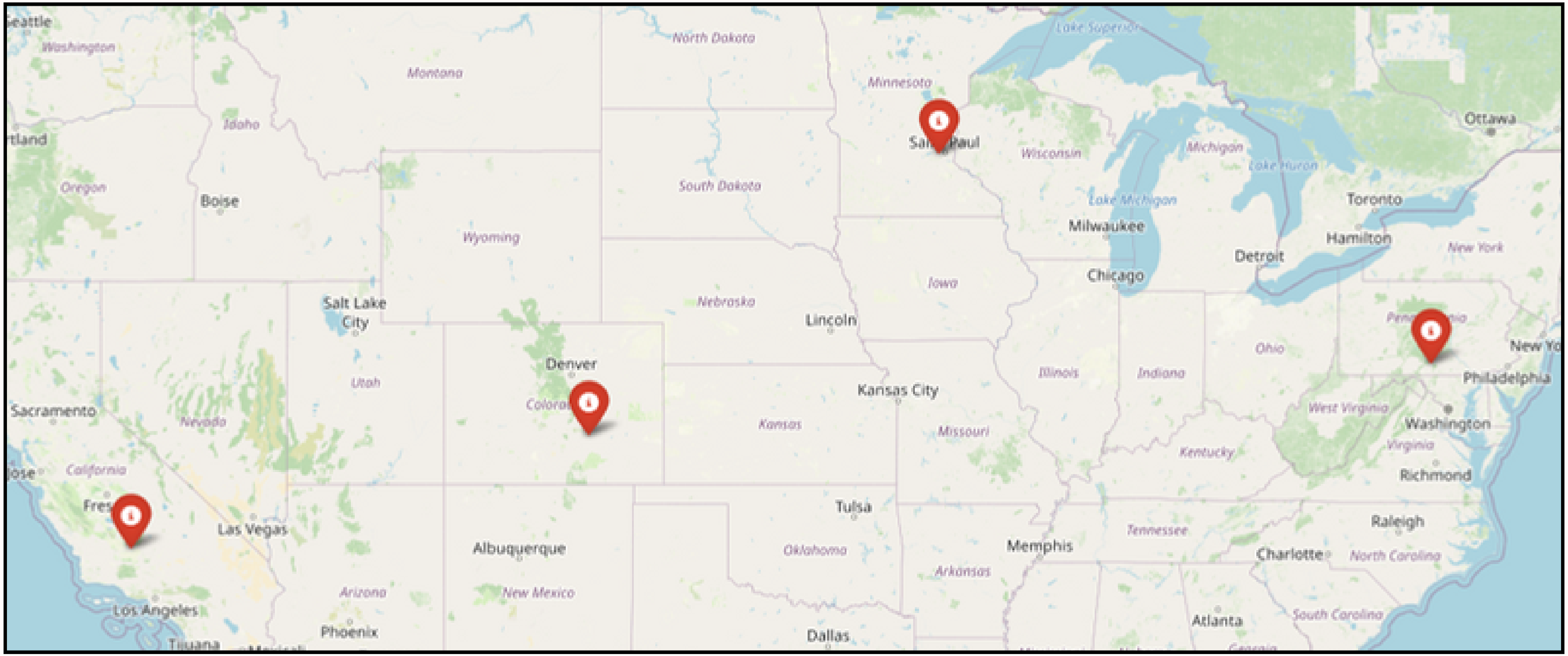
Strategic Expansion: Optimizing Target's Distribution Center Network

EXISTING DISTRIBUTION CENTERS LOCATIONS

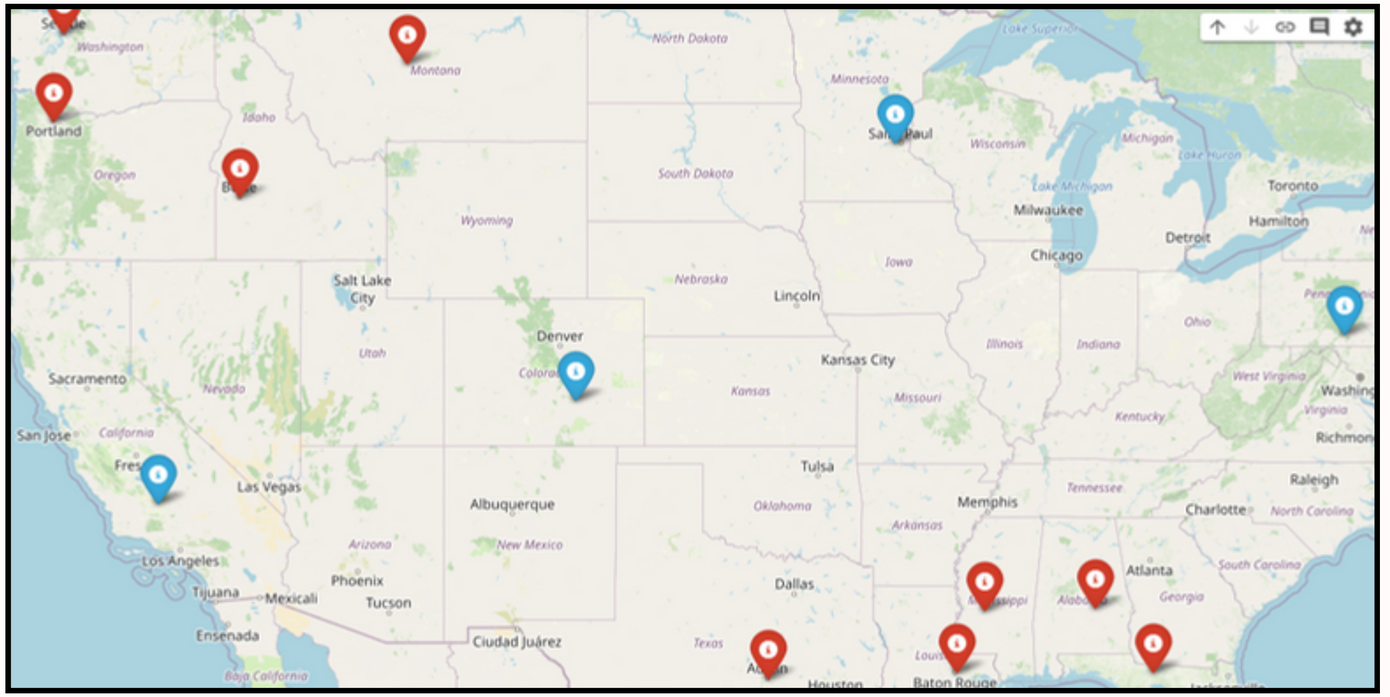


INEFFICIENCY OF EXISTING DC'S

Untouched states of United States

Handling inventory load

Reduced guest satisfactiion and increased shipping cost



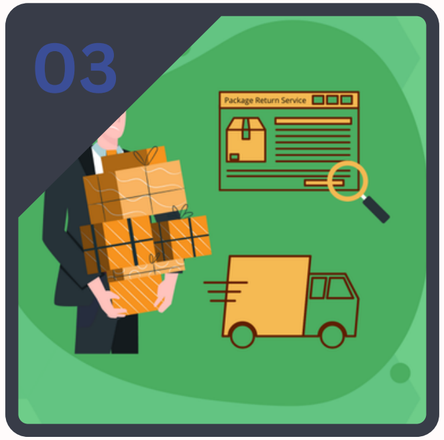
Problem Statement

Target seeks to enhance its supply chain efficiency by strategically expanding its distribution center network. To achieve this, they need to analyze guest order data to determine the optimal number and locations of distribution centers for efficient order fulfillment and cost reduction.

FOCUS AREAS

- High guest concentration
- Frequency of the orders placed

- Average oder frequency
- Product preferences



- Returns rate
- Minimum shipping Distance

- Timely delivery of goods

WORK FLOW


Cognizance
EDA
Clustering
Analysis
Obtain DC's locations
Infer

ATTRIBUTES

Guest Profile

- Guest ID
- Geographical location
- Guest segmentation

Order and Product Preferences

- Order frequency
- Average order value
- Books
- Clothing
- Beauty
- Home appliances
- Electronics

Behavioural Response

- Shipping preference
- Returns rate
- Promotional response

STATES EXTRACTION

	guest_id	latitude	longitude	state
0	G50029	41.934523	-114.793454	Nevada
1	G39532	28.411406	-98.053030	Texas
2	G93168	37.861983	-79.324797	Virginia
3	G34201	33.101184	-115.570116	California
4	G69749	49.022389	-123.155575	British Columbia
...	...	...	...	...
49995	G94802	44.079450	-90.836179	Wisconsin
49996	G85276	49.211234	-93.865861	Ontario
49997	G81954	38.680922	-88.385192	Illinois
49998	G70741	47.677203	-106.121727	Montana
49999	G88038	28.962244	-114.066342	Baja California

50000 rows × 5 columns

	State	Count
0	Texas	3187
1	Montana	2186
2	California	2167
3	Nevada	1507
4	New Mexico	1489
5	Arizona	1441
6	Oregon	1355
7	Wyoming	1346
8	Michigan	1345

48 rows x 2 columns

EDA

NULLS COUNT

```
guest_id                0
geographic_loc          0
order_frequency         0
avg_order_value         0
Electronics             0
Clothing                0
Home Appliances         0
Books                   0
Beauty                  0
returns_rate            0
guest_segmentation      0
shipping_preferences    0
promotional_response    0
state                   0
country                 0
latitude                0
longitude               0
dtype: int64
```

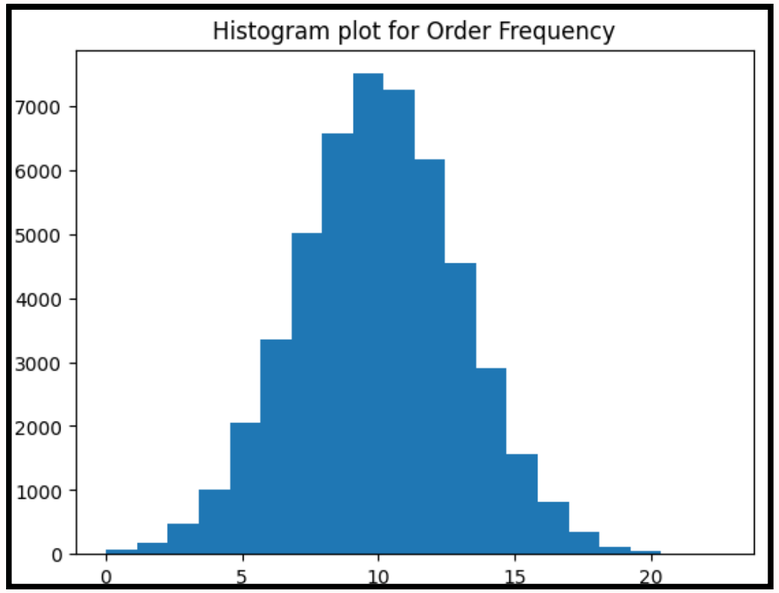
OUTLIERS IN THE DATA

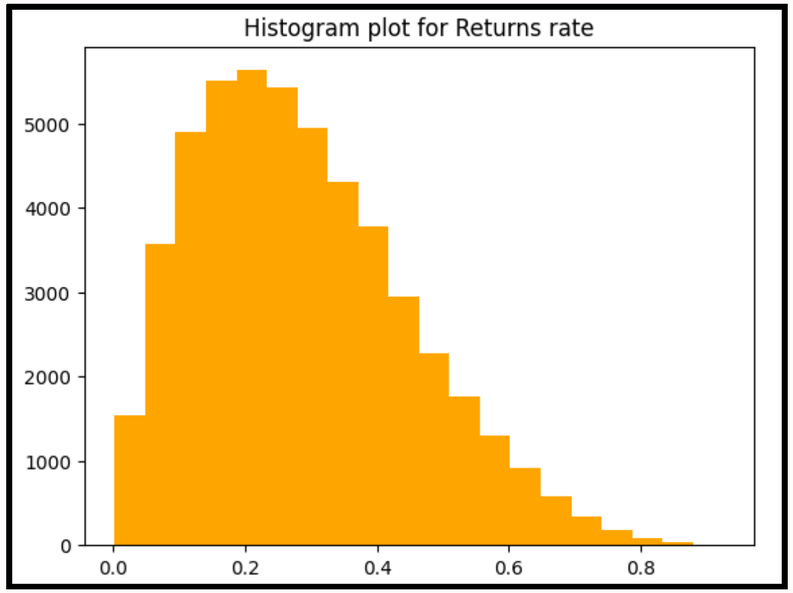
	order_frequency	returns_rate	avg_order_value
55	14.916133	0.097920	266.897583
208	20.816772	0.455954	8.922011
216	10.008862	0.251418	207.617078
404	12.219164	0.289620	190.775848
419	13.290622	0.811543	7.112426
...	...	...	...
49671	13.501575	0.229033	178.257545
49766	19.004620	0.346332	14.114750
49834	14.316574	0.191154	194.738049
49848	12.796308	0.214384	517.967595
49952	12.878852	0.850635	32.406200

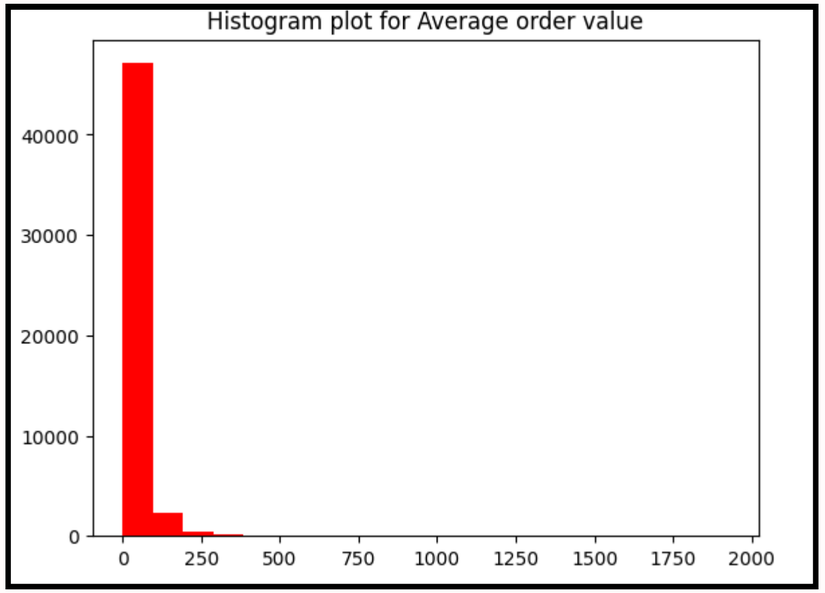
1172 rows × 3 columns

DATA DESCRIPTION & VISUALISATION

NUMERICAL ATTRIBUTES

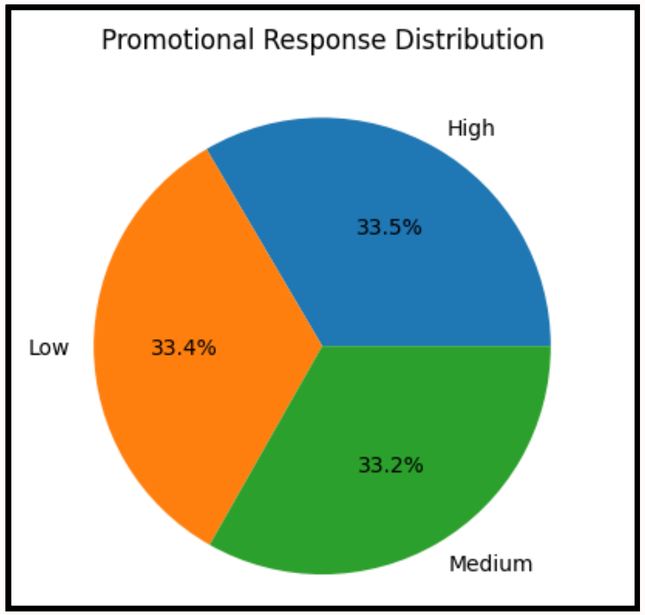


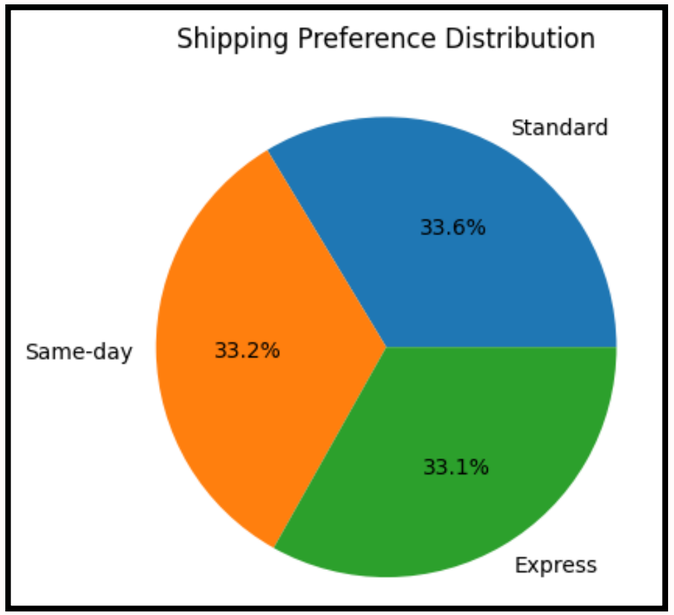


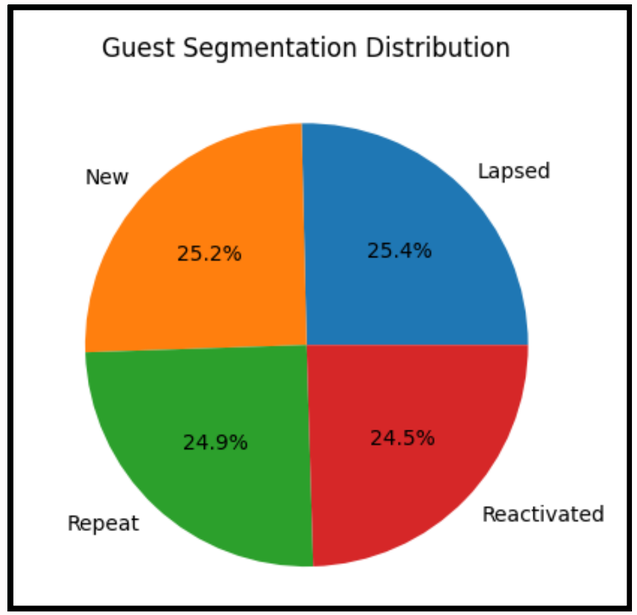


DATA DESCRIPTION &VISUALISATION

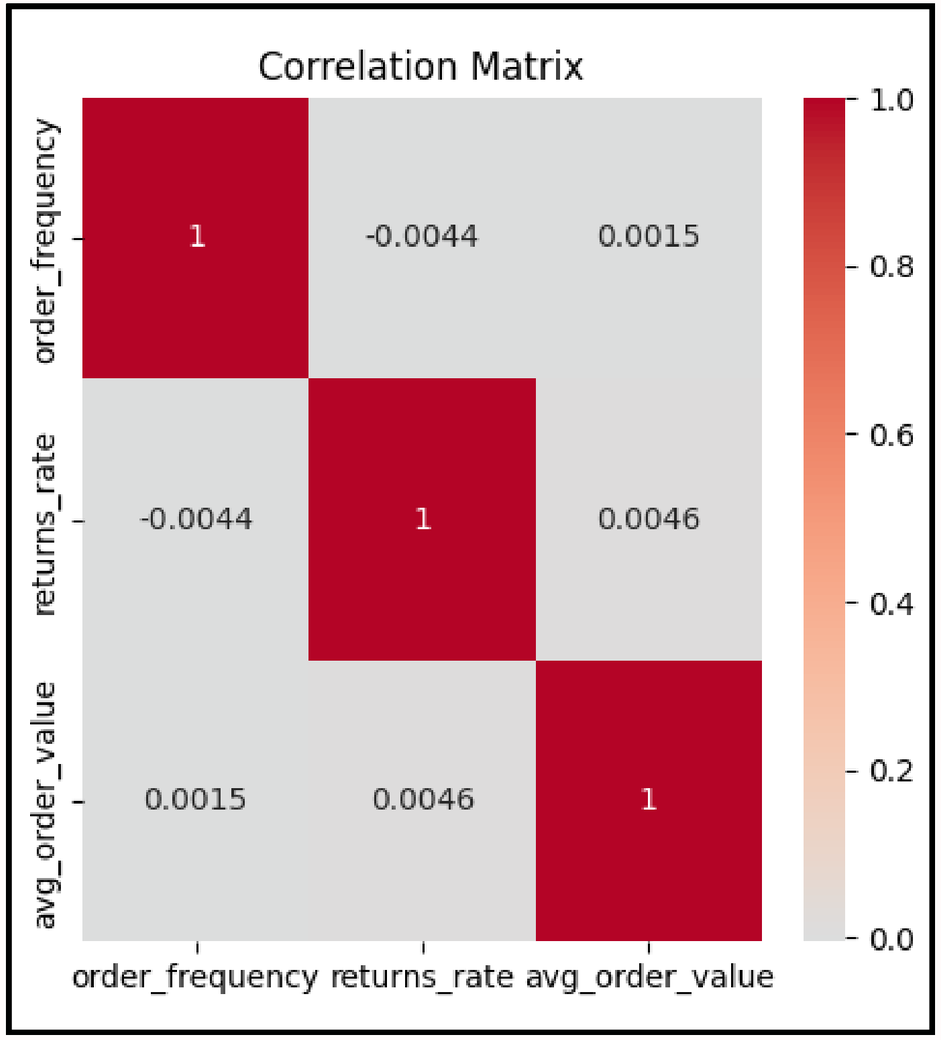
CATEGORICAL ATTRIBUTES

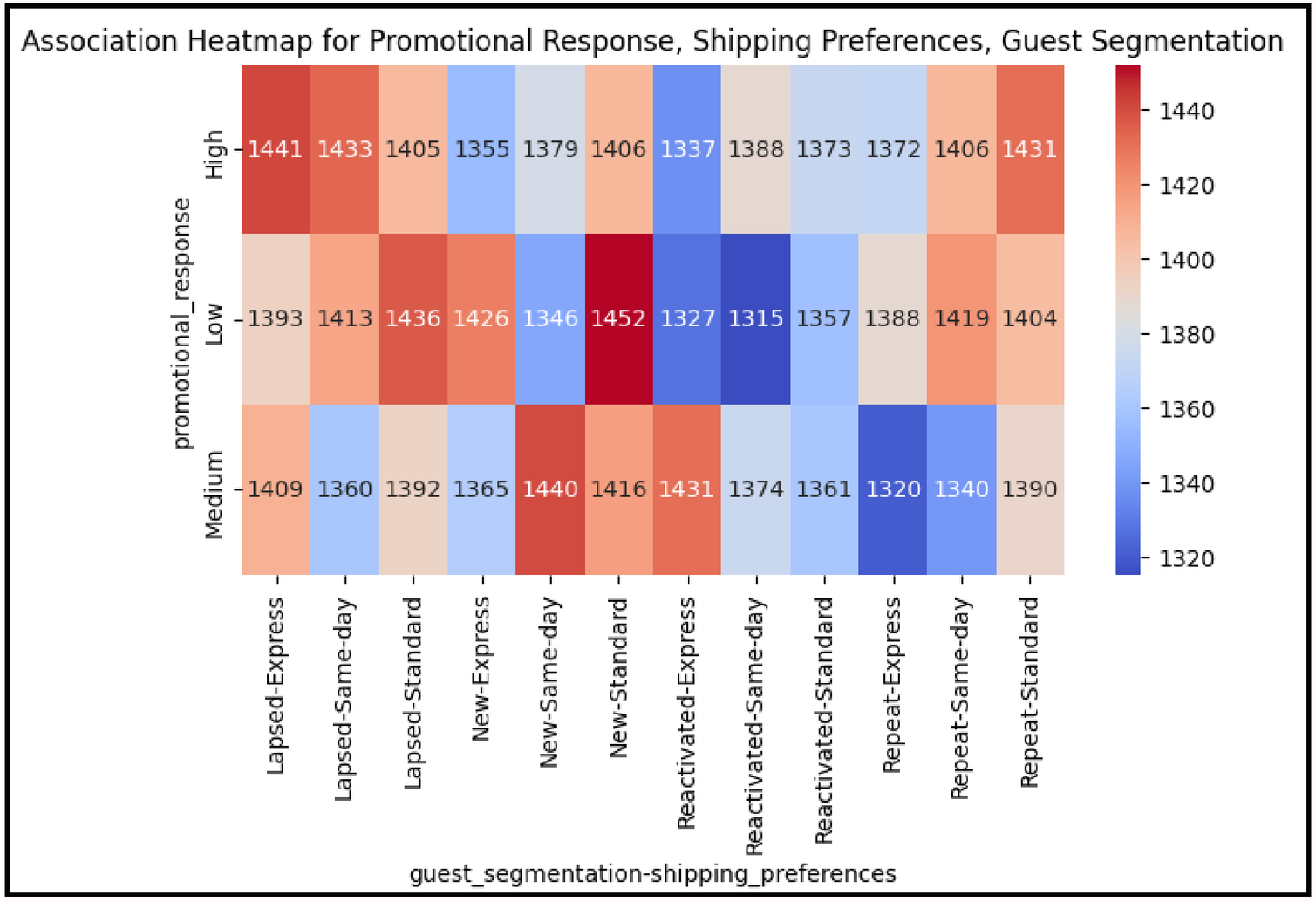






CORRELATION AMONG THE ATTRIBUTES





CLUSTERING

Clustering can help in finding distribution center locations by providing insights into customer behavior and geographic concentrations

Various clustering approaches used:

DB SCAN:

- Density based spacial clustering algorithm
- Identifies regions of high customer concentration based on their geographic locations.

KMeans clustering:

Selected features

Latitude	Longitude
Order Frequency	Returns rate

- Segmenting customers based on these quantitative metrics.
- Identify regions with similar purchasing patterns.

KPrototype Clustering

K-Prototype Clustering is effective for this dataset because it contains both numerical and categorical features

The features selected for clustering:

Numerical

Order Frequency

Returns rate

Categorical

Guest segmentation

Shipping Preferences

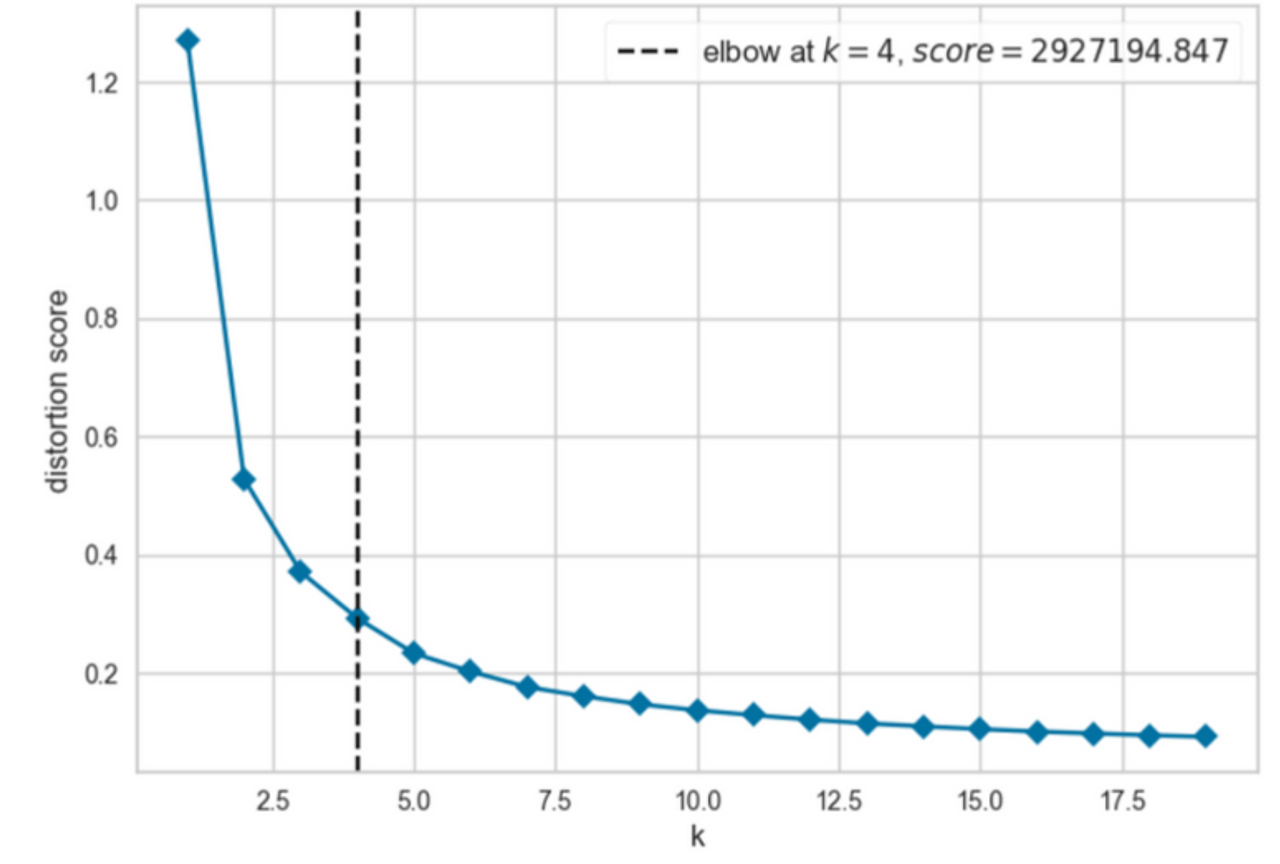
Promotional Response

Geographical

Latitude

Longitude

ELBOW CURVE

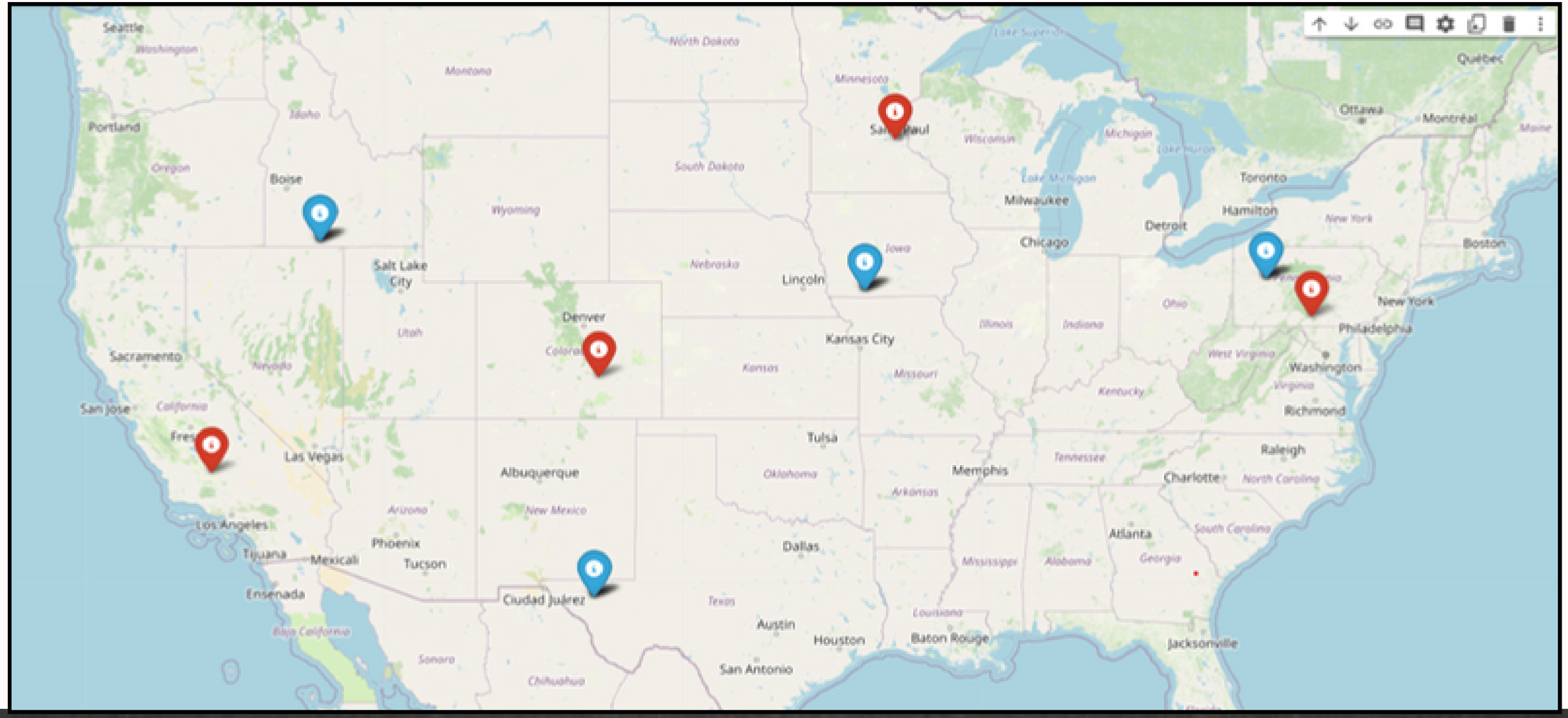


From the elbow curve the optimal number of clusters(k) is 4.

DCs' VISUALIZATION

Existing Distribution centers

New proposed Distribution centers

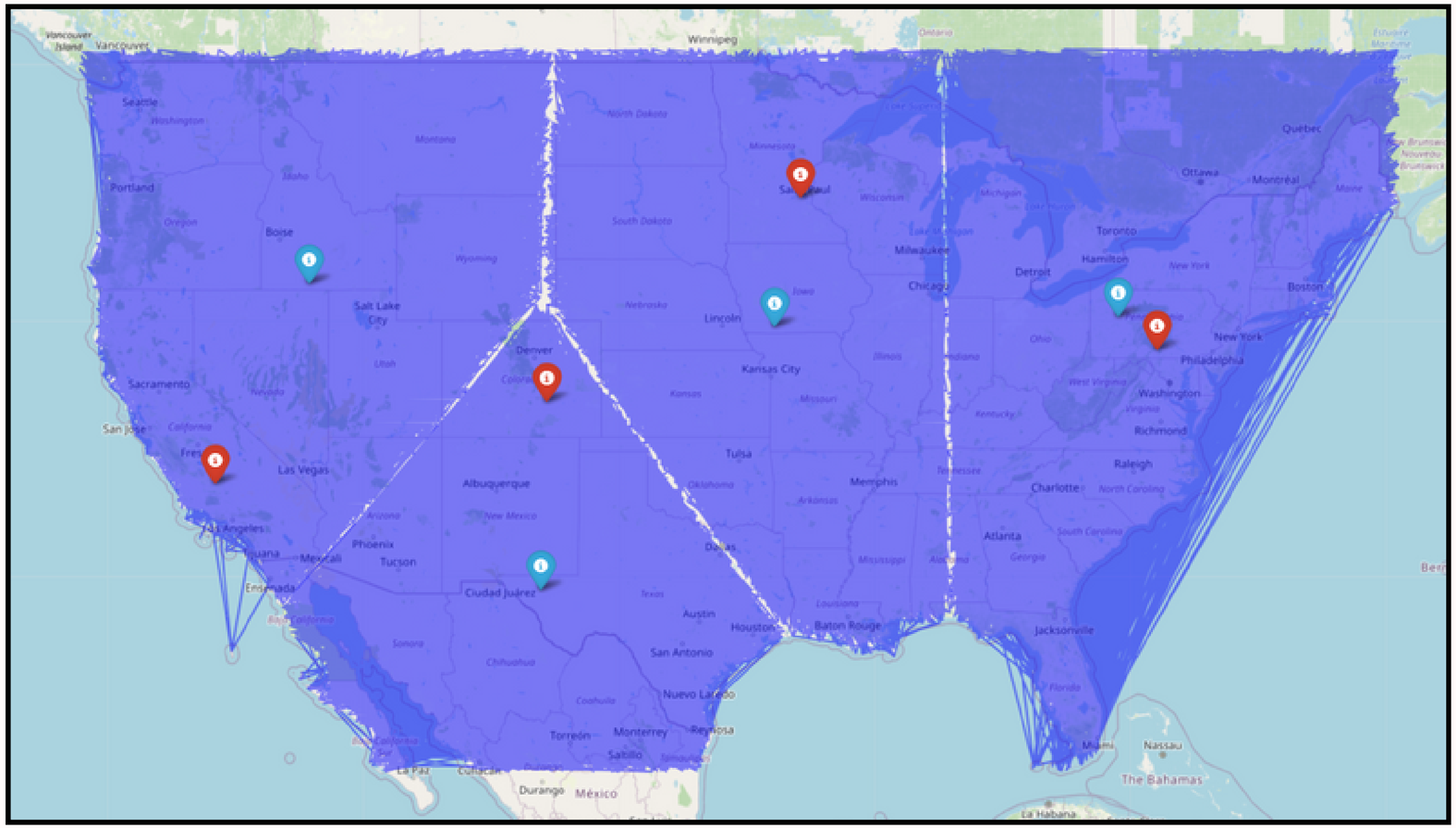


DC Boundaries

Pennsylvania - 12997 guests
Iowa - 13304 guests
Idaho - 12894 guests
Texas - 10506 guests

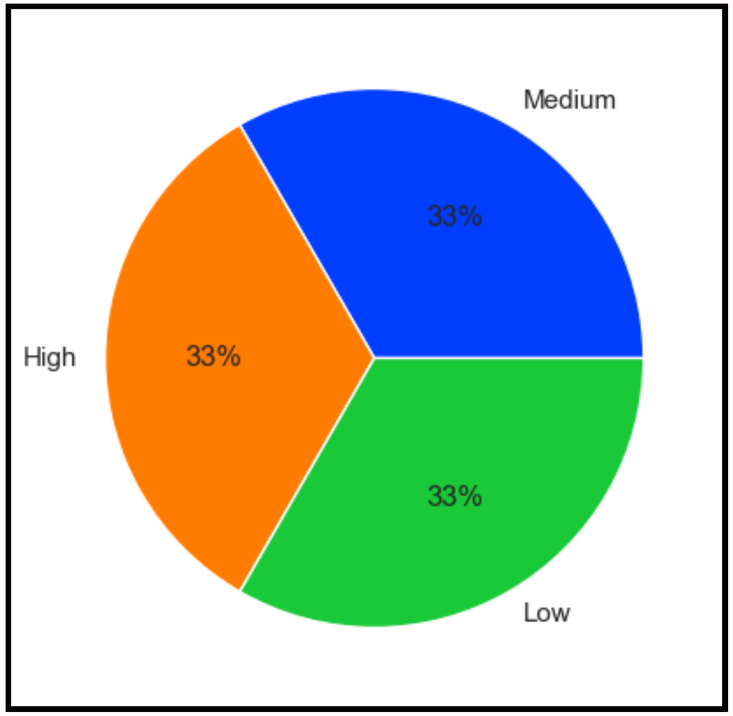
Determining DC locations

- Uniform distribution of Data
- Guest Concentration

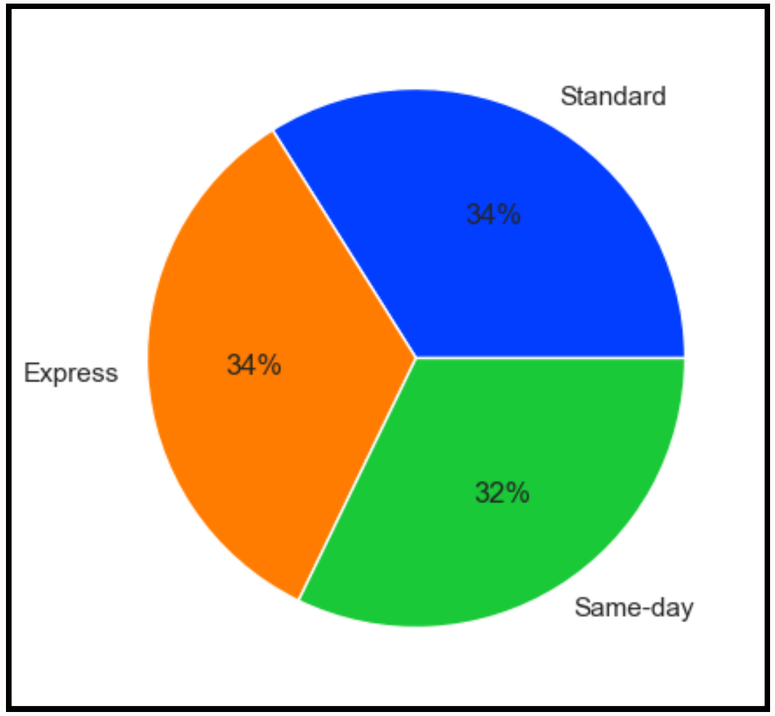


DC1 - Pennsylvania

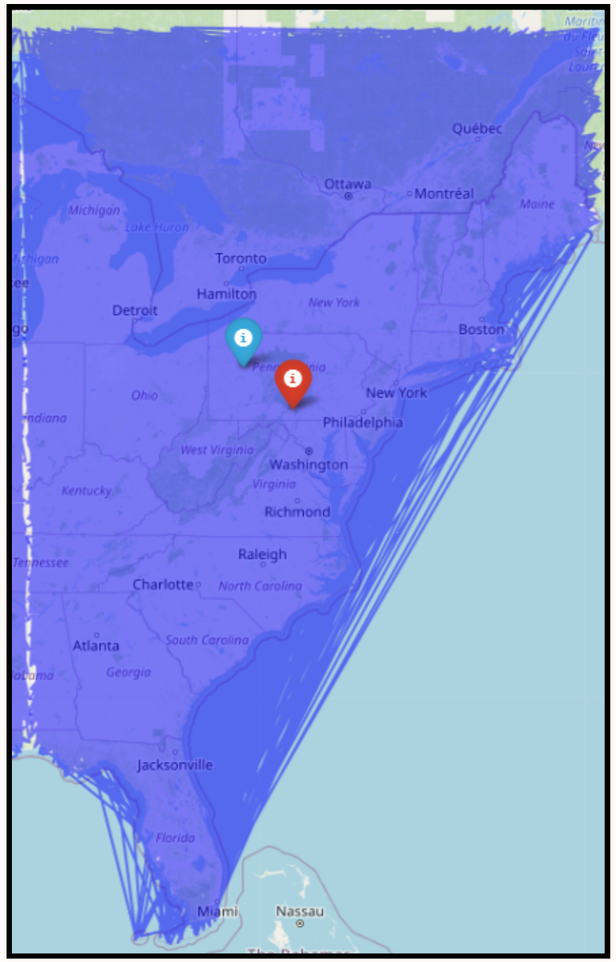
- High number of guests of nearly 5000 from the North Eastern part of United States of America.
- To distribute the load of 12997 members.



Promotional Response



Shipping Preference



Highly populated states

New York - 2744

Ohio - 2712

Michigan - 993

Florida - 844

Distance from already present DC

352.32 km

438.49 km

748.43 km

1354.16 km

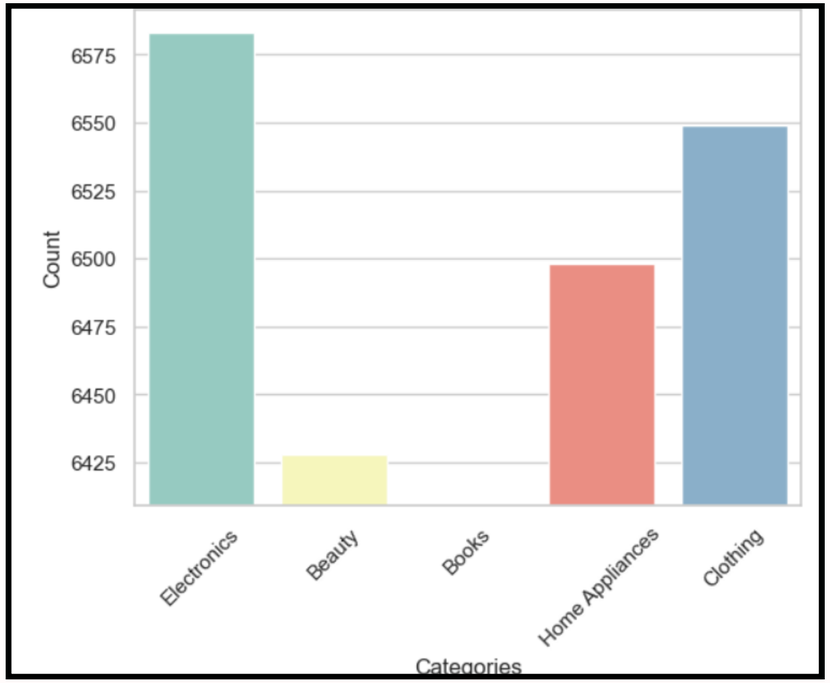
Distance from newly obtained DC

356.55 km

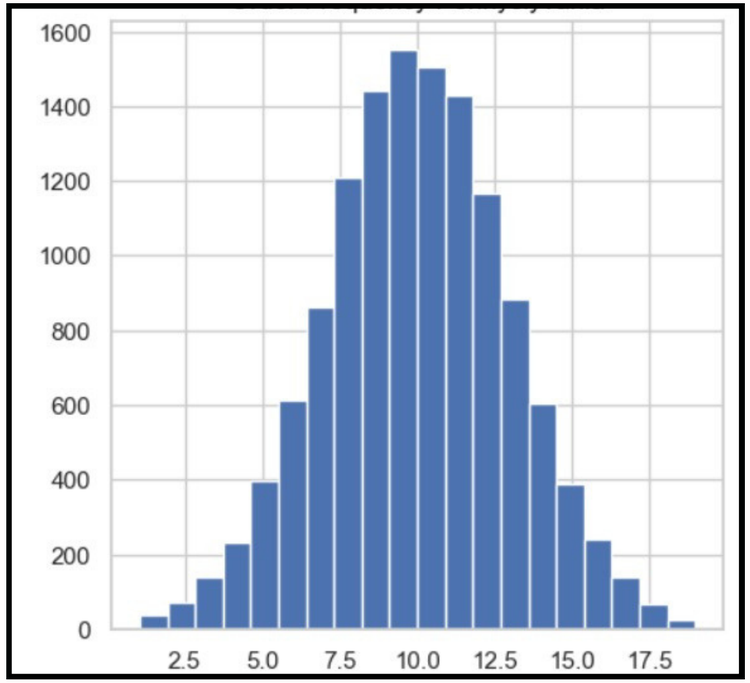
296.16 km

557.70 km

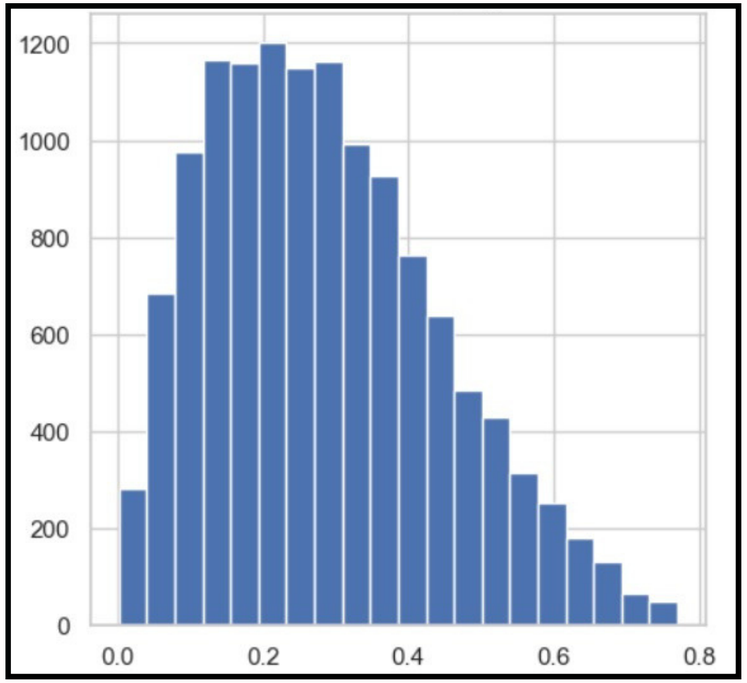
1430.28 km



Item preferences



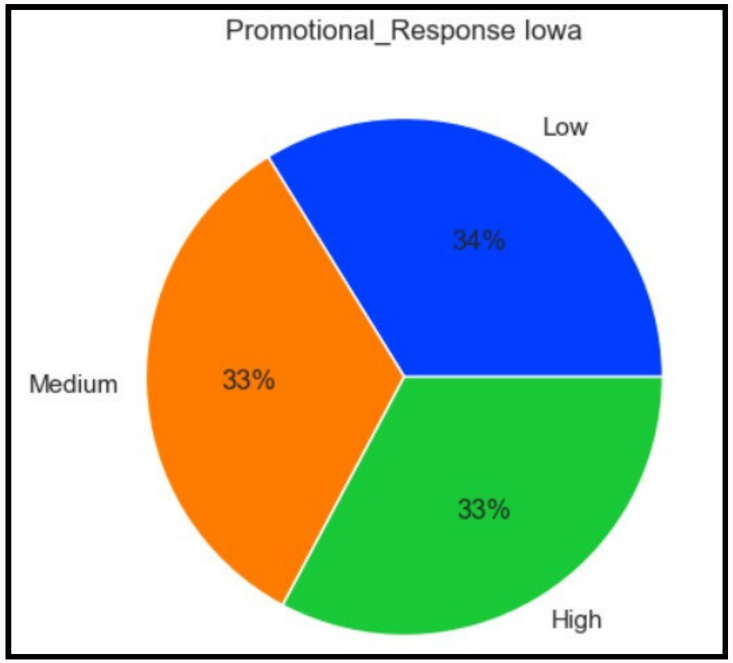
Order Frequency



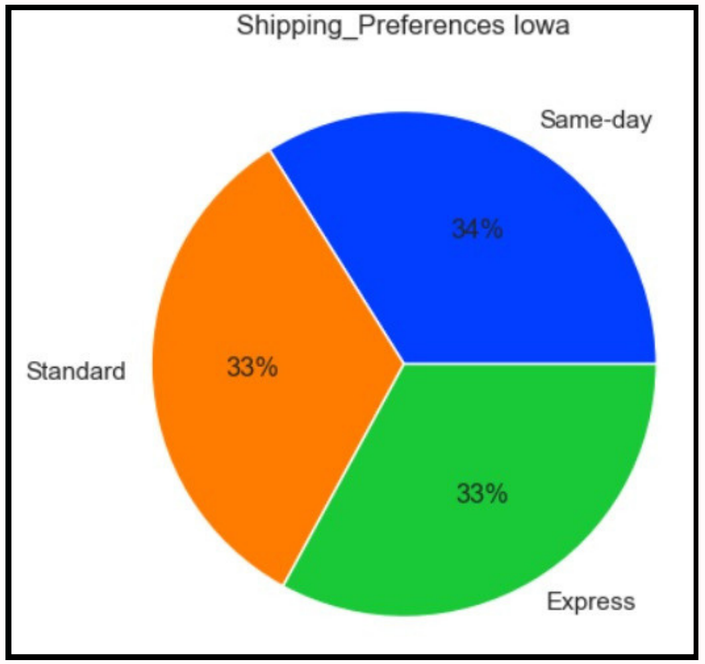
Returns Rate

DC2 - Iowa

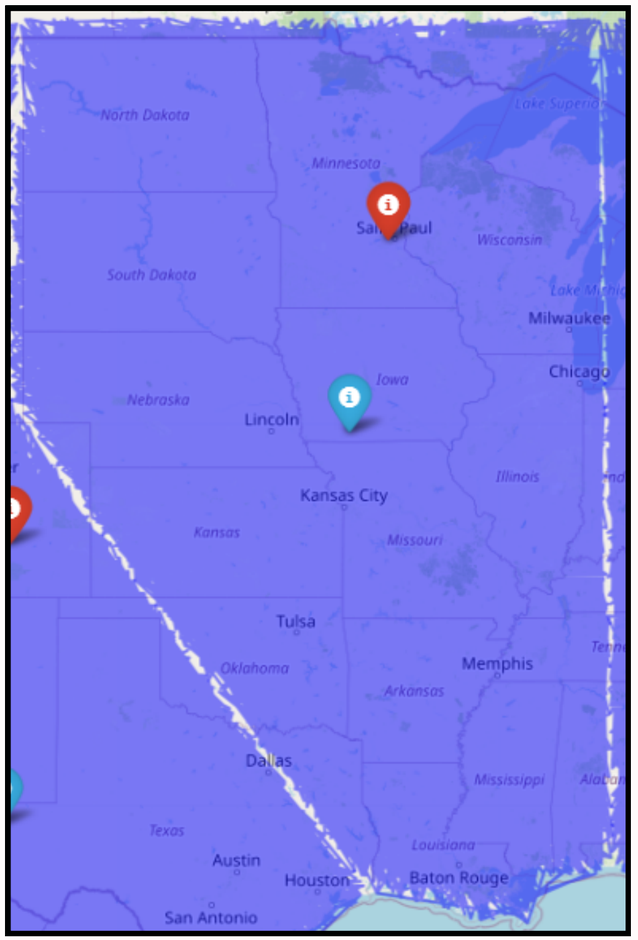
- Iowa helps to serve the guests in central and southern United states of America



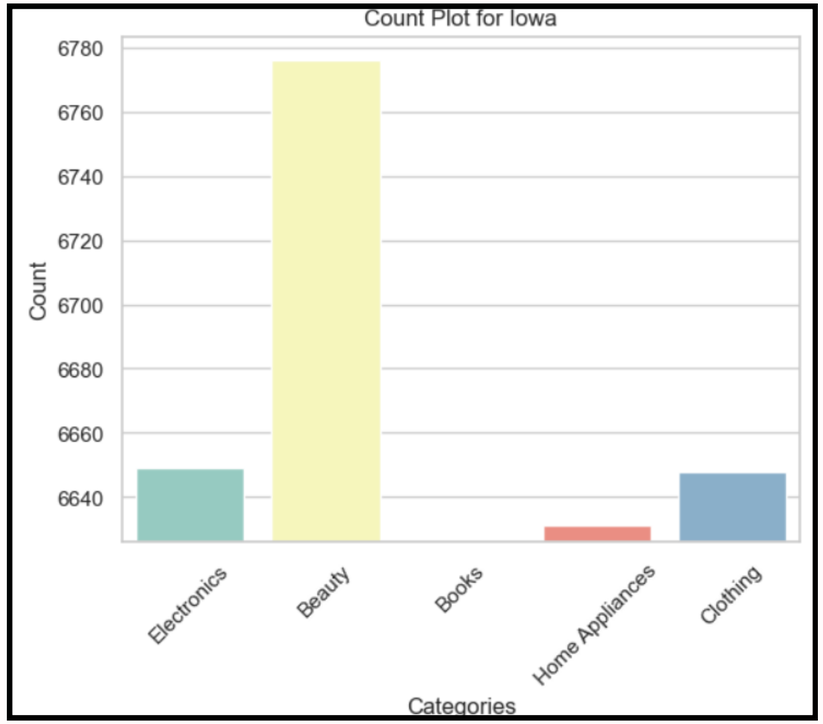
Promotional Response



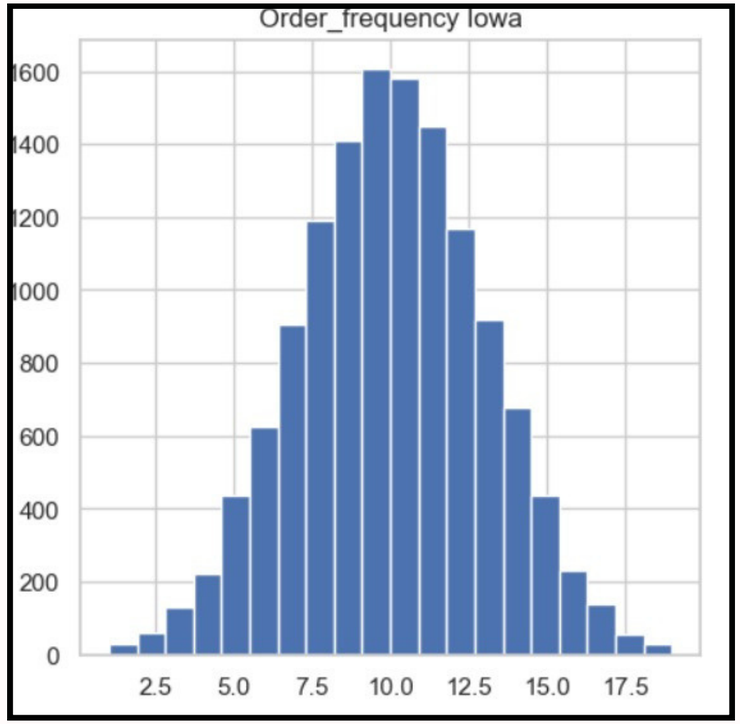
Shipping Preference



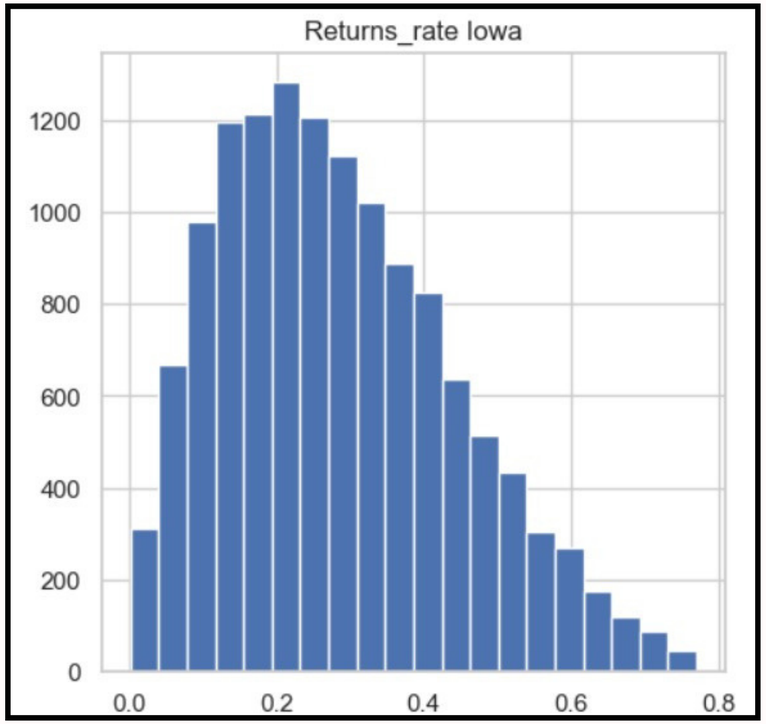
Highly populated states	Distance from already present DC	Distance from newly obtained DC
Nebraska - 1094	654.12 km	458.92 km
South Dakota - 1017	557.15 km	628.97 km
North Dakota - 1001	609.99 km	873.67 km
Kansas - 999	826.15 km	409.01 km
Wisconsin - 904	277.46 km	574.17 km



Item preferences



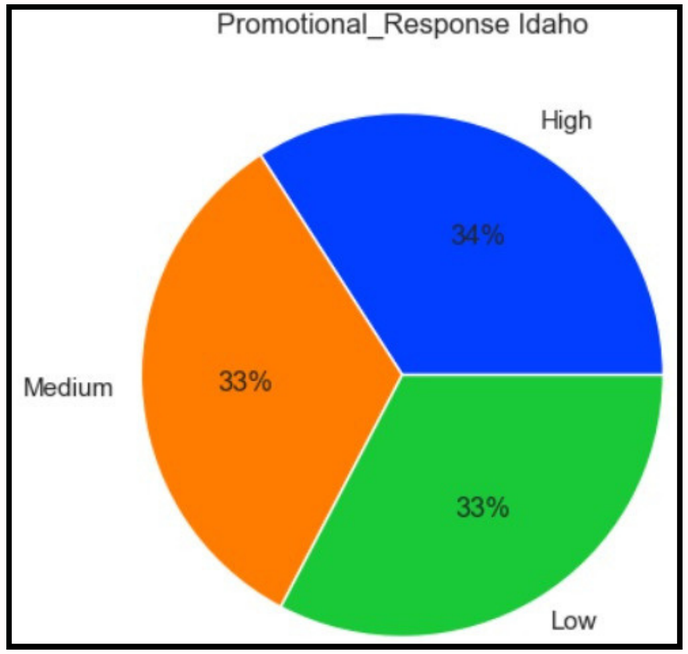
Order Frequency



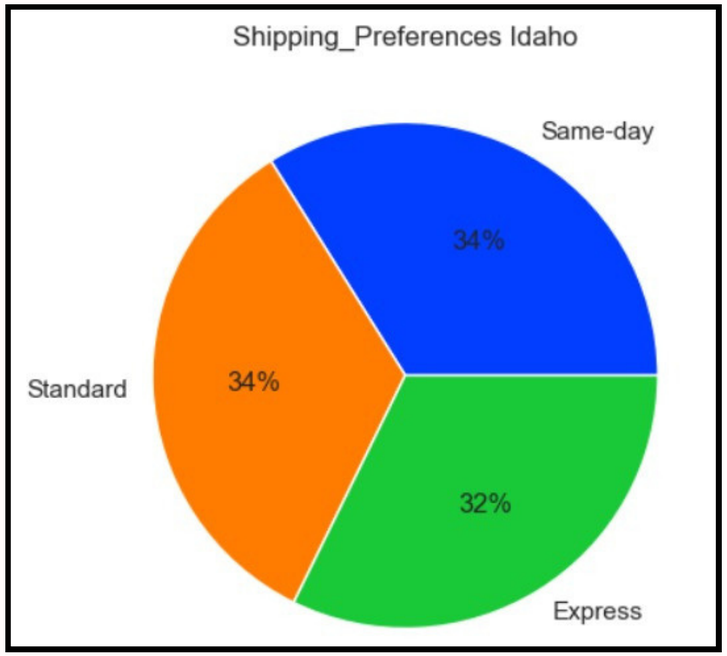
Returns Rate

DC3 - Idaho

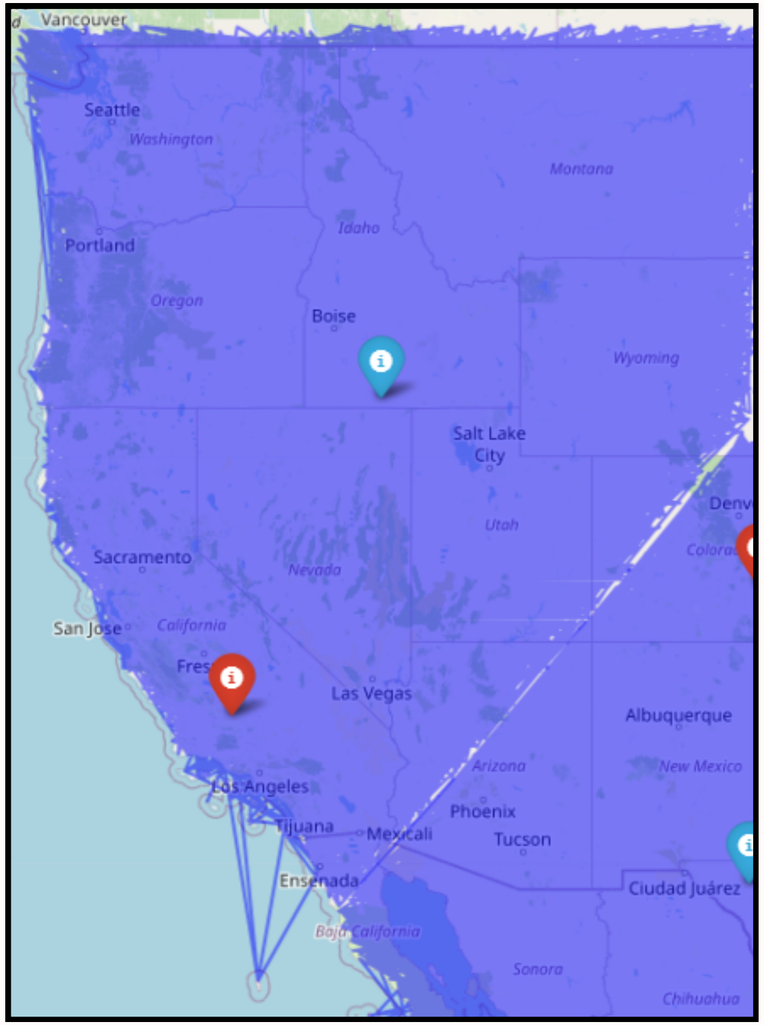
- Idaho covers the north-west part of United States of America
- California far away from the highly populated states



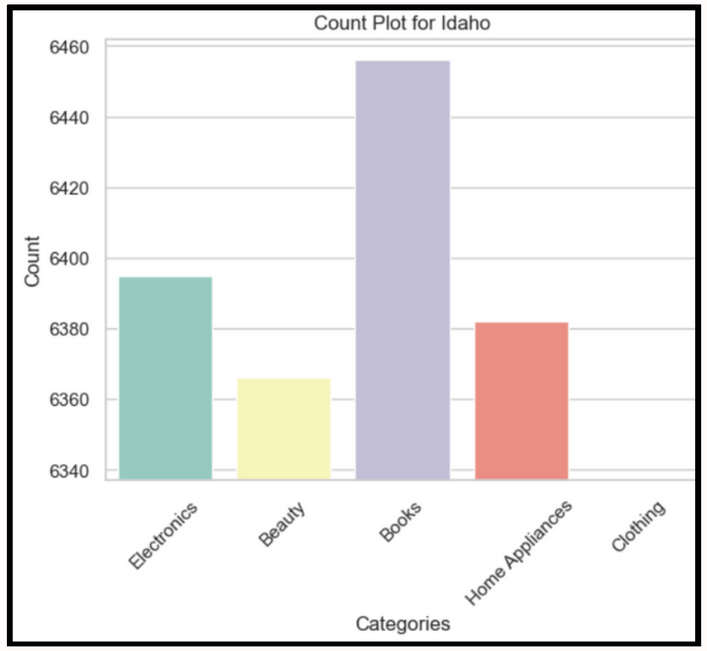
Promotional Response



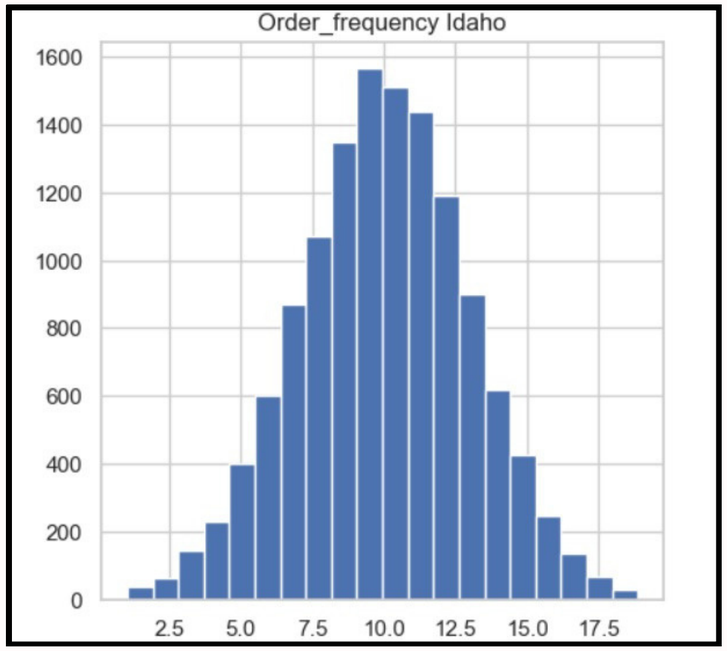
Shipping Preference



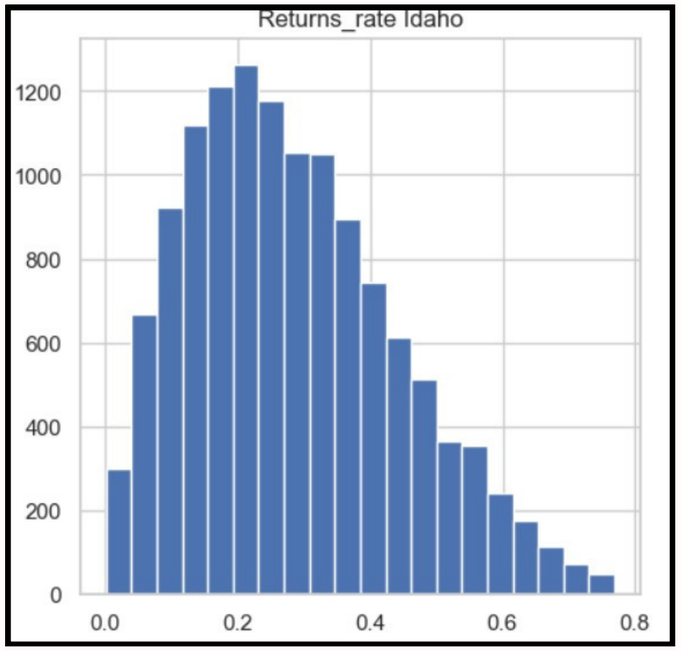
Highly populated states	Distance from already present DC	Distance from newly obtained DC
Montana - 2132	1506.07 km	673.86 km
Nevada - 1507	484.74 km	352.82 km
Oregon - 1355	949.24 km	487.21 km
Wyoming - 1247	1289.50 km	589.82 km
Utah - 1065	787.64 km	413.10 km
Washington - 1010	1341.68 km	729.07 km



Item preferences



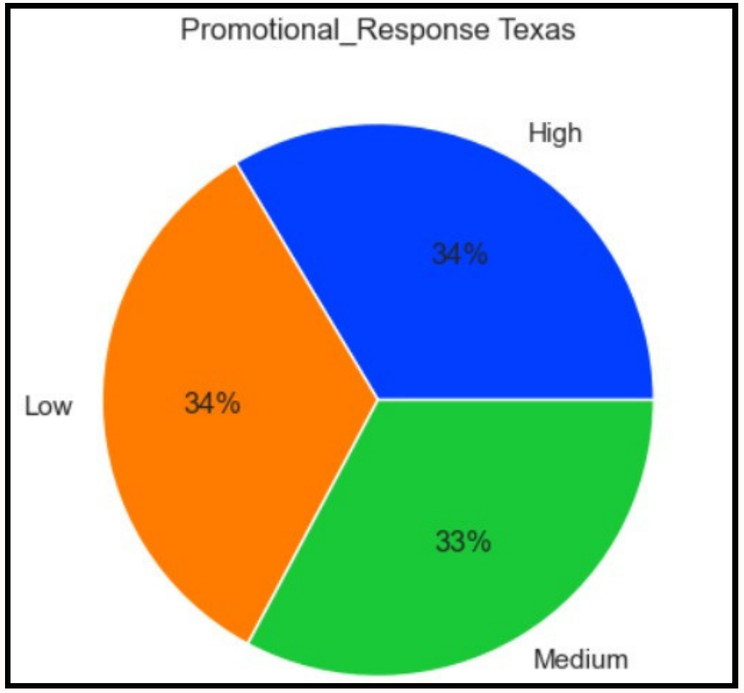
Order Frequency



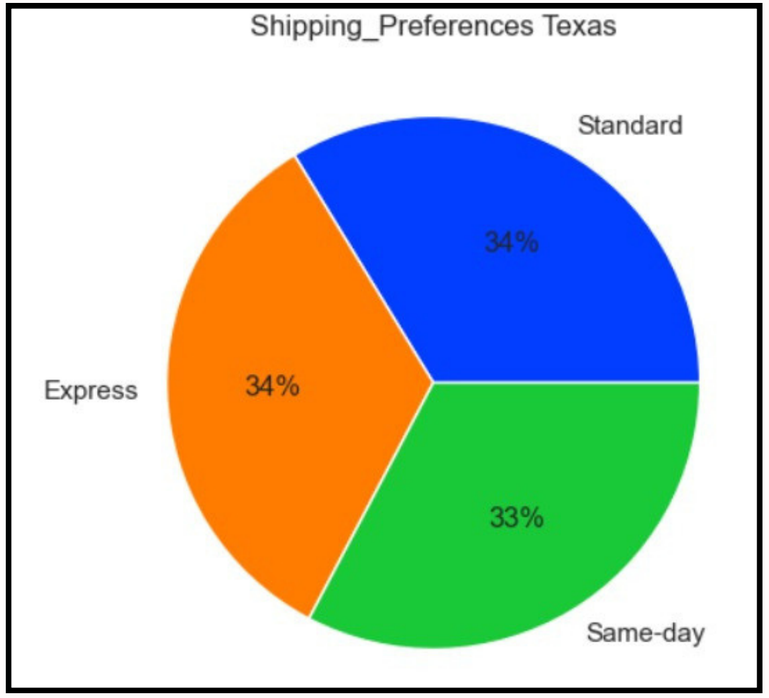
Returns Rate

DC4 - Texas

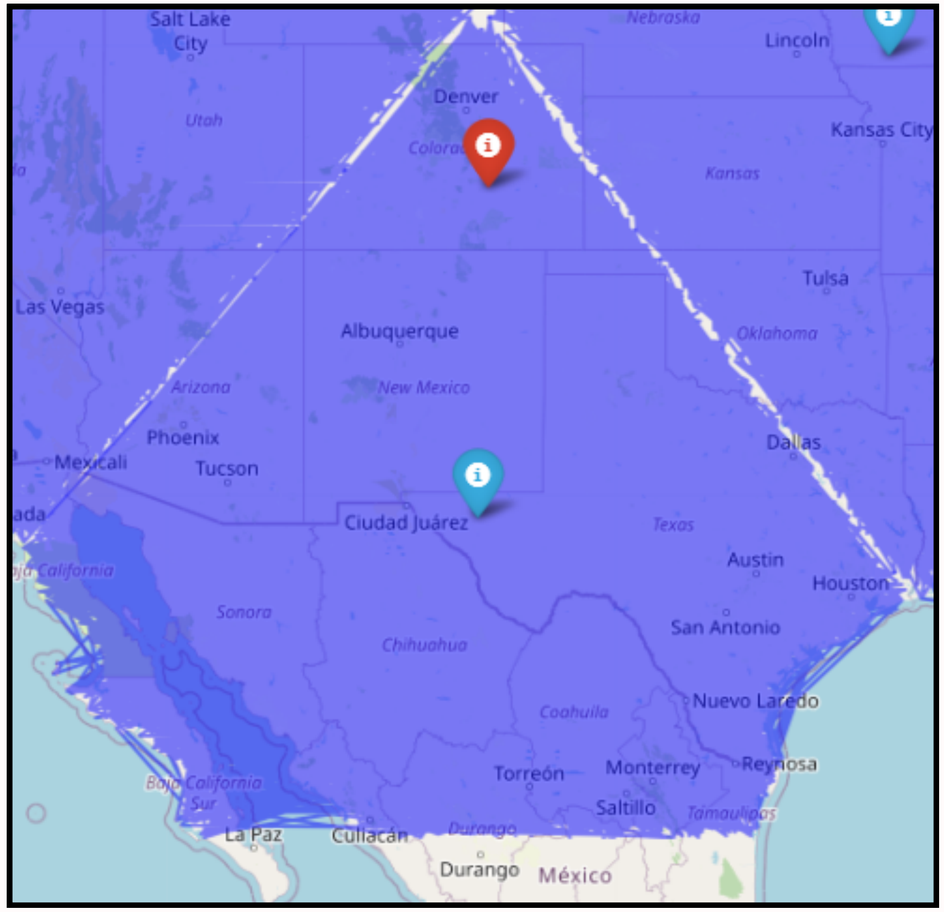
- Texas covers the southern part of United States of America
- Texas alone has a concentration of 2887 guests



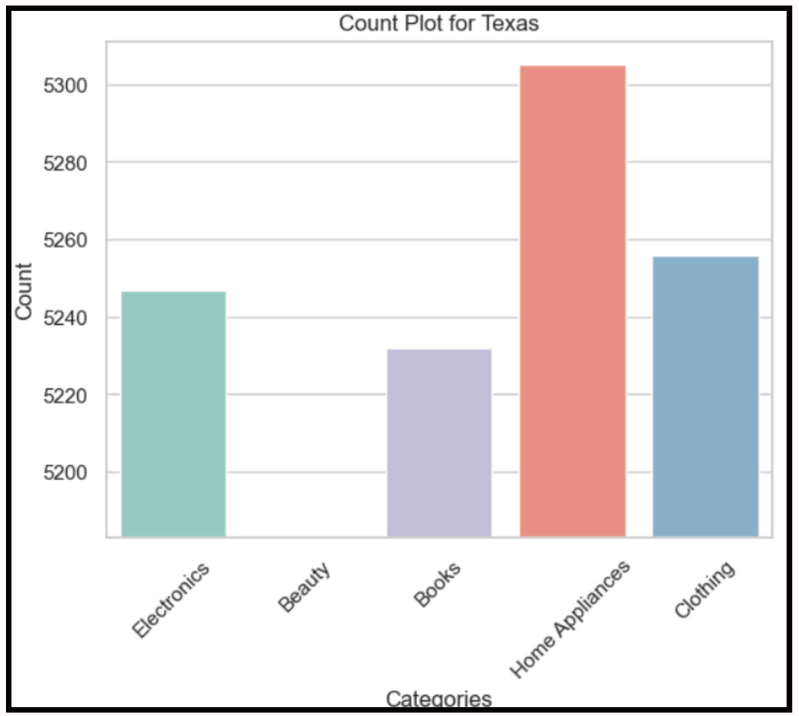
Promotional Response



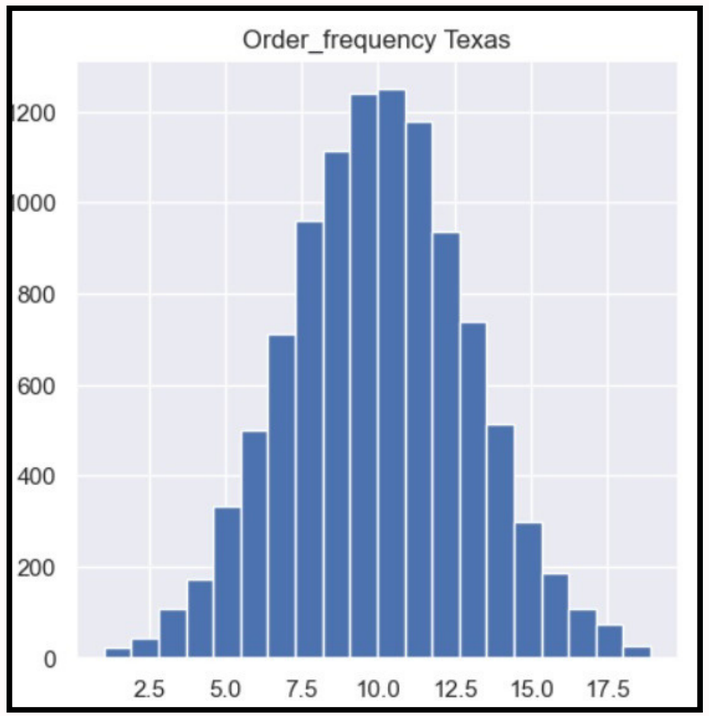
Shipping Preference



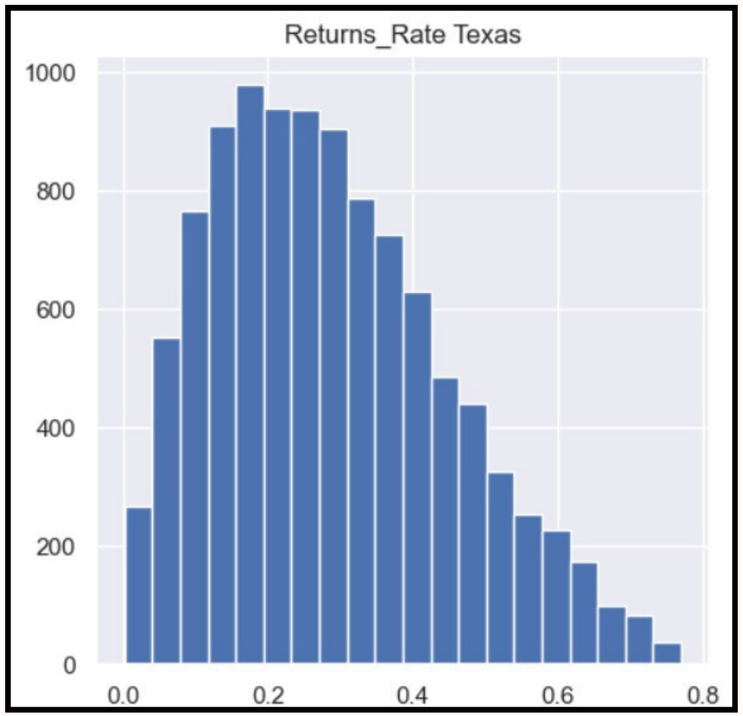
Highly populated states	Distance from already present DC	Distance from newly obtained DC
Texas - 2887	871.56 km	462.51 km
Arizona - 1922	767.14 km	636.99 km
New Mexico - 1489	445.51 km	352.27 km



Item preferences



Order Frequency



Returns Rate

Highlights

- Uniformly distributed data
- Clustering: K-Prototype
- 4 Distribution Centers: Pennsylvania, Iowa, Idaho and Texas
- Pennsylvania: Eastern regions of America
- Iowa: Central America
- Idaho: North-west regions of America
- Texas: Southern regions of America
- Uniform data distribution in all clusters
- Enhanced load distribution, strategic inventory management
- Timely delivery and returns services, improved customer experience

Team

B SITA SWAPNIKA

B ASRITHA

G AISHWARYA

K BHAVITHA

B SREEVIDYA

Mentor

Mr. Aviral Maheswari

"Balancing the science of logistics with the art of guest delight - that's the essence of a successful strategy to locate distribution centers."

THANK YOU!