**Project Scope Document**

Prime Product Analyzer

Project Description  
This project aims to develop a Python-based data pipeline that accesses real-time ecommerce data from an Amazon API available on RapidAPI.com. The pipeline will extract data related to product search optimization and other relevant ecommerce metrics, process this data for analysis, and provide actionable insights that can inform product positioning and sales strategies on Amazon.

Project Objectives

* Data Acquisition: Efficiently pull real-time data using the Amazon API to capture current trends and activities.
* Data Processing: Implement robust data transformation techniques to convert raw data into a usable format for analysis.
* Data Analysis: Analyze processes data to identify key trends, patterns, and insights related to product performance on Amazon.
* Automation: Automate the data fetching and processing tasks to ensure the pipeline operates continuously without manual intervention.
* Scalability: Design the pipeline to handle increased data volumes as API usage and data fetching activities scale up.

Deliverables

1. Data Extraction Module
2. Data Transformation Module
3. Data Loading Module
4. Documentation
5. Testing Suite

Technical Requirements

Programming Language: Python

Libraries and Frameworks:

* Requests or Httpx for API interactions.
* Pandas and NumPy for data manipulation
* PySpark for handling larger data sets if necessary

APIs:

* Amazon real-time data API from RapidAPI.com

Data Storage:

* Use of SQL or NoSQL databases, or flat files for the processed data.