100 Worlds Beyond Our Sun: A Data-Driven Search for Earth 2.0

Raminder S. Kalsi

Etiwanda High School

Independent Research Project

Self-Directed

October 1, 2025

Abstract

Inspired by my interest in astrobiology, I used a NASA dataset that includes more than 38,000 confirmed and candidate planets. This study looks at the top 100 exoplanets that may be habitable. Planets were chosen through data cleaning and filtering according to the orbital period, planet radius, equilibrium temperature, surface temperature, and stellar insolation—all of which are important factors for habitability. The statistical analysis brought up several candidates with Earth-like characteristics. These findings support the ongoing search for an "Earth 2.0" and may influence the selection of targets for future research.

Keywords: exoplanets, habitable, NASA dataset, Earth 2.0, analysis

Introduction

In order to select the top 100 habitable worlds, this study analyzes the data from the NASA Exoplanet Archive, which consists of over 38,000 planets. The study takes into account 5 main features that are crucial for habitability: the orbital period, planet radius, equilibrium temperature, surface temperature, and stellar insolation. Python was used for cleaning data, filtering, and statistical analysis. These factors are essential for identifying the planets with Earth-like conditions and helping us focus on the ones that resemble Earth most closely and have potential to support life. The results aim to help scientists decide which planets be studied more closely in the future, potentially revealing a true “Earth 2.0”. While astronomers continue their hunt in the stars for signals of life, the question remains: Which distant planets might be like our own?

Method

Dataset:

This study used NASA’s Exoplanet Archive, which contains over 38,000 confirmed and candidate exoplanets. Data analysis conducted by using Python in both Jupyter Notebook and Google Colab, which allowed for smooth, interactive coding and easy visualization throughout the process. To ensure the data was accurate and most relevant, the dataset was filtered and cleaned:

- Planets containing missing or null values in necessary fields such as planet radius, orbital period, and stellar temperature were filtered out.
- Incomplete or poorly constrained entries were removed, ensuring that only exoplanets with sufficient data for evaluation were retained.

Assessments and Measures:

The main variables looked at were orbital period, planet radius, equilibrium temperature, surface temperature, and stellar insolation. Earth-like conditions were used as a baseline, but the study also considered other planetary traits to reflect the different environments that could potentially support life. After cleaning and filtering the data, these variables were examined to get a better sense of the selected exoplanets' characteristics. To dig deeper, visual tools like scatter plots and histograms were used to spot patterns, relationships, and outliers in the data, which helped make sense of the planets' and stars' properties.

Habitability Ratings:

A table was created using Python to assign a habitability score for each planet relative to Earth's conditions. This score provided a way to compare exoplanets based on how similar their key features are to those of Earth, serving as a reference for potential habitability and scientific interest. The score provided was divided into three different categories, Category 1, Category 2, and Category 3.

Category 1. Planetary conditions with a score of 7–10 out of 10 are highly habitable. This exoplanet shares many characteristics with Earth, such as stable temperatures, favorable orbital period, and size that is similar to Earth's. This planet has a high potential to support life as we know it. Finding a planet that closely resembles Earth in important ways is extremely rare, indicating why only one attained this achievement.

Category 2. Planets in this category have a score of 4 to 6.9 out of 10, indicating moderate habitability. Although these exoplanets have some characteristics in common with Earth, they may differ in one or more important areas, like size or temperature range. Even though they are not perfect, they still have room for more research and, in some cases, could sustain life.

Category 3. Planets in this category have a habitability score of less than 4 out of 10, indicating low habitability. There are notable distinctions between these exoplanets such as extreme temperatures or unsuitable sizes and more.

*Kepler 442b***.** As a category 1 planet, Kepler-442b orbits a relatively calm K-type star and is located in the constellation Lyra, approximately 1,200 light-years away. It completes a revolution every 112 days and is marginally larger than Earth, with a radius that is roughly 1.34 times larger. Scientists believe it may have temperatures appropriate for liquid water because it is located in the habitable zone of its star and receives energy from the Sun that is comparable to that received by Earth. The combination of Kepler-442b's Earth-like size, steady orbit, and star's low radiation is what makes it one of the most promising candidates so far.

TOI-2095 c. As a Category 2 planet, TOI-2095 c is a super-Earth orbiting an M-dwarf star about 150 light-years away in the constellation Aries. With an orbital period of approximately 21 days and an estimated radius of approximately 1.4 times that of Earth, it falls within its star's habitable zone, receiving a moderate level of stellar energy. Its parent red dwarf produces flares and radiation that will create difficulty for the atmosphere to remain stable. While TOI-2095 c has some Earth-like aspects, it is not yet confirmed as habitable through more detailed atmospheric data. Nevertheless, it is one of the more promising M-dwarf candidates.

K2-216 b. This Category 3 planet orbits a K-type main-sequence star approximately 125 light-years distant in the Taurus constellation. The exoplanet is rocky, with a radius slightly more than 1.75 Earths, and it orbits in just a little more than two days, placing it very close to its host star. As a result of this close orbit, it has extreme stellar radiation and hot surface temperatures, making it uninhabitable by the standards of Earth. Despite its low habitability, K2-216 b is valuable in the study of the formation of rocky planets in hostile stellar environments.

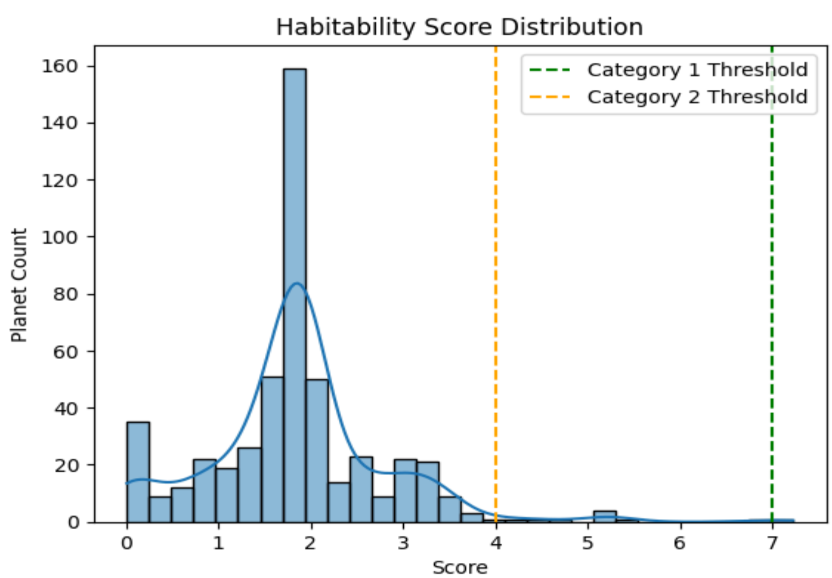
Results

Overview of Findings:

After filtering and reviewing over 3,000 exoplanets, 100 were identified as meeting the habitability criteria defined in this study. These candidates exhibit:

- Equilibrium temperatures ranging from 198.6 K to 2463 K
- Radii between 0.51 and 11.58 Earth radii

Such values suggest a wide variety of planetary environments. While many fall outside the "Goldilocks zone," some demonstrate conditions that could be crucial for supporting life. However, confirmation of their habitability would require further observation, especially regarding atmospheres and surface conditions.



(Figure 1) shows the relationship between equilibrium temperature and overall habitability score among the top 100 candidates.

Outcome 1: Habitability Score Distribution:

Each planet was assigned a score from 0 to 10, comparing how closely it aligned with Earth's known life-supporting characteristics.

- Category 1 (High Habitability): 1 planet
  - Scored ≥ 7.0; very Earth-like in key parameters.
- Category 2 (Medium Habitability): 8 planets
  - Scored between 4.0 and 6.99; Earth-like in some aspects, differs in some.
- Category 3 (Low Habitability): 90+ planets
  - Scored < 4.0; one or more traits are significantly un-Earth-like.

This scoring system highlights the rarity of planets that exhibit ideal combinations of features.

Outcome 2: Parameter Trends

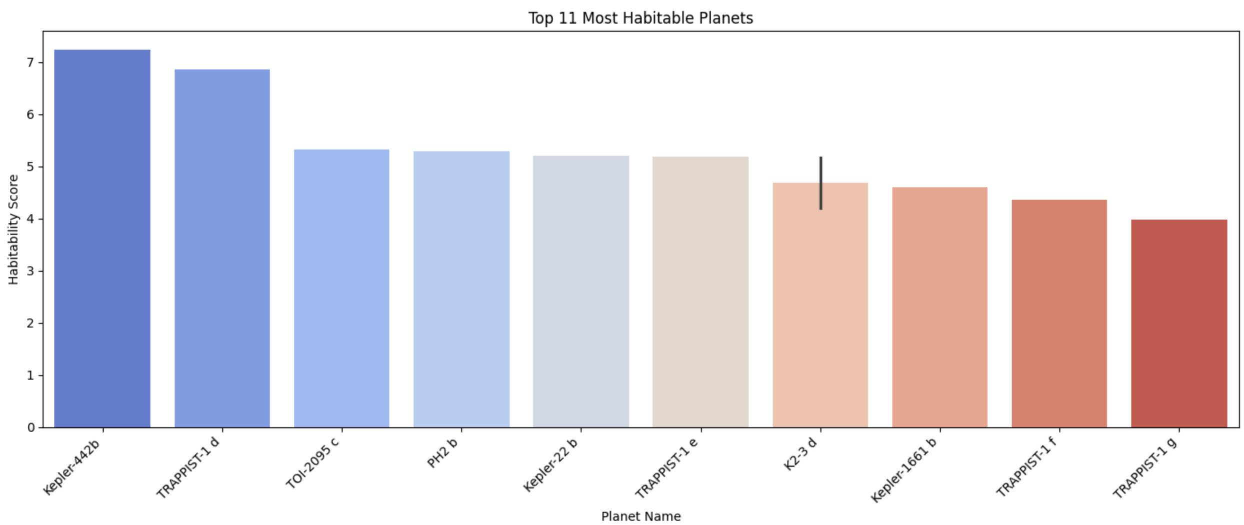
Habitability scores were calculated using a weighted formula based on planetary and stellar features that strongly influence habitability. The weights were assigned as follows:

Parameter	Weight	Rationale
Planet Radius	25%	Indicates likelihood of a rocky surface and atmospheric retention
Stellar Insolation	25%	Directly affects climate and potential for liquid water
Eccentricity	20%	Stable, low-eccentricity orbits reduce temperature extremes

Orbital Period	20%	Used as a proxy for distance when flux data is missing
Temperature	10%	Based on surface and equilibrium temperature (if available), indicating the possibility of the presence of water, which is crucial for life to thrive as we know it.

These weights reflect the relative importance of each factor in supporting life as we understand it. Earth was used as the baseline (10/10 in each category), and all other planets were scored in comparison.

Top Habitability Candidates:



(Figure 2) Ranks the top 10 exoplanets by total habitability score.

The following chart shows the 10 most habitable exoplanets identified in this study. These candidates scored highest across all weighted categories and most closely match Earth's balance of size, temperature, and stable orbit. Common characteristics among top scorers include:

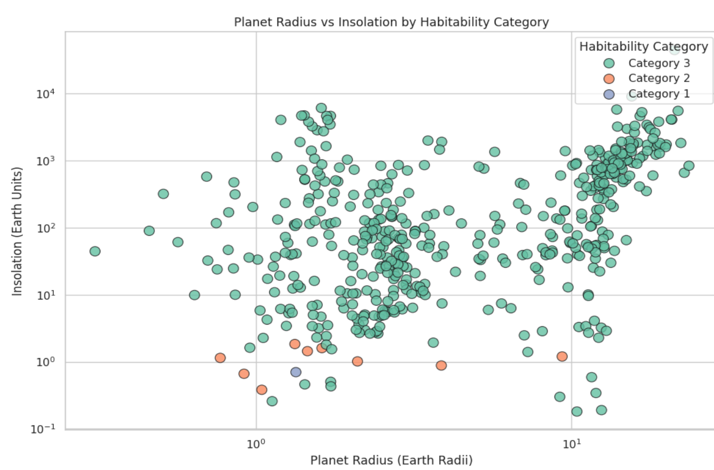
- Temperatures consistent with potential for liquid water

- Near-circular orbits (low eccentricity)
- Radii close to Earth's
- Energy received from their host stars comparable to Earth's

Discussion

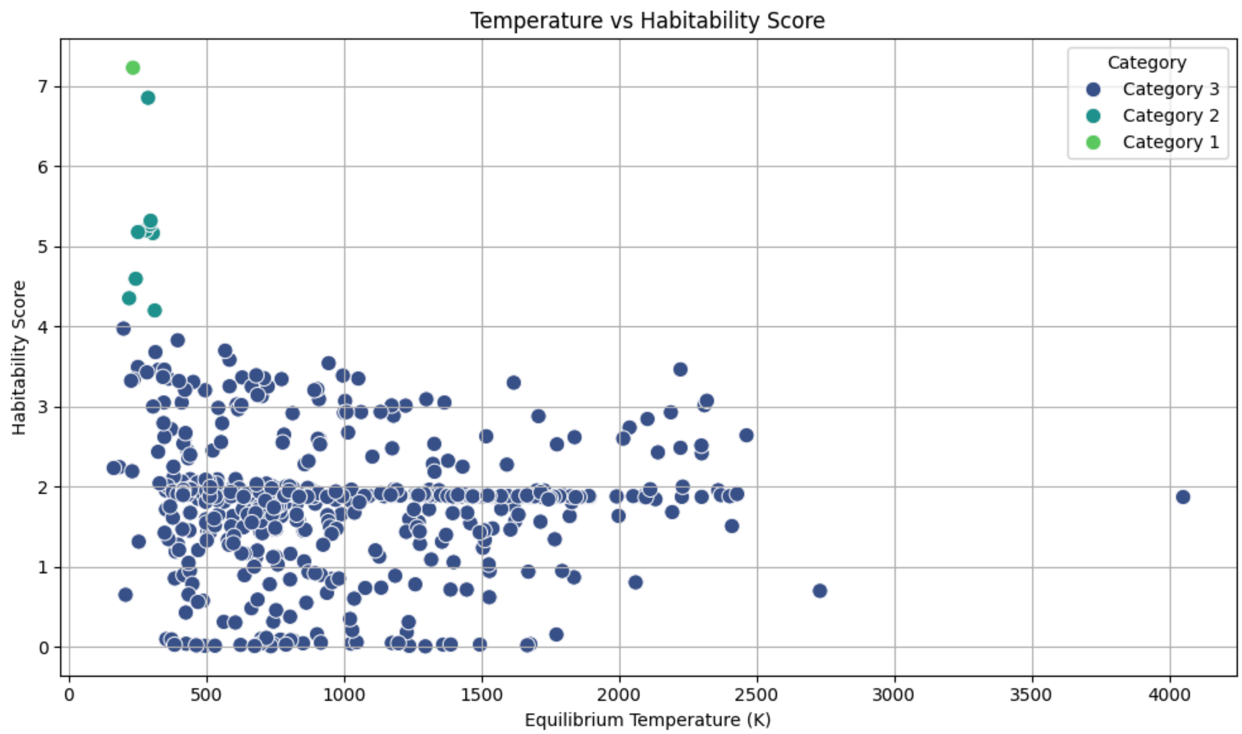
These findings demonstrate that a considerable fraction of exoplanets exhibit Earth-like features—particularly in size, energy input, and orbital behavior—within the zone of potential habitability. The ranking scheme in this study used features most strongly linked to the ability of the presence of liquid water and stable climates. Habitable status cannot, however, be determined from physical properties. Most of the top-rated planets lack essential information on the atmospheres' composition, surface pressure, and biosignatures that would be needed to confirm life-friendly conditions. Therefore, while the scoring system is a useful filter, it has to be viewed as a starting point, not a conclusion.

Parameter Interactions:



(Figure 3) displays the relationship between planetary radius and stellar insolation among the top-scoring candidates.

1. **Planet Radius vs. Stellar Insolation:** This graph shows that smaller, rocky planets tend to cluster at moderate levels of stellar flux—suggesting that planets similar in size to Earth are more likely to lie within a favorable energy range.



(Figure 1) shows the relationship between equilibrium temperature and overall habitability score among the top 100 candidates.

2. **Temperature vs. Habitability Score:** This plot reveals a clear peak in habitability score around Earth-like equilibrium temperatures (roughly 250–300 K). Scores drop off drastically at extreme temperatures, reaffirming the importance of optimal temperatures to life as we know it.

These trends support the logic behind the weighted scoring model and emphasize that even among promising candidates, the balance between parameters is critical.

Limitations

While this study presents a filtered and analyzed list of potentially habitable exoplanets, several limitations should be acknowledged:

1. Data Limitations

The dataset from the NASA Exoplanet Archive contained missing, estimated, or uncertain values for key parameters such as radius, mass, and orbital characteristics. These gaps may affect the accuracy of comparisons and overall scoring. For example, K2-3 d appeared multiple times due to duplicate entries, affecting ranking clarity.

2. Measurement Uncertainty

Many planetary characteristics are based on indirect observations and carry significant uncertainty. These values—especially for surface temperature, eccentricity, or insolation—may vary as newer, more precise measurements become available.

3. Assumptions in Filtering and Scoring

The habitability scoring model relied on a predefined set of observable parameters:

- Radius (25%)
- Stellar insolation (25%)
- Eccentricity (20%)
- Orbital period (20%)
- Temperature (10%)

While this framework highlights Earth-like features, it does not account for several critical aspects of habitability, such as atmospheric composition, surface conditions, magnetic field strength, or planet type (e.g., terrestrial vs. gas giant). This could result in higher scores for planets that are not realistically habitable.

4. Lack of Atmospheric and Time-Series Data

This study analyzed static planetary characteristics only. Key insights from time-series data—such as light curves from transits and more—were not included, which limits the depth of the habitability evaluation.

5. Methodological Constraints

Although Python and its libraries allowed for great data cleaning and visualization, the analysis did not use more advanced approaches such as machine learning, probabilistic modeling, or other simulations. These techniques could enhance future investigations by handling uncertainty and complexity more effectively.

Conclusion

This study examined 100 exoplanets marked as good prospects for habitability based on their planetary and stellar parameters. With a weighted scoring system based on Earth's values, the planets were graded into three categories of habitability. While only one achieved high habitability scores, many gave encouraging signs that require further investigation. The results highlight the importance of planetary radius, stellar insolation, orbital period, orbital eccentricity, and temperature in determining potential habitability.

Although promising, this study is the first step that relies on accessible data and assumptions. Additional work with input from atmospheric composition, surface conditions, and orbital stability over long timescales will be required to make habitability estimates as accurate as possible. The planets mentioned are the main targets for additional observations with next-generation telescopes and missions searching for life outside our solar system.

In summary, this research is contributing to the search for worlds that can support life by providing an evidence-based method of determining the potential of exoplanets for supporting life. Future advancements in observation technology will have to confirm these distant worlds and their potential for supporting life.

References

1. "NASA Exoplanet Archive." *NASA/IPAC Exoplanet Science Institute*, California Institute of Technology, https://exoplanetarchive.ipac.caltech.edu. Accessed 1 June 2025.
2. "Welcome to Python.org." Python Software Foundation, https://www.python.org. Accessed 22 June 2025.
3. "NumPy." *NumPy.org*, https://numpy.org. Accessed 22 June 2025.
4. "pandas - Python Data Analysis Library." *pandas.pydata.org*, https://pandas.pydata.org. Accessed 22 June 2025.
5. "What Makes a Planet Habitable?" *Sellers Exoplanet Environments Collaboration (SEEC)*, NASA Goddard Space Flight Center, https://seec.gsfc.nasa.gov/what_makes_a_planet_habitable.html. Accessed 22 June 2025.
6. Schulze-Makuch, Dirk, et al. "Requirements and Limits for Life in the Context of Exoplanets." *Proceedings of the National Academy of Sciences*, vol. 110, no. 48, 2013, pp. 21310–21315. https://doi.org/10.1073/pnas.1304212111. Accessed 1 June 2025.
7. Meadows, Victoria S., and Rory K. Barnes. "Factors Affecting Exoplanet Habitability." *Handbook of Exoplanets*, edited by Hans J. Deeg and Juan C. Belmonte, Springer, 2018, pp. 57–80. https://doi.org/10.1007/978-3-319-55333-7_57. Accessed 1 June 2025.
8. Lissauer, Jack J. "Habitable Zone." *Encyclopaedia Britannica*, Encyclopaedia Britannica, Inc., https://www.britannica.com/science/habitable-zone. Accessed 1 June 2025.
9. "Exoplanet Habitability: Criteria & Conditions." *Vaia*, 2024, https://www.vaia.com/en-us/explanations/physics/astrophysics/exoplanet-habitability/. Accessed 1 June 2025.

10. “Google Colaboratory.” Google Research, https://colab.research.google.com/. Accessed 22 June 2025.