SDAYA Data Science Camp

Human Resources Analytics Job Changing

Submit to: Mr.Ali El-kassas

Done by:
Arwa Alolyani , Ghaidaa Althobaiti

HR Analytics Job Changing

In our study, we dealt with data divided into training data and test data.

The training data was divided into training and verification

The test data contains 2129 rows and 13 columns, the validation contains 4790 rows and 14 columns, and the training set contains 14368 rows and 14 columns.

This data suffers from different class's, this problem will be dealt with by using the classification module.
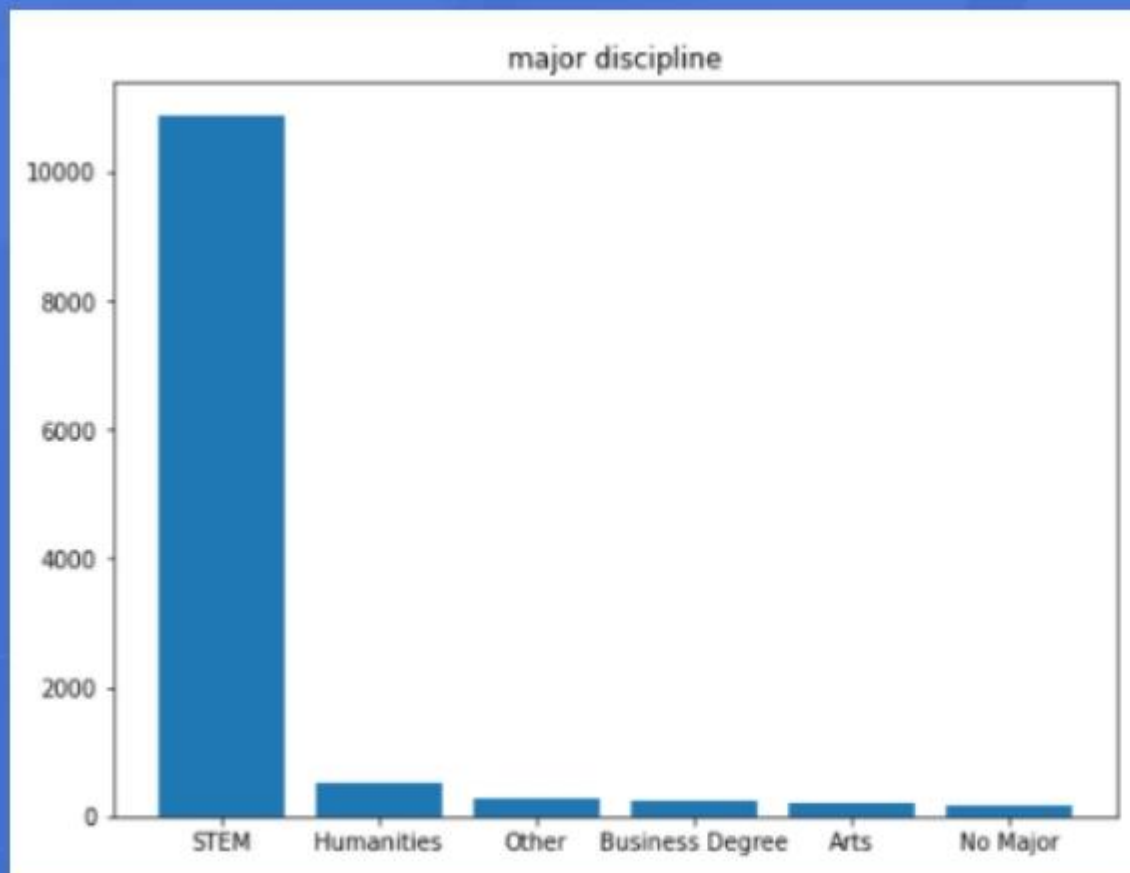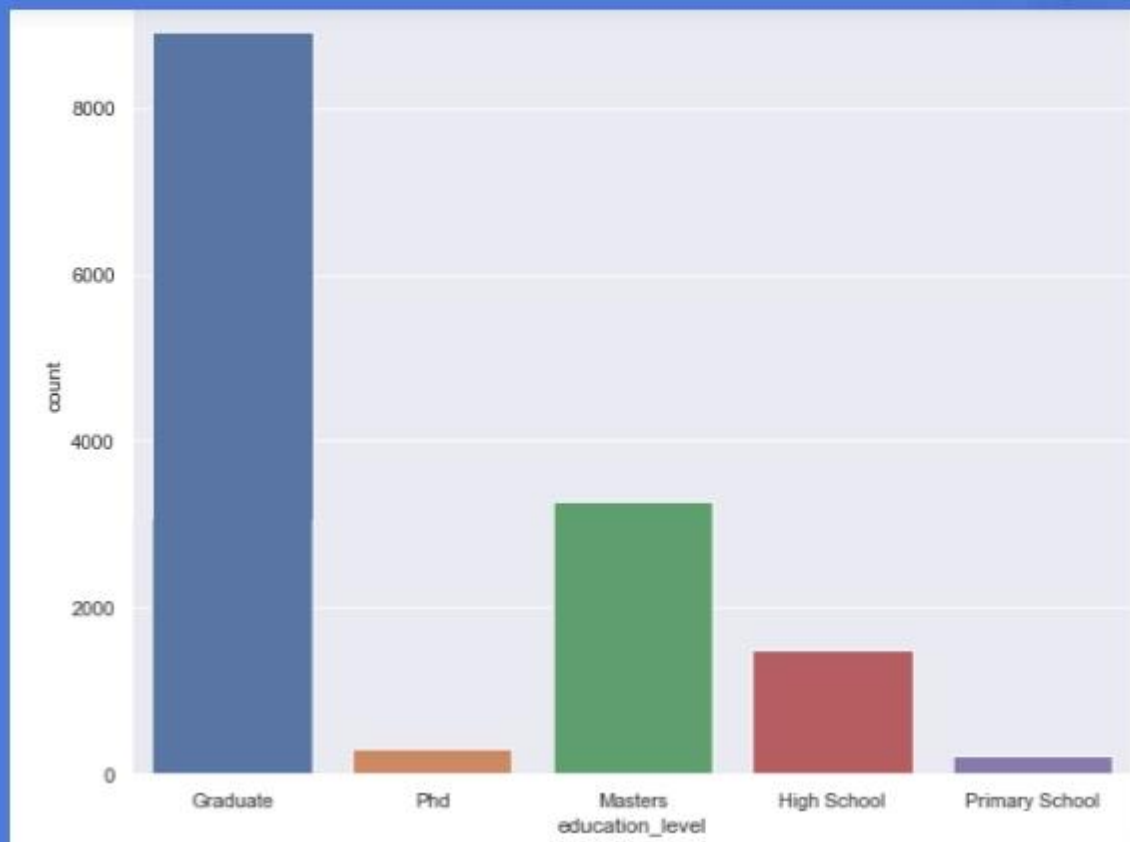
Processes :

In the beginning, we divided the data and identify its content by extracting data, analyzing it and then visualizing it. Then the data was cleaned up by removing outliers and null values, modifying some column values, adding new columns describing the employee's job level, experience and education level.

The following was extracted:

- How many years of relevant experience does the candidate have?
- What is the educational level most in demand for jobs?
- What is the most employed gender?
- What is the sex that has the most training hours?
- How many educational levels does your company have?
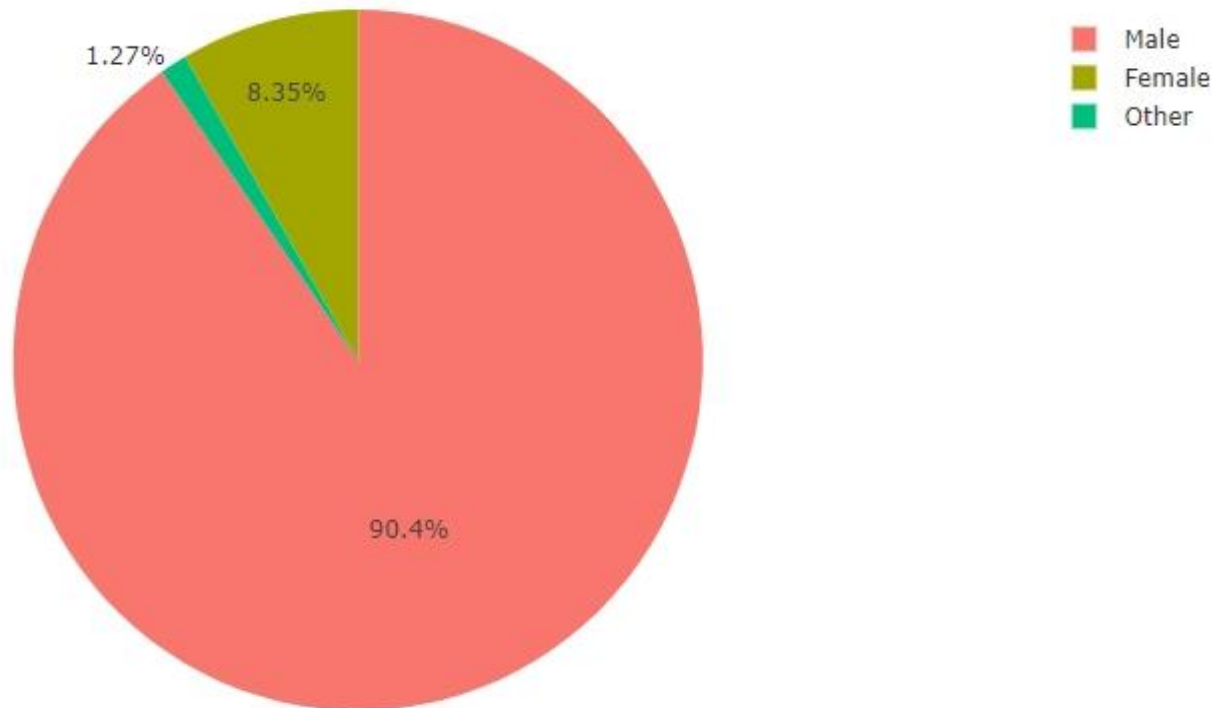- **Problem Statement:** Predict the probability of a candidate looking for a new job
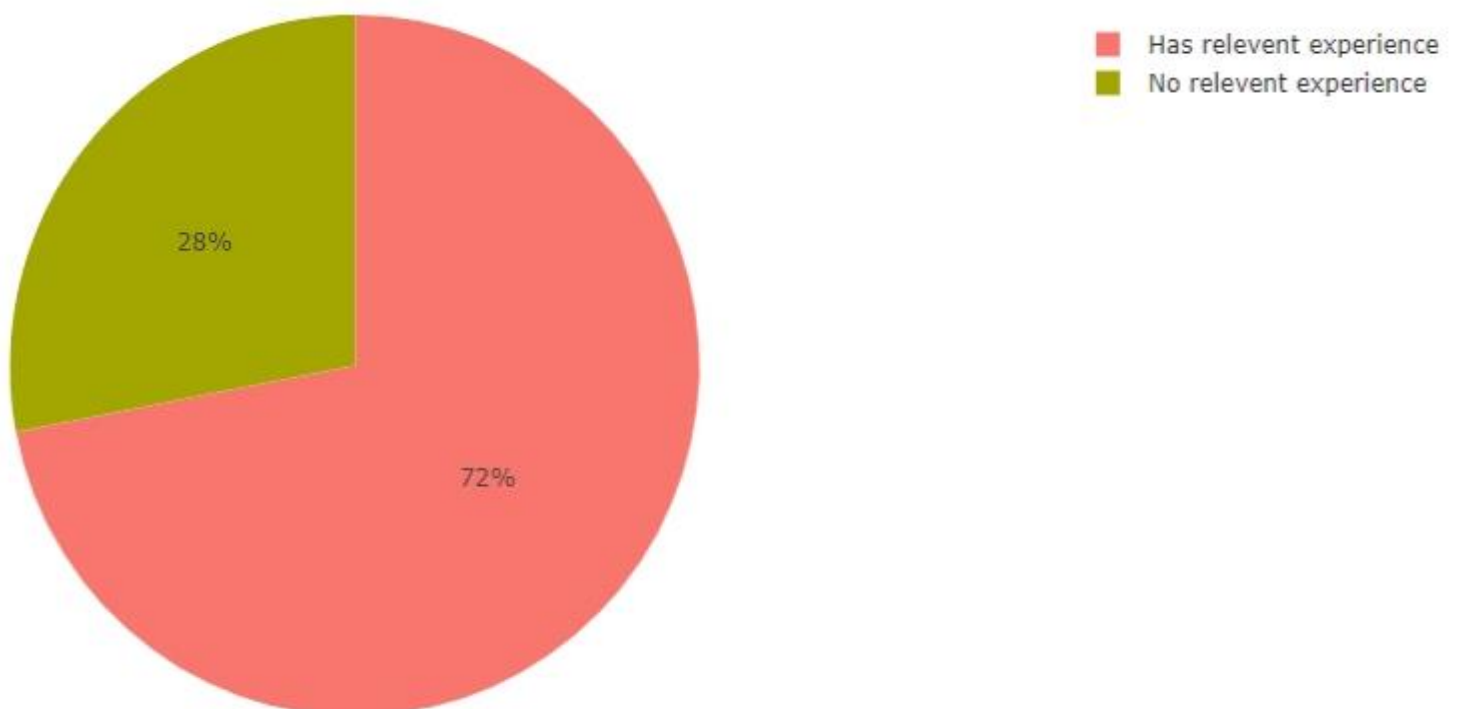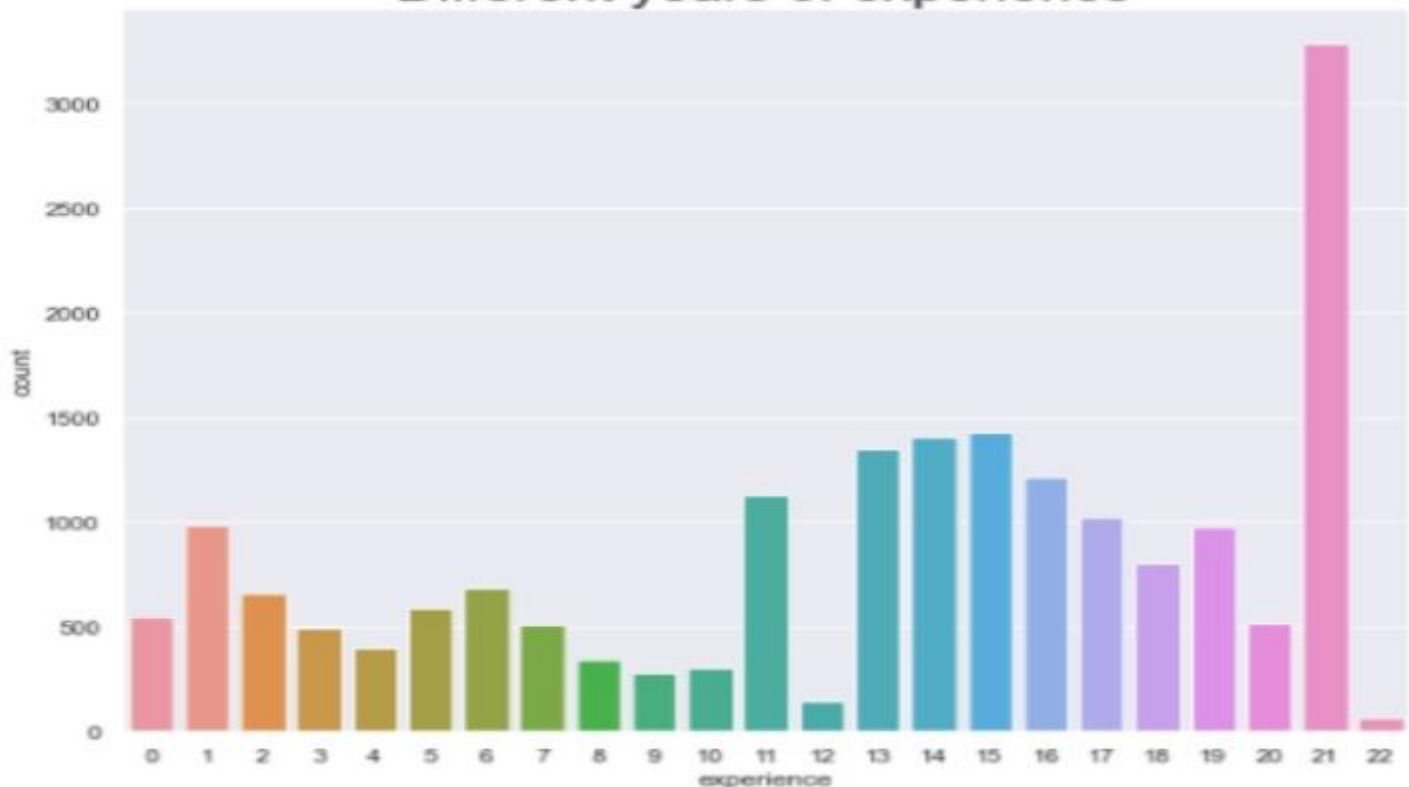
Data visualization :

Data visualization :

gender

Legend:
- Male
- Female
- Other

1.27%
8.35%
90.4%

Relevent Experience

Legend:
- Has relevent experience
- No relevent experience

28%
72%

Data visualization :

Different years of experience

The score for kNN is
Training: 82.25%
Test set: 76.20%

The score for logistic regression is
Training: 76.58%
Test set: 76.55%

Looking for New Job Percentage

No 24.93%

yes 75.07%