Project Submission – Phase 1

Subject: CS6.302 Software Systems Development

Project 6: Data Visualization – ETL

1. Team Details:

a. Project Name: Indian Import and Export Data Visitation Portal
b. Team members:
   i. Aashrey Jain (Roll number: 2024202012)
   ii. Ashima Mathur (Roll number: 2024201073)
   iii. Deepesh Vendoti (Roll number: 2024204013)
   iv. Kanika Aapan (Roll number: 2024201021)
   v. Khooshi Popat (Roll number: 2024201069)

2. Scope:

Through this project, we aim to provide a dashboard that will help the user in visualizing India's import and export data between the years 2017 and 2023. We want to provide the option for visualizing the cumulative trends over this period as well as year-wise data.

Some of the visual representations that we want to provide to the user are:

a. A line graph showing the absolute value of total imports and exports for each year across all countries and commodities.
b. A doughnut chart showing the proportion of imports and exports by commodity across countries. This will tell us about the commodities that are most/least imported/exported. We can have an option to visualize the data by value or by quantity.
c. A doughnut chart showing the proportion of imports and exports by country across commodities. This will tell us about the countries that trade the most/least with India. Similar to the previous chart, we can have an option to visualize the data by value or by quantity.
d. A chart to show the commodities that contribute the highest value per unit quantity to India's imports and exports.

We also want to provide an option for the user to see the raw data based on the filters applied by them, on a separate screen.

We are not considering building an authentication system on top of the visualization framework.

The dataset that we have chosen can be found under the following catalogues of the Government of India datasets:

a. Import data: https://www.data.gov.in/catalog/principal-commodity-wise-import
b. Export data: https://www.data.gov.in/catalog/principal-commodity-wise-export

Due to the Covid-19 pandemic, the data is not available for the period 2020-21 so we will skip that year. The datasets contain the import and export data (value in Million USD and quantity in Kgs) for each commodity and country that trades with India.

3. System Design

The project requires full stack development to provide seamless visualization of the underlying data. The underlying components are:

a. Front End: The front-end of the application will be built using HTML, CSS, and ReactJS, along with external JavaScript libraries for implementing supporting functionalities. It will consume the REST API implemented in the backend to fetch the data from the database and then display it pictorially using ChartJS.

b. Back End: The back-end portion of the application will be written in ExpressJS running on Node.js. It will be an API server that fetches data from the database tables and returns it to the front-end in JSON format. All queries to the database will be handled by the backend.
c. Database: We will need a SQL database to store and query our data. Since we need to perform complex queries like GROUP BY and AGGREGATE, we will prefer to use a SQL database like MySQL. It will act as the reporting database for our application.

Proposed Tech Stack:

a. Front End: HTML, CSS, ReactJS with other JavaScript libraries.
b. Back End: ExpressJS server running on Node.js in a Linux environment.
c. Database: MySQL.

4. Stakeholders

These are the stakeholders for this project:

a. Team members: Aashrey, Ashima, Deepesh, Kanika and Khooshi.
b. Project supervisor: Dr. Sai Anirudh Karre
c. Data source: https://www.data.gov.in

The following would be the beneficiaries of a visualization portal like this one:

a. Importers and Exporters
b. Manufacturers and Industries
c. Supply Chain Industries and Analysts
d. Market, Media and Trade Analysts
e. Researchers and Academicians
f. Policy & Research institutions

5. Use Cases

The primary objective of the portal is to provide a data visualisation portal to the beneficiaries of the Import/Export data. It is intended to be a one-stop solution so that stakeholders can access ongoing trends in the data in a clean and concise manner, while gaining insights that are useful to them. The following specific cases can be considered:

a. Importers of goods can use the charts to identify market opportunities, assess risks, optimize supply chains, and make data-driven strategic decisions.
b. Exporters of good can identify opportunities, forecast demand, optimize pricing, manage risks, optimize supply chains, and make informed business decisions.
c. By analysing data, market and trade analysts can identify market trends, assess trade policies, analyse supply chains, and support economic analysis.
d. By making decisions based on data, manufacturers and industries can forecast raw material availability and explore new market opportunities for their products.
e. This tool would also help traders in identifying potential trade barriers and supply chain disruptions.
f. By analysing data, researchers can identify market trends, assess trade policies, and support economic research.
g. The data visualization tool can be used to suggest policy-level changes depending upon the trends and insights.
h. Policy research institutions can leverage import-export data visualization tools to conduct in-depth analysis of global trade patterns, inform policy development, and evaluate the effectiveness of trade policies.
i. Users can analyse regional integration of businesses, promote sustainable development, and address social and development issues.
j. Supply chain managers and stakeholders can use the data visualization on our portal for capacity-planning and optimization of shipping routes and containers.

6. Delivery Timeline

We want to deliver the MVP version of the application by the Phase 2 evaluation. The final version of the project would be ready in time for the final evaluation in November/December. The detailed timeline is:

a. By end of September: Ready the designs for the UI/UX and finalize the data pipeline and table structure. Also finalize the REST API structure for the backend.
b. By first week of October: Coding begins. The frontend developers start designing the components and implementing the UI. The data is loaded to the database tables and ready for querying.
c. By second week of October: The MVP version of the project is ready. The backend is ready to accept basic queries from the frontend. Stored procedures are written to the database.
d. By end of October: The database is completely ready for the final version, along with the stored procedures. The backend is completely ready to interact with the database and all endpoints work as expected.
e. By second week of November: The final version of the project is ready with all intended functionalities.
f. Till evaluation: Testing of the project for all use cases and implementation of any good-to-have features.

We aim to deliver the following as a part of the MVP:

a. Absolute visualization of the import data over the years in the form of a line graph.
b. Tabular visualization of the raw import data with application of filters obtained from the user.

7. Work Distribution

We have identified the following roles that will need to be fulfilled by the team for the successful delivery of the project within the set deadlines:

a. Front End: Two team members will be involved in the creation of the UI/UX that will form the front-end of the application.
b. Back End: One member will be responsible for developing the back-end of the project using ExpressJS and Node.js
c. Database: One member will be responsible for creating and managing the data pipeline that will read the dataset, store it in tables and respond to the queries generated by the backend.
d. UI Design and Documentation: One member will be responsible for designing the UI/UX of the dashboards. This would involve creation of design documents and UI screens that will be implemented by the front-end developers. This role also involves writing end user-facing documentation.

We are yet to distribute these roles among our team.