CAPSTONE PROJECT

Exploring California Commute Trends
A Lens for Bahrain Transport Future

Prepared by: Mariam Azam
Data Analytics Immersive Bootcamp Fellow, General Assembly

June - July | 2025

DOCUMENTATION

Lead Instructors:
Juan Carlos OrozcoRey
Ali El-Kassas

Submitted on: 1st July 2025

Project Agenda

- ✓ Data Dictionary

- ✓ Data Handling Summary

- ✓ Introduction

- ✓ Problem Statement

- ✓ Project Approaches

- ✓ Target Audience

- ✓ Dataset Overview

- ✓ Analytical Questions

- ✓ Key Findings

- ✓ Recent Trends: California vs. Bahrain

- ✓ Limitations & Future Opportunities

- ✓ Recommendations for Bahrain

- ✓ References

Data Dictionary

Columns: 25

Column name	Description
ind_definition	Indicator definition (String)
reportyear	Year of the report (Integer) (2000, 2005-2006, 2006-2010, 2008-2010)
race_eth_code	Race/ethnicity code (Integer)
race_eth_name	Race/ethnicity name (String) 8 ethnicities (Latino, White, Asian, AfricanAm, AIAN, NHOPI, Multiple, Other) + Total
geotype	Geographic type (String) CA (California), RE (Region), CO (County), PL (Places), CT (Census Tracts)
geotypevalue	Geographic type value (String)
geoname	Geographic name (String) Based on Geotype

Column name	Description
county_name	County name (String)
county_fips	County FIPS code (Integer)
region_name	Region name (String)
region_code	Region code (Integer)
mode	Mode of transportation (String)
mode_name	Mode of transportation name (String) 8 modes (Worked at home, Bicycle, Car/truck/van drive alone, car/truck/van carpool, Car/truck/van total, Public Transport, Walking)
pop_mode	Population using the mode of transportation (Integer)
pop_total	Total population (Integer)

Column name	Description
percent	Percent of population using the mode of transportation (Float)
LL95CI_percent	Lower 95% confidence interval for the percent of population using the mode of transportation (Float)
UL95CI_percent	Upper 95% confidence interval for the percent of population using the mode of transportation (Float)
percent_se	Standard error for the percent of population using the mode of transportation (Float)
percent_rse	Risk ratio for California residents against that region or county (Float)
CA_decile	Decile for the overall population size in California against a race/ethnicity or geographic area type (Integer)
CA_RR	Risk ratio for California residents against that region or county (Float)
version	Version number for identifying changes in the dataset (Integer)

Data Handling Summary

Column Names	Were Self-explanatory. No changes were made.
Data Types	Columns were in the right data format. No changes were made.
Null check	Key columns needed for analysis had nulls. region_name: 33% (were compensated for using column geotype & geoname) population: 52% (blanks were excluded for analysis)
Duplicate-check	No duplicates found.
Dropped Columns	Columns on page 4 of this report were not used for analysis. However, they were not dropped.
Additional Columns	No new columns were added.
Additional data	Dataset with information on household income was merged with the main dataset in Power BI. It had two columns: county_name and household income.

Introduction

Commuting to work is a key part of daily life and shapes access to jobs, services, and opportunities. How people commute reflects patterns of access, affordability, and equity.

Bahrain 2030 Vision: Transport Highlights

The Kingdom emphasizes the need for a sustainable, inclusive, and efficient transportation system to support economic growth, reduce congestion, and improve quality of life.

California: Case Study

California presents a compelling case with its:
• Diverse population and wide range of commuting modes across regions and ethnicities
• Persistent challenges in transit inclusiveness and accessibility, despite being a highly developed state

Problem Statement

Challenges in commuting contribute to mobility gaps, road congestion, safety risks, longer commute times, and growing inequities across regions, income levels, and racial groups.

Approaches

• Analyze commuting trends across California regions and ethnicities
(2000, 2005-2010)
• Integrate income as an external factor to understand *why* choices differ
• Explore the overtime shifts in commutimg choices
• Extract insights applicable to Bahrain's transportation goals

Target Audience

1. Government Authorities
2. Public Transport Companies
3. Research & Academic

Dataset Overview

Two Datasets were used.

Data Sources: U.S. Census Bureau, Decennial Census & American Community Survey
Data Size: 336K Rows
Data Period: 2000 | 2005 - 2010

Key Features of Main Dataset:
• Report Year
• Geography (Region | County | Places)
• Ethnicities (8 Ethnicities)
• Commute Modes (7 Commute Modes)
• Population

Key Features of Supplementary Dataset:
• County Names
• $ Household Income

Note: *Data Dictionary & Data Handling Summary is provided at the end of the report.*

Sample for Online Voices on Bahrain Public Transit

Discussion Boards on Reddit platform were reviewed. Around 60-80 comments were used to extract keywords that capture the public perception of Bahrain Public Transport.

Data Period:
Ranging from 2 years ago – 2 months ago.

Technical Analysis | Key Findings

Geographical Trends

What are the most common commuting modes used in California?
How do they vary by region?
How does public transport usage correlate with population size?

Key Takeaways:

• Driving alone is the most dominant commute mode across all regions.
• Larger population size doesn't necessarily translate to higher transit use.

Demographic Trends

Which ethnic groups are most and least likely to use public transportation?

Key Takeaways:

• African Americans have the highest public transit usage; White populations have the lowest.
• Transit use varies by ethnicity but remains generally low across groups.

Socioeconomic Factor

• How do income levels influence commute choices?

Key Takeaways:

• Driving alone is the leading commute mode across all income levels.
• Contrary to expectations, low- and middle-income groups do not rely heavily on public transit. Many opt to walk instead.

Overtime Shift

• Which commuting modes have declined overtime?
• How have commuting patterns shifted for Public Transportation?

Key Takeaways:

• Carpooling and walking have declined over the decade.
• Public transit use has remained stagnant, showing no significant improvement.

Recent Trends: California vs. Bahrain

California

• Latest trends show that transit ridership declined due to rising car ownership, ride-hailing services, and residential shifts.
• An increase in the population who work remotely has been noticed after the pandemic.

Bahrain

• Significant infrastructure improvements in recent years (e.g., expansion of bus routes, terminal upgrades).
• However, public perception reflects concerns around comfort, heat, accessibility, and service frequency.
• Transit remains underused due to limited awareness and walking infrastructure.

Limitations | Future Considerations

• **Local Data**: No access to real commuting data from Bahrain. Research-based insights were shared
• **Period Gap**: The dataset spans 2000 to 2010, with a gap from 2001–2004. It presents an opportunity for future research to explore changes during this intermediate period and build a more continuous understanding of commuting patterns."
• **Crucial Factors**: E.g. distance, time, are crucial to understand why commute choices are made
• **Commuter Experience**: E.g. comfort, satisfaction, or complaints with public transit services

Recommendations | Takeaways for Bahrain Transport Future

• Gather national commuting data and incorporate richer external factors (e.g., regional aspects, occupation, income levels) for Bahrain-specific insights
• Assess the needs of different population groups (e.g., students, low-income workers) to ensure transit is accessible to these diverse groups
• Improve walking conditions around residential and commercial areas with safe & shaded Infrastructure
• Use outreach and education to build confidence in transit systems and reduce car dependence, directly aligning with Bahrain's transportation goals

References

✓ *California Commuting Mode choice*. (2023, January 12). Kaggle. https://www.kaggle.com/datasets/thedevastator/california-commuting-mode-choice-from-2000-2010

✓ *Commute mode*. (n.d.). Bureau of Transportation Statistics.

   https://www.bts.gov/browse-statistical-products-and-data/state-transportation-statistics/commute-mode

✓ Center for Jobs. (2018, January 3). *California commuters continue to choose single occupant vehicles - Center for jobs.* https://centerforjobs.org/ca/special-reports/california-commuters-continue-to-choose-single-occupant-vehicles

✓ *'Over 26,000 new cars hit Bahrain roads every year' | THE DAILY TRIBUNE | KINGDOM OF BAHRAIN*. (2022, October 27). DT News. https://www.newsofbahrain.com/bahrain/85168.html

✓ *National Development: Bahrain Vision 2030 | Invest in Bahrain*. (2024, August 29). Invest in Bahrain. https://www.bahrainedb.com/about-us/national-development-strategy

✓ Staff, U. I. (2025, June 28). *What's Driving California? The Past, Present, and Future(s) of Transportation in the Golden State*. UCLA Institute of Transportation Studies. https://www.its.ucla.edu/publication/whats-driving-california/

✓ Provider map. (n.d.). California Mobility Marketplace. https://www.camobilitymarketplace.org/provider-map.html

THANK YOU