Slide 1: Pair Trades Definition

Slide 2: Select the universe of securities

We selected a basket of 1,500 liquid US Equities and analyze relative performance over the past 2 years.

We supplement price history with Market Capitalization and Credit Score data to provide additional descriptive parameters for the analysis.

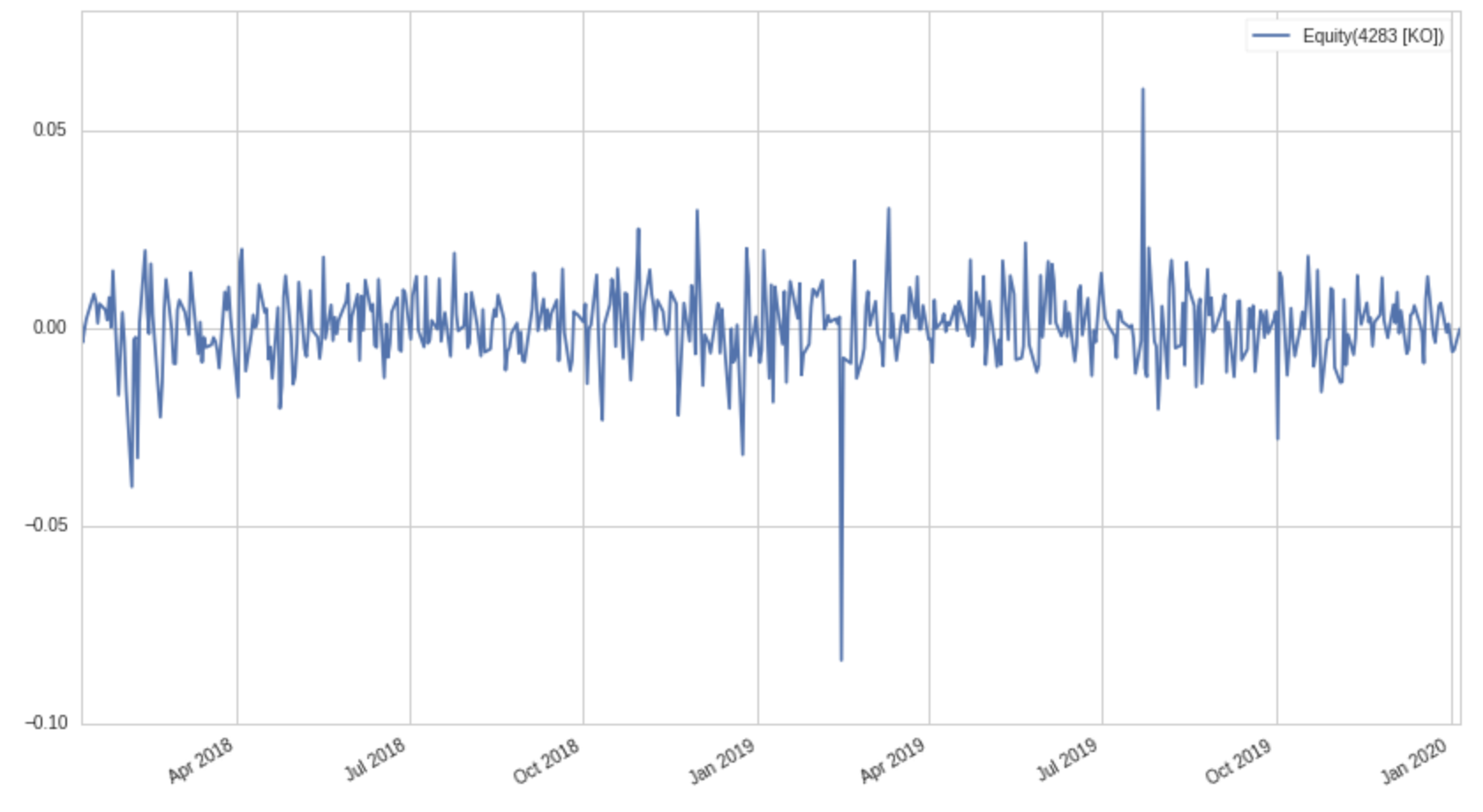
We remove stocks of companies that do not have a single dominant income stream or do not have complete data for the analysis period.

A screenshot of a cell phone

Description automatically generated

Slide 3: Calculate Daily % Returns

Next we calculate Daily % Returns from the Price History of each stock to create a normalized measure of performance.

Daily % Return for Stock KO

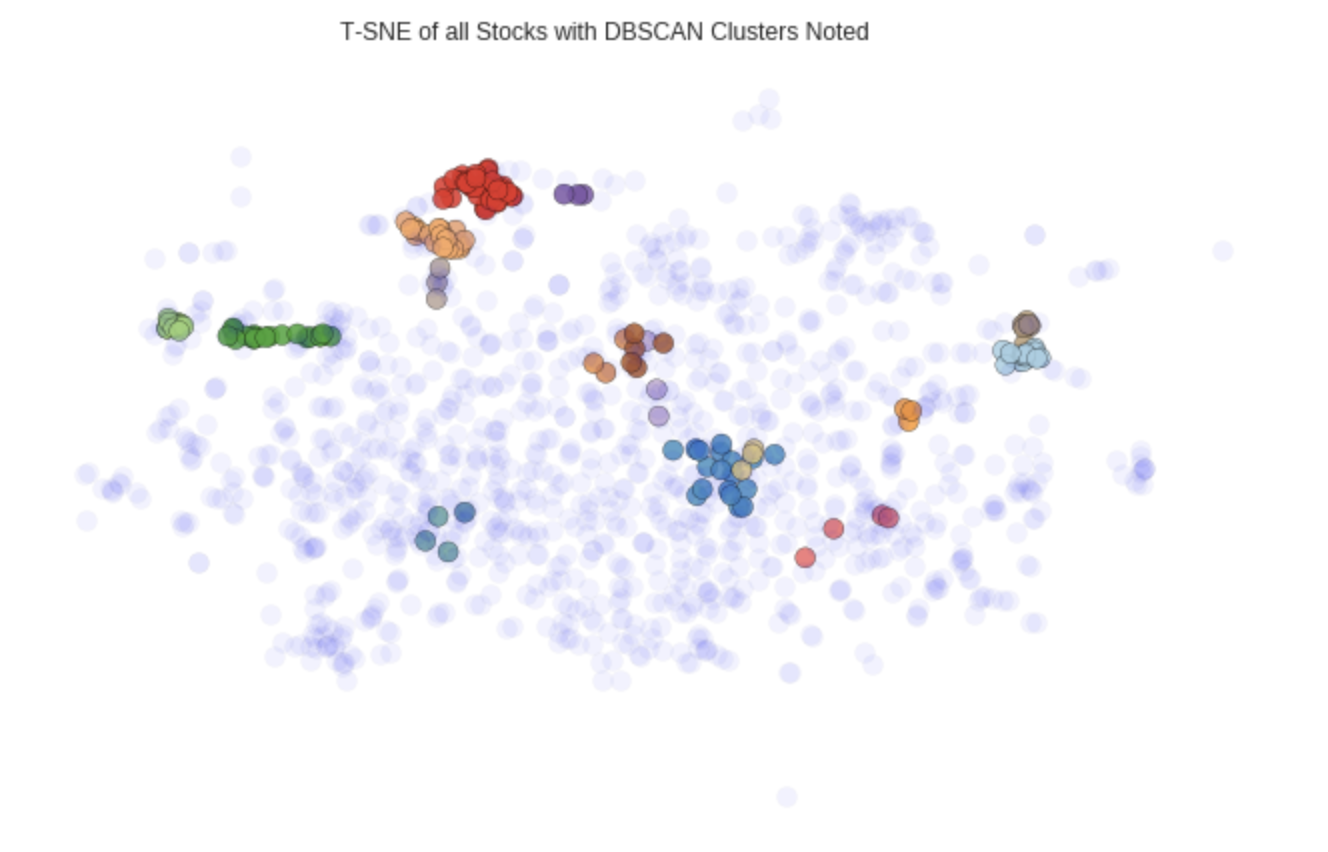
Slide 4: Identify Clusters of related securities

We use the scikit-learn DBSCAN unsupervised machine learning algorithm to identify clusters within the data.

DBSCAN is an alternative to KMeans but has advantages in that it does not require the number of expected clusters to be defined in advance and it ignores stocks that do not neatly fit into clusters.

After data cleaning the original 1,500 securities is reduced to 1,364. There are 929,566 pair trades possible and 16 clusters are detected.

We use T-SNE to visualize the high dimension data into 2 dimensions.



Slide 5: Cluster Analysis

We examine the number of stocks identified in each cluster. Cluster 9 includes 3 securities,   
including KO (Coca Cola), PEP (Pepsi).

A close up of a logo

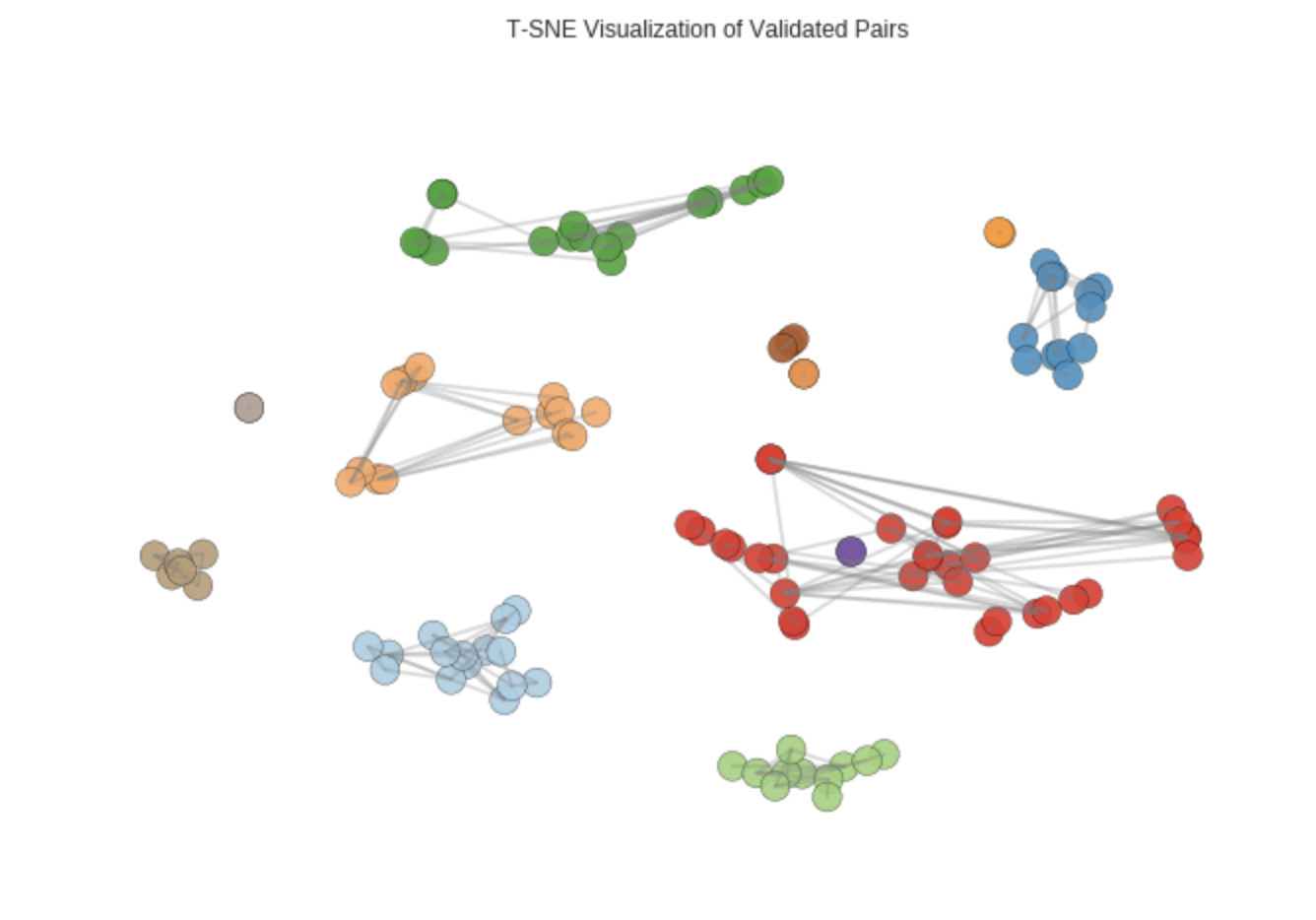
Description automatically generated

We observe that Proctor & Gamble (Red) is currently outperforming relative to Coca Cola (Blue) and the analysis indicates this relationship is likely to revert to historic mean.

A close up of a map

Description automatically generated

Slide 6: Finally we evaluate the clusters for Cointegrated Pairs and reduce the original 900K possible trades down to 143.

A screenshot of a cell phone

Description automatically generated