definition of unemployment used in the USA?

Descriptive Statistics and Graphical Summaries: 19/20 inclusion of time series plot and nice discussion of trends over time with their real world context - comparison of the two data sets - boxplots are nice to highlight means, medians, IQR etc

Statistical Methods: 10/15 the test types used are not actually specified or the individual test assumptions - code in appendices is needed to know which tests are applied - test choice of prop.test is correct but justification could be better

Results and Conclusions: 8/20 correct results and interpretation of p-value - very little discussion about interpretation in terms of results and context

Conclusions using additional data: 3/10 same as above

RScript and additional data files: 20/20

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7