

# Abstract

The gender pay gap remains a contentious and highly 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.

## 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 an unstudied topic.

To reason about this, we had to define 3 methods to attack this problem.

*Describe 1st hypothesis here*

*Describe 2st hypothesis here*

Third, we define **The Cost of Being a Woman** as the amount of disparity between salaries that a women loses (or potentially earns) compared to men per month. It's the amount a man with her skills, title, and years of experience would be expected to earn. The analysis aims to provide a confidence interval with  $\alpha = 95\%$  to find this cost.

## Description of the dataset

[Moha does this]

## Methods

### 1st hypothesis

### 2nd hypothesis

### The Cost of Being a Woman

We can define the cost of begin a woman as following

Cost := Expected salary based on objective factors — Current Salary

To estimate the cost of being a woman, we need to find the objective factors that should dictate one's salary and use them to predict their salaries regardless of gender.

The objective factors used in this study are:

1. Years of Experience
2. Title
3. Level

We used a multiple regression model to predict each woman's expected salary. This model captures how salary is determined for male employees based on their qualifications, creating a baseline for equitable pay.

The idea behind this is to have the explanatory variables all the columns in the dataset except for gender and train this model solely on men data. Therefore, this model treats the entire workforce as men. The trained model is applied to the female workforce to find their expected salary based on objective factors.

Now, we use the current salary given in the dataset, and use this distribution to create a confidence interval.

*Show CI image here*