Yatsa Case Study

Study Objectives

1

To find the average distance vehicles travel in a day, week, or month? Or over weekdays vs weekends?

Study Objectives

1

To find the average distance vehicles travel in a day, week, or month? Or over weekdays vs weekends?

2

To determine the most common destinations for vehicles and how often they travel to those destinations

Study Objectives

1

To find the average distance vehicles travel in a day, week, or month? Or over weekdays vs weekends?

2

To determine the most common destinations for vehicles and how often they travel to those destinations

3

To find any patterns or trends in driver behaviour, such as speeding

Study Objectives

1

To find the average distance vehicles travel in a day, week, or month? Or over weekdays vs weekends?

2

To determine the most common destinations for vehicles and how often they travel to those destinations

3

To find any patterns or trends in driver behaviour, such as speeding

4

To see trends in the types of trips made by vehicles.

02

Data loading and preprocessing

Data description

The data was provided in JSON format:

1. assets.json - has information on the assets tracked by Yatsa
2. trips.json - has information on trips made by assets as tracked by Yatsa

Loading the data

JSON ---->> Python Dictionary ---->> DataFrame

Attributes overview

Trips Data

4804 entries, 15 columns

Missing values

origin.odometer - all values
destination.odometer - all values
origin.address - 1
destination.address - 1

Assets Data

34 entries, 4 columns

Missing values

Year - 45%
Make - 38%

01

Data cleaning

Make copies of the data frames, results:

assets. copy and trips. copy

02

Data cleaning

Retain only 'id' and 'make' columns of the assets copy to later merge with the trips DataFrame

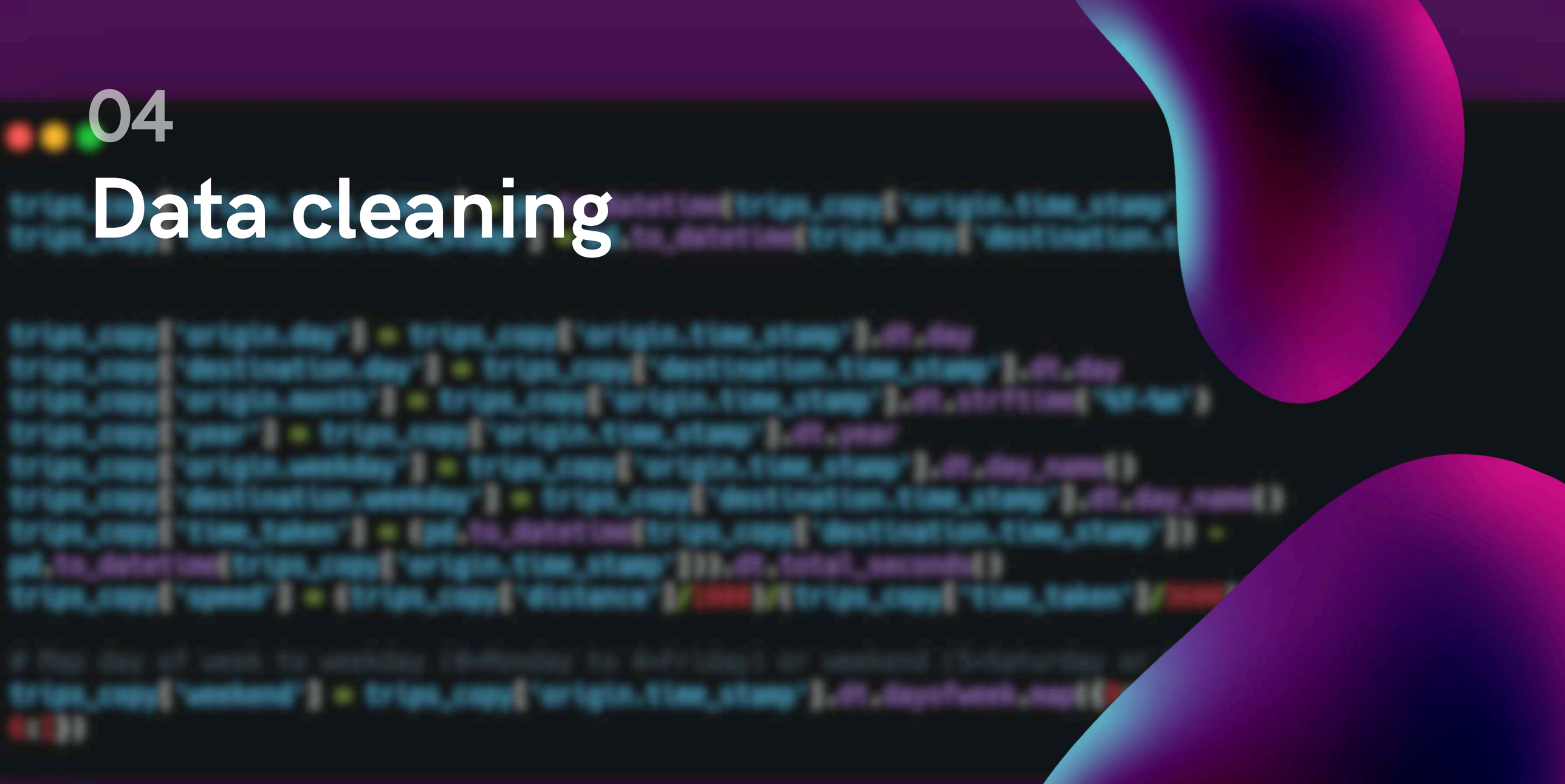
03

Data cleaning

Drop the row containing null values in the '**destination.address**' and '**origin.address**' columns

04

Data cleaning



Convert **time_stamp** columns to **datetime** data type and create new columns from the **origin.time_stamp** and **destination.time_stamp**

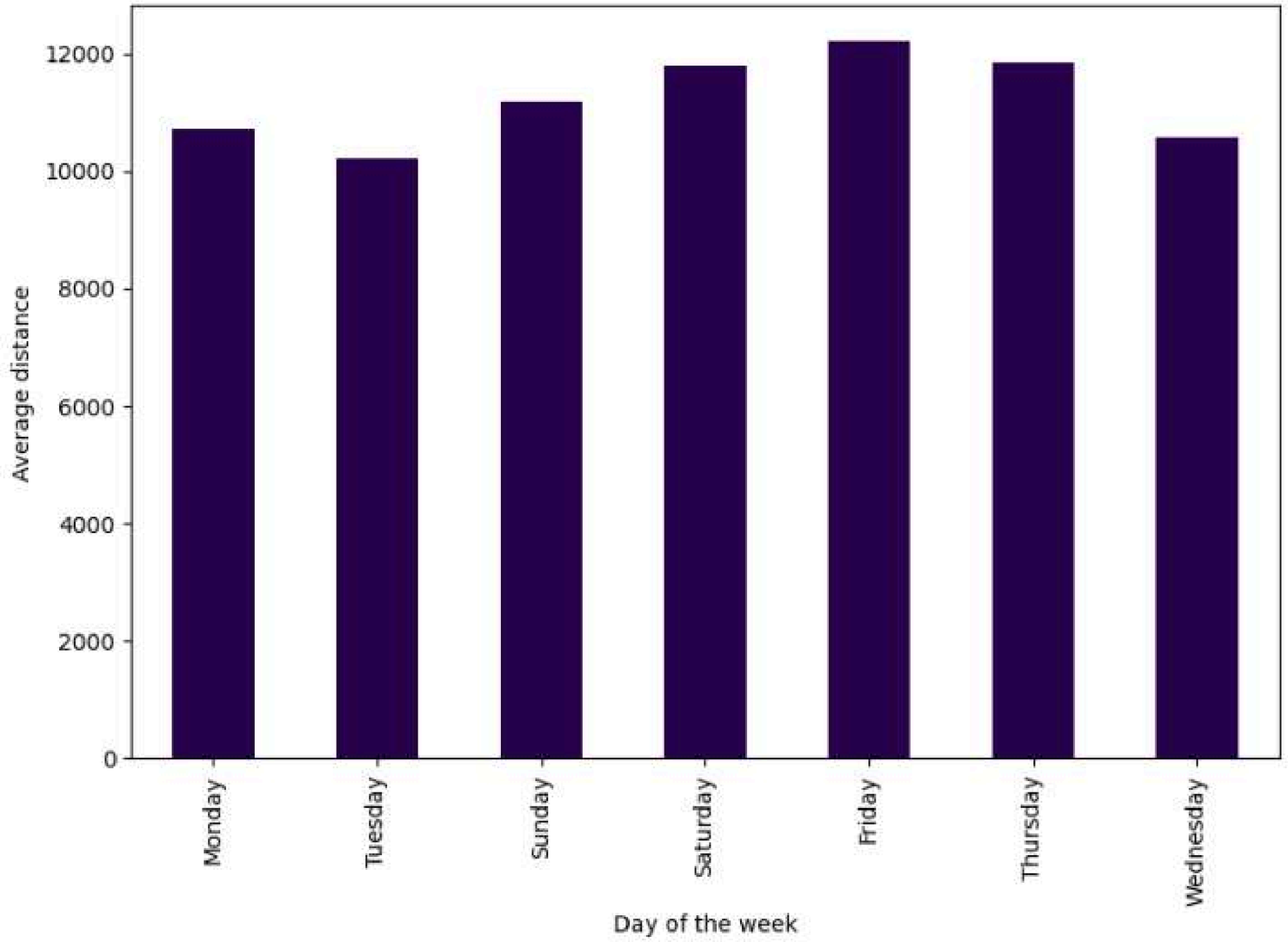
Merging the data frames

Merge the trips and assets dataframes
Retain useful columns

03

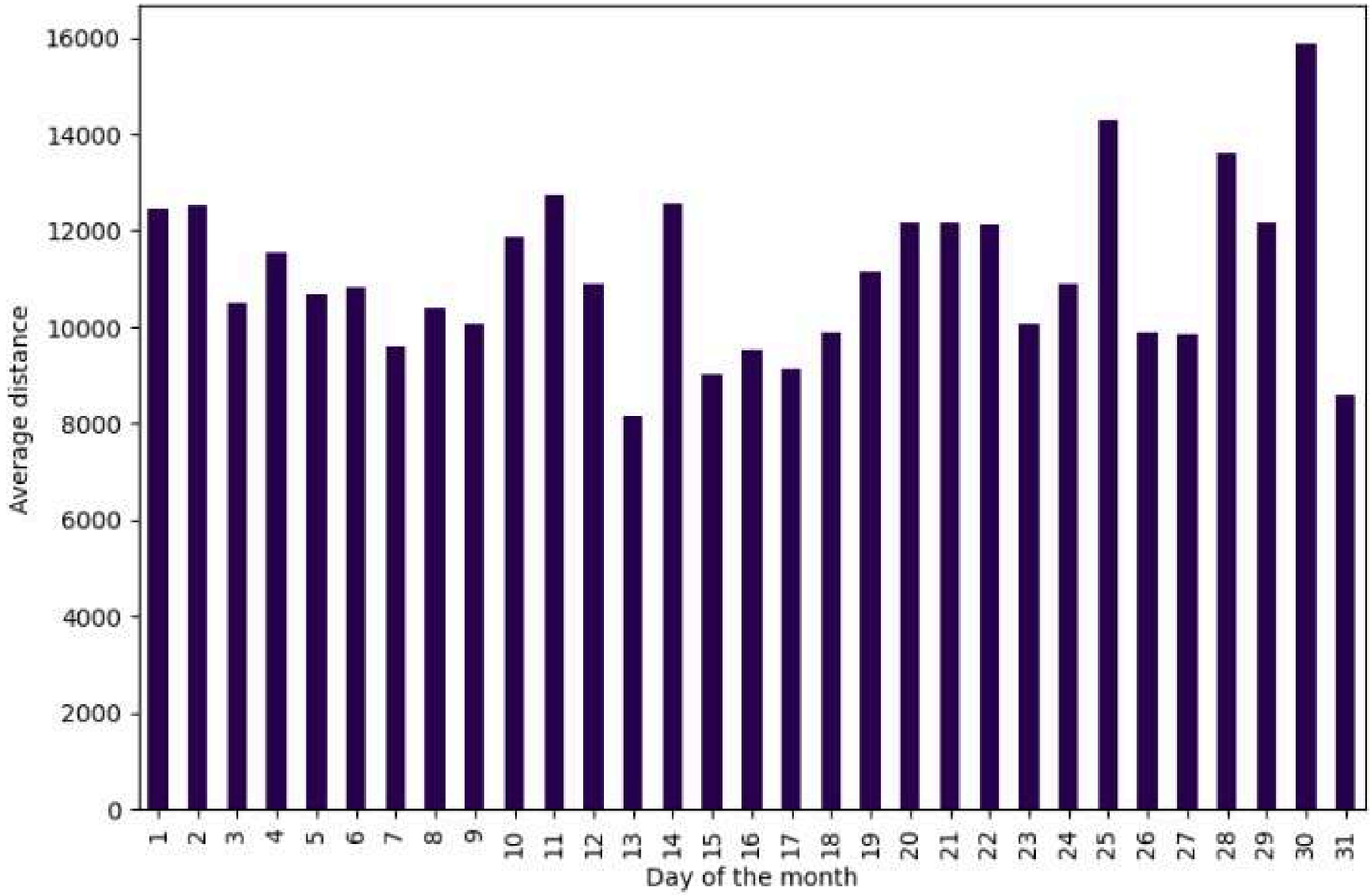
Exploration & Key findings

Average Distance Covered per Day of the Week



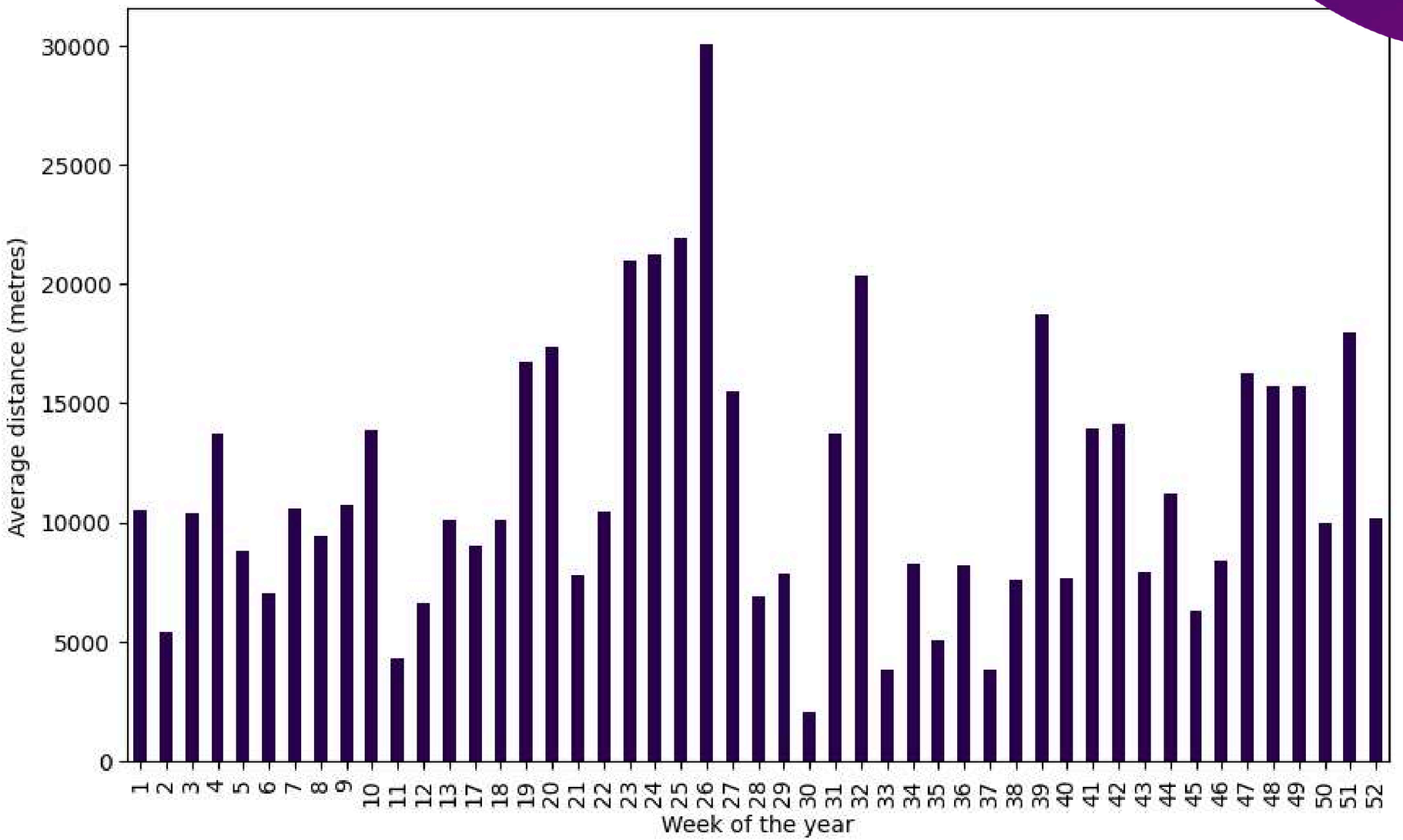
- Fridays have the highest average distance covered
- Followed closely behind by Saturdays and Thursdays
- Tuesday and Wednesday record the lowest average distance

Average Distance Covered per Day of the Month



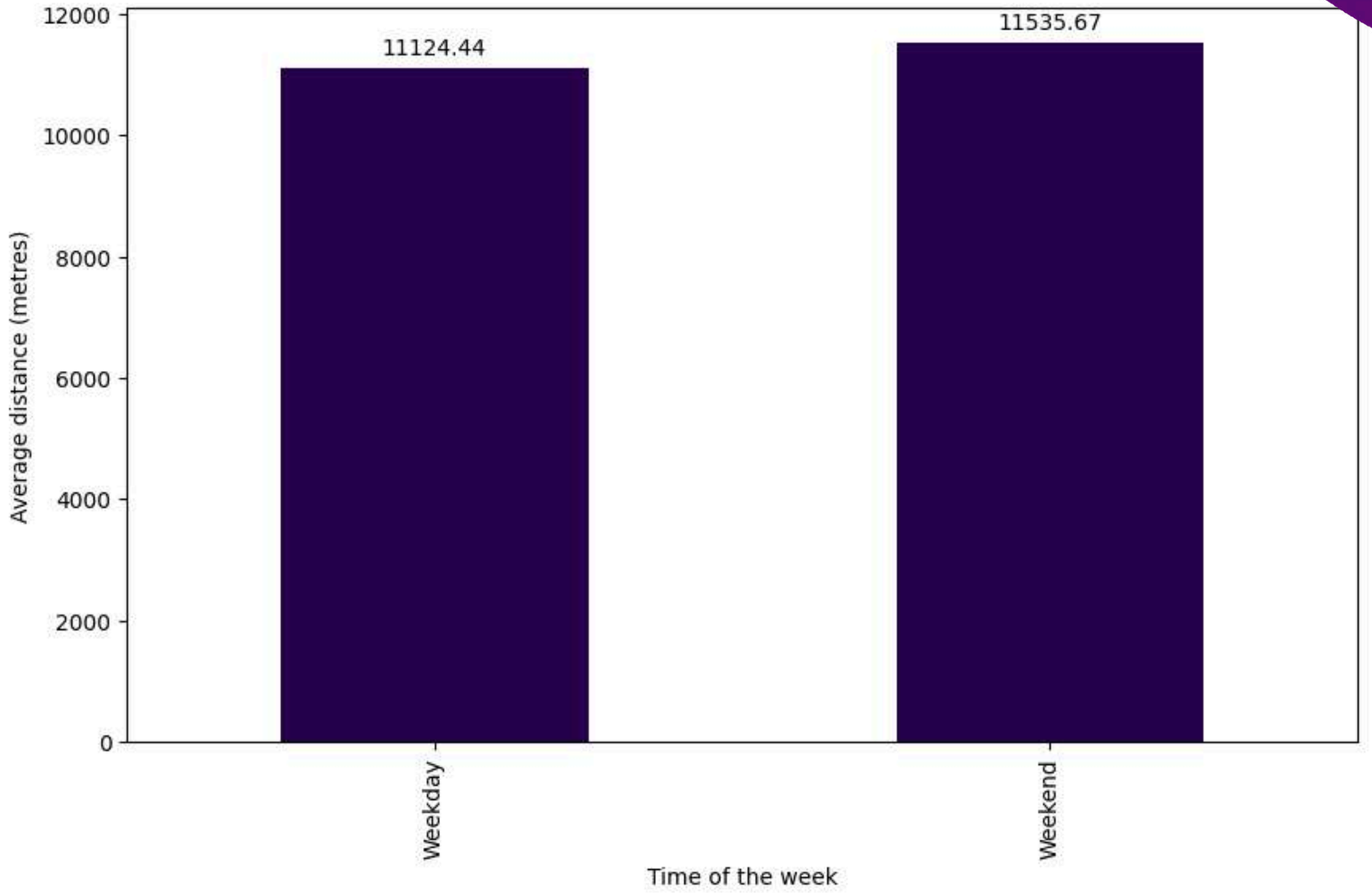
- Most people travelled on the 30th day of the month
- A good number also travelled on the 25th and the 28th days
- The least travelled days of the month were the 13th and 31st

Average distance covered per week



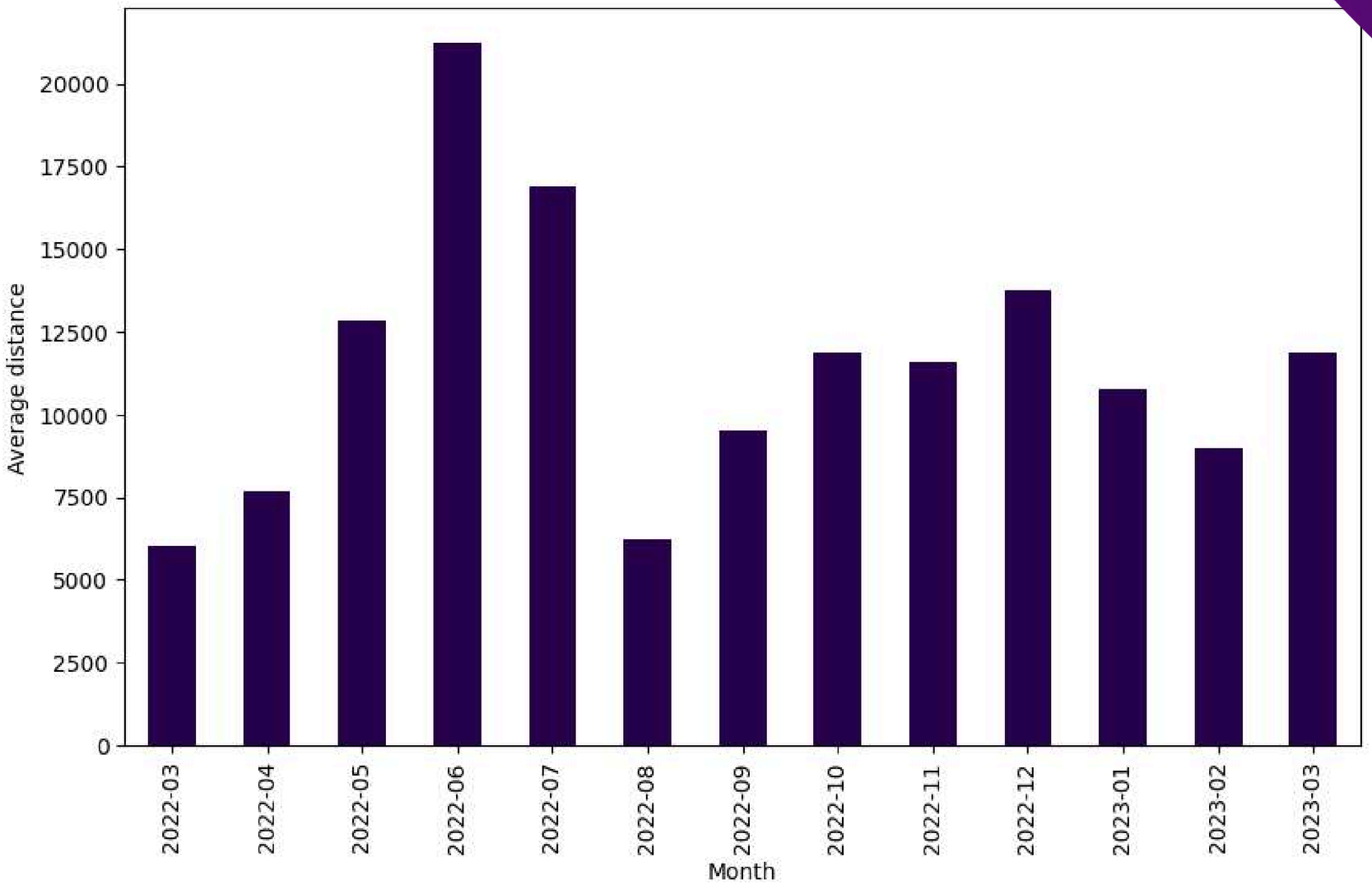
- The 26th week had the highest average distance
- Most distance is covered from the 23rd to the 26th week
- The 30th week had the least average distance covered

Average distance covered on Weekdays vs Weekends



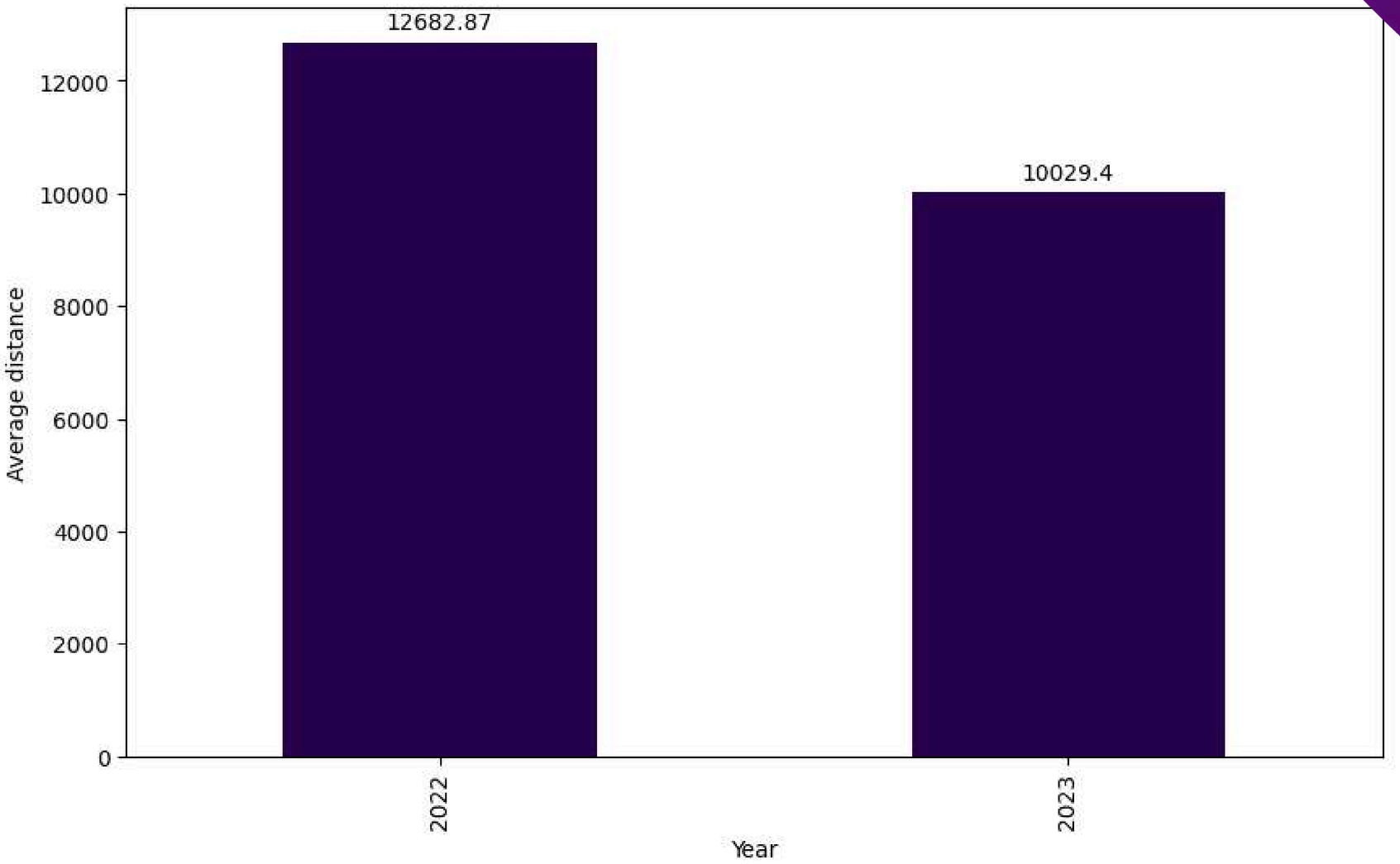
- More distance was covered on the weekends

Average distance covered per month

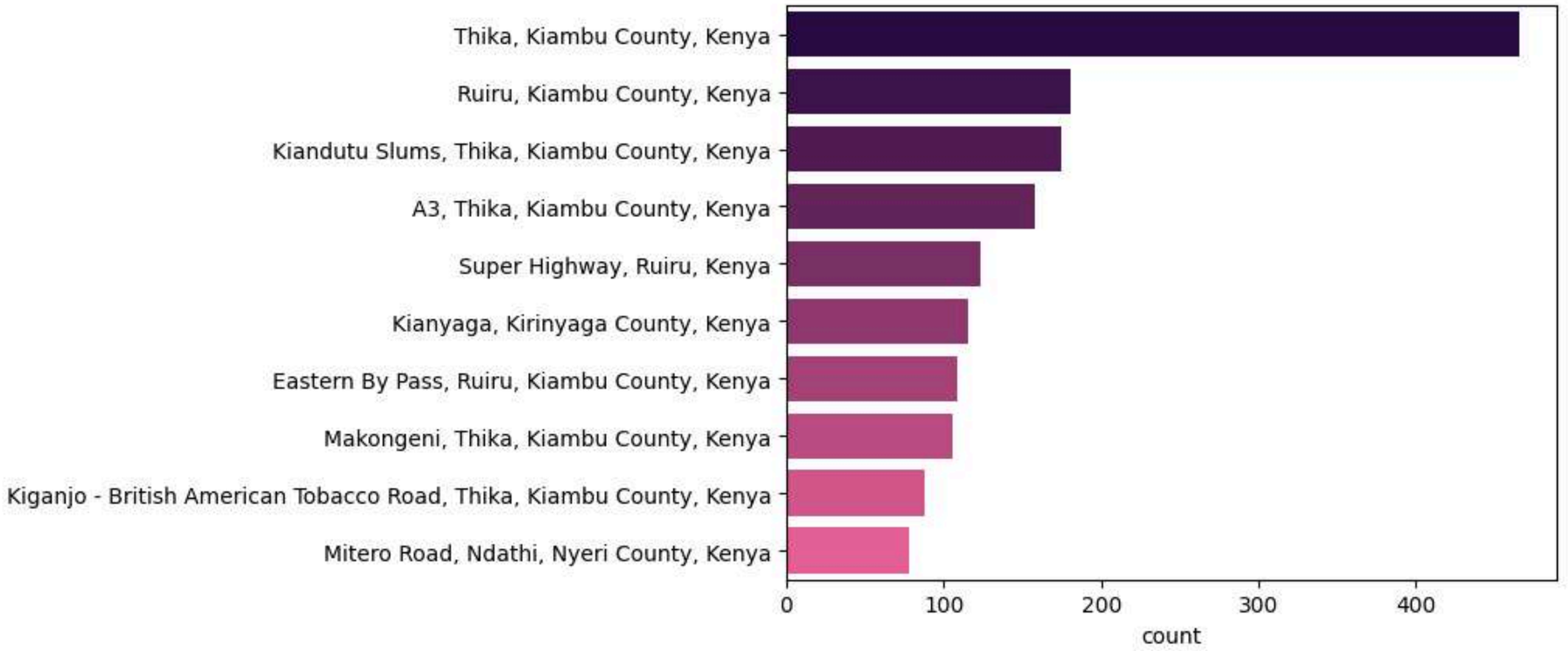


- Highest average distance was recorded in June 2022
- Least average distance was in March and August the same year
- Longest average distance was recorded between Mar and Dec 2022

Average distance covered per year

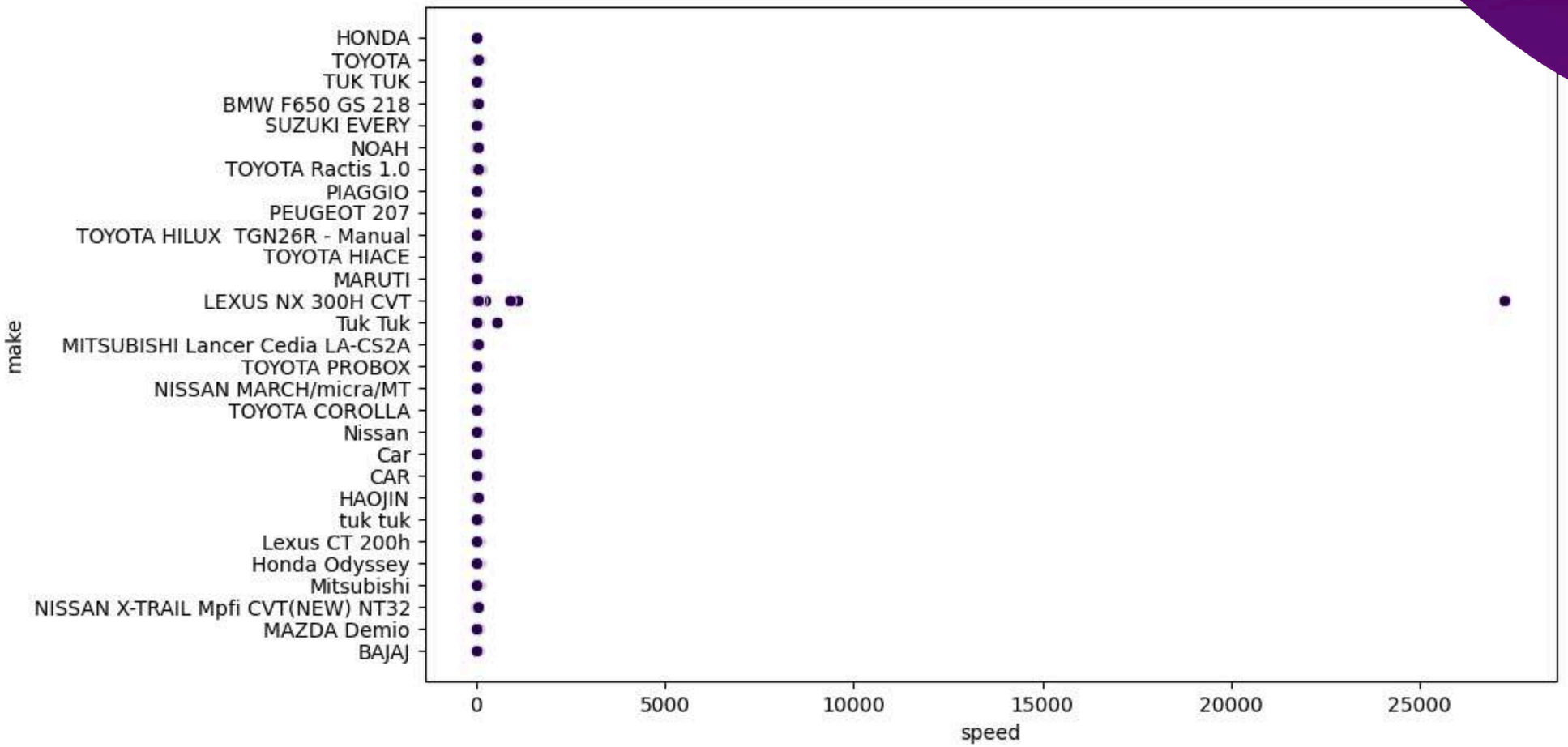


Most Common Vehicle Destinations



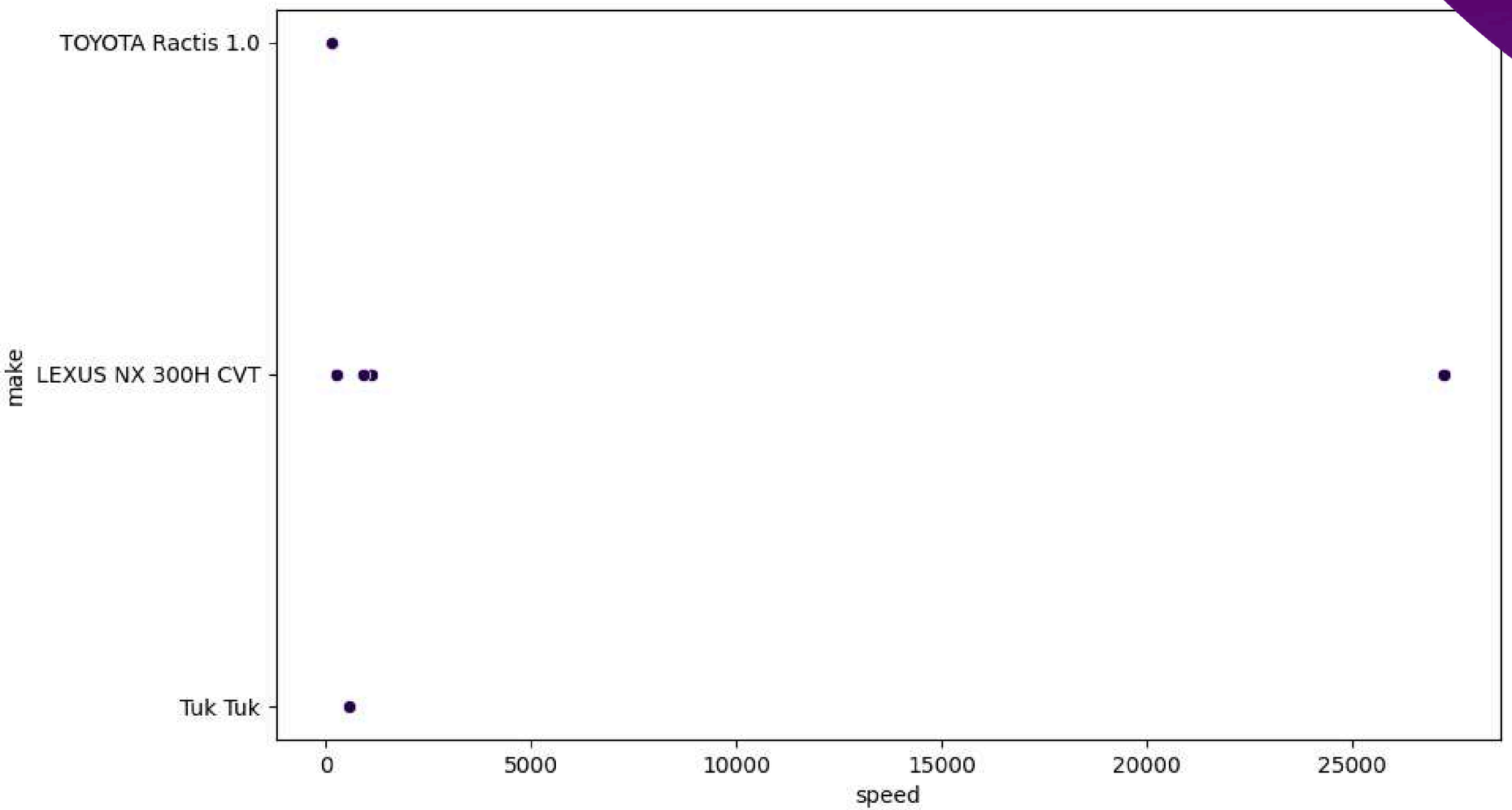
- Kiambu county is the most common destination
- Thika is the most popular destination, followed by Ruiru

Trends in Driver Behaviour: Speeding



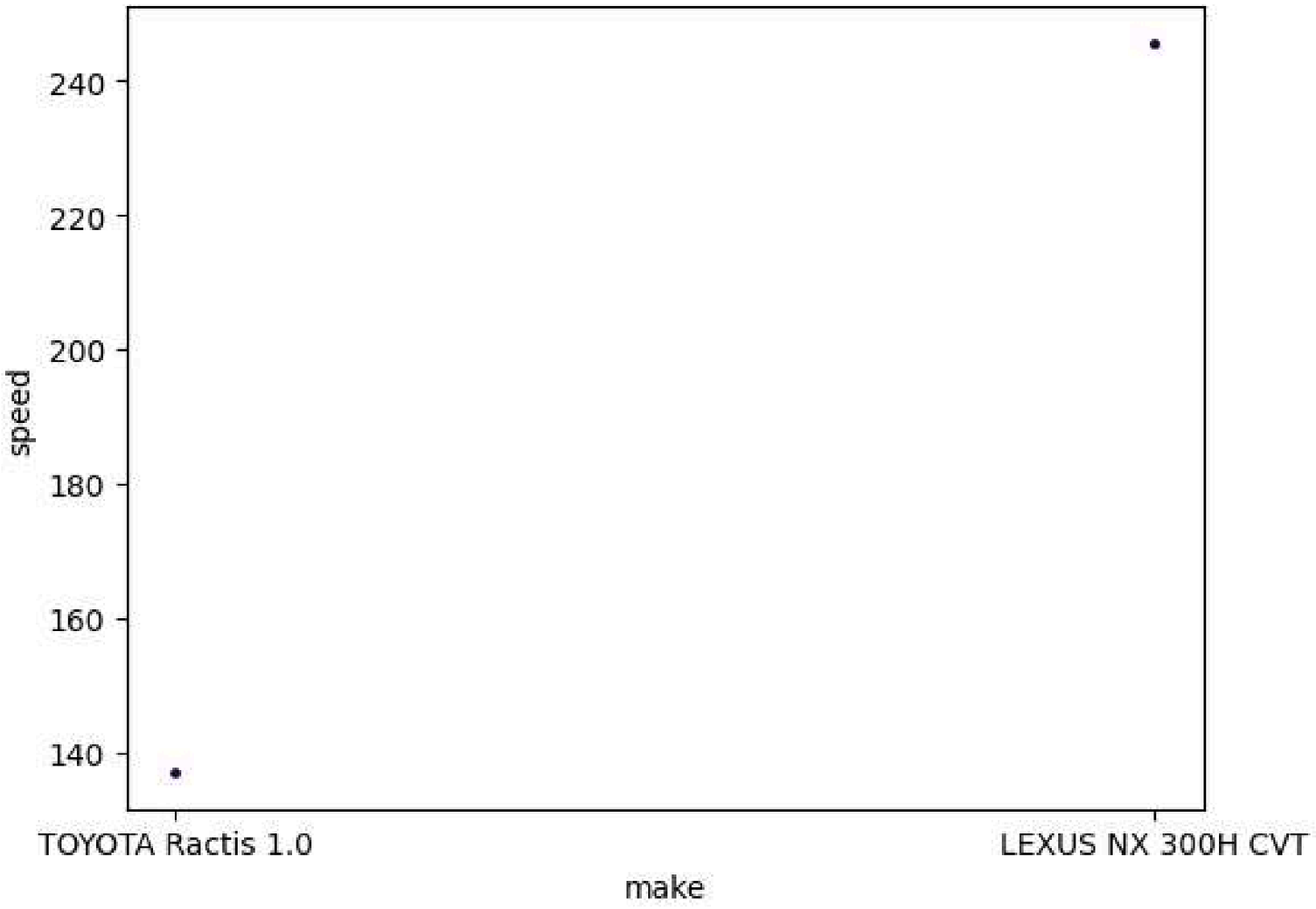
- Lexus NX had the highest average speed (> 200 km/hr)

Trends in Driver Behaviour: Speeding



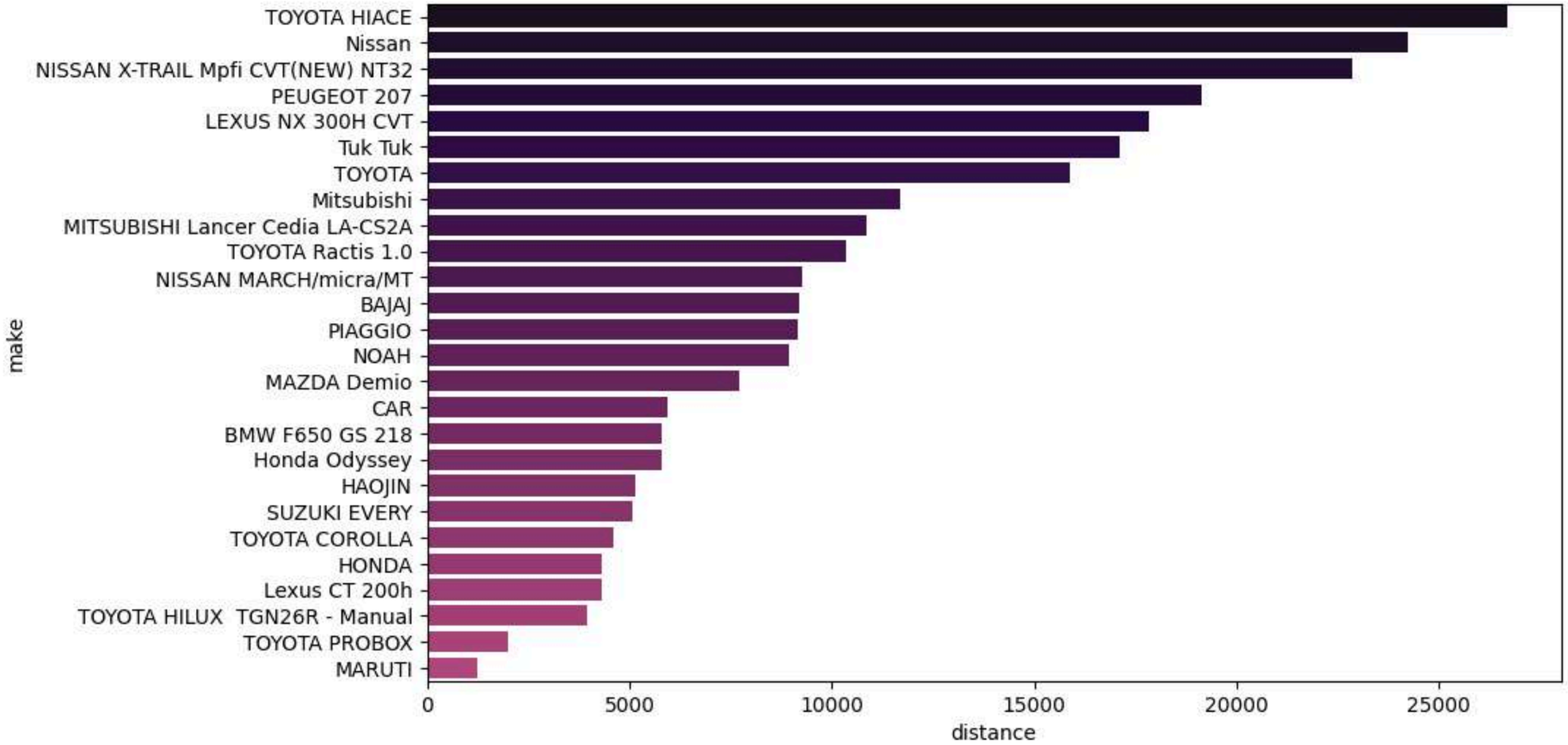
- Lexus NX had the highest average speed (> 200 km/hr)

Trends in Driver Behaviour: Speeding



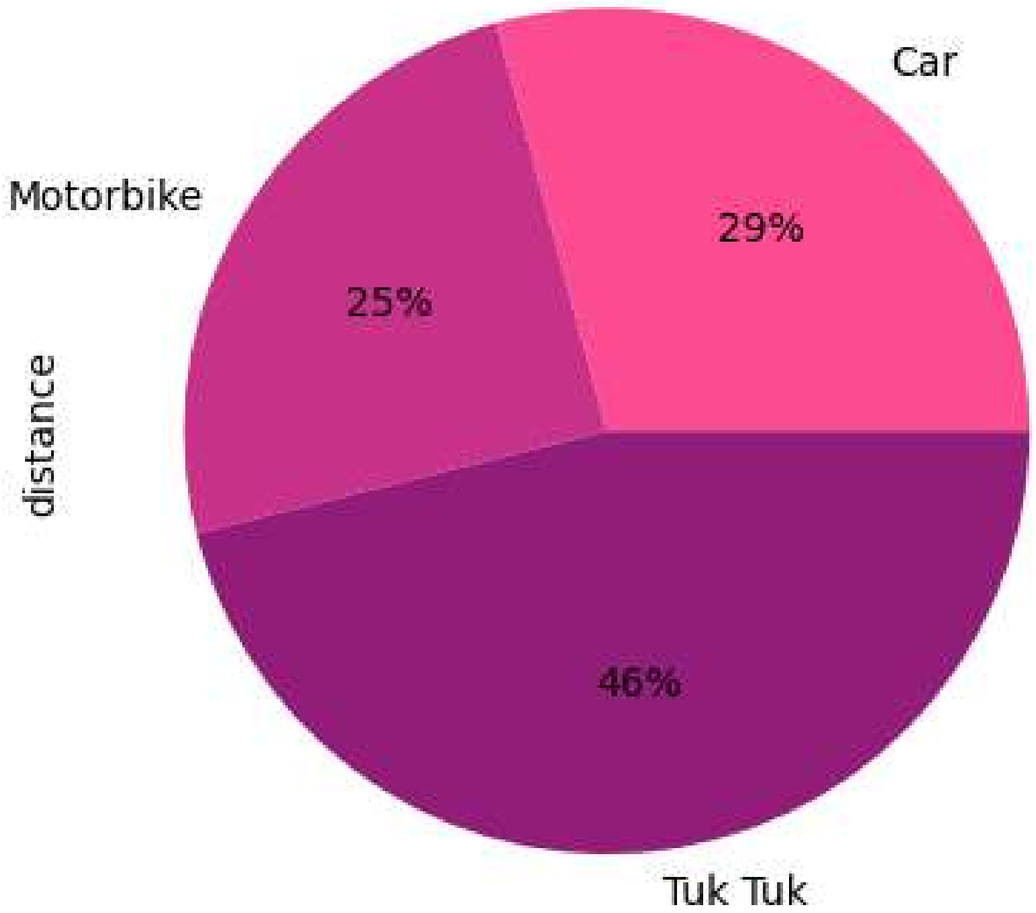
Lexus NX had the highest average speed over 240 km/hr.

Trends in types of trips made by vehicles: Distance



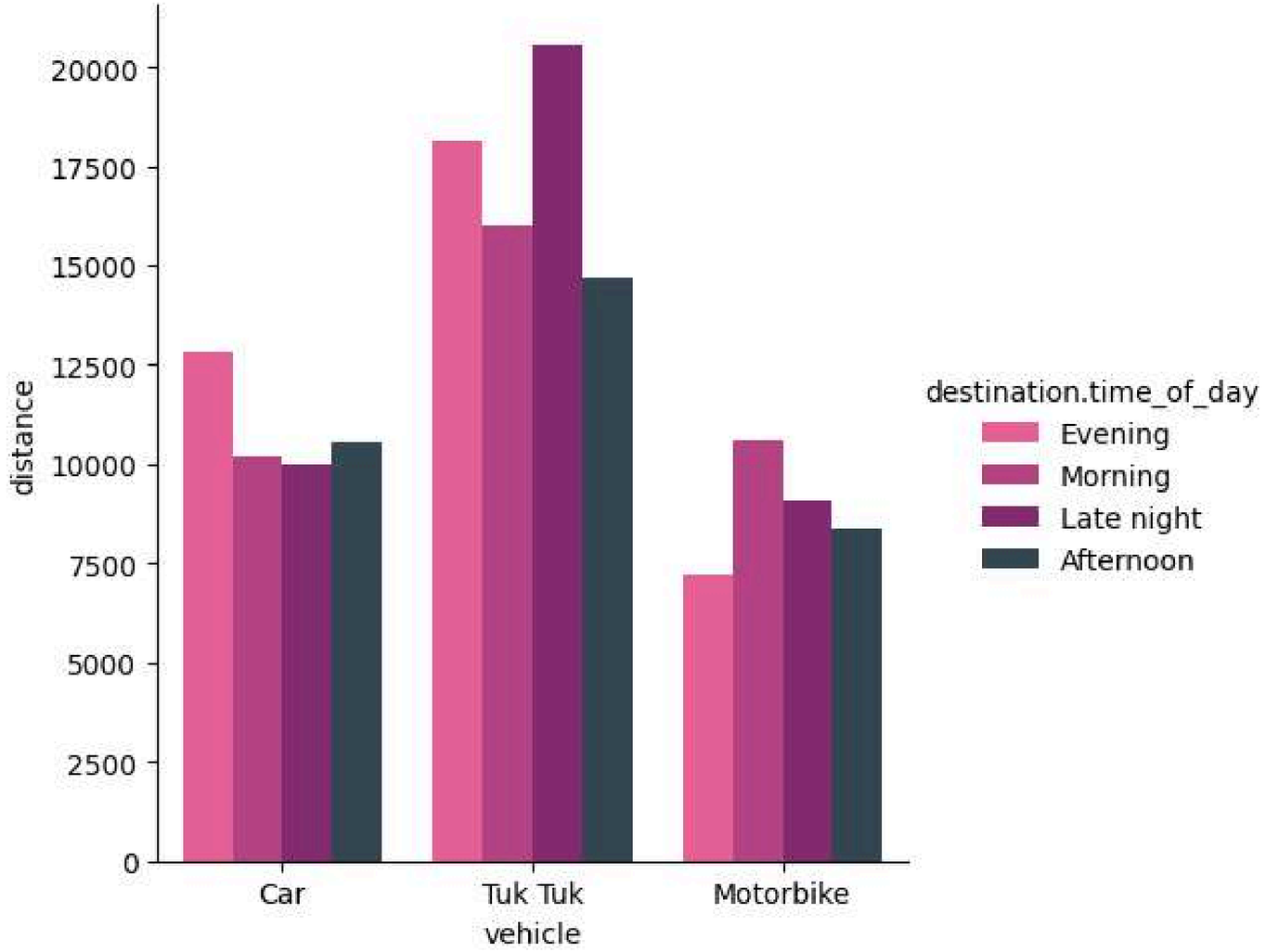
- Toyota Hiace had the highest average distance covered
- Tuk-tuks also recorded a high average distance
- Maruti had the least average distance covered

Trends in types of trips made by vehicles: Distance



- Tuk-tuks cover most of the total distance (46%)

Trends in types of trips made by vehicles: Distance & Time of Day



- Tuk-tuks cover most of the total distance
- Interestingly most of it late at night

Key Findings

1

Most trips are made over the weekend, Friday being the day with the highest number of trips. Most trips are also made on the 25th, 26th and 30th implying most trips are made during the end month

Key Findings

1

Most trips are made over the weekend, Friday being the one with the highest number of trips. Most trips are also made on the 25th, 26th and 30th implying most trips are made during the end month

2

Kiambu, Kirinyaga and Meru counties appear the most on top 10 destinations with Thika being the most popular destination

Key Findings

1

Most trips are made over the weekend, Friday being the one with the highest number of trips. Most trips are also made on the 25th, 26th and 30th implying most trips are made during the end month

2

Kiambu, Kirinyaga and Meru counties appear the most on top 10 destinations with Thika being the most popular destination

3

Lexus NX drivers had the highest average speed

Key Findings

1

Most trips are made over the weekend, Friday being the one with the highest number of trips. Most trips are also made on the 25th, 26th and 30th implying most trips are made during the end month.
Most trips made in June, least in August

2

Kiambu, Kirinyaga and Meru counties appear the most on top 10 destinations with Thika being the most popular destination

3

Lexus NX drivers had the highest average speed

4

Tuk Tuks are second in place with the highest average distance covered after Toyota Hiace, most of the trips made late in the night.

Recommendations

1

Focus on popular destinations: Since Kiambu, Kirinyaga, and Meru counties appear the most on top 10 destinations with Thika being the most popular destination, you can prioritize marketing efforts towards these areas to attract more customers.

Recommendations

1

Focus on popular destinations: Since Kiambu, Kirinyaga, and Meru counties appear the most on top 10 destinations with Thika being the most popular destination, you can prioritize marketing efforts towards these areas to attract more customers.

2

Monitor and incentivize safe driving: Since Lexus NX drivers had the highest average speed, it's important to monitor driver behavior and incentivize safe driving practices. You can consider implementing a driver rating system that rewards safe driving behaviors such as obeying traffic rules and driving within speed limits.

Recommendations

1

Focus on popular destinations: Since Kiambu, Kirinyaga, and Meru counties appear the most on top 10 destinations with Thika being the most popular destination, you can prioritize marketing efforts towards these areas to attract more customers.

2

Monitor and incentivize safe driving: Since Lexus NX drivers had the highest average speed, it's important to monitor driver behavior and incentivize safe driving practices. You can consider implementing a driver rating system that rewards safe driving behaviors such as obeying traffic rules and driving within speed limits.

Limitations

1

The dataset given on the assets and their models was quite small and collected over a short period of time. So it might be inaccurate to suggest trends based on this dataset alone

Limitations

1

The dataset given on the assets and their models was quite small and collected over a short period of time. So it might be inaccurate to suggest trends based on this dataset alone

2

Having data on average fuel consumption per make/model would also have been really great for the exercise.

thank you!