Armen Tamrazyan, Hmayak Paravyan

Course Project

Analysis of how education level, job-education alignment, and wage levels vary across demographic groups in Yerevan.

Statistics A (CS108)

American University of Armenia

10/05/2025

Outline:

1. Introduction
2. Research question
3. Data Description
4. Exploratory Data Analysis (EDA)
5. Methodology
6. Discussion
7. Conclusion
8. Appendix

Introduction

The labor market in Yerevan, like many urban centers, is influenced by a complex interplay of factors, including education, job availability, and economic conditions. Understanding how education level affects wage levels and whether individuals' jobs match their qualifications is crucial for addressing employment disparities and fostering inclusive economic growth. This research aims to examine how job-education alignment and education levels influence wages in Yerevan, the capital of Armenia. The findings of this study provide insights into the structural issues that may affect employment opportunities and wages, particularly for individuals with different levels of education.

Research question

How do employment rates, wage levels, and job-education alignment vary across demographic groups in Yerevan? In particular, does having a job that matches one's educational qualification lead to better wages, and how is this influenced by the level of education attained?

The importance of the question addressed:

- The labor market in Yerevan is not experienced equally by all residents. Employment opportunities, income levels, and the alignment of jobs with educational qualifications can vary significantly based on demographic factors such as age, education level, and other socio-economic characteristics.
- By investigating how these factors influence job market outcomes, this research aims to highlight disparities and inform policies that can promote greater equity in employment. Understanding these patterns will provide insights into how to create a more inclusive labor market, where access to fair wages and job opportunities is not determined by background or educational attainment.

Data Description

We use the 2023 anonymized **Labour Force Survey (LFS)** microdata for Armenia obtained from the Statistical Committee Republic of Armenia (Armstat), which contains about 27,000 individual-level observations collected across all regions of the country.

First of all, among 190 attributes in our dataset, we filtered only 11 + 1 = 12 attributes represented in the table below.

Variable Name	Type	Role
ID_mem	Numeric (Integer)	Unique identifier for each individual
Region	Categorical	Demographic grouping (e.g., Yerevan)
Age Range	Categorical	Age group classification
Education level	Categorical	Highest level of education attained
current job/ job sphere	Categorical	Job sphere or current occupation
Exact Salary	Numeric (Float)	Monthly salary in local currency (AMD)
Salary in Range	Categorical	Salary range (added)
Education Match Status	Categorical	Whether the job matches educational qualification (Yes/No)
Full Time or Part Time	Categorical	Employment type (full-time or part-time)
Worked Hours Per Week in Last Month	Numeric (Float)	Number of hours worked per week
Steps to Find Job	Categorical	Steps undertaken to find a job (if available)
Obstacles to Find Job	Categorical	Barriers or obstacles to finding a job

Filtrations

In Excel:

1. The 2023 anonymized Labour Force Survey (LFS) microdata for Armenia consists of approximately 27,000 rows (representing individuals) and around 190 attributes. However, for the purpose of this research, only a subset of the data was necessary.

Therefore, we conducted a manual filtration process, removing irrelevant attributes that were not directly related to our research questions. Additionally, since the data was coded, i.e, Region was named "A3", Yerevan was named "1", after filtration, we changed these codes into their meaningful names.

2. Since the focus of this study is on Yerevan, we filtered the dataset to include only individuals residing in Yerevan, excluding data from other regions. This process ensured that our analysis was specifically focused on the labor market characteristics and trends within Yerevan, aligning with the objectives of our research.

In R:

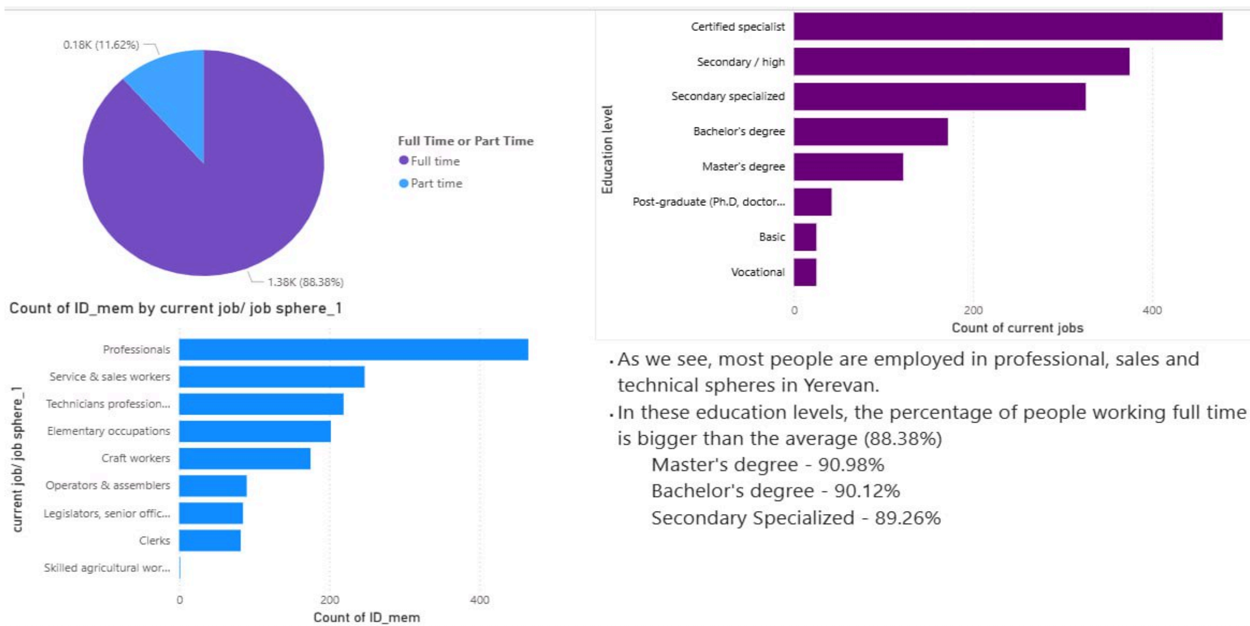
1. Assigning Salary Ranges to each row: While we had the exact salary and did not have its range, we assigned the range of that salary in the Salary Range column for later analysis.
2. Filtered only those rows where Work Sphere is not null: to get the people who work in our sample
3. Removed Unnecessary Columns: Obstacles to finding a job and ways to finding a job are mentioned only by people who do not work, so for that, we created a separate table in R called filtered_sample (for EDA) by removing those columns from our final sample.

Exploratory Data Analysis (EDA)

In the **Power BI** dashboard (including summary stats)

There are two dashboards:

1. **Employment & Job Type Dashboard**



Included Visuals:

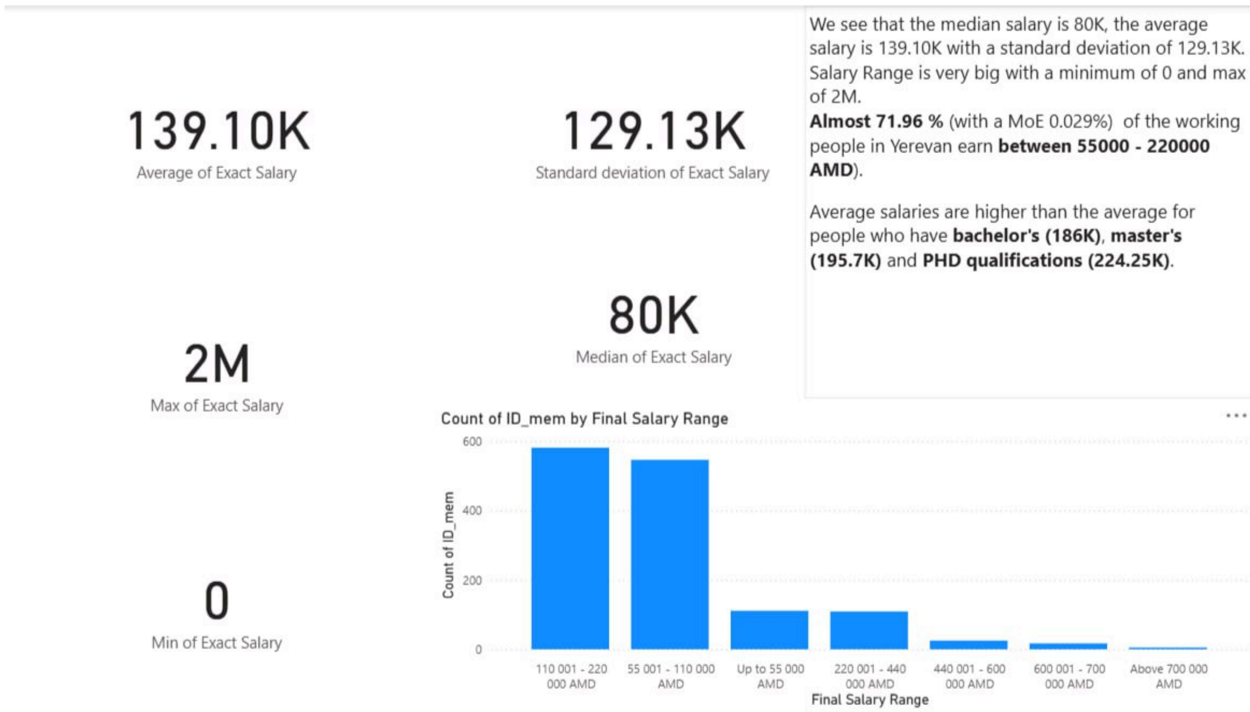
Pie-Chart (comparison of percentages of full-time vs part-time workers),

Stacked Bar Chart (the number of people in each sphere and the number of people in each education level)

Purpose:

Discovers patterns in which spheres most people are employed, and at what education level we have a bigger percentage of people working full time than the overall full-time working percentage.

2. Summary Statistics for Salary



Included Visuals:

Cards(showing mean, standard deviation, median, minimum, and maximum salary values in our dataset)

Clustered Column Chart (shows age ranges and their corresponding count of people earning in that range)

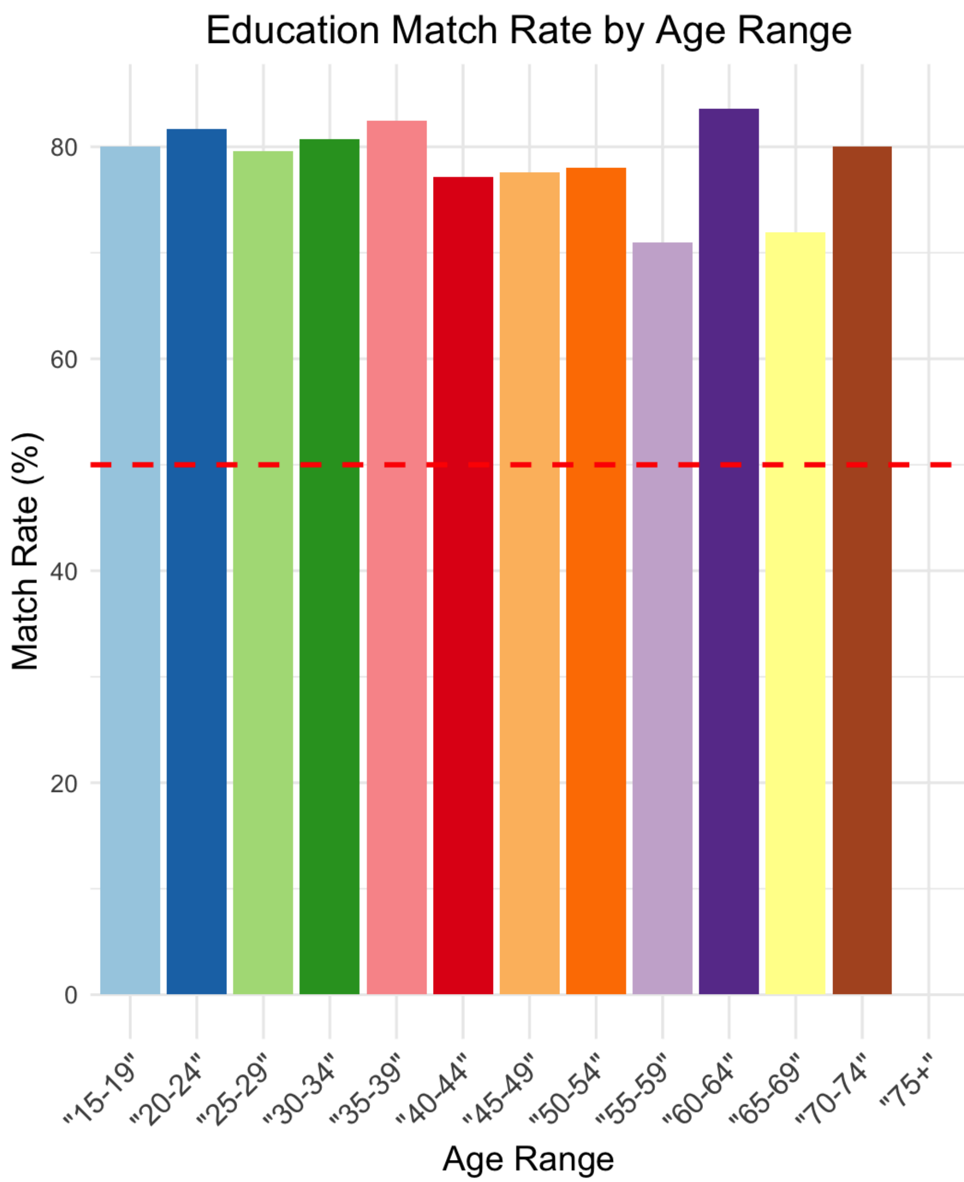
Purpose:

Understand general information about the Yerevan population's salaries. By filtering with the specific education levels which education levels have a mean salary higher than the overall mean salary. Made 99 percent CI for people working in a specific salary range.

Armen Tamrazyan
Hmayak Paravyan 10/05/2025

In R

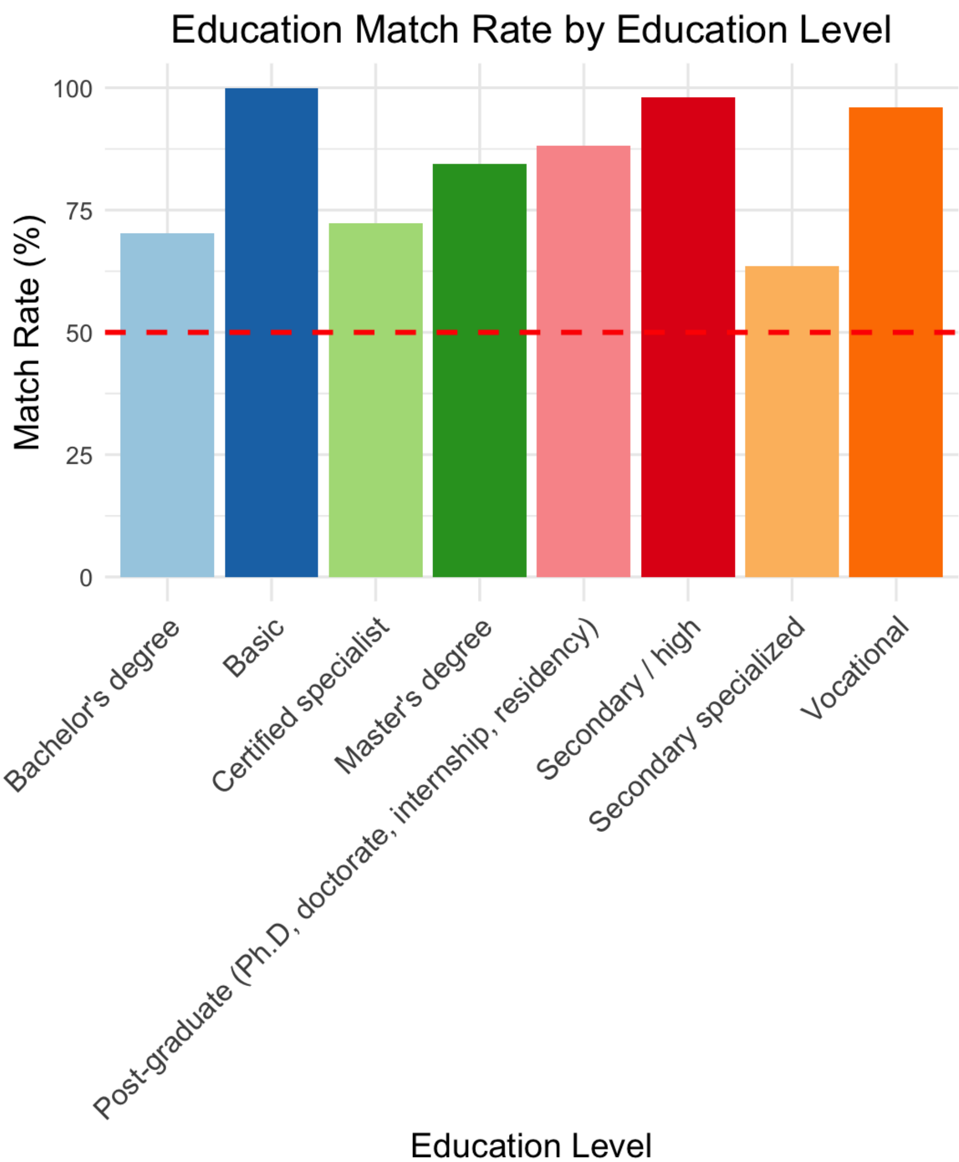
1. **"Match with Age Range"**:



- **Visual**: This plot shows the **job-education match rate** across different **age ranges**.
- **Insight**: This visualization helps to explore how the **job-education match** rate changes across various age groups. Understanding whether older or younger age groups are more likely to have jobs that match their education can provide insights into job market trends, age-related challenges, and potential barriers to aligning education with employment. If younger age groups have a higher match rate, it may suggest better alignment between education and job opportunities early in one's career.

  Why it’s important: This is directly related to the research question because it answers how **job-education alignment** varies across **age groups**, and helps analyze whether certain age groups face more challenges in securing jobs that align with their educational qualifications.

2. **"Match with Education Level"**:

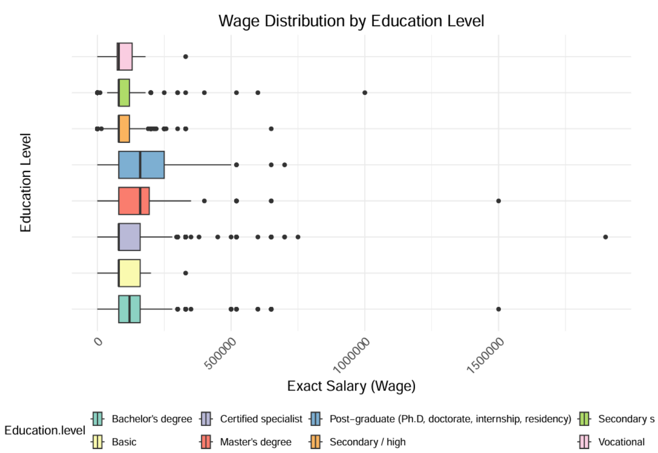


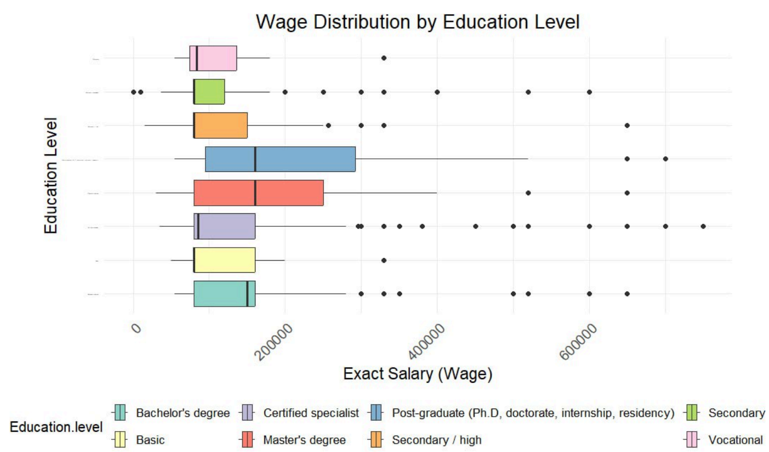
- **Visual**: This plot displays the **job-education match rate** across different **education levels**.

- **Insight**: This visualization is crucial to answering whether higher **education levels** (e.g., master's degree, PhD) are associated with better **job-education**

alignment. A higher match rate for people with advanced degrees may indicate that those with more specialized education are more likely to find jobs in their field of study.

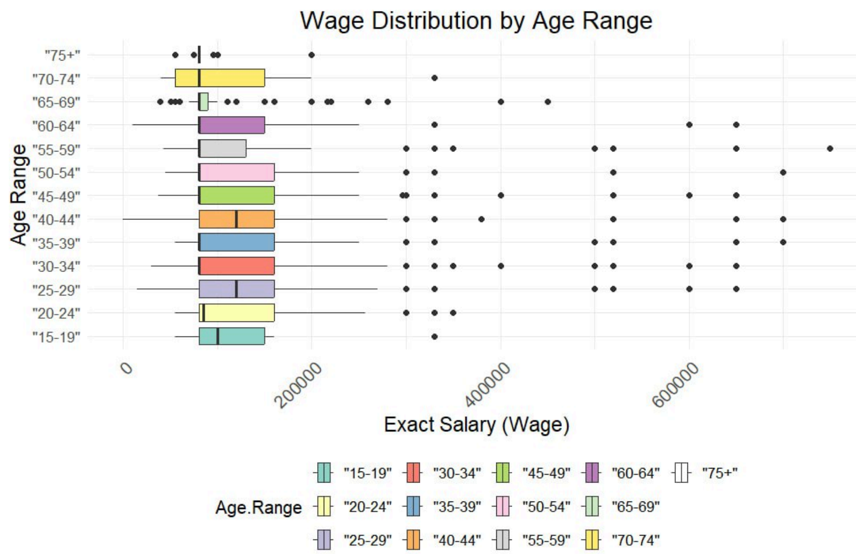
Why it's important: By comparing the job match rates across different education levels, we can see if education plays a role in better alignment between education and job opportunities.





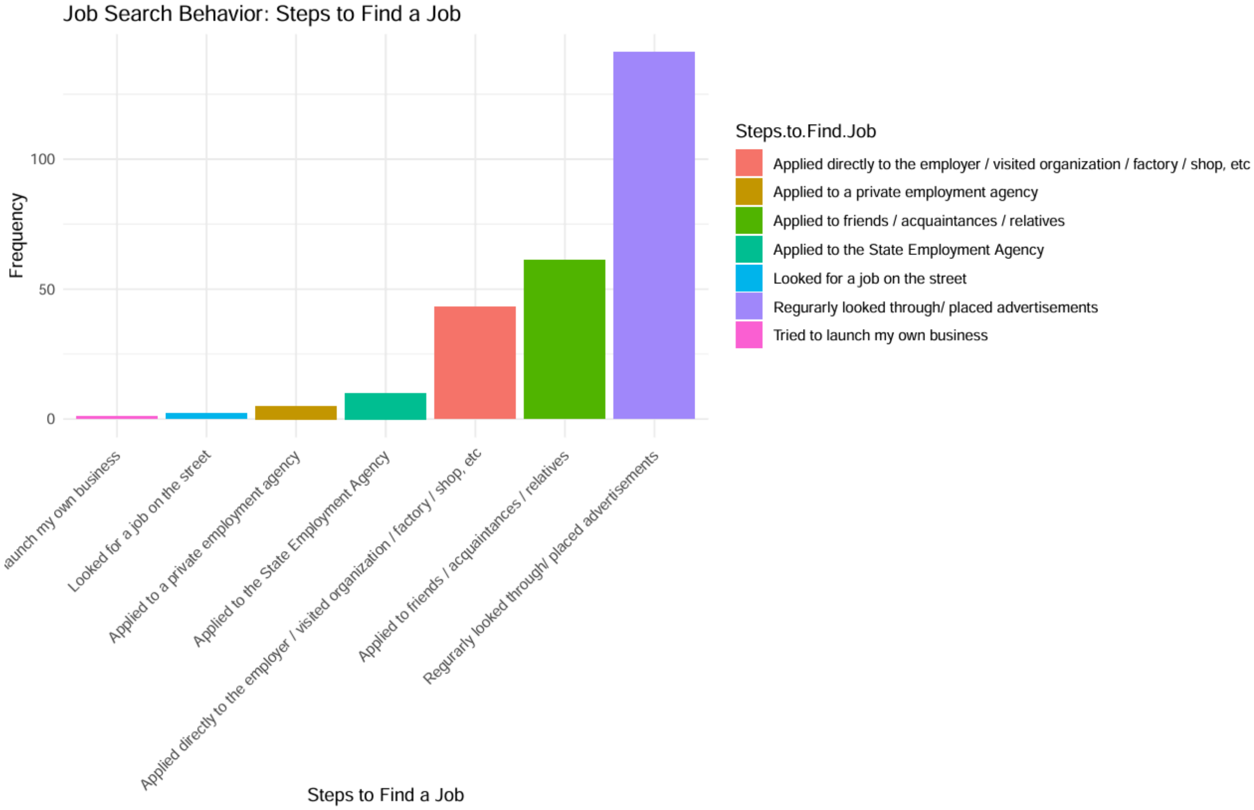
3. **"Wage by Education Level"**:
   - **Visual**: This plot shows the **wage distribution** across various **education levels**.
   - **Insight**: If people with higher education levels (such as a Master's or PhD) tend to earn more, it suggests that education is a strong factor in wage determination.
   - **Why it's important**: This plot helps us assess whether **wages** increase with **education** and if people with more education earn higher salaries. It will also highlight the relationship between **education** and **wages**, and how **job-education alignment** might affect this.

4. **"Wage by Age Level"**:



- **Visual**: This plot shows the **wage distribution** across different **age ranges**.
- **Insight**: This visualization can help answer whether **wages** increase with **age** (and potentially experience). It can also help us understand how **wages** change as individuals get older, potentially reflecting more years of work experience and better job matches as they age
- **Why it's important**: It helps us answer if there is a **correlation between age and wages**, and whether older age groups are earning more, possibly due to better **job-education alignment** or greater experience.

5. **"Job Search Behaviour (Unemployed Edition)"**:



- **Visual**: This plot shows the different **steps to find a job** (e.g., applied to a private agency, applied directly to the employer, etc.).

- **Insight**: This visualization helps in understanding the various job search strategies employed by people. The methods chosen (e.g., job search through advertisements, applying directly to employers) may be influenced by the education level, job-education alignment, and the **demographic group** (e.g., age).

- **Why it's important**: It's important because it helps to answer how **job search behavior** varies across demographic groups, which ties into the broader question of how **employment rates** and **job-education alignment** vary by demographic factors.

Outliers:

After creating the visualizations for **"Wage by Education Level"** and **"Wage by Age Level"**, we observed several **outliers** according to the **boxplots**. However, **four specific outliers** stood out as they were far removed from the rest of the data points, significantly influencing the clarity and **insightfulness** of the graphs.

These extreme values were so far away from the rest of the data that they distorted the representation of wage distributions and made it difficult to discern meaningful patterns. To address this issue, we removed these four extreme points from the graphs in the second image, which provides a clearer and more accurate view of the overall wage distribution.

The second set of visuals, excluding these extreme outliers, offers a more **representative** and **insightful** analysis of wage variations across different **education levels** and **age groups**, ensuring that the key trends are more visible and interpretable.

Methodology

Choice of Methods

1. **Confidence Intervals (CI):**

   CIs were chosen to estimate the **mean wage** and **job-education alignment**, as they provide a **range** within which the true population parameter is likely to fall. This method accounts for sample variability and helps assess whether the differences in wages between education groups are statistically significant, offering more reliable estimates than point estimates alone.

2. **Hypothesis Testing:**

    Hypothesis testing was used to assess whether **job-education alignment** leads to higher wages. Specifically, we tested whether the mean wage of individuals whose job matches their qualifications is significantly higher than those whose job does not. The **Chi-square** test was applied to determine if there is a relationship between **education level** and **job-education match status**.

Confidence Intervals (CI):

1. CI for the proportion of people working in Yerevan who report a job matching their education.
    - **Purpose**: This method is used to estimate the **true proportion** of people in **Yerevan** whose jobs align with their **educational qualifications**. By calculating the **confidence interval (CI)** for the proportion, we can determine the range within which the true population proportion lies, accounting for sample variability. This provides a precise estimate of how common job-education alignment is in the labor market in Yerevan.
    - **Relevance**: Understanding **job-education alignment** is crucial for answering the question of whether **higher educational attainment** leads to better **job fit** in Yerevan. The proportion gives us insight into the **prevalence** of this alignment, which can inform whether having a job that matches one's education is widespread across various demographic groups (such as **age** and **education level**). This can help policymakers, educators, and employers assess whether **education policies** or **labor market interventions** are needed to improve job-education matching.

Sample size: 1566

Proportion with matching education: 0.7861

99% Confidence Interval: [0.75934, 0.8128]

Meaning that, we can be 99 percent confident that 75.93% - 81.28% of working people of Yerevan work in their specialized spheres.

2. CI for the Mean Wage by Education Level
    - **Purpose**: The **confidence interval (CI)** for the **mean wage** by education level aims to estimate the average wage for each education level group (e.g., high school, bachelor's degree, master's, PhD) with a specific level of confidence. This CI provides a range within which the true population mean wage for each education group is likely to fall.
    - **Relevance**: This approach is essential for understanding how **wages** vary across different **education levels**. By calculating the CI, we can assess whether differences in wages between educational groups are statistically significant, offering valuable insights into how **higher education** might be associated with **higher wages**. This directly addresses the research question regarding the impact of **education level** on **wages** in Yerevan, shedding light on whether obtaining higher education leads to better earnings.

Bachelor's degree: [159145.43, 212946.08]

Basic: [89097.76, 142902.24]

Certified specialist: [139347.39, 165676.82]

Master's degree: [163339.03, 228092.16]

Post-graduate (Ph.D, doctorate, internship, residency): [170949.8, 277550.2]

Secondary / high: [102097.75, 114137.9]

Secondary specialized: [96269.81, 113552.56]

Vocational: [82482.99, 133017.01

Meaning that for each education level, we have its corresponding lower and upper bounds for the mean among the whole Yerevan population.

Hypothesis Testing:

1. Test whether the mean wage of those whose job matches their qualification is higher than those whose job doesn't.
   - **Purpose**: This hypothesis test, likely a **t-test**, compares the **mean wage** between two independent groups: individuals whose **job matches their qualification** (Yes) and those whose **job does not match** (No). The test aims to determine if there is a **statistically significant difference** in the average wage between these two groups. Specifically, we are testing whether having a job that aligns with one's educational qualifications leads to a higher mean wage.
   - **Relevance**: This test is directly aligned with the research question of whether **job-education alignment** contributes to **higher wages** in Yerevan. By comparing the mean wages between individuals who have jobs that match their qualifications and those who do not, we can assess whether job-education alignment plays a role in wage disparities. This insight can help inform policies and practices to improve the **match between education and employment**, potentially leading to better wages and job satisfaction for the workforce.

Mean salary for 'Yes' group: 145465

Mean salary for 'No' group: 117035

H0: mean_yes <= mean_no

H1: men_yes > mean_no

Using the formula for calculating the t-value for two means:

t-statistic: 4.327

Degrees of freedom (df): 779.2

P-value: 0.00000855

So, reject the null hypothesis: The mean salary for 'Yes' is significantly greater than 'No' as the t-statistic is in the rejection region for 95 percent confidence and P-value is a very small number . **Therefore, on average, people who work in their specialized spheres earn more than those who do not.**

2. Test for independence between education level and job-education match status.
   - **Purpose**: The **Chi-Square Test of Independence** is used to determine whether there is a **statistically significant relationship** between **education level** and **job-education match status** (whether the individual's job aligns with their educational qualifications). This test examines whether the distribution of job-education match status is dependent on education level or if both variables are independent of each other.

- **Relevance**: This test is essential because it directly answers the research question of whether **education level** influences **job-education alignment**. If a significant association is found, it would suggest that individuals with higher education levels are more likely to have jobs that match their qualifications. This insight can inform policies aimed at improving **job-market outcomes** and ensuring that higher education leads to better job fits, thereby potentially enhancing labor market efficiency and job satisfaction.

Chi-Squared Value: 163.841,

df = 7,

p-value = 0.000000000000000004985 which is smaller than 0.05 (for 95% confidence). This means that there is a **statistically significant association** between **education level** and **education match status**. Therefore, **we reject the null hypothesis, meaning that the education level and education match status are not independent.**

Education Level	No	Yes	**Total**
Bachelor's degree	51	121	172
Basic	0	25	25
Certified specialist	133	346	479
Master's degree	19	103	122
Post-graduate (Ph.D, doctorate, internship, residency)	5	37	42
Secondary / high	7	368	375
Secondary specialized	119	207	326
Vocational	1	24	25

Total	335	1231	1566

Discussion

Answer to the Research Question:

The analysis shows that **education level** and **job-education match status** are **significantly related** in Yerevan. This suggests that individuals with higher education levels are more likely to have jobs that align with their educational qualifications. The results support the hypothesis that **higher education leads to better job-education alignment**, which in turn may contribute to **higher wages**. The findings from the Chi-Square Test of Independence indicate that education level plays a critical role in whether a person's job matches their qualifications.

Practical Significance in Context:

This relationship between education level and job-education match status has practical implications for both policymakers and employers in Yerevan. If higher education levels are associated with better job-education alignment, then enhancing access to education, particularly higher education, may improve job market outcomes. **Employers** may also benefit from this insight, as hiring individuals whose qualifications match their jobs could lead to greater job satisfaction, efficiency, and productivity.

Limitations & Potential Biases:

Regional Focus:

The analysis is limited to **Yerevan**, which aligns with the specific focus of the research question. While the findings provide valuable insights into the job-education alignment within

Yerevan, the results may not apply to other regions of Armenia, where labor market dynamics, educational opportunities, and job availability may differ. However, this regional focus allows for a deeper understanding of the labor market trends specific to the capital.

Data Representation:

The analysis only includes individuals who are employed and excludes those who are unemployed or underemployed, which could introduce **selection bias**. This limits the ability to understand how job-education alignment impacts the entire workforce, including those not currently employed or in jobs that don't require education.

Conclusion

Conclusion:

The analysis demonstrates that education level is significantly associated with both job-education alignment and wage levels in Yerevan, showing that individuals with higher education earn higher wages, those whose jobs match their qualifications earn significantly more, and education level and job-education match status are dependent.

Recommendations & Next Steps:

To enhance **job-education alignment**, it is crucial for policymakers to prioritize initiatives that bridge the gap between educational qualifications and available job opportunities. Efforts should be directed at strengthening the connection between **education curricula** and the needs of the labor market. Additionally, **regional disparities** in job-education alignment should be further explored, with particular attention to areas outside Yerevan, where education-job mismatches may be more pronounced. Future research could also focus on identifying the

underlying causes of wage inequality and **underemployment** to develop more targeted interventions for improving labor market outcomes.

Appendix

Armstat - The 2023 anonymized Labour Force Survey (LFS) microdata for Armenia

Books used: Statistics: Concepts and Applications by David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Chapter ten: Inferences about a population proportion pp. 339 - 365

Armstat - the number of people working in Armenia's regions over different years