

DataCo Regional Performance Analysis and Supply Chain Optimization

Business Analytics - Sekolah Data Pacmann

Outline

- Introduction
- Topic, Dataset, and Tools
- Problem Statements, Goals, Metrics, and Scenario
- Data Cleaning, Feature Engineering
- Data Analysis, Business Insights and Hypothesis Test
- Dashboard
- Conclusion and Future Work
- References

Introduction

Introduction

A well-functioning supply chain is crucial for businesses to provide quality products and satisfy their customers. As such, supply chain management presents both challenges and opportunities for businesses to optimize their processes and improve their bottom line. My project, "DataCo Regional Performance Analysis and Supply Chain Optimization," focuses on analyzing DataCo's regional performance and offering recommendations for enhancing their supply chain management.

Through this project, I aim to gain a better understanding of the complexities of supply chain management while honing my analytical skills. Ultimately, I hope to provide DataCo with actionable insights to improve their supply chain management and better serve their customers.

Topic, Dataset, and Tools

Topic, Dataset, and Tools

- Topic:
 - Supply Chain Optimization
- Dataset:
 - [DataCo Smart Supply Chain For Big Data Analysis | Kaggle](#)
- Tools:
 - Data cleaning, feature engineering, and data analysis: Python (Pandas, Altair, Plotly, SciPy)
 - Dashboarding: Python (Dash and Plotly), Datawrapper, Render









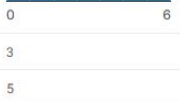
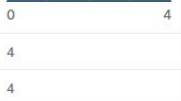
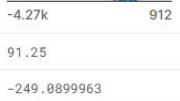
Topic, Dataset, and Tools

- Dataset Preview:

The dataset used for this analysis comes from DataCo Global's supply chain operations, consisting of 180,519 rows and 53 columns. It provides a comprehensive view of various aspects of the company's supply chain, including shipping, order details, customer information, product categories, geographical data, and performance metrics. The dataset covers order date placed between January 1, 2015 and January 31, 2018. This wealth of information allows for a thorough examination of sales, order volume, and delivery efficiency across diverse regions and markets.

DataCoSupplyChainDataset.csv (95.91 MB) Download Fullscreen More

Detail Compact Column 10 of 53 columns

Type	# Days for shipping...	# Days for shipmen...	# Benefit per order	# Sales per customer	Delivery
DEBIT	38%				Late delivery
TRANSFER	28%				Advance s
Other (61341)	34%				Other (39)
DEBIT	3	4	91.25	314.6400146	Advance
TRANSFER	5	4	-249.0899963	311.3599854	Late del
CASH	4	4	-247.7799988	309.7200012	Shipping
DEBIT	3	4	22.86000061	304.8099976	Advance
PAYMENT	2	4	134.2100067	298.25	Advance
TRANSFER	6	4	18.57999992	294.980011	Shipping
DEBIT	2	1	95.18000031	288.4200134	Late del
TRANSFER	2	1	68.43000031	285.1400146	Late del
CASH	3	2	133.7200012	278.5899963	Late del
CASH	2	1	132.1499939	275.3099976	Late del
TRANSFER	6	2	130.5800018	272.0299988	Shipping
TRANSFER	5	2	45.68999863	268.7600098	Late del
TRANSFER	4	2	21.76000023	262.2000122	Late del
DEBIT	2	1	24.57999992	245.8099976	Late del
TRANSFER	2	1	16.38999939	327.75	Late del
DEBIT	2	1	-259.5799866	324.4700012	Late del
PAYMENT	5	2	-246.3600006	321.2000122	Late del

Problem Statements, Goals, Metrics and Scenario

Problem Statements, Goals, Metrics, and Scenario

- Problem Statement:
 - The company needs to analyze sales, order volume, and delivery efficiency across different regions and markets to identify opportunities for growth and areas that need improvement. A comprehensive dashboard will enable stakeholders to make informed decisions regarding region expansion, resource allocation, and process optimization.
- Goals:
 - Develop a dashboard that provides insights into regional sales performance, order volume, and delivery efficiency.
 - Identify high-performing and underperforming regions to optimize resource allocation.
 - Discover potential markets for expansion and improvement.
 - Enable data-driven decision-making for supply chain management.

Problem Statements, Goals, Metrics, and Scenario

- Metrics:
 - Total sales by region and market
 - Total order by region and market
 - Total profit
 - Average delivered days
 - Average orders by region and market
 - OTIF rate
 - Sales by product category within each region
- Scenario:
 - Assume that the company is planning to growth its operations into underperforming regions and wants to optimize its supply chain processes in high performing regions. The management team needs a data-driven approach to identify high-potential markets and areas for improvement in current operations.

Data Cleaning and Feature Engineering

Data Cleaning and Feature Engineering

Before proceeding to this stage, I carefully examined the columns needed for my project and selected the ones I deemed sufficient for the analysis and dashboard creation. Out of 53 columns, I chose 16. After selecting these columns, I performed data cleaning tasks such as reformatting the column names to a uniform snake case format. Since almost all the country names were in Spanish and not commonly known, I replaced them with English or their internationally recognized names. I also replaced some region names to be more internationally recognized.

After the data cleaning, I proceeded to feature engineering, where I extracted information from existing columns and added it to new ones. I calculated sales per item, determined the difference in shipping days between actual and scheduled, and extracted the year, month, and date from the order date to facilitate filtering and analysis later on. The following images shows the before and after of the dataframe after it has been cleaned and transformed.

Data Cleaning and Feature Engineering

Before:

```
print(main_df.head())
```

Type	Days for shipping (real)	Days for shipment (scheduled)	
0 DEBIT	3	4	
1 TRANSFER	5	4	
2 CASH	4	4	
3 DEBIT	3	4	
4 PAYMENT	2	4	

Benefit per order	Sales per customer	Delivery Status	
0 91.25	314.64	Advance shipping	
1 -249.09	311.36	Late delivery	
2 -247.78	309.72	Shipping on time	
3 22.86	304.81	Advance shipping	
4 134.21	298.25	Advance shipping	

Late_delivery_risk	Category Id	Category Name	Customer City	
0 0	73	Sporting Goods	Caguas	
1 1	73	Sporting Goods	Caguas	
2 0	73	Sporting Goods	San Jose	
3 0	73	Sporting Goods	Los Angeles	
4 0	73	Sporting Goods	Caguas	

Customer Country	Customer Email	Customer Fname	Customer Id	Customer Lname	
0 Puerto Rico	XXXXXXXXXX	Cally	20755	Holloway	
1 Puerto Rico	XXXXXXXXXX	Irene	19492	Luna	
2 EE. UU.	XXXXXXXXXX	Gillian	19491	Maldonado	
3 EE. UU.	XXXXXXXXXX	Tana	19490	Tate	
4 Puerto Rico	XXXXXXXXXX	Orli	19489	Hendricks	

Customer Password	Customer Segment	Customer State	Customer Street	
0 XXXXXXXXXX	Consumer	PR	5365 Noble Nectar Island	
1 XXXXXXXXXX	Consumer	PR	2679 Rustic Loop	
2 XXXXXXXXXX	Consumer	CA	8510 Round Bear Gate	
3 XXXXXXXXXX	Home Office	CA	3200 Amber Bend	
4 XXXXXXXXXX	Corporate	PR	8671 Iron Anchor Corners	

Customer Zipcode	Department Id	Department Name	Latitude	Longitude	
0 725.00	2	Fitness	18.25	-66.04	
1 725.00	2	Fitness	18.28	-66.04	
2 95125.00	2	Fitness	37.29	-121.88	
3 90027.00	2	Fitness	34.13	-118.29	
4 725.00	2	Fitness	18.25	-66.04	

After:

```
print(selected_df.head(5))
```

order_date	order_region	order_country	order_city	
0 2015-01-01 00:00:00	Central America	Mexico	Mexico City	
1 2015-01-01 00:21:00	South America	Colombia	Dos Quebradas	
2 2015-01-01 00:21:00	South America	Colombia	Dos Quebradas	
3 2015-01-01 00:21:00	South America	Colombia	Dos Quebradas	
4 2015-01-01 01:03:00	South America	Colombia	Dos Quebradas	

market	sales	order_item_quantity	shipping_date	
0 Latin America	299.98	1	1/3/2015 0:00	
1 Latin America	199.99	1	1/4/2015 0:21	
2 Latin America	250.00	5	1/4/2015 0:21	
3 Latin America	129.99	1	1/4/2015 0:21	
4 Latin America	199.92	4	1/6/2015 1:03	

days_for_shipping_real	days_for_shipment_scheduled	category_name	
0 2	4	Camping & Hiking	
1 3	4	Water Sports	
2 3	4	Women's Apparel	
3 3	4	Men's Footwear	
4 5	4	Indoor/Outdoor Games	

delivery_status	late_delivery_risk	shipping_mode	order_status	
0 Advance shipping	0	Standard Class	CLOSED	
1 Advance shipping	0	Standard Class	PENDING_PAYMENT	
2 Advance shipping	0	Standard Class	PENDING_PAYMENT	
3 Advance shipping	0	Standard Class	PENDING_PAYMENT	
4 Late delivery	1	Standard Class	CLOSED	

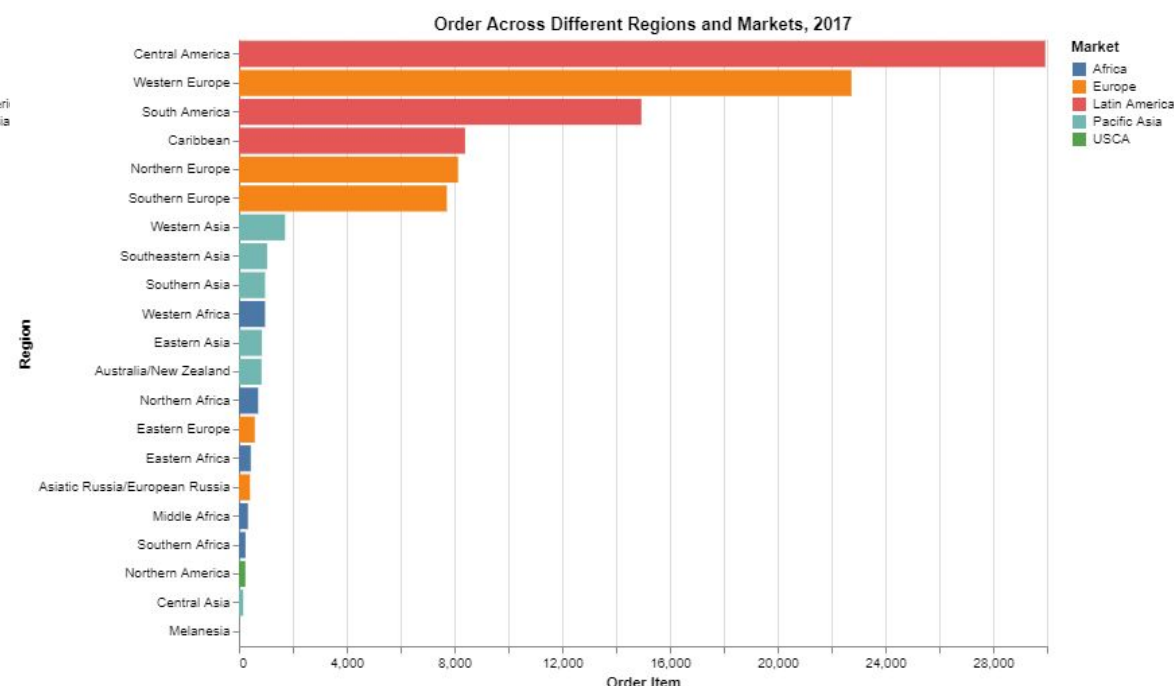
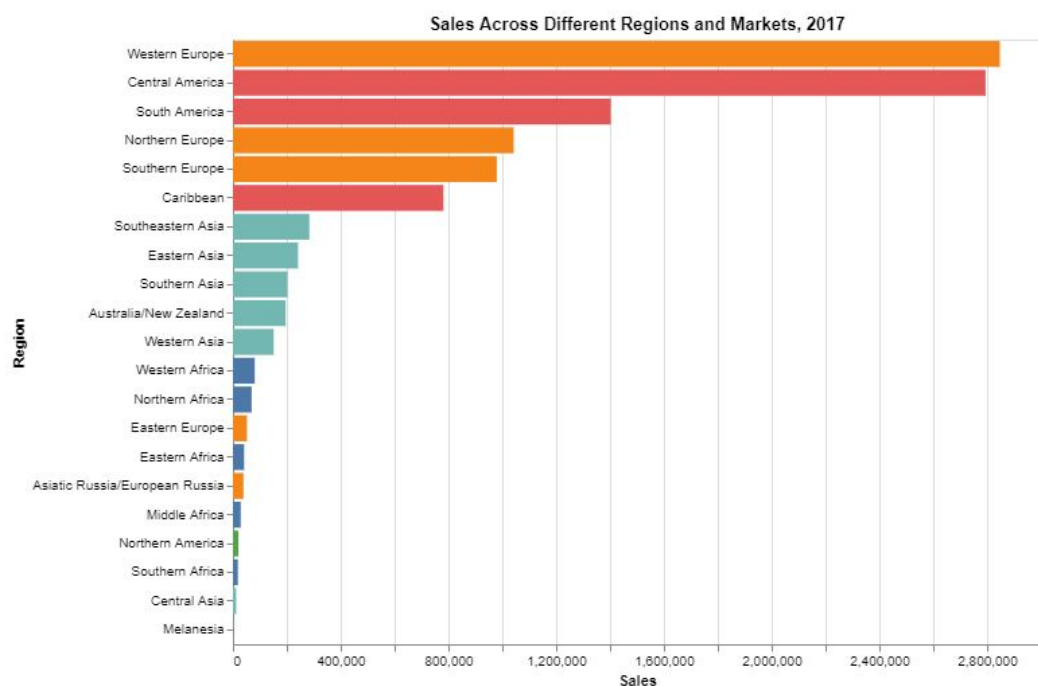
order_profit_per_order	sales_per_item	shipping_days_difference	
0 88.79	299.98	2	
1 91.18	199.99	1	
2 68.25	50.00	1	
3 36.47	129.99	1	
4 33.59	49.98	-1	

order_year	order_month	order_year_month_date	
0 2015	January	2015-01-01	
1 2015	January	2015-01-01	
2 2015	January	2015-01-01	
3 2015	January	2015-01-01	
4 2015	January	2015-01-01	

Data Analysis and Business Insights

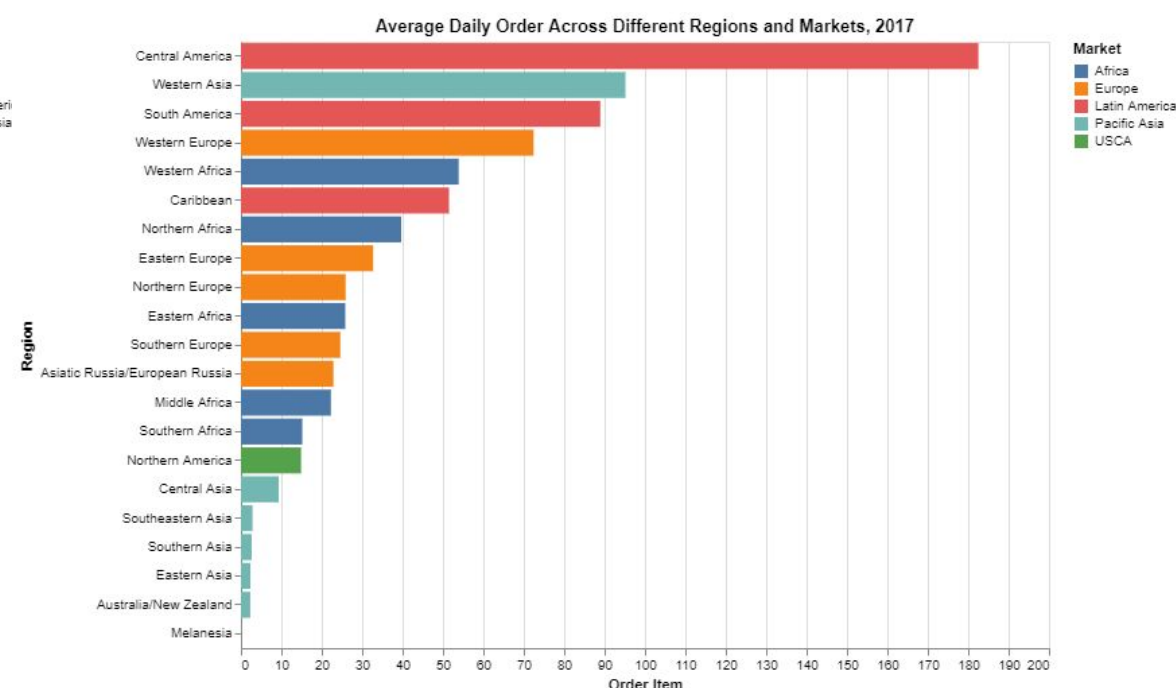
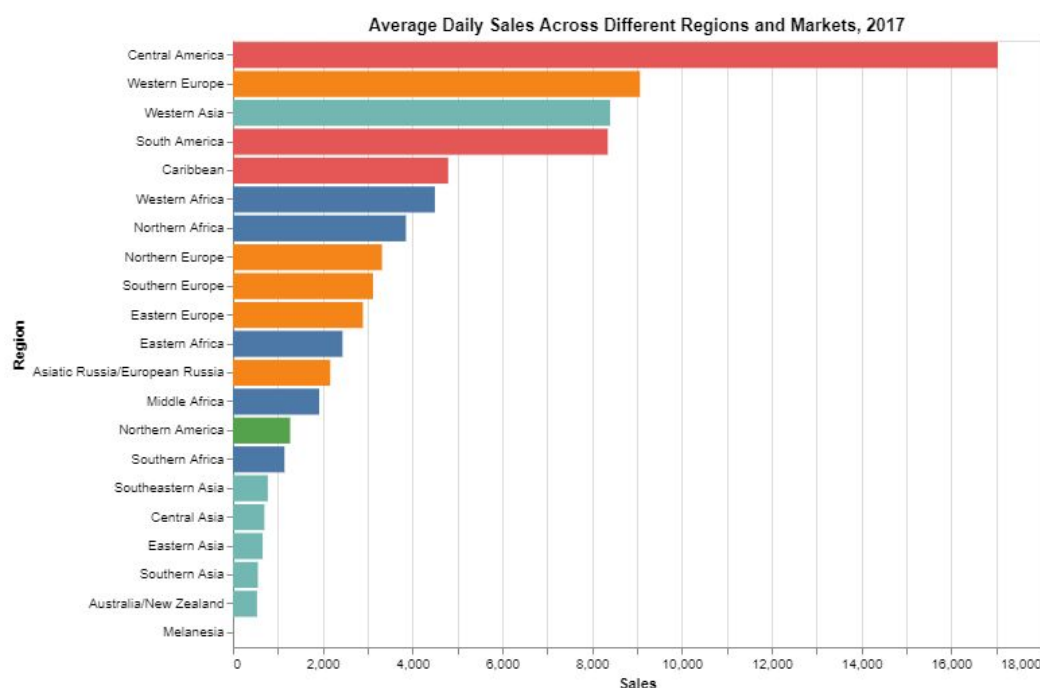
Data Analysis and Business Insights

- Sales and Order Analysis
 - Visualization Sales and Order Volume



Data Analysis and Business Insights

- Sales and Order Analysis
 - Visualization Daily Average Sales and Order



Data Analysis and Business Insights

- Sales and Order Analysis
 - Insights
 - ❖ Western Europe, Central America, and South America are the top three regions in terms of order volume and sales.
 - ❖ Melanesia, Central Asia, and Southern Africa have the lowest order volume and sales.
 - ❖ Western Asia has a high average daily sales value but a low average daily order volume, suggesting high-value orders.
 - ❖ Europe and Latin America markets have the highest sales and order volumes.
 - ❖ Africa has the lowest sales and order volumes among all markets.

Data Analysis and Business Insights

- Sales and Order Analysis
 - Recommendations
 - ❖ Focus on increasing sales and order volume in the low-performing regions such as Melanesia, Central Asia, and Southern Africa. This may involve implementing targeted marketing campaigns or improving the product offerings in these regions.
 - ❖ Maintain and strengthen the position in high-performing regions like Western Europe, Central America, and South America. Keep a close eye on the competition and make sure to continue delivering a great customer experience.
 - ❖ Investigate the high-value orders in Western Asia and identify opportunities to further capitalize on these sales. This may involve identifying the best-selling products and offering promotions or bundle deals to boost sales.
 - ❖ Explore the reasons behind the low order volume in Northern America, despite being part of the USCA market. Consider targeted strategies to increase order volume in this region.

Data Analysis and Business Insights

- Delivery Efficiency Analysis
 - Visualization On Time Delivery Rate Across Market and Shipping Mode



Data Analysis and Business Insights

- Delivery Efficiency Analysis
 - Visualization Average Days for Shipping (Real vs Scheduled) Across Market



Data Analysis and Business Insights

- Delivery Efficiency Analysis

- Insights

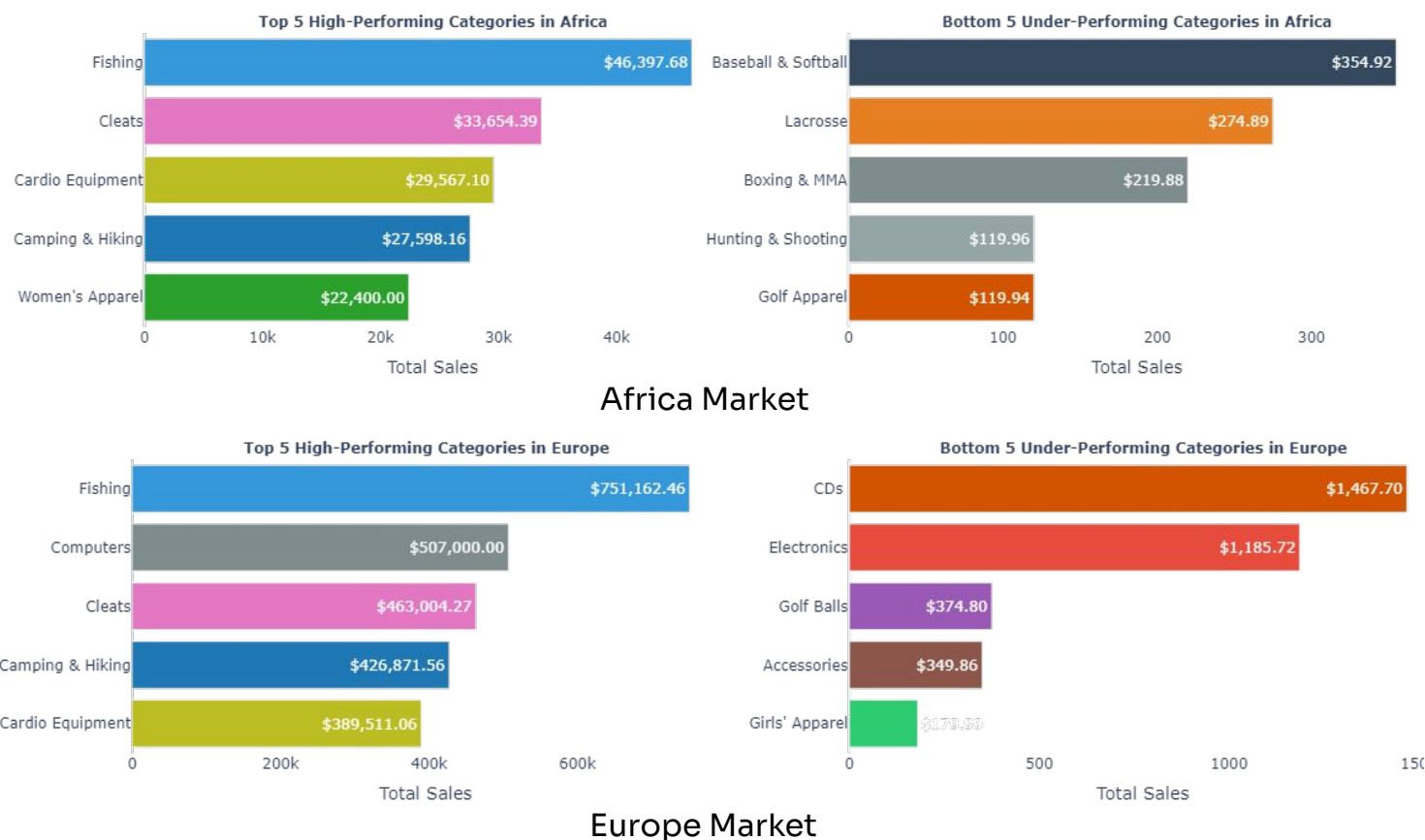
- ❖ First Class shipping has a 0% on-time delivery rate across all markets, indicating significant inefficiencies.
 - ❖ Same Day shipping has the highest on-time delivery rate in the USCA market (80%), but it's lower in other markets.
 - ❖ Standard Class shipping generally performs better than Second Class shipping in terms of on-time delivery, with the best performance in the Africa market (62.96%).
 - ❖ First Class shipping takes 2 days on average, while it is scheduled for 1 day, indicating a discrepancy between the actual and scheduled shipping time.
 - ❖ Same Day shipping is faster than scheduled across all markets, with the best performance in the USCA market (0.20 days on average).
 - ❖ Second Class shipping generally takes longer than scheduled, with the biggest discrepancy in the Pacific Asia market (4.07 days on average vs. 2 days scheduled).
 - ❖ Standard Class shipping meets the scheduled shipping time in Africa, Latin America, and Pacific Asia markets, but it's slightly faster in the Europe and USCA markets.

Data Analysis and Business Insights

- Delivery Efficiency Analysis
 - Recommendations
 - ❖ Investigate and address the issues causing the 0% on-time delivery rate for First Class shipping across all markets. This may involve analyzing the causes of delays and implementing process improvements or working closely with shipping partners.
 - ❖ Monitor and improve the shipping performance for Second Class shipping, as it consistently takes longer than scheduled. Consider collaborating with shipping partners to identify bottlenecks and develop solutions to reduce shipping time.
 - ❖ Maintain the good performance of Standard Class shipping in terms of meeting or exceeding the scheduled shipping time, and aim to replicate the faster-than-scheduled performance observed in Europe and USCA markets across all regions.

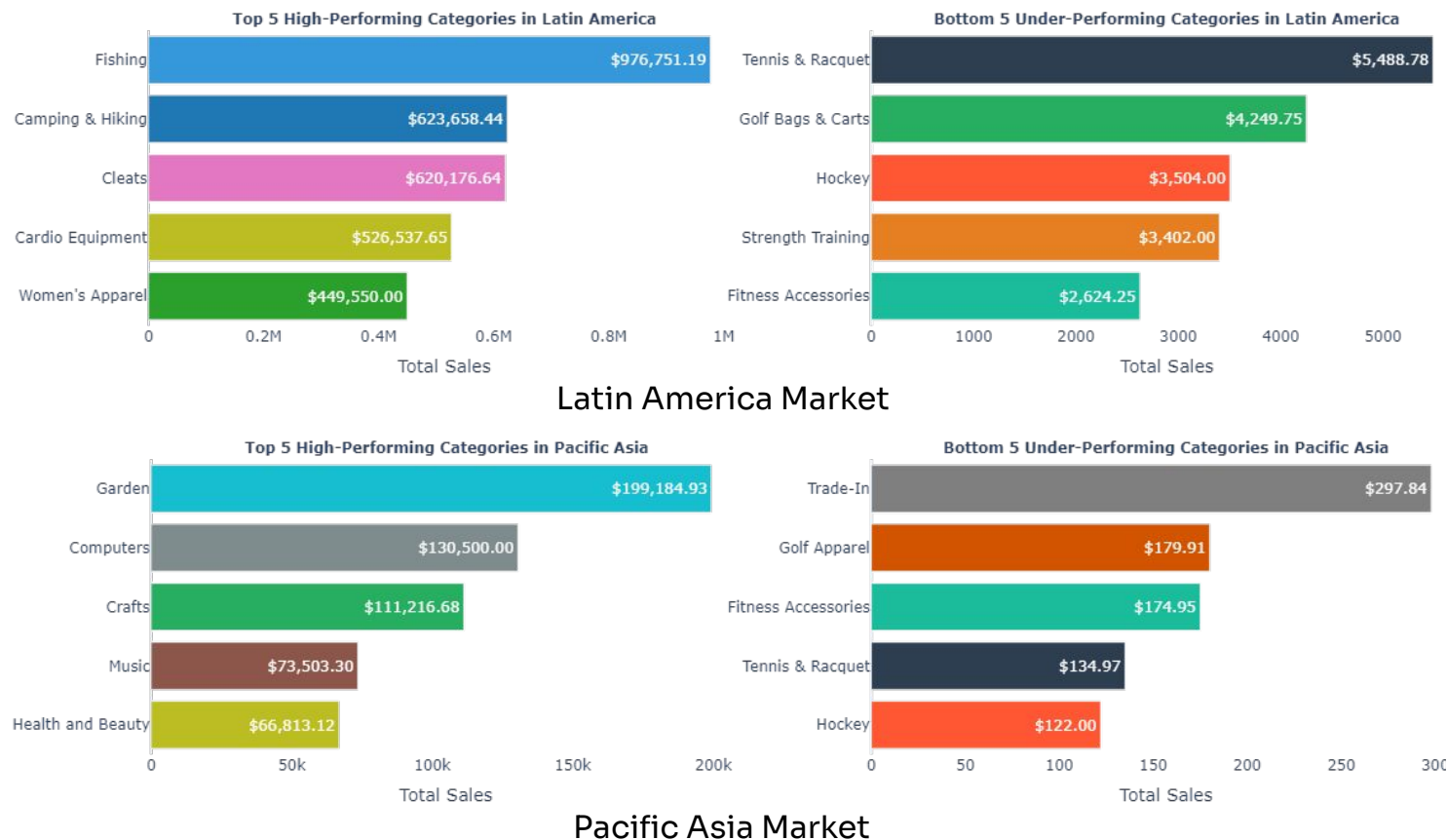
Data Analysis and Business Insights

- Products Performance Analysis
 - Visualization High-Performing and Under-Performing Products Across Markets



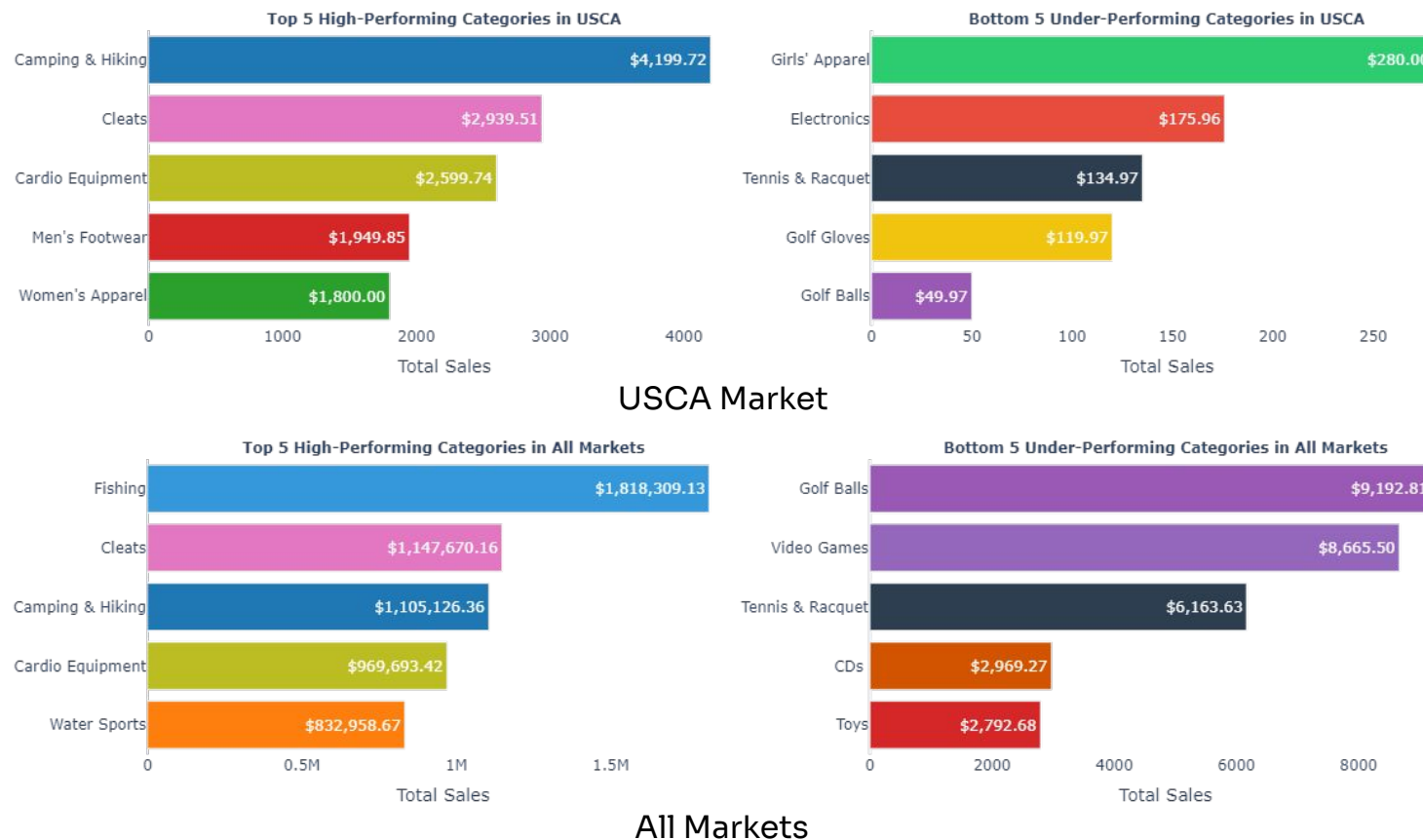
Data Analysis and Business Insights

- Products Performance Analysis
 - Visualization High-Performing and Under-Performing Products Across Markets



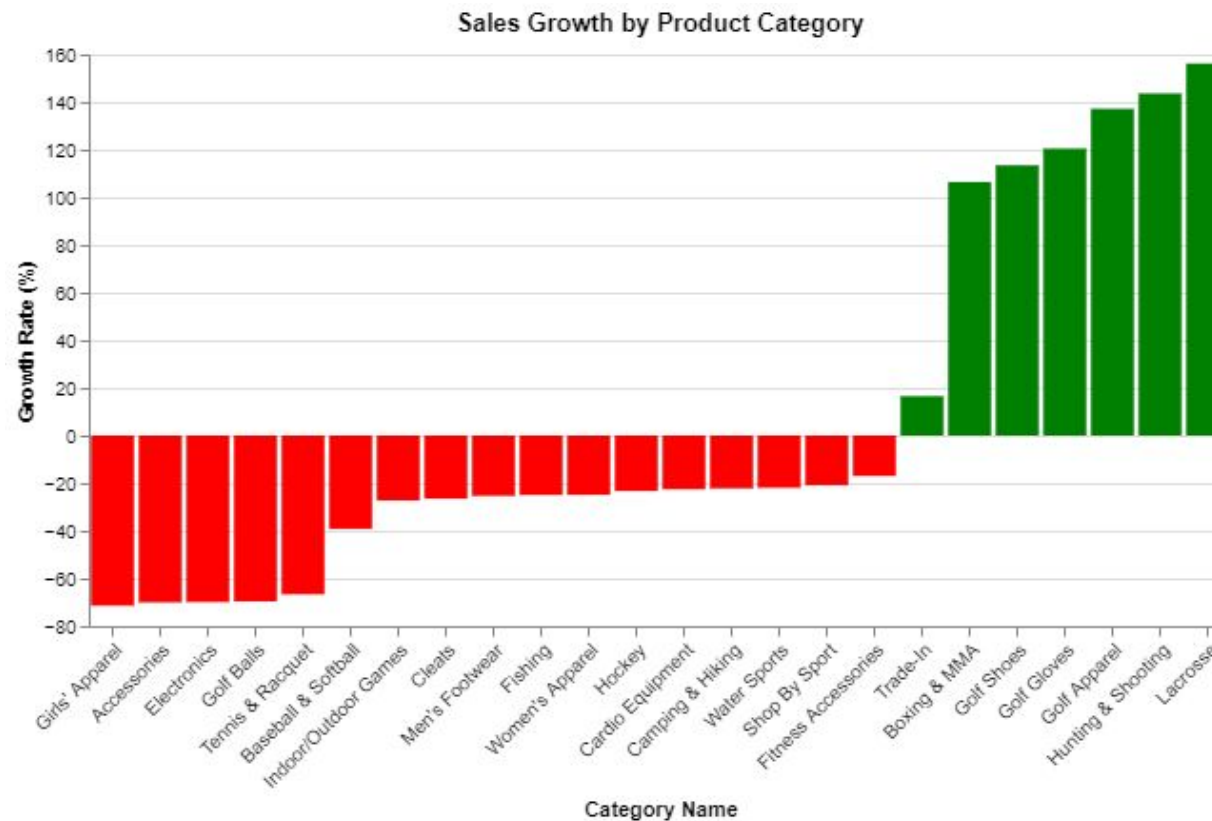
Data Analysis and Business Insights

- Products Performance Analysis
 - Visualization High-Performing and Under-Performing Products Across Markets



Data Analysis and Business Insights

- Products Performance Analysis
 - Visualization Sales Growth by Product Category



Data Analysis and Business Insights

- Products Performance Analysis

- Insights

- ❖ Africa: Fishing, Cleats, Cardio Equipment, Camping & Hiking, and Women's Apparel are the top-selling products. Golf Apparel, Hunting & Shooting, Boxing & MMA, Lacrosse, and Baseball & Softball are the least-selling products.
- ❖ Europe: Fishing, Computers, Cleats, Camping & Hiking, and Cardio Equipment are the top-selling products. Girls' Apparel, Accessories, Golf Balls, Electronics, and CDs are the least-selling products.
- ❖ Latin America: Fishing, Camping & Hiking, Cleats, Cardio Equipment, and Women's Apparel are the top-selling products. Fitness Accessories, Strength Training, Hockey, Golf Bags & Carts, and Tennis & Racquet are the least-selling products.
- ❖ Pacific Asia: Garden, Computers, Crafts, Music, and Health and Beauty are the top-selling products. Hockey, Tennis & Racquet, Fitness Accessories, Golf Apparel, and Trade-In are the least-selling products.
- ❖ USCA: Camping & Hiking, Cleats, Cardio Equipment, Men's Footwear, and Women's Apparel are the top-selling products. Golf Balls, Golf Gloves, Tennis & Racquet, Electronics, and Girls' Apparel are the least-selling products.
- ❖ The Boxing & MMA, Golf Apparel, Golf Gloves, Golf Shoes, Hunting & Shooting, Lacrosse, and Trade-In categories experienced positive sales growth between 2016 and 2017.
- ❖ The Accessories, Baseball & Softball, Camping & Hiking, Cardio Equipment, Cleats, Electronics, Fishing, Fitness Accessories, Girls' Apparel, Golf Balls, Hockey, Indoor/Outdoor Games, Men's Footwear, Shop By Sport, Tennis & Racquet, Water Sports, and Women's Apparel categories experienced negative sales growth between 2016 and 2017.

Data Analysis and Business Insights

- Products Performance Analysis
 - Recommendations
 - ❖ Focus on promoting and expanding the top-selling product categories in each market to capitalize on their popularity and maximize revenue.
 - ❖ Investigate the reasons for low sales in the bottom 5 product categories in each market. Consider product improvements, better marketing strategies, or discontinuing underperforming products to optimize the product mix.
 - ❖ Analyze the factors contributing to positive sales growth in the Boxing & MMA, Golf Apparel, Golf Gloves, Golf Shoes, Hunting & Shooting, Lacrosse, and Trade-In categories, and apply these strategies to other product categories to promote growth.
 - ❖ Address the issues leading to negative sales growth in various categories, such as product quality, pricing, marketing, or distribution channels, to improve sales performance and ensure overall business growth.

Data Analysis and Business Insights

- Hypothesis Testing on Shipping Modes
 - Objective:
 - ❖ To determine if there are significant differences in delivery efficiency (e.g., on-time delivery rate) across different shipping modes.
 - Hypothesis:
 - ❖ Null Hypothesis (H_0): There is no significant difference in on-time delivery rates among shipping modes.
 - ❖ Alternative Hypothesis (H_1): There is a significant difference in on-time delivery rates among shipping modes.
 - Method:
 - ❖ One-way ANOVA test is used in this case because it is a suitable method to test for significant differences among the means of multiple groups, while controlling for the increased risk of Type I errors that would arise from multiple pairwise comparisons.
 - Test Result:
 - ❖ Reject the null hypothesis: There is a significant difference in on-time delivery rates among shipping modes.

Data Analysis and Business Insights

- Hypothesis Testing on Shipping Modes
 - Insights:
 - ❖ The test result indicates that the on-time delivery rates differ significantly among the different shipping modes.
 - Recommendations:
 - ❖ Investigate the causes of the significant differences in on-time delivery rates among shipping modes.
 - ❖ Address inefficiencies in shipping modes with lower on-time delivery rates to improve overall delivery performance.
 - ❖ Prioritize shipping modes with higher on-time delivery rates and explore the possibility of replicating their success factors across other shipping modes.

Dashboard

DataCo Regional Performance Dashboard

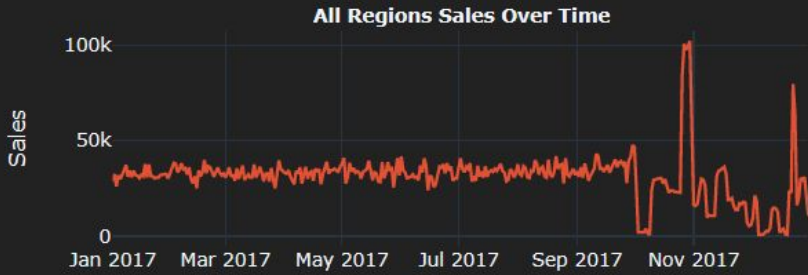
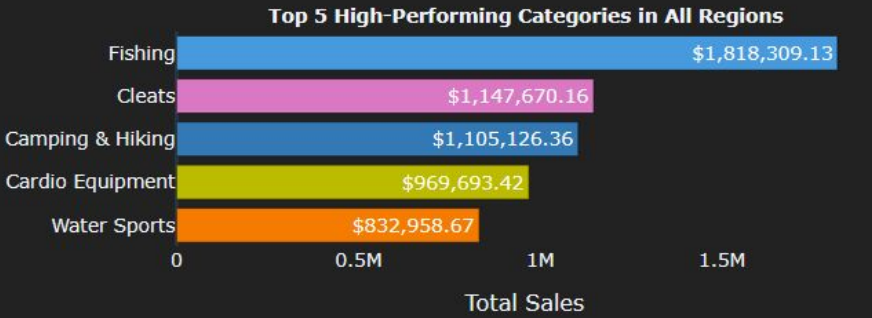
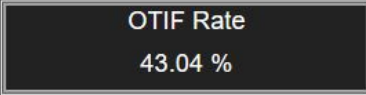
Select Year:

2017

Select Region:

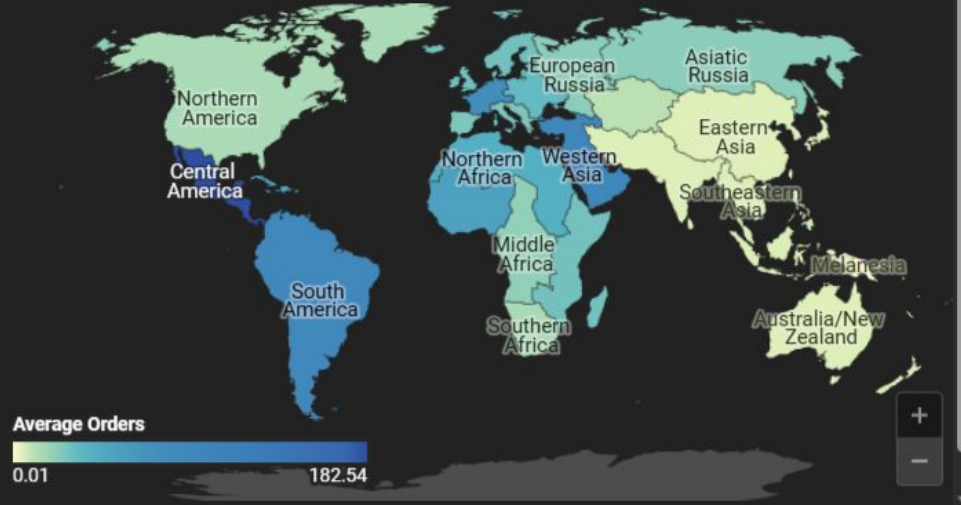
All Regions

Our dashboard offers a comprehensive exploration of sales, order volume, and delivery efficiency across diverse regions and markets. Through a range of visualizations, such as scorecards, choropleth maps, and bar charts, we can gain valuable insights on growth opportunities and areas for improvement. With tailored visualizations, we have the ability to filter data by year and region, allowing for in-depth analysis and a deep understanding of regional performance and supply chain optimization.

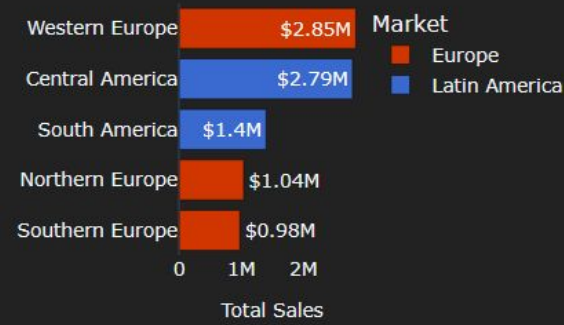


Daily Average Orders by Regions

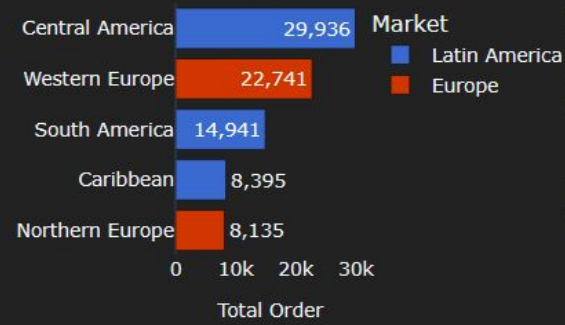
2017



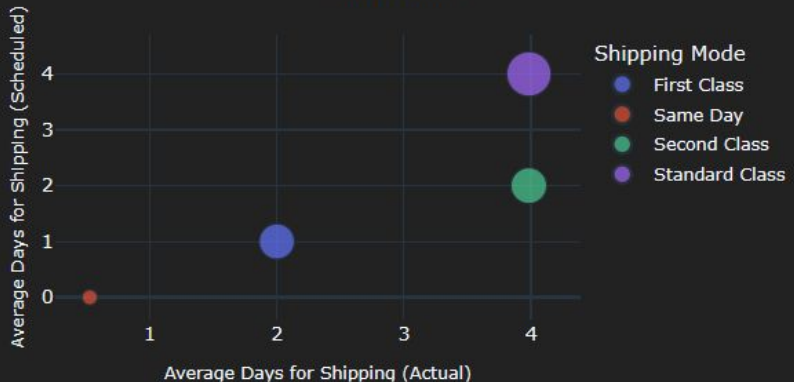
Top 5 High-Performing Regions by Total Sales



Top 5 High-Performing Regions by Total Order



Relationship between Shipping Days and Late Delivery Risk All Regions



Dashboard

The primary users of the dashboard will be supply chain managers, regional managers, and company executives responsible for strategic decision-making and performance evaluation. The dashboard will provide insights into regional sales performance, order volume, and delivery efficiency, enabling stakeholders to make informed decisions regarding market expansion, resource allocation, and process optimization.

Built on top of Dash and deployed on Render, this dashboard provides a comprehensive exploration of sales, order volume, and delivery efficiency across various regions and markets. Visit this link to try out the dashboard: <https://bit.ly/dataco-dashboard>

(Note: It may take a few minutes for the page to load initially, as I am using Render's free tier. The delay is not indicative of a broken site; please be patient and give it a moment to load)

Conclusion and Future Work

Conclusion

In conclusion, this analysis of the DataCo's data has provided valuable insights into sales performance, order trends, delivery efficiency, product performance, and shipping modes. Based on these insights, we propose the following recommendations and future work to improve the company's operations and enhance customer satisfaction:

- Sales, Order, and Shipping Mode Analysis:
 - Focus on high-performing markets and target marketing efforts during high-demand months to maximize sales growth.
 - Investigate reasons for low sales numbers in the USCA market and develop strategies to boost sales in this region.
 - Prioritize and allocate resources towards more efficient shipping modes based on the significant differences in on-time delivery rates found through hypothesis testing.
 -
- Delivery Efficiency Analysis:
 - Address issues causing the 0% on-time delivery rate for First Class shipping and improve the shipping performance for Second Class shipping.
 - Maintain the good performance of Standard Class shipping and aim to replicate faster-than-scheduled performance across all regions.
- Product Performance Analysis:
 - Promote top-performing products while identifying ways to improve sales of underperforming product categories.

Future Work

- Analyze customer demographics and preferences for personalized marketing campaigns and product recommendations.
- Investigate the impact of external factors on the company's sales and operations.
- Perform in-depth analysis of the supply chain to identify opportunities for optimization and cost reduction.

By addressing these recommendations and considering future work, the company can enhance its performance, improve customer satisfaction, and ensure a more reliable and efficient supply chain.

Reference

- Pacmann's Business Analytics Class Material
- Design and Analysis of Experiments - Montgomery, D. C. (2017)

Thank You

Alvin Noza P.H