PROJECT PROPOSAL

1. The database contains data from Sao Paulo's metro, rail, and bus routes available up to the current date from their web portal https://www.sptrans.com.br/. The data of interest consists of 10 files; agency, calendar, fare_attributes, fare_rules, frequencies, routes, shapes, stop_times, stops, trips.

   [Primay Keys Underlined, Foreign Keys Color Coded: # = Referencing Field]

   Agency - This table only contains 1 tuple with the agency_id and related data

   Calendar - contains days of the week and start/end dates for services by service_id.
   service_id,
   monday,
   tuesday,
   wednesday,
   thursday,
   friday,
   saturday,
   sunday,
   start_date,
   end_date

   Fare_attributes contains data on the various available fare types and their prices
   fare_id,
   price,
   currency_type,
   payment_method,
   transfers,
   transfer_duration

   Fare_rules relates fares with the routes they correspond to.
   fare_id#,
   route_id#,
   origin_id,
   destination_id,
   contains_id

Frequencies contains data on the interval between departures at different points in the day specific to a given “trip”.

trip_id#,
start_time,
end_time,
headway_secs

Routes gives information about routes, such as name, description, type and identification color.

route_id,
agency_id,
route_short_name,
route_long_name,
route_type,
route_color,
route_text_color

Shapes models trip paths using a series of geo-location points.

shape_id,
shape_pt_lat,
shape_pt_lon,
shape_pt_sequence,
shape_dist_traveled

Stop_times specifies the planned arrival and departure times for each stop of an itinerary and what order they are visited.

trip_id#,
arrival_time,
departure_time,
stop_id#,
stop_sequence

Stops has the information of each stop, including an id, name, and coordinates.

stop_id,
stop_name,
stop_desc,
stop_lat,
stop_lon

Trips contains data for routes, each of which can contain multiple “trips”. The data includes trip names and directions in relation to the routes and the service id/s that provide each route.

```
route_id#,
service_id#,
trip_id,
trip_headsign,
direction_id,
shape_id#
```

2. The basic functions of our web application are to aid users as they plan to explore the city of Sao Paulo. To start, the users will be able to create their own logins and set their own passwords in order to access, save, modify, and delete their trip plans. With the database in the backend of the website, users will be able to input data to plan simple trips with their desired starting and ending destination, the backend will query the database find relevant information, and then connect back to the user to display the results.

3. We hope that our application will be able to help users with preliminary, simple trip planning in Sao Paulo because most visitors will not choose to rent a car. We will try to address these Through the database and a user interface, we hope to be able to have users be able to input their starting location and their desired location to see possible bus routes that are available.

4. One in a Brazilian Trip

We are envisioning a similar, yet humble, version of old-school Google Maps or MapQuest that would present the user with three input boxes: one for starting location, one for ending location, and one for the desired leave time. From these input boxes, we will convert the user entered data into a database query to return any and all information matching that data. If no relevant route exists, we will return text that says there are none available. Since our dataset has built in coordinate location, we would like to try to be able to have the user find the closest stop to them.

In addition, through queries and data analysis, we want to learn which roads and routes are the busiest, so we can alert users that these routes may experience higher volumes of traffic and the next available bus at their stop. We are going to try to have these values as a page view on our website, so the users can refer to these values as needed.

5. Our project uses the Sao Paulo transportation dataset. This dataset contains information about the bus company, days of operations, and fare details. It also contains real data on the planned bus route, interval of departure, stop times, and more. Our group thought we could use this information to create an application that could help navigate the city.

Our goal is to create a website that can help users traverse the Brazilian city. The website will contain information regarding the closest stop to the user, alongside the next available bus. Additionally, the site will allow the user to log in to access other features. For instance, they can save trip ideas or create a new path for themselves. This will allow people to get around the city in a way that's best for them. We also plan on providing badges when users achieve milestones, such as visiting a certain number of stops or attending different lines. The app also won't include alternative modes of transportation, such as driving. These latter features encourage users to take the bus instead of other modes of transportation, which could indirectly benefit the environment. Thus, this app is a useful and fun tool for navigating Sao Paulo in a more environmentally friendly way.

6. Public transportation is vital to communities. It reduces the community's carbon footprint, provides low cost and accessible travel, reduces the cars on the road to free up urban spaces and reduce parking lots, and much more. However, unless you are familiar with an area or know where to look, it can be overwhelming and difficult to plan trips. A population that would benefit from this would be tourists! They are not in town long enough to be intimately comfortable with the timetables and resources, so we are hoping to bridge that gap with our application Braz1lian Trip (A One in a Brazilian Trip).

Using the public transportation dataset, we hope to convert important information on schedules, stops, and frequency into a database and connect a user interface to help users plan routes in order to make the most out of their trip to Sao Paulo, Brazil. We are imagining an application that allows users to create accounts and through inputting data in fields, we will compile their entered data to a query, and help them plan their routes to visit their desired areas of the city.

7. Our project is useful because we are connecting users to information on public transportation options in Sao Paulo. It is important to promote public transportation because it allows for accessible travel. Environmentally, it makes a considerable difference because it reduces the amount of cars on the road and thus the pollution from gasoline. Another benefit of less cars on the road is more space efficient with the use of roads and rails. It allows for safer pedestrian travel and other second order effects. Public transportation is also scalable and technologically-driven to be more efficient and safer for users.

Our product is very similar to Google Maps, MapQuest, and other map software because we are using a database about public transportation. In these instances, the user is able to input information and the database in the background is running relevant queries and parsing information to output. Where ours is different is that we allow users to create their own accounts to save and preserve data that they can access later. It is also a small niche website that is not trying to promote the use of cars or benefiting the automotive industry and unlike the other applications, our application will not default to car travel.

8. Our project uses the Public Transportation Service database, which is found on Kaggle. This data pertains to the transportation system of Sao Paulo, a major city in Brazil. Many of the tables contain different forms of company identification or representation. The three files that perform this task are the agency, far_rules, fare_attributes, and calendar files. The agency file is used to model the agency that manages the vehicles. The fair_rules file is used to state which route each type of bus is employed in, such as CPTM in route “CPTM L10”. fare_attributes compliments this data well by providing information on cost for each type of vehicle, alongside what currency is accepted. Lastly, the calendar file displays the days in which the service is run, alongside the period of operation. If the company works on that day, a 1 is placed under that attribute.

The dataset also contains information about the routes themselves. First, the routes file identifies and describes each bus route, with each tuple having a name and other important descriptions. Next, The trips file describes the direction and shape of a route that a bus may have taken multiple times. The frequencies dataset compliments trips by providing the times at which these trips start and finish. Lastly, the shapes, stop_times, and stops files share details about the route itself. Shapes define the trip path of each bus using ordered points as markers. Stops gives the name and location of each area where the bus stops, and stop_times provides the arrival, departure, and order of these stops.

9. Nate- 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 stages 4 and 5 to see what other functionalities you want to provide to the users. You should include:
CRUD (Create, Read, Update, Delete)
The website will allow users to interact with the Sau Paulo Public Transportation Service database to retrieve information necessary to navigate the system including what routes are nearby, when service is available, and how much fare they can expect to pay. Users will be able to create an account to allow them to save information such as favorite routes or trips, and to “plan a trip” by adding various stops. The application can then advise the user which routes to take to each stop, approximately how long the itinerary will take, and show the total fare cost for the entire planned trip. Users will also be able to use a search function to locate a stop they would like to visit, find out which routes go to that stop, and how much the fare will be.

A potential added functionality if we are able to access user geolocation data may allow the application to suggest the closest stop for the route the user is seeking. There is also potential

to “gamify” the experience by tracking which stops a user has navigated to and awarding “badges” for meeting certain criteria such as visiting x number of stops or traversing each “color” of route.

10. A UI-mockup can be found in the Sao_Paulo_UI_Mockup pdf.

11. **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.

- ☐ Angela - Base Website 6/26
    - ☐ Paths
    - ☐ Views
    - ☐ URLs
- ☐ Nate - User Creation/Management 6/30
    - ☐ Create an Account
    - ☐ Login/Password
    - ☐ Ankit - Forgot Password (CRUD)
    - ☐ Encrypted Store in Database
- ☐ Shon- Relational Design 7/07
    - ☐ Primary Keys
    - ☐ Foreign Keys
    - ☐ Datatypes
- ☐ Ankit - Creating Database
    - ☐ Commands to create Tables
    - ☐ Inserting data from dataset
    - ☐ Creating fake user data
- ☐ Ankit - Start Location/End Location User Input
- ☐ Badges (Triggers)
    - ☐ Angela- Medals for Itinerary Notes
    - ☐ Nate- Rainbow Badge
    - ☐ Shon - TBA
    - ☐ Ankit - TBA
- ☐ User Itinerary
    - ☐ User “Home” page of their notes
    - ☐ Itinerary Database