# **Vehicle Dataset Analysis Report**

## **Introduction**

This report presents an analysis of a vehicle dataset containing 10,000 entries across 25 features. The dataset captures vehicle information, GPS coordinates, vehicle status, and other operational details, offering insights into vehicle usage patterns, departmental activity, and operational statuses.

## **1. Dataset Overview**

### **Key Features**

* **Vehicle No:** Unique identifier for each vehicle.
* **Vehicle Name:** Placeholder (all entries are missing).
* **Company:** Department associated with the vehicle.
* **Latitude & Longitude:** GPS coordinates.
* **Branch:** Operating branch of the vehicle.
* **Gpsactualtime & Datetime:** Timestamps for the vehicle's GPS.
* **Vehicletype, Status, Devicemodel, Speed, Imeino:** Vehicle technical and status details.
* **Odometer:** Distance traveled by the vehicle.
* **Driver First Name & Last Name:** Names of the assigned driver (mostly missing).
* **Power, Ignition Status (Ign), External Voltage, Angle:** Indicators of vehicle performance and operation.

## **2. Data Quality Check**

### **Summary Statistics:**

* **Latitude & Longitude:** GPS coordinates span a reasonable range indicating data accuracy (Mean Lat: 19.41, Long: 72.82).
* **Speed:** Ranges from 0 to 90 km/h, with a mean of 5.31 km/h. This suggests many vehicles are either stationary or moving slowly.
* **External Voltage:** Averages around 16.7V, suggesting normal operation of vehicle electronics.

### **NaN (Missing) Values:**

* **Vehicle Name, Driver First Name & Last Name, Immobilize State, SOS, and some other categorical features** have missing or incomplete data.
* All numeric fields (Latitude, Longitude, Speed, Odometer, etc.) are complete with no NaN values.

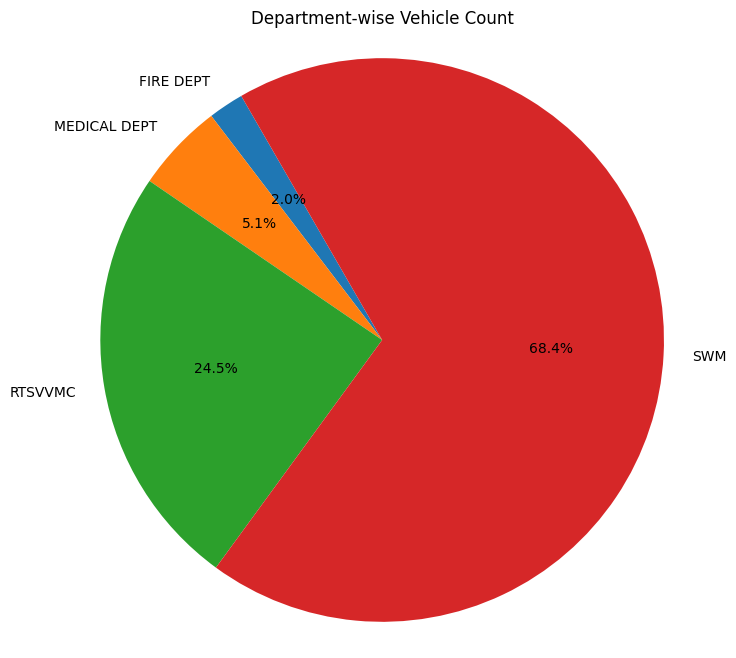
### **Duplicates:**

Duplicates have been removed to ensure data quality. No rows with all NaN values were found.

## **3. Feature Exploration**

### **3.1. Department-wise Vehicle Distribution**

Analyzing the distribution of vehicles across different departments reveals the following:



|  |  |
| --- | --- |
| **Department** | **Vehicle Count** |
| FIRE DEPT | 501 |
| MEDICAL DEPT | 951 |
| RTSVVMC | 2896 |
| SWM | 5652 |

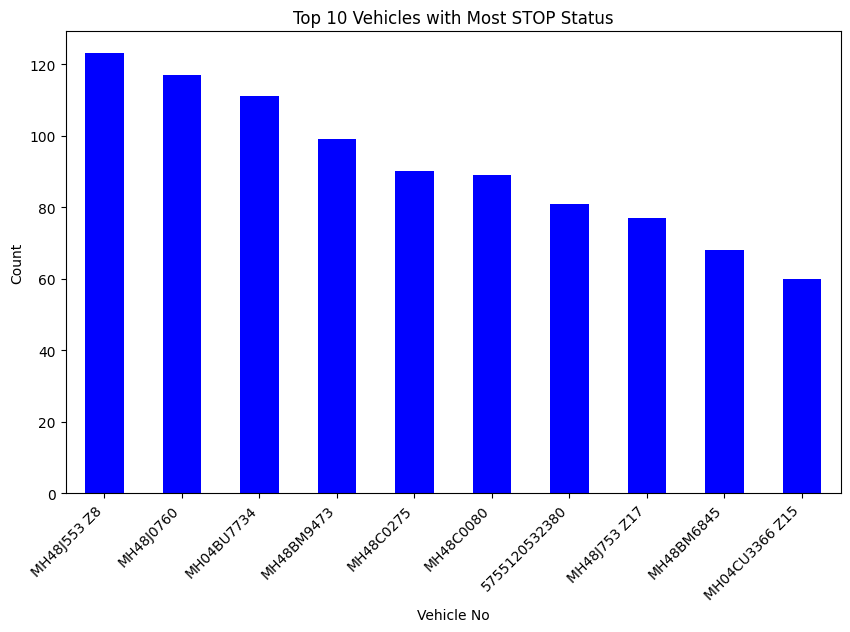
* **Observation:** The SWM (Solid Waste Management) department has the highest number of vehicles, while the FIRE DEPT has the least.

### **3.2. Vehicle Status Analysis**

The dataset includes vehicle status indicators such as **STOP**, **RUNNING**, **Ignition ON**, and **Ignition OFF**.

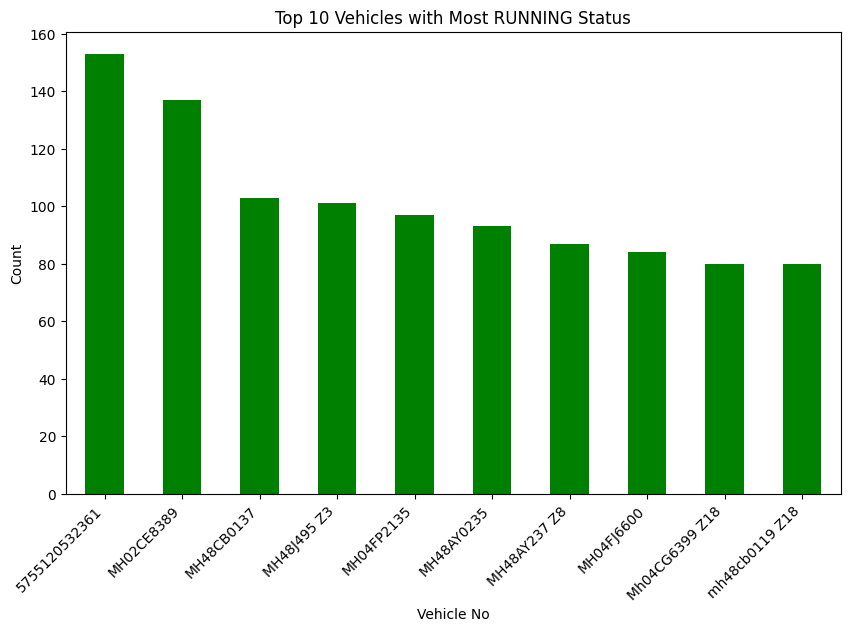
#### **Top 10 Vehicles with Most STOP Status**

* **Vehicle with most STOP status:** MH48J553 Z8 with 123 times
* Below is a bar plot visualizing the top 10 vehicles with the most STOP status:



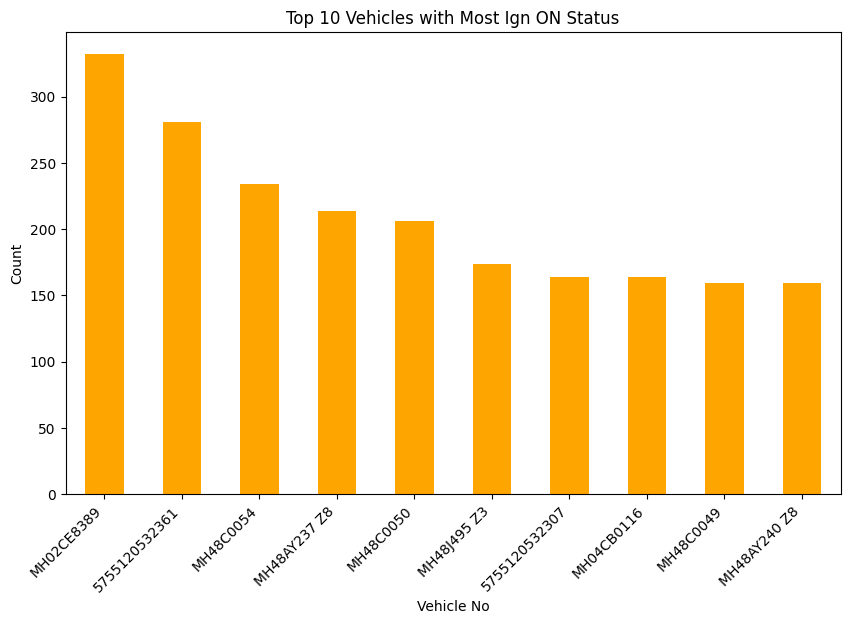
#### **Top 10 Vehicles with Most RUNNING Status**

* **Vehicle with most RUNNING status:** MH02CE8389 with 137 times (its second as first is just a number)
* Below is a bar plot visualizing the top 10 vehicles with the most RUNNING status:

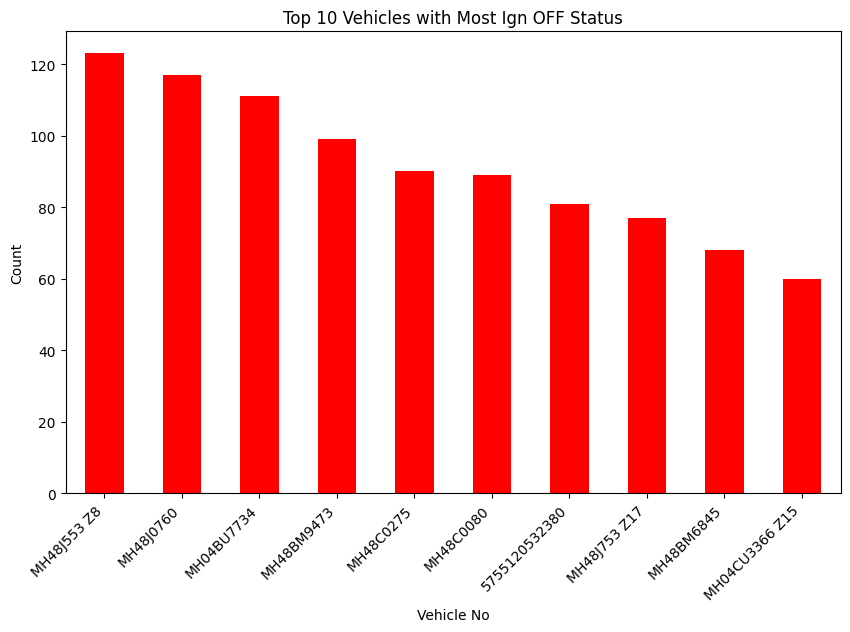


#### **Ignition Status Analysis**

* **Vehicle with most Ignition ON:** MH02CE8389 with 332 times



* **Vehicle with most Ignition OFF:** MH48J553 Z8 with 123 times



### **3.3. Most and Least Frequent Vehicles**

* **Most frequent vehicle:** MH02CE8389

This vehicle appears most often in the dataset.

* **Least frequent vehicle:** MH48CB1642

This vehicle appears the least in the dataset.

## **4. Conclusions & Insights**

* **Vehicle Usage:** The **SWM** department manages the most vehicles, followed by **RTSVVMC**. This likely reflects the nature of SWM's operations, requiring a larger fleet.
* **STOP vs. RUNNING:** Vehicle MH48J553 Z8 stops frequently 123 times, while MH02CE8389 runs frequently 137 times, indicating patterns of usage that may require further investigation.
* **Ignition Status:** Ignition patterns reflect vehicle usage trends, with some vehicles being regularly powered on (MH02CE8389).