**Project No: 2**

**Project Title: Facility Location Optimization**

**Prepared by: Rose Mary Jose**

**Date: 23-06-23**

Prepared for “Thoughtware Training Private Limited”, under Guidance of its CEO Mr. Pattabhi Raman. The project will be subject to further research, modification and exclusive use of “Thoughtware Training Private Limited”

**Objective of the project**

Facility location problem (FLP) is a classic optimization problem. The problem involves selecting the best location for a facility such that the transportation cost and other related costs, and the distance is minimized. The problem is commonly used in logistics and supply chain management since the efficient and effective movement of goods from raw material sites to processing facilities, component fabrication plants, finished goods assembly plants, distribution centres, retailers and customers is critical in today’s competitive environment. So, it is important to optimize the location of distribution centres, warehouses, and other related facilities based on geographical demands, facility costs and transportation distance.

The aim of this project is to find an optimized facility location whose load-distance column is minimized. Also, the optimized facility should have minimum distance with the customer’s location. This could be able to find the optimal location when the load of customer location is changed as necessary.

**Dataset**

The dataset considered consist of the **4 columns and 24 rows**, i.e., the dataset consists of information of 24 customer locations

customer\_location: the location of the customers is given as city names.

Latitude: the latitude of respective customer location is given.

Longitude: the longitude of the respective customer location is given.

Load(Ton): the demand of each customer location is given in ton units.

## importing necessary libraries  
import numpy as np  
import pandas as pd  
import io  
import seaborn as sns  
import matplotlib.pyplot as plt  
import warnings  
warnings.filterwarnings("ignore")

## reading the dataset  
data=pd.read\_excel('Facility\_location.xlsx')  
data

customer\_location Latitude Longitude Load(Ton)  
0 Kochi 9.973543 76.296246 7  
1 Chennai 13.083592 80.255685 10  
2 Bangalore 12.993423 77.568437 15  
3 Panaji 15.491269 73.828136 3  
4 Pune 18.525901 73.859873 5  
5 Hyderabad 17.402453 78.492485 7  
6 Indore 22.727118 75.862472 4  
7 Lucknow 26.855947 80.953482 5  
8 Patna 25.600270 85.143725 3  
9 Mumbai 19.092550 72.901081 15  
10 Delhi 28.712359 77.209895 15  
11 Surat 21.173907 72.825038 7  
12 Jaipur 26.924920 75.771320 5  
13 Ahmedabad 23.026477 72.593458 8  
14 Kolkata 22.578843 88.350782 12  
15 Bhopal 23.281345 77.392141 6  
16 Visakhapatnam 17.703284 83.230988 7  
17 Faridabad 28.419490 77.297439 3  
18 Noida 28.548644 77.389449 6  
19 Nagpur 21.160302 82.571020 3  
20 Mirzapur 25.139028 82.560377 7  
21 Amrister 31.639363 74.870460 9  
22 Ranchi 23.348542 85.311022 4  
23 Guwahati 26.118912 91.713853 8

The dataset is given above.

**EDA**

The shape of the dataset considered is (24, 4). i.e., 24 rows and 4 columns.

shape of dataset: (24, 4)  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 24 entries, 0 to 23  
Data columns (total 4 columns):  
 # Column Non-Null Count Dtype   
--- ------ -------------- -----   
 0 customer\_location 24 non-null object   
 1 Latitude 24 non-null float64  
 2 Longitude 24 non-null float64  
 3 Load(Ton) 24 non-null int64   
dtypes: float64(2), int64(1), object(1)  
memory usage: 896.0+ bytes  
None

Info of the dataset returns the information about the data frame. It includes the number of columns, column labels, datatypes, non-null count of the columns, memory usage, and range index. The dataset consists of 24 rows and 4 columns, in which 1 is of dtype object, 1 is of int64, and 2 of float64.

Describe function is used to get the summary statistics of the dataset.

Describe function of the numerical column summarizes the central tendency, dispersion and shape of the dataset’s distribution. It returns the count, mean, standard deviation, minimum value, maximum value, 25th, 50th, and 75th percentile values. From this it is observable that the demand of the customer location is varied. It ranges from 3 to 15. From this it can be concluded that different places have different demand the facility location need to be located somewhere near to the high demand place. As the demand of places changes the facility location also get changed.

Latitude Longitude Load(Ton)  
count 24.000000 24.000000 24.000000  
mean 22.063395 78.927036 7.250000  
std 5.610915 5.120331 3.767712  
min 9.973543 72.593458 3.000000  
25% 18.320246 75.546105 4.750000  
50% 22.876798 77.390795 7.000000  
75% 26.303171 82.563037 8.250000  
max 31.639363 91.713853 15.000000

**Data Preparation**

The dataset considered is clean since there is no missing values and the data doesn’t contain duplicate values.

**Model Selection**

The objective of the project is to find the optimal location whose load distance is minimum. Optimal facility location problem is an optimization problem in supply chain management. Optimization in supply chain management refers to the process of finding the best possible solution or configuration that maximized efficiency, minimizes costs, and improves overall performance within the supply chain network. It involves using mathematical modelling, algorithms, and analytical techniques to optimize various aspects of the supply chain, including inventory levels, transportation routes, production schedules, facility locations, and resource allocation. There are different types of optimization problem in supply chain management such as, cost minimization, service level maximization, resource utilization, risk mitigation, sustainability etc.

To achieve optimization in supply chain management, various techniques and tools are employed, such as linear programming, network optimization, simulation, predictive analytics, and machine learning. These methods help in analyzing large amount of data, considering multiple constraints and objectives, and finding the best possible solutions that optimize the supply chain operations.

Here the problem is to find the optimal location whose load distance is minimum. In order to do the optimization firstly the mathematical formulation is to be done.

**Mathematical Formulation**

Mathematical formulation refers to the process of representing the problem mathematically using variables, constraints, and an objective function. From the type of objective function optimization technique is chosen. There are different types of optimization technique such as linear programming, integer programming, non-linear programming etc.

Decision variables: these are the variables that represent the quantities or decisions to be optimized. They are typically denoted by symbols and can take on specific values later.

Constraints: they represent the limitations or restrictions on the decision variables. They define the feasible region of solutions. Constraints can be mathematical equations or inequalities that the decision variables must satisfy.

Objective function: it defines the goal of the optimization problem. It quantifies the measure to be maximized or minimized. It is typically a mathematical expression involving the decision variables.

Parameters: they are the fixed value given in the problem and influence the optimization process. These are known values and do not change during the optimization.

The variables in this facility location problem are the optimal latitude and the optimal longitude of the output facility location.

In this problem there are no constraints.

The objective function:

D = distance between optimal location and the customer locations in the data set.

L = load(Ton) of each customer location.

The objective function is given by,

Min

Where i is the range of number of customer locations in the dataset.

**Optimization technique**

Since the mathematical formula considered in this problem is a non-linear, for optimization modelling the Pyomo library in python is used. Pyomo, python based open-source optimization modelling language, is a powerful optimization modelling language that allows users to easily create, solve, and analyze mathematical models in python. Pyomo allows to choose variety of solvers, both open-source and commercial. In this project ‘glpk’ solver is used in pyomo to find the optimal facility location.

!pip install pyomo

%%capture  
import sys  
import os  
  
if 'google.colab' in sys.modules:  
 !pip install idaes-pse --pre  
 !idaes get-extensions --to ./bin  
 os.environ['PATH'] += ':bin'

import pyomo.environ as pyo  
from pyomo.environ import \*  
from pyomo.opt import SolverFactory

After loading the data and doing necessary data preparations, model is to be build. Model is to be initialized and then decision variables are created. A ConcreteModel() is created since the data is provided at the moment.

## making the list of all columns of the dataframe  
customers = data['customer\_location'].tolist()  
latitudes = data['Latitude'].tolist()  
longitudes = data['Longitude'].tolist()  
loads = data['Load(Ton)'].tolist()  
num\_customers = len(customers)

# Define the Pyomo model  
model = pyo.ConcreteModel()

For the facility location optimization, the optimal latitude and optimal longitude are the decision variables, they are added using the var() function.

# Variables  
model.lat\_optimal = pyo.Var(domain=pyo.Reals)  
model.lon\_optimal = pyo.Var(domain=pyo.Reals)  
model.y = pyo.Var(customers, domain=pyo.Binary)

Then objective function is added. For the facility location optimization, the objective function is to minimize the load\*distance of the customer location with the facility location. For that distance is calculated for between each customer location and the optimal facility location. That distance is multiplied with respective loads and the minimum is calculated.

# Objective  
model.obj = pyo.Objective(expr=sum(loads[i] \* 111.12\*pyo.sqrt((latitudes[i] - model.lat\_optimal)\*\*2 + (longitudes[i] - model.lon\_optimal)\*\*2)for i in range(num\_customers)),sense=pyo.minimize)

The final step is to solve the model created using the ‘ipopt’ solver. The value of the created variables is found and the final output is accessed using print function.

# Solve the model using the Ipopt solver  
opt = SolverFactory('ipopt')  
results = opt.solve(model)

The value of the created variables, optimal latitude and optimal longitude and the value of the objective function is:

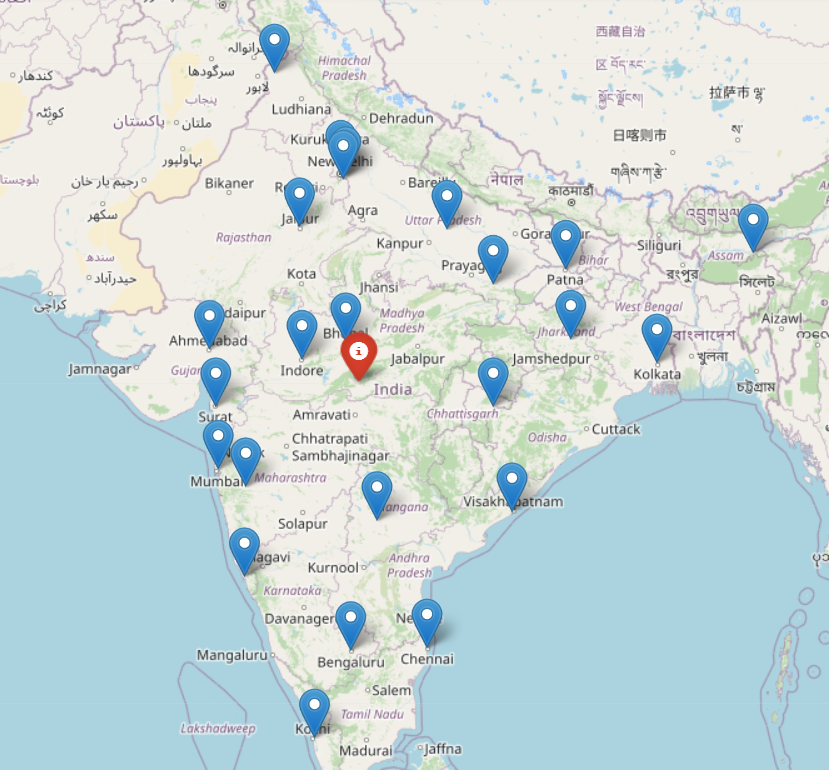
# Access the optimal latitude and longitude values  
optimal\_facility\_latitude = model.lat\_optimal.value  
optimal\_facility\_longitude = model.lon\_optimal.value

min\_load\_distance = model.obj()

Optimal Facility Latitude: 22.060033318577428  
Optimal Facility Longitude: 77.8418949340904  
Sum of Load \* Distance: 142294.84376304114

For the given dataset the optimal facility location is found out to be Chhuri, Madhya Pradesh (22.060033,77.84189), with a load distance of 142294.84 TonKM.

Using the folium library of python, the customer locations and the optimal facility location is plotted in map.



**Model Validation**

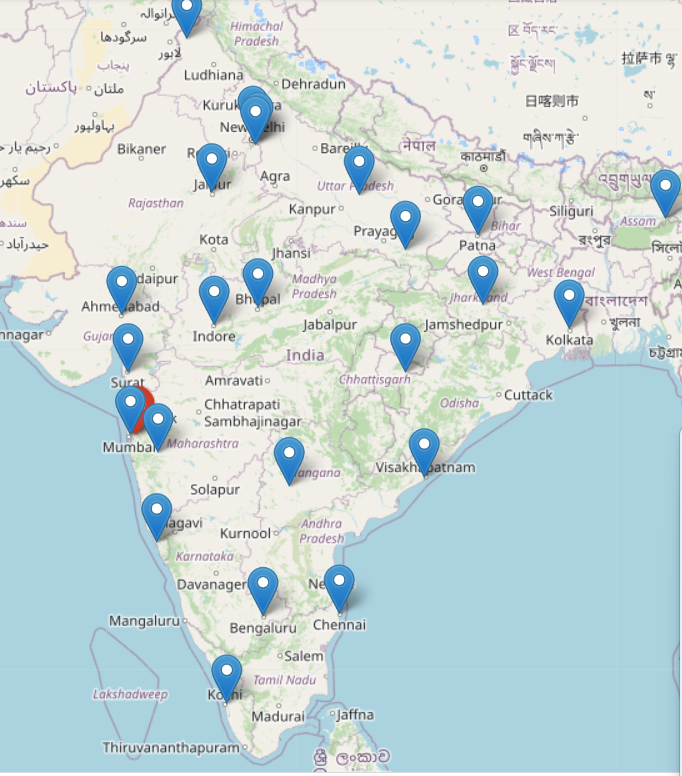
Model validation of this optimization is done by changing the load of a city and observing whether the change in the load of a city changes the facility location. It is done by changing the load of Mumbai to 100Ton from 15 Ton in order to see if location changes with respect to the load. The new dataset created is used by pyomo library for model building and ipopt solver is used to find the optimal latitude and longitude along with the load distance.

Similar steps used in earlier facility location optimization is used here. And the obtained output is given below.

Optimal Facility Latitude: 19.17785606246781  
Optimal Facility Longitude: 73.10209659048975  
Sum of Load \* Distance: 169817.85622576645

For the dataset with load at Mumbai as 100, the optimal facility location is found out to be Kalyan, Maharashtra (19.17785,73.10209) with a load distance of 169817.856 TonKM.

Using the folium library of python, the customer locations and the optimal facility location is plotted in map.



From the map it is clear that the facility location is near Mumbai where the load is maximum. So, the model created can be used for finding the optimal facility location whose load distance is minimum.

**Conclusion**

By optimizing facility locations, companies can achieve benefits such as reduced transportation costs, improved response times, enhanced customer service, minimized inventory levels, and increased overall efficiency in the supply chain. it helps businesses align their physical infrastructure with the demands of the market, leading to improved competitiveness and profitability.

By optimizing the supply chain, companies can achieve benefits such as reduced costs, improved customer satisfaction, increased efficiency, better resource utilization, enhanced competitiveness, and improved overall performance. Optimization enables businesses to make informed decisions, optimize their processes, and adapt to changing market conditions to stay competitive in a dynamic business environment.