A logo of a globe with a graduation cap

Description automatically generated**A logo with dots and lines

Description automatically generated**

**Names :**

**Supervisor** : Dr Amal Mahmoud

student project :

* 1- Bishoy victor Nabil
* 2- Ayman saadan
* 3-Ebram Effat
* 4-Amgad Abdelhamid Hamza

Abstract

The "Job Scope" project explores compensation trends and job market characteristics within the Egyptian technology sector. Utilizing data scraped from a job listing website, this analysis aims to uncover critical insights into salary distributions, job roles, business focus, and diversity across the tech industry. By examining various factors like years of experience, education level, and company location, the project provides a comprehensive view of compensation dynamics and market needs. The results are intended to guide job seekers, employers, and policymakers in making informed decisions regarding employment and compensation in Egypt's tech industry.

Introduction

The technology sector in Egypt has been experiencing significant growth, with increasing demand for skilled professionals across various roles. As the job market evolves, understanding compensation trends and the factors influencing salaries becomes crucial for job seekers, employers, and decision-makers. The "Job Scope" project was initiated to provide a data-driven analysis of compensation and job roles in the Egyptian technology sector, leveraging data scraped from a popular job website.

This project aims to address several key questions, such as:

* What is the average compensation across different business markets, job roles, and locations?
* Are there disparities in compensation based on experience, education, or gender?
* How does the work setting (remote, hybrid, in-office) impact compensation and job availability?

By answering these questions, the project offers valuable insights that will help both individuals and organizations navigate the evolving job market landscape.

Methodology

***Data Description***

The dataset used in this analysis was scraped from a website that lists technology sector jobs in Egypt. It contains information about job titles, business markets, years of experience (YOE), education requirements, compensation (total and ranges), and company locations. Additional variables include work settings (remote, hybrid, or in-office), gender representation, and business focus (e.g., software, finance).

The dataset's key columns that were scraped include:

* **Timestamp**
* **Job title**
* **Degree**
* **Business market**
* **Programming language**
* **Business focus**
* **Business line**
* **Industries**
* **Is egp**
* **Years of experience (YOE)**
* **Total compensation in EGP**
* **Work setting (remote, hybrid, in-office)**
* **Company location**
* **Gender**
* **Degree requirement**

**Before dividing the table into 7 tables(no.of columns is 1650):**

* **Table 1: job**

**- job\_id**

**- Company\_ID**

* **Title**
* **MinCompensationRange**
* **MaxCompensationRange**
* **WorkSetting**
* **Yoe**
* **TotalCompensationEgp**
* **IsEgp**

* **Table 2: Calendar**

**- date**

* **Year**
* **Quarter**
* **MonthName**
* **MonthNum**
* **Weekday**
* **Weeknum**

* **Table 3 : Company**
* **Company\_Id**
* **Businessmarket**
* **BusinessSize**
* **BusinessFocus**
* **Businessline**
* **Location**

* **Table 4 : Industries**
* **ID\_In**
* **Industry**

* **Table 5 : Programming Language**
* **ID\_pL**
* **Programming Language**

* **Table 6 : Employee**
* **Emp\_ID**
* **Gender**
* **Level**
* **Degree**

* **Table 7 : Data**
* **Timestamp**
* **ID**
* **ID\_Employee**
* **ID\_job**

*Tools*

The following tools were used to perform the analysis:

* **Excel**: Initial data cleaning, preprocessing, and organization of the scraped data.
* **Python**: Advanced data manipulation, outlier detection, and statistical analysis.
* **SQL**: For querying, filtering, and joining data from different sources during the analysi
* **Tableau**: For visualizing key insights and creating dashboards to represent trends in compensation, job roles, and market characteristics.

***Method***

1. **Data Collection and Preprocessing**: The data was scraped from the website using automated scraping scripts. The collected data was first cleaned in Excel, where missing values and duplicates were handled, and categorical variables were encoded.
2. **Exploratory Data Analysis (EDA)**: Python was used to conduct EDA, where key variables like compensation, job titles, business size, and location were analyzed. Outliers in salary data were identified and handled to prevent skewed results.
3. **Statistical Analysis**: Python and SQL were employed to calculate averages, medians, and standard deviations of compensation based on factors like YOE, education, and location. This analysis helped identify patterns in compensation distribution.
4. **Visualization**: Tableau was used to create interactive dashboards that visualize compensation trends, gender pay gaps, job role distribution, and work setting preferences across different industries and company sizes.

Contribution

The "Job Scope" project offers several key contributions to understanding the Egyptian technology job market:

1. **Compensation Insights**: The project reveals average compensation by business market, industry, job title, and work setting, helping professionals assess their market value.
2. **Job Role Distribution**: By analyzing common job roles and their relationship with business focus, the project highlights the most in-demand roles and programming languages in the tech sector.
3. **Diversity and Inclusion**: The analysis uncovers gender pay disparities and representation issues, providing actionable insights for improving diversity in the industry.
4. **Work Setting Trends**: The project explores how work settings (remote, hybrid, in-office) influence compensation and job availability, shedding light on the evolving preferences of companies and employees in the tech sector.
5. **Geographical Analysis**: By comparing compensation across regions, the project identifies high-paying areas and provides insights into regional differences in job market dynamics.

**Appendix**

The appendix contains:

* **Data Source**: A detailed description of the website from which the data was scraped, along with a summary of the scraping process.
* **Code Snippets**: Examples of Python, SQL queries, and Tableau dashboards used in the analysis.
* **Glossary of Terms**: Definitions for terms such as "business market," "YOE," and "business focus" to clarify the scope of analysis.
* **Limitations**: Discussion of any limitations faced during data collection and analysis, such as missing values or the scope of the dataset.