2024 Vehicle Data Set

A Data Cleaning Project

06/20/2024

Contents

[The Vehicle Dataset 2024 2](#_Toc169792152)

[Dataset Variables 2](#_Toc169792153)

[Data Cleaning 2](#_Toc169792154)

[Cleaning Findings 2](#_Toc169792155)

[Mitigation Methods 2](#_Toc169792156)

[Summary of Outcomes 2](#_Toc169792157)

[Export Clean Dataset 2](#_Toc169792158)

[Limitations 2](#_Toc169792159)

[Impact of Limitations 2](#_Toc169792160)

[Sources 2](#_Toc169792161)

# The Vehicle Dataset 2024

The "Vehicle Dataset 2024" dataset was found through Kaggle.com and is available under the ODC Attribution License. <https://www.kaggle.com/datasets/kanchana1990/vehicle-dataset-2024>

## Dataset Variables

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable Name** | **Data Type** | **Description** | **Variable Example** |
| **name** | Text | The vehicle's full name, including make, model, and trim. | "2024 Jeep Wagoneer Series II" |
| **description** | Text | A brief vehicle description, often including key features and selling points. | "White Knuckle Clearcoat 2023 Dodge Durango Pur…" |
| **make** | Categorical | The vehicle manufacturer (Ford, Toyota, BMW). | "Jeep" |
| **model** | Categorical | The model name of the vehicle. | "Wagoneer" |
| **type** | Categorical | The type of vehicle. (New, Used). | "New" |
| **year** | Continuous | The year the vehicle was manufactured | "2024" |
| **price** | Discrete | The price of the vehicles is in USD. | "74600.0" |
| **engine** | Text | Details about the engine, including type and specifications | "24V GDI DOHC Twin Turbo" |
| **cylinders** | Discrete | The number of cylinders in a vehicle's engine. | "6.0" |
| **fuel** | Categorical | The type of fuel the vehicle uses. (Gasoline, Diesel, Electric) | "Gasoline" |
| **mileage** | Continuous | The vehicle's mileage. | "32.0" |
| **transmission** | Categorical | The transmission type. (Automatic, Manual) | "8-Speed Automatic" |
| **trim** | Categorical | The trim level of the vehicle indicating different feature sets or packages | "Series II" |
| **body** | Categorical | The vehicle's body style. (SUV, Sedan, Pickup Truck) | "SUV" |
| **doors** | Discrete | The number of doors on the vehicle | "4.0" |
| **exterior\_color** | Categorical | The exterior color of the vehicle | "White" |
| **interior\_color** | Categorical | The interior color of the vehicle | "Global Black" |
| **drivetrain** | Categorical | The vehicle's drivetrain. (All-wheel Drive, Front-wheel Drive | "Four-Wheel Drive" |

# Data Cleaning

## Cleaning Findings

## Mitigation Methods

## Summary of Outcomes

## Export Clean Dataset

# Limitations

## Impact of Limitations

# Sources

Kanchana1990. (2024, May 29). Vehicle Dataset 2024. *Kaggle.com*. Retrieved May 31, 2024, from <https://www.kaggle.com/datasets/kanchana1990/vehicle-dataset-2024>