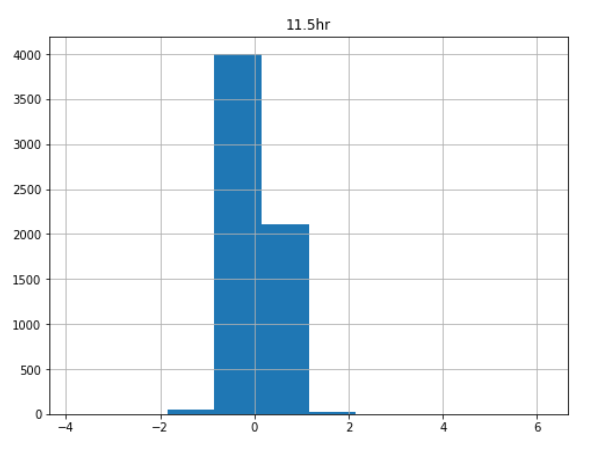
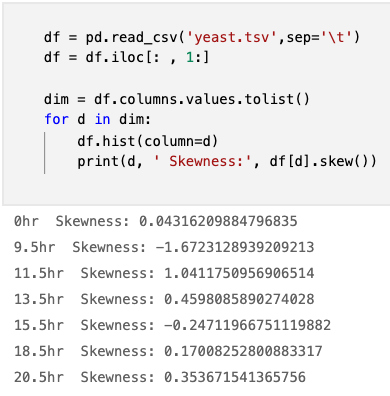
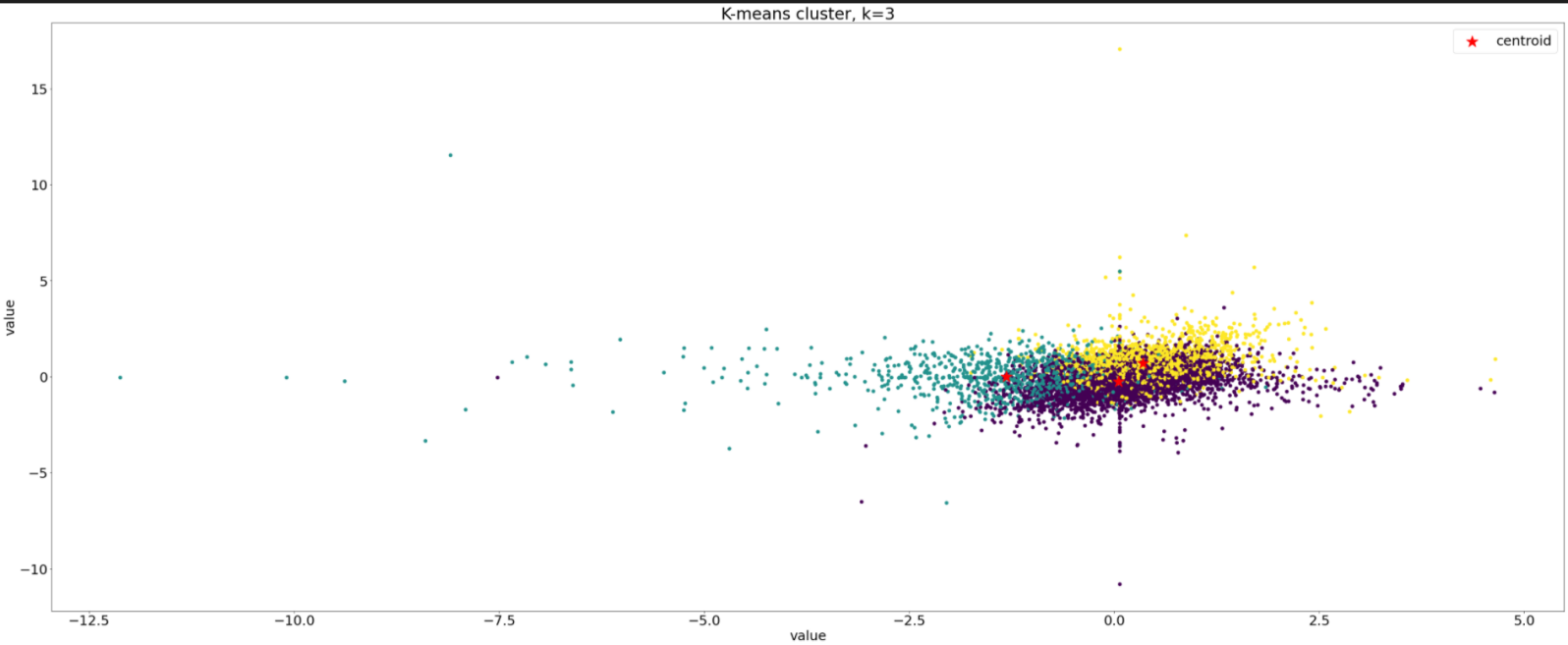
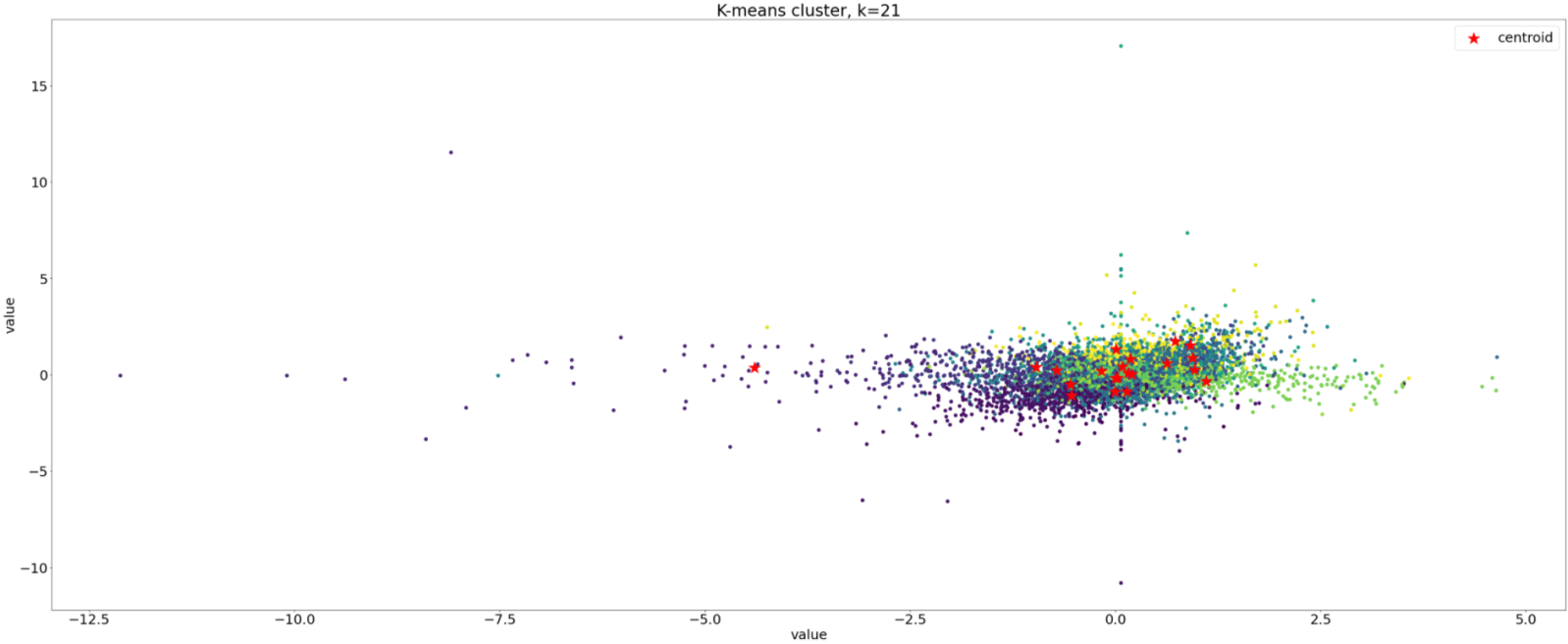
For data preprocessing: As we can see, the data is skewed.

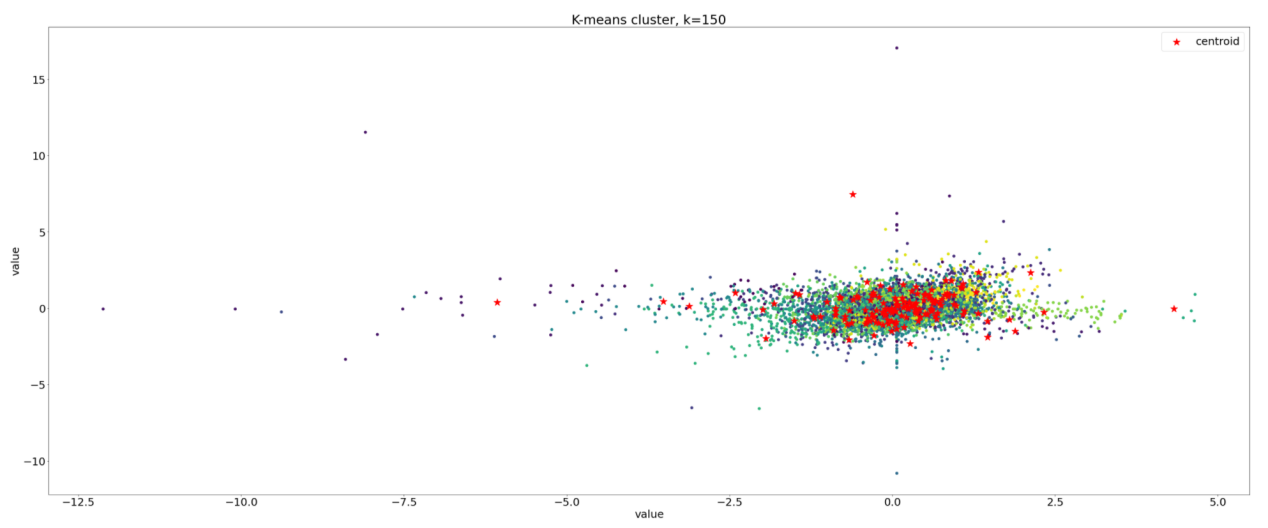


Result: as k increases, outliers tend to get a centroid and form cluster

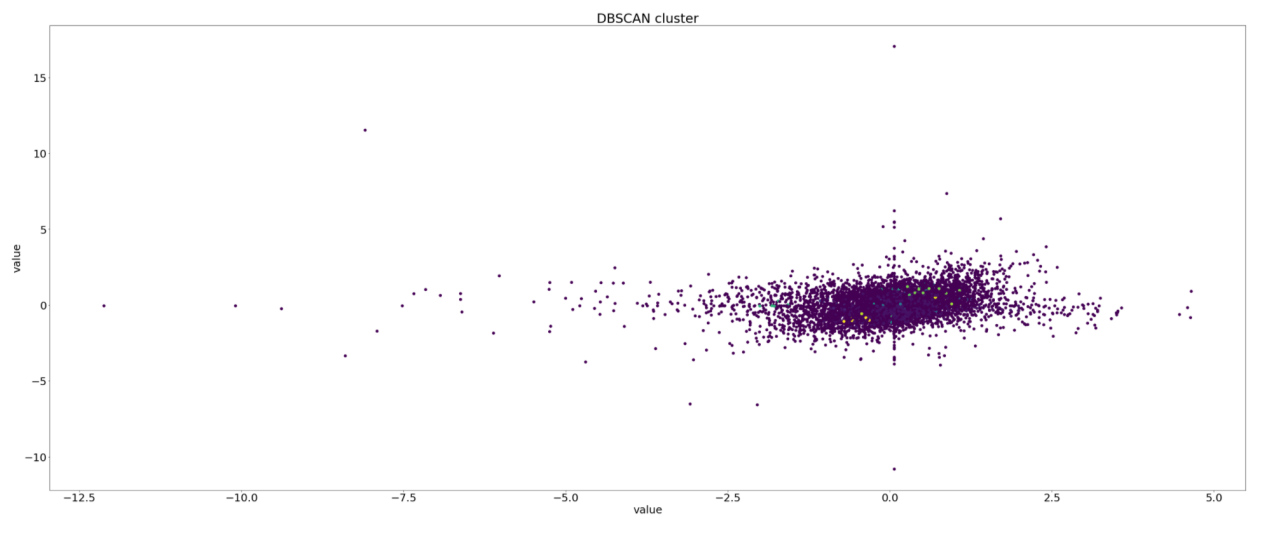




When k = 150



DBSCAN results



PCA visualization

