Process

Motivation

According to the United Nations, Earth's population is projected to keep rising for the next 60 years, reaching a peak of about **10.3 billion** in the **mid-2080s**. With this continuous growth, Earth will **face increasing challenges** related to resources, sustainability, and population management. As demands for food, water, and energy increase, and as climate change and environmental degradation threaten the sustainability of ecosystems, the question arises: how can humanity ensure its long-term survival? This question becomes more urgent in light of reality: humans are consuming earth's resources at an alarmingly high rate. Forests are being cleared faster than they can be regrown, freshwater supplies are dwindling in many regions across the globe, and non-renewable energy sources such as fossil fuels—which currently make up **82% of global energy consumption**—are being rapidly depleted. Even renewable resources, such as wind, solar, and geothermal energy, are being **pushed to their limits** by growing demand and economic constraints. With a growing population consuming more than the planet can sustainably provide, Earth, on its own, may not be enough to support humanity's needs. This leads us to the problem: **How can we sustain humanity?** Through the use of the Exoplanet archive, we will explore what factors can get us closer to this new home and what are potential candidates for a home.

Methodology

- We began by exploring the database to identify **key variables** relevant to planetary habitability and examined their relationships. Columns with a high proportion of missing values were removed to ensure data quality.
- With the remaining data, we performed an inner join to create a unified table free of null entries. Using **pandas**, we engineered new quantitative features, such as planetary density, and introduced **categorical classification** fields based on scientifically informed thresholds to better capture indicators of habitability.
- To **visualize and refine** our feature selection, we generated interactive histograms of density and applied color-coded "hues" based on researched benchmarks to distinguish terrestrial from gaseous planets. We then identified the most informative variables and computed summary statistics to analyze central tendencies and variability.
- To consolidate different variables, we **standardized the units** and created a standard unit distance based on these variables from Earth's standard units.
- **Histograms** were used to further examine the distributions of these variables, allowing us to **filter** out planets that fell outside the desired ranges. Finally, we integrated the curated dataset, representing planets within a reasonable distance and meeting our **habitability criteria** into an API for display on the website.

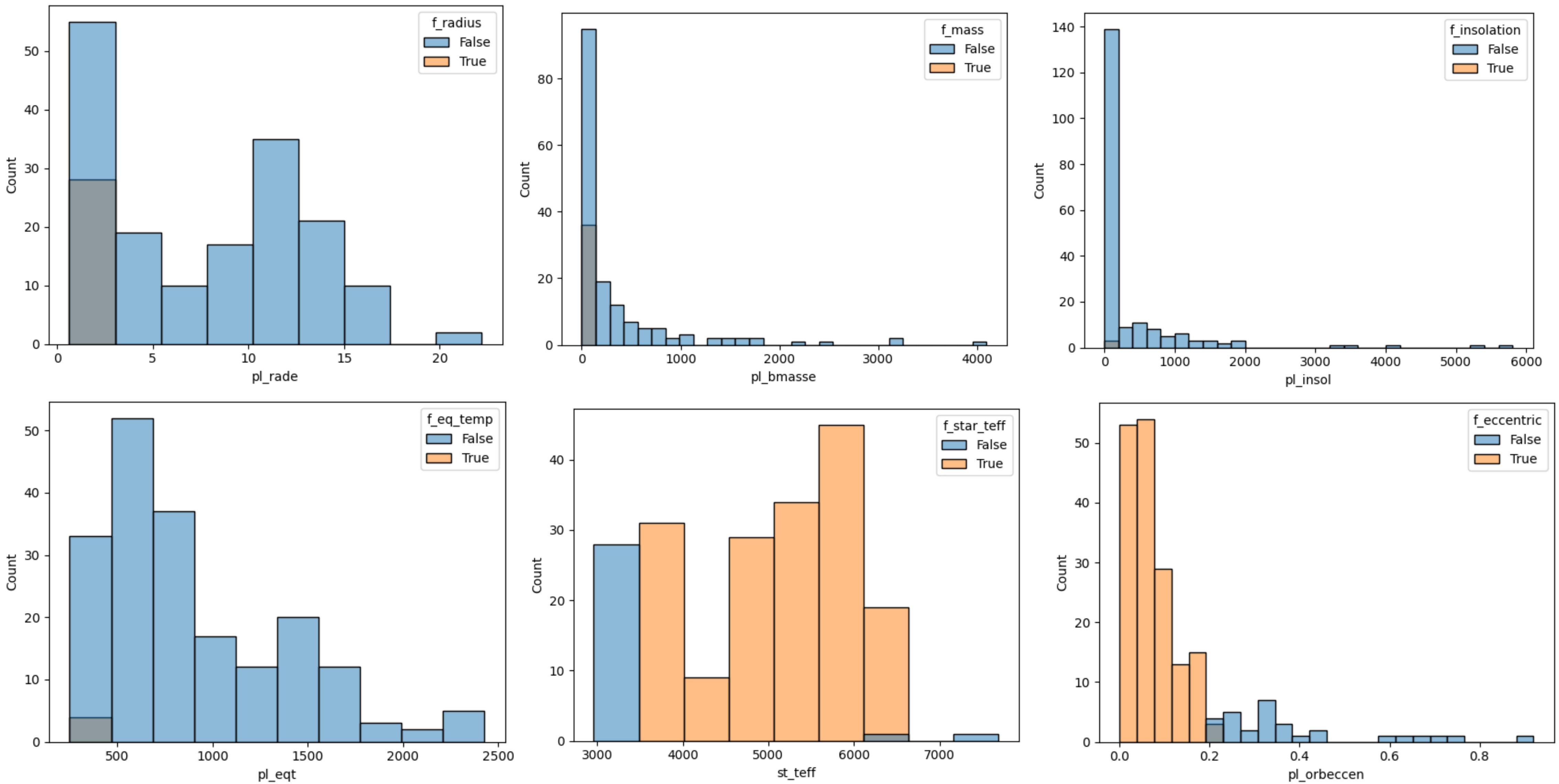
6 Variables

- **Planetary Radius (pl_rade):** The radius of the planet as a ratio to Earth's radius. Dictates gravity, atmospheric pressure, and the state of water on a planet.
- **Planetary Mass(pl_masse):** The mass of the planet as a ratio to Earth's mass. Determines the ability to maintain a magnetic field, terrestrial state, and maintaining an atmosphere.
- **Stellar Insolation(pl_insol):** The amount of light the planet receives compared to Earth per a square unit on average. This indicated the climate on the planet.
- **Equilibrium Temperature(pl_eqt):** Estimate of average temperature based on star distance and insolation. Can help to determine surface temperature and cooling or reflection.
- **Stellar Effective Temperature(st_teff):** The surface temperature, flares, and wavelength. Can help determine the radiation stability.
- **Orbital Eccentricity(pl_orbeccen):** The sensitivity of the seasons and climate through the orbit. Lower values can give a more stable bodies of water and climate, preventing huge swings or collapses.

Typical Units	True Habitable Ranges
Earth radii (R⊕)	Planet's radius is between **0.5 and 1.6 R⊕**
Earth masses (M⊕)	Planet's mass is between **0.2 and 5 M⊕**
Earth insolation units (S⊕)	Stellar energy received is **0.35–1.75 × Earth's**
Kelvin (K)	Equilibrium temperature is **180–310 K**
Kelvin (K)	Host star's effective temperature is **3500–6500 K**
Dimensionless orbital eccentricity	Orbital eccentricity is **< 0.2**

Deciphering Between Habitability

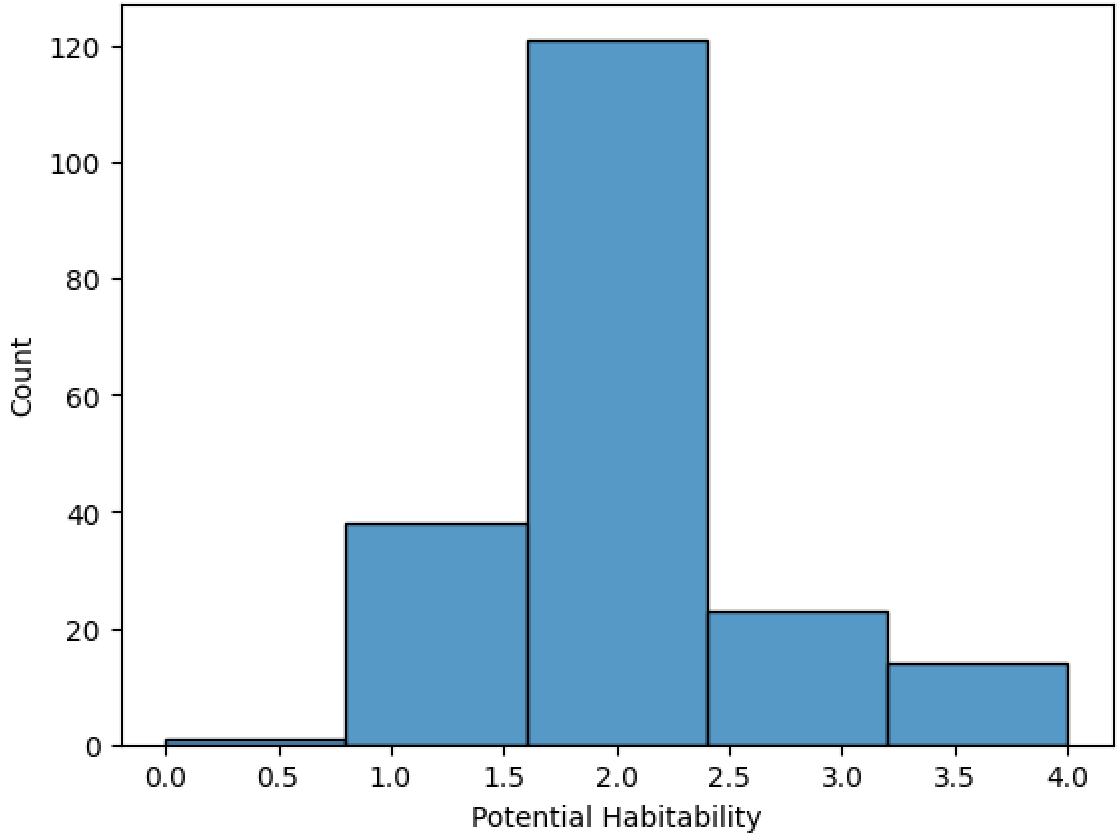
Variable	True	False
Planetary Radius (pl_rade)	169	28
Planetary Mass(pl_masse)	161	36
Stellar Insolation(pl_insol)	194	3
Equilibrium Temperature(pl_eqt)	193	4
Stellar Effective Temperature(st_teff)	167	30
Orbital Eccentricity(pl_orbeccen)	167	30


f_radius
False
True
Count
pl_rade
f_mass
False
True
Count
pl_bmasse
f_insolation
False
True
Count
pl_insol
f_eq_temp
False
True
Count
pl_eqt
f_star_teff
False
True
Count
st_teff
f_eccentric
False
True
Count
pl_orbeccen

Fields Attained

Fields	Count
0	1
1	38
2	121
3	23
4	14

HOW MANY PLANETS PASSED THE CONDITIONS?



Introducing Second Earth

Results

Through our desired variables we were able to create a set of planets that are more habitable than majority of the planets found in archive, saving researchers time from sorting through the thousands of planets. Furthermore, we have the basis to create a classification model, the habitable score decision tree or regression model for a habitation score. Overall, we built a way to conveniently and clearly display direction for discovery, accelerating the process sustaining human life.