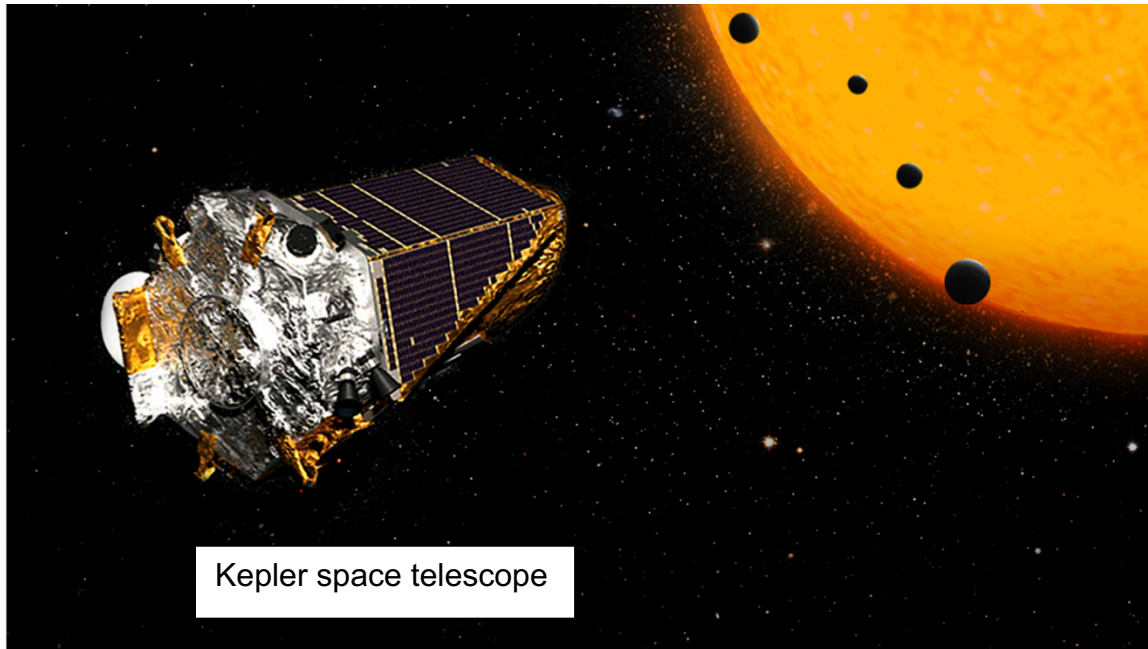


# ***Visualization of Kepler Exoplanet Search Results***



Kepler space telescope

**Author:** Md Redyan Ahmed

**Affiliation:** Indian Institute of Science Education and Research- Kolkata

**Supervisor:** Khan Muhammad Saqiful Alam

**Affiliation:** National University of Singapore

## THEORY AND CONTEXT:

All of the planets in our solar system orbit around the sun. Planets that orbit around other stars are called exoplanets. The Kepler space observatory is a Nasa-build satellite that was launched in 2009. The telescope is dedicated to search for exoplanets in star systems besides our own, with the ultimate goal of possibly finding other habitable planets besides our own. The original mission ended in 2013 due to mechanical failures, but the telescope has nevertheless been functional since 2014 on a "k2" extended mission.

Kepler had verified 1284 new exoplanets as of May 2016. As of October, 2017 there are over 3000 confirmed exoplanets total (using all detection methods, including ground-based ones). The telescope is still active and continues to collect new data on its extended mission.

## Content:

This dataset is a cumulative record of all observed kepler "objects of interest" — basically, all of the approximately 10,000 exoplanet candidates kepler has taken observations on.

This dataset has an extensive data dictionary, which can be accessed [here](#). Highlightable columns of note are:

- **kepoi\_name:** A KOI is a target identified by the Kepler Project that displays at least one transit-like sequence within Kepler time-series photometry that appears to be of astrophysical origin and initially consistent with a planetary transit hypothesis
- **kepler\_name:** [These names] are intended to clearly indicate a class of objects that have been confirmed or validated as planets—a step up from the planet candidate designation.
- **koi\_disposition:** The disposition in the literature towards this exoplanet candidate. One of CANDIDATE, FALSE POSITIVE, NOT DISPOSITIONED or CONFIRMED.
- **koi\_pdisposition:** The disposition Kepler data analysis has towards this exoplanet candidate. One of FALSE POSITIVE, NOT DISPOSITIONED, and CANDIDATE.
- **koi\_score:** A value between 0 and 1 that indicates the confidence in the KOI disposition. For CANDIDATES, a higher value indicates more confidence in its disposition, while for FALSE POSITIVES, a higher value indicates less confidence in that disposition.

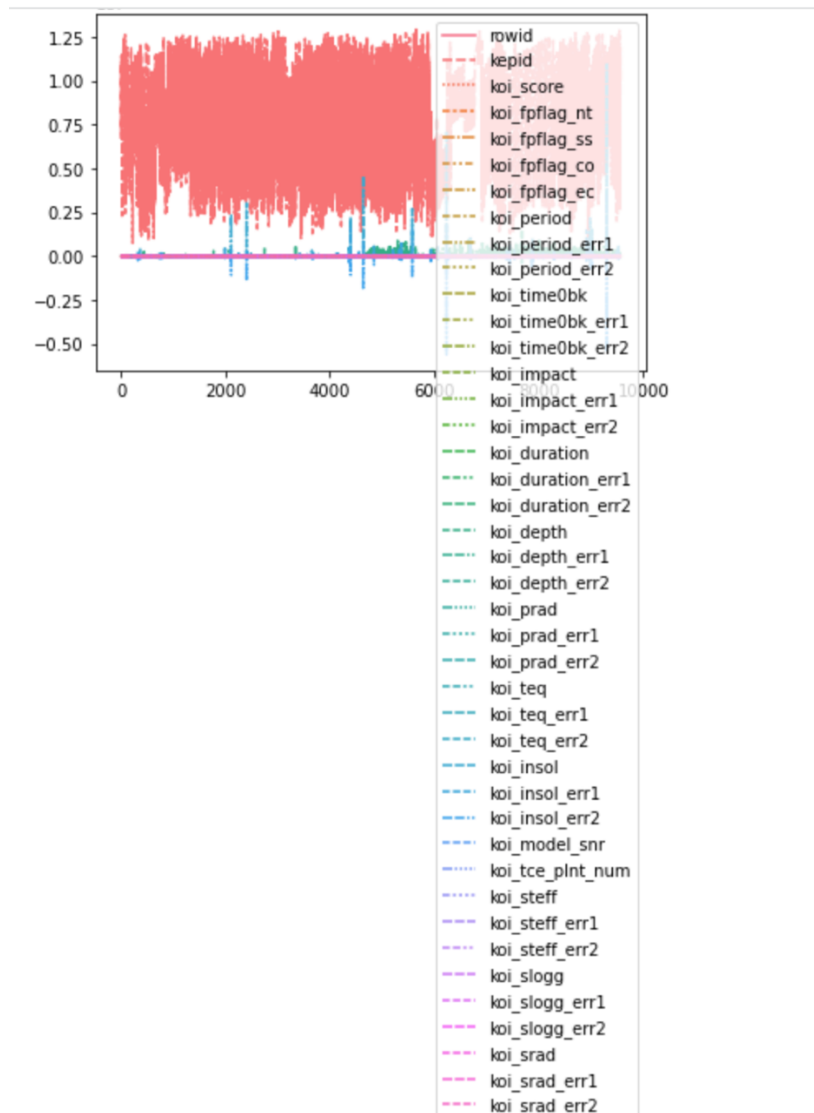
### Data Visualization:

Firstly, let's take a glimpse at the data :

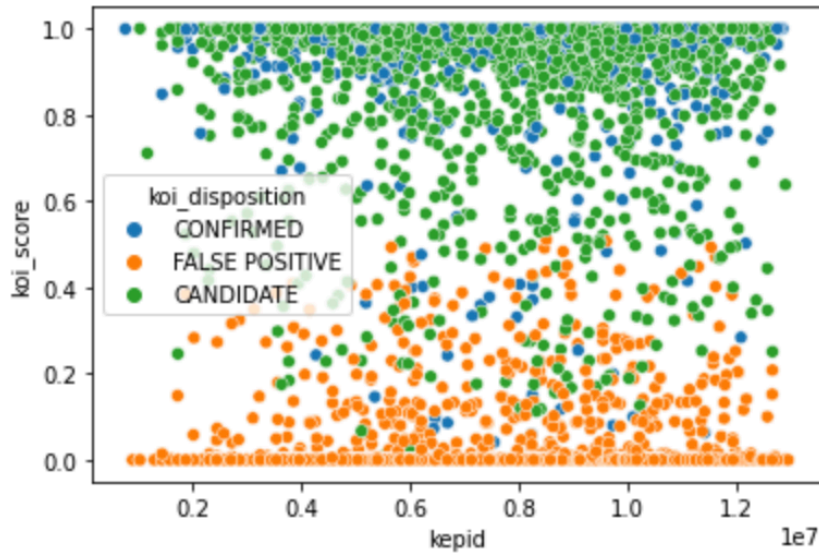
	rowid	kepid	kepoi_name	kepler_name	koi_disposition	koi_pdisposition	koi_score	koi_fpflag_nt	koi_fpflag_ss	koi_fpflag_co
0	1	10797460	K00752.01	Kepler-227 b	CONFIRMED	CANDIDATE	1.000	0	0	0
1	2	10797460	K00752.02	Kepler-227 c	CONFIRMED	CANDIDATE	0.969	0	0	0
2	3	10811496	K00753.01	NaN	FALSE POSITIVE	FALSE POSITIVE	0.000	0	1	0
3	4	10848459	K00754.01	NaN	FALSE POSITIVE	FALSE POSITIVE	0.000	0	1	0
4	5	10854555	K00755.01	Kepler-664 b	CONFIRMED	CANDIDATE	1.000	0	0	0

\*\*\* All columns and rows of data aren't available in this image.

Now, let's see how all the data of the csv file look in **lineplot**:

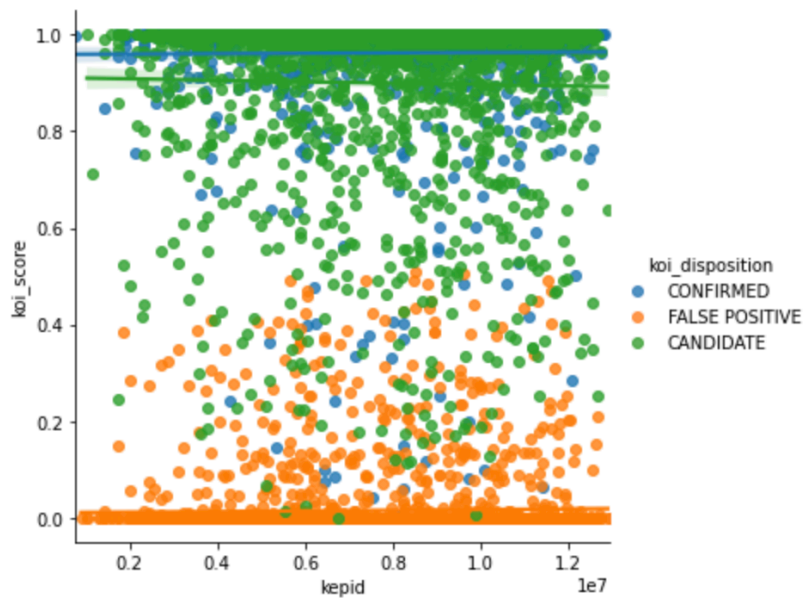


Now, we are going to visualize **CONFIRMED**, **FALSE POSITIVE** and **CANDIDATE** exoplanets in **scatterplot** via **koi\_disposition** :



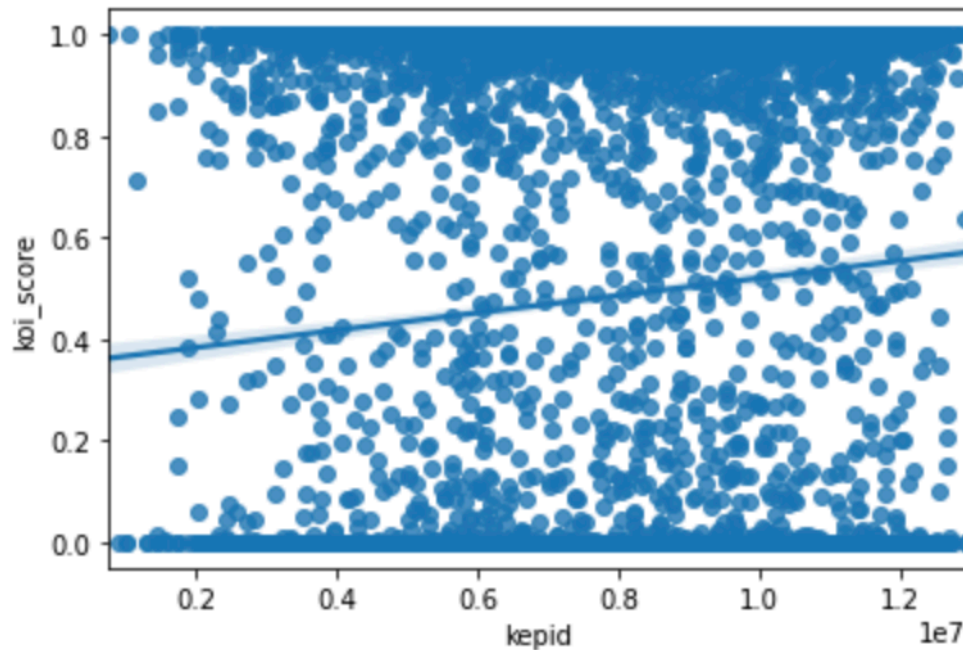
This **scatterplot** reflects that number of **CANDIDATE exoplanets** are higher in number in comparison with FALSE POSITIVE and CONFIRMED exoplanets in this dataset.

Let's take a look at the **lineplot** now :



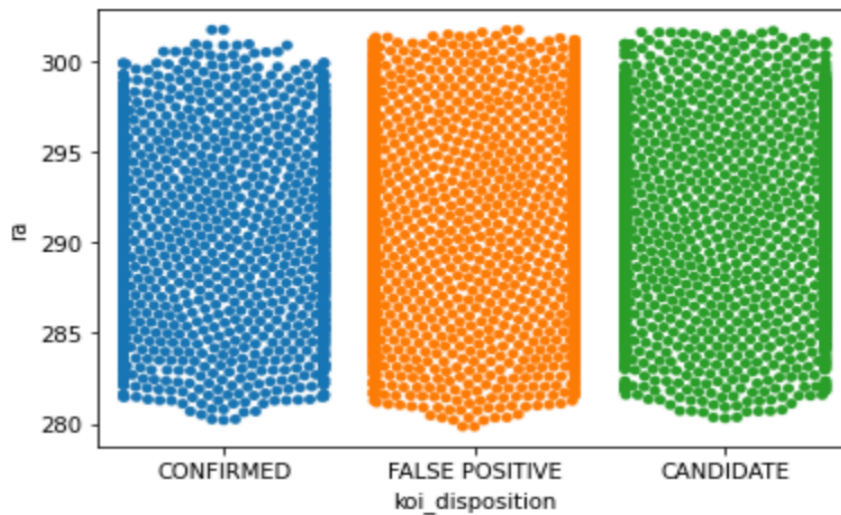
In this lineplot, it's crystal clearly reflected that number of CANDIDATE exoplanets are **higher** in numbers in comparison with FALSE POSITIVE and CONFIRMED exoplanets in this dataset.

Let's try to visualize **koi\_score** and **kepid** via **regression plot** now:



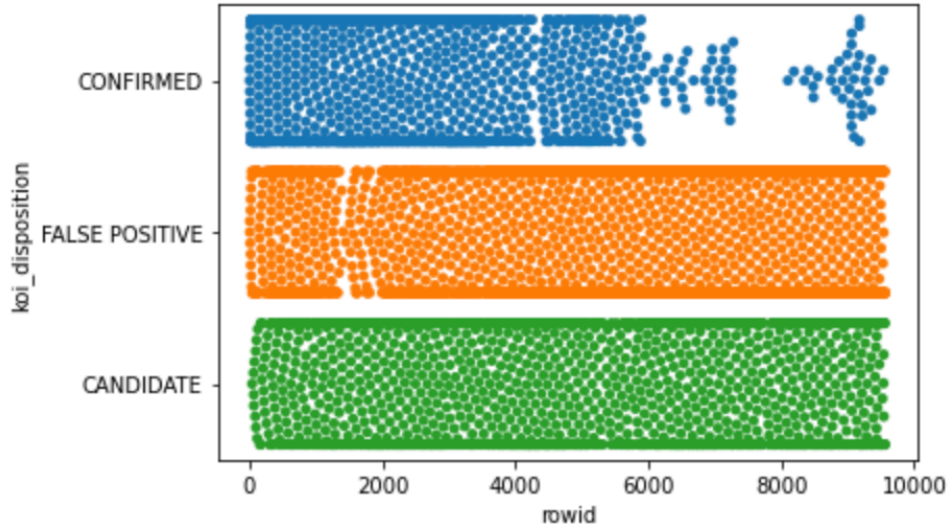
In this **regression plot**, the **koi\_scores** are mostly more than 0.4 and nearby to 1.0

Let's take a look at the **swarmplot** now with respect to **ra** and **koi\_disposition**:



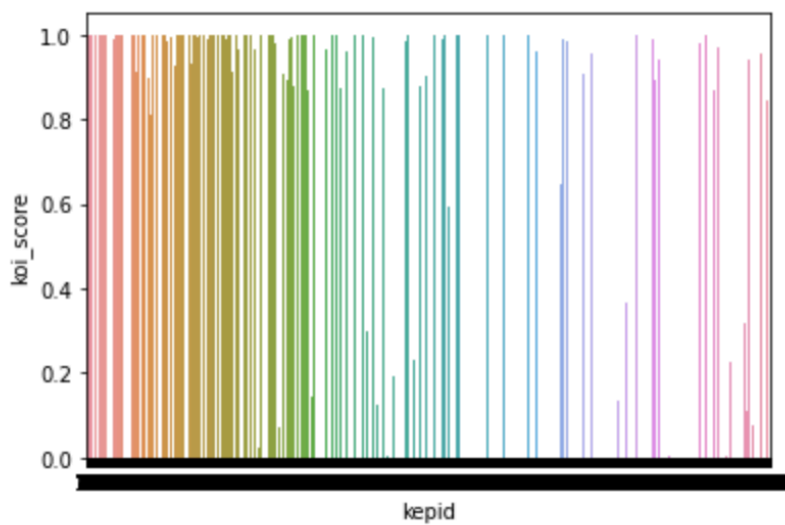
We can see a categorical scatterplot with non-overlapping points of **CANDIDATE** , **FALSE POSITIVE** and **CONFIRMED** exoplanets with respect to **ra** and **koi\_disposition**.

Let's take a look at the **swarmplot** now with respect to koi\_disposition and rowid:



In this **swarmplot**, the number of **CONFIRMED** exoplanets tend to be less (in comparison with others) with rowid more than 6000. .

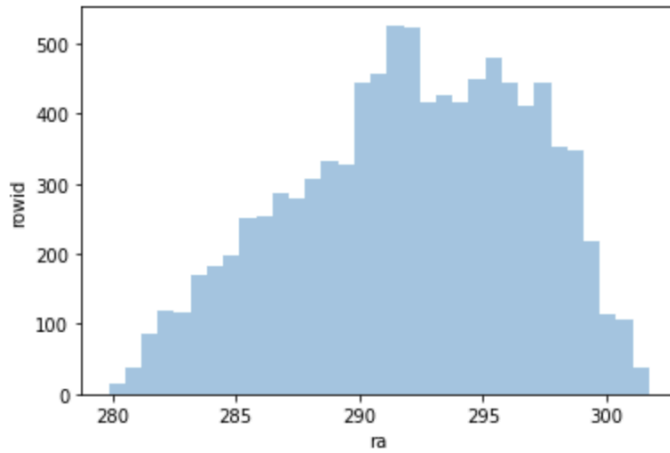
Let's try to visualize Koi\_scores via **barplot**:



In this **barplot**, most of the kepids have koi\_score 1.0

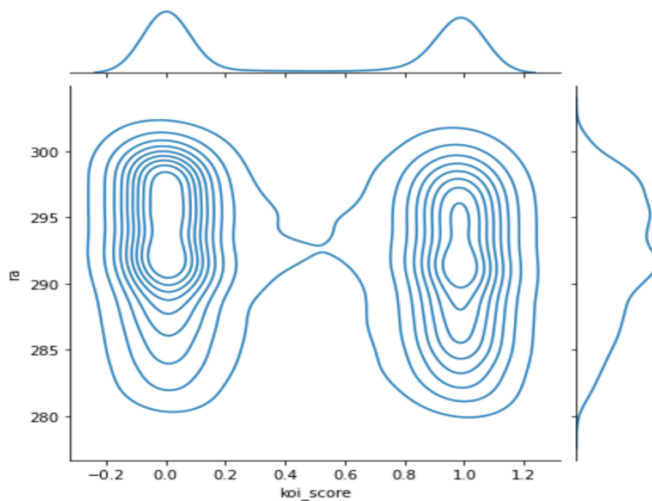


In this below-mentioned histogram, we see that the rowid value is 500 when the ra is ranged between 290-295



In the below-mentioned jointplot,

- the curve at the top of the figure is a KDE plot for the data on the x-axis(Koi\_score) and
- the curve on the right of the figure is a KDE plot for the data on the y-axis(ra)
- **plot** of two variables **with** bivariate and univariate graphs are shown.





## **Methodology:**

All the codes to visualize this data have been shared in github.

## **Acknowledgements:**

This dataset was published as-is by NASA. You can access the original table [here](#).