

NAME – JAIDEEP KUMAR SINGH SECTION - K23FA

ROLL NO – 62 REGNO – 12317127

COURSE CODE – INT375

COURSE TEACHER – SANDEEP KAUR

PYTHON CA 2 PROJECT

**Exoplanet Data Analysis**

The ongoing exploration of exoplanets represents one of the most exciting frontiers in modern astrophysics. This project presents a comprehensive data-driven analysis of a large exoplanet catalog, designed to address fundamental questions about the nature, diversity, and distribution of planets beyond our solar system.

Our investigation is structured around six core objectives, each targeting a specific aspect of exoplanet science. We will begin by evaluating the efficacy of various exoplanet detection methodologies, comparing their yields and biases to understand how our observational techniques shape our view of exoplanetary systems. We will then proceed to a detailed characterization of exoplanet sizes, constructing a statistical profile of planetary radii to assess the prevalence of different planet types and inform estimates of Earth-like planet occurrence rates.

A quantitative assessment of the contribution of each discovery method to the total exoplanet count will provide insight into the relative importance of different observational strategies. Furthermore, we will reconstruct the timeline of exoplanet discoveries, visualizing the exponential growth of detections and connecting it to advancements in astronomical instrumentation and data analysis. We will also examine the spectral characteristics of exoplanet host stars, categorizing them and identifying trends that may reveal preferred stellar environments for planet formation. Finally, we will delve into the relationship between stellar effective temperature and exoplanet radius, seeking to uncover potential correlations that could illuminate the processes of planet formation and orbital migration. Through rigorous data analysis and compelling visualizations, this project aims to provide a valuable contribution to the collective effort to understand the place of our solar system within the broader context of the galaxy and the search for potentially habitable worlds

.

# zv‘’ Objective 1: What Types of Stars Host the Most Exoplanets?

# This objective categorizes the types of stars known to have exoplanets, highlighting the most frequent host star classifications.

Code:

plt.figure(figsize=(12, 8))

text = " ".join(star for star in df['st\_spectype'].dropna().astype(str))

wordcloud = WordCloud(width=800, height=400, background\_color='white', colormap='turbo').generate(text) # 'turbo' colormap

plt.imshow(wordcloudinterpolation='bilinear')

plt.axis("off")

plt.title('Stellar Types Hosting Exoplanets', fontsize=14)

plt.tight\_layout(pad=0)

plt.show()

A close up of letters

AI-generated content may be incorrect.**OUTPUT:**

# Vddddsdddobjective 2:  What are the Most Common Exoplanet Discovery Methods?

# This objective identifies the primary techniques used to detect exoplanets and compares their prevalence in current discoveries.

# Code:  plt.figure(figsize=(10, 6))

# df['discoverymethod'].value\_counts().sort\_values().plot(kind='barh', color=sns.color\_palette('husl', len(df['discoverymethod'].unique())), fontsize=10) # 'husl' palette

# plt.title('Most Common Exoplanet Discovery Methods', fontsize=14)

# plt.xlabel('Number of Exoplanets', fontsize=12)

# plt.ylabel('Discovery Method', fontsize=12)

# plt.tight\_layout()

# plt.show()

# 

# Output:

# Objective 3: How Do Exoplanet Sizes Compare?

# This objective visualizes the range of exoplanet radii to illustrate the variety of planet sizes found beyond our solar system.

# Code: plt.figure(figsize=(8, 6))

# sns.violinplot(y=df['pl\_rade'].dropna(), color='springgreen') # 'springgreen' color

# plt.title('Distribution of Exoplanet Sizes (Earth Radii)', fontsize=14)

# plt.ylabel('Exoplanet Radius (Earth Radii)', fontsize=12)

# plt.yscale('log')

# plt.tight\_layout()

# plt.show()

# 

# Objective 4: What is the Proportion of Exoplanets Discovered by Each Method?

# This objective determines the relative contribution of each discovery method to the total number of exoplanets detected.

# Code: plt.figure(figsize=(12, 6)) # Increased width for bar chart labels

# method\_counts = df['discoverymethod'].value\_counts()

# sns.barplot(x=method\_counts.index, y=method\_counts.values, palette='viridis') # Bar chart

# plt.title('Number of Exoplanets Discovered by Each Method', fontsize=14)

# plt.xlabel('Discovery Method', fontsize=12)

# plt.ylabel('Number of Exoplanets', fontsize=12)

# plt.xticks(rotation=45, ha='right', fontsize=10)

# plt.tight\_layout()

# plt.show()

# A graph with a number of explanatory text AI-generated content may be incorrect.

# Objective 5: How Many Exoplanets Have Been Discovered Each Year?

# This objective tracks the progress of exoplanet detection over time, showing the growth of discoveries year by year.

# Code: plt.figure(figsize=(10, 6))

# years = df['disc\_year'].value\_counts().sort\_index()

# plt.plot(years.index, years.values, marker='o', color='tomato') # 'tomato' color

# plt.fill\_between(years.index, years.values, color='tomato', alpha=0.2)

# plt.title('Exoplanets Discovered Per Year', fontsize=14)

# plt.xlabel('Discovery Year', fontsize=12)

# plt.ylabel('Number of Exoplanets Discovered', fontsize=12)

# plt.grid(axis='y', alpha=0.7)

# plt.tight\_layout()

# plt.show()

# 

# A screenshot of a computer AI-generated content may be incorrect.Screenshot of workspace :

# A screen shot of a computer AI-generated content may be incorrect.