# Application Flow Diagram Explanation and Structure

## Flow Diagram Components

### 1. User Interaction Flow

- \*\*Landing Page:\*\*

- Users visit the landing page to sign up or log in.

- After authentication, they are redirected to their dashboard.

- \*\*Dashboard:\*\*

- Users can view previously generated synthetic datasets.

- They can upload new datasets to generate synthetic datasets or view existing scorecards.

- \*\*Scorecard Viewer:\*\*

- Displays interactive visualizations (e.g., bar charts, scatter plots) comparing original and synthetic datasets.

---

### 2. Backend and Database Flow

- \*\*Dataset Upload:\*\*

- Users upload a dataset via the dashboard.

- The dataset is temporarily stored for processing.

- Validation ensures the dataset format is correct.

- \*\*Synthetic Data Generation:\*\*

- The backend processes the uploaded dataset using SDV or similar tools.

- Synthetic datasets are generated based on user configurations (e.g., specific columns or row count).

- The synthetic dataset is saved for future access.

- \*\*Statistical Comparison:\*\*

- Original datasets are benchmarked with descriptive statistics (mean, median, variance, etc.).

- A scorecard is created comparing the original dataset and the synthetic dataset.

- Original datasets are deleted after processing, but scorecards are saved.

- \*\*User Storage:\*\*

- Synthetic datasets and scorecards are stored persistently in the database, linked to the user’s account.

---

### Flow Summary

1. \*\*User Uploads a Dataset\*\*

- Input: Dataset in CSV, XLSX, or JSON format.

- Output: Synthetic dataset saved; original dataset deleted; scorecard generated and saved.

2. \*\*Synthetic Dataset Generation\*\*

- Backend processes the dataset using SDV.

- Synthetic dataset is saved in the database for user access.

3. \*\*Scorecard Generation\*\*

- Backend computes statistical benchmarks (e.g., mean, variance, correlation).

- Scorecard compares the original dataset to the synthetic dataset.

4. \*\*User Dashboard\*\*

- Users can:

- Download previously generated synthetic datasets.

- View interactive scorecards.

- Upload new datasets.

---

## Diagram Legend

1. \*\*Frontend (Green):\*\*

- Landing page → Dashboard → Scorecard Viewer.

2. \*\*Backend (Blue):\*\*

- API Endpoints: `/upload`, `/generate`, `/scorecard`.

3. \*\*Database (Orange):\*\*

- Tables: Users, Synthetic Datasets, Scorecards.

---

## Next Steps

1. Use this structure to create a visual flow diagram using tools like Lucidchart, Draw.io, or a similar tool.

2. If you need further clarification or adjustments, let me know!