Exoplanet Dataset Description

This dataset is a comprehensive collection of confirmed exoplanets along with detailed information about their host stars and discovery details. Exoplanets are planets that exist outside our solar system, and their study helps scientists understand planetary formation, orbital behavior, and the possibility of life beyond Earth. The dataset is compiled from reliable astronomical missions and peer-reviewed scientific sources, making it suitable for academic, research, and internship-level projects.

Dataset Overview

The dataset contains approximately 39,000 records, where each record represents a confirmed exoplanet. It includes nearly 289 attributes that cover planetary characteristics, host star properties, and discovery-related information. Multiple astronomical identifiers such as Gaia, HIP, HD, and TIC are included to allow accurate cross-referencing with other space missions and star catalogs.

Planetary Characteristics

Planet-related attributes describe the physical and orbital properties of exoplanets. These include planet mass, radius, density, orbital period, semi-major axis, eccentricity, equilibrium temperature, and the amount of stellar radiation received. These parameters are important for understanding whether a planet is rocky, gaseous, or potentially Earth-like.

Host Star Information

The dataset provides detailed information about host stars, which play a major role in determining planetary environments. Host star attributes include stellar mass, radius, luminosity, effective temperature, age, metallicity, surface gravity, and distance from Earth. These properties help researchers analyze the relationship between stars and their planets and evaluate habitability conditions.

Discovery and Observation Details

Discovery-related information includes the year of discovery, detection method such as transit, radial velocity, or direct imaging, and details about the telescope or observatory used. Publication dates and data release indicators are also provided, ensuring transparency and reliability of the dataset.

Importance and Applications

This dataset is highly valuable for data analysis and machine learning applications. It can be used to study trends in exoplanet discoveries, classify planets based on physical properties, and build predictive models for planetary habitability. Due to its large size and wide range of features, the dataset is well suited for internship projects, academic research, and beginner-to-intermediate level data science work in astronomy.