REPORT: Exploratory Data Analysis (EDA) on the Flicr8k Dataset

The project utilizes the Flickr8k Image Dataset, a benchmark dataset for image description tasks. It consists of 8,092 images, with each image accompanied by up to five human-written captions. The motivation for using this specific dataset is to train an image captioning model. This task is a classic problem in deep learning that combines computer vision to extract features from an image and natural language processing to generate a textual description of the image content. The project's goal is to predict a new caption for an image by training the model on the features and corresponding captions from the dataset. a tabular dataset likely used for image-to-text tasks. The dataset consists of 40,455 entries with two distinct features:images and captions.

Image feature: This column has 8,091 unique values. The top 5 most frequent images each appear 5 times.

Caption feature: This column is highly diverse, with 40,201 unique captions.The top 5 most frequent captions are: (1) Two dogs playing in the snow (7 times). (2) two dogs play together (6 times). (3) A dog runs through the grass. (6 times). (4) A dog swimming with a stick in its mouth. (6 times). (5) A person on a bmx bike (6 times).

Visualization and Exploratory Data Analysis (EDA):