

Analysis & Report

Summary

This Jupyter Notebook demonstrates data analysis using Python with an integrated SQL workflow, retrieves data using pandas with SQL queries, and processes it for insights. Key steps include:

- Establishing a connection to a relational database Data warehouse.
- Executing SQL queries within Python to extract customer and sales data.
- Performing exploratory data analysis (EDA) on sales trends, cost distribution, and customer behavior.
- Visualizing insights with plots to support data-driven decisions.

Overall, this notebook provides a complete workflow — from database extraction to data cleaning, analysis, and visualization — highlighting practical techniques for business analytics.

Problem Statement

Businesses generate large volumes of customer, product, and sales data, but raw data alone does not provide actionable insights. To make informed decisions, companies need to understand customer behavior, product performance, and overall sales trends.

This analysis aims to address key business questions, such as:

- Who are the most valuable customers?
- How do sales and order trends change over time?
- Which products drive the most revenue and customer engagement?
- What is the distribution of costs and spending patterns?
- How can customers be segmented based on their purchase and product preferences?

By integrating SQL queries with Python-based data analysis, this notebook provides a structured approach to transform raw transactional data into meaningful insights on both customers and products, enabling data-driven business decisions.

Objectives

The main objectives of this analysis are:

- 1. **Database Integration**
 - Connect to a relational database and extract customer, product, and sales data using SQL queries.
- 2. **Data Cleaning & Transformation**
 - Prepare the extracted datasets for analysis by handling missing values, formatting fields, and creating derived metrics.
- 3. **Exploratory Data Analysis (EDA)**
 - Analyze customer purchasing behavior, product performance, sales trends, and cost distribution.
- 4. **Segmentation & KPIs**
 - Segment customers into categories (e.g., VIP, regular, new) and calculate key metrics such as total sales, order frequency, recency, and average order value.
- 5. **Visualization & Insights**
 - Build visualizations to highlight patterns and trends, making insights clear and actionable for business decision-making.

Through these objectives, the notebook aims to demonstrate a complete workflow that converts raw customer and product data into meaningful insights for business intelligence.

Setup & Configuration

View Dataframes

	customer_key	customer_id	customer_number	first_name	last_name	country	\
0	1	11000	AW00011000	Jon	Yang	Australia	
1	2	11001	AW00011001	Eugene	Huang	Australia	
2	3	11002	AW00011002	Ruben	Torres	Australia	
3	4	11003	AW00011003	Christy	Zhu	Australia	
4	5	11004	AW00011004	Elizabeth	Johnson	Australia	

	marital_status	gender	birthdate	create_date
0	Married	Male	1971-10-06	2025-10-06
1	Single	Male	1976-05-10	2025-10-06
2	Married	Male	1971-02-09	2025-10-06
3	Single	Female	1973-08-14	2025-10-06
4	Single	Female	1979-08-05	2025-10-06

	product_key	product_id	product_number	product_name	category_id	\
0	1	601	BB-7421	LL Bottom Bracket	CO_BB	
1	2	602	BB-8107	ML Bottom Bracket	CO_BB	
2	3	603	BB-9108	HL Bottom Bracket	CO_BB	
3	4	478	BC-M005	Mountain Bottle Cage	AC_BC	
4	5	479	BC-R205	Road Bottle Cage	AC_BC	

	category	subcategory	maintenance	cost	product_line	start_date
0	Components	Bottom Brackets	Yes	24	Unknown	2013-07-01
1	Components	Bottom Brackets	Yes	45	Unknown	2013-07-01
2	Components	Bottom Brackets	Yes	54	Unknown	2013-07-01
3	Accessories	Bottles and Cages	No	4	Mountain	2013-07-01
4	Accessories	Bottles and Cages	No	3	Road	2013-07-01

	order_number	product_key	customer_key	order_date	shipping_date	\
0	S043697	80	10769	2010-12-29	2011-01-05	
1	S043698	36	17390	2010-12-29	2011-01-05	
2	S043699	36	14864	2010-12-29	2011-01-05	
3	S043700	47	3502	2010-12-29	2011-01-05	
4	S043701	36	4	2010-12-29	2011-01-05	

	due_date	sales_amount	quantity	price
0	2011-01-10	3578	1	3578.0
1	2011-01-10	3400	1	3400.0
2	2011-01-10	3400	1	3400.0
3	2011-01-10	699	1	699.0
4	2011-01-10	3400	1	3400.0

Exploratory Data Analysis

◆ Dimensions Exploration

1. Explore All Customer Countries

Goal: Identify all unique countries where customers are located to understand geographic coverage.

Chart: None (tabular output)

EDA Type: Categorical (geographic dimension)

Structure: Display a distinct list of countries from the customer dataset.

Insight: Customers are spread across 6, highlighting an international customer base.

2. Explore All Product Categories ("Major Divisions")

Goal: Understand the product hierarchy by exploring unique combinations of category, subcategory, and product names.

Chart: None (tabular output)

EDA Type: Categorical (product hierarchy)

Structure: Remove duplicate rows based on category, subcategory, and product name to identify distinct product entries.

	category	subcategory	product_name
240	Accessories	Bike Racks	Hitch Rack - 4-Bike
277	Accessories	Bike Stands	All-Purpose Bike Stand
3	Accessories	Bottles and Cages	Mountain Bottle Cage
4	Accessories	Bottles and Cages	Road Bottle Cage
294	Accessories	Bottles and Cages	Water Bottle - 30 oz.
..
230	None	None	LL Mountain Pedal
233	None	None	LL Road Pedal
231	None	None	ML Mountain Pedal
234	None	None	ML Road Pedal
236	None	None	Touring Pedal

[295 rows x 3 columns]

Insight: Total unique products in Products dataset: 295

◆ Date Exploration

3. First and Last Order Dates (Time Range)

Goal: Determine the time range of available sales data to define the analysis period.

Chart: None (date output)

EDA Type: Temporal (time range identification)

Structure: Extract the earliest (first) and latest (last) order dates from the dataset after removing null values.

Insight: We have sales data spans from 2010-12-29 to 2014-01-28, covering ~3.1 years.

4. Oldest and Youngest Customer Ages

Goal: Understand the age distribution of the customer base by identifying the oldest and youngest customers.

Chart: None (numeric output)

EDA Type: Demographic (age analysis)

Structure: Extract minimum and maximum birthdates from the dataset, convert them to ages, and calculate the oldest and youngest customer ages in years.

- Insight:
- The oldest customer is 109 years old (born on 1916-02-10).
 - The youngest customer is 39 years old (born on 1986-06-25).

◆ Measure Exploration

5. Tabular Report of Key Business Measures

Goal: Summarize essential business metrics to establish a high-level overview of sales, products, and customers.

Chart: None (tabular KPI report)

EDA Type: Descriptive (summary statistics)

Structure: Calculate key measures including:

- Total sales revenue
- Total items sold
- Average selling price
- Total and distinct orders
- Total and distinct products
- Total and distinct customers (including those who placed orders)

Out[11]:

	Measure	Value
0	Total Sales	29,356,250
1	Total Items Sold	60,423
2	Average Selling Price	486.09
3	Total Orders	27,659
4	Total Products	295
5	Total Customers	18,484
6	Customers Who Placed Orders	18,484

Insight: This report consolidates the core KPIs of the business, providing a snapshot of overall sales performance, product variety, and customer engagement.

EDA Visualizations

◆ Magnitude Analysis

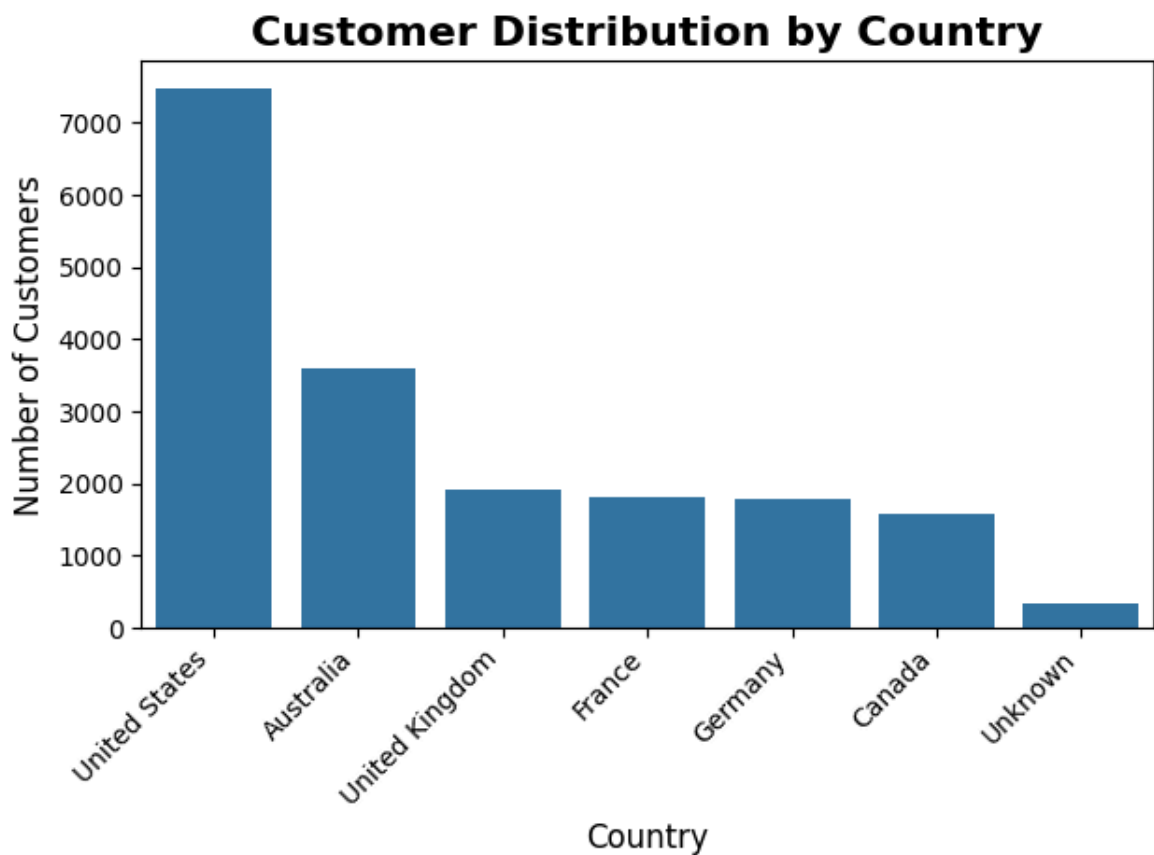
6. Customer Distribution by Country

Goal: Analyze how customers are distributed geographically to identify major markets.

Chart: Bar chart

EDA Type: Categorical (geographic distribution)

Structure: Group customers by country, count distinct customers, and visualize the distribution with a bar chart. Value labels are displayed above each bar for clarity.



Insight:

- United States leads with 7,482 customers, the highest among all countries.
- The dataset includes customers from 6 distinct countries.
- A total of 18,484 customers are distributed across all countries.
- Note: 337 customers have an 'Unknown' country, indicating potential data quality issues.

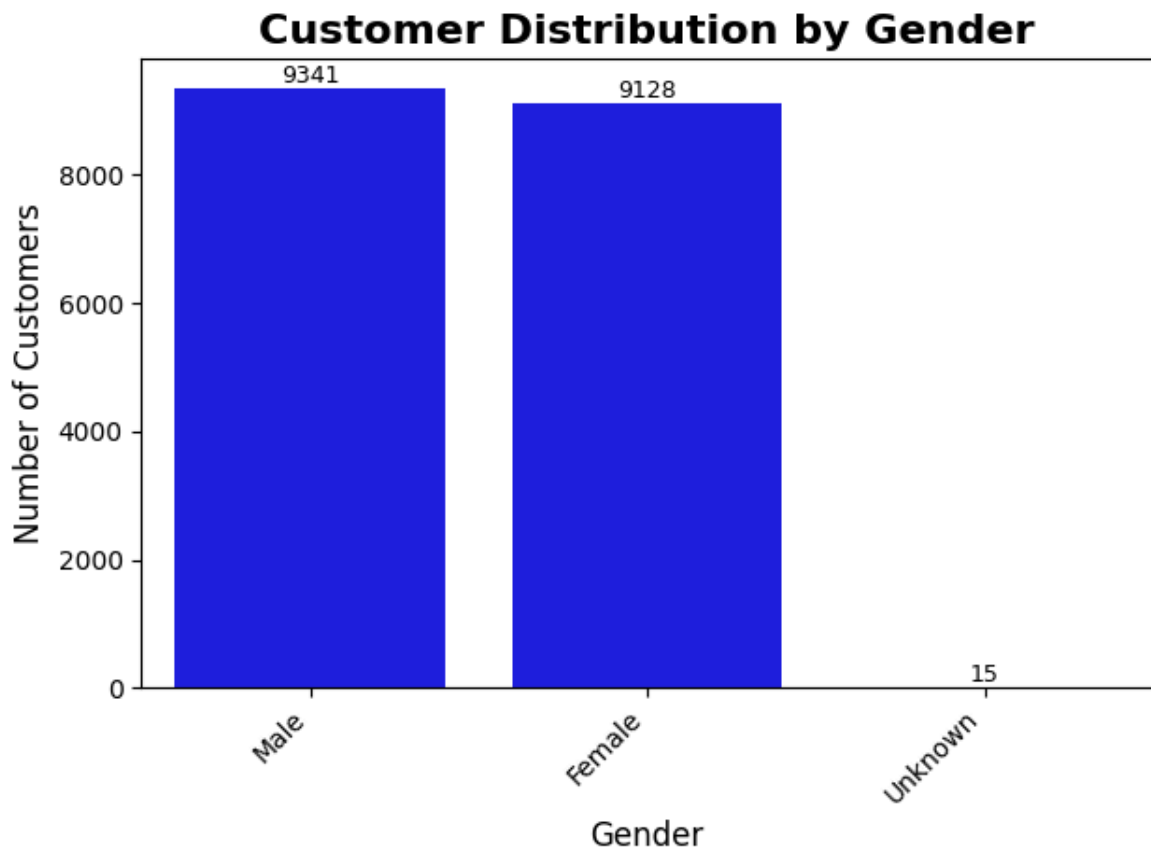
7. Customer Distribution by Gender

Goal: Understand the demographic breakdown of customers by gender to identify dominant customer segments.

Chart: Bar chart

EDA Type: Categorical (demographic distribution)

Structure: Group customers by gender, count the number of customers in each group, and sort in descending order. A bar chart is used to visualize the distribution, with value labels placed above each bar for clarity.



Gender Distribution Insights:

- Male is the dominant gender category with 9,341 customers (50.5% of all customers).
- Female is the second most common with 9,128 customers.
- The ratio of Male to Female customers is 1:1
- The dataset contains 3 distinct gender categories.
- Total of 18,484 customers are distributed across all gender categories.
- Note: 15 customers (0.1%) have unknown/unspecified gender data.
- The gender distribution is relatively balanced between the top categories.

8. Products Distribution by Category

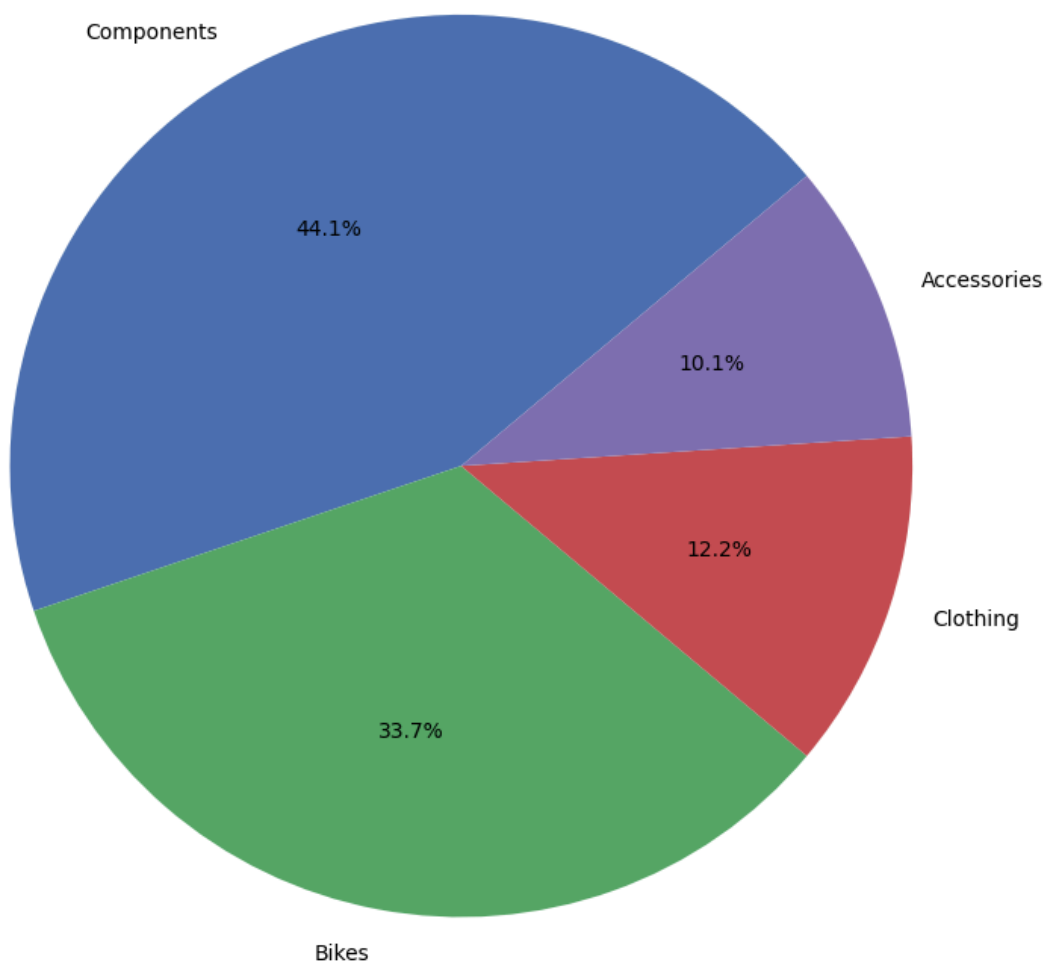
Goal: Analyze the spread of products across different categories to identify which categories dominate the product portfolio.

Chart: Pie chart

EDA Type: Categorical (category-level distribution)

Structure: Group products by category, count the total products in each, and visualize the proportion of categories using a pie chart with percentage labels for clarity.

Products Distribution by Category



Product Category Distribution Insights:

- Components dominates the product catalog with 127 products (44.1% of all products).
- Bikes is the second largest category with 97 products (33.7%).
- Accessories has the smallest representation with 29 products (10.1%).
- The product portfolio spans 4 distinct categories with a total of 288 products.
- The top 3 categories account for 259 products (89.9% of the total catalog).
- The distribution is moderately concentrated, indicating focused specialization.

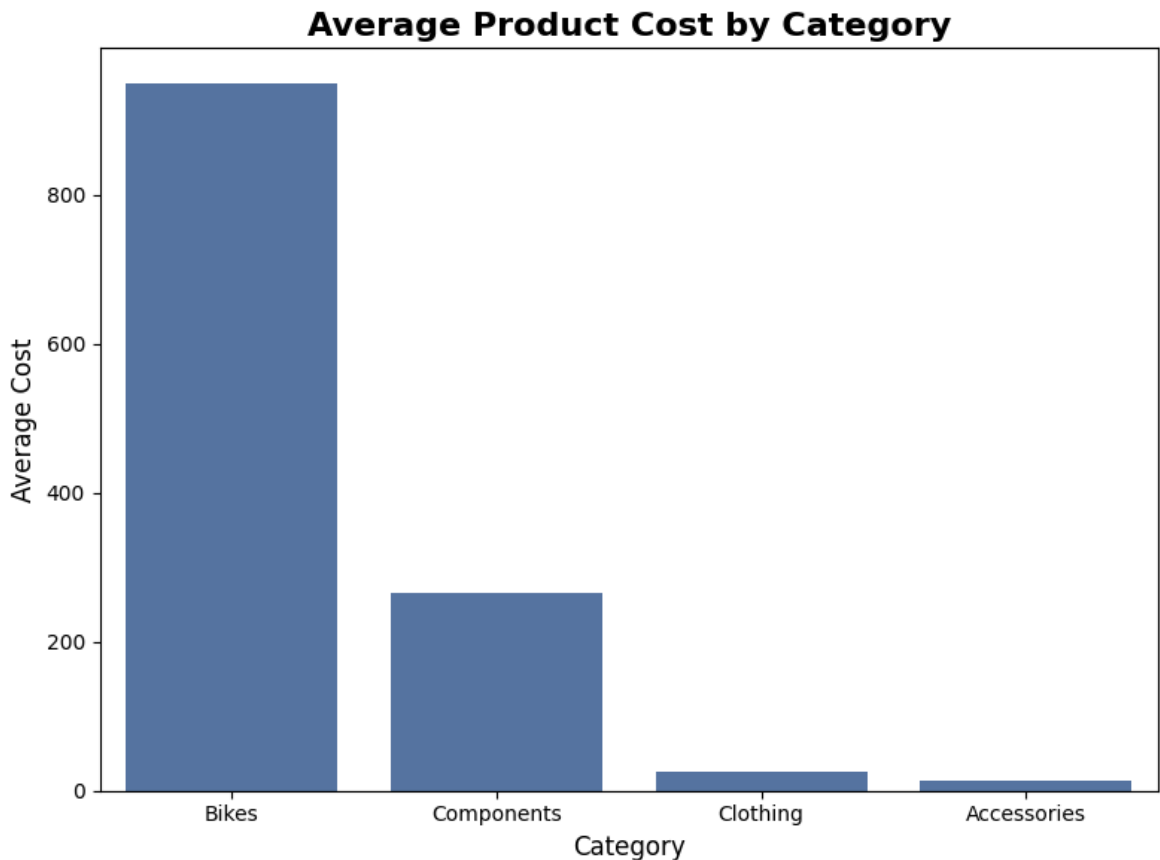
9. Average Product Cost by Category

Goal: Compare the average cost of products across different categories to identify higher-value and lower-value product groups.

Chart: Bar chart

EDA Type: Numerical (mean cost comparison by category)

Structure: Group products by category, calculate the mean product cost within each category, and sort them in descending order. A bar chart is used to visualize and compare the average cost across categories.



Key Insights:

- Most expensive: Bikes (\$949.44)
- Most affordable: Accessories (\$13.17)
- Overall average: \$313.03
- Price difference: 72.1x between highest and lowest categories

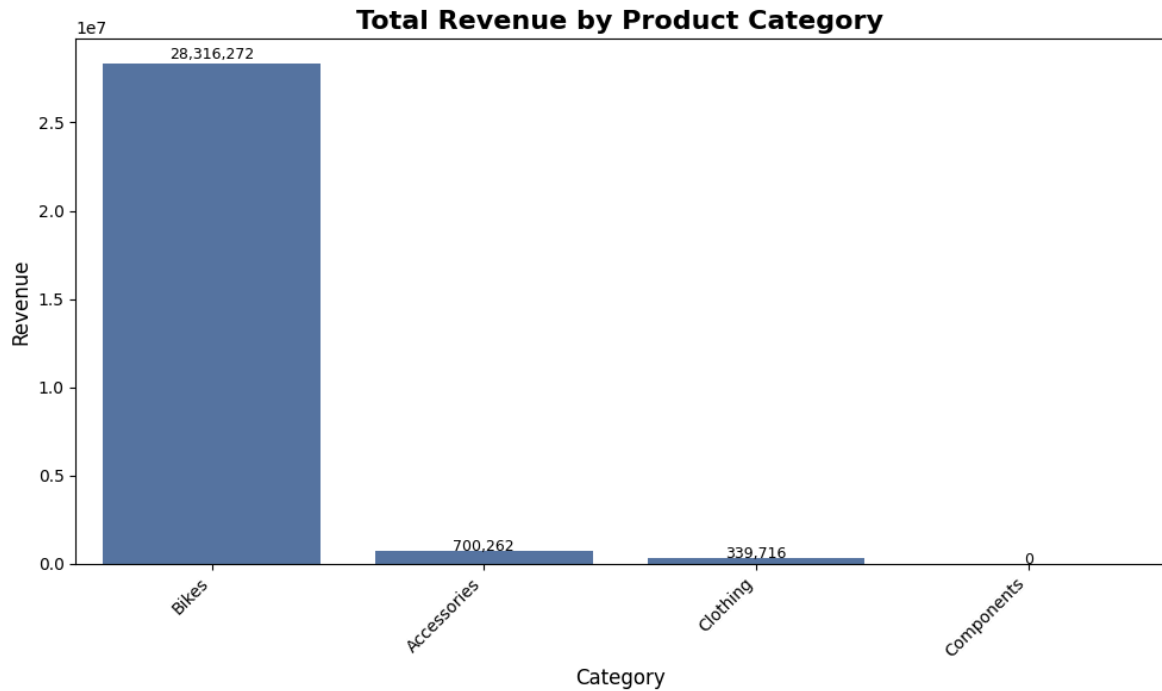
10. Total Revenue by Product Category

Goal: Evaluate which product categories generate the highest revenue to identify key revenue drivers.

Chart: Bar chart

EDA Type: Numerical (aggregated revenue by category)

Structure: Merge products and sales data on `product_key` using a left join, then group by product category and calculate the total revenue (`sum` of sales amount). Sort the results in descending order and visualize them with a bar chart, displaying value labels above each bar for clarity.

**Key Insights:**

- Top performer: Bikes (\$28,316,272 - 96.5%)
- Lowest performer: Components (\$0)
- Total revenue: \$29,356,250 across 4 categories

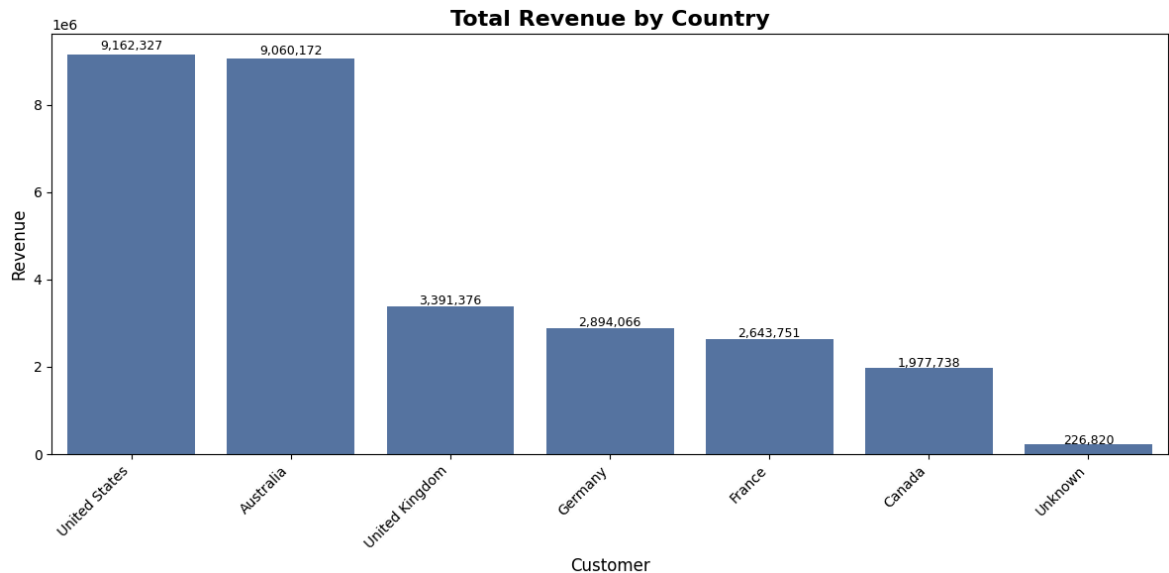
11. Total Revenue by Country

Goal: Identify the geographic distribution of revenue to highlight which countries contribute most to overall sales.

Chart: Bar chart

EDA Type: Numerical + Categorical (revenue aggregated by geographic category)

Structure: Merge customers with sales data on `customer_key` using a left join, then group by country and calculate the total revenue (`sum` of sales amount). Sort the countries in descending order of revenue and visualize with a bar chart, including value labels above each bar for better readability.

**Key Insights:**

- Top market: United States (\$9,162,327 - 31.2% of total)
- Second market: Australia (\$9,060,172 - 30.9% of total)
- Geographic reach: 7 countries generating \$29,356,250 total revenue
- Average per country: \$4193750.00
- Top 5 markets: \$27,151,692 (92.5% of total revenue)
- Strong markets (1.5x average): 2 countries above \$6290625.00
- Market range: 40.4x difference between strongest and weakest markets
- Growth opportunity: Limited strong markets suggest expansion potential
- Market performance: Range from \$226820.00 (Unknown) to \$9162327.00 (United States)

◆ Ranking Analysis

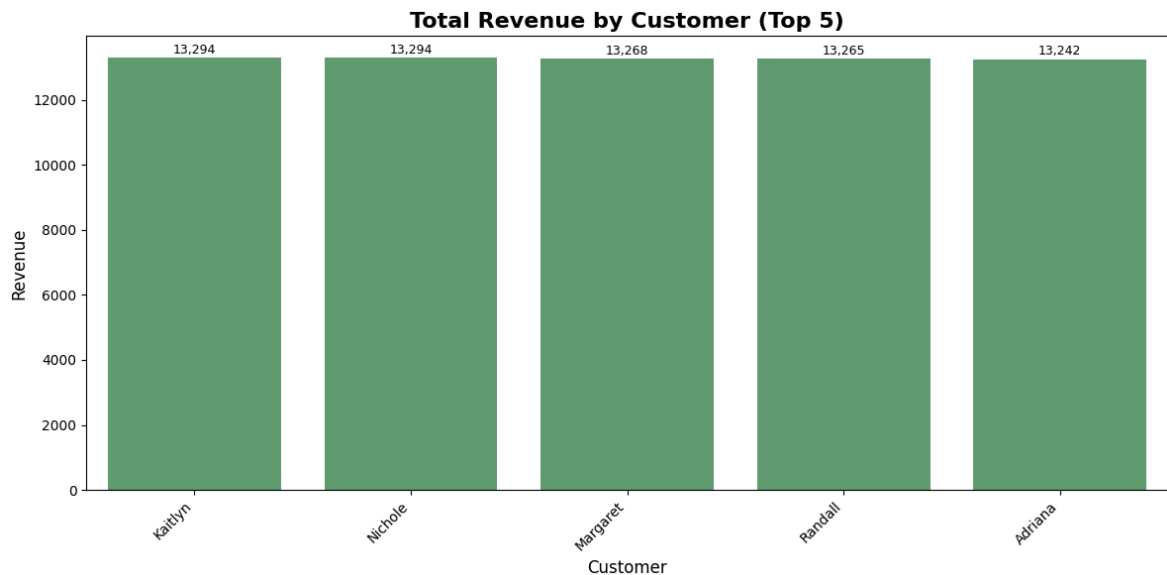
12. Total Revenue by Customer (Top 5)

Goal: Identify the top individual customers contributing the highest revenue to analyze customer-level value.

Chart: Bar chart

EDA Type: Numerical + Categorical (revenue aggregated at customer level)

Structure: Group sales data by customer (`customer_key` , `first_name` , `last_name`) and calculate total revenue (`sum` of sales amount). Sort customers in descending order of revenue and plot the top 5 with a bar chart, adding value labels above each bar for clarity.

**Key Insights:**

- Top customer: Kaitlyn Henderson (\$13,294 - 0.0% of total)
- Second place: Nichole Nara (\$13,294 - 0.0% of total)
- Top 5 contribution: \$66,363 (0.2% from 18484 customers)
- Average customer revenue: \$1588.20
- VIP customers (3x average): 1933 customers generating above \$4764.59
- Revenue concentration: 5155 customers (27.9%) generate 80% of revenue
- Revenue diversification: Well-distributed customer base reduces concentration risk
- Performance range: 1.0x difference between 1st and 5th place customers
- Total customer base: 18484 customers

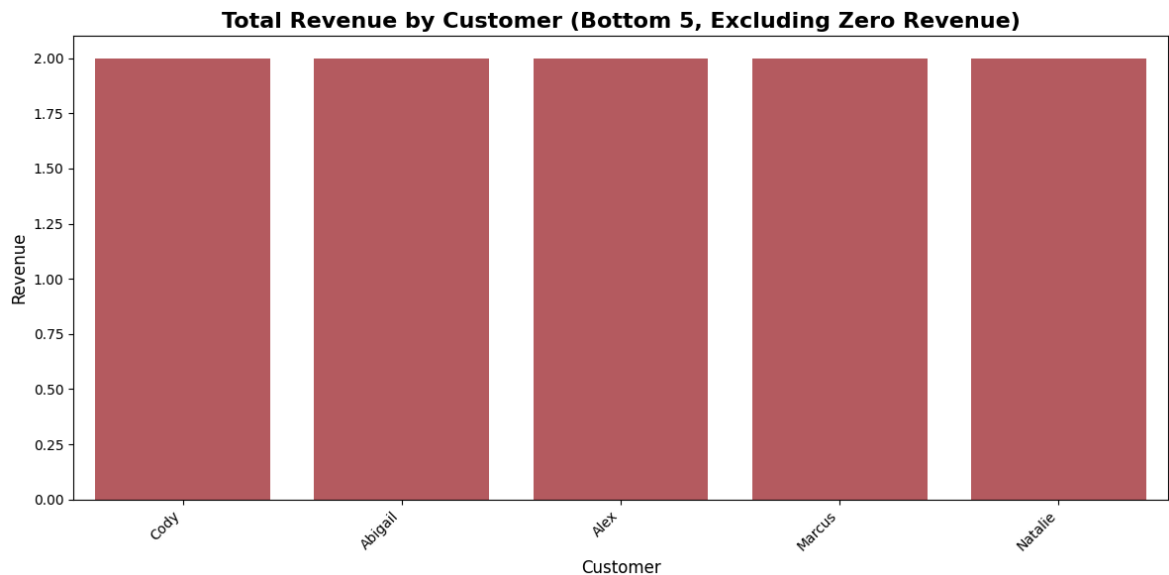
13. Total Revenue by Customer (Bottom 5, Excluding Zero Revenue)

Goal: Highlight the lowest revenue-generating customers (excluding those with zero revenue) to analyze weak or underperforming customer segments.

Chart: Bar chart

EDA Type: Numerical + Categorical (revenue aggregated at customer level)

Structure: Filter out customers with zero revenue, then sort remaining customers in ascending order by total revenue. Select the bottom 5 customers and visualize with a bar chart to compare their contribution levels.

**Key Insights:**

- Lowest performer: Cody Sanders (\$2.00)
- Second lowest: Abigail Morris (\$2.00)
- Bottom 5 contribution: \$10.00 (0.00% of total revenue)
- Average customer revenue: \$1588.20
- Performance gap: Bottom customer generates 794.1x less than average
- Low-value customers: 9306 customers (50.3%) below \$397.05
- Bottom 10% segment: 1848 customers generate \$20781.00 (0.1% of revenue)
- Customer development opportunity: Large segment of low-value customers suggests potential for engagement programs
- Revenue impact: Bottom customers have minimal revenue contribution but represent relationship-building opportunities
- Total active customers: 18484 generating revenue

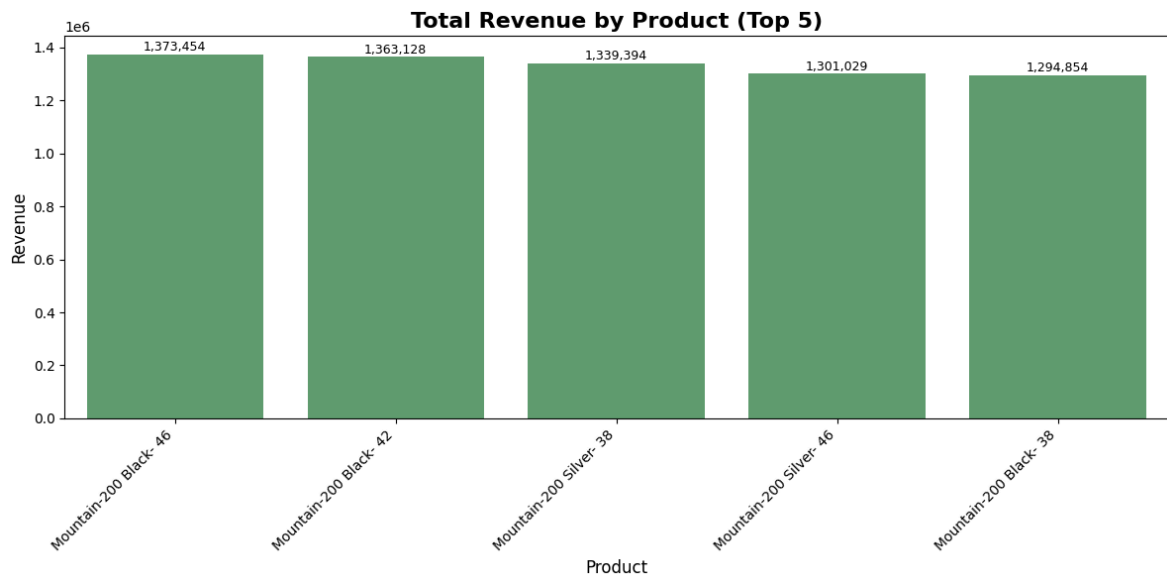
14. Total Revenue by Product (Top 5)

Goal: Identify the top-performing products that contribute the highest revenue to the business.

Chart: Bar chart

EDA Type: Numerical + Categorical (revenue aggregated at product level)

Structure: Group sales data by product name and calculate total revenue (`sum` of sales amount). Sort products in descending order of revenue and plot the top 5 using a bar chart, with value labels above each bar for clarity.

**Key Insights:**

- Top performer: Mountain-200 Black- 46 (\$1,373,454 - 4.7% of total)
- Second place: Mountain-200 Black- 42 (\$1,363,128 - 4.6% of total)
- Top 5 contribution: \$6,671,859 (22.7% from 295 products)
- Average product revenue: \$99,513
- Star products (2x average): 38 products generating above \$199,025
- Revenue concentration: 35 products (11.9%) generate 80% of revenue
- Portfolio shows good diversification - revenue well distributed across products
- Performance range: 1.1x difference between 1st and 5th place

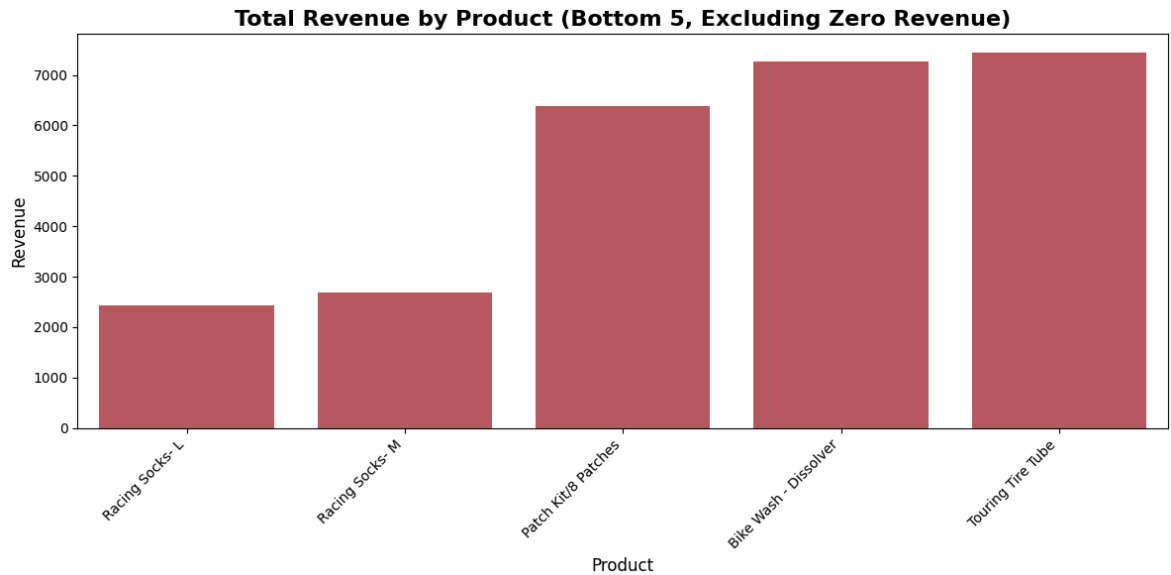
15. Total Revenue by Product (Bottom 5, Excluding Zero Revenue)

Goal: Identify the lowest revenue-generating products (excluding those with zero sales) to analyze underperforming items in the product portfolio.

Chart: Bar chart

EDA Type: Numerical + Categorical (revenue aggregated at product level)

Structure: Group sales data by product name and calculate total revenue (`sum` of sales amount). Filter out products with zero revenue, then sort in ascending order and select the bottom 5. Visualize these products with a bar chart to highlight their limited contribution.

**Key Insights:**

- Lowest performer: Racing Socks- L (\$2430.00)
- Bottom 5 contribute: \$26206.00 (0.09% of total revenue)
- Total active products: 130 generating revenue

◆ Change over Time

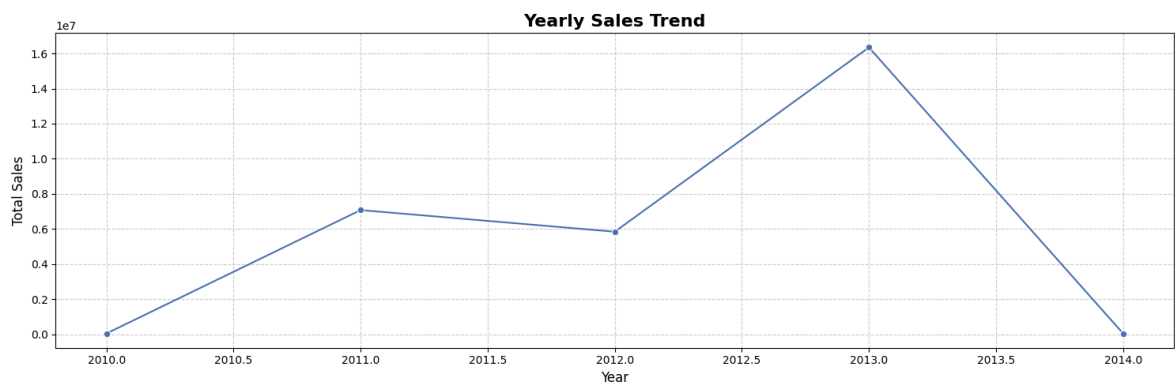
16. Yearly Sales Trend

Goal: Analyze sales performance over the years to identify long-term growth patterns or declines.

Chart: Line chart

EDA Type: Time series (yearly aggregated sales)

Structure: Convert `order_date` to datetime, extract the year, and group sales data by year. Aggregate total sales (`sum` of sales amount) for each year and plot a line chart with markers to visualize year-over-year sales trends.



Yearly Sales Trend Insights:

- Peak performance: 2013 with \$16,344,878 in sales.
- Lowest performance: 2010 with \$43,419 in sales.
- Overall growth: +5.1% from 2010 to 2014.
- Best year-over-year growth: 2011 with +16194.9% increase.
- Steepest decline: 2014 with -99.7% decrease.
- Recent trend: strong upward trajectory in sales performance.

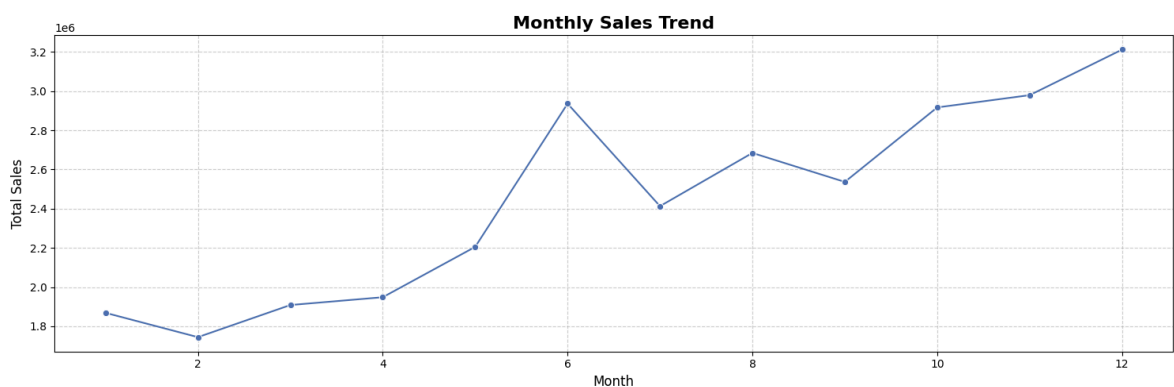
17. Monthly Sales Trend

Goal: Examine sales patterns across months to detect seasonality and recurring demand cycles.

Chart: Line chart

EDA Type: Time series (monthly aggregated sales)

Structure: Convert `order_date` to datetime, extract the month component, and group sales data by month. Aggregate total sales (`sum` of sales amount) for each month and plot a line chart with markers to visualize month-to-month sales variation.



Monthly Sales Trend Insights:

- Peak month: Dec with \$3,211,396 in sales.
- Lowest month: Feb with \$1,744,517 in sales.
- Seasonal variation: 1.8x difference between peak and low months.
- Strongest quarter: Q4 (\$9,107,059) vs weakest Q1 (\$5,521,450).
- Business shows moderate seasonality with clear monthly patterns.

◆ Cumulative Analysis

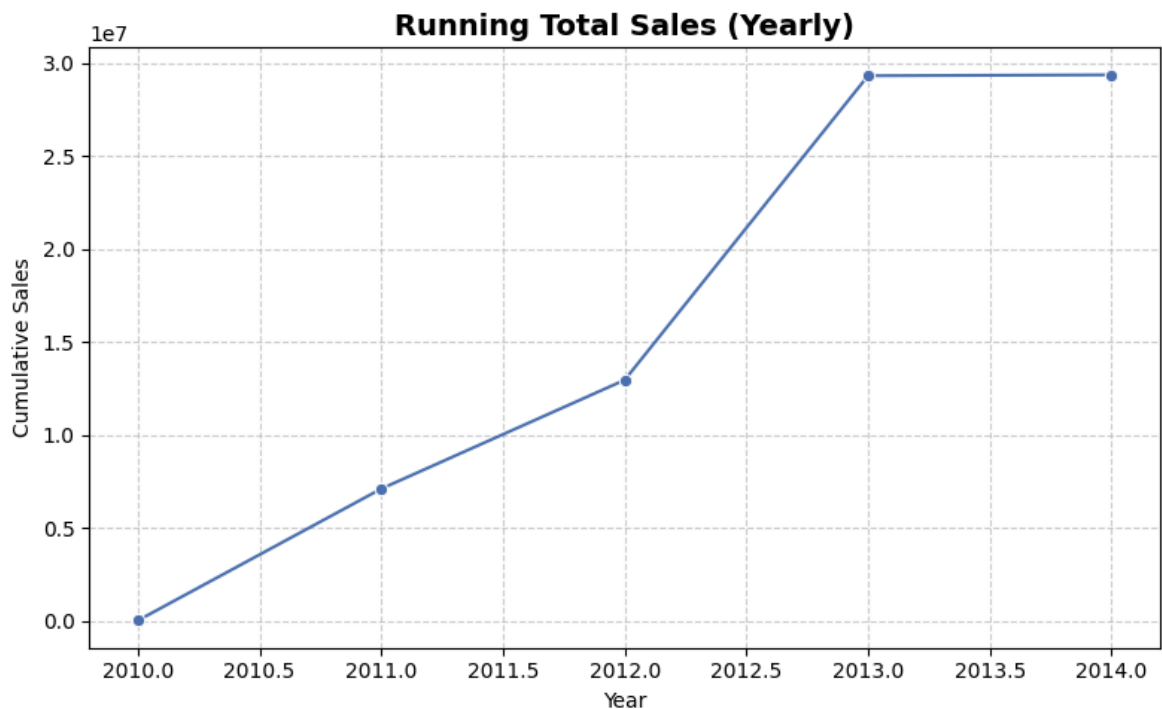
18. Running Total Sales (Yearly)

Goal: Measure cumulative sales growth over time to evaluate long-term performance and overall revenue progression.

Chart: Line chart

EDA Type: Time series (yearly cumulative sales)

Structure: Extract the year from `order_date` , aggregate total sales (`sum` of sales amount) per year, and compute the cumulative running total using `cumsum()` . Plot the cumulative sales trend across years with a line chart to highlight overall revenue growth trajectory.



Running Total Sales (Yearly) Insights:

- Cumulative sales reached \$29,351,258 over 4.0 years (2010-2014).
- Best performing year: 2013 contributed \$16,344,878.
- Lowest performing year: 2010 contributed \$43,419.
- Major milestone: Reached \$10,000,000 in cumulative sales by 2012.
- Growth pattern: Business shows accelerating growth trajectory.
- Average annual contribution: \$5,870,252 per year.

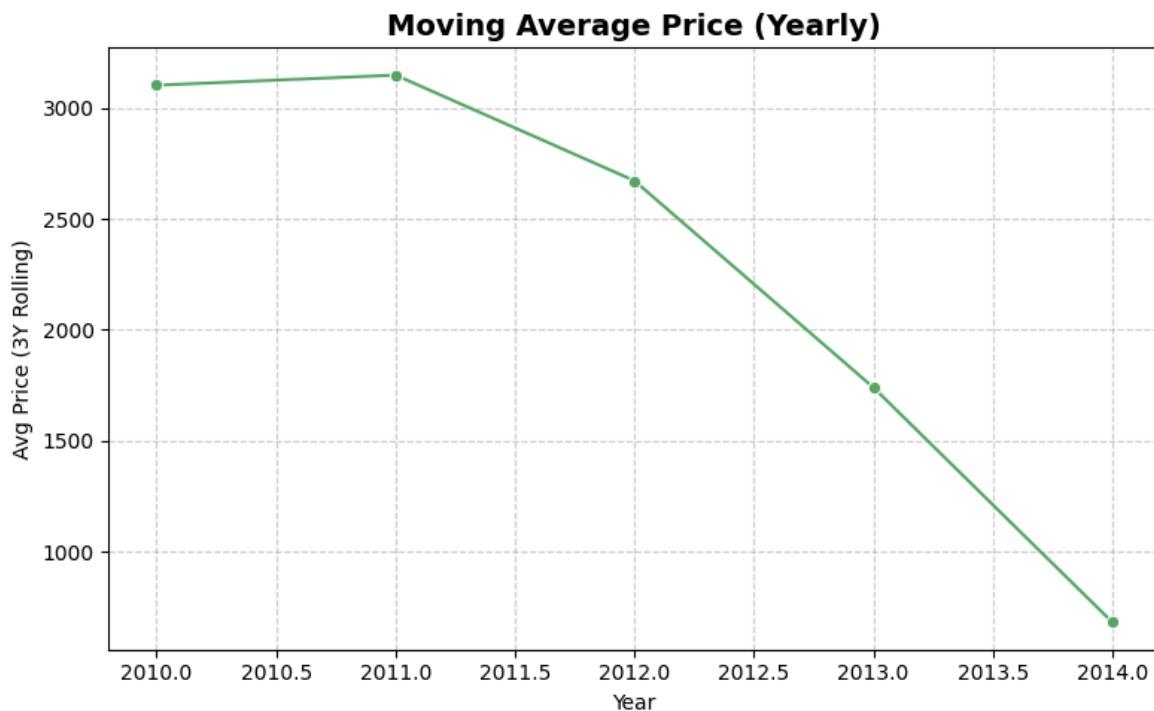
19. Moving Average Price (Yearly)

Goal: Smooth out yearly fluctuations in product prices to observe long-term pricing trends.

Chart: Line chart

EDA Type: Time series (rolling average of yearly prices)

Structure: Aggregate the average product price (`mean` of price) by year, then compute a 3-year rolling moving average to reduce volatility and highlight overall pricing trends. Plot the moving average as a line chart with markers for clarity.



Moving Average Price (Yearly) Insights:

- Price evolution: \$3101.36 (2010) to \$684.23 (2014), -77.9% change.
- Peak pricing: \$3147.04 in 2011, lowest at \$684.23 in 2014.
- Price volatility: 46.3% coefficient of variation over the period.
- Recent trend: Prices are declining based on 3-year moving average.
- Pricing strategy shows volatile patterns with smoothed fluctuations.

◆ Performance Analysis

20. Product Sales Segmentation Across Years

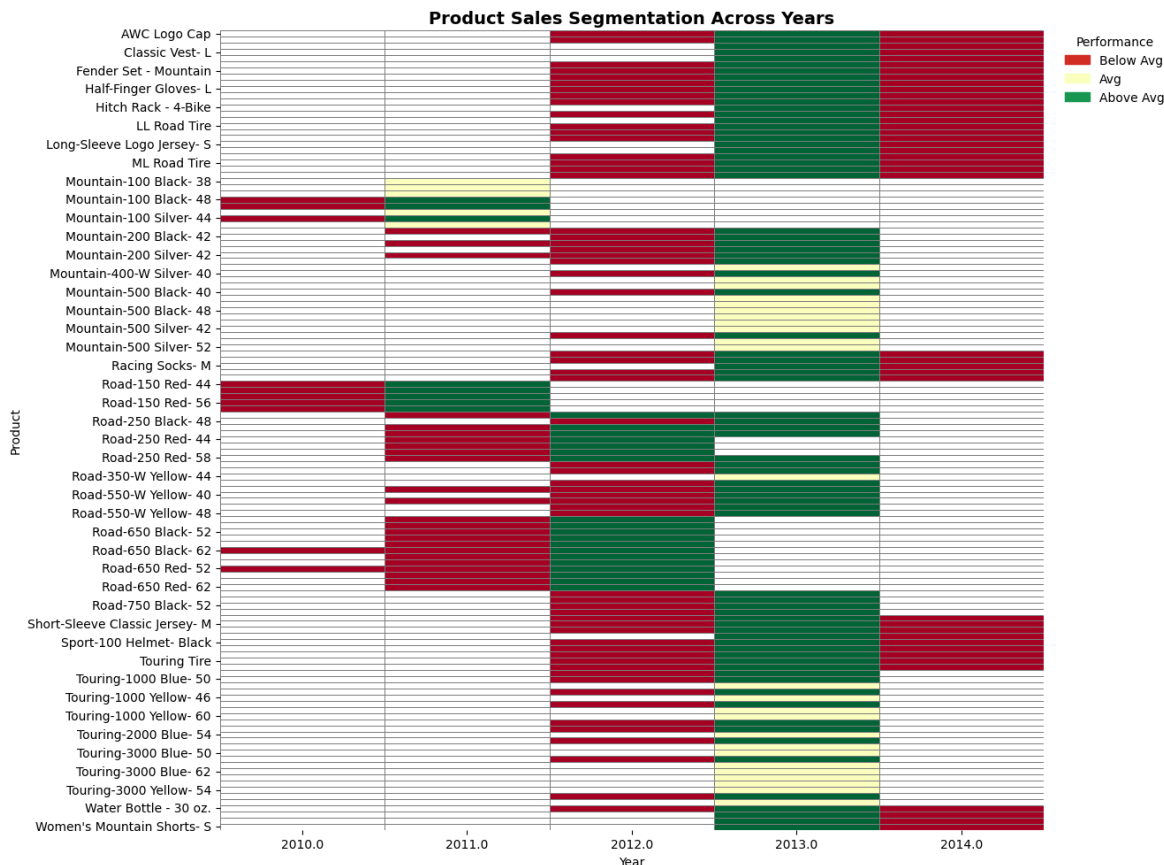
Goal: Compare each product's yearly sales performance against its own historical average to identify products performing above or below expectations.

Chart: Heatmap

EDA Type: Time series + Categorical (product-level performance segmentation)

Structure:

1. Merge sales and product data on `product_key`.
2. Convert `order_date` to datetime and extract year.
3. Aggregate yearly sales for each product.
4. Compute the average sales per product across all years.
5. Calculate the difference between yearly sales and average sales, classifying results as *Above Avg*, *Avg*, or *Below Avg*.
6. Visualize the segmentation using a heatmap with a custom legend indicating performance categories.



Product Sales Segmentation Heatmap Insights:

- Performance distribution: 104 above average (38.4%), 137 below average (50.6%).
- Best performing year: 2013 with 75.5% of products performing above average.
- Strong performers: 4 products consistently exceed average (>60% of years).
- Underperformers: 37 products consistently fall below average (>60% of years).
- Performance variability: 89 products show mixed year-to-year performance.
- Portfolio insight: Heatmap reveals 130 products across 5 years with clear performance patterns.

◆ Part to Whole Analysis

21. Sales Distribution by Category

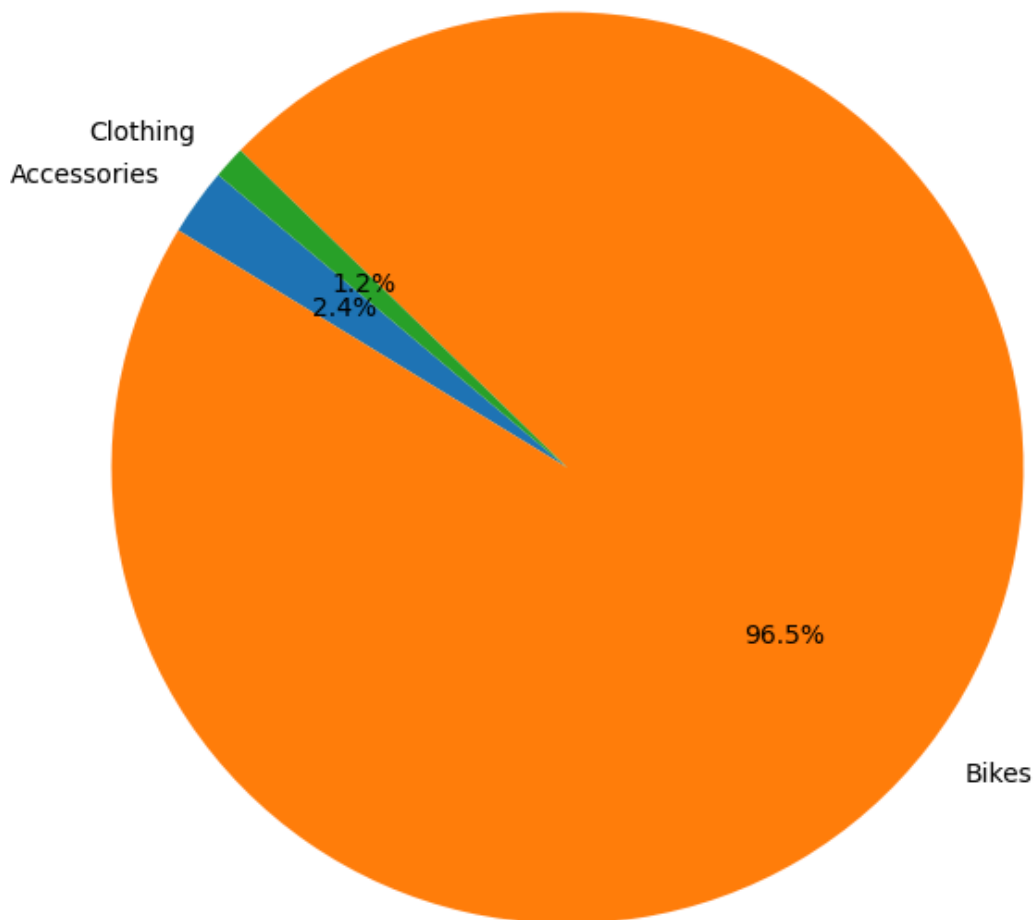
Goal: Analyze the proportion of total sales contributed by each product category to identify dominant categories in overall revenue.

Chart: Pie chart

EDA Type: Categorical + Numerical (sales aggregated by category)

Structure: Group sales data by product category, calculate total sales (`sum` of sales amount), and visualize the contribution of each category as percentages in a pie chart.

Sales Distribution by Category



Key Insights:

- Dominant category: Accessories (2.4% of total sales)
- Second largest: Bikes (96.5% of total sales)
- Market composition: 3 categories with \$29,356,250 total sales
- Average per category: \$9,785,417
- Top 3 categories: 100.0% of total market share
- Major categories (>15% share): 1 categories contributing 96.5%
- Market structure: Balanced distribution across categories
- Category range: From 1.2% (Clothing) to 2.4% (Accessories)
- Performance gap: 2.1x difference between largest and smallest categories
- Portfolio focus: Limited major categories suggest concentrated business model

◆ Data Segmentation

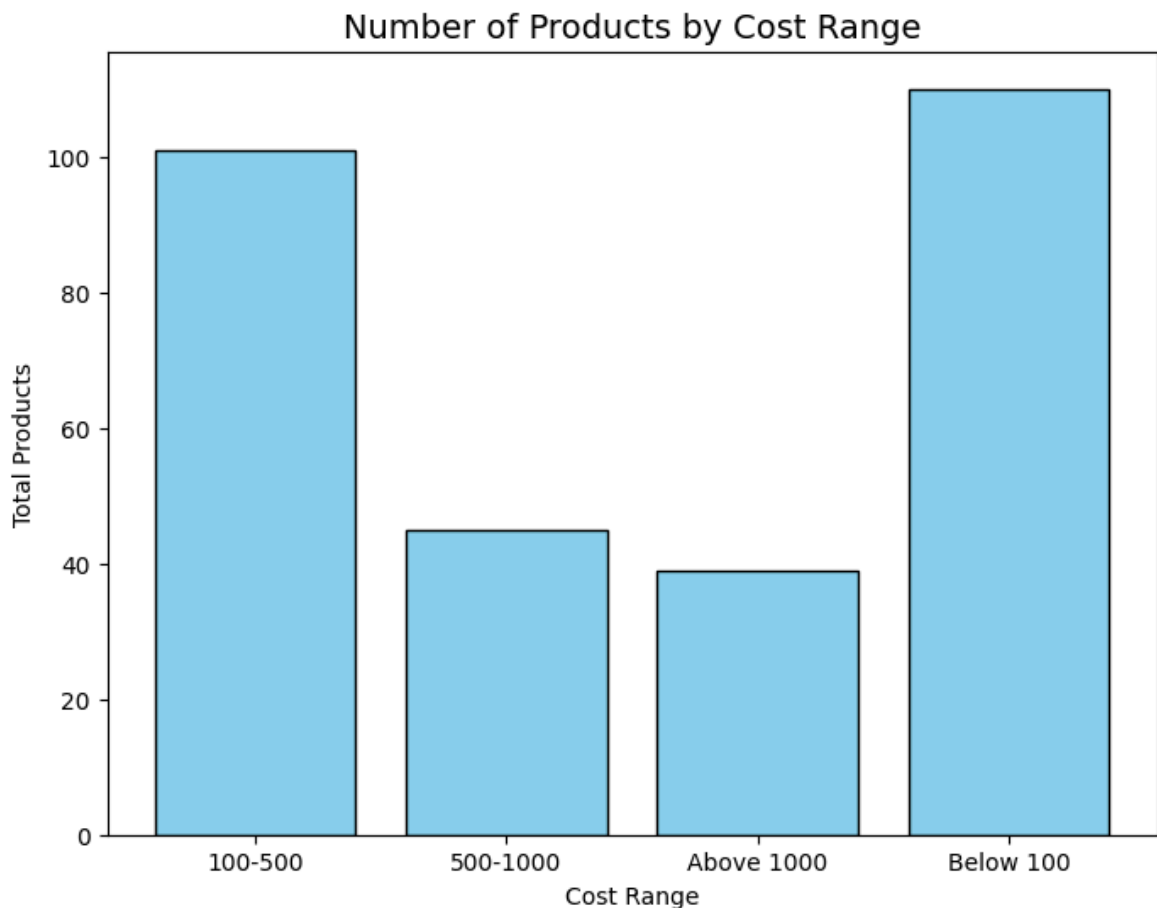
22. Number of Products by Cost Range

Goal: Categorize products based on their cost to understand the distribution of low, mid, and high-priced items.

Chart: Bar chart

EDA Type: Categorical (cost range)

Structure: Define cost ranges and create a new `cost_range` column in the products DataFrame. Group products by cost range and count the total products in each category. Visualize the distribution using a bar chart with labeled axes and clear titles.



Key Insights:

- Dominant price range: Below 100 (110 products - 37.3%)
- Total product portfolio: 295 products across 4 price ranges
- Average products per range: 73.8
- Budget segment (<\$100): 110 products (37.3%)
- Mid-range segment (\$100-\$1000): 146 products (49.5%)
- Premium segment (>\$1000): 39 products (13.2%)
- Market positioning: Balanced portfolio across price segments
- Portfolio distribution: Well-distributed across multiple price ranges
- Core market: 146 products (49.5%) in mainstream price ranges

23. Customer Segments

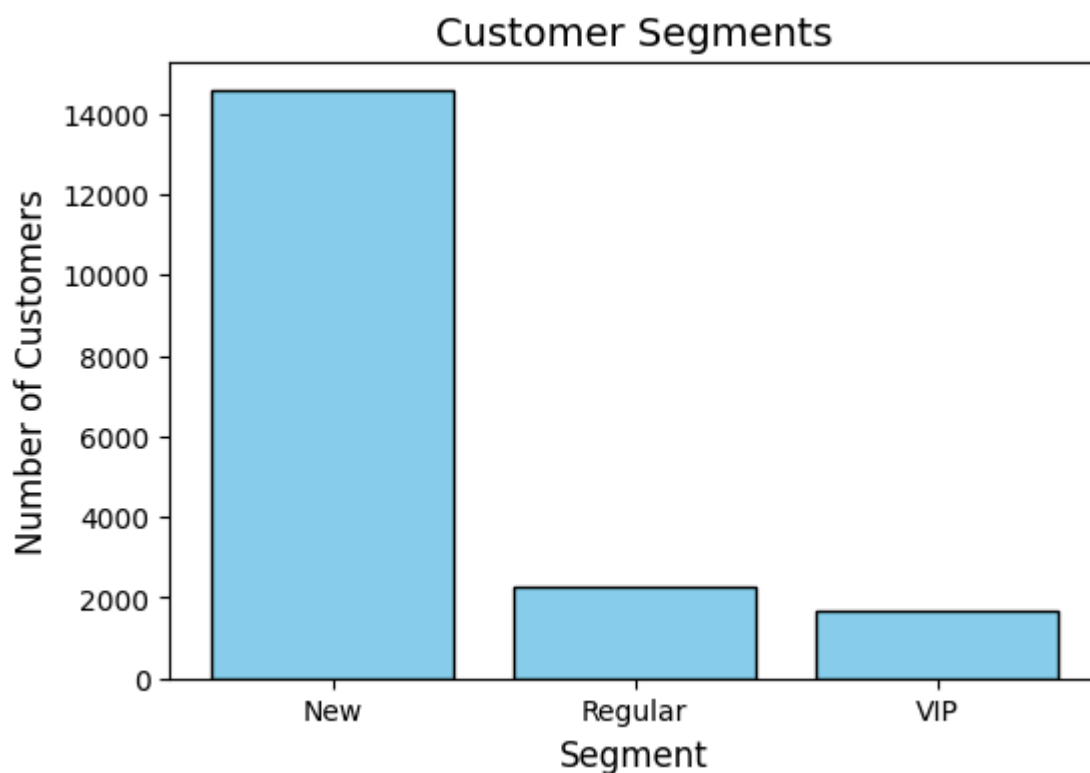
Goal: Segment customers based on their spending and engagement duration to identify VIP, regular, and new customers.

Chart: Bar chart

EDA Type: Categorical (customer segmentation)

Structure:

1. Merge sales and customer data on `customer_key`.
2. Calculate total spending and first/last order dates for each customer.
3. Compute customer lifespan in months.
4. Group customers into three segments based on their spending behavior:
 - **VIP:** Customers with at least 12 months of history and spending more than \$5,000.
 - **Regular:** Customers with at least 12 months of history but spending \$5,000 or less.
 - **New:** Customers with a lifespan less than 12 months.
5. Count the total number of customers in each segment.
6. Visualize the distribution using a bar chart with segment names on the X-axis and customer counts on the Y-axis.



Key Insights:

- Customer base composition: 18484 customers across 3 segments
 - VIP customers: 1661 (9.0%) - High value, loyal customers
 - Regular customers: 2261 (12.2%) - Established, moderate spenders
 - New customers: 14562 (78.8%) - Recent or low-engagement customers
 - VIP segment value: \$10,809,530 (36.8% of total revenue)
 - Average VIP spending: \$6,508 per customer
 - Average Regular spending: \$3,411 per customer
 - Customer maturity: 3922 customers (21.2%) with 12+ month relationship
 - Retention focus: Limited VIP base suggests need for loyalty programs
 - Growth phase: High proportion of new customers indicates expansion or recent customer acquisition
 - Upgrade potential: 2261 Regular customers could be developed into VIP status
-

Customer Metrics Report

Purpose:

- This report consolidates key customer metrics and behaviors

Highlights:

1. Segments customers into categories (VIP, Regular, New) and age groups.

2. Aggregates customer-level metrics:

- total orders
- total sales
- total quantity purchased
- total products
- lifespan (in months)

3. Calculates valuable KPIs:

- recency (months since last order)
- average order value
- average monthly spend

	customer_key	customer_number	customer_name	age_year	total_orders	\
0	1	AW00011000	Jon Yang	54	3	
1	2	AW00011001	Eugene Huang	49	3	
2	3	AW00011002	Ruben Torres	55	3	
3	4	AW00011003	Christy Zhu	52	3	
4	5	AW00011004	Elizabeth Johnson	46	3	

	total_sales	total_quantity	total_products	first_order	last_order	\
0	8249	8	8	2011-01-19	2013-05-03	
1	6384	11	10	2011-01-15	2013-12-10	
2	8114	4	4	2011-01-07	2013-02-23	
3	8139	9	9	2010-12-29	2013-05-10	
4	8196	6	6	2011-01-23	2013-05-01	

	lifespan_months	customer_segment	age_group	recency_months	\
0	28	VIP	50 and above	150	
1	35	VIP	40-49	143	
2	26	VIP	50 and above	152	
3	29	VIP	50 and above	150	
4	28	VIP	40-49	150	

	avg_order_value	avg_monthly_spend
0	2750	295
1	2128	182
2	2705	312
3	2713	281
4	2732	293

◆ Export Final Customer Report Table as a CSV File for Further Analysis and Dashboarding

✓ Customer Report CSV Generated: customer_report.csv

Product Metrics Report

Purpose:

- This report consolidates key product metrics and behaviors.

Highlights:

1. Segments products by revenue to identify High-Performers, Mid-Range, or Low-Performers.
2. Aggregates product-level metrics:
 - total orders
 - total sales
 - total quantity sold
 - total customers (unique)
 - lifespan (in months)
3. Calculates valuable KPIs:

- recency (months since last order)
- average order revenue (AOR)
- average monthly spend

	product_key	product_name	category	subcategory	cost	\
0	4	Mountain Bottle Cage	Accessories	Bottles and Cages	4	
1	5	Road Bottle Cage	Accessories	Bottles and Cages	3	
2	6	Mountain-500 Black- 40	Bikes	Mountain Bikes	295	
3	7	Mountain-500 Black- 42	Bikes	Mountain Bikes	295	
4	8	Mountain-500 Black- 44	Bikes	Mountain Bikes	295	

	total_orders	total_sales	total_quantity	total_customers	first_sale	\
0	2025	20340	2034	2004	2012-12-28	
1	1712	15408	1712	1700	2012-12-28	
2	48	25920	48	48	2012-12-30	
3	49	26460	49	49	2013-01-09	
4	58	31320	58	58	2013-01-05	

	last_sale	lifespan_months	recency_months	avg_order_revenue	\
0	2014-01-28	13	141	10	
1	2014-01-25	13	141	9	
2	2013-12-13	12	143	540	
3	2013-12-25	12	142	540	
4	2013-12-25	12	142	540	

	avg_monthly_revenue	product_segment
0	1565	Mid-Range
1	1185	Mid-Range
2	2160	Mid-Range
3	2205	Mid-Range
4	2610	Mid-Range

◆ Export Final Product Report Table as a CSV File for Further Analysis and Dashboarding

✓ Product Report CSV Generated: customer_report.csv

◆ Export Final HTML Report

✓ HTML Report Generated: report.html