**SYNOPSIS REPORT ON**

Data Scraping and Visualisation with Python

*Submitted in partial fulfillment of the requirements for the award of the degree of*

**BACHELOR OF COMPUTER APPLICATIONS**

****

**Batch: 2021 - 24**

***Under the Guidance of Submitted By***

Dr. Santosh Kumar Singh Parth Verma 01428402021

Anmol Sahni 01528402021

**Bosco Technical Training Society**

**Don Bosco Technical School, Okhla Road, New Delhi-110025**

**9643868820, 8527787221, 011-41033889**

**Affiliated to**

****

**Guru Gobind Singh Indraprastha University**

### Sector - 16C, Dwarka, New Delhi - 110078 (India)

### Phone: +91-11-25302170, Fax: +91-11- 25302111

### E-Mail Id: ggsipu.pr@rediffmail.com, pro@ipu.ac.in

Table of Contents

[1. Problem Statement 3](#_Toc147262940)

[2. Objectives and Scope 4](#_Toc147262941)

[3. Methodology 5](#_Toc147262942)

[4. Hardware and Software Used 6](#_Toc147262943)

[5. Resources and Limitations 7](#_Toc147262944)

[6. Testing Methodology 8](#_Toc147262945)

[7. Conclusions 9](#_Toc147262946)

# **Problem Statement**

**Problem**

* Lack of Analysis and visualisation when dealing with online social networking sites for working professionals seeking job opportunities. Data scraping and visualization project can solve various problems across different domains such as Market Research and Competitor Analysis.

**Objectives and Scope**

* The primary objective of a data scraping and visualization project is to extract, process, and analyse data from concerned websites or other data repositories, and then present the insights gained from that data in a visual and meaningful way.
* To Develop a python program to take user inputs from user and perform data scraping on the basis of data inputted by the user.
* Scraping data with optimum speed and visualise the data scraped in the form of insights and visualisation which convey meaningful data.

**Methodology**

**Data Collection:** Gather relevant data from one or more sources to support analysis and decision-making.

**Data Processing:** Clean, transform, and structure the raw data for analysis.

**Data Analysis:** Extract meaningful insights, trends, patterns, and correlations from the data.

**Data Visualization:** Present the analysed data in a visual format that is easy to understand and interpret

**Decision Support:** Enable users to make informed decisions based on the visualized data and insights.

**Automation:** Streamline the process of data collection, analysis, and visualization through automation, where applicable.

**Hardware and Software Used**

* **Hardware Requirements:**
  + RAM: 4 Gb minimum
  + Storage: 1 Gb minimum
  + Internet speed over 30mbps
* **Software Requirements:**
  + Operating system platform
  + Database management system- Any SQL server
  + Programming languages – Python
  + Scraping libraries - Beautiful Soup, Selenium
  + Data analysis tools –Matplotlib, Seaborn, Pandas, NumPy,
  + Framework Support: Django
  + IDEs: Python IDE

**Testing Methodology**

Following are the testing techniques used

* Manually inspect and validate the scraped data from different sources to ensure accuracy, completeness, and integrity.
* Verify that the data processing and transformation procedures meet expectations.
* Review data visualizations (e.g., charts, graphs, dashboards) to ensure they accurately represent the data and provide meaningful insights.
* Check for visual and design consistency.
* Verify that users can customize, interact with, and understand the data visualizations.
* Monitor response times, system resource usage, and scalability.

Completion Criteria: The testing process is considered complete when all identified test cases have been executed, and no further defects or bugs are found.

**Conclusions**

* In conclusion, the data scraping and visualization project aims to empower users with valuable insights from diverse data sources. By following systematic manual testing methodologies, we can ensure the accuracy, functionality, and user-friendliness of the application and visualizations. With rigorous testing and continuous improvement, the project not only meets its objectives but also provides a powerful tool for informed decision-making, efficient data analysis, and enhanced user experiences.