

Are people really struggling to find a job?*

Adrian Ly

March 25, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

1	Introduction	2
2	Data	2
2.1	Data Measurement	2
2.2	Data Analysis	2
3	Model	2
3.1	Model set-up	3
3.2	Model justification	4
4	Results	4
5	Discussion	8
5.1	First discussion point	8
5.2	Second discussion point	8
5.3	Third discussion point	8
5.4	Weaknesses and next steps	8
6	References	9

*Code and data are available at: [LINK](#)

1 Introduction

2 Data

2.1 Data Measurement

2.2 Data Analysis

3 Model

The goal of our modeling strategy, utilizing multiple linear regression, is multifaceted:

- **Understanding Relationships:** To understand the relationship between several independent variables, such as education level and age groups, and the dependent variable, which in this context is the youth unemployment rates. The model is designed to isolate the effect of each variable on unemployment rates.
- **Quantifying Impact:** To quantify the impact of each predictor. Regression coefficients provide a numerical value that represents the expected change in the dependent variable for a one-unit change in an independent variable, all else being equal.
- **Predictive Analysis:** To develop a predictive framework that can be employed to estimate or forecast the outcome variable, in this case, unemployment rates, when provided with new data for the predictors.
- **Inferential Statistics:** To conduct hypothesis testing to ascertain whether the relationships observed within the data are statistically significant and not merely due to random variation.
- **Policy Decision Support:** To offer empirical evidence that may inform policy decisions. For instance, if a correlation is found between higher education levels and increased unemployment, it might indicate the necessity for policies that better align education with employment opportunities.
- **Data-Driven Insights:** To extract insights from the data that could lead to an enhanced understanding of labor market dynamics, such as how different age groups' unemployment rates interact.
- **Modeling Assumptions Testing:** To validate the assumptions underlying the regression model, ensuring the reliability and accuracy of the model's predictions.
- **Confounding Variable Control:** To account for potential confounders that might influence the relationship between education level and unemployment rates. The inclusion of relevant variables aims to minimize estimation bias.

- **Sensitivity Analysis:** To comprehend the model's sensitivity to changes in inputs, which assists in evaluating the robustness of the model.

This strategy is designed to produce a model that is both explanatory and predictive, offering a current snapshot of the relationships in the data as well as a tool for future forecasting. Such a model can be utilized to guide decision-making processes, or pinpoint where interventions may be effective.

3.1 Model set-up

$$\begin{aligned}\text{Youth Unemployment} = & \beta_0 + \beta_1 \cdot \text{Education Level} \\ & + \beta_2 \cdot \text{Reference Period} \\ & + \beta_3 \cdot \text{Adult Unemployment} \\ & + \beta_4 \cdot \text{Senior Unemployment} + \epsilon\end{aligned}$$

Where:

- **Youth Unemployment** is the dependent variable we are trying to predict. These are generally categorized as people who are between the ages of 15-24.
- **Education Level** is an independent variable representing the level of education. (e.g., high school, undergraduate, postgraduate)
- **Reference Period** is another independent variable that represents the time frame of the data collection (e.g., year).
- **Adult Unemployment** and **Senior Unemployment** are independent variables representing unemployment rates for adult and senior age groups, with people ages ranging between 25-44 and above 45, respectively.
- β_0 is the y-intercept, representing the expected value of **Youth Unemployment** when all the independent variables are 0.
- $\beta_1, \beta_2, \beta_3, \beta_4$ are the coefficients for each independent variable, representing the change in **Youth Unemployment** for a one-unit change in the respective independent variable, holding all other variables constant.
- ϵ represents the error term, accounting for the variability in **Youth Unemployment** not explained by the model.

This model aims to understand how various factors contribute to youth unemployment. The coefficients $\beta_1, \beta_2, \beta_3, \beta_4$ provide insights into the relationship between each independent variable and youth unemployment:

- β_1 tells us how changes in the education level might affect youth unemployment, holding other factors constant. A positive coefficient suggests that higher education levels are associated with higher youth unemployment rates, while a negative coefficient suggests the opposite.

- β_2 gives insight into how changes over the reference period (e.g., years) impact youth unemployment, which can help identify trends over time.
- β_3 and β_4 help understand how unemployment rates in other age groups (adults and seniors) are related to youth unemployment. This can reveal whether higher unemployment rates in these groups are associated with higher or lower youth unemployment rates.

The intercept β_0 provides the baseline level of youth unemployment when all other variables are zero, which may not always be a practical scenario but is necessary for the model’s mathematical formulation. The error term ϵ captures the model’s limitations and any random factors affecting youth unemployment that aren’t included in the model.

3.2 Model justification

The utilization of a multiple linear regression model for analyzing the relationship between people’s unemployment status, age groups, and education level over time is substantiated by various rationales. Primarily, the model adeptly accommodates multiple independent variables such as education level and age group, each potentially exerting influence on the dependent variable, unemployment status. This capacity allows for an assessment of the individual impact of each predictor while simultaneously controlling for the influence of others. Such a model not only facilitates the quantification of the relationships between these factors and unemployment—via coefficients that express the extent of change in response per unit change in a predictor (with other variables held constant)—but also enables hypothesis testing on the parameters. This testing is instrumental in determining whether the associations between predictors and the response are statistically significant, thereby offering insights into the meaningfulness of these relationships.

4 Results

Our results are summarized in Table 1.

Table 2: Percentage Change in Youth Unemployment by education level (2019-2023)

Education Level	Unemployment Rate (2019)	Unemployment Rate (2023)	Percentage Change (%)
0 to 8 years	17.5	18.6	6.29
Above bachelor’s degree	10.3	9.3	-9.71
Bachelor’s degree	7.7	8.5	10.39
High school graduate	10.9	11.2	2.75

Table 2: Percentage Change in Youth Unemployment by education level (2019-2023)

Education Level	Unemployment Rate (2019)	Unemployment Rate (2023)	Percentage Change (%)
Postsecondary certificate or diploma	6.7	6.4	-4.48
Some high school	16.8	16.8	0.00
Some postsecondary	9.8	9.1	-7.14
Total, all education levels	10.7	10.8	0.93
University degree	8.0	8.6	7.50

Table 3: Percentage Change in Youth Unemployment by education level (2022-2023)

Education Level	Unemployment Rate (2022)	Unemployment Rate (2023)	Percentage Change (%)
0 to 8 years	15.0	18.6	24.00
Above bachelor's degree	8.8	9.3	5.68
Bachelor's degree	7.6	8.5	11.84
High school graduate	10.6	11.2	5.66
Postsecondary certificate or diploma	6.7	6.4	-4.48
Some high school	14.9	16.8	12.75
Some postsecondary	8.5	9.1	7.06
Total, all education levels	10.1	10.8	6.93
University degree	7.7	8.6	11.69

Table 1: Regression Analysis: Impact of Education Level and Age on Youth Unemployment

	(1)
(Intercept)	328.01 (368.38)
education_levelAbove bachelor's degree	5.14*** (1.40)
education_levelBachelor's degree	4.14** (1.37)
education_levelHigh school graduate	−0.41 (1.16)
education_levelPostsecondary certificate or diploma	1.24 (1.29)
education_levelSome high school	−0.95 (1.10)
education_levelSome postsecondary	−2.20+ (1.28)
education_levelTotal, all education levels	4.55*** (1.25)
education_levelUniversity degree	4.27** (1.37)
reference_period	−0.16 (0.18)
adult_unemployment	2.19*** (0.31)
senior_unemployment	−0.20 (0.37)
Num.Obs.	45
R2	0.923
R2 Adj.	0.898
Log.Lik.	−78.222
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001	

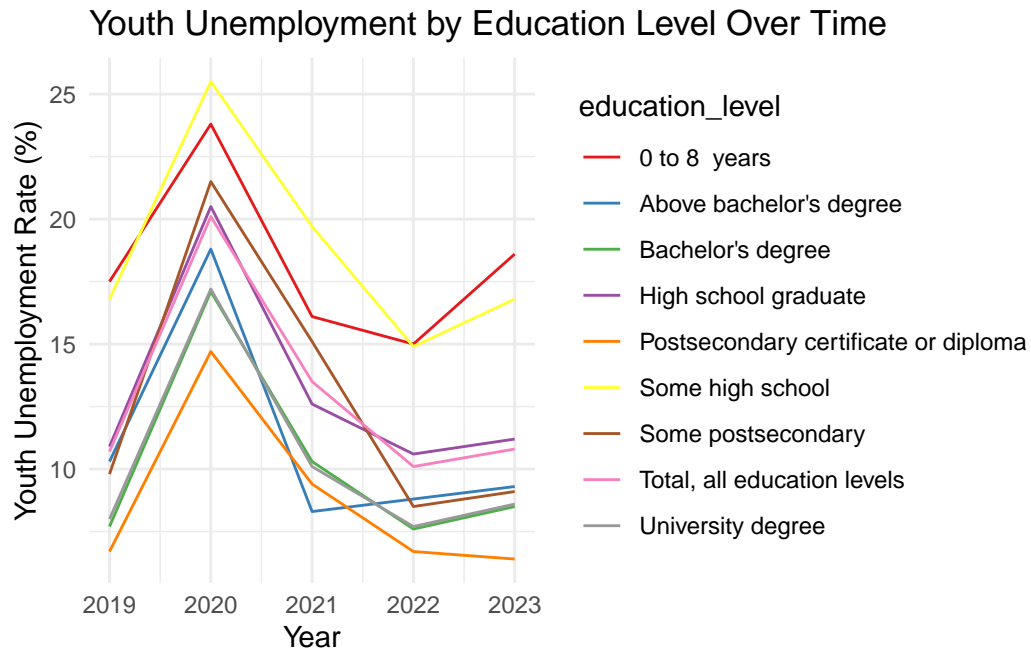


Figure 1: Youth Unemployment Rates

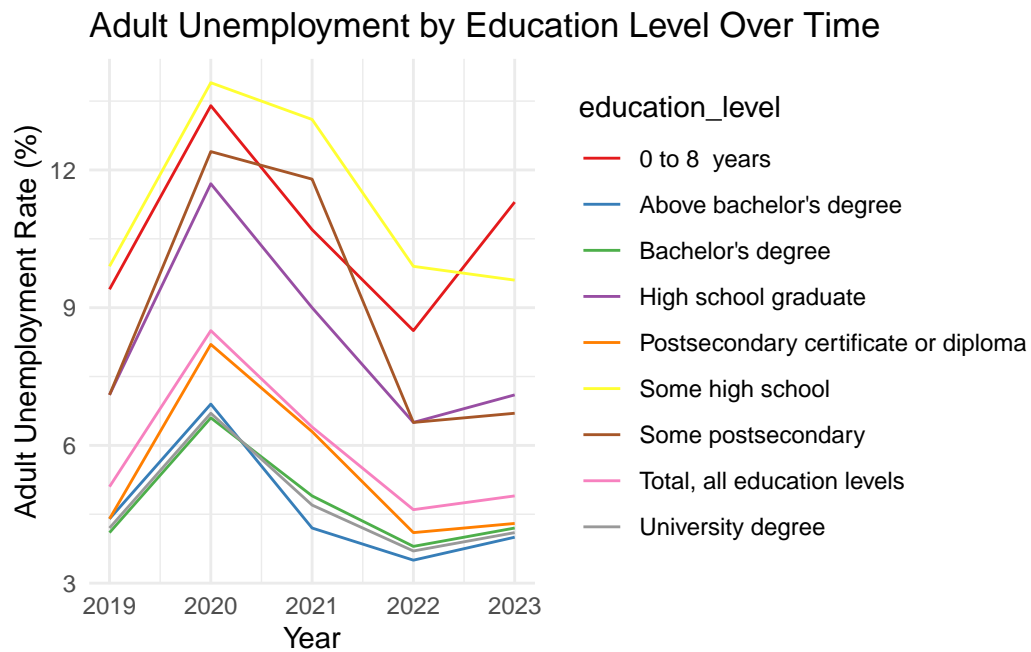


Figure 2: Adult Unemployment Rates

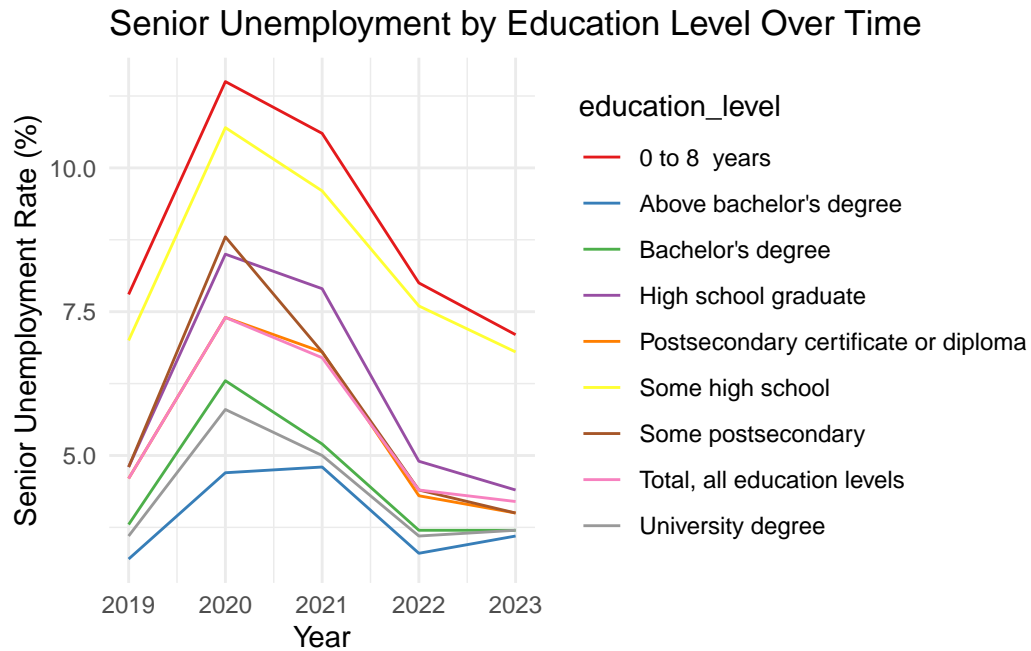


Figure 3: Senior Unemployment Rates

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

6 References