

Exploratory data analysis for Labor market

The first project in SDAIA Bootcamp T5 Data Science & Artificial intelligence

Abstract

The main purpose of this analytical report is to provide clear and approximate insights about the Saudi labor market in the period 2019-2020 and to elicit the most in-demand jobs during this period and the extent of the impact of the Covid-19 pandemic on the labor market. The first stage in labor market data analysis is to clean, prepare and handle the data. The second stage is to apply feature selection to the labor market data framework and reduce synthetic processing so that we achieve the objectives of this analysis. Finally, provide a clear visualization to make it easier to understand the data and extract ideas from it.

Note: This information has been extracted from Bayt.com.

Design

This project arose from questions about a rumor by many that the Saudi labor market was negatively affected during the Corona pandemic, and many students were wondering about the most demanded jobs in the future that serve the vision of the Kingdom 2030. male and female.

Data

The dataset contains 3,469 entries with 26 features, and each row contains an ID that identifies each job posting.

Algorithms

- 1- Performed data cleaning.
- 2-Displays the Columns, Types, Rows and the memory used by the dataframe.
- 3- Rename Columns because not valid in python.
- 4- Detecting Missing Values.
- 5- Handling with missing values EX: where gender is null = M/F, Degree is null put the Mode.
- 6- Add new feature (based on old feature)
- 7- Visualized the resulted clean data.

Tools

There are the technologies and libraries that I use while working on the project : Python , Jupyter Notebook.

Data manipulation: Pandas.

Plotting: Seaborn and Matplotlib.

Communication

The most important representation in the project.

