SEED

خوشحال خیبر پختونخوا

atomcamp

Assignment No.3

Course Instructor

Sir Yahya Bajwa

Teaching Assistant

Miss Nimra

Batch-02 DS

Submitted By: Muhammad Hamza

Data Science & AI Boot Camp

"Case Study of Pakistan's Startup"

Topic:

“A Data-Driven Exploration of Pakistan's Startup”

Objective:

The objective of this case study is to analyze and derive insights from a dataset containing information about startups based in Pakistan and their monthly sales. The goal is to gain a deeper understanding of the startup ecosystem, identify trends, and draw meaningful conclusions that can inform business strategies.

Problem Statement:

The startup dataset for Pakistan has messy data that needs cleaning, making it tough to get useful insights. Figuring out trends, startup locations, and performance is tricky due to errors. Fixing these issues is important for making smart decisions in Pakistan's startup scene.

Problem Analysis:

In our data-driven exploration, we will:

- Clean and Standardize Data:
- Gain Startup Overview
- Explore Founding Dates and Trends
- Conduct Geographical Analysis
- Perform Sales Analysis
- Analyze Sales Trends

Datasets:

- **Provided Dataset:**
  Pakistan-startup-census - Google Sheets
- **Cleaned Dataset:**
  Pakistan-Startup-Census-Done

Methodology:

The dataset contains information about startups, including their names, locations, taglines, categories, and founding dates. In our exploration of Pakistan's startup data, we followed a simple yet effective approach. We used If statements and logical thinking to clean the data, ensuring accuracy. Additionally, tools like pivot tables, VLOOKUP, and filters were employed to organize and validate our findings. Whether it was identifying the total startups, analyzing sales trends, or determining the oldest startup, we applied these methods to verify our results. This combination of logic and Excel functionalities strengthened the reliability of our conclusions and provided a clear understanding of the startup landscape.

Assignment Tasks:

1. **Data Cleaning:**
   We have extracted city names and standardize the 'Year' column. Merged Monthly Sales Sheets using VLOOKUP and remove errors, create a 'City' column referencing the first sheet & removed the respective duplicates and kept the unique one.

2. **Startup Overview:**
   How many Startups are in the data?
   Answer:
   There was a total of ***"399"*** startups in the dataset.

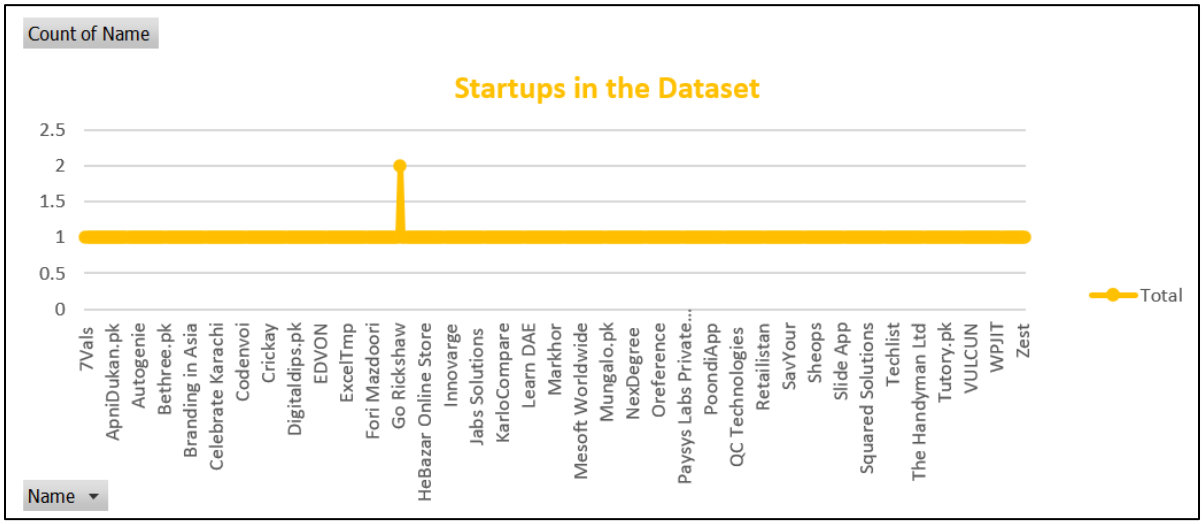

Figure 2.1

Which city has the most startups?
Answer:
"Karachi" is the one that has the most startups.

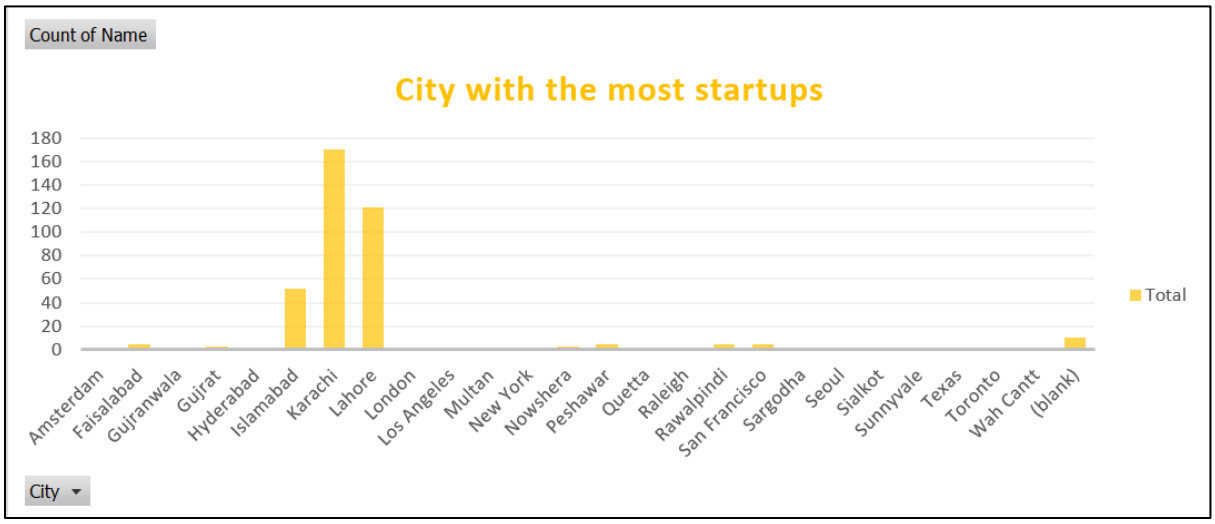

Figure 2.2

3. Founding Dates:

Which is the oldest startup(s) in the dataset? (if multiple name all)

Answer:

"Orient Electronics (1957)" is the oldest startup in the dataset. The table verifying the insight is given as following.

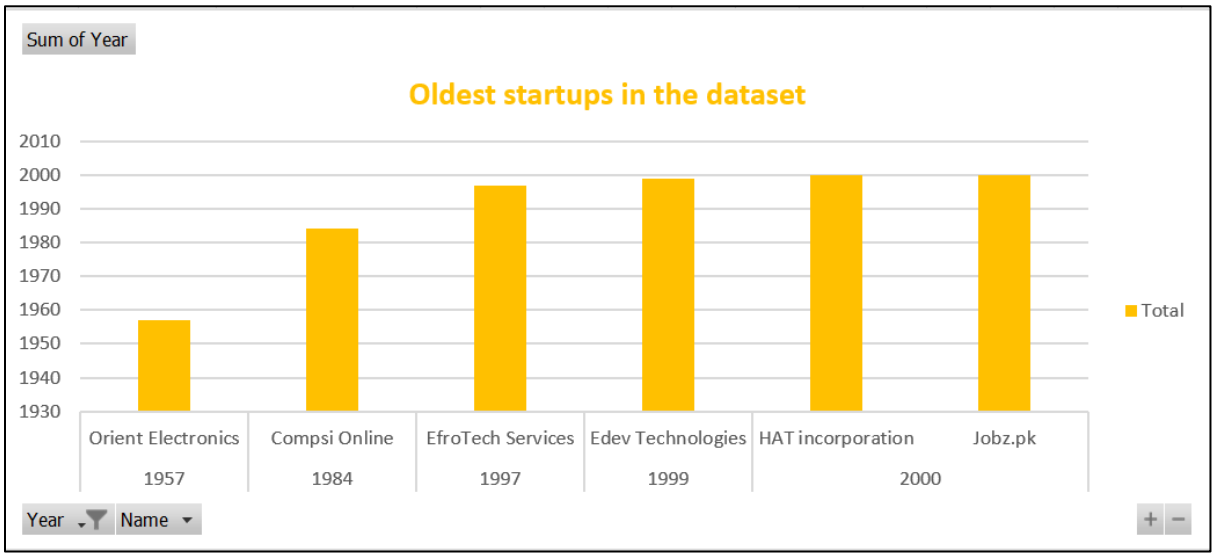

Figure 3.1

4. Founding Trends:

In which year most startups were founded?

Answer:

"2016" is the year in which most of the startups were founded in the dataset

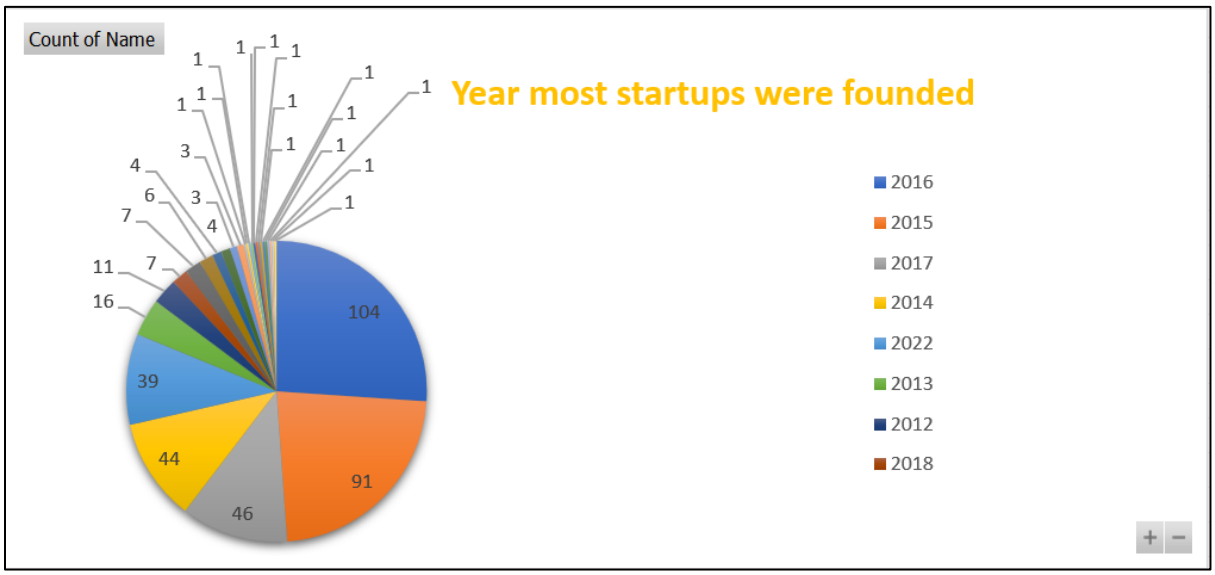

Figure 4.1

5. **Geographical Analysis:**

From the previous question, identify the city with the highest number of startups founded in that specific year.

Answer:

"Karachi" is the city with the highest number of startups founded in that particular year. The table verifying the insight is given as following

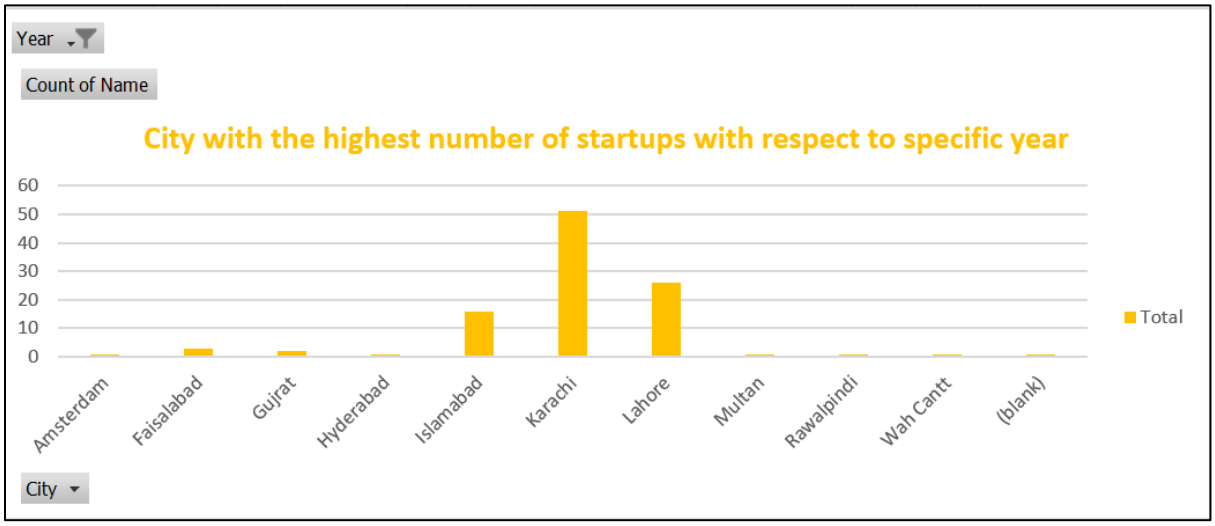

Figure 5.1

6. **Sales Analysis:**

Identify the month with the highest total sales across all startups.

Answer:

Startups made the most sales in "***October"*** compared to any other month.

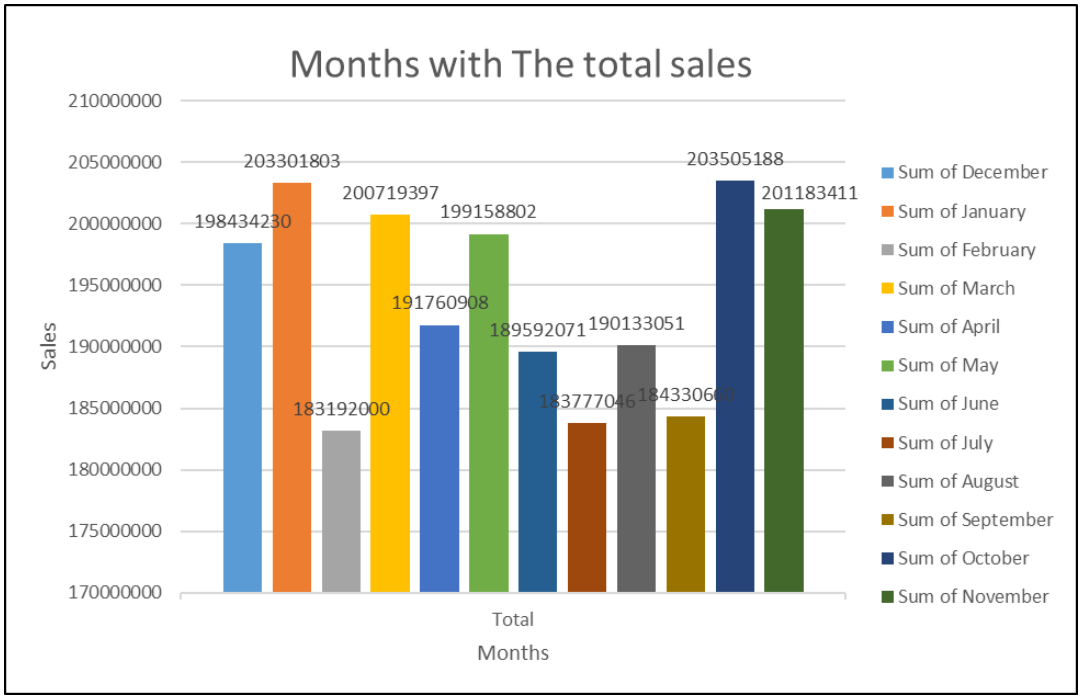

Figure 6.1

Which month had the highest average sales?

Answer:

On average, startups made the most sales in ***"October".*** The table verifying the insight is given as following

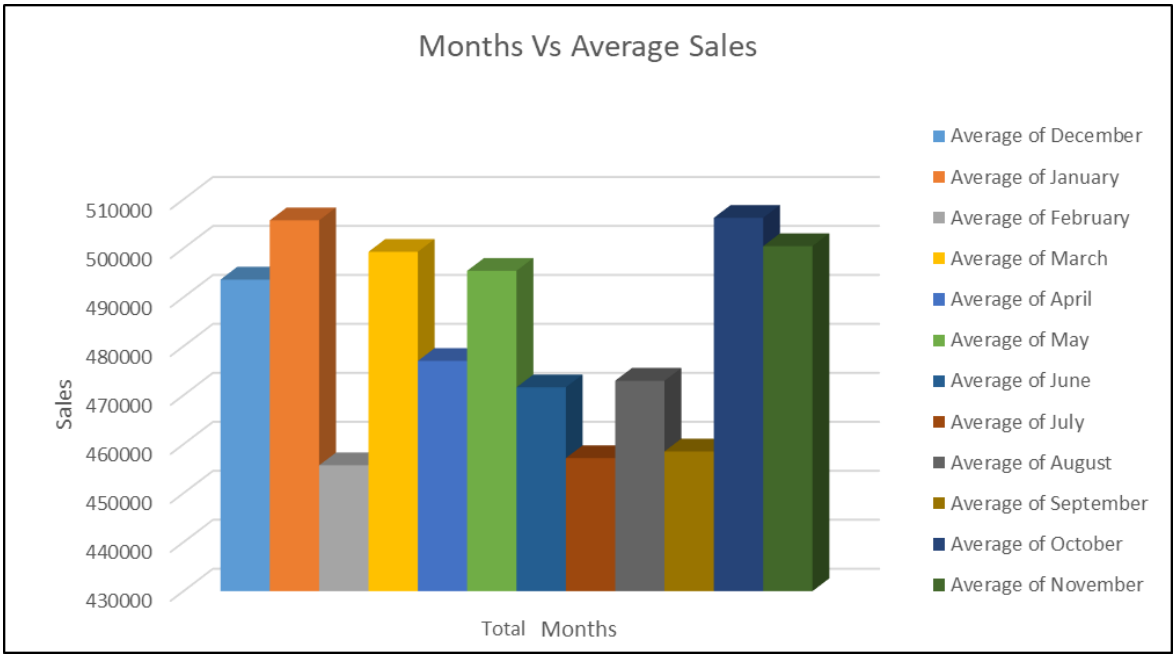

Figure 6.2

7. City-wise Sales:

Identify the city with the highest average sales in July.

Answer:

"Seoul" demonstrated the highest average sales performance among all cities in the July dataset.

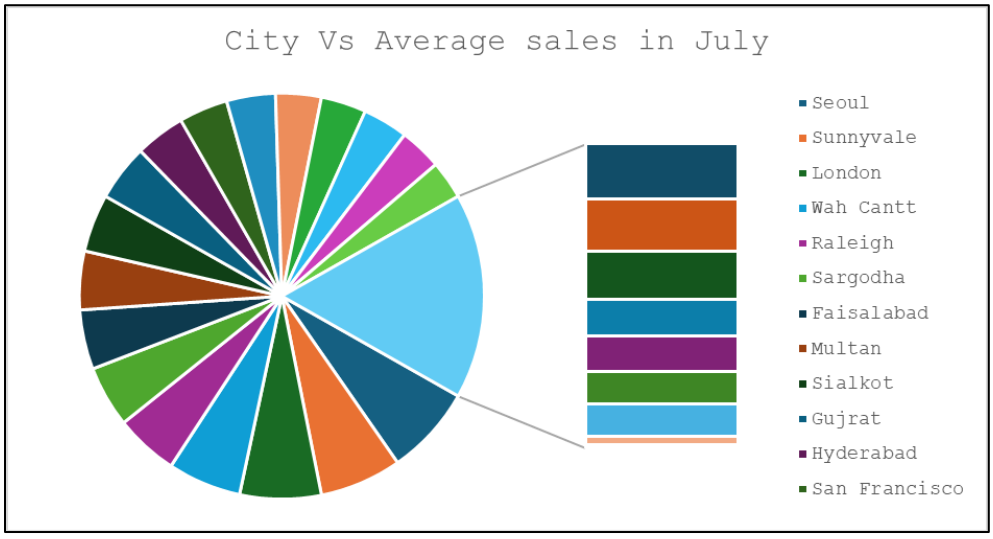

Figure 7.1

Determine the city with the highest total sales for the complete year.

Answer:

Comprehensive analysis of annual sales data designates "***Karachi***" as the city with the highest total sales. The table verifying the insight is given as following

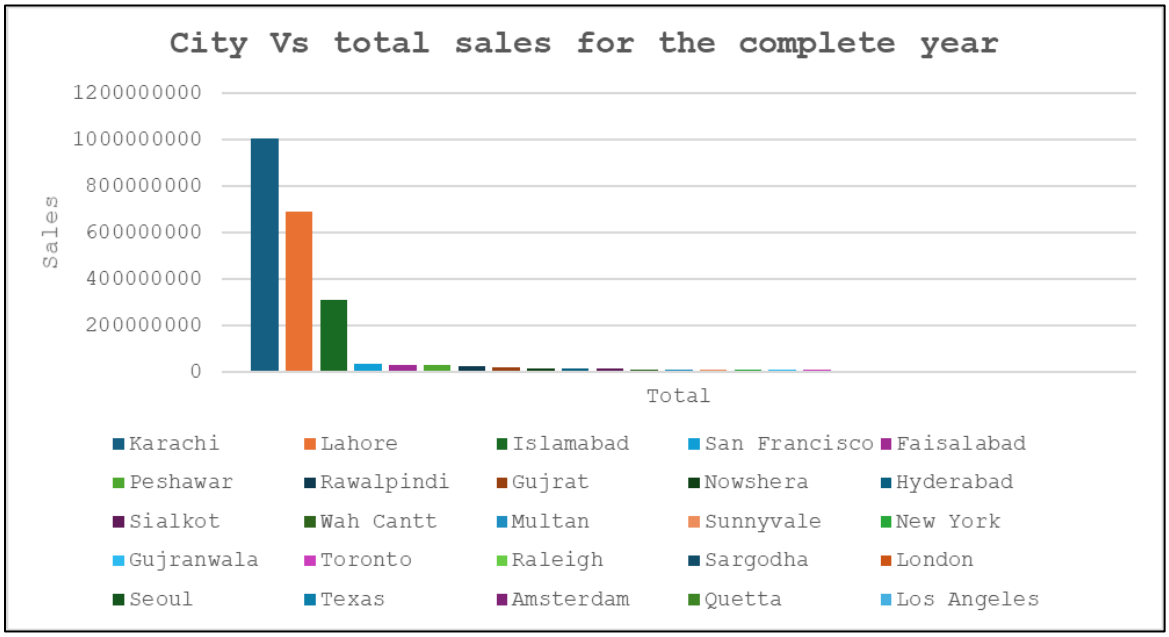

Figure 7.2

8. **Startup Performance:**

Identify the startup with the highest total sales for the complete year.

Answer:

Analysis of annual sales data reveals "***Lecture Hunt***" as the clear leader in total sales, surpassing all competitors. The Chart verifying the insight is given as following

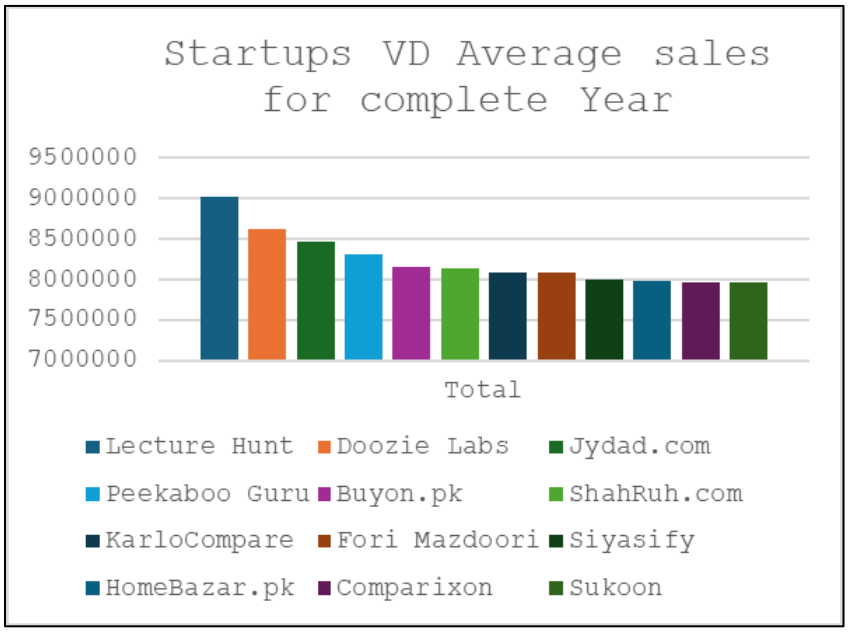

Figure 8.1

Determine the startup with the highest average sales per month for the complete year.
Answer:

"Lecture Hunt" is the one having the highest average sales per month for the complete year. The Chart verifying the insight is given as following:

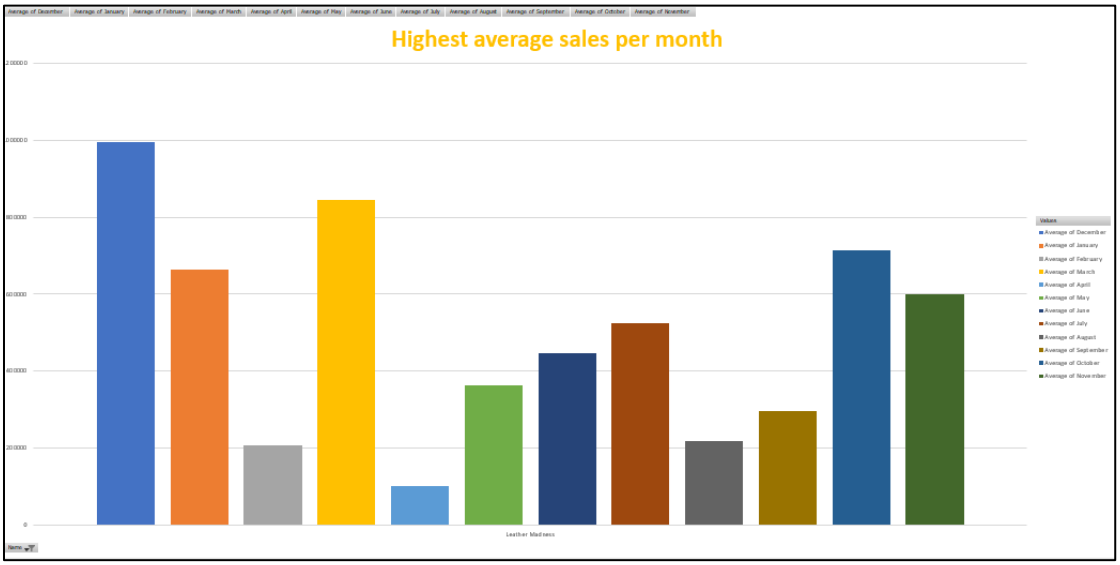

Figure 8.2

9. Lowest Sales:

Identify the city with the minimum total sales for August.
Answer:

"Toronto" is the city having the minimum total sales for august. The Chart verifying the insight is given as following

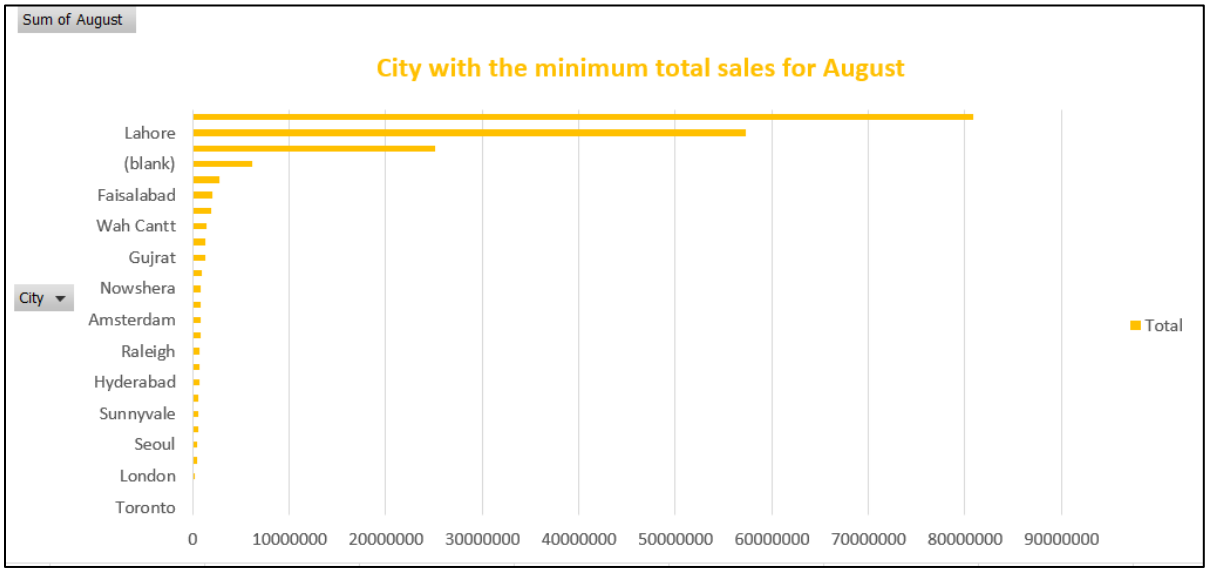

Figure 9.1

10. Sales Trends:

Determine the month with the lowest average sales.

Answer:

"February" is the month with the lowest average sales. The Chart verifying the insight is given as following:

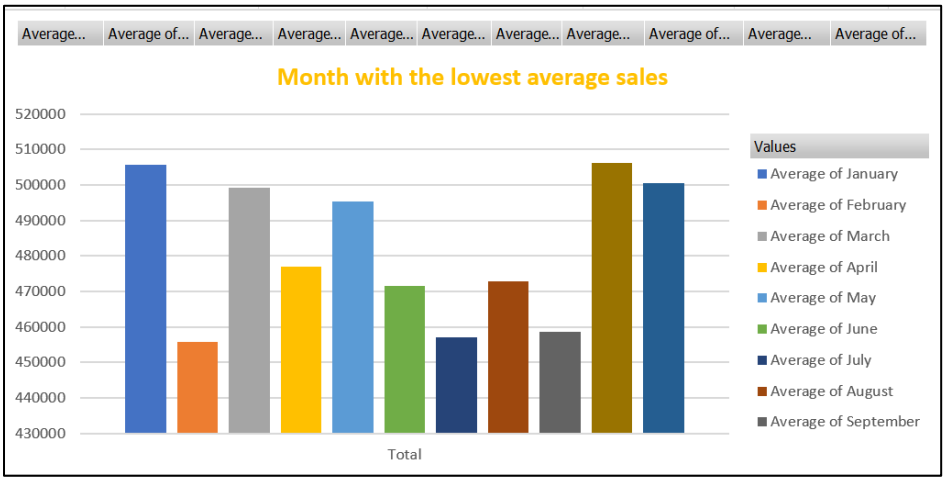

Figure 10.1

Identify startups with the highest monthly sales for two or more months.

Answer:

"Merayi" is the startup with the highest monthly sales for two or more months.

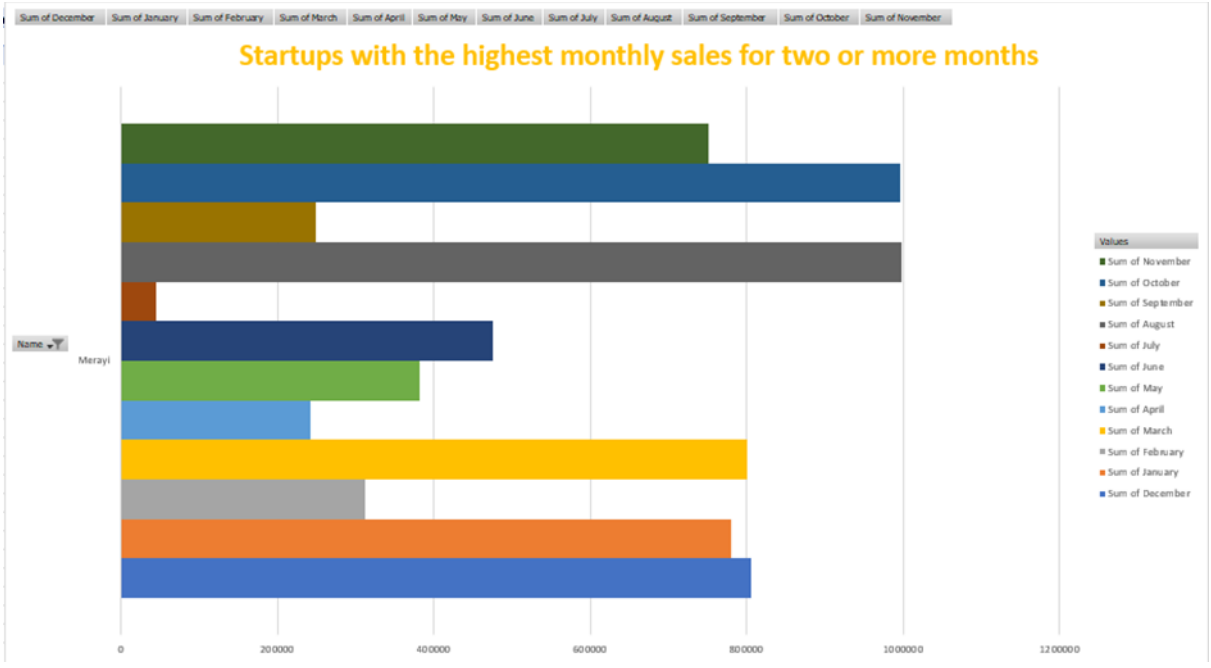

Figure 10.2

Conclusion:

In summary, our data-driven exploration of Pakistan's startup landscape uncovered crucial insights despite initial data challenges. Karachi emerged as a prominent hub, hosting the highest number of startups and demonstrating substantial sales. The year 2016 marked a significant surge in startup establishments. "Lecture Hunt" excelled with the highest total sales, while "Leather Hunt" maintained the highest average sales per month. Addressing data issues proved pivotal for informed decision-making in Pakistan's dynamic startup scene. This analysis provides actionable insights for strategic planning, emphasizing the importance of data quality in shaping effective business strategies.
