Project One Analysis

For my project I decided to do data analytics on a NASA exoplanet dataset. This dataset includes information on detected exoplanets from various observatories. For example, some of the information included is: exoplanet mass, atmospheric composition, parent star mass, year discovered, etc. From looking at the dataset I decided to answer three different question:

1. How has the number of exo-planet discovered changed over time?
2. Are all discoveries made from a couple different observatories?
3. Is there a specific discovery method preferred?

To filter the data, I decided to remove all the extra columns of the data that I needed. I ended up keeping around 10 columns. I then decided to add an extra column that had a range of the number of samples I had to use as an x-axis for when I needed to plot.

Chart, scatter chart

Description automatically generatedThe first thing I decided to do was to get a general idea about the planets in the data set like mass and density. The average planet discovered in the data set was a Super Jupiter with 595 Earth masses. The smallest planet was .02 earth masses which would make it smaller than our Earth. After finding the information about the planets I decided to graph the masses and densities of the exoplanets. Unsurprisingly I saw that the mass and density have a positive correlation. I also learned that the data for planet mass/density was actually pretty spread out. Lots of different masses.

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Afterwards, I used different count functions to help answer my original questions. I learned that:

1. 2016 had the most discoveries made, and there is a pattern that more discoveries were made after 2010 than before 2010. The number of discovered exoplanets does increase as the years increase (for the most part)
2. The preferred method of discovering exoplanets was the transit method. The biggest amount of discoveries were made through this method, and the most facilities use this method as well
3. About 80% of the exoplanets discovered in this dataset came from the Kepler mission/telescope
4. The Kepler mission also used the most amount of discovery methods, just like K.W Observatory
5. Kepler only lasted 9 years on the mission and different observatories lasted different times. Not all missions lasted the same amount as I had thought.