Reminders…

- Requirements:
  - Answer all the questions in this document
  - When complete, download a PDF copy of this document for future reference. You can also save it in your Google Drive.
  - Don't know how to download it as a PDF? You can find more information about downloading this by clicking here.

Content

Scrub

You are being provided with data about sales at a store. You can access this data by clicking here [ADD LINK] and clicking on Use Template in the upper right corner.

This data has issues that need to be fixed before it can be used for your data analysis project. some issues may include:

- Duplicate records
- Missing values
- Obviously wrong values

Using Google Sheets, review the data, find the issues, and clean the dataset for your use in this project.

What are some errors you identified in this dataset? Identify the column or row where possible. For example: “Duplicate transaction data on rows 45 and 46.” You do not need to include every error, but include at least three.

1. Several columns contained blanked rows(e.g., `Customer_ID`, `Product_Name`, `Size`, and `Product_Category`,`Product_Line`).
2. The “Duplicate check” column indicated repeated order numbers, but these represented multiple items from the same order.
3. Some product categories and sizes were missing or inconsistent.

How did you fix the errors that you identified in the previous question? For example: “Deleted the duplicate transaction data on row 46.”

Replaced blanked rows with “Unknown” placeholders to avoid data loss while still acknowledging missing information.

Verified that repeated order numbers were not true duplicates and kept them.

Checked for invalid or blank values and confirmed all numeric and date fields were properly formatted.

No rows were deleted; data integrity was maintained.

Explore

Using the spreadsheet tools in Google Sheets, explore the data. You are encouraged to use spreadsheet functions like AVERAGE and CORREL as well as SQL queries like ORDER BY and LIMIT.

When you have used these tools, create a chart that highlights a relationship you discovered in the data. For instance, you might create a bar chart that shows sales of a particular item in different months of the year to showcase how well it sells in warmer months.

What spreadsheet functions did you use and what results did you get? For example: “The AVERAGE of the “sales” column was $35.55.” You do not need to include every function you used, but include at least three.

Used `SUM(Line_Total)` to calculate **Total Sales = $16,406.14**.

Used `AVERAGE(Price)` and `MEDIAN(Price)` to find **Average Price = $26.03** and **Median Price = $25.48**.

Used `MIN(Price)` and `MAX(Price)` to find **Price Range = $10.80 – $39.55**.

Used `CORREL(Price, Quantity)` to find the correlation between price and quantity (**–0.0095**, showing no relationship).

Used `SUMIFS(Line_Total, Product_Line, "cat")` and `SUMIFS(Line_Total, Product_Line, "dog")` to calculate **Cat Sales = $9,030.16** and **Dog Sales = $7,375.98**.

What SQL queries did you use and why? For example: “Used the AND clause to see sales that were for a certain item and above a certain price” You do not need to include every query you used, but include at least three.

Used a **SELECT** query to review sales by product line:

```
SELECT Product_Line, SUM(Line_Total) FROM data GROUP BY Product_Line
```

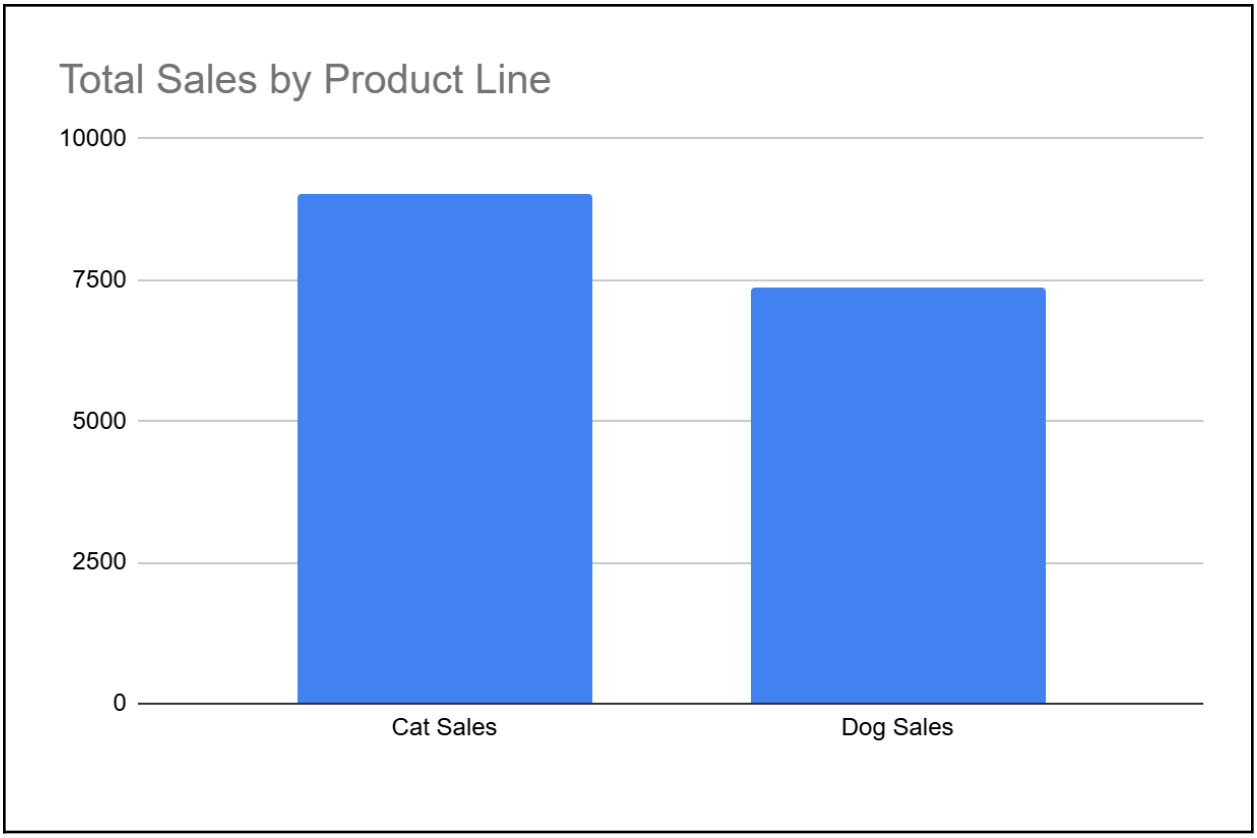
Used a **WHERE** clause to check orders above a certain amount:

```
SELECT * FROM data WHERE Price > 30
```

Used an **ORDER BY** clause to sort items by total sales:

```
SELECT Product_Name, SUM(Line_Total) FROM data GROUP BY Product_Name
ORDER BY SUM(Line_Total) DESC
```

Copy and paste at least one chart into this document that was created from the dataset.



Total Sales by Product Line — Cat products slightly outperform dog products in overall revenue.

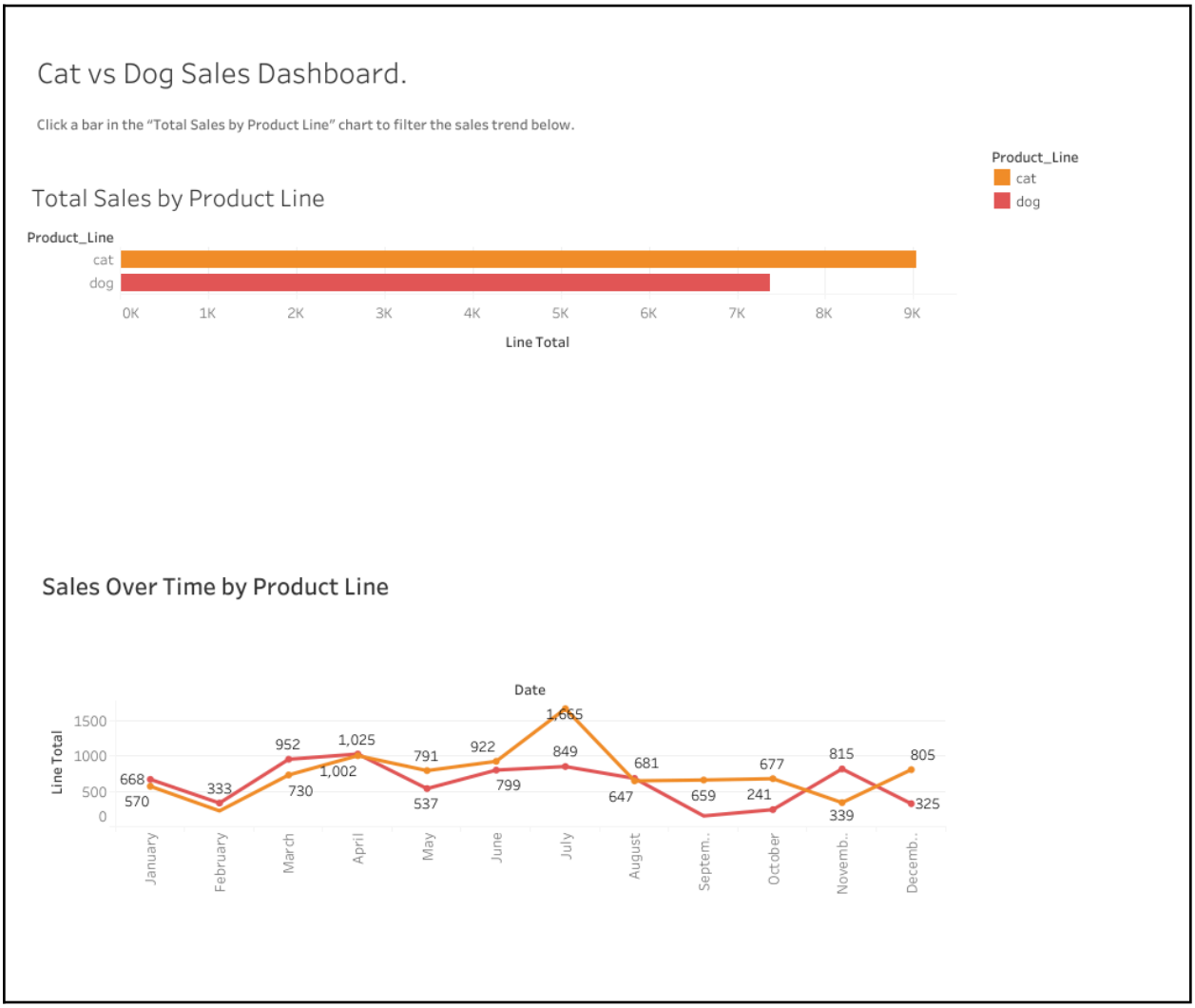
Visualize

Using the dataset you have scrubbed and explored, create a dashboard with at least two charts and at least one interaction.

Copy and paste the URL for your published Tableau Public dashboard

https://public.tableau.com/app/profile/andrejs.verhovods/viz/AndrejsV_CatsvsDogsSalesDash/CatvsDogSalesDashboard_

Copy and paste an image of the dashboard downloaded from Tableau Public



This dashboard shows total and monthly sales for cat and dog product lines. Users can click on a product line to filter the monthly sales trend. The analysis reveals that cat products have slightly higher total sales overall, with peaks during the middle months of the year.