

Title

São Paulo Navigator: A Data-Driven Journey Through Ratings and Routes

Project Summary

In this project, we will build a public transportation information web tool that allows users to view and plan public transport routes, calculate fares, add and view reviews for each route, and see historical trends in transportation data over the past decades. We will be using public transit data from the TA recommended Sao Paulo, Brazil public transport dataset. We will be using languages and frameworks like SQL, React, Flask, Django and more to design our web app. Ultimately, our goal is to aggregate and present public transport information in an accessible and user friendly manner.

1. Describe what data is stored in the database. (Where is the data from, and what attributes and information would be stored?)

This dataset stores comprehensive information about the public transformation system of Sao Paulo, Brazil.

The data directly comes from the Sao Paulo transit system and is provided in a GTFS format, which is great for the application we are about to create.

The dataset contains many different CSV files, starting with a routes table, fare_rules, fare_attributes, stops, trips, stop_times, and many more.

With our project in mind, our rough draft of the schema would consist of the following:

Search Functionality: For this feature, we would combine the Calendar table provided as well as the routes table, shapes, trips csv files. These tables utilize multiple attributes for example:

(shape_id,"shape_pt_lat","shape_pt_lon","shape_pt_sequence","shape_dist_traveled"). This tuple comes from the shapes csv file and could be used for help with mapping and calculation. For the search functionality, we would also include the trips and stops tables in order to choose when we want to leave depending on a particular stop.

For Fare Calculation, we would have two tables, Fare_Rules and Fare_attributes. Fare_rules would consist of 5 attributes: fare_id, route_id, origin_id, destination_id, contains_id. These attributes would help us associate the fare depending on location, destination, and commonality. Fare_attributes would be another table that consists of, fare_id, price, currency_type, payment_method, transfers, transfer_duration. Here we could create a relation between the two different tables with regards to fare_id, and if that table has that specific fare_id, it would have a certain price tag on it.

For Reviews and Accessibility information, our plan is to create a new table, one that can be updated by the user. This table would be called the Reviews table and it would consist of the user_review, like, dislike, comment, user_id, etc. This table is not associated with the data set, but it is a new feature that could help improve any issues that can occur.

2. What are the basic functions of your web application? (What can users of this website do? Which simple and complex features are there?)

Our web application has four main functionalities and might include additional components:

Functionality -

1. Search Functionality - Route Planner and visualizing the route
 - a. Complex Features -
 - i. Visualizing the route on the map
 - ii. Integrating other functionalities onto the planner page
 - b. Simple Features -
 - i. Finding the route between user inputted source and destination
2. Fare Calculator- Allow use to input up to n destinations and get fare (Yahav)
 - a. Complex features -
 - i. Extend route planning to find the least costly route, and recommend it to the user.
 - b. Simple Features -
 - i. Calculation of the fare
3. Reviews and Accessibility information - Add extra information for specific routes (PJ)
 - a. Complex Features -
 - i. Validating and Process Reviews
 - ii. Updating the database

4. Historical Data Analysis - Find trends in historical data for routes (Danish)

3. What would be a good creative component (function) that can improve the functionality of your application? (What is something cool that you want to include? How are you planning to achieve it?)

We believe that our main features are very robust and extensive already. However, one creative component that can improve the functionality of our application is definitely the use of better front-end animations. A good user experience is always key to have a successful application. If a user looks at a website and observes a poor design only made with HTML, that specific user will not enjoy using that website. We want to achieve this by the end of our project, since this isn't a very important part of functionality.

4. **Description** of an application of your choice. State as clearly as possible what you want to do. What problem do you want to solve, etc.?

Our goal is to build a web application that aggregates and presents transit information in a very accessible and user friendly manner. We want to eliminate the use of multiple sources to plan a trip (one for fares, one for reviews, one for visualization etc.). Having an all-in-one app saves the user time and effort. Our map and route visualization will further enhance the accessibility of the application. The component of analyzing routes and fare trends based on historic data adds a component of data-driven for the user to understand how the routes they are using are changing on yearly/monthly/weekly basis. Furthermore, large, high quality and robust data ensures an error-free and reliable application.

5. **Usefulness.** Explain as clearly as possible why your chosen application is useful. Make sure to answer the following questions: Are there any similar websites/applications out there? If so, what are they, and how is yours different?

This application is useful for the common public transport user in Sao Paulo. It provides route planning capabilities that are backed by historical data and fare calculations. In general, it enables users to receive more accurate information, something that apps such as Google Maps may not do as well.

There are similar applications like São Paulo Metro App/System. Our application differentiates itself by adding a component of historical trends and collecting reviews from users.

1. Trends - Allows the user to see how a route/train/place behaves at certain points of time visually thereby adding his/her planning.
2. Add and view reviews - The reviews exist to add credibility to the routes and transit options present in the application

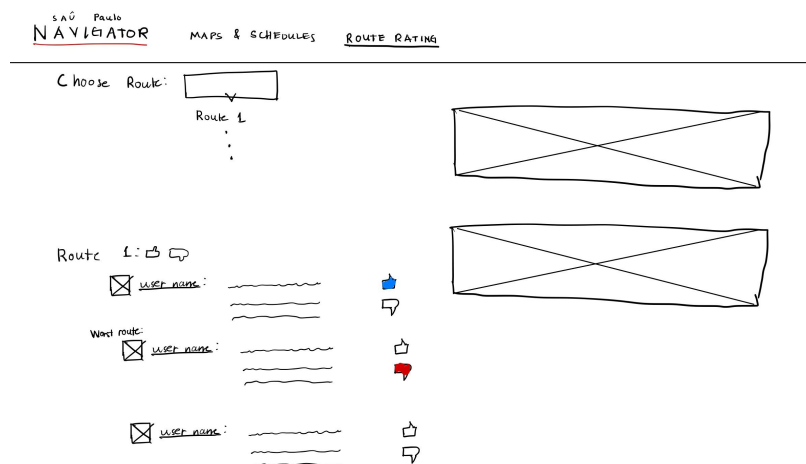
6. Realness. Describe what your data is and where you will get it.

- i. Our data is available on Kaggle
(<https://www.kaggle.com/datasets/mateuscco/sao-paulo-transportation-service>). It has a usability score 8.3/10 indicating that is a fairly complete and credible dataset
- ii. The dataset consists of 10 sub-datasets - fares, routes, trips, frequencies, calendar etc. We will be attempting to use all or a mix of these dataset to achieve our goal.

7. Description of the **functionality that your website offers.** This is where you talk about what the website delivers. Talk about how a user would interact with the application (i.e., things that one could create, delete, update, or search for). Read the requirements for stage 4 to see what other functionalities you want to provide to the users. You should include:

1. **A low-fidelity UI mockup:** What do you imagine your final application's interface might look like? A PowerPoint slide or a pencil sketch on a piece of paper works!

Route Rating Website UI mockup:



SÃO Paulo
NAVIGATOR MAPS & SCHEDULES

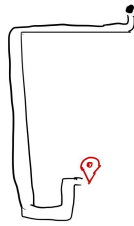
Trip Planner

Efficient Routes:

- 1
- 2

Lowest Fare

\$ \$ \$



SÃO Paulo
NAVIGATOR MAPS & SCHEDULES

Trip Planner

From
São Paulo

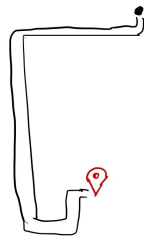
To
X X X

When
• •

Trip Options

- Least time
- Fewest Transfers
- Least Price

Navigator



2. Project work distribution: Who would be responsible for each of the tasks or subtasks?

List of the person responsible for which exact functionalities in section 6. Explain how backend systems will be distributed across members. Be as specific as possible as this could be part of the final peer evaluation metrics.

Search Functionality - Route Planner and visualizing the route: Khushal

Fare Calculator: Yahav

Reviews and Accessibility information: PJ

Historical Data Analysis - Find trends in historical data for routes: Danish