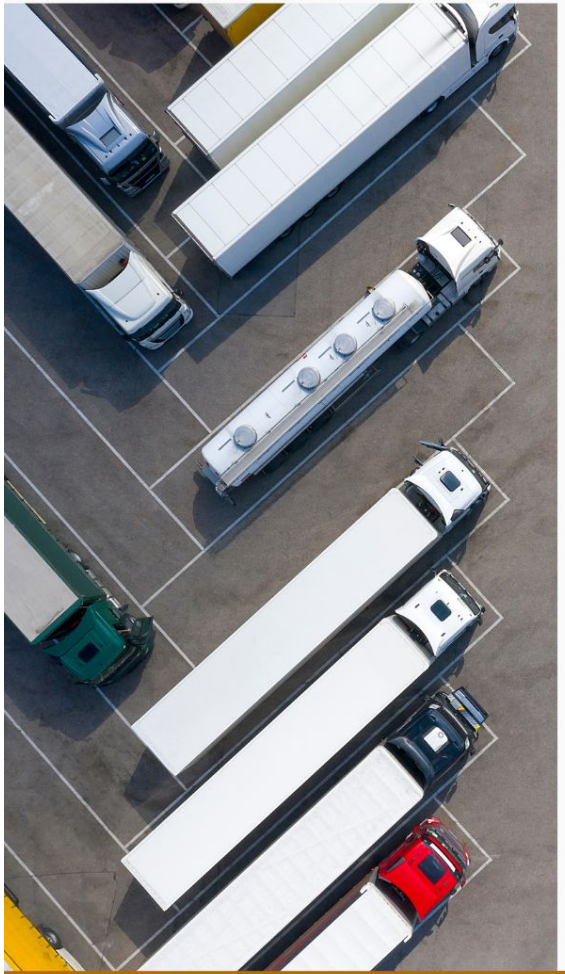
SmartHop

Final Presentation

Team Geography

Members: Tanveer | Anshika | Vishal | Yousif | Saswati | Vaishnav |

University of Maryland 1856 Robert H. Smith School of Business



AGENDA

01 **Executive Summary**

02 **Data Overview**
Snowflake extraction and Relationship Diagram

03 **Data Manipulation**
Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

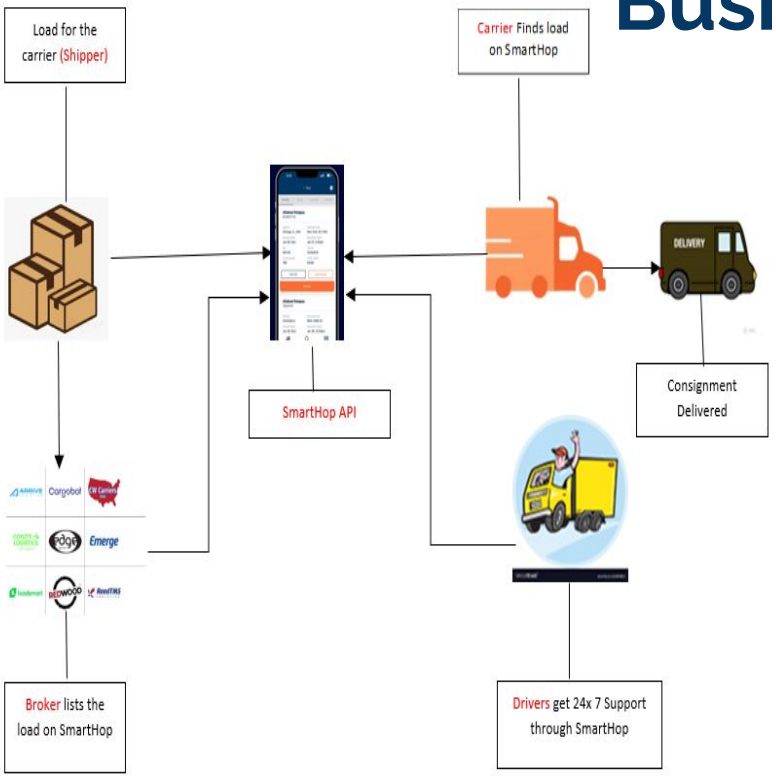
04 **Exploratory Data Analysis**
Key Observations

05 **Findings**

06 **Data Modelling**

07 **Recommendations**

Business Problem



- Unclear impact of carrier's geographical location and risk appetite impacts SmartHop's performance:
  - Market Volatility in terms of volume
  - Varying profitability based on preference in load selection and markets

Objective

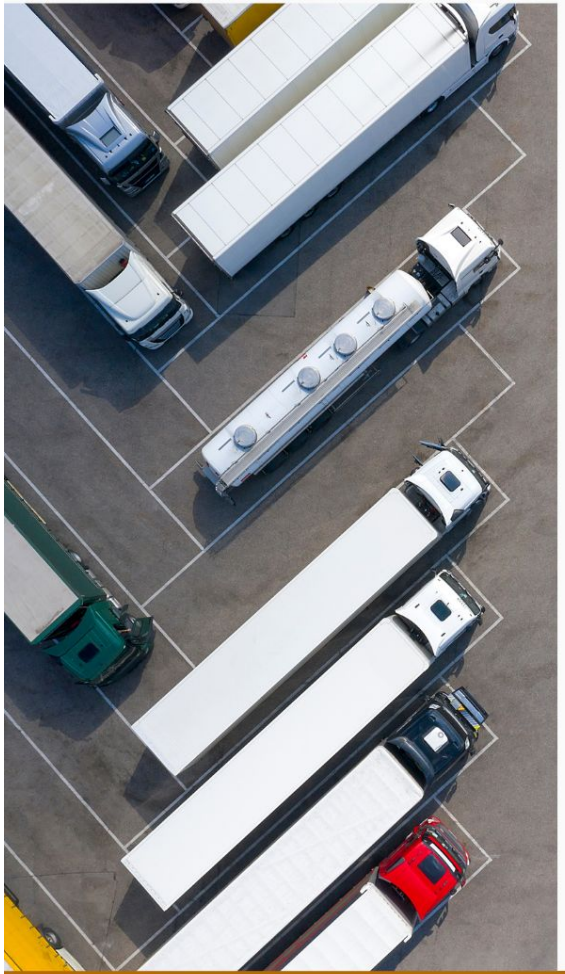
- Recommend strategies - routes, load choice or carrier behavior to improve the performance of carriers based on geographical analysis.

SmartHop

Executive Summary

TOP FINDINGS

ROBERT H. SMITH SCHOOL OF BUSINESS



AGENDA

01 Executive Summary

02 **Data Overview**
Snowflake extraction and Relationship Diagram

03 **Data Manipulation**
Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

04 **Exploratory Data Analysis**
Key Observations

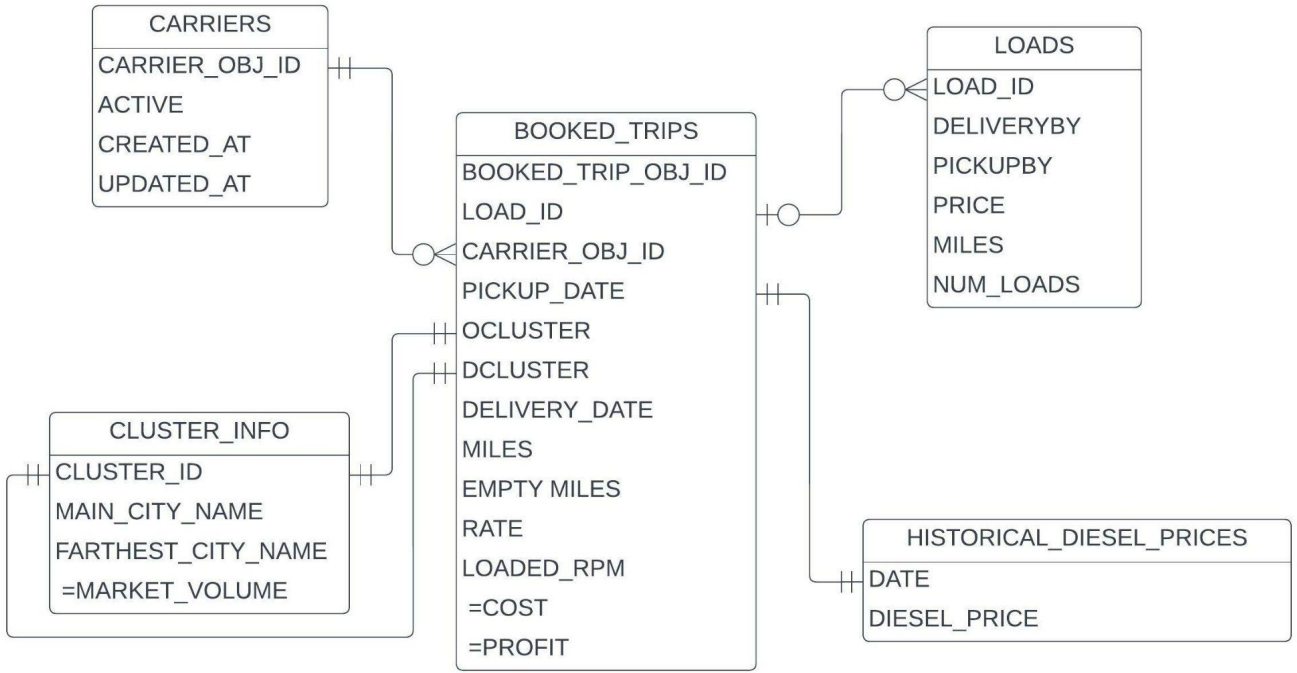
05 **Findings**

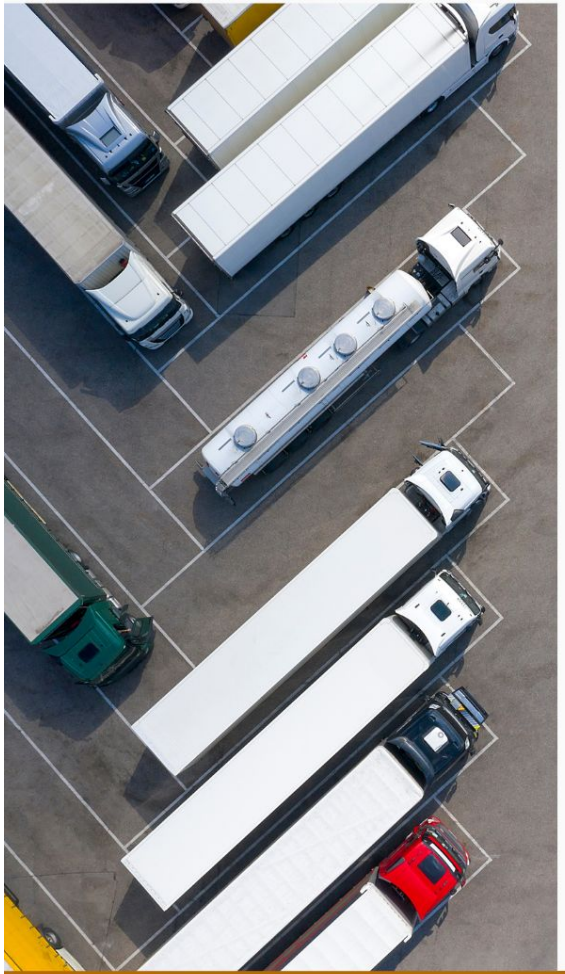
06 **Data Modelling**

07 **Recommendations**

SmartHop

Relationship Diagram





AGENDA

01 Executive Summary

02 Data Overview

Snowflake extraction and Relationship Diagram

03 Data Manipulation

Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

04 Exploratory Data Analysis

Key Observations

05 Findings

06 Data Modelling

07 Recommendations

Data Overview

46,000
Booked Trips

Transactional data of booked trips done over the last few years

- Missing origin and destination clusters
- Missing rates or miles
- Cost or profit is not recorded

18,000
Carriers

Master data of carriers in the SmartHop database

- Do not include carriers headquarter
- Carriers Obj ID is the primary key

70
Clusters

Customer clusters identified by SmartHop with information on main city and coordinates

- Each cluster has a single main city
- No identifier of dead market or high market

Data Extraction (SQL)

45,000,000+
Load
Available load to carriers for each cluster, date and offered rate

- Some load are duplicated
- Too large data to download

18,000,000+
Load
Available load to carriers for each cluster, date and offered rate

- Some load are duplicated
- Too large data to download

Deduplicating

1. Aggregate (Origin, Destination, Pickup Date, Comment, Client)
2. Keep latest load updated

Partial Download

1. Utilize deduplicated data set
2. Download selective dates

Data Extraction (SQL)

700,000+ Load

Available load to carriers for each cluster, date and offered rate

- Some load are duplicated
- Too large data to download

300,000+ Load

Available load to carriers for each cluster, date and offered rate

- Some load are duplicated
- Too large data to download

Summarized load

1. Filter deduplicated data with booked trip date, and origin or destination
2. Aggregate data to available origin, destination, avg miles, avg rate

Final Dataset

- Available routes in each date
- Average rate of the route
- Average miles of the route
- Number of loads available in route

Data Cleaning

26,700

Booked Trips

Transactional data of booked trips done over the last few years

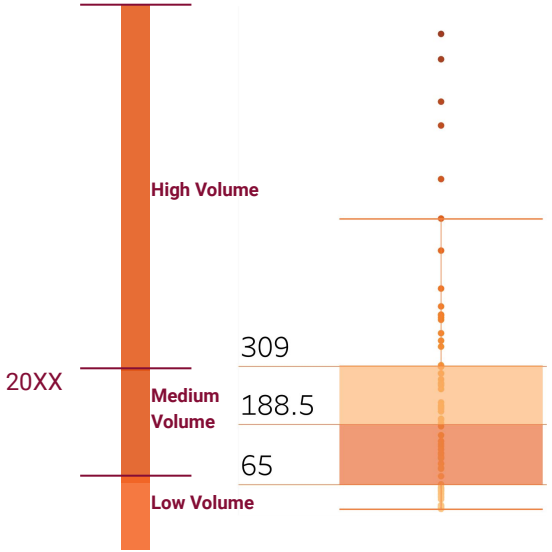
- Cleansed from blank records in key features
- Cleaned from key outliers

- Removed blank records in
  - **DCluster**
  - **OCluster**
  - **Rate**
  - **Miles**
  - **Carrier Object ID**
- Removed outliers in rate and miles.
- Excluded cancelled status trips.
- Removed records with “Test” in brokers.

Feature Engineering

CREATING NEW VARIABLES:

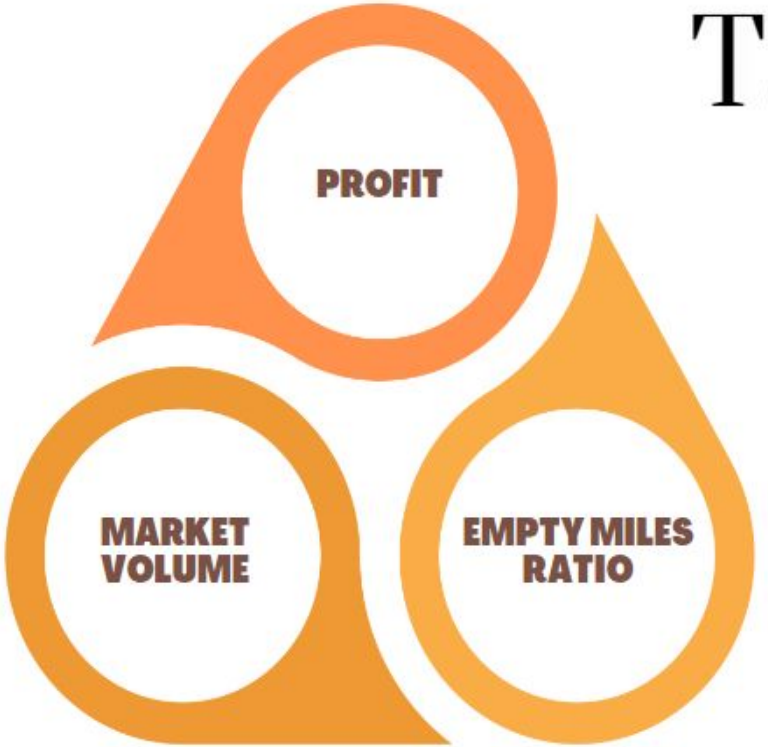
- **Cost**
  - Referenced the provided cost per mile by SmartHop
  - **Historical diesel gallon price to cost per mile**
- **Profit**
  - Calculated by subtracting cost from RPM_RAW*Total Miles
- **Cluster classification**
  - Each cluster was classified into High, Medium or Low volume cluster based on the number of trips that were made to the cluster



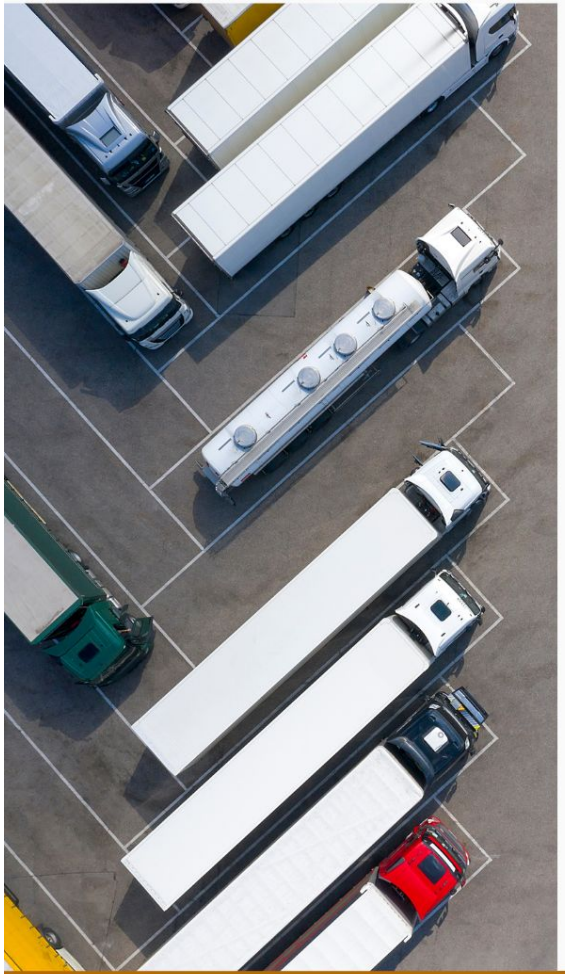
Box plot of Market Volume

SmartHop

Target Variable



- PROFIT OF THE TRIP MADE
- VOLUME OF TRIPS OF THE CLUSTER
- RATIO OF EMPTY MILES TO TOTAL MILES.



AGENDA

01 Executive Summary

02 Data Overview

Overview and Relationship Diagram

03 Data Manipulation

Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

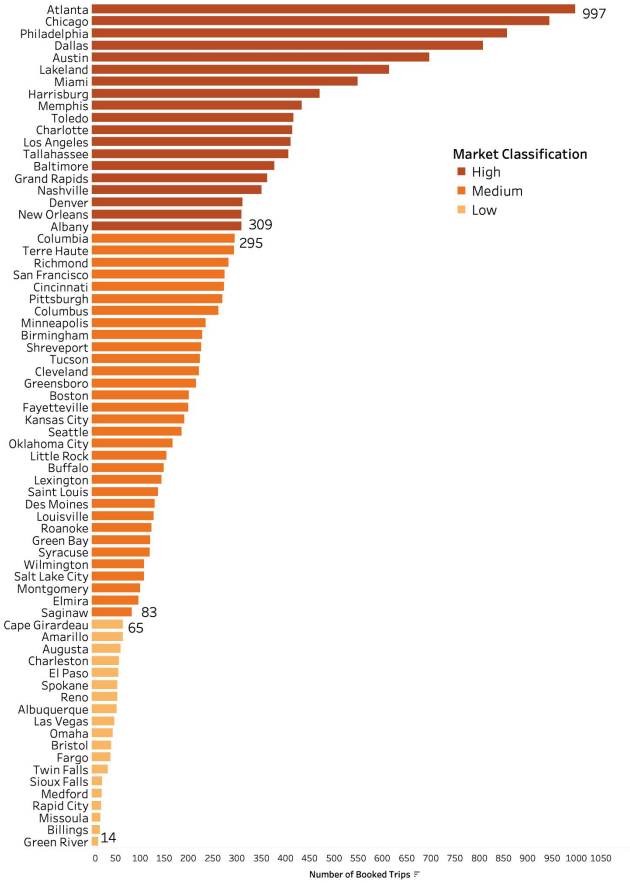
04 Exploratory Data Analysis

Key Observations

05 Findings

06 Data Modelling

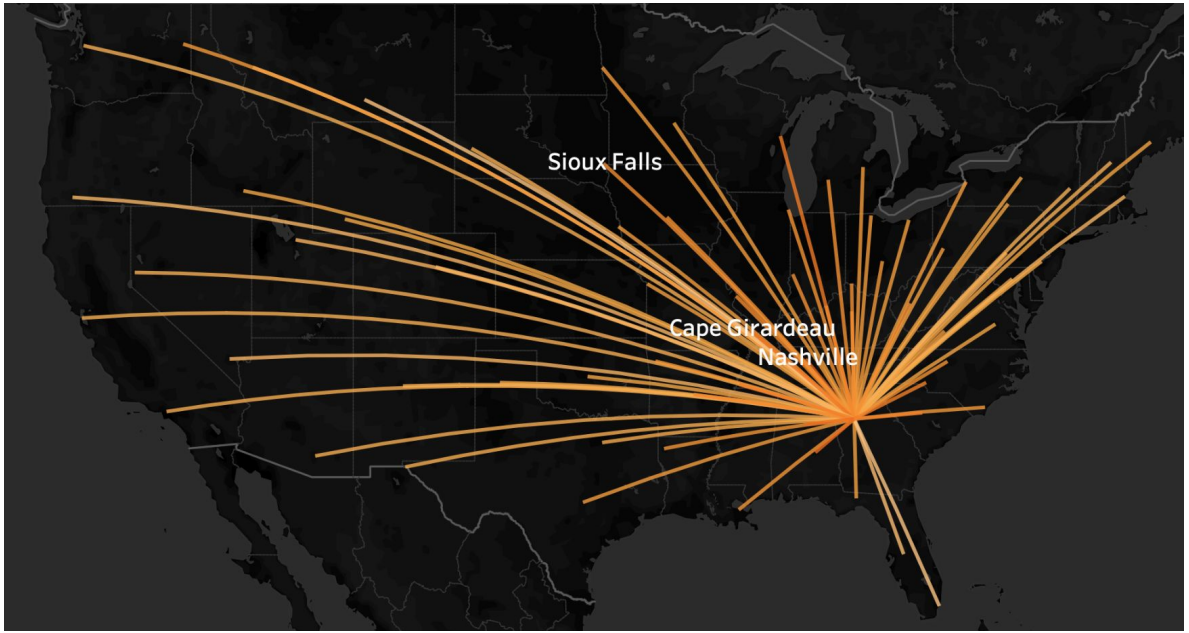
07 Recommendations



Market Volume Classification per Destination City

- Atlanta is leading the High Volume Destination Market with a total of 997 booked trips.
- Columbia is Medium Volume Destination Market's leader with a total of 295 booked trips.
- Green River has the lowest volume in terms of loads getting delivered to the city with a total of only 14 booked trips.

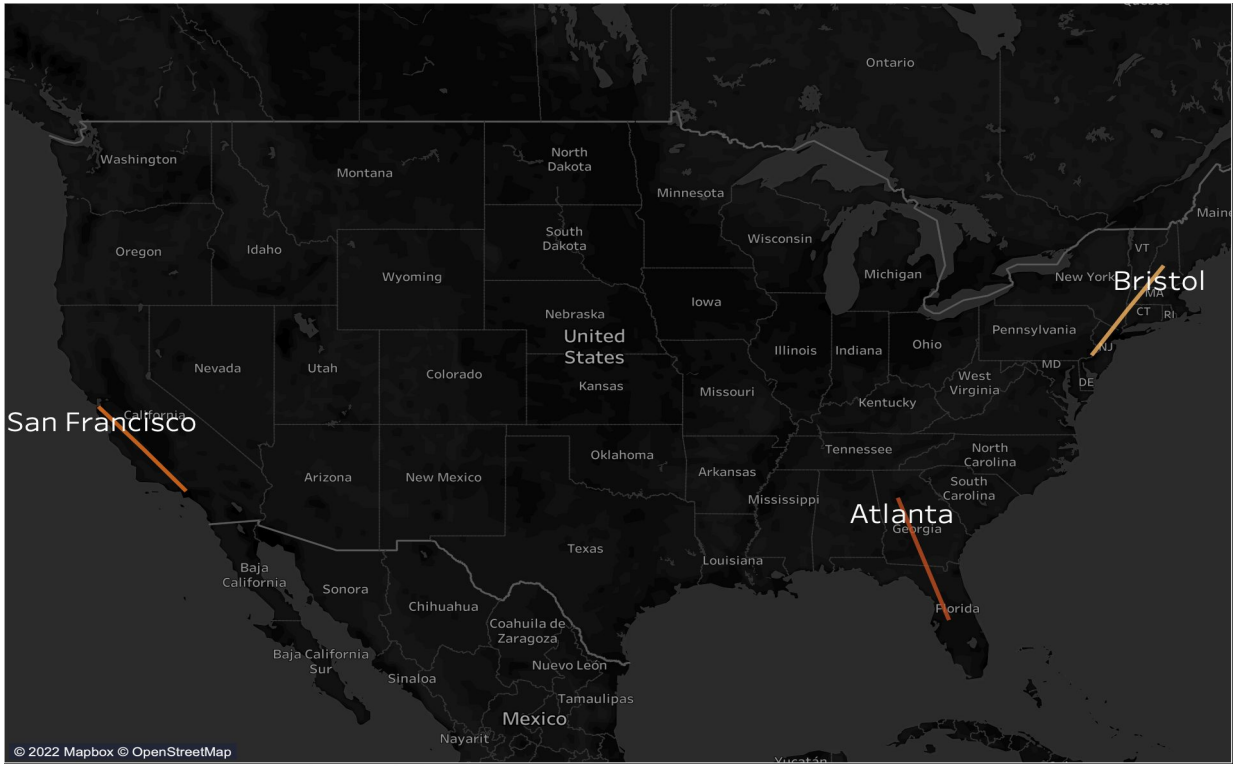
Best Routes ending at Atlanta (highest market)



Best Routes in terms of average profit per mile –

- Long Trip:
  Sioux Falls – Atlanta
- Medium Trip:
  Cape Girardeau – Atlanta
- Short Trip:
  Nashville – Atlanta

Popular Routes for different Market Volumes



- High Volume:

  Atlanta – Lakeland

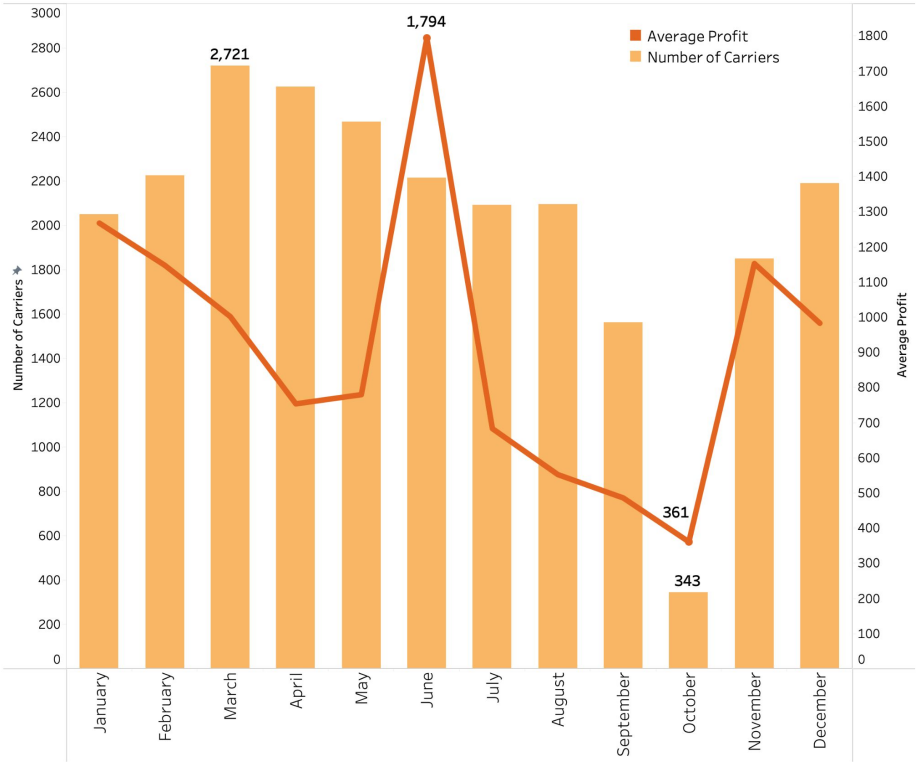
- Medium Volume:

  San Francisco – Los Angeles

- Low Volume:

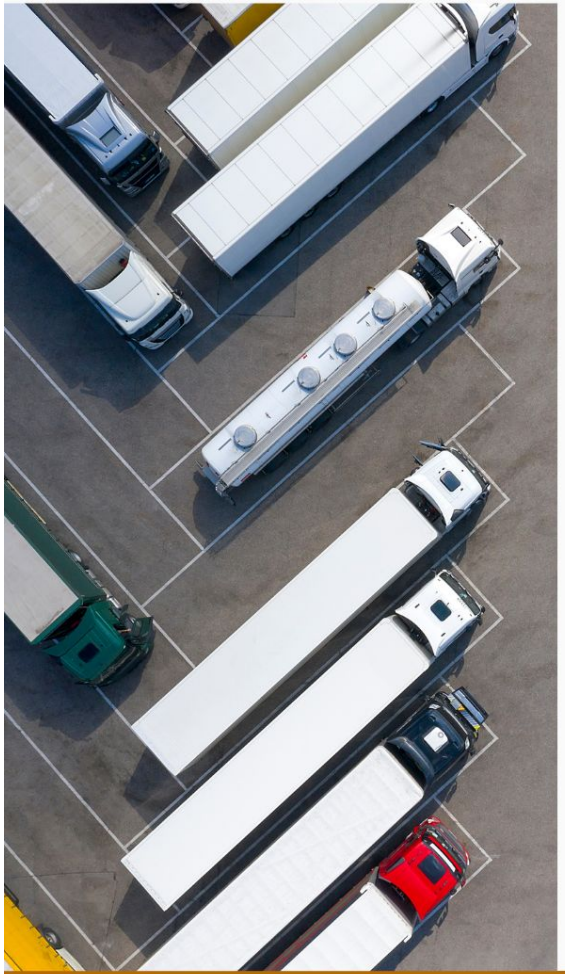
  Bristol – Philadelphia

Monthly Count v/s Profit of Carriers



Are the most booked up months also the most profitable?

- June had the highest average profit.
- March had the highest number of carriers doing business with SmartHop.
- October had the lowest average profit as well as the lowest number of carriers doing business with SmartHop.



AGENDA

01 **Executive Summary**

02 **Data Overview**

Snowflake extraction and Relationship Diagram

03 **Data Manipulation**

Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

04 **Exploratory Data Analysis**

Key Observations

05 **Findings**

06 **Data Modelling**

07 **Recommendations**

SmartHop

Analysis Roadmap

Identify carrier routes and profitability	Analyze dead markets	Identify correlations to carriers behavior	Incorporate insight into recommendation
- **Network analysis** Routes and profitability of the carriers' booked trips - **Cluster analysis** Profitability of carriers based on cluster volume.	- **Premium analysis** Analysis of inherit premium or cost associated to dead market	- **Behavior analysis** Compute the correlations between identified features to interpret carriers' behaviors	- **Profit analysis and recommendations** Summarize the accumulated insights to improve the profitability of carriers

SmartHop

CARRIER ANALYSIS

HIGHEST COUNT OF BOOKED TRIPS

- No Base Clusters
- FEBRUARY AND MARCH are most busy months
- 13.5% Empty mile ratio

Top 10 Carriers

HIGHEST PROFITABLE CARRIERS

- LOW Empty miles ratio
- Taking shorter trips
- Operating in high RPM areas
- Operates in only 5 months of the year

No prominent base cluster

Carrier ID: 613bd4ca73f348105897b822

Two Profitable carriers

Characteristics	Carrier-bfa	Carrier-d2f
Average profit/mile	1.39	1.93
Empty miles ratio	8%	2%
Operating months/year	5 months	5 months
Trip length	Short	Short
Amount of trips	Medium	High
Preferred region	California	Pennsylvania

SmartHop

LOSS ANALYSIS

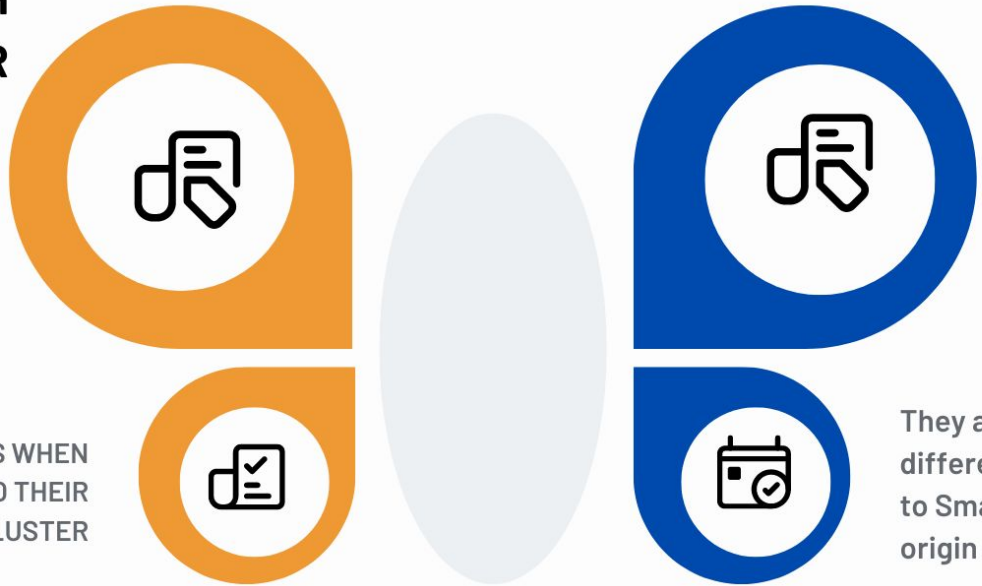
Analyzing carriers which made Unprofitable trips

CARRIERS WITH BASE CLUSTER

Taking losses after profitable trips to prepare for big gains

CARRIERS OVERALL

Sudden losses with no set patterns.



TAKING LOSS WHEN COMING BACK TO THEIR BASE CLUSTER

They are delivering in a different load and coming back to Smarthop with a different origin cluster.

Recommendations

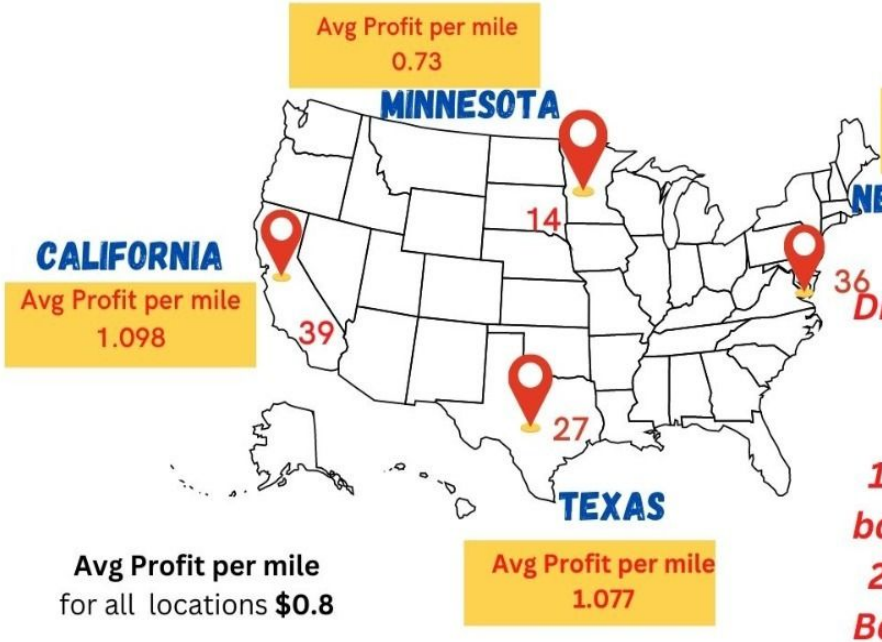
Suggest carriers good deals when coming back to base cluster

Monitor carrier's date of booking to find gaps in their trips as potential opportunites of business.

ROBERT H. SMITH SCHOOL OF BUSINESS

SmartHop

Geographical Analysis- Region



INSIGHTS

Different Average profits for geographically.

Carrier Behavior

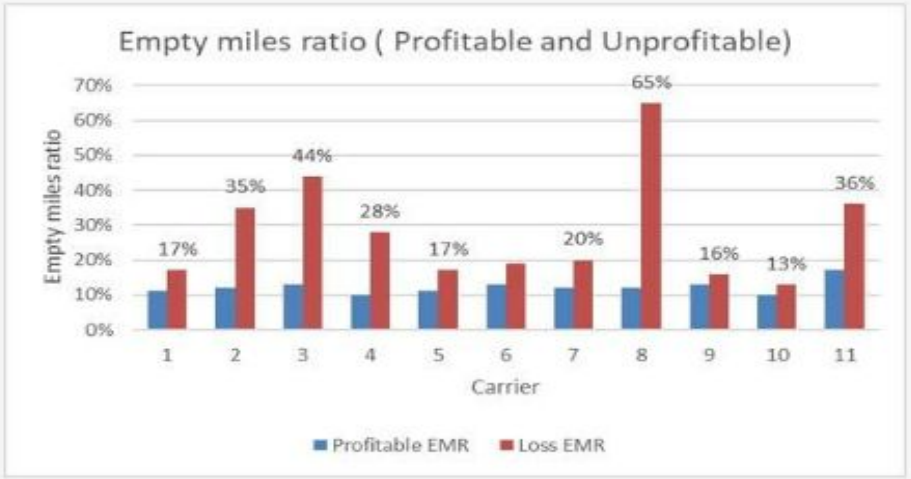
1. Regional carriers heavily operate within base clusters.

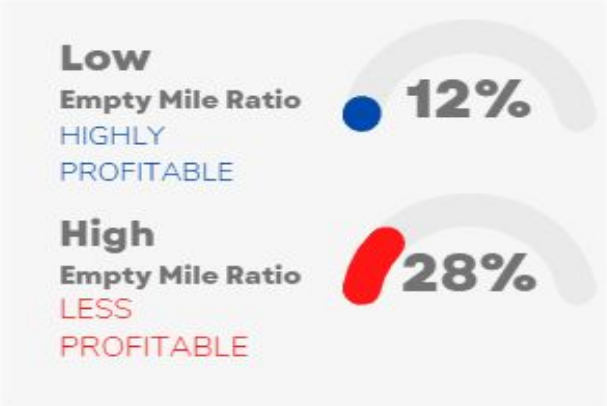
2. All the trucks under them come back to Base cluster- even if there are huge losses.

Recommend trips within the clusters.

Carrier Analysis

When a driver from these Top carriers takes a load, he can compare the Empty miles ratio to see if the trip is going to be Less/More Profitable.



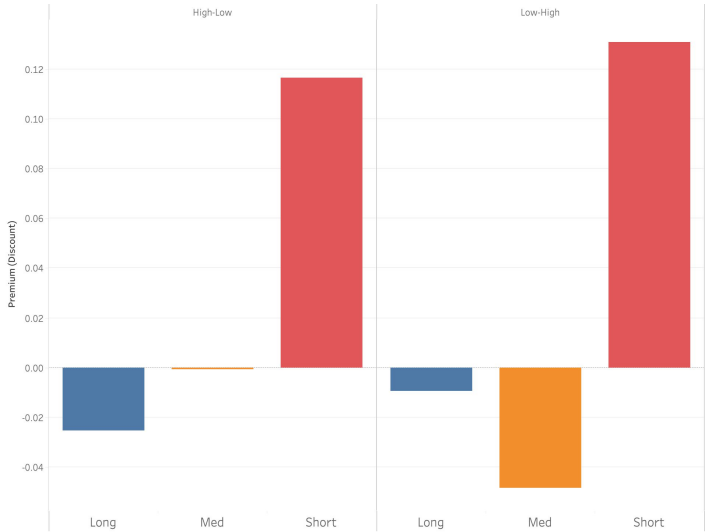


- **Majority of these carriers do not have a base cluster.**
- **Average profit obtained by the carrier is lesser than the average load profit offered.**

ROBERT H. SMITH SCHOOL OF BUSINESS

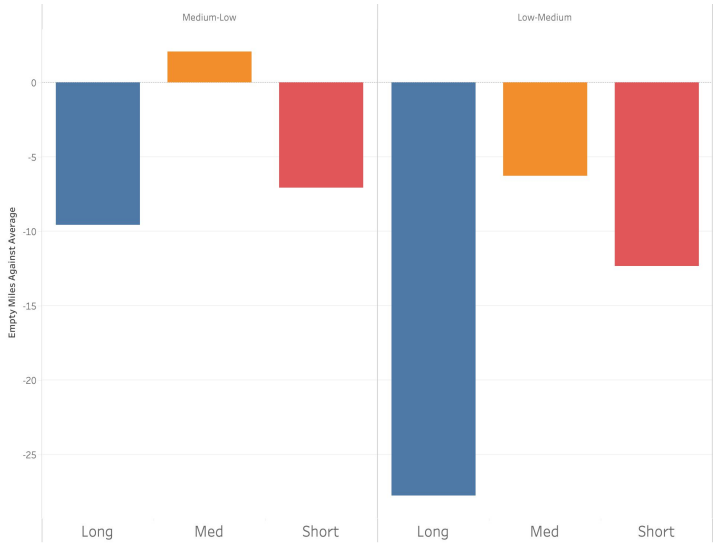
Premium Analysis (High Volume markets)

- Trips to and from low markets are more profitable in short distance trips.
- Empty miles in trips from and to low markets are below average (lower burn).



Premium Analysis (Medium Volume markets)

- Long distance trips from and to dead market cost more compared to average.
- Empty miles between dead and medium market are below average. (lower burn)



Premium Analysis

Market	To	From	Net
H -> L -> H	0.89	0.84	0.01
H -> M -> H	0.90	0.94	0.10
H -> H -> H	0.85	0.85	-0.04

Overall average is $**0.87 PPM**

Low Risk High Value

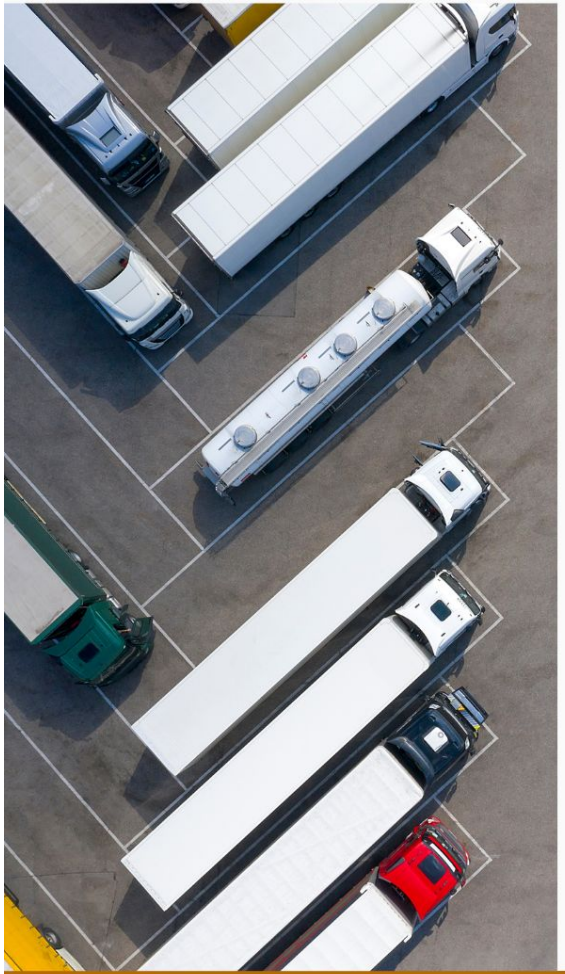
Market	To	From	Net
L -> L -> L	0.89	0.89	0.04
L -> M -> L	0.89	1.01	0.16
L -> H -> L	0.84	0.89	0.01

Trips between all market and medium provides highest reward to medium risk

Low Risk Medium Value

Market	To	From	Net
M -> L -> M	1.01	0.89	0.16
M -> M -> M	0.97	0.97	0.20
M -> H -> M	0.94	0.90	0.10

Medium Risk High Value



AGENDA

01 Executive Summary

02 Data Overview

Snowflake extraction and Relationship Diagram

03 Data Manipulation

Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

04 Exploratory Data Analysis

Key Observations

05 Findings

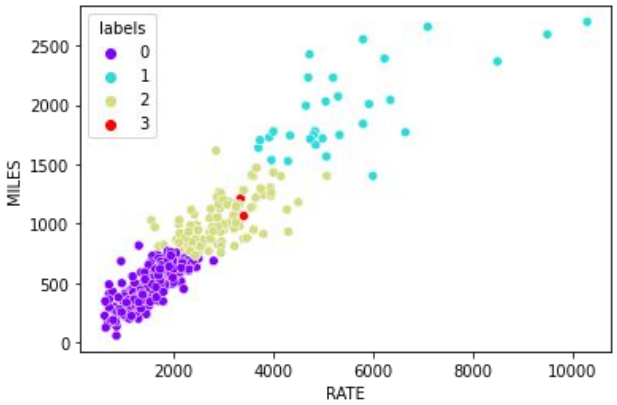
06 Data Modelling

07 Recommendations

SmartHop

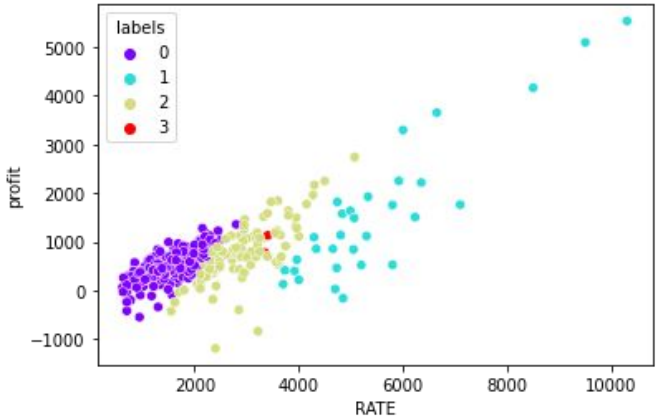
Carrier's Segmentation (using ML)

Rate v/s Mile Clusters



4
UNIQUE
CLUSTERS

Profit v/s Mile Clusters



- Clustering done with both varying and fixed costs
- K means and hierarchical algorithms used for clustering
- Ideal number of clusters is 4 based on elbow plot, silhouette score and dendrogram.
- 4 distinct clusters identified with base preference as length of trip

ROBERT H. SMITH SCHOOL OF BUSINESS

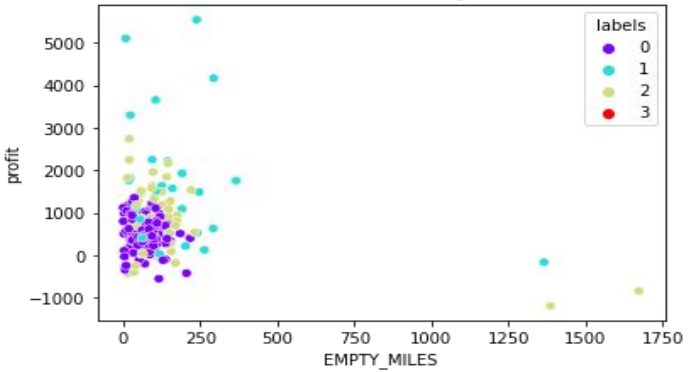
Cluster Overview

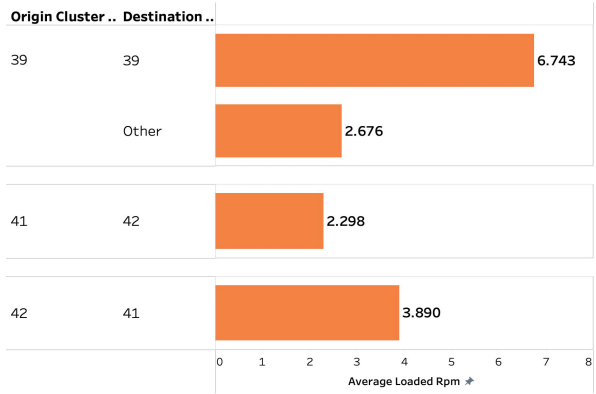
labels	MILES	EMPTY_MILES	RATE	LOADED_RPM	cost	profit	market_high	market_medium	market_low	labels
0	482.327781	70.525917	1534.807792	3.698983	1043.751314	491.056479	18.423810	10.395238	0.971429	210
1	1962.854257	180.840577	5494.862966	2.868764	3922.706716	1572.156250	3.967742	2.000000	0.580645	31
2	1016.231897	120.850887	2892.415278	3.099759	2118.839806	773.575472	17.554455	10.336634	1.782178	101
3	1027.502109	128.925018	3042.457943	3.186822	2167.539931	874.918013	471.333333	320.000000	57.333333	3

Unique elements of each cluster

- Cluster 0 Prefers shorter trips, high empty mile ratio, low profit per trip
- Cluster 1 Prefers longer trips, lowest empty mile ratio, high profit per trip
- Cluster 2 The average cluster with average empty miles and profit
- Cluster 3 Three clusters with high number of trips made by all three

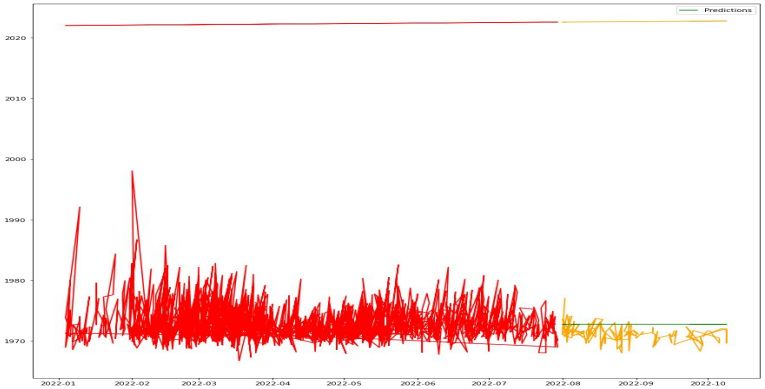
Cluster Insights

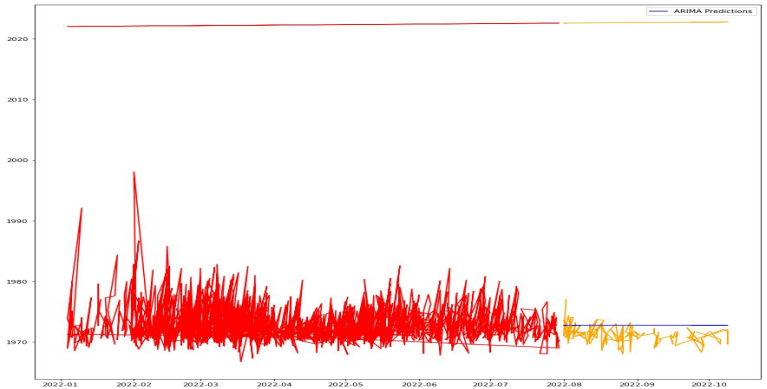




- **Cluster 0:**
  - Short trip preference is high risk as it is highly dependent on empty miles and consecutive trips rate per mile
  - Should try to haul within clusters in a high rpm cluster or should haul from high to medium or low markets to increase profit
- **Cluster 1**
  - Profit is dependent on empty miles and costs.Lower average rpm,so should focus on getting highest rpm possible for more profit

Changes in Profit Generation





Based on Training of top 10 carriers (based on no of trips) for 7 months and then if the Trend followed then

Profits Expected=$154422.09 (Based on ARMA and ARIMA)
Actual Profit=$66945.65

This corresponds to a loss of nearly $90000 and this is only for trips made by only 10 carriers.

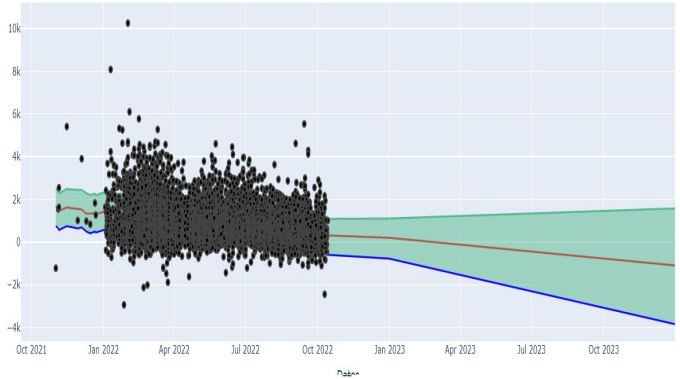
Reasons for the Drop in Profits:

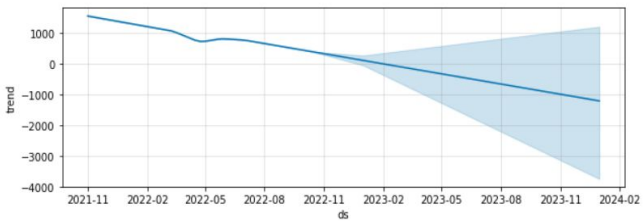
- Few of these profit generating Carriers left the SmartHop platform during the summer and they contributed to burgeoning profits in the months where we trained the data.
- The drivers that were left had more occurrences of EMPTY_MILES in test period than in the first few months.
- The high and medium markets did not change their LOADED_RPM value to a significant extent.
- External factors: Increase in Gasoline Prices, Localized Players(50000 new Trucking Companies emerged in last 1 year)

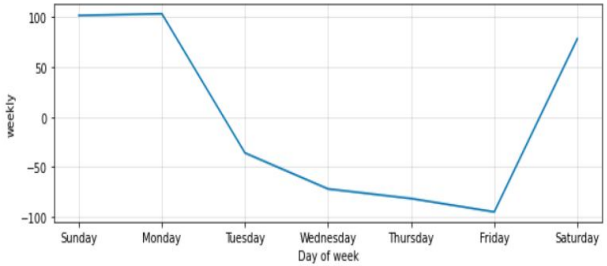
SmartHop

Do carrier preferences affect profits?

Profit Forecasting

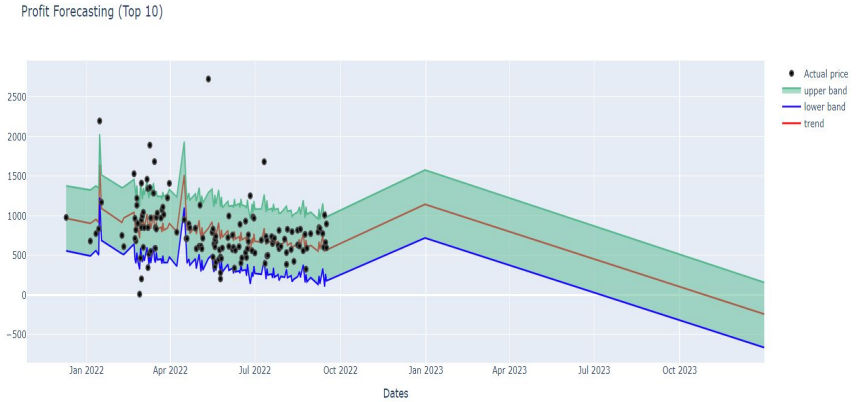


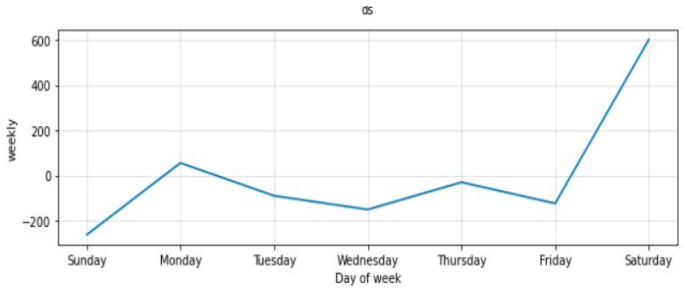


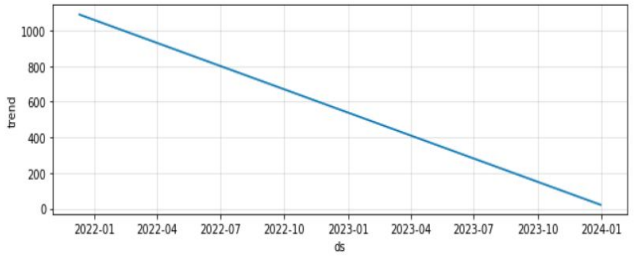


- The Profit generation has seen a slow decline over time and mostly profits are less $4000.
- It is forecasted that that by start of 2024 the profits per trip will be halved.
- Note: Most carriers are willing to haul on **Weekends** which shows more profits (as more trips generally implies more profits). So, many might be part-time Carriers.

Profit Forecasting (Top 10)

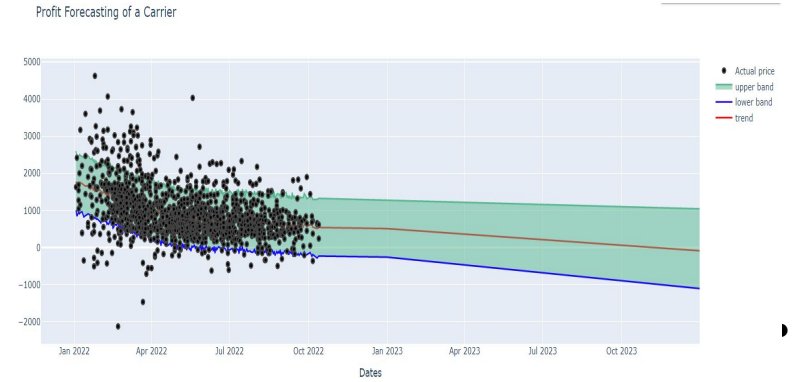


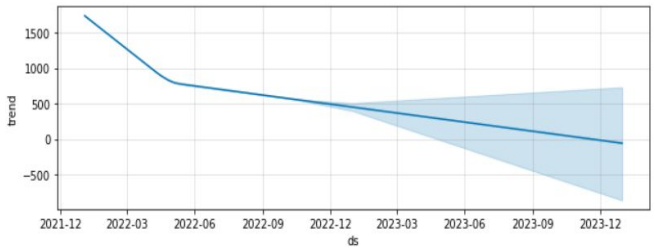


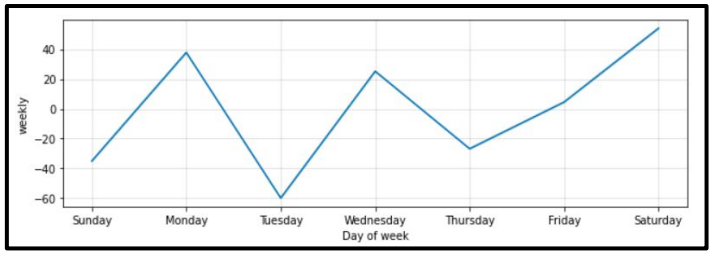


- The Profit generation shows an oscillating pattern based on **Market Conditions and Varied carrier preferences.**
- The forecast shows that if the situation stays the same, then profits will face a decline in the next 2 years even for these profitable carriers.
- It is worth noting that most carriers are willing to haul on any days and not much preferential over the weeks. So, their primary occupation is **Truck Hauling**.

Profit Forecasting (Top Carrier)

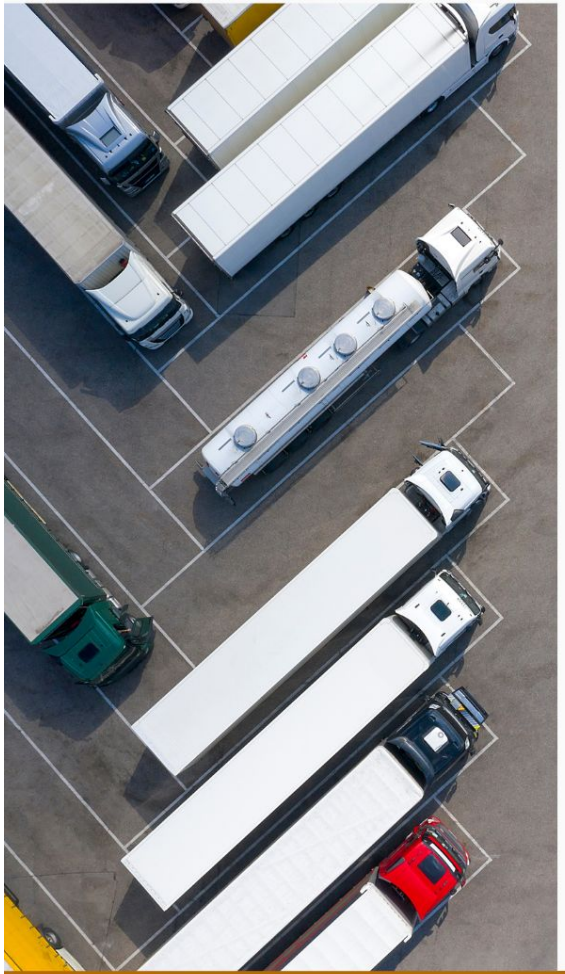






- The carrier has done most trips which yielded in a profit that was more than the average profit. But of late the profits have been declining.
- The trucker, in his 9 months with the company has made around 1400 trips.
- The carrier is expected to make majority of future profitable trips.

Carrier Behavior-
The carrier prefers short trips more with empty miles which are quite less than average.



AGENDA

01 **Executive Summary**

02 **Data Overview**

Snowflake extraction and Relationship Diagram

03 **Data Manipulation**

Data Extraction, Data cleaning, Feature Engineering, Target Variables Identification

04 **Exploratory Data Analysis**

Key Observations

05 **Findings**

06 **Data Modelling**

07 **Recommendations**

SmartHop

RECOMMENDATIONS

WHEN SUGGESTING TRIPS

For higher customer retention,
monitor profits for main carriers and procide them personalized benefits.

01 PROFIT PREDICTIONS

Introduce Cookies on the website to monitor the locations for the carriers, time of inactivity (for the days that they dont book any trips) to segment them and target specifically.

02 ANALYTICS TRACKING

04 WEBSITE OPTIMIZATIONS

Provide the carriers a list of suggestible loads even if its not their preference with creative infographics to suggest- High profit or Recent routes travelled or less/ distance/ quick trip, etc.

03 CUSTOMER SEGEMENTATION

- Base Cluster or not
- Hauling preference- long or short

ROBERT H. SMITH SCHOOL OF BUSINESS

Thanks

We will love to answer your questions!!