

King Hussein School Of Computing Sciences

Data Visualization Project

**Project: Main factors associated with hiring at UN agencies**

Group 5

**Prepared by:**

Alexandra Jaberi

Rakan Armoush

Raghad Joudeh

Omar Attawneh

Abdel Rahman Yaghi

Heba Najdawi

Sondos Ali (20200438)

**Supervised by:**

Dr. Osama Abdel Hay

**Contributions:**

Alexandra Jaberi

Rakan Armoush

Raghad Joudeh

Omar Attawneh

Abdel Rahman Yaghi

Heba Najdawi

Sondos Ali: have worked on designing and writing the report with the help of Alexandra.

**Contents**

[**1.** **Introduction** 3](#_Toc124104220)

[**2.** **Exploratory Analysis** 3](#_Toc124104221)

[**2.1** **Data Inconsistency** 3](#_Toc124104222)

[**2.2** **Miscalculation** 4](#_Toc124104223)

[**2.3** **Missing data** 4](#_Toc124104224)

[**3.** **Explanatory Analysis** 4](#_Toc124104225)

**Table of Figures**

[Figure 1. Percentages of recruited people based on the fluency of the languages 5](#_Toc124588112)

[Figure 2. The effect of region on recruitment\*\* 6](#_Toc124588113)

[Figure 3. Relation between age and number of employees related to that age 6](#_Toc124588114)

[Figure 4. The relation between the Age of the employee and the level of recruitment 7](#_Toc124588115)

[Figure 5. The relation between the entity type and age with gender 8](#_Toc124588116)

# **Introduction**

The purpose for this project is to give the human resources department three primary insights that can be driven from the dataset retrieved from The United Nations Agencies. The dataset provided contains two sub-datasets. The first, recruitment, contains those already hired at the UN. The second, named roster, contains candidates registered in the roster talent pool.

This Project works on studying and analyzing data to visualize it with specific tools as ( Python ), the Analytics will be simplified and visualized in an accurate way which will get us to the potential factors that influenced the outcome of the interest.

In addition, this project will provide on how an experiment can be designed to investigate the most critical factor.

# **Exploratory Analysis**

Exploratory analysis is what you do to understand the data and figure out what might be noteworthy or interesting to highlight to others.

The main problems found in the datasets given are listed below:

* 1. **Data Inconsistency**

Data inconsistency is when the same data exists in different formats in multiple tables. Unfortunately, data inconsistency can lead to unreliable and/or meaningless information.

After exploring the data, some data points had been misnamed leading to having more unique values than mentioned. The best solution for this problem is to replace the misnamed values to the nearest of the mentioned possibilities.

The features where inconsistency in the data have been found from both datasets are the following:

1. Jo type, included in recruitment dataset
2. Region, included in recruitment dataset
3. Entity type, included in recruitment dataset
4. Job Family, included in roster dataset
5. Posting Year, included in recruitment dataset

# **Miscalculation and Errors in Numbers**

Numbers are considered dangerous and sensitive to calculations. Miscalculations in features that depend on other features are common natural mistakes when filling data. This problem can be solved by checking the calculations for the independent features and comparing the results to dependent features. After finding the errors, we used the built-in function *dataframevalue.loc[rows, columns]* in python to locate and refill the errors with the right results.

* 1. **Missing data**

# **Explanatory Analysis**

**References**

Excerpt From Storytelling with Data: A Data Visualization Guide for Business Professionals Nussbaumer Knaflic, Cole.

## 3.1 Data Visuals and Graphs

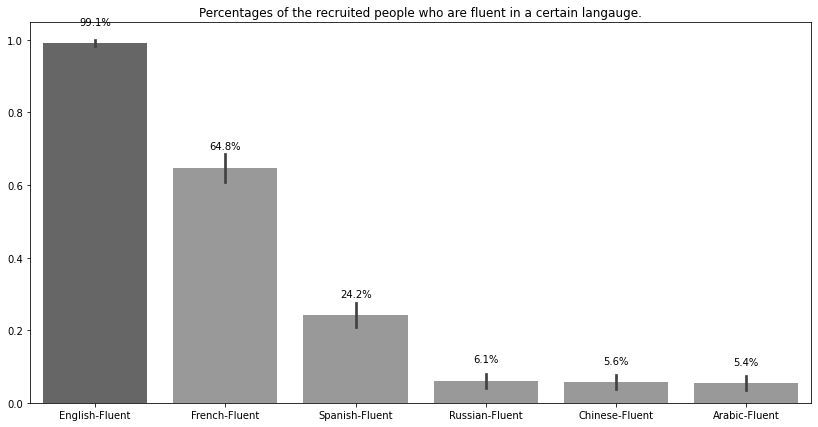


Figure 1. Percentages of recruited people based on the fluency of the languages

It is shown clearly that english language is the most spoken among candidates that have been recruited with a percentage of 99.1%. French is the second most spoken language with a percentage of 64.8%. Followed by spanish language with a percentage of 24.2% of the people recruited are fluent in. Russian, Chinese and Arabic languages got the least percentages with 6.1%, 5.6%, 5.4% of people fluently speaking them respectively.

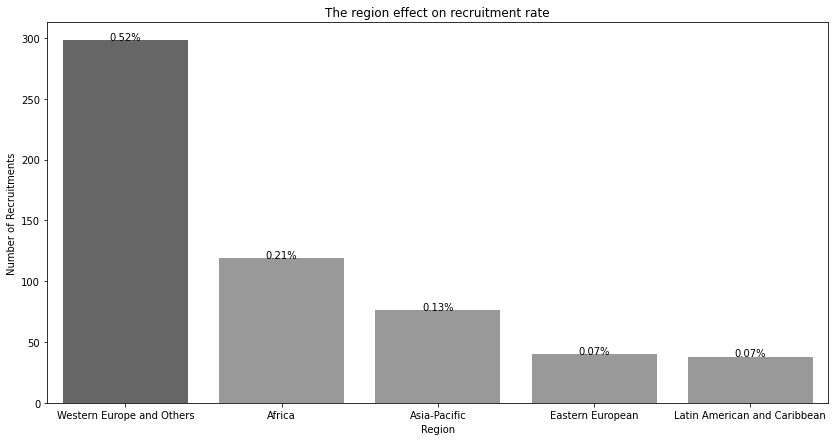


Figure . The effect of region on recruitment\*\*

Figure 2 shows the relationship between recruitment and the region of the applicant, it is shown clearly that the number of hired applicants from Western Europe are significantly higher than the other regions with a percentage of 52%, Africa got the second higher number of recruitment with a percentage of 21% , asia pacific 13% , regions eastern european, latin, have the same percentages 7%.

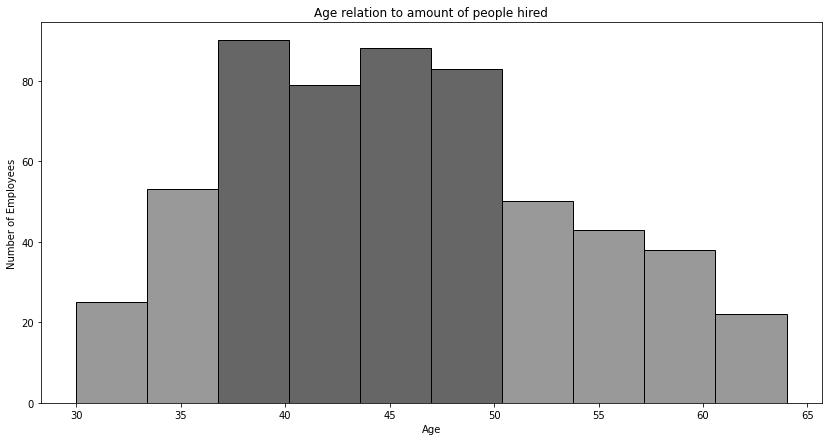


Figure . Relation between age and number of employees related to that age

As viewed in Figure 3, the relation is between Age and Number of Employees, where the data is normally distributed and the highest recruited individuals were in range from 36 to 50, this shows that they have a better chance of being recruited. Comparing it to figure 4, there is an obvious relationship between Age and The Level of Professionalism, where the older the age, the higher the level. This might indicate that even if there are a lot of employees recruited in the mid range of age, they are nearly distributed in levels 3 and 4. While in level 5, where the older gets recruited in, shows that the older in age have more experience than the younger. However there are four potential outliers in level P-3 which indicates that not all employees with higher age are recruited at a higher level.

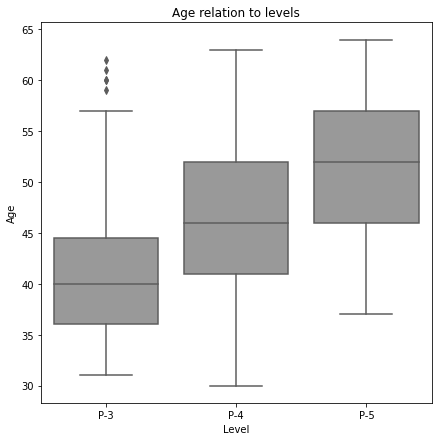


Figure . The relation between the Age of the employee and the level of recruitment

In Figure 4, all levels of professionalism are normally distributed, but unlike level 4 and 5, level 3 is compressed (has a tight range of age) with the range of age. Surprisingly, in level 3 there are potential outliers according to the age of recruited employees having ages from 58 to 63.

## 3.2 Central Statistical Relationships and Correlations

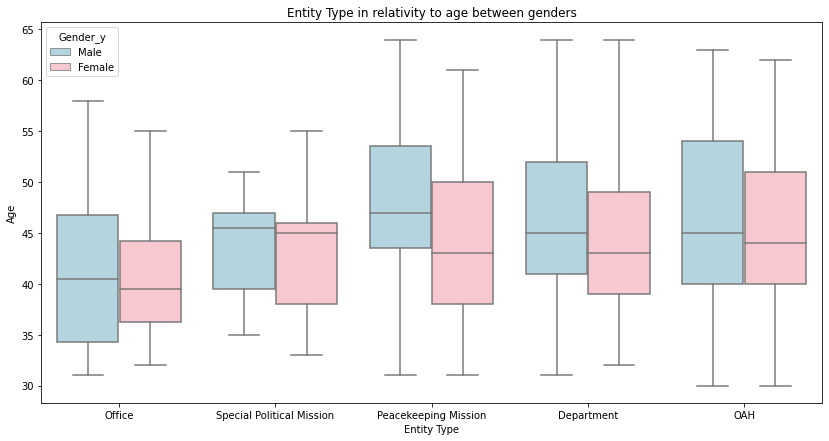


Figure . The relation between the entity type and age with gender

In addition to the interpretation of Figure 3 and 4, Figure 5 adds the Entity type and the Gender of employees. Figure 5 shows that males are usually more employed than females in all types of entities. Here the Special Political Mission Entity has the least number of employees, 45 being the average ages of males and 44 being the average ages of females. While OAH has the most number of employees, 45 being the average ages of males and 43 being the average ages of females.

The figure also shows that unlike all entities that have the range of males’ ages are slightly greater than females’ ages, but the males in Peacekeeping Mission Entity are notably older than the females, 47 being the average ages of males and 42 is the average ages of females.