Project Outline: Web Scraping using Python.

- **Background of the study:**
  - ➢ The background of the study for web scraping lies in the growing importance of data in the modern world. Data has become a critical asset for businesses and organizations, and the ability to collect, analyze, and interpret data has become a competitive advantage. However, web scraping provides a solution to this problem by automating the process of data collection and making it easier to access, analyze and automate.

- **Research Problem:**
  - ➢ One of the primary challenges in web scraping is dealing with the vast amount of unstructured data available on the internet. This data is often scattered across multiple websites and can be in different formats, making it difficult to collect and analyze.

- **Research Questions:**
  - ➢ What are the most effective techniques for web scraping and processing large amounts of unstructured data from the internet?
  - ➢ How can the impact of web scraping on the websites being scraped be minimized to prevent performance issues or increased server load?
  - ➢ How can web scraping be used to track market trends and gather competitive intelligence to inform business strategies?

- **Data Set:**
  - ➢ There are various types of data sets that can be used for web scraping using Kaggle website.
  - ➢ E-Commerce product data – Amazon, eBay and Walmart.
  - ➢ Job Posting data – Job title, location, salary, and company information.
  - ➢ Sports data – Game scores, Player Statistics and Team ranking.

- **Analysis:**
  - ➢ One of the key benefits of web scraping is its ability to collect large amounts of data in a relatively short amount of time. This data can be used for a variety of purposes, such as market research, price monitoring. Web scraping can also be used to automate data collection processes, saving time and resources for businesses and organizations.

- **Findings:**
  - Pricing information: Web scraping can be used to collect pricing data from e-commerce websites, which can be used to analyze price trends, identify competitors, and inform pricing strategies.

- **Discussion:**
  - One of the main benefits of web scraping is its ability to automate data collection, saving time and resources for businesses and organizations. This can be particularly useful for collecting data from large numbers of websites or for monitoring websites on a regular basis.

- **Limitations:**
  - Incomplete or inaccurate data: Web scraping tools can encounter errors, such as incomplete or missing data, which can affect the quality of the data set. Additionally, web scraping can sometimes encounter technical challenges.