Python Data analysis report of SOCIAL\_MEDIA\_DATASET

🎯 Objective:

As a data analyst at Entri Marketing Agency, the goal was to analyze a dataset of 5,000 simulated social media posts to help optimize engagement using data cleaning, visualization, and object-oriented programming techniques.

---

✅ Part A: Data Ingestion & Exploration

Imported the dataset using Pandas after mounting Google Drive in Colab.

Displayed the first 5 and last 3 records to understand data structure.

Used .info() to inspect data types and memory usage.

Calculated percentage of missing values in each column to assess data quality.

---

✅ Part B: Data Wrangling

Removed low-engagement posts (Likes < 50) using boolean indexing.

Filled missing Hashtag entries with 'NoHashtag'.

Converted Date\_Posted to proper datetime format for analysis.

---

✅ Part C: Object-Oriented Programming

Created a SocialPost class implementing encapsulation for likes and comments.

Defined a method get\_engagement() to compute total engagement.

Created a subclass Author with an additional author\_rating property.

Implemented polymorphism by overriding the display() method in the subclass.

Demonstrated usage by creating objects and calling methods dynamically.

---

✅ Part D: Data Visualization

Line Chart: Showed daily post frequency using matplotlib.

Stacked Bar Chart: Compared average likes, shares, and comments by category.

Heatmap: Used seaborn to visualize correlations among numeric variables (e.g., Likes, Shares, Comments).

---

✅ Part E: Advanced Analysis

Used .iloc[] to access the 3rd row’s Likes.

Used .loc[] to filter Fashion category posts with more than 1000 Likes.

---

📊 Tools & Libraries Used

Pandas for data handling

Matplotlib & Seaborn for visualization

Python OOP for structuring data post logic

---

🔚 Conclusion:

This assignment demonstrated a full-cycle data analysis process — from loading and cleaning data to visualizing insights and modeling data using object-oriented programming. It helped improve understanding of Python for real-world data applications in social media analytics