Dataset Description

ExoHabitAI: Exoplanet Archive - TESS Mission

---

1. Overview

This dataset contains comprehensive scientific observations of exoplanets and their host stars from NASA's TESS (Transiting Exoplanet Survey Satellite) mission. Each record represents a confirmed or candidate exoplanet with detailed measurements of planetary properties, stellar characteristics, orbital parameters, and detection metadata.

Dataset Specifications:

- **Format:** CSV (Comma-Separated Values)
- **Records:** ~39,000 exoplanet observations
- **Features:** ~289 columns
- **Granularity:** One row = one exoplanet
- **Source:** https://exoplanetarchive.ipac.caltech.edu/

---

2. Feature Categories

a) Planetary Identifiers

Column	Description	Example
pl_name	Official planet name	TOI-700 d
hostname	Host star name	TOI-700
tic_id	TESS catalog ID	150428135
pl_letter	Planet designation in system	d
gaia_id	Gaia catalog ID	Gaia DR2 123456789
disc_year	Discovery year	2020

disc_facility	Discovery facility	TESS

b) Planetary Physical Properties

Column	Description	Units	Usage
pl_rade	Planet radius	Earth radii ($R_\oplus$)	Size classification
pl_bmasse	Planet mass	Earth masses ($M_\oplus$)	Type determination
pl_dens	Bulk density	g/cm³	Composition inference
pl_eqt	Equilibrium temperature	Kelvin	Habitability assessment
pl_insol	Stellar flux received	Earth flux	Energy input
pl_orbeccen	Orbital eccentricity	0.0 - 1.0	Orbit stability

Classification Guidelines:

- **Rocky planets:** $R < 1.5\ R_\oplus$, density > 4 g/cm³
- **Super-Earths:** $1.5 < R < 2.5\ R_\oplus$
- **Neptune-like:** $2.5 < R < 6\ R_\oplus$, low density
- **Gas giants:** $R > 6\ R_\oplus$ or $M > 50\ M_\oplus$
- **Potentially habitable:** 250 K < Temperature < 350 K, 0.3 < Flux < 1.5

c) Orbital Parameters

Column	Description	Units	Significance
pl_orbper	Orbital period	Days	Orbit duration
pl_orbsmax	Semi-major axis	AU	Star-planet distance
pl_trandep	Transit depth	ppm	Detection signal strength
pl_trandur	Transit duration	Hours	Transit geometry
pl_orbincl	Orbital inclination	Degrees	Viewing angle
pl_imppar	Impact parameter	Stellar radii	Transit path

Key Insights:

- Transit detection requires inclination ≥ 85°

- Shorter periods = closer, hotter planets
- Semi-major axis determines received stellar flux

d) Host Star Properties

Column	Description	Units	Classification Use
st_teff	Effective temperature	Kelvin	Spectral type
st_rad	Stellar radius	Solar radii ($R_\odot$)	Star size
st_mass	Stellar mass	Solar masses ($M_\odot$)	Evolutionary stage
st_lum	Luminosity	Solar luminosity	Energy output
st_met	Metallicity [Fe/H]	Dex	Chemical composition
st_logg	Surface gravity	$\log_{10}(cm/s^2)$	Stellar density

Stellar Types:

- **M-dwarf:** Teff < 4000 K (most common exoplanet hosts)
- **K-type:** Teff 4000-5200 K
- **G-type (Sun-like):** Teff 5200-6000 K
- **F-type:** Teff 6000-7500 K

e) Detection Metadata

Column	Description	Values
disposition	Confirmation status	"Confirmed" / "Candidate"
disposition_source	Confirming authority	Institution/Survey
pl_refname	Reference publication	ADS bibcode
rv_flag	Radial velocity detection	0/1
tran_flag	Transit detection	0/1
ttv_flag	Transit timing variation	0/1

Quality Indicators:

- **Confirmed + Multiple methods:** Highest confidence
- **Confirmed + Single method:** Medium confidence

- **Candidate:** Awaiting confirmation

---

3. Data Quality

a) Missing Data Patterns

Common Missing Values:

- **Planet mass (~60%):** Often unavailable for transit-only detections
- **Density:** Missing when mass or radius unavailable
- **Eccentricity:** Difficult to measure for long-period planets
- **Some stellar parameters:** Incomplete for faint stars

b) Data Validation

Physical Consistency Checks:

- Density should be 0.1 - 30 g/cm³
- Temperature should correlate with stellar flux
- Orbital period should match semi-major axis (Kepler's laws)
- Transit depth should match radius ratio

---

4. Interpretation Caveats

- **Mass-radius degeneracy:** Similar sizes can have different compositions
- **Atmospheric assumptions:** Equilibrium temperature assumes no atmosphere
- **Stellar activity:** Can introduce measurement noise
- **Multi-planet interactions:** Complex gravitational dynamics in some systems

---

5. Quick Reference

a) Planet Size Classification

Category	Radius Range	Mass Range	Example
Earth-sized	0.8-1.25 $R_\oplus$	0.5-2 $M_\oplus$	Earth, Venus

Super-Earth	1.25-2 $R_\oplus$	2-10 $M_\oplus$	Kepler-452b
Mini-Neptune	2-4 $R_\oplus$	10-20 $M_\oplus$	GJ 1214b
Neptune-like	4-6 $R_\oplus$	10-50 $M_\oplus$	Neptune
Jupiter-like	>6 $R_\oplus$	>50 $M_\oplus$	Jupiter

b) Stellar Classification

Type	Temperature (K)	Color	Common Hosts
M	<3,900	Red	Proxima Centauri
K	3,900-5,200	Orange	Epsilon Eridani
G	5,200-6,000	Yellow	Sun
F	6,000-7,500	White	Procyon

c) Habitability Criteria

- **Temperature:** 250-350 K (liquid water range)
- **Stellar flux:** 0.3-1.5 Earth flux
- **Radius:** 0.8-1.5 $R_\oplus$ (rocky planets)
- **Orbital eccentricity:** <0.3 (stable conditions)
- **Star type:** K or G preferred (stable, long-lived)

---

6. Glossary

AU: Astronomical Unit - Earth-Sun distance (~150 million km)
Equilibrium Temperature: Expected surface temperature from energy balance
Flux: Stellar energy received per unit area
Habitable Zone: Orbital region where liquid water could exist
Inclination: Orbital tilt relative to observer's line of sight
Radial Velocity: Star's motion due to planet's gravitational pull
Semi-major Axis: Average orbital distance
Transit: Planet passing in front of its star
Transit Depth: Brightness decrease during transit

Document Version: 1.0
Last Updated: January 12, 2025
Prepared by: Dheeraj M