WEBSITE ANALYSIS AND VISUALIZATION

Hemanth Aradhya BR*1, Manjunath B*2

*1Student, Department Of Master Of Computer Application, Maharaja Institute Of Technology Mysore, Mysore, Karnataka, India.

*2Assistant Professor & HOD, Department Of Master Of Computer Applications, Maharaja Institute Of Technology Mysore, Mysore, Karnataka, India.

ABSTRACT

Website analysis and visualization using SEMRush Data provides analysis of a website's performance, including domain analysis such as the volume of organic traffic, the number of backlinks and referring domains, a keyword overview, the domain's organic keywords, and organic competitors etc.

I.   INTRODUCTION

On the two websites isocrates.com and madtechbi.com, the investigation was conducted. We obtain a significant amount of data when we integrate the website with the SEMRush tool. By using the Application Programming Interface (API), raw semrush data can be obtained without logging into the software directly. We obtain the Using python we pull the required data through API.

We can analyses the unstructured data we get from the SEMRush API to learn more about it, and we can use pandas to turn it into structured format. Next, we used the data we had just collected to generate some insights and then displayed the data using a visualizations tool. After visualization, it offers comprehensive knowledge of our websites isocrates.com and madtechbi.com. We also need to comprehend what doesn't operate well for our business. We can determine whether we are accomplishing our aims and objectives and where we may improve our needs by using data visualization.

II.   METHODOLOGY

SEMRush: The tool is frequently used for keyword research and gathering information on internet rankings, including indicators like search volume and cost per click (CPC). The programme additionally gathers keyword data from the Google and Bing search engines.

Python: Python is a general-purpose, high-level, interpreted programming language. Code readability is prioritized in its design philosophy, which makes heavy use of indentation.

Python uses garbage collection and has dynamic typing. It supports a variety of programming paradigms, including procedural, object-oriented, and functional programming as well as structured programming (especially this). Due to its extensive standard library, it is frequently referred to as a "batteries included" language. Python API is used to obtain historical and current data.

Application Programming Interfaces, or APIs, are servers that may be accessed using programs to retrieve and deliver information. The primary purpose of APIs is to obtain data. We must submit a request to an API in order to get data from it. The way API requests operate is just the same as when you ask an API server for data, and the server answers.

Pandas: Pandas is a Python module used for statistical analysis and data modification. It is an open-source package that helps a various types of data processing activities and is quick and simple to use. These cover wrangling, statistical analysis, merging, and several other processes.

Data Frame objects with inbuilt indexing that is quick and effective for manipulating data. Tools for reading and writing data between in-memory data structures and many formats, including text and CSV files, Microsoft Excel spreadsheets, and the quick HDF5 format. Gain automated label-based alignment in calculations and quickly transforms jumbled data into an ordered shape with intelligent data alignment and integrated management of missing data.

Data sets can be easily reshaped and pivoted. Slicing of enormous data sets with clever label-based algorithms, elaborate indexing, and sub setting. Data structures' size mutability can be achieved by adding and removing columns. Using a robust group by engine to combine, divide, and perform operations on data sets while aggregating or processing data. Joining and merging of data sets with high performance.

Jupyter Notebook: The most recent web-based interactive development environment for code, data, and notebooks is JupyterLab. Users may create and arrange workflows in data science, scientific computing, computational journalism, and machine learning using the interface's flexibility. A modular structure encourages expansions to increase and improve functionality. The notebook provides a web-based application ideal for recording the entire computation process, including designing, documenting, and running code as well as conveying the results. This expands the console-based approach to interactive computing in a qualitatively new direction. Two elements are combined in the Jupyter notebook. A web application is a browser-based tool for interactively creating documents that incorporate mathematical calculations, explanatory text, and their rich media output. A representation of all material viewable in the web application, including calculation inputs and outputs, explanatory language, math, graphics, and rich media representations of objects are included in notebook papers.

Plotly: An interactive, open-source plotting toolkit for Python, plotly provides over 40 different chart types for a variety of statistical, financial, geographic, scientific, and three-dimensional use-cases. Plotly, a Python tool that is built on top of the Plotly JavaScript library (plotly.js), allows users to create stunning interactive web-based visualizations that can be viewed in Jupyter notebooks, saved to standalone HTML files, or used as a component of web applications that are entirely written in Python and served using Dash. To distinguish it from the JavaScript library, the plotly Python library is frequently called "plotly.py."

III.   MODELING AND ANALYSIS
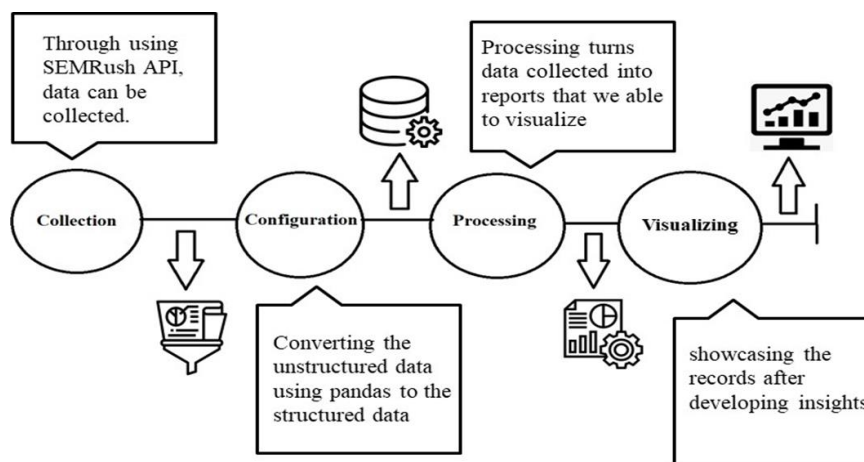
IV.   RESULTS AND DISCUSSION

Test Cases

Test Case No	Test case	Required Input	Expected Output	Actual Output	Test pass/Fail
1	Domain Overview	Request parameter + API Key	Historical and current information on the volume of organic keywords and organic traffic	As expected	Pass
2	Site Audit	Request parameter + API Key	Get an overview of the audit including the number of found issues errors, warnings, and notices	As expected	Pass
3	Visibility	Request parameter + API Key + select period	Domain visibility over the selected period	As expected	Pass
4	Organic Overview	Request parameter + API Key + Keyword	This information includes the number of new and lost keywords	As expected	Pass

5	Domain Organic Search Keywords	Request parameter + API Key + Keyword	Top 10 organic keywords which brings user to our domain	As expected	Pass
6	Competitors in Organic Search	Request parameter + API Key	Competitor domains that appear in organic search results	As expected	Pass
7	Domain v/s Domain	Request parameter + API Key+domain1+domain2	Enables users to compare two domains based on factors	As expected	Pass
8	Phrase Questions	Request parameter + API Key + Keyword	Displays a list of keyword-related phrase questions	As expected	Pass

V.   CONCLUSION

A decent page performance is crucial for modern Website. Poor page performance will result in fewer visitors, which may result in a loss in business. So website analysis is a great option that gives you the crucial knowledge you'll need to enhance your website. We can increase the web exposure of our domain and learn some marketing lessons. A site audit helps to assess a website's page performance and lets us know how many faults and warnings are there, allowing us to quickly remedy the problem. You may gain more knowledge about the visitors to your website by viewing the data, which is one of its key benefits. Analytics reveals how users arrive at a website, how long they stay there, and how much traffic it receives.

VI.   REFERENCES

[1]   www.w3shools.com
[2]   https://pandas.pydata.org/
[3]   https://isocrates.com/
[4]   https://madtechbi.com/
[5]   https://developer.semrush.com/api/v3/analytics/
[6]   https://plotly.com/python/getting-started/#plotly-charts-in-dash