Effective Analysis for Automating Fleet Management Based on Vehicle Performance Factors

B. M. I. W. G. Gurusinghe
179462D

Faculty of Information Technology.
University of Moratuwa
2021

Effective Analysis for Automating Fleet Management Based on Vehicle Performance Factors

B. M. I. W. G. Gurusinghe
179462D

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfilment of the requirements of the Master of Science in Information Technology.

2021

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text, and a list of references is given.

Name of Student: B. M. I. W. G. Gurusinghe

Signature of the Student: _____

Date:

Supervised by:

Name of Supervisor: Mr. Chaman Wijesiriwardhana

Signature of Supervisor: _____

Date:

Dedication

I dedicate my dissertation work to my family and friends who have supported and encouraged me throughout this project.

Acknowledgement

I am using this opportunity to express my gratitude to everyone who supported me throughout this research. I am thankful for their aspiring guidance, invaluable constructive criticism, and friendly advice during the project work. I am sincerely grateful to my lecturer Mr. Chaman Wijesiriwardhana for sharing truthful and illuminating views on several issues related to the project. I would like to express my sincere gratitude to Dr. M. Firdhous for feeding me knowledge and guidance when doing my research. Special thanks to my sister Jayani Gurusinghe and best friends Thaveesha Gamage, Amasha Perera, and all the batch mates of the M.Sc. in IT degree program who have supported me in every possible way.

Abstract

Various fleet management systems are used by companies where transportation is the main business function. Fleet management applications are used for vehicle tracking purposes and mechanic diagnostic purposes to troubleshoot problems related to vehicles. Fleet management (FM) is used in vehicle maintenance as well as improving vehicle and driver performance, speed control, fuel management, and many other different areas to improve the business operation in different aspects. The usage of FM systems improves fleet productivity and efficiency, reduces transportation costs, and minimizes vehicle investment risk. For businesses like cash in transit, fleet systems are used to monitor operations and to ensure security. This research proposes a systematic procedure to the Cash-In-Transit (CIT) division of the AB Securitas (Private) Limited. The CIT division receives job requests from customer locations of different geographical areas. Based on daily job requirements, routes for vehicle operations are identified, and resources for each route are allocated by considering different factors such as productivity, cost reduction, secure transport, vehicle condition, the distance between customer locations, vehicle availability, and profit. Since the CIT operating unit uses both company-owned and hired vehicles for their operations, there should be a scientific process to assign both vehicles profitably. However, in order to allocate the vehicles to the daily jobs effectively and cost efficiently, several factors need to be considered. When it comes to allocating vehicles to daily jobs, vehicle performance is a key factor to be considered. It is obvious that vehicles with higher performance need to be allocated to the most appropriate jobs and routes to gain a considerable profit. The proposed solution aims to overcome the issues and unjustified matters identified in the current manual process and reduce the complexity of the allocation method. Further, this solution provides a detailed analysis of the impact of vehicle breakdowns on vehicle performance since it is a key factor that needs to be considered in order to allocate the right vehicle to the most appropriate job accurately.

Keywords: Fleet management, Transport productivity, Vehicle performance, Vehicle breakdowns

Table of Content

Introduction 1

1.1 Prolegomena 1
1.2 Background and Motivation 1
1.3 Problem Statement 2
1.4 Aim 3
1.5 Objectives 3
1.6 The Approach of the System 4
1.7 Structure of the Thesis 4
1.8 Summary 5

Review of literature 6

2.1 Introduction 6
2.2 Existing Developments in Fleet Management 6
2.3 Existing Process Analysis 17
2.4 Functional Requirements 18
2.5 Non Functional Requirements 19
2.6 Problem Definition 19
2.7 Summary 20

Adopted Technologies & Tools 21

3.1 Introduction 21
3.2 PHP Server Side Scripting Language 21
3.3 APACHE2 Web Server 22
3.4 Maria DB 23
3.5 Adopted Software Tools 24
3.6 LARAVAL 25
3.7 Summary 26

Analysis for Detecting Breakdown Factors 27

4.1 Introduction 27
4.2 Data Mining 27
4.2.1 Supervised Learning 27
4.2.2 Unsupervised Learning 28
4.2.3 Semi-Supervised Learning 28
4.3 Why Linear Regression 29
4.4 Description of Data 30
4.5 Data Pre-processing 31
4.6 Descriptive Analysis 33
4.7 Advanced Analysis 37
4.8 Model Summary 44
4.8.1 Multiple Linear Regression Model 44
4.8.2 ANOVA 47
4.9 The Prediction Accuracy of the Model 47
4.10 Summary 48
Design and Implementation of the System 49
5.1 Introduction 49
5.2 Analysis of the System 49
5.3 Top Level Architecture 50
5.4 Design Architecture 55
5.4.1 Process Models 55
5.4.2 Process Models System Architecture 56
5.5 System Modules and Functionalities 58
5.6 Main Web Application 59
5.7 Overall Solution 60
5.8 Implementation of Database 61
5.9 Summary 61
Assessment of the Fleet Management System 62
6.1 Introduction 62

6.2 Evaluation of the Fleet Management System 62
6.3 Summary 63
Conclusion and Future Work 64
7.1 Introduction 64
7.2 Achievement 65
7.3 Limitations and Future Works 66
7.4 Summary 67
References 68
Appendixes 72
Appendix A – User Interfaces. 72
Appendix B – Coding Parts 87
Appendix C - Test Cases 90

List of Figures

Figure 1 Re-coding the variable 'Model'.....32
Figure 2 Attribute Construction – BPM.....32
Figure 3 Histogram for BPM.....33
Figure 4 Bar plot for Make.....34
Figure 5 Bar plot for Model.....34
Figure 6 Bar plot for Fuel Category.....35
Figure 7 Bar plot for Engine Capacity.....36
Figure 8 Histogram for Age.....36
Figure 9 Scatter Plot for Maintenance expenses.....39
Figure 10 Scatterplot for Fuel Usage.....39
Figure 11 Residuals vs Predicted Values.....42
Figure 12 Normal P-P Plot.....43
Figure 13 Sum of square values.....48
Figure 14 Top-level architecture.....50
Figure 15 data flow diagram.....51
Figure 16 entity-relationship diagram.....52
Figure 17 System Class Diagram.....53
Figure 18 Use Case Diagram.....53
Figure 19 Fleet Management System DB Structure.....54

Figure 20 Waterfall model 55
Figure 21 MVC Architecture 57
Figure 22 an overview of task management and daily plan module 59
Figure 23 Main Web application 60

List of Tables

Table 1 Categorical and Continuous Variables 31
Table 2 Pearson Correlation Values 37
Table 3 Pearson Correlation of BPM and independent variables 38
Table 4 Durbin Watson Value of the model 40
Table 5 VIF values for independent variables 41
Table 6 Centered Leverage Value 43
Table 7 Model Summary 44
Table 8 Standardized Coefficients 45
Table 9 ANOVA table 47

Chapter 1

Introduction

1.1 Prolegomena

An overview of the research is given in this chapter. It describes the research background and briefly explains the proposed solution. The research aim and objectives are outlined afterward. The outline of the remaining chapters is concluded in the chapter conclusion.

1.2 Background and Motivation

Fleet management is indispensable for businesses that rely on transportation as their main business function [1] and it subsumes different functions, such as vehicle and fleet driver management, fuel management, vehicle maintenance and finance, vehicle tracking and diagnostics, monitoring performances, etc. to ensure that the business has a reliable, efficient and cost-effective transport system which meets business goals. Fleet management is also important to improve fleet productivity whilst also reducing costs and increasing profits. [2] When it comes to cash-in-transit, ensuring security is extremely important.

AB Securitas (Private) Limited is a leading company that is formed to focus on the different aspects and demands of the security Industry. Their Cash in Transit (CIT) Division handles the physical transportation of cash between locations as required by the client. This research proposes a systematic procedure to the CIT division of the AB Securitas (Private) Limited to develop an automated system for the allocation of vehicles and drivers for their daily operations. There are 95 vehicles in the CIT division's fleet relating to CIT operations and more than 750 locations on daily operations. The CIT operating unit uses both company-owned vehicles and hired vehicles for their operations. When assigning a vehicle and driver for an operation there are many factors to be considered.

The profitability is based on the type of vehicle (company-owned or hired), the distance of routes, condition of the vehicle, availability of own vehicles, most profitable way to hire vehicles from suppliers, security factors should consider when assigning a vehicle, and driver for operation.

Currently, all the reservations are done manually daily. No supervision takes place during the process and therefore, some mistakes and unjustifiable matters occur while the operations are planned. Sometimes hired vehicles are assigned for long-distance runs. The company-owned vehicles are kept for a long period for repairs as well as getting hired vehicles additionally. Sometimes company-owned vehicles which are not in good condition and vehicles which were released after repairs are assigned for long-distance runs, and this results in vehicle breakdowns while on the run. On the other hand, high-performing vehicles that are cost-effective are allocated for short distances, while low-performing vehicles are assigned to long distances which ultimately results in lower profits. Sometimes the hired vehicles are allocated depending on the supplier's requirement and not based on the company's profitability. In this situation, the company is not able to identify and cut down unnecessary expenses. Also, it is essential to provide protected and secured transportation in cash in transit. Therefore, the daily rotation of vehicles and drivers is required in this process. With the complexity of this situation, one person cannot control the whole process of vehicle allocation and driver allocation properly without a systematic procedure. Therefore, an automatic system for allocating vehicles and drivers for daily operations has been proposed for the company.

1.3 Problem Statement

In the analyzing process of CIT division operations, the researcher has found various unsolved problems related to the process, such as not having a proper procedure to manage the hiring process, long-duration to detect vehicle and driver availability, not having a proper method to collect related data, issues in managing rotation for security concerns and complexity in managing the process. Allocating the owned and hired vehicles to the most appropriate route is the most critical process of this system since there are several serious

conditions to be considered before allocating them. Among those factors, vehicle performance should be given a higher priority because when it comes to vehicles, performance is a highly related factor to be considered. So to gain high effectiveness and accuracy by automating the fleet management process, we need to do a detailed analysis regarding the performance factors of the vehicle since it directly affects the key process of the system. Due to these issues, AB Securitas (Private) Limited management team has requested to build a whole new automated decision support system to manage their daily operations. Therefore, a solution is proposed for a proper system by eliminating the above-mentioned issues.

1.4 Aim

This research aims to propose a web-based automated system using related web technologies to manage CIT processes accurately and effectively and acquire centralized control in the CIT division.

1.5 Objectives

i. Identify the current CIT division process and the problems of the existing process.
ii. Identify the most accurate, profitable method of assigning vehicles and drivers to reduce costs and ensure secure transportation.
iii. Reduce the complexity of the current process and data gathering by introducing a centralized design.
iv. Analyzing the key vehicle performance factors to increase the accuracy and reduce the complexity of the vehicle allocating process

1.6 The Approach of the System

The proposed system is a web-based application. The web application would be developed using PHP server-side scripting language and other related web-based technologies and tools such as scripting languages, PHP frameworks, and MySQL database management. The application is run on an Apache2 web server.

In this application, the system user is able to add daily job requests received from customers and generate route details relevant to the daily operations. The system will generate the vehicle requirement for the date and manage company-owned vehicles and hired vehicles based on that requirement. Supplier vehicles are hired systematically and profitably allocated to routes. Company-owned vehicles are also allocated in a cost-effective, efficient manner based on the performance, and rotation of vehicles and drivers are also considered. Finally, the system will generate the daily allocation report to managing all the operations of the CIT division.

1.7 Structure of the Thesis

The thesis structure is organized as follows. The research overview and the background of the research describe in chapter 1. Further is gives a brief explanation of the proposed solution.

There are seven main chapters in this report. Existing knowledge and present learning of the study is organized under chapter 2. Chapter 3 is about the used technologies and tools for developing the solution. The statistical analysis of detecting breakdown factors included in chapter 4. The Design and implementation of the system and the evaluation of the system is described in Chapter 5 and 6. Finally, the conclusion and suggestions for future work are presented in chapter 7.

1.8 Summary

By analyzing the current CIT daily operation handling process of AB Securitas (Private) Limited, the issues related to the existing process were identified. To overcome these issues and to cut down the unnecessary costs, a systematic process was designed and proposed as an automated system. The necessity of detailed analysis has been identified to maximize the effectiveness and accuracy while reducing the complexity of the vehicle allocating process of the proposed automated fleet management system. By using the proposed design the company would be able to find solutions for current issues and reduce complexity and raise more profit.

Chapter 2

Review of literature

2.1 Introduction

An overview of the research, a brief explanation of the proposed solution, and a description of the research background are given in the previous chapter. This chapter describes identifying key variables of the process, factors that affect the performance of vehicles, the required data, how data mining techniques are used in fleet management. It also describes suitable methodologies to be followed when implementing the presented design and the early developments and background of recent and past developments. For this purpose, a review of past researches has been presented in this chapter.

2.2 Existing Developments in Fleet Management

Transportation is a complex field that involves several players, decision levels, uncertainties, and considerable capital expenditure. To remain competitive and effective, this industry relies on a large amount of data, sophisticated models, and optimization techniques, as well as information technology. [3]

Fleet management is a function that allows companies operating in the transportation industry, to remove or minimize the risks associated with vehicle investment while improving efficiency, productivity, and reducing their overall transportation costs. The purpose of fleet management is to track and manage the activities and conditions of vehicles and drivers using digital means. This helps the transportation and logistics industries keep abreast of how their fleets perform while giving ideas on how to optimize their usage [4]. Fleet management supports these businesses to carry out their daily operations smoothly, improve performance, and keep operating costs at a minimum.

Fleet management includes a range of functions such as vehicle maintenance, driver management, speed management, tracking and diagnostics, fuel management, health and safety management, and vehicle re-marketing. [5] It enables businesses to accomplish a series of specific tasks in the management of any or all aspects relating to a fleet of vehicles operated by a company. Fleet management software should be able to manage processes, tasks, and events related to vehicles and drivers. With the use of fleet management systems, the company is able to acquire more control over drivers and vehicles, save the cost of unjustified allocations of vehicles, monitor vehicle and driver performance, reduce the time spent on predictions, save on fuel consumption, prevent delays and optimize daily operations. Increase in revenue per mile traveled, increase in revenue per trip, reduction in vehicle idling time, availability of different cargo handling and other equipment, reduction of operating costs, availability of drivers, possibilities for additional usage of vehicles during trips, the quantity of cargo per single delivery, evidence of the points of start and end of delivery, stock management, evidence of the time of delivery, reduction of fuel consumption are some of the intentions and objectives of the improvement of the operation of the transport companies after introducing fleet management systems for their companies.[9]

Many industrial and security companies need to keep track of and optimize their fleet. Planning, controlling, monitoring, and management of various activities in transportation sectors are considered to be extremely important. Information and communication are becoming unavoidable factors in the optimization of transport processes. Various technologies applicable for this purpose are available and different types of systems are developed based on the requirements. Those can be found in two forms as online systems and offline systems. [7]

Fleet Management is a data-driven industry. The fleet management systems can be based on dynamic data (real-time data) or non-dynamic data. These data can be vehicle-centric, behavioral, operational, or administrative. Therefore the fleet management systems generate a profuse amount of data. Currently, fleet systems entail database information

(offline or non-dynamic) software and telematics and GPS tracking (online or dynamic) software.

Dynamic data systems are mostly used in vehicular positioning systems. These systems are widely used to track vehicle positions, monitor traffic, enable remote diagnostics of the process, and also for inter-vehicle communication. GPS (Global Positioning System) modules, LBS (Location-Based Services), in-vehicle embedded systems, image capturing techniques,[6] Onboard Diagnostic system(OBD) interfaces, Radio Frequency Identification (RFID) readers [7] are used as positioning systems. In dynamic positioning systems, network-based positioning and terminal-based positioning are used as positioning methods. In network-based positioning, location tracking is done using the base station network. In the terminal method, the location is calculated by the user device itself from signals received from base stations.

Non-dynamic data-based or offline systems store and analyze collected data from vehicles. These systems should be able to deal with a higher number of vehicles with fewer overheads. The characteristics of these applications include running a well-structured database, guaranteeing fast and easy information access, guaranteeing data protection, improving dispatch, presenting good interactivity and flexibility. These systems can be used to optimize vehicles' and drivers' distribution, statistical analysis for fleet management optimization, report the mileage, fuel consumption, route rotations, driver performance, driver assignment, vehicle maintenance problems, and driver status. [7]

In many cases, the systems are designed as online, real-time data-based systems with the use of techniques such as Global Positioning System (GPS) and Location-Based Services (LBS).

The Internet Multimedia Subsystem (IMS) proposed in [14] is a server-based IMS application where a telecommunication architecture is used. The system assists transport companies to improve their fleet management by tracking vehicle position or by defining the route of movement. This application allows transport companies to monitor their vehicles by continuously updating the location of the vehicle which is constantly visible

on the web interface. For this purpose, existing mobile devices, or other specialized devices which have the ability to use SMS (Short Message Service) and GPRS (General Packet Radio Services) data channels are used.

Hamed Saghaei also discusses an approach for controlling and monitoring a fleet management system in his research [1] using a GPS/GLONASS (Global Positioning System/Global Navigation Satellite System) Tracker. He presents his solution as a web-based software called PayaRadyab. The design of this system aims to support the decision-making process of transportation based business organizations by tracking the locations of the vehicles, monitoring drivers' movement, and by evaluating every event of their fleet in real-time. PayaRadyab can capture the exact position of the desired vehicle on the different server-based maps and provide detailed reports, traveled path, fuel consumption rate, speed limits, and other necessary information. With the use of GPS/GLONASS-based automatic vehicle locators called Rad100, the system can provide positioning with high accuracy. Rad100 is able to determine the vehicle's coordinates and transfer them via GPRS or SMS of the GSM network towards a defined web server. The GPS information picked up by Rad100 can be dealt with active trackers and provide a live picture of where the vehicle is located.

The fleet management system which is presented in [15] collects driver and vehicle event data and allows fleet managers to view reports by processing and analyzing these data. Fleet managers are able to monitor vehicular and driver event data and evaluate driver performance and correct or reward driver behavior. In this research different aspects such as web-based applications, image capturing techniques, and vehicular embedded systems are considered to achieve the expected objectives. One aspect is the web-based application that provides a portal and user interface. The web-based application allows users to view pages showing vehicle and driver event details, integrated event "snapshots", pre-event and post-event videos, and maps, along with all of the above together in various selectable combinations. [15] Another aspect discussed in the research is configuring capture devices in fleet vehicles to record driver and vehicle event data and transmit it into a format for analysis by an end-user. For this purpose, the system has used the vehicle-mounted device

to capture images from the front of a vehicle and to record data associated with a particular driver and vehicle in the fleet.

The fleet management system describes in [16] is consists of a central system that is based on a server system that consists of several databases and application servers. The system functions as a user system including user computers that are connected to the server system and the user software. The user software can be a stand-alone software with a network database connection or web application. According to [16] web-based solutions are dominant because of the significant advantages and rapid development of the web industry and related technologies. The user system presented in [16] focuses on vehicle maintenance, vehicle tracking and diagnostic, fuel management, driver management, health, and safety management, and data displaying, querying, and reporting. This system handles a large amount of data. The system process and analyze these data to generate journey logs, vehicle parameter reports, and histograms, driver style classification, fuel consumption, and event logs. These outputs can be queried for specific vehicles, drivers (or groups) in specific periods. With the use of the system, transport business organizations can expect to decrease in the overall operating costs of the company, increase in vehicle utilization and optimal route planning, the necessity of the replacement and supports the logistics solutions, improve the fuel consumption, improve the efficiency of transport, decrease of extra charges caused by the delays.

According to research [8], a big fleet cannot be managed using an online tracking system. Statistical analysis of data stored in a database is useful for fleet management optimization.

The routing system is also the main component in a fleet management system which supports organizations in the selection of the routes, optimization of the number of vehicles, control of drivers, and flexibility of the carrier to possible user requirements. [9]

The organizations have identified the benefits of implementing such a measure in monitoring the working hours of the vehicle, the possibility of rerouting the vehicle, and the possibility of using the system for vehicle routing. [9]

According to [9] several important elements such as analysis of the existing condition of the fleet and its organization, defining the priorities in fleet management within the company, analysis of the possibility for adjustments in the software to specific characteristics of the company, and analysis of the possibility for further upgrading of the system in the future should be considered when introducing an information technology support system for a fleet organization. When designing a fleet management system several operating characteristics should be considered such as fleet size, operating range, routing variability, and time sensitivity. [10]

Fleet Size

The size of the fleet is defined by the number of units in the fleet. Size is categorized as small (1 – 19 vehicles), medium (20-99 vehicles), large (100-499 vehicles), and very large (500 or more vehicles).

Operating Range

This represents the operating range of the fleet. Fleets that operate within a single city are categorized as "Local fleets". Fleets that operate beyond their local areas, but remain primarily within a single region of the country are categorized as "Regional Fleets". Fleets whose operating ranges extend across multiple regions are categorized as "National Fleets". [10]

Routing Variability

Fleets can have either fixed or variable routines. Fixed-route fleets operate on the same routes every day and typically, have regular pick-up and delivery schedules. By contrast, variable-route carriers are subject to frequent changes in routing and scheduling.

Time Sensitivity

The urgency of the operations handled by the fleet vehicles is referred to as "Time Sensitivity". This factor should be considered by time-sensitive carriers who are expected to make deliveries according to strict time schedules in their fleet systems.

Suggested decision factors are maximizing revenue per mile traveled, maximizing revenue per trip, minimizing operating costs, vehicle and equipment availability, driver availability, driver service hours, fuel conservation, etc. These decision factors support route planning and dispatching to ensure the fleet vehicles and drivers are allocated to maximize the satisfaction of the system requirements.

The factor, maximizing revenue per mile is used by organizations that carefully plan routes and dispatch schedules minimizing total mileage and minimizing mileage between two locations. This is mainly considered by organizations that have adopted automated routing, dispatching, and load planning systems. [10]

Maximization of revenue per trip is important to organizations that are seeking to improve per-trip revenue regardless of trip distance and who are less concerned about choosing routes that minimize mileage. [10]

Equipment availability is an important fleet management decision factor for organizations that handle large volumes of long-distance, line-haul shipments that are often dispersed across wide geographic areas. Vehicle availability is very important for small and medium fleets when assigning vehicles to a single route or shipment. Fleet management decisions such as sequenced routing, handling several shipments on a single run, hiring outside from the fleet can be made using this factor. [10]

The minimization of overall fleet operating costs is an important consideration in fleet management decisions. Routine maintenance, conserving fuel, engine performance, driving patterns, and other factors help to control fleet costs. [10]

Driver service hour is also a significant factor to be considered in fleet management decisions. Drivers mustn't exceed legal hours-of-service limits [10] and if routing distances are different, drivers should not be repeatedly allocated for the same distances. Rotation for driver allocation can be used to avoid such situations.

Fuel conservation is also very important in fleet systems. Selecting routes to minimize mileage is one way for carriers to conserve fuel. [10]

Assigning vehicles with high fuel economy for long distances and vehicles with low fuel economy to short distances can help in conserving fuel. However, optimizing the performance of the vehicles results in a gain in fuel efficiency.

Vehicle performance is described based on power performance, fuel economy, breakdowns, durability, and safety. [12] According to [11] vehicle speed, air/fuel ratio, intake airflow, Revolutions per minute(rpm), engine load, accelerator pedal position, lambda sensor voltage, catalyst temperature, and absolute load parameters are identified as key variables.

Performance Monitoring

Performance monitoring is a systematic approach to evaluating vehicle capabilities and limitations. The process involves improving productivity and competence and is used to solve problems related to the performance of the fleet.

Vehicle breakdowns have a high impact on vehicle performance. A vehicle breakdown is a type of unplanned incident where a vehicle fails during operation on a roadway and is forced to stop [18]. According to research [17], vehicle breakdown is the mechanical failure that impedes or even completely stops the vehicle's operation. Flat battery, fuel pressure problems, faulty electrical wiring, tire puncture, driver error, and many other reasons can cause a vehicle breakdown. It is an unusual event, and many drivers are uncertain about how to respond to the failure. This can lead to unsafe behavior and, in some cases, secondary incidents. [18]

A review of past studies indicates that breakdowns are the major proportion of road incidents. Therefore to improve traffic incident management strategies various incident duration analyses and prediction models have been developed in past researches [17]. If generally categorized the methodologies applied in previous literature most can be identified as Regression models including linear regression (Garib et al., 1997; Peeta et al., 2000; Valenti et al., 2010), quantile regression (Khattak et al., 2016), Fuzzy system model (Dimitriou and Vlahogianni, 2015; Kim and Choi, 2001; Wang et al., 2012), Classification tree method (Knibbe et al., 2006; Zhan et al., 2011), Artificial neural network (Lee and

Wei, 2010; Pereira et al., 2013), Bayesian network (Ozbay and Noyan, 2006; Shen and Huang, 2011), Support vector machine (Zong et al., 2013) [17]

In the research [17] the incident data drawn from different sources, including closed-circuit television (CCTV), InterCAD Emergency Messaging283System (ICEMS), towing logs, police reports, manual and operator entries from the Greater Sydney Metropolitan Area in Australia for 4.5 years from January 2012 to June 2016 used as the data source. The historical incident dataset consisted of over 240,000 records of unplanned incidents. Breakdown records were 29% of the whole dataset and there are 73,084. After preprocessing the data few lack-related data records were removed and not considered for the analysis, and the final dataset consisted of 72,764 breakdown records.

The factor analysis of the research [17] is done using the factor analysis method in SPSS statistical software. Each category of independent variables is analyzed separately and a linear dimension reduction approach is used to resolve the multicollinearity issue because a relatively high correlation exists among some variables. To define probabilistic loadings for independent variable analysis a Varimax Rotation with Kaiser Normalization is conducted. For all the factors the Variance Inflation Factor (VIF) was calculated and the maximum value found was 5 which suggesting that multicollinearity is not an issue.

The modeling results consist of the estimated coefficients, p-values, and the overall goodness of fit, including log-likelihood [17]. In the final comparison, the most appropriate FP (log-logistic) and LC(log-normal) models are reported. LC modal had a better fit than the FP model, with all factors significant when FP has only 4 out of 9 factors significant.

Vehicles aged less than 10 years had a negative relation with breakdowns and vehicles aged greater than 10 years had a positive impact to have quick repairs on the roadside. Furthermore heavy vehicles had a positive effect and rainy days are observed to have a negative impact on breakdown duration.

The case study “Analysis of Vehicle Breakdown Frequency” is done in New South Wales, Australia to identify critical variables that affect the number of reported vehicle breakdowns.[18] a random-effect negative binomial regression model is used to find out

the impact of different independent variables on the number of breakdowns. For the analysis historical incident data for 4.5 years are used. 90,159 records of reported on-road vehicle breakdowns in the dataset. 50.8% of breakdowns involved light vehicles, 24.8% involved heavy vehicles and the remaining observations did not specify the vehicle type [18]. There were correlations over time and space which can violate the assumptions of ordinary least squares regression and misestimate the errors on the model coefficients. Therefore random-effect (RE) and random-parameter (RP) models are used to resolve these correlations. The. By including all the potential independent variables, a simple multiple linear regression model was built and the multicollinearity issue was addressed by calculating the variance inflation factors (VIFs), and the variables with VIFs exceeding 10 were omitted. To explore model specification then a reverse stepwise approach was used. From the VIF procedure, all the variables of interest were included in the model fitting [18]. The least significant variables were dropped and AkaikeInformation Criterion (AIC) and Bayesian Information Criterion (BIC), were evaluated. The coefficients, standard errors, z-values, and continuous variables' elasticities of the significant variables are included in the final modal results.

The interventions to manage the impact of breakdowns proposed by this research are educate drivers about vehicle maintenance procedures, encourage them to read the vehicle manuals thoroughly, conducting routine maintenance and minor repair procedures during the vehicle registration process. Also, it identified some vehicle makes and models that have earned reputations for reliability and fuel economy of some products, which could be required to identify each model's vulnerability to breakdown[18].

Driver performance is very important in fleet management, thus it impacts fuel economy, fleet safety, vehicle utilization, compliance scores, customer satisfaction, and ultimately the success of the business. Different driving styles can be identified by monitoring drivers and those different styles can impact the reduction in fuel costs, prevention of accidents, and reduction in vehicle maintenance cost.

Vehicle and driver performance monitoring with detailed information helps to make better decisions. Driver performance can be calculated using variables such as fuel consumption,

driver idling profile, speeding, rapid acceleration, hard braking, and utilization patterns. Driver performance measurements are based on both quantitative and qualitative data. [13] Amount of fuel, high variation in speed and acceleration pedal, the average amount of accidents, experience in years, and quantitative measurements and capability, safety, and reliability are qualitative measurements used in fleet systems.

Different motor carriers have different fleet characteristics and fleet management objectives. Every motor carrier is unique. Thus, the design of a fleet system should be customer-driven, and logistics-oriented and require diversification as well as flexibility [10].

The critical elements involved in the routing/dispatching process are information used for routing/dispatching, fleet size, amount of driver employment, operational approach, delivery parameters that the fleet provides, parameters associated with performance, and hiring information. [10]

Operating characteristics described in [10] were considered in this research at the designing phase of the proposed automated fleet management system. The AB Securitas has a medium-scaled fleet that operates as a Regional Fleet. Their operations are time-sensitive and the fleet is identified as a fixed-route fleet. Therefore, the system is designed as a non-dynamic database system because there is no requirement for real-time route generation. Also in daily job plan generation, the time factor was considered when allocating task locations and vehicles for the generated routes.

Vehicle availability is a critical decision factor because AB Securitas has a medium-scaled fleet. Often due to the unavailability of company-owned vehicles caused as a result of repairs and other task allocations, the company has to hire additional vehicles from outside suppliers to meet their demand. The system should consider fulfilling daily vehicle requirements by hiring the required amount of vehicles appropriately based on some factors such as supplier trust and quality of the service provided.

Maximization of revenue per trip is extremely important for the company since they are using hired vehicles to fulfill their vehicle requirement. It is essential to manage both

company-owned and hired vehicles in such a way that guarantees a better advantage for the company rather than the supplier. Therefore, the analysis had to be carried out to figure out which routes maximize the profit in order to allocate company-owned vehicles and which routes minimize the loss in order to allocate hired vehicles.

The minimization of overall fleet operating costs is also an important factor to be considered in designing the system. The system is designed to plan routes in such a way as to reduce vehicle maintenance costs, conserve fuel, and reduce costs associated with risk. Vehicle performance is also the main factor to be considered when allocating vehicles for routes. Therefore, the system was designed to filter and allocate vehicles based on their performance.

The system is designed to manage drivers who are employed at the AB Securitas fleet ensuring that they have been allocated to daily operations in a way where the company receives the drivers' maximum service. The allocations are planned in such a way that the hours-of-service limit of drivers is not exceeded and drivers are not assigned to the same job tasks and are allocated with fair rotations.

The decision factors and operational characteristics are important in deciding the applicable fleet management techniques and in constructing the system design and development logic.

2.3 Existing Process Analysis

The CIT unit currently handles all the operations manually. Gathering data, analyzing data, route generating, resource allocating, and route rotation all managed by a manual process. Due to this complexity, many unjustified, unprofitable, and insecure allocations can be identified and there is no centralized controlling facility. This process needs to collect much data from different units. Therefore, it needs more individuals to handle the process, so operation cost is high and this process consumes more time. By reviewing the existing process functional and non-functional requirements identified for the proposed system.

2.4 Functional Requirements

Functional requirements are the requirements that the system should satisfy once developed and what the system is offering. It describes system behavior and its inputs and outputs.

- Own vehicle management module
  the system contains company-owned vehicle data and needs to manage vehicle status. The module is required to check vehicle availability and get a hiring requirement.
- Supplier management module
  Manages supplier data. The module is required to manage the hiring process and generate a supplier allocation report.
- Supplier vehicle management module
  Manage supplier vehicle data.
- Driver management module
  Manage driver data and required for driver allocation.
- Customer management module
  Manage customer data.
- Customer locations management module
  Manage customer locations data. Required in the daily route generation process.
- Special tasks for the fleet management module
  Manage special tasks assigned for the fleet vehicles. Required in vehicle availability generation.
- Fleet vehicle repair records management module
  Records accident/breakdown details of fleet vehicles. Required in vehicle availability generation.
- Route module
  Manage route data. Required in daily route generation and daily vehicle requirement generation.
- Route city management module
  Manage a route and city relationship. Required in daily route generation.

- Task plan management module
  Manage daily tasks.
- Route plan module
  Report daily job results, required in resource rotation.

2.5 Non Functional Requirements

Non-functional requirements are the quality attributes such as performance, usability, security, and reliability that can be used to judge the operation of the developed system. Non-functional requirements do not evaluate the direct services delivered to the user and are used to evaluate system capabilities and constraints rather than its behavior. The system's effectiveness and failures judge based on non-functional requirements.

- Because it contains the CIT data, the system should be secure from hackers, intruders
- The accuracy and reliability of system-generated reports should be high.

2.6 Problem Definition

Various unsolved issues such as not having proper connectivity between the different data sources, taking a long time to detect routes, taking a long time to rotate resources and complexity, not allocating the right vehicles for the routes, have been identified in the literature review. Our attempt was to develop a web-based software tool to overcome the above circumstances by identifying the right vehicles by their performance factors.

2.7 Summary

A comprehensive literature review on the fleet management solutions and techniques acquired and how data mining techniques are used in fleet systems is presented in this chapter. The technologies and tools used for the implementation process of our solution will discuss in the next chapter.

Chapter 3

Adopted Technologies & Tools

3.1 Introduction

The technologies and software tools used for implementing the automated fleet management system and for analyzing the breakdown factors are described in this chapter. In developing the client-side and server-side of the system, HTML5, JavaScript, and PHP languages were mainly used. Maria DB Database Management System and Laravel eloquent were used in data handling. To run the application, Apache2 Server was used as a webserver. Laraval 6 PHP framework was used in developing the system. The SPSS statistical software is used for the breakdown factor analysis.

3.2 PHP Server Side Scripting Language.

PHP stands for Hypertext Preprocessor [15]. PHP is an open-source, general-purpose server-side scripting language that can be embedded with HTML. PHP is created by Rasmus Lerdorf in 1995 for Zend technologies and the current stable version is PHP 7.4.6 [15]. PHP is suited for web development and used to create dynamic web pages.
PHP has many advantages and some of them are listed below.

- PHP is open-source. PHP can be downloaded freely without any cost.
- Availability of free PHP frameworks, tools, and many built-in libraries that can be used to implement different functionalities.
- Platform independence. PHP supports all main operating systems and can be easily integrated with other programming languages and different databases.

- It can be embedded with HTML through PHP special tags. Both HTML and PHP can be written in one file to display dynamic data and to reduce the number of codes.
- Because PHP executes on the server, the view does not display the PHP source code.
- PHP can communicate using different network protocols such as (HTTP, IMAP, POP3, SNMP) through a network.
- Performance and security are high.
- A stable programming language. Bugs and issues are fixed over time in different versions.
- Flexible, simple, and easy to learn.
- Strong support for object-oriented programming.

To develop the web-based fleet management system we have used PHP scripting language by considering the above advantages.

3.3 APACHE2 Web Server

Apache HTTP Server is a free, open-source web server software developed by Apache software foundation initially in 1995. Apache is operatable in different platforms including Linux, Windows, and Unix. [19]

Apache is a module-based software that includes multiple customizable modules for different purposes such as security, load balancing, rewriting, caching, and compressing. Following are some main modules that extend the core functionality of Apache [19]

- mod_ssl - Secure Sockets Layer and Transport Layer Security support module.
- mod_proxy - Proxy module
- mod_rewrite - URL rewriting module

- mod_log_config - Custom log files
- mod_ext_filter – Filtering module

Apache is the most popular and widely used web server software because it has many advantages when compared with other server software.

Apache is a general-purpose HTTP server. [20] There are different modules such as mod_php, mod_python, mod_perl, etc. in Apache that enable the functionality to work with different programming languages such as PHP, Python, Perl, Tcl. [20]

Apache is highly scalable, IPv6-compatible, and supports HTTP/2. [19]

3.4 Maria DB

MariaDB is a free, open-source most popular relational database server developed by Maria DB Foundation initially in 2009. Since it is a fork of MySQL, MariaDB is highly compatible with MySQL [21]. MariaDB is open-source. The closed source features introduced in MySQL enterprise edition are freely available in MariaDB open-source version. Also, MariaDB publicly releases all bugs and development plans. They provide test cases for all fixed bugs that oracle does not provide with MySQL.

Faster and improved performance in character set conversions, shut down time, debugging, and overall speed. It supports system versioned tables and its queries can access both historic and current data. MariaDB runs on different operating systems, to provide advanced clustering, uses a standard and popular querying language, and supports a wide variety of programming languages. [22]

3.5 Adopted Software Tools

GRUNT - GRUNT is a JavaScript task runner created by Ben Alman. Grunt wrote in Node.js and distributed via NPM (Node Package Manager). Frequent tasks such as compiling, unit testing, minifying, and linting can be automated using the Grunt tool. Custom tasks can be defined in a separate file known as a Gruntfile. Those tasks can be run using a command-line interface. The user is able to use plugins in the Grunt ecosystem. [24]

Advantages of GRUNT

- Consistency, effectiveness, efficiency, repeatability due to automated/predefined task queries.
- Access to many predefined plugins.
- Allows developers to build their plugins and publish them to npm so that others can also use them in their developments.

SPSS – SPSS is a software package that provides advanced statistical analysis, text analysis, big data integration, data preprocessing features, open-source extensibility, and a machine learning algorithm library. SPSS is widely used for both qualitative and quantitative data analysis by market researchers and survey companies.

Bivariate statistics such as t-test, ANOVA, Means, Correlation, Nonparametric tests, Bayesian

Prediction, Linear regression, Descriptive statistics, Descriptive Ratio Statistics, Factor analysis, cluster analysis, and many statistics included in the SPSS package. Data can be import to SPSS from many types of files and can use to generate reports, charts, and plots of distributions and trends and also to conduct complex statistical analyses. SPSS is user-friendly, scalable, and flexible and with simple instructions, perform complex data manipulation and analysis. [24]

Microsoft Excel: Microsoft Excel is a leading spreadsheet program use for calculations, data storing and visualization, creating graphs, charts, pivot tables, and plots. It also provides a Visual Basic macro programming language feature. Excel also use as an analysis tool and a graphing tool.

3.6 LARAVAL

Laravel is a web framework written in PHP. Laravel is free, open-source, and based on symphony which is also a free PHP framework that contains reusable components. Laravel follows Model-View-Controller architecture. [25]

Laravel has many features:

Bundles: Laravel provides modular packaging systems with bundled features that can automate the installing, upgrading, removing process of the framework. [25]

Composer: A dependency manager to add Laravel specific PHP packages from a repository. [26]

Eloquent ORM: Eloquent object-relational mapping is an advanced PHP implementation to create relationships between database objects. Eloquent ORM presents a "Model" to interact with each database table. [25]

Query builder: An interface to create and run database queries by using a set of classes and methods. This uses PDO parameter binding to protect from SQL injections and work with many database systems. [25]

Schema builder: Schema builder enables database table manipulation using the "Schema" class. [26]

Migrations: Laravel migrations easily manage database schema with schema builder. It provides database version control features. Allows to change database layout and update the state. [25]

Seeding: Laravel seeding allows to import data to a database using the "Seed" class. [25]

Blade: Blade is the template engine that provides the view of the application. [25]

Artisan: Artisan is the Laravel command-line interface driven by the Symfony console component. It provides several commands to create and process the Laravel project. [25]

Packages: Laravel composer offers many ready to use packages such as Cashier, Scout, Dusk, Envoy, Sanctum, Horizon. [25]

Core Extensions: To customize or to replace core components Laravel provides many extensions. [25]

Security: Laravel provides many security features such as authentication, authorization, email verification, encryption, hashing to ensure security within the application.

CSRF Protection: Protects the application from cross-site request forgery which prevents unauthorized commands from being performed by an unauthenticated user. [26] CSRF protection middleware validates every user request made to the application using an automatically generated CSRF token.

3.7 Summary

An automated fleet management system is developed by using PHP server-side scripting language, Laravel framework. This system maintains customer, customer locations, company-owned and supplier-owned vehicles, driver, special tasks, repair, breakdowns master files that hold relevant master data. Therefore MariaDB database management system is used as the system database to store the relevant data. To run the application Apache2 Server is used as a webserver. For the breakdown factor analysis, the SPSS software package and MS Excel spreadsheets are used. The next chapter describes the statistical analysis done to identify the relationship between vehicle features and the breakdown of the vehicle.

Chapter 4

Analysis for Detecting Breakdown Factors

4.1 Introduction

The technologies and tools used to conduct the breakdown factor analysis and to implement the proposed application are explained in Chapter 3. This chapter describes how the data analysis was used to solve the research problem by analyzing the collected relevant fleet data and obtain a result that can be used in the system development.

4.2 Data Mining

Data mining can be defined as the process of extraction of previously unknown, nontrivial, hidden, and potentially useful knowledge from raw data. Data mining is also known as Knowledge Discovery of Data (KDD). There are several steps to be followed in data mining that includes clearly identifying the domain, creating the target dataset, data cleaning and pre-processing, data reduction, selecting the data mining methodology or algorithm, and interpreting identified patterns and knowledge. There are different specific types of learning models in machine learning such as supervised learning, unsupervised learning, and semi-supervised learning. Based on the dataset and the research problem, an algorithm should be trained using a specific learning model.

4.2.1 Supervised Learning

Supervised learning uses labeled datasets to train the algorithm in order to classify data or predict outcomes. In supervised learning, there are several input variables (x) and an output variable (Y) and the algorithm generates the mapping function between the input and output. With the use of the mapping function, it can predict the output variable for a new input data set.

There are two main types of supervised learning; Classification and Regression.

Classification: Used to accurately assign test data into specific categories when the output variable is categorical. KNN algorithm, Decision Trees, Naive Bayes, Random forest, and Bagging are some commonly used classification algorithms.

Regression: Used to understand the relationship between the dependent and independent variables. Linear regression is used when the output variable is numerical while logistic regression is used when the output variable is categorical. Polynomial regression is also another popular regression algorithm.

4.2.2 Unsupervised Learning

Unsupervised learning is used to analyze and cluster unlabeled data sets when only input variables are present with no output variable. Unsupervised learning can be classified as clustering and association and is useful when the result is unknown.

Clustering: Discover the inherent groupings in the data. The clustering model looks for similar data and groups them together. K-means clustering algorithm is an example of a clustering algorithm.
Association: Used to discover rules that describe large data portions. eg. Apriori Algorithm

4.2.3 Semi-Supervised Learning

Used for a large number of unlabeled data points with a small subset of labeled data. This is a mixture of supervised and unsupervised techniques and useful when there is a high volume of data from which it is difficult to extract relevant features clearly.

4.3 Why Linear Regression

Linear regression is a supervised learning technique used to reproduce the output value by processing historical data. Regression analysis describes the relationships between a set of variables. The dependent variable is the main attribute that should be predicted and the independent variable(s) are the attributes that have an impact on the dependent variable. The dependent variable should be a continuous variable and the independent variable(s) can be either categorical or continuous. Linear models build a linear relationship between the input and output parameters and produce an equation to represent the relationship between each independent variable and the dependent variable by using coefficient values. The generated equation can be used in further predictions on new data. Regression analysis considers which attributes matter the most to the model, which have the least impact and which variables should be ignored, and how each variable interacts with each other, to generate the outcome. There are two types of linear regression, simple linear regression, and multiple linear regression. If there is a single independent variable to predict the dependent variable, simple linear regression is used and if there are two or more independent variables, multiple linear regression is used to predict the dependent variable. Regression is suitable to predict an effect and to determine the strength of the predictors.

The analysis in this research is based on historical and labelled data with multiple continuous and categorical variables. The main objective was to identify the relationship between the performance of a vehicle (represented by breakdowns of a vehicle in the data) and the other factors relevant to a vehicle. The dependent variable is not categorical and there is no requirement in classifying data into specific classes. Therefore, regression was chosen instead of classification. The correlation method is also used to identify the associations between two continuous variables but the regression makes a distinction between independent and dependent variables that can be used in future predictions which are not achieved by the correlation method. Therefore, linear regression is chosen instead of correlation. The dataset includes multiple independent variables, thus multiple linear regression is the applicable method for this analysis.

4.4 Description of Data

Existing data of the vehicle fleet were collected for this research study. The dataset consisted of 105 entries and the variables were as follows.

V_No: Unique vehicle number, given to each vehicle.
Id: Special identification, given to each vehicle.
Registered_Year: The vehicle registered year.
Make: The brand of the vehicle.
Model: The name of the vehicle product.
Fuel_Category: The energy source of the vehicle.
Engine_Capacity: The measurement of the total volume of fuel and air that can be pushed through the cylinders in the engine which is measured by cubic centimeters.
Year_Manufactured: The production year of the vehicle.
Maintenance_Expenses_2020: The amount required to maintain the vehicle throughout the year 2020 (in Rupees).
Fuel_Usage_2020: The average fuel usage of the vehicle within a day in Rupees.
Fuel_Economy: The distance travelled per unit volume of fuel, measured as kilometers per liter.
Milage_a_Day_2020: The average distance that the vehicle has travelled within a day (in kilometers).
Milage_2020_Days300: The average distance that the vehicle has travelled within the year (in kilometers).
No_of_Services: Number of times the vehicle has completed its periodic services to ensure it is roadworthy.
Breakdowns: The number of times that the vehicle had mechanical failures.
Age: The number of the years in between the calendar year and the model year.
Battery_Replace_Delayed_in_Days: The days delayed to install or replace the vehicle battery.
Tire_Replace_Delayed_km: The number of kilometers travelled after exceeding the distance that should replace the tires of the vehicle.

Categorical Variables	Continuous Variables
V_No	Fuel_Economy
Id	Milage_a_Day_2020
Registered_Year	Milage_2020_Days300
Make	No_of_Services
Model	Breakdowns
Fuel_Category	Age
Engine_Capacity	Battery_Replace_Delayed_in_Days
Year_Manufactured	Tire_Replace_Delayed_km
	Maintenance_Expenses_2020
	Fuel_Usage_2020

Table 1 Categorical and Continuous Variables

4.5 Data Pre-processing

Data pre-processing is an essential step in the model development process. Data pre-processing is used to transform raw data into useful and efficient formats. The data can be in different formats and there might be issues with the collected data due to human errors, limitations, or flaws in the data collection process. Therefore, preparing incomplete or inconsistent raw data for further processing using data pre-processing is a major step in the data mining process. The main techniques associated with data pre-processing are data cleaning, data transformation, and data reduction.

Data cleaning involves handling missing data, noisy data, and duplicates of the data. Data transformation involves discretization, attribute selection, and normalization. Data reduction is handled by techniques like data cube aggregation, numerosity reduction, and dimensionality reduction.

The collected data set includes records for each vehicle in the fleet but the data set is not huge and complex. The data set does not include missing values or noisy data. Also, there are no duplicate entries. If similarities are found between two independent variables those variables are removed from the dataset because that can cause multicollinearity. Thus, the manufactured_year variable was removed from the dataset because there is a relationship between the age of the vehicle and the manufactured_year.

Some data in the dataset has been recorded for better organization of data. Fuel Category and Make attributes were recoded into integer values from 1 to 2 and Model recoded into integers from 1 to 10.

```
DATASET ACTIVATE DataSet6.
RECODE MODEL ('CR42'='1') ('DA17A'='2') ('E24'='3') ('L300'='4') ('LH162'='5') ('LH171'='6')
    ('LH200'='7') ('QD32'='8') ('VWE25'='9') ('LH172'='10').
EXECUTE.
```

Figure 1 Re-coding the variable 'Model'

Attribute construction is used to create new attributes from the available attributes. The breakdowns attribute is not a fair selection to compare each vehicle because the distance travelled within the year is different for each vehicle. Therefore a new variable named Breakdown per Mileage (BPM) was computed to provide a better meaning.

```
DATASET ACTIVATE DataSet1.
COMPUTE BPMFinal=(Breakdowns / Milage2019Days300)*10000.
EXECUTE.
```

Figure 2 Attribute Construction – BPM

4.6 Descriptive Analysis

A descriptive analysis was done to get an understanding of the variables and the relationships among them.

The following histograms and bar plots were used for this purpose.

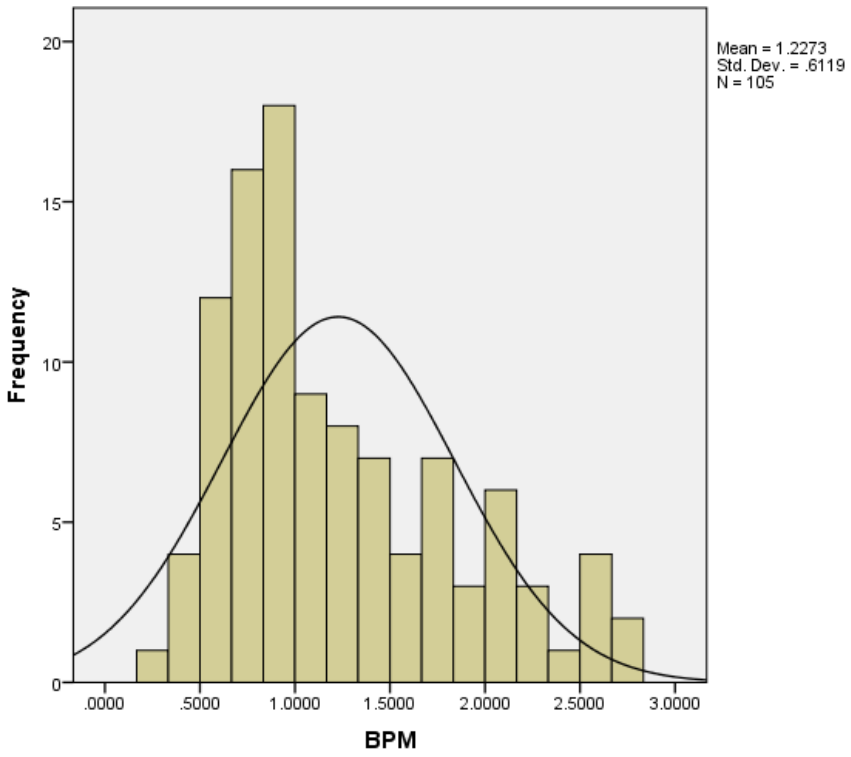


Figure 3 Histogram for BPM

The above histogram depicts the distribution of the variable BPM and it was observed to be slightly skewed to the right.

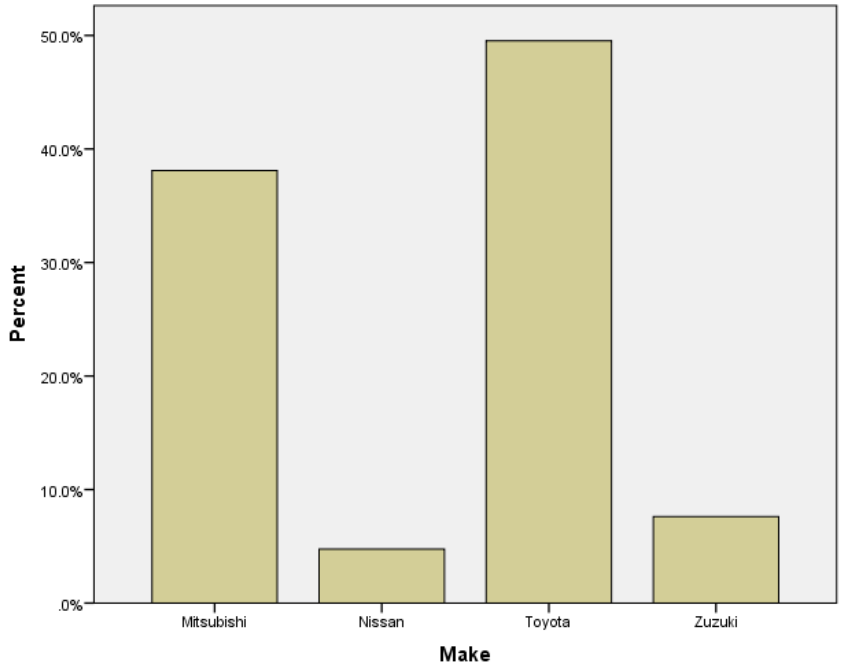


Figure 4 Bar plot for Make

The above bar plot depicts the distribution of the variable Make and the highest percentage was observed for Toyota while the lowest percentage was observed for Nissan.

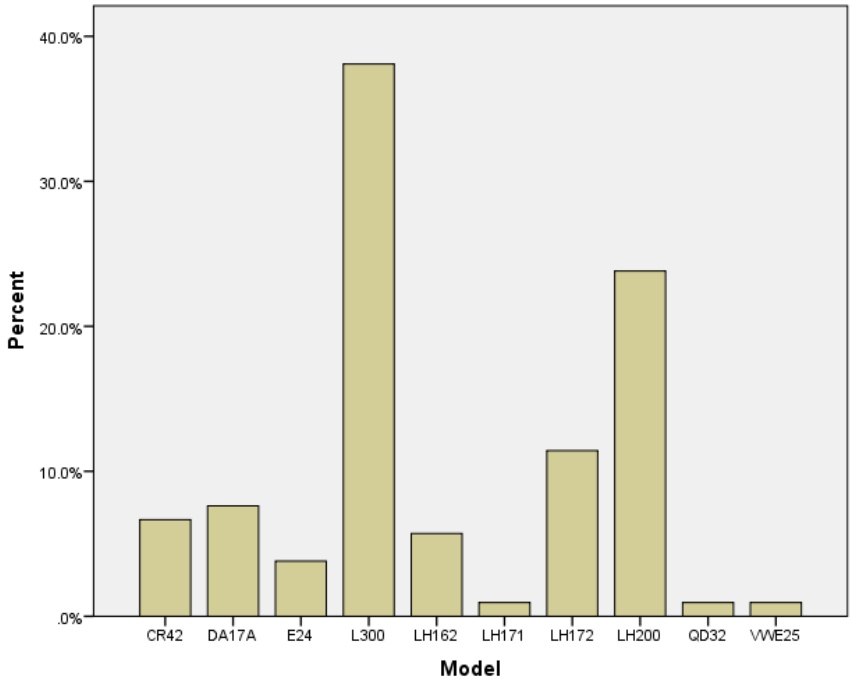


Figure 5 Bar plot for Model

The above bar plot depicts the distribution of the variable Model and the highest percentage was observed for L300 while the lowest percentages were observed for LH171, QD32, and VWE25.

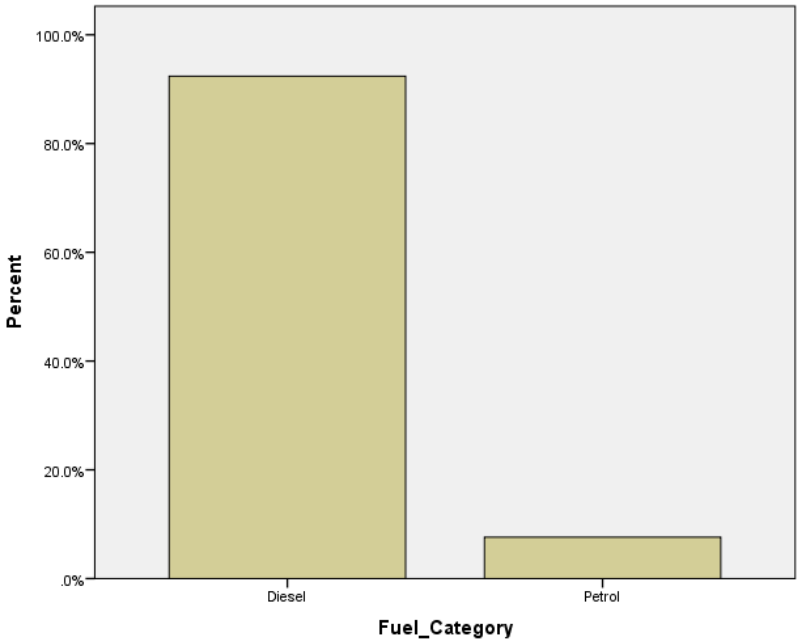


Figure 6 Bar plot for Fuel Category

The above bar plot depicts the distribution of the variable Fuel Category and it was observed that a majority of the vehicles use Diesel as the source of fuel.

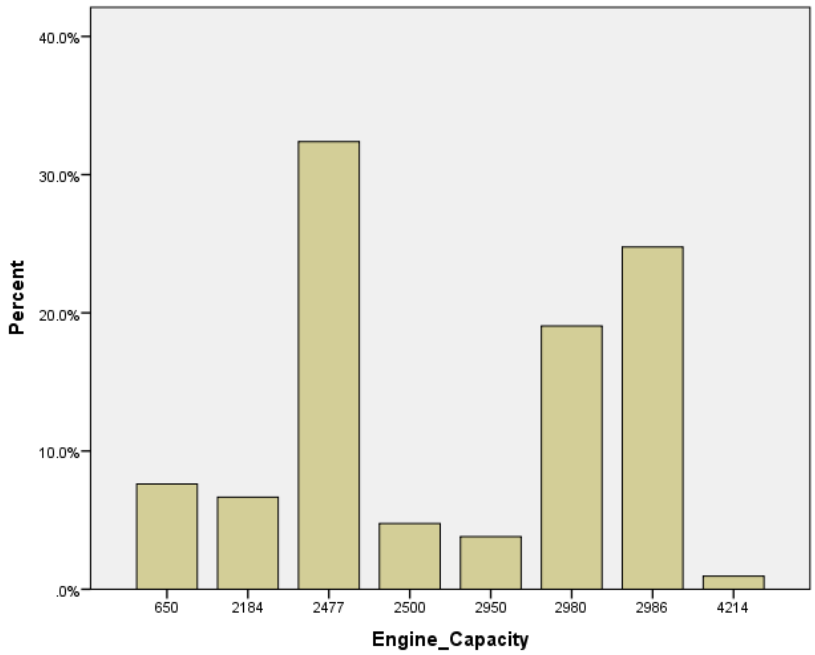


Figure 7 Bar plot for Engine Capacity

The above bar plot depicts the distribution of the variable Engine Capacity and the highest percentage was observed for 2477CC while the lowest percentage was observed for 4214CC.

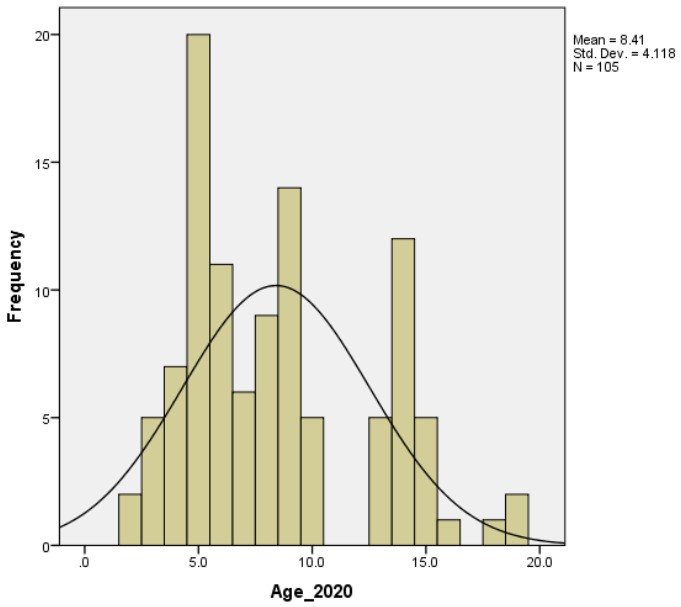


Figure 8 Histogram for Age

The above histogram depicts the distribution of the variable Age and it was observed to be slightly skewed to the right.

The Pearson correlations were calculated between BPM and the remaining numeric variables.

		Tire_Replace_Delayed_km	Maintenance_Expenses_2020	Fuel_Usage_2020	Fuel_Economy	No_of_Services_5000KM	Age_2020	Battery_Replace_Delayed_in_Days
BPM	Pearson Correlation	.707**	-.020	-.437**	-.611**	-.679**	.638**	.574**
	Sig. (2-tailed)	.000	.838	.000	.000	.000	.000	.000
	N	105	105	105	105	105	105	105

Table 2 Pearson Correlation Values

Positive correlations above 0.5 were observed for Tire_Replace_Delayed_km, Battery_Replace_Delayed_in_Days, and Age with BPM. Negative correlations above 0.5 (less than -0.5) were observed for Fuel economy and No_of_Services_5000KM with BPM. However, low correlation values were observed for Maintenance expenses and Fuel usage with BPM.

4.7 Advanced Analysis

Multiple linear regression is used to explain the relationship between one dependent variable and two or more independent variables. Further, multiple linear regression models can explain the value of the dependent variable at a certain value of independent variables.

Out of the 105 observations in the data set, 80% were used to build the model while 20% were used to validate the model. Hence, the model was developed using 84 observations. The developed model was applied to the remaining 21 observations in the validation set and the predicted values were obtained. Then the sum of squares of the difference between the predicted and actual was calculated and this was used as a measure to assess the model.

Before proceeding to the analysis, it is essential to identify if the data can be analyzed using multiple linear regression. Therefore before proceeding to the analysis, the data should be evaluated against the assumptions under multiple linear regression. These assumptions must be satisfied to generate a valid result by regression analysis.

To apply multiple linear regression, the dependent variable should be continuous and there should be two or more independent variables that can be either categorical or continuous.

Between the dependent variable and each independent variable, there should be a linear relationship. The linear relationship can be tested by creating scatterplots and partial regression plots or calculating Pearson correlation. The relationships between BPM and the independent variables were calculated by using Pearson correlation and to further understand the correlation of low degree correlations, scatter plots were drawn.

		Tire_Replace_Delayed_km	Maintenance_Expenses_2020	Fuel_Usage_2020	Fuel_Economy	No_of_Services_5000KM	Age_2020	Battery_Replace_Delayed_in_Days
BPM	Pearson Correlation	.707**	-.020	-.437**	-.611**	-.679**	.638**	.574**
	Sig. (2-tailed)	.000	.838	.000	.000	.000	.000	.000
	N	105	105	105	105	105	105	105

Table 3 Pearson Correlation of BPM and independent variables

Positive correlations above 0.5 were observed for Tire_Replace_Delayed_km, Battery_Replace_Delayed_in_Days, and Age with BPM.

Negative correlations above 0.5 (less than -0.5) were observed for Fuel economy and No_of_Services_5000KM with BPM. However, low correlation values were observed for Maintenance expenses and Fuel usage with BPM.

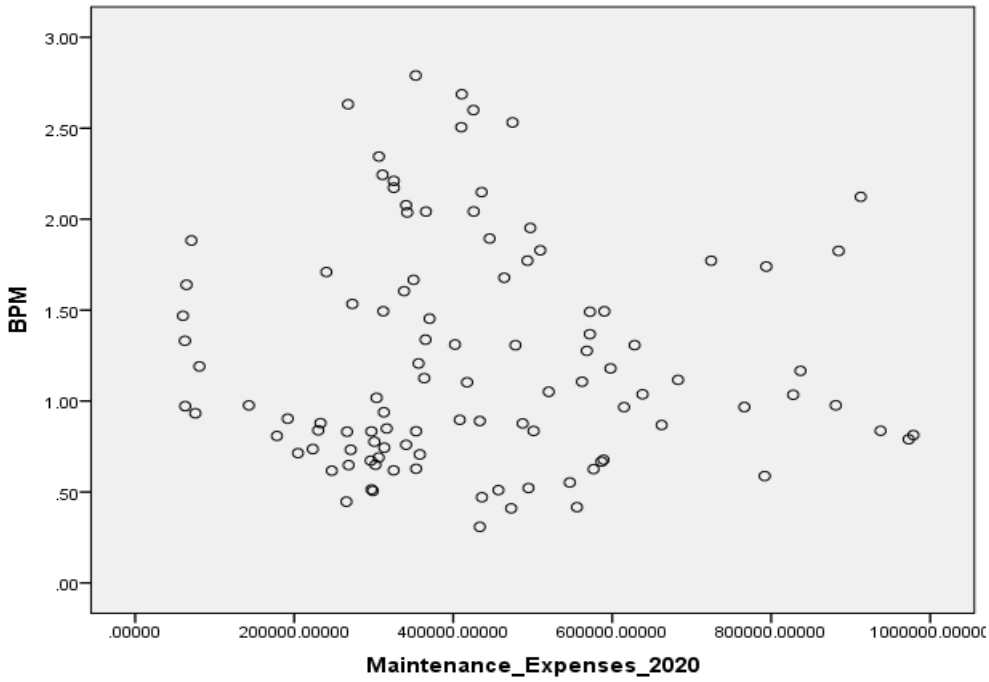


Figure 9 Scatter Plot for Maintenance expenses

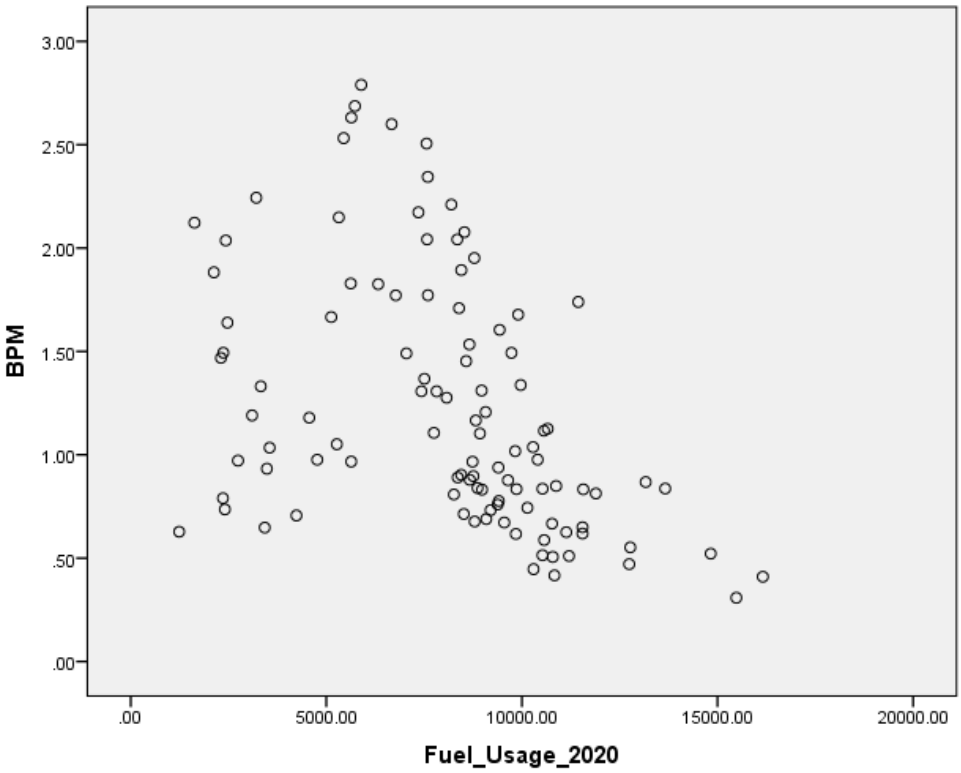


Figure 10 Scatterplot for Fuel Usage

Thus those two variables were transformed by applying data discretization. [30] Binning method is used to arrange these data into bins by creating binning intervals. Maintenance expenses and Fuel usage were arranged into bins as ‘high’ and ‘low’ by using the median value.
(high = greater than median and low = lower than the median).

The dataset should be tested for autocorrelation which measures the relationship between a variable's past values and its current value. The autocorrelation can be tested by Durbin-Watson statistics. The result of the Durbin-Watson statistic should be between 0 and 4. A positive autocorrelation is represented by values 0 to 2 and negative autocorrelation is represented by values 2 to 4. Value 2.0 represents that there is no autocorrelation detected in the dataset. [29]

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.937[a]	.878	.869	.2237531	1.628

Table 4 Durbin Watson Value of the model

A value of 1.628 was observed under the Durbin-Watson statistic and a positive autocorrelation is identified for the model. Therefore the autocorrelation assumption was not violated by the model.

The data should be tested for multicollinearity to identify whether there is a relationship among the independent variables of the model. The similarities between independent variables can cause correlations [30]. A good model ideally should not have multicollinearity. Multicollinearity can be tested by using the VIF value for each variable. The VIF values should lie between 1 to 10 and if it is less than 1 or greater than 10, it is evident that multicollinearity exists. The VIF values for each variable in the model lies between 1-10 and therefore, the model does not violate the multicollinearity assumption. [29]

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	-.329	.461		-.713	.478		
	Model	-.023	.011	-.093	-2.137	.036	.834	1.199
	Fuel_Usage_2020	.207	.071	.168	2.908	.005	.473	2.116
	Age_2020	.074	.010	.466	7.687	.000	.430	2.323
	Battery_Replace_Delayed_in_Days	.001	.000	.123	2.651	.010	.731	1.368
	Tire_Replace_Delayed_km	7.411E-5	.000	.179	2.698	.009	.357	2.799
	No_of_Services_5000KM	-.018	.003	-.355	-5.348	.000	.358	2.790

Table 5 VIF values for independent variables

The model should show homoscedasticity. Homoscedasticity can be tested by creating a scatter plot of regression standardized residuals versus regression standardized predicted values. The scatterplot should not show any defined shape such as a parabola or a distinctive pattern. The scatterplot of residuals against predicted values for the model does not show any clear pattern and therefore, the model is not violating the homoscedasticity assumption.

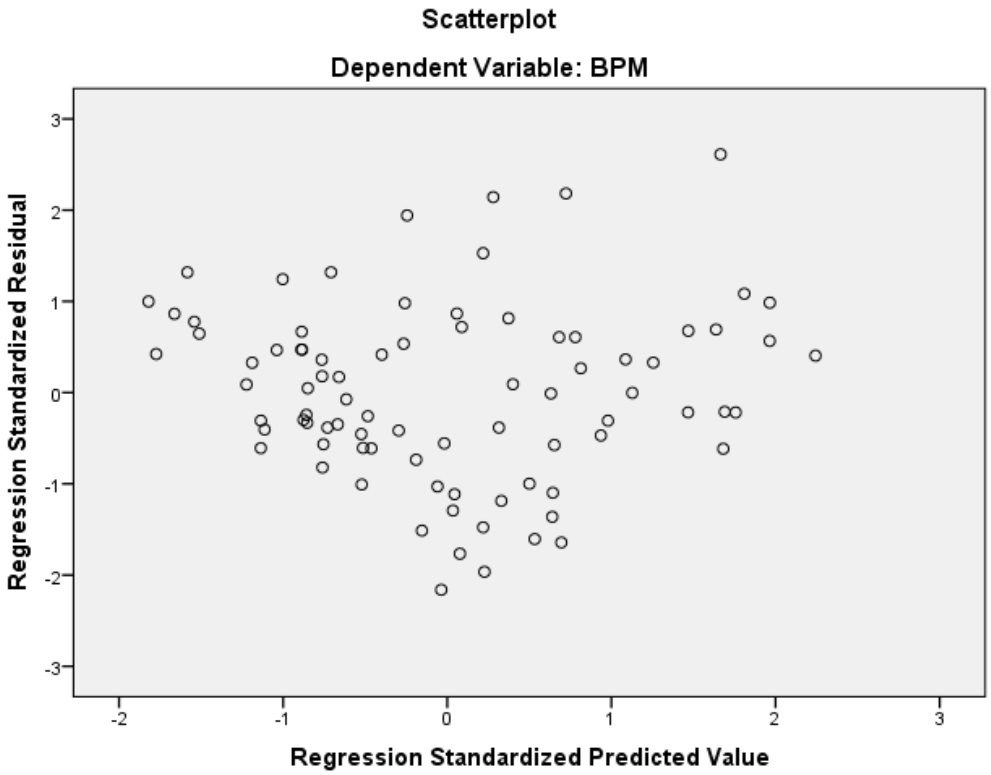


Figure 11 Residuals vs Predicted Values

The data set should be tested for any unusual points which can cause different impacts on the regression line. These unusual points are outliers, high leverage points, or highly influential points which can reduce the predictive accuracy of the model. The outliers and leverage points in the model are tested using Centred Leverage Value. The centered leverage value is calculated as 3 * (1+k) / n where k is the number of independent variables and n is the number of data records. [29] The calculated value for centered leverage value is (3 * (1+6) / 84) = 0.25 and no observations have leverage values above 0.25.

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Stud. Residual	-2.252	2.781	.000	1.006	84
Deleted Residual	-.5248783	.6625655	.0001110	.2353843	84
Stud. Deleted Residual	-2.315	2.913	.001	1.021	84
Mahal. Distance	1.913	12.109	5.929	2.419	84
Cook's Distance	.000	.148	.013	.023	84
Centered Leverage Value	.023	.146	.071	.029	84

Table 6 Centered Leverage Value

Finally, it should be tested whether the residuals (errors) are approximately normally distributed. This can be tested by Normal P-P Plot and Normal Q-Q Plot of the standardized residuals. The relationship should be approximately linear without exception. It can be seen from the below plot that the observations for the model lie approximately in line with the 45-degree line and therefore this assumption is not violated.

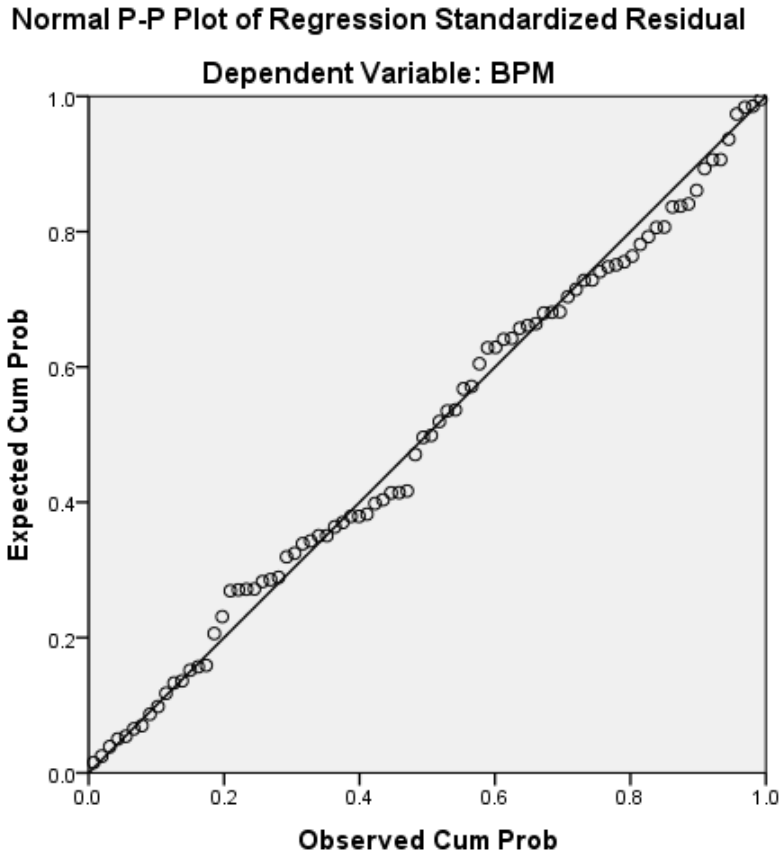


Figure 12 Normal P-P Plot

4.8 Model Summary

R squared value measure for the linear regression model and percentage of the R square. It indicates the degree to which percentage of the variance of the dependent variable is explained by independent variables. R square value measures the strength of the relationship between created model and the dependent variable. [28]

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.937[a]	.878	.869	.2237531	1.628

a. Predictors: (Constant), No_of_Services_5000KM, Age_2020, Model, Battery_Replace_Delayed_in_Days, Fuel_Usage_2020, Tire_Replace_Delayed_km
b. Dependent Variable: BPM

Table 7 Model Summary

According to the regression analysis, the R squared value was recorded as 87.8% and the adjusted R squared value was recorded as 86.9%. When looking at the R squared value, it shows 87.8% of the variance of breakdown per mileage explained by vehicle model, fuel usage, number of services, age, and battery replace delayed in days, and tire replaced delayed in km. As per the value of R squared, this model is well fitted and is suitable for prediction. Furthermore, 12.2% of the variance of break down per mileage has been explained by other influencing factors which are not covered by the current study.

4.8.1 Multiple Linear Regression Model

In the model developed first, all the p-values except for variables make, fuel category, engine capacity, maintenance expenses, and fuel economy were less than 0.05 significance level. It implies vehicle model, fuel usage, number of services, age, battery replacement delayed in days, and tire replacement delayed in km are significant variables to the breakdown per mileage and should be represented in the final regression model.

Therefore, the final model was developed as follows.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.329	.461		-.713	.478		
	Model	-.023	.011	-.093	-2.137	.036	.834	1.199
	Fuel_Usage_2020	.207	.071	.168	2.908	.005	.473	2.116
	Age_2020	.074	.010	.466	7.687	.000	.430	2.323
	Battery_Replace_Delayed_in_Days	.001	.000	.123	2.651	.010	.731	1.368
	Tire_Replace_Delayed_km	7.411E-5	.000	.179	2.698	.009	.357	2.799
	No_of_Services_5000KM	-.018	.003	-.355	-5.348	.000	.358	2.790

a. Dependent Variable: BPM

Table 8 Standardized Coefficients

$$Y= X_0+X_1+X_2 + X_3+ X_4+ X_5+ X_6+\varepsilon$$

$$Y=-0.329-0.093X_1+0.168X_2-0.355X_3+ 0.466X_4+0.123X_5+0.179+\varepsilon$$

The unstandardized beta (ß), this value represents the slope of the line between the dependent variable and the independent variable. The standardized beta coefficient indicates the direction of the relationship and impact on the dependent variable. This standardized beta coefficient lies between -1 to +1. If it is close to -1 or +1 it implies a strong association with the dependent variable and independent variables. Variables of model and number of services have negative impact and other impacted variables such as

fuel usage, age, battery replace delayed in days and tire replace delayed in km have a positive impact on breakdown mileage per day.

The intercept term indicates the value which the dependent variable takes when values of all independent variables are zero.

Beta value for the number of services is -0.355 indicates that increasing one unit of the number of services causes to decrease in the breakdown per mileage by 0.355 units while other independent variables remain constant which means that breakdown per mileage and number of services has an inverse relationship.

On the other hand, the beta value for fuel usage is 0.168 indicates that increasing one unit of fuel usage causes an increase in the breakdown per mileage by 0.168 units while other independent variables remain constant which means that breakdown per mileage and fuel usage has a positive relationship.

Further, the beta value for the age of the vehicle is 0.466 indicates that increasing one unit of vehicle age causes to increase in the breakdown per mileage by 0.466 units while other independent variables remain constant which means that breakdown per mileage and age of vehicle has a positive relationship.

Furthermore, the beta value for battery replace delayed in days is 0.123 indicates that increasing one unit of battery replace delayed in days causes to increase the breakdown per mileage by 0.123 units while other independent variables remain fixed which means that breakdown per mileage and battery replace delayed in days has a direct relationship.

Finally, the beta value for tire replace delayed km is 0.179 indicates that increasing one unit of tire replace delayed in km causes to increase the breakdown per mileage by 0.179 units while other independent variables remain constant which means that breakdown per mileage and tire replace delayed in km has a direct relationship.

4.8.2 ANOVA

ANOVA which is short for "Analysis of Variance" involves the calculation of the number of parameters which are all summarized in a table. In order to measure the overall significance of the model, p-value can be used. The table of analysis of variance was presented to ensure the overall significance of the model as follows.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	27.847	6	4.641	92.701	.000[b]
	Residual	3.855	77	.050		
	Total	31.702	83			

Table 9 ANOVA table

The significance value on the ANOVA table should be less than the confidence level value (α value) for the model to be concluded as a well-fitted one. According to the above table, the p-value is equal to 0.000 and it illustrates that the overall model is well fitted.

As observed in the above table, the residual sum of squares of the model was observed to be 3.855 which is an acceptable indication of the accuracy of the model.

4.9 The Prediction Accuracy of the Model

Regression models' accuracy and performance can be measured in many ways using cross validation or split validation. Cross validation is used to estimate the skill of the created model. To estimate the skill of the model on new data, the K-fold cross validation procedure is used. In this method, a value is selected for the single parameter called K, and data is randomly shuffled into k groups. Validation is done by taking each group as a holdout set and other groups as the training set. The Split validation is validating the model on fixed parts of training and test data. [31] In this research split validation method was applied to test the model accuracy. The model created by using the training data set after applying prepossessing methods and assumptions are validated by using the test data set.

The model was applied to the test sample of 21 observations which was separated before developing the model and predicted values were generated as shown below. Finally, the sum of squares of the difference between the predicted and actual was calculated. The sum of squares value obtained for the test data set was 3.79. The value is comparatively small and therefore it can be assumed that the model is suitable to be used for predictions.

)	Battery_Replace_Delayed_in_Days	Tire_Replace_Delayed_km	BPM	StandardError	PredictedValue	SumofSquare
18	324.0	19797.0	2.0368	.09	2.47	.19
9	309.0	12261.0	1.1162	.12	1.06	.00
15	477.0	17276.0	2.6867	.09	2.47	.05
19	398.0	17935.0	2.1226	.09	2.63	.25
10	165.0	12304.0	1.0342	.13	1.24	.04
16	286.0	19001.0	1.6661	.08	2.48	.66
7	214.0	16066.0	.9671	.04	1.39	.18
5	223.0	13389.0	.4166	.07	.24	.03
5	165.0	16443.0	.8491	.06	.91	.00
9	61.0	16471.0	.6277	.07	1.26	.40
9	190.0	18477.0	1.0512	.06	1.81	.57
9	262.0	13728.0	.6494	.07	.88	.05
7	28.0	13124.0	.7899	.09	.89	.01
5	145.0	15112.0	.6479	.06	.96	.10
4	216.0	15272.0	1.8254	.07	1.49	.12
9	50.0	15572.0	.5879	.06	.74	.02
19	120.0	18684.0	1.3312	.10	2.25	.84
9	135.0	17881.0	1.1792	.09	1.42	.06
7	186.0	15150.0	1.4937	.05	1.43	.00
5	276.0	12759.0	.7063	.10	1.17	.22
7	234.0	15222.0	.8331	.05	.76	.00

Figure 13 Sum of square values

4.10 Summary

This chapter presented a detailed analysis of factors that affecting to the vehicle breakdown. The chapter includes the data used for the analysis, data pre-processing, descriptive analysis, advanced analysis, and the summary of a built modal. The overall design of the developed system and the details related to the system implementation is described in the next chapter.

Chapter 5

Design and Implementation of the System

5.1 Introduction

This chapter will describe the detailed and specific requirements of the project and it explains the initial situation of the system concerning requirement analysis, project goals, and the way the design of the system is developed based on the requirements. Further, the implementation process and factors related to system development are explained.

5.2 Analysis of the System

The research hypothesis is building an automated fleet management system using web technologies to manage the process of the CIT division by acquiring centralized control.

Users

The users of this system are the operational staff of the CIT division. Only authorized users can manage the data of the system.

Input

Data relevant to customers, customer locations, routes, vehicles, drivers, repairs, and breakdowns were the initial master data of the system. Those data are added to the application through the modules created in the admin panel.

Output

The system output is the daily process result. The formatted result is obtained using database queries. The output contains useful data that supports decision-making and analysis.

Process

The system provides different modules with an administration interface. The modules facilitate data insertion, deletion, manipulation, data updating and generates appropriate detailed output for daily routes and resource allocations using the data inserted into the system.

Features

The system manipulates the data inserted into the system and produces accurate output by eliminating the complexity of the manual process. The system effectively manages the CIT division's daily process. Also, the system provides user-friendliness, easy accessibility, and centralized control.

5.3 Top Level Architecture

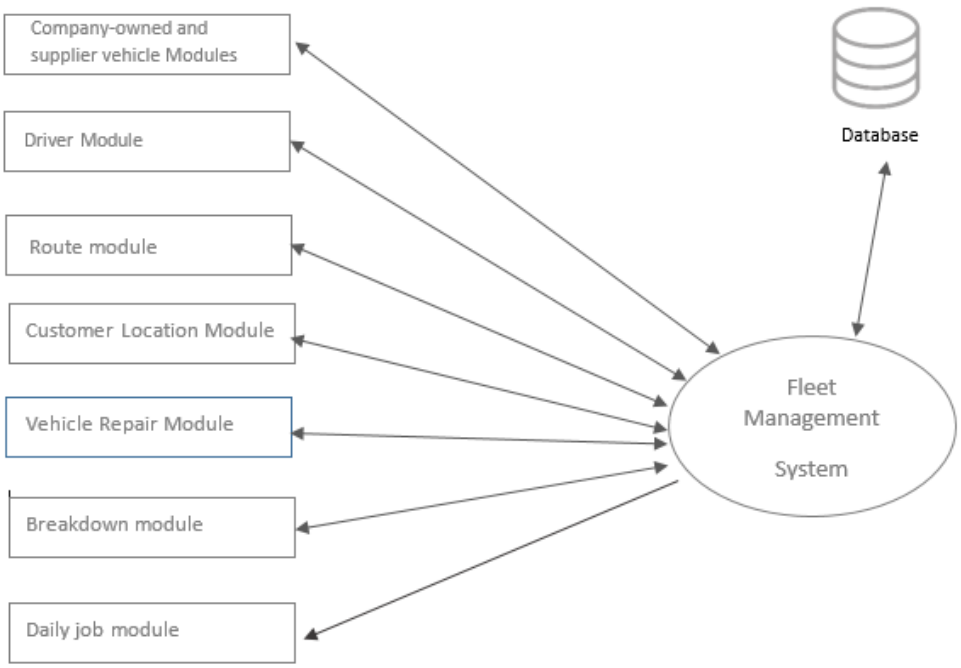


Figure 14 Top-level architecture

The system data flow diagram:

The visual representation of the system processes and data flow with those processes is provided by the data flow diagram. It also illustrates what the system inputs and outputs are.

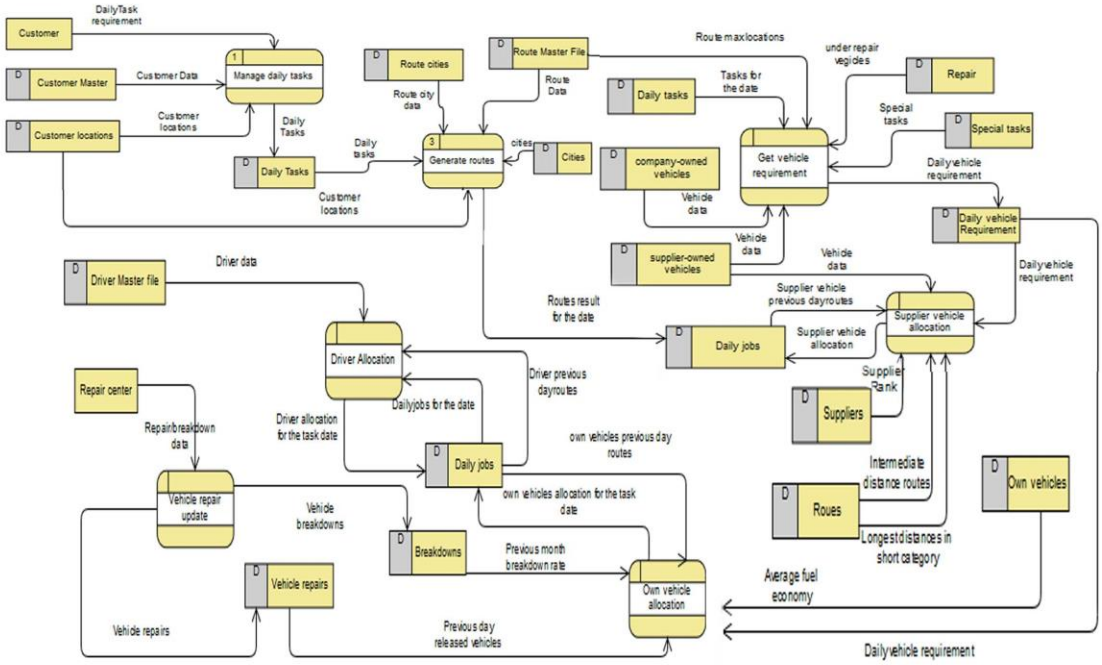


Figure 15 data flow diagram

The system entity-relationship diagram:

The ER diagram represents system entities and relationships between these entities. ERD is a high-level conceptual data model diagram used to analyze data requirements in database designing.

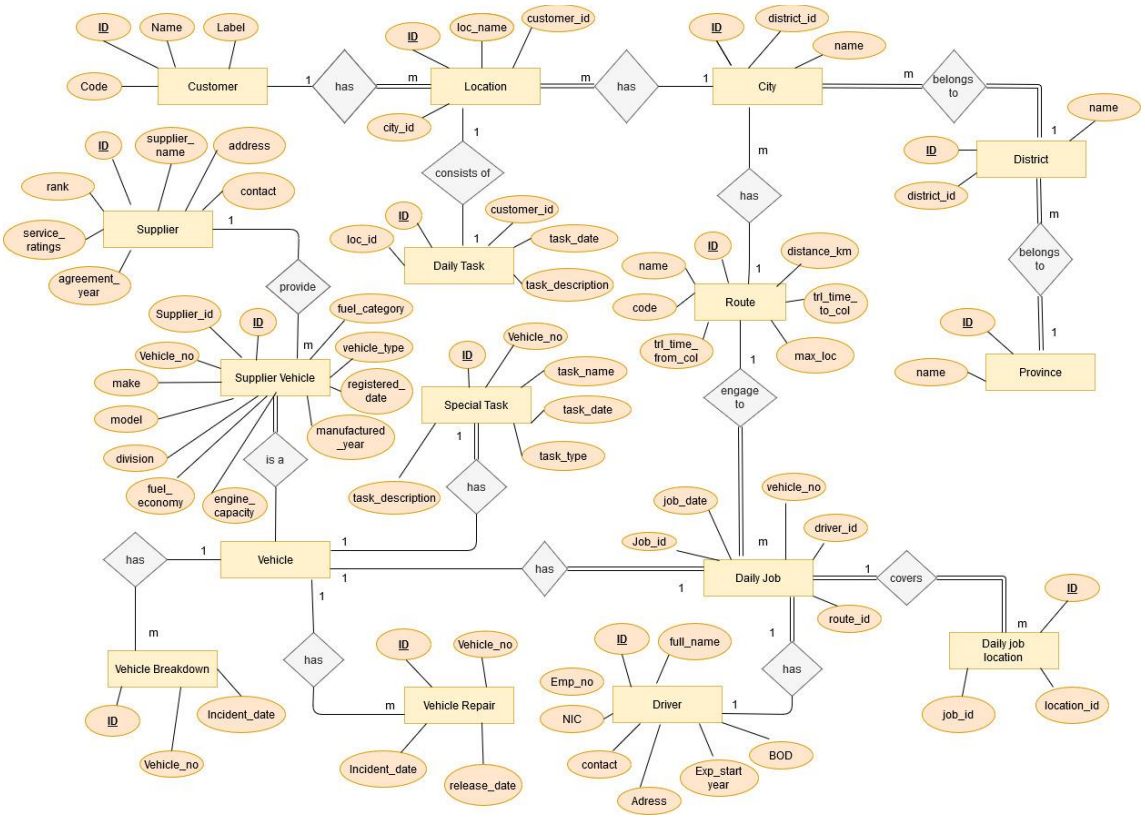


Figure 16 entity-relationship diagram

The system class diagram:

Represents main elements and interactions of the system by classes which help to determine the static relationships between them.

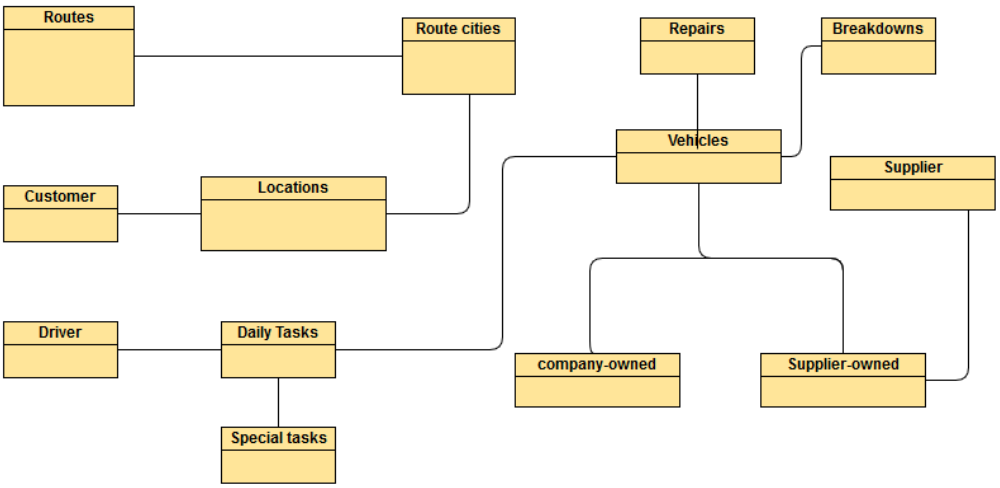


Figure 17 System Class Diagram

The system use case diagram:

Use case diagrams represent the users and their interactions with the system and describe the actions (use case) of the system.

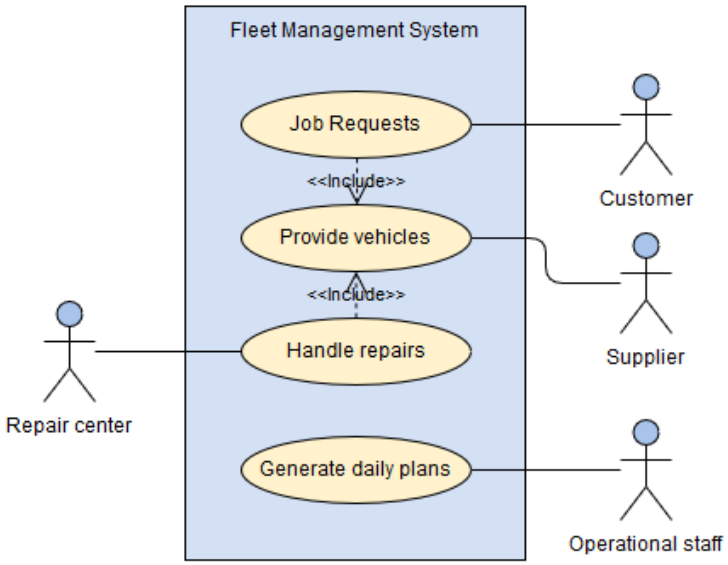


Figure 18 Use Case Diagram

An overview of the database structure is provided by the figure below.

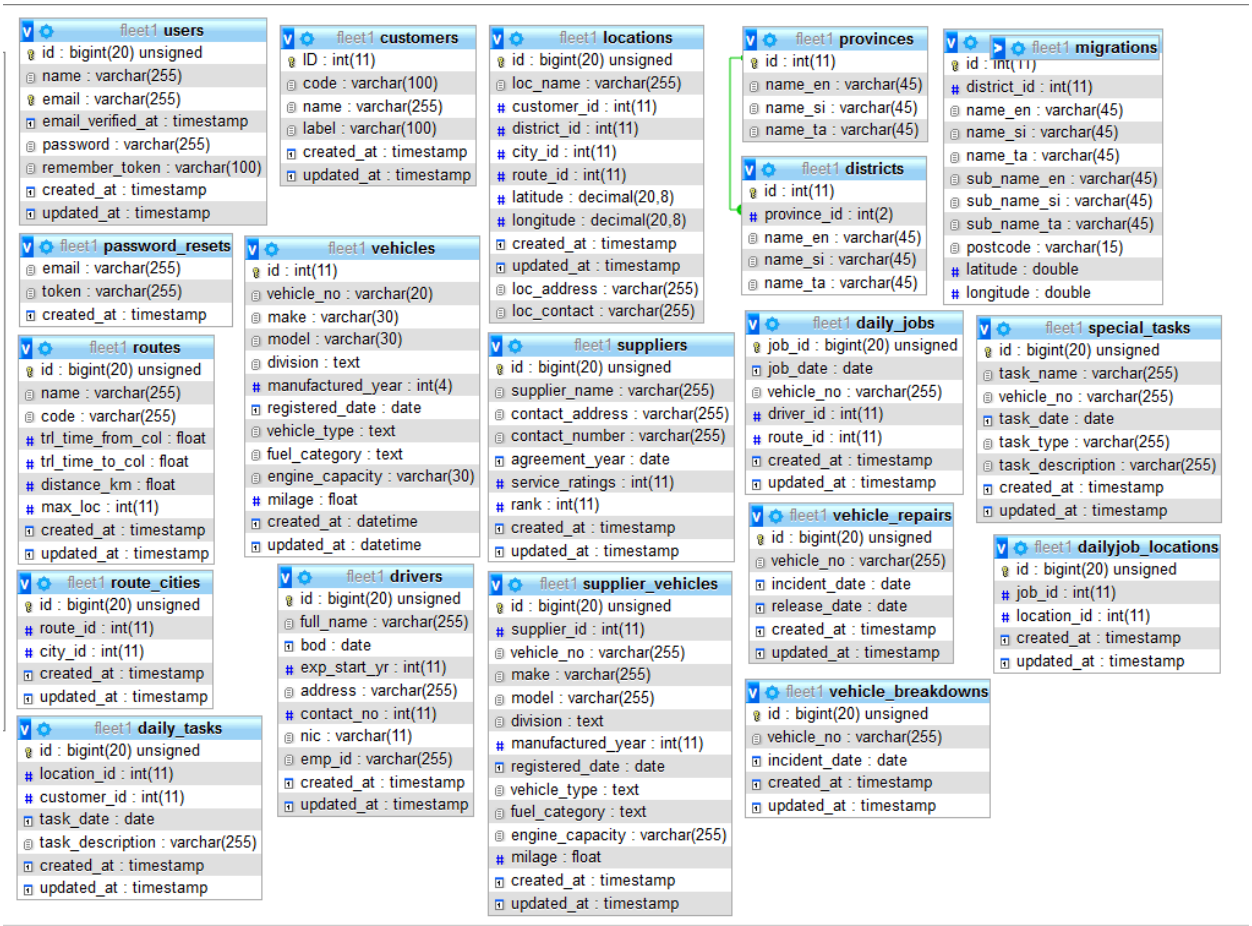


Figure 19 Fleet Management System DB Structure

5.4 Design Architecture

5.4.1 Process Models

Numerous processes can be used in software development. Waterfall Process, Iterative Process, Incremental Process, Agile Software Development, Extreme Programming, Rational Unified Process, and Rapid Application Development are the most commonly used process models.

Software development involves a set of methodical tasks and activities. The development includes steps such as analysis, design, implementation, testing, delivery, and maintenance. It is described as an SDLC (Software development life cycle or System development life cycle).

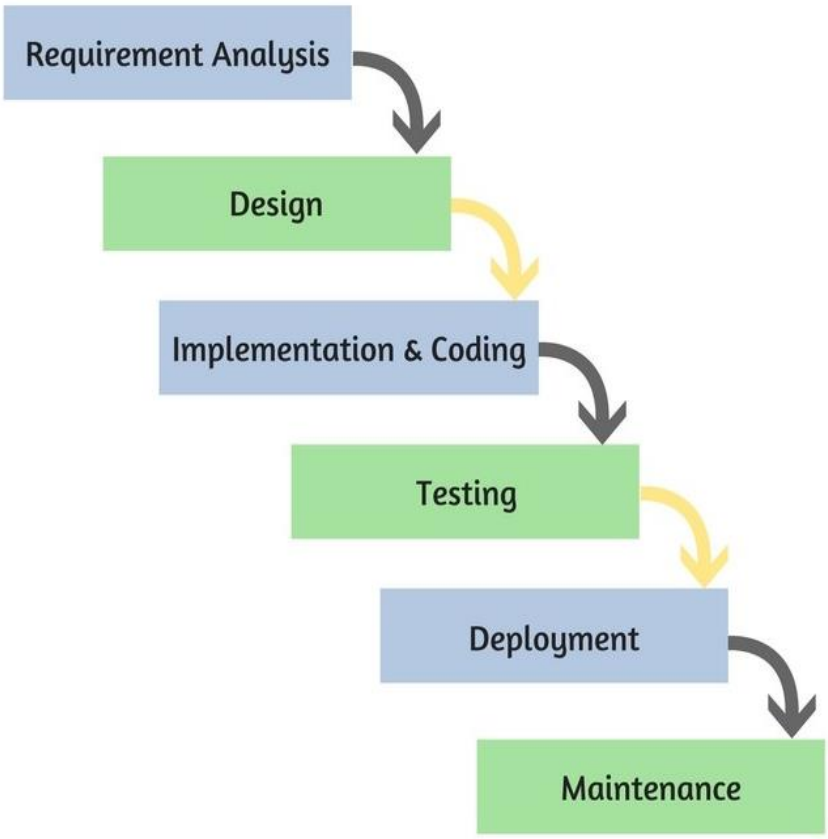


Figure 20 Waterfall model

Object-Oriented Analysis and Design

Analyzing system logical components and their interactions and designing the system based on the analyzed object models is described as Object-Oriented Analysis and Design (OOAD). This focuses on objects and their properties, methods, and data. OOAD is used in incremental and iterative development. It has two phases of activities as object-oriented analysis (OOA) and object-oriented design (OOD). The OOA has five stages defined as,

Identify the objects.
Organize the objects.
Define object interactions between each other.
Analyze the external behavior of the objects.
Analyze the internal behavior of the objects.

OOD is creating a design for the system that was planned in the analyzing stage and identifying suitable methods to implement it. OOA is theoretically planning ideas and generating concepts to identify what should be created, and OOD is how it is created. [27]

5.4.2 Process Models System Architecture

The MVC architecture was used for this system. MVC architecture representing 'model, view, and controller' is shown in figure 21. The model creates interaction with the database, and the view is responsible for creating interfaces. The controller is responsible for updating what's presented in the view and updating data through the model according to the received web requests.

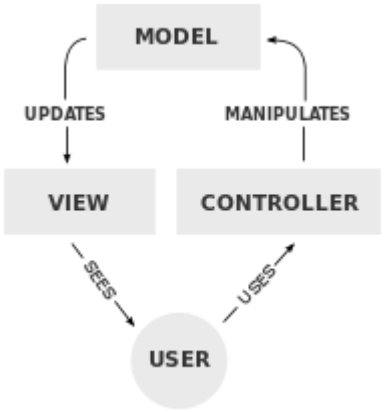


Figure 21 MVC Architecture

The Laravel framework is largely based on the MVC architecture and it facilitates separating the business logic (Model) from presentation (View) under the control through the (Controller). MVC framework of Laravel is a collection of libraries. All of these libraries perform some important activities and work together for the accomplishment of a common goal.

In the Laravel framework, the Model is a class with properties that match the columns in the database table and provide an interface with the data storage.

Controllers manage the business logic of the applications and are responsible for completing user actions.

The view is the presentation layer of the MVC architecture and Laravel uses Blade as its templating engine that is used in creating the user interface.

5.5 System Modules and Functionalities

Usually, operational staff of the CIT division who are responsible to plan the fleet operations for a given task date is given access to the system. These back-end users are able to access the system to extract information and update system data.

In the customer management module, customer data can be added to the system. Updating and deleting customer data is also managed in this module. Customer locations are managed by the locations module. There are two separate modules to manage company-owned vehicles and supplier-owned vehicles. Driver data are managed by the driver management module. Routes can be created by the route module and cities for the route are created in route city modules. Daily tasks are updated in the daily tasks module and vehicle assignments for special tasks are done by the special task management module. Vehicle repair and breakdown data are managed by separate modules.

Route plan for the task date with allocated resources can be generated by the Daily job plan module. In the daily plan module, routes, and route locations according to the job requests received from customers are generated for the given task date. The module also manages vehicle allocation and driver allocation for the generated routes. Company-owned and supplier-owned vehicles are allocated for routes based on profit. The condition of the vehicle and the performance of the vehicle is also considered in the allocation process. The module uses vehicle repair details, breakdown details, and fuel economy data for this. The module is also responsible for resource rotation management.

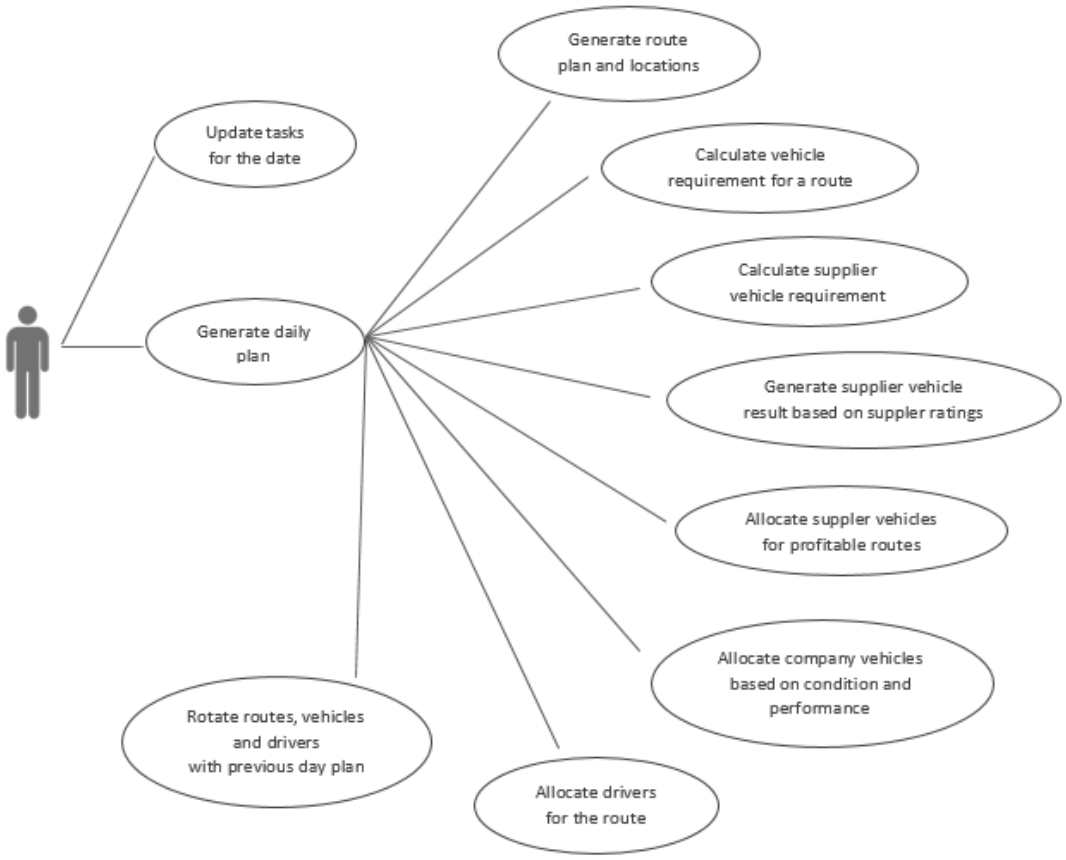


Figure 22 an overview of task management and daily plan module

5.6 Main Web Application

The main web application contains a few modules such as route management, task management, and daily job plan module, customer module, vehicle module, etc.

The data layer performs Create, Read, Update, and Delete (CRUD) operations on the database. Web services fetch data from the database and the presentation layer represents

fetched data in the view. Report generation areas and other interfaces are included in the presentation layer.

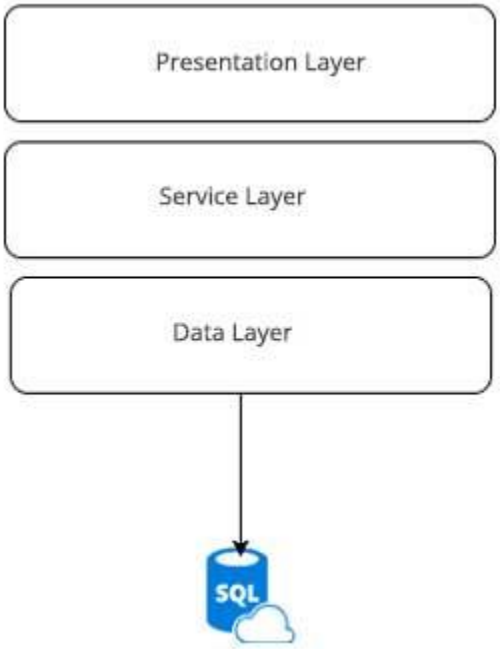


Figure 23 Main Web application

5.7 Overall Solution

For the development, as the main programming technology PHP 7 server-side scripting language was used with the Laraval 6 framework. The Maria DBMS was used as the database server and Apache 2 was used as a webserver to run this solution. Based on the requirement and the factors identified in the background research, the system was developed as a web-based application.

To avoid reinventing the wheel, a suitable, reliable, and secure Laravel based admin panel that provides the theming layers and basic extensible modules such as user management is utilized for the application. Laravel eloquent and SQL client was used to work with Maria DB. Since the web application is PHP-based, PhpStrom 10 was used as the integrated development environment (IDE).

5.8 Implementation of Database

Maria DB server was used to implement the database, and to communicate with database modal, Laravel eloquent ORM was used.

5.9 Summary

This chapter described the design of the system along with its overall architecture with general analysis. It also explains the design models used in application development, the system entities, components, users, classes, and their interconnections. Further, it explains how the system is implemented based on the design. This system consists of different modules to manage related data and to obtain the required output. This system was developed as a web-based system using PHP, MARIA DB, and APACHE2. After implementation, the system should be evaluated by conducting a proper testing mechanism. The next chapter discusses how the system and its functionality were evaluated in the testing phase.

Chapter 6

Assessment of the Fleet Management System

6.1 Introduction

The design and implementation of the fleet management solution are mainly discussed in chapter 5. This chapter mainly focuses on the testing phase of the system development process. Therefore, this chapter contains details related to the assessment of the developed system. It mainly evaluates whether the objectives of developing the system were achieved. To check whether the system generated expected outputs, the developer has created and executed test cases. By using different testing methods the accuracy in each technique is measured**.**

6.2 Evaluation of the Fleet Management System

Software quality assurance (SQA) is the process that ensures that the developed software meets and complies with the defined or standardized quality specifications. Software quality assurance is also considered an ongoing process within the software development life cycle (SDLC). In this context, therefore, during various stages of the development cycle, the software which was developed for AB Securitas was tested.

When considering the project objectives, the following mainly focused on the system.

- Identify the present CIT division process.
- Identify the problems of the existing process
- Identify the accurate profitable way of assigning vehicles and drivers to reduce costs and ensure secure transportation.
- Reduce the complexity of the current process and data gathering by introducing a centralized design.

The first two objectives mainly link to the pre-steps of the system development. They have been accomplished at requirement gathering. The last two objectives link with the implementation process. Evaluating that process was achieved by executing test cases (Refer to Appendix C).

Fleet management-related operations include many sub-processes. To run these sub-processes smoothly the company requires a well-functioning system. The current process has some drawbacks such as,

- Complexity.
- Unjustified, unprofitable decisions made.
- Less efficiency for information processing.
- Need more individuals to handle the process.
- Less security.
- No centralized control.

To address these issues, this research develops automated fleet management. In order to evaluate the accuracy of the developed system, the test cases (Refer Appendix C) are executed in the testing process. These test cases are related to the main functions of the developed system. The accuracy of the developed system was evaluated by running test cases.

6.3 Summary

This chapter mainly focuses on evaluating the developed fleet management system. Using different test cases, the functionality of the system has been evaluated. Test cases were developed covering different aspects of the system and to evaluate the accuracy of every module of the system. The next chapter discusses the conclusion of this project, limitations, and future work under this research area.

Chapter 7

Conclusion and Future Work

7.1 Introduction

Fleet management systems are a used to manage the fleet and its other resources effectively and improve fleet operations' profitability by reducing costs and risks associated with fleet investments. This chapter discusses the overall achievement summary of the implemented fleet management system, limitations, and describes available findings for further development. This chapter also includes a critical discussion and assessment of the project.

AB Securitas is a prominent company that focuses on the different aspects and demands of the security industry in Sri Lanka. Cash-in-transit is their main business area and they had identified different issues related to their current business process in cash-in-transit task management and were willing to enhance their business by working with new ideas and technologies. That was the commencement of the development of this automated decision support fleet management system. Since vehicle performance directly affects the key process of the system, a detailed analysis regarding the performance factors of the vehicle wa s conducted to gain high effectiveness and accuracy in the vehicle allocation process. Therefore a statistical analysis wa s done to identify the factors that affect the most to a vehicle breakdown. From the performed analysis, the relationships between vehicle features and the breakdowns were identified.

The produced system supports operational staff at the cash-in-transit division to manage their daily job tasks and generate daily plan results. The automated system handles all the sub-functions associated with the fleet management activities in the CIT division.

The solution mainly helps the CIT division to manage their daily activities and provides all the features to assist in saving time by reducing complexity.

This solution provides the facilities to create daily tasks based on their customer job requests. Once tasks are entered all other processes such as vehicle allocation, driver allocation, route rotation, filtering, and sorting of vehicles are handled by the system.

7.2 Achievement

The automated fleet management system is a suitable solution to replace the current manual process in the organization. This results in decreasing complexity and effort, improving profit and efficiency, better allocations in resources with the most appropriate routes, reducing the workload of operational staff, and increasing motivation in the staff of the CIT division.

The output of this system has provided an automated solution for different sub-processes related to fleet management such as route generation, vehicle and driver data sorting, and filtering, supplier vehicle handling, allocation of daily vehicles for profit, allocation of daily vehicles by considering vehicle condition, ensuring secure transport at route generation stage, driver allocation and rotation. This system supports the company in planning daily operations in several aspects by presenting a centralized control.

The main goals of the project were managing the daily process of the fleet in the CIT division and supporting the decision-making process to enhance the company revenue.

The fleet management system could be used to achieve the following objectives.

- Generate daily routes based on job requests.
- Allocate resources for all operations efficiently and profitably.
- Reduce the complexity of the process and reduce the workload of the operational staff.
- Successful completion of the database and design stage of the system.
- The user-friendly interface of the system was developed.
- Generate a daily plan report for a given task date that satisfies all requirements.

7.3 Limitations and Future Works

The following difficulties were encountered during the time period up to the dissertation submission date and most of the difficulties were encountered at the start of the development of the system.

The requirement gathering stage was difficult due to the conflicting requirements by the different levels of user and management level concepts. Due to security reasons, actual data gathering was not authorized by the company. The company was not willing to disclose its customer and job location details, especially their cash transport routes and actual vehicle, driver information. Therefore, test data was developed to evaluate the system functionality based on their actual data patterns.

Some difficulties encountered during the development stage of the system are listed below.

- Receiving information from different systems of the company.
- The requirement targets the CIT division and its fleet management's process, but different data should be obtained to get the final result.
- Real-time Data Synchronization process.

Most errors were successfully resolved. Also, this system helps to learn how to fix errors and find solutions working about them.

We are suggesting to improve the own vehicle allocation module by using the outcome of the breakdown factor analysis when defining allocation rules. Currently, the allocation rules are defined using the breakdown per mileage (BPM) data and from the statistical analysis, we identified what are the significant variables that relate to the BPM. Therefore that knowledge can be used for further analysis and can be used to improve the developed system modules of vehicle allocation.

Also this system can upgrade to interact with the other systems in the company to improve functionality by getting a more centralized view and to synchronize data

over other systems and the system can be improved to visualize generated routes by using a map application programming interface and further to generate optimized routes. Also, this system can further expand to track and monitor vehicle and driver performance data in a wide area to improve fleet efficiency.

7.4 Summary

This chapter considers the overall achievement of the project. It also describes limitations & future work related to this project. By integrating a fleet management system to manage fleet operations, the companies are able to gain many advantages in different aspects of their business process. Fleet management influences both vehicle and driver management and support in measuring performance. A well-structured and well-designed system will increase fleet productivity and efficiency and raise company profits in different areas. A good software solution can reduce fleet risk, reduce wastage, and utilize resources properly. The software designed and implemented for AB Securitas (Pvt) Ltd is focused to achieve the requirements and objectives identified in the feasibility study. The system was developed as a solution to overcome the issues identified in the CIT division of the company. We suggest enhancing the system with the knowledge obtained from the statistical analysis. Also to interact with other divisions' systems to get more centralized control and expanding the system to increase the fleet efficiency by monitoring vehicle and driver performance using different variables.

References

[1] Hamed Saghaei (2016) *Design and Implementation of a Fleet Management System Using Novel GPS/GLONASS Tracker and Web-Based Software* Shahrekord

[2] “How Does Fleet Management Work.”[Online].
Available: https://www.expertmarket.com/fleet-management/how-does-fleet-management-work
[Accessed: 12-Mar-2020]

[3] “Fleet Management and Logistics” [Online]. Available: https://books.google.lk/books?hl=en&lr=&id=KXjdBwAAQBAJ&oi=fnd&pg=PP9&dq=fleet+management+software&ots=w3dcuj6LUG&sig=9w0K7tqUqWnKkE-c_lKqudEXCEA&redir_esc=y#v=onepage&q=fleet%20management%20software&f=false [Accessed: 14-Feb-2020]

[4] “What is the Purpose of Fleet Management” [Online] Available: https://financesonline.com/what-is-the-purpose-of-fleet-management/ [Accessed: 14-Mar-2020]

[5] “Fleet Management - Wikipedia.” [Online]. https://en.wikipedia.org/wiki/Fleet_management [Accessed: 14-Mar-2020]

[6] “Dynamic fleet management.” [Online] Available: https://books.google.lk/books?hl=en&lr=&id=8LhdeoxYzLAC&oi=fnd&pg=PA1&dq=fleet+management+research+papers&ots=VyA2IF0L3_&sig=3ePf5Wm0FKcx6woLAj86hlHr6VM&redir_esc=y#v=onepage&q=fleet%20management%20research%20papers&f=false
[Accessed: 16-Mar-2020]

[7] Ahmad Aljaafreh, Majdi Khalel, Islam Al-Fraheed, Kafa Almarahleh, Rwan Al-Shwaabkeh, Saja Al-Etawi and Waed Shaqareen (2011) *Vehicular Data Acquisition System for Fleet Management* Tafila

[8] Hans M. Molin, Mission Viejo, Ken D., Irvine, (2013) *Vehicular fleet management system and methods of monitoring and improving driver performance in a fleet of vehicles* California

[9] Kristijan Rogic, Brani Sla V Sutic, Go Ran Kolaric (2008) *Methodology of Introducing Fleet Management System* Zagreb

[10] Commercial Vehicle Fleet Management and Information Systems [Online] Available: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjY7_zAypXpAhUKbisKHfMXDOMQFjAAegQIAxAB&url=https%3A%2F%2Frosap.ntl.bts.gov%2Fview%2Fdot%2F3811%2Fdot_3811_DS1.pdf%3F&usg=AOvVaw1kVlCp9CSbFhG4AqoXyBNh [Accessed: 20-Mar-2020]

[11] Adriano Alessandrini, Francesco Filippi, Fernando Ortenzi. [2012] *Consumption Calculation of Vehicles Using OBD Data* Florida.

[12] Liangfei Xu, Minggao Ouyang, Li Jianqiu, Jianfeng Hua. [2011] *Influence of powertrain parameters on vehicle performance of a fuel cell/battery city bus*. Beijing.

[13] Iulian Carpatorea [2017] *Methods to Quantify and Qualify Truck Driver Performance.* Halmstad. www.hh.se/hup

[14] Samir Causevic, Adisa Haskovic, Belma Memić, Alem Čolaković. (2012) *IMS application as fleet tracking solution.* Sarajevo. Technics Technologies Education Management.

[15] “Hypertext Preprocessor - Wikipedia.” [Online]. https://en.wikipedia.org/wiki/PHP [Accessed: 21-Mar-2020]

[16] Gáspár, P., Szalay, Zs., Aradi, Sz. (2014) *Highly Automated Vehicle Systems*, BME MOGI Budapest.

[17] Sai Chanda, Zhuolin Lia, Vinayak V. Dixita,b, S. Travis Waller (2021) *Examining the macro-level factors affecting vehicle breakdown duration*, rCITI Australia.

[18] Sai Chanda, Zhuolin Lia, Vinayak V. Dixita,b, S. Travis Waller (2020) *Analysis of Vehicle Breakdown Frequency*, NSW Australia.

[19] "Apache HTTP Server - Wikipedia." [Online]. https://en.wikipedia.org/wiki/Apache_HTTP_Server [Accessed: 25-Mar-2020]

[20] "What is Apache. " [Online] https://www.hostinger.com/tutorials/what-is-apache. [Accessed: 25-Mar-2020]

[21] "MariaDB - Wikipedia. " [Online]. https://en.wikipedia.org/wiki/MariaDB [Accessed: 25-Mar-2020]

[22] "MariaDB Vs MySQL Features. " [Online] https://mariadb.com/kb/en/mariadb-vs-mysql-features/ [Accessed: 25-Mar-2020]

[23] "Grunt" [Online] https://en.wikipedia.org/wiki/Grunt_(software) [Accessed: 25-Mar-2020]

[24] "Importance and benefits of SPSS in research" [Online] https://www.writeneed.com/post/2017/07/05/importance-and-benefits-of-spss-in-research [Accessed: 25-Mar-2020]

[25] "Laravel - Wikipedia" [Online]. https://en.wikipedia.org/wiki/Laravel [Accessed: 25-Mar-2020]

[26] "Laravel" [Online] https://laravel.com

[27] "Object-Oriented Analysis and Design: What is it and how do you use it?" [Online] https://airbrake.io/blog/design-patterns/object-oriented-analysis-and-design

[28] “Interpret R Squared Regression” [Online]
https://statisticsbyjim.com/regression/interpret-r-squared-regression/

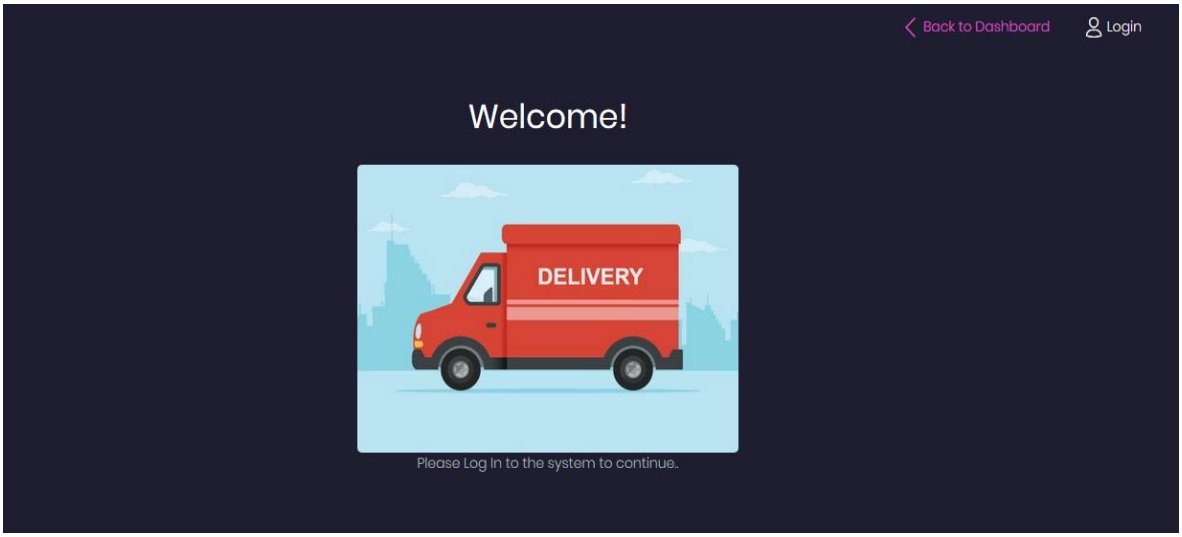
[29] “Further regression in SPSS” [Online]
https://www.sheffield.ac.uk/polopoly_fs/1.531431!/file/MASHRegression_Further_SPSS.pdf

[30] “Multiple Regression Analysis using SPSS Statistics” [Online]
https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php

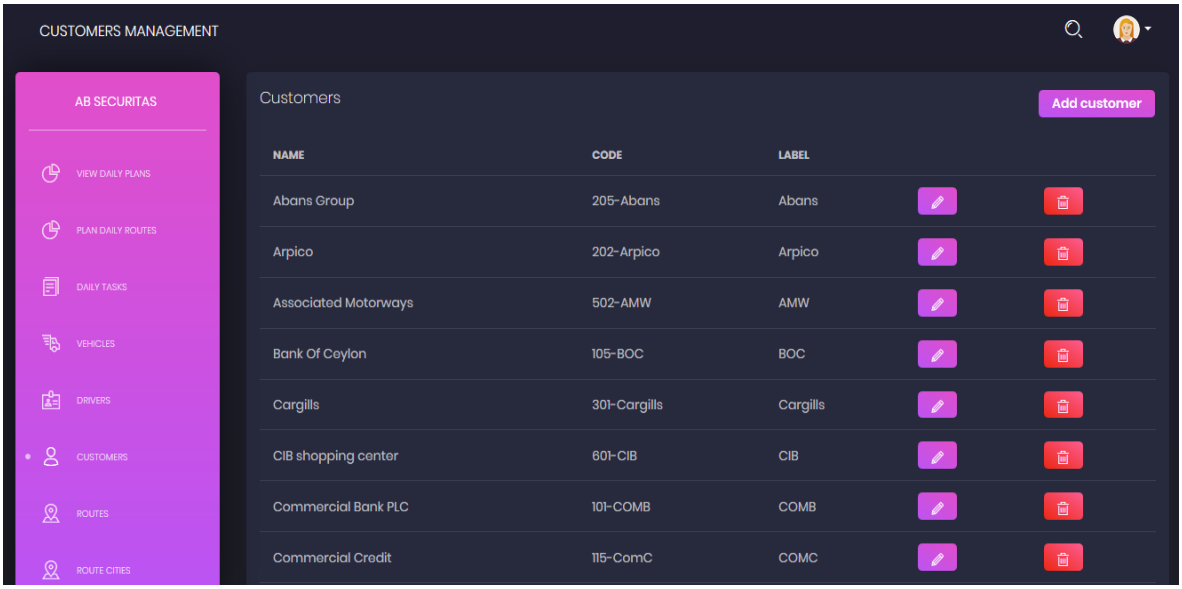
[31] How to test validate a regression model
https://www.researchgate.net/post/How_to_test_validate_a_regression_model

Appendixes

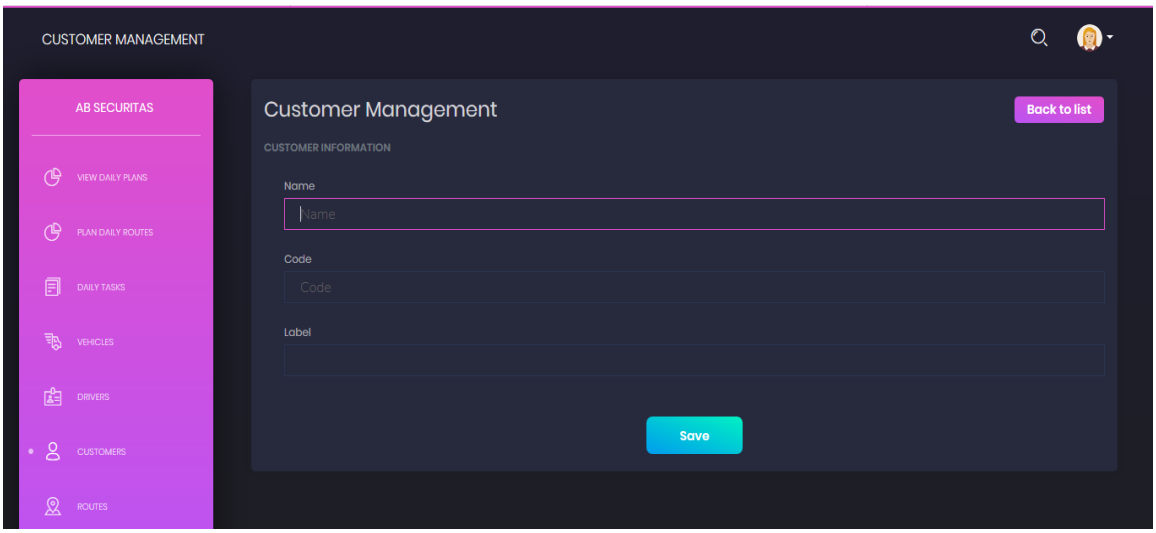
Appendix A – User Interfaces.



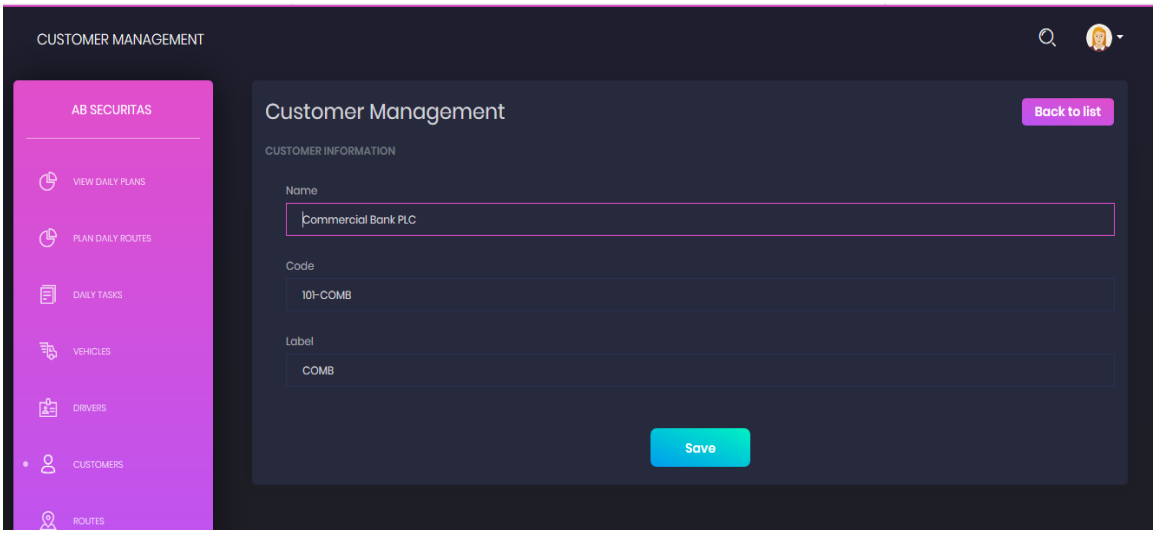
Welcome Screen



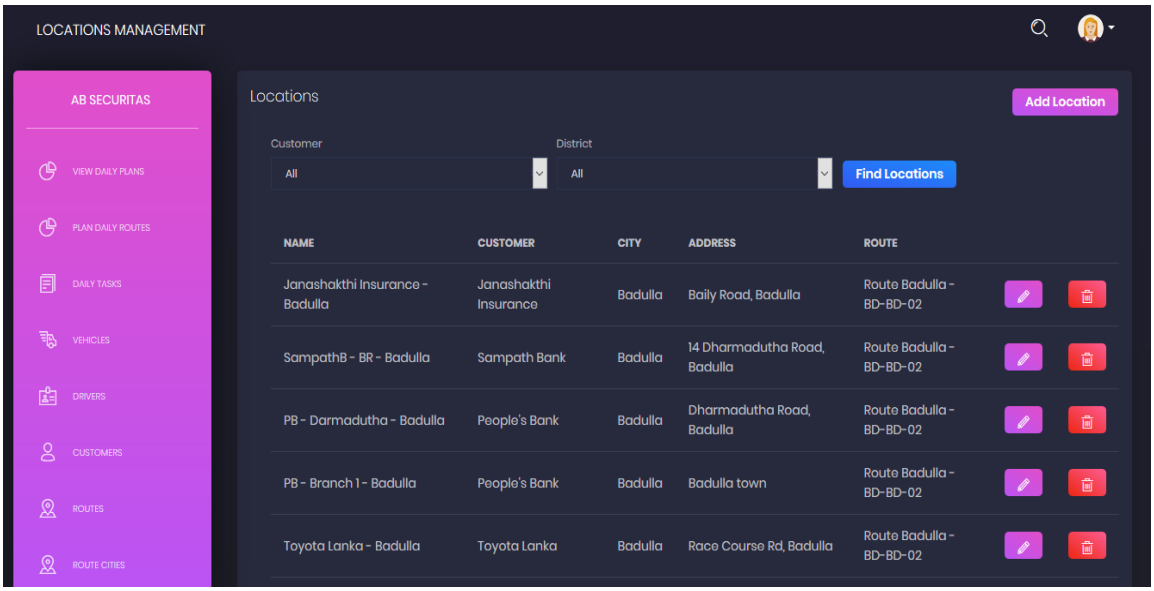
Customer Management



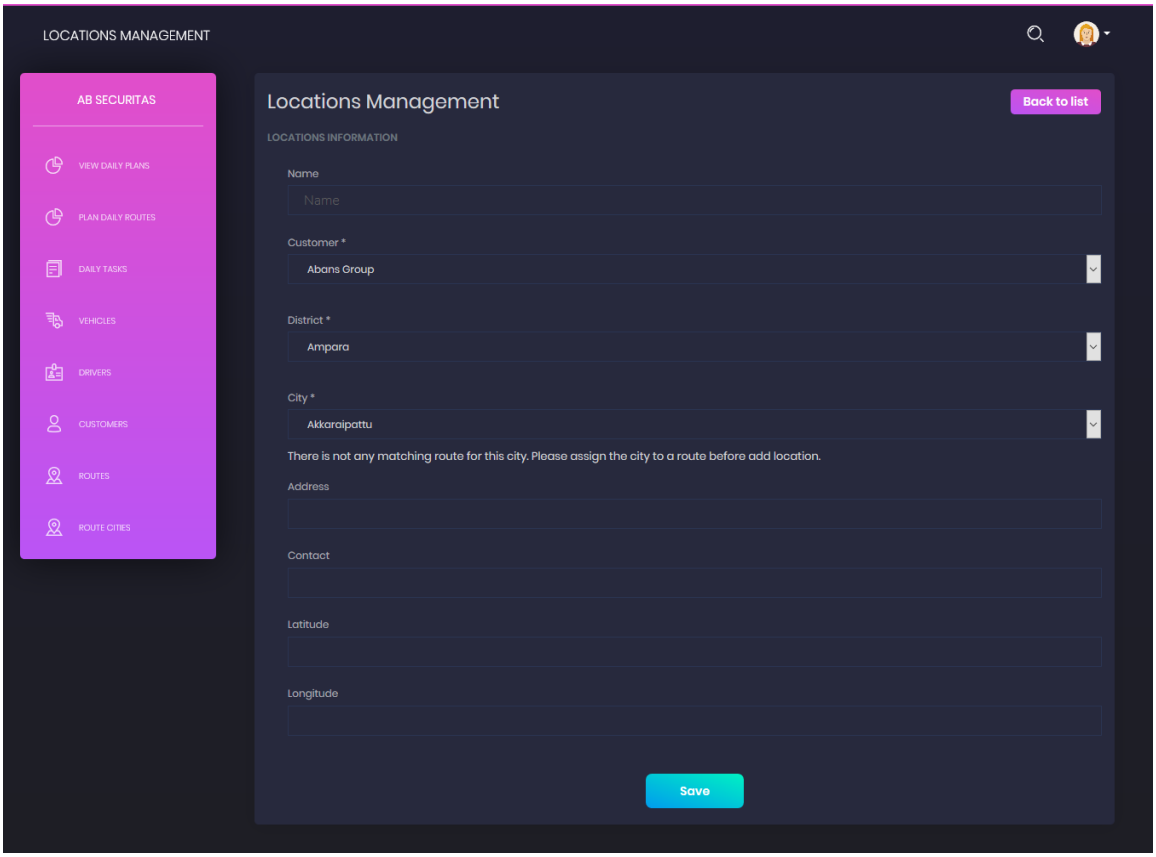
Create Customer



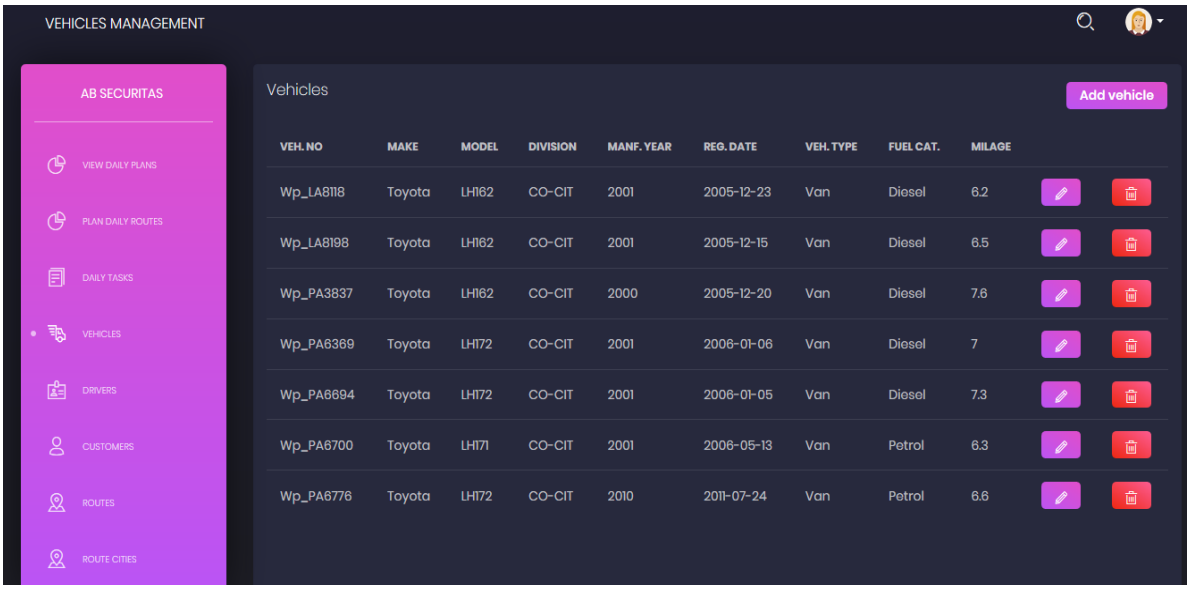
Update Customer



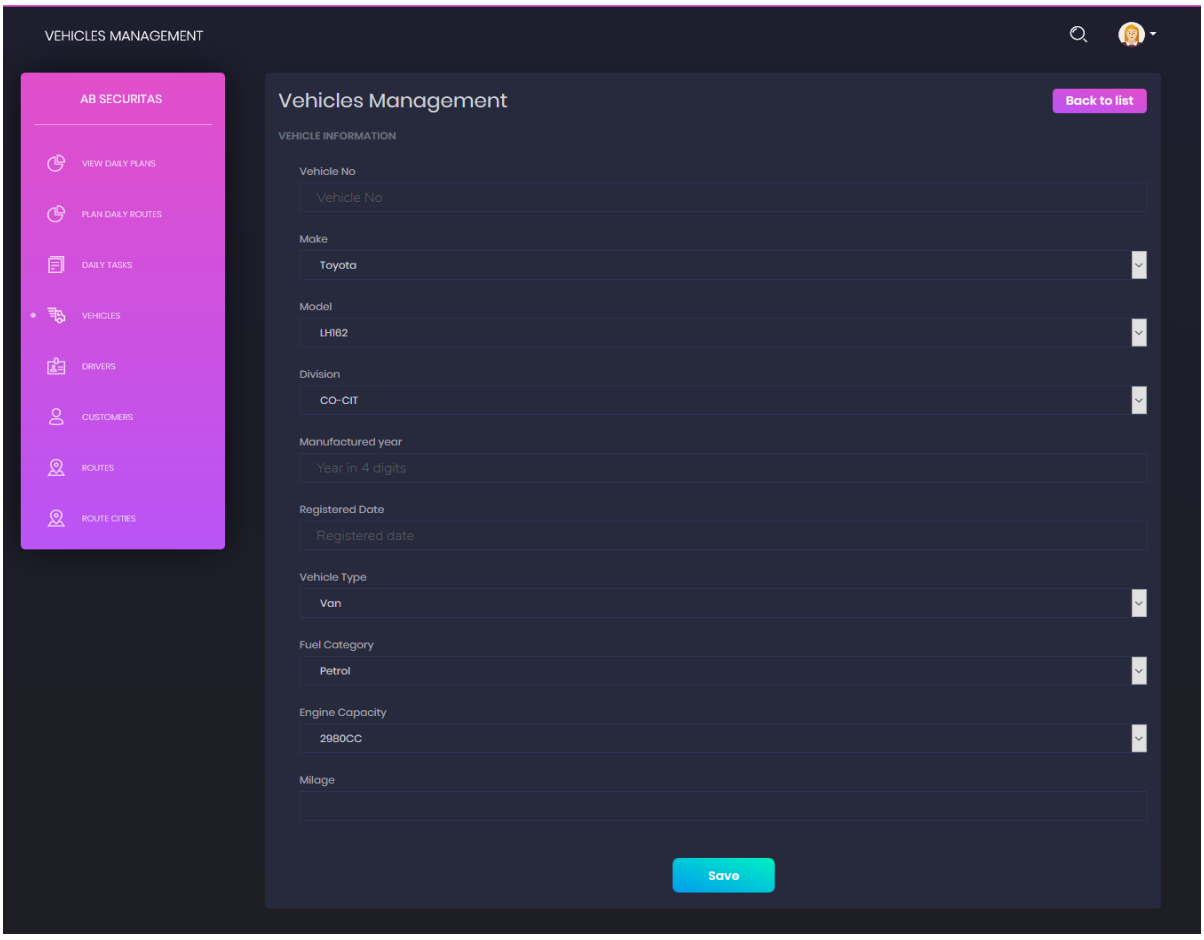
Customer Locations Management



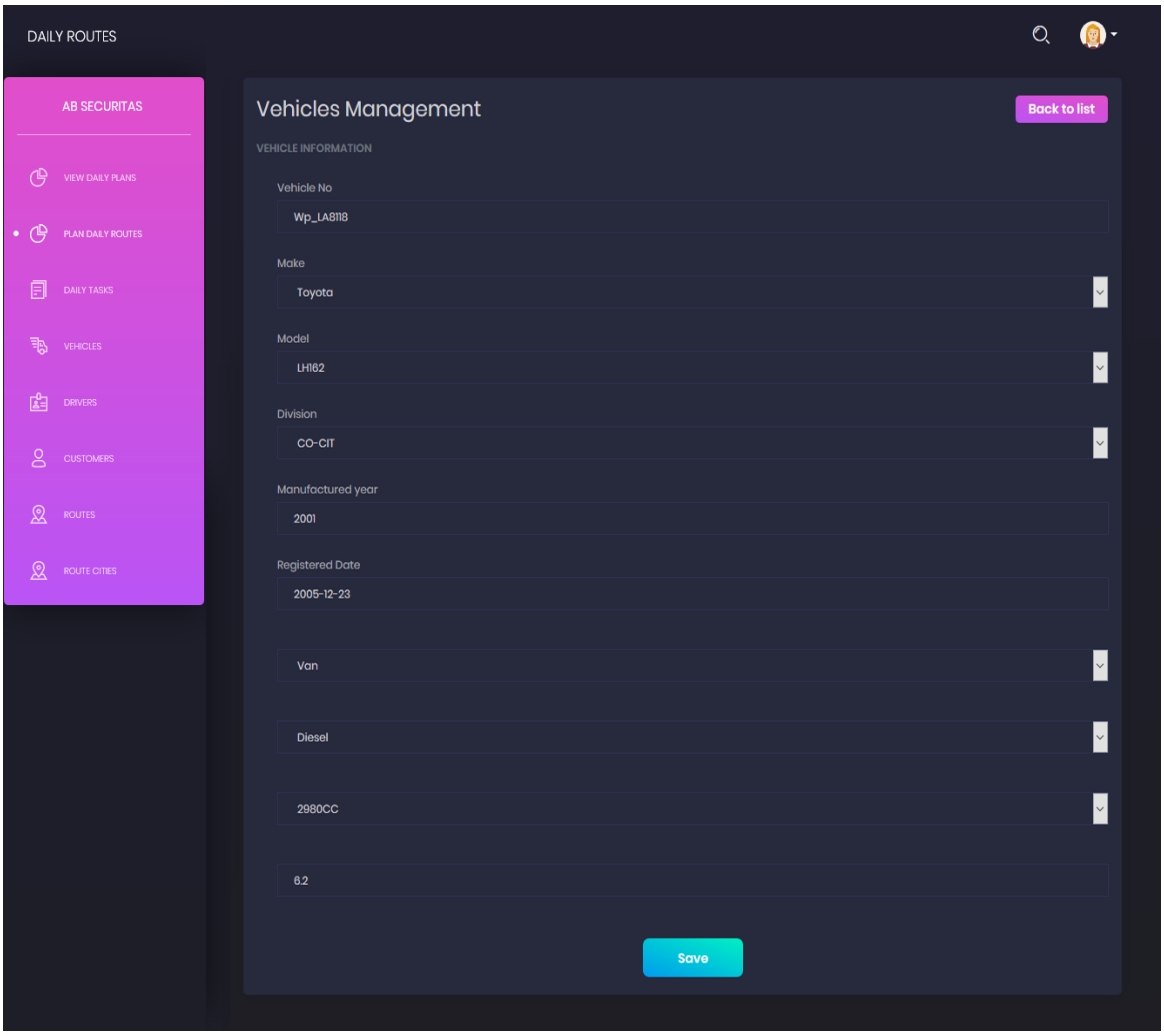
Create Customer Locations



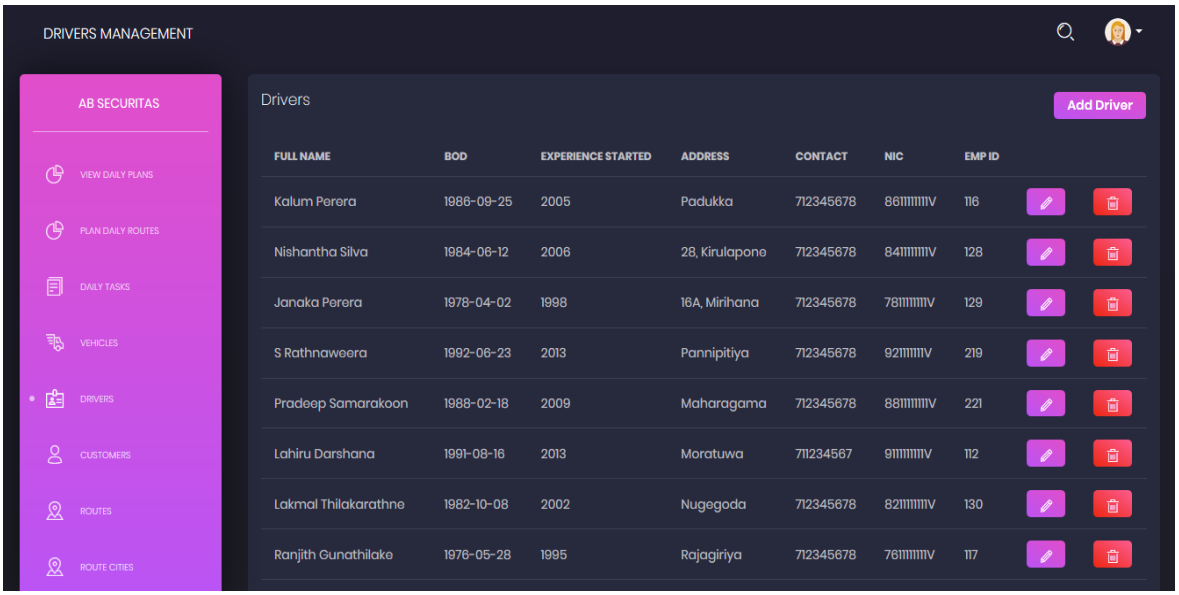
Vehicle Management



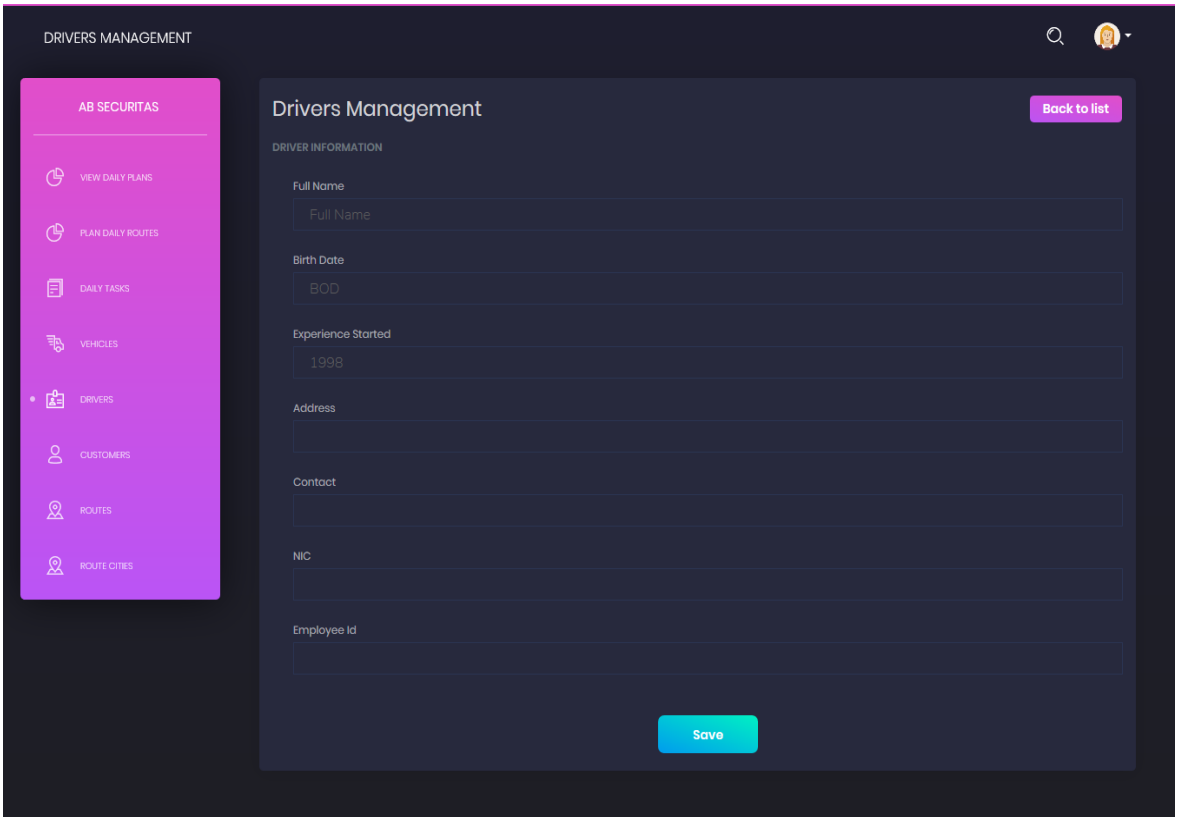
Create Vehicle



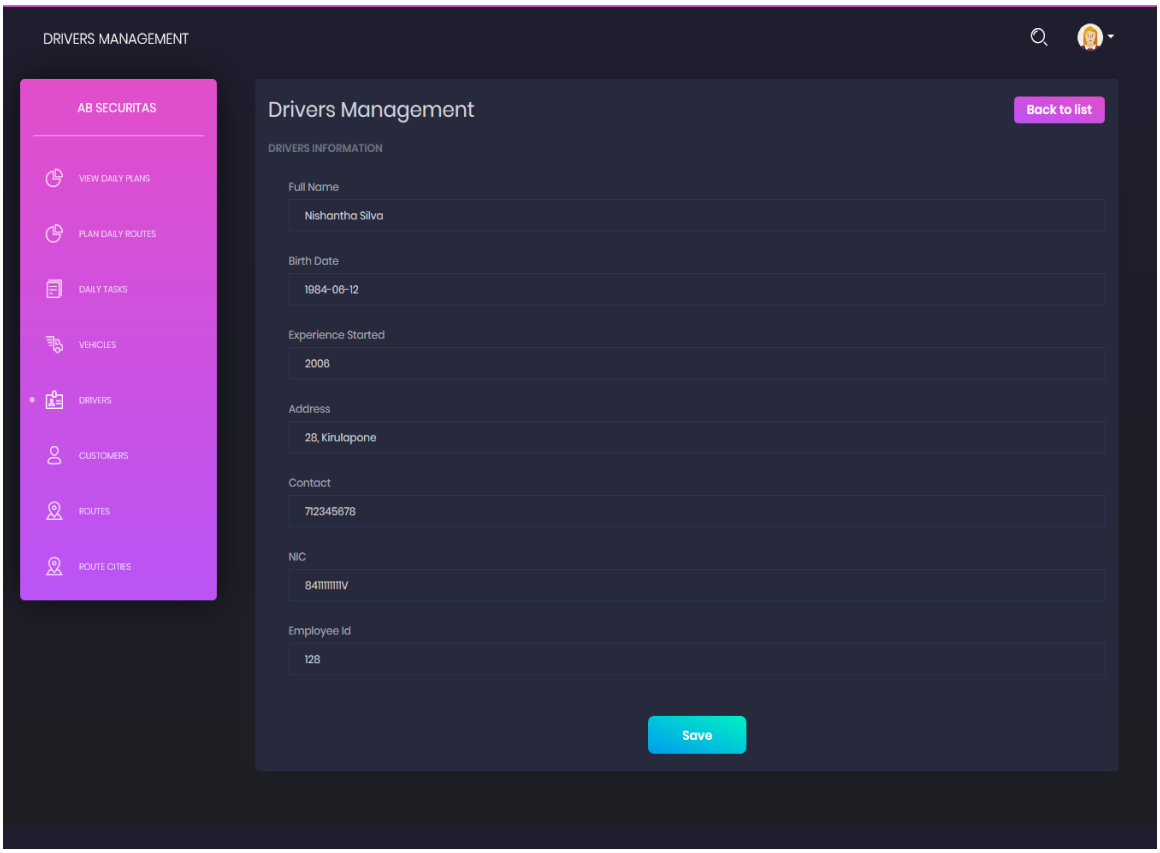
Edit Vehicle



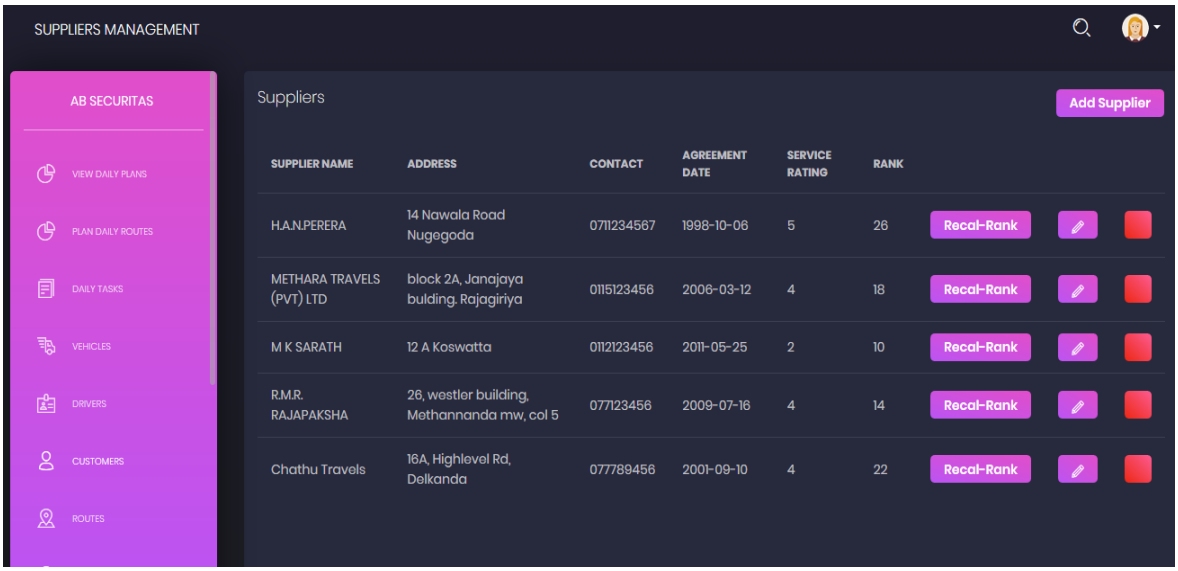
Driver Management



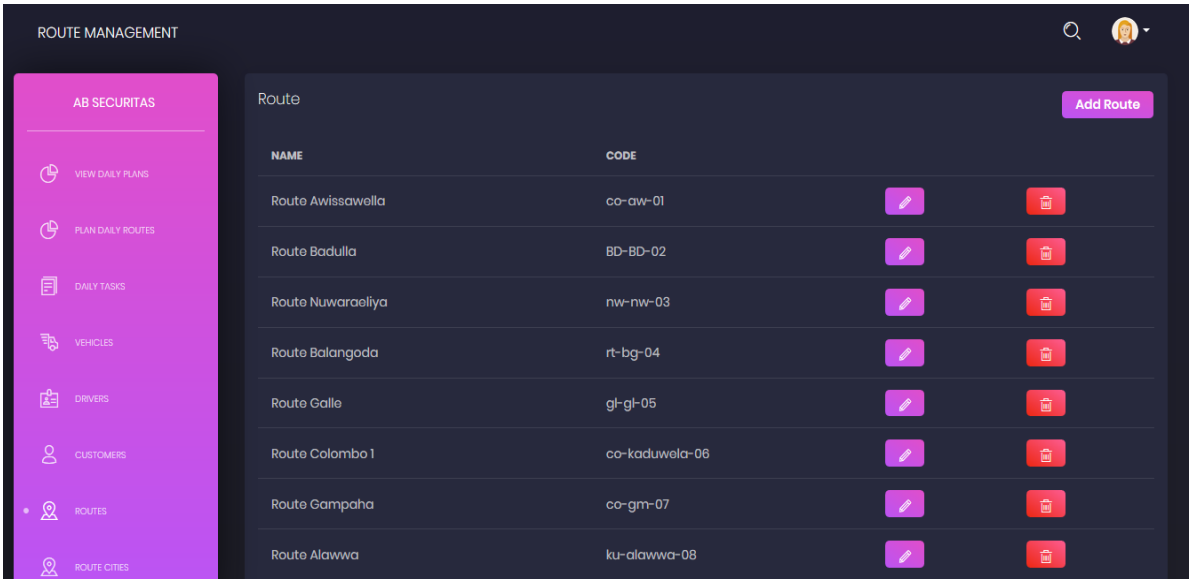
Create Driver


DRIVERS MANAGEMENT
AB SECURITAS
VIEW DAILY PLANS
PLAN DAILY ROUTES
DAILY TASKS
VEHICLES
DRIVERS
CUSTOMERS
ROUTES
ROUTE CITIES
Drivers Management
Back to list
DRIVERS INFORMATION
Full Name
Nishantha Silva
Birth Date
1984-06-12
Experience Started
2006
Address
28, Kirulapone
Contact
712345678
NIC
841111111V
Employee Id
128
Save

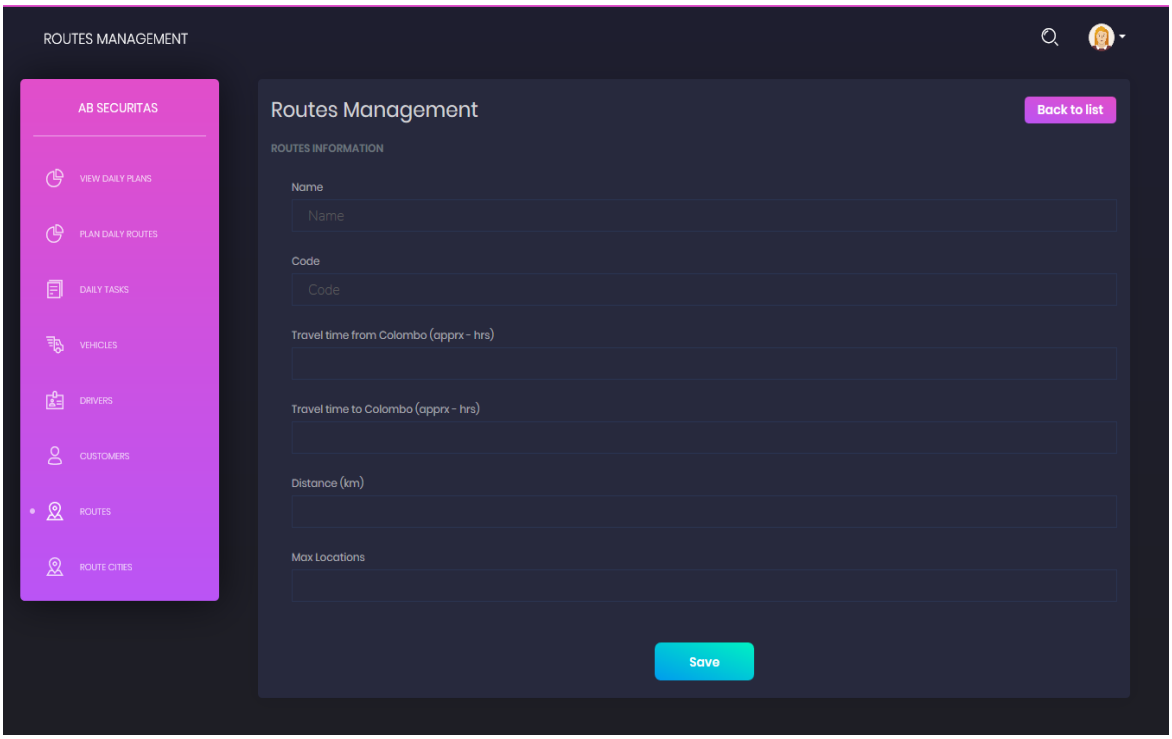
Edit Driver



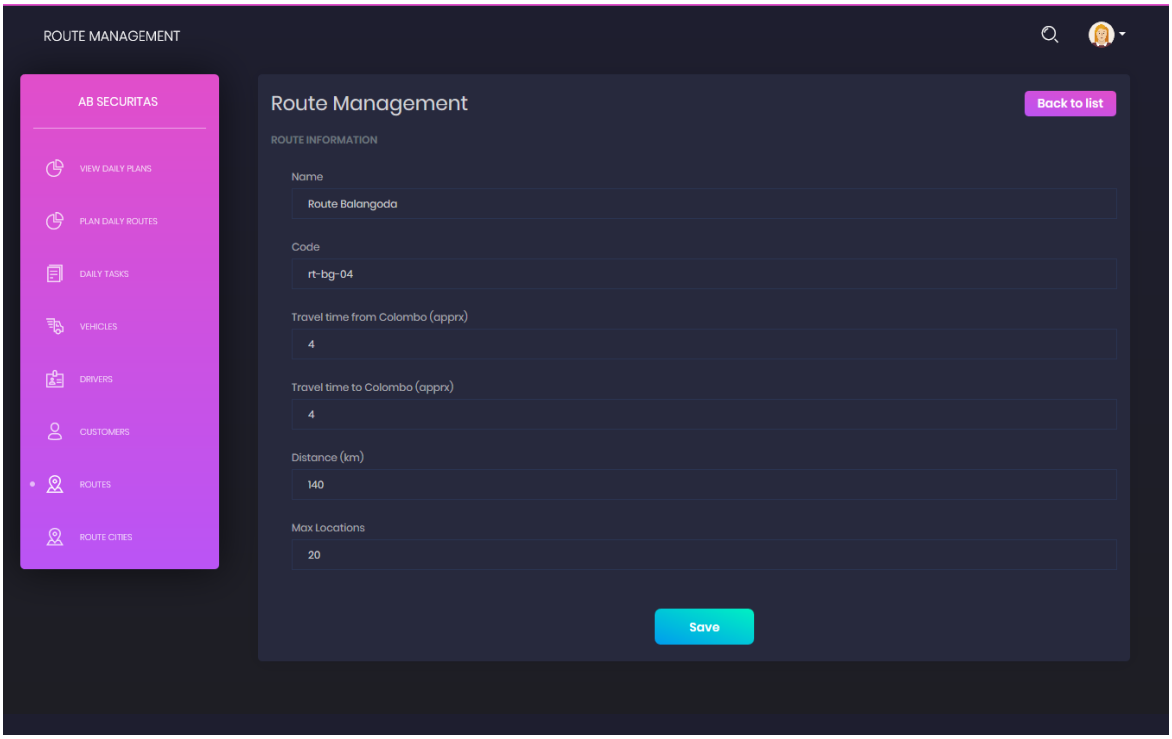
Supplier Management



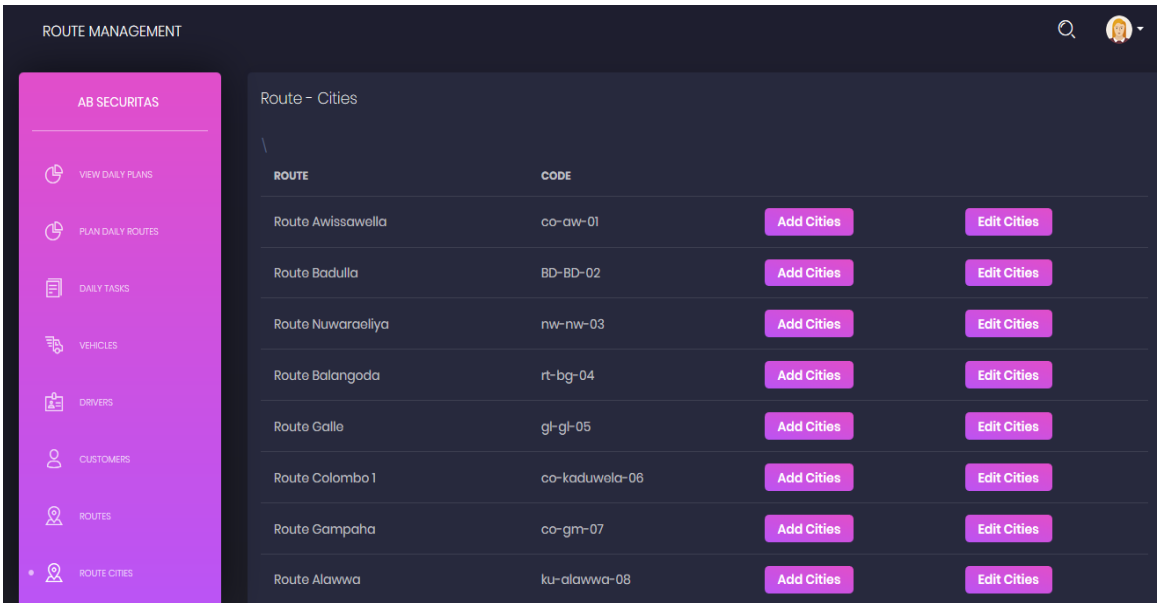
Route Management



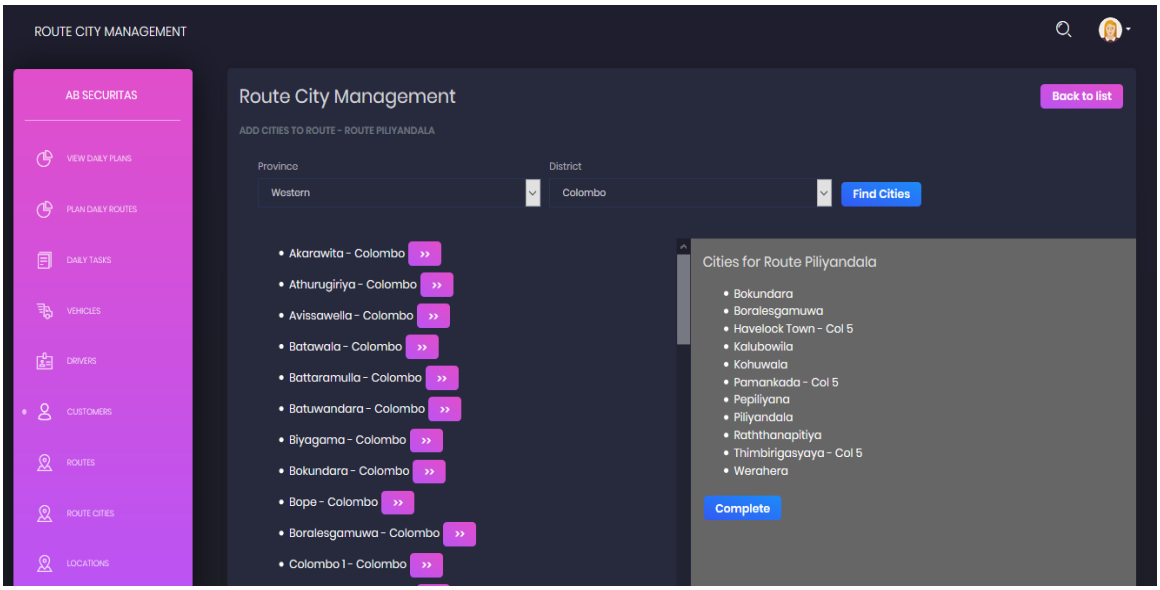
Create Route



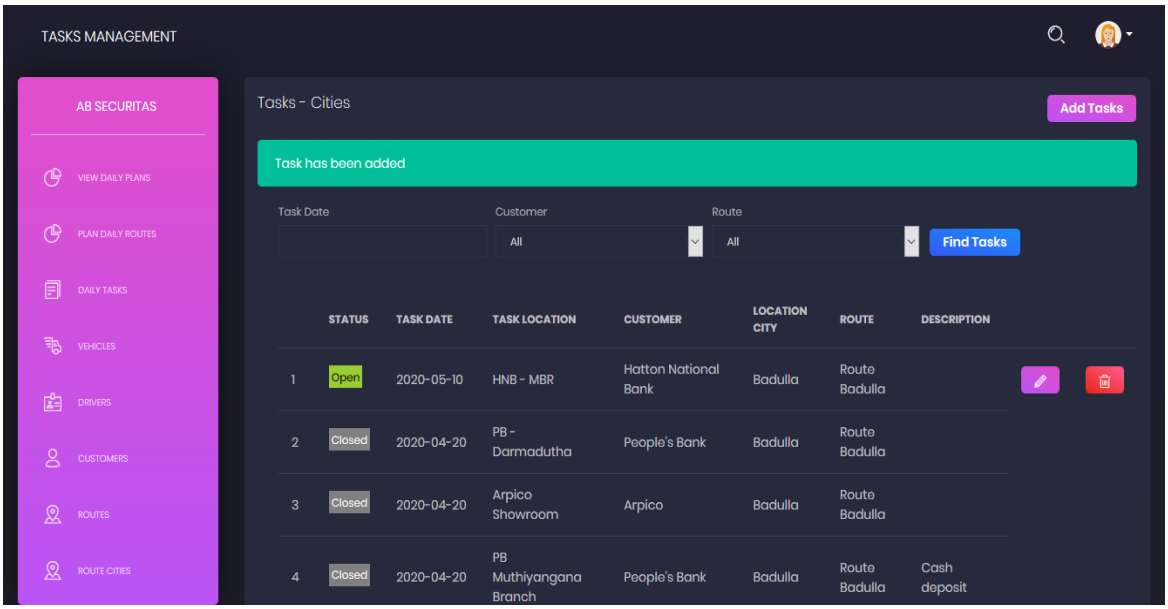
Edit Route



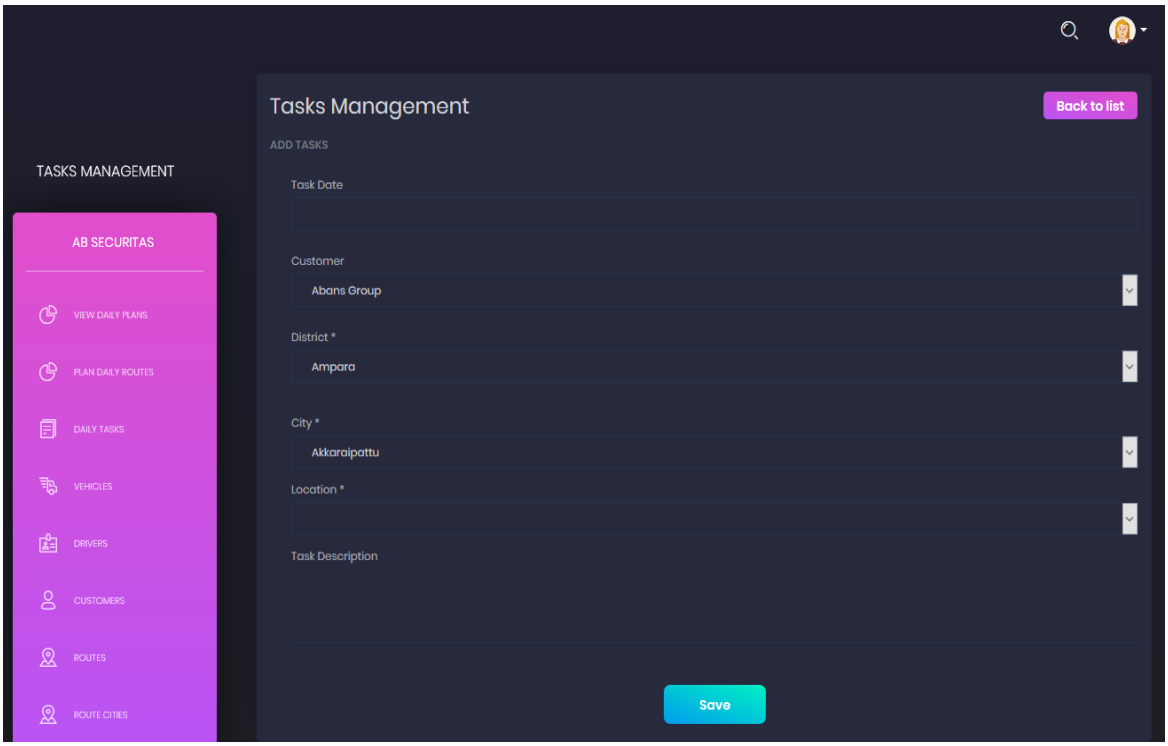
Route-City Management



Route-City Create Interface



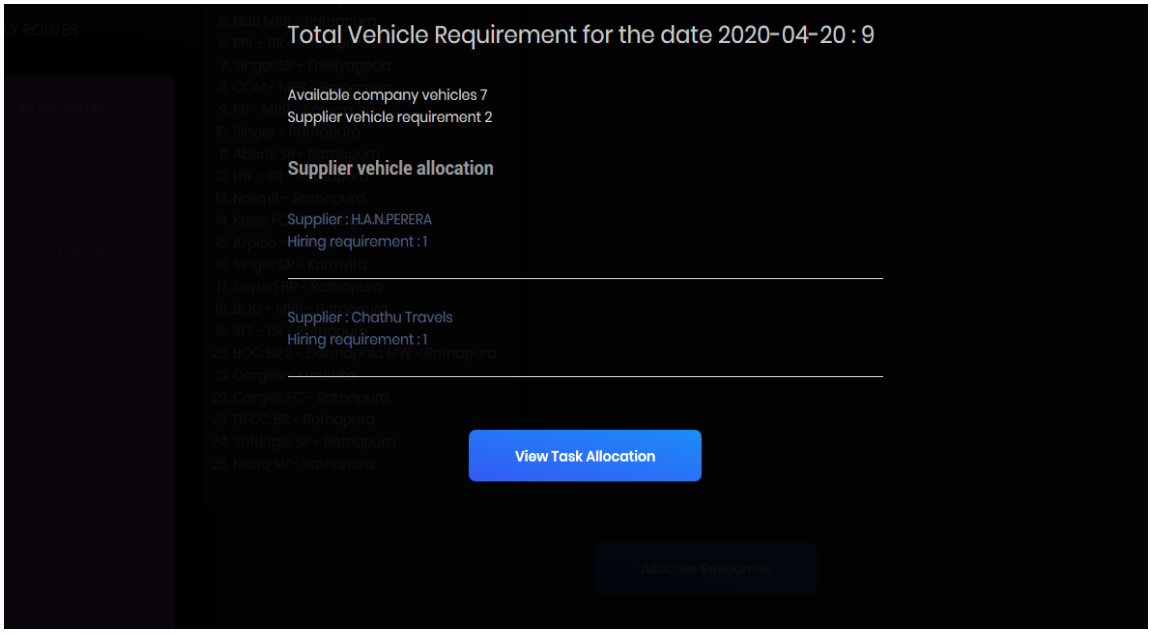
Task Management



Create Task

Route Generation – Processing Interface (Corrections are needed to continue)

Route Generation – Processing Interface (Corrections are done)



Vehicle Requirement Summary

Route Plan with Allocated Resources

Appendix B – Coding Parts

```php
use App\SupplierVehicle;
use Illuminate\Http\Request;
use DB;
use App\Customer;
use App\Location;
use App\DailyTasks;
use App\Suppliers;
use App\DailyJobs;
use App\DailyjobLocation;
use App\Route;
use App\Driver;
use DateTime;

class TaskRoutesProceedController extends Controller
{
    public function index(Request $request)
    {
        $task_routes = DB::table('daily_tasks')
            ->select("routes.name", "routes.max_loc"
                ,DB::raw("(GROUP_CONCAT(locations.loc_name SEPARATOR ',')) as `loc_list`")
                ,DB::raw("(GROUP_CONCAT(cities.name_en SEPARATOR ',')) as `city_list`")
                ,DB::raw("COUNT(locations.loc_name) as count_row")
            )
            ->join('locations', 'daily_tasks.location_id', '=', 'locations.id')
            ->join('routes', 'locations.route_id', '=', 'routes.id')
            ->join('cities', 'locations.city_id', '=', 'cities.id')
            ->where('task_date',$request->filterdate)
            ->groupBy('routes.name')
            ->get();
        return view( view: 'daily-routes/proceed', ['task_routes'=>$task_routes, 'taskDate'=>$request->filterdate]);
    }
```

Get Task Routes

```php
$summery .= "<div><h4 class='sm-title'>Supplier vehicle allocation</h4></div>";

if ($need_sup_veh_count <= $veh_sup_rank_0) {
    $allocation_sup_0 = $need_sup_veh_count;
} else {
    $allocation_sup_0 = $veh_sup_rank_0;
    $filled_need_sv = $filled_need_sv + $veh_sup_rank_0;
    $remain_veh = $need_sup_veh_count - $filled_need_sv;

    $i = 1;
    while ($filled_need_sv < $need_sup_veh_count) {
        if ($remain_veh <= ${"veh_sup_rank_$i"}) {
            ${"allocation_sup_$i"} = $remain_veh;
        } else {
            ${"allocation_sup_$i"} = ${"veh_sup_rank_$i"};
        }
        $filled_need_sv = $filled_need_sv + ${"allocation_sup_$i"};
        $remain_veh = $need_sup_veh_count - $filled_need_sv;
        $i++;
    }
}

$suppliers = Suppliers::orderBy('rank', 'desc')->limit($i)->get();
$sup_count = $suppliers->count();

$sup_index = 0;
foreach ($suppliers as $supplier) {
    $summery .= "<div>Supplier : " . $supplier->supplier_name . "</div>";
    $summery .= "<div>Hiring requirement : " . ${"allocation_sup_$sup_index"} . "</div>";
    $summery .= "<div class='sep'></div>";
```

Supplier vehicle allocation

```
    $avl_vehicles = DB::select('SELECT vehicle_no FROM vehicles
    WHERE vehicle_no NOT IN
    (SELECT vehicle_no FROM vehicle_repairs where release_date >= "'.$task_date.'"
    UNION SELECT vehicle_no FROM special_tasks WHERE task_date = "'.$task_date.'")');

    $avl_vehicles_array = array();
    $av = 0;
    foreach ($avl_vehicles as $avl_vehicle){
        $avl_vehicles_array[$av] = $avl_vehicle->vehicle_no;
        $av++;
    }


    //9.Get vehicles released after repairs and have high breakdown rate
    $own_alloc_one_select_veh = array();
    $sv_ind = 0;

    $veh_released_yd = DB::select('SELECT vehicle_no FROM vehicle_repairs where release_date = "'.$prev_date.'"')

    foreach ($veh_released_yd as $veh_released_yd_single){
        $own_alloc_one_select_veh[$sv_ind] = $veh_released_yd_single->vehicle_no;
        $sv_ind++;
    }

$brdn_period_start = date( format: 'Y-m-d', strtotime( time: '-30 days'));

$high_bd_rate_vehicles = DB::table('vehicle_breakdowns')
    ->select("vehicle_breakdowns.vehicle_no"
        ,DB::raw("(GROUP_CONCAT(vehicle_breakdowns.incident_date SEPARATOR ',')) as `breakdwn_list`")
        ,DB::raw("COUNT(vehicle_breakdowns.vehicle_no) as brkdwn_count")
    )
    ->where('incident_date','>=',$brdn_period_start)
    ->groupBy('vehicle_breakdowns.vehicle_no')
    ->get();

foreach ($high_bd_rate_vehicles as $high_bd_rate_vehicle){
    $brdwn_count = $high_bd_rate_vehicle->brkdwn_count;
    if($brdwn_count >= 3){
        $own_alloc_one_select_veh[$sv_ind] = $high_bd_rate_vehicle->vehicle_no;
        $sv_ind++;
    }
}

$own_alloc_one_prev_runs = array();

foreach ($own_alloc_one_select_veh as $own_alloc_one_selected){

    $rl_veh_record = DailyJobs::where('vehicle_no',$own_alloc_one_selected)
        ->where('job_date',$prev_date)
        ->join('routes', 'daily_jobs.route_id', '=', 'routes.id')
        ->select('daily_jobs.vehicle_no', 'routes.id as route_id', 'routes.distance_km as distance')
        ->get();

    if($rl_veh_record->isEmpty()) { //if vehicle has not allocated for a run yesterday
        $data = array(
            "vehicle_no"=>$own_alloc_one_selected,
```

Vehicle Allocation and Rotation

Appendix C - Test Cases

ID	Function	Testing Procedure	Expected Output	Test Priority
1	Create customer location.	Create a location where the city is not assigned for a route.	It should display a notification that there is not any matching route.	High
2	Edit a task.	Edit a task which is operation has completed.	Completed tasks should not be editable.	High
4	Generate daily routes	Generate route for a task date.	Routes list should be displayed according to the task locations.	High
5	Check vehicle requirements.	Check vehicle requirement for a task date.	The vehicle requirement for a route should be correctly matched with the work that can cover in that route.	Medium
6	Create the daily plan.	Trying create the daily plan once the report is generated for the task date.	Once the report is generated the continue button should be disabled.	High
7	Create the daily plan.	Try to create the daily plan with the minimum number of locations for a route.	An error should display allocating a vehicle for the route without sufficient job requests is not profitable.	Medium
8	Create the daily plan.	Try to create the daily plan with tasks that cannot cover by one vehicle but not enough tasks to allocate multiple vehicles.	An error should display allocating multiple vehicles is not profitable.	High
9	Check supplier vehicle requirement	Check supplier vehicle requirements with supplier rank.	Supplier vehicle requirement match with the supplier ranks.	Medium
10	Check supplier vehicle allocation	Check supplier vehicle allocation with allocated routes.	Supplier vehicles should not be allocated to the most profitable route range.	High
11	Check supplier vehicle allocation	Check supplier vehicle allocation with previous day routes.	Supplier vehicles should not be allocated to the same route.	Medium
12	Check own vehicle allocation	Check own vehicles which are released from repair center recently.	Recently repair released vehicles should not be allocated for long runs	Medium

13	Check company-owned vehicle allocation	Check company-owned vehicle allocation which has the high breakdown rate for last month.	Vehicle recently repaired multiple times should not be allocated for long runs.	Medium
14	Check company-owned vehicle allocation.	Check company-owned vehicle allocation with vehicle fuel economy.	Vehicles that have high fuel economy rate should be allocated for long runs.	Medium
15	Check own vehicle allocation	Check own vehicle allocation with vehicle fuel economy.	Vehicles have low fuel economy rate should not be allocated for long runs	Medium
16	Check own vehicle allocation	Check own vehicle allocation with previous day routes.	Own vehicles should not be allocated to the same route.	Medium
17	Check driver allocation	Check driver allocation with previous day routes.	Driver should not be allocated to the same route.	Medium