**Bike Stores Sales Analysis**

***Approach:***

1. **Ask**: Identify a specific question or business problem to solve
2. **Prepare**: Collect relevant data from reliable sources and manage the data
3. **Process**: Clean the data by handling missing values, duplicates, errors, etc
4. **Analyze**: Perform exploratory data analysis (EDA) to identify patterns and outliers
5. **Visualize**: Create clear, impactful visualizations that highlight key findings
6. **Share**: Communicate the results effectively

***Objective:***

To analyze the sales activities within the company and provide actionable insights into sales trends from 2016 to 2018. The analysis will focus on the following key areas:

1. **Sales Trends**: Identify patterns and fluctuations in sales volume over the 2016–2018 period
2. **Revenue Analysis**: Break down revenues by region, store, product category, and brand to pinpoint high-performing areas
3. **Customer Insights**: Generate a list of top customers based on purchase volume and revenue contribution
4. **Sales Rep Performance**: Identify top-performing sales representatives based on sales volume and revenue generated

***Process:***

**I. SQL:**

1. **Load the data into Microsoft SQL Server:**

Import the data carefully into the appropriate tables and schema, ensuring everything is set up correctly for querying

1. **Ensure all the data is loaded properly with no errors in between:**

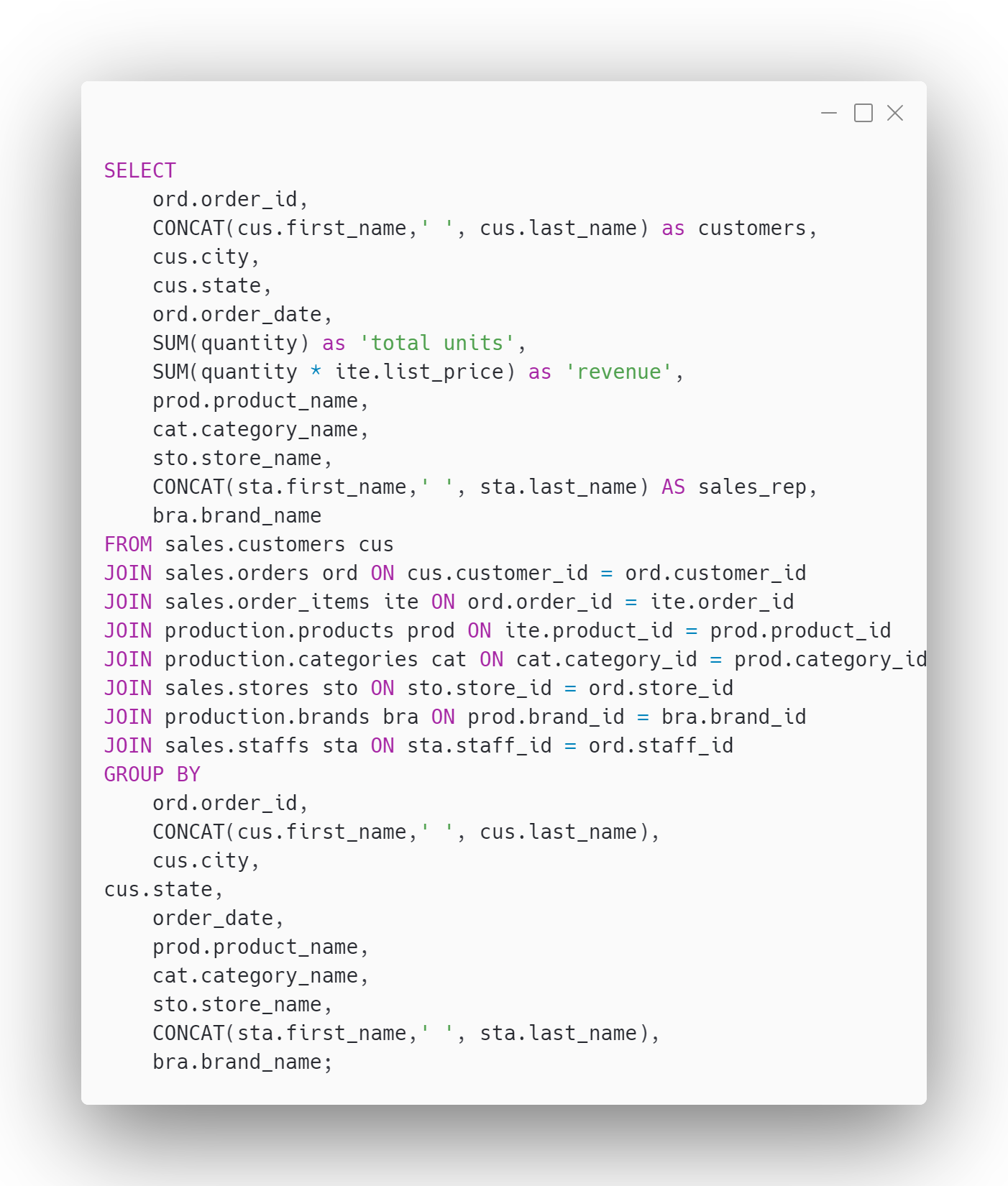
Verify that the data transfer is complete, with no missing rows or errors like data type mismatches or key violations

1. **Spend some time familiarizing yourself with the data:**

Explore table structures, relationships to get an overview of the data

1. **Write an SQL query to consolidate all important parameters into a single table while ensuring the project objectives are met:**

Build a well-structured query that combines relevant columns using joins and aggregations, optimized for accuracy and performance. The final query:



**II. Excel:**

1. **Establish a connection from SQL to Excel and import the SQL queried data:**

Navigate to the Data tab in Excel, select "New Query" from the database options, and choose "From SQL Server Database" Then, input the exact SQL query you wrote to pull the necessary data into Excel for further analysis

1. **Keep the project objectives in mind while creating the required tables and graphs:**

Focus on delivering the specific data and visualizations requested by the client, ensuring that the tables are structured as required and the graphs are created to represent the metrics effectively for analysis

1. **Create required tables and graphs for visualization:**

Use pivot tables and charts for key metrics like trends, comparisons, or summary statistics. Ensure graphs align with project goals, such as sales volume trends or regional performance insights.

1. **Make a dashboard in Excel with Slicers to make the graphs interactive:**

Create a dynamic dashboard in Excel by incorporating slicers, which allow users to filter data visually and interact with the graphs seamlessly. This enhances usability and provides an engaging way to analyse data

1. **Validate and polish the final dashboard:**

Test slicers for proper interactivity and ensure they control all relevant graphs. Add titles, labels, and colour schemes to enhance usability and readability. Hide Gridlines and Graph borders to make the dashboard aesthetically pleasing.

**III. Tableau:**

1. **Connect Tableau to MS Excel:**

* Establish a connection between Tableau and the MS Excel workbook previously worked on
* Verify the data by checking the columns in the workbook for errors, missing data, or incorrect mappings

1. **Verify the Data Integrity:**

* Ensure all data fields have been properly imported and mapped in Tableau
* Check for issues like null values or data type mismatches and correct them if necessary

1. **Create Graphs and Charts for Objectives:**

* Begin by creating a new worksheet for each chart or graph required to meet the project objectives
* Example: For "Identify patterns and fluctuations in sales volume over the 2016–2018 period," create line graphs or bar charts that display trends over time
* Use appropriate measures and dimensions to explore data patterns and achieve actionable insights

1. **Prepare Summary Worksheets for the Dashboard:**

* Create additional worksheets that will serve as summary tabs for the dashboard
* Example: Include metrics like total revenue, total units sold, total customers, and total orders, formatted to appear in a single row on the dashboard

1. **Build and Align the Dashboard:**

* Combine the prepared worksheets into a single dashboard layout
* Ensure graphs, charts, and summary tabs are aligned and proportionally sized

1. **Add Interactivity with Slicers:**

* Include slicers (filters) such as "Year" and "State" in the dashboard
* Ensure these slicers dynamically update all connected graphs and metrics to reflect user selections

1. **Final Touches to the Dashboard:**

* Polish the dashboard by adding titles, legends, and tooltips for clarity
* Adjust the colour schemes and alignments to make the visuals more appealing
* Test the dashboard for functionality, ensuring all slicers and charts behave as intended

***Insights:***

***1. Sales Trends***

* Mount Vernon led city-wise revenue at $117,010.21, showing strong customer loyalty.
* Ballston Spa followed with $110,065.34, indicating potential for further growth.
* San Angelo, Baldwinsville, and Howard Beach maintained stable revenue between $100K and $110K, with room for expansion.
* New York dominated state-wise revenue with $5.83M, followed by California ($1.79M) and Texas ($962.6K), where growth opportunities exist.

***2. Revenue Analysis***

* **Stores**: Baldwin Bikes led store-wise revenue at $5.83M, followed by Santa Cruz Bikes ($1.79M) and Rowlett Bikes ($962.6K).
* **Categories**: Mountain Bikes generated $3.03M, followed by Road Bikes ($1.85M) and Cruisers ($1.11M). Electric Bikes ($1.02M) show emerging potential.
* **Products**: Trek Slash 8 27.5 - 2016 was the top seller with $615,998.46, followed by Trek Conduit+ - 2016 ($434,998.55) and Surly Straggler 650b - ($253,829.49).

***3. Top Customers:***

|  |  |
| --- | --- |
| Marcelene Boyer | $29,38,889 |
| Venita Daniel | $28,87,353 |
| Genna Serrano | $9,52,722 |
| Mireya Copeland | $8,37,424 |
| Kali Vargas | $5,16,695 |
| Layla Terrell | $4,45,906 |

***4. Top Salesmen:***

|  |  |
| --- | --- |
| Pamelia Newman | $37,802 |
| Abby Gamble | $37,501 |
| Sharyn Hopkins | $37,139 |
| Lyndsey Bean | $35,858 |
| Emmitt Sanchez | $34,504 |
| Melanie Hayes | $34,391 |
| Debra Burks | $30,646 |
| Elinore Aguilar | $29,662 |
| Corrina Sawyer | $29,215 |
| Shena Carter | $27,619 |

***Strategies:***

1. **Top Stores**:

* **Baldwin Bikes**: Invest in exclusive promotions and loyalty programs to sustain dominance.
* **Santa Cruz Bikes**: Introduce targeted marketing campaigns to boost visibility and revenue growth.
* **Rowlett Bikes**: Focus on community engagement and localized promotions to capitalize on untapped potential.

2. **Product Categories**:

* **Mountain Bikes**: Launch seasonal discounts and upgrade options to maintain dominance.
* **Electric Bikes**: Expand product lines and offer test rides to tap into the growing demand.

3. **Top Products**:

* Promote top-performing products like **Trek Slash** with bundled offers and extended warranties to boost customer retention.

4. **Top Customers**:

* Offer **tier-based loyalty programs** (e.g., gold, platinum levels) with escalating benefits like discounts, free servicing, or exclusive previews.
* Introduce a **referral bonus system**, rewarding top customers for bringing in new clients

5.**Top Salesmen**:

* Implement a **tiered commission structure** with higher incentives for surpassing sales targets.
* Recognize achievements through **"Employee of the Month" awards** and public accolades.
* Offer **career development opportunities**, like leadership training or certifications, to build loyalty and skillsets.

***Challenges Faced:***

1. **MSSQL Server Connection Issue:**

* **Problem:** Faced "Error Locating Server/Instance Specified" while attempting to connect to MSSQL Server using a specific server name
* **Initial Checks:** Verified the server's name and instance in SQL Server Management Studio (SSMS)
* **Troubleshooting:** Examined SQL Server Configuration Manager for enabled protocols like TCP/IP and ensured SQL Server Browser service was running
* **Additional Steps:** Used ODBC Data Source Administrator to confirm the data source settings and checked the port being used for connections
* **Solution:** Discovered the issue stemmed from the incorrect server’s name format; logging in with "DESKTOP-T9AF2BR" instead of "DESKTOP-T9AF2BR\SQLEXPRESS" successfully established the connection

1. **MSSQL to Excel Database Connection:**

* **Problem:** Encountered difficulties linking the MSSQL database to Excel due to missing features in Excel 2016 Student version, such as PowerPivot
* **Challenges:** Noticed that certain advanced connection options available in newer Excel versions were unavailable
* **Initial Attempts:** Tried alternative methods like manually adding a new ODBC data source using "ODBC Driver 17 for SQL Server" to create a pathway between Excel and MSSQL
* **Discovery:** Realized that "Get Data" in newer Excel versions is equivalent to "New Query" in older ones like Excel 2016
* **Solution:** Successfully connected to the database using "New Query" while adapting to the limitations of the older Excel version

1. **Shared Data Cache in Pivot Tables:**

* **Problem:** Multiple PivotTables using the same data source shared a single data cache, causing changes in one PivotTable to affect others
* **Impact:** This limitation prevented the ability to independently display different metrics in separate PivotTables
* **Cause:** PivotTables were automatically linked to the same data cache because they shared the same data source
* **Exploration:** Experimented with altering cache-sharing settings and considered duplicating the data source to create independent caches
* **Solution:** Used the PivotTable and PivotChart Wizard (shortcut: Alt+D+P) to generate separate data caches, allowing for greater flexibility in metrics presentation

1. **Slicer Report Connections Issue:**

* **Problem:** Could not access "Report Connections" for slicers like "State" and "Store Name," even after inserting slicers from the original table (Sheet 1)
* **Reason:** Slicers must be directly tied to a pivot table or chart to enable the option
* **Solution:**
* Go to the Dashboard (Sheet 3)
* Select the chart with the desired field (e.g., "State")
* Insert the slicer while the chart is active
* Link the slicer to other charts using "Report Connections

1. **Creating a parameter on Tableau to view Top ‘n’ Customers:**

* **Create Parameter:**
* Right-click in the Data Pane > Create Parameter > Name it (e.g., "Top N Customers")
* Set Integer as the type and assign a default value
* **Create Set**:
* Right-click **Customer Name** > **Create > Set**
* In the **Top** tab, select **By Field**, choose your parameter, and click **OK**
* **Apply Set to Filter**:

Drag the set to the **Filters** shelf to filter based on the parameter

* **Enable Parameter Control**:

Right-click the parameter > **Show Parameter** for dynamic input

* **Adjust & Test**:

Enter a value in the parameter control to display the top 'N' customers

1. **Creating a Summary Tile for Total Revenue, Customers, Orders, and Units Sold:**

* **Problem:**  
  Faced difficulties adding a summary tile, particularly with placements and alignment.
* **Challenges:**  
  Struggled to consolidate multiple metrics into a single, visually organized tile.
* **Solution:**
  1. Created separate sheets for each parameter (e.g., total revenue, customers).
  2. On the dashboard, inserted a horizontal container to organize the sheets.
  3. Added the sheets into the container and used the center-align option to ensure proper placement.

1. **Adding a Custom Box to Select Top 'N' Customers (Dynamic Selection):**

* **Problem:**  
  Needed to create a dynamic selection for displaying the top 'N' customers but was unsure of the steps.
* **Solution:**
  1. **Create a Parameter:**
     + Right-click the Data Pane → Create Parameter.
     + Name it *Top N Customers*.
     + Set the default value to 10 and the data type to Integer.
  2. **Create a Set:**
     + Right-click the *Customer* field → Create → Set.
     + Name it *Top Customers*.
     + In the *Top* tab, select “By Field” and use the *Top N Customers* parameter.
  3. **Apply the Set as a Filter:**
     + Drag the *Top Customers* set to the Filters shelf.
  4. **Show the Parameter Control:**
     + Right-click the *Top N Customers* parameter → Select *Show Parameter*.
  5. Adjusted and tested by inputting various values in the parameter control.

1. **Aligning Sheets on the Dashboard and Ensuring a Pleasing Colour Palette:**

* **Problem:**  
  Struggled with cramped sheet placement, uneven alignment, and inconsistent colours on the dashboard.
* **Challenges:**  
  Aligning multiple visualizations while ensuring aesthetics was time-consuming.
* **Solution:**

1. Alignment and Spacing:

* + Used horizontal and vertical containers to organize sheets systematically.
  + Adjusted padding for consistent spacing and added inner padding to prevent content from touching edges.

2. Optimize Sheet Sizes:

* + Set each sheet’s size to *Fit Width* or *Entire View* to avoid overcrowding.
  + Dynamically resized containers to ensure no element overlapped with others.

3. Colour Coordination:

* + Chose 3–5 complementary colours using tools like Adobe Colour.
  + Applied a consistent theme for headings and labels across all sheets.