

## Practicing Data Visualization and Cutting Data I:

1. In Kaggle, go to the following notebook template that has been created for you:  
<https://www.kaggle.com/code/austinhinkel/gettingstartedwithtng50data/notebook>.  
Click the three dots at the top right and select copy and edit notebook. We'll be using data from a cosmological simulation called Illustris TNG-50, which traces the positions of stars and dark matter particles over cosmic time.
2. The first couple of cells should be completed for you. The data will include columns for the X, Y, and Z positions of the particles, as well as a column denoting the particle as a star ('s') or a dark matter particle ('d'). A graph should also be created for you via the third cell of code. Copy and paste this code to also plot x vs z and y vs z.
3. Now let's create a new DataFrame object called starData. Use the following code to keep only the particles identified as stars:

```
starData = data.loc[data.particleType=='s', :]  
#The .loc[] method takes two arguments:  
# The first argument is a cut that keeps only particles that are stars (particleType = 's')  
# The second argument (:) is not changing any of the columns.
```
4. Do something similar for a new DataFrame object called darkMatterData.
5. Create the following plots for the Dark Matter particles (be sure to label your axes!):
  - a. X vs. Y
  - b. Y vs. Z
  - c. X vs. Z
6. Create corresponding plots for the stars. Do your results look different? What can you conclude about the distribution of dark matter and stars? How are they the same? How are they different?
7. Now let's take a closer look at the X vs. Y plot for the stars. Use the plt.xlim([low, high]) and plt.ylim([low, high]) commands to zoom in on the central galaxy (You will have to figure out the numbers to plug in for the low and high values of X and Y). Estimate the diameter of the galaxy and describe what you see.