Mini Project Data Analysis

Project Report

Iswara Jay Jr.

Outline :

1. Background
2. Dashboard Overview
3. Product Analysis Visualization Properties
4. RFM Analysis Visualization Properties
5. Shipper Analysis Visualization Properties

Background

Working from the previous mini Data Engineering project (Part 2a & 2b)we want to do several analysis case on Northwind dataset, as follows :

Product Analysis

Covering our most popular products traded, and what is our top customers prefer.

RFM Analysis

Helps us understand our consumer-base better, and their classification characteristics.
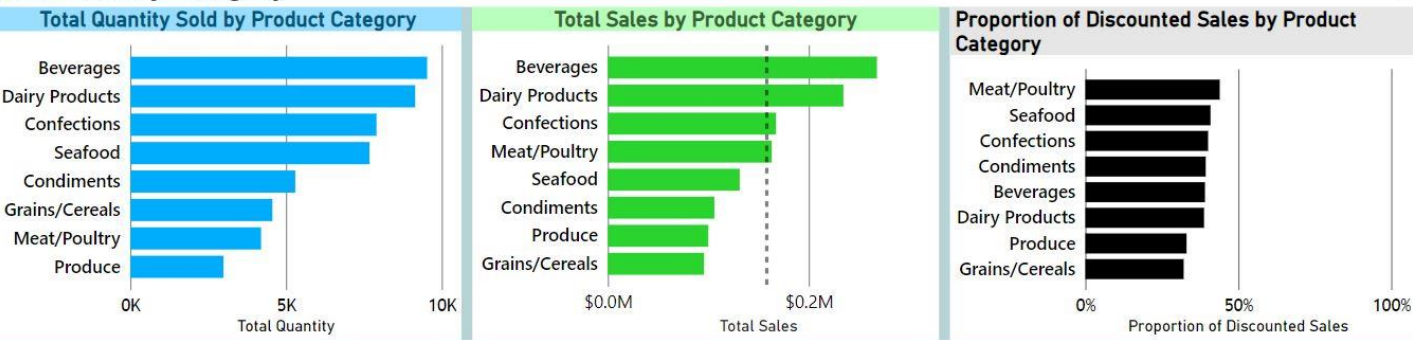
Shipper Analysis

Covering how our current shippers performance compares with each other.

Resulted in a PowerBI dashboard that is the revised version of those found in 2b. This presentation will explains more details regarding the design choices of created dashboard.
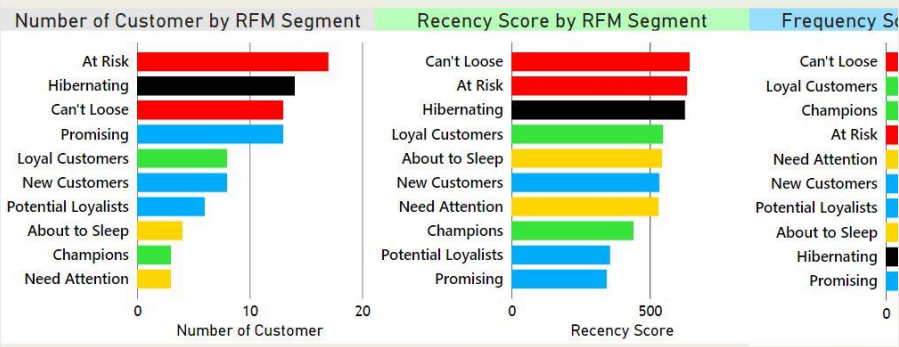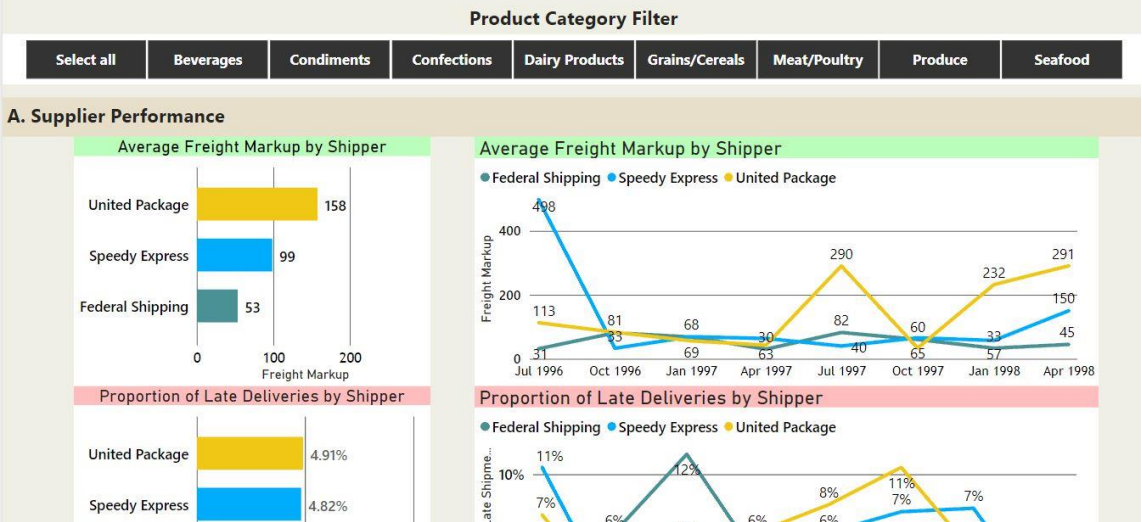
Dashboard Overview

Product Analysis

Main Properties :

- Static Data
- Page 1 : Visualization for Product Analysis
- Page 2 : Visualization for RFM Analysis
- Page 3 : Visualization for Supplier Analysis

RFM Analysis

Supplier Analysis

Product Analysis Visualization Properties

Section A : Metrics by Category

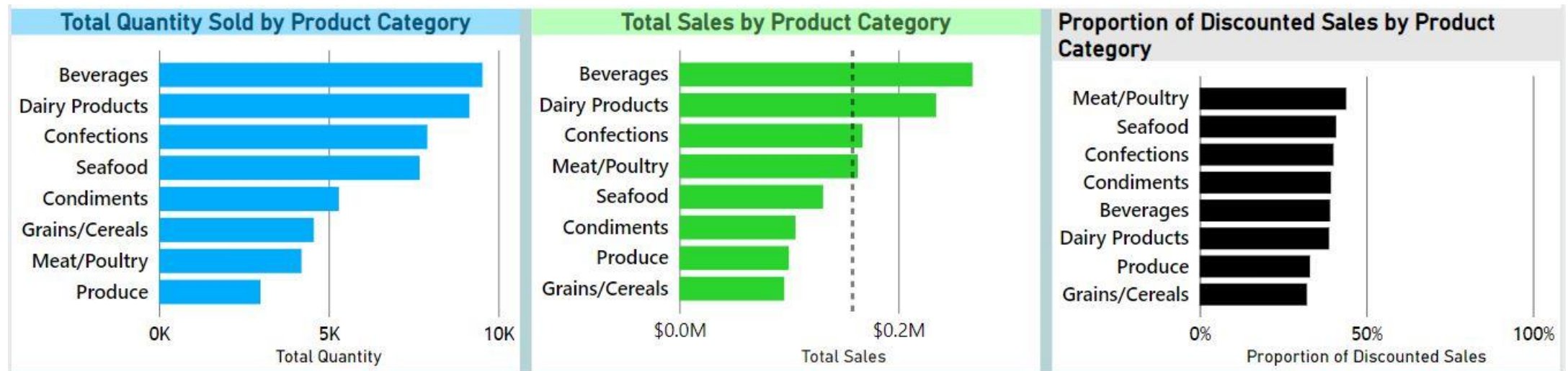Used to gives an overview of important sales metrics by product category.

Color Choice :
1. **Blue** for total quantity sold
2. **Green** for total sales
3. **Black** for proportion of order with discount

Metrics color assignment is kept consistent through the page to makes it **more intuitive.**

Important to view whether proportion of discounted orders **correlates with other metrics** (like total quantity sold)

Bar chart to visualize difference between categorical variables

Total Quantity Sold by Product Category

Category	
Beverages	
Dairy Products	
Confections	
Seafood	
Condiments	
Grains/Cereals	
Meat/Poultry	
Produce	

0K   5K   10K
Total Quantity

Total Sales by Product Category

Beverages
Dairy Products
Confections
Meat/Poultry
Seafood
Condiments
Produce
Grains/Cereals

$0.0M $0.2M
Total Sales

Proportion of Discounted Sales by Product Category

Meat/Poultry
Seafood
Confections
Condiments
Beverages
Dairy Products
Produce
Grains/Cereals

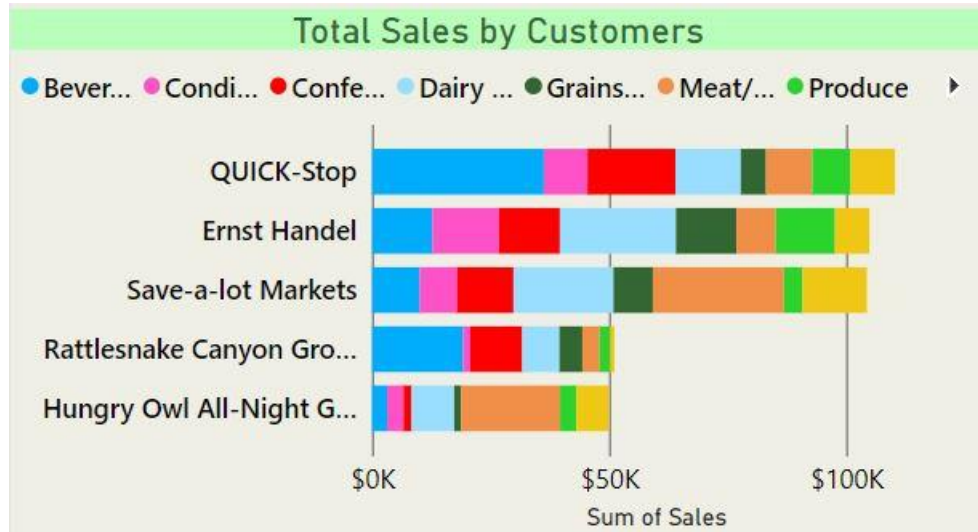0%   50%   100%
Proportion of Discounted Sales

Product Analysis Visualization Properties

Section B : Metrics Filtered by Product Category

Used to gives an overview of important various types of sales metrics filtered using product category, including :

1. **Top 5 Customers** with the highest sales of a certain product categories.

Sales = Green

Category color are mapped by similarity :
- Drinks = Blue
- Processed foods = Red/Pink
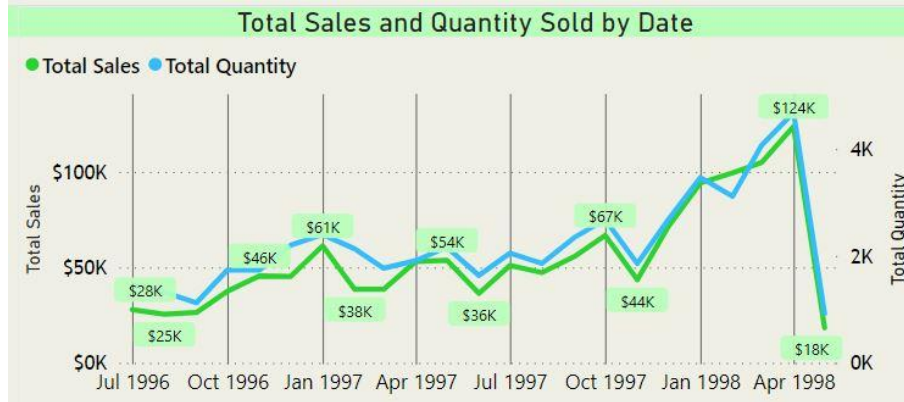- Plant-based Fresh = Green
- Animal-based Fresh = Yellow

Stacked Bar : To observe difference between categorical values (customer) and to get what our most valuable customers food basket looks like

Product Analysis Visualization Properties

Section B : Metrics Filtered by Product Category

**2. Historical sales and quantity sold** of a certain product categories.

Sales = Green; Quantity sold = Blue

More details given for Sales Series rather than Total Quantity since it's more integral from business standpoint

Line Chart: To observe how sales & total quantity develop over a period of time

3. Sales metrics of products which are the member of a certain product categories.
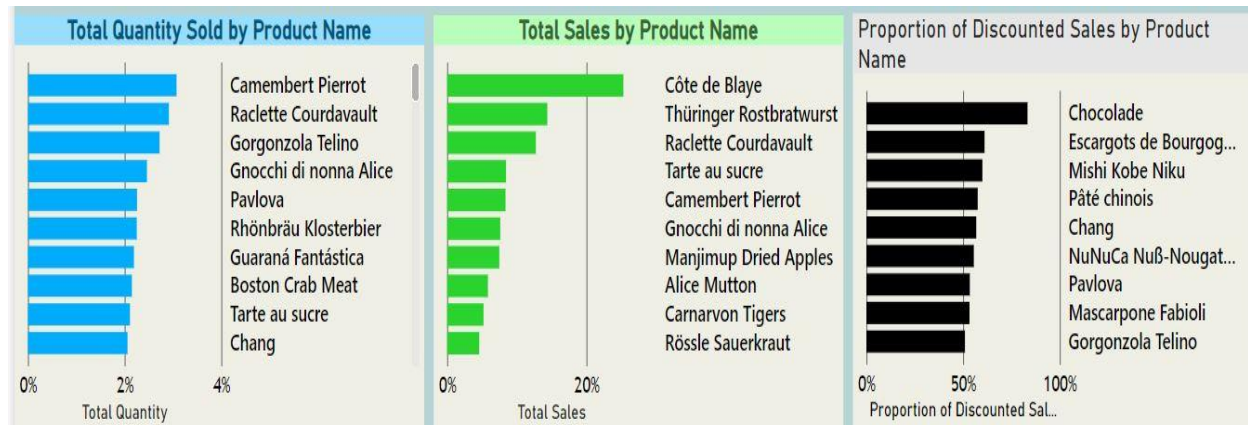

We know how a certain brand performs compared to another from the same category, and whether that is caused by more discount or not

Bar Chart: To observe difference between categorical values. Y-axis on the right for better readability

RFM Analysis Visualization Properties

Gives breakdown of RFM analysis, consisted of several visualization categories as follows :

1. Metrics Overview by RFM Segment
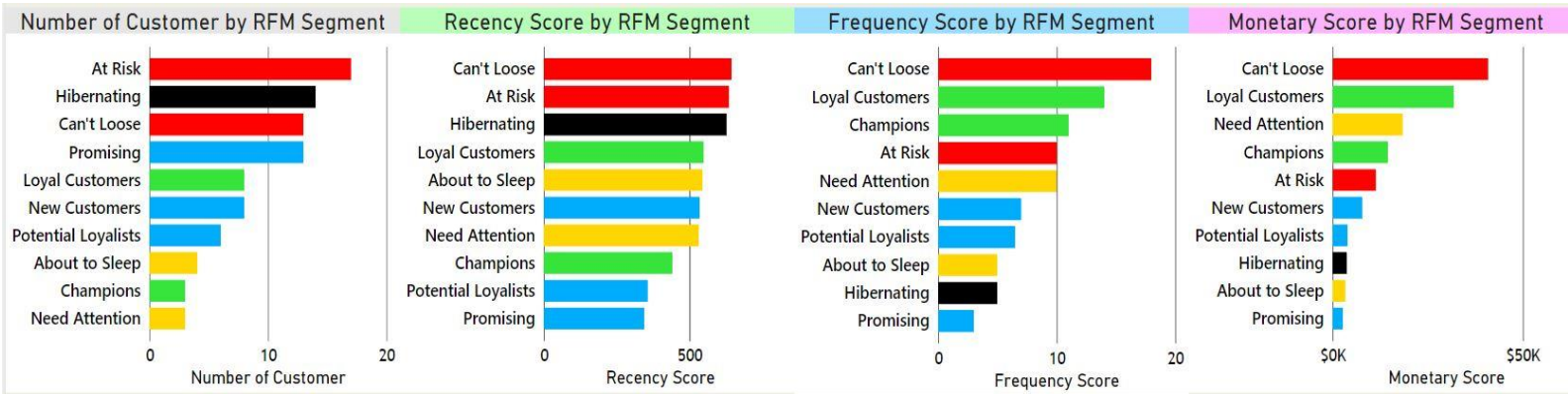
To know how each segment differs from one another in Recency, Frequency, and Monetary

Bar colors are mapped by categories :
- Red : Really important customers that's need to be addressed quickly
- Yellow : Red but with lower priority
- Blue : Promising new customers
- Green : Overall still healthy
- Black : Inactive

Headers are color coded to better differentiate between metrics

Bar Chart: To observe difference between categorical values.

2. Matrix Table of All Customers

Used to get a more detailed view on each customer name, can be filtered with 3 main segment groups :

- High Value : Red Bars
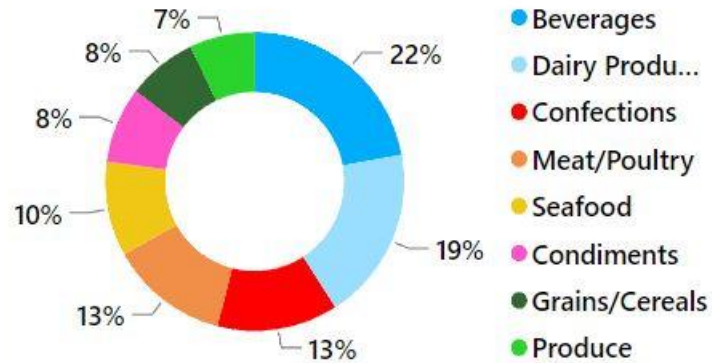- Promising : Blue + Green
- Other : Yellow + Black
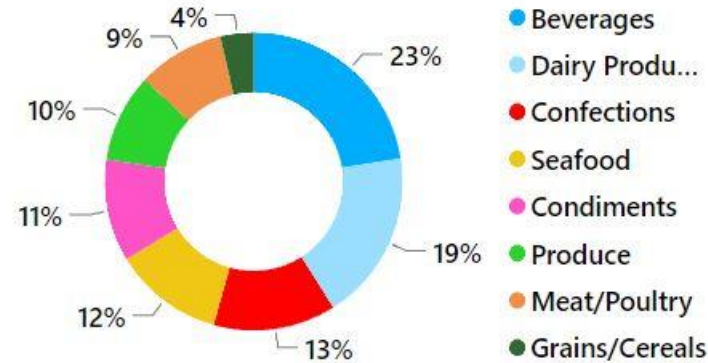
RFM Analysis Visualization Properties

3. Sales Distribution of High Value & Promising Customers

Since these two groups are the one that needs to be given more attention, we need to know what is their preference in general. Sales distribution is one of the simple way to do so.


Category color code is the same as visualization in Product Analysis for better continuity

Pie Chart: To observe composition by categorical variables in a single group/category.

Supplier Analysis Visualization Properties

Overall can be filtered by product categories for better breakdown of a certain group of products, **since each supplier has radical performance change for a certain product categories.**

Section A : Supplier Performance
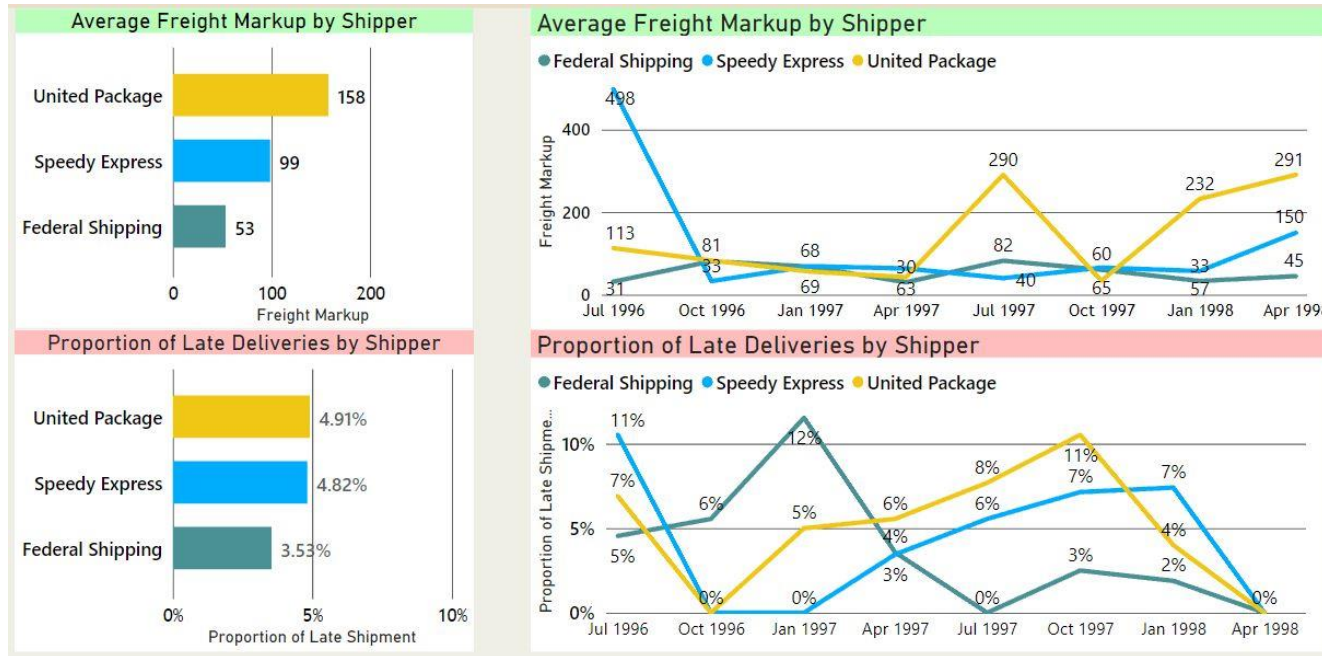
Line chart to view historical development of markup and late deliveries, so we know whether our suppliers are improving or not

Green Header : Higher is better
Red Header : Lower is better

Each supplier is kept at a consistent color code

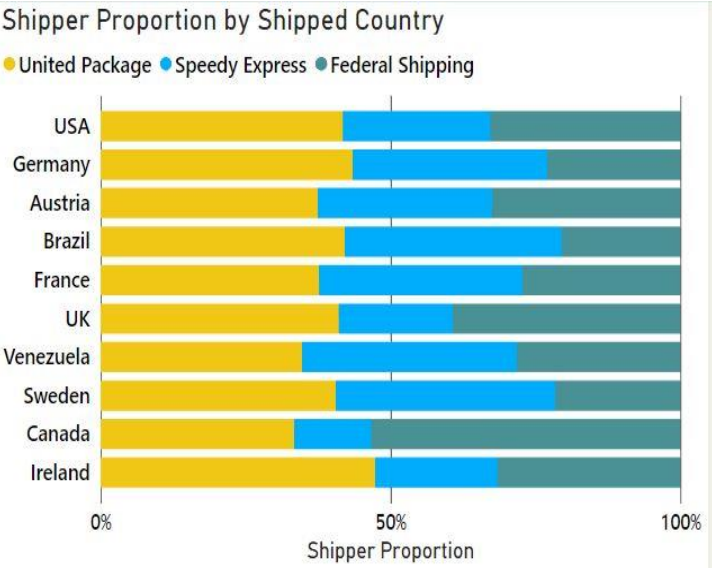For general overview of Markup (how profitable) and late deliveries proportion (how competent)

Section B : Supplier Utilization

To get information on how much a supplier is being utilized to ship a certain product category.

Stacked bar of profit let us know which country of destination contributes the most to our business

While Shipper proportion let us know which shipper did we use the most for every destination.

If we are using a shipper with bad performance (low markup and/or high percentage of late delivery) at significant proportion to our top destination, then we need to improve out supply chain.