Downtown LA to Riverside Driving Traffic Pattern Project Description

DSCI510 Project

Spring 2023

Shuhang Feng

Introduction

The purpose of this project is to analyze the traffic patterns on highways most frequently taken between Downtown/Irvine and Riverside, focusing on I-10, I-60, SR-71, and SR-91. By studying these highways, we aim to better understand the average speeds, congested areas, peak travel times, and the impact of accidents on traffic flow. As a student at USC and a part-time employee near LAX, I personally use these routes often, and this project will help me make better-informed decisions about my daily commute.

Motivation Surrounding Project Topic

Traffic congestion in Southern California has become a significant issue for commuters like me, and its impact on our daily lives and the local economy cannot be ignored. Understanding traffic patterns and identifying problematic areas can help both individuals and policymakers make informed decisions about infrastructure improvements, traffic management, and transportation policies to alleviate congestion and improve overall traffic flow. By creating a model tailored to my personal commute, I hope to better understand the factors affecting my travel times and explore potential solutions for a more efficient journey.

Brief Description of Data Sources

We will utilize two primary data sources for this project:

a. Performance Monitoring System (PeMS) - Provided by the California Department of Transportation (Caltrans), this real-time traffic speed data covers major highways and freeways

throughout California. We will focus on the Los Angeles region, specifically the highways mentioned above.

b. https://data.lacity.org/Public-Safety/Traffic-Collision-Data-from-2010-to-Present/d5tf-ez2w CSV data set: Los Angeles Traffic Collision Data. The City of Los Angeles provides data on traffic collisions that occurred within the city from 2010 to present in CSV format. This data includes information on the location, time, and individuals. By analyzing this data, we can identify high-traffic areas with high collision rates and develop predictive models to help prevent future accidents.

There is a map visualization tool for traffic-collision-data, I inspected and found out that the accident count on east LA (approximately east of i605) has a limited number of only 65, while central and west LA has an accident count of 590k, which I assume the data for these places are not collected specifically. Thus I decided not to use this dataset since it did not include all data I need to analyze the traffic conditions from Downtown LA to riverside.

Analysis Performed

Analysis consists of the following steps:

i. Data Collection - Scrape traffic speed data from PeMS and accident data from LA Collision Data.

ii. Data Preprocessing - Clean and organize the collected data, focusing on the relevant highways and time periods.

iii. Speed Analysis - Calculate average speeds for each lane of each highway during specific time intervals and identify congested areas and peak travel times.

iv. Accident Analysis - Analyze the frequency and distribution of accidents on the target highways and assess the impact of accidents on traffic speeds and congestion.

v. Correlation and Predictive Modeling - Explore the relationship between average speeds, accidents, and other factors such as weather and events. Develop predictive models to estimate future traffic patterns and congestion levels for my daily commute.

Conclusions Drawn

As the visualization graph shows, the lane on the left has the highest average speed.