Section - A

A) 1) Initially there are 15 X 15 X 4 dimensions in the input image. The first convolution layer has 5 X 5 X 4 X 1, Padding = 1, Stride = 1. The output layer is ((15 - 5 + 2*1)/1 + 1 X (15 - 5 + 2*1)/1 + 1 X 1) which gives the output (13 X 13 X 1).

When we pass it through the max pooling layer Kernel = 6 X 6 X 1, stride = 2, the output dimensions are ((13 - 6)/2 + 1 X (13 - 6)/2 + 1, 1), which gives (6 X 6 X 1).

The final convolution layer has the following dimensions, Kernel = 5 X 4 X 3 X 1 , Padding = 2, Stride = 2. Its output layer is ((6 - 5 + 2*2)/2 + 1 X (6 - 4 + 2*2)/2 + 1 X 1) which gives the output (3 X 4 X 1).

2) Pooling is used to reduce the dimensionality of the input feature maps. It also keeps the model generalizable by preventing overfitting of data, and prevents models from learning too specific details of the data in a region and instead encourages the CNN to retain the most important features. It also adds translation invariance to the network.

3) In the first convolution layer, There are 160 learnable parameters. In the first convolution layer, there are 5 X5 X 4 X 1 = 100 parameters, in the max pool layer there are no learnable parameters, in the final convolution layer there are 5 X3 X4 X 1 = 60. Since there are no biases, there are 160 total learnable parameters.

B) No, it is not possible for the k means algorithm to revisit a previous configuration. This is because, after each iteration, the inertia, that is the sum of squares of distances of each sample from its centroid decreases. After a certain number of iterations, it is minimized thus converging the algorithm. To go back to a previous configuration would mean an increase in its inertia, which is not possible. If it

remains exactly the same, that means the algorithm has converged, thus proving why it must converge in a finite number of iterations.

C) Yes, a neural network can be used to model K nearest neighbors. The number of neurons in the hidden layer is equal to the number of features, the number of neurons in the output layer is equal to the number of classes to denote the probability of each class. We can use one hidden layer. The activation function to use will be softmax for multiclass classification.

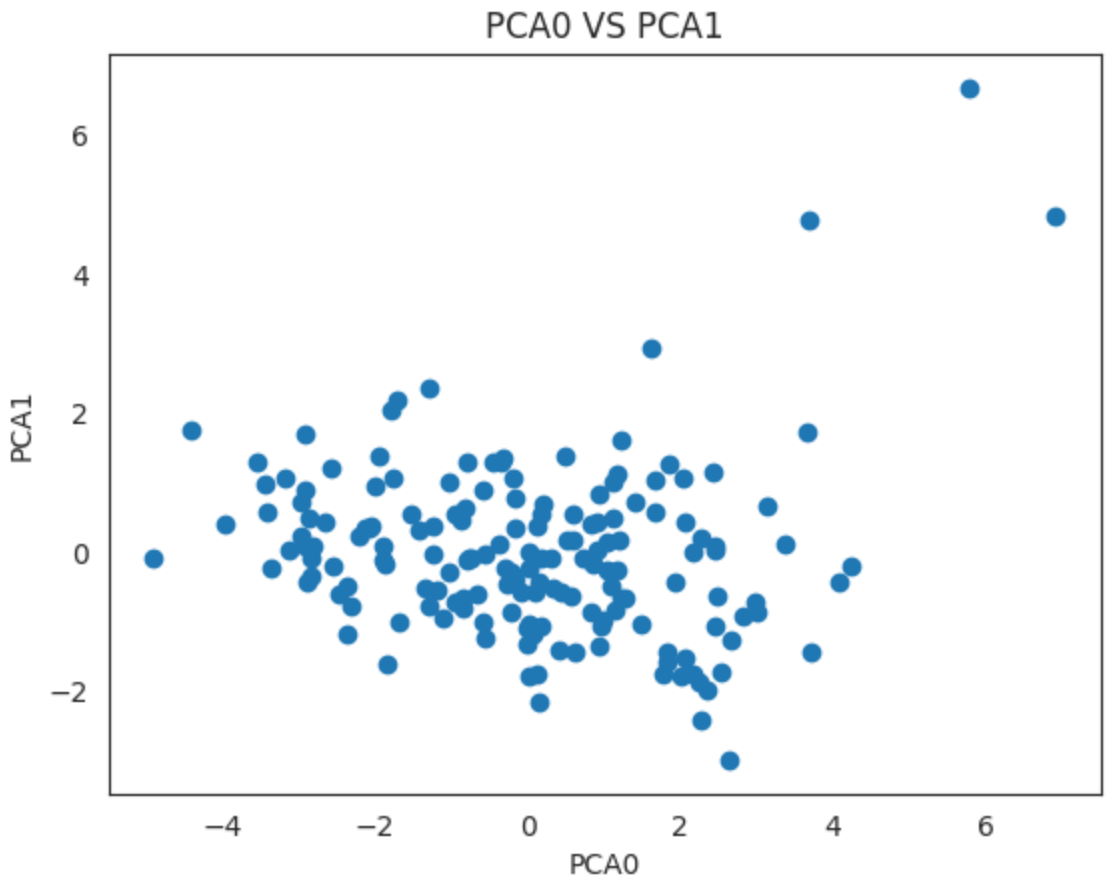
D) Linear kernels simply perform linear operations on the input data, which can be represented as $O=\sum(W \cdot X)+b$, where O is the output, X is the input, W is the weight and b is the bias. Non linear kernels add an activation function on top of this, like for example ReLU or sigmoid. This makes the non linear kernel more complex and allows the model to learn more complex patterns, while the linear kernels can allow the model to study only linear relationships between weights, input and the output.

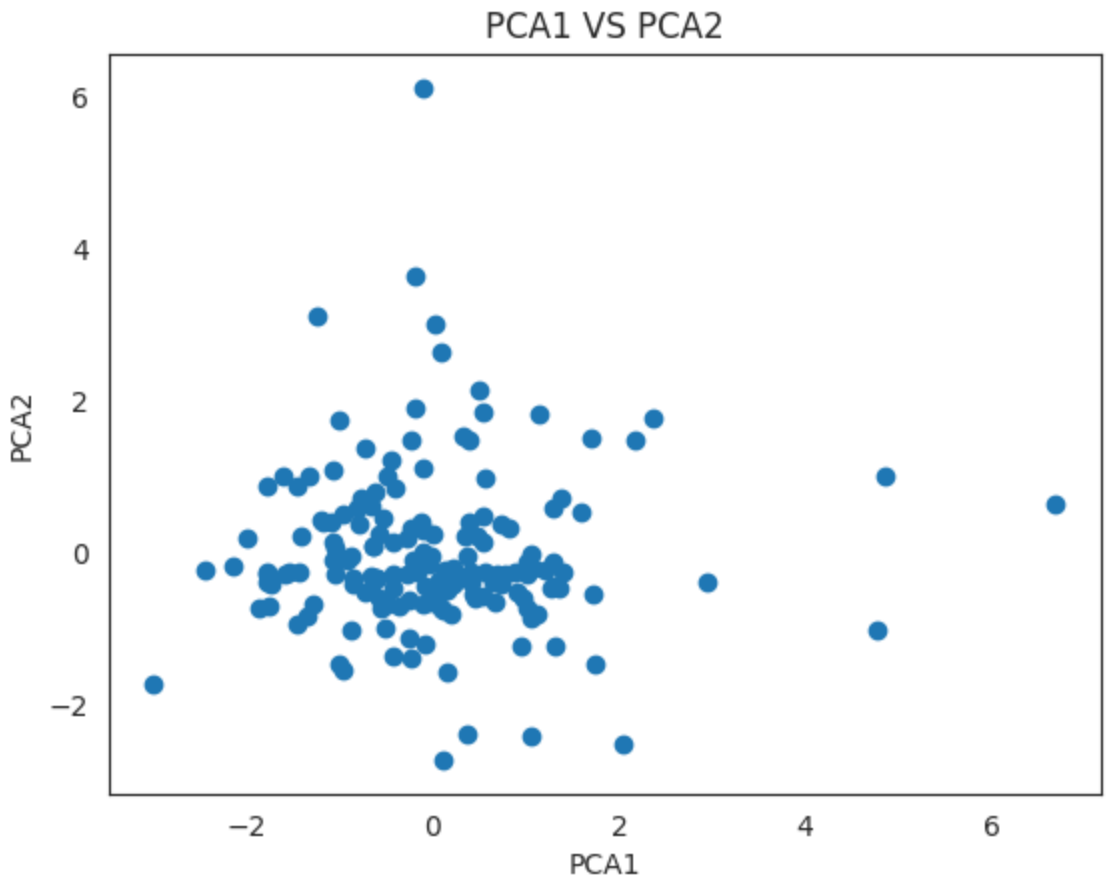
Section - C

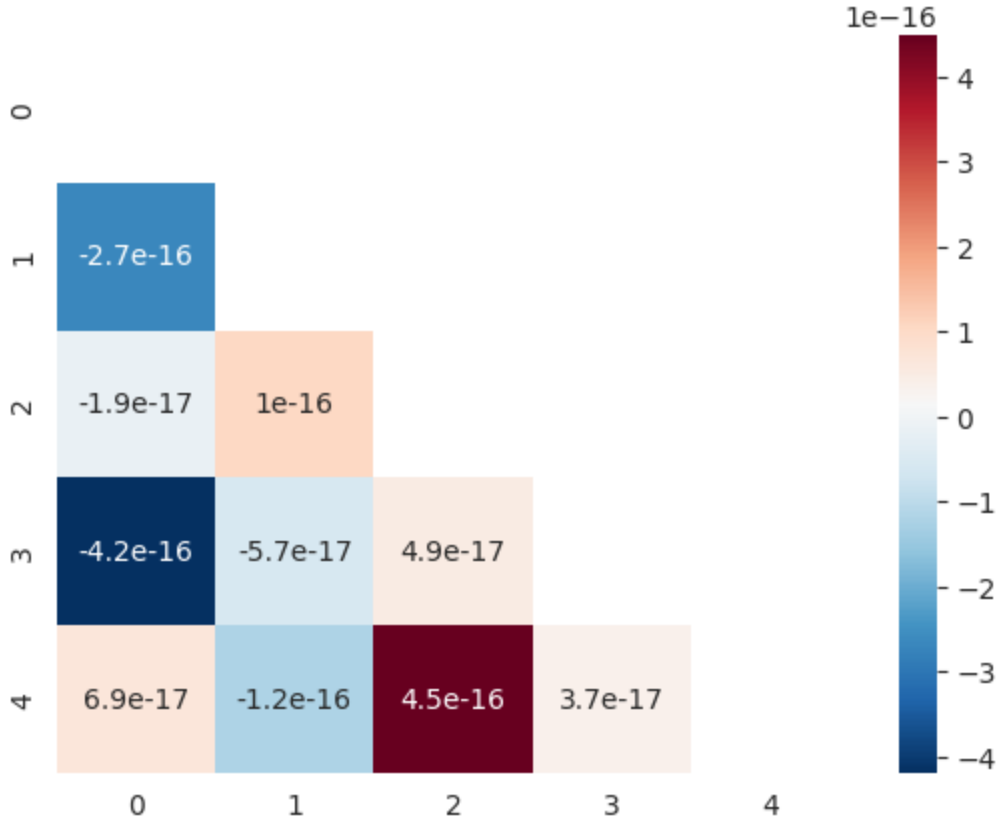
1) Here is the result of the exploratory analysis.

#	Column	Non-Null Count	Dtype
0	country	167 non-null	object
1	child_mort	167 non-null	float64
2	exports	167 non-null	float64
3	health	167 non-null	float64
4	imports	167 non-null	float64
5	income	167 non-null	int64
6	inflation	167 non-null	float64
7	life_expec	167 non-null	float64
8	total_fer	167 non-null	float64
9	gdpp	167 non-null	int64

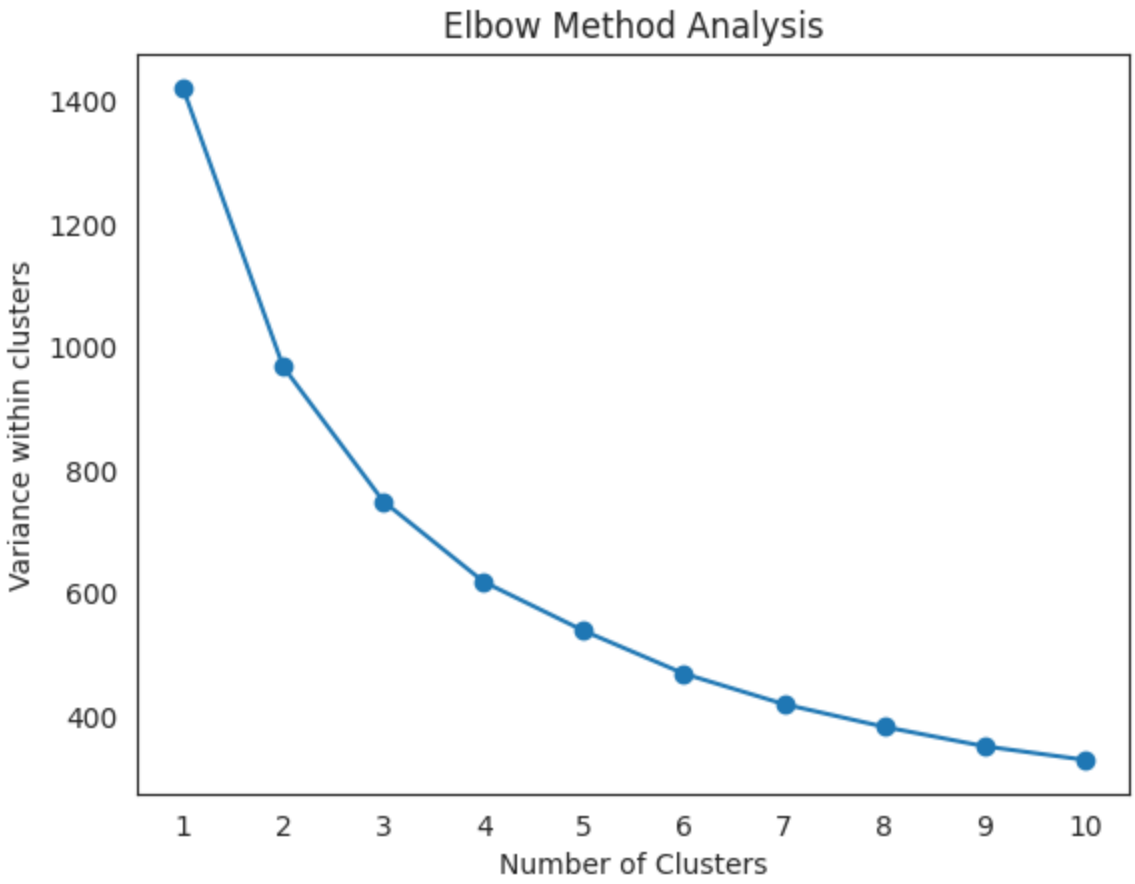
2) We decided on 5 features as optimal number of components after performing PCA as it retains around 94.53% of the variance.

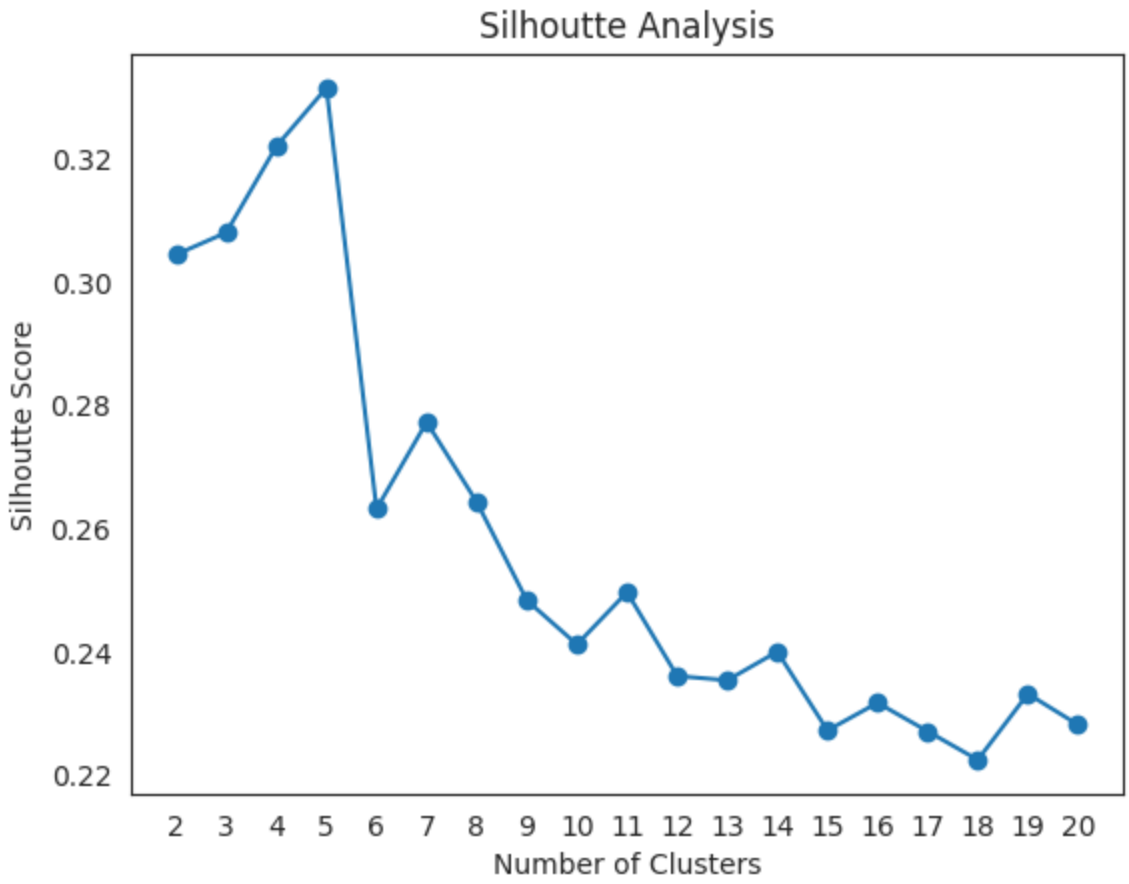

PCA0 VS PCA1
PCA1
PCA0
6
4
2
0
−2
−4
−2
0
2
4
6


PCA1 VS PCA2
PCA2
PCA1
6
4
2
0
−2
−2
0
2
4
6



3) Elbow analysis and Silhouette analysis narrowed the number of clusters to either 3 or 4. I chose 3 as the number of clusters so keep the model easy to interpret.

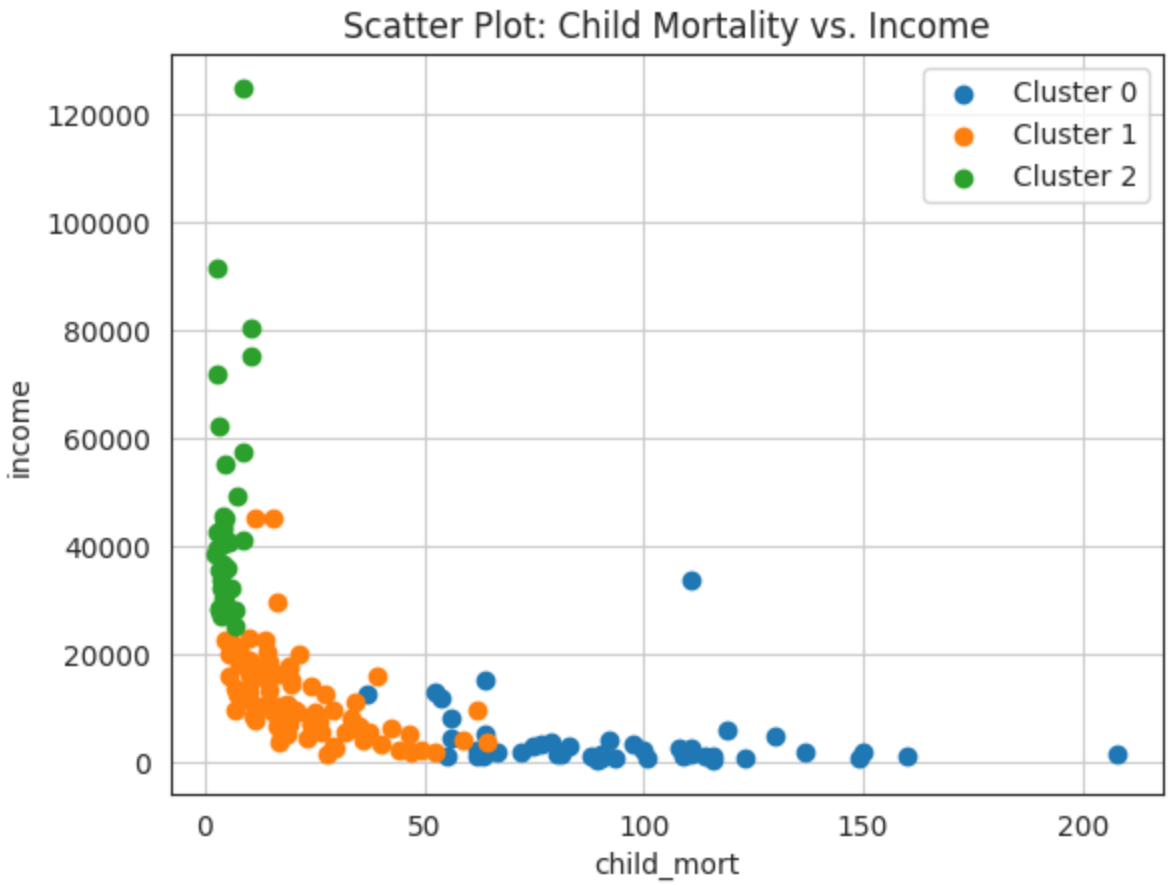

Elbow Method Analysis
Variance within clusters
1400
1200
1000
800
600
400
1
2
3
4
5
6
7
8
9
10
Number of Clusters



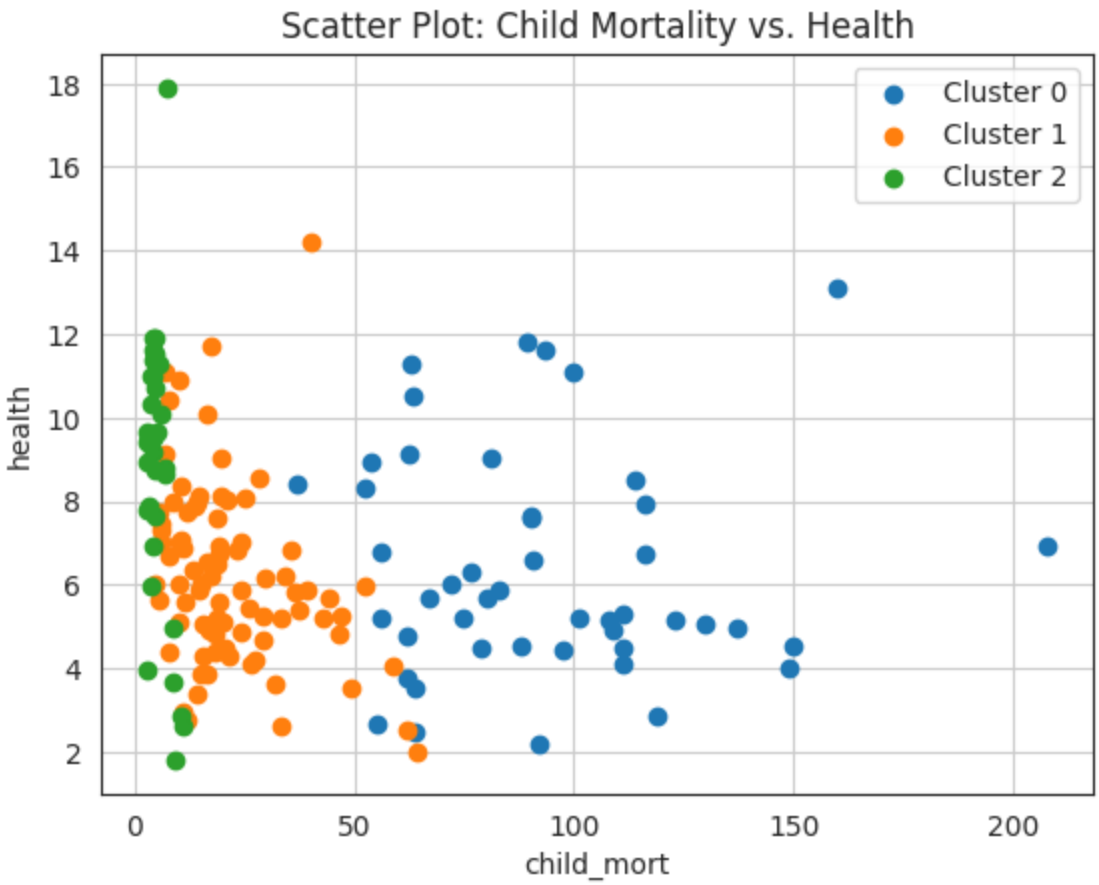
4) These are the clusters that we formed

```
Cluster 0
['Afghanistan' 'Angola' 'Benin' 'Botswana' 'Burkina Faso' 'Burundi'
 'Cameroon' 'Central African Republic' 'Chad' 'Comoros' 'Congo, Dem. Rep.'
 'Congo, Rep.' "Cote d'Ivoire" 'Equatorial Guinea' 'Eritrea' 'Gabon'
 'Gambia' 'Ghana' 'Guinea' 'Guinea-Bissau' 'Haiti' 'Iraq' 'Kenya'
 'Kiribati' 'Lao' 'Lesotho' 'Liberia' 'Madagascar' 'Malawi' 'Mali'
 'Mauritania' 'Mozambique' 'Namibia' 'Niger' 'Nigeria' 'Pakistan' 'Rwanda'
 'Senegal' 'Sierra Leone' 'South Africa' 'Sudan' 'Tanzania' 'Timor-Leste'
 'Togo' 'Uganda' 'Yemen' 'Zambia']
Cluster 1
['Albania' 'Algeria' 'Antigua and Barbuda' 'Argentina' 'Armenia'
 'Azerbaijan' 'Bahamas' 'Bangladesh' 'Barbados' 'Belarus' 'Belize'
 'Bhutan' 'Bolivia' 'Bosnia and Herzegovina' 'Brazil' 'Bulgaria'
 'Cambodia' 'Cape Verde' 'Chile' 'China' 'Colombia' 'Costa Rica' 'Croatia'
 'Dominican Republic' 'Ecuador' 'Egypt' 'El Salvador' 'Estonia' 'Fiji'
 'Georgia' 'Grenada' 'Guatemala' 'Guyana' 'Hungary' 'India' 'Indonesia'
 'Iran' 'Jamaica' 'Jordan' 'Kazakhstan' 'Kyrgyz Republic' 'Latvia'
 'Lebanon' 'Libya' 'Lithuania' 'Macedonia, FYR' 'Malaysia' 'Maldives'
 'Mauritius' 'Micronesia, Fed. Sts.' 'Moldova' 'Mongolia' 'Montenegro'
 'Morocco' 'Myanmar' 'Nepal' 'Oman' 'Panama' 'Paraguay' 'Peru'
 'Philippines' 'Poland' 'Romania' 'Russia' 'Samoa' 'Saudi Arabia' 'Serbia'
 'Seychelles' 'Solomon Islands' 'Sri Lanka'
 'St. Vincent and the Grenadines' 'Suriname' 'Tajikistan' 'Thailand'
 'Tonga' 'Tunisia' 'Turkey' 'Turkmenistan' 'Ukraine' 'Uruguay'
 'Uzbekistan' 'Vanuatu' 'Venezuela' 'Vietnam']
Cluster 2
['Australia' 'Austria' 'Bahrain' 'Belgium' 'Brunei' 'Canada' 'Cyprus'
 'Czech Republic' 'Denmark' 'Finland' 'France' 'Germany' 'Greece'
 'Iceland' 'Ireland' 'Israel' 'Italy' 'Japan' 'Kuwait' 'Luxembourg'
 'Malta' 'Netherlands' 'New Zealand' 'Norway' 'Portugal' 'Qatar'
 'Singapore' 'Slovak Republic' 'Slovenia' 'South Korea' 'Spain' 'Sweden'
 'Switzerland' 'United Arab Emirates' 'United Kingdom' 'United States']
```

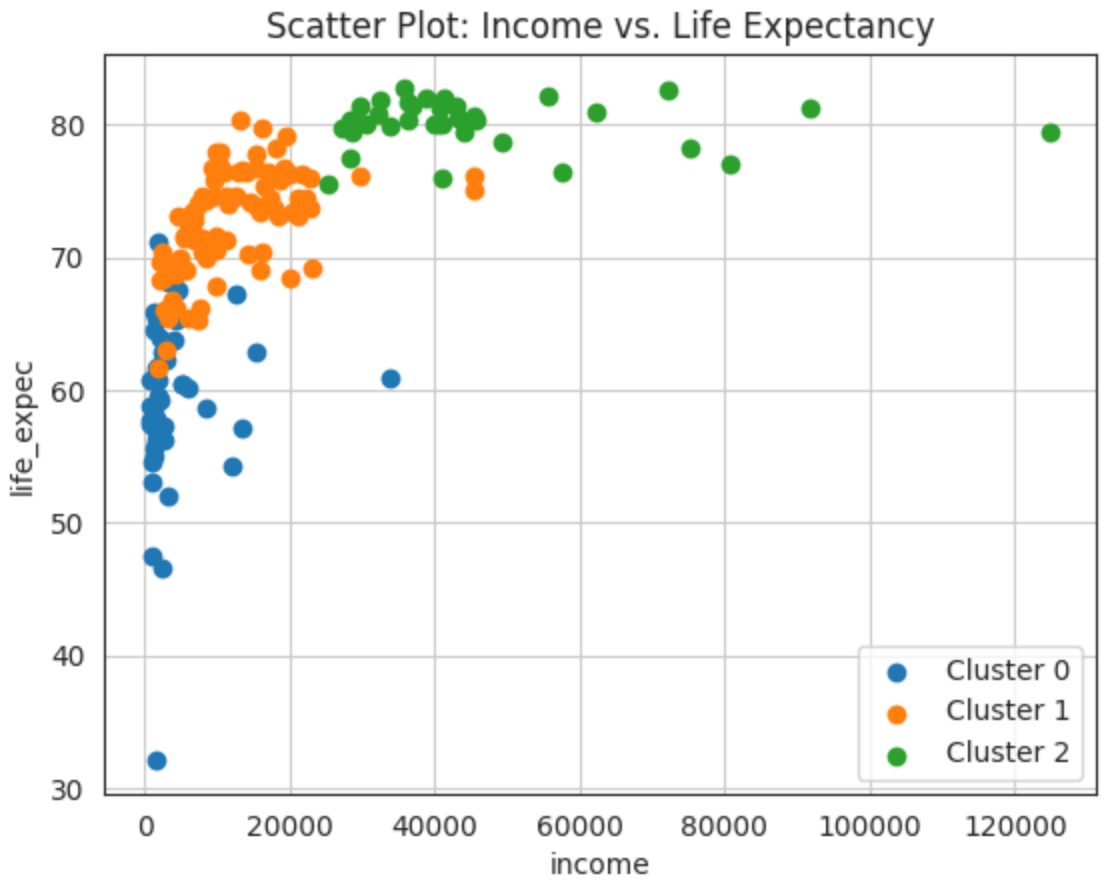
5) We made 10 scatter plots and found the following results:
    A) Cluster 0 has countries with high child mortality, low income, high fertility and low life expectancy.
    B) Cluster 1 had lower child mortality, higher income, lower fertility and higher life expectancy.
    C) Cluster 2 had lowest child mortality, highest income, lowest fertility and higher life expectancy.
    D) We did not find any specific clusters forming for explaining variation in health, export, import and inflation. This could mean that we need more clusters for doing so.
    E) Child mortality and income showed an Negative correlation.



F) Child mortality and health did not show a direct correlation.



G) Income and health showed a positive correlation.



6) In broad terms, cluster 0 represents underdeveloped countries, cluster 1 represents developing countries and cluster 2 represents developed countries based on their child mortality, income, life expectancy and fertility rate.