DV Lab

2) Connecting to Data Source – Database and Tableau Joins

Steps to Connect to Excel in Tableau

1. Open Tableau → Click **Connect** → Choose **Microsoft Excel** under "To a File."
2. Browse and select **sample_superstore.xls**.
3. Drag the sheets **Orders**, **Returns**, and **People** to the workspace.

Joins in Tableau

1. Inner Join

- Join **Orders** and **Returns** using **Order ID**.
- Displays only matching records from both sheets.

Steps:

1. Drag **Orders** to the canvas.
2. Drag **Returns** to the right of **Orders**.
3. In the Join configuration, select **Inner Join** and specify the key as **Order ID**.

---

2. Left Join

- Includes all records from **Orders** and matching ones from **Returns**.
- Unmatched rows in **Returns** show **NULL**.

Steps:

1. Change the join type to **Left Join**.

---

3. Right Join

- Includes all records from **Returns** and matching ones from **Orders**.

- Unmatched rows in **Orders** show **NULL**.

Steps:

1. Change the join type to **Right Join**.

---

4. Full Outer Join

- Combines all records from **Orders** and **Returns**, filling unmatched rows with **NULL**.

Steps:

1. Change the join type to **Full Outer Join**.

---

6) Querying Data from CSV - Query Editor, Connecting the data from the Excel Source, Clean, Transform the data.

Here's a simplified version of the steps to get data from different sources in Power BI:

1. **Open Power BI Desktop**: Click on "Blank Report."
2. **Get Data**: Go to the Home tab, click "Get Data," and choose your data source (e.g., Excel, CSV, Oracle).
3. **Select File**: Choose the file with the "Table" icon and open it.
    - Use the Navigator window to either:
        - **Load the table directly**, or
        - **Select Transform Data** to make changes in the Power Query Editor.

Transforming Data from an Excel File:

1. **Check Data Types**: Review the automatically applied data types and change them if needed.
2. **Remove Unnecessary Columns**:
    - Right-click on unwanted columns (e.g., "DeliveryMethod," "WebsiteURL") and select "Remove."

- Use "Choose Columns" from the ribbon and uncheck unnecessary columns (e.g., “PhoneNumber,” “FaxNumber,” “PrimaryContact”).

3. **Split Columns**:
   - Right-click on “ProvinceCity” → "Split Column" → By Delimiter (use custom delimiter ‘(‘).
   - Split again by delimiter ‘)’ to separate the Province and City. Delete the column with only ‘)’.
   - Rename columns: "CityProvince1" to "City" and "CityProvince2" to "Province."
4. **Handle Duplicate Columns**:
   - Delete one of the “CustomerGroup” columns.
   - Split “CustomerName” by delimiter and remove the ‘)’ in the “CustomerName 2” column.
   - Rename columns: "CustomerName1" to "CustomerName" and "CustomerName2" to "Head Office."
5. **Handle Null Values**:
   - Add a new column called “Head Office Status” using conditional logic:
     - “if Head Office equals null then No Head Office else Head Office Exists.”
6. **Rename Table**: Change the table name to “Customer.”

Transforming Data from a Text/CSV File:

1. **Get Data**: Click "Get Data → Text/CSV" and select the `invoice.txt` file.
2. **Transform Data**: Verify PowerBI identifies headers. If needed, change the data type of the “Sales” column to “Fixed decimal number.”
3. **Remove Unnecessary Columns**: Delete the “TotalChillerItems” column (contains only ‘0’ values).
4. **Merge Date Columns**:
   - Select the “Day,” “Month,” and “Year” columns (Ctrl + click).
   - Right-click and choose "Merge."

- Set a custom separator (‘/’) and name the new column “Date.”
- Change the new column’s data type to “Date.”

5. **Create Relationship**: PowerBI automatically creates a relationship between the “Customer ID” in the “Customer” table and “Customer Code” in the “InvoiceData” table, which can be checked in Model View.

7) Creating Reports & Visualizations - Different types of charts, Formatting charts with Title, Colors.

1. **Canvas Background**:
   - Go to “Visualizations → Format your report page.”
   - Click on “Canvas Background,” choose a color, and reduce Transparency to see the color.
2. **Add Title**:
   - Click on “Text Box” from the ribbon.
   - Type “Customer & Invoice,” make it bold, select a font color, and center it on the canvas.
3. **City-wise Total Sales Visual**:
   - Go to “Visualizations → Build Visual.”
   - Choose “Stacked Bar Chart.”
   - Drag “City” to the Y-axis and “Sales” to the X-axis.
   - Use “Focus Mode” for detailed city sales tooltips.
4. **Province-wise Sum of Sales Visual**:
   - Go to “Visualizations → Build Visual.”
   - Choose “Stacked Column Chart.”
   - Drag “Province” to the Y-axis and “Sales” to the X-axis.
5. **City Sales Interaction**:
   - Click on “Texas” in the province chart to highlight only the cities in Texas and view their sales.

6. **KPI Card for Sales**:
    - Ensure no charts are selected.
    - Go to “Visualizations → Build Visual.”
    - Choose “Card” and drag “Sales” onto it to show the KPI.
7. **Year-wise Sales with Slicer**:
    - Go to “Visualizations → Build Visual.”
    - Choose “Slicer” and select “Date → Date Hierarchy,” then choose only “Year.”
    - Change visual style to “Vertical List” if needed.
    - Select a year from the slicer and interact with the province chart to see filtered sales data in the KPI Card.

8) Dashboards - Filters in Power BI, Formatting dashboards.

Here’s a concise guide for setting up Filters and Formatting in Power BI using **HR Data.xlsx**:

Dataset Setup:

1. **Load Data**:
    - Open Power BI → Blank Report → Get Data → Excel Workbook.
    - Select **Table 1** from the Navigator and choose “Transform.”
2. **Add Conditional Column**:
    - Create a new column named **Attrition Count**.
    - Change its datatype to **Whole Number**.
3. **Apply Changes**:
    - Click **Close & Apply**.

Canvas Design:

1. Background Color:

- Go to **Visualizations → Format your report page → Canvas Background**.

- Choose a color and adjust transparency.

2. Title:

- Add a **Text Box**.
- Type **"Filters and Formatting Dashboard"**.
- Style: Bold, Underline, Center-aligned, Font size adjustment, and add a background color (via **Effects → Background Color**).

Visualizations:

3. KPI Card:

- Go to **Visualizations → Build Visual → Card**.
- Drag **Employee Count** into the card.
- Rename title to **"Overall Employees"** (via **Format your visual → General → Title (ON)**).
- Customize background and text color.
- Turn **OFF the Category Label** in the Visual Tab.

4. Pie Chart:

- Select **Pie Chart** from the Visualizations pane.
- Drag **Department** to the **LEGEND** and **Attrition Count** to **VALUES**.
- Rename the title to **"Department-wise Attrition"** (via **Format Visual → General → Title**).
- Customize the background.

5. Stacked Column Chart:

- Select **Stacked Column Chart** from the Visualizations pane.
- Drag **Age Band** to the **X-Axis**, **Employee Count** to the **Y-Axis**, and **Gender** to the **LEGEND**.
- Rename title to **"Age & Gender-wise Employee Count"**.

6. Donut Chart:

- Select **Donut Chart** from the Visualizations pane.
- Drag **Gender** to the **LEGEND** and **Attrition Count** to **VALUES**.
- Rename the title to **"Gender-wise Attrition Count"**.

7. Slicer:

- Select **Slicer** from the Visualizations pane.
- Drag **Department** to the slicer's **Field** section.
- Rename title to **"Department"**.
- Format:
  - Go to **Visual → Style → Tile**.
  - Enable **"Show – Select All"** under **Selection**.
- Add "Filters on this visual" for **Department**.

Now you have a fully designed and formatted dashboard for HR data in Power BI!

9 Analysis of revenue in sales dataset:

i) Create a choropleth map (fill the map) to spot the special trends to show the state which has the highest revenue.

ii) Create a line chart to show the revenue based on the month of the year.

iii) Create a bin of size 10 for the age measure to create a new dimension to show the revenue.

iv) Create a donut chart view to show the percentage of revenue per region by creating zero access in the calculated field.

v) Create a butterfly chart by reversing the bar chart to compare female & male revenue based on product category.

vi) Create a calculated field to show the average revenue per state & display profitable & non-profitable state.

Here's a step-by-step guide to creating the **Revenue Sales Dashboard** using **Revenue Sales Data.xlsx**:

---

1. Load the Dataset

- Open Power BI → **Blank Report** → **Get Data** → Select **Excel Workbook**.
- Select the dataset and load it directly (no transformation needed).

---

2. Filled Map

- Add a **Filled Map** to the canvas.
- **Drag Fields**:
  - Drag **State** to **LOCATION**.
  - Drag **State** to **LEGEND** (to color-code states).
  - Drag **Revenue** to **TOOLTIP** (to view revenue details by hovering).
- **Customize Tooltip**:
  - Select **"Maximum"** for **Revenue** in the TOOLTIP field.
- The map will now highlight states, with the highest revenue visible in the tooltip.

---

3. Line Chart for Revenue by Month

- Add a **Line Chart** to the canvas.
- **Drag Fields**:
  - Drag **Revenue** to **Y-AXIS**.
  - Drag **Month** to **X-AXIS**.
- The chart now shows revenue trends by month.

---

4. Stacked Column Chart for Revenue by Age Group

- Right-click on **Customer Age** in the Data Pane → Choose **New Group**.
- **Create Bin**:
  - Set **Bin Size** to **10** → Click **OK**.
  - A new field named **Customer Age (Bin)** will appear.
- Add a **Stacked Column Chart**.
- **Drag Fields**:

    - Drag **Customer Age (Bin)** to **X-AXIS**.
    - Drag **Revenue** to **Y-AXIS**.
- The chart now shows revenue grouped by age ranges.

---

5. Donut Chart for Revenue Percentage by Region

- Add a **Donut Chart** to the canvas.
- **Drag Fields**:
    - Drag **State** to **LEGEND**.
    - Drag **Revenue** to **VALUES**.
- **Customize Visual**:
    - Go to **Format Visual → Visual → Detail Labels → Options**.
    - Set **Position → Inside** to show the percentage of revenue inside the chart.

---

6. Butterfly Chart for Revenue Comparison by Gender

- Add a **Stacked Column Chart** to the canvas.
- **Drag Fields**:
    - Drag **Product Category** to **X-AXIS**.
    - Drag **Revenue** to **Y-AXIS**.
    - Drag **Customer Gender** to **LEGEND**.
- The chart will display male vs. female revenue comparisons for each product category.

---

7. Calculate Average Revenue per State

Create a Measure:

1. Go to **Home → New Measure**.
2. Add this formula:

```
AveragePerState = AVERAGEX(VALUES(SalesTable[State]), CA
```

```
LCULATE(SUM(SalesTable[Revenue])))
```

3. Add a **Table** visual to the canvas.
4. **Drag Fields**:
    - Drag **State** to **COLUMNS**.
    - Drag **AveragePerState** to **COLUMNS**.

Create a New Column for Profit Status:

1. Go to **Modeling → New Column**.
2. Add this formula:

```
ProfitStatus = IF(SalesTable[AveragePerState] > 1000, "P
rofitable", "Not-Profitable")
```

3. Drag **ProfitStatus** to the **COLUMNS**.

---

8. Title and Background

- Add a **Text Box**:
    - Type: **"Revenue Sales Dashboard"**.
    - Style: Bold, Center-aligned, and apply your desired font size and color.
- Set the **Canvas Background**:
    - Go to **Format your report page → Canvas Background**.
    - Choose a color and adjust transparency.

Final Output

Your dashboard should include:

- **Filled Map** highlighting states with maximum revenue.
- **Line Chart** for monthly revenue trends.
- **Stacked Column Chart** for age-grouped revenue.
- **Donut Chart** showing revenue percentages by region.
- **Butterfly Chart** comparing male and female revenue by product category.

- **Table** for average revenue and profitability status by state.

Let me know if you need any further assistance!

11) Analysis of HR Dataset:

i)Create KPI to show employee count, attrition count, attrition rate, attrition count, active employees, and average age.

ii) Create a Lollipop Chart to show the attrition rate based on gender category.

iii) Create a pie chart to show the attrition percentage based on Department Category- Drag department into colours and change automatic to pie. Entire view, Drag attrition count to angle. Label attrition count, change to percent, add total also, edit label.

iv) Create a bar chart to display the number of employees by Age group,

v) Create a highlight table to show the Job Satisfaction Rating for each job role based on employee count.

vi) Create a horizontal bar chart to show the attrition count for each Education field Education field wise attrition – drag education field to rows, sum attrition count to col,

vii) Create multiple donut chart to show the Attrition Rate by Gender for different Age group.

Step 1: Load and Rename Dataset

1. Open Power BI → **Blank Report** → **Get Data** → Select **CSV** and load the dataset.
2. Rename the table to **HR** by double-clicking the table name in the Fields pane.

Step 2: Create KPIs

Set the **Canvas Background** and title:

- Go to **Format your report page → Canvas Background**, and set a color of your choice.

- Add a **Text Box** and type "HR Dashboard" as the title. Style it as desired (bold, center-aligned, colored, etc.).

a) KPI for Employee Count

1. Create a New Measure:

```
Employee Count = COUNT(HR[EmployeeNumber])
```

2. Add a **Card Visual** to the canvas.
3. Drag and drop the **Employee Count** measure into the **Fields**.

b) KPI for Attrition Count

1. Create a New Measure:

```
Attrition Count = COUNTROWS(FILTER('HR', HR[Attrition] =
"Yes"))
```

2. Add another **Card Visual**.
3. Drag and drop the **Attrition Count** measure into the **Fields**.

c) KPI for Attrition Rate

1. Create a New Measure:

```
Attrition Rate = DIVIDE([Attrition Count], [Employee Cou
nt], 0) * 100
```

2. Add another **Card Visual**.
3. Drag and drop the **Attrition Rate** measure into the **Fields**.

d) KPI for Active Employees

1. Create a New Measure:

```
Active Employees = [Employee Count] - [Attrition Count]
```

2. Add another **Card Visual**.
3. Drag and drop the **Active Employees** measure into the **Fields**.

e) KPI for Average Age

1. Create a New Measure:

```
Average Age = AVERAGE(HR[Age])
```

2. Add another **Card Visual**.
3. Drag and drop the **Average Age** measure into the **Fields**.

Step 3: Lollipop Chart for Attrition Rate by Gender

1. Add a **Line & Stacked Column Chart** to the canvas.
2. Drag Fields:
   - Drag **Gender** to **X-AXIS**.
   - Drag **Attrition Count** to **COLUMN Y-AXIS**.
   - Drag **Attrition Rate** to **LINE Y-AXIS**.
3. Customize the chart as needed.

Step 4: Pie Chart for Attrition Percentage by Department

1. Add a **Pie Chart** to the canvas.
2. Drag Fields:
   - Drag **Department** to **LEGEND**.
   - Drag **Attrition Count** to **VALUES**.
3. Customize the chart:
   - Go to **Format Visual → Detail Labels → Position → Inside**.
   - Change the labels to show **percentages** and **total**.

Step 5: Bar Chart for Employee Count by Age Group

1. Create a Bin:
   - Right-click on **Age** → Select **New Group**.
   - Rename the group as **Age Group** and set **Bin Size** to **10** → Click **OK**.
2. Add a **Stacked Bar Chart** to the canvas.

3. Drag Fields:
    - Drag **Age Group** to **Y-AXIS**.
    - Drag **Employee Count** to **X-AXIS**.

Step 6: Highlight Table for Job Satisfaction by Job Role

1. Add a **Matrix** visual to the canvas.
2. Drag Fields:
    - Drag **Job Role** to **ROWS**.
    - Drag **Job Satisfaction** to **COLUMNS**.
    - Drag **Employee Count** to **VALUES**.
3. Customize the matrix:
    - Apply conditional formatting to highlight cells based on **Employee Count**.

Step 7: Horizontal Bar Chart for Attrition by Education Field

1. Add a **Stacked Bar Chart** to the canvas.
2. Drag Fields:
    - Drag **Education Field** to **Y-AXIS**.
    - Drag **Attrition Count** to **X-AXIS**.

Step 8: Multiple Donut Charts for Attrition Rate by Gender and Age Group

Donut Chart 1:

1. Add a **Donut Chart** to the canvas.
2. Drag Fields:
    - Drag **Gender** to **LEGEND**.
    - Drag **Attrition Rate** to **VALUES**.
3. Apply a filter:
    - Drag **Age Group** to **Filters on this visual** and select **Under 30**.

Donut Chart 2:

1. Duplicate the first Donut Chart.
2. Modify the filter:
    - Change the **Age Group** filter to **Under 50**.

Final Output

Your HR Dashboard should include:

1. KPIs for **Employee Count**, **Attrition Count**, **Attrition Rate**, **Active Employees**, and **Average Age**.
2. A **Lollipop Chart** for Attrition Rate by Gender.
3. A **Pie Chart** for Attrition Percentage by Department.
4. A **Bar Chart** for Employee Count by Age Group.
5. A **Highlight Table** for Job Satisfaction by Job Role.
6. A **Horizontal Bar Chart** for Attrition Count by Education Field.
7. Two **Donut Charts** showing Attrition Rate by Gender for different Age Groups.

Let me know if you need further clarification!

12 Analysis of Amazon Prime Dataset:

i) Create a Donut chart to show the percentage of movie and tv shows

ii) Create a area chart to shows by release year and type

iii) Create a horizontal bar chart to show Top 10 genre

iv) Create a map to display total shows by country

v) Create a text sheet to show the description of any movie/movies.

vi) Build an interactive Dashboard.

Here's a concise version of your instructions for building the Power BI dashboard:

Data Transformation

1. Load **"Amazon-Prime-Titles.csv"** into Power BI.
2. Rename "listed in" → **"Genre"**.
3. Remove columns: **Duration, Rating, date_added, cast, director**.
4. Rename the table to **"Amazon"**.
5. Close & Apply.

Visualizations

1. **Donut Chart (Movies vs. TV Shows)**
   - Drag **Type** to **Legend**.
   - Drag **Show ID** to **Values** (set to **Count**).
   - Filter Type for "Movie" and "TV Show" only.
2. **Area Chart (Release Year & Type)**
   - Drag **Release Year** to **X-Axis**, **Type** to **Y-Axis**, and **Legend**.
3. **Horizontal Bar Chart (Top 10 Genres)**
   - Drag **Genre** to **Y-Axis** and **Title** to **X-Axis**.
   - Apply a **Top N filter** (set to 10).
4. **Map (Total Shows by Country)**
   - Drag **Country** to **Location** and **Show ID** to **Tooltip** (set to **Count**).
5. **Text Table (Descriptions)**
   - Drag **Title** and **Description** into a **Table** visual.

Dashboard

1. Add visuals to the canvas with a title: **"Amazon Prime Titles Dashboard"**.
2. Add slicers for **Release Year** and **Type** for interactivity.
3. Style with a background and theme for a professional look.

This provides the same output in a simplified way. Let me know if more changes are needed!