Planning Document: Translink Data Parser

Dataset and Task

The dataset consists of two major components: static data and live data.

The static data includes several key files: `routes.txt` , which contains information on all the routes operated by Translink; `stops.txt` which provides details about all the bus stops in Queensland; `trips.txt` which assigns a unique trip ID to each trip, making it easier to identify specific trips; `stop_times.txt` which records every stop made by any bus route; and `calendar.txt` , which provides the necessary data to validate bus schedules within specific time ranges.

For the live data, there are two useful JSON files that are cached in real time. The first is `trip_updates.json` , which contains detailed information about the bus schedules, including departure times, arrival times, delays, and vehicle IDs. The second is `vehicle_positons.json` ,which provides real-time information on the locations of the vehicles.

Task: Build a fully-functional TransLink data parser that provides a summary of all Route stops.

Important Fields and Joins

1. `route_id` This key field is used to join the `routes.txt` and `trips.txt` datasets, allowing us to filter `route_id` based on specific routes.

2. `trip_id` This key is crucial for joining `trips.txt` and `stop_times.txt` and allow us to filter all the `stop_id`

3. `stop_id` This field is used to join `stop_times.txt` with `stops.txt` to get detailed information about each stop.

Steps to Perform the Task

1. Join `routes.txt` and `trips.txt` on `route_ id` and join `stop_times.txt` and

`stops.txt` on `stop_id`.

2. Wait for the user to prompt the detail of the trip information.

3. Use the `calendar.txt` to validate the availability of the bus route.

4. Use `service_id`, `trip_id` from the static data, along with user-provided details, to retrieve real-time information from the cached live data.

5. Display the result using `console.table()`.

Reference

1. OpenAI. (2024, August 29). Code for converting POSIX to formatted date-time, formatted time to POSIX, and class definitions for CSV processing generated by ChatGPT (GPT-4o). Retrieved from https://chat.openai.com/