

A close-up photograph of a black car's windshield and roof. A bright yellow price tag with the number '9499' is visible on the windshield. The car is parked in a lot, with other vehicles blurred in the background.

Cars Data Set

Team Name - Velocity

Team Members - Daniel, Alston, Eugene

Introduction

Welcome, everyone. Today, We're excited to present our analysis of the "CarsData" dataset. Our goal is to dive into this dataset, uncover insights, and share our findings with you. Throughout this presentation, we'll provide an overview of the dataset, address research questions, showcase data visualization, explore SQL queries, and highlight key findings.

Let's dive in and explore the amazing world of automotive data together.



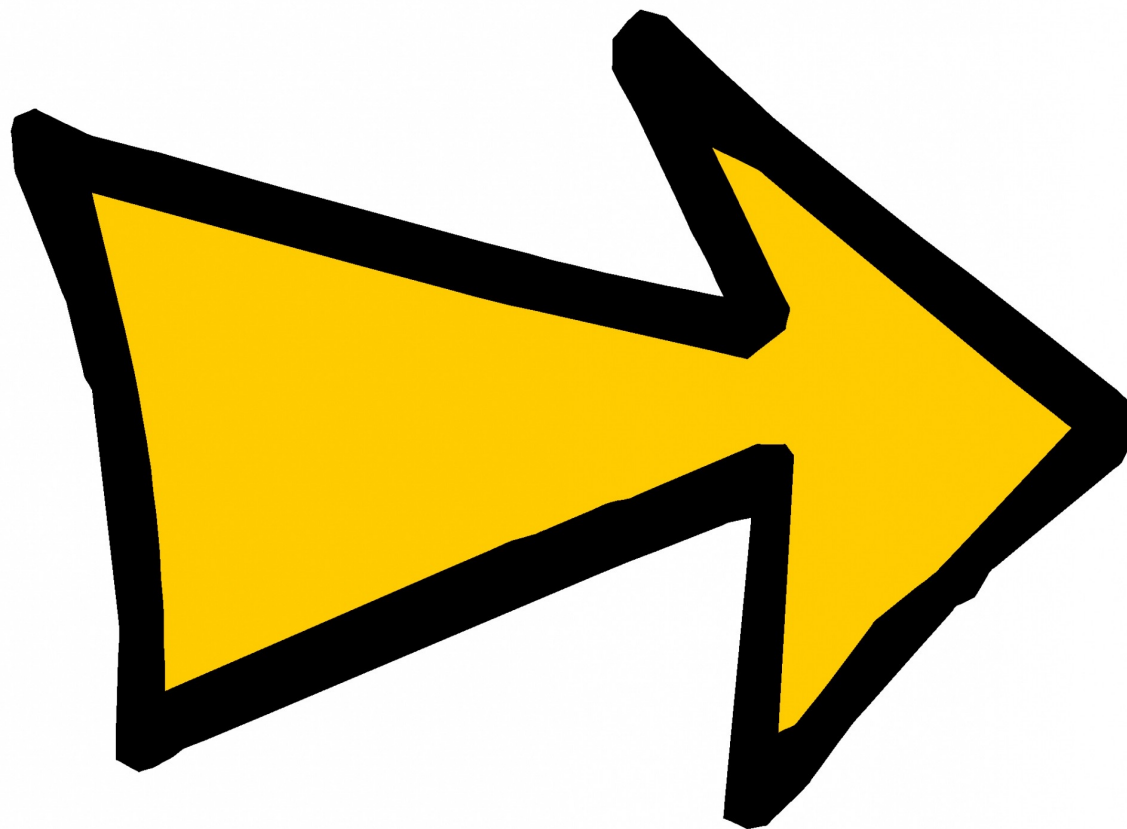
Dataset Description

- Overview
 - Comprehensive vehicle information database.
 - Includes make, model, year, fuel type, transmission, price
 - Sourced from Kaggle, URL – [Vehicle Dataset from Kaggle](#)
- Reliability
 - Data reliability hinges on Kaggle's credibility
 - Acknowledges potential issues: incomplete lists, outdates data
- Emphasizes thorough verification and cleaning
- Data Cleaning
 - Addressed missing, duplicate, and inconsistent data
 - Improved data quality for meaningful analysis
 - Steps included handling missing data, removing duplicates, standardizing formats

Dataset Overview

A complete overview of the car world is presented in an "CarsData" data set, providing a variety of features for each vehicle model. It contains a number of columns that provide important information on the used vehicles included in this data set.

These Columns include:



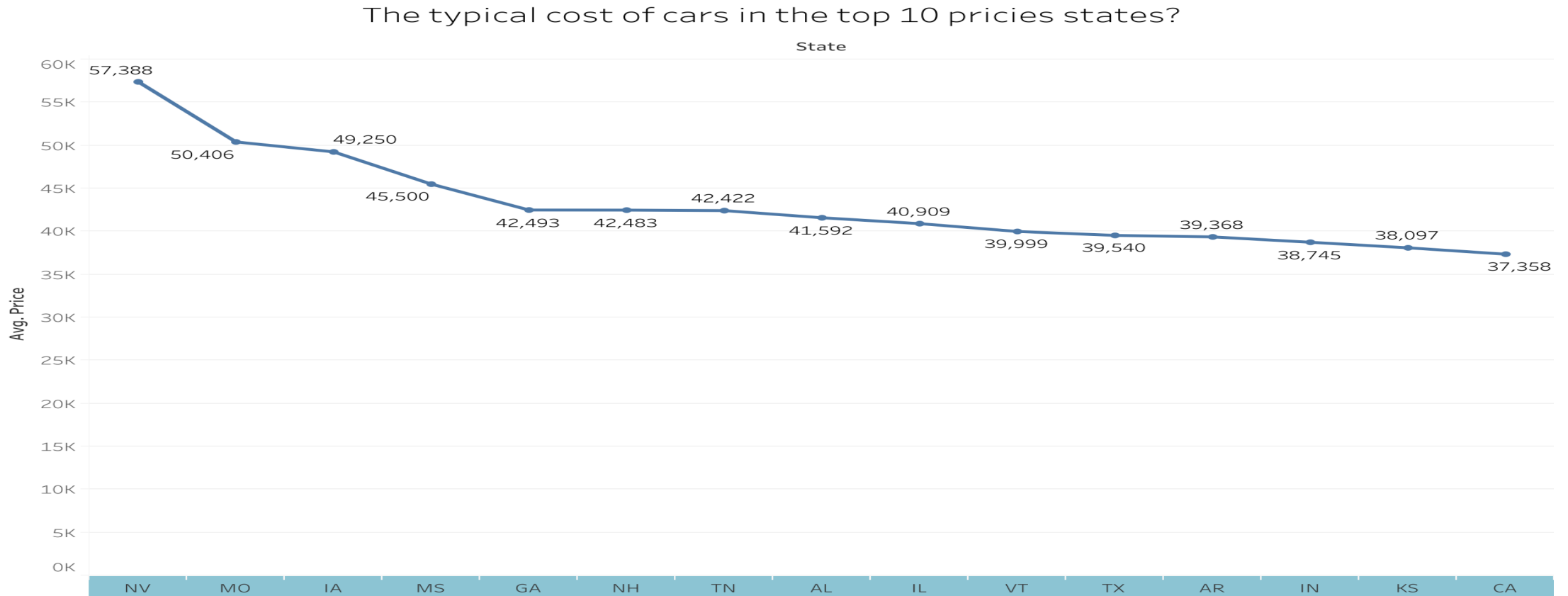
Dataset Overview

- **Make/Model:** Represents the brand/model
- **Price:** Indicate the price of each car model.
- **Year:** Denotes the manufacturing year of the car.
- **Fuel Type:** Specifies the type of fuel the car uses. (eg. Gasoline, Electric, E85 Flex Fuel)
- **Transmission:** Type of transmission systems.
- **State:** Indicates the State where the car is located.
- **InteriorColor:** Represents the Interior color of the car.
- **Used/New:** Indicates weather the car is used or new.
- **SellerType:** Indicates the type of seller.
- **Min MPG:** Denotes the min miles per gallon
- **Max MPG:** Maximum max miles per gallon
- **Drivetrain:** Types of drivetrain systems (eg, FWD, RWD, AWD)

This databank provides valuable information on different aspects of cars, such as their specifications, cost and location. It provides a powerful source of information for analyzing and exploring, given the number of entries covering various car models.

Data Visualizations 1

Line Chart- What are average car prices in the top 15 priciest states?

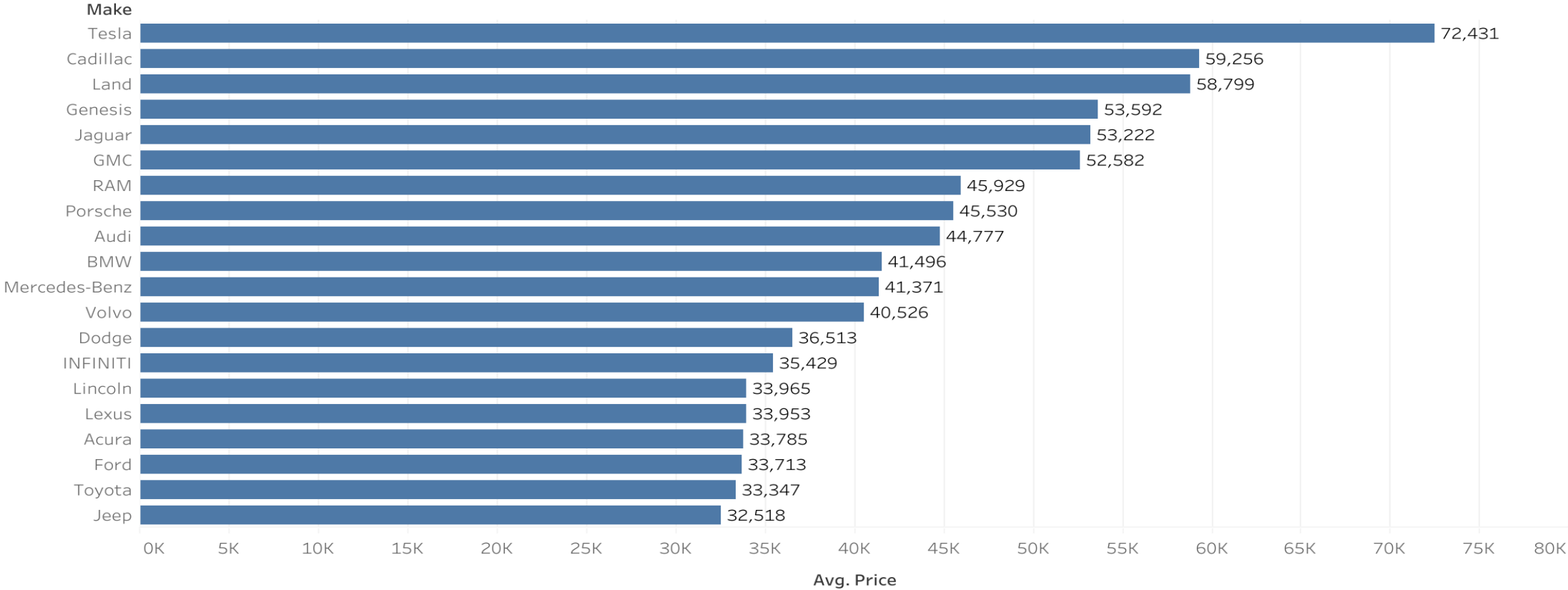


The bar chart provides a comprehensive view of the average distribution of car prices within the top 10 states with the highest prices. By visually comparing the pricing patterns across these regions, it enables the identification of variations in pricing, thereby providing insights into the relative affordability of cars in different areas within the top 10. It's important to note that this chart represents a snapshot of average prices and does not depict pricing trends over a specific period of time.

Data Visualizations 2

Bar Chart - Which 20 Car Models have the highest average price?

Bar Chart: Top 20 Car Models with the Highest Average Prices



The bar chart visually represents the average prices of the top 20 car models, offering insights into pricing trends and model popularity. This visualization aids in identifying which car models command higher or lower prices in the market, providing valuable information for consumers and industry stakeholders.

SQL Queries & Results;

1. What is the difference in Average Max mileage (MPG) between used cars with automatic and manual transmissions?

Average Max Mileage Variation by Transmission Type:

```
SELECT Transmission, AVG(MaxMPG) AS AverageMPG
FROM CarsData1
GROUP BY Transmission
```



Transmission	AverageMPG
1-Speed A/T	77.0000
1-Speed Automatic	113.0000
10-Speed Automatic	24.0000
4-Speed Automatic	22.3333
5-SPEED A/T	22.0000
5-Speed Automatic	21.3158
5-Speed Manual	33.5000
6-Speed A/T	29.3333
6-Speed Automatic	27.9861
6-Speed Automatic with Auto-Shift	31.0000
6-Speed Automatic with Overdrive	31.0000
6-Speed M/T	23.0000
6-Speed Manual	29.0000
7-Speed A/T	28.0000
7-Speed Automatic	25.3000
7-Speed Automatic with Auto-Shift	28.5263
8 speed tiptronic s automatic	24.0000
8-SPEED A/T	27.8000
8-Speed Automatic	27.1933
8-Speed Automatic with Auto-Shift	32.6667
9-SPEED A/T	29.0000
9-Speed Automatic	26.6667
9g tronic automatic	29.0000
A/T	30.0000
Automatic	26.8696

2. What are the three most common interior colors in the used cars?

Top 3 Interior Colors Among Used Cars:

```
SELECT InteriorColor, COUNT(*) AS Count
FROM CarsData1
GROUP BY InteriorColor
ORDER BY Count DESC
Limit 3;
```



InteriorColor	Count
Black	205
Gray	38
Beige	17

Data Visualizations 3 and SQL

Line Chart -How has the price trend evolved for Acura over the past ten years?

```
SELECT Year, AVG(CAST(REPLACE(Price, '$', '') AS DECIMAL(10,2))) AS Average_Price
```

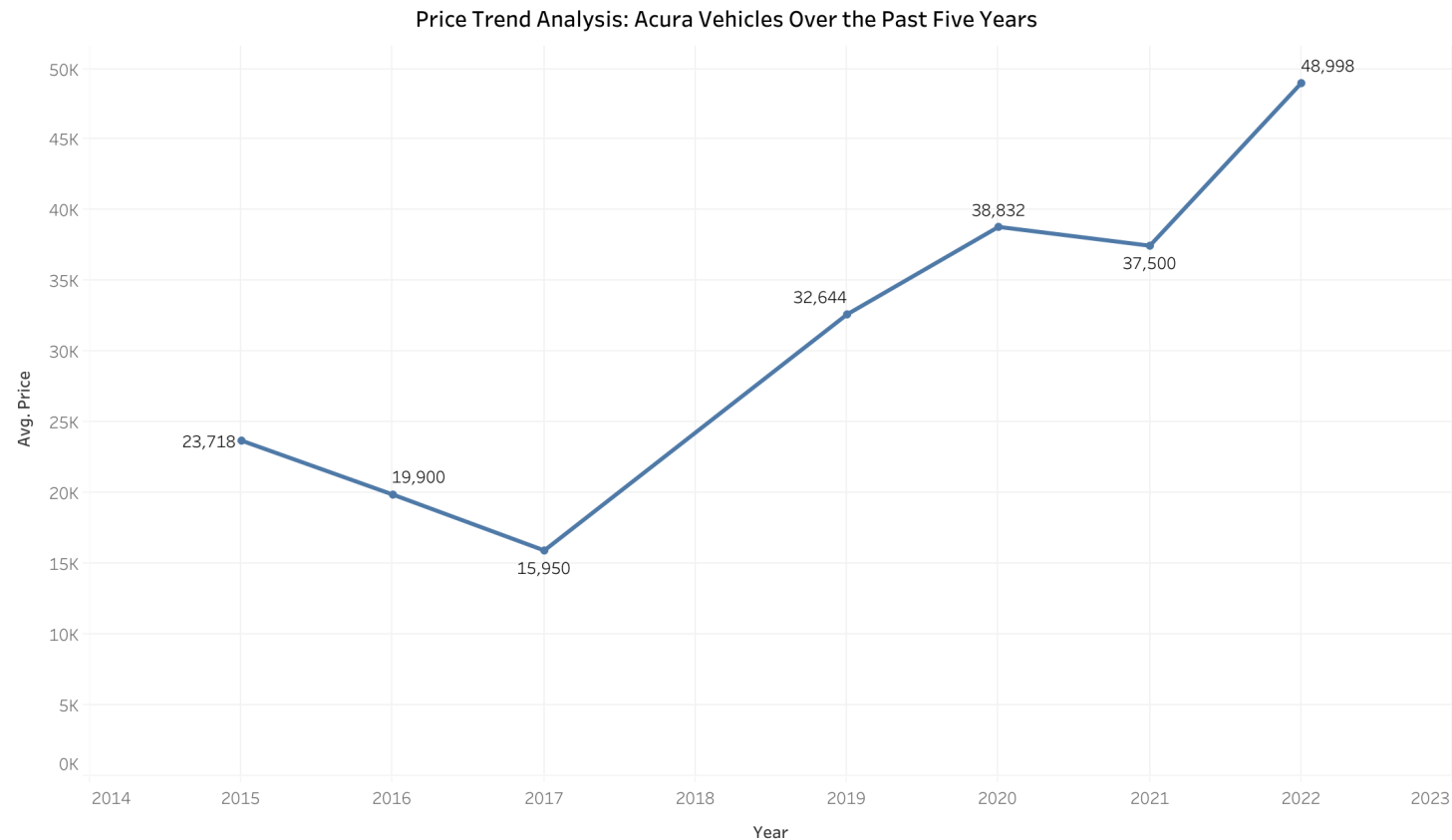
```
FROM CarsData1
```

```
WHERE Make = 'Acura' AND Year >= YEAR(CURRENT_DATE()) - 7
```

```
GROUP BY Year
```

```
ORDER BY `CarsData1`.`Year`
```

Year	Average_Price
2022	48.000000
2021	37.500000
2020	38.000000
2019	32.000000
2017	15.000000



This visualization offers a clear depiction of how Acura vehicle prices have changed over time, aiding in understanding pricing patterns and trends.

Additional Research Questions

- Which cars of states have sold the most per capita? (find population)
- What's the difference in price between cars with automatic and manual transmissions?
- How does the price of front-wheel drive vehicles compare to that of all-wheel drive vehicles?
- How has the average price of vehicles changed over the years?

Conclusion

- What We Learned:
 - We've been exploring a wide variety of used car data, discovering insights into their features, price and more.
 - Identified trends in fuel types, transmission type and mileage variations across a number of vehicle models.
- Challenges Faced:
 - To ensure accuracy and consistency, the dataset is cleaned up and standardized.
 - In the case of variables such as fuel type and transmission, addressing missing data and inconsistencies.
- Future Steps:
 - Further analysis on the correlation between car features and prices.
 - Research into regional trends in car preference and price.
 - In order to analyse the automotive market in greater detail, it is necessary to integrate further data sets.

The End

Thank you!

