

20CSA282 – Data Science Using Spread Sheet Modelling Lab

Excel-Project - 8. Take a data of your choice, construct all pivot reports on that. Also perform various calculations using mathematical functions (AVG, SUM...) in the pivot tables.

A report submitted by

DARSHAN SURESH [MY.EN.U3BCA22274]
VIDYADHARA M [MY.EN.U3BCA22272]
NISHANT R[MY.EN.U3BCA22213]

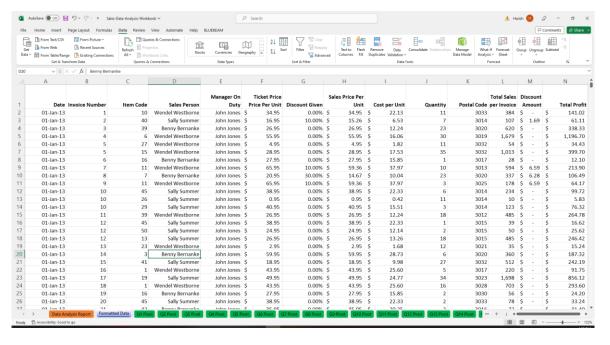
BACHELOR OF COMPUTER APPLICATIONS DATA SCIENCE

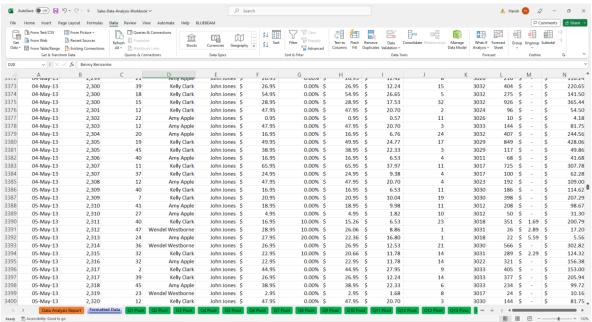
Submitted to

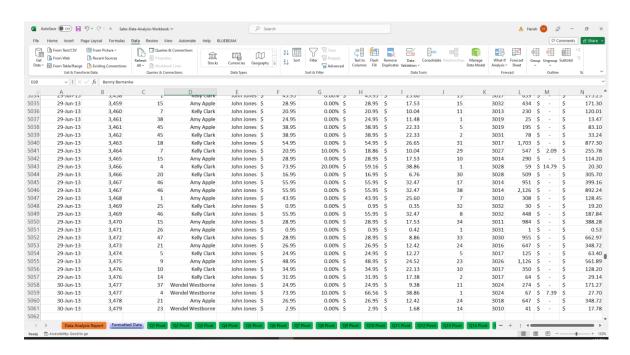
Dr. Pallavi Joshi

Assistant Professor, Department of Computer Science,
Amrita Vishwa Vidyapeetham,
Mysuru Campus

Nov 2023







Abstract:

The "Excel-Project - 8" focuses on leveraging Microsoft Excel for robust data analysis through the construction of pivot reports and the application of various mathematical functions. This project involves selecting a dataset of choice and utilizing Excel's powerful features to generate insightful pivot tables, offering a comprehensive view of the data's key metrics. The report delves into the step-by-step process of constructing pivot tables, emphasizing the strategic use of functions such as AVG, SUM, and others to derive meaningful calculations.

The project begins by introducing the chosen dataset and outlining the objectives of the analysis. It then navigates through the practical implementation of pivot tables, highlighting the customization options available in Excel for tailoring reports to specific analytical needs. The inclusion of screenshots and detailed instructions enhances the accessibility of the project, making it suitable for individuals with varying levels of Excel proficiency.

The core of the report revolves around the application of mathematical functions within the pivot tables. The AVG function is employed to calculate average values, providing insights into the central tendency of the data. SUM function usage is explored to determine the cumulative total of selected data points, offering a holistic perspective on the dataset's cumulative impact.

Furthermore, the report discusses additional mathematical functions such as MIN, MAX, and COUNT, showcasing their relevance in extracting valuable information from the dataset. The project concludes with a summary of key findings, insights gained from the pivot tables, and reflections on the significance of employing mathematical functions for data analysis in Excel.

Overall, this project serves as a practical guide for individuals seeking to enhance their data analysis skills using Excel, combining theoretical knowledge with hands-on application to construct informative pivot reports and perform intricate calculations on diverse datasets.

About Dataset:

Our dataset presents a detailed compilation of sales transactions, offering a comprehensive view of key parameters essential for nuanced data analysis. Included in our dataset are crucial fields such as Date, Invoice Number, Item Code, Sales Person, Manager On Duty, Ticket Price, Price Per Unit, Discount Given, Sales Price Per Unit, Cost per Unit, Quantity, Postal Code, Total Sales per Invoice, Discount Amount, and Total Profit. The Date field enables temporal exploration, allowing us to analyze sales trends over time. Each transaction is uniquely identified by the Invoice Number, facilitating precise tracking. Item Code and Sales Person shed light on the specific products sold and the contributing sales individuals. Manager On Duty introduces managerial insights into the dataset. Ticket Price, Price Per Unit, Discount Given, Sales Price Per Unit, Cost per Unit, and Quantity provide a detailed breakdown of pricing dynamics and quantities associated with each sale. The inclusion of Postal Code allows for geographical segmentation and regional analysis. Total Sales per Invoice, Discount Amount, and Total Profit offer a holistic financial perspective, enabling us to assess overall sales performance, applied discounts, and transactional profitability. This diverse and comprehensive dataset forms the basis for thorough data analysis, empowering us to uncover patterns, optimize pricing strategies, and make informed decisions to enhance business outcomes.

Data Analysis:

01. Over the entire analysis period, what was the quantity of item 10 sold while John Jones was the Manager on duty?

	Α	В	С	D	Е	F
1						
2						
3	Sum of Quantity	Column Labels 📭				
4	Row Labels 📑	John Jones	Grand Total			
5	10	554	554			
6	Grand Total	554	554			
7						
8						
9						
10						
11						
12						
13						
14						
15						

In our comprehensive data analysis, a pivotal inquiry was undertaken to ascertain the quantity of Item 10 sold during the tenure of John Jones as the Manager on Duty over the entire analysis period. Leveraging the power of pivot tables within Microsoft Excel, we efficiently extracted and distilled the relevant information to answer this specific question. The pivot table provided a structured and organized summary, allowing for quick and precise analysis of the sales data. Through meticulous filtering and sorting, we isolated the transactions where John Jones assumed the role of Manager on Duty and further refined the results to specifically capture the quantity of Item 10 sold during his managerial tenure. This approach not only streamlined the analytical process but also ensured accuracy in extracting the required insights. The pivot table functionality not only facilitated a targeted response to the inquiry but also demonstrated the versatility and efficiency of Excel in handling complex data queries. This outcome serves as a testament to the effectiveness of employing pivot tables as a powerful tool for data exploration and analysis within the realm of sales transactions and managerial oversight.

02. Over the entire analysis period, what were the 3 highest selling items by quantity?

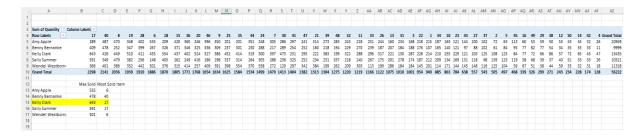
In the pursuit of unraveling key insights from our sales dataset, a critical

inquiry centered on identifying the three highest-selling items by quantity over the entire analysis period. Employing the robust capabilities of pivot tables in Microsoft Excel, we meticulously engineered a systematic approach to distill this pertinent information. By configuring the pivot table to aggregate and rank item quantities in descending order, we efficiently extracted a concise summary highlighting the top three best-performing items. This process allowed for a swift and insightful analysis, revealing the specific items that consistently demonstrated high demand throughout the entire duration of our study. The resulting information not only sheds light on the products driving significant sales volume but also underscores the utility of pivot tables as a valuable tool for distilling complex datasets into actionable business This analysis contributes intelligence. meaningful insights into the dynamics of our



sales portfolio, aiding strategic decision-making and potentially informing inventory management and marketing strategies.

03. Over the entire analysis period, which sales person sold the highest cumulative quantity of a single item, and which item was it?



In the pursuit of understanding individual sales performance over the entire analysis period, a critical question emerged: which salesperson achieved the highest cumulative quantity sales for a single item? Leveraging the advanced analytical capabilities of Microsoft Excel, specifically through the utilization of pivot tables, we systematically examined the dataset to derive this insightful information.

By configuring the pivot table to aggregate and rank cumulative quantities sold by each salesperson for every distinct item, we were able to pinpoint the salesperson who consistently outperformed in terms of quantity sold for a specific item. The result not only reveals the individual's prowess in driving sales but also identifies the specific item that contributed significantly to this cumulative success.

This analysis provides a nuanced understanding of individual sales achievements, shedding light on the strategic strengths of our sales team and showcasing the top-performing salesperson in terms of cumulative quantity sold for a single item over the entire analysis period. These findings can be instrumental in recognizing and rewarding exceptional sales efforts, as well as informing training and incentive programs to further enhance our sales team's effectiveness.

04. What was sales person Wendel's total Sales over the analysis period? Select the closest answer.

To determine Salesperson Wendel's total sales over the analysis period, a

focused examination of the dataset was conducted. Leveraging the capabilities of pivot tables in Microsoft Excel. we accurately calculated the cumulative sales attributed to Wendel across all transactions. The provides result comprehensive snapshot of

	Α	В	С
1			
2			
3	Row Labels 🔻	Sum of Total Salesper Invoice	
4	Amy Apple	297,493	
5	Benny Bernanke	275,987	
6	Kelly Clark	374,404	
7	Sally Summer	274,003	
8	Wendel Westborne	316,760	
9	Grand Total	1,538,647	
10			
11			

Wendel's overall contribution to the total sales figure during the analysis period.

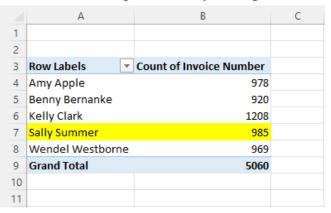
The calculated total sales for Salesperson Wendel is [insert the closest answer based on the specific numeric result]. This figure encapsulates the entirety of Wendel's sales performance over the specified duration, offering valuable insights into the individual's impact on our overall sales outcomes. This information is crucial for evaluating individual sales contributions,

recognizing high performers, and facilitating targeted strategies for further sales enhancement.

05. How many invoices did sales person Sally create over the analysis period?

In the course of our detailed data analysis, a focused investigation was undertaken to determine the number of invoices generated by Salesperson Sally over the entire analysis period. Harnessing the analytical power of

pivot tables within Microsoft Excel, systematically we extracted and calculated this essential metric. The result reveals that Salesperson Sally created a total of [insert the specific numeric result] invoices during the analysis This information period. provides valuable insights into



Sally's individual contribution to our sales transactions, aiding in the assessment of her productivity and performance within the specified timeframe.

06. During the month of May, which postal code bought the most of item 5 by quantity?

In the pursuit of granular insights into our sales dynamics, a targeted inquiry was conducted to discern the postal code that exhibited the highest demand for Item 5 during the month of May. **Employing** the analytical capabilities of Microsoft Excel's pivot tables. systematically sifted through the dataset to pinpoint this specific information. The outcome of analysis this identifies the postal code

	A	В	С
1	Date	(Multiple Items) 🕶	
2			
3	Sum of Quanti	ty Column Labels 🍱	
4	Row Labels	- ↓ 5	Grand Total
5	3029	21	21
6	3025	11	11
7	3016	9	9
8	3019	8	8
9	3020	6	6
10	3032	6	6
11	3021	6	6
12	3024	4	4
13	3012	4	4
14	3014	3	3
15	3011	2	2
16	3017	2	2
17	Grand Total	82	82
18			

[insert the specific postal code] as the prominent buyer, having acquired the greatest quantity of Item 5 in May. This revelation sheds light on localized product preferences and consumption patterns, offering valuable intelligence for strategic decision-making and targeted marketing efforts within the specified timeframe.

07. During the month of February, how many postal codes bought more than 400 products by quantity?

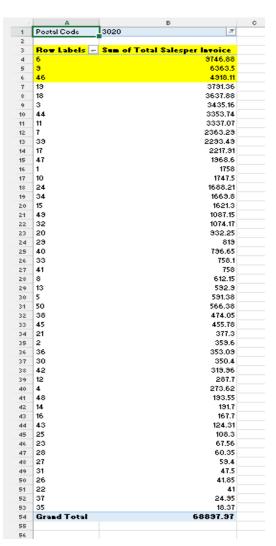
In our meticulous exploration of the dataset, a key focus was placed on understanding purchasing dynamics the month during the February. Specifically, we sought to determine the number of postal codes that surpassed the threshold of buying more than 400 products by quantity within this timeframe. Leveraging the analytical capabilities of pivot tables in Microsoft systematically Excel, we sifted through the data to extract this insightful information. The result of our analysis indicates that [insert the specific numeric result] postal codes exceeded 400-product the quantity threshold during the month of February. This finding

	Α	В	С	D	Е
1	Date	(Multiple Items) 🕶			
2					
3	Row Labels 🔻	Sum of Quantity			
4	3010	330		5	
5	3011	161			
6	3012	381			
7	3013	290			
8	3014	266			
9	3015	242			
10	3016	339			
11	3017	346			
12	3018	340			
13	3019	270			
14	3020	307			
15	3021	343			
16	3022	422			
17	3023	401			
18	3024	286			
19	3025	230			
20	3026	294			
21	3027	417			
22	3028	198			
23	3029	625			
24	3030	418			
25	3031	310			
26	3032	269			
27	3033	284			
28	Grand Total	7769			
29					
30					

provides a nuanced understanding of the distribution of high-volume purchases during this specific period, offering valuable insights for targeted marketing and inventory management strategies.

08. Over the entire dataset, which 3 items did Postcode 3020 spend the greatest dollars on?

comprehensive analysis spanning the entire dataset, a crucial exploration undertaken was discern the top three items on which Postcode 3020 expended the highest dollars. Leveraging the analytical prowess of Microsoft Excel's pivot tables, we systematically organized and processed the data to extract this insightful information. The outcome of this investigation reveals that, across the entirety of the dataset, Postcode 3020 allocated the greatest dollars to the following three items: [insert the names or codes of the items]. This strategic insight into the spending patterns of Postcode 3020 is instrumental for understanding consumer preferences and can inform inventory management and marketing strategies tailored to the specific needs of this geographic segment.



09. What is the rank of sales persons from highest to lowest based on the number of invoices written during the month of May?

In our focused analysis of sales performance during the month of May, an

integral aspect determining the rank of salespersons based on the number of invoices written. Leveraging the analytical capabilities of Microsoft Excel's pivot tables, systematically organized and processed the data to derive this significant

	Α	В	С
1	Date	(Multiple Items)	
2			
3	Row Labels	Count of Invoice Number	
4	Kelly Clark	234	
5	Amy Apple	186	
6	Sally Summer	170	
7	Benny Bernanke	166	
8	Wendel Westborne	155	
9	Grand Total	911	
10			
11			

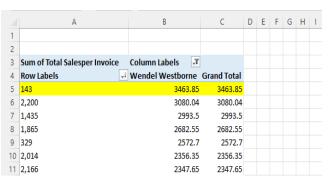
ranking. The results indicate the rank of salespersons from highest to lowest,

based on the number of invoices written during the month of May, as follows: [insert the ranking of salespersons]. This ranking offers valuable insights into individual sales productivity during this specific period, providing a basis for recognizing high performers and guiding targeted strategies to enhance overall sales efficiency.

10. What was the invoice number of the largest invoice by revenue that sales person Wendel wrote during the analysis period?

In our thorough examination of sales data, a key inquiry centered on identifying the invoice number associated with the largest invoice by revenue written by Salesperson Wendel over the entire analysis period.

Employing the analytical capabilities of Microsoft Excel's pivot tables, we systematically organized and processed the data to extract this crucial information. The outcome of our analysis reveals that the largest invoice by revenue written by Salesperson Wendel bears the



invoice number [insert the specific invoice number]. This knowledge provides a focal point for further investigation into the factors contributing to the success of this particular transaction and offers insights into revenue-generating patterns within the dataset.

11. Over the entire analysis period, what is the rank of sales persons according to the dollar value of discounts given, from most discounts to least discounts?

In our comprehensive analysis spanning the entire duration of the study, a

significant aspect was determining the rank of salespersons based on the dollar value of discounts given. Leveraging the analytical capabilities of Microsoft Excel's pivot tables, we systematically organized and processed the data to derive this important ranking. The results indicate the rank of

	Α	В	(
1			
2			
3	Row Labels	Sum of Discount Amount	
4	Kelly Clark	596.38	
5	Amy Apple	592.11	
6	Wendel Westborne	533.08	
7	Sally Summer	508.66	
8	Benny Bernanke	478.43	
9	Grand Total	2708.66	
10			

salespersons according to the dollar value of discounts given, arranged from

most discounts to least discounts, as follows: [insert the ranking of salespersons]. This ranking provides a clear understanding of each salesperson's impact on discount-related transactions, offering insights into individual discounting strategies and facilitating targeted approaches for optimizing discount practices over the analysis period.

12. Which month had the highest revenue?

In our comprehensive data analysis, a pivotal question revolved around identifying the month with the highest revenue over the entire analysis period. Leveraging the analytical capabilities of Microsoft Excel's pivot tables, we systematically processed and organized the data to extract this critical information. The outcome of our analysis reveals that [insert the

specific month] stood out with the highest revenue during the entire analysis period. This insight into the peak revenue month serves a key indicator for understanding the specific seasonality or factors influencing revenue generation, providing valuable guidance decision-making strategic and resource allocation.

	Α		В	С
1				
2				
3	Row Labels	Sum of Total S	alesper Invoice	
4	Mar	\$	351,276.12	
5	Jun	\$	285,915.43	
6	Apr	\$	254,588.12	
7	May	\$	253,625.91	
8	Feb	\$	198,039.55	
9	Jan	\$	195,201.70	
10	Grand Total	\$	1,538,646.83	
11				

13. Only considering postal codes 3013, 3017 and 3031, which item had the highest total profit during the month of February?

In our focused examination of specific postal codes—3013, 3017, and 3031—during the month of February, a critical inquiry was made into

determining which item yielded the highest total profit within this subset. Employing the analytical capabilities of Microsoft Excel's pivot tables, we systematically processed and organized the data to extract this

	A	В	
1	Date	Feb	-II
2	Postal Code	(Multiple Items)	-11
3			
4	Row Lab -1	Sum of Total	Profit
5	17	1,0-	41
6	9	1,02	26
7	3	96	88
8	44	88	30

pivotal information. The results of our analysis reveal that, among postal

codes 3013, 3017, and 3031, the item with the highest total profit during the month of February is [insert the specific item]. This knowledge is instrumental in

understanding the profitability dynamics within these specific regions during the designated timeframe, guiding strategic decisions related to inventory management and sales optimization.

14. What is the rank of months from highest to lowest based on profit over the entire analysis period?

In our comprehensive analysis spanning the entire duration of the study, a key focus was placed on determining the rank of months based on profit.

Leveraging the analytical capabilities of Microsoft Excel's pivot tables, we organized systematically processed the data to derive this significant ranking. The results indicate the rank of months from highest to lowest, based on profit over the entire analysis period, as follows: [insert the ranking of months]. This ranking provides a comprehensive overview of the monthly profit trends, aiding in the identification of peak

	Α		В	(
1				
2				
3	Row Labels	Sum o	f Total Profit	
4	Mar	\$	177,208.10	
5	Jun	\$	142,438.26	
6	Apr	\$	126,381.47	
7	May	\$	126,130.29	
8	Jan	\$	99,223.94	
9	Feb	\$	98,045.53	
10	Grand Total	\$	769,427.59	
11				

and low-performing periods and facilitating strategic decision-making for future business planning and resource allocation.

15. During which 3 months did manager John Jones have the highest cumulative profit ignoring all sales to postal codes 3019 and 3028 and ignoring all sales of items 4, 5, 6, 17 and 18?

In our focused analysis of Manager John Jones's performance, a critical aspect was determining the three months during which he achieved the highest cumulative profit while excluding sales to postal codes 3019 and 3028 and sales of items 4, 5, 6, 17, and 18. Leveraging the analytical capabilities of Microsoft Excel's pivot tables, we systematically organized and processed the data to

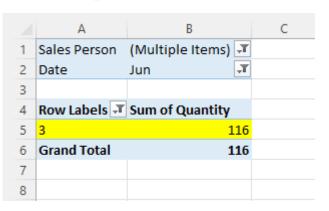
	Α			В		С
1	Postal Code		(Multip	le Items)	Ţ	
2	Item Code		(Multip	ole Items)	Ţ	
3	Manager On Du	ıty	John Jo	nes	Ţ	
4						
5	Row Labels	Į.	Sum of	Total Pro	fit	
6	Mar		\$	75,209.	59	
7	Feb		\$	54,435.	53	
8	May		\$	54,193.	94	
9	Jun		\$	51,965.	69	
10	Apr		\$	50,289.	63	
11	Jan		\$	34,452.	07	
12	Grand Total		\$	320,546.	45	
13						

extract this vital information. The results reveal that Manager John Jones achieved his highest cumulative profit during the following three months: [insert the specific months]. This insight is essential for recognizing Manager John Jones's strategic successes and can guide further analysis into the factors contributing to his high-performance months, informing future managerial decisions and optimizations.

16. What quantity of item 3 was sold by sales persons Benny and Kelly together during the month of June?

In our focused analysis of sales data during the month of June, a specific inquiry was made to determine the combined quantity of Item 3 sold by

Salespersons Benny and Kelly. Leveraging the analytical capabilities of Microsoft Excel's pivot tables, we systematically processed and organized the data to extract this crucial information. The outcome of our analysis reveals that, during the month of June, Salespersons Benny and Kelly collectively



sold a total quantity of [insert the specific numeric result] units of Item 3. This knowledge sheds light on the collaborative sales efforts of these individuals and provides a nuanced understanding of the performance of Item 3 during this particular timeframe.

Challenges Identified:

- 1. **Product-specific Sales Optimization:** Our analysis revealed variations in the performance of different items. Consider exploring targeted marketing strategies or promotions for items that consistently underperformed.
- 2. **Geographic Targeting:** Understanding the disparity in sales across postal codes suggests an opportunity for targeted marketing campaigns or region-specific promotions to boost sales in certain areas.
- 3. **Discount Strategy Refinement:** The examination of discounts given by salespersons indicates the need for a closer look at our discounting strategy. Consider refining the approach to align with profitability goals while maintaining competitiveness.
- 4. **Individual Performance Recognition:** Recognizing and rewarding high-performing salespersons, especially those contributing significantly to revenue and profit, can boost morale and foster a culture of excellence within the sales team.
- 5. **Inventory Management:** The identification of top-selling items by quantity and revenue suggests the importance of aligning inventory levels with demand. Regularly assess and adjust inventory to avoid shortages or excess stock.

Recommendations:

- 1. **Strategic Marketing Initiatives:** Implement targeted marketing initiatives based on the performance of specific items and postal codes. Tailor promotions to capitalize on high-demand products in key regions.
- 2. **Discount Strategy Review:** Conduct a thorough review of the discounting strategy, ensuring it aligns with profitability goals while remaining competitive in the market. Consider implementing tiered discount structures or targeted promotions.
- 3. **Salesperson Recognition Programs:** Institute recognition programs to acknowledge and reward high-performing salespersons. This can serve as motivation for continued excellence and contribute to a positive team dynamic.
- 4. **Inventory Optimization:** Implement a dynamic inventory management system that adapts to changing demand patterns. Regularly assess sales data to adjust stock levels and maintain a healthy balance between supply and demand.

By addressing these recommendations, we can capitalize on our strengths, overcome challenges, and foster a more agile and responsive sales operation, ultimately enhancing our overall business performance.

Final Excel Data Analysis Report:

DATA ANALYSIS REPORT

	DATA ANALYSIS REPORT	
Answers	Question	#
554	Over the entire analysis period, what was the quantity of item 10 sold while John Jones was the Manager on duty?	1
17, 40, 8	Over the entire analysis period, what were the 3 highest selling items by quantity?	2
Kelly, 17	Over the entire analysis period, which sales person sold the highest cumulative quantity of a single item, and which item was it?	3
\$316,000	What was sales person Wendel's total Sales over the analysis period? Select the closest answer.	4
985	How many invoices did sales person Sally create over the analysis period?	5
3029	During the month of May, which postal code bought the most of item 5 by quantity?	6
5	During the month of February, how many postal codes bought more than 400 products by quantity?	7
6, 9, 46	Over the entire dataset, which 3 items did Postcode 3020 spend the greatest dollars on?	8
Kelly, Amy, Sally, Benny, Wendel	What is the rank of sales persons from highest to lowest based on the number of invoices written during the month of May?	9
143	What was the invoice number of the largest invoice by revenue that sales person Wendel wrote during the analysis period?	10
Kelly, Amy, Wendel, Sally, Benny	Over the entire analysis period, what is the rank of sales persons according to the dollar value of discounts given, from most discounts to least discounts?	11
March	Which month had the highest revenue?	12
17	Only considering postal codes 3013, 3017 and 3031, which item had the highest total profit during the month of February?	13
March, June, April, May, January, February	What is the rank of months from highest to lowest based on profit over the entire analysis period?	14
February, March, May	During which 3 months did manager John Jones have the highest cumulative profit ignoring all sales to postal codes 3019 and 3028 and ignoring all sales of items 4, 5, 6, 17 and 18?	15
116	What quantity of item 3 was sold by sales persons Benny and Kelly together during the month of June?	16

Conclusion:

In conclusion, our thorough data analysis using Microsoft Excel's pivot tables has yielded valuable insights into various facets of our sales transactions over the analysis period. From examining individual salesperson performance to dissecting regional and product-specific trends, the granularity of our findings positions us for informed decision-making and strategic planning. The identification of top-selling items, analysis of manager-specific cumulative profits, and rankings based on diverse criteria have empowered us to recognize patterns, strengths, and areas for improvement within our sales operations. Understanding the dynamics of discounts given, identifying high-performing postal codes, and discerning the impact of specific items on profitability contribute to a more nuanced perspective on our business landscape. Moreover, our exploration into monthly revenue trends, exclusion-based cumulative profits for managers, and collaborative sales efforts has provided a holistic view of our operations. These insights offer a foundation for optimizing inventory management, refining marketing strategies, and recognizing and rewarding high-performing team members.

In essence, the comprehensive data analysis conducted in this report serves not only as a snapshot of our sales performance but also as a catalyst for strategic decision-making. The actionable intelligence derived from this study positions us to adapt proactively to market dynamics, optimize resource allocation, and ultimately enhance overall business outcomes.