

LinkedIn jobs postings Exploratory Data Analysis



Raghad Althunayan
13 November 2021

Introduction

Many job applicants wish to benefit as much as possible from the jobs published on LinkedIn. In this project, I will analyze exploratory data for LinkedIn job postings based in Saudi (Dec 2019 to Mar 2021) to answer some of the questions.

Question/need:

- 1- What are the companies that may hire T5 Bootcamp graduates?
- 2- How many companies posted a job post on LinkedIn?
- 3- What's the average working experience years required?
- 4- What is the peak month of job posting?

Dataset

The LinkedIn job posting dataset represents the name of position and company, numbers of experience. Data was collected by data engineer Majed Al Hulayel.

I will do the analysis on the period from Dec 2019 to Mar 2021.

Dataset Description:

- 13 Columns
- 294762 Rows

Field Name	Description
position_id	Represents id unique number of position
Position	Represents the name of position
Company	Represents the name of the company that posted the position
Location	Represents the company location
Date	Represents the date (MM-DD-YY)
Level	Represents the position level
Job_Function	Represents field of activity
Industry	Represents the economic classification of main business activity in the company
Description	Represents the description of the position
Link	The link of job position



linkedin_id	Represents LinkedIn user unique number
company_linkedin_url	The link of company
experience	Represents the number of working years required

Algorithms

For data cleaning and pre-processing, I will start by deleting the duplicate records and check if there are any null values then drop them. I will visualize the result and try to find any correlations such as the relationship between the positions and experiences.

Tools

- Technologies: SQL, SQLite , Python, Jupyter notebook
- Libraries: Numpy, Pandas, Matplotlib, Seaborn