



Fleet Vehicle Maintenance Analysis

Don Wilson
Vikash Kumar

Problem Statement



The company has some information on past failures, as well as on fleet usage and maintenance.

As the data team, we are here to offer a data-driven approach.

More specifically, we want to use the information we have to answer the following questions:

- **What are the most common factors behind these failures?**
- **Which fleets are most likely to fail?**

Data Set Overview



- **Usage_Data**

Information about how the fleets have been used in the past, as measured by the number of miles collected at certain points in time

- **Maintenance_Part_Consumption**

Stores which parts were serviced or replaced, by asset, and the reason

- **Maintenance_Failure**

Indicates whether or not a part had a failure (not always available)

Data Transformation



- **What?**

The process of converting data from one format or structure into another format or structure

- **Why?**

Reduce size - Raw data is often too lengthy and difficult to analyze

High Level View - Transformed Data gives more clarity on general behavior and aspects of data

Comparative Study- enables easy collaboration with related Data Structures

- **How?**

Aggregation - reduce occurrences aggregating with sum, mean, count, first, last etc.

Data Cleaning -Duplicate Removal, Missing Value Treatment, Outlier dealing

Feature Engineering - extracts more information from existing data

Transformation of Usage Data Table

Raw Data Structure

shape: 22513 * 3

	Asset	Time	Use
0	A403193	5	31194.65
1	A403193	17	31223.54
2	A403193	56	31362.71
3	A403193	124	31701.44
4	A403193	144	31724.49

Cons of Raw Data

- All observations of all Assets are present, which can't make any insight

Pros of Transformed Data

- Each Asset has only one entry- hence high level view of the Asset is obtained
- Age & Distance of first & last maintenances – gives view of usage & durability of the Asset
- Average Distance & Time between services indicates the performance

Transformed Data Structure

shape: 1894*10

Asset	Age_initial	Age_lastknown	Fleet_serviced_count	Distance_initial	Distance_lastknown	Time_in_service	Distance_service	Distance_by_fleet	Time_by_fleet
000204	46	722	15	31449.65	33212.76	676	1763.11	117.540667	45.066667
000270	435	617	3	26378.56	27036.77	182	658.21	219.403333	60.666667
000463	30	616	18	30451.54	31894.00	586	1442.46	80.136667	32.555556
000495	6	695	16	30851.25	32540.25	689	1689.00	105.562500	43.062500

Removal of Time Column from Maintenance Part Consumption ?

- Maintenance_Part_Consumption has repeated entry of Time details present in Usage_Data
- Multiple Time values for multiple entries is not desirable, but Service level only required
- Some Time values are even slightly mismatching with Usage_Data, which is not going to make any impact in analysis

Owing to these reasons Time Column is Removed

- **Example:** Distribution & Interdependency of Time for Asset A891025 in both tables

Usage_data

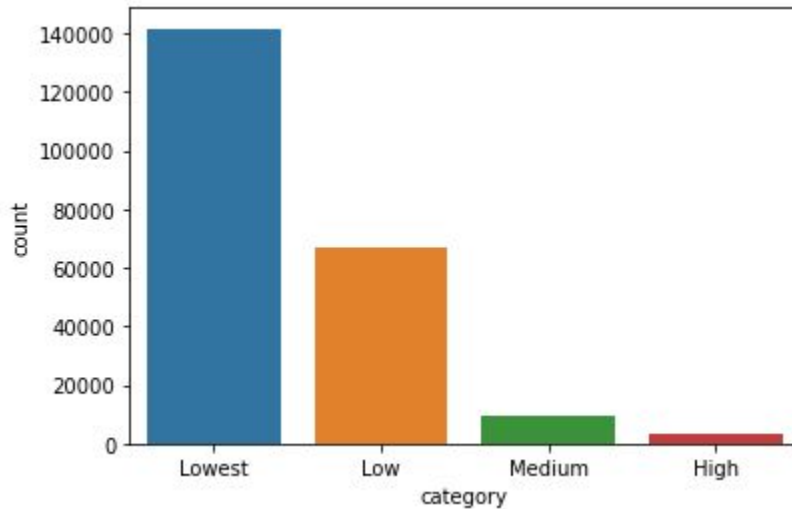
Asset	Time	Use
A891025	620	25852.80
A891025	650	25951.13
A891025	660	25965.42

Maintenance_part_consumption

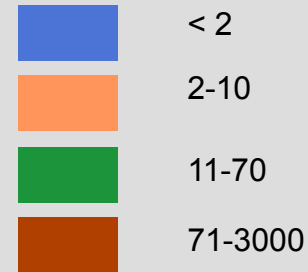
Asset	Time	Reason	Part	Quantity
A891025	620	R193	P780033	1
A891025	649	R193	P838794	1
A891025	669	R193	P476637	3
A891025	669	R193	P977544	1
A891025	670	R193	P467150	1
A891025	670	R193	P413461	1

Distribution of the quantity of parts

- More than 60% of the cases, only 1 quantity of a Part are used
- Negligible cases which require more than 1000 quantities of a Part

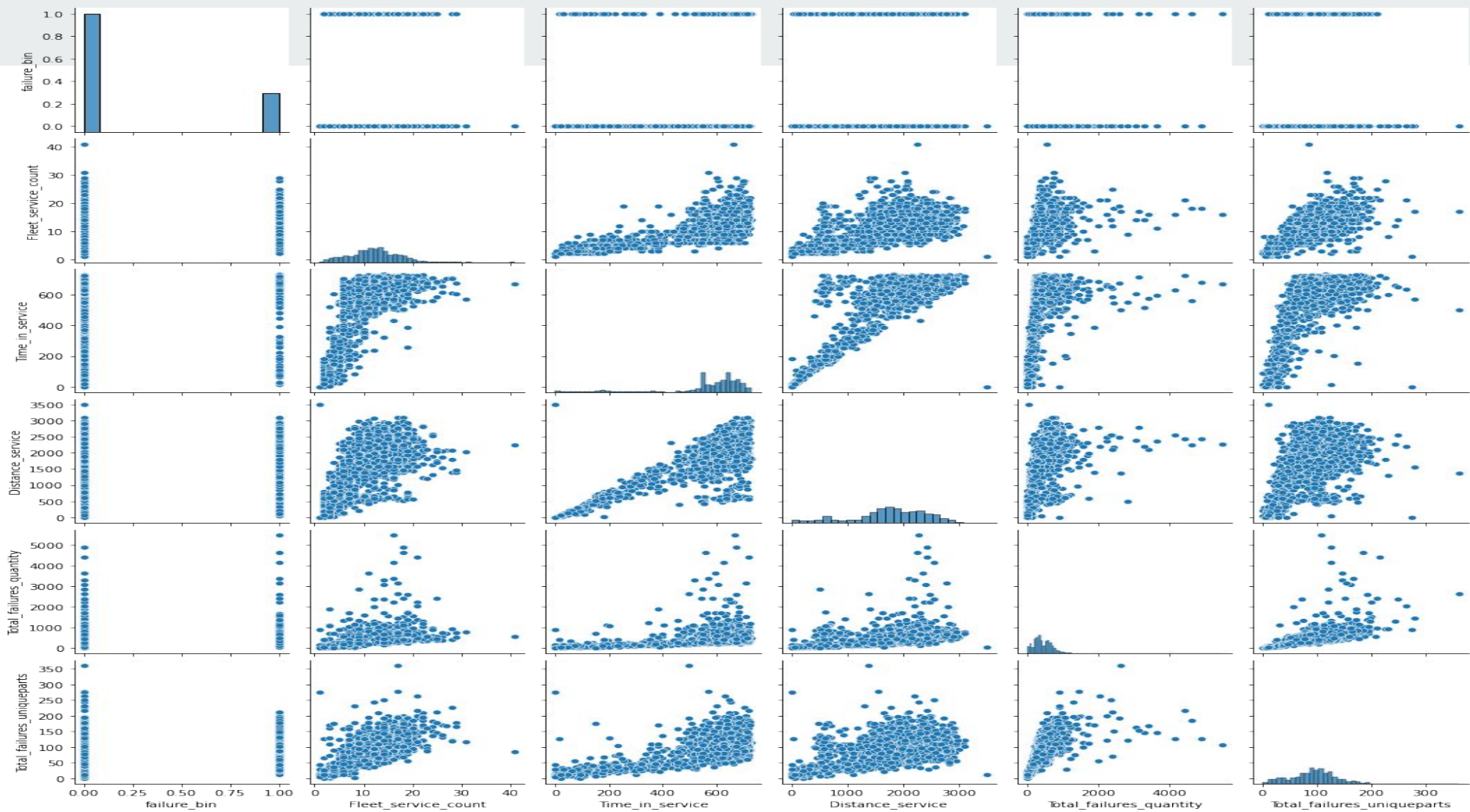


Quantity of Parts

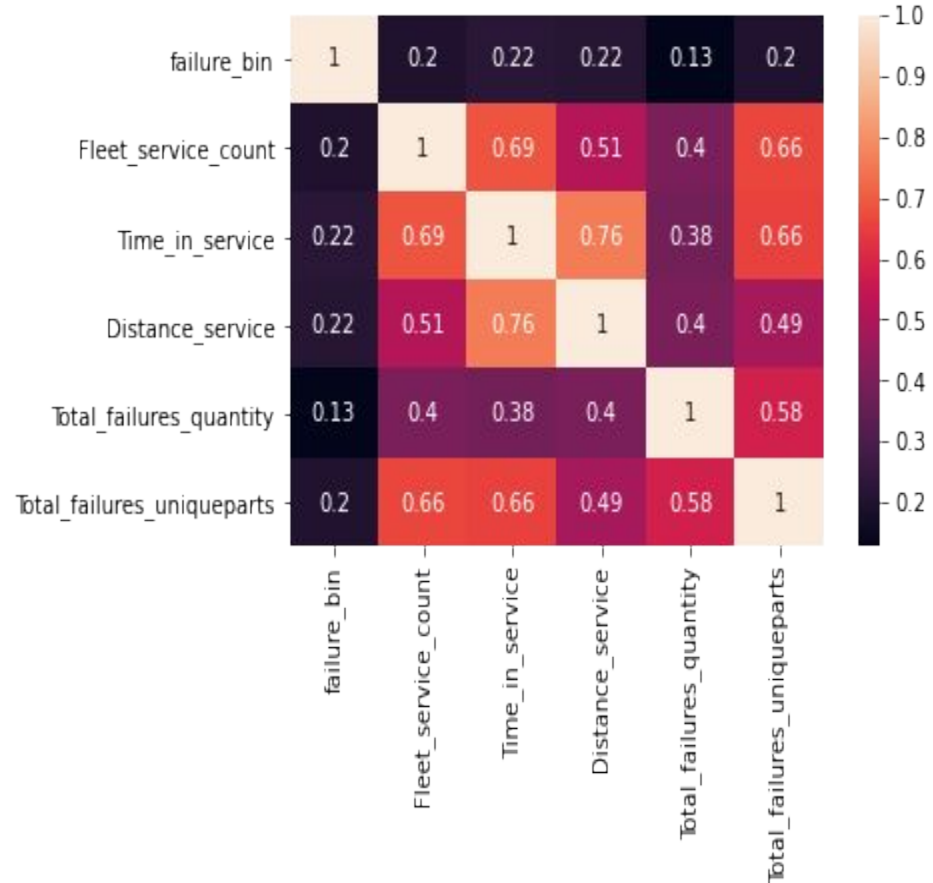




**Perform Bivariate analysis on Time_in_service,
Distance_service, Fleet_serviced_count,
Total_failures_quantity,
Total_failures_uniqueparts**



- Distance service and time in service is highest correlated.
- Time in Service is highly correlated with fleet service count and total failures in unique parts .
- Fleet service count is highly correlated to total Failures in unique part.



Conclusion



- R193, R565, R707 are the top three reason for failures.
- R193 and R565 are the top two reason of failure which require highest unique parts.
- A998987 is the asset which spend more effort on maintenance & and due to major cause R193
- Majority of time change in parts was not required just maintenance occur.
- Asset which had a breakdown require more average quantity of Unique parts.
- The more distance travelled by assets the more time spent on maintenance, It also mean that more no of unique parts will be required.

Thank You