SUMMER INTERNSHIP REPORT

Submitted By

MARISHWARAN A (211722118034)

In partial fulfillment for the award of the degree

Of

BACHELOR OF TECHNOLOGY

In

Computer and Communication Engineering

RAJALAKSHMI INSTITUTE OF TECHNOLOGY

CHENNAI 600 124

ANNA UNIVERSITY

CHENNAI: 600 025

SEPTEMBER 2025

ANNA UNIVERSITY

CHENNAI:600 025

BONAFIDE CERTIFICATE

Certified that this summer internship report **Data Analytics (Power BI)** is the bonafide work of "**MARISHWARAN A - 211722118034** " who carried out the internship work.

SIGNATURE

DR.E.GANESH Ph.D,

HEAD OF THE DEPARTMENT

Computer and Communicatioin Engineering

Rajalakshmi Institute of Technology

Kuthambakkam

Chennai 600124

SIGNATURE

Mrs.SUSHMITHA S M.E ,Ph.D,

INTERNSHIP CO ORDINATOR

Computer and Communication Engineering

Rajalakshmi Institute of Technology

Kuthambakkam

Chennai 600124

CERTIFICATE:

CLOUD COUNSELAGE
IT & MANAGEMENT CONSULTING & SERVICES

EXPERIENCE LETTER

To Whomsoever It May Concern

Date: 10-May-2025

This letter is to certify that **Marishwaran A** has successfully completed internship with **Cloud Counselage Pvt. Ltd.** under our '**Internship Program'** in association with our **Gift-A- Career Foundation** for a duration of **240 hours.**

During this engagement, **Marishwaran A** has abided by the company policies, attended various industry-specific training sessions, and successfully submitted **Data Analytics(powerBI)** project deliverables by following the best practices and project management practices.

Throughout the internship **Marishwaran A** has worked ethically, followed the instructions, performed necessary research, and worked with minimum supervision. The delivery of the project demonstrates their domain knowledge and skills, a structured approach to problem solving, ability to follow instructions, and inclination to work hard.

This association has been beneficial for us, and we wish **Marishwaran A** all the success in their future endeavours.

For **CLOUD COUNSELAGE PVT. LTD.,**

CLOUD COUNSELAGE PRIVATE LIMITED MUMBAI

Subhi Shildhankar

Subhi Shildhankar
Co-Founder & Director (HR)

ACKNOWLEDGEMENT

I am personally indebted to many who had helped me during the course of this internship work. My deepest gratitude to the **God Almighty.**

I greatly and profoundly thankful to our beloved Chairperson **Dr. (Mrs.) Thangam Meganathan**, **M.A., M.Phil., Ph.D.,** and Chairman **Thiru. S. MEGANATHAN, B.E., F.I.E.,** for facilitating me with this opportunity. My sincere thanks to our respected Director **Dr. R. Sundar. Ph.D.,** for his consent to take up this internship work and make it a great success.

I am also thankful to our Principal **Dr. R. Maheswari. Ph.D.,** for his never-ending encouragement that drives me towards innovation.

I am thankful to our Head of the Department ***DR.E.GANESH*,** and all the faculty members for their valuable teachings and suggestions.

From the bottom of my heart with profound reference and high regards, I would like to thank **Subhi Shildhankar** who have been the pillar of this project without whom I could not have been able to complete the project successfully.

I hereby declare that all of the information that I have provided here is correct to the best of my knowledge

ABSTRACT

TITLE:Vehicle data analysis using Power BI

- In our recent project, we undertook a comprehensive analysis of vehicle data on behalf of ZF Company. The initial phase involved the systematic collection of data from the Vahan portal, a reliable source for vehicle-related information. This data compilation served as the foundational step in our analytical journey, providing us with a diverse and extensive dataset to work with.

- Following the data acquisition, our focus shifted to the critical stage of data transformation. This process involved refining and structuring the raw data to ensure its compatibility with the analytical tools and systems we intended to employ. Through meticulous cleaning, formatting, and integration, we aimed to enhance the overall quality and utility of the dataset, setting the stage for meaningful insights.

- The primary objective assigned to our team was the development of a comprehensive dashboard based on the processed vehicle data. This dashboard was envisioned to serve as a user-friendly interface, offering ZF Company a consolidated and visual representation of key metrics and trends derived from the analyzed data. By leveraging advanced visualization techniques, our goal was to empower decision-makers within the company with actionable insights, facilitating informed strategic decisions and optimizing operational efficiency.

- The resultant dashboard not only presented real-time snapshots of critical metrics but also facilitated trend analysis over time, providing a dynamic tool for continuous monitoring and strategic planning. Ultimately, our endeavor was not merely to present data visually but to empower ZF Company with a robust analytical toolset that translates data into actionable intelligence, fostering agility and adaptability in an ever-evolving automotive landscape.

- In conclusion, our multi-faceted approach encompassed data collection, transformation, and the ultimate creation of a tailored dashboard to meet the specific needs of ZF Company. The amalgamation of these stages aimed to provide a holistic and streamlined solution, unlocking the potential within the gathered vehicle data for enhanced business intelligence and strategic planning.

S.no	CONTENTS	Page no

Learning Objectives/Internship Objectives

- **Mastery of Data Collection Techniques**:
  Develop proficiency in extracting and collecting relevant vehicle data from authoritative sources, with a specific focus on the Vahan portal.

- **Data Transformation and Cleaning Skills:**
  Acquire hands-on experience in transforming raw data into a structured and usable format through comprehensive cleaning, formatting, and integration processes.

- **Dashboard Development Proficiency:**
  Gain expertise in creating dynamic dashboards that effectively communicate complex vehicle data trends and metrics, using visualization tools to enhance accessibility and user experience.

- **Advanced Analytics Competence:**
  Acquire a deep understanding of advanced analytics techniques, such as trend analysis, anomaly detection, and correlation identification, to extract meaningful insights from diverse vehicle data dimensions.

- **Professional Application of Analytical Tools:**
  Familiarize oneself with industry-standard analytical tools and technologies, applying them proficiently to analyze and visualize vehicle data for practical business applications in the automotive sector.

- **Business Intelligence Application:**
  Understand how to apply business intelligence principles to enhance organizational decision-making processes, leveraging the insights derived from vehicle data to drive informed and effective strategies.

- **Continuous Monitoring and Adaptability:**
  Learn to design and implement systems for continuous monitoring of key metrics, fostering adaptability and agility in response to evolving trends in the automotive landscape.

PROJECT FY 23-TIMELINE

Week No	Date	Activities
Week_1	11/4/2023	First meeting with ZF-discussed about vahan portal
Week_1	12/4/2023	State wise individual data collection-2 States per day
Week_1	13/04/2023	State wise individual data collection-2 States per day
Week_1	17/04/2023	State wise individual data collection-2 States per day
Week_2	18/04/2023	State wise individual data collection-2 States per day
Week_2	19/04/2023	State wise individual data collection-2 States per day
Week_2	20/04/2023	State wise individual data collection-2 States per day
Week_2	21/04/2023	State wise individual data collection-2 States per day
Week_2	22/04/2023	State wise individual data collection-2 States per day
Week_2	24/04/2023	State wise individual data collection-3 States per day
Week_3	25/04/2023	State wise individual data collection-3 States per day
Week_3	26/04/2023	State wise individual data collection-3 States per day
Week_3	27/04/2023	State wise individual data collection-3 States per day
Week_3	28/04/2023	State wise individual data collection-3 States per day
Week_3	29/04/2023	State wise individual data collection-3 States per day
Week_4	3/5/2023	Second meeting with ZF-discussed about data structure
Week_4	8/5/2023	Second meeting with ZF-discussed about combined format
Week_5	9/5/2023	Transformation of individual data into combined dataset-1 State per day
Week_5	10/5/2023	Transformation of individual data into combined dataset-1 State per day
Week_5	11/5/2023	Transformation of individual data into combined dataset-1 State per day
Week_5	12/5/2023	Transformation of individual data into combined dataset-1 State per day

Week_5	13/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_5	15/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	16/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	17/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	18/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	19/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	20/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_6	22/05/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	20/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	21/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	22/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	23/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	24/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_7	26/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	27/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	28/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	29/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	30/06/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	1/7/2023	Transformation of individual data into combined dataset-1 State per day
Week_8	3/7/2023	Transformation of individual data into combined dataset-1 State per day
Week_9	11/7/2023	Transformation of individual data into combined dataset-1 State per day
Week_9	12/7/2023	Transformation of individual data into combined dataset-1 State per day
Week_9	13/07/2023	Transformation of individual data into combined dataset-2 State per day
Week_9	14/07/2023	Transformation of individual data into combined dataset-2 State per day

Week_9	15/07/2023	Transformation of individual data into combined dataset-2 State per day
Week_9	17/07/2023	Transformation of individual data into combined dataset-2 State per day
Week_10	18/07/2023	State wise(RTO) individual data collection-5 States per day
Week_10	19/07/2023	State wise(RTO) individual data collection-5 States per day
Week_10	20/07/2023	State wise(RTO) individual data collection-6 States per day
Week_10	21/07/2023	State wise(RTO) individual data collection-6 States per day
Week_10	22/07/2023	State wise(RTO) individual data collection-6 States per day
Week_10	24/07/2023	State wise(RTO) individual data collection-6 States per day
Week_11	25/07/2023	Sample Dashboard Creation
Week_11	26/07/2023	Sample Dashboard Creation
Week_11	27/07/2023	Sample Dashboard Creation
Week_12	18/08/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_12	19/08/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_12	21/08/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_13	22/08/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_13	23/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_13	24/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_13	25/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_13	26/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_13	28/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_14	29/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_14	30/08/2023	Transformation of individual data into combined dataset(RTO)-2 State per day
Week_14	31/08/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_14	1/9/2023	Transformation of individual data into combined dataset(RTO)-1 State per day

Week_14	2/9/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_14	4/9/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_15	20/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_15	21/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_15	22/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_15	23/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_15	25/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_16	26/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_16	27/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_16	29/09/2023	Transformation of individual data into combined dataset(RTO)-1 State per day
Week_16	30/09/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_16	3/10/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_16	4/10/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_17	5/10/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_17	6/10/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_17	7/10/2023	Transformation of individual data into combined dataset(RTO)-1/2 State per day
Week_17	9/10/2023	Data visualization and Insights
Week_17	10/10/2023	Data visualization and Insights

1. INTRODUCTION

1. Introduction:

The contemporary automotive landscape is undergoing transformative shifts driven by technological advancements and the growing emphasis on data-driven decision-making. In this context, our project, conducted in collaboration with ZF Company, focuses on the comprehensive analysis of vehicle data sourced from the Vahan portal. The Vahan portal serves as a rich repository of diverse vehicle-related information, presenting an opportune gateway to unravel insights that can shape strategic decisions within the automotive industry.

As technology continues to redefine the automotive sector, the ability to harness and interpret data becomes paramount. ZF Company, a prominent player in the automotive industry, recognizes the potential of data-driven insights for optimizing operations, enhancing product offerings, and staying competitive in a dynamic market. This project seeks to bridge the gap between raw data and actionable intelligence, culminating in the development of a dynamic dashboard tailored to meet the specific needs of ZF Company. The synthesis of data analysis and visualization is anticipated to empower decision-makers within the organization, facilitating a proactive approach to challenges and opportunities within the automotive ecosystem.

1.1 Problem Statement:

In terms of data analysis, the scope extends beyond mere aggregation, exploring diverse dimensions such as usage patterns, geographical variations, and performance metrics. The project aims to provide a holistic view of the automotive landscape, empowering ZF Company with actionable intelligence derived from a granular examination of the data.

1.2 Scope of this project:

The significance of this project lies in its potential to revolutionize how ZF Company leverages vehicle data for strategic decision-making. A streamlined dashboard, infused with insightful visualizations, can serve as a catalyst for informed choices, driving operational efficiency and fostering innovation. By distilling complex data into accessible intelligence, the project is poised to enhance ZF Company's competitiveness in a rapidly evolving automotive market.

Furthermore, the project's significance extends to its contribution to the broader discourse on the integration of data analytics in traditional industries. The lessons learned from this endeavor can serve as a blueprint for other companies navigating the intersection of data and automotive manufacturing, fostering a culture of continuous improvement and adaptability.

1.3 Significance:

Furthermore, the project's significance extends to its contribution to the broader discourse on the integration of data analytics in traditional industries. The lessons learned from this endeavor can serve as a blueprint for other companies navigating the intersection of data and automotive manufacturing, fostering a culture of continuous improvement and adaptability.

1.4 Objective:

The overarching objective of this project is the development of a dynamic dashboard that encapsulates the key insights derived from the analysis of vehicle data sourced from the Vahan portal. This includes mastering data collection techniques, honing data transformation and cleaning skills, and applying advanced analytics to unravel nuanced trends within the dataset. The specific objectives include:

Efficient Data Collection: Develop proficiency in extracting and collecting relevant vehicle data from the Vahan portal, ensuring a comprehensive dataset for analysis.

Data Transformation: Acquire hands-on experience in transforming raw data into a structured and usable format through systematic cleaning, formatting, and integration processes.

Dashboard Development: Gain expertise in creating a user-friendly dashboard that visually communicates critical metrics and trends derived from the analyzed vehicle data.

Advanced Analytics: Develop competence in applying advanced analytics techniques, such as trend analysis, anomaly detection, and correlation identification, to extract meaningful insights from the dataset.

Strategic Decision Support: Enable decision-makers at ZF Company with actionable intelligence, translating data insights into informed strategies for operational and strategic planning.

1.5 Research Design:

The research adopts a mixed-methods approach, combining quantitative analysis of vehicle data with qualitative assessments of dashboard usability. The data collection phase involves systematic extraction from the Vahan portal, followed by comprehensive transformation to ensure compatibility with analytical tools. The development of the dashboard incorporates user feedback and iterative design processes to optimize user experience and information accessibility.

The research design process for vehicle data analysis involves a structured sequence of steps to ensure that the data collected from vehicles can be transformed into meaningful insights for decision-making. The process begins with problem identification, where the objective of the study is clearly defined, such as monitoring driving behavior, predicting maintenance needs, or optimizing fuel consumption. Once the research objective is fixed, the next step is to **formulate research questions and hypotheses**. For example, a hypothesis could be that "higher engine temperature readings correlate with increased fuel consumption."

The second stage involves **data collection design**. Vehicle data can be acquired from various sources including On-Board Diagnostics (OBD-II) ports, Internet of Things (IoT) sensors, GPS modules, and telematics systems. At this stage, decisions are made regarding the type of data (speed, RPM, fuel usage, braking patterns, location), the sampling frequency, and the method of data transmission (wired logging or wireless streaming to the cloud). Care is taken to ensure the reliability, accuracy, and security of collected data.

The third stage is **data preparation**, where raw vehicle data is cleaned and pre-processed. This includes handling missing values, filtering noise, synchronizing sensor timestamps, and converting signals into usable formats. For example, GPS coordinates may be converted into speed or distance metrics, while sensor voltages may be translated into engineering units. This step ensures consistency and improves the quality of subsequent analysis.

Next comes **data analysis and model design**. Various analytical methods such as descriptive statistics, time-series analysis, clustering, regression, or machine learning algorithms are applied depending on the research goals. For instance, predictive models can be trained to forecast component failures, while clustering methods can classify driver behavior patterns. The choice of analytical technique depends on the research questions, data type, and expected outcomes.

The fifth stage is **interpretation and validation**. The results obtained are compared with the initial hypotheses and validated through cross-testing, benchmarking, or expert evaluation. This ensures that the patterns or predictions identified are both statistically significant and practically relevant.

Finally, the research design process ends with **reporting and implementation**, where findings are documented, visualized using graphs or dashboards, and recommendations are provided. These insights can be used by manufacturers for product improvement, by fleet managers for operational efficiency, or by drivers for safety enhancement.

1.6 Limitation:

Despite the comprehensive nature of this research, certain limitations exist. The accuracy and reliability of insights heavily depend on the quality of data available from the Vahan portal. External factors such as changes in regulatory frameworks or unexpected shifts in the automotive industry may impact the relevance of certain findings. Additionally, the scope of the dashboard may be constrained by resource limitations and technological constraints.

The transformation of heterogeneous data into structured formats involves assumptions and approximations, which can introduce errors. Derived features such as fuel efficiency or acceleration may not always capture real-world conditions accurately due to variations in vehicle type, road conditions, and driving behavior.

Processing and analyzing large volumes of vehicle data requires significant computational resources. In cases where high-frequency sensor data is collected continuously, the system may face challenges in real-time analysis due to processing overhead, memory constraints, or network latency in cloud integration.

Vehicle data, particularly when linked with locations, poses risks to user privacy. Without robust encryption and access control, there is a possibility of sensitive data being exposed or misused. The project may not fully address advanced security protocols, limiting its deployment in real-world scenarios.

The absence of benchmark datasets for comparison may hinder the validation of the system's accuracy. Moreover, real-world testing may be limited by time, resources, or access to diverse driving conditions, affecting the reliability of the evaluation.

2. DATA COLLECTION

2.1 Data Collection from Vahan Portal:

For the ambitious vehicle data analysis project undertaken for ZF Company, the initial phase involved the meticulous collection of data from the Vahan portal. This authoritative repository provides a wealth of information related to vehicles, serving as a foundational resource for our comprehensive study.

2.2 Categorization of Vehicles:

In shaping the analytical framework, we strategically categorized vehicles into distinct segments, including goods vehicles, public service vehicles, and four-wheelers. This segmentation provides a nuanced understanding of the diverse automotive landscape, allowing for targeted analysis and insights generation.

2.3 Subcategories for In-Depth Analysis:

Within each major category, a granular approach was adopted through the inclusion of subcategories such as LMV (Light Motor Vehicles), MMV (Medium Motor Vehicles), HMV (Heavy Motor Vehicles), LPV (Light Passenger Vehicles), MPV (Medium Passenger Vehicles), HPV (Heavy Passenger Vehicles), LGV (Light Goods Vehicles), MGV (Medium Goods Vehicles), HGV (Heavy Goods Vehicles), and more. This detailed subcategorization ensures a comprehensive examination of various vehicle types, facilitating a nuanced analysis.

2.4 State-Wise and RTO-Wise Data Collection:

To provide a holistic perspective, data collection extended beyond a mere national overview. Both state-wise and RTO (Regional Transport Office)-wise data were meticulously gathered, encompassing all Indian states and union territories. This multi-dimensional approach allows for a deep dive into regional variations, facilitating insights into localized trends and patterns.

2.5 Temporal Scope: 2012-2022:

The temporal scope of the data collection spans a decade, from 2012 to 2022. This extensive timeframe enables the identification of long-term trends, cyclical patterns, and the impact of significant events or policy changes on the automotive landscape over the years.

2.6 Data Enrichment for Comprehensive Analysis:

Beyond the raw numbers, efforts were directed towards data enrichment. This included incorporating additional contextual information, such as demographic factors, economic indicators, and regulatory changes. By enriching the dataset, the analysis gains depth, allowing for a more robust interpretation of the trends observed.

2.7 Ensuring Data Quality and Accuracy:

Recognizing the critical importance of data quality, stringent measures were implemented to ensure accuracy and reliability. Data validation processes were employed to identify and rectify discrepancies, guaranteeing the integrity of the dataset used for subsequent analysis.

2.8 Ethical Considerations:

In the pursuit of data collection, ethical considerations were paramount. Adherence to privacy regulations and responsible data handling practices were rigorously maintained to uphold the confidentiality and integrity of the information sourced from the Vahan portal.

DATA COLLECTION FROM VAHAN PORTAL:

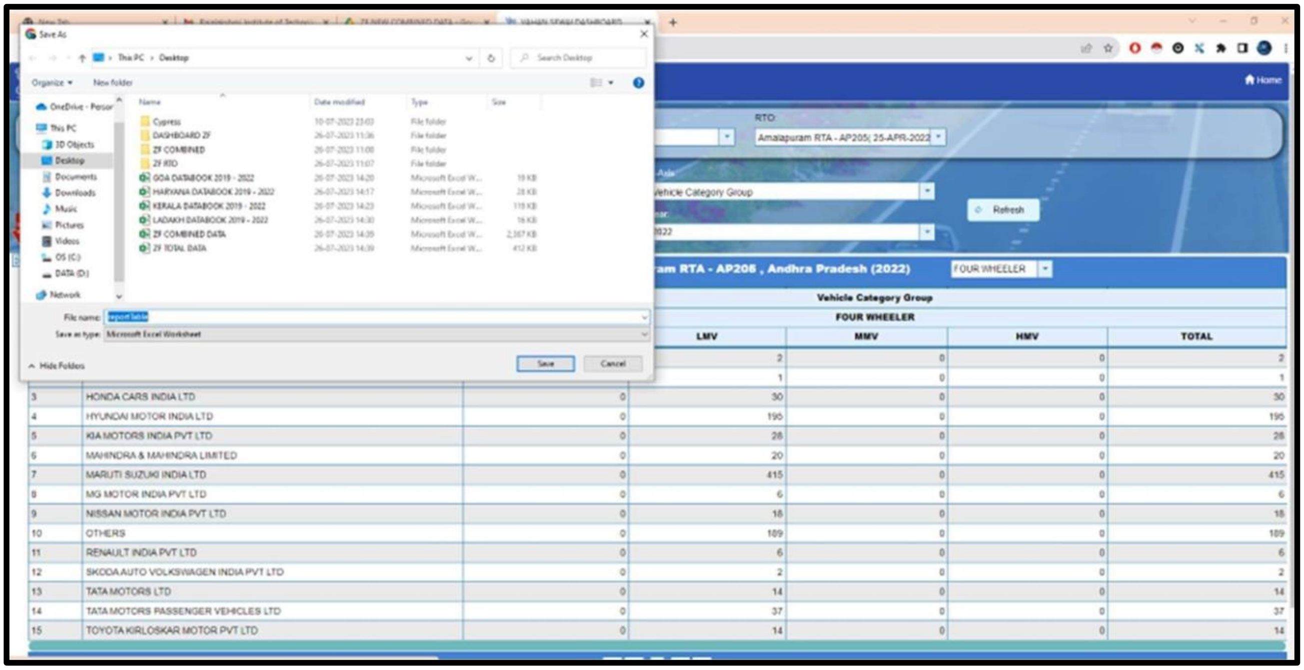


Fig. 2.4.1

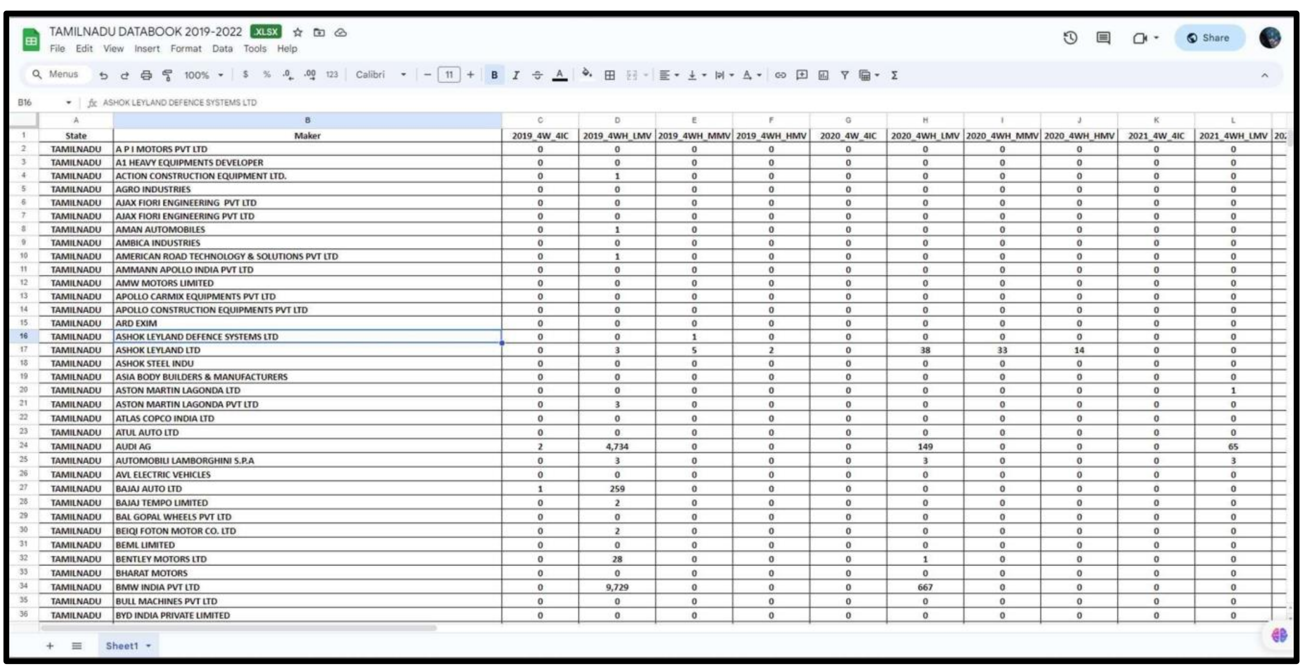


Fig. 2.4.2

3. DATA TRANSFORMATION

3. DATA TRANSFORMATION

In the initial phase of our project for ZF Company, data collection from the Vahan portal played a pivotal role in establishing the foundation for our comprehensive vehicle data analysis. The Vahan portal, renowned for its rich repository of vehicle-related information, served as a robust source for our research endeavors.

Following the meticulous data collection process, our attention shifted towards the crucial stage of data transformation. This involved a multifaceted approach aimed at refining the raw data and optimizing its usability for subsequent analytical processes. One of the primary steps involved the systematic removal of duplicates within the dataset. This critical process ensured that the data was streamlined, eliminating redundancy and enhancing the accuracy of our analyses.

Simultaneously, we addressed the presence of null values within the dataset. Rigorous data cleaning techniques were applied to identify and subsequently handle or remove instances of missing or incomplete data. This step was paramount in ensuring the integrity and completeness of our dataset, contributing to the reliability of the insights derived from subsequent analyses.

Furthermore, the transformation process extended to the formatting of the data into a suitable and standardized structure. This step involved organizing the data in a manner that aligns with the analytical tools and systems employed in the subsequent phases of our project. Standardized formats facilitate a seamless integration of the data into visualization tools, thereby enhancing the efficiency and accuracy of our analytical processes.

By undertaking these comprehensive data transformation measures, our objective was not only to refine the quality of the dataset but also to lay the groundwork for a robust and efficient analytical workflow. The removal of duplicates and null values, coupled with data formatting, set the stage for meaningful analyses that would contribute significantly to the development of the envisioned dashboard for ZF Company.

In conclusion, the transformation of data collected from the Vahan portal involved a meticulous process of removing duplicates and addressing null values, followed by formatting the data into a suitable structure. These steps were pivotal in enhancing the quality and usability of the dataset, ensuring that subsequent analyses and the development of the dashboard would be founded on accurate, complete, and well-organized information.

The next stage is **feature extraction and transformation**. Here, raw variables are converted into higher-level features that are more meaningful for analysis. For example, acceleration is derived from changes in speed over time, fuel efficiency is calculated as distance traveled per unit of fuel, and harsh braking events are identified from rapid deceleration patterns.

Similarly, categorical variables like “gear position” may be encoded numerically for compatibility with machine learning algorithms.

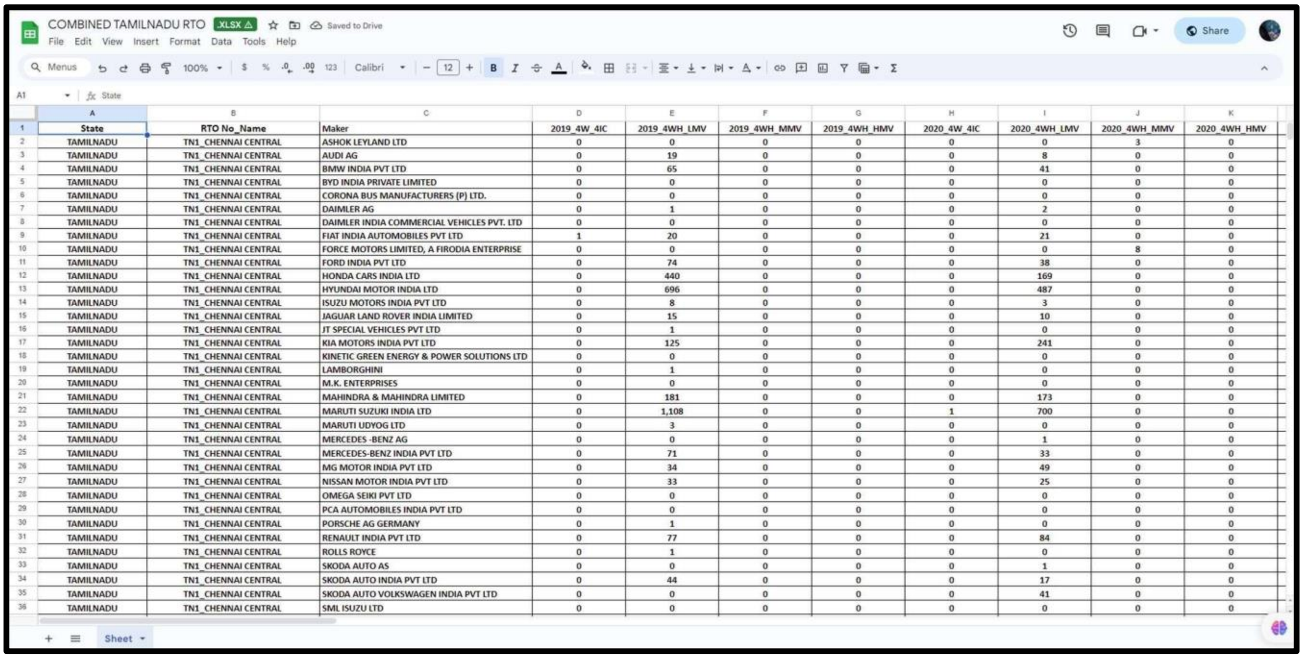
Following this, **dimensionality reduction and aggregation** are often applied. Large-scale vehicle datasets may contain redundant or highly correlated variables. Techniques such as Principal Component Analysis (PCA) or clustering can be used to reduce complexity without losing essential information. Aggregation processes also summarize data over time, such as average fuel consumption per trip or daily distance covered, which makes the dataset easier to analyze.

Another important part of the transformation process is **data integration**. Vehicle data often comes from multiple heterogeneous sources, such as onboard sensors, GPS trackers, weather APIs, and road traffic systems. Integration ensures that these diverse inputs are merged into a single unified dataset for comprehensive analysis.
Finally, the transformed dataset undergoes **validation and formatting**.

The processed data is checked against predefined rules or domain knowledge to ensure correctness and reliability. It is then stored in structured formats like SQL databases, cloud data warehouses, or time-series databases, making it readily accessible for downstream analytics, visualization, or predictive modeling.

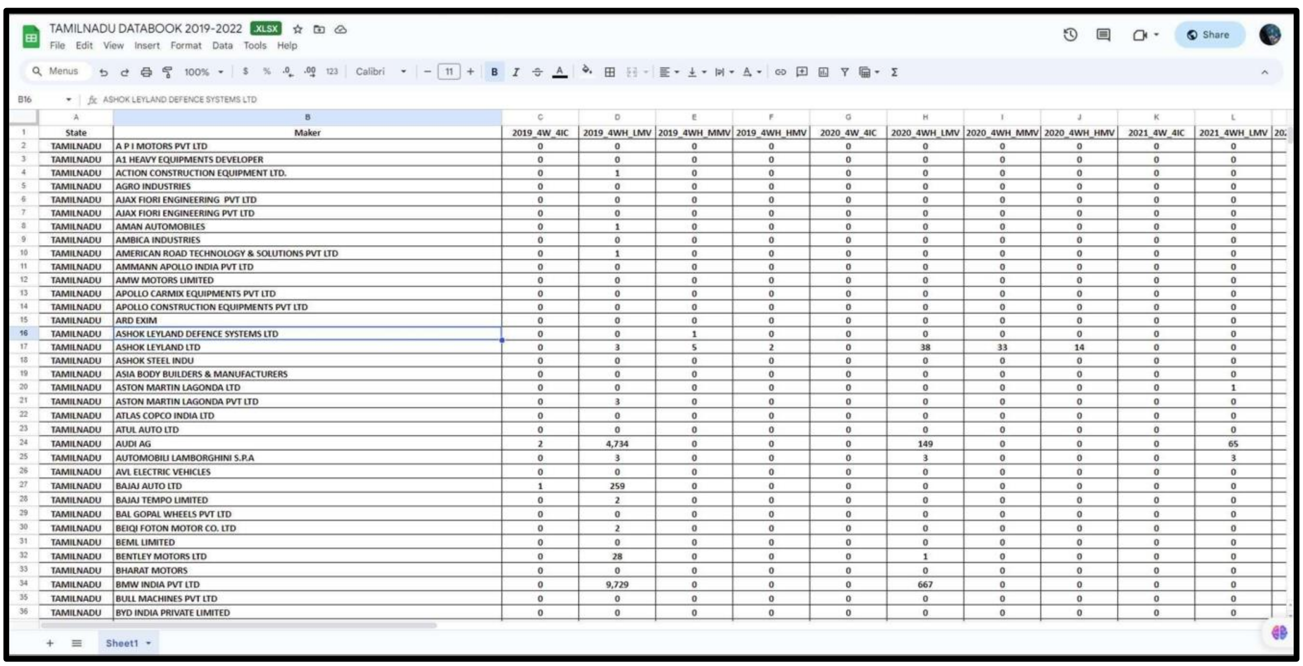
The data transformation process in vehicle data analysis involves cleaning, normalization, feature extraction, aggregation, integration, and validation. By systematically converting raw vehicle signals into meaningful, structured datasets, this process enables accurate modeling, insightful analysis, and reliable decision-making in automotive research and applications.

DATA TRANSFORMATION



State	RTO No_Name	Maker	2019_4W_4IC	2019_4WH_LMV	2019_4WH_MMV	2019_4WH_HMV	2020_4W_4IC	2020_4WH_LMV	2020_4WH_MMV	2020_4WH_HMV
TAMILNADU	TN1_CHENNAI CENTRAL	ASHOK LEYLAND LTD	0	0	0	0	0	0	3	0
TAMILNADU	TN1_CHENNAI CENTRAL	AUDI AG	0	19	0	0	0	8	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	BMW INDIA PVT LTD	0	65	0	0	0	41	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	BYD INDIA PRIVATE LIMITED	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	CORONA BUS MANUFACTURERS (P) LTD.	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	DAIMLER AG	0	1	0	0	0	2	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	DAIMLER INDIA COMMERCIAL VEHICLES PVT. LTD	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	FIAT INDIA AUTOMOBILES PVT LTD	1	20	0	0	0	21	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	FORCE MOTORS LIMITED, A FIRODIA ENTERPRISE	0	0	0	0	0	0	8	0
TAMILNADU	TN1_CHENNAI CENTRAL	FORD INDIA PVT LTD	0	74	0	0	0	38	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	HONDA CARS INDIA LTD	0	440	0	0	0	169	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	HYUNDAI MOTOR INDIA LTD	0	696	0	0	0	487	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	ISUZU MOTORS INDIA PVT LTD	0	8	0	0	0	3	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	JAGUAR LAND ROVER INDIA LIMITED	0	15	0	0	0	10	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	JT SPECIAL VEHICLES PVT LTD	0	1	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	KIA MOTORS INDIA PVT LTD	0	125	0	0	0	241	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	KINETIC GREEN ENERGY & POWER SOLUTIONS LTD	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	LAMBORGHINI	0	1	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	M.K. ENTERPRISES	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MAHINDRA & MAHINDRA LIMITED	0	181	0	0	0	173	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MARUTI SUZUKI INDIA LTD	0	1,108	0	0	1	700	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MARUTI UDYOG LTD	0	3	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MERCEDES -BENZ AG	0	0	0	0	0	1	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MERCEDES-BENZ INDIA PVT LTD	0	71	0	0	0	33	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	MG MOTOR INDIA PVT LTD	0	34	0	0	0	49	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	NISSAN MOTOR INDIA PVT LTD	0	33	0	0	0	25	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	OMEGA SEIKI PVT LTD	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	PCA AUTOMOBILES INDIA PVT LTD	0	0	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	PORSCHE AG GERMANY	0	1	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	RENAULT INDIA PVT LTD	0	77	0	0	0	84	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	ROLLS ROYCE	0	1	0	0	0	0	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	SKODA AUTO AS	0	0	0	0	0	1	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	SKODA AUTO INDIA PVT LTD	0	44	0	0	0	17	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	SKODA AUTO VOLKSWAGEN INDIA PVT LTD	0	0	0	0	0	41	0	0
TAMILNADU	TN1_CHENNAI CENTRAL	SML ISUZU LTD	0	0	0	0	0	0	0	0

Fig. 3.1



State	Maker	2019_4W_4IC	2019_4WH_LMV	2019_4WH_MMV	2019_4WH_HMV	2020_4W_4IC	2020_4WH_LMV	2020_4WH_MMV	2020_4WH_HMV	2021_4W_4IC	2021_4WH_LMV
TAMILNADU	A P I MOTORS PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	A1 HEAVY EQUIPMENTS DEVELOPER	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ACTION CONSTRUCTION EQUIPMENT LTD.	0	1	0	0	0	0	0	0	0	0
TAMILNADU	AGRO INDUSTRIES	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AJAX FIORI ENGINEERING PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AJAX FIORI ENGINEERING PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AMAN AUTOMOBILES	0	1	0	0	0	0	0	0	0	0
TAMILNADU	AMBICA INDUSTRIES	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AMERICAN ROAD TECHNOLOGY & SOLUTIONS PVT LTD	0	1	0	0	0	0	0	0	0	0
TAMILNADU	AMMANN APOLLO INDIA PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AMW MOTORS LIMITED	0	0	0	0	0	0	0	0	0	0
TAMILNADU	APOLLO CARMIX EQUIPMENTS PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	APOLLO CONSTRUCTION EQUIPMENTS PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ARD EXIM	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ASHOK LEYLAND DEFENCE SYSTEMS LTD	0	0	1	0	0	0	0	0	0	0
TAMILNADU	ASHOK LEYLAND LTD	0	3	5	2	0	38	33	14	0	0
TAMILNADU	ASHOK STEEL INDU	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ASIA BODY BUILDERS & MANUFACTURERS	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ASTON MARTIN LAGONDA LTD	0	0	0	0	0	0	0	0	0	1
TAMILNADU	ASTON MARTIN LAGONDA PVT LTD	0	3	0	0	0	0	0	0	0	0
TAMILNADU	ATLAS COPCO INDIA LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	ATUL AUTO LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	AUDI AG	2	4,734	0	0	0	149	0	0	0	65
TAMILNADU	AUTOMOBILI LAMBORGHINI S.P.A	0	3	0	0	0	3	0	0	0	3
TAMILNADU	AVL ELECTRIC VEHICLES	0	0	0	0	0	0	0	0	0	0
TAMILNADU	BAJAJ AUTO LTD	1	259	0	0	0	0	0	0	0	0
TAMILNADU	BAJAJ TEMPO LIMITED	0	2	0	0	0	0	0	0	0	0
TAMILNADU	BAL GOPAL WHEELS PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	BEIQI FOTON MOTOR CO. LTD	0	2	0	0	0	0	0	0	0	0
TAMILNADU	BEML LIMITED	0	0	0	0	0	0	0	0	0	0
TAMILNADU	BENTLEY MOTORS LTD	0	28	0	0	0	1	0	0	0	0
TAMILNADU	BHARAT MOTORS	0	0	0	0	0	0	0	0	0	0
TAMILNADU	BMW INDIA PVT LTD	0	9,729	0	0	0	667	0	0	0	0
TAMILNADU	BULL MACHINES PVT LTD	0	0	0	0	0	0	0	0	0	0
TAMILNADU	BYD INDIA PRIVATE LIMITED	0	0	0	0	0	0	0	0	0	0

Fig. 3.2

4. DATA VISUALIZATION

4. DATA VISUALIZATION

The process of transforming raw data into a suitable format is a pivotal step that sets the stage for meaningful analysis and interpretation. With the data now structured and refined, the focus shifts towards the crucial phase of visualization, a powerful technique aimed at extracting actionable insights from complex datasets. Visualization not only enhances the accessibility of information but also provides a dynamic platform for uncovering patterns, trends, and correlations within the data.

In this stage of the data analysis journey, various visualization techniques are employed to present information in a comprehensible and intuitive manner. Graphs, charts, and interactive dashboards become instrumental tools in translating the transformed data into visually appealing representations. Through strategic visualization, complex relationships and trends that may be challenging to discern in raw data become readily apparent, enabling stakeholders to grasp the significance of patterns and make informed decisions.

One key advantage of visualization is its ability to communicate information efficiently across diverse audiences within an organization. Decision-makers, analysts, and other stakeholders can quickly assimilate insights presented visually, facilitating a shared understanding of the data's implications. Interactive dashboards, in particular, offer a user-friendly interface that allows users to explore data dynamically, drilling down into specific details and adjusting parameters to gain a comprehensive view of the information.

The choice of visualization techniques depends on the nature of the data and the objectives of the analysis. Common visualizations include bar charts, line graphs, pie charts, scatter plots, and heatmaps, each serving a unique purpose in highlighting specific aspects of the data. Additionally, more advanced techniques such as 3D visualizations, geospatial mapping, and network diagrams may be employed to capture complex relationships and spatial patterns.

Beyond the aesthetic appeal, the true value of visualization lies in its capacity to uncover insights that drive strategic decision-making. Whether identifying performance trends, recognizing anomalies, or understanding the impact of variables, visualization transforms data into actionable intelligence. This transformative process not only aids in immediate decision-making but also lays the foundation for continuous monitoring and adaptation in response to evolving trends.

In conclusion, the visualization phase represents a critical juncture in the data analysis process. By harnessing the power of visual representation, organizations can unlock the latent potential within their data, empowering stakeholders with a deeper understanding of complex relationships and trends. As the data is brought to life through visuals, it becomes a compelling tool for strategic decision-making, fostering a data-driven culture within the organization and contributing to overall business intelligence and success.

Fig. 4.1

5. POWER-BI DASHOARD

5. DATA VISUALIZATION

In this project, the Power BI tool stands as the linchpin of our analytical endeavors, serving as the conduit through which we transform raw data into actionable insights. Power BI's robust capabilities in data visualization empower us to create a series of dynamic and interactive dashboards. These dashboards, meticulously designed and tailored to the specific needs of our analysis, serve as powerful tools for unraveling complex datasets and extracting meaningful insights.

5.1 Power BI as the Analytical Engine:

Power BI's versatility allows us to seamlessly integrate diverse datasets, facilitating a holistic view of the data landscape. Leveraging its intuitive interface, we employ various data visualization techniques to represent intricate relationships, trends, and patterns within the data. The tool's real-time updating feature ensures that our dashboards evolve synchronously with the dynamic nature of the underlying data.

5.2 Dashboard Diversity for Comprehensive Analysis:

Our utilization of Power BI extends beyond mere visualization; we harness its capabilities to craft a multitude of dashboards, each designed to scrutinize specific facets of the dataset. From geographical distribution maps to trend analysis timelines, our dashboards cater to diverse analytical needs. This diversity ensures a comprehensive exploration of the dataset, enabling stakeholders to glean insights from various perspectives.

5.3 Analytical Depth Through Visualization:

The power of visualization in data analysis cannot be overstated. Power BI enables us to create visually appealing representations that transcend the complexity of raw data. Through the use of charts, graphs, and interactive elements, our dashboards transform abstract numbers into tangible insights, fostering a deeper understanding of the dataset and facilitating informed decision-making.

5.4 Interactive Dashboards for User Engagement:

Our approach to dashboard creation extends beyond static representations. Leveraging Power BI's interactive features, our dashboards invite user engagement. Stakeholders can drill down into specific data points, filter information based on parameters of interest, and extract on-the-fly insights. This interactive element not only enhances the user experience but also empowers stakeholders to actively participate in the analytical process.

The design of the interactive dashboard focuses on data visualization and user interactivity. Vehicle parameters such as speed, fuel efficiency, mileage, engine temperature, braking patterns, and trip histories are represented through graphs, charts, gauges, and maps. These visualizations allow users to quickly identify patterns, trends, and anomalies.

5.1 Insights Extraction and Decision Support:

The ultimate goal of our Power BI-driven dashboards is to extract actionable insights. By distilling complex data into digestible visualizations, we provide decision-makers with a clear and concise understanding of critical trends and patterns. The tool's capabilities enable us to go beyond surface-level observations, delving into the nuances of the dataset to uncover insights that drive strategic decision-making.

insight extraction is to move beyond descriptive analysis (what happened) toward predictive and prescriptive analysis. Predictive models can forecast events such as fuel depletion, battery failure, or component wear, whereas prescriptive analytics can recommend specific actions, such as scheduling preventive maintenance or altering driving habits to extend vehicle life.

5. DATA TRANSFORMATION

Fig. 5.1

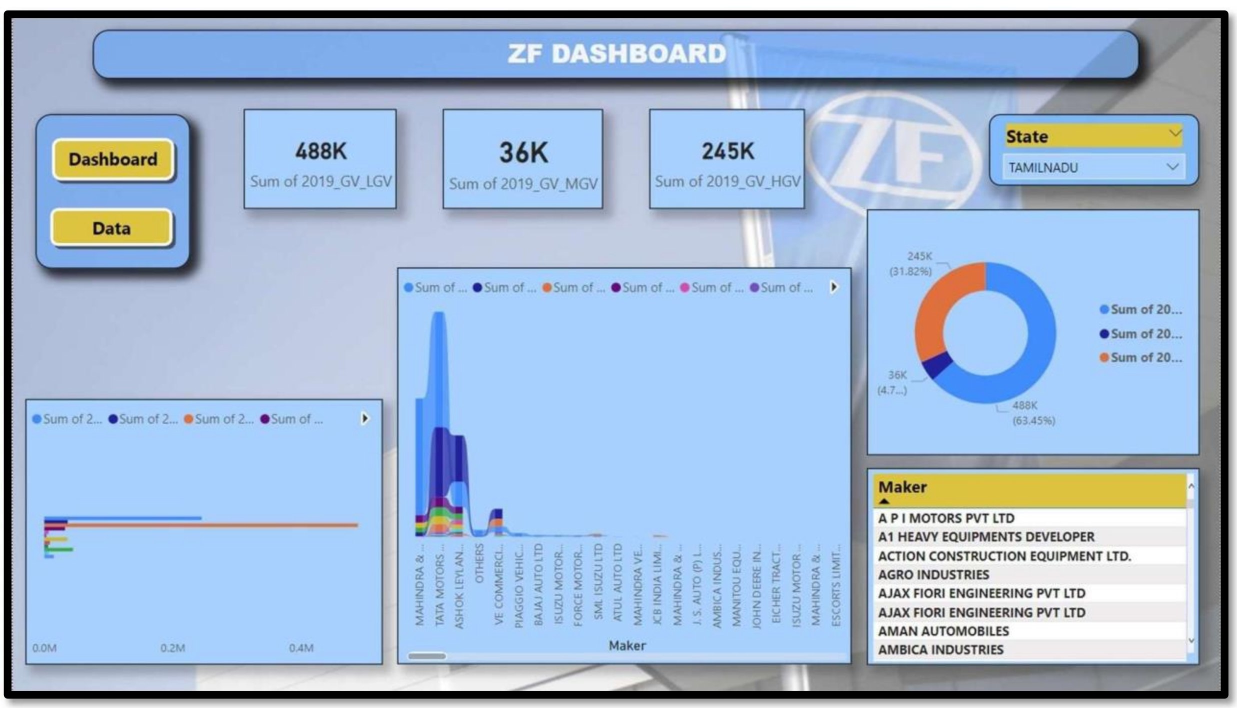
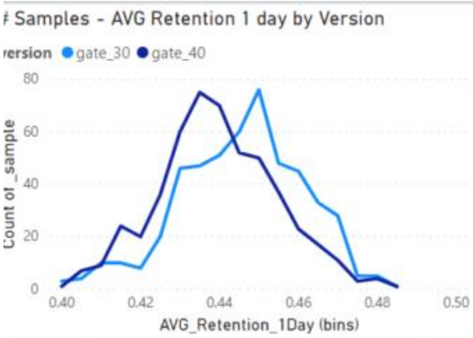
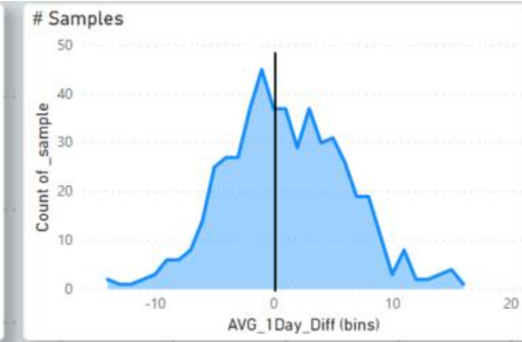
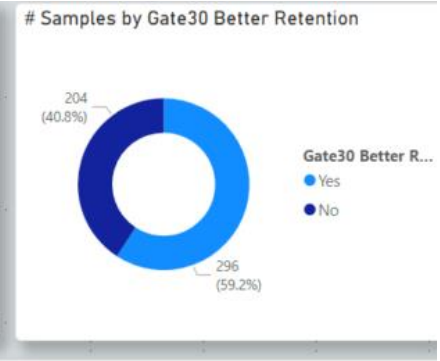


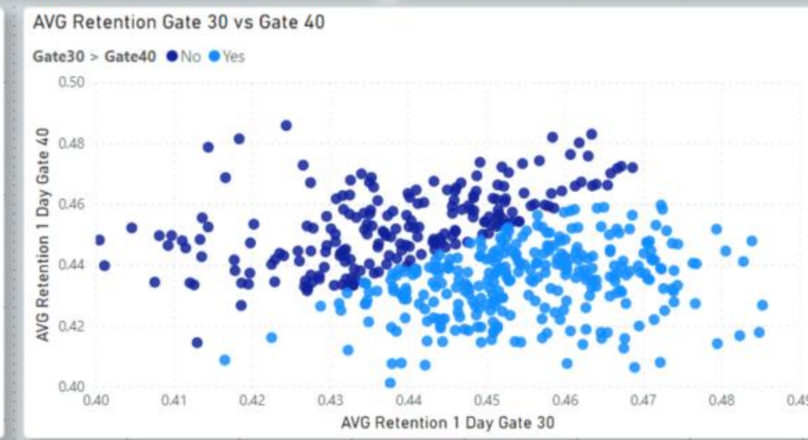
Fig. 5.2







sample	AVG 1 Day Retention Gate30	AVG 1 Day Retention Gate40	% AVG 1 Day Diff
1	0.4479	0.4582	-2.25
2	0.4509	0.4179	7.89
3	0.4547	0.4638	-1.95
4	0.4301	0.4520	-4.86
5	0.4229	0.4346	-2.70
6	0.4548	0.4403	3.29
7	0.4311	0.4275	0.85
8	0.4460	0.4470	-0.22
9	0.4602	0.4399	4.63
10	0.4580	0.4507	1.61
11	0.4513	0.4194	7.62
Total	**0.4481**	**0.4421**	**1.49**



6. CONCLUSION

6. CONCLUSION

In the culmination of this expansive project, our journey through the intricacies of vehicle data analysis for ZF Company has been marked by a meticulous blend of methodical data handling, advanced analytics, and innovative visualization techniques. The utilization of the Power BI tool served as the keystone, enabling us to transcend the inherent complexity of raw data and orchestrate a symphony of insights that resonate with strategic significance.

Our initial foray into data collection from the Vahan portal laid the foundation for a rich and diverse dataset, providing us with a comprehensive palette upon which our analytical brushstrokes could unfurl. The subsequent phase of data transformation, facilitated by Power BI's robust capabilities, ensured that the raw data underwent a metamorphosis into a refined and structured entity, ready for in-depth analysis.

The creation of a multitude of dashboards emerged as the pièce de resistance of our analytical endeavor. These dashboards, meticulously tailored to address specific dimensions of the data, became powerful lenses through which we scrutinized usage patterns, geographical distributions, and performance metrics. The diversity inherent in our dashboard design facilitated a holistic exploration of the dataset, unveiling nuanced insights that transcended the boundaries of conventional data analysis.

The dynamic and interactive nature of our dashboards, driven by Power BI's prowess, transformed the analytical process into an engaging and participatory experience. Stakeholders were not merely passive observers; they became active agents, navigating through the data landscape, drilling down into specific points of interest, and extracting on-the-fly insights. This interactive element not only enhanced user engagement but also empowered decision-makers with a profound understanding of the dataset.

As we navigated through the project's objectives, our analytical depth reached new heights. The amalgamation of data visualization techniques, advanced analytics, and the dynamic capabilities of Power BI unearthed trends, correlations, and anomalies that eluded traditional analysis methods. The insights extracted from this synergy provide ZF Company with a nuanced understanding of its automotive landscape, fostering a strategic foresight that is essential in the ever-evolving automotive industry.

In conclusion, this project transcends the boundaries of a mere data analysis endeavor. It stands as a testament to the transformative power of technology, analytics, and strategic vision. The partnership between meticulous data handling and the innovative capabilities of Power BI has not only met the project objectives but has elevated the discourse around data-driven decision-making. As we step away from this project, we leave behind not just dashboards and insights but a legacy of analytical excellence and a blueprint for future endeavors in the dynamic realm of data analytics.