Data Science and Analytics Practical Class Manual

Introduction

Welcome to the practical class on data science and analytics. In this class, you will learn how to use APIs for data collection and, with the help of Python, gather, transform, and store information locally. We will use the DataForSEO platform and the Google Maps API to explore real-world use cases relevant to your region, Ghana.

1. About DataForSEO

What is DataForSEO?

DataForSEO is a comprehensive platform that provides access to various data APIs specifically designed for SEO, digital marketing, and data analysis. It offers a range of APIs, including SERP API, Keyword Data API, Traffic Analytics API, and more. These APIs allow you to collect data from search engines, social media platforms, and other sources to gain insights and make informed decisions.

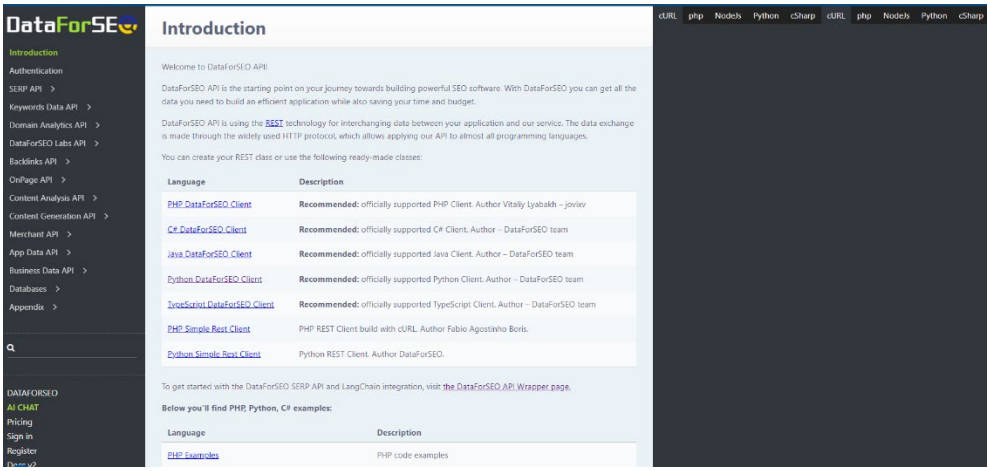
Key Features of DataForSEO

- **Wide Range of APIs**: Provides access to multiple APIs for different data collection needs.
- **Real-Time Data**: Offers real-time data collection for up-to-date insights.
- **Easy Integration**: Designed for easy integration with various programming languages, including Python.
- **Comprehensive Documentation**: Detailed documentation to help users understand and implement the APIs effectively.

2. DataForSEO Documentation

Overview of the Documentation

The DataForSEO documentation provides extensive information on how to use their APIs. The documentation is structured to guide users through various stages of implementation, from authentication to using specific endpoints. Here are the key sections of the documentation:



1. **Introduction**
2. **Authentication**
3. **API Endpoints** (SERP API, Keywords Data API, Domain Analytics API, etc.)
4. **Libraries and Examples**

Accessing the Documentation

You can access the DataForSEO documentation by visiting the DataForSEO Documentation.

Key Sections Explained

Introduction

The introduction provides a general overview of the DataForSEO API, including its purpose and the benefits of using it for data collection and analysis.

Authentication

Authentication is a crucial step to ensure secure access to the DataForSEO APIs. The documentation provides detailed instructions on how to authenticate your requests using API keys. You will need to include your API key in the headers of your HTTP requests.

API Endpoints

DataForSEO offers several APIs, each with specific endpoints. Here are some of the main APIs you will find in the documentation:

- **SERP API**: For accessing search engine results data.
- **Keywords Data API**: For obtaining keyword-related data.
- **Domain Analytics API**: For analyzing domain data.
- **DataForSEO Labs API**: For experimental and lab-based data analysis tools.

- **Backlinks API**: For retrieving backlink data.
- **OnPage API**: For on-page SEO data.
- **Content Analysis API**: For content analysis metrics.
- **Content Generation API**: For generating SEO-friendly content.
- **Merchant API**: For merchant data and analytics.
- **App Data API**: For mobile app data.
- **Business Data API**: For business-related data.
- **Databases**: For accessing various database resources.

Each API has specific endpoints, request parameters, and response formats detailed in the documentation.

Libraries and Examples

DataForSEO provides client libraries in various programming languages to help you get started quickly. Here are some recommended clients:

- **PHP DataForSEO Client**
- **C# DataForSEO Client**
- **Java DataForSEO Client**
- **Python DataForSEO Client**
- **TypeScript DataForSEO Client**

Additionally, the documentation includes code examples to demonstrate how to use these clients effectively. Go on 'Python DataForSEO Client' for more examples.

Practical Exercise Instructions

Objective

In this exercise, you will learn how to use the DataForSEO API to collect data from the Google Maps API, process the data, and save the output as a raw JSON file. You will be provided with a Python script and a flowchart to help you understand the code flow.

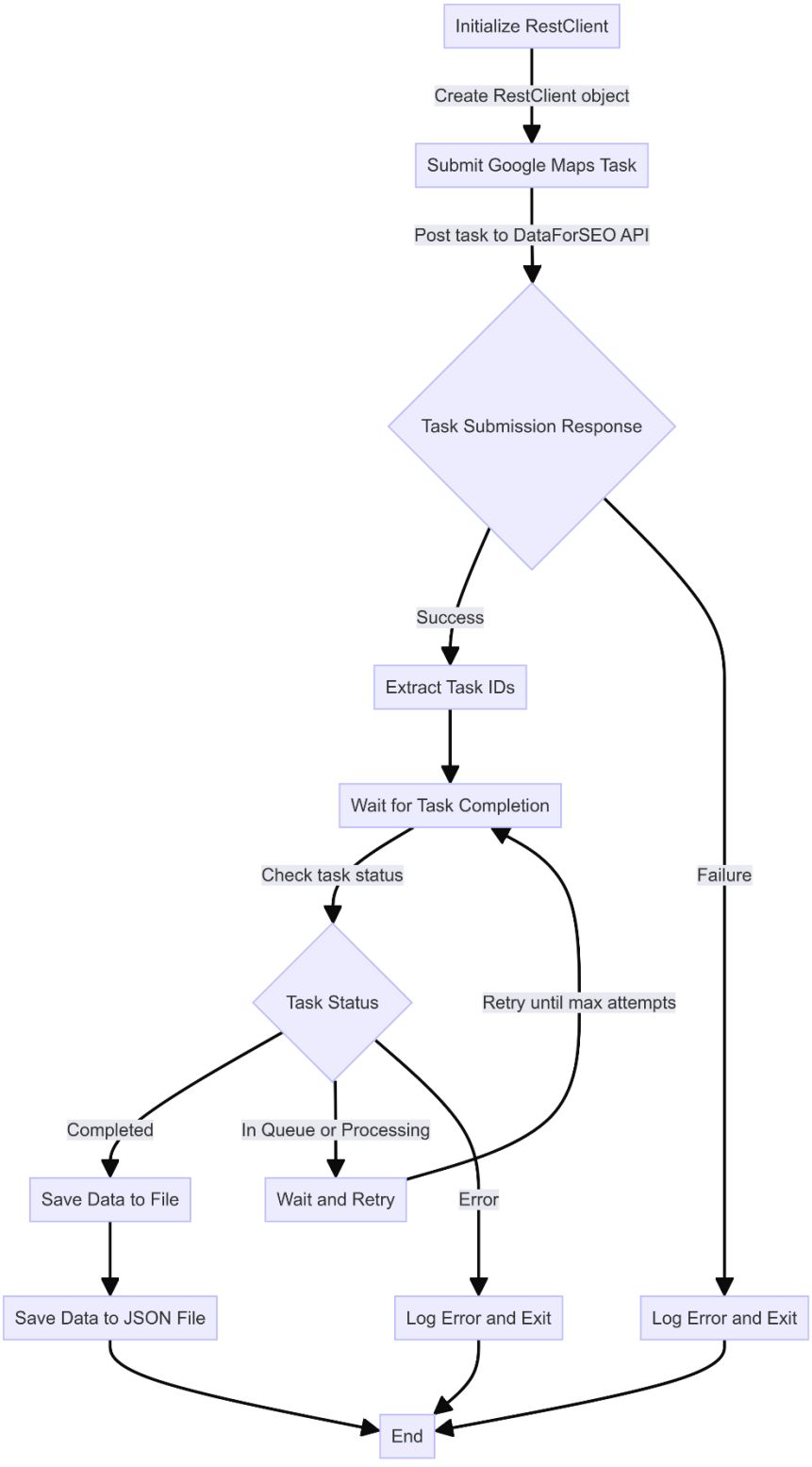
Steps

1. **Download the Python Script**
    - You will be provided with a Python script file named ***GoogleMapsEndPoint_raw.py*** .
    - Download and save the file to your local machine.
2. **Examine the Python Script**
    - Open the data_collection.py file in your preferred code editor (e.g., VSCode, PyCharm, or any text editor).

- Carefully read through the script to understand how it works. Pay attention to the following sections:
  - Initialization of the RestClient class.
  - Creation of a RestClient object.
  - Submission of a Google Maps task.
  - Handling the task submission response.
  - Checking the task status and handling retries.
  - Saving the data to a JSON file.

3. **Match the Code with the Flowchart**
   - Refer to the provided flowchart (attached below) which illustrates the flow of the script.
   - Match each step in the flowchart to the corresponding code in the ***GoogleMapsEndPoint_raw.py*** file.
   - This exercise will help you understand the sequence of operations and the decision-making process in the script.

Flowchart


Initialize RestClient
Create RestClient object
Submit Google Maps Task
Post task to DataForSEO API
Task Submission Response
Success
Failure
Extract Task IDs
Wait for Task Completion
Check task status
Task Status
Retry until max attempts
Completed
In Queue or Processing
Error
Save Data to File
Wait and Retry
Save Data to JSON File
Log Error and Exit
Log Error and Exit
End