**CHAPTER 1**

**INTRODUCTION**

**1.1 Context**

This project has been done as a part of my course for the **BS Information Technology** at Bahaudin Zakariya University Multan. I had Six month to fulfil the requirements to succeed the module. Every three weeks, a meeting was organized to show and report my progress and fix the next objectives.

**1.2 Motivations**

Being extremely interested in everything having a relation with the Data Science, the independent project was a great occasion to give me the time to learn and confirm my interest for this field. The fact that we can handle a million of data and take useful insights from it. We can use Data Mining in Finance, Medicine, almost everywhere. That is why I decided to conduct my project around the Data Mining.

**1.3 Idea**

This project was motivated by my desire to investigate the descriptive analysis field of Data Mining since it allows to approach big data which is a very hot topic. Following my previous experience where it was about airline data descriptive analysis, I applied the same idea with technologies and try to figure out which is:

* What are the top programming languages in demand?
* What are the top database skills in demand?
* What are the popular Development Platforms?
* What are the popular Web frameworks?

**1.4 Sources**

Because I truly think that sharing sources and knowledges allow to help others but also ourselves, the sources of the project are available at the following link:

https://github.com/jonygujjar/Final-Year-Project

Feel free to give me your point of view or ideas for anything you want. I used ipython notebook which is very useful to understand the entire process of my project since you can follow each step with the corresponding code.

**1.5 Purpose**

To keep pace with changing technologies and remain competitive, the organization regularly analyses data to help identify future skill requirements. That is what my project is doing. This project will evaluate some of the top trending technologies that will be useful not only for the organization but also for the tech professionals by different ways. Here are some of the main processes where my project will be helpful.

* Organization Hiring Process.
* Organization System Upgradation or Modifications.
* Skill development Process.

**1.Hiring process**

Hiring Process always play an important role in the future of every organization. By the hired person, the direction any organization has been determined. In hiring process, the skill set or required skills have been listed for the job postings. In this job posting the required skills are identified by management of the organization from which the future of the organization is determined. This skill set is very important. So, in order to determined which skill set is trending now and have long lasting positive effect on organization. There is always a problem in selecting this skill set which are not only effective but also are future proof.

That is what my project is solving this problem for the organizations which provides a list of top trending technologies available for hiring process.

**2.System upgradation**

System upgradation or new system installation is also important for the proper and efficient working of an organization. Multiple technologies are available to implement in existing and new system. So, in order to decide which technology should be implemented for maximum performance and efficiency. For example, there are multiple databases are available for information system, but we have to choose the one which should be cost efficient and future proof so that organization do not need to change it after some time period. Here comes my project to solve the mystery which proposed the top trending databases and other technologies where you can choose one from a short-listed trending technologies according to your organization needs.

**3.Skill development process**

Every organization pays a lot for **skill development process**. This is not the case of only organization but also for a single person who have a desire to develop new skills. There are a lot more technologies are available in which only some are more effective and provide a trust of future proofing. To determine which skills are trending and have a secure profitable future needs complex research and even we can go to a wrong direction which in last represent a loss in term of money and time. To ensure which skill is desired and have a profitable future we need a companion that provides a accurate set of data upon which we can made a right decision. That is what my project is doing it provides a list of top trending technologies where you or your organization can invest.

**1.6 Resources**

There are multiple resources are used in my project from data wrangling process to the data visualization process which are listed below:

* **Python** (Python is an interpreted high-level general-purpose programming language). Link: [https://www.python.org](https://www.python.org/)
* **Pandas** (pandas is a software library written for the Python programming language for data manipulation and analysis). Link: <https://pandas.pydata.org>
* **NumPy** (NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays). Link: <https://numpy.org>
* **Matplotlib** (Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy). Link: [https://matplotlib.org](https://www.python.org/)
* **Dash (**Dash is a python module which is used for creating web dashboards from python program.) Link: https://dash.plotly.com/
* **Scikit-learn** (Scikit-learn is a free software machine learning library for the Python programming language.) Link: <https://scikit-learn.org/stable/.org>
* **Plotly** ( Plotly's Python graphing library makes interactive, publication-quality graphs.) Link: https://plotly.com/python/
* **Folium** (Folium is a Python library used for visualizing geospatial data.) Link: <http://python-visualization.github.io/folium/>
* **Jupyter Notebook** (Project Jupyter is a project and community whose goal is to "develop open-source software, open-standards, and services for interactive computing across dozens of programming languages".)

Link: https://jupyter.org/

* **Dash Bootsrap** (dash-bootstrap-components is a library of Bootstrap components for Plotly Dash, that makes it easier to build consistently styled apps with complex, responsive layouts.) Link: https://dash-bootstrap-components.opensource.faculty.ai/