**Project Plan: E-Commerce Analytics and Product Monitoring System**

Introduction

This document is a detailed Project Plan for fyp23048. The Project Plan will comprise a Project Background, a Project Objective, a Project Method, and a Project Schedule. The Project Plan aims to provide a clear view of the reasons the topic was chosen, what the project consists of, how the project will be implemented, and the different milestones of the project.

Project Background

E-commerce refers to buying and selling goods and services on the Internet. With its worldwide reach, convenience, and efficiency, E-commerce has gained huge significance over the past decade. For this project, we will limit the term "E-commerce" to third-party sellers on online shopping platforms. Third-party sellers have become increasingly popular in recent years due to the low difficulty of generating profit. With the established infrastructure of online shoppers, sellers gain an immediate customer base. They can reduce marketing costs while scaling up quickly, thus gaining popularity.

To maximize profit margins, third-party sellers require a ginormous amount of accurate and updated data to keep track of product performances, as well as gain insights into other competitor products. Since most third-party sellers are small businesses, it is often unlikely for them to design, create, and maintain software and programs to suit their needs. Currently, they can only do that by either manually tracing and collecting data, or subscribing to paid services to collect data. However, these subscription services provide inaccurate data due to the use of approximation instead of directly extracting data. In addition, they do not implement features for users to save wanted products and automate the tracking process. Therefore, it is believed that by creating a project that 1) provides accurate, real-time data, 2) automates the product tracking process, and 3) provides useful analyzing tools, users can increase efficiency and have a competitive advantage over other competitors in the market.

There are dozens of E-commerce solutions in the market that target third-party sellers. The most popular software is Helium 10. Helium 10 is the leading all-in-one software platform trusted by over 2,000,000 e-commerce entrepreneurs, agencies, and mainstream global brands to power their business growth on Amazon by delivering accurate, data-driven solutions.[1] Unfortunately, except for some fields that are directly accessible through Amazon's API, all other data that users pay to look for are only approximates calculated by AI and Machine Learning models. [2]  The proposed project is planned to fill such a gap by providing actual data instead of approximated data. It is expected to be designed as an all-around support software that covers all aspects that require data instead of a search engine for specific data, which is what most current software is offering.

Project Objectives

In the current E-commerce ecosystem, third-party sellers face various challenges in managing their business. They lack efficient tools and real-time data to make decisions effectively. Current market subscription services can only partially aid them, but the incomplete tools and design further complicate the user experience. Third-party sellers will greatly benefit from a new ecosystem that boosts data accuracy and working efficiency.

The objective of the project is as follows:

1. One-stop E-commerce Solution

To develop a holistic solution that provides comprehensive coverage for third-party sellers to enhance decision-making processes and competitiveness in the marketplace.

1. Integrated Analytical Tools

To develop analytical tools tailored to suit the needs of third-party sellers to present data in a useful and meaningful way to speed up the analysis and decision-making process.

1. Real-time Data Extraction

To develop programs and functions to ensure that data extracted are always exact and retrieved in real-time. This integration ensures that insights from collected data are always accurate.

1. Automated Product Tracking

To implement an automated system that allows users to save, monitor, and track products and collect desired data in the most efficient way. This can eliminate the need for manpower, thus boosting overall working efficiency.

1. Scalability and Evolution

To build a platform that can respond to market and structural changes. New features and updates can be implemented with only minor changes to the entire platform.

1. User-Centric Design

To design user interfaces that is intuitive and user-friendly.

The benefits brought by the proposed project are as follows:

1. Holistic Business Management

Third-party sellers can conduct daily operations on one single platform that provides all essential tools.

1. Accurate Data

Real-time data extraction guarantees up-to-date and accurate information, thus providing accurate insights to third-party sellers.

1. Optimized Decisions

With tailored analytical tools and accurate data, better decisions can be made. This leads to better business strategies, and ultimately higher profitability.

1. Enhanced Efficiency

The one-stop solution and automation of product tracking eliminates the use of manpower and improves productivity.

1. Improved User Experience

A user-centric design allows users of all technological backgrounds to fully utilize the potential of the platform.

Project Methodology

The proposed project is to develop a web-based application using Retool, a user-friendly platform that allows users to build effective business software. [3] This application will include three principal functionalities:

Search Page: A dedicated function that enables users to search specific product attributes instantly.

Automated Worker: A system designed to autonomously monitor selected products and capture relevant data.

Analytical Page: A dedicated space equipped with tools to visualize and interpret the collected data.

The project will focus on products listed on Amazon, the world’s leading multinational technology company focusing on e-commerce. [4]  To obtain the necessary data, web scraping techniques will be employed Octoparse, a software for automatic data extraction [5], and Scrapy, a free and open-source web-crawling framework written in Python. [6] After data retrieval, it will be systematically stored in a PostgreSQL database, a powerful, open-source object-relational database system. [7] Integration with PostgreSQL is seamless as the database gets instantiated automatically upon setting up a Retool account.

The feasibility of the project is as follows:

1. Technical feasibility

Retool is a proven platform for building web-based applications. It also offers easy integration with PostgreSQL. Octoparse and Scrapy are both known for web scraping. They are both compatible with scraping Amazon sites. PostgreSQL is a reliable open-source database system that has been available for decades.

1. Economic feasibility

Other than Retool and Octoparse, all software is free. For Octoparse, the free version has all features that the project requires. For Retool, depending on the method the application will be exported, a paid license might be needed. This can be compensated by the Virtual Fund of the FYP.

1. Operational feasibility

The main features of the application are designed with the user in mind. It is hoped to be user-friendly and can maximize efficiency. Thus, the usability of the project is fulfilled. As for the backend, with the automated linkage of the PostgreSQL database to Retool, it is expected that connection problems will be eliminated. For web scraping, it is expected that codes and programs will be adjusted whenever the page structure of Amazon sites changes.

1. Legal feasibility

The web scraping process will complete adhere to the terms of service and the robot.txt file of Amazon. For data storage, no personal identifiable information will be collected or stored without explicit consent.

1. Time feasibility

It is suggested that the web-based application can be developed and launched within six months without significant unforeseen challenges. Schedules and key milestones will be set up to ensure that the progress is keeping up with the timeline set.

Project Schedule and Milestones

The projected timeline is as follows:

October: Research and Planning

Research will be conducted to define exact implementation methods of all functions. Documents will be created to document findings and how different tools will be used in detail so that findings can be used in later stages.

November: Tool Setup and Create Web Scraping functions

All necessary tools will be downloaded or set up. Web Scraping functions that will be used in “Search Page” and “Automated Worker” will be created.

December: Database Setup and UI Design

Class diagrams for showing the relationship between tables will be created. The database will be set up accordingly with demo data. UI Design for all three pages will be drafted.

January: Develop “Search Page” and “Automated Worker”

The frontend UI will be created for both pages. Both backend and frontend logics will then be created and linked to corresponding elements of the pages.

February: Develop “Analytical Page”

The same process done in January will be repeated. The “Analytical Page” is expected to be more sophisticated, thus will require more time to implement.

March: Testing and make improvements

The Alpha version shall be developed and tested by this stage. Improvements will be made to ensure consistency, smoothness, and aesthetics. The Beta version and Final version will be developed and published by the end of the month.

April: Deployment

The application will be deployed. Post-project review will be conducted to gather learnings for future uses.

The projected project milestones are as follows:

October: Documentation of research findings completed

November: Octoparse and Scrapy implemented, Retool account set up

December: Class Diagrams and UI designs developed, PostgreSQL database set up

January: “Search Page” and “Automated Worker” pages completed

February: “Analytical Page” completed

March: Alpha and Beta versions completed, Final version under development

April: Final version completed and published

Conclusion

In this comprehensive project plan, various areas are discussed. It is believed that the E-Commerce Analytics and Product Monitoring System will be a great supplement to third-party sellers on online shopping platforms. The system aims to address the need for accurate, real-time data and user-friendly features to aid with users' working efficiency. It is hoped that by following the proposed timeline, this project will be completed successfully with a positive impact on the E-commerce ecosystem.

References

[1] Www.g2.com. [Online]. Available: https://www.g2.com/products/helium-10/reviews. [Accessed: 01-Oct-2023].

[2] Helium10.com. [Online]. Available: https://kb.helium10.com/hc/en-us/articles/360007907533-How-Does-Helium-10-Get-Its-Data-. [Accessed: 01-Oct-2023].

[3] “Retool,” Retool.com. [Online]. Available: https://retool.com/. [Accessed: 01-Oct-2023].

[4] Wikipedia contributors, “Amazon (company),” Wikipedia, The Free Encyclopedia, 30-Sep-2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Amazon\_(company)&oldid=1177975754.

[5] Getapp.com. [Online]. Available: https://www.getapp.com/all-software/a/octoparse/. [Accessed: 01-Oct-2023].

[6] Wikipedia contributors, “Scrapy,” Wikipedia, The Free Encyclopedia, 27-Sep-2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Scrapy&oldid=1177348947.

[7] “PostgreSQL,” PostgreSQL, 01-Oct-2023. [Online]. Available: https://www.postgresql.org/. [Accessed: 01-Oct-2023].