University of Science and Technology
School of computa�onal science and ar�ficial intelligence
DSAI 202: Data governance

Name: Bosy Ayman

ID: 202202076

Sec: 2

Dataset name: Google Play Store

Dataset link:https://www.kaggle.com/datasets/lava18/google-play-store-apps/data

Dataset description

1- One or two paragraphs describing your dataset

The dataset described here contains information about apps available on the Google Play Store.

Each row in the dataset represents an individual app and includes various attributes such as category, rating, size, and more. These details offer a comprehensive view of the app landscape on the Google Play Store, enabling developers and analysts to uncover trends, preferences, and opportunities within the Android ecosystem.

2- Create a table for ALL the columns in your dataset similar to the following table:

It Containts duplicate values ,missed values and impropriate data type

Column name	Column profiling	Column dependencies	Notes
App	Text	**None**	Name of the app

Category	categorical	**None**	**There are 33 categories** **Most installed one is Family**
Rating	Numerical	**None**	The rate is out of 5.0 -It has missed values -Contains outliers
Reviews	Numerical	**None**	**Number of people rated the app** **Related to the rating**
Size	Text	**None**	**It's not numeric**
Installs	Text	**None**	**Related to the rating**
Type	categorical	**None**	**Contains missed values** **Most of the apps are free**

Price	Numerical	**Type**	**Most of the apps are free**
Content Rating	Categorical	**None**	**The age group the app is targeted at.** **80% is for everyone** **11% Teen** **8% other**
Genres	Text	**Category**	**An app can belong to multiple genres (apart from its main category).** **Tools 8%** **Entertainment 6%** **Other 86%**
Last Updated	Date	**None**	**The app must be updated regularly to match the android update**
Current Ver	Text	**None**	**Contains missed values**

Android Ver	Text	None	Contains missed values Most used android version is 4.1