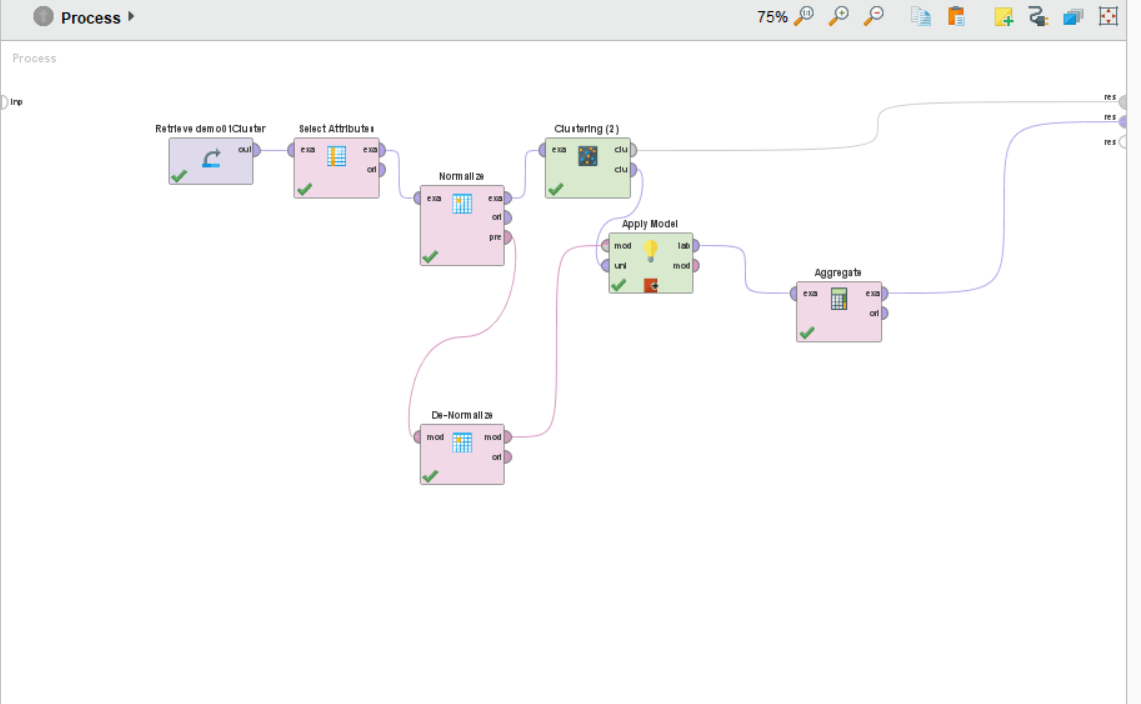
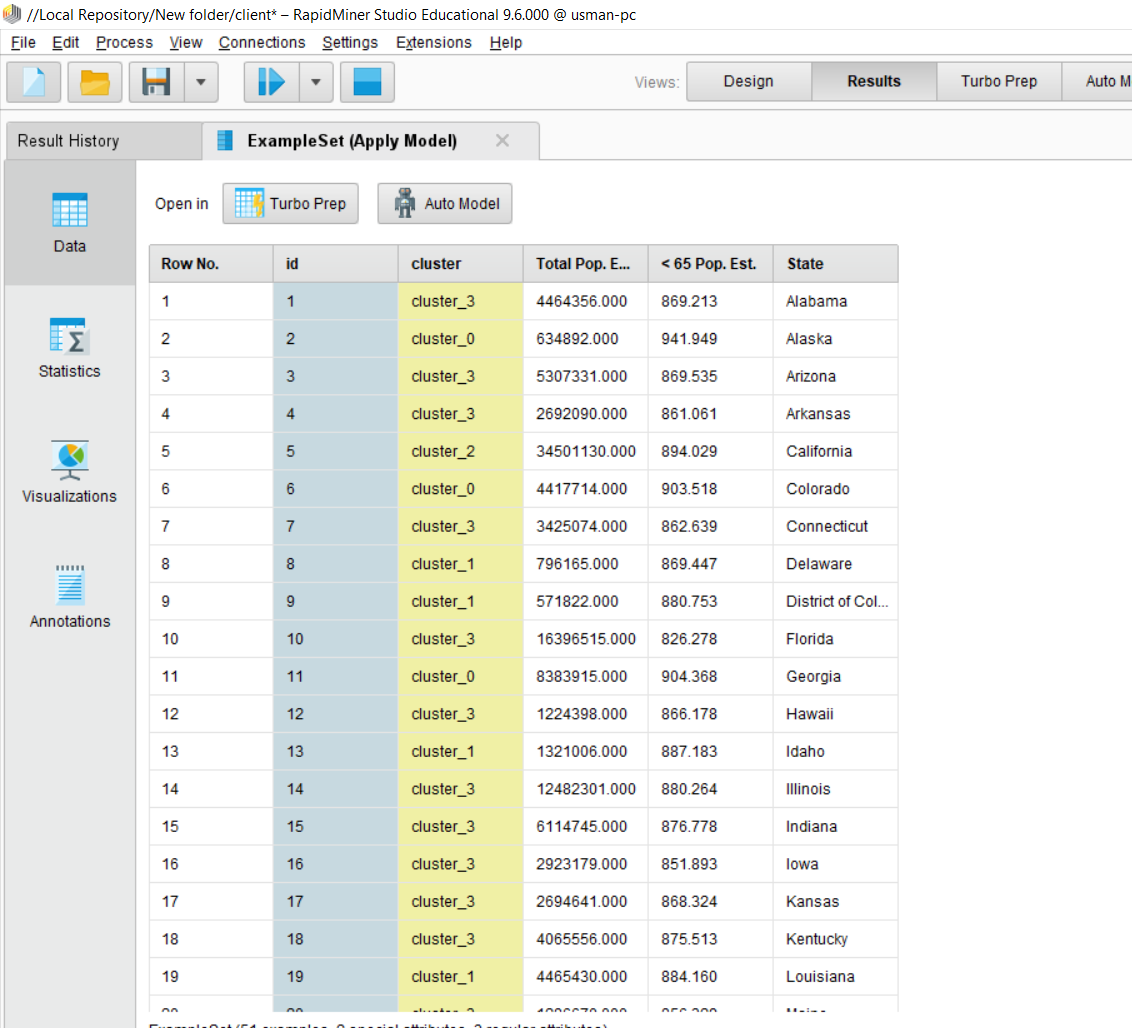
# K-mean using RapidMiner

# Process:



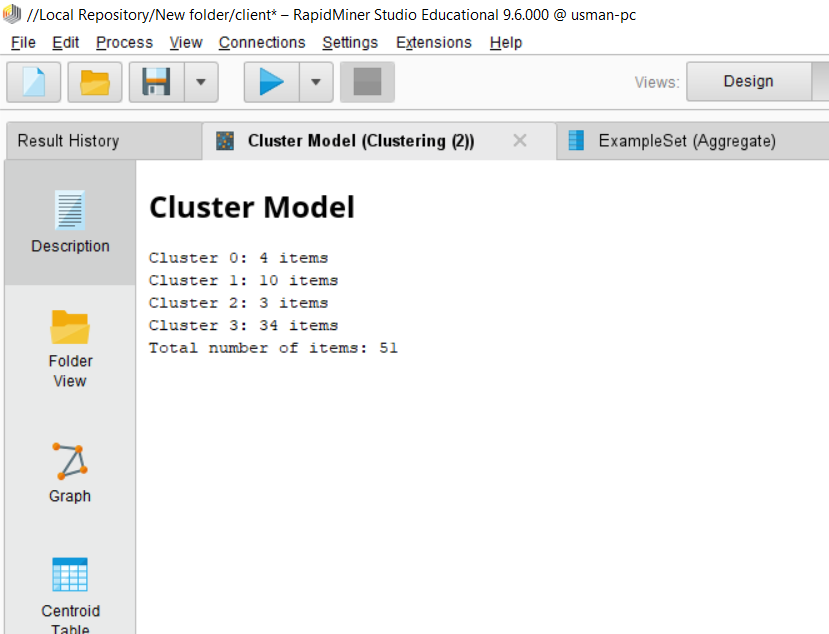
# Output



# Analysis of results

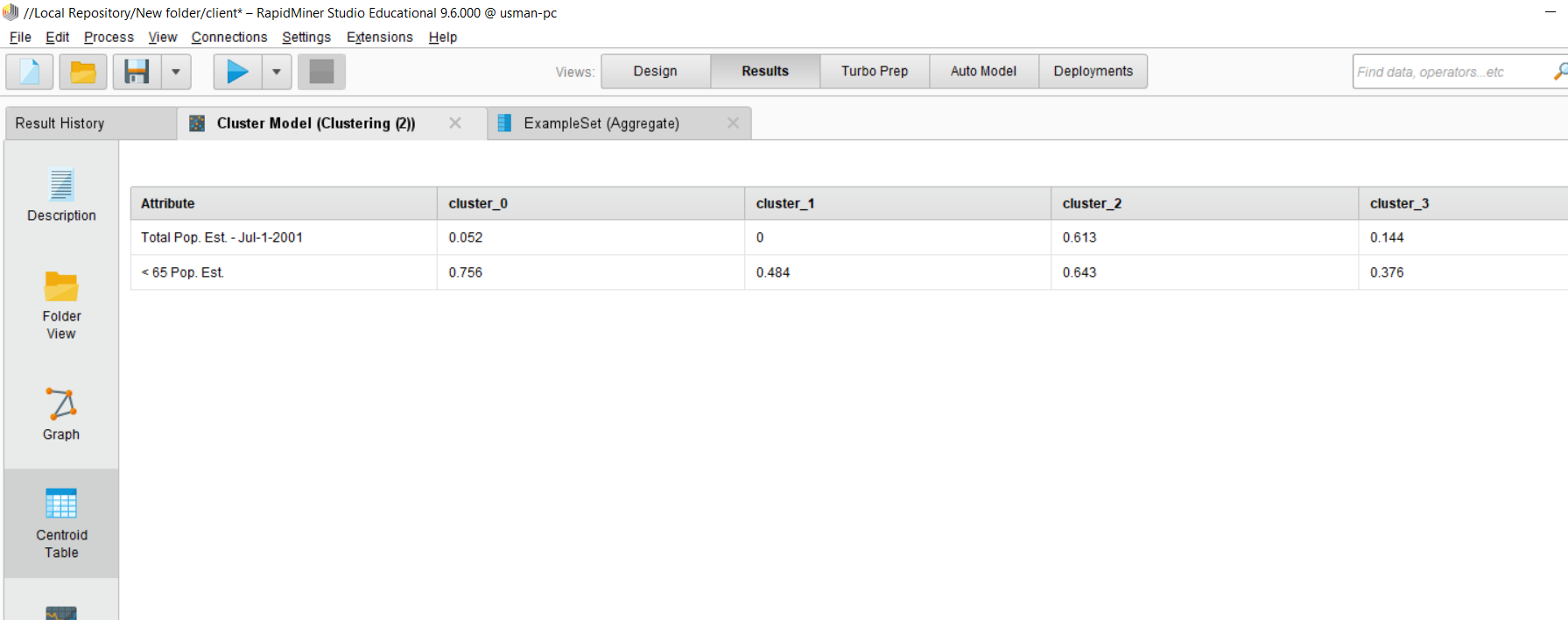
## How many states are in each cluster?

The no. of states against each cluster is mentioned in the screenshot



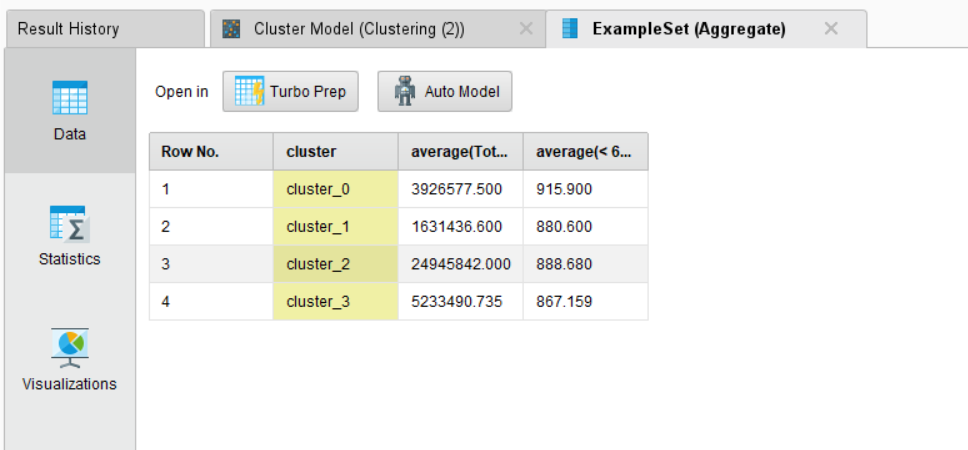
## What are the centroids of the 4 clusters generated?

In the screenshot, the centroids are mentioned as found in the centroid table.



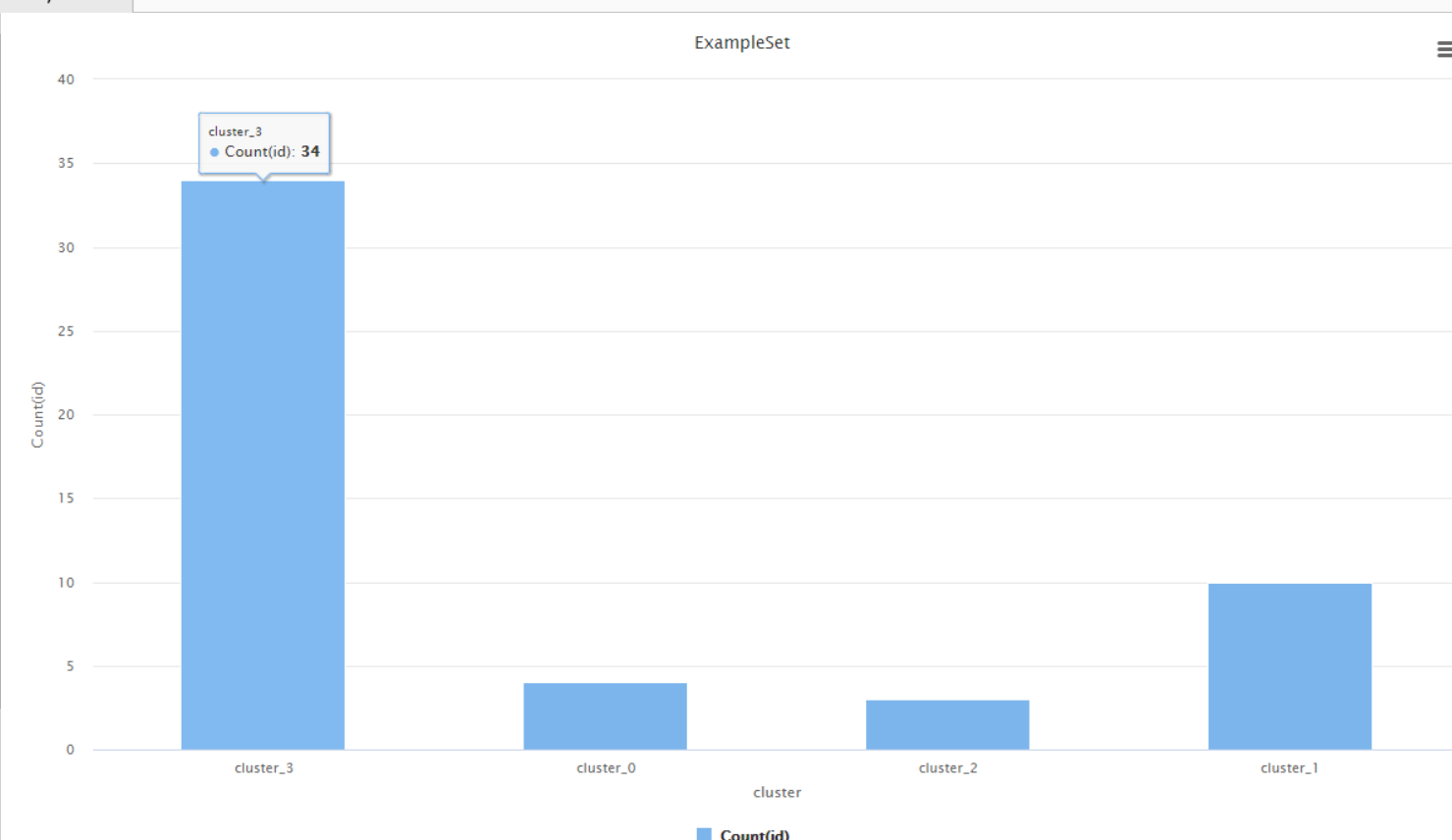
# Translate the centroid values into real values?

In order to find the centroid values in the original scale, we applied a denormalization on the output of K-Means and fed it into a model and then aggregated the centroid values which gives the following real values.



## Which country belongs to each cluster?

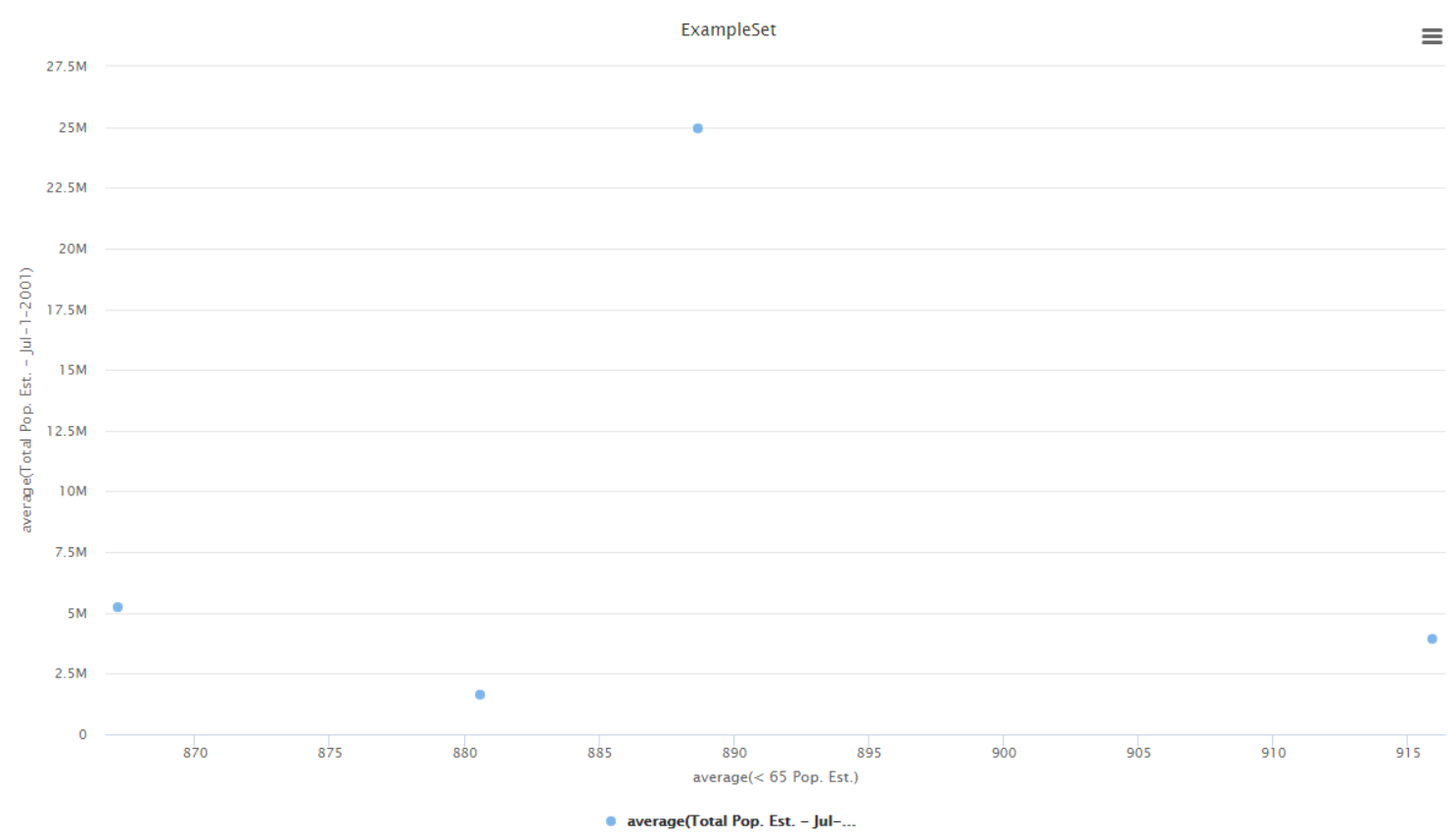
Below is the count of countries that belong to each cluster, I have generated this graph after summing the state ids over each cluster



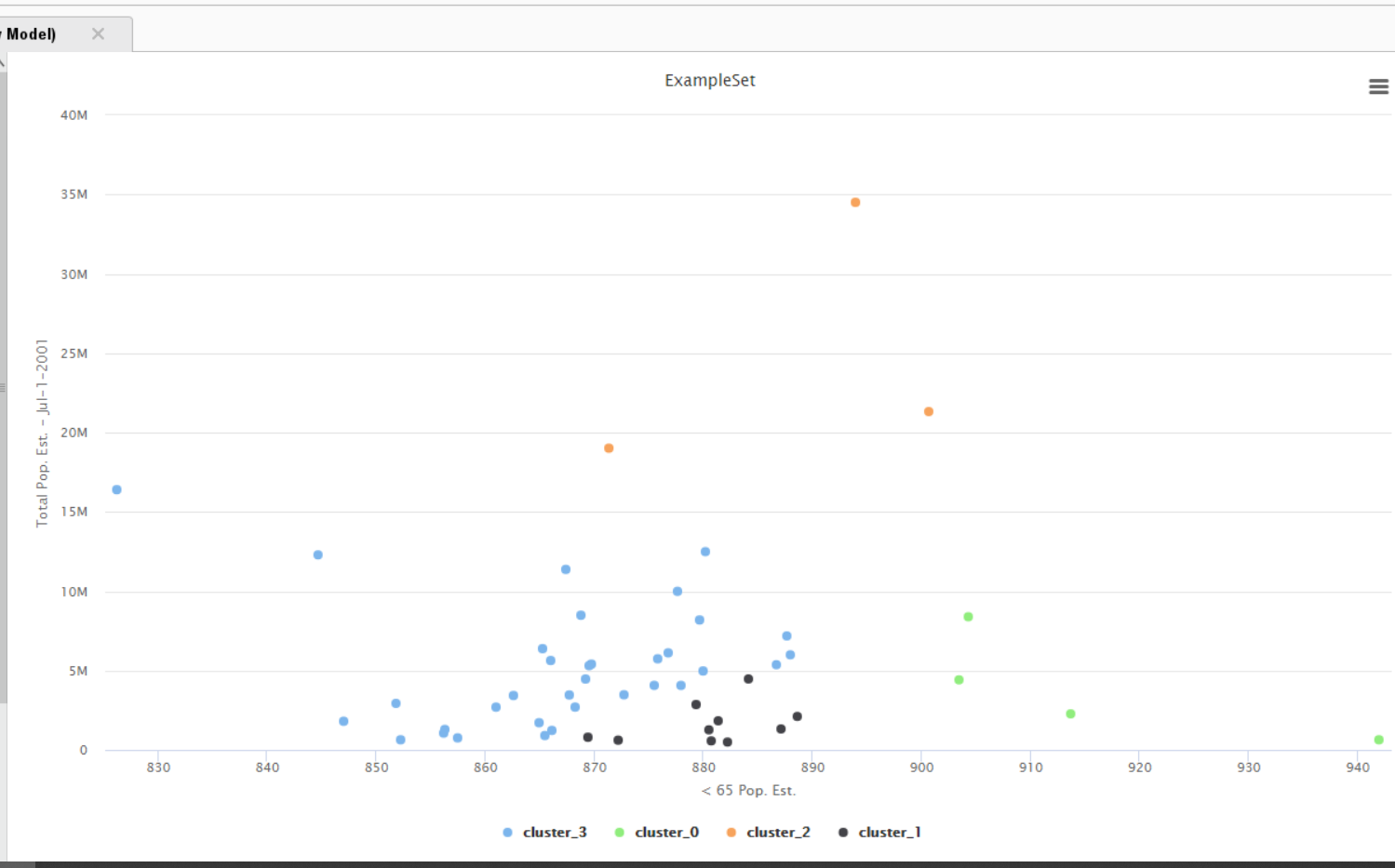
Names of coutries belonging to each cluster could be found in the following output, The clustor number is mentioned against every state.



## Add a scatter plot of the clusters generated (using the visualizations option)



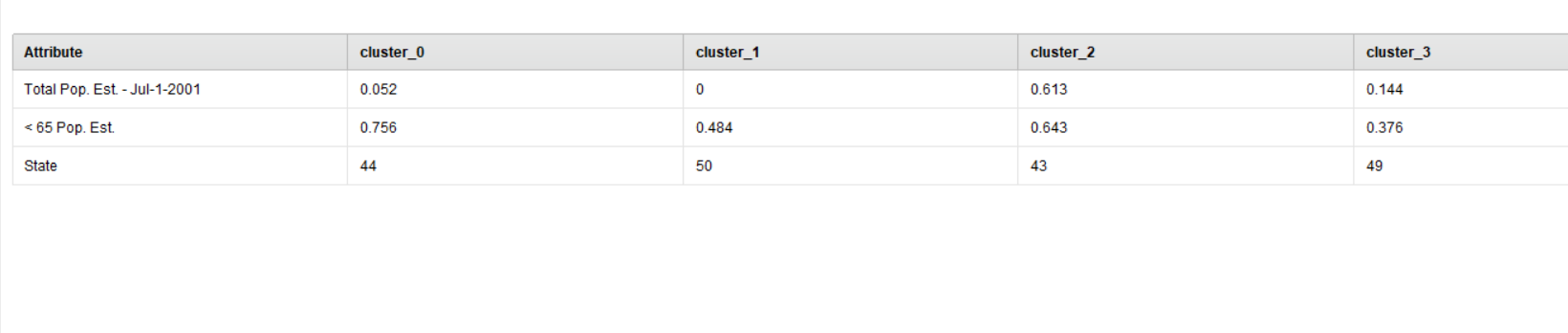
This scatter plot is regarding the original data set centroids



This is the visualization of Scatter plot; it is scarce because of lesser no. of examples (just 51 in number)

## What characterize each of the clusters (in relation to the two parameters used for clustering)? (i.e. explain the characteristics of the countries that belong to each cluster).

The centroid table State Ids are

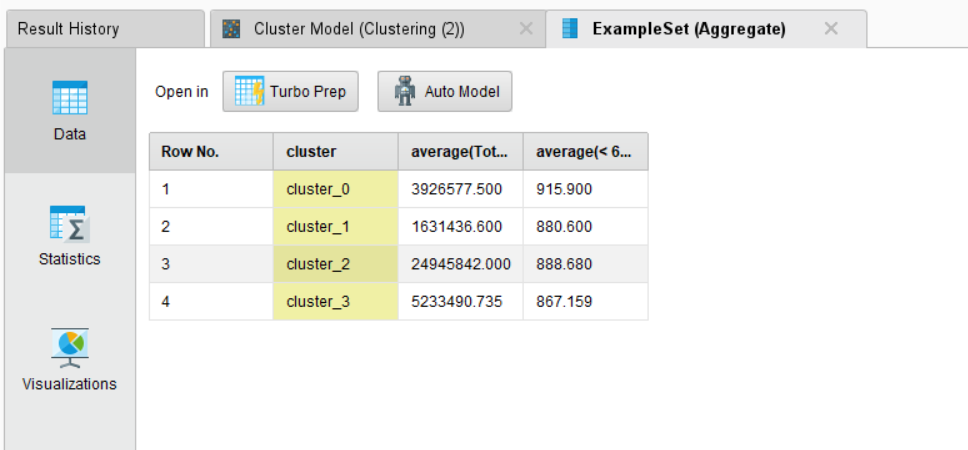


ID 44 corresponds to **Texas** state in table

ID 50 corresponds to **Wisconsin** state in table

ID 43 corresponds to **Tennessee** state in table

ID 49 corresponds to **West Virginia** state in table



Centroid normalized values

**Characteristics**

* **Countries having total pop close to** 3926577.50 **and <65 pop close to** 915 **(similar to Texas) belong to Cluster 0**
* **Total pop close to** 1631436.600 **and <65 pop close to** 880.599 **(similar to Wisconsin) belong to 1**
* **Total pop close to** 24945842.000 **and <65 pop close to** 888.679 **(similar to Tennessee) belong to 2**
* **Total pop close to** 5233490.735 **and <65 pop close to** 867.158 **(similar to West Virginia) belong to 3**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |