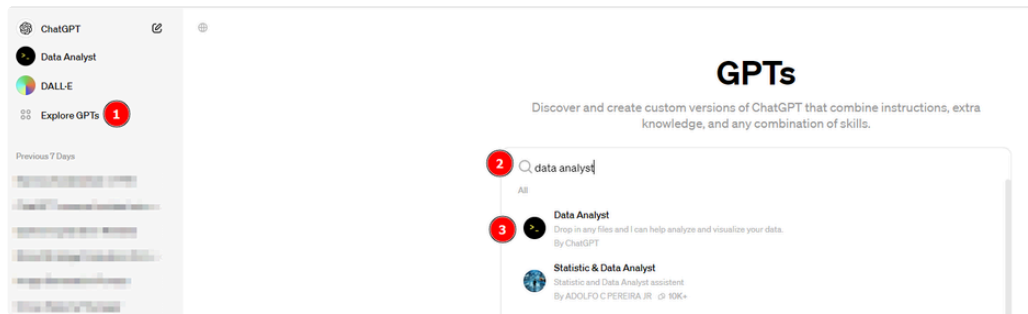


ChatGPT Data Analyst

[How to get to the Data Analyst GPT](#)

[ChatGPT Data Analyst Example](#)

How to get to the Data Analyst GPT



ChatGPT Data Analyst Example

[Make a graph of how each sales team is performing compared to the rest. Use data_dictionary.csv to understand the structure of the data](#)

First, I'll read the `data_dictionary.csv` file to understand the structure of the data and see how the datasets relate to each other. Let's start by looking at the contents of this file.

The data dictionary outlines the structure for various tables:

1. Accounts Table:

- Contains company-related details such as `account` (name), `sector`, `year_established`, `revenue`, `employees`, `office_location`, and `subsidiary_of`.

2. Products Table:

- Contains product-related details such as `product` (name), `series`, and `sales_price`.

3. Sales Teams Table:

- Contains details about sales agents and their management hierarchy, including `sales_agent`, `manager`, and `regional_office`.

4. Sales Pipeline Table:

- Holds data on sales opportunities including `opportunity_id`, `sales_agent`, `product`, `account`, `deal_stage`, `engage_date`, `close_date`, and `close_value`.

Next, I'll look into how each sales team is performing. For this analysis, I'll need to:

- Combine data from the `sales_teams` and `sales_pipeline` tables.
- Focus on `close_value` as a measure of performance.
- Summarize the data by each sales team (i.e., by `manager` or `sales_agent`).

Let's start by loading and examining the contents of the `sales_teams.csv` and `sales_pipeline.csv` files.

Now that we have the `sales_teams` and `sales_pipeline` data loaded, we can see the following structures:

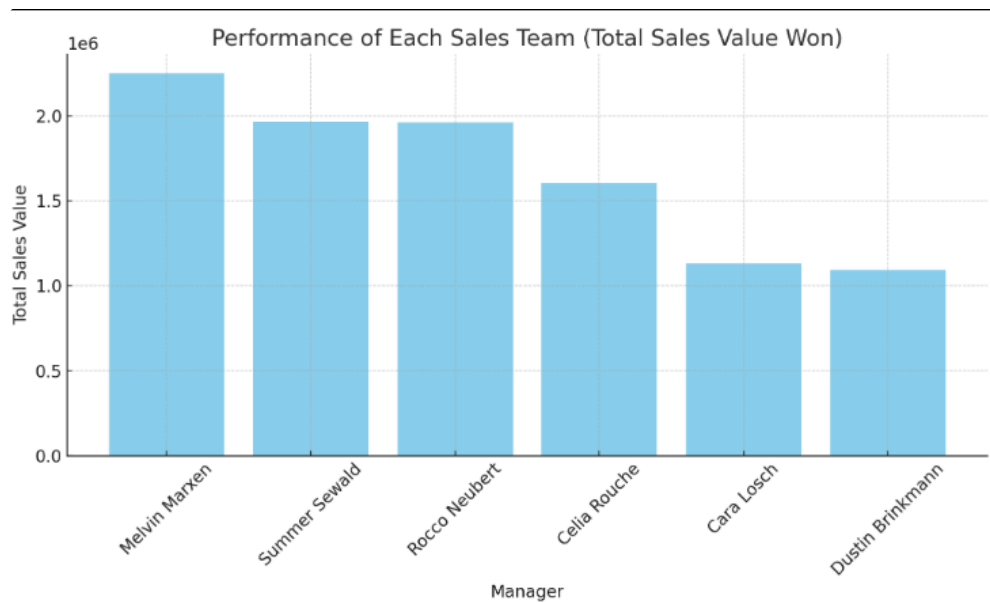
- Sales Teams Table:** Lists sales agents, their managers, and regional offices.

- **Sales Pipeline Table:** Provides details on each sales opportunity, including the agent handling it, the product involved, the client account, the stage of the deal, and the deal value.

The next step is to merge these tables to align the performance metrics (like `close_value`) with the respective teams. We'll focus on `sales_agent` to join these tables and then aggregate the performance by manager to see how each sales team is doing.


Let's proceed with merging the data and calculating the total sales per manager.

The data shows the total sales (`close_value`) achieved by each manager's team. Next, I will create a bar graph to visually compare the performance of these sales teams. Let's generate that graph now.



Here is the bar graph showing the performance of each sales team, represented by the total sales value achieved under each manager. As you can see, Melvin Marxen's team leads with the highest sales, followed closely by Summer Sewald and Rocco Neubert.

i Click the **View Analysis** button to see the Python code and other behind-the-scenes data after a chat response

Here is the chart displaying the top 5 selling product categories by total revenue. This visualization should help you see which categories are generating the most revenue. **View analysis** of monetary value. If you need further analysis or adjustments to the chart, just let me know! [-] 

[Make a graph of percentage of opportunities won by each sales team](#)

Data Analyst

To create a graph showing the percentage of opportunities won by each sales team, we'll need to:

1. Calculate the total number of opportunities managed by each manager's team.
2. Determine how many of these opportunities resulted in a win.
3. Compute the percentage of won opportunities for each team.

Let's start by calculating these metrics.

Now that we have the percentage of opportunities won by each manager's team, I'll plot this data on a bar graph to visually compare the performance across teams. Let's generate that graph.



Here's the bar graph illustrating the percentage of opportunities won by each sales team, managed under different managers. Rocco Neubert's team stands out with the highest win rate, followed by Cara Losch and Summer Sewald. This view helps in understanding not just the volume of sales, but also the efficiency in closing deals successfully.

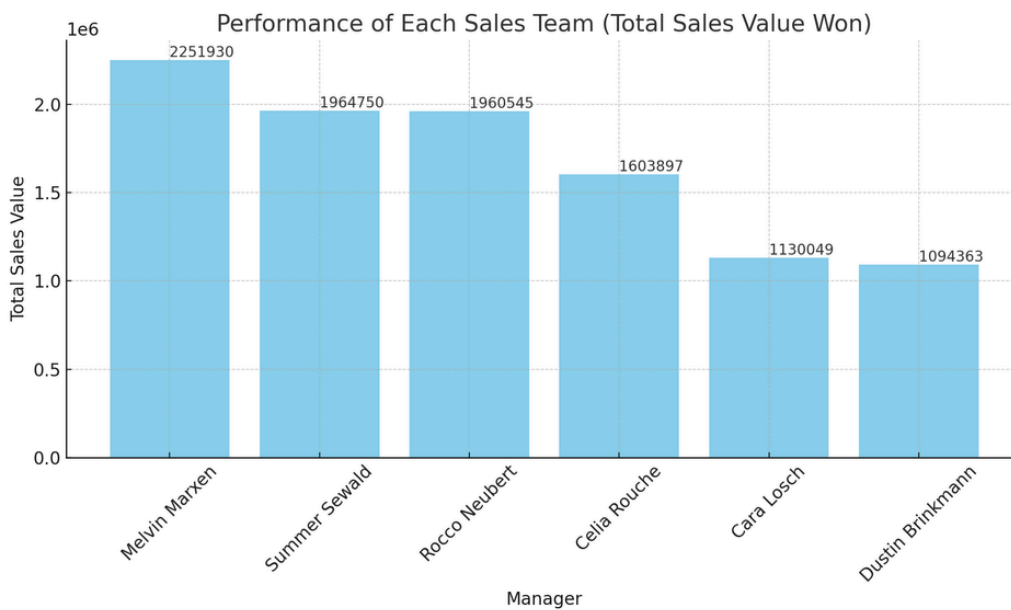
If there's anything more you'd like to explore or another aspect of the data you're curious about, just let me know!

You

[Add data labels to both charts](#)

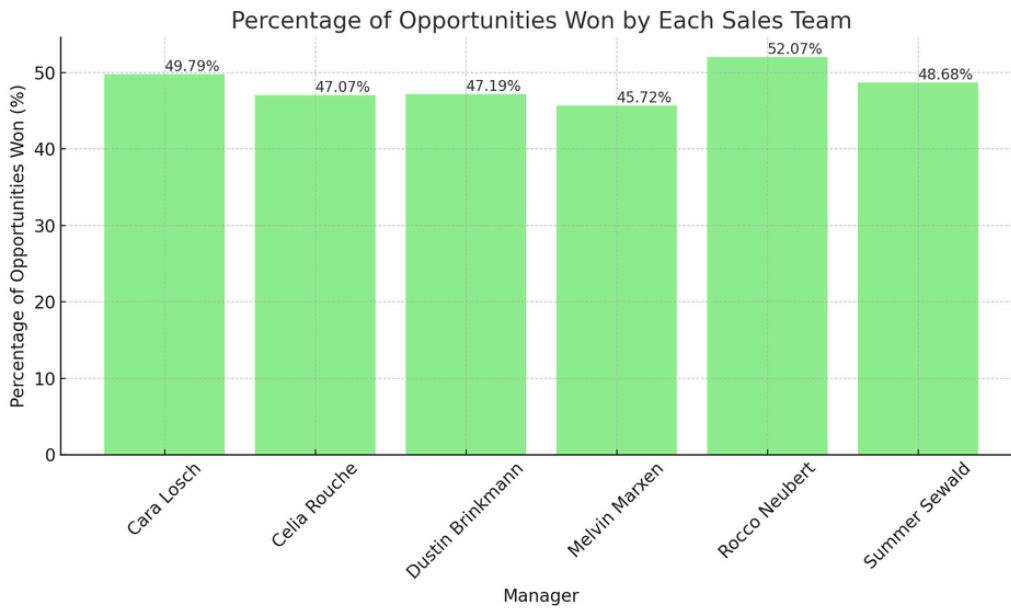
Data Analyst

Let's update both charts to include data labels that display the numerical values directly on the graphs. I'll start by revising the chart that shows total sales value by each sales team, and then I'll update the chart showing the percentage of opportunities won. Let's do that now.



Here's the updated bar graph showing the performance of each sales team, with data labels now added to display the total sales value directly on the bars.

Next, I'll update the graph for the percentage of opportunities won by adding data labels to that as well. Let's do that.



The chart showing the percentage of opportunities won by each sales team has been updated with data labels, providing clear visibility into the exact percentages for each team.

These enhancements should provide a comprehensive view of the sales teams' performances. If there's anything else you'd like to adjust or analyze further, just let me know!

You

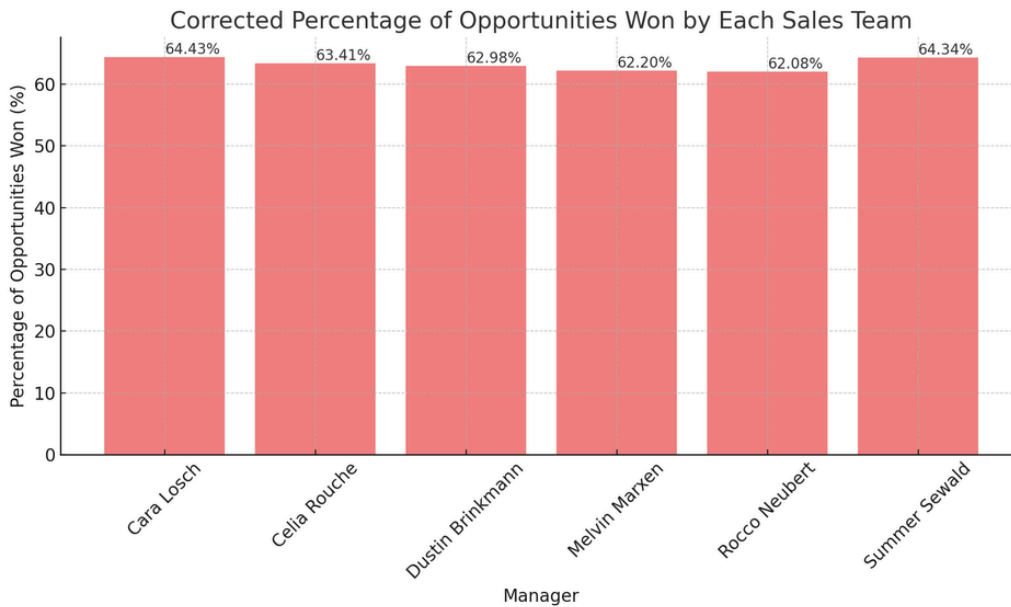
Your percentage of opportunities won chart doesn't look correct. Only calculate the percentage based on deals Won and Lost. Ignore other deal_stage values

Data Analyst

Let's correct the calculation of the percentage of opportunities won by excluding any deal stages other than "Won" and "Lost". This will provide a more accurate representation of the success rate for each team.

I'll adjust the data filtering and then recalculate the percentages accordingly. Let's proceed with that update.

The percentages of opportunities won have been recalculated to consider only "Won" and "Lost" deal stages. Now, let's update the bar graph with this corrected data and include data labels for better readability.



Here's the updated bar graph showing the corrected percentages of opportunities won by each sales team, now accurately reflecting only deals that were either "Won" or "Lost". The data labels on the graph provide the exact percentages.

This should give a more precise view of each team's effectiveness in closing deals. Let me know if there's anything else you'd like to modify or explore further!

You

Can you identify any quarter-over-quarter trends?

Data Analyst

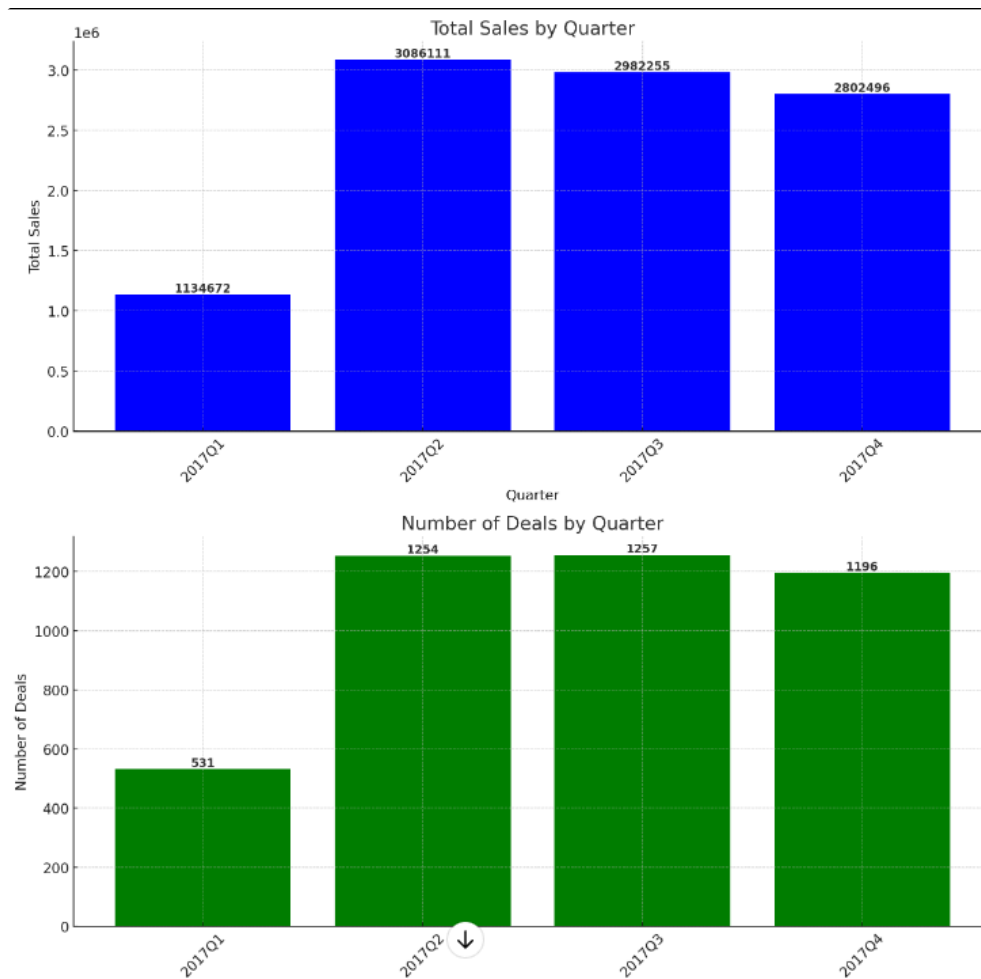
To identify quarter-over-quarter trends in the sales data, we'll need to analyze changes in sales performance over time using the close dates from the sales pipeline data. Here's what I'll do:

1. **Extract the year and quarter from the close dates.**
2. **Aggregate the total sales value and number of won deals by quarter and year.**
3. **Create visualizations to display these trends.**

Let's start by extracting the year and quarter from the close dates and preparing the data for analysis.

We have the sales data aggregated by quarter for the year 2017, showing total sales and the number of deals won for each quarter. Now, I'll create visualizations to show these quarter-over-quarter trends in both the total sales and the number of deals closed.

Let's visualize this data to observe the trends.



Here are the quarter-over-quarter trends for 2017:

- **Total Sales by Quarter:** There's a significant rise in total sales from Q1 to Q2, followed by a slight decline in Q3 and further in Q4. This suggests a peak in sales performance mid-year.
- **Number of Deals by Quarter:** The number of deals won follows a similar pattern to total sales, increasing substantially from Q1 to Q2, then slightly declining through Q3 and Q4.

These trends could reflect seasonal effects, sales initiatives, or market changes. Understanding these patterns can help in planning and forecasting for future periods.

If you'd like to delve deeper into specific aspects of these trends or extend the analysis to other years or metrics, just let me know!

You

[For number of deals by quarter, I'm getting different results in PowerBI. I see 647 for Q1 for example](#)

Data Analyst

Let's troubleshoot the discrepancy in the "Number of Deals by Quarter" count that you're seeing between my analysis and what you've observed in PowerBI. It could be due to several reasons such as data filtering differences, or the exact method used to count deals.

I'll first re-check our current method of counting deals by quarter. Let's verify that the date conversions and aggregations were performed correctly, focusing on ensuring the count of 'Won' deals matches your observations. We'll specifically look into the first quarter of 2017.

According to the data processing here, there are indeed 531 deals that were marked as "Won" in Q1 2017. If PowerBI is showing a different number (647), it could be due to one or more of the following reasons:

1. **Data Filtering Differences:** Make sure that both datasets are filtering on the same criteria. For instance, is PowerBI possibly counting deals that were closed but not marked as "Won" (e.g., deals that are in a pending state)?
2. **Deal Stage Definitions:** Verify that the definition of a "won" deal is the same in both analyses.
3. **Date Handling:** Double-check how the dates are formatted and interpreted in PowerBI. Sometimes, date parsing can lead to different results if the settings or formats are inconsistent.

To align these numbers, you might want to confirm the filters and conditions applied in PowerBI. If needed, I can help with more specific checks or analyses based on additional data or criteria you may have.

You

Re-do the number of deals per quarter, including all deals that closed in that quarter

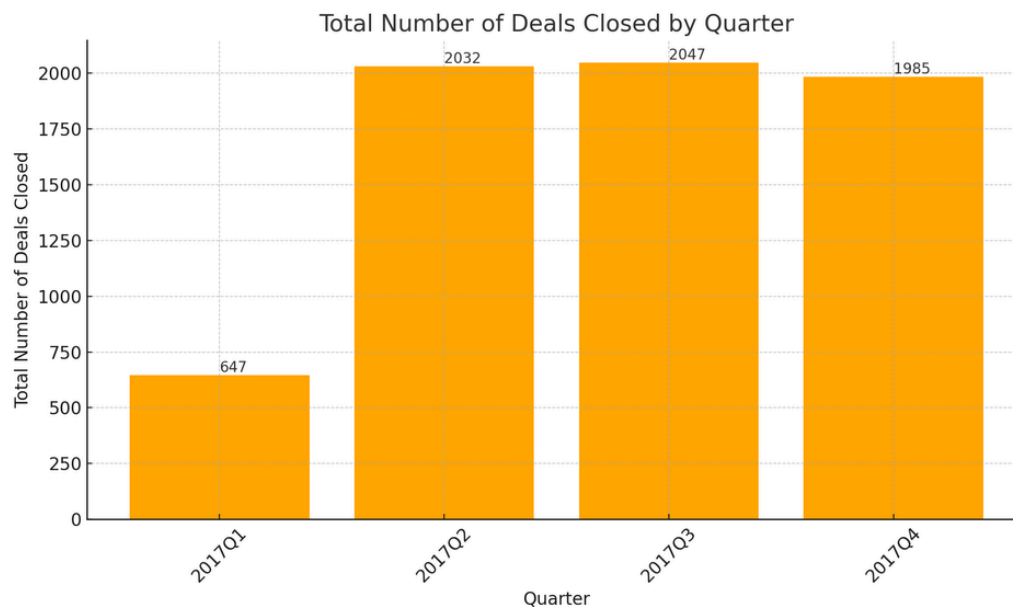
Data Analyst

Sure, let's revise our approach to count all deals that closed in each quarter, regardless of the deal stage. This will give us the total number of deals concluded in each quarter.

I'll calculate the count of all closed deals by quarter and year. Let's proceed with that.

The revised counts are now showing for each quarter of 2017, including all closed deals, regardless of their final deal stage. It looks like there were 647 deals closed in Q1 2017, which matches the number you mentioned from PowerBI.

Next, I'll create a bar chart to visualize these total deals per quarter. Let's plot this.



Here's the bar graph showing the total number of deals closed each quarter for the year 2017. The data labels directly on the bars provide the exact counts. This visualization confirms the closure activities over the quarters, with a steady increase from Q1 to a peak in Q3, and a slight decrease in Q4.

If there's anything more specific you'd like to analyze or another angle you'd like to explore, feel free to let me know!

You

Do any products have better win rates?

Data Analyst

To determine if any products have better win rates, we'll analyze the win rate of each product based on the sales pipeline data. We'll calculate the win rate as the ratio of deals won to the total number of deals (won and lost) for each product.

Here's how we'll proceed:

1. **Filter the data to include only 'Won' and 'Lost' deals** since these are the ones that contribute to a win or loss outcome.
2. **Group the data by product** and calculate the number of won and lost deals for each.
3. **Compute the win rate for each product.**

Let's calculate this now.

Here are the win rates for each product, calculated as the percentage of deals won out of the total deals that were either won or lost:

- **MG Special** has the highest win rate at approximately **64.84%**.
- **GTX Plus Pro** follows closely at about **64.30%**.
- **GTX Basic** and **GTXPro** also have win rates above 63%.

Products like **MG Advanced** and **GTK 500** have slightly lower win rates, around 60%.

This analysis reveals that some products tend to have higher success rates in closing deals, which might indicate greater market acceptance or better alignment with customer needs.

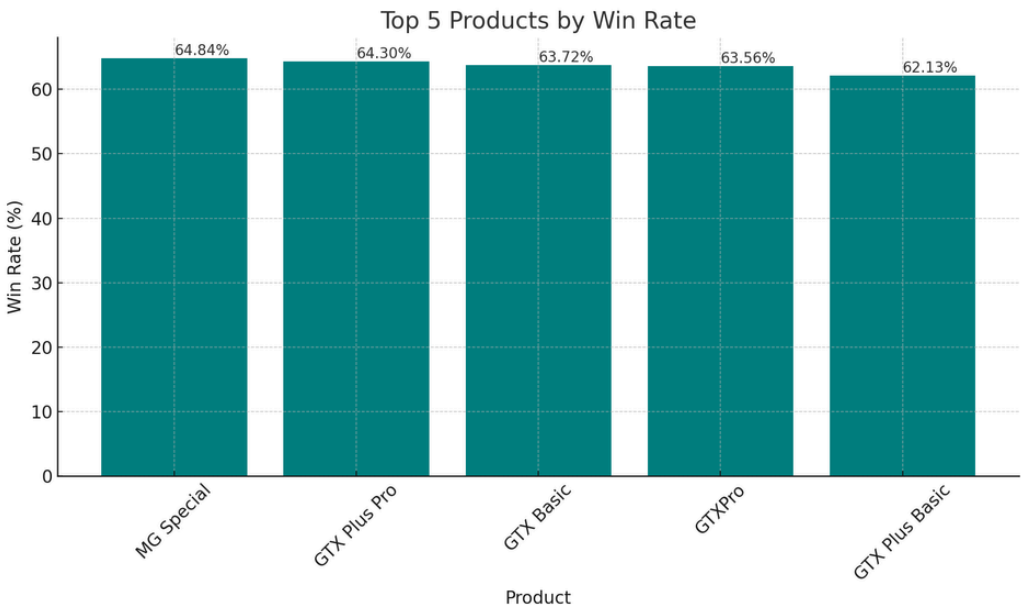
If you'd like further analysis on why certain products perform better or additional insights into any specific product, let me know!

You

[Make a graph of the top 5 products](#)

Data Analyst

Let's create a graph to visually represent the win rates of the top 5 products with the highest win rates. This will provide a clearer view of which products are performing best in terms of closing deals. Here we go:



Here is the bar graph showing the win rates of the top 5 products. The graph clearly illustrates how products like MG Special and GTX Plus Pro lead in success rates when closing deals, with win rates exceeding 64%.

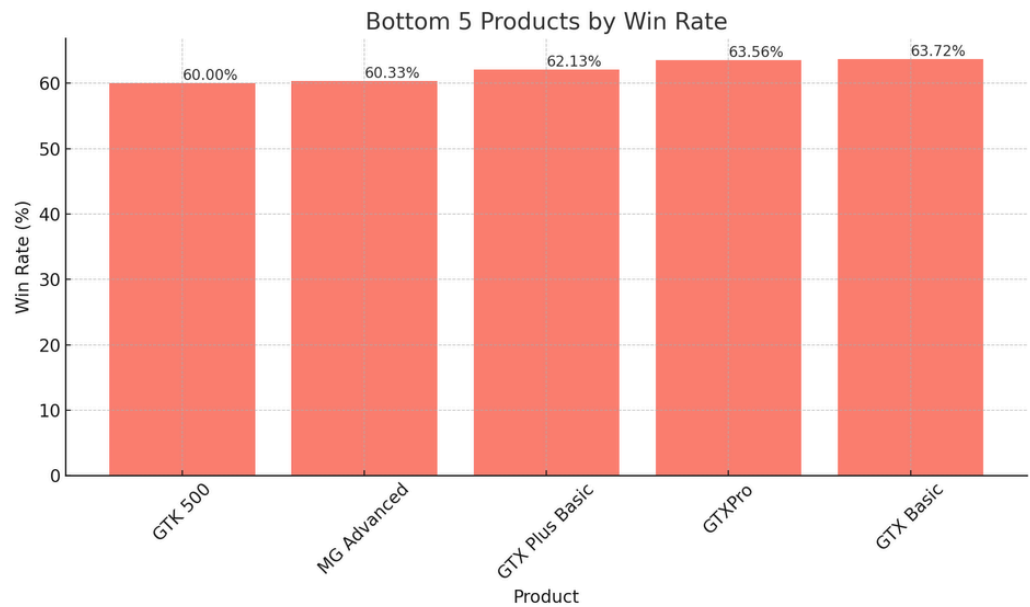
This visualization helps in identifying which products might warrant more focus in sales strategies or further investigation into their market strengths. If you have any more questions or need additional data analyzed, feel free to ask!

You

Show a graph of the bottom 5 products

Data Analyst

Let's visualize the win rates for the bottom 5 products, those with the lowest success rates in closing deals. This can help identify products that might require further marketing support, adjustments in sales tactics, or product enhancements. Here's the graph for the bottom 5:



Here's the bar graph showcasing the win rates for the bottom 5 products. This visual helps to clearly identify products like GTK 500 and MG Advanced, which have the lowest success rates, both hovering around 60%.

Analyzing these products further could provide insights into potential areas for improvement or adjustments in strategy. If there's more detailed analysis you'd like to perform on these products or additional aspects to explore, just let me know!

You

Are any sales agents lagging behind?

Data Analyst

To determine if any sales agents are lagging behind in terms of their sales performance, we'll analyze the win rates and total sales values for each sales agent. This will give us an idea of who might be underperforming compared to their peers.

Here's how we'll proceed:

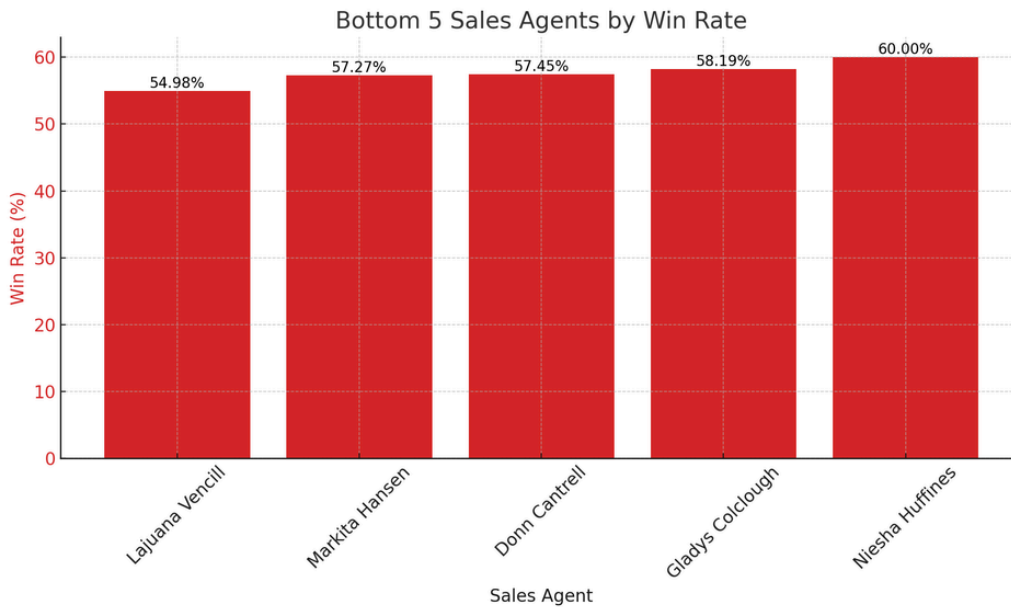
1. **Calculate the win rate for each sales agent** (number of deals won divided by the total number of deals that were either won or lost).
2. **Calculate the total sales value for each sales agent** to see their contribution in terms of revenue.

Let's start by calculating these metrics.

Here are the metrics for each sales agent, including their win rates and total sales values. For a clearer view of those who might be lagging behind, we'll plot this data, focusing on win rates and total sales values.

Since there are many agents, I'll visualize the bottom 5 in terms of win rate to identify those who might need additional support or training. Let's generate that graph.

The bar graph highlights the bottom 5 sales agents by win rate, providing a clear view of who might be underperforming compared to their peers. If you want to discuss strategies to help these agents or if you need further analysis on other metrics, just let me know!



You

Count how many sales agents each manager has with a win rate of 65% or better

Data Analyst

To determine how many sales agents each manager has who are performing with a win rate of 65% or better, I'll filter the dataset to include only those agents, group by manager, and count the number of such agents for each manager.

Let's perform this analysis.

Each manager has the following number of sales agents with a win rate of 65% or better:

- **Dustin Brinkmann:** 3 agents
- **Cara Losch:** 2 agents
- **Celia Rouche:** 2 agents
- **Rocco Neubert:** 2 agents
- **Summer Sewald:** 2 agents

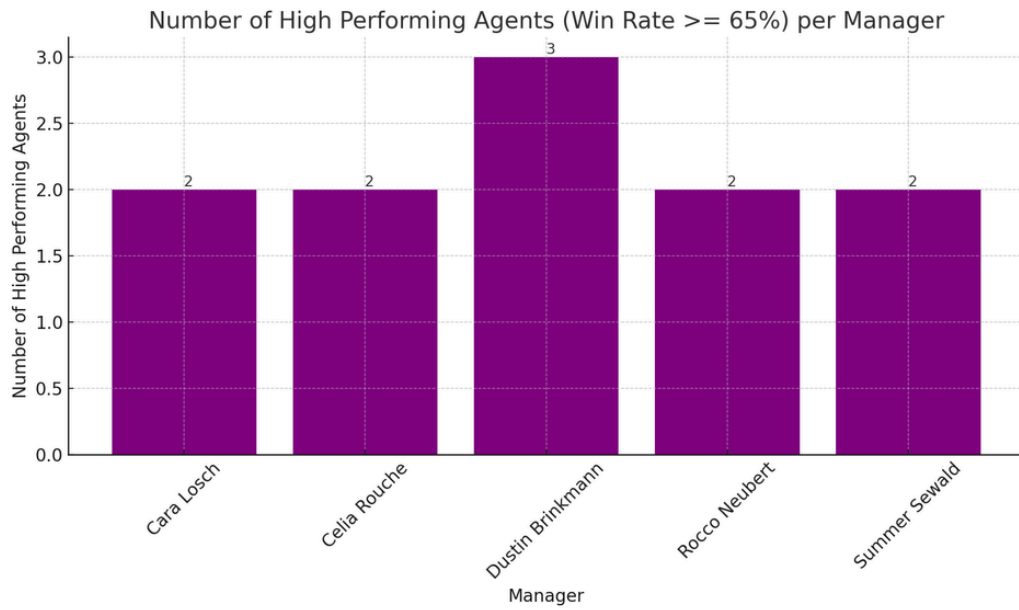
This information could be helpful in assessing the strength of each manager's team and potentially in identifying best practices that could be shared across teams. If you need more detailed analysis or further breakdowns, feel free to ask!

You

Make a graph of that

Data Analyst

Let's create a bar graph to visualize the number of high-performing sales agents (with a win rate of 65% or better) for each manager. This will provide a clear comparison of managerial performance in terms of nurturing top talent. Here's the graph:



The bar graph above illustrates the number of high-performing sales agents (with a win rate of 65% or better) managed by each manager. Dustin Brinkmann leads with three high-performing agents, while the other managers each have two.

This visualization can help in understanding managerial effectiveness in developing successful sales teams. If there are any more aspects you'd like to explore or other insights you need, just let me know!