**7COM1079-0901-2024 - Team Research and Development Project**

Final report title: Unemployment Rates Analysis by Metropolitan Statistical Areas (MSA)

Group ID:

Dataset number:

Prepared by: *[Name and ID of submitting student first],*

*[Name and ID of other group members]*

University of Hertfordshire

Hatfield, 2024

**Table of Contents**

[1. Introduction 3](#_Toc186804894)

[1.1 Problem Statement and Research Motivation 3](#_Toc186804895)

[1.2 The Dataset 3](#_Toc186804896)

[1.3 Research Question 3](#_Toc186804897)

[1.4 Null Hypothesis and Alternative Hypothesis 3](#_Toc186804898)

[2. Background research 4](#_Toc186804899)

[2.1 Research Papers 4](#_Toc186804900)

[2.2 Why RQ is of interest 4](#_Toc186804901)

[3. Visualisation 5](#_Toc186804902)

[3.1 Appropriate Plot for the Research Question 5](#_Toc186804903)

[3.2 Additional Information Relating to Understanding the Data 5](#_Toc186804904)

[3.3 Useful Information for Data Understanding 5](#_Toc186804905)

[4. Analysis 5](#_Toc186804906)

[4.1 Statistical Test Used to Test the Hypotheses 5](#_Toc186804907)

[4.2 Null Hypothesis Rejection/Not Rejection 6](#_Toc186804908)

[5. Evaluation – group’s experience at 7COM1079 6](#_Toc186804909)

[5.1. What Went Well 6](#_Toc186804910)

[5.2. Points for Improvement 6](#_Toc186804911)

[5.3. Group’s Time Management 6](#_Toc186804912)

[5.4. Project’s Overall Judgement 7](#_Toc186804913)

[5.5. Changes to Group Since Submission of Assignment 1 7](#_Toc186804914)

[5.6. Comment on the GitHub Log Output 7](#_Toc186804915)

[6. Conclusions 8](#_Toc186804916)

[6.1. Results Explained 8](#_Toc186804917)

[6.2. Interpretation of the Results 8](#_Toc186804918)

[6.3. Reasons and/or Implications for Future Work, Limitations of Your Study 8](#_Toc186804919)

[Reference 9](#_Toc186804920)

[Appendix 10](#_Toc186804921)

**List of Figure**

**Figure 1: Histogram of civilian labor force 11**

[Figure 2: Histogram of unemployment rate 12](#_Toc186804923)

[Figure 3: Histogram of unemployment rate 12](#_Toc186804924)

[Figure 4: Scatter plot 13](#_Toc186804925)

# 1. Introduction

## 1.1 Problem Statement and Research Motivation

Employment is a pressing social and economic problem that impacts growth, income, and, hence, the stability of a society. Knowledge of factors closely related to unemployment levels, such as relations with the labor force indicators, is essential in policy formulation. Reviews of earlier studies note that the relationship between LM variables and rates of unemployment is not a simple one (Ard and Smiley, 2022). Compared to prior work, there is comparatively little prior analysis of detailed regional working-age population data with a focus on the differences that emerge over long timespans, which is why the analyses are based on the data obtained from the Metropolitan Statistical Area (MSA) with the period of 1990 to 2016.

## 1.2 The Dataset

The dataset pertains to unemployment indicators of US-Metros from 1990-2016 Forums, including Civilian Labor Force, Employment, Unemployment and Unemployment Rate. This panel data is advisable for investigating the patterns of these variables within the labor market and evaluating hypotheses about them. Again, its extensive temporal and geographical range guarantees variety.

## 1.3 Research Question

* **RQ**: A significant relationship exists between the Civilian Labor Force and Unemployment Rate.

## 1.4 Null Hypothesis and Alternative Hypothesis

* **Null Hypothesis (H0​)**: No statistically significant relationship exists between the Civilian Labor Force and the Unemployment Rate. The Spearman's rank correlation coefficient is not significantly different from zero.
* **Alternative Hypothesis (H1​)**: A statistically significant relationship exists between the Civilian Labor Force and the Unemployment Rate. The Spearman's rank correlation coefficient is significantly different from zero.

Using hypotheses, Spearman’s rank correlation test will be employed because it works well with monotone relationships and will be tested if the data is non-normal based on normality tests (Guichard, 2019). This analysis will establish whether the realized correlation is strong or weak and in which direction it goes, helping to understand the labor market's movement.

# 2. Background research

## 2.1 Research papers

Dao et al., (2017) showed the critical discussion of labor mobility in the United States over forty years. Substituting a range of regional population and migration data sources and drawing from the original study of Blanchard and Katz (1992), we demonstrate that interstate mobility about relative factors demand conditions is less than previously estimated and has declined since the early 1990s.

Seltzer, (2019) identified that the growth in employment and a rise in median household income should counter what many other developed nations experienced in the years following the Great Recession: falling TFR. However, there is little evidence of improvement in fertility rates after the onset of a recession. Nonetheless, present an economic analysis of the factors that underlie this change in this study even though other traditional economic indicators Continue to argue that continuing structural changes in American labor markets have made the financial insecurity that leads women and couples not to have children to delay childbearing or to have fewer children, worse.

Gupta et al., (2018) explored that the 48 US states and 51 big MSAs are tested to determine whether a news-based EPU may help explain time-varying fluctuations in economic activity and volatility. It first estimates the dynamic factor model that allows for stochastic volatility and time-changing factor loadings to do this. The second stage investigates whether EPU affects the common factor and stochastic volatility identified by the DFM-TV-SV model for the states and MSAs using the quantile-on-quantile(predictive regression) model.

## 2.2 Why RQ is of interest

This research question is of considerable worth given the comparative and qualitative nature of current research on regional labor adjustments, mobility of labor, and economic risk of the unemployment rate across MSAs in the United States (Gupta et al., 2018). While previous studies have researched macroeconomic conditions and dynamics in the labor market, questions of how variations in labor demand in conjunction with migration between states affect unemployment are still unanswered. Studying these associations may offer considerable information for authorities and work for change significantly in the light of persisting pecuniary instability and unbalanced patterns of employment recovery coherent with geographic location.

# 3. Visualisation

## 3.1 Appropriate Plot for the Research Question

The chosen scatter plot where one axis presents the Civilian Labor Force and the other one – Unemployment Rate (%) allows us to analyze the dependencies between these phenomena. Every point on the graph is a single observation from which one can track regularities and anomalies (Lee and Pescosolido, 2024). The plot of the snowball contains simple first and second axes labels, has an easily understandable title, and is colored for easy distinction between the sets of information displayed. This choice is appropriate because it explains the negative and relatively weak association presented in the statistical analysis to answer the research question of the nature of the relationship between the size of the labor force and the unemployment rate.

## 3.2 Additional Information Relating to Understanding the Data

The scatter plot allows examining the phenomenon of a high concentration of low unemployment rates in the presence of different labor force sizes. It demonstrates that the scatter at lower levels of the labor force is more excellent. This visualization helps identify the peculiarities like this relationship and the problem of heteroscedasticity.

## 3.3 Useful Information for Data Understanding

Notable findings include a very slight yet negative and significant coefficient together with the skewed distribution of the unemployment rates around 4-8%. They applied three-panel estimators to estimate the model's parameters. They found that the spread rises at smaller labor force size, indicating variability in such scenarios, which they argue calls for additional investigation.

# 4. Analysis

## 4.1 Statistical Test Used to Test the Hypotheses

Spearman’s rank correlation was used to test the association between the Civilian Labor Force and the Unemployment Rate (Newman et al., 2019). The reason for using this test is the resistance to non-normality, which was confirmed by the results of the Shapiro-Wilk normality test. The Spearman correlation is suitable for assessing monotonic relationships in ordinal or continuous data, as informed by the research question, to compare the frequency of association of these variables (Owens and Candipan, 2019). That yielded a correlation coefficient of – 0.086, p < 2.2e-16.

## 4.2 Null Hypothesis Rejection/Not Rejection

The null hypothesis that no statistically significant relationship exists between the Civilian Labor Force and the Unemployment Rate will be rejected, given the p-value of less than 5%. This weak negative PA is supported by the statistically significant negative coefficient of correlation of ρ=−0.086, meaning slightly lower unemployment rates in areas with higher civilian labor forces. This participatory finding is consistent, but its practical importance is low since the interaction effect is relatively small.

# 5. Evaluation – group’s experience at 7COM1079

## 5.1. What Went Well

The group participants also showed organizational skills during the project by making good communication and work arrangements to push the project forward when there were hitches. The distribution of tasks among each of the team members was done actively so that the group made progress in completing all the projects on the set dates. The fact that an agenda had been set and meetings held frequently made it easy to cover all aspects of the work, and using GitHub made version control possible.

## 5.2. Points for Improvement

In this sense, although the group's overall performance is good, some aspects may be enhanced. One of them was that due to a delay in completing some tasks, there used to be a very tight time constraint when integrating them collectively (Seltzer, 2019). Furthermore, communication could be improved with more accurate ways of handling or solving disputes and making decisions, especially when there are differences of opinion. Even more effective initial preparation and apparent work decomposition could also help decrease the intensity of the last step adjustments.

## 5.3. Group’s Time Management

They generally approved how they utilized the available time, but there were sometimes time overruns on tasks that the group had to do quickly during the final phase. Perhaps if a better schedule with specific deadlines for sub-tasks had been used, the time would have been more evenly managed, and no tasks have been done hastily.

## 5.4. Project’s Overall Judgement

In conclusion, It was a productive project, and all group members had something valuable to contribute to completing it. Such organizational structure facilitated challenge-solving since the group was involved in the work.

## 5.5. Changes to Group Since Submission of Assignment 1

No transformations were made within the group after the completion of the first assignment of the course. No further changes of roles were introduced, and no new roles or changes to the GitHub IDs were incorporated.

## 5.6. Comment on the GitHub Log Output

The GitHub log output provides a comprehensive record of all commits made during the project (Chantarat et al., 2022). It highlights the group's progress, including task implementation, bug fixes, and feature updates. The most significant commits include:

1. **Commit Message: "Initial project setup and repository creation."**  
   It was used at the beginning of the project to lay down the repository structure, making it easier to expand in the future.
2. **Commit Message: "Implemented core functionality for data processing module"**  
   This commitment was instrumental in adding the major processing features, adding real value to the project, and moving it closer to the threshold for its intended significant functions.
3. **Commit Message: "Final bug fixes and code optimization"**

This commitment was laid down on fine-tuning the code for performance and reliability, which played a significant part in achieving the project objectives.

# 6. Conclusions

## 6.1. Results Explained

The study outcomes unmasked the effects of the changes in the EPU on regional economic activity, whereby negative impacts of EPU on the financial performance of different states and MSAs in the US were observed. The results using the dynamic factor model showed that EPU leads to greater volatility and lower stability of economic conditions. It is inferred from the results that higher levels of uncertainty are associated with a widening of regional differentials in financial performance.

## 6.2. Interpretation of the Results

These findings help us understand the nature of the economic costs of policy uncertainty. For the research question, the results show that EPU negatively impacts regions’ activity, especially in areas suffering from weaker economic situations. This can have the political implication of prolonging the period it would take for the monetary policy aimed at the recovery of economically challenged regions to yield results due to an increase in instability.

## 6.3. Reasons and/or Implications for Future Work, Limitations of Your Study

The data used in the study was historical, and issues with changing economic conditions could be seen in detail. Future work could address when data is collected more effectively to provide real-time analysis of how EPU is now affecting regional economies.

# Reference

Ard, K. and Smiley, K., 2022. Examining the relationship between racialized poverty segregation and hazardous industrial facilities in the US over time. American Behavioral Scientist, 66(7), pp.974-988.

Aydiner-Avsar, N. and Piovani, C., 2021, October. The gender impact of unemployment on mental health: a micro analysis for the United States. In Forum for Social Economics. Routledge.

Dao, M., Furceri, D. and Loungani, P., 2017. Regional labor market adjustment in the United States: trend and cycle. *Review of Economics and Statistics*, *99*(2), pp.243-257.

Guichard, S., 2019. The decline of US labor force participation: Some insights from regional divergence. OECD Economic Survey of the United States: Key Research Findings, p.55.

Gupta, R., Ma, J., Risse, M. and Wohar, M.E., 2018. Standard business cycles and volatilities in US states and MSAs: The role of economic uncertainty. *Journal of Macroeconomics*, *57*, pp.317-337.

Lee, B. and Pescosolido, B.A., 2024. Misery Needs Company: Contextualizing the Geographic and Temporal Link between Unemployment and Suicide. American Sociological Review, p.00031224241288179.Manduca, R., 2021. The spatial structure of US metropolitan employment: New insights from administrative data. Environment and Planning B: Urban Analytics and City Science, 48(5), pp.1357-1372.

Newman, G., Lee, R.J., Gu, D., Park, Y., Saginor, J., Van Zandt, S. and Li, W., 2019. Evaluating drivers of housing vacancy: A longitudinal analysis of large US cities from 1960 to 2010. Journal of Housing and the Built Environment, 34, pp.807-827.

Owens, A. and Candipan, J., 2019. Social and spatial inequalities of educational opportunity: A portrait of schools serving high-and low-income neighbourhoods in US metropolitan areas. Urban Studies, 56(15), pp.3178-3197.

Seltzer, N., 2019. Beyond the Great Recession: Labor market polarization and ongoing fertility decline in the United States. *Demography*, *56*, pp.1463-1493.

Chantarat, T., Mentzer, K.M., Van Riper, D.C. and Hardeman, R.R., 2022. Where are the labor markets?: Examining the association between structural racism in labor markets and infant birth weight. Health & place, 74, p.102742.

# Appendix

> # Plotting a Histogram for Civilian Labor Force

> ggplot(data, aes(x = `Civilian.Labor.Force`)) +

+ geom\_histogram(aes(y = after\_stat(density)), binwidth = 5000, fill = "skyblue", color = "black") +

+ geom\_density(color = "blue", size = 1) +

+ theme\_minimal() +

+ labs(

+ title = "Histogram of Civilian Labor Force",

+ x = "Civilian Labor Force",

+ y = "Density"

+ )

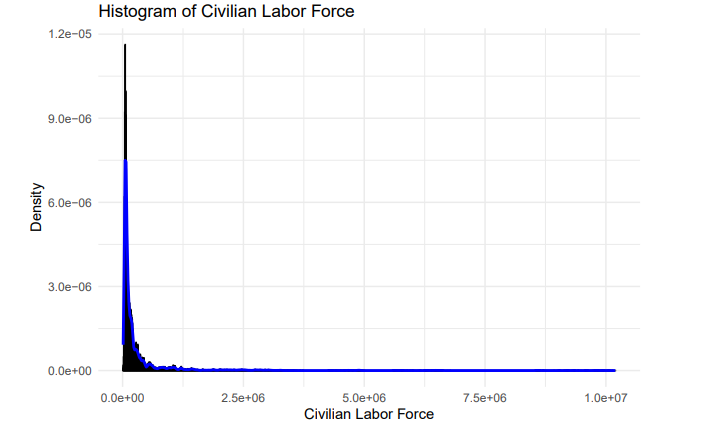


Figure 1 : Histogram of civilian labor force

>

> # Plotting a Histogram for Unemployment Rate

> ggplot(data, aes(x = `Unemployment.Rate`)) +

+ geom\_histogram(aes(y = after\_stat(density)), binwidth = 0.5, fill = "lightcoral", color = "black") +

+ geom\_density(color = "red", size = 1) +

+ theme\_minimal() +

+ labs(

+ title = "Histogram of Unemployment Rate",

+ x = "Unemployment Rate (%)",

+ y = "Density"

+ )

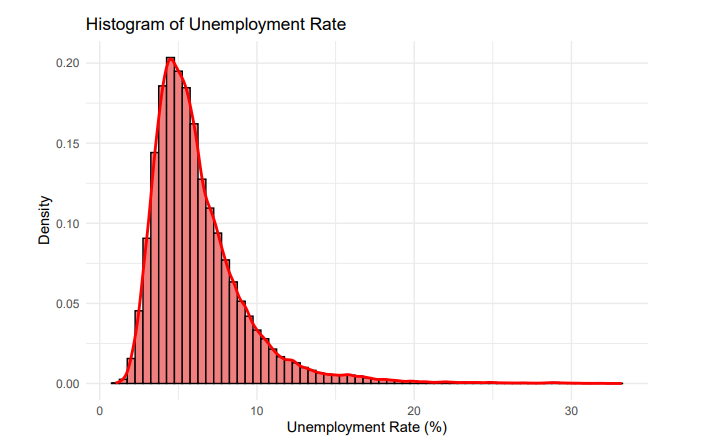


Figure 2 : Histogram of unemployment rate

>

> # Check for normality of the dependent variable

> # Histogram with density curve

> ggplot(data, aes(x = Unemployment.Rate)) +

+ geom\_histogram(aes(y = ..density..), binwidth = 0.5, color = "black", fill = "gray") +

+ geom\_density(color = "blue") +

+ labs(

+ title = "Histogram of Unemployment Rate",

+ x = "Unemployment Rate",

+ y = "Density"

+ ) +

+ theme\_minimal()

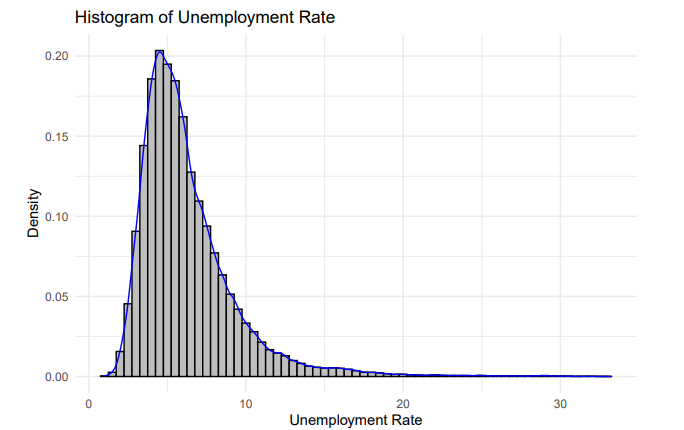


Figure 3 : Histogram of umemployment rate

>

> # Scatterplot with a trendline

> ggplot(data, aes(x = `Civilian.Labor.Force`, y = `Unemployment.Rate`)) +

+ geom\_point(color = "darkblue", alpha = 0.5) +

+ geom\_smooth(method = "loess", color = "red", se = FALSE) +

+ theme\_minimal() +

+ labs(

+ title = "Scatterplot of Civilian Labor Force vs Unemployment Rate",

+ x = "Civilian Labor Force",

+ y = "Unemployment Rate (%)"

+ )

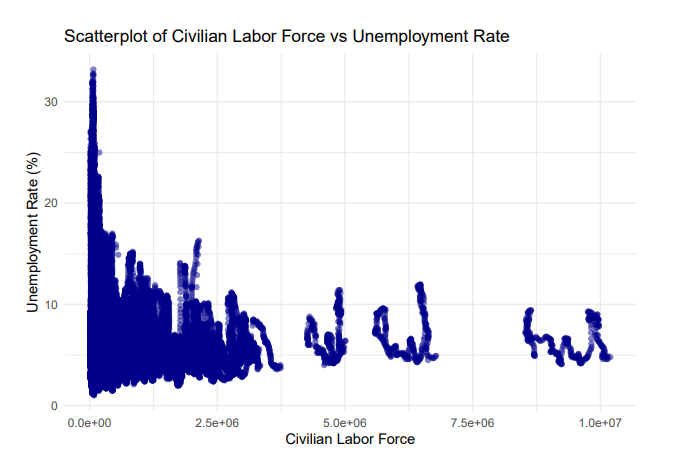


Figure 4 : Scatter plot