**Datasets before and after cleansing**

**Here are the key steps we can follow for the report:**

**1. Data Overview and Differences**

* Summarize the structure of both datasets (e.g., number of rows, number of columns, types of data).
* Identify any major changes between the datasets, such as:
  + **Null or missing values**: Any rows with missing values in fields like Rating, Reviews, Size, etc.
  + **Inconsistent formats**: For example, dates not following a uniform format or inconsistent numeric formats (like in the Installs column).
  + **Duplicates**: Whether there are duplicate entries and how they have been addressed.

**2. Pre-Cleansing Issues**

* Highlight common issues before data cleansing, such as:
  + Invalid entries (e.g., ratings with NaN).
  + Incorrect or inconsistent data types (e.g., numerical data stored as strings).
  + Outliers (e.g., unusually high or low ratings/reviews).

**3. Data Cleansing Process**

* Document the steps taken to cleanse the data:
  + Handling of missing or invalid data.
  + Formatting adjustments for consistency (e.g., uniform date formats, number formats).
  + Removal of duplicate rows.

**4. Post-Cleansing Data Analysis**

* Show how the data looks after cleansing:
  + Reduced number of rows (if duplicates or invalid entries were removed).
  + Uniform data formats.
  + Any additional transformations that have been applied to make the dataset more usable.

**5. Comparative Insights**

* Compare statistics from both datasets, such as:
  + **Average ratings, reviews, and installs** before and after cleansing.
  + **Distribution of categories** and whether they have been affected.
  + Any impact on the **data types** and number of non-null values in each column.