



جامعة مصر للمعلوماتية
EGYPT UNIVERSITY
OF INFORMATICS

Egypt University of Informatics
Computing and Information Sciences
Data Analysis Course

A Study on the Gender Pay Gap in Egypt's Tech Market

Supervised by:

Dr. Mohamed Taher Alrefaie

Submitted by:

Mohamed Elmosalamy 23-101283

Omar Mohammad 23-101288

Omar Hesham 23-101302

Mohamed Kotb 24-101222

2025-05-24

Contents

1	Abstract	3
2	Introduction	3
3	Research Questions	3
4	Hypothesis	4
	4.1 Difference in mean salaries	4
	4.2 Controlled difference in mean salaries	4
5	Population of Interest	4
6	Dataset	4
7	Analysis	5
8	Hypothesis Testing Steps	5
	8.1 Difference in mean salaries	5
	8.2 Controlled difference in mean salaries	5
	8.3 Cost of Being a Woman	5
9	Conclusion	6
	9.1 Difference in mean salaries	6
	9.2 Controlled difference in mean salaries	6
	9.3 The Cost of Being a Woman	7
10	Any Potential Issues	7

1 Abstract

The gender pay gap remains a contentious and widely debated issue worldwide. This analysis investigates whether women earn less than their peers in **Egypt's tech industry**.

The study aims to quantify salary differences between genders using publicly available data, offering insights into whether a significant gender-based wage gap exists.

2 Introduction

The gender pay gap has been a global issue across labor markets. While considerable research has been done in Western countries, limited data-driven analysis exists in the Middle East. In Egypt, this remains a largely unstudied topic.

To address this, we define three methods of investigation:

1. Assessing the gender pay gap

The first hypothesis tests for the presence of a gender pay gap in Egypt's tech industry by conducting a **Welch's independent t-test** on male and female salaries. This analysis aims to answer: **"Is there a pay gap between men and women in Egypt's tech industry?"**

2. [Placeholder for Hypothesis 2]

3. Estimating The Cost of Being a Woman

The Cost of Being a Woman is defined as the monthly salary disparity a woman faces compared to a man with identical skills, title, and experience. We aim to provide a 95% confidence interval for this cost.

3 Research Questions

1. Is there a statistically significant difference in mean salaries between genders?
2. After controlling for years of experience and other contributing factors, does the pay gap persist?
3. What is the cost of being a woman—how much does a woman gain or lose per month compared to an equally qualified man?

4 Hypothesis

4.1 Difference in mean salaries

- Null hypothesis (H_o): There is no significant pay gap between men and women.
- Alternative hypothesis (H_a): There is a significant pay gap between men and women.
- Significance level (α): 0.05

4.2 Controlled difference in mean salaries

- Null hypothesis (H_o): After controlling for contributing factors (experience, title, level, etc.), there is no significant gender pay gap.
- Alternative hypothesis (H_a): A significant pay gap persists even after controlling for these factors.
- Significance level (α): 0.05

5 Population of Interest

All professionals working in Egypt's tech field.

6 Dataset

The dataset used comes from the [Egyptian Tech Market Survey API](#), conducted in 2024.

Dataset columns include:

- **Gender:** Male, Female
- **Degree:** Bachelor's degree (Yes, No)
- **Title:** Professional title (e.g., Data Analyst, Scrum Master)
- **Level:** Professional level (e.g., Junior, Senior, Team Lead)
- **YearsOfExperience:** Number of years in the tech field
- **Salary:** Monthly salary in EGP
- **IsEgp:** Currency used (EGP, foreign, hybrid)
- **ProgrammingLanguages:** Languages the subject can write
- **BusinessMarket:** Scope (Local, Regional, Global)
- **BusinessSize:** Company size (Start-up, SME, Large)
- **WorkSetting:** Working environment (Office, Remote, etc.)
- **CompanyLocation:** City/state of the company

Sample size: **2649**

7 Analysis

TODO MOHA -> Insert visualizations and summaries from Python notebook

8 Hypothesis Testing Steps

8.1 Difference in mean salaries

We used **Welch's Independent T-Test**:

1. Hypotheses defined (see [Section 4.1](#))
2. Significance level set to 0.05
3. Data cleaned and prepared
4. Used `scipy.stats.ttest_ind(equal_var=False)`
5. Decision made based on resulting p-value vs. alpha

8.2 Controlled difference in mean salaries

For this, we applied the **Blinder-Oaxaca decomposition**:

1. Hypotheses defined (see [Section 4.2](#))
2. Significance level set to 0.05
3. Data cleaned and outcome/explanatory variables defined
4. Ran group regressions, decomposed results using `statsmodels`
5. Interpreted contributions of each factor (explained/unexplained)

8.3 Cost of Being a Woman

Cost is defined as:

Cost := Expected Salary based on objective factors – Actual Salary

Objective factors used:

- Years of Experience
- Title
- Level

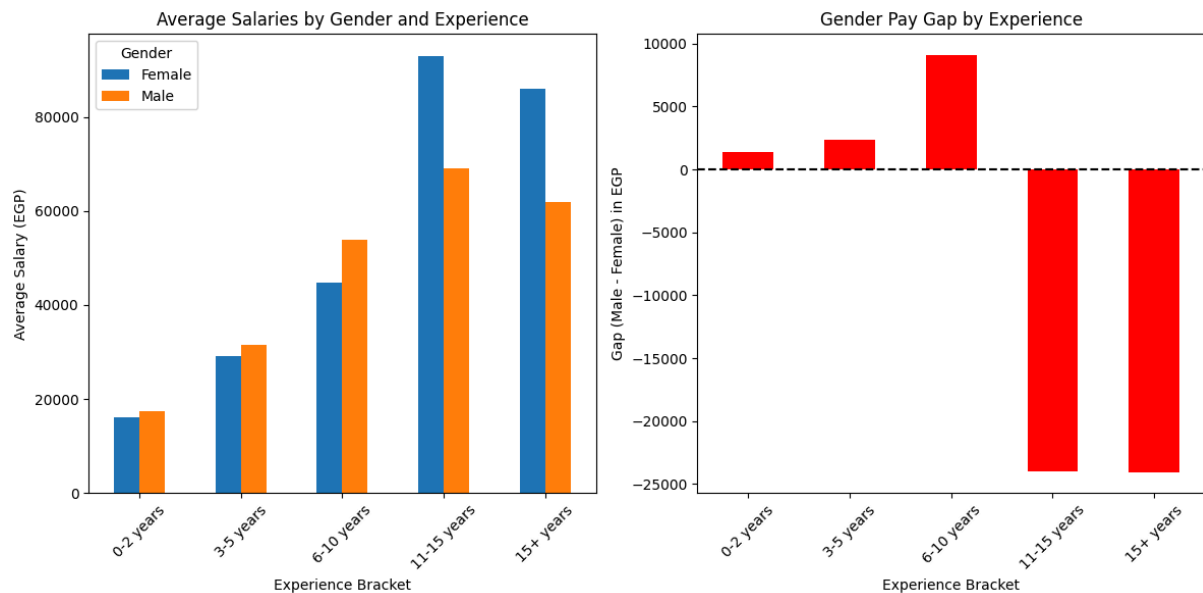
We trained a regression model on male data using all variables except gender, then applied it to female employees to estimate expected salaries. This allowed us to construct a 95% confidence interval around the cost of being a woman.

9 Conclusion

9.1 Difference in mean salaries

P-value: 0.2668 — well above the 0.05 threshold. We fail to reject the null hypothesis. No significant evidence of a gender-based pay gap exists.

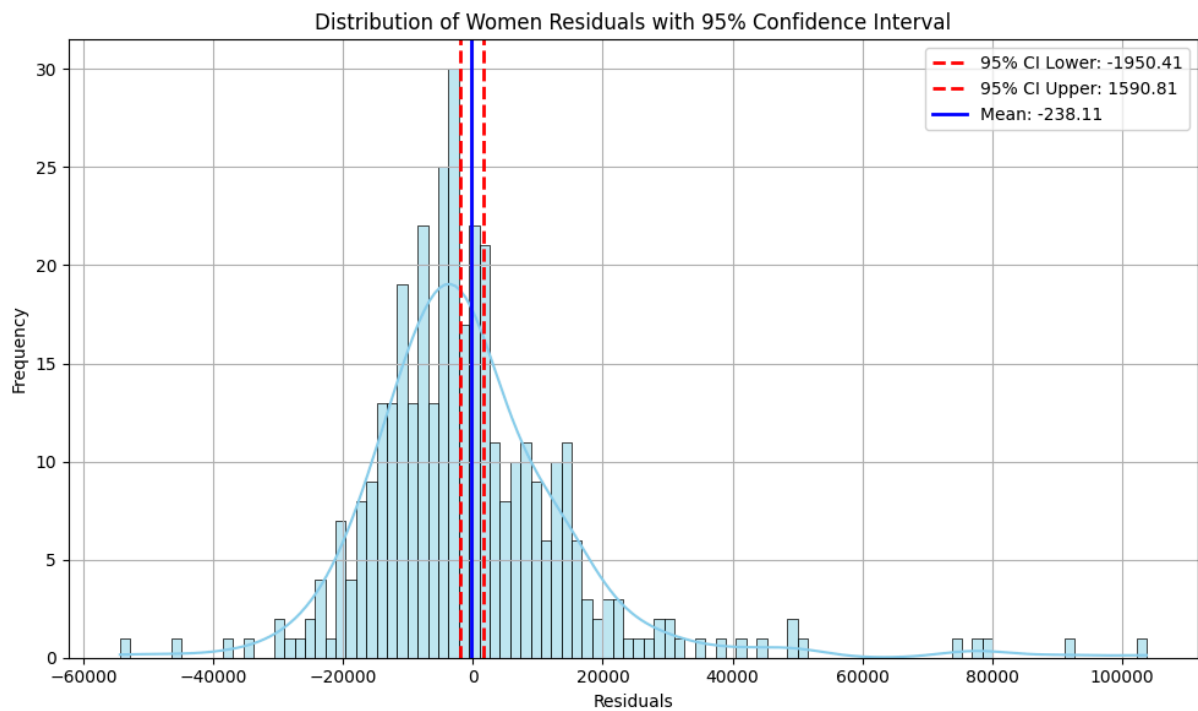
9.2 Controlled difference in mean salaries



Insights:

- Average salaries rise with experience for both genders
- Men earn slightly more in the first 10 years
- After 10 years, women's average salaries exceed those of men

9.3 The Cost of Being a Woman



The 95% confidence interval ranges from **-1950.41 to 1590.81 EGP**. Since zero lies within this interval, it suggests that the “cost of being a woman” may, statistically, be zero.

10 Any Potential Issues

TODO MOHA -> Discuss outliers and scarcity of female records in the dataset