Project Charter

**Project Name:** Light Duty Vehicles Retail Sale Analysis

**Department:** Retail Sale

**Focus Area:** Fuel Consumption Ratings and Carbon Dioxide Emission

**Product/Process:** Data Analysis

**Prepared By**

| **Document Owner(s)** | **Project/Organization Role** |
| --- | --- |
| Arushi Gupta | Project Lead |
| Anik Roy | Team Member |

**Project Code Version Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Change Description** |
| 1.0 | 11/15/22 | Anik Roy | * Business problem statement * Import the dataset * Exploratory descriptive analysis |
| 2.0 | 11/15/22 | Mansi Sinha | * Exploratory descriptive analysis |
| 3.0 | 11/15/22 | Divyanshi Soni | * Dataset description * Data Wrangling |
| 5.0 | 11/17/22 | Arushi Gupta | * Code review * Documentation |
| 7.0 | 11/18/22 | Mansi | ● Exploratory descriptive analysis  ● Initial model setup |
| 8.0 | 11/21/22 | All team members | * Model local deployment |
| 9.0 | 11/21/22 | All team members | * Multiple model set up to compare and pick an optimal model |
| 11.0 | 11/22/22 | Arushi | * Code Review |
| 12.0 | 11/22/22 | All team members | * Web App deployment * User manual |

**TABLE OF CONTENTS**

[**PROJECT PURPOSE**](#_heading=h.39kk8xu) **4**

[**PROJECT EXECUTIVE SUMMARY**](#_heading=h.48pi1tg) **4**

[**PROJECT OVERVIEW**](#_heading=h.184mhaj) **5**

[**PROJECT SCOPE**](#_heading=h.279ka65) **5**

[4.1 Goals and Objectives](#_heading=h.meukdy) 5

[4.2 Project Deliverables](#_heading=h.243i4a2) 6

[4.3 Project Duration (Start date: 11/15/2022 - End date: 11/24/2022 )](#_heading=h.38czs75) 6

[**PROJECT CONDITIONS**](#_heading=h.415t9al) **7**

[5.1 Project Assumptions](#_heading=h.2gb3jie) 7

[5.2 Project Constraints](#_heading=h.3rdcrjn) 8

[**PROJECT STRUCTURE APPROACH**](#_heading=h.4bewzdj) **8**

[**PROJECT TEAM ORGANIZATION PLANS**](#_heading=h.15phjt5) **8**

[**PROJECT REFERENCES**](#_heading=h.3jd0qos) **9**

[**APPROVALS**](#_heading=h.2xn8ts7) **10**

# PROJECT PURPOSE

With the accelerated growth of urbanization, environmental issues caused by transportation have been challenging due to the significant negative impact on climate change. Moreover, 20 to 30% of global greenhouse gases (GHG) are emitted from passenger and freight transportation, and 75% of total carbon dioxide emissions originate from passenger cars. Estimating and visualizing fuel consumption and exhaust emissions are critical for quantifying the energy cost and air pollution caused by transportation, as well as detailing emission control strategies.

# PROJECT EXECUTIVE SUMMARY

* Goals

To provide an insight into vehicle fuel consumption and carbon dioxide emission through a series of rigorous data analytics and machine learning.

* Objectives

To estimate carbon dioxide emission based on different factors.

* Scope

The scope of this problem is to identify what factors affect the carbon dioxide emission and fuel consumption of the vehicle.

* Assumptions

The dataset is accurate.

* Timeline

Our goal is to finish the project in 10 days, from November 15th, 2022 to November 24th, 2022.

* Approach

We will start by doing the exploratory descriptive analysis on the dataset. We will then extract the necessary data for the model application and do data preprocessing. Parallelly, domain and model research will be done to make the best decisions on the data preprocessing. We will then build the model on train data to predict on test data using different classification algorithms. After choosing the best algorithm among all, the final model will be tuned to make it better. Finally, a Web App would be deployed to predict the output.

# PROJECT OVERVIEW

With the accelerated growth of urbanization, environmental issues caused by transportation have been challenging due to the significant negative impact on climate change. Moreover, 20 to 30% of global greenhouse gases (GHG) are emitted from passenger and freight transportation , and 75% of total carbon dioxide emissions originate from passenger cars. Estimating and visualizing fuel consumption and exhaust emissions are critical for quantifying the energy cost and air pollution caused by transportation, as well as detailing emission control strategies

# PROJECT SCOPE

## 4.1 Goals and Objectives

|  |  |
| --- | --- |
| **Goals** | **Objectives** |
| * Explore the need for analyzing the amount of CO2 produced by various cars. | * Research about the various reasons for difference in CO2 emissions in various cars. |
| * Understand the Dataset | * Research about the various models that are optimal for float data. |
| * Perform the Exploratory Data Analysis on the dataset and then do feature selection. | * Research about the features needed for applying such models. |
| * Perform descriptive analysis | * Research various ways to visualize the data. |
| * On the basis of previous analysis determine the possible ways models can be created to predict CO2 emissions. | * Research various ways to preprocess the data. |
| * Choose best model based on model metrics | * Research about the best model metric to be used to compare models. |
| * Tune the hyperparameters of the model | * Research about the various ways to perform hyperparameter tuning. |
| * Model deployment | * Research about how to deploy the Web App. |

## 

## 4.2 Project Deliverables

| **Milestone** | **Deliverable** |
| --- | --- |
| 1. Explore the dataset | Clear documentation on data and data visualization |
| 2. Data cleansing | Clear documentation on the data EDA and sentiment analysis |
| 3. Data preprocessing | Clear documentation on how the tweet text is altered before tokenization |
| 4. Model selection | Clear documentation on how the best fit model was chosen and tuned |
| 5. Model deployment | Web App deployment for client’s access |

## 4.3 Project Duration (Start date: 11/15/2022 - End date: 11/24/2022 )

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Milestone** | **Date Estimate** | **Deliverable(s) Included** | **Confidence Level** |
| Understanding the problem statement | 11/15/2022 to 11/16/2022 | * Prepare task lists and notes on problem understanding | High |
| Preparing the project charter | 11/16/2022 to 11/18/2022 | * Goals and Objectives * Project deliverables * Deliverables out of scope * Project duration | High |
| Understanding the dataset | 11/15/2022 to 11/16/2022 | * Dataset inspection * Features and label identification * Perform exploratory descriptive analysis | High |
| Data preparation | 11/17/2022 to 11/18/2022 | * Prepare the initial dataset for modeling using feature engineering | Medium |
| Getting the best model | 11/18/2022 to 11/20/2022 | * Perform different modeling techniques such as Multinomial Logistic Regression, Naïve Bayes Classifier, Random Forest, Decision Tree * Compare metrics to choose the best model among all | Medium |
| Model evaluation | 11/21/2022 to 11/22/2022 | * Evaluate results obtained on the test data by the final model chosen * Tune the model to obtain better results | Medium |
| Model deployment | 11/22/2022 to 11/23/2022 | * Create a User Manual for the client to test | Medium |
| Evaluation | 11/24/2022 | * Evaluation of the User Manual by the Project Manager, Stakeholders and Client | Medium |

# PROJECT CONDITIONS

## 5.1 Project Assumptions

* It is assumed that in the categorical columns the spellings of each category are correct.
* We combined several data in categorical analysis, assuming they behave the same way.
* We are assuming that the training data is well rounded to fit our model such that it is right fitted.
* We are assuming that FUELCONSUMPTION\_COMB can uniquely identify CO2 Emissions and not using the CITY and HWY values separately.

## 

## 5.2 Project Constraints

## The project constraint is to minimize the margin of error in predicting the carbon dioxide emissions.

## 

# PROJECT STRUCTURE APPROACH

The project is to be implemented using the Multiple Linear Regression process.

# PROJECT TEAM ORGANIZATION PLANS

|  |  |  |
| --- | --- | --- |
| **Project Team Role** | **Project Team Member(s)** | **Responsibilities** |
| Project Management | Arushi Gupta | * Project Charter Review * Code Review * Documents Review |
| Data Handling | Entire Team | * Data Research * Data Understanding |
| Data Preparation / EDA | Entire Team | * Data Cleansing * Data Visualization * Data Balancing |
| Model Building | Anik Roy  Divyanshi Soni | * Model Research * Model Building * Model Tuning |
| Model Deployment | Anik Roy | * Model Deployment |
| Documentation | Anik Roy | * Error Document * User Manual |
| Arushi Gupta  Divyanshi Soni | * Project Charter * Technical Document |
| Mansi Sinha | * Project Progress Document |

# APPROVALS

**Prepared by** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Manager

**Approved by** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Sponsor