Paper Assignment #3

Fall Stat Measuring & Modeling Data

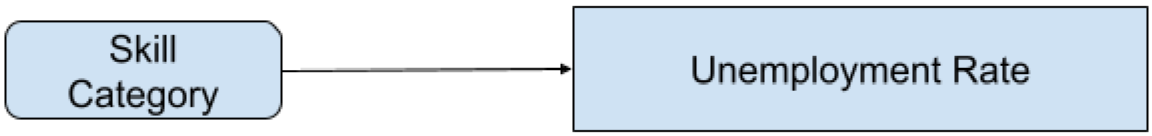
Examining the association between skill category and unemployment rates in the U.S

This study will investigate the relationship between skill categories and unemployment rates in the U.S. and look at how economic impacting forces like covid 19, and technological changes, such as AI and automation, can impact workers differently based on their education and skill levels. Specifically, this analysis focuses on whether low-skilled workers experience consistently higher ranges and higher general unemployment rates compared to medium- and high-skilled workers. The study uses the cpsaat07 dataset from the Bureau of Labor Statistics (BLS), examining unemployment trends from 2017 to 2023. Although direct data on AI adoption is unavailable, this study aims to identify patterns that could reflect broader inequalities in how technological changes affect employment.

Theory

Existing research suggests that AI and automation disproportionately affect low-skilled workers due to their concentration in routine and manual tasks. One article describes AI as "routine-biased technological change on steroids," emphasizing how AI adds intelligence to automation tools and replaces people in routine and increasingly nonroutine cognitive tasks (Tyson & Zysman, 2022, p. 256). Historical trends in automation reveal that technological shifts often disproportionately benefit high-skilled workers, while low-skilled roles are increasingly at risk. This is known as capital-skill complementarity, and demonstrates how technological advancements favor high-skilled workers, exacerbating inequalities in the labor market (Ernst, Merola, & Samaan, 2019, p. 10). Research also shows that job polarization is occurring, with middle-skill occupations declining while high-skill roles expand as a share of total employment (Tyson & Zysman, 2022, p. 256). Low-skilled workers are at the greatest risk of job displacement, particularly in sectors like manufacturing, retail, and logistics. Some researchers argue that AI may also create new jobs through productivity gains, potentially offsetting job losses and resulting in a net positive impact on employment (Petropoulos, 2018, p. 121). The economic impact of the COVID-19 pandemic further exposed differences in the labor market. In 2020, 80% of job losses were among the lowest quarter of wage earners, disproportionately affecting Black, Hispanic, and Asian American and Pacific Islander (AAPI) workers compared with white workers (Gould, 2021).Gould's analysis used the Economic Policy Institute's (EPI) Current Population Survey (CPS) Extracts, which split the data into two periods: before the pandemic and during the pandemic downturn. This method allowed for detailed analysis of job losses across occupations, wage levels, and demographic groups, highlighting disparities tied to systemic inequalities. This uneven distribution of job losses is tied to occupational segregation, where low-wage workers are overrepresented in industries like leisure and hospitality and underrepresented in higher-paid management professions, making them more vulnerable during economic downturns (Gould, 2021). These findings show how systemic inequalities intersect with labor market trends, leaving marginalized groups vulnerable to both economic shocks and technological changes. The inequality in job opportunities also raises concerns about marginalized groups, who frequently lack access to the resources and education needed to benefit from technological advancements. These communities are less likely to participate in the development of AI, further creating cycles of inequality. This study examines how unemployment rates vary among skill categories (low, medium, and high) between 2017 and 2023, exploring whether these trends align with theoretical predictions about the disproportionate impacts on low-skilled workers due to economic and technological changes.

Path Diagram



Skill Category:

This is the independent variable and is a categorical variable with three levels:

Low-Skilled: This category includes individuals who have either not completed high school or have only a high school diploma.

Medium-Skilled: This category includes individuals who have completed some college education but did not earn a bachelor's degree, as well as those with an associate's degree.

High-Skilled: This category includes individuals with a bachelor's degree or higher (e.g., master's, doctoral, or professional degrees).

Unemployment Rate:

This is the dependent variable and is a numerical variable that measures the percentage of people within each skill category who are unemployed. It represents the economic vulnerability of each skill level and allows for comparisons of unemployment trends across the three groups.

Year:

This is an independent variable and is a numerical variable representing the year for each data point in the analysis. The years range from 2017 to 2023, providing a time frame to examine unemployment trends and determine whether certain economic events, like the COVID-19 pandemic, disproportionately impacted specific skill categories.

Hypothesis

The study is that there’s a relationship between skill category and unemployment, and that the lower skilled category is more likely to experience higher unemployment rates.

Measurement

This study looks at the relationship between skill categories (low, medium, and high) and unemployment rates in the U.S. from 2017 to 2023. The data comes from the Bureau of Labor Statistics (BLS) cpsaat07 dataset, and there’s a data set for each year. To prepare the data for analysis, I cleaned and organized the raw datasets. Empty rows and unnecessary columns were removed because of how the data was originally formatted. Column names, such as “less than high school” and “high school,” were renamed. Education levels were then combined into a single column called "education" and grouped into three skill categories: low, medium, and high. A "year" column was added to link each row to its respective year, making it possible to track trends over time. These steps helped organize the dataset so I could focus on the relationship between skill categories and unemployment rates. Skill categories are the independent variable and were divided into three groups:

- Low-skilled workers: People who didn’t finish high school or only have a high school diploma.
- Medium-skilled workers: People who attended college but didn’t complete a bachelor’s degree or have an associate’s degree.
- High-skilled workers: People with a bachelor’s degree or higher, like a master’s or doctorate.

The dependent variable is the unemployment rate, which is the percentage of people unemployed within each skill category. This helps compare how unemployment affects each group differently. The year is another variable in the study, representing the time frame from 2017 to 2023. This allows for an analysis of unemployment trends over time, including during events like the COVID-19 pandemic.

Statistical Analysis

This study analyzes the relationship between skill categories (low, medium, and high) and unemployment rates in the U.S. from 2017 to 2023. The data was examined using visualizations and statistical tests to identify trends and differences between the skill categories. Descriptive statistics showed that unemployment rates ranged from as low as 1.7% to as high as 11.7%, with low-skilled workers consistently facing higher unemployment rates compared to medium- and high-skilled workers. The average unemployment rate was 4.23% across the years, but low-skilled workers had much more variation in their rates than the other groups. A line graph was created to show unemployment trends over time for each skill category. The graph revealed that low-skilled workers consistently had the highest unemployment rates throughout the period, while high-skilled workers had the lowest. All skill categories saw a sharp rise in unemployment in 2020 due to the economic impacts of the COVID-19 pandemic. Although unemployment rates began to decline after 2020, low-skilled workers continued to face the greatest challenges, with rates remaining higher than those of medium- and high-skilled workers. A box plot comparing unemployment rates across skill categories confirmed these findings, showing that low-skilled workers not only had higher unemployment rates but also experienced greater variability. High-skilled workers, on the other hand, had consistently low unemployment rates with little variation. Additionally, a histogram demonstrated the broader distribution of unemployment rates among low-skilled workers compared to the other groups. To determine whether the differences in unemployment rates across skill categories were statistically significant, an ANOVA (Analysis of Variance) test was conducted. The results showed a statistically significant difference between the groups, with a p-value of 3.44e-05, well below the 0.001 significance level. This indicates that the observed differences in unemployment rates between the skill categories are very unlikely to have occurred by chance.

Figure 1

```
summary(education_unemployment)
```

```
##   education          unemployment_rate      year          skill_category
##  Length:42          Min.   : 1.700    Min.   :2017   High Skill  :14
##  Class :character   1st Qu.: 2.550    1st Qu.:2018   Medium Skill:14
##  Mode  :character   Median : 3.700    Median :2020   Low Skill   :14
##                     Mean   : 4.229    Mean   :2020
##                     3rd Qu.: 5.500    3rd Qu.:2022
##                     Max.   :11.700    Max.   :2023
```

```
summary(education_unemployment$unemployment_rate)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.700   2.550   3.700   4.229   5.500  11.700
```

Figure 2

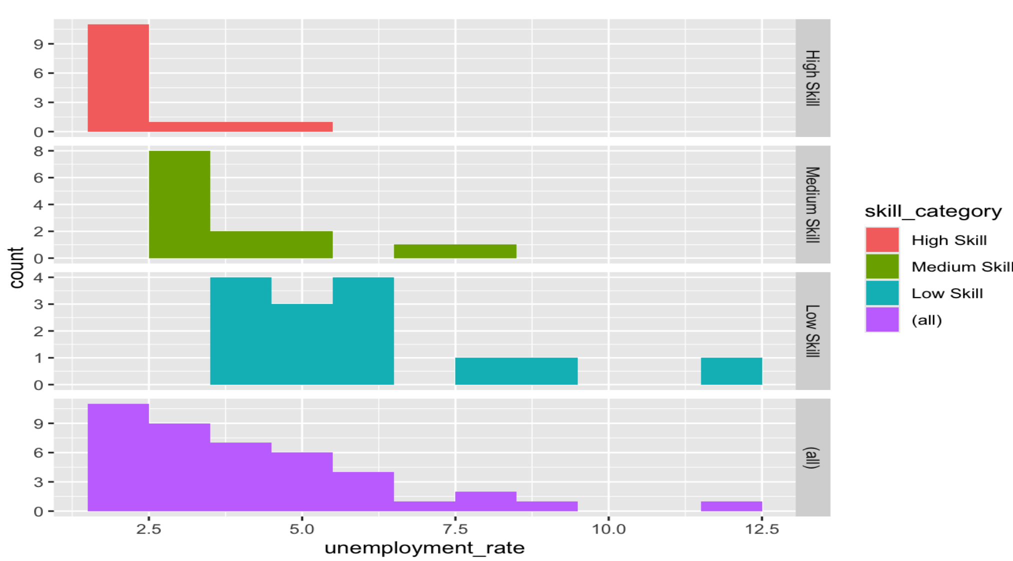


Figure 3

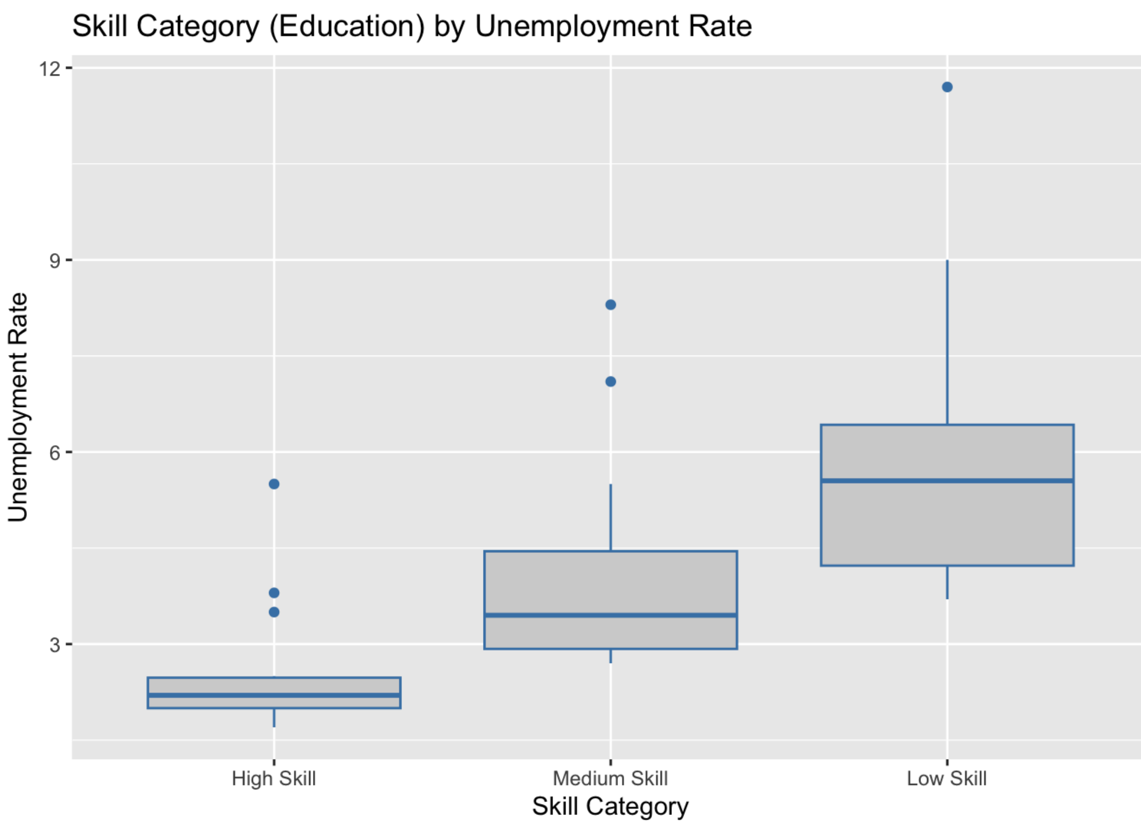
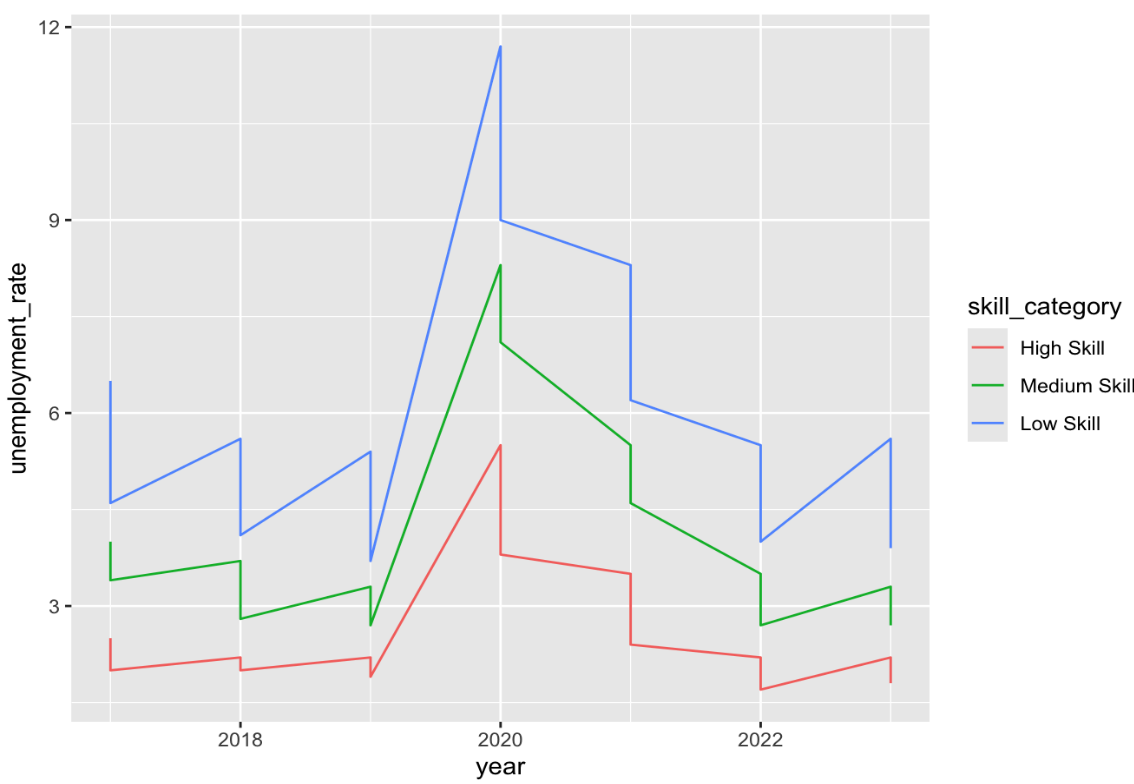


Figure 4



Conclusion

The results of this study show a significant relationship between skill categories and unemployment rates in the U.S. from 2017 to 2023. The findings highlight that low-skilled workers consistently face higher unemployment rates and greater variability compared to medium- and high-skilled workers. During the COVID-19 pandemic in 2020, all skill categories experienced a sharp rise in unemployment, with low-skilled workers having the highest slope in increase. Figure 4 shows how unemployment rates vary across skill categories over time, emphasizing the persistent challenges faced by low-skilled workers, especially during economic disruptions. The box plot in Figure 3 provides additional evidence, showing that low-skilled workers not only experience higher unemployment rates on average but also have the widest range of rates. Figure 2 shows the broader distribution of unemployment rates among low-skilled workers compared to medium- and high-skilled workers. The ANOVA test confirms that these differences are statistically significant, with a p-value of 3.44e-05, meaning that the observed disparities in unemployment rates between skill categories are unlikely to have occurred by chance. These findings align with existing research suggesting that low-skilled workers are more economically vulnerable due to their concentration in routine and manual jobs, which are more susceptible to economic and technological disruptions. Future research could benefit from incorporating more quantitative data on the adoption of AI and automation to examine how these technological shifts might disproportionately impact low-skilled workers. This study's findings suggest that low-skilled workers are particularly at risk, as they seem to face the worst consequences during periods of economic change. A deeper understanding of how emerging technologies will affect job opportunities and stability for different skill categories is essential for developing policies and strategies that can reduce these risks and promote equity in outcomes in the labor market.

References

Gould, E. (2021, May 20). Low-wage, low-hours workers were hit hardest in the COVID-19 recession: The State of Working America 2020 Employment Report. Economic Policy Institute. https://www.epi.org/publication/swa-2020-employment-report/

Tyson, L. D., & Zysman, J. (2022). Automation, AI & Work. Daedalus, 151(2), 256–271. https://www.jstor.org/stable/48662040

Dellot, B., & Balaram, B. (2018). MACHINE LEARNING. RSA Journal, 164(3 (5575)), 44–47. https://www.jstor.org/stable/26798354

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What Can Machines Learn and What Does It Mean for Occupations and the Economy? AEA Papers and Proceedings, 108, 43–47. https://www.jstor.org/stable/26452702

Cardenas, S., & Vallejo-Cardenas, S. F. (2019). Continuing the conversation on how structural racial and ethnic inequalities affect AI biases. 2019 IEEE International Symposium on Technology and Society (ISTAS). https://doi.org/10.1109/istas48451.2019.8937853

Petropoulos, G. (2018). The impact of artificial intelligence on employment. https://www.researchgate.net/profile/Amy-Healy/publication/325796290_How_to_escape_the_low_learning_trap_in_a_runaway_labour_market/links/5b3f7edaaca27207851e87fd/How-to-escape-the-low-learning-trap-in-a-runaway-labour-market.pdf#page=142

Ernst, Ekkehardt, Rossana Merola, and Daniel Samaan. "Economics of Artificial Intelligence: Implications for the Future of Work." *IZA Journal of Labor Policy*, vol. 9, no. 4, 2019. https://sciendo.com/article/10.2478/izajolp-2019-0004