Startup Funding Analysis Report

Prepared By: Ayush Katkar
Email: ayushkatkar3002@gmail.com
Tools Used: Python (Jupyter Notebook), Power BI
Dataset: Indian Startup Funding Dataset (2018–2021)

1. Introduction

The Indian startup ecosystem has grown rapidly over the years, attracting major global investments. This project focuses on analyzing the Startup Funding dataset using Python to uncover insights such as funding trends, top investment sectors, leading cities, and investor contributions. The objective is to transform raw data into meaningful insights through exploratory data analysis and visualization.

---

2. Dataset Description

Dataset Source: Startup Funding Data (India)
Dataset Format: CSV
Total Records: Includes startups from multiple sectors receiving funding from different investors across several years

Main Columns Used:

- Startup Name
- Industry Vertical
- Sub-Vertical
- City/Location
- Investors
- Investment Type

- Amount in Crore
- Date

---

3. Tools and Technologies Used

- Python Programming
- Pandas
- NumPy
- Matplotlib
- Seaborn

---

4. Data Preprocessing

To ensure clean and accurate data analysis, the following preprocessing steps were performed:

- Handling missing values in key columns such as Amount and City
- Converting funding Amount from INR format into numeric format
- Removing duplicate records
- Standardizing city names (example: Bangalore and Bengaluru combined)
- Extracting year from the Date column for time-series analysis

5. Python Code Used for Data Cleaning and EDA

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('startup_funding.csv')

df['Amount'] = df['Amount'].replace(',', '', regex=True)
df['Amount'] = df['Amount'].astype(float)

df['City'] = df['City'].replace({
    'Bangalore': 'Bengaluru',
    'Delhi': 'New Delhi',
    'Gurgaon': 'Gurugram',
    'Hyderabad': 'Hyderabad'
})

df['Date'] = pd.to_datetime(df['Date'], errors='coerce')
df['Year'] = df['Date'].dt.year

df = df.dropna(subset=['Amount', 'City'])
```

6. Data Analysis and Visualization

```
sector_funding = df.groupby('Industry
Vertical')['Amount'].sum().sort_values(ascending=False).head(10)

plt.figure(figsize=(12,6))
sector_funding.plot(kind='bar')
plt.title("Top 10 Funded Sectors in India")
plt.xlabel("Sector")
plt.ylabel("Total Funding (in Crore)")
plt.show()
```

Insight: Bengaluru and Mumbai emerged as the top startup hubs in India.

```
year_funding = df.groupby('Year')['Amount'].sum()

plt.figure(figsize=(10,5))

plt.plot(year_funding.index, year_funding.values, marker='o')
```

```
plt.title("Year-wise Funding Trend")

plt.xlabel("Year")

plt.ylabel("Total Funding (in Crore)")

plt.grid()

plt.show()
```

Insight: Funding shows strong growth with visible major peaks in succeeding years.

```
top_investors =
df.groupby('Investors')['Amount'].sum().sort_values(ascending=False).head(10)

plt.figure(figsize=(10,5))

top_investors.plot(kind='bar')

plt.title("Top Investors in Indian Startups")

plt.xlabel("Investors")

plt.ylabel("Total Funding (in Crore)")

plt.show()
```

Insight: Major global and domestic investors contributed significantly to startup growth.

7. Insights and Findings

- Indian startups experienced a huge growth in investment over recent years.
- Bengaluru, Mumbai, and New Delhi are the leading startup hubs.
- FinTech, E-commerce, and Technology are the most funded sectors.
- A few strong investors dominate a major share of the total funding.
- Investment trends indicate global market confidence in Indian startups.

---

8. Conclusion

The analysis highlights that India has become one of the strongest startup ecosystems in the world, supported by strong technological innovation and large investment inflow. Continuous

improvement in funding and sector diversification is helping in economic development and job creation. This project provided a data-driven understanding of funding distribution and market preferences in the Indian startup landscape.
