

# Exoplanet Dataset - Data Description

---

## Dataset Overview

- Source: NASA Exoplanet Archive
- File Type: CSV
- Total Rows: 39,315
- Total Columns: 289
- Domain: Confirmed exoplanets and host star properties
- Note: Multiple rows may exist for the same planet. Use default\_flag = 1 for primary records.

## Planet Identification Details

- rowid – Unique row identifier
- pl\_name – Planet name
- hostname – Host star name
- pl\_letter – Planet designation
- hd\_name, hip\_name – Catalog identifiers
- tic\_id – TESS Input Catalog ID
- gaia\_dr2\_id, gaia\_dr3\_id – Gaia mission identifiers

## System Information

- sy\_snum – Number of stars in the system
- sy\_pnum – Number of planets
- sy\_mnum – Number of moons
- cb\_flag – Circumbinary system indicator

## Discovery Information

- `discoverymethod` – Detection method
- `disc_year` – Year of discovery
- `disc_facility` – Observatory/facility
- `disc_telescope` – Telescope used
- `disc_instrument` – Instrument used
- Detection flags: `rv_flag`, `tran_flag`, `micro_flag`, `ima_flag`

## Orbital Parameters

- `pl_orbper` – Orbital period (days)
- `pl_orbsmax` – Semi-major axis (AU)
- `pl_orbeccen` – Orbital eccentricity
- `pl_orbincl` – Orbital inclination
- Includes uncertainty and limit columns (`err1`, `err2`, `lim`)

## Planet Physical Properties

- `pl_rade` – Radius (Earth radii)
- `pl_bmasse` – Mass (Earth masses)
- `pl_dens` – Density
- `pl_eqt` – Equilibrium temperature
- `pl_insol` – Insolation flux

## Host Star Properties

- `st_teff` – Effective temperature
- `st_rad` – Stellar radius
- `st_mass` – Stellar mass
- `st_lum` – Luminosity
- `st_met` – Metallicity
- `st_logg` – Surface gravity

## **Data Characteristics**

- Contains measurement uncertainties
- Includes upper/lower limit flags
- Multiple entries per planet possible
- Missing values present in some physical parameters
- Mixed data types in catalog identifiers