Classification and Geo-location Analysis of LPG Consumers

A

Project Report

submitted in partial fulfillment of the
requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE

Specialization in

Oil And Gas Informatics

By :

Name	Roll No	Branch
Diwakar Kumar	R970216025	CSE OGI
Gaurav Mishra	R970216028	CSE OGI
Abhishek Suman	R970216004	CSE OGI

Under the guidance of

Dr. Kingshuk Srivastava
Assistant Professor (SG)
Department of Informatics

UPES
UNIVERSITY WITH A PURPOSE

Department of Informatics
School of Computer Science
UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
Bidholi, Via Prem Nagar, Dehradun, Uttarakhand
2019-20

CANDIDATES DECLARATION

We hereby certify that the project work entitled **Classification and Geo-location Analysis of LPG Consumers** in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science And Engineering with Specialization in Oil and Gas Informatics and submitted to the Department of Informatics at School of Computer Science, University of Petroleum And Energy Studies, Dehradun, is an authentic record of our work carried out during a period from **August, 2019** to **December, 2019** under the supervision of **Dr. Kingshuk Srivastava, Assisstant Professor(SG), Department of Informatics** .

The matter presented in this project has not been submitted by us for the award of any other degree of this or any other University.

Name of Student(s)	Diwakar Kumar	Gaurav Mishra	Abhishek Suman
Roll No.	R970216025	R970216028	R970216004

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Date: 12 December 2019)   **(Dr. Kingshuk Srivastava)**
                                                                                     Project Guide

Dr. Thipendra Pal Singh

Head
Department of Informatics
School of Computer Science
University of Petroleum And Energy Studies
Dehradun - 248001 (Uttarakhand)

ACKNOWLEDGEMENT

Name of Student(s)	Diwakar Kumar	Gaurav Mishra	Abhishek Suman
Roll No.	R970216025	R970216028	R970216004

ABSTRACT

Humanity faces a unique and far reaching challenge. Our energy needs are growing as a result of continued population increases, economic growth, and individual fuel/energy consumption. At the same time, emissions from fuel wood and fossil fuels, the main energy source for heating in homes and powering our economies, are contributing to climate change and affecting the local air quality. Liquefied Petroleum Gas(LPG) is used as fuel for thousands of applications. In developing countries, the main benefits of LPG is in helping people to switch from unsustainable biomass use to a clean and safe cooking fuel. LPGs domestic uses can never be ignored. It has played a revolutionary role when it comes to changing the face of domestic fuels used for heating and cooking. This project consists of two modules. First part classifies different LPG consumers using geographical map. And second part will depict the minimum cost path of distributorship.

Keywords: LPG, Consumers, Geo-location, Classification.

TABLE OF CONTENTS

Contents

LIST OF FIGURES

List of Figures

1 Introduction

Analysis refers to breaking a whole data into smaller chunks for individual examination. Data analysis is a methodological approach of inspecting, filtering, transforming, and designing a model with the goal of extracting useful information from the data for better decision making. There are a several data analysis methods including data mining, text analytics, business intelligence and data visualization which implements analytical and statistical tools for gaining insight of useful information. Data analysis is a process for obtaining raw data and converting it into useful information for better decision-making and extracting conclusions from big data by users and analysts as well as it provides in-depth insight into data.

Classification is the process of identifying the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification predictive modelling is the task of approximating a mapping function from input variables to discrete output variables.

Across the world, LPG industry is continuously facing strong competition all over the world. LPG sector is constantly facing challenges from the market. The systematic delivering and high service quality is to be systemize for maintaining customer allegiance. With these crowded competitive business environments, utilities have to spread his business processes in every area. All over the world, gas prices are one of the major issues. The reduced prices in global market can binds the customers. According to Hovland (2007)[1], Some managing parameters like maintenance of service orders, construction work orders, day to day customer service request are getting difficult with the use of old tools of management. Pati et al. (2010)[2] explained that in business environment employees attitude is very much necessary for increasing organizational performances. In India, for establishing service quality standard of LPG several surveys have to be organised to obtain the problems associated with the service quality evaluation and customer satisfaction of LPG industries for getting cost effective solutions. In some parts of India, performance standard is up to the mark of customer satisfactory level but on other hand there are some places where customers are not satisfied with the LPG industry. Pricing and distribution is the basic problems faced by Indian customers. The gas distribution problem occurs due to the lack of interest among the consumers.

2 Background Study

Here is the conclusion of some of the reference paper that we review to make our project better and to know more technologies that we can use in our system.

- In this paper[3], Liquefied Petroleum Gas is described as fuel for thousands of applications. In developing countries, the main benefits of LPG is in helping people to switch from unsustainable biomass use to a clean and safe cooking fuel. It also depicts that LPGs domestic uses can never be ignored. It also states that HPCL commenced marketing of LPG under the brand name "HP GAS". The data were fully based on Primary & Secondary data collection.

- In this paper[2], they have shown how today's LPG market complains registration is done through computerised process by the customers and the customers are becoming more sensitive to LPG gas problems. It also shows how customers have the right to select their suppliers of LPG as well as good service quality. For understanding these complex processes and developing tools, factor analysis has been conducted. On the basis of questionaries survey, they have also shown the service satisfaction of Indian consumers are studied and suggestions are provided to improve service quality.

- In this paper[1], the study seeks among others to investigate the magnitude of noted disruptions that contribute to the shortage of LPG and the extent to which the LPG usage categories influence the length and frequency of shortage. A survey methodology was deployed for a total of 35 LPG retailers and 120 consumers in the Ashanti Region of Ghana. Results of the study indicated that the Technical and Transportation disruption was ranked by retailers and consumers as major factors that disrupt the supply of LPG.

3 Problem Statement

One can't understand the customer behavior by looking at the given dataset. Analysis is required to make probable understanding of dataset. It will be easy to visualize a pattern rather than textual representation.

4 Objective

To classify and analyze the consumption pattern of different consumers based on their geographical location.

5 Methodology

Step 1: Understanding the concept of classification and population analysis.

Step 2: Gathering and filtering datasets.

Step 3: Understanding different dataset attributes.

Step 4: Identifying mathematical formulations used for classification and analysis.

Step 5: Designing flowcharts and algorithm.

Step 6: Coding and Visualization.

Step 7: Report.

6 Design

The design part consists of different phases:

1. Data Collection: Selecting appropriate and relevant data is one of the major parts of this project. For data collection, we have collected all the dataset from HINDUSTAN PETROLEUM CORPORATION LTD., Mumbai. The dataset comprises of different attributes like latitude and longitude of different LPG consumers, distributor code, region code, suppliers name etc.

2. Understanding and wrangling of datasets: The collected datasets have different outliers which needs to be removed for better analysis and visualization. Understanding of attributes of dataset is performed at this phase. Data wrangling is process of transforming and mapping raw data form into another format.

3. Understanding different libraries: For this project we were using different python libraries which are pandas, numpy, math, folium, etc.

- Panda- pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with structured (tabular, multidimensional, potentially heterogeneous) and time series data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, real world data analysis in Python.

- Numpy-NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.

  It is the fundamental package for scientific computing with Python. Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data.

- Folium -olium is a Python Library that can allow us to visualize spatial data in an interactive manner, straight within the notebooks environment many (at least myself) prefers. The library is highly intuitive to use, and it offers a high degree of interactivity with a low learning curve.

4. Implementation: For the implementation part, we are using python programming language. Some mathematical formulas are used for calculation of geolocation distance of an LPG consumer from its distributorship.

s = 6378.1 * acos(sin(slat)*sin(elat) + cos(slat)*cos(elat)*cos(slon1 - elon))

where, s is distance
slat and slon are distributors latitude and longitude
elat and elon are consumers latitude and longitude

From the calculated distance we will categorize different LPGs consumers.
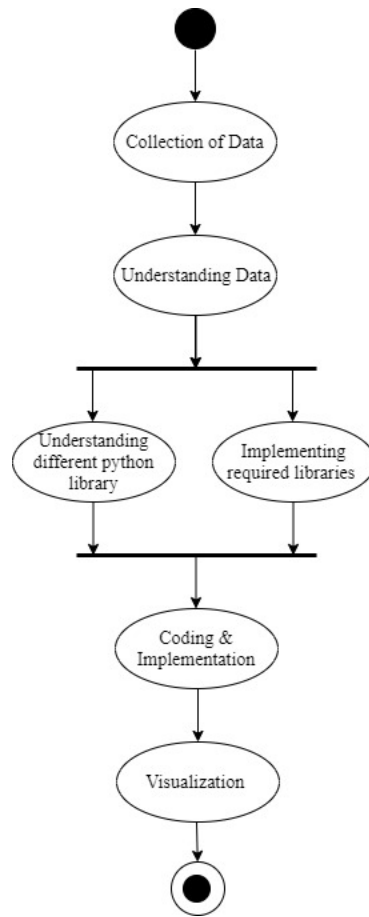
5. Visualisation.

Figure 1: *Activity Diagram*

7 Implementation

7.1 Psuedocode

import folium, pandas library

7.1.1 Reading csv file and storing its attributes in list

Initializing initial parameters
data=read_excel("EZYGas.xlsx")
d=data.set_index("ConsId")
lat=list(data["Latitude"])
lon=list(data["Longitude"])
id=list(data["ConsId"])
declare d as list()
for each id i do
set slat as 18.5723333 radians
set slon as 73.8911667 radians
set elat as radians(lat[i])
set elon as radians(lon[i])
calculate distance as dist
print distance is dist
append dist in d

4

7.1.2 Creating map feature group for presenting multiple markers

Initialize map with location, zoom_start, tiles
create map as "myMap"
Use location and popup as Marker to map using add_child
for each lat,lon,d in zip:
fg.add_child(folium.CircleMarker(location = [lat, lon], radius = 4 , popup =str(d)+"Km", fill_color= "red", fill_opacity = 0.7))

7.1.3 Adding markers to map

map.add_child(fg)

7.1.4 Saving the map

map.save("EZYMap2.html")

7.2 Output Screen

Figure 2: *Output 1*

5
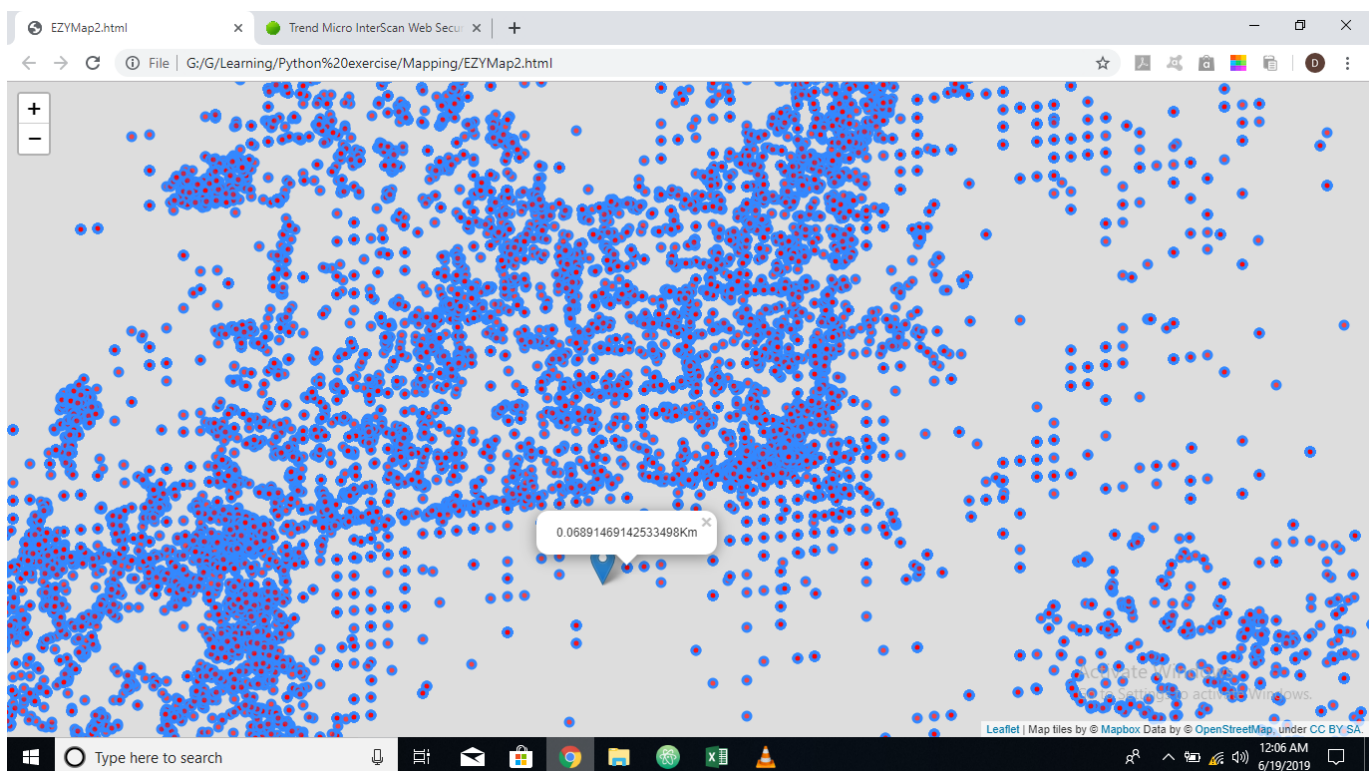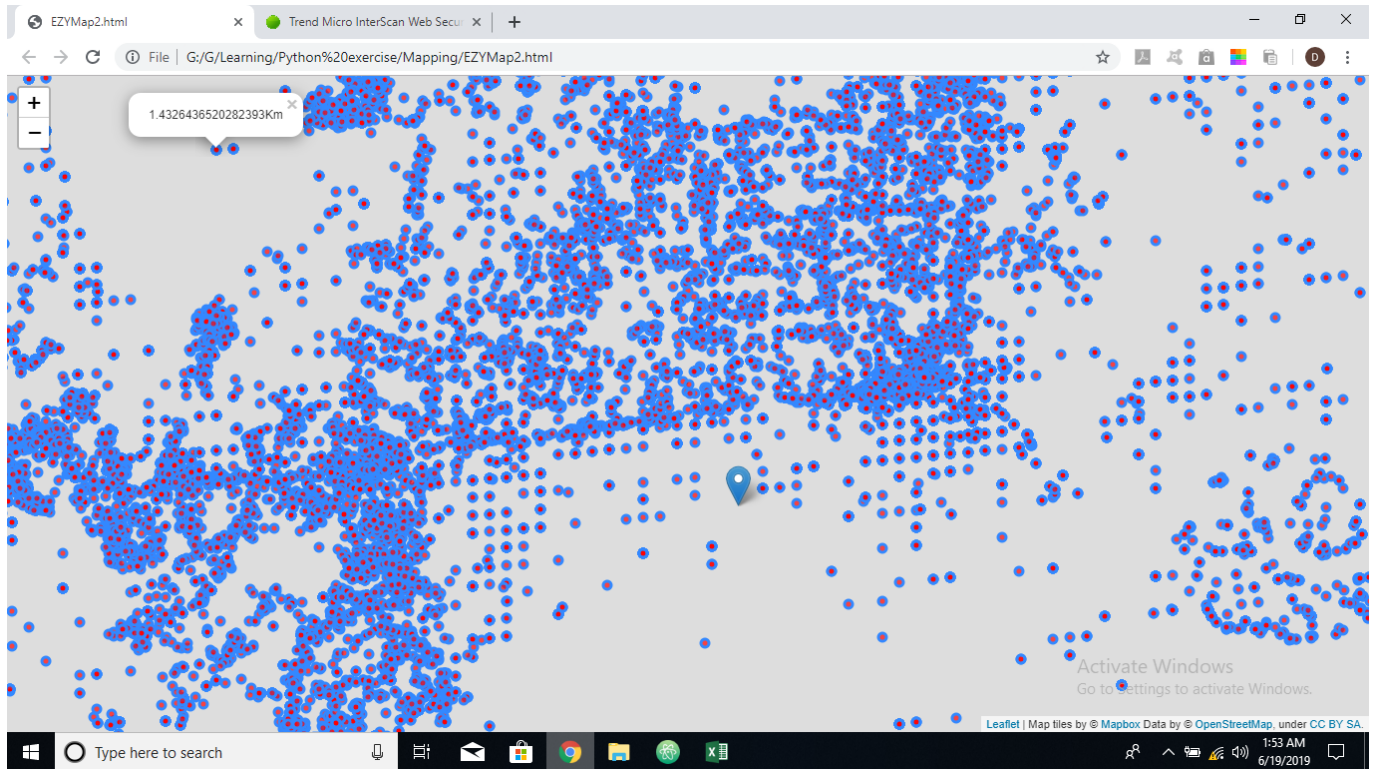
Figure 3: *Output 2*
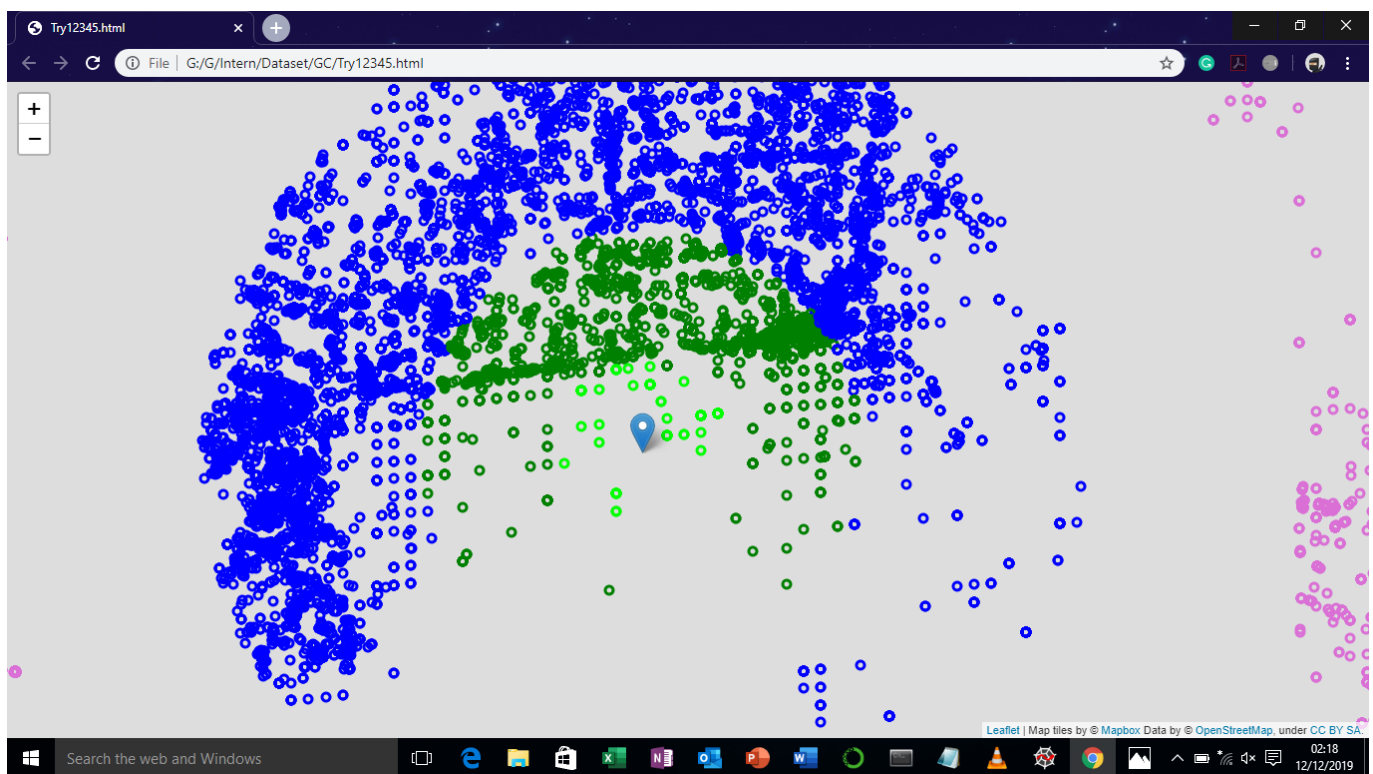
Figure 4: *Color Analysis of Consumer*

7.3 Result Analysis

- First, dataset of consumers which falls under a single distributorship is identified and filtered. After that consumers are mapped on the map interface.
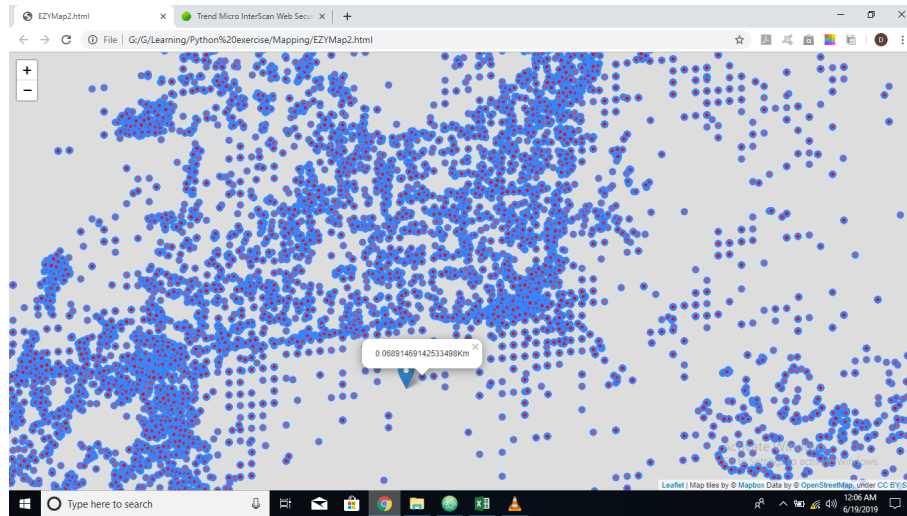

Figure 5: *Visualization of Dataset*

- Distance of consumers from distributorship is calculated using the latitude and longitude position of consumers and distributorship. Geolocation of distributorship [18.5723333, 73.8911667]

Figure 6: *Distance Calculation*

- Based on the distance calculated, consumers are clustered into coloured groups.
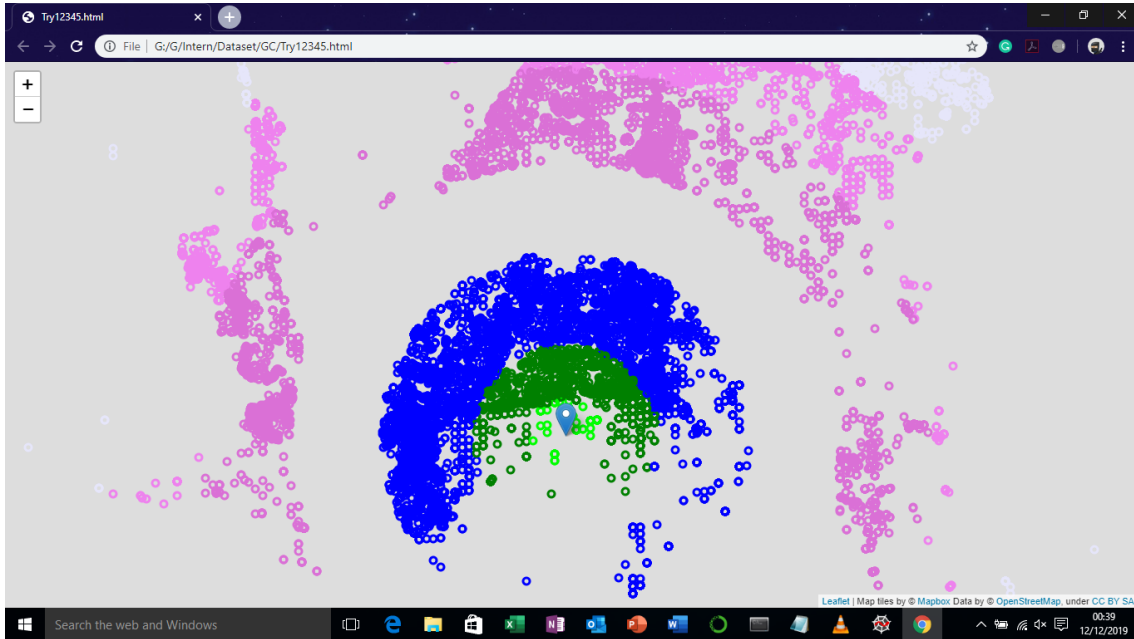
Figure 7: *Clustering of Consumer*

- Based on the distance from the distributorship, minimum cost path is determined.
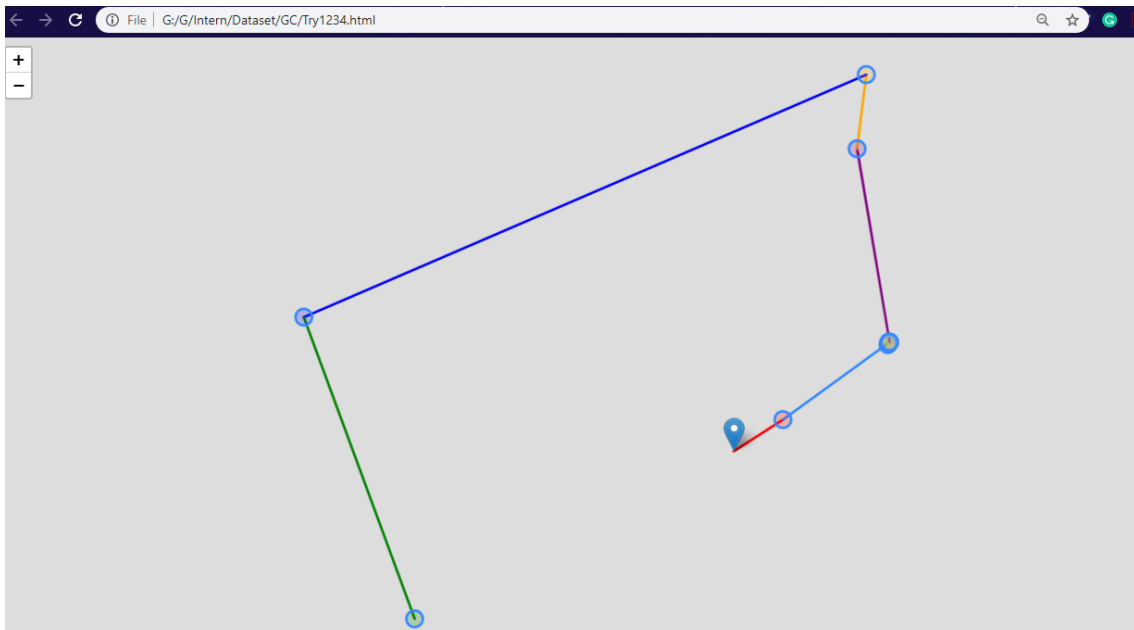
Figure 8: *Minimum Cost Path*

8   Conclusion and Future Scope

Liquefied Petroleum Gas (LPG) is increasingly becoming the preferred choice of fuel in the world. The increase in patronage could be attributed to its affordability, efficiency and environmental friendliness. Like other energy sources, there are challenges that impede the smooth supply of LPG resulting in shortage. Several factors are involved which ultimately affects the efficiency of the entire Supply Chain system. The perennial LPG shortage is an uncomfortable experience that needs urgent attention. Therefore, the study seeks among others to investigate the magnitude of noted disruptions that contribute to the shortage of LPG and the extent to which the LPG usage categories influence the length and frequency of shortage.This project classifies different LPG consumers using geographical map and finally the minimum cost path of distributorship is identified. Now-a-days data interpretation and analysis is one of the crucial parts of business insights. This project depicts the colour-based classification of consumers which falls under the same range of distance fromm distributorship. In last, minimum cost path of travelling is calculated and analysed. This project can be further implemented using Graphical User Interface which makes interaction between the user or analyst and the system environment more interactive and efficient. For graphical user interface implementation, user can implement it using Tkinter library. For more accuracy we can also use altitude. Suppose two people living in the same building have same latitude and longitude position then in that case, the analyst or user will use altitude to differentiate between two users.

References

[1] D. Asamoah, R. Amoakohene, and E. Adiwokor, "Analysis of Liquefied Petroleum Gas (LPG) Shortage in Ghana : A Case of the Ashanti Region," vol. 3, no. 5, pp. 89–98, 2015.

[2] V. Yadav, S. Mullick, and S. Satapathy, "Domestic Gas Delivery and Service Satisfaction of," vol. III, no. Vii, pp. 266–274, 2014.

[3] K. District, "A STUDY ON CONSUMER ATTITUDE TOWARDS DOMESTIC LPG GAS WITH SPECIAL REFERENCE TO KRISHNAGIRI DISTRICT," vol. 4, no. 3, pp. 3–5, 2015.