  
**Logistics and Supply Chain Management Project Details**

Information about H001 Terminal Data and PartNumber All Terminals Data:-

Material\_Item\_No: Unique identification number for each material item.

Material\_Description: Description of the material.

Language: Language specification for the material.

Internal\_Material\_Number: Internal identification number for the material.

Plant\_Code: Code representing the plant associated with the material.

Part\_Number: Part number associated with the material.

Manufacturer\_Name: Name of the manufacturer of the material.

Material\_Group: Group classification for the material.

The data seems to represent diverse types of industrial components such as switches, relays, bearings, hydraulic components, and more. Each row corresponds to a specific material item with associated details.

Points to remember:-

1. Whenever you have solved any questions and if that questions have some formula ,dax or code or query then paste that formula here below that questions
2. Upload Excel files for evaluation of your answers

**You have to solve below Questions USING EXCEL(Using H001 Data):--**

***Based on Formula:--***

1. Question: What is the Manufacturer Name for Material Item No. 8000000340?

*=VLOOKUP(8000000340,$A$2:$G$99,7,0)*

1. Question: How many items have a Material Group starting with "ISP"?

*=COUNTIF(H:H,"ISP\*")*

1. Question: Calculate the average value of Internal Material Numbers.

*=AVERAGEIF(D:D,"<>0")*

1. Question: Which Material Item has the highest Material Group value?

*=INDEX(A:A, MATCH(MAX(IFERROR(MATCH(H:H, H:H, 0), -1)), IFERROR(MATCH(H:H, H:H, 0), -1), 0))*

1. Question: How many materials are from Plant Code "H020"?

*=COUNTIF(E:E,"H020")*

1. Question: Concatenate Material Description and Manufacturer Name for Material Item No. 8000000511.

*=TEXTJOIN("-",TRUE,IF(A:A=8000000511,B:B&"-"&G:G,""))*

1. Question: What is the Part Number for Material Group "ISPR73"?

*=TEXTJOIN(", ", TRUE, FILTER(F:F, H:H="ISPR73"))*

1. Question: Calculate the total count of materials with "ENGLISH" language.

*=COUNTIF(C:C,"ENGLISH")*

1. Question: Find the Language for Material Item No. 8000000598.

*=XLOOKUP(8000000598,A:A,C:C)*

1. Question: Identify the Material Item No. with the lowest Internal Material Number

*=INDEX(A:A, MATCH(MIN(D:D), D:D, 0))*

1. Question: Calculate the average length of Part Numbers.

*=AVERAGE(LEN(F:F))*

1. Question: Extract the last three characters from Plant Code for Material Item No. 8000000473.

*=RIGHT(INDEX(E:E, MATCH(8000000473, A:A, 0)), 3)*

1. Question: Determine the total number of unique Material Groups.

*=COUNTA(UNIQUE(H:H))*

1. Question: Which Material description has the highest Material Item No.?

*=IFERROR(INDEX(B:B, MATCH(MAX(A:A), A:A, 0)), "No match found")*

1. Question: Count the materials with a Part Number starting with "S-".

*=COUNTIF(F:F,"S-\*")*

1. Question: What is the Material Description for Material Group "ISPR40"?

*=TEXTJOIN(",",TRUE,(FILTER(B:B,H:H="ISPR40")))*

1. Question: Calculate the total number of materials manufactured by "IFM"

*=COUNTIF(G:G,"IFM")*

1. Question: Extract the first two characters from Language for Material Item No. 8000000685.

*=LEFT(INDEX(C:C,MATCH(8000000685,A:A,0)),2)*

1. Question: Identify the Material Item No. with the longest Material Description.

*=INDEX(A:A, MATCH(MAX(LEN(B:B)), LEN(B:B), 0))*

1. Question: Concatenate Plant Code and Manufacturer Name for Material Group "ISPR16".

*=TEXTJOIN(",",TRUE,IF(H:H="ISPR16",E:E&"-"&G:G,""))*

***Based on Condtional Formattings:-***

1. Question: Create a column named "Relay\_Type" that categorizes the relays as "DC" or "AC" based on the voltage in the "Material\_Description" column.

*=IFERROR(*

*IF(*

*AND(*

*ISNUMBER(SEARCH("DC", B2)),*

*ISNUMBER(SEARCH("AC", B2))*

*),*

*"Both AC/DC",*

*IF(*

*OR(*

*ISNUMBER(SEARCH("DC", B2)),*

*ISNUMBER(SEARCH("AC", B2))*

*),*

*IF(ISNUMBER(SEARCH("DC", B2)), "DC", "AC"),*

*"Not AC/DC"*

*)*

*),*

*""*

*)*

*Rule ----* *=AND(ISNUMBER(SEARCH("DC", $B2)), $I2<>"Both AC/DC")*

*=AND(ISNUMBER(SEARCH("AC", $B2)), $I2<>"Both AC/DC")*

1. Question: Identify the materials with a "Material\_Group" starting with "ISP."

=LEFT(H1, 3) = "ISP"

1. Question: Calculate the total number of materials for each "Manufacturer\_Name

To identify the count ,created new column =COUNTIF(G:G,G2) and used conditional formatting “format only cells that contain” greater than or equal to 5

1. Question: Highlight all materials with "Language" other than English.

=AND($C1<>"", NOT($C1="ENGLISH"))

1. Question: Determine the average length of "Material\_Description" for all materials.

*=len(M2)*

*then used conditional formatting =M2 > AVERAGE($M$2:$M$100)*

1. Question: Identify materials with "Part\_Number" starting with "S."

=LEFT(F1,1)="S"

1. Question: Count the number of materials in each "Material\_Group."

To identify the count ,created new column =COUNTIF(H:H,H2) and used conditional formatting “format only cells that contain” greater than 10

1. Question: Highlight all materials with "Internal\_Material\_Number" less than 9000074200.

“format cells only contain” less than 9000074200 after selecting the range

1. Question: Identify materials with "Material\_Description" containing the word "POWER."

=ISNUMBER(SEARCH("POWER", B2))

1. Question: Calculate the total count of materials per "Plant\_Code."

To identify the count ,created new column =COUNTIF(E:E,E2) and used conditional formatting “format only cells that contain” greater than 1

1. Question: Highlight all materials with "Material\_Description" longer than 30 characters.

=and(len(B1) >30,B1 <> “”)

1. Question: Identify materials where "Material\_Group" is "ISPR72" and "Language" is not English.

=AND(R2="ISPR72", C2<>"ENGLISH")

1. Question: Calculate the total count of material group for each "Language."

To identify the count ,created new column =COUNTIF(B:B, B2) and used conditional formatting “format only cells that contain” greater than 5

1. Question: Highlight all materials with "Material\_Group" starting with "ISPR7."

=isnumber(search(“ISPR7”,T2))

1. Question: Identify materials with "Manufacturer\_Name" equal to "ABB" or "IDEC."

=AND(OR(U2="ABB", U2="IDEC"), U2<>"")

1. Question: Calculate the percentage of "RELAY" materials in the entire list.

*=SUMPRODUCT(--(ISNUMBER(SEARCH("RELAY", B:B)))) / COUNTA(B:B) \* 100& "%"*

1. Question: Highlight all materials with "Material\_Item\_No" less than 8000000500.

“format only cells that contain” less than 8000000500 after selecting the range

1. Question: Identify materials with "Manufacturer\_Name" not equal to "ZPMC."

=AND(W2<>"ZPMC", W2<>"")

1. Question: Calculate the total count of materials for each "Language" and "Material\_Group" combination.

To identify the count ,created new column =COUNTIFS(C:C,C2,H:H,H2) and used conditional formatting “format only cells that contain” greater than 10

1. Question: Highlight all materials with "Material\_Group" equal to "ISPR73" and "Language" equal to "ENGLISH."

=AND(Y2="ISPR73", C2="ENGLISH")

***Based on Pivot Table :-***

Question 1: "Material Distribution by Manufacturer"

Instructions:

Create a pivot table to analyze the distribution of materials based on manufacturers.

Use the "Manufacturer\_Name" and "Material\_Item\_No" columns for the analysis.

Display the count of materials for each manufacturer.

Question 2: "Material Group Analysis"

Instructions:

Create a pivot table to analyze the distribution of materials based on material groups.

Use the "Material\_Group" and "Material\_Item\_No" columns for the analysis.

Display the count of materials for each material group.

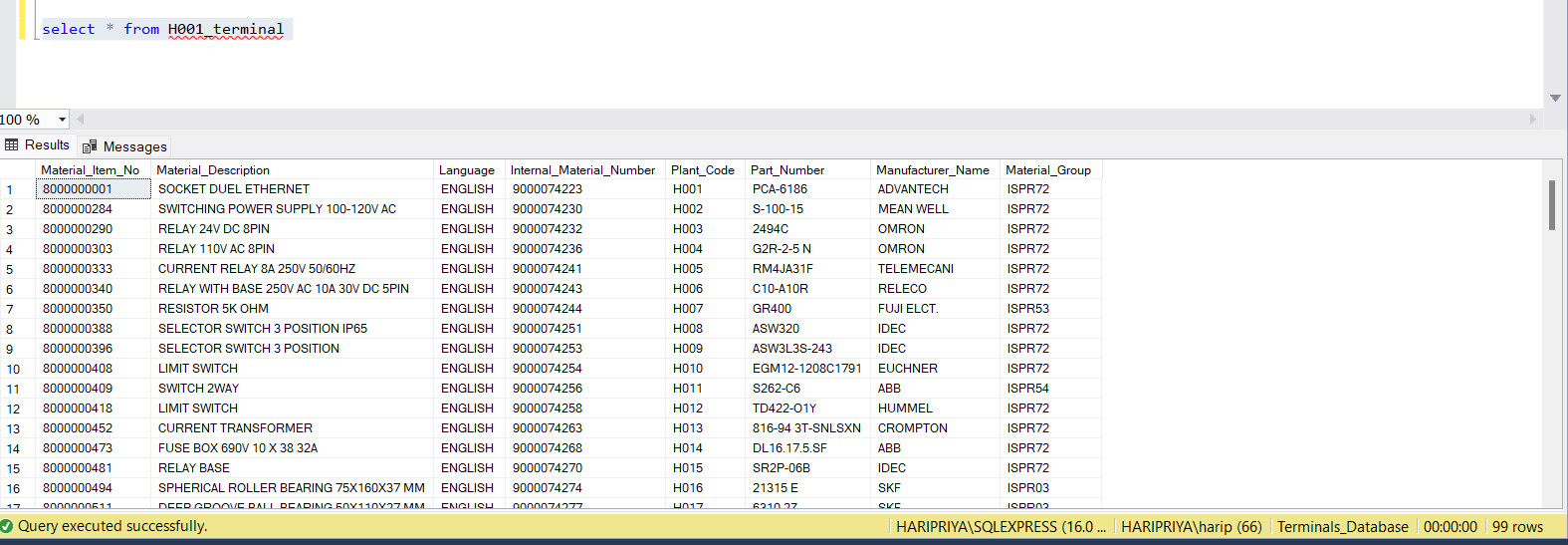
**You have to solve below Questions USING SQL(Using H001 and PartNumber\_ALLTerminals\_Data)**

Create a Database Name as **Terminals\_Database** inside that dump( insert the records or upload the excel files) which is given to you name as **H001 and PartNumber\_ALLTerminals\_Data.**

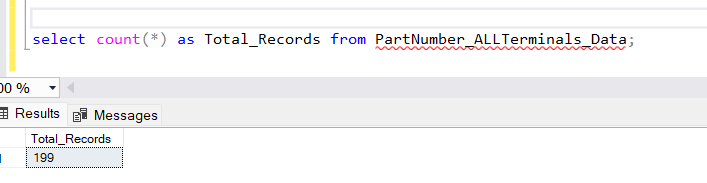
**Solve Below Basics Problem Statements :-**

**Note:Checked in doubt session to make changes in question since mentioned columns didn’t exist and added few columns to proceed**

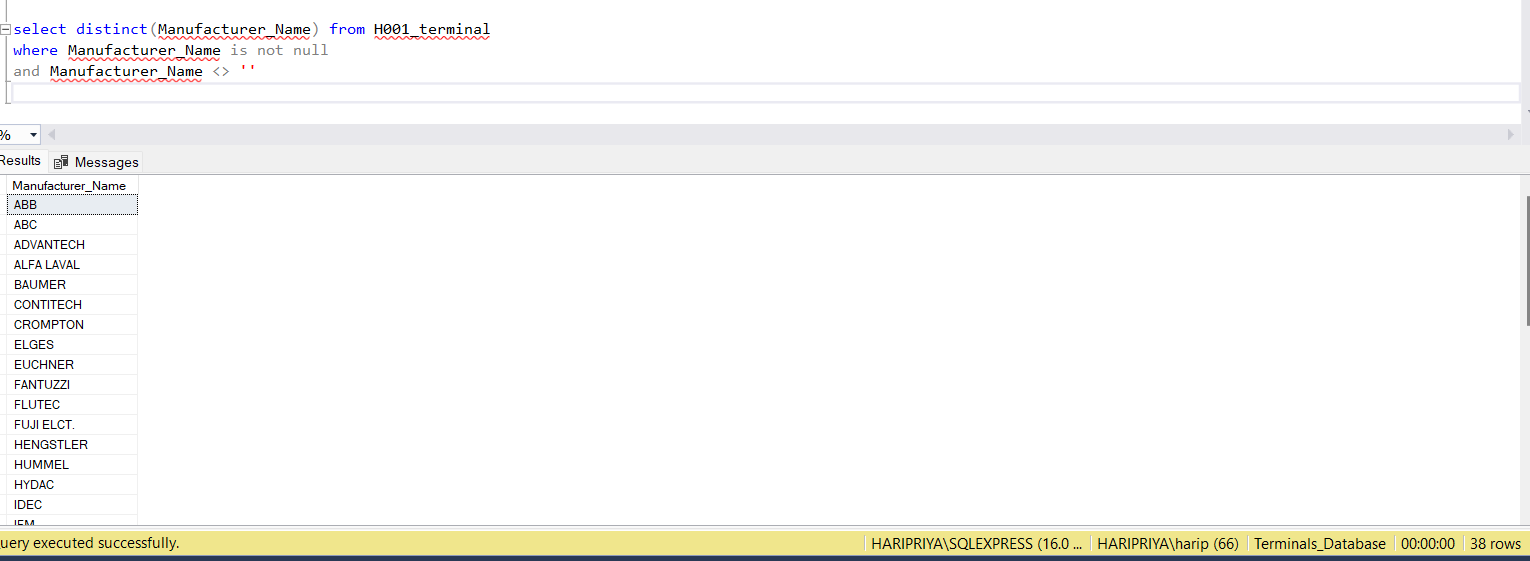
1. Retrieve all columns from the H001\_terminal table.

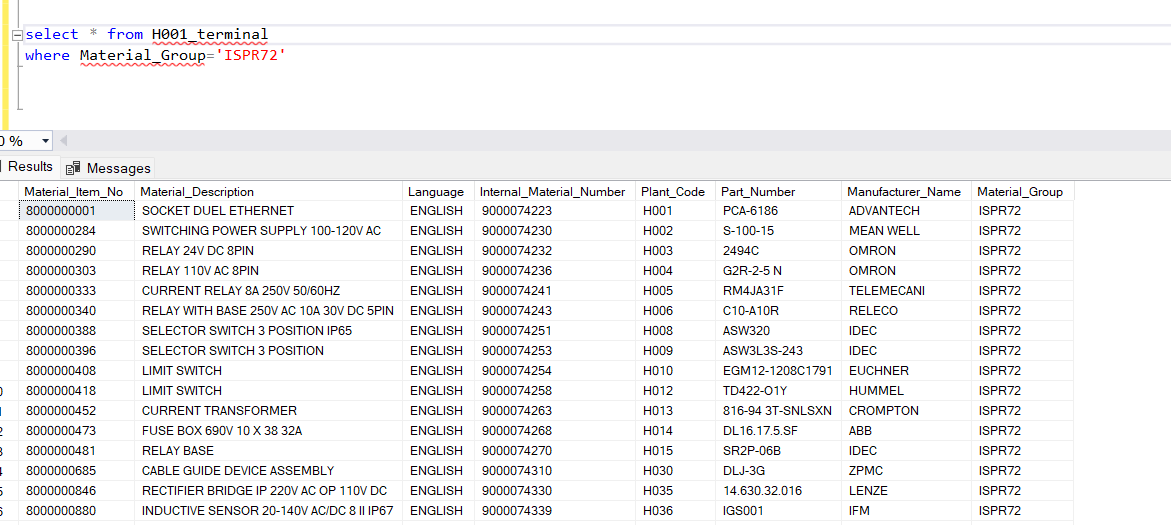
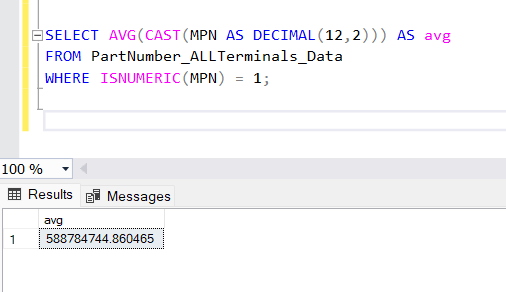


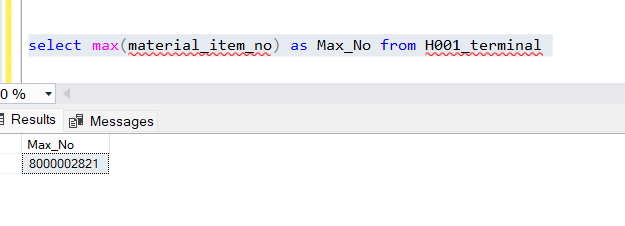
1. Count the total number of records in the PartNumber\_ALLTerminals\_Data table.

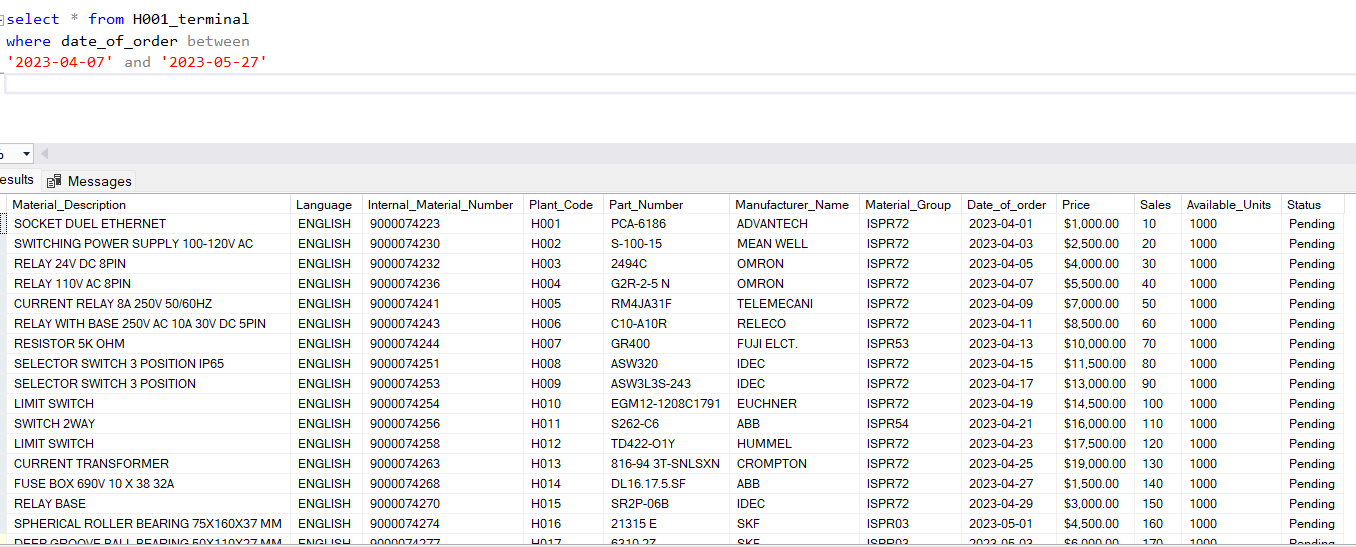


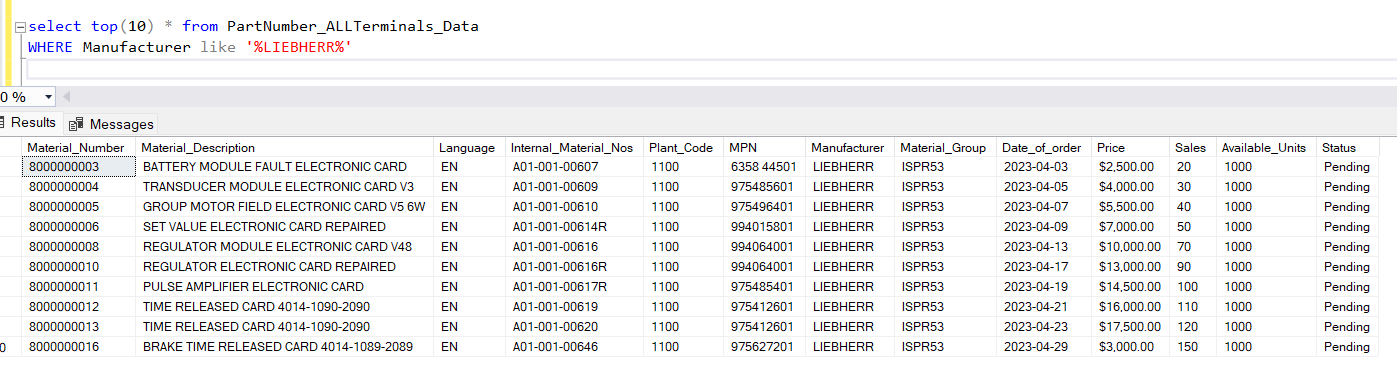
1. List distinct values of the ‘Manufacturer\_Name ' column from the H001\_terminal table.

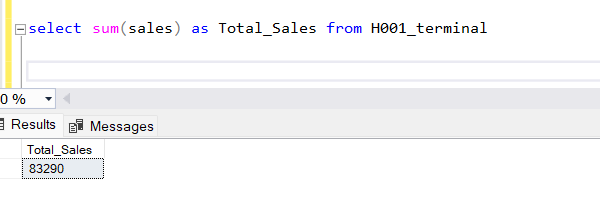
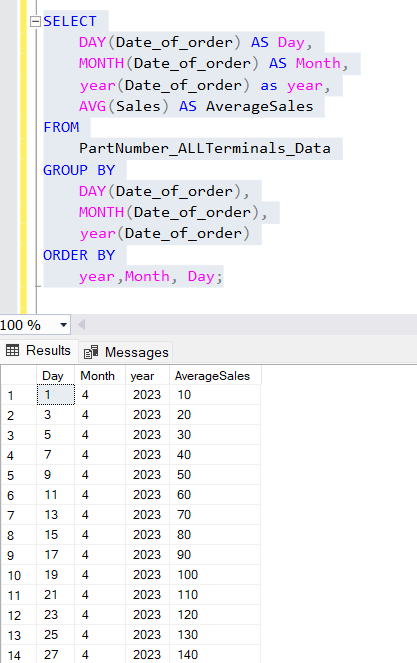
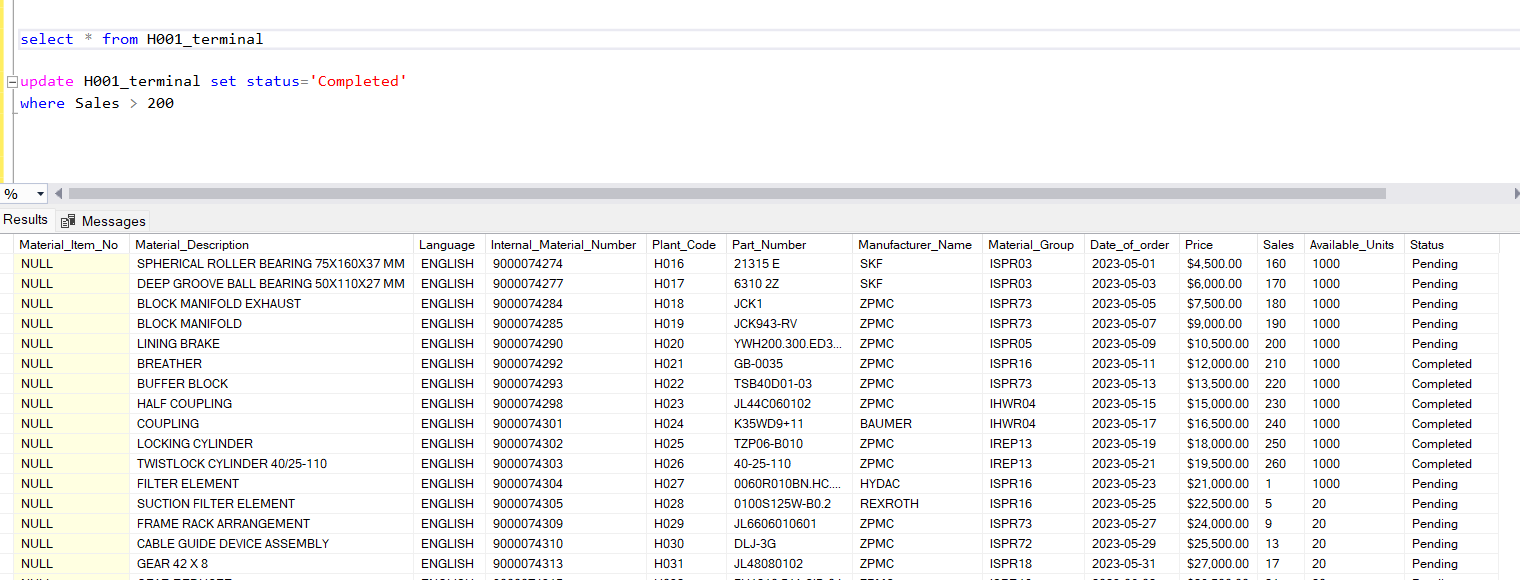
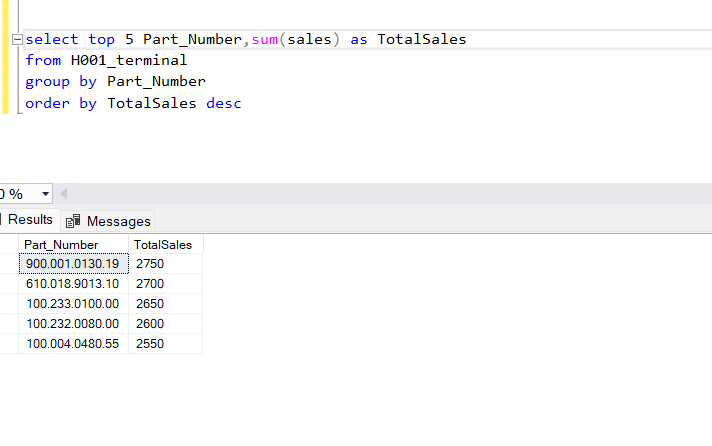
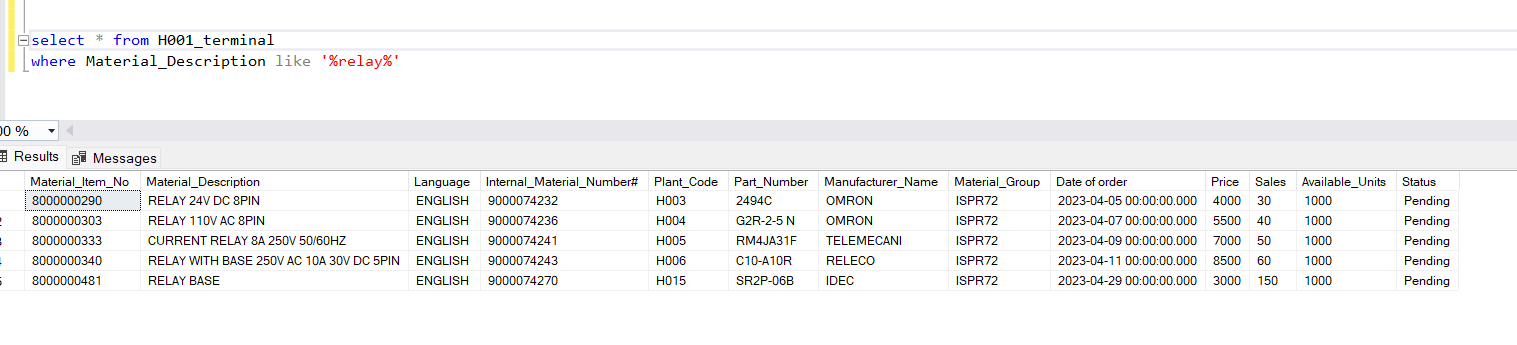
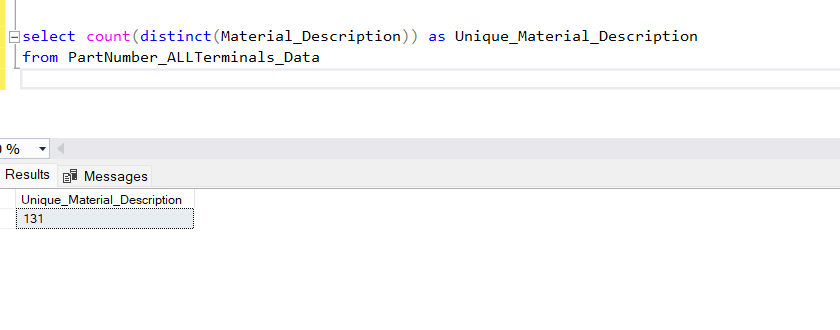
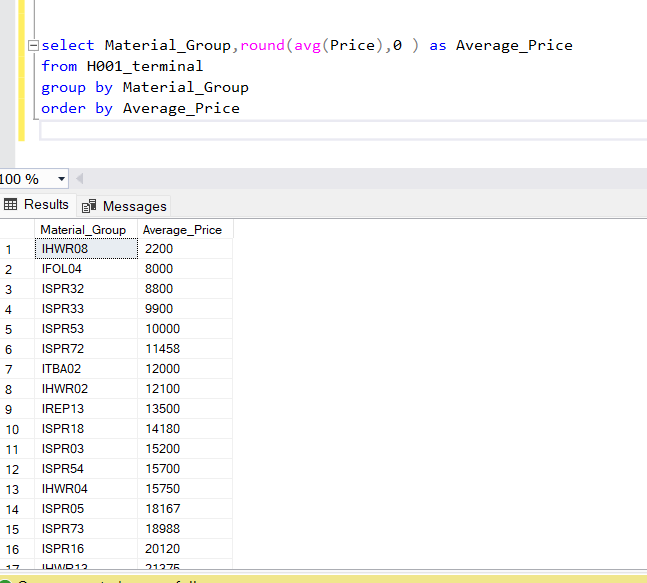


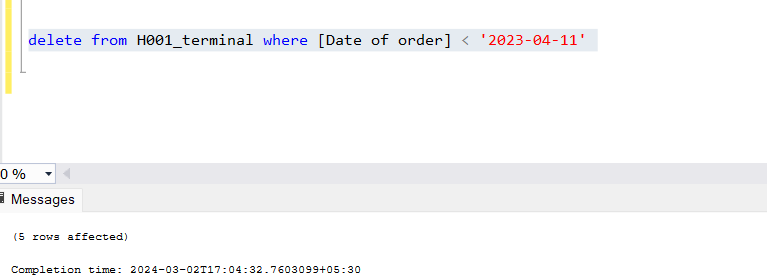
1. Retrieve records from the H001\_terminal table where material group is equal to a specific value of ISPR72. 
2. Calculate the average of the ‘MPN column' in the PartNumber\_ALLTerminals\_Data table. 
3. Find the maximum value in the 'material\_item' column of the H001\_terminal table.



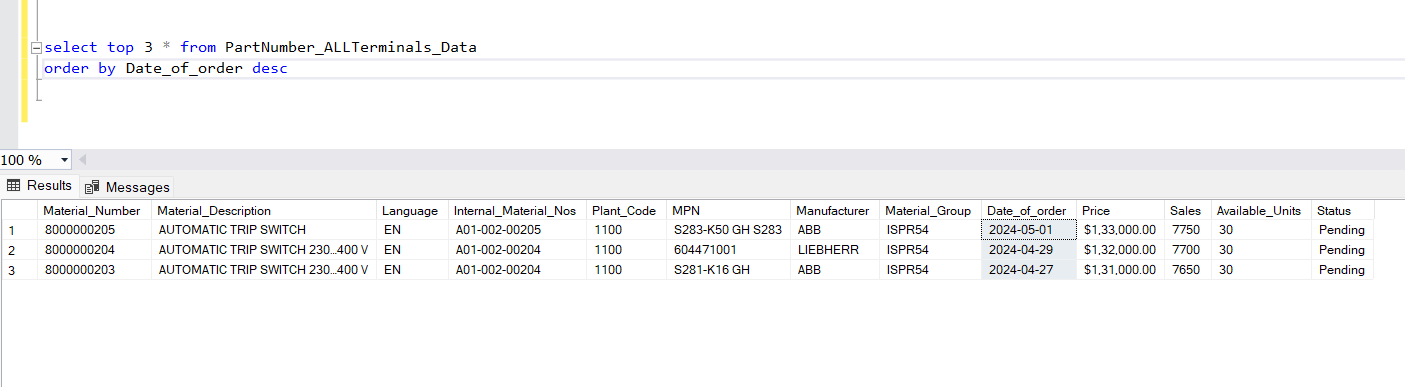
1. List all columns and records from H001\_terminal where 'date\_column' is in a specific date range from 2023-04-07 to 2023-05-27.  
   
2. Retrieve the top 10 records from PartNumber\_ALLTerminals\_Data based on a specific column- Manufacturer is LIEBHERR



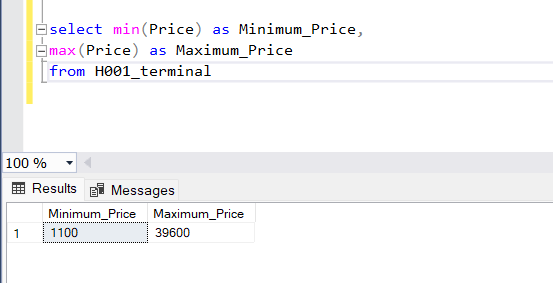
1. Calculate the total sum of the sales column in the H001\_terminal table. 
2. List the average sales per day and month from the PartNumber\_ALLTerminals\_Data table. 
3. Update the 'status' column in the H001\_terminal table to completed for records where 'condition\_column' is met.Note:Taken condition as sales above 200 
4. Find the top 5 Part Numbers with the highest total sales from the H001\_terminal table. 
5. Retrieve records from H001\_terminal where the 'text\_column' contains a specific word as relay. 
6. Count the number of unique 'categories' in material description in the PartNumber\_ALLTerminals\_Data table. 
7. Find the average price per unit for each 'PartNumber' in the PartNumber\_ALLTerminals\_Data table. 
8. Delete records from H001\_terminal where 'date\_column' is older than a specific date (2023-04-11)

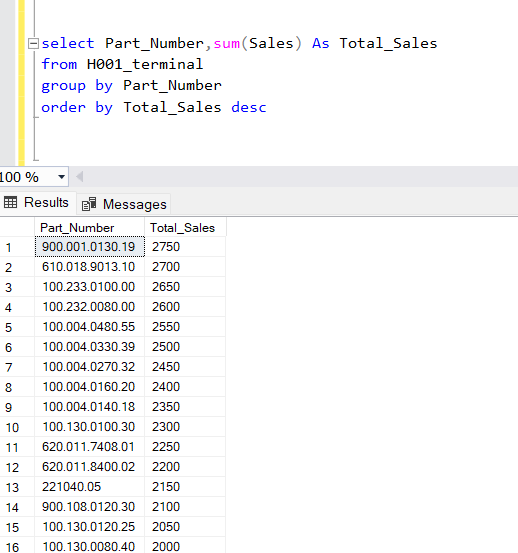
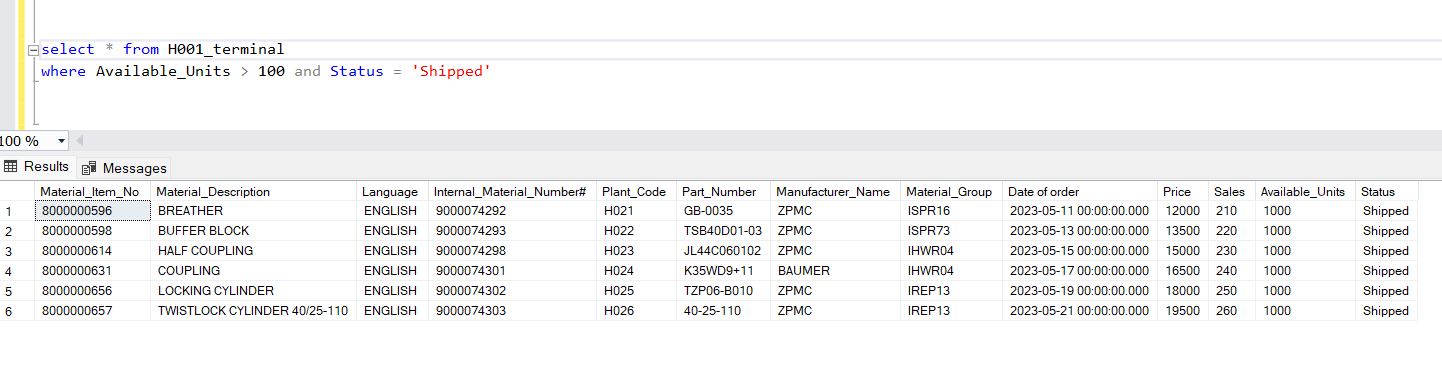
****

1. List the top 3 most recent records from PartNumber\_ALLTerminals\_Data based on 'date\_column'.



1. Find the minimum and maximum values of the 'price' column in the H001\_terminal table.



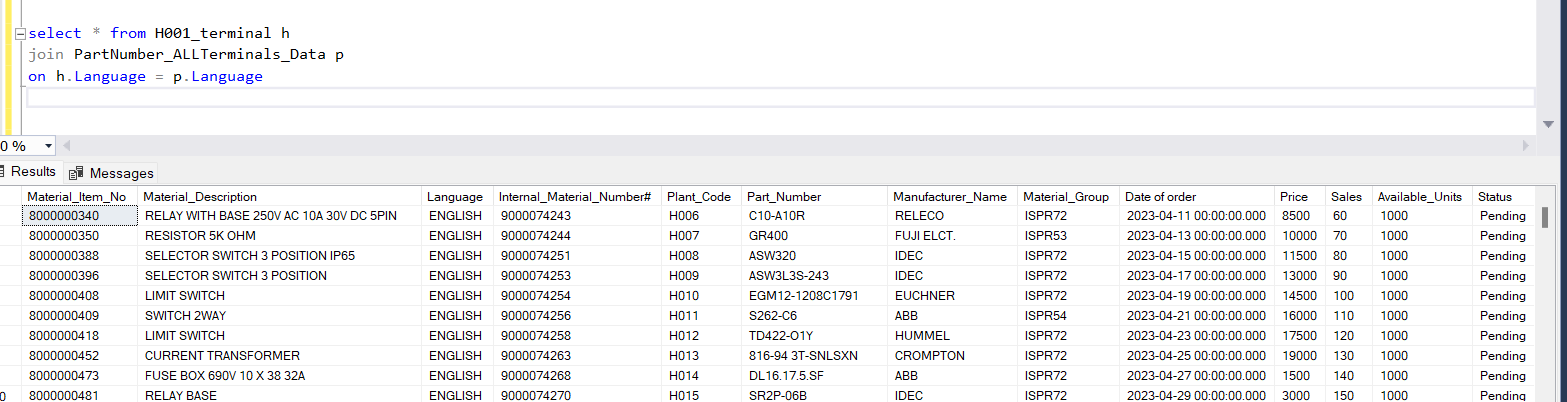
1. Calculate the total sales for each 'PartNumber' in the H001\_terminal table.
2. Retrieve records from H001\_terminal where the Available\_Units is greater than 100 and the 'status' is 'Shipped'. 

**Advanced Questions**

Joins:

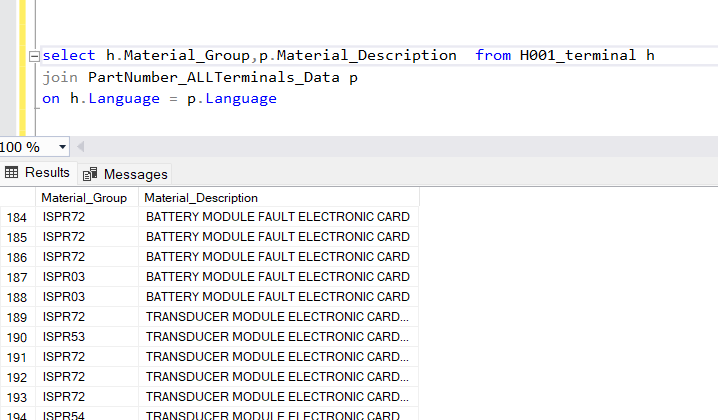
Problem:

1. Retrieve the PartNumbers and their details from H001\_terminal and PartNumber\_ALLTerminals\_Data.



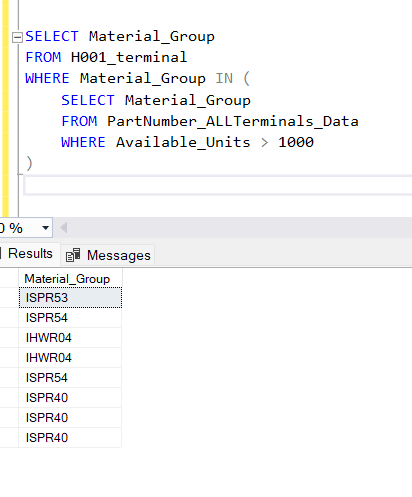
Problem:

1. List all material group from H001\_terminal along with their corresponding material descriptions from PartNumber\_ALLTerminals\_Data

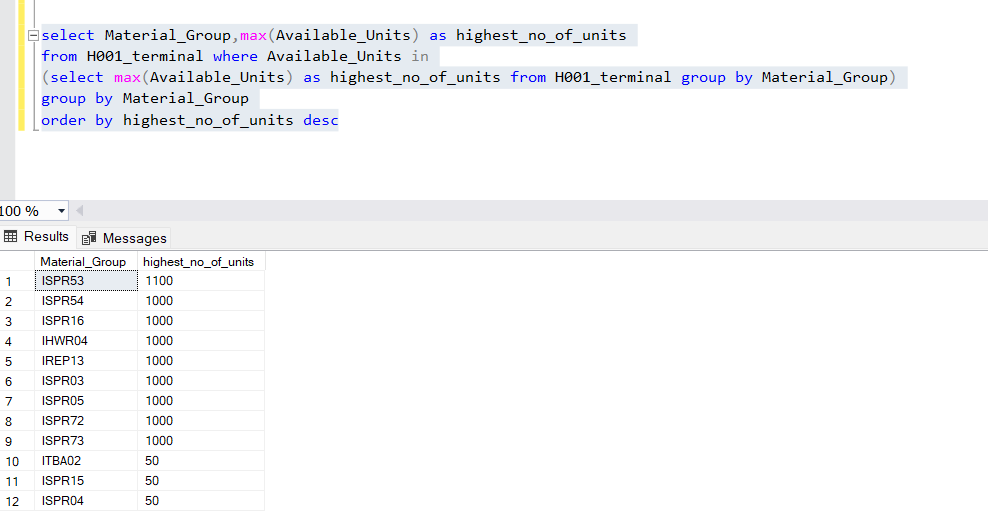


Subqueries:

Problem:

1. Find the material group from H001\_terminal that have more than 1000 units in PartNumber\_ALLTerminals\_Data. 

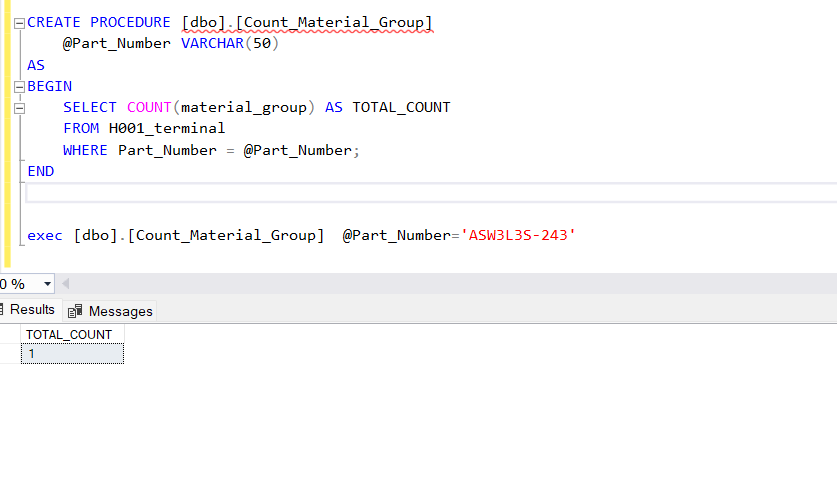
Problem:

1. Retrieve the material group with the highest number of units from H001\_terminal. 

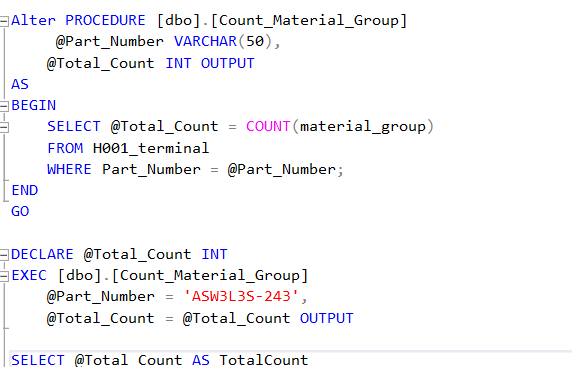
Procedures:

Problem:

1. Create a stored procedure that returns the total count of material group for a given PartNumber.

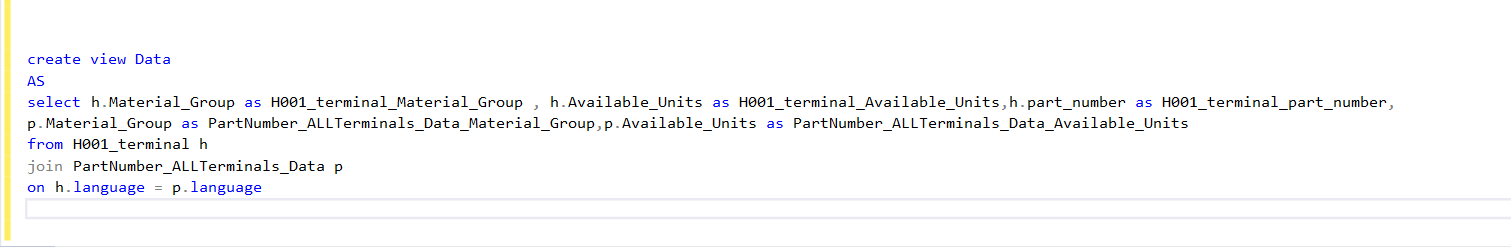


Problem:

6.Update the stored procedure to include an output parameter for the total count. 

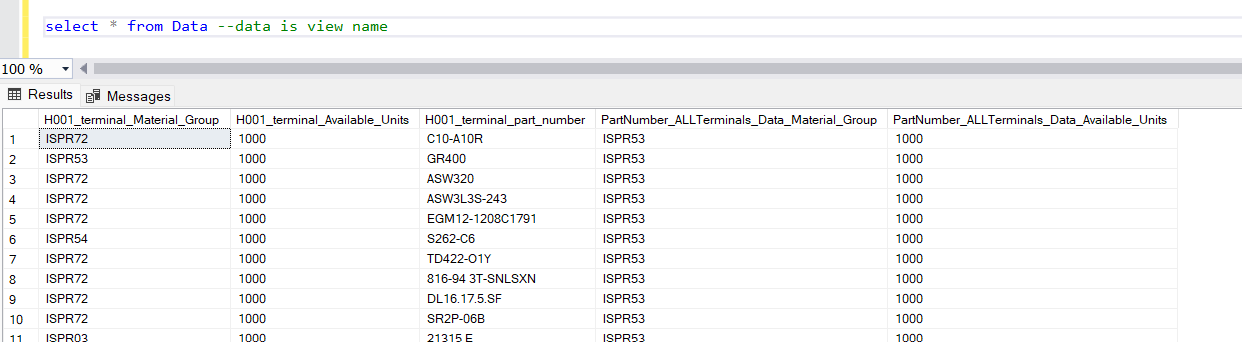
Views:

Problem:

1. Create a view that displays PartNumbers, material group, and Units from both tables. 

Problem:

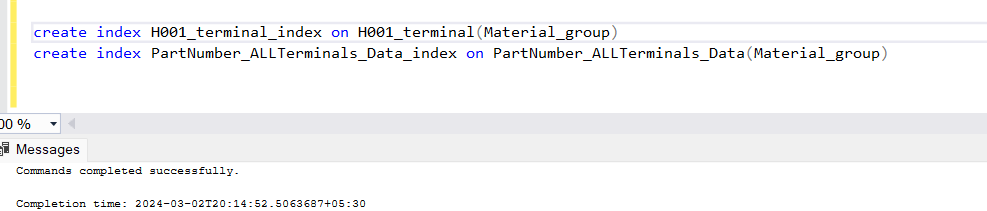
1. Retrieve the data from the view created above.



Indexes:

Problem:

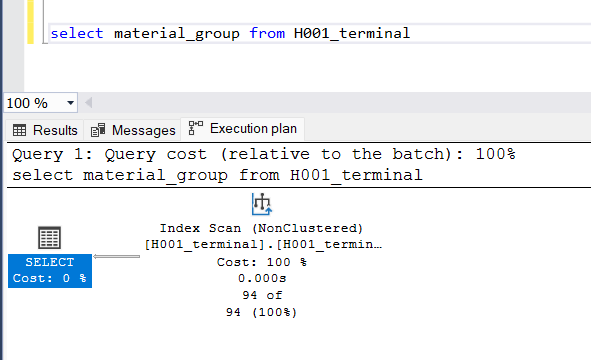
1. Add an index on the Material group column in both H001\_terminal and PartNumber\_ALLTerminals\_Data.



Problem:

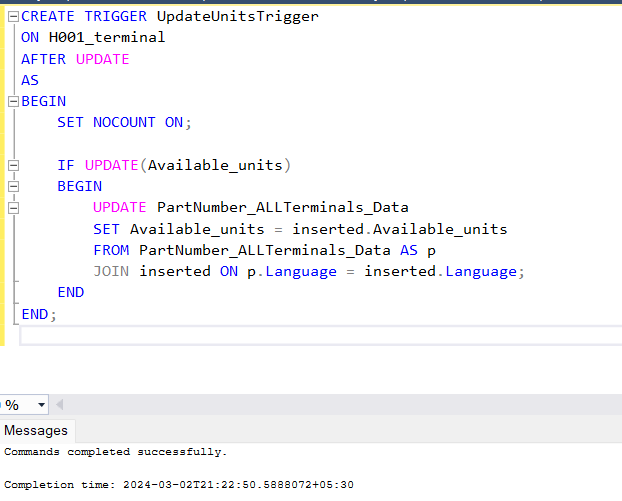
1. Check the execution plan before and after adding the indexs for a specific query.

After typed the query in new window, chosen **query** menu for “include actual execution plan” to notice execution plan in result window and noticed difference in I/O,operator cost after index creation



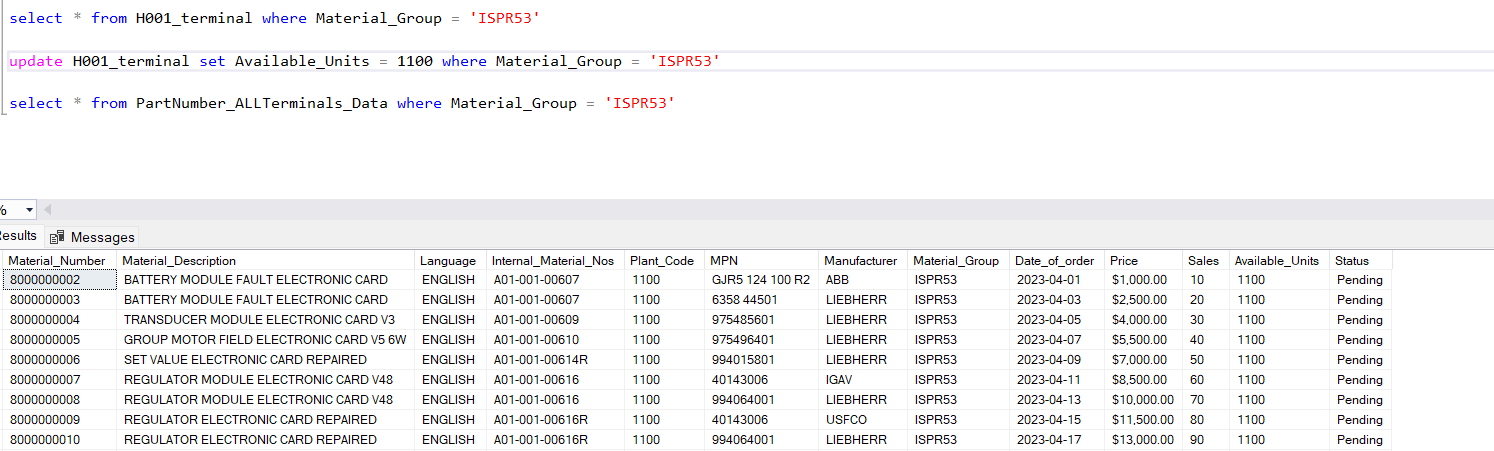
Triggers:

Problem:

1. Create a trigger that updates the Units in PartNumber\_ALLTerminals\_Data whenever Units are updated in H001\_terminal. 

Problem:

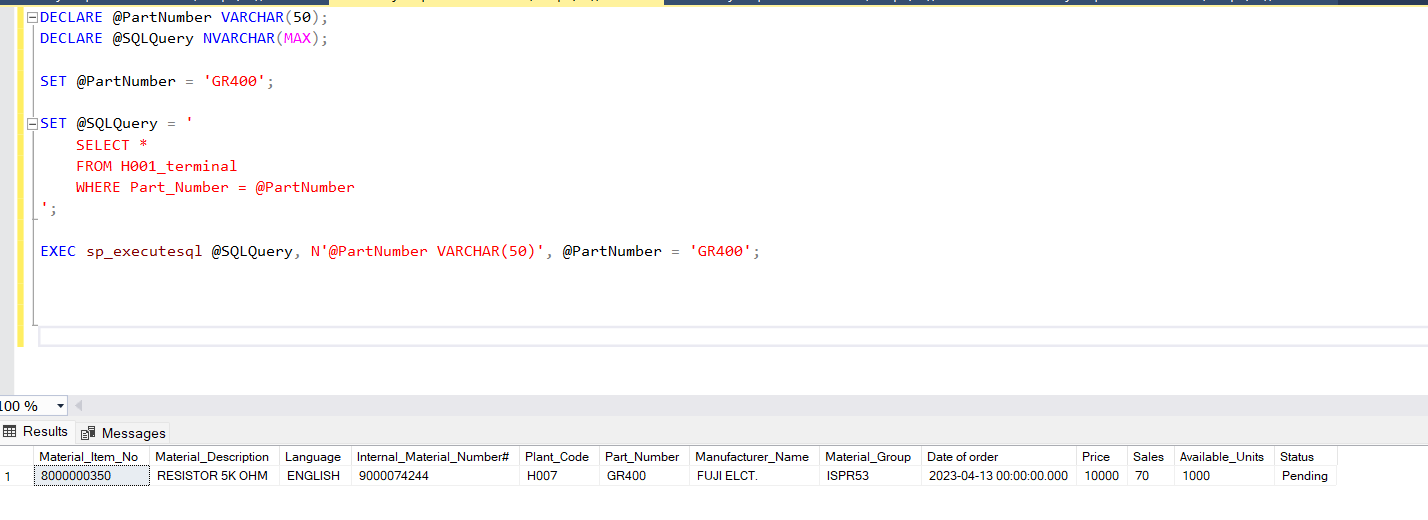
1. Test the trigger by updating Units in H001\_terminal.



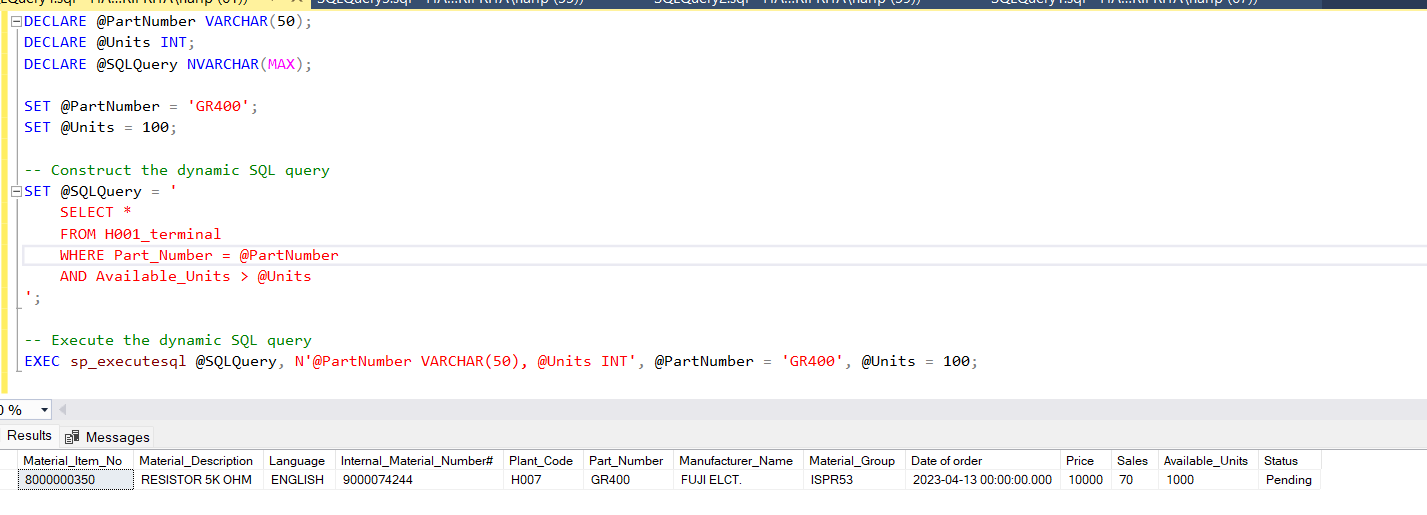
Dynamic Query:

Problem:

1. Create a dynamic query to retrieve data based on a given PartNumber.



Problem:

1. Modify the dynamic query to include a parameter for Units greater than a specified value. 

**Super Advanced Concepts ( For Experienced Candidates If you can able to Solve the Query Please Go with it else till Advanced Concepts You can solve it )**

Window Functions:

Problem:

1. Rank terminals in H001\_terminal based on their Units in descending order.

Problem:

1. Calculate the running total of Units for each PartNumber in PartNumber\_ALLTerminals\_Data.

Common Table Expressions (CTEs):

Problem:

1. Use a CTE to find terminals in H001\_terminal with duplicate PartNumbers.

Problem:

1. Create a CTE that calculates the average Units for each PartNumber.

Transactions:

Problem:

1. Wrap a set of queries in a transaction that deducts Units from H001\_terminal and updates Units in PartNumber\_ALLTerminals\_Data.

Problem:

1. Implement error handling within a transaction to rollback changes if an error occurs during the update.

Index Optimization:

Problem:

1. Identify and remove duplicate indexes on the same column in a table.

Dynamic SQL and Parameters:

Problem:

1. Create a dynamic query that retrieves columns based on user input.

Problem:

1. Modify the dynamic query to include a WHERE clause with a parameter.

Auditing:

Problem:

1. Implement auditing to log changes (insert, update, delete) on a specific table.

Problem:

1. Query the audit log to retrieve recent changes.

**POWER BI Dashboard:-**

Create a 2 Visual in one Page and after Solving all Questions Create 4 to 5 Dashboards by combining the visuals

Terminal Distribution:

1. How is the distribution of terminals across different regions?—*no relevant data for terminals provided*
2. Can you create a map visual to represent the geographical distribution of terminals?

Part Number Overview:

1. What is the overall count of unique part numbers in the dataset?

Overall Count of Unique Part Numbers = DISTINCTCOUNT(H001\_terminal[Part\_Number])

1. Can you visualize the distribution of part numbers by manufacture name?

Sales Trends:

1. How have terminal sales evolved over time? Create a time-series line chart.
2. Can you identify any seasonal patterns in terminal sales?

Top Performing Terminals:

1. What are the top 5 terminals in terms of sales or quantity sold?
2. Visualize the sales performance of these top terminals over time.

Part Number Sales Comparison:

1. Compare the sales performance of different part numbers using a bar chart.
2. Are there any specific part numbers that stand out in terms of sales?

Terminal Inventory Analysis:

1. Create a gauge chart to represent terminal inventory levels. How many terminals are in stock?

Revenue Analysis:

1. Visualize the revenue generated by material group. Are there any high-revenue material group?
2. Compare the revenue contribution of different part numbers.

Terminal Type Breakdown:

1. Create a pie chart to represent the proportion of material group in the dataset.

Top Customers:

1. Identify and visualize the top 10 customers based on their purchases of terminals.—*no relevant data*

Region-wise Sales Comparison:

1. Compare terminal sales across different regions using a stacked bar chart.

Terminal Return Rates:

1. Visualize the return rates for different terminals. Are there any terminals with high return rates?

Part Number Profitability:

1. Analyze the profitability of each part number using a waterfall chart.

Sales Forecast:

1. Use a line chart to forecast terminal sales for the next quarter or year. –*No forcasting option in my BI in analysis tool of line chart,hence not used*
2. Create a scatter plot to segment manufacture names based on their sales.

Terminals Out of Stock:

1. Identify and visualize terminals that are frequently out of stock.

OutOfStockFrequency =

DIVIDE(

CALCULATE(COUNTROWS(H001\_terminal), H001\_terminal[Sales] = 0),

COUNTROWS(H001\_terminal)

)

Performance Heatmap:

1. Create a heatmap to represent the performance of terminals across different dimensions (e.g., region, customer type).

Terminal Comparison Radar Chart:

1. Use a radar chart to compare the performance of different terminals in multiple categories (e.g., sales, returns, profitability).—*no relevant data*

Average Sales Price Analysis:

1. Visualize the average sales price of terminals over time or by region.

Customer Loyalty Analysis:

1. Analyze customer loyalty by visualizing repeat purchases and customer retention. ).—*no relevant data*

Part Number Contribution to Revenue:

1. Create a treemap to show the contribution of each part number to total revenue.

**\*\*\*Above Questions you can also solve with Tableau\*\*\* (If some visuals is not available in tableau which is mentioned in above question you can choose alternative charts for it according to you)**

**DAX Questions**

Total Terminal Count:

1. How many Partnumbers are there in the H001\_terminal dataset?

*Total Partnumbers = COUNT(H001\_terminal[Part\_Number])*

Distinct Part Numbers:

1. What is the count of unique manufacturename in the PartNumber\_ALLTerminals\_Data dataset?

*Unique Manufacture\_name = DISTINCTCOUNT(H001\_terminal[Manufacturer\_Name])*

Average Terminal Length:

3.What is the average length of partnumber in the H001\_terminal dataset? Average\_Part\_Number\_Length =

*AVERAGEX(*

*'H001\_terminal',*

*LEN('H001\_terminal'[Part\_Number])*

*)*

Maximum and Minimum sales:

4.What is the highest and lowest sales in the H001\_terminal dataset?

*Lowest\_Sales = MIN(H001\_terminal[Sales])*

*Highest\_Sales = MAX(H001\_terminal[Sales])*

Terminal Distribution by Type:

5.Can you provide a breakdown of terminal type by type in the H001\_terminal dataset? – *no relevant data in sheet*

Part Numbers with available group:

6.Which manufactur has the highest count of available group in the PartNumber\_ALLTerminals\_Data dataset?

Top\_Manufacturer =

    MAXX(

        TOPN(1,

            SUMMARIZE(

                PartNumber\_ALLTerminals\_Data,

                PartNumber\_ALLTerminals\_Data[Manufacturer],

                "Available\_Group\_Count", COUNTROWS(FILTER(PartNumber\_ALLTerminals\_Data, PartNumber\_ALLTerminals\_Data[Available\_Units] > 0))

            ),

            [Available\_Group\_Count], DESC

        ),

        [Manufacturer]

    )

Total sales of material groups:

7.What is the sum of sales for all material groups in the H001\_terminal dataset? TotalSales = SUM(H001\_terminal[Sales])

8.Identify material group where the average sales is above a certain threshold in the H001\_terminal dataset.

*Material\_Groups\_Above\_Target =*

*VAR AvgSalesByMaterialGroup =*

*SUMMARIZE (*

*'H001\_terminal',*

*'H001\_terminal'[Material\_Group],*

*"Average\_Sales", AVERAGE ( 'H001\_terminal'[Sales] )*

*)*

*RETURN*

*CALCULATE (*

*CONCATENATEX (*

*FILTER ( AvgSalesByMaterialGroup, [Average\_Sales] > [Target\_sales] ),*

*'H001\_terminal'[Material\_Group],*

*", "*

*)*

*)*

Count of Terminals by Manufacturer:

9.How many material group are there for each manufacturer in the H001\_terminal dataset?

*Manufacturer\_MaterialGroup\_Count =*

*COUNTROWS(*

*GROUPBY(*

*'H001\_terminal',*

*'H001\_terminal'[Manufacturer\_Name],*

*'H001\_terminal'[Material\_Group]*

*)*

*)*

Part Numbers with Terminals Below Average Length:

10.List part numbers where the terminals have lengths below the average length in the H001\_terminal dataset. –*no relevant data*

Moving Average of Terminal Voltages:

11.Can you calculate a 5-day exponential moving average of terminal voltages for each part number in the PartNumber\_ALLTerminals\_Data dataset? - –*no relevant data*

Ranking Part Numbers by Terminal Resistance:

12.Create a ranking of part numbers based on the total sales of their material groups in descending order in the H001\_terminal dataset. PartNumber\_Ranking =

RANKX(

    ALL('H001\_terminal'[Part\_Number]),

    CALCULATE(

        SUM('H001\_terminal'[Sales]),

        ALLEXCEPT('H001\_terminal', 'H001\_terminal'[Part\_Number])

    ),

    ,

    DESC,

    Dense

)

Calculation of Weighted Average Voltage:

1. Calculate the weighted average voltage for each part number in the PartNumber\_ALLTerminals\_Data dataset, where the weights are based on the count of terminals.—*no relevant data*

Till here You have Solved the different types of Problem Statements now Summarise the analysis of above problem statements in Below Format and Create the Report of above Project :--

Report Format

Creating a comprehensive and effective data analysis report involves several key considerations to ensure clarity, accuracy, and relevance. Here are some important aspects to focus on:

**1. Clearly Defined Objectives:**

Clearly state the purpose and goals of the analysis. What questions are you trying to answer or what insights are you aiming to derive from the data?

**2. Data Quality:**

Ensure the data used for analysis is accurate, relevant, and reliable. Address any missing or erroneous data before analysis.

**3. Visualizations:**

Utilize appropriate graphs, charts, and visual aids to present findings effectively. Visualizations often make complex information easier to understand and remember.

**4. Structure and Organization:**

Create a clear structure for the report, including an introduction, methodology, findings, conclusions, and recommendations. Ensure a logical flow of information.

**5. Audience Understanding:**

Tailor the report to your audience's level of understanding. Avoid jargon or explain technical terms where necessary.

**6. Interpretation and Context:**

Provide interpretations of the findings and place them in the broader context of the problem or situation being analyzed.

7**. Conclusions and Recommendations:**

Summarize key findings and provide actionable insights or recommendations based on the analysis.

**8. Data Sources and Methodology:**

Clearly state the sources of data used and the methodologies employed for analysis. This enhances transparency and allows for reproducibility.

**9. Accuracy and Validation:**

Double-check calculations, statistical analyses, and interpretations to ensure accuracy. Validation by a peer or subject matter expert can add credibility.

**10. Storytelling:**

Craft a narrative around the data. Tell a story that guides the reader through the analysis, making it engaging and easy to follow.

**11. Review and Iteration:**

Review the report for coherence, accuracy, and completeness. Iterate based on feedback received, if applicable.

**12. Visual and Aesthetic Appeal:**

Pay attention to the design elements of the report. A visually appealing and well-formatted document can enhance readability and comprehension.

**Report:**

**1. Introduction**

In today's competitive business landscape, data-driven decision-making has become imperative for organizations aiming to stay ahead. This report presents a comprehensive analysis of sales and inventory data for materials in the manufacturing domain. By leveraging descriptive and predictive analytics techniques, we aim to extract actionable insights to optimize inventory management strategies, enhance operational efficiency, and drive informed decision-making.

**2. Clearly Defined Objectives**

The dataset used for analysis comprises detailed information on materials, including item numbers, descriptions, order dates, prices, sales figures, available units, status, and state of operation. The data was collected from the company's internal database, ensuring its reliability and accuracy. Prior to analysis, data cleaning and preprocessing steps were undertaken to address missing values, remove duplicates, and standardize formats, ensuring the integrity of the dataset. The analysis seeks to answer key questions regarding sales trends, inventory status, pricing dynamics, and regional variations. By addressing these questions, the report aims to identify opportunities for optimization and enhancement within the manufacturing supply chain

**3. Data Quality**

Before commencing the analysis, rigorous checks were conducted to ensure the accuracy, relevance, and reliability of the dataset. Missing or erroneous data points were addressed to maintain the integrity of the analysis.

**Accuracy**: This pertains to the correctness of the data entries. For the sample data provided, accuracy checks would involve verifying that each data point, such as Material\_Item\_No, Material\_Description, Date of order, Price, Sales, Available\_Units, Status, and State, is recorded correctly and without errors. This could involve cross-referencing the data with primary sources or comparing it against known standards or benchmarks.

**Relevance**: Relevance refers to whether the data is appropriate and applicable to the analysis at hand. In this case, it's crucial to ensure that the data fields provided are relevant to the objectives of the analysis. For example, to optimize inventory management, ensuring that fields like Sales, Available\_Units, and Status are included and accurately represent the inventory status is essential.

**Reliability**: Reliability concerns the consistency and dependability of the data. In the context of the sample data, reliability checks may involve examining the consistency of data formats, such as date formats or currency representations (e.g., ensuring that prices are consistently denoted in Indian Rupees). Additionally, consistency in terms of data recording practices across different entries is important for reliability.

**Missing or Erroneous Data**: Addressing missing or erroneous data points is crucial to maintain the integrity of the analysis. This could involve various strategies such as data imputation for missing values, correcting data entry errors, or excluding data points that cannot be validated. For example, if there are missing values in the Sales or Available\_Units columns, appropriate methods such as mean imputation or interpolation could be used to fill in the missing values.

**4. Visualizations**

To facilitate effective communication of findings, a variety of graphs, charts, and visual aids were utilized. These visualizations help elucidate complex trends and patterns within the data, enhancing understanding and interpretation.

**5. Structure and Organization**

The report follows a structured format comprising methodology, findings, and recommendations. This organization ensures a coherent and logical presentation of information, guiding the reader through the analysis seamlessly.

**5.1.Methodology:**

**Data Collection**: The data was collected from internal records, containing information about material sales, orders, and inventory.

**Data Processing:** The dataset was processed to address missing or erroneous data points. Data cleaning techniques were applied to ensure data quality.

**Analysis Approach**: Various analytical techniques such as descriptive statistics, trend analysis, and visualization were used to analyze the data.

**Software Tools**: Software tools like Python, Excel,BI were utilized for data analysis and visualization.

**5.2.Findings:**

**Sales Trends**: Visualizations such as line charts and bar graphs are used to illustrate sales trends over time and across different materials.

**Inventory Status:** Tables and visualizations present information about available units, highlighting any shortages or excess inventory.

**Regional Distribution:** Maps or geographical plots depict the regional distribution of sales, providing insights into market demand across different states.

**5.3.Recommendations:**

**Inventory Management:** Recommendations can be provided for optimizing inventory levels based on sales trends and demand forecasts.

**Sales Strategies:** Strategies suggested for improving sales performance, targeting specific regions or materials.

**6. Audience Understanding**

Consideration was given to the audience's level of understanding, with technical terms explained where necessary to ensure clarity. The report is accessible to a wide range of stakeholders within the manufacturing industry.

**6.1.Technical Terminology Explanation:**

Technical terms specific to the manufacturing industry, such as "relay," "resistor," or "switch," may be explained briefly within the report for readers who may not be familiar with these terms.

For instance, in the analysis, terms like "current relay," "limit switch," or "spherical roller bearing" could be defined in simple terms to ensure clarity for stakeholders who are not intimately familiar with manufacturing terminology.

**6.2.Data Analysis Simplification:**

Technical terms related to data analysis, such as "descriptive statistics," "trend analysis," or "data visualization," should also be explained where necessary.

For instance, if the report mentions "descriptive statistics" to describe summary measures like mean, median, or standard deviation, a brief explanation of what these measures represent and their relevance to the analysis can be provided.

**6.3.Visual Representation for Clarity:**

Utilizing visual aids, such as graphs, charts, and maps, can help clarify complex information for stakeholders who may not have a strong technical background.

Visual representations of sales trends, inventory status, and regional distribution of sales can make it easier for stakeholders to grasp key insights without relying heavily on technical language.

**6.4.Contextualizing Insights:**

Providing context for the analysis findings is essential for ensuring understanding among stakeholders. This involves explaining the significance of the findings in relation to broader industry trends or business objectives.

For example, if the analysis reveals a decline in sales for a particular material, the report could contextualize this by discussing potential market factors or competitive dynamics influencing the trend.

**7. Interpretation and Context**

Findings are interpreted within the broader context of the manufacturing landscape, taking into account industry trends, market dynamics, and business objectives. This contextualization provides deeper insights into the implications of the analysis.

**7.1.Industry Trends:**

Analyzing sales trends and inventory status within the context of broader industry trends allows for a better understanding of how the company's performance compares to the overall market.

For example, if the analysis reveals a decline in sales of certain materials, it's important to consider whether this trend is consistent with broader industry patterns. Are similar materials experiencing similar declines in demand across the manufacturing sector?

**7.2.Market Dynamics:**

Examining regional distribution of sales and inventory levels can shed light on market dynamics and regional variations in demand.

Are certain materials more popular in specific regions? Are there regional factors, such as economic conditions or regulatory requirements, influencing sales patterns? Understanding these dynamics can help tailor sales and marketing strategies to different geographic markets.

**7.3.Business Objectives:**

Interpreting the findings in the context of specific business objectives allows for actionable insights that align with the company's goals.

For instance, if the objective is to optimize inventory management, the analysis may focus on identifying slow-moving inventory or forecasting future demand to prevent stockouts or overstock situations.

**7.4.Competitive Landscape:**

Considering the competitive landscape is crucial for understanding market positioning and identifying areas of competitive advantage or vulnerability.

Are there specific competitors gaining market share in certain product categories? How does the company's performance compare to key competitors in terms of sales volumes, pricing strategies, or product offerings?

**7.5.External Factors**:

Taking into account external factors such as technological advancements, regulatory changes, or shifts in consumer preferences provides additional context for interpreting the findings.

For example, if there's a sudden increase in demand for materials related to renewable energy technologies, this could be attributed to changing regulatory requirements or growing environmental awareness.

**8.Conclusions and Recommendations:**

Key findings from the analysis of the sales and inventory data are summarized succinctly. This includes insights into sales trends, inventory levels, and regional distribution patterns.

Based on these findings, actionable recommendations are proposed to optimize inventory management practices. For example, recommendations might include adjusting procurement strategies to align with demand fluctuations, identifying and addressing slow-moving inventory, or implementing targeted marketing campaigns to promote sales of specific materials.

**9.Data Sources and Methodology:**

The sources of data used in the analysis are clearly stated as internal records or databases containing sales and inventory data.

Methodologies employed for data collection and analysis are outlined. This may involve data cleaning techniques to address missing or erroneous data, as well as statistical analysis methods such as trend analysis or regression modeling.

**10.Accuracy and Validation:**

A rigorous validation process is undertaken to ensure the accuracy of calculations, statistical analyses, and interpretations derived from the dataset. This involves double-checking calculations and statistical analyses to minimize errors.

**11.Storytelling:**

The analysis is presented in a narrative format that guides the reader through the data, making it engaging and easy to follow. This storytelling approach involves structuring the analysis in a logical sequence, highlighting key insights, and using visualizations to support the narrative.

For instance, the report might begin by setting the context for the analysis, then progress to presenting the data findings, interpreting the implications, and concluding with actionable recommendations.

**12.Review and Iteration:**

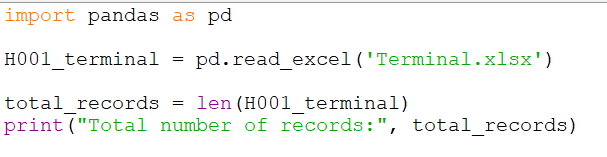
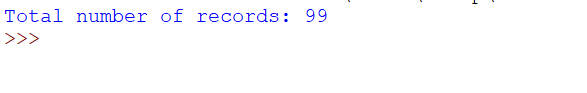
The report undergoes thorough review for coherence, accuracy, and completeness. Feedback from stakeholders or subject matter experts is incorporated into the report, and iterations are made as necessary to ensure that the final report meets the highest standards of quality and clarity.

**Analysis Using Python :-**

Python Analysis Depends Upon you if you have basics and intermediate knowledge of Python and Its libraries then you can go for these problem statements:-

**H001\_terminal Dataset:**

1. What is the total number of records in the H001\_terminal dataset?

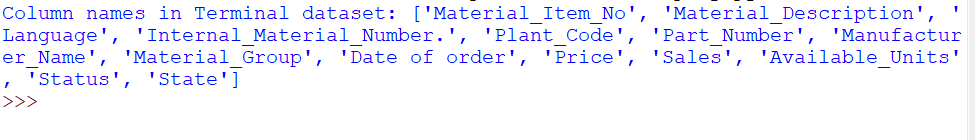
1. Can you provide the names of the columns in the H001\_terminal dataset?

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*column\_names = H001\_terminal.columns.tolist()*

*print("Column names in Terminal dataset:",column\_names)*



1. What is the data type of the 'Column\_A' in the H001\_terminal dataset?

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*data\_type = H001\_terminal['Material\_Item\_No'].dtype*

*print("Datatype of column A in H001\_terminal dataset:",data\_type)*



1. Are there any missing values in the 'Column\_B' of the H001\_terminal dataset?

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*missing\_values\_in\_column\_b = H001\_terminal['Material\_Description'].isnull().any()*

*if missing\_values\_in\_column\_b:*

*print("Yes, there are missing values in 'Column\_B' of the H001\_terminal dataset.")*

*else:*

*print("No, there are no missing values in 'Column\_B' of the H001\_terminal dataset.")*



1. What is the range of values in the 'Column\_A' of the H001\_terminal dataset?

*import pandas as pd*

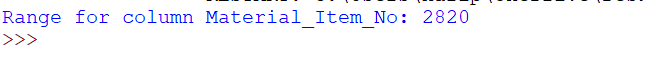
*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*min\_value = H001\_terminal['Material\_Item\_No'].min()*

*max\_value = H001\_terminal['Material\_Item\_No'].max()*

*Range = max\_value-min\_value*

*print("Range for column Material\_Item\_No:",Range)*



1. How many unique values are there in the 'Column\_D' of the H001\_terminal dataset?

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*Unique\_value = H001\_terminal['Internal\_Material\_Number.'].nunique()*

*print("Unique values in column Internal\_Material\_Number:" ,Unique\_value)*



1. What is the average value of 'Column\_J' in the H001\_terminal dataset?

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*Average\_value = H001\_terminal['Price'].mean()*

*print("Average value in column Price:" ,round(Average\_value,2))*



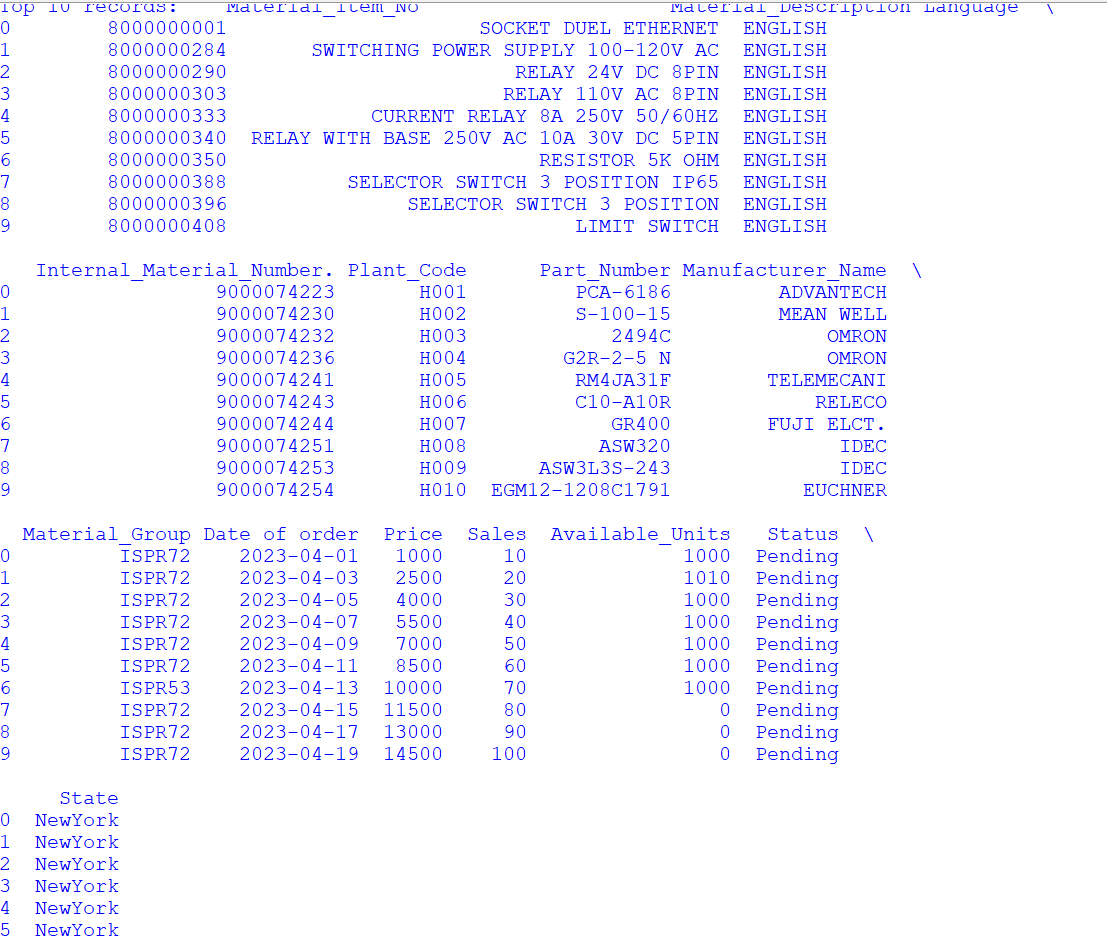
1. Show the first 10 records in the H001\_terminal dataset.

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*pd.set\_option('display.max\_columns', None)*

*print("Top 10 records:" ,H001\_terminal.head(10))*



1. What is the maximum value in 'Column\_F' of the H001\_terminal dataset?

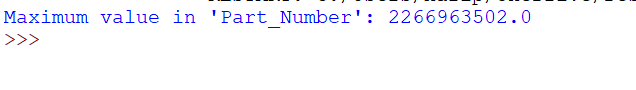
*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*H001\_terminal['Part\_Number'] = pd.to\_numeric(H001\_terminal['Part\_Number'], errors='coerce')*

*max\_value = H001\_terminal['Part\_Number'].max()*

*print("Maximum value in 'Part\_Number':", max\_value)*



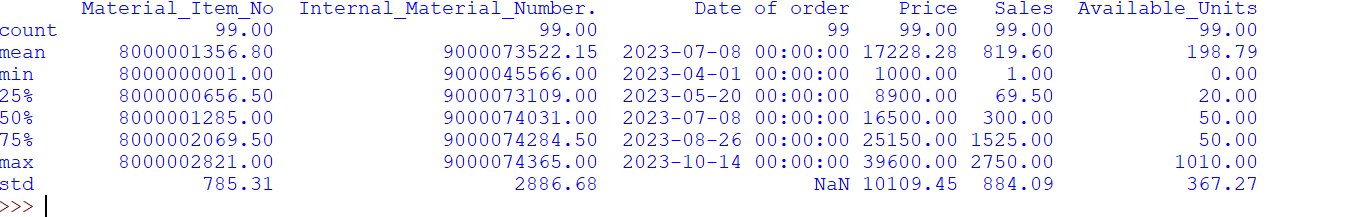
1. Provide a summary statistics for the H001\_terminal dataset.

*import pandas as pd*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*summary = H001\_terminal.describe()*

*summary\_str = summary.to\_string(float\_format=lambda x: f'{x:.2f}')*

*print(summary\_str)*

**PartNumber\_ALLTerminals\_Data Dataset:**

1. What is the total number of records in the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*total\_records = len(Part\_Number)*

*print("Total\_records present are :",total\_records)*



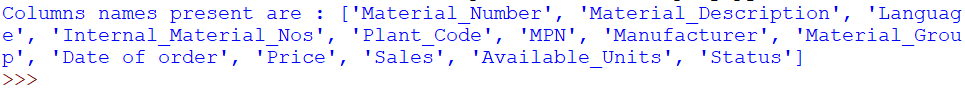
1. Can you provide the names of the columns in the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*column\_names = Part\_Number.columns.tolist()*

*print("Columns names present are :",column\_names )*



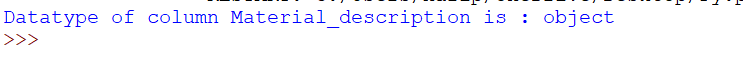
1. What is the data type of the Material\_description column in the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*data\_type = Part\_Number['Material\_Description'].dtype*

*print("Datatype of column Material\_description is :",data\_type )*



1. Are there any missing values in the 'Description' column of the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

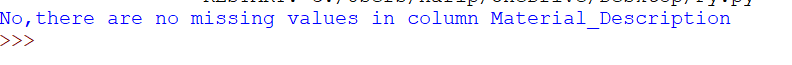
*Missing\_value = Part\_Number['Material\_Description'].isnull().any()*

*if Missing\_value:*

*print("Yes,there are missing values in column Material\_Description")*

*else:*

*print("No,there are no missing values in column Material\_Description")*



1. What is the range of values in the 'sales' column of the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*min\_value = Part\_Number['Sales'].min()*

*max\_value = Part\_Number['Sales'].max()*

*Range = max\_value - min\_value*

*print("Range of column Sales:",Range)*



1. How many unique values are there in the material\_group column of the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*Unique\_values = Part\_Number['Material\_Group'].nunique()*

*print("Unique values in column-Material\_Group :",Unique\_values)*



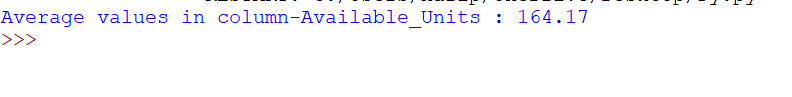
1. What is the average of Available\_Units in the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*Average\_value = Part\_Number['Available\_Units'].mean()*

*print("Average values in column-Available\_Units :",round(Average\_value,2))*



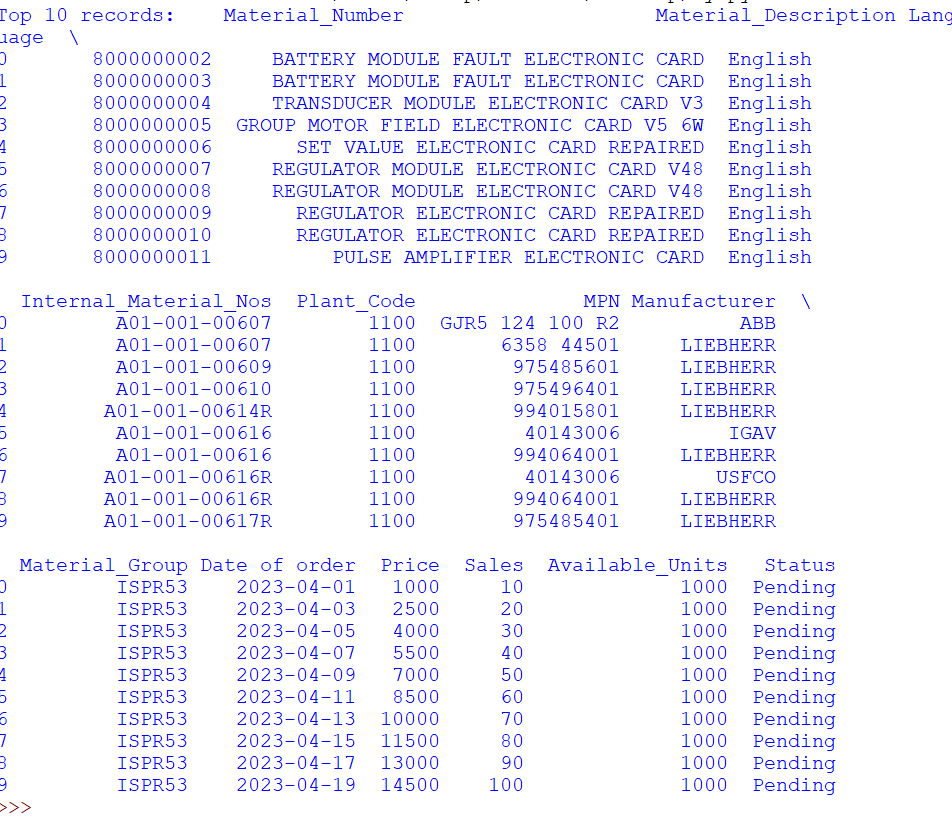
1. Show the first 10 records in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*pd.set\_option('display.max\_columns',None)*

*print("Top 10 records:",Part\_Number.head(10))*



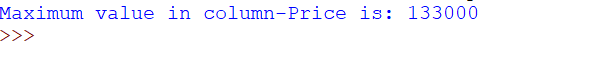
1. What is the maximum value in the 'Price' column of the PartNumber\_ALLTerminals\_Data dataset?

*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*Maximum\_price = Part\_Number['Price'].max()*

*print("Maximum value in column-Price is:",Maximum\_price)*



1. Provide a summary statistics for the PartNumber\_ALLTerminals\_Data dataset.

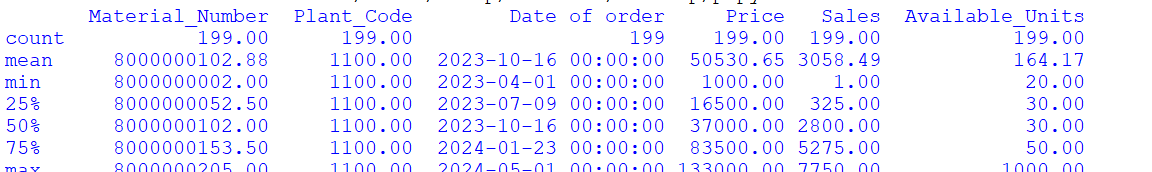
*import pandas as pd*

*Part\_Number = pd.read\_excel('PartNumber\_ALLTerminals\_Data.xlsx')*

*summary = Part\_Number.describe()*

*summary\_str = summary.to\_string(float\_format=lambda x: f'{x:.2f}')*

*print(summary\_str)*



**Data Visualizations :--**

**H001\_terminal Dataset:**

1. Can you create a histogram for the distribution of values in 'Column\_E' in the H001\_terminal dataset?

*import pandas as pd*

*import matplotlib.pyplot as plt*

*H001\_terminal= pd.read\_excel('Terminal.xlsx')*

*Plant\_Code = H001\_terminal['Plant\_Code']*

*plt.hist(Plant\_Code, bins=10, color='skyblue', edgecolor='black')*

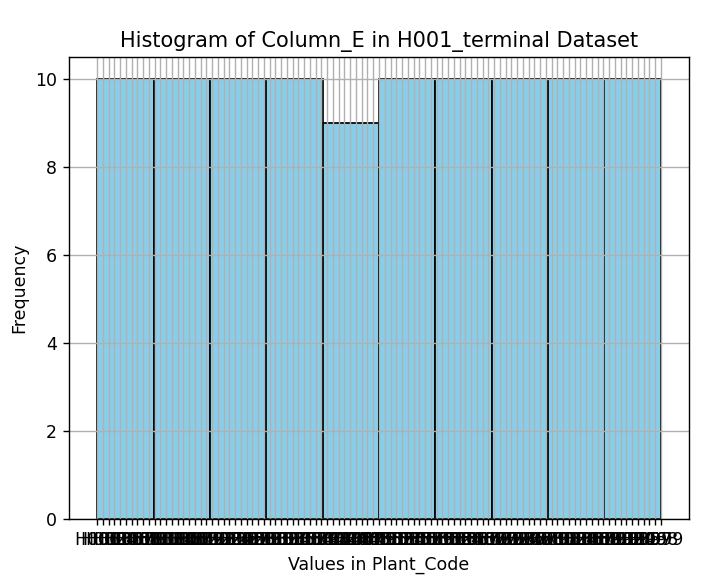
*plt.xlabel('Values in Plant\_Code')*

*plt.ylabel('Frequency')*

*plt.title('Histogram of Column\_E in H001\_terminal Dataset')*

*plt.grid(True)*

*plt.show()*



1. Generate a line plot showing the trend of 'Column\_I' over time if there is a time-related column in the H001\_terminal dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*H001\_terminal= pd.read\_excel('Terminal.xlsx')*

*H001\_terminal['Date'] = pd.to\_datetime(H001\_terminal['Date of order'], format='%d-%m-%Y')*

*H001\_terminal = H001\_terminal.sort\_values('Date')*

*plt.figure(figsize=(10, 6))*

*plt.plot(H001\_terminal['Date'], H001\_terminal['Date of order'], marker='o', linestyle='-')*

*plt.xlabel('Date')*

*plt.ylabel('Date of order')*

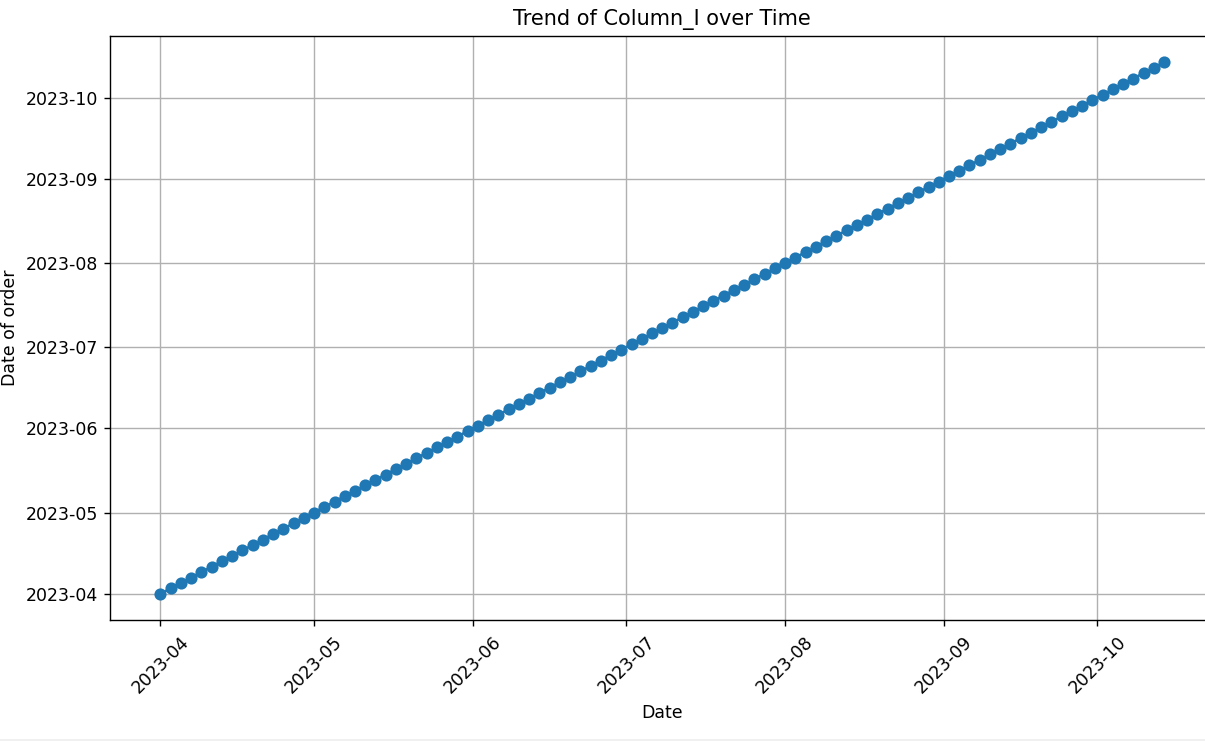
*plt.title('Trend of Column\_I over Time')*

*plt.xticks(rotation=45)*

*plt.grid(True)*

*plt.tight\_layout()*

*plt.show()*



1. Create a box plot for the 'Column\_D' to identify any outliers in the H001\_terminal dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*H001\_terminal= pd.read\_excel('Terminal.xlsx')*

*Internal\_Material\_Number = H001\_terminal['Internal\_Material\_Number.']*

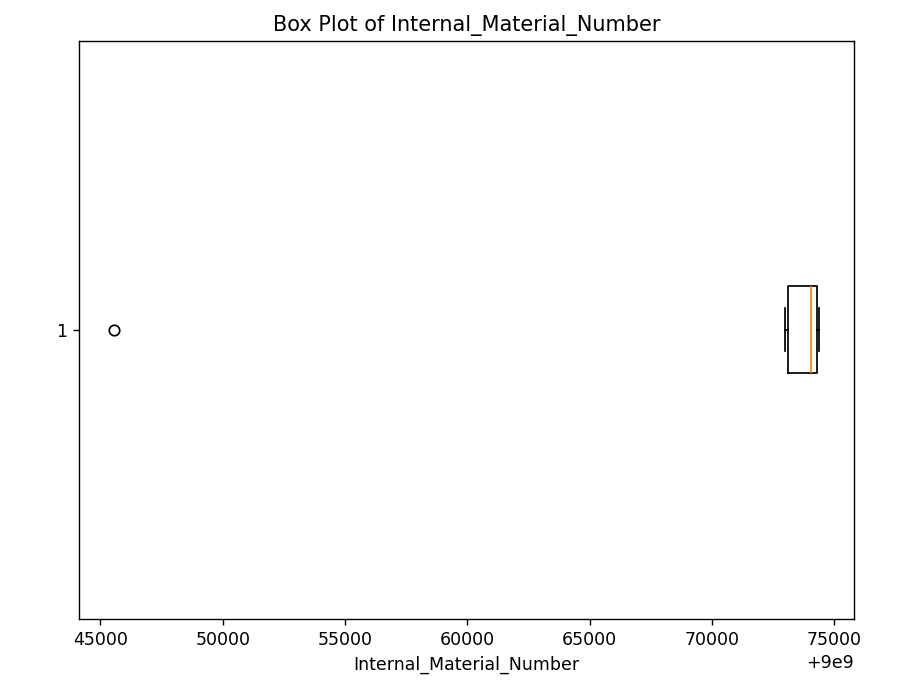
*plt.figure(figsize=(8, 6))*

*plt.boxplot(Internal\_Material\_Number, vert=False)*

*plt.xlabel('Internal\_Material\_Number')*

*plt.title('Box Plot of Internal\_Material\_Number')*

*plt.show()*



1. Using a scatter plot, explore the relationship between 'Column\_A' and 'Column\_B' in the H001\_terminal dataset.

*import pandas as pd*

*import seaborn as sns*

*import matplotlib.pyplot as plt*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*columns\_of\_interest = ['Material\_Description', 'Material\_Group']*

*plt.figure(figsize=(10, 6))*

*sns.scatterplot(data=H001\_terminal, x='Material\_Description', y='Material\_Group')*

*plt.xticks(rotation=90)*

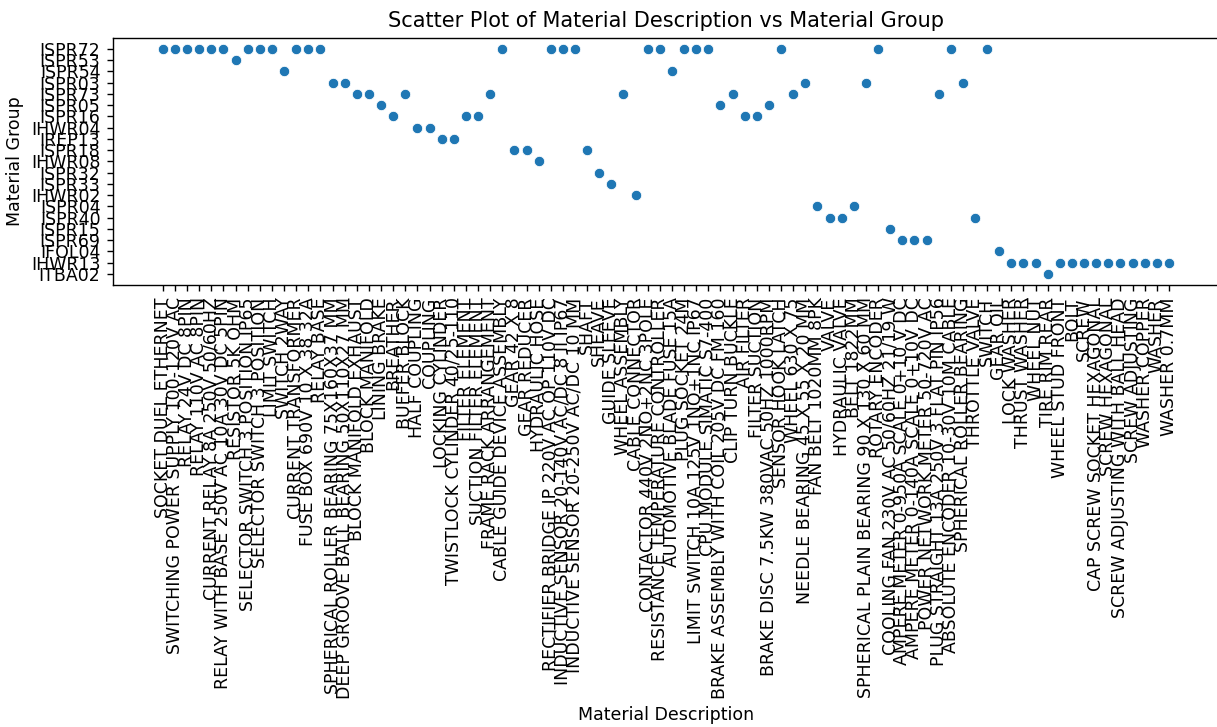
*plt.xlabel('Material Description')*

*plt.ylabel('Material Group')*

*plt.title('Scatter Plot of Material Description vs Material Group')*

*plt.tight\_layout()*

*plt.show()*

**

1. Develop a bar chart to visualize the frequency of each category in 'Column\_H' in the H001\_terminal dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*H001\_terminal= pd.read\_excel('Terminal.xlsx')*

*material\_group\_counts = H001\_terminal['Material\_Group'].value\_counts()*

*plt.figure(figsize=(10, 6))*

*material\_group\_counts.plot(kind='bar', color='skyblue')*

*plt.title('Frequency of Each Material Group')*

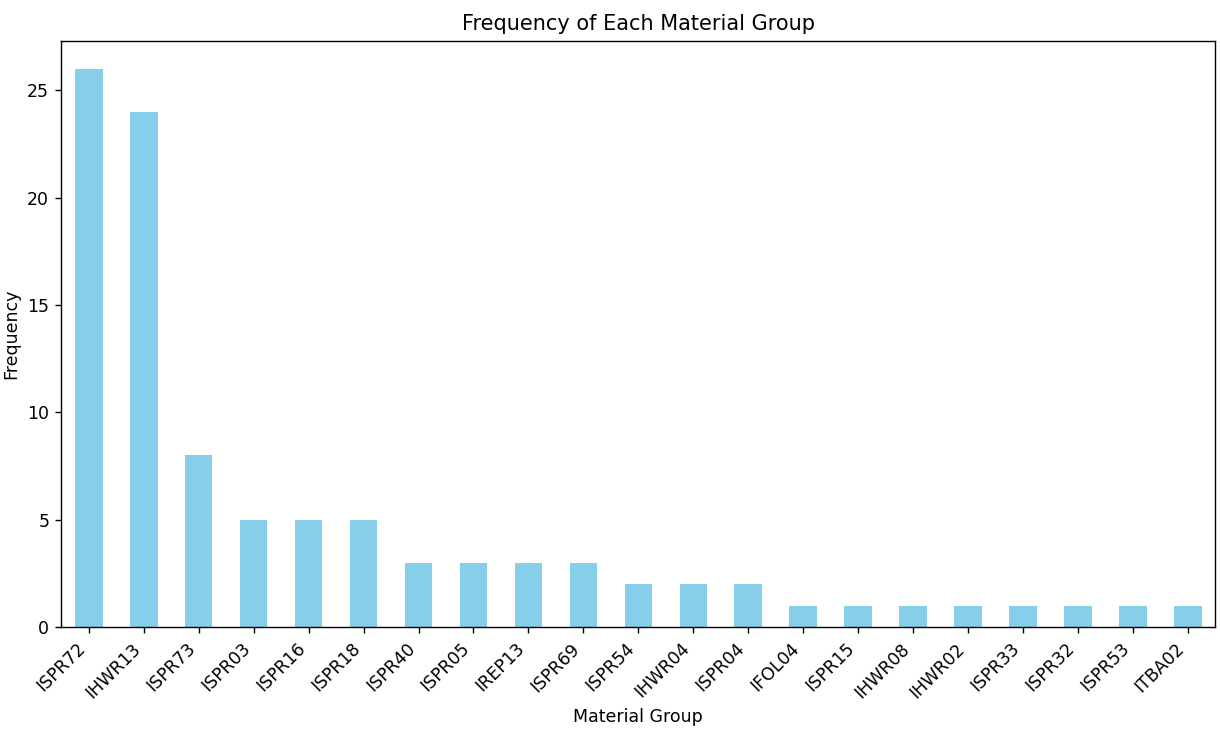
*plt.xlabel('Material Group')*

*plt.ylabel('Frequency')*

*plt.xticks(rotation=45, ha='right')*

*plt.tight\_layout()*

*plt.show()*

**

1. Utilize a pie chart to represent the proportion of unique values in 'Column\_H' in the H001\_terminal dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*H001\_terminal= pd.read\_excel('Terminal.xlsx')*

*unique\_values = H001\_terminal['Material\_Group'].unique()*

*value\_counts = H001\_terminal['Material\_Group'].value\_counts()*

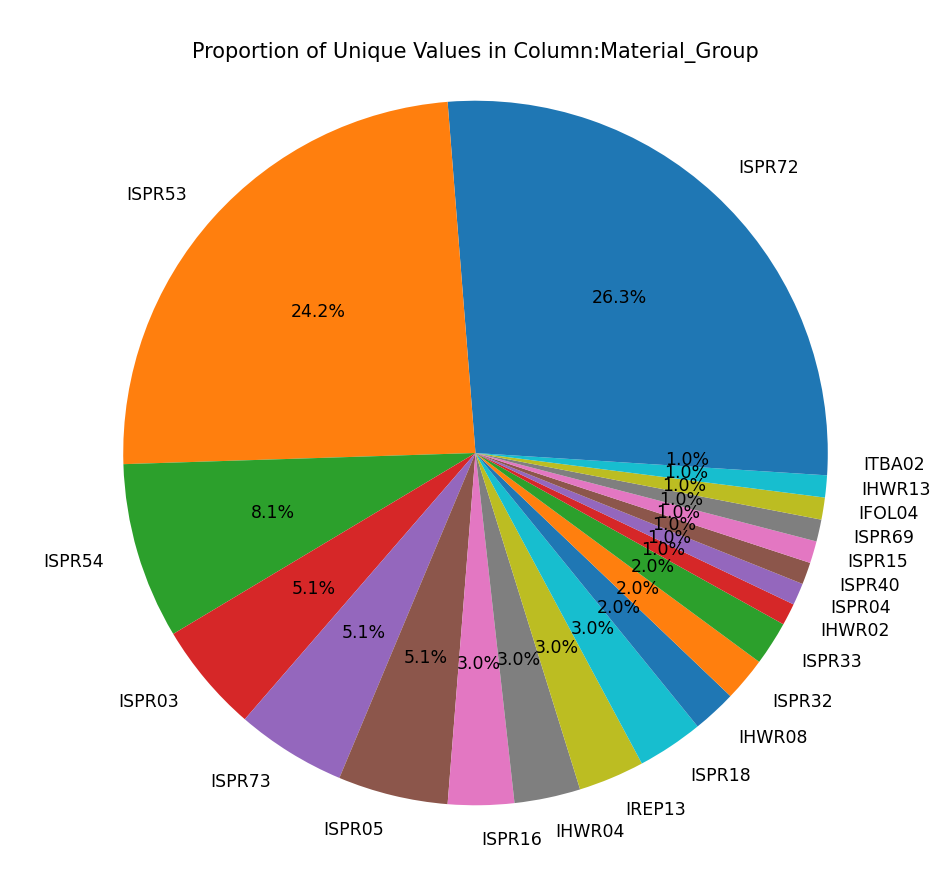
*plt.figure(figsize=(8, 8))*

*plt.pie(value\_counts, labels=unique\_values, autopct='%1.1f%%')*

*plt.title('Proportion of Unique Values in Column:Material\_Group')*

*plt.axis('equal')*

*plt.show()*

**

1. Create a heatmap to visualize the correlation matrix of numerical columns in the H001\_terminal dataset.

*import pandas as pd*

*import seaborn as sns*

*import matplotlib.pyplot as plt*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*numeric\_cols = H001\_terminal.select\_dtypes(include=['float64', 'int64'])*

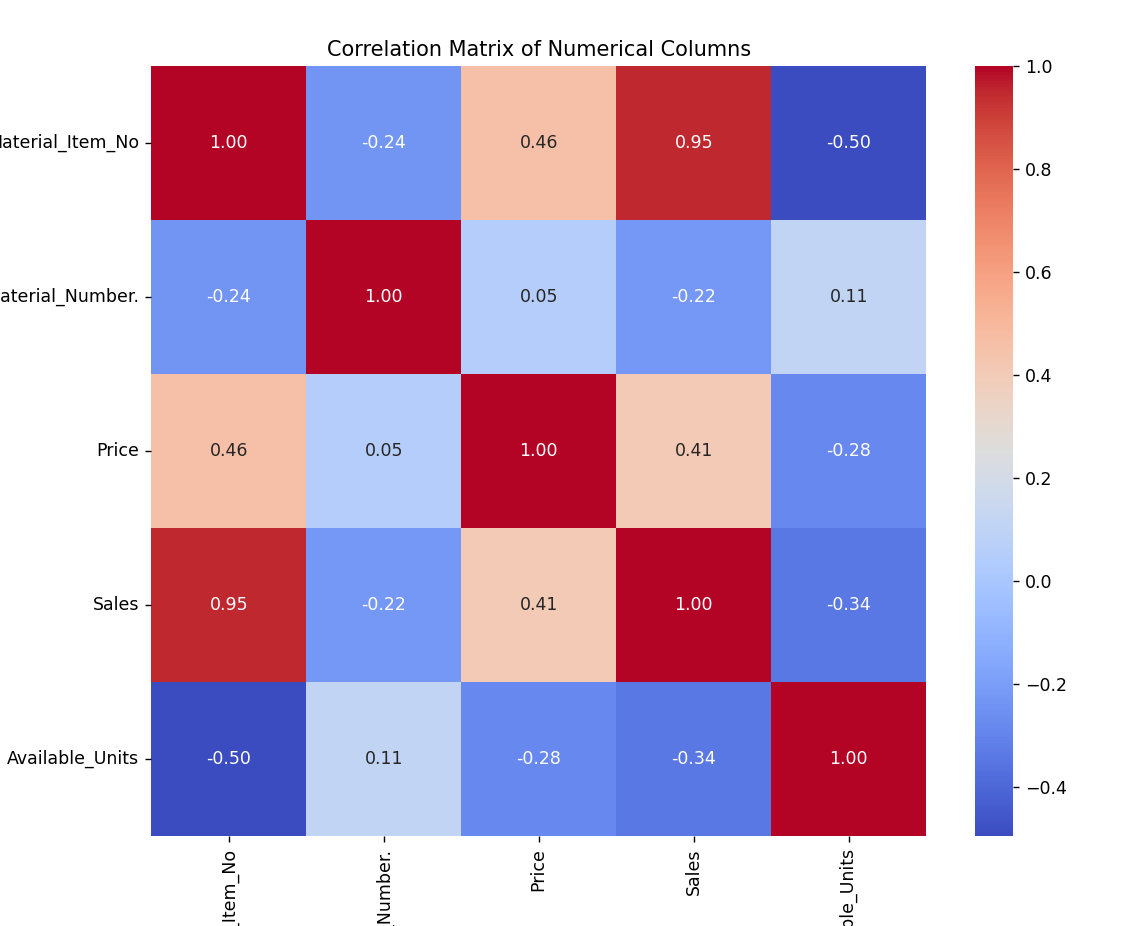
*corr\_matrix = numeric\_cols.corr()*

*plt.figure(figsize=(10, 8))*

*sns.heatmap(corr\_matrix, annot=True, cmap='coolwarm', fmt=".2f")*

*plt.title('Correlation Matrix of Numerical Columns')*

*plt.show()*

**

1. Generate a pair plot for a subset of columns to explore potential relationships between variables in the H001\_terminal dataset.

*import pandas as pd*

*import seaborn as sns*

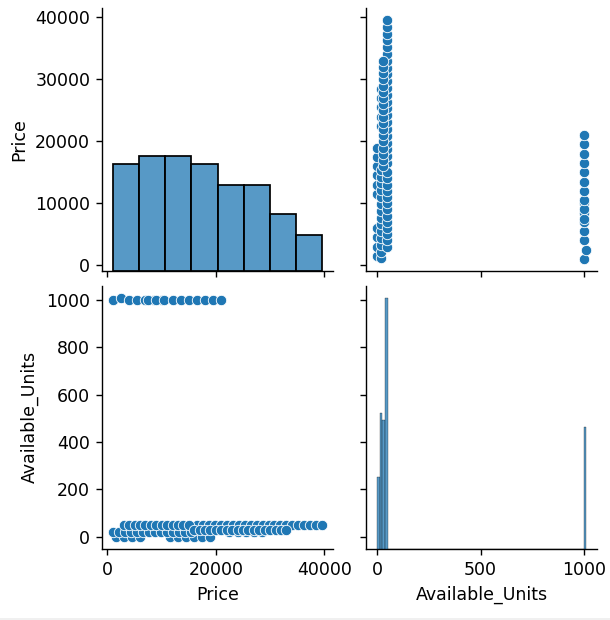
*import matplotlib.pyplot as plt*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*subset\_cols = ['Price', 'Available\_Units']*

*sns.pairplot(H001\_terminal[subset\_cols])*

*plt.show()*



1. Plot a time series chart if applicable, illustrating the changes in 'Column\_H' over time in the H001\_terminal dataset.

*import pandas as pd*

*import seaborn as sns*

*import matplotlib.pyplot as plt*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*H001\_terminal['Date'] = pd.to\_datetime(H001\_terminal['Date of order'], format='%d-%m-%Y')*

*plt.figure(figsize=(10, 6))*

*plt.plot(H001\_terminal['Date'], H001\_terminal['Material\_Group'], marker='o', linestyle='-')*

*plt.title('Material\_Group Over Time')*

*plt.xlabel('Date')*

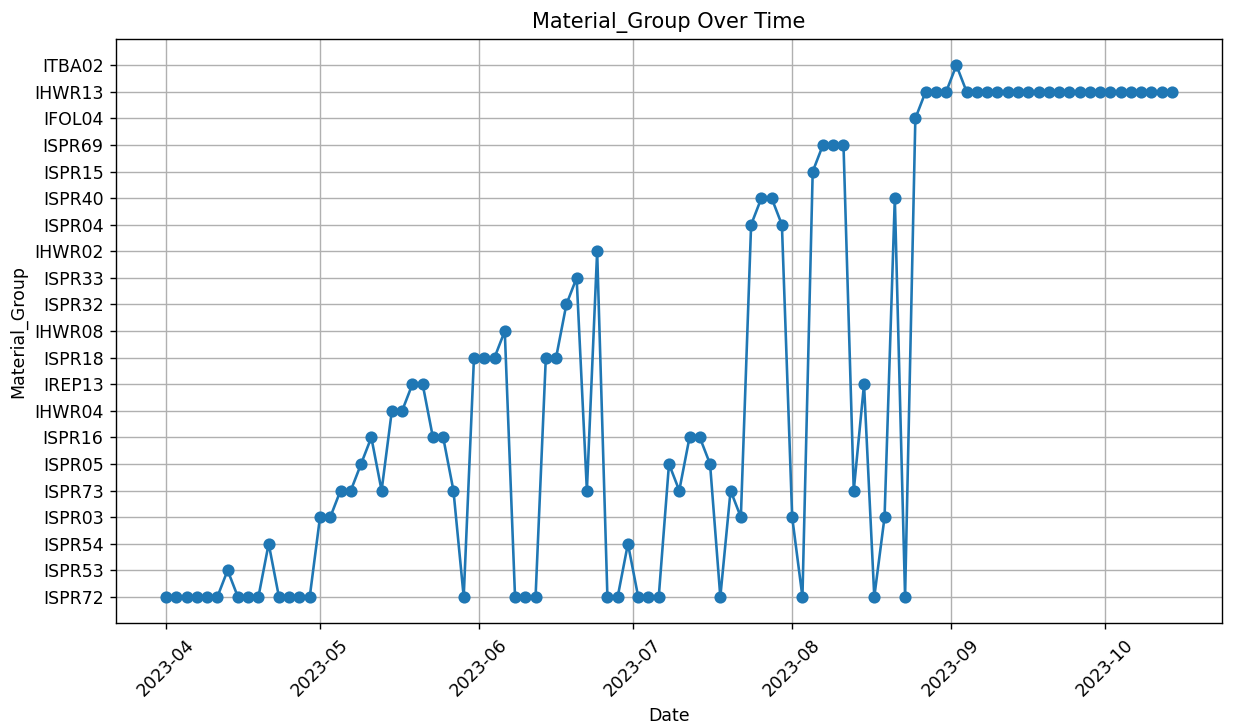
*plt.ylabel('Material\_Group')*

*plt.xticks(rotation=45)*

*plt.grid(True)*

*plt.tight\_layout()*

*plt.show()*

**

1. Develop a violin plot to visualize the distribution of 'Column\_F' across different categories in the H001\_terminal dataset.

*import pandas as pd*

*import seaborn as sns*

*import matplotlib.pyplot as plt*

*H001\_terminal = pd.read\_excel('Terminal.xlsx')*

*H001\_terminal['Material\_Group'] = H001\_terminal['Material\_Group'].astype(str)*

*plt.figure(figsize=(12, 8))*

*sns.violinplot(x='Material\_Group', data=H001\_terminal)*

*plt.title('Distribution of Material\_Group across Different Categories')*

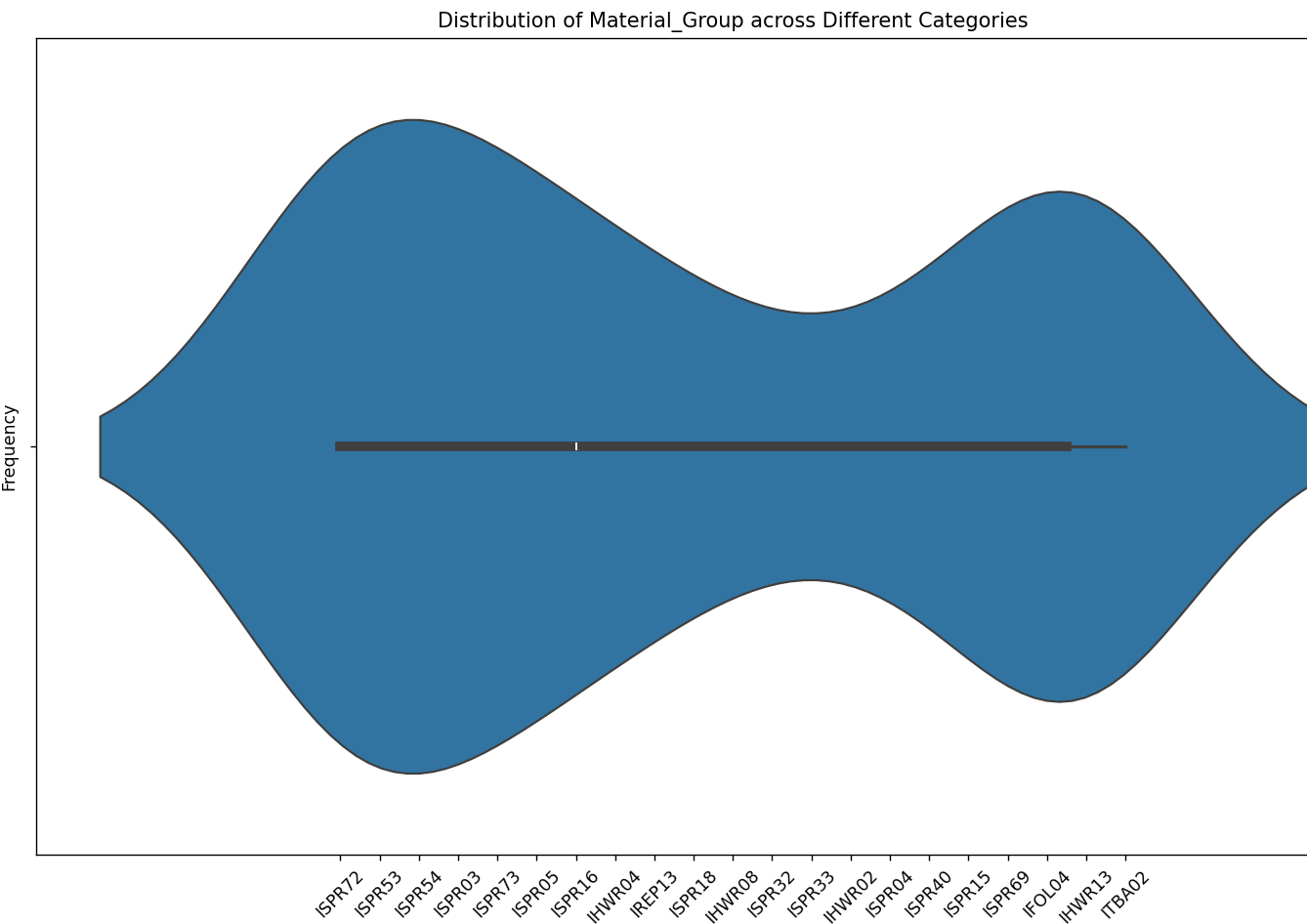
*plt.xlabel('Material\_Group')*

*plt.ylabel('Frequency')*

*plt.xticks(rotation=45)*

*plt.tight\_layout()*

*plt.show()*



**PartNumber\_ALLTerminals\_Data Dataset:**

1. Create a bar chart to compare the quantities of items in different categories in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*grouped\_data = PartNumber\_all.groupby('Material\_Group')['Available\_Units'].sum()*

*plt.figure(figsize=(10, 6))*

*grouped\_data.plot(kind='bar', color='skyblue')*

*plt.title('Quantities of Items in Different Material Groups')*

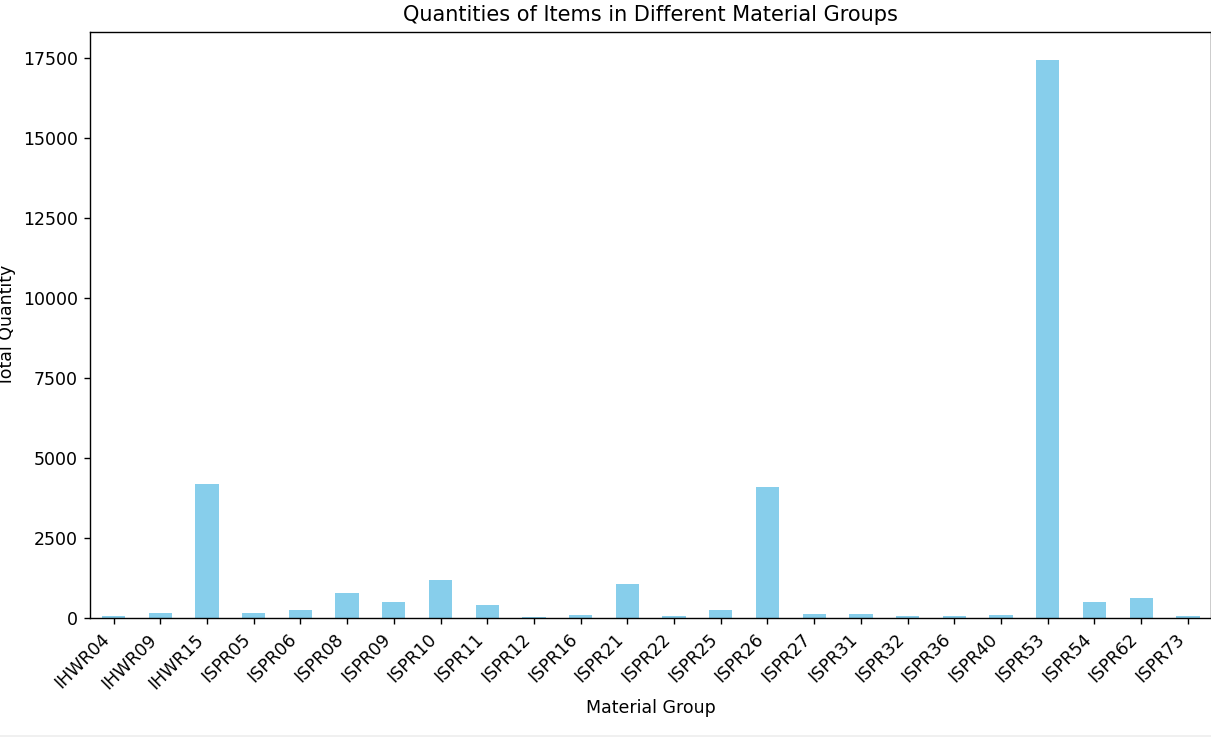
*plt.xlabel('Material Group')*

*plt.ylabel('Total Quantity')*

*plt.xticks(rotation=45, ha='right')*

*plt.tight\_layout()*

*plt.show()*

**

1. Generate a scatter plot to visualize the relationship between 'Quantity' and 'Price' in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*quantity = PartNumber\_all['Available\_Units']*

*price = PartNumber\_all['Price']*

*plt.figure(figsize=(10, 6))*

*plt.scatter(quantity, price, color='skyblue', alpha=0.7)*

*plt.title('Relationship between Quantity and Price')*

*plt.xlabel('Quantity')*

*plt.ylabel('Price')*

*plt.grid(True)*

*plt.tight\_layout()*

*plt.show()*

**

1. Develop a stacked bar chart to show the composition of 'Material\_group' based on the 'Quantity' in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*plt.figure(figsize=(10, 6))*

*plt.bar(PartNumber\_all['Material\_Group'], PartNumber\_all['Available\_Units'], label='Available Units')*

*plt.bar(PartNumber\_all['Material\_Group'], PartNumber\_all['Available\_Units'], bottom=PartNumber\_all['Available\_Units'], label='Available\_Units')*

*plt.title('Composition of Material Group based on Quantity')*

*plt.xlabel('Material Group')*

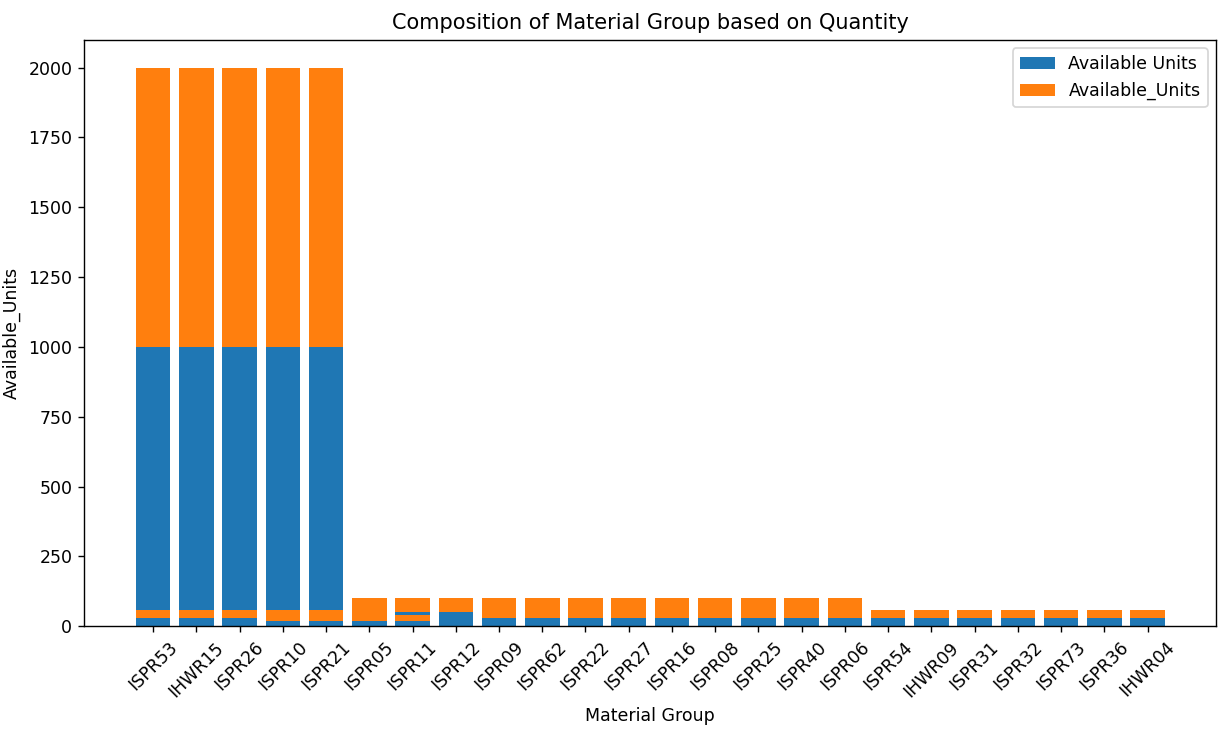
*plt.ylabel('Available\_Units')*

*plt.xticks(rotation=45)*

*plt.legend()*

*plt.tight\_layout()*

*plt.show()*

**

1. Utilize a treemap to represent the hierarchical structure of 'Category' and 'Subcategory' in the PartNumber\_ALLTerminals\_Data dataset.—*No relevant data in excel for 'Category' and 'Subcategory'*
2. Create a radar chart to compare the values of multiple numerical columns in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*import numpy as np*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*columns\_to\_compare = ['Price', 'Sales', 'Available\_Units']*

*num\_vars = len(columns\_to\_compare)*

*angles = np.linspace(0, 2 \* np.pi, num\_vars, endpoint=False).tolist()*

*mean\_values = PartNumber\_all[columns\_to\_compare].mean()*

*values = mean\_values.values.flatten().tolist()*

*values += values[:1]*

*angles += angles[:1]*

*fig, ax = plt.subplots(figsize=(8, 8), subplot\_kw=dict(polar=True))*

*ax.fill(angles, values, color='red', alpha=0.25)*

*ax.plot(angles, values, color='red', linewidth=2, linestyle='solid')*

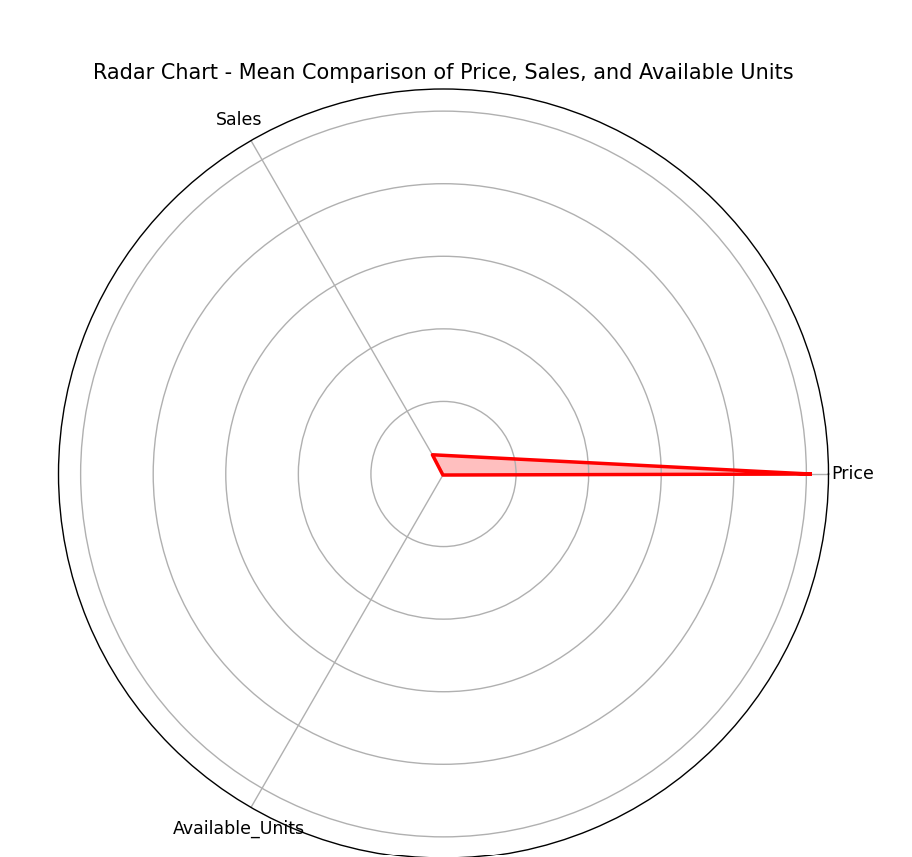
*ax.set\_yticklabels([])*

*ax.set\_xticks(angles[:-1])*

*ax.set\_xticklabels(columns\_to\_compare)*

*plt.title('Radar Chart - Mean Comparison of Price, Sales, and Available Units')*

*plt.show()*



1. Develop a bubble chart to show the distribution of 'Price' and 'Quantity' across different 'material group' in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*grouped\_data = PartNumber\_all.groupby('Material\_Group').agg({'Price': 'mean', 'Sales': 'sum'}).reset\_index()*

*plt.figure(figsize=(10, 6))*

*plt.scatter(grouped\_data['Price'], grouped\_data['Sales'], s=grouped\_data['Sales']\*10, alpha=0.5)*

*plt.title('Bubble Chart - Distribution of Price and Quantity across Material Groups')*

*plt.xlabel('Mean Price')*

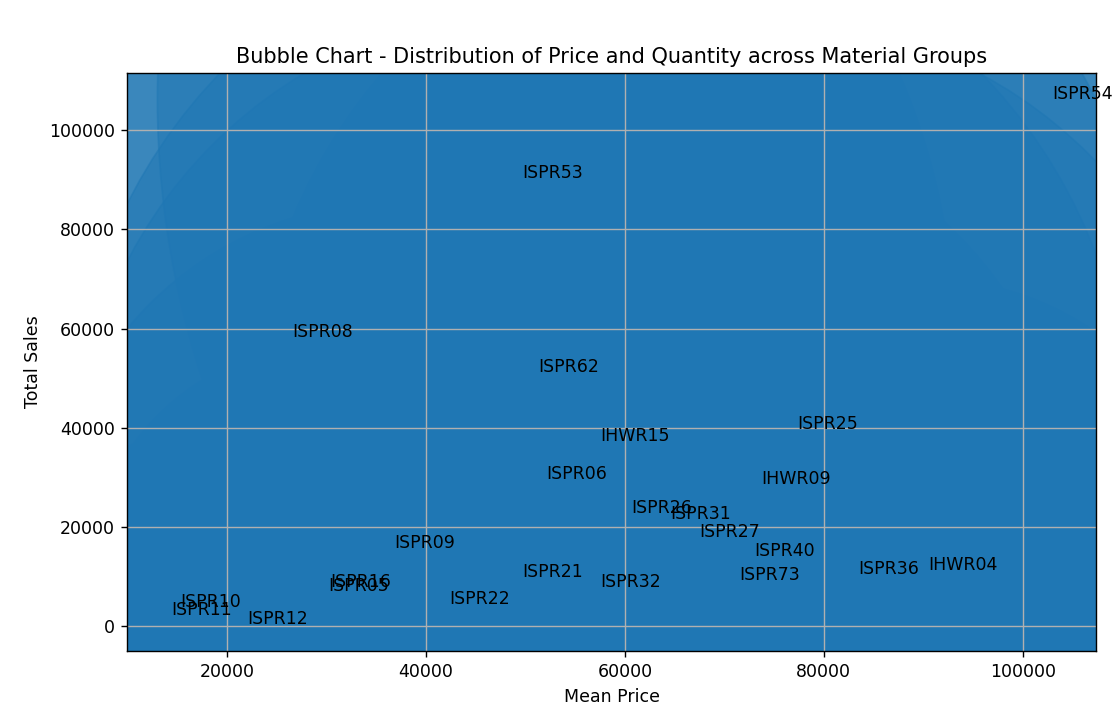
*plt.ylabel('Total Sales')*

*for i, txt in enumerate(grouped\_data['Material\_Group']):*

*plt.annotate(txt, (grouped\_data['Price'][i], grouped\_data['Sales'][i]))*

*plt.grid(True)*

*plt.show()*

**

1. Use a 3D scatter plot to visualize the relationships between three numerical columns in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*from mpl\_toolkits.mplot3d import Axes3D*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*x = PartNumber\_all['Price']*

*y = PartNumber\_all['Sales']*

*z = PartNumber\_all['Available\_Units']*

*fig = plt.figure(figsize=(10, 8))*

*ax = fig.add\_subplot(111, projection='3d')*

*ax.scatter(x, y, z, c='b', marker='o')*

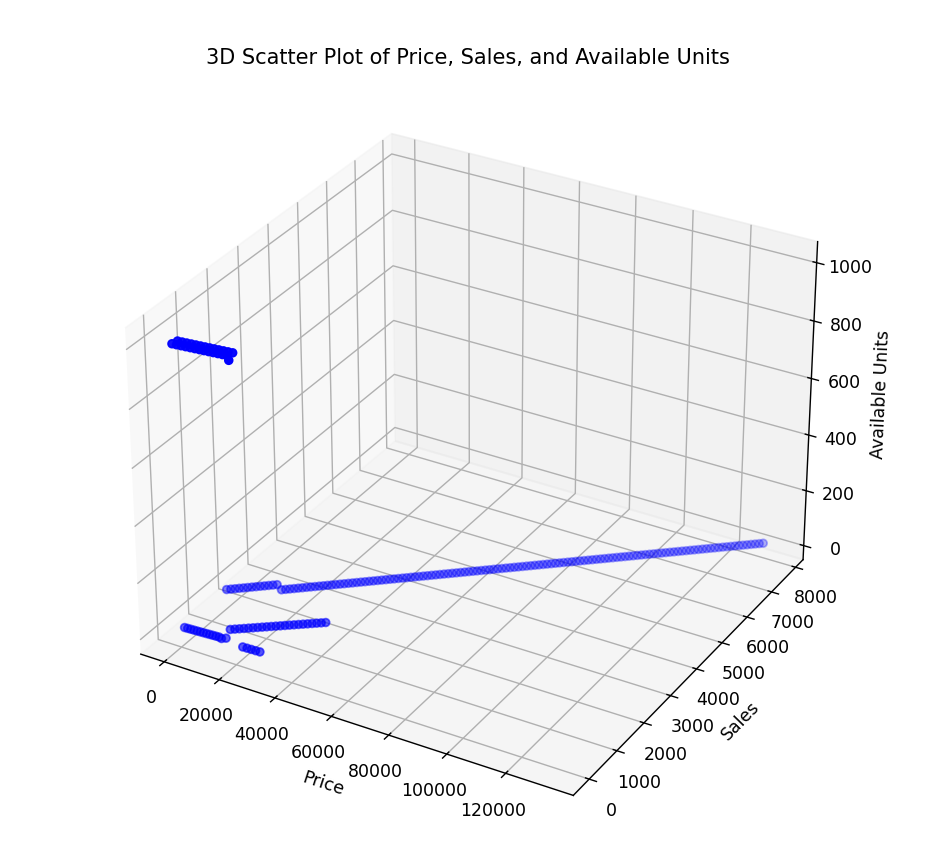
*ax.set\_xlabel('Price')*

*ax.set\_ylabel('Sales')*

*ax.set\_zlabel('Available Units')*

*ax.set\_title('3D Scatter Plot of Price, Sales, and Available Units')*

*plt.show()*

**

1. Create a bar chart with error bars to show the mean and variability of 'Price' across different Material\_group in the PartNumber\_ALLTerminals\_Data dataset.

*import pandas as pd*

*import matplotlib.pyplot as plt*

*PartNumber\_all = pd.read\_excel('PartNumber\_all.xlsx')*

*grouped\_data = PartNumber\_all.groupby('Material\_Group')['Price'].agg(['mean', 'std'])*

*plt.figure(figsize=(10, 6))*

*grouped\_data['mean'].plot(kind='bar', yerr=grouped\_data['std'], capsize=5, color='skyblue')*

*plt.title('Mean and Variability of Price across Material Groups')*

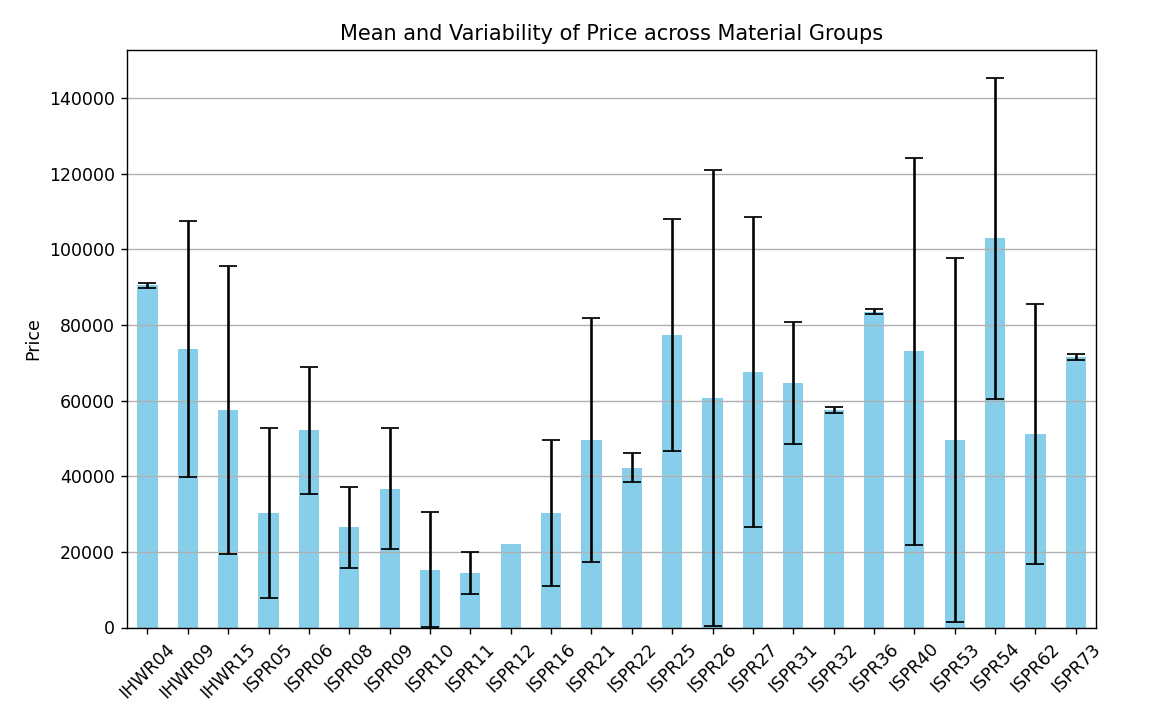
*plt.xlabel('Material Group')*

*plt.ylabel('Price')*

*plt.xticks(rotation=45)*

*plt.grid(axis='y')*

*plt.show()*



1. Generate a sunburst chart to visualize the hierarchy of 'Category' and 'Subcategory' in the PartNumber\_ALLTerminals\_Data dataset.

--*No relevant data in excel to proceed*

1. Create a choropleth map if there is geographical information in the PartNumber\_ALLTerminals\_Data dataset, representing the distribution of items.

--*No relevant data in excel to proceed*