

Question/Problem statement:

The Human Resource in Company ABC wants to recruit newly graduated students from the KSA universities to fill available positions in the country. The manager asked to make a study on the data of the graduated students at the universities including all majors in the university. And prepare a report of the findings.

Questions:

- Number of BSc., master's degree, Diploma, and PHD graduate Students according to the years?
- Number of students according to the region?
- Number of graduate students according to the gender of study?
- Sorting graduate students according to the university?
- Sorting number of students according to the field of study?

Data Description:

The data set is about number of enrolled university students by field of study in Saudi Arabia 2013-2018 The dataset provided by Saudi open data

website: [http:// data.gov.sa/ar](http://data.gov.sa/ar)

Type of data	TEXT, INT
Size of data	15*67983
Rows	67983
Columns	15
Name columns used	(Year, number of students, Region, University Name, field, Specialization, gender, Academic level, Nationality, Education type, type of educational institution,)
Name columns do not need	(,educational level, type of university, state, general field)

Tools:

- Programs: SQLite, Jupyter
- Libraries: Pandas, NumPy, Seaborn and Matplotlib.
- Functions: group by
- Plots: pie plot, Histogram

MVP Goal:

The aim of this project is to anticipate students who are most qualified for job in a field of study. Knowing the majors that have the most the number of students