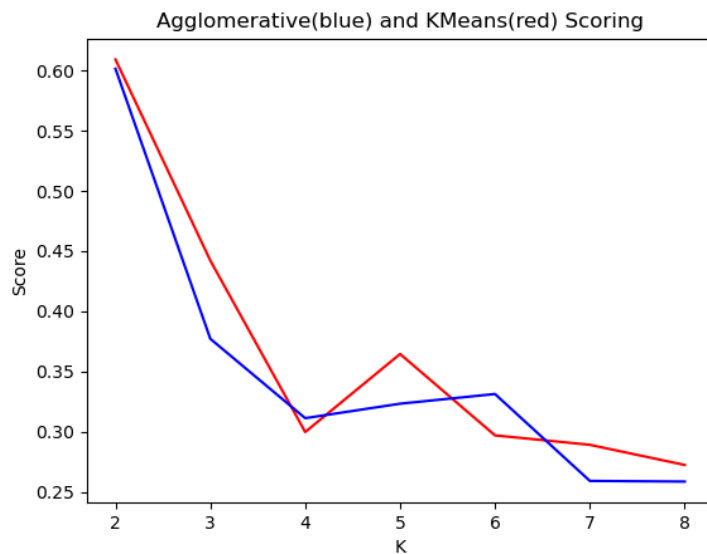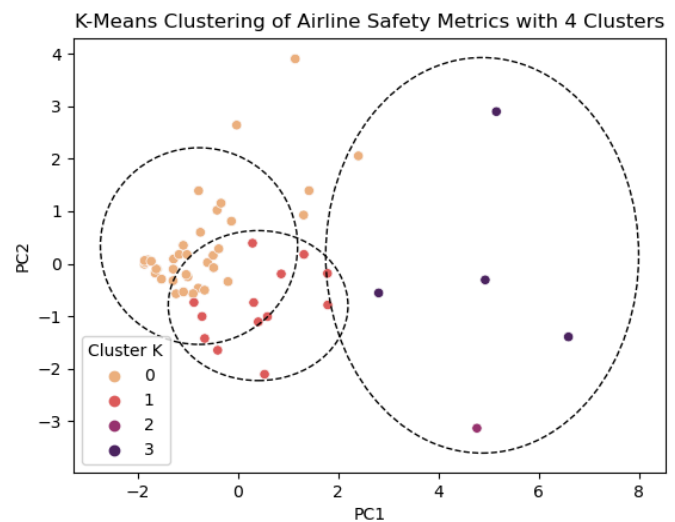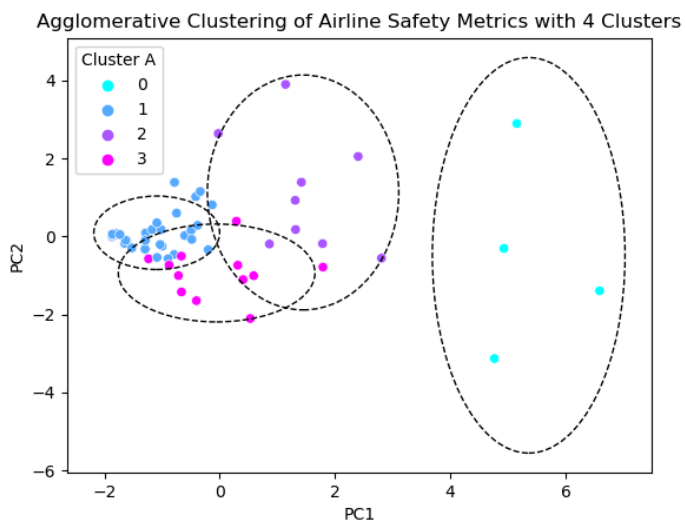# Difference Between Agglomerative and KMeans Approach

## Scoring

To determine which clusters to use for each model, evaluation was done using the elbow method.



From this visualization, we can see that both models begin to perform worse when K = 4. Using this, both Agglomerative and KMeans were built with 4 clusters.

## Model Visualizations



(Note that KMeans only has 3 'groupings' as cluster 2 has a single element)

## KMeans Results

| | Cluster K | incidents_85_99 | fatal_accidents_85_99 | fatalities_85_99 | incidents_00_14 | fatal_accidents_00_14 | fatalities_00_14 | incidents_mean | fatal_mean | fatalities_mean |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 3.605263 | 0.947368 | 26.842105 | 2.947368 | 0.447368 | 49.052632 | 3.276316 | 0.697368 | 37.947368 |
| 1 | 1 | 8.384615 | 3.076923 | 315.076923 | 3.615385 | 0.769231 | 42.923077 | 6.000000 | 1.923077 | 179.000000 |
| 2 | 2 | 76.000000 | 14.000000 | 128.000000 | 6.000000 | 1.000000 | 88.000000 | 41.000000 | 7.500000 | 108.000000 |
| 3 | 3 | 20.000000 | 8.000000 | 262.750000 | 16.500000 | 2.250000 | 149.750000 | 18.250000 | 5.125000 | 206.250000 |

Safest: Cluster 0

Most Dangerous: Cluster 3

## Agglomerative Results

| | Cluster A | incidents_85_99 | fatal_accidents_85_99 | fatalities_85_99 | incidents_00_14 | fatal_accidents_00_14 | fatalities_00_14 | incidents_mean | fatal_mean | fatalities_mean |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 35.000000 | 9.750000 | 238.750000 | 15.250000 | 2.000000 | 166.000000 | 25.125000 | 5.875000 | 202.375000 |
| 1 | 1 | 3.096774 | 0.741935 | 14.129032 | 2.580645 | 0.225806 | 14.032258 | 2.838710 | 0.483871 | 14.080645 |
| 2 | 2 | 10.444444 | 3.222222 | 128.888889 | 6.222222 | 2.000000 | 179.111111 | 8.333333 | 2.611111 | 154.000000 |
| 3 | 3 | 6.000000 | 2.583333 | 311.833333 | 2.833333 | 0.333333 | 33.166667 | 4.416667 | 1.458333 | 172.500000 |

Safest: Cluster 1

Most Dangerous: Cluster 0

Using the above clusters, we can classify the 'safest' airlines by the airlines that are present in both KMeans Cluster 0 and Agglomerative Cluster 1. These airlines are:

'Condor', 'TACA', 'Qantas*', 'Philippine Airlines', 'Sri Lankan / AirLanka', 'Aeromexico*', 'SAS*', 'Aerolineas Argentinas', 'Finnair', 'Virgin Atlantic', 'Xiamen Airlines', 'Singapore Airlines', 'Cathay Pacific*', 'Hawaiian Airlines', 'Alitalia', 'Lufthansa*', 'Air New Zealand*', 'Air Canada', 'El Al', 'Royal Air Maroc', 'British Airways*', 'Alaska Airlines*', 'LAN Airlines', 'Gulf Air', 'Southwest Airlines', 'All Nippon Airways', 'COPA', 'Aer Lingus', 'Austrian Airlines', 'KLM*', 'TAP - Air Portugal'

Similarly, we can classify the 'most dangerous' airlines as those present in both KMeans Cluster 3 and Agglomerative 0. These are:

'United / Continental*', 'American*', 'Delta / Northwest*'