Data Glacier

Your Deep Learning Partner

IT Salary Survey Analysis

Virtual Internship

Esraa Sultan

28-11-2021

Background
- Case study -

Background:

• XYZ is a company that wants to attract technical staff to the technology department, it wants to know the expected salary rate according to major that is need it.

• Objective :

 Provide actionable insights to help XYZ firm in identifying the right salary for each major.

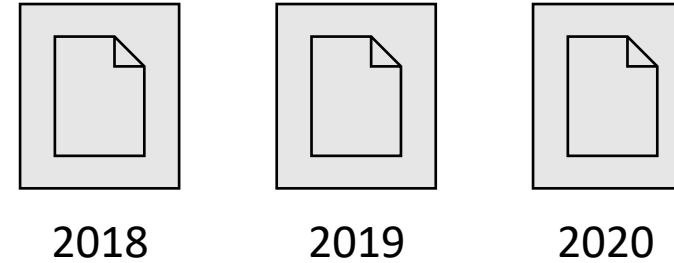The analysis has been divided into three parts:
• Deep cleaning data.
• Data understanding.
• Forecasting the most programming language used recently.
• Finding the highest salary for which major.
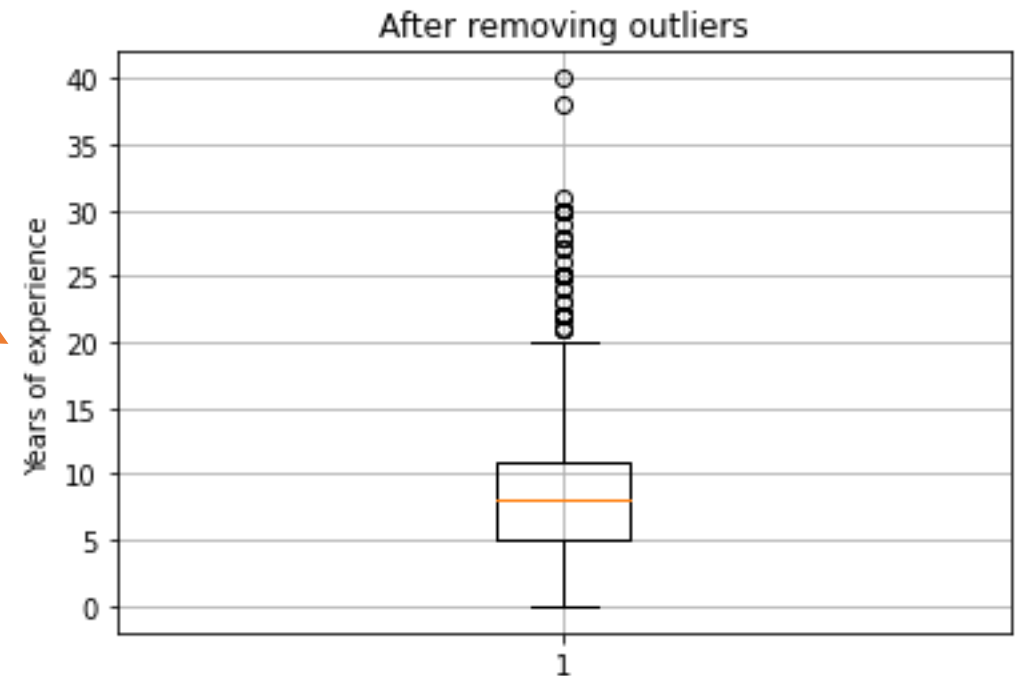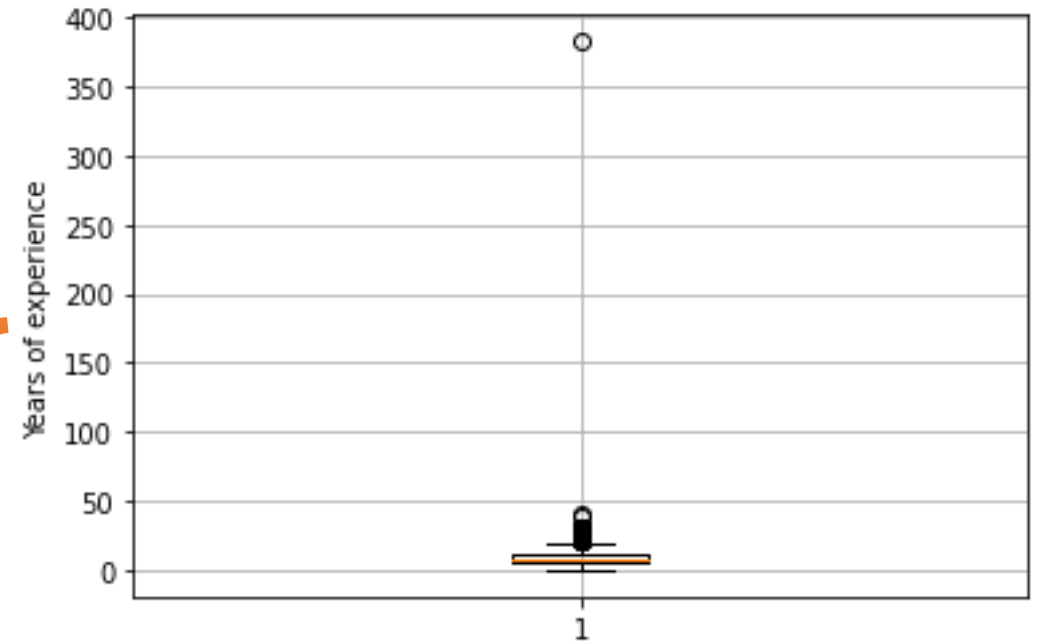
Data Exploration

2018   2019   2020

- 3 files (for each year one file).
- Timeframe of the data: 2018, 2019 and 2020.
- 19 Features(After Cleaning).
- Total data points :2958 (After Cleaning).

Assumptions:
- All the missing data in numerical features full with 0.
- All the missing data in categorical features full with MISSING.
- All the anomalies have been converted to these two values.

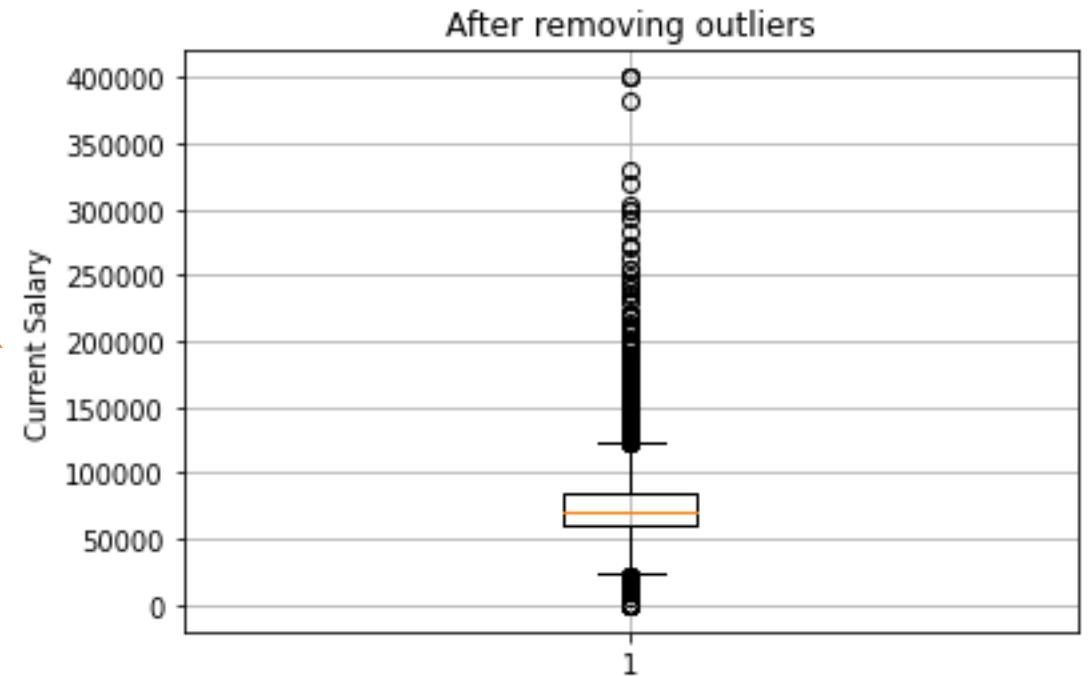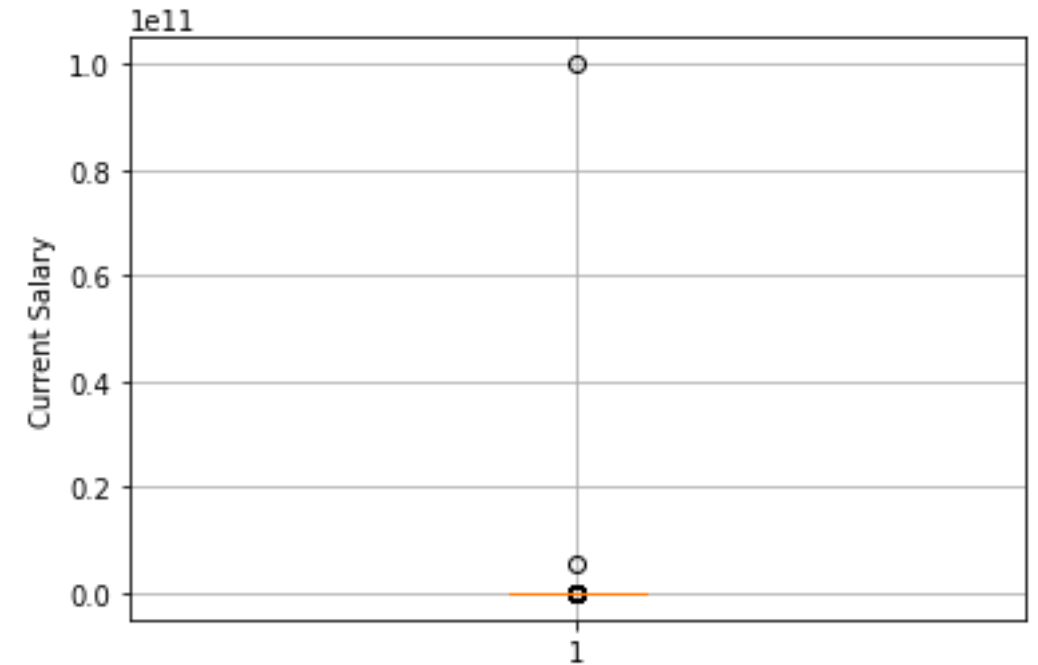# Outliers detection and treatment

We found outliers for the years of the experience
Has a value =383 which is not possible value.
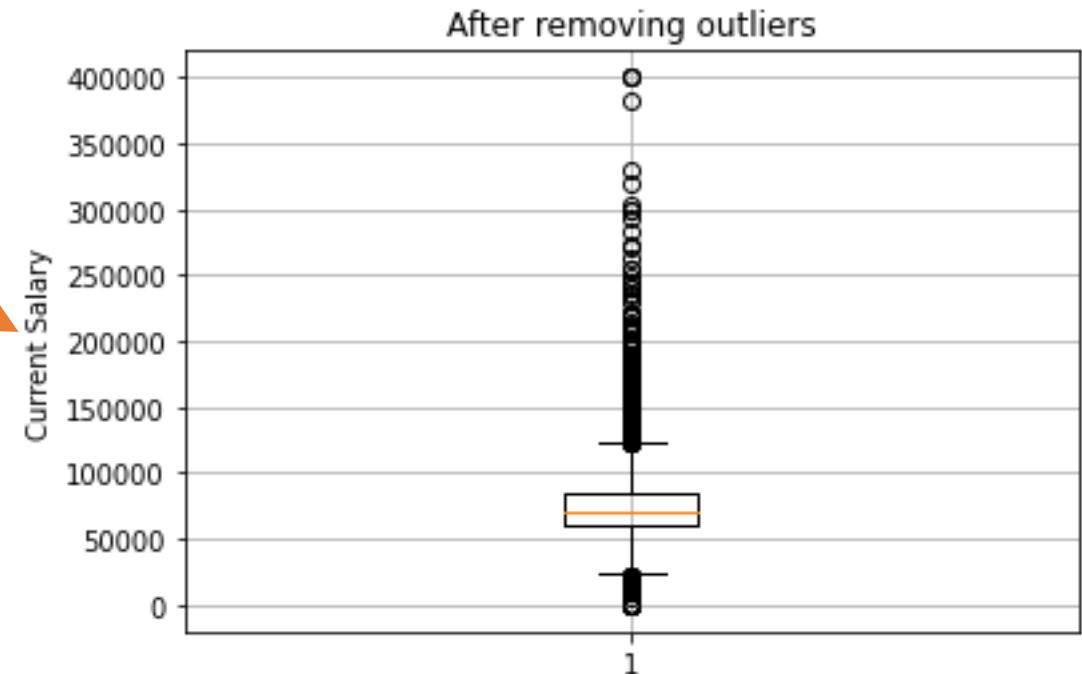

After removing outliers

Outliers detection and treatment

We found outliers for the current Salary
Has a value =99999999999.0 which is not possible value.

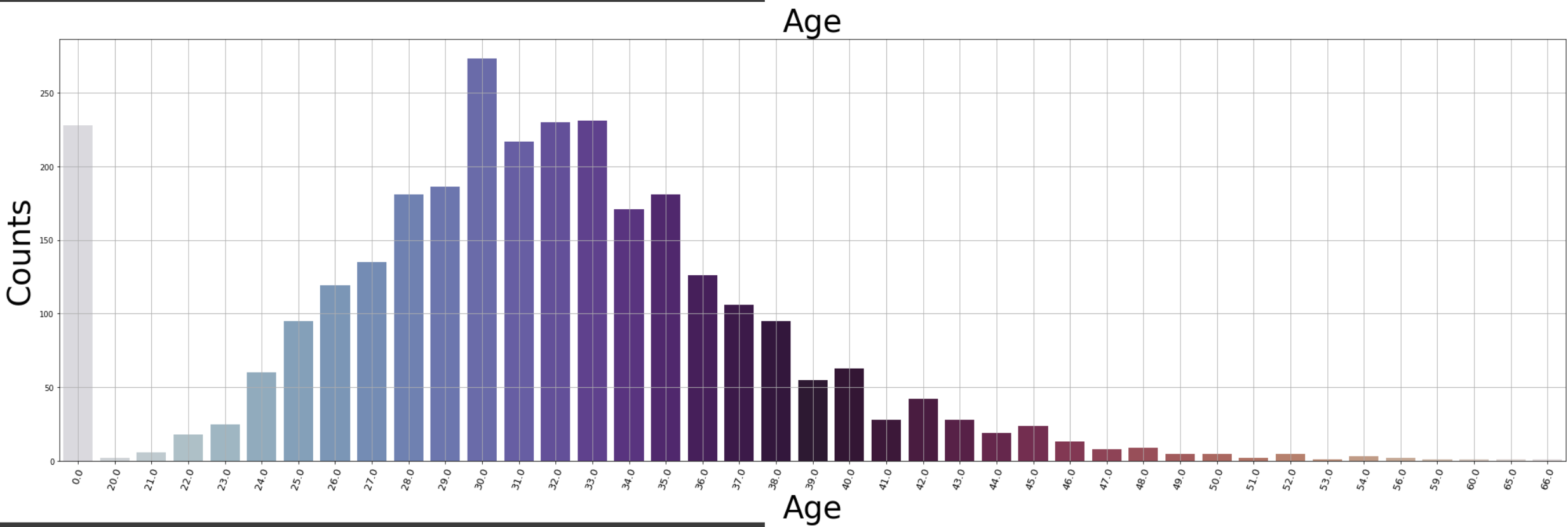# Outliers detection and treatment

We found outliers for Salary one year ago
Has a value = 780000.0 which is not possible
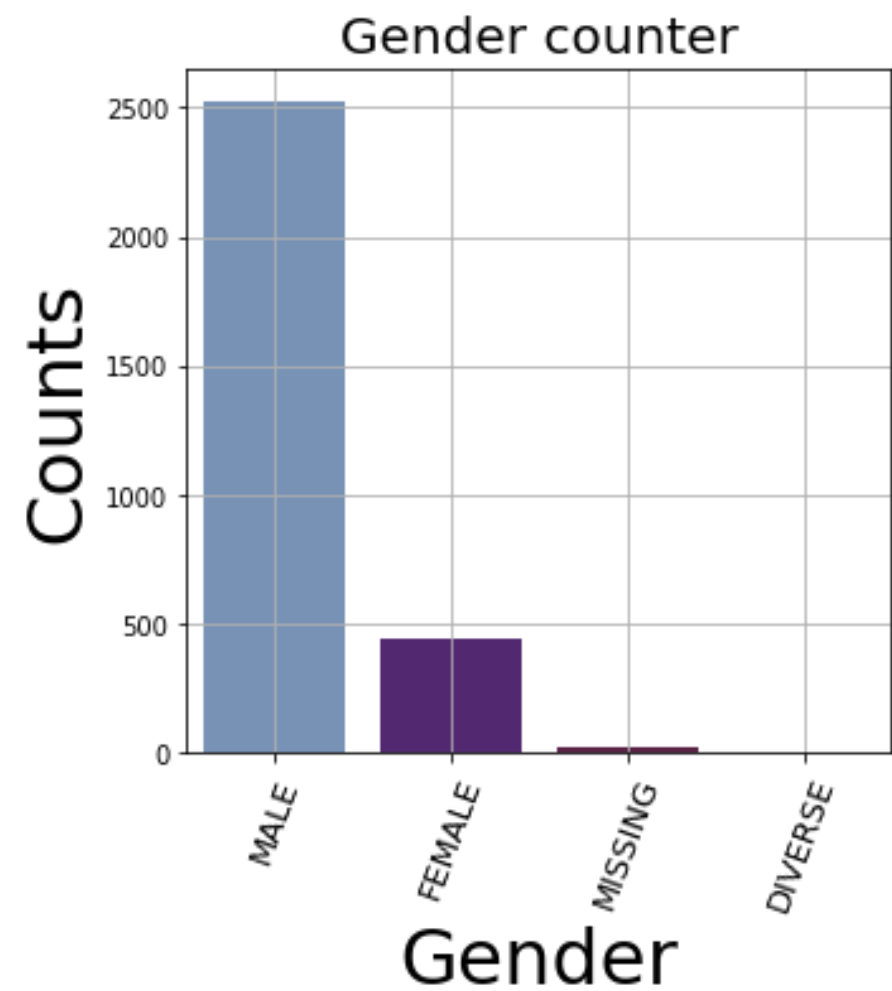value.

After removing outliers

Age

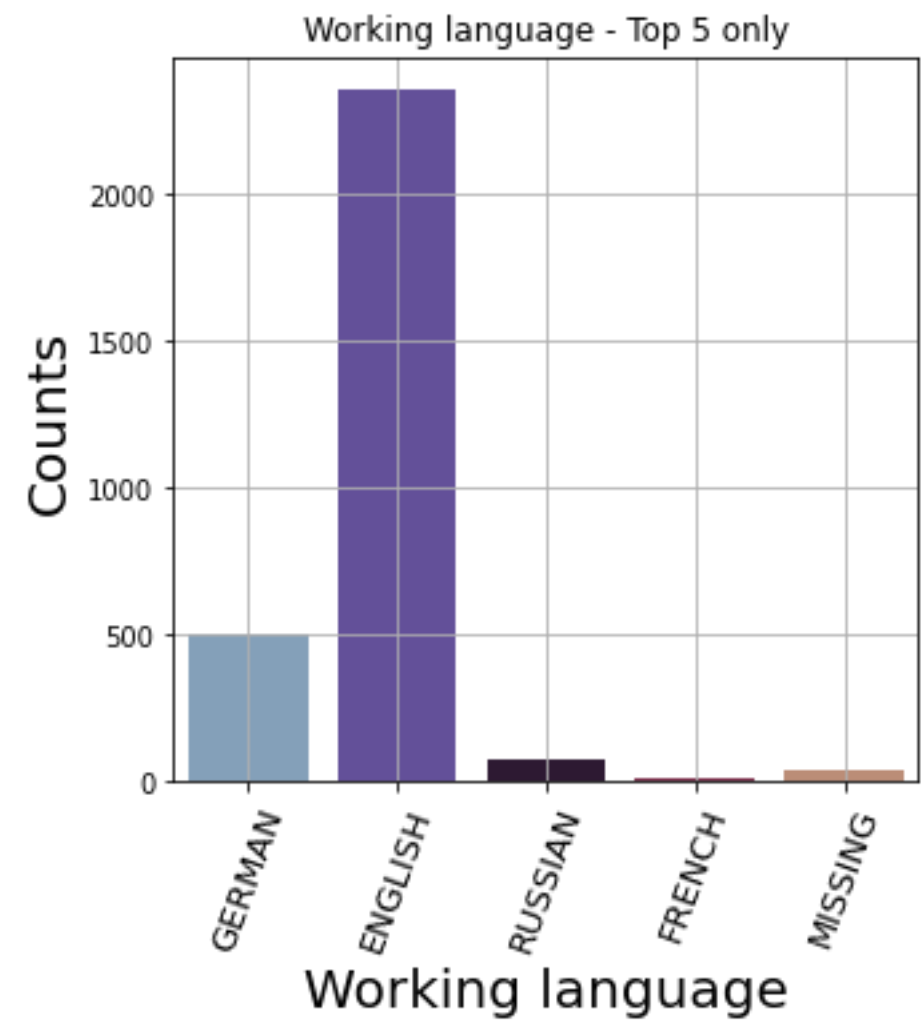The range of most age of the people they sharing of the survey between 24-40 years.

Gender

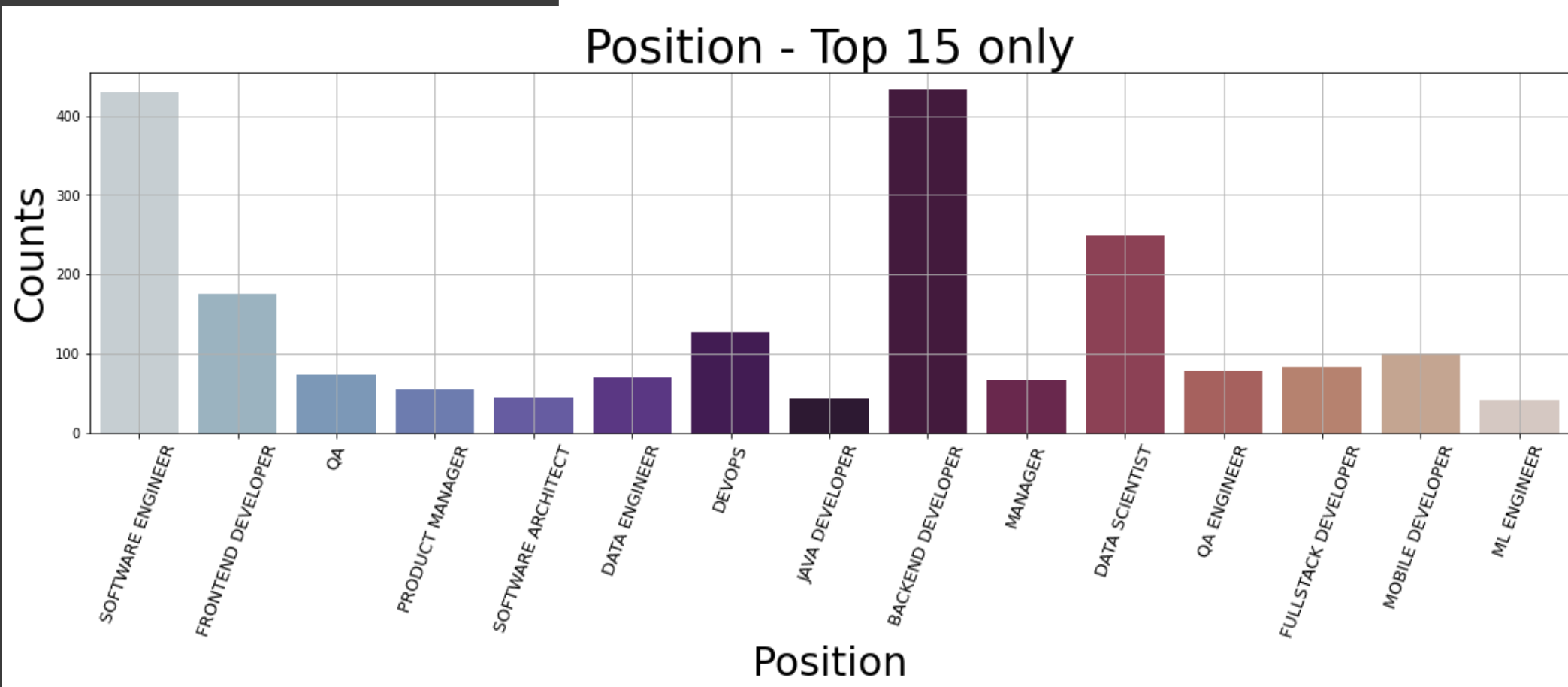The participation of male prevails more than female.

Working language

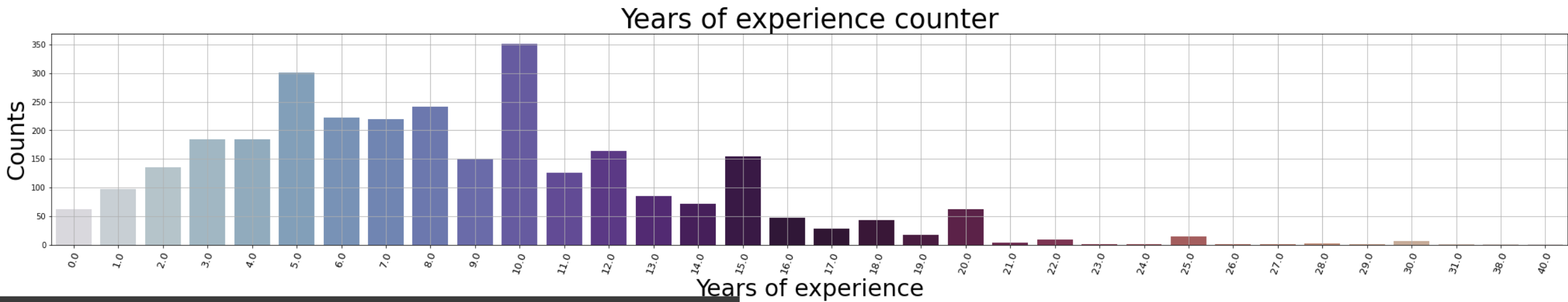Most of those who work in the IT communicate in English language.

Working language - Top 5 only

Position

The Software Engineer then Backend Developer the two most popular majors.

The ML Engineer is the less one because it is new field.

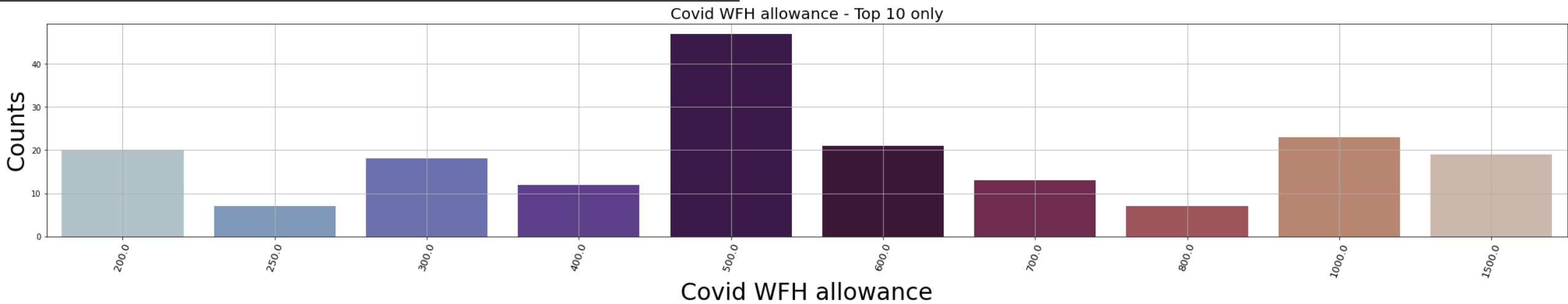Data Glacier
Your Deep Learning Partner

Years of experience

The range of years of experience of the people they working in IT field between 24-40 years.
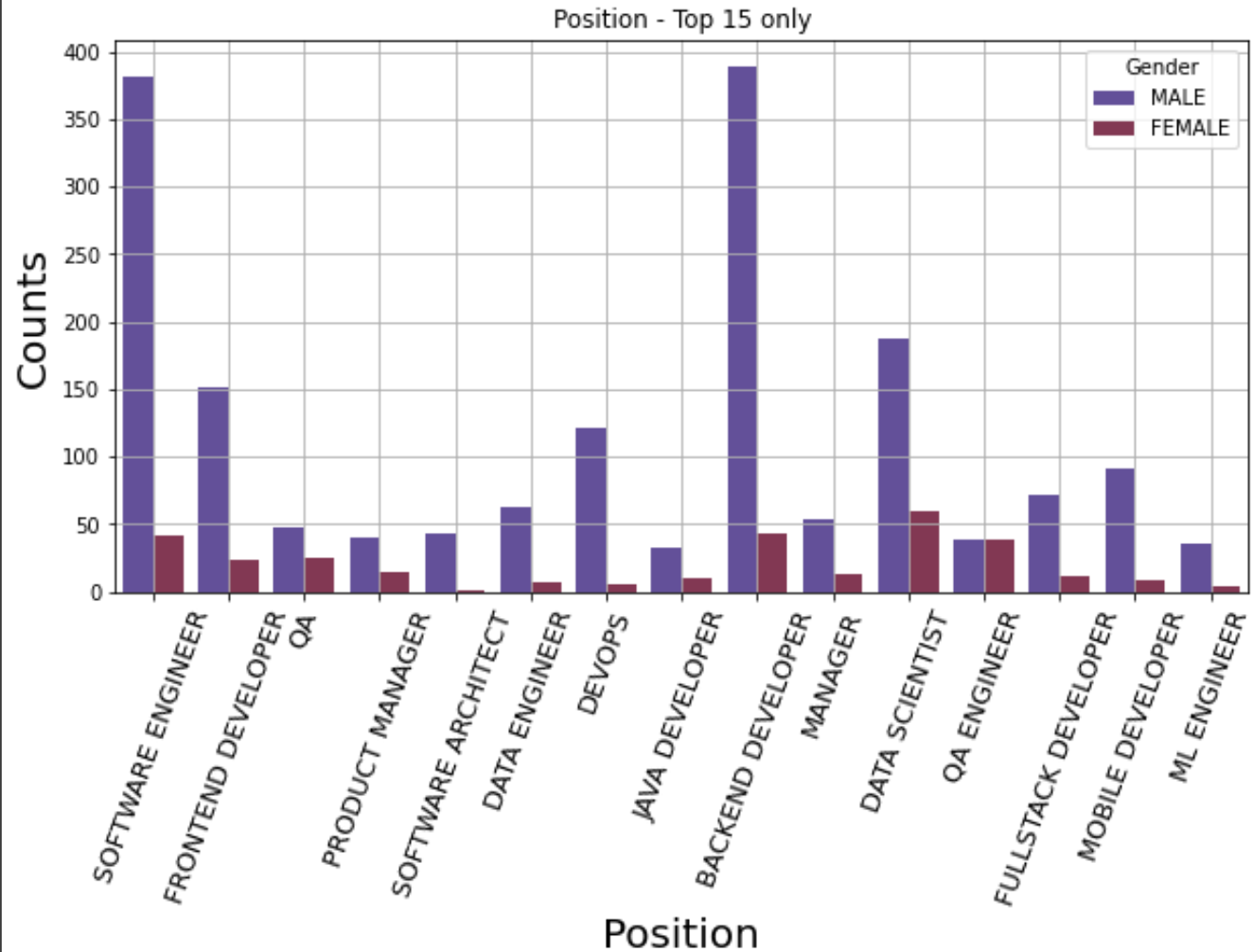
Covid WFH allowance

Most companies have tried to maintain their employees by give them   allowance of during the period of remote work in the Covid pandemic
The most of them give 500$ as allowance.
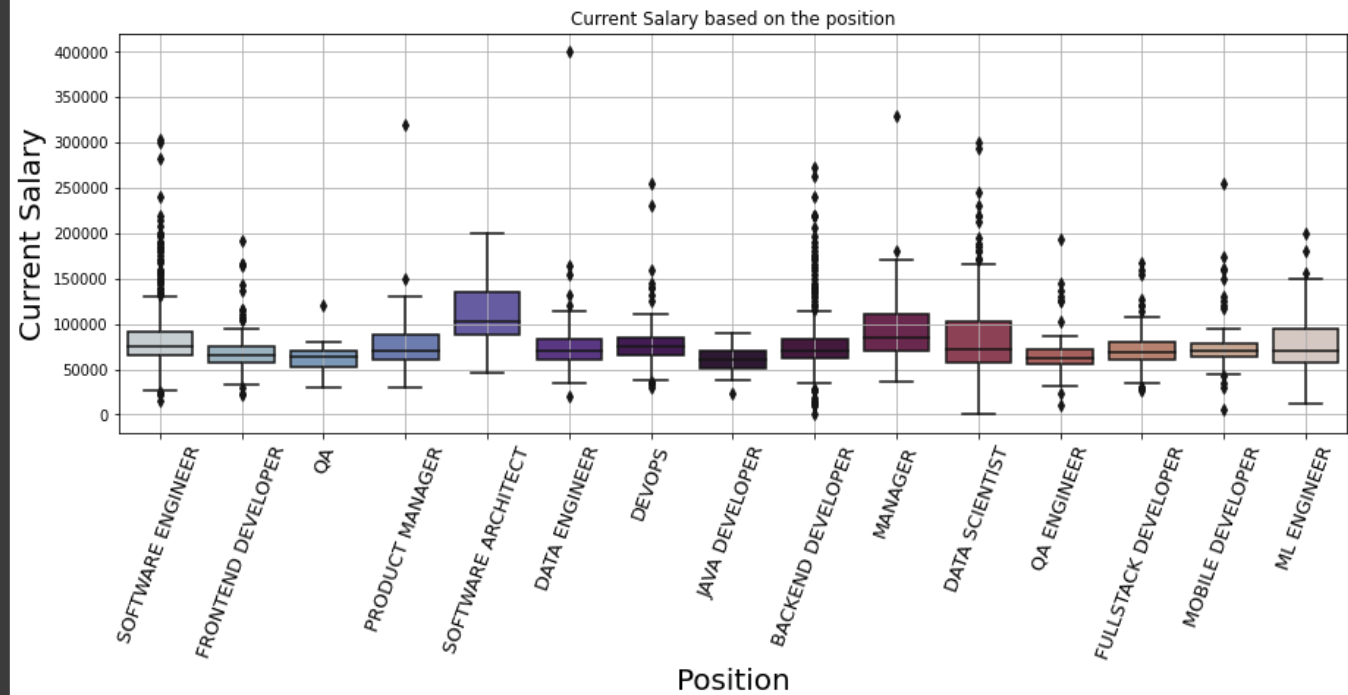

Covid WFH allowance - Top 10 only

Gender and Position

Because the number of males is more than females in the survey , their number in each positions also increases.

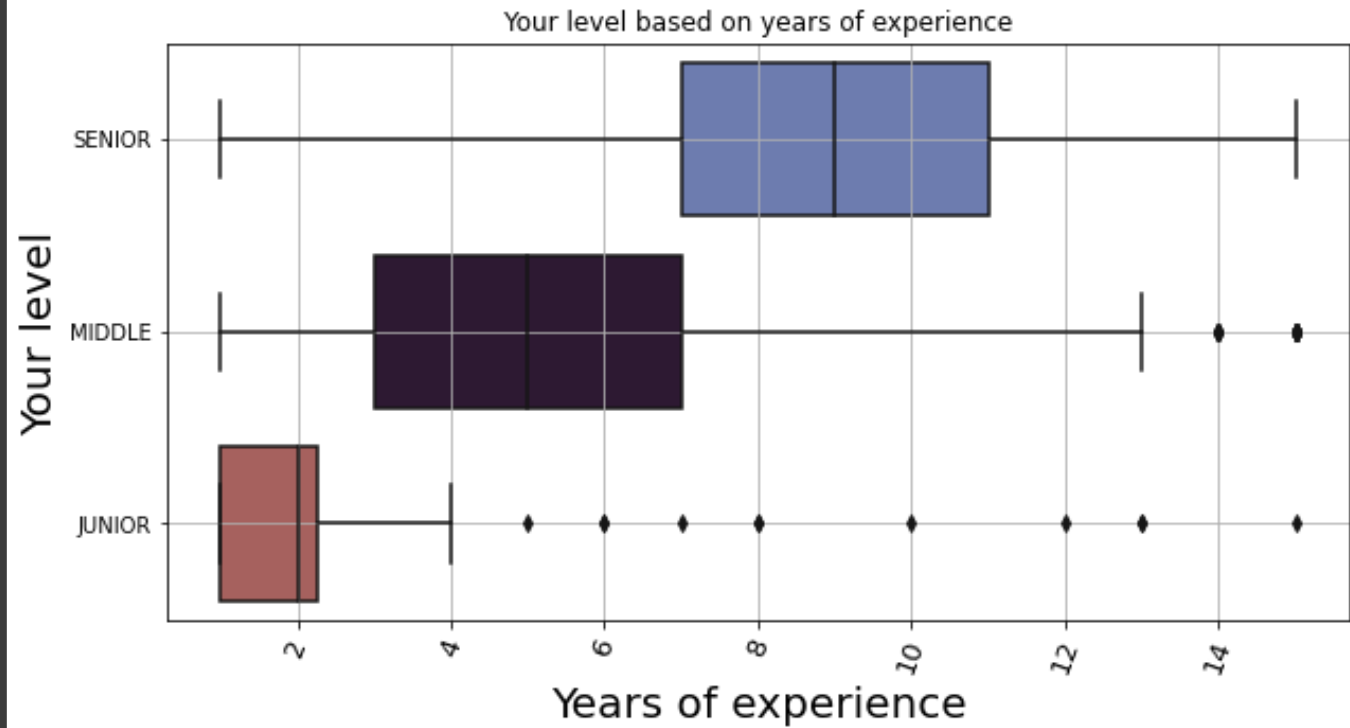**Data Glacier**
Your Deep Learning Partner

Salary by Position

The Software Engineer, Backend Developer, Data Science and Data Engineer has the most highest Salary.

level and Experience

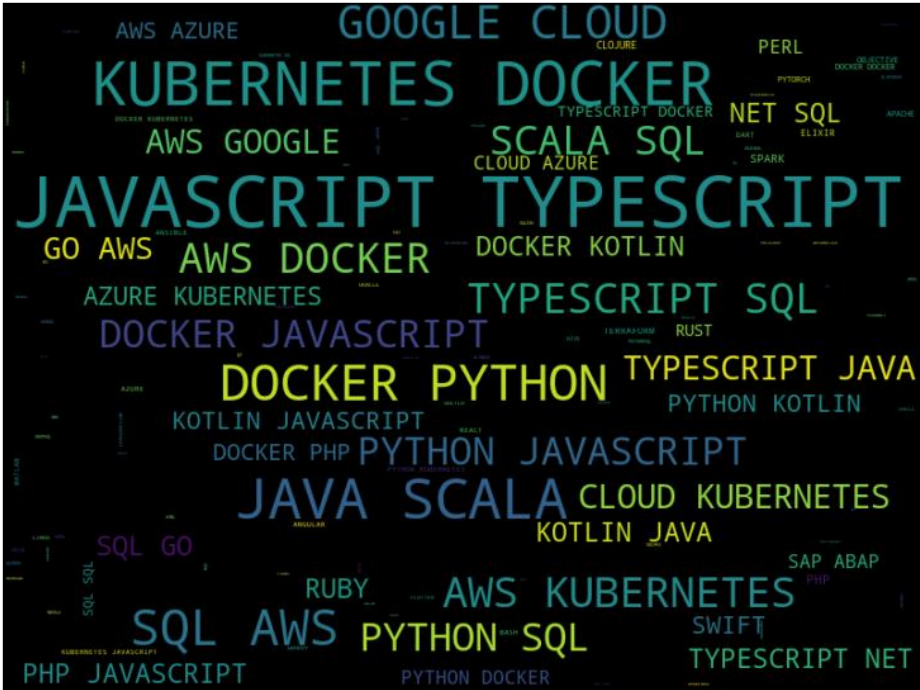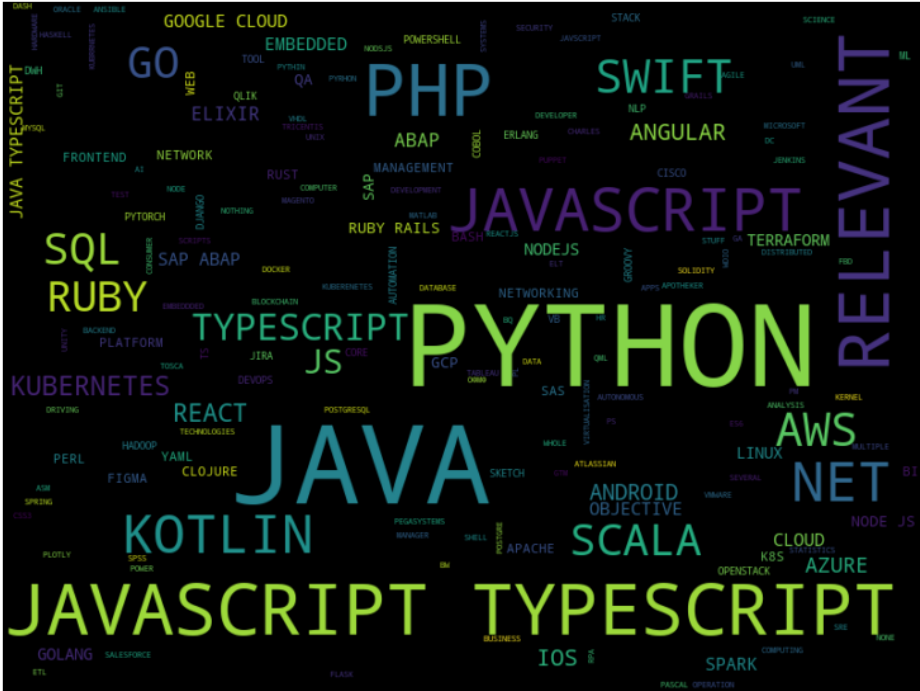The higher the level, the more years of experience, which is expected

Salary and Experience

The more years of experience, the higher the salary, which is expected.

Current Salary based on years of experience

Programing Language

The python, java, javascript and SQL the most also the highest used.
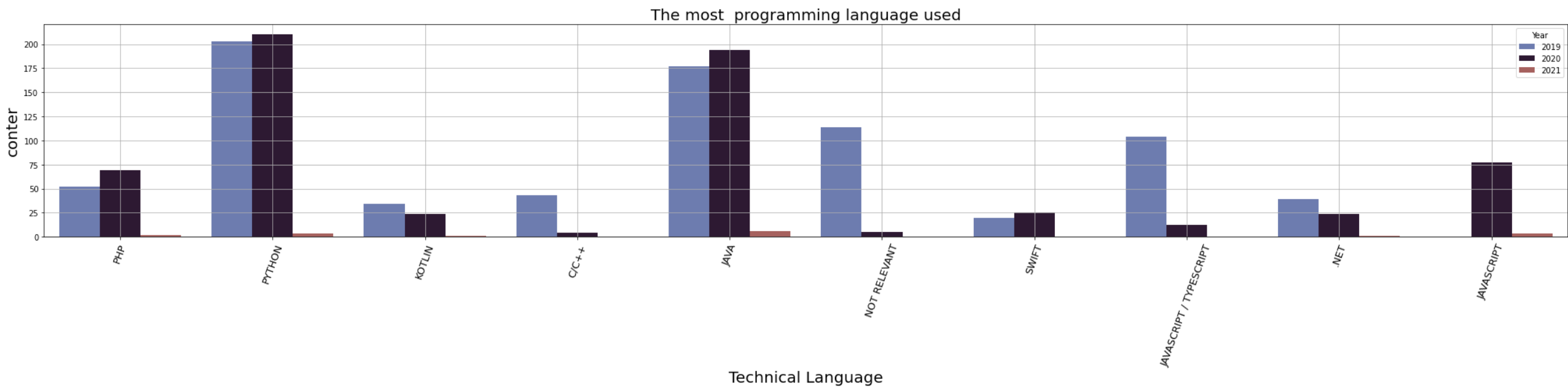
Company and Position

The products and startup the most also the highest companies type in IT field.

The Software Engineer, Backend Developer, ,Frontend Developer and Full Stack most also the highest position frequently.
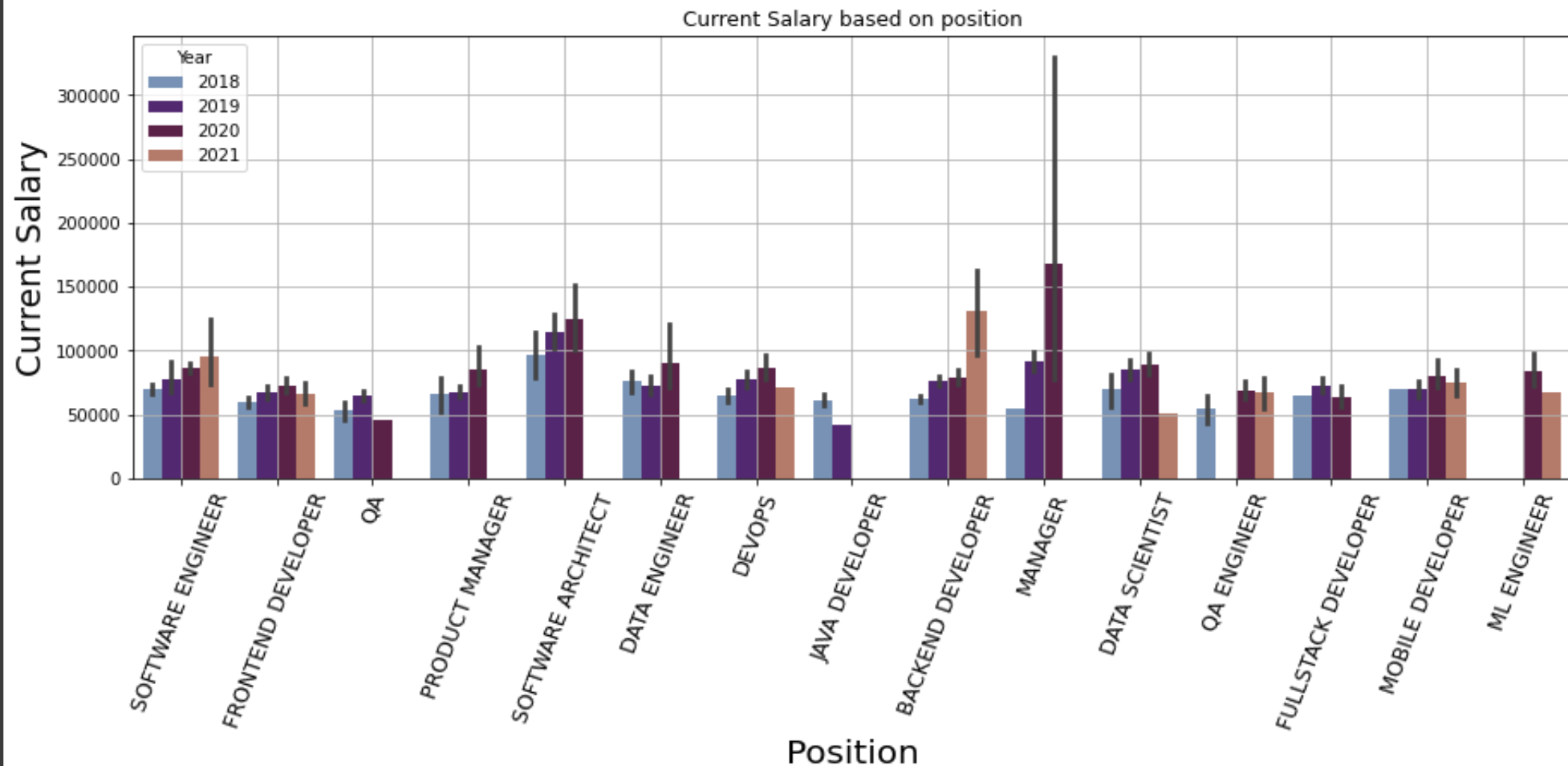
Data Glacier
Your Deep Learning Partner

Recommendations #1

The python and java the two most also the highest used recent years.

The most programming language used

Recommendations #2

The Software Engineer then Backend Developer the two most also the highest Salary recent years.

The Manager has highest salary in 2020 but there is no data for 2021.

Current Salary based on position

Thank You