

# Project 1: Analysis of job market data to explore the open positions related to the jobs in the field of data or machine learning in a particular region.

## A.Data collection method:

-> Job market data was collected from <https://www.linkedin.com/> website for

the post of “Data Scientist in USA” using octoparse.

-> Octoparse is a modern visual web data extraction software. Both experienced and inexperienced users would find it easy to bulk-extract information from websites with it. For most scraping tasks, no coding is needed. Octoparse supports Windows XP, 7, 8, and 10. It works well for both static and dynamic websites, including those web pages using Ajax. To export the data, there are various data formats of your choice like CSV, EXCEL, HTML, TXT, and databases. Octoparse simulates human operations to interact with web pages.

1	Job Titles	Salary	Location	Employment Ty	Industry
2	Data Analyst	\$90,000	Atlanta, GA	Full Time	Industrial Automation
3	ML engineer	\$1,10,000	Boston, MA	Full Time	IT Services
4	Jr Data Analyst	\$85,000	Chicago, IL	Full Time	IT Services
5	Data Analyst	\$78,000	Palo Alto, CA	Full Time	IT Services
6	Sr Data Scientist	\$1,20,000	Boston, MA	Full Time	IT Services
7	Jr Data Scientist	\$93,000	Houston, TX	Full Time	IT Services
8	Sr Data Scientist	\$96,000	Austin, TX	Full Time	IT Services
9	Data Scientist	\$1,08,000	Atlanta, GA	Full Time	IT Services
10	Data Scientist	\$1,02,000	Seattle, WA	Full Time	Industrial Automation
11	Data Engineer	\$98,000	Tucson, AZ	Full Time	IT Services
12	Sr Data Engineer	\$1,15,000	Boston, MA	Full Time	IT Services
13	AI engineer	\$90,000	Austin, TX	Full Time	IT Services
14	Devops Engineer	\$96,000	San Francisco, CA	Full Time	IT Services
15	Data Analyst	\$91,000	Bellevue, WA	Full Time	IT Services
16	Data Engineer	\$88,000	Atlanta, GA	Full Time	IT Services
17	Data Scientist	\$1,06,000	Bellevue, WA	Full Time	IT Services
18	ML engineer	\$83,000	San Francisco, CA	Full Time	Management Consulting
19	Data Engineer	\$79,000	Seattle, WA	Full Time	IT Services
20	Data Analyst	\$86,000	Tucson, AZ	Full Time	IT Services

## Sample Data

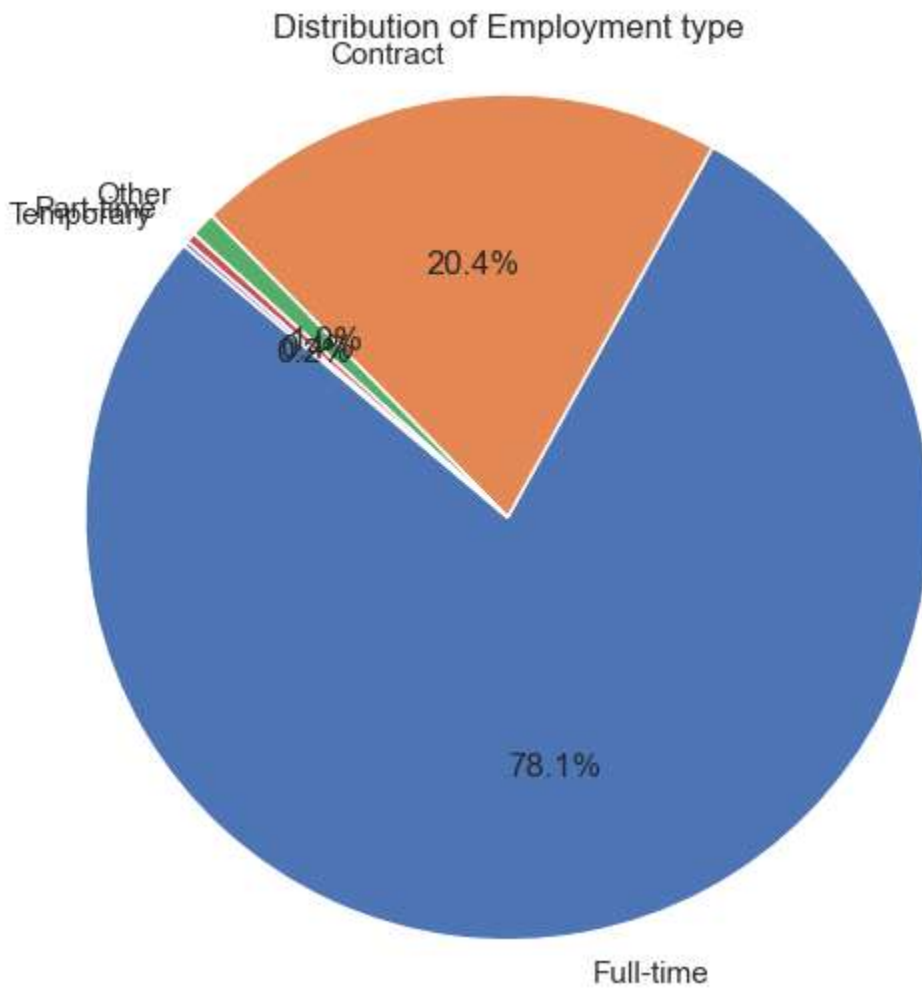
-> Extracted data is more than 800 in number but after cleaning the data got 400 in number.

-> I have done Data Visualization with some sort of data

## B. Market Data Visualization :

### 1. Employment Type:

-> There are 5 to 6 type of Employments which are Full time, temporary, Contract, Part Time etc. In that, Full time employment have majority of the people and contract based people are second highest in this and Others contain in very few percentage.



---

## 2. Titles of Positions Available:

-> There are various positions are available like Data Scientist, Data Analyst, Machine Learning Engineer, Devops Engineer etc. In this, Data Scientist has highest number of Positions available.

The plot displays the frequency of various job titles. The y-axis represents 'Frequency' from 0 to 80. The x-axis represents 'Titles' with labels rotated 45 degrees. The plot shows a dense distribution of titles, with a prominent peak for 'Data Scientist' and 'Machine Learning Engineer'.

the skillset it required. Because I was good at Mathematics and I started learning Python because of its various libraries which are very easy to use.

-> These are the Relevant skills required for Machine Learning Engineers--:

:: Applied Mathematics.

:: Neural Network Architectures.

:: Physics.

:: Data Modeling and Evaluation.

:: Advanced Signal Processing Techniques.

:: Natural Language Processing.

:: Audio and video Processing.

:: Reinforcement Learning.