**Exploratory Data Analysis (EDA) Report — Personality Dataset**

**🧾 Dataset Overview**

* The dataset has **2,900 rows** and **8 columns**.
* It focuses on finding whether a person is **Introvert** or **Extrovert** based on social behavior.

**📊 Types of Data**

* **Numerical columns** (like time alone, events attended, friends, etc.): 5 columns
* **Categorical columns** (like stage fear, drained feeling, personality): 3 columns
* Personality is either **Extrovert** or **Introvert**.

**📉 Summary of Numbers**

* People show different behaviors. For example:
  + Some spend more time alone.
  + Some go out or post more often.
  + Friend circle sizes vary from 0 to 15.

**⚠️ Missing Data**

* Almost all columns have missing values.
* For example:
  + Post\_frequency missing in 65 rows
  + Stage\_fear missing in 73 rows
* These missing values should be **handled properly**.

**🧹 Data Cleaning and Changes**

* Some missing values were filled (e.g., using the mean).
* Some rows with missing data were **removed**.
* A few columns were **renamed or deleted** for clarity.
* **New columns were added**, like:
  + Social\_Score: default 0
  + Social\_friend\_circle: categorizes as High/Medium/Low friend circle
  + A column showing if friend circle increased from the last row

**🔎 Analysis and Exploration**

* Data was sorted and filtered in different ways.
* Duplicate rows were found and removed.
* Some values were checked using simple conditions and logic.
* Small datasets were merged and saved as files (CSV, JSON, SQLite).

**📈 Interesting Findings**

**Personality Patterns:**

* **Extroverts**:
  + Post more often
  + Attend more social events
  + Have more friends
  + Mostly **no stage fear**
* **Introverts**:
  + Spend more time alone
  + Have fewer friends
  + Mostly **have stage fear**

**Pivot Table Results:**

* Most **Introverts** (1,299) have stage fear.
* Most **Extroverts** (1,338) do **not** have stage fear.

**Conclusion**

* The dataset gives useful information to understand **Introverts vs Extroverts**.
* **Stage fear** and **friend circle size** are big clues.
* Data needs some cleaning (fix missing and wrong values).
* You can build **prediction models** using this data.
* This EDA is a great first step before deep analysis or machine learning.