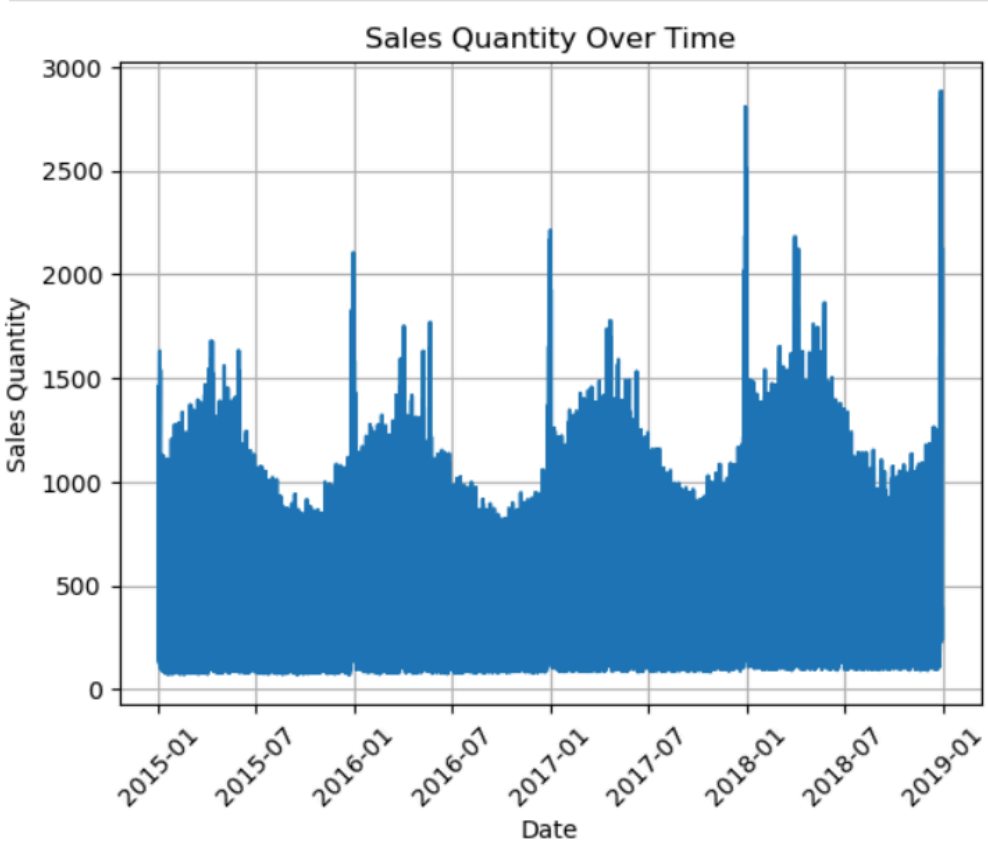
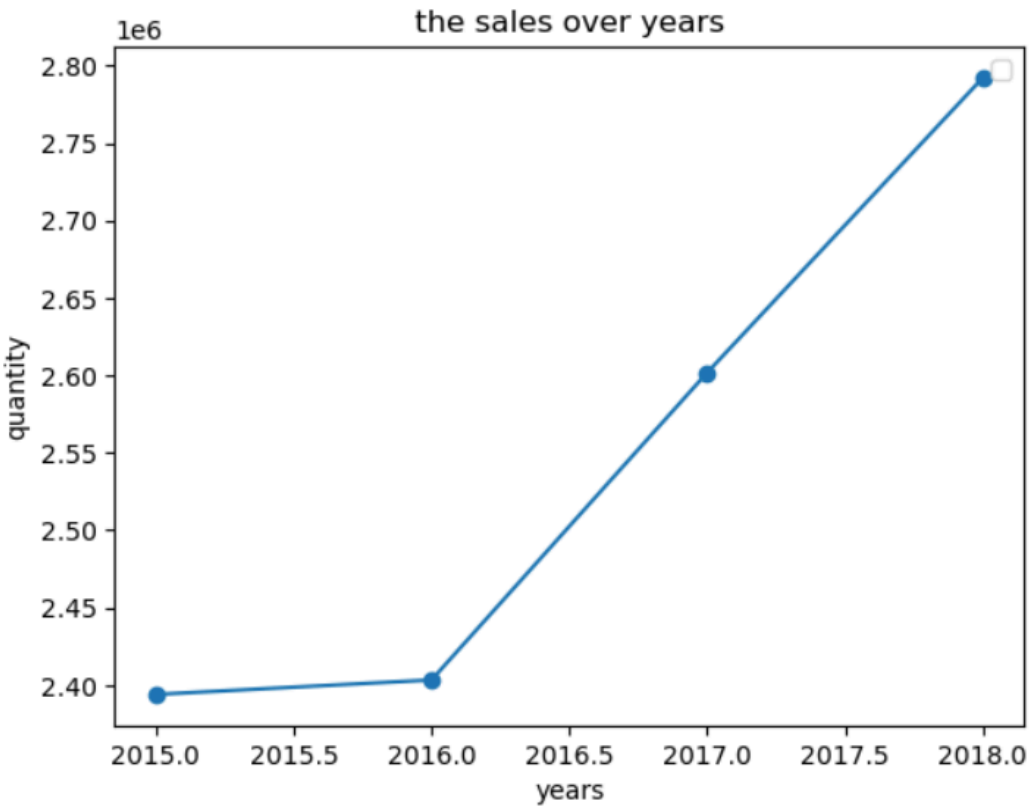



This line chart represents the sales quantity over time. The x-axis shows the date, ranging from early 2015 to late 2019, while the y-axis represents the number of units sold.

Observations:

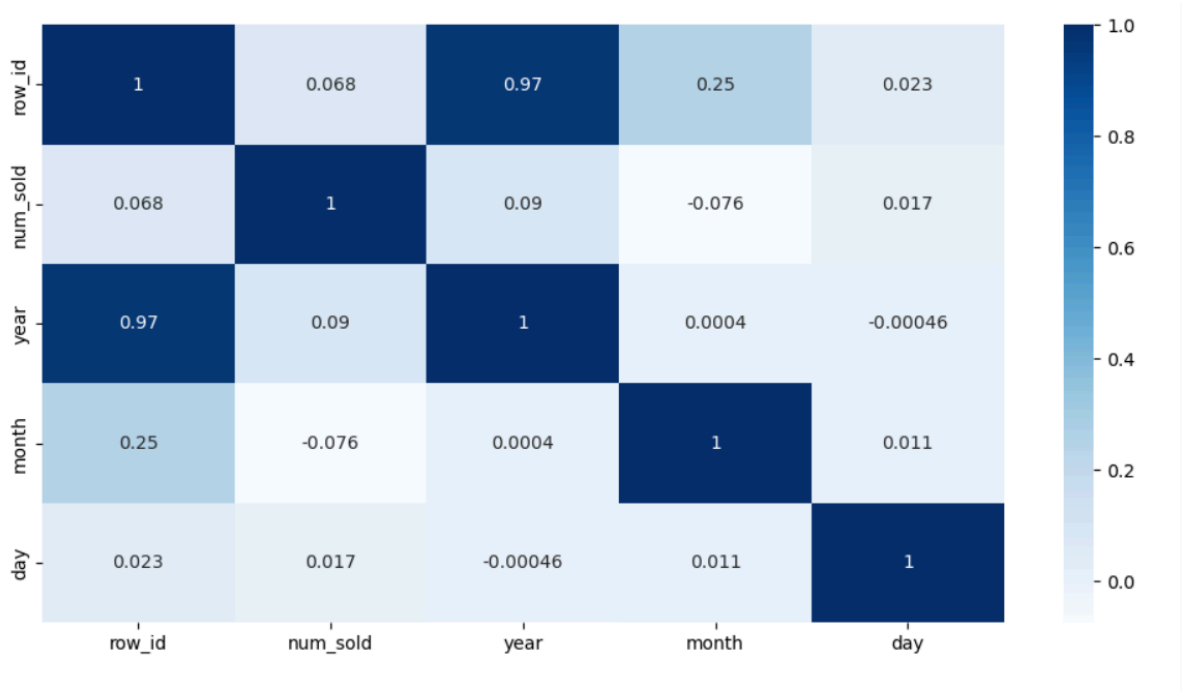
- The sales data exhibits a repeating seasonal pattern, with peaks occurring at regular intervals.
- These spikes suggest periodic increases in sales, which could be due to promotional events, holidays, or seasonal demand.
- The trend remains relatively stable over the years, with no significant overall increase or decrease in sales.
- The grid and rotated x-axis labels enhance readability, making it easier to observe sales trends over time.




This line chart illustrates the total sales quantity over different years. The x-axis represents the years from 2015 to 2018, while the y-axis indicates the number of units sold.

**Observations:**

- The total sales remained nearly constant between 2015 and 2016.
- A significant increase in sales occurred from 2016 to 2017, continuing into 2018.
- The upward trend suggests a strong growth in sales over the years, possibly due to increased market demand, better marketing strategies, or expansion of the business.
- The markers on the line help highlight the sales values for each year.



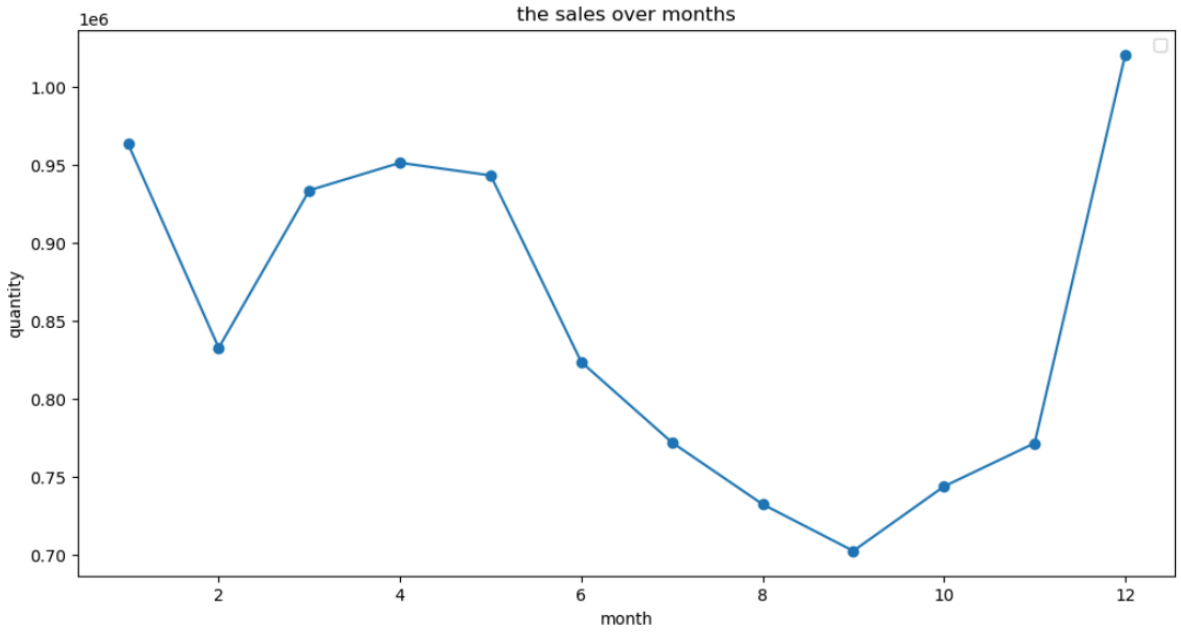

This heatmap represents the correlation between different numerical variables in the dataset. The values range from -1 to 1, where:

- **1 (dark blue)** indicates a perfect positive correlation, meaning two variables increase together.
- **0 (white/light blue)** means no correlation between the variables.
- **Negative values (if present)** suggest an inverse relationship, meaning one variable increases while the other decreases.

**Observations:**

- There is a **strong correlation (0.97) between "year" and "row_id"**, which suggests that data entries are sequentially recorded over time.
- The **sales quantity ("num_sold") has a weak correlation with "year" (0.09) and "month" (-0.076)**, meaning that sales numbers do not strongly depend on these time attributes.
- The **month variable has a slight positive correlation (0.25) with "row_id"**, indicating that sales patterns might repeat monthly to some extent.

This heatmap helps in understanding how different variables in the dataset are related, which can be useful for predictive modeling and data analysis.

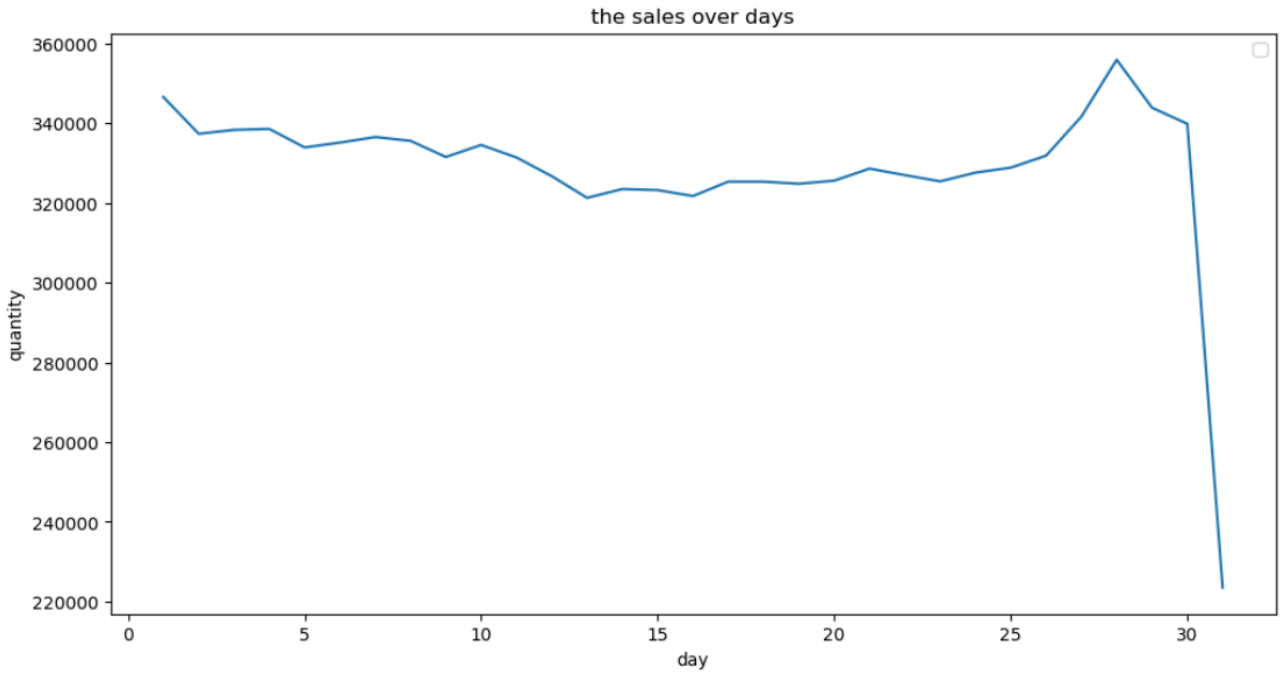



This line chart illustrates the total sales quantity over different months of the year. The x-axis represents the months from 1 to 12, while the y-axis indicates the number of units sold.

## Observations:

- Sales started high in January but dropped significantly in February.
- A recovery occurred from February to May, reaching a peak in April.
- Sales gradually declined from June to September, hitting the lowest point in September.
- From October onward, sales began to rise again, with a sharp increase in December, marking the highest sales volume.
- The markers on the line help highlight the sales values for each month.

This trend suggests seasonal fluctuations in sales, possibly influenced by consumer demand, marketing efforts, or seasonal trends.

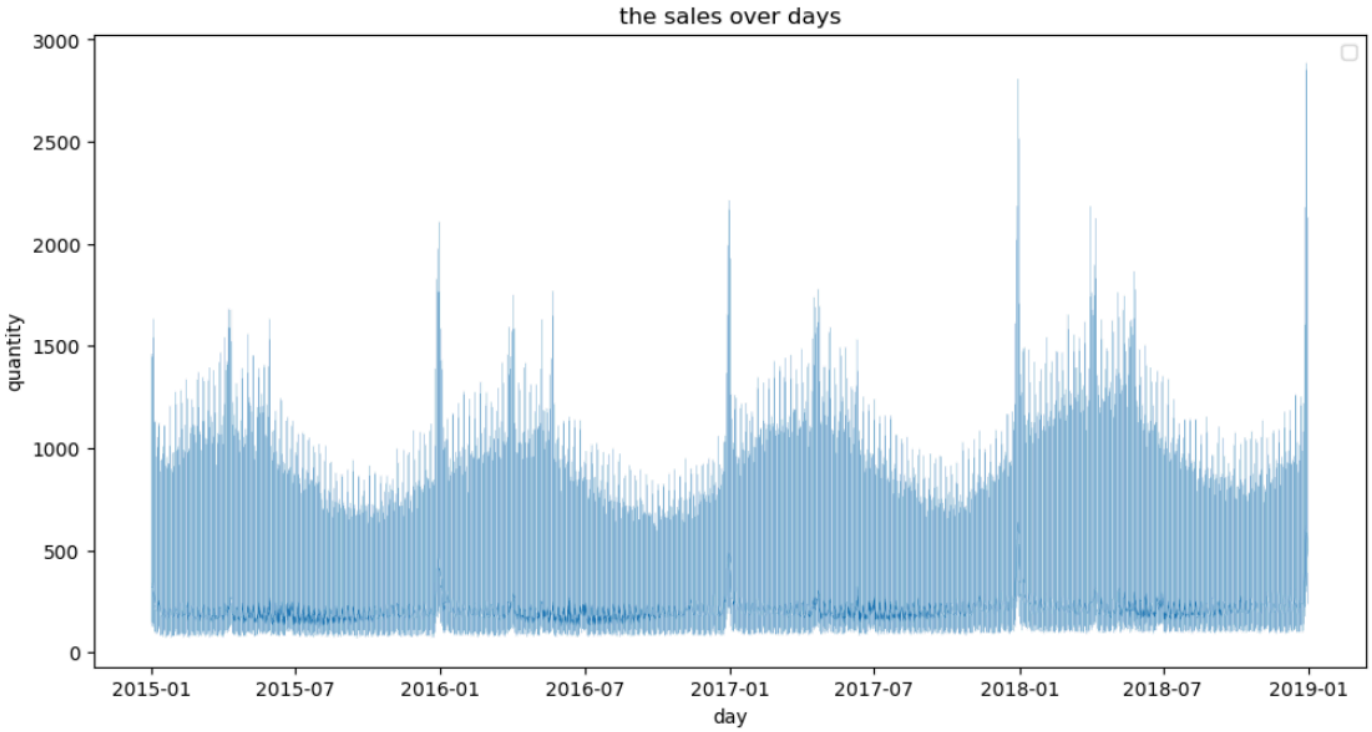



This line chart illustrates the total sales quantity over different days of the month. The x-axis represents the days from 1 to 31, while the y-axis indicates the number of units sold.

## Observations:

- Sales started at a high level at the beginning of the month.
- A slight decline was observed in the first few days, followed by fluctuations in sales throughout the month.
- Sales remained relatively stable in the middle of the month, with minor increases and decreases.
- Towards the end of the month, sales peaked significantly, reaching the highest point around day 28.
- A sharp drop in sales occurred on the last day, which might indicate missing or incomplete data for that period.

This trend suggests possible variations in daily demand, promotional activities, or seasonal shopping behavior.

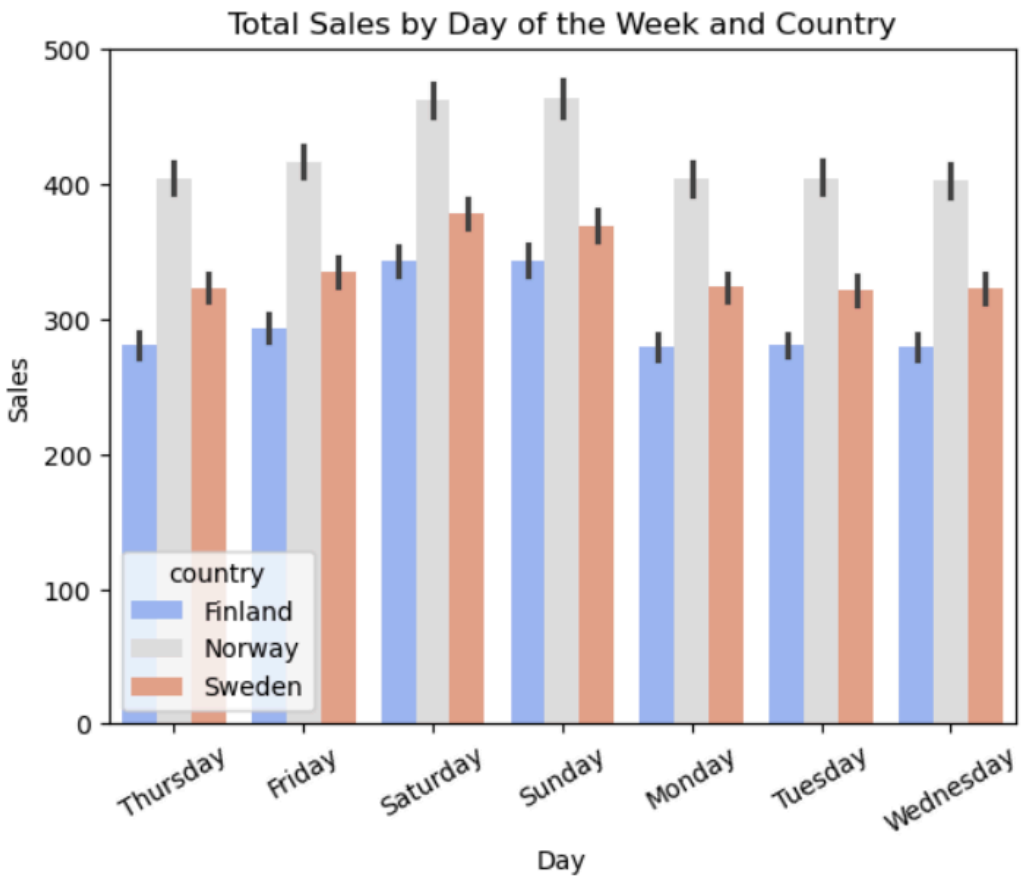



This line chart illustrates the daily sales quantity over multiple years. The x-axis represents the timeline from 2015 to early 2019, while the y-axis indicates the number of units sold per day.

## Observations:

- The sales data shows a clear seasonal pattern, with periodic peaks and declines occurring roughly once a year.
- There are significant spikes in sales at regular intervals, likely corresponding to special events, promotions, or holiday seasons.
- The sales tend to increase gradually over several months, reach a peak, and then decline before the cycle repeats.
- Despite fluctuations, the overall sales trend appears stable across the years, with no drastic long-term growth or decline.

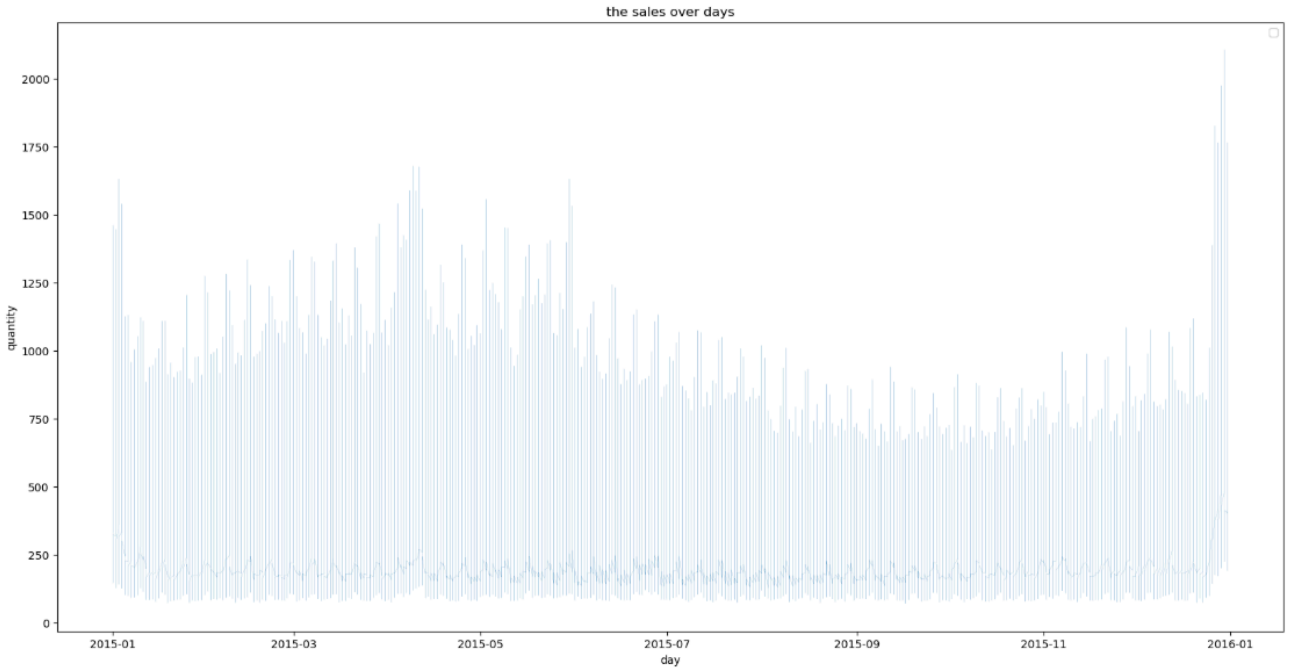
This trend suggests a recurring sales cycle, potentially influenced by consumer behavior, seasonal demand, or company strategies like annual promotions.




This bar chart illustrates the total sales by day of the week for three different countries: **Finland, Norway, and Sweden**.

## Observations:

- The x-axis represents the days of the week from **Thursday to Wednesday**.
- The y-axis indicates the total **sales**.
- Each day has three bars, each representing a different country:
    - **Finland (blue)**
    - **Norway (gray)**
    - **Sweden (orange)**
- **Norway consistently has the highest sales** across all days, followed by Sweden, while Finland has the lowest sales.
- **Sales peak on Saturday and Sunday** for all three countries, with Norway reaching its highest sales on these days.
- Sales drop slightly on **Monday and Tuesday** but remain relatively stable.
- **Error bars** on top of each bar indicate the variation or uncertainty in sales data. This suggests that sales tend to be **higher during the weekend**, which might be due to increased shopping activity on Saturdays and Sundays. Norway has the strongest sales performance throughout the week, while Finland has the lowest.

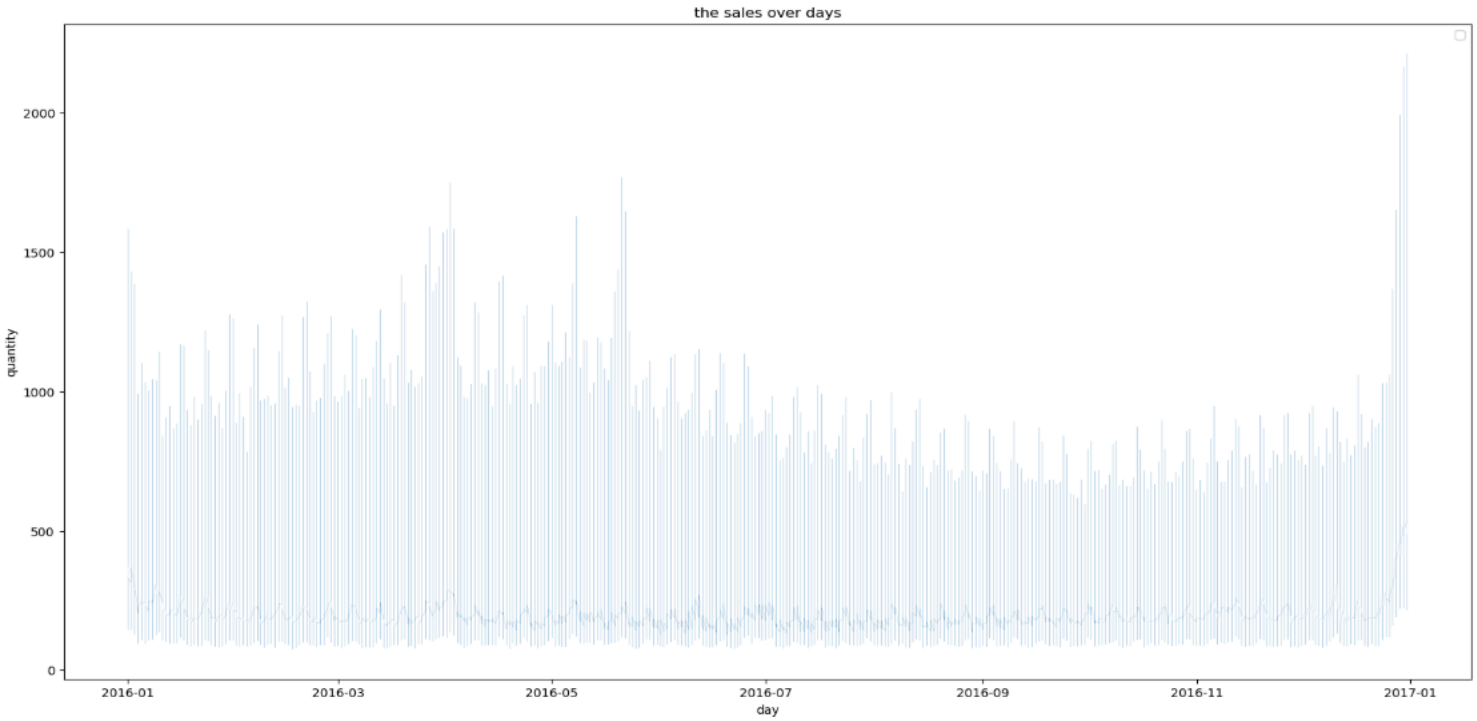



This line chart illustrates the daily sales quantity over the course of a year. The x-axis represents the timeline from early 2015 to early 2016, while the y-axis indicates the number of units sold per day.

## Observations:

- The sales data exhibits **daily fluctuations** with periodic peaks and declines, suggesting a **seasonal pattern**.
- There are noticeable **spikes in sales at the beginning and end of the year**, likely corresponding to special events, promotions, or holiday seasons.
- Sales tend to **increase gradually over several months**, reach a peak, and then decline before repeating the cycle.
- Despite these fluctuations, the overall sales trend remains relatively stable, with no drastic long-term increase or decrease.

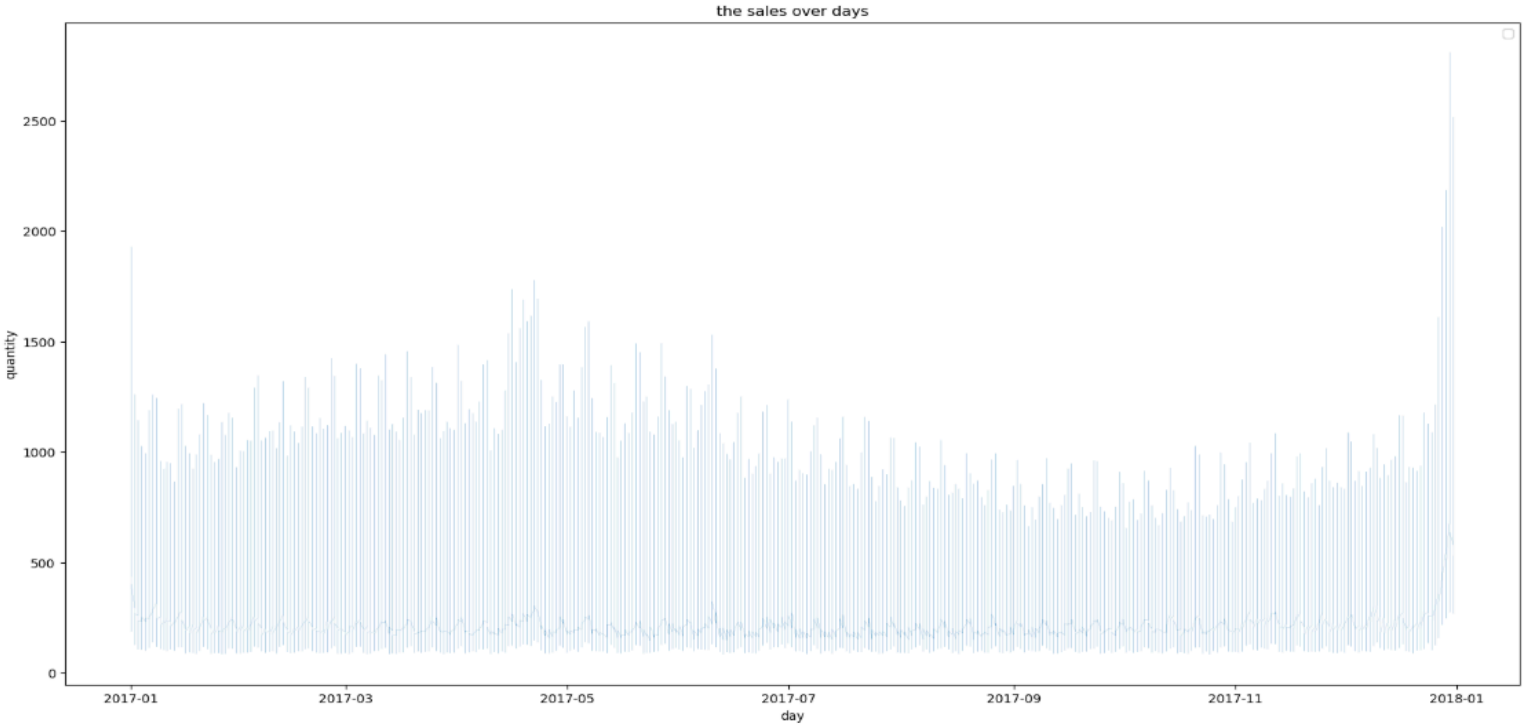
This trend suggests a **recurring sales cycle**, potentially influenced by consumer behavior, seasonal demand, or company strategies such as annual discounts and promotions.




This line chart illustrates the daily sales quantity over the year 2016. The x-axis represents the timeline from January 2016 to early January 2017, while the y-axis indicates the number of units sold per day.

## Observations:

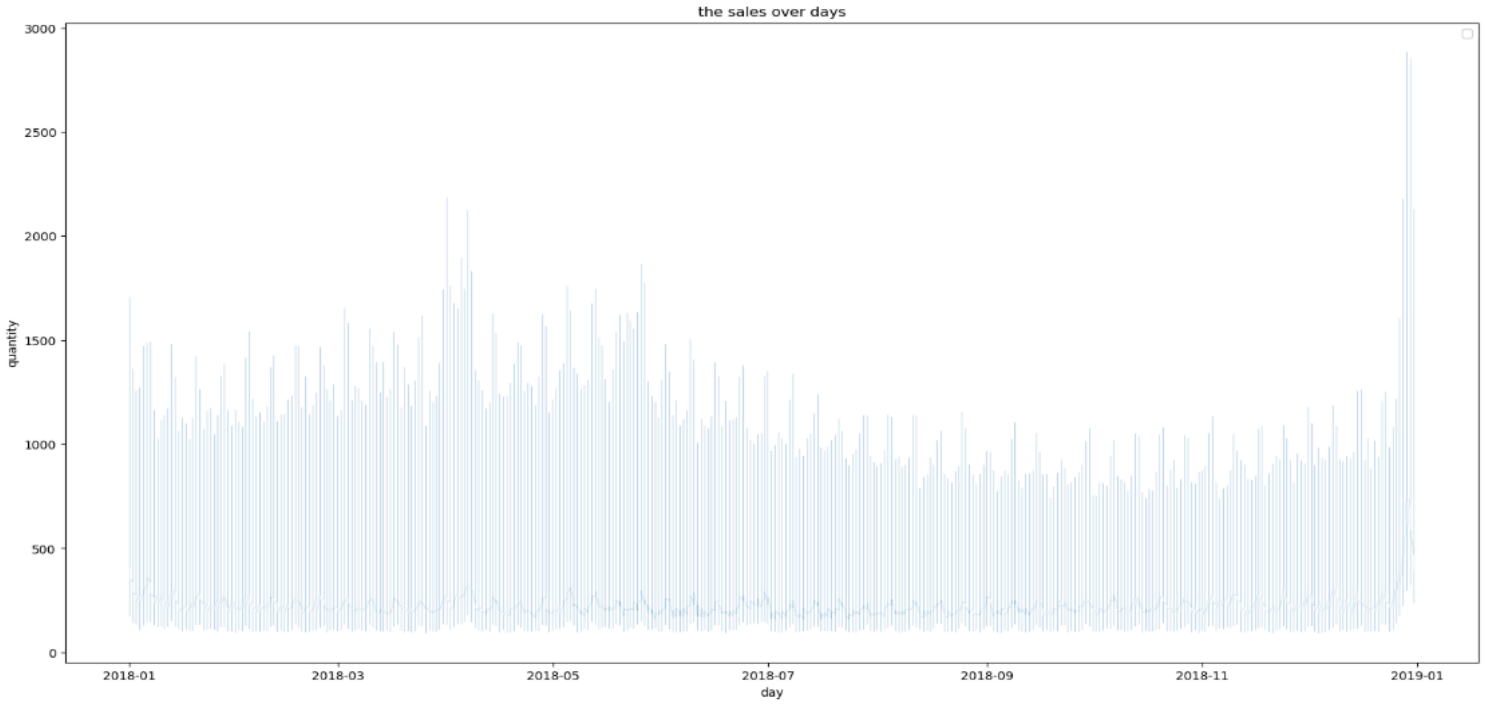
- The sales data follows a **seasonal pattern**, with fluctuations in sales volume throughout the year.
- There are significant **spikes in sales at the beginning and end of the year**, likely linked to special promotions, holidays, or major shopping events.
- The sales trend shows a gradual **increase in sales volume during the first half of the year**, followed by a decline in the later months before another sharp increase in December.
- The overall trend remains **consistent**, indicating a **recurring sales cycle** influenced by consumer demand and external factors like seasonal promotions or market strategies.




This line chart illustrates the daily sales quantity over the year 2017. The x-axis represents the timeline from January 2017 to early January 2018, while the y-axis indicates the number of units sold per day.

## Observations:

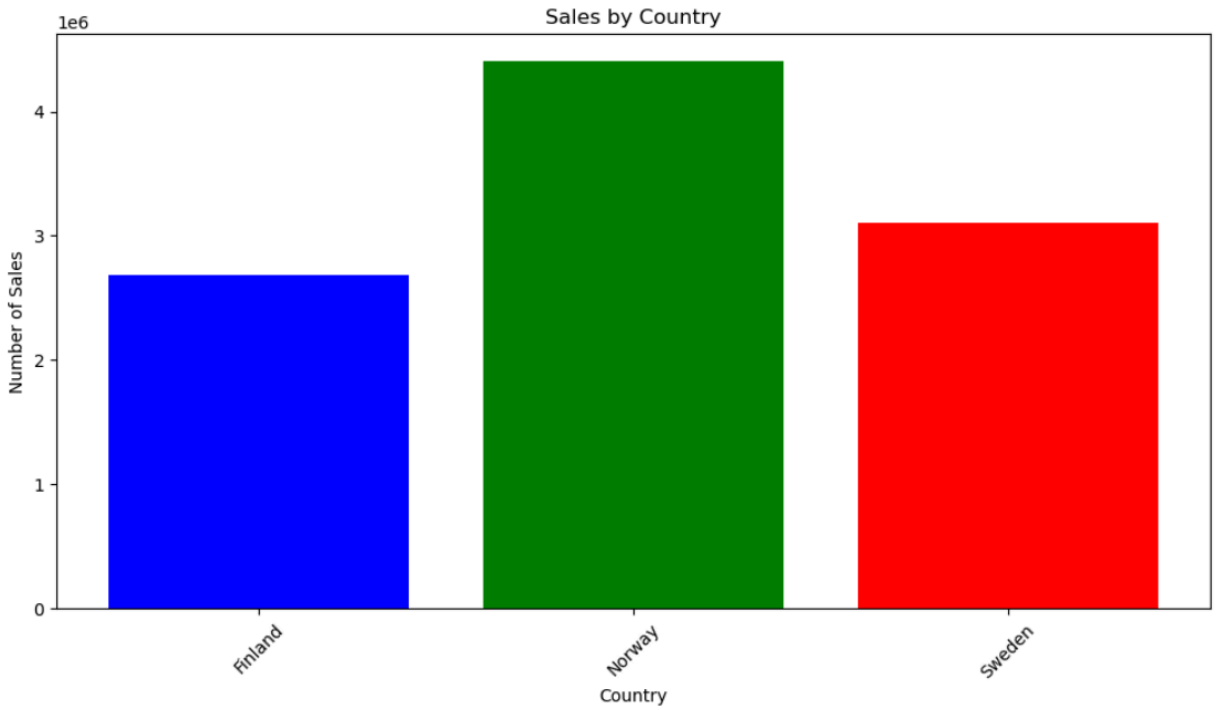
- The sales data follows a **recurring seasonal pattern**, with peaks and dips throughout the year.
- There are **sharp spikes in sales at the start and end of the year**, suggesting the influence of major promotional events or holiday seasons.
- Sales show **gradual growth in the first half of the year**, followed by a decline in mid-year, and then another sharp rise toward December.
- The fluctuations and cycles are similar to previous years, reinforcing a **consistent sales pattern** influenced by consumer behavior and seasonal demand.




This line chart illustrates the daily sales quantity over the year 2018. The x-axis represents the timeline from January 2018 to early January 2019, while the y-axis indicates the number of units sold per day.

## Observations:

- The **seasonal sales pattern** continues, with periodic peaks and declines occurring throughout the year.
- Sales reach their **highest points at the beginning and end of the year**, suggesting strong performance during holiday seasons or major promotions.
- There is **an increase in sales during the first half of the year**, followed by a gradual decline in mid-year and a sharp rise towards December.
- This trend aligns with previous years, reinforcing a **consistent annual cycle** likely influenced by consumer behavior, promotional strategies, and seasonal demand.

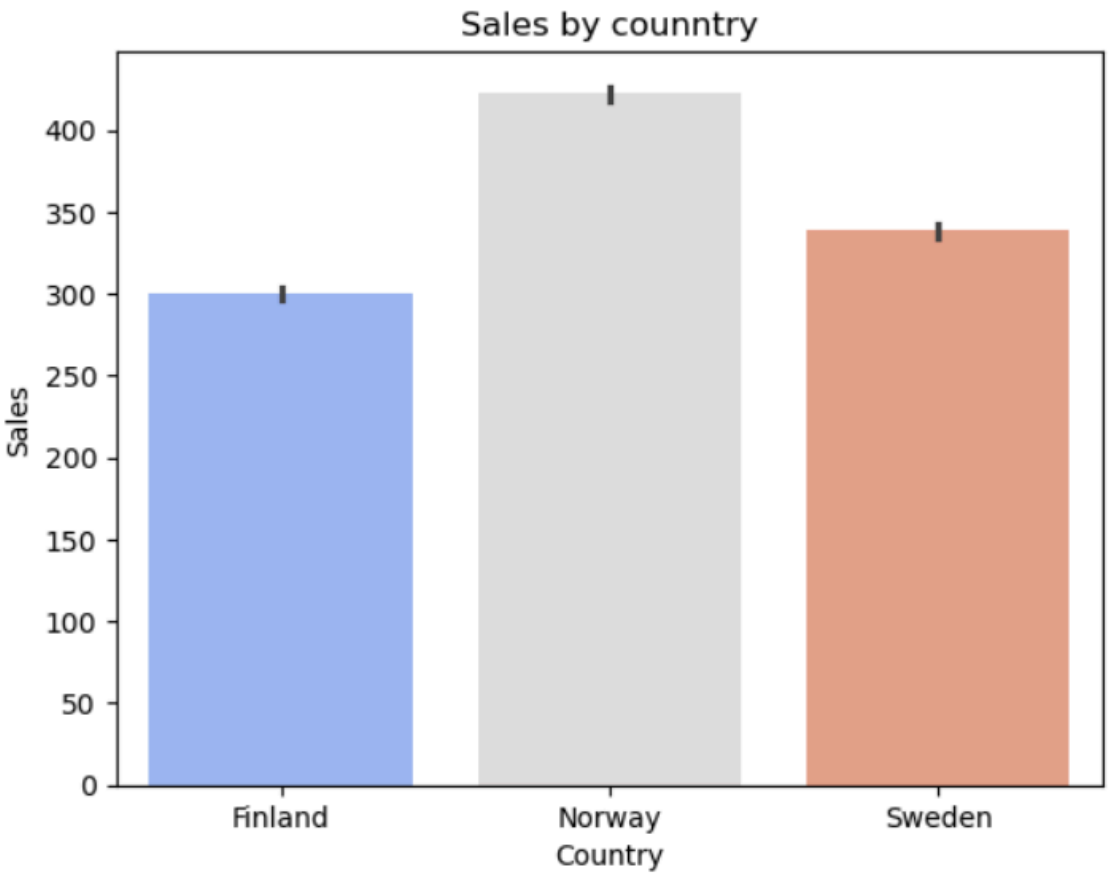



This bar chart represents **sales by country** for Finland, Norway, and Sweden.

## Observations:

- **Norway** has the highest number of sales, exceeding **4.5 million**.
- **Sweden** follows with sales slightly above **3 million**.
- **Finland** has the lowest sales, around **2.7 million**.
- The bars are color-coded for distinction:
    - **Finland (blue)**
    - **Norway (green)**
    - **Sweden (red)**

This visualization suggests that Norway is the **strongest market** among the three countries, while Finland has the **lowest performance** in sales.

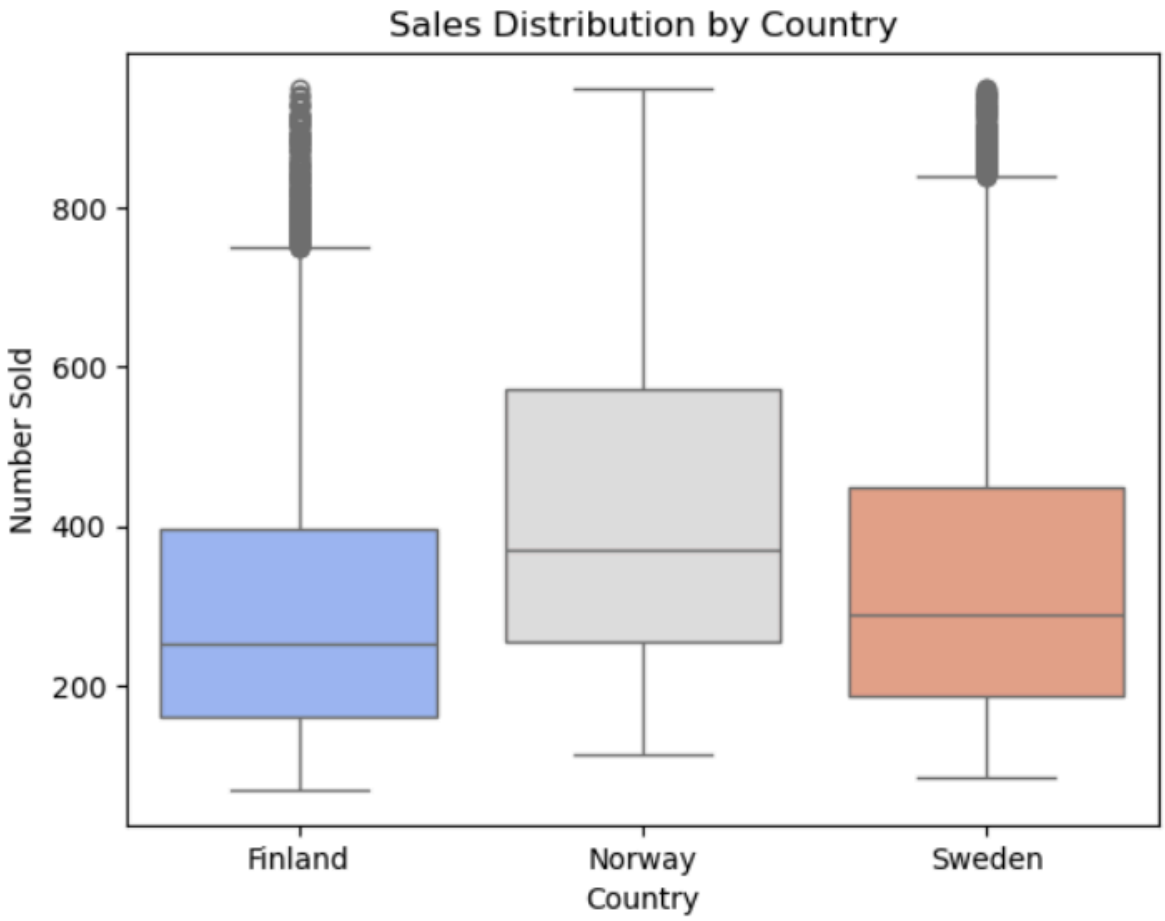



This bar chart represents **sales by country** for Finland, Norway, and Sweden.

## Observations:

- **Norway** has the highest number of sales, exceeding **400**.
- **Sweden** follows with sales around **340**.
- **Finland** has the lowest sales, around **300**.
- The bars are color-coded for distinction:
    - **Finland (light blue)**
    - **Norway (gray)**
    - **Sweden (light brown)**
- There are **small error bars**, indicating some variability in the data.

This visualization suggests that Norway is the **leading market**, followed by Sweden and then Finland.

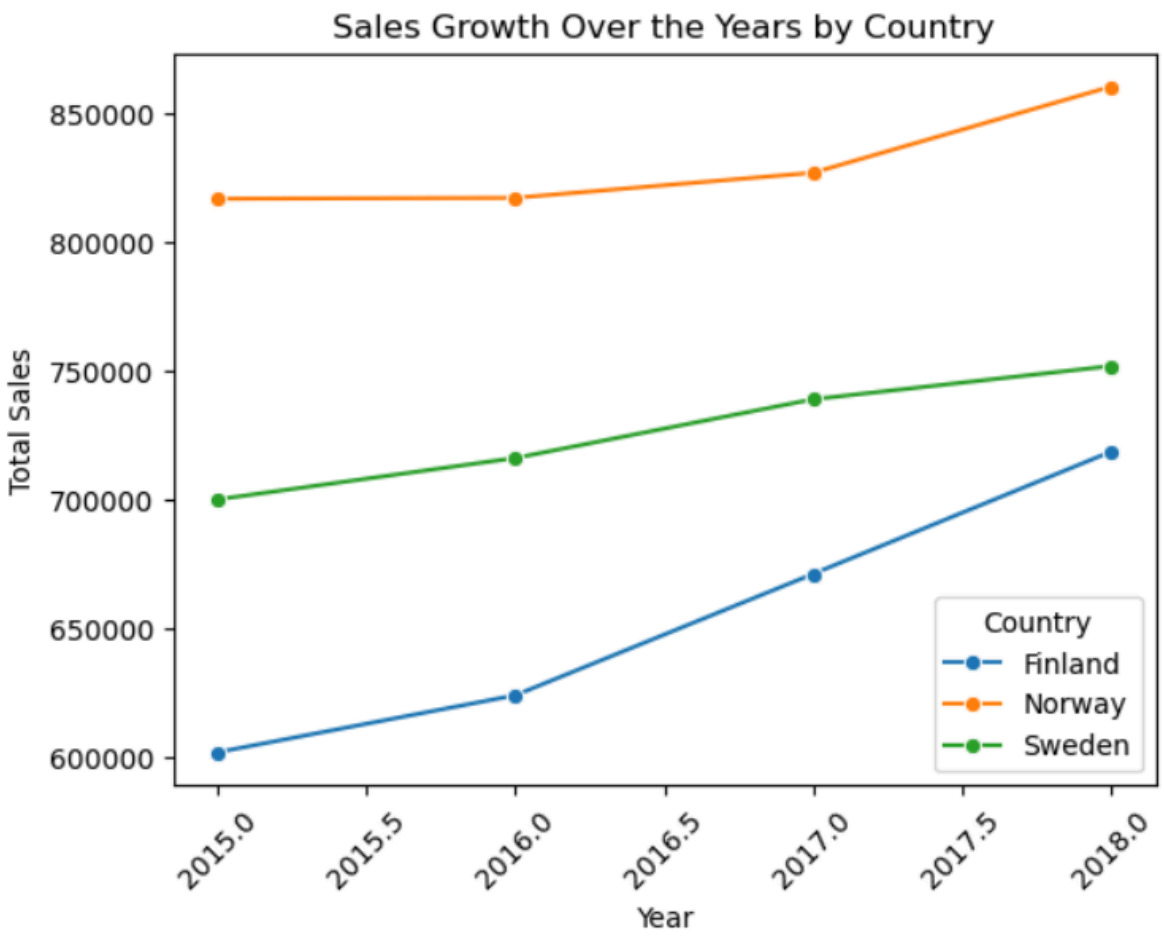



This is a **box plot** representing the **sales distribution by country** for **Finland, Norway, and Sweden**.

## Observations:

- **Norway** has the highest median sales among the three countries.
- **Sweden and Finland** have similar medians, but Sweden has a wider spread in sales distribution.
- **Outliers** (small circles above the whiskers) are present in **Finland and Sweden**, indicating that some data points are significantly higher than the rest.
- **Norway has the widest interquartile range (IQR)**, meaning the sales data is more spread out compared to Finland and Sweden.

This visualization suggests **Norway's sales are generally higher** and more varied, while **Finland and Sweden have more outliers and a smaller range of typical sales values**.

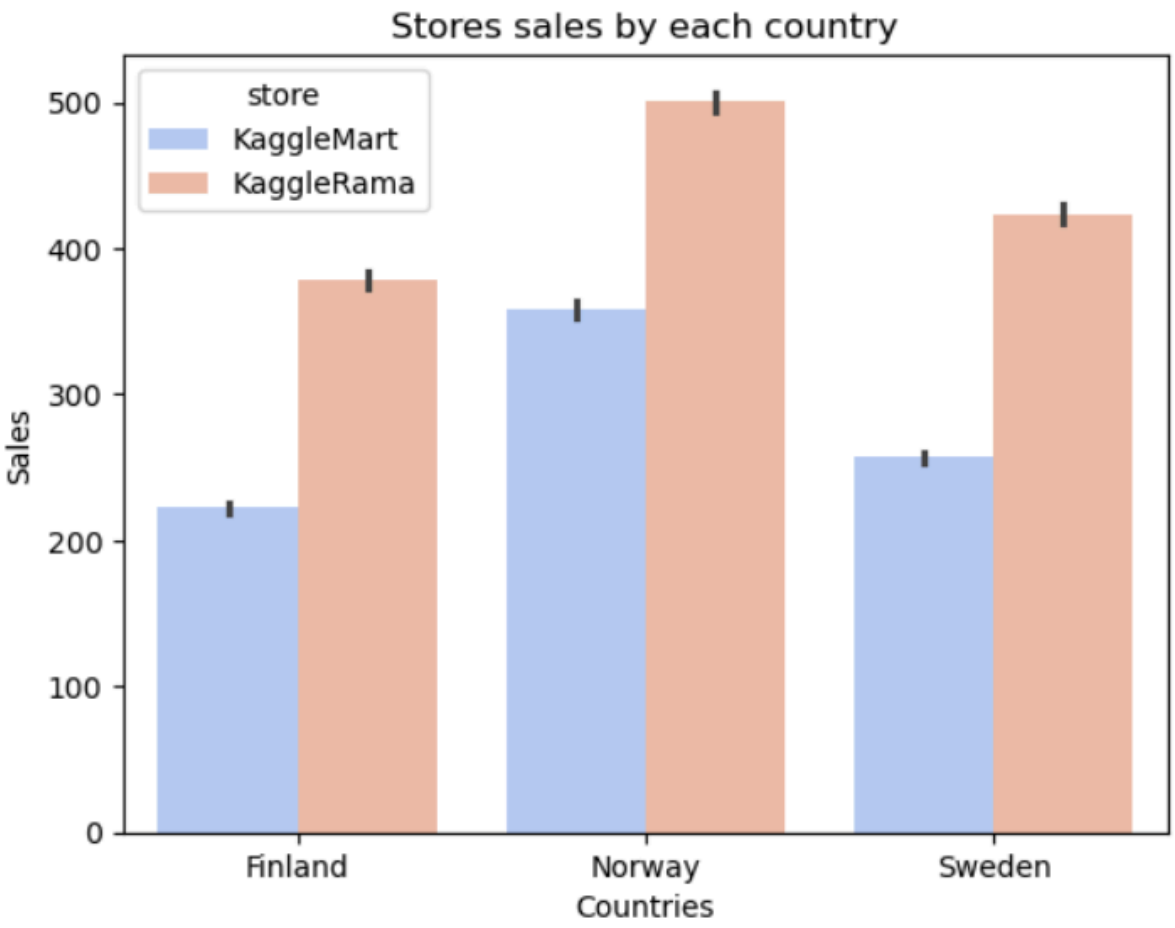



This is a **line chart** showing **sales growth over the years** for **Finland, Norway, and Sweden** from **2015 to 2018**.

## Observations:

- **Finland (blue line)** shows the **steepest growth**, with total sales increasing significantly over the years, especially from **2016 to 2018**.
- **Norway (orange line)** started with the **highest sales** but had **slower growth** until **2017**, after which there was a noticeable increase.
- **Sweden (green line)** has a **steady and moderate increase** in sales, growing consistently each year.

This suggests that **Finland's sales are accelerating rapidly**, while **Norway maintains a leading position with steady growth** and **Sweden follows a stable upward trend**.

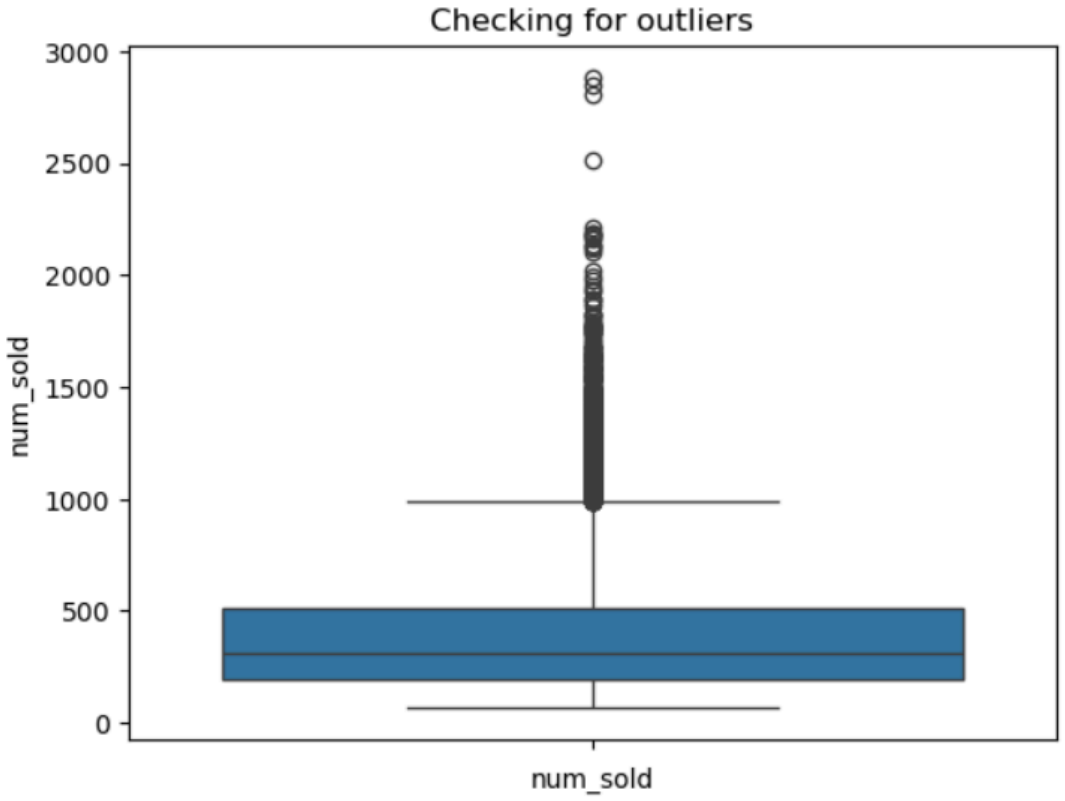



This **bar chart** displays **store sales by country**, comparing two stores: **KaggleMart (blue)** and **KaggleRama (peach)** across **Finland, Norway, and Sweden**.

## Observations:

- **KaggleRama** consistently **outperforms KaggleMart** in all three countries.
- In **Finland**, KaggleRama's sales are significantly higher than KaggleMart's.
- In **Norway**, both stores have high sales, but **KaggleRama leads** by a noticeable margin.
- In **Sweden**, the trend is the same—KaggleRama has higher sales than KaggleMart.

This suggests that **KaggleRama is performing better overall** across all countries, while **KaggleMart has lower but steady sales**.

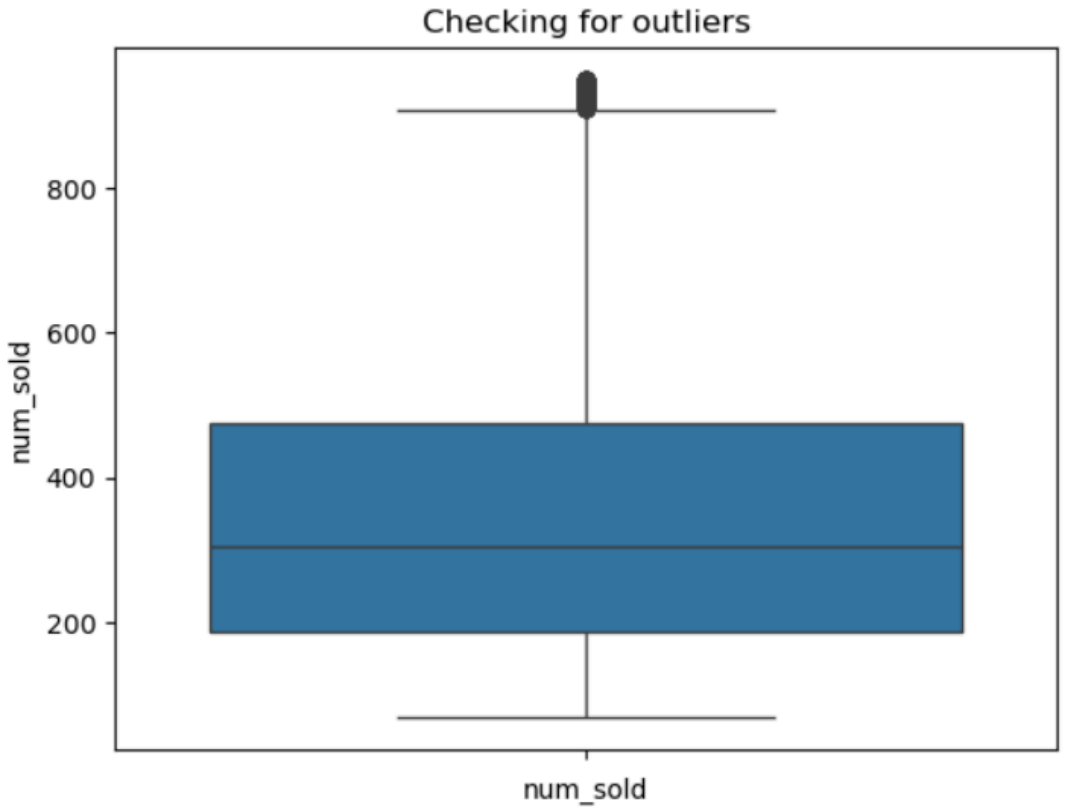



This **box plot** is used for **checking outliers** in the **num_sold** variable.

## Observations:

- The **box (IQR region)** shows that most sales values are **between approximately 100 and 500**.
- The **whiskers** extend to around **1000**, indicating the upper limit before considering values as outliers.
- There are **numerous outliers above 1000**, with some exceeding **2500-3000** sales.
- The **presence of many high-value outliers** suggests that some products or events cause exceptionally high sales.

This visualization helps identify **extreme values that might need further analysis** or **potential filtering** based on business needs.



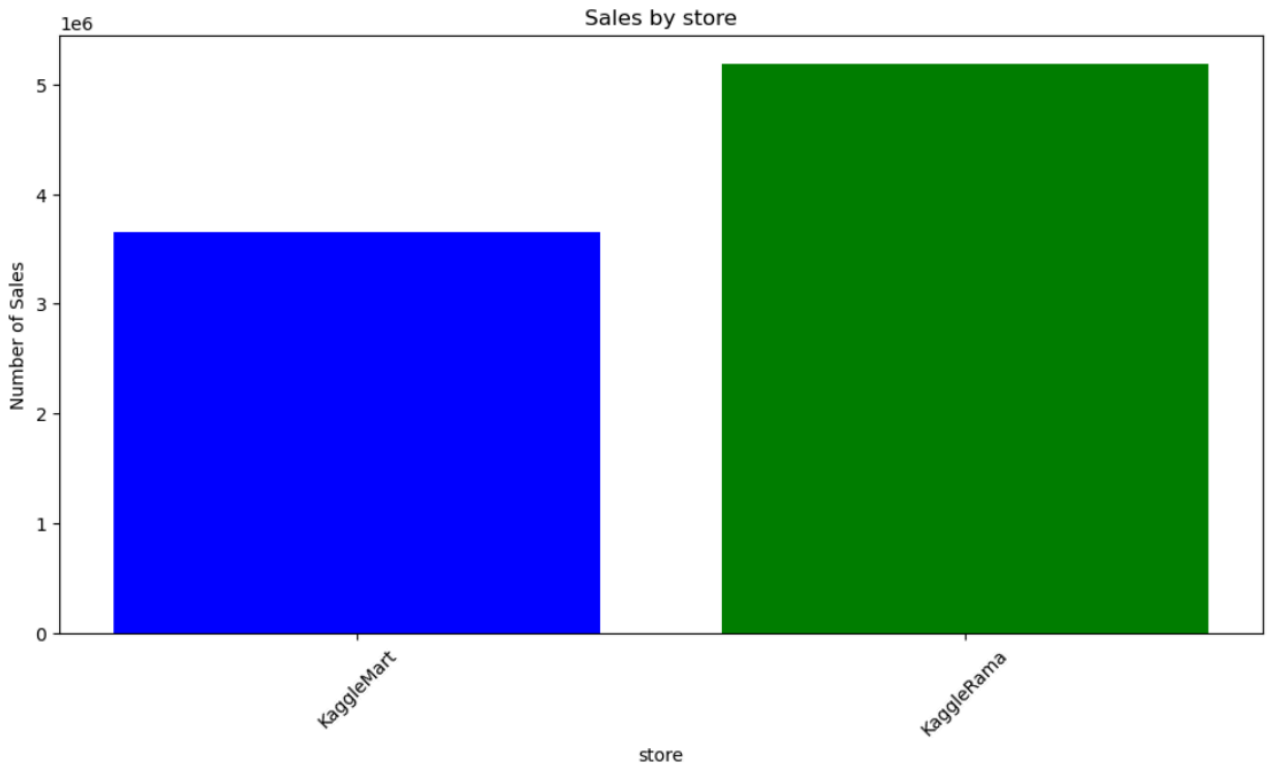

This **box plot** again checks for **outliers** in the **num_sold** variable.

## Comparison with Previous Box Plot:

- The **distribution appears much cleaner**, with **fewer extreme outliers**.
- The **interquartile range (IQR)** is between approximately **100 and 500**, similar to the previous plot.
- The **whiskers extend to nearly 900**, with **only a few outliers above that**.
- This suggests that **previous extreme outliers (above 2500-3000) have been removed or filtered**.

## Possible Interpretation:

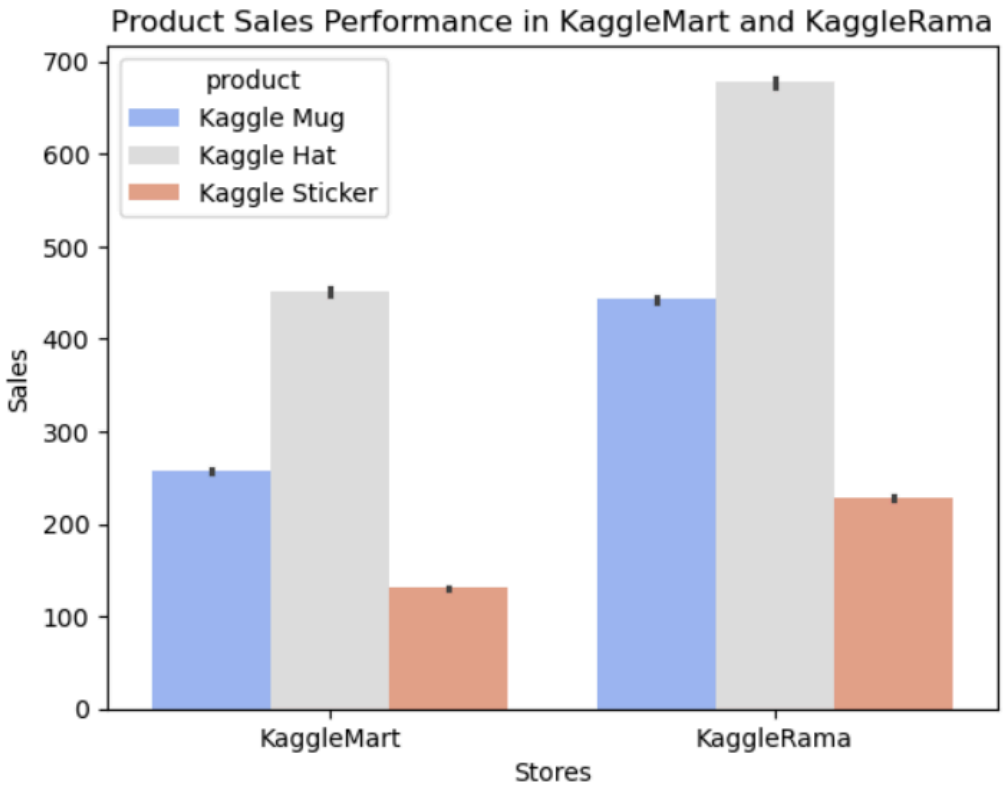
- This indicate **data cleaning** or **outlier removal based on a threshold**.
- The remaining **outliers above 900** are still worth investigating but are far fewer than before.
- The **data is now more balanced** and likely provides **better insights without extreme values skewing the analysis**.




This bar chart represents the total number of sales for each store.

## Observations:

- The **KaggleMart store (blue bar)** has approximately **3.8 million sales**.
- The **KaggleRama store (green bar)** has over **5 million sales**, making it the **better-performing store**.
- KaggleRama's sales are **significantly higher** than KaggleMart's, showing a strong performance difference.
- This visualization helps compare the overall sales performance of both stores and could guide further analysis into the reasons behind the variation in sales.




This bar chart illustrates the sales performance of different products in KaggleMart and KaggleRama.

## Observations:

- **Kaggle Hats** (gray bars) have the highest sales in both stores, with **KaggleRama selling around 700** and **KaggleMart around 450**.
- **Kaggle Mugs** (blue bars) perform **better in KaggleRama (~450 sales) than in KaggleMart (~250 sales)**.
- **Kaggle Stickers** (orange bars) have the lowest sales, though KaggleRama still outperforms KaggleMart.
- The higher sales in KaggleRama across all products suggest **stronger customer demand** or **better marketing strategies** in that store.

This visualization helps compare product performance across stores and could inform stock or marketing decisions.