**STEPS FOR DATSET IN POWER BI**

**1) SALES DATASET :-**

**Step 1: Import data**

**\*\*Launch Microsoft Power BI Desktop to allow you to create the visualizations.**

**\*\*Import the Adventure Works sales dataset called Adventure Works sales dataset.csv. This is the dataset you want to visualize.**

**\*\*Step 2: Create interactive filters**

**One method for filtering data is a slicer—a visual tool that enables users to filter data interactively within a report.**

**\*\*To add a slicer for the product category, on the Home tab, select Slicer.**

**\*\*The Slicer visual on the Home tab.**

**\*\*Drag the Product Category field from the Fields pane to the slicer visual. This will create an interactive filter based on product categories.**

**---Product Category**

**\*\*Repeat the process to create another slicer, this time for the Payment Method field.**

**\*\*For each slicer, add a title by selecting the Format tab and choosing Title from the options.**

**\*\*Format tab of visualizations**

**\*\*Drag the positioning of the slicers to the right side of the report area.**

**Step 3: Create a stacked column chart**

**\*\*Add a stacked column chart to your report. From the Fields pane, drag the Order Total onto the Y-Axis, Product Size onto the X-Axis, and Product Category onto the Legend.**

**\*\*The stacked column chart icon in visualizations**

**\*\*Set the chart title to Order Total by Product Size and Product Category for 2023.**

**\*\*Drag the stacked column chart visual to the left side of the Report area (next to the interactive slicer filters).**

**Step 4: Create a donut chart**

**\*\*Add a donut chart for Order Total Share by Product Category.**

**--The donut chart icon in visualizations**

**\*\*Set the chart title to Order Total Share by Product Category for 2023.**

**\*\*Drag the donut chart visual below the stacked column chart (created in Step 3).**

**Step 5: Create a table**

**\*\*Add a table showing all product names with their corresponding order totals.**

**--The table visual**

**\*\*Drag the table visual to the right side of the Report area, below the interactive filters (created in Step 2).**

**2)Combine file (sales.customer feedback,inventory) :-**

**--Instructions**

**\*\*Download the Inventory, Sales and Customer Feedback CSV data files to your computer.**

**\*\*Right-click on the Sales Data CSV file and choose Open with > Excel. This will open the file in Excel.**

**\*\*Repeat the step 2 process to open the other two files.**

**\*\*Sales Dataset: What is the total sales value per product?**

**\*\*First, you are presented with the Sales dataset. It needs to be more organized. It has missing TransactionIDs, SalesAmounts, and text in place of numbers. The analysis question to answer is: “What is the total sales value per product?” To answer this, you'll have to fill in missing SalesAmounts, transform textual entries into numbers, and then aggregate the sales by product. The transformations you implement here will directly feed into the sales trends that Adventure Works will soon uncover.**

**\*\*Step 1: Addressing Missing TransactionIDs**

**\*\*Select the header of the TransactionID column to select the column.**

**\*\*On the Excel ribbon, in the Data tab, select the Filter button.**

**\*\*Select the drop-down arrow on the TransactionID column header and uncheck the box next to Blanks.**

**\*\*Ensure all entries in the column are treated as numerical by setting the data type to Number.**

**\*\*Step 2: Dealing with SalesAmount anomalies**

**\*\*Select the header of the SalesAmount column to select the column.**

**\*\*Manually go through the column to identify and replace non-numeric values like four-hundred. Delete the existing value and type 400 (or the appropriate numeric value).**

**\*\*Replace any missing SalesAmounts with 0 so that these entries are not excluded from the analysis. To replace null or blank cells with 0, press CTRL + H to bring up the Find and Replace dialog box. The Sales data is now clean and consistent, ready for further analysis in Power BI.**

**\*\*Inventory Dataset: What is the stock level for each bike category?**

**\*\*Next up is the Inventory dataset. Here, you notice missing ProductNames and Categories, and inconsistencies in RestockingFrequency. The task is to figure out: What is the stock level for each bike category? To discover this, you need to standardize Category and fill in missing product details. The transformed data will subsequently support improved inventory management strategies at Adventure Works.**

**--Step 1: Addressing Missing ProductName**

**\*\*Similar to handling missing TransactionIDs in Sales data, you will filter out any rows with missing ProductNames, ensuring your analysis only includes complete records. Select the ProductName column.**

**--Go to the Data tab and select Filter.**

**\*\*Select the drop-down arrow on the ProductName column and deselect Blanks. Click OK.**

**--Step 2: Handling Missing Category**

**\*\*Next, you start evaluating the missing Category data. Select the Category column.**

**\*\*Manually check the column and evaluate the missing entries.**

**\*\*Replace any missing with No Category Provided, press CTRL + H to bring up the Find and Replace dialog box.**

**--Step 3: Standardizing RestockingFrequency**

**\*\*Inconsistencies in RestockingFrequency can also hinder accurate analysis. Select the RestockingFrequency column.**

**--Press CTRL + H to open the Find and Replace dialog box.**

**\*\*In the Find what field, type 30 d, in the Replace with field, type 30 days, and select Match entire cell contents from the Options menu. Select Replace All. With the Inventory data clean and standardized, you can easily answer the business question What is the stock level for each bike category? when loaded into Power BI.**

**--Customer Feedback Dataset: What is the average feedback score for each product?**

**\*\*Finally, you check the Customer Feedback dataset. Here, FeedbackScores are irregular, dates are erroneous, and some feedback entries are missing altogether. Your analysis question is: What is the average feedback score for each product? This question requires you to correct the irregularities in FeedbackScores, fix erroneous dates, and determine how to treat missing feedback entries. These transformations will give Adventure Works the knowledge it needs to enhance customer satisfaction.**

**--Step 1: Addressing Missing FeedbackID**

**\*\*Begin by filtering out any rows with missing FeedbackIDs in the customer feedback data. Similar to missing TransactionIDs and ProductNames, missing FeedbackIDs can render the corresponding data points anonymous and potentially unreliable. Select the FeedbackID column.**

**--On the Data tab, select Filter.**

**\*\*Select the drop-down arrow on the FeedbackID column and deselect Blanks. Click OK.**

**--Step 2: Dealing with inconsistent FeedbackScore**

**\*\*Deal with inconsistencies in the FeedbackScore column by setting the data type to Number, ensuring all scores are treated as numerical values.**

**\*\*Manually go through the column to identify and replace non-numeric values like four-comma-seven. To do this, double-click on the cell with the non-numeric value, delete the existing value and type 4.7 (or the appropriate numeric value).**

**--Step 3: Invalid dates in the FeedbackDate column**

**Next, address erroneous dates in the FeedbackDate column. Select the FeedbackDate column.**

**\*\*Set the data type to Date on the ribbon, which will force Excel to interpret all entries as dates. To change the Regional Date Format, select the FeedbackDate column and navigate to More Number Formats from the dropdown menu on the ribbon. Select Custom then select the mm/dd/yyyy date format.**

**\*\*Manually check the column to identify and replace non-date values like five May 2023. To do this, double-click on the cell with the non-numeric value, delete the existing value and type 05 May 2023 (or the appropriate numeric value). Evaluation of the Customer Feedback data is now complete and prepared for future analysis in Power BI.**

**3)BIKES :-**

**Instructions**

**Create a new Word document and name it Product launch analysis. Follow the prompts below to complete the exercise.**

**Step 1: Identify the stakeholders**

**Understanding who the stakeholders are in a given scenario is a critical step in any data analysis project. In this case, you identify stakeholders (individuals or groups) within the company that will be directly affected by, or have a direct interest in, the sales trends and customer preferences that your analysis will reveal. These include the sales, marketing, and product development teams.**

**Now, answer the following questions related to stakeholder identification:**

**Briefly explain why any two of the stakeholder groups at Adventure Works may have an interest in the insights related to the sales performance of various products.**

**Which of the stakeholders need to know about customer preferences when it comes to Adventure Works’ products? Why is this information important for them? (Note: Customer preferences refer to customer wants, needs, and expectations regarding products and/or services.)**

**Step 2: Identify and consider the data sources**

**Download the Sales.csv file below and open it in Microsoft Excel.**

**This dataset contains valuable information that will be key to your analysis. The spreadsheet contains various fields, including:**

**Product ID**

**Product Category**

**Product Subcategory**

**Product Name**

**Product Description**

**Product Price**

**Product Weight**

**Product Size**

**Order ID**

**Customer ID**

**Order Date**

**Order Status**

**Order Quantity**

**Order Total**

**Payment Method**

**Note: Excel will not automatically format the Sales.csv document. To improve the readability of the data, you need to split the columns in Excel by delimiter, as follows:**

**Launch Microsoft Excel.**

**To open the CSV file, select File > Open. Navigate to the directory where your CSV file is saved. If you can't find your CSV file in the list, make sure you change the file type dropdown to All Files (.) or Text Files (\*.txt, \*.csv). Select your file, then the Open button.**

**When you open a CSV file directly in Excel, it might not automatically detect the semicolon delimiter. If the data isn't split properly, you need to use the Text Import Wizard. If it doesn't start automatically upon opening the file, you can manually start it by selecting the column with your data, followed by Data > Text to Columns. Then, select Delimited > Next. Uncheck any delimiters that are automatically selected, check Other and then enter the semicolon character (;) in the blank textbox. A preview of your data should display. Select Next.**

**In this optional step, you can set specific data formats for the columns. For general purposes, you can just leave the data format as General. Select Finish. Your data should now display as in the screenshot below.**

**Excel dataset screenshot with Adventure Works' sales data separated into columns by field type with field names highlighted.**

**Although the sales dataset is the only data source provided in this exercise, business data ordinarily comes from multiple sources. In the current case study, additional relevant data sources may include customer details, marketing data, supply chain data, feedback data, and website analytics. As a result, part of the data analysis process may be integrating data from different sources to get a complete picture.**

**Answer the following questions to guide your understanding of data sources in the case study:**

**What kind of information can you gather from the provided dataset in relation to product, order, and customer details?**

**How can you use the data fields provided to understand sales trends and customer preferences? Briefly discuss your suggested approach for gaining insight into these two aspects separately.**

**Explain how the other data sources within Adventure Works might be relevant to your analysis or add value to your insights.**

**Step 3: Data import and cleaning**

**Having identified your stakeholders and your data sources, the next step is to import your data and ensure it is clean and ready for analysis. Keep in mind that data often contains errors, duplicates, or missing values. This can compromise the integrity of your analysis, for example, resulting in Adventure Works making poor decisions related to the product launch based on incorrect data.**

**As a part of the data analysis process, you must evaluate the collected data to determine whether cleaning is required. For example, you may identify data inconsistencies or errors. Ensuring the data is clean and prepared for analysis is important for Adventure Works, as your analysis will help shape this launch in many ways, including marketing plans, sales strategies, and even final product changes.**

**Answer the following questions related to the data cleaning process:**

**Name and briefly describe an Excel function that can help you clean the Product Name, Product Category, and Product Size columns.**

**For the Adventure Works sales dataset, indicate the most appropriate Excel data type for the content of each column. For example, Product ID: Number.**

**Explain how you could exclude the missing entry in the Product Description column in the Adventure Works sales dataset at the source using Excel.**