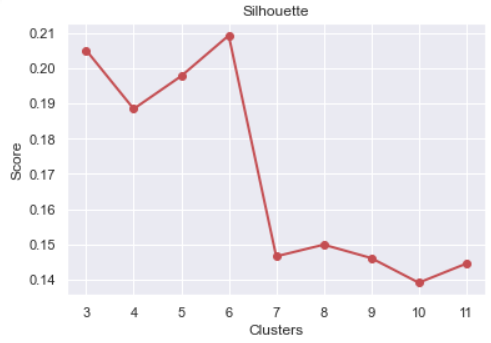
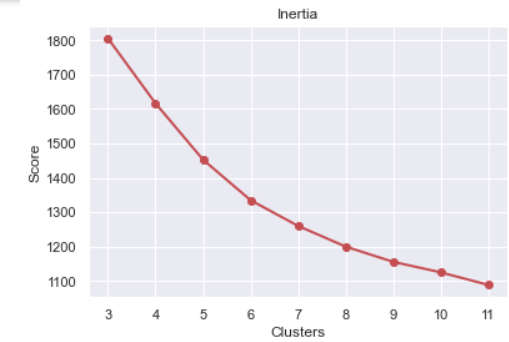
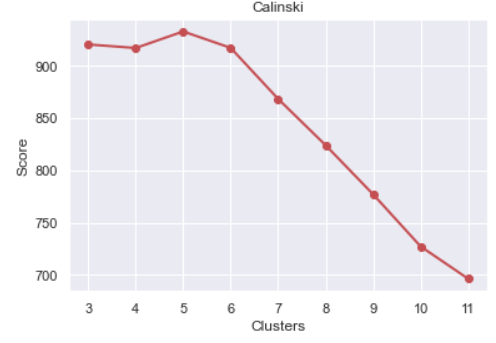
MSDS-422-SEC56 Quick Assignment KMeans Daniel Zaremba

**KMeans Data Observations**





**Notes & Observations**

After changing the random state number a couple times i.e the starting point to see if the clusters were consistent and looking at the above plots I settled on k = 6. The clusters could not be used to predict Loss Amount. Cluster 0 is the riskiest and Cluster 5 is the least risky. Cluster 0 has the highest TRUNC\_IMP\_DEROG (derogatory marks on their credit report) while cluster 5 has the highest TRUNC\_IMP\_YOJ (years on job) which makes sense.

