CASE STUDY

SUDDEN SURGE IN SALES

Author: Jeremy Alí Chávez Sequeira

In collaboration with: Microsoft & Coursera

Date: 15/12/2025

Summary

I. Introduction 1
II. Case study 2
III. Analysis of the problem 3
IV. Solution 4
Part 1: Inspecting the dataset 4
Part 2: Carrying out the time series analysis 5
Part 3: Examining the influencers for the peaks 6
Part 4: Gaining insights from the product attributes 7
Part 5: Conclusions of the analysis 9
V. Recommendations 11
VI. References 14

I. Introduction

This case study is taken from Microsoft's *Data Analysis and Visualization with Power BI* course, which is the fifth of eight courses in the "Microsoft Power BI Data Analyst" program on Coursera. The data model for the exercise is provided as a PBIX file (Power BI Desktop report).

The activity is titled "**Exercise: Explaining the increase**", and it's part of Module 4 – Section 2: *Analytical tools in Power BI*, which aims to equip the learner with the strategies and techniques necessary to properly carry out a time series analysis. It focuses on applying the "Analyze" feature, selecting the appropriate visual types (Line, Area and Scatter charts), and using supporting tools such as reference lines and error bars.

This exercise requires the analyst to use the proper types of charts along with the Analyze feature to be able to answer the question of the analysis and produce a report that supports the conclusions. The Analyze feature provides AI-generated correlations directly from the dataset, offering potential explanations for increases or decreases in the variable under study throughout the time series. Line and scatter charts are especially suitable for visualizing data over time, as they help the analyst identify trends, correlations, and potential outliers.

Overall, this case study offers solid analytical practice by simulating a common real-world scenario: investigating the cause of anomalies within a defined time window. It tests the analyst's ability not only to extract meaningful insights, but also to select effective visualizations that clearly communicate findings to stakeholders.

II. Case study

“The marketing department at Adventure Works has identified an unexpected surge in bike sales on two distinct days in the previous month. Once they notified the senior management team of this sharp increase, the information caught the eye of the CEO. They have informed the marketing department that they are keen to identify the contributing factors behind this growth so that the company can capitalize on it and propel the business forward.

The manager of the marketing department has asked the analytics department to determine the reason for the surge in bike sales. It is important that the contributing factors are identified quickly as they may be time-specific and if so, the company would need to move quickly to take advantage of them and generate more business. After a brief discussion with your manager Jamie, you both feel that Microsoft Power BI’s Analyze feature would be the most effective tool to rapidly generate visualizations that could uncover the driving forces behind the sales surge.”

III. Analysis of the problem

The objective was to analyze Adventure Works' sales dataset to identify the key factors that caused two significant peaks in Sales Amount during the month of March. This occurrence leads to the central question of the case study:

"**What product or customer attributes explain the unusual increase in sales on those two days**?"

To address this question, a time series analysis was conducted using line, area, and scatter charts since they provide the clearest visibility of trends, fluctuations and outlier behavior. The product attributes considered in this evaluation are *Product Size*, *Product Category*, and *Product Price*, while customer attributes include *Customer Location / Country* and *Customer Age*.

The Analyze feature in Power BI—specifically the *Explain the Increase* option— was used to obtain automated insights and possible correlations generated directly from the dataset. Additionally, Power BI's *Automatically find clusters* feature was also applied to detect underlying patterns and identify any outliers that could contribute to understanding the observed peaks.

IV. Solution

Part 1: Inspecting the dataset

A thorough inspection of the dataset was carried using both the Table View and the Model View in Power BI. The data model contains three tables: *Customer*, *Order*, and *Sales*. In the Model View, the *Order* and *Customer* tables show a One-to-One relationship, and the *Customer* and *Sales* tables also appear with a One-to-One cardinality.

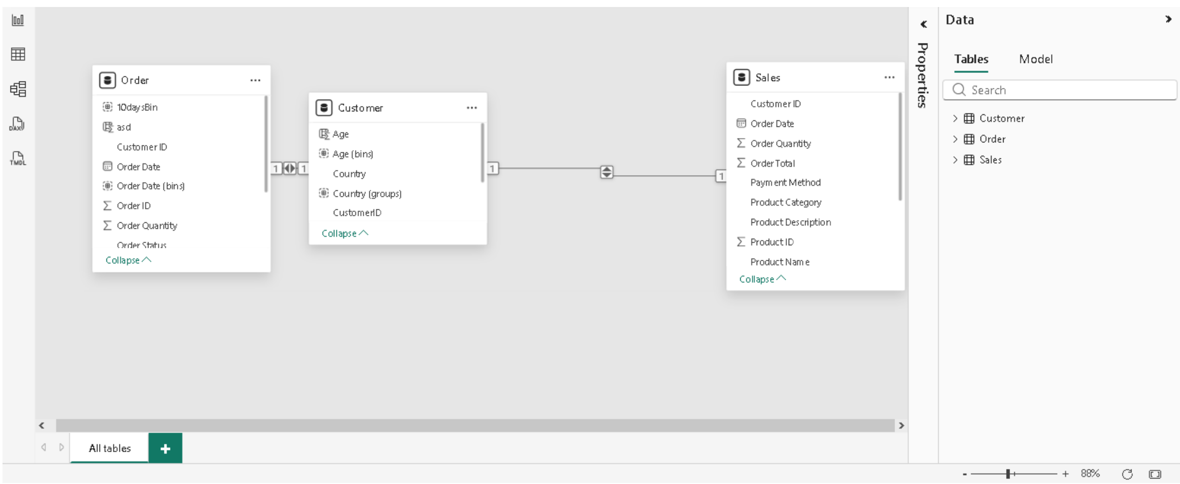
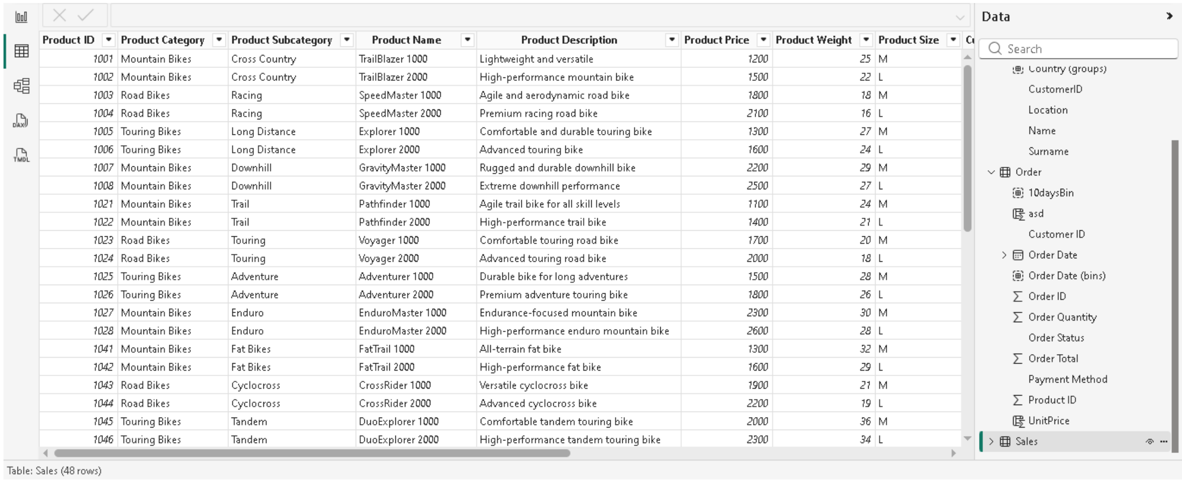


Fig 1. Model View for the dataset of the case study

Examining the Table View provides the full list of attributes available in each table. We can derive all product qualitative attributes from the *Sales* table, including *Product Category*, *Product Subcategory*, *Product Name*, *Product Size* and *Product Price*. This table also includes important transactional fields such as *Order Date* and *Order Total*, which are necessary for the time series analysis.

The *Order* table provides additional fields such as *Order Quantity* and *Payment Method*, while the *Customer* table contains demographic attributes like *Customer Age* and *Customer Country*. These attributes will later serve to segment the data and evaluate their potential influence on the sales peaks identified in the time series.



Product ID	Product Category	Product Subcategory	Product Name	Product Description	Product Price	Product Weight	Product Size
1001	Mountain Bikes	Cross Country	TrailBlazer 1000	Lightweight and versatile	1200	25	M
1002	Mountain Bikes	Cross Country	TrailBlazer 2000	High-performance mountain bike	1500	22	L
1003	Road Bikes	Racing	SpeedMaster 1000	Agile and aerodynamic road bike	1800	18	M
1004	Road Bikes	Racing	SpeedMaster 2000	Premium racing road bike	2100	16	L
1005	Touring Bikes	Long Distance	Explorer 1000	Comfortable and durable touring bike	1300	27	M
1006	Touring Bikes	Long Distance	Explorer 2000	Advanced touring bike	1600	24	L
1007	Mountain Bikes	Downhill	GravityMaster 1000	Rugged and durable downhill bike	2200	29	M
1008	Mountain Bikes	Downhill	GravityMaster 2000	Extreme downhill performance	2500	27	L
1021	Mountain Bikes	Trail	Pathfinder 1000	Agile trail bike for all skill levels	1100	24	M
1022	Mountain Bikes	Trail	Pathfinder 2000	High-performance trail bike	1400	21	L
1023	Road Bikes	Touring	Voyager 1000	Comfortable touring road bike	1700	20	M
1024	Road Bikes	Touring	Voyager 2000	Advanced touring road bike	2000	18	L
1025	Touring Bikes	Adventure	Adventurer 1000	Durable bike for long adventures	1500	28	M
1026	Touring Bikes	Adventure	Adventurer 2000	Premium adventure touring bike	1800	26	L
1027	Mountain Bikes	Enduro	EnduroMaster 1000	Endurance-focused mountain bike	2300	30	M
1028	Mountain Bikes	Enduro	EnduroMaster 2000	High-performance enduro mountain bike	2600	28	L
1041	Mountain Bikes	Fat Bikes	FatTrail 1000	All-terrain fat bike	1300	32	M
1042	Mountain Bikes	Fat Bikes	FatTrail 2000	High-performance fat bike	1600	29	L
1043	Road Bikes	Cyclocross	CrossRider 1000	Versatile cyclocross bike	1900	21	M
1044	Road Bikes	Cyclocross	CrossRider 2000	Advanced cyclocross bike	2200	19	L
1045	Touring Bikes	Tandem	DuoExplorer 1000	Comfortable tandem touring bike	2000	36	M
1046	Touring Bikes	Tandem	DuoExplorer 2000	High-performance tandem touring bike	2300	34	L

Table: Sales (48 rows)

Fig 2. Table View for the dataset of the case study

Part 2: Carrying out the time series analysis

To obtain an initial overview of the sales behavior for the month, two visualizations were created in the Report View. The first one is a line chart showing the daily sales trend for March, which serves as the primary tool to identify fluctuations, patterns, and potential anomalies. The second visualization is a scatter chart, used to explore how *Product Price* correlates with the *Sum of Sales Amount* throughout the month. Since scatter charts require a dimension with a sufficient number of distinct values to display meaningful point distribution, an inspection of the dataset confirmed that *Product Name* was the most appropriate attribute to use.

The resulting visualizations are shown below:

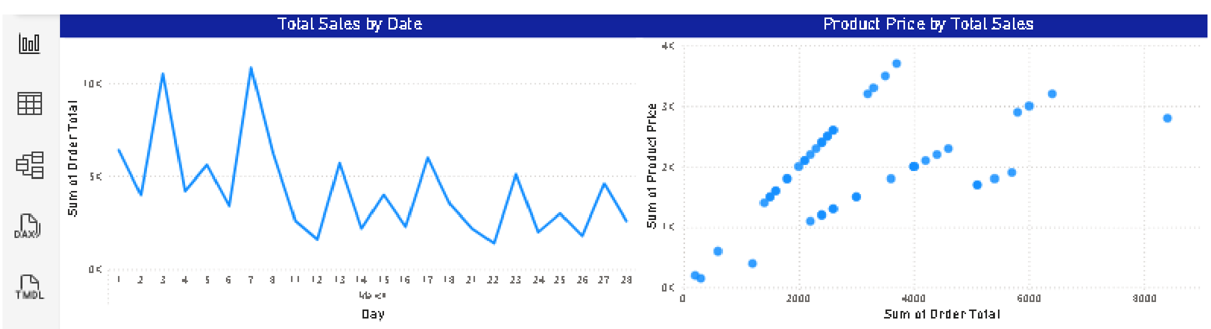


Fig 3. Line and scatter chart for the time series analysis

From the line chart, we can clearly identify two sudden spikes in sales occurring on March 3 and March 7. The scatter chart also reveals a noticeable outlier: the sales behavior of the *AerosSpeed 1000* bike. Overall, the scatter plot suggests a positive correlation between product price and total revenue, which aligns with expected behavior and helps set the foundation for the next steps of the analysis.

Part 3: Examining the influencers for the peaks

Before using the *Analyze* feature, a clustering technique was applied to gain additional context for the automated analysis. The *Automatically find clusters* option was selected from the scatter chart, specifying three clusters as the input. The resulting visualization was as follows:

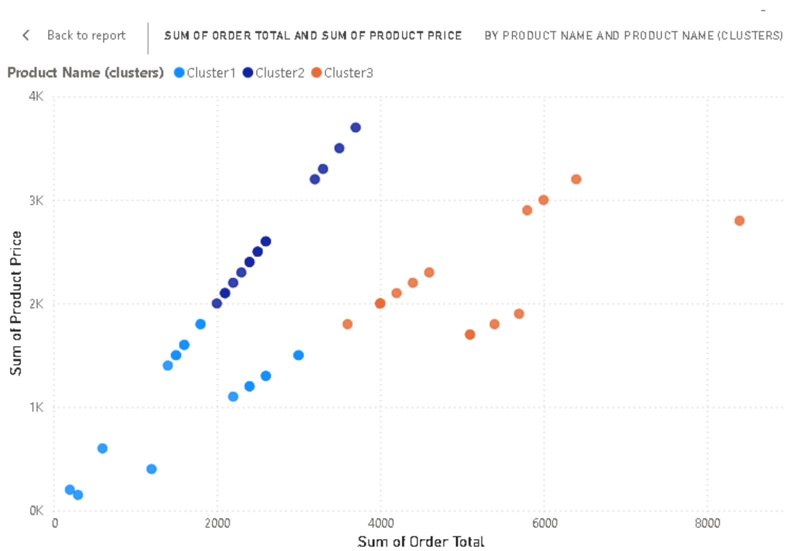


Fig 4. Resulting scatter chart with the automatically found clusters as the legend

The three resulting clusters showed the following characteristics:

- **Cluster 1:** low-priced products with low sales
- **Cluster 2**: medium-to-high-priced products with moderate sales
- **Cluster 3:** medium-to-high-priced products with high sales

These clusters provided an initial segmentation that would help in interpreting the insights generated by Power BI.

With this context in place, the *Explain the increase* feature was run for both March 3 and March 7. The top five AI-generated correlations were noted for each day. Across both analyses, three product-related attributes repeatedly appeared as key influencers of the sales peaks: *Product Size*, *Product Category*, and the Cluster assignment of *Product Name*. These recurring attributes provide a strong indication of the factors contributing to the sudden increases in sales on both dates.

Part 4: Gaining insights from the product attributes

Three additional visualizations were created to draw conclusions about the key product attributes influencing the sales peaks.

Since the *Product Size* field contains only two distinct values (L and M), a donut chart was selected to clearly display the proportion of total revenue (Sum of Order Total) contributed by each size. This visualization helps identify whether one size dominated the sales spikes:

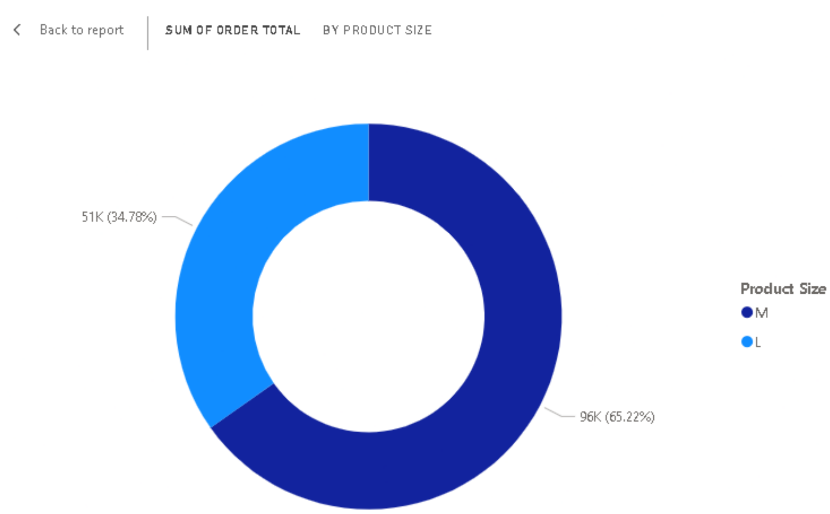


Fig 5. Donut chart for Total Revenue by Product Size

We can infer from the resulting chart that M-sized bikes account for the majority of sales in March.

Next, to compare the distribution of total sales across the previously identified clusters, a Treemap visualization was used. This chart provides an intuitive comparison of the relative contribution of each cluster to the overall revenue:

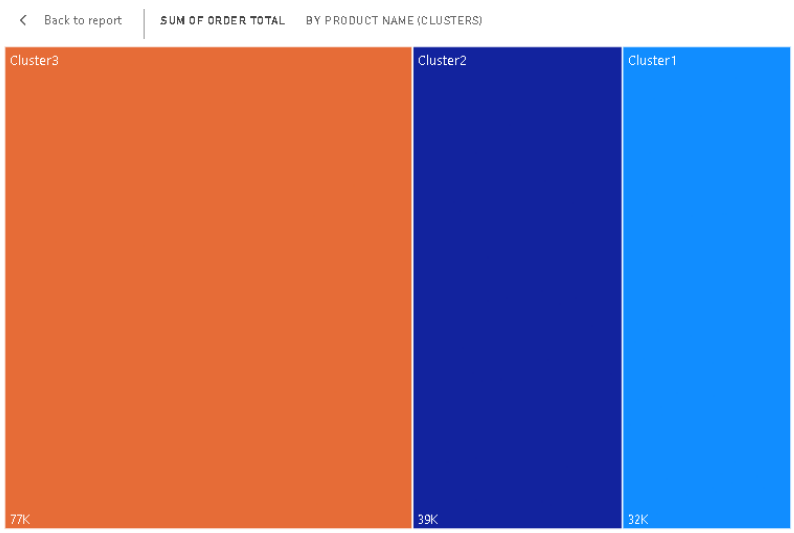


Fig 6. Treemap for Total Revenue by Product Name Clusters

Finally, to examine how each *Product Category* individually contributed to total sales, a waterfall chart was created. This visualization allows us to observe how each category incrementally builds up (or reduces) the total revenue, making it suitable for evaluating category-level influence during the observed peaks.

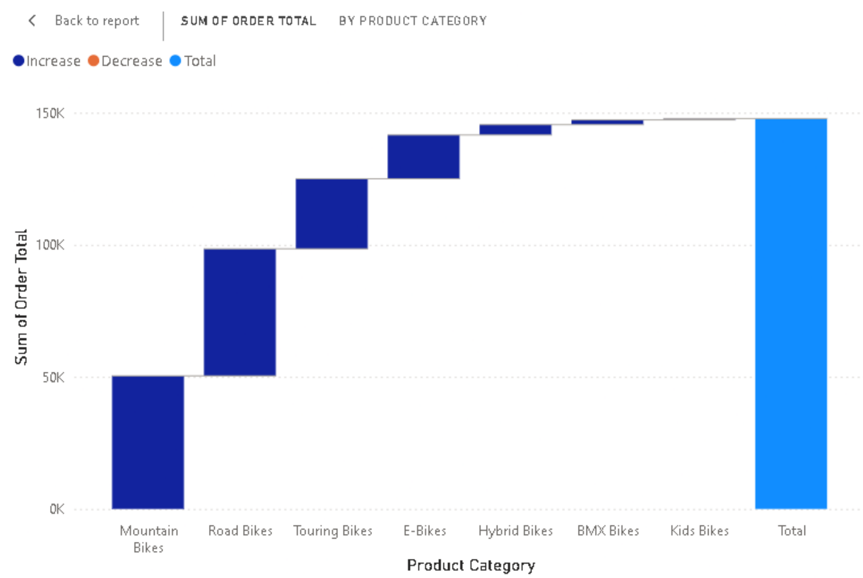


Fig 7. Waterfall chart of Total Revenue by Product Category

Part 5: Conclusions of the analysis

Based on the examination of the key product attributes and their influence on the peak sales days, we can confidently conclude that the sudden surge in bike sales was driven by the following factors:

- **A strong customer preference for M-sized bikes**, which consistently represented the largest share of total sales throughout the month.
- **A high concentration of sales among some mid-to-high-priced products (Cluster 3)**, indicating that some products (like the *Aerospeed 1000*) in the medium-to-high price range drove a substantial portion of the revenue during the peak days.
- **A dominant contribution from Mountain and Road Bikes**, which emerged as the most purchased categories and heavily influenced the spikes observed on March 3 and March 7.

These combined insights provide a clear understanding of the product-level drivers behind the sales increases during the identified peak days.

The final report page is shown below:

Fig 8. Final report page for the case study

V. Recommendations

The next actions are recommended for Adventure Works, following the results of the analysis:

1. Focus inventory and marketing efforts on high-performing categories

Given that Mountain Bikes and Road Bikes produced the highest Order Totals during both peak days, the company should ensure optimal inventory levels for these categories and develop targeted marketing strategies to maintain and further increase demand. Seasonal promotions, feature placements, and product bundles may reinforce their already strong performance.

2. Prioritize medium-to-high-priced, high-demand products

The clustering analysis showed that products within *Cluster 3* (medium-to-high-priced products with high sales) were major contributors to the revenue spikes. To capitalize on their performance, the company should:

- Feature these items more prominently on online listings and in-store displays
- Closely monitor inventory to prevent stock-outs
- Analyze purchasing behavior around these products to identify potential upselling or cross-selling opportunities

3. Optimize inventory for high-demand M-sized products

Because M-sized bikes consistently represented the majority of total sales, Adventure Works should proactively secure adequate stock levels for this size, especially during anticipated high-demand periods. Ensuring availability reduces the risk of missed sales, increases conversion rates, and strengthens marketing efforts centered around customer-preferred products.

4. Leverage identified patterns for predictive planning

The attributes that most influenced the sales peaks—*Product Size*, *Product Category*, and *Cluster* assignments—provide a strong foundation for predictive demand models. Incorporating these variables into forecasting tools can help the company anticipate future spikes and prepare operationally in terms of inventory, logistics, staffing, and targeted promotions.

5. Reevaluate low-performing clusters

Products within *Cluster 1* (low-priced products with low sales) demonstrated minimal contribution to revenue. Adventure Works should:

- Assess whether these products should be discounted, repositioned, or phased out
- Investigate potential causes of underperformance, such as low visibility, limited customer interest, or narrow profit margins
- Evaluate whether rebranding or bundling could improve their performance before making discontinuation decisions

6. Conduct deeper analysis of customer segments

Some customer-related attributes (e.g., *Customer Surname*) provided limited explanatory value during the analysis. To gain deeper insight into customer behaviors behind the sales spikes, it is recommended to segment customers based on:

- Geographic region
- Payment method
- Purchase frequency
- Customer Lifetime Value (CLV)

This additional segmentation can help identify strategic customer groups that contribute disproportionately to revenue and can guide personalized marketing and retention initiatives.

VI. References

Microsoft. (2023). *Exemplar: Explaining the Increase* [Course excerpt]. In *Data Analysis and Visualization with Power BI*. https://www.coursera.org/learn/data-analysis-and-visualization-with-power-bi/supplement/DF1lm/exemplar-explaining-the-increase