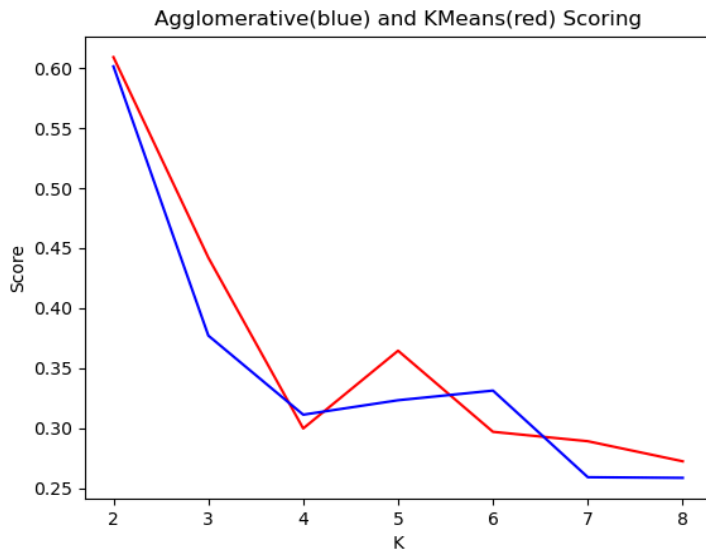


Difference Between Agglomerative and KMeans Approach

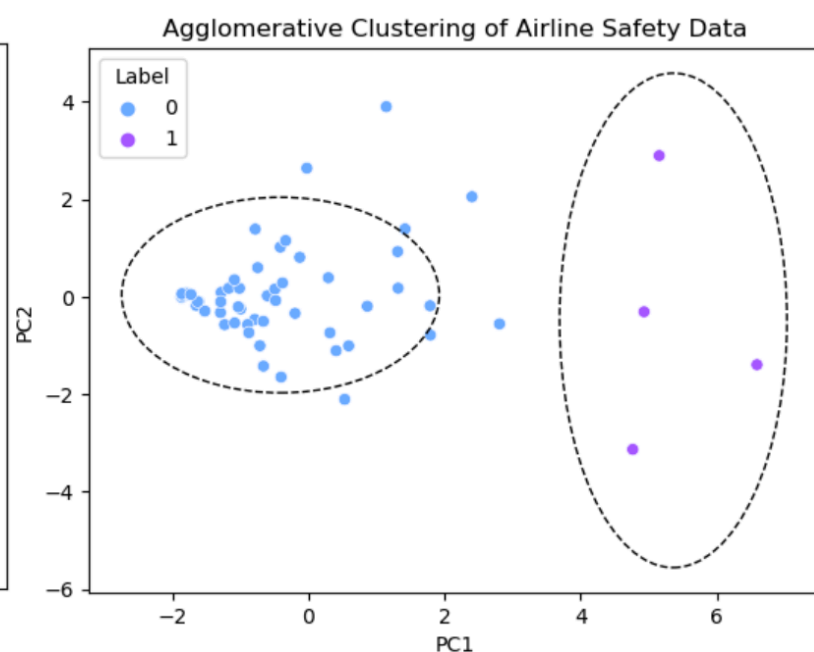
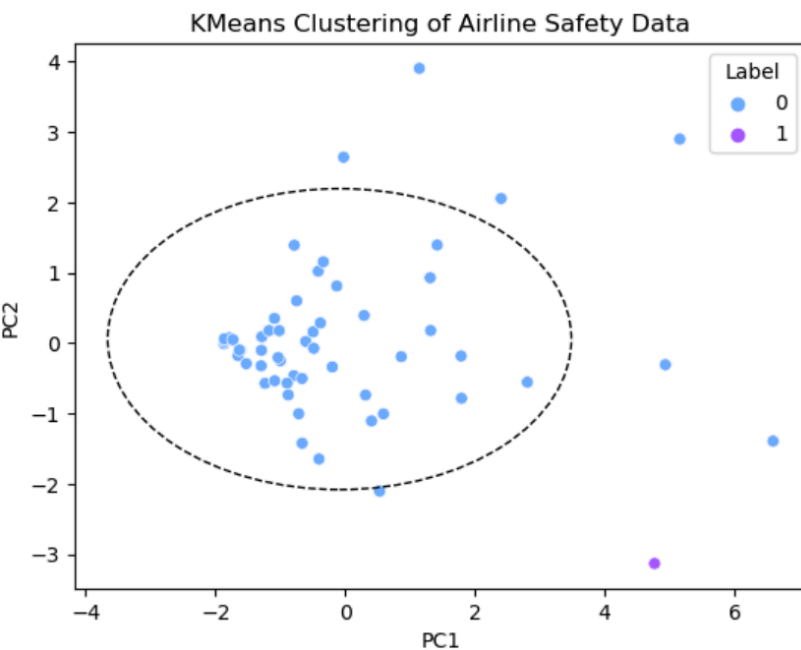
Scoring

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 perform the best when $k = 2$. This seems to be intuitive as it could cluster the data into 'safe' and 'dangerous'.

Model Visualizations



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	5.927273	1.963636	112.127273	4.090909	0.654545	54.927273	5.009091	1.309091	83.527273
1	1	76.000000	14.000000	128.000000	6.000000	1.000000	88.000000	41.000000	7.500000	108.000000

Safest: Cluster 1

Most Dangerous: Cluster 0

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	5.038462	1.596154	102.692308	3.269231	0.557692	47.019231	4.153846	1.076923	74.855769
1	1	35.000000	9.750000	238.750000	15.250000	2.000000	166.000000	25.125000	5.875000	202.375000

Safest: Cluster 0

Most Dangerous: Cluster 1

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

'Ethiopian Airlines', 'Aeromexico*', 'US Airways / America West*', 'Saudi Arabian', 'Lufthansa*', 'TAM', 'Xiamen Airlines', 'All Nippon Airways', 'Egyptair', 'Southwest Airlines', 'South African', 'Kenya Airways', 'Air India*', 'TACA', 'Pakistan International', 'Turkish Airlines', 'Alitalia', 'Air New Zealand*', 'LAN Airlines', 'SWISS*', 'Austrian Airlines', 'Air France', 'Finnair', 'Philippine Airlines', 'Garuda Indonesia', 'El Al', 'KLM*', 'Hawaiian Airlines', 'Aerolineas Argentinas', 'SAS*', 'Cathay Pacific*', 'Japan Airlines', 'British Airways*', 'Condor', 'TAP - Air Portugal', 'Thai Airways', 'Alaska Airlines*', 'Singapore Airlines', 'Qantas*', 'Iberia', 'Royal Air Maroc', 'Virgin Atlantic', 'Gulf Air', 'China Airlines', 'Air Canada', 'Korean Air', 'Sri Lankan / AirLanka', 'Aer Lingus', 'COPA', 'Malaysia Airlines', 'Vietnam Airlines', 'Avianca'

Similarly, we can classify the 'most dangerous' airlines as those present in both KMeans Cluster 1 and Agglomerative 1. These are:

'Aeroflot*'

