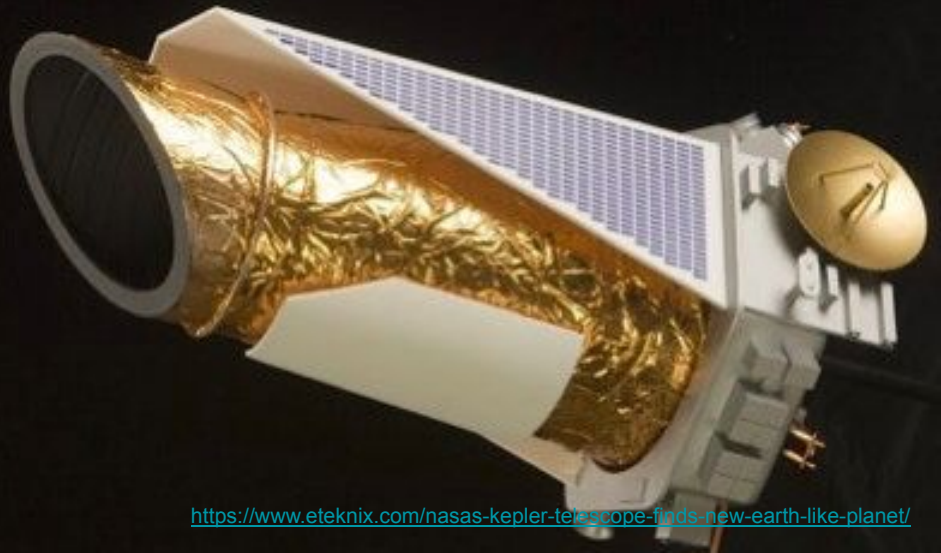
Project: Finding Flaring Stars in Kepler Data

Background:
The Kepler telescope measured the brightness of ~150,000 stars from 2009-2013. While the primary purpose was to find exoplanets, flares were identified around ~4000 stars. Studying this flare activity is of particular interest to scientists so they can better understand how a stars activity might affect the habitability of the planets orbiting it. We were presented with time series data on the brightness of these stars with the intent of identifying the flare activity in them.

Goals:

1. Work towards replicating Davenport's method
2. Develop new methods to identify flares using existing R libraries

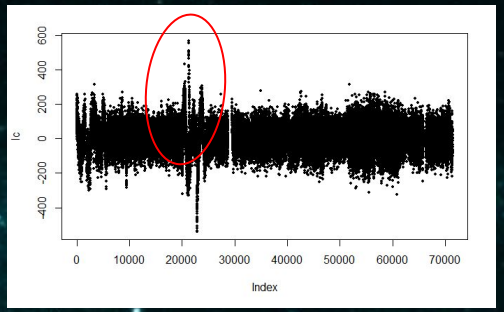


https://www.eteknix.com/nasas-kepler-telescope-finds-new-earth-like-planet/

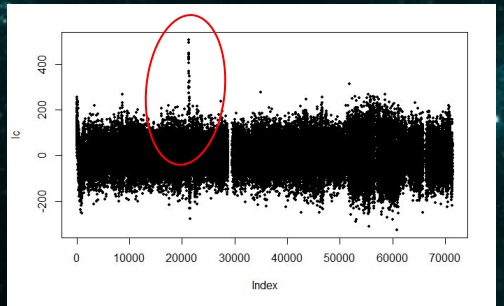
Task 1: Repeating Davenport's Method for Flare Detection

Previous Work:

In 2016, Davenport released his catalogue of over 4000 flaring stars. Since repeatability is an important part of science, our sponsor asked us to reproduce Davenport's original method for flare detection in R (Davenport's was done in Python).



Noise in the signal makes analysis of the flares difficult



Filtering leads to a clearer image of the flare activity

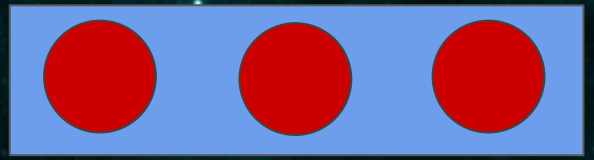
Method:

Light curve

↓ Median Smoothing, Polynomial Fitting

Filtered light curve

↓ Split into continuous segments



↓ More Smoothing, Sine Fitting

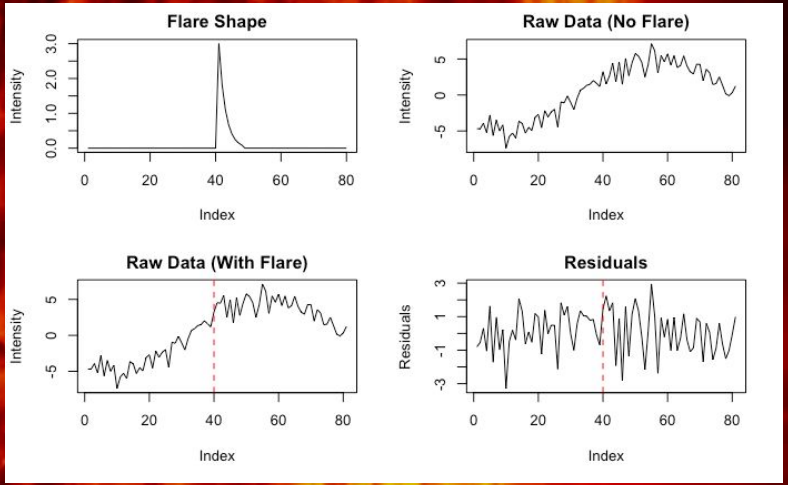
Change point analysis

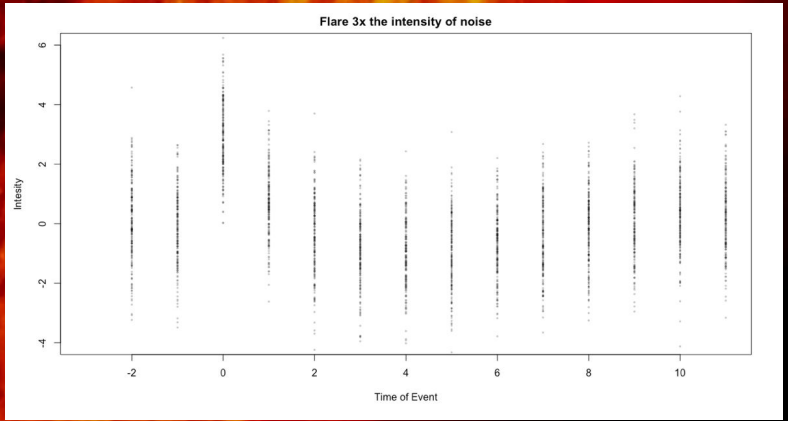
https://www.forbes.com/sites/quora/2017/11/30/how-cold-is-outer-space/#37e501527f11

Task 2: Novel Methods for Flare Detection Using R

Summary:

- Started with ARIMA residuals from previous research.
- Explored several outlier / changepoint detection libraries. With the intent of using dynamic time warping to score identified outliers for similarity to flare events.
- Generated flares in synthetic dataset to identify the shape after ARIMA.
    - In the example in the top right, the flare was 3x the sigma of the added error. After ARIMA it was equal to the variance causing it to appear as random noise. After repeating for 1000 we could generate synthetic profiles of these flare events after ARIMA.
- Proceeded with outlier detection using change point analysis, however the outlier distribution was symmetric with a slightly negative bias which is opposite of what was expected and provided no predictive value for the number of flares.
- More research needs to be done on exactly how these flare profiles project themselves in ARIMA before the residuals cans be used to predict stellar activity.





https://www.princeton.edu/news/2018/07/30/princeton-space-meet-isis-heading-touch-sun