Flag Study

Table of Contents

Table of Figures

Table of Tables

Introduction

Religion in national symbols can often be found in national anthems of flags. In this project we will try to classify the national flags according to their characteristics. It's commonly known that a national flag is designed with specific meanings for its colors and symbols. The colors of a national flag may be worn from the people of a nation to show their patriotism, or the design of a national flag may be altered after the occurrence of important historical events. In the first part of the project we will try to predict the religion of a country according to their flag characteristics. Although, in the second part, we will try to create groups of flags that have common characteristics, clusters.

Dataset

The dataset contains various details of countries and their flags (in this dataset there are flags of countries that are not still exist). In more details, the dataset contains **194 countries** and **30 variables-attributes** for each country. We have split the dataset into tree tables according to the type of the variable. In the Table 1 - Geographical Characteristics we have variables that are related to the geographical characteristics of the flag, in the Table 2 – Colour Characteristics we have the variables related to the colours of the flag and in the Table 3 - Geometrical Characteristics we have all the geometrical characteristics (e.g. shapes, lines).

Variable Name	Variable Type	Description
name	Character	Name of the country concerned
landmass	Factor (6 levels)	Geographical Continent
zone	Factor (4 levels)	Geographic quadrant, based on Greenwich and the Equator
area	Numerical	Area in thousands of square km
population	Numerical	Population in round millions
language	Factor (8 levels)	Language
religion	Factor (5 levels)	Religion

Table 1 - Geographical Characteristics

Variable Name	Variable Type	Description
colours	Numerical	Colours Number of different colours in the flag
red	Factor (2 levels)	0 if red absent, 1 if red present in the flag
green	Factor (2 levels)	0 if green absent, 1 if green present in the flag
blue	Factor (2 levels)	0 if blue absent, 1 if blue present in the flag
gold	Factor (2 levels)	0 if gold absent, 1 if gold present in the flag
white	Factor (2 levels)	0 if white absent, 1 if white present in the flag
black	Factor (2 levels)	0 if black absent, 1 if black present in the flag
orange	Factor (2 levels)	0 if orange absent, 1 if orange present in the flag
mainhue	Factor (8 levels)	Predominant colour in the flag
topleft	Factor (8 levels)	Colour in the top-left corner (moving right to decide tie-breaks)
topright	Factor (8 levels)	Colour in the bottom-left corner (moving left to decide tie-breaks)

Table 2 – Colour Characteristics

Variable Name	Variable Type	Description
bars	Numerical	Number of vertical bars in the flag
stripes	Numerical	Number of horizontal stripes in the flag
circles	Numerical	Number of circles in the flag
crosses	Numerical	Number of (upright) crosses
saltires	Numerical	Number of diagonal crosses
quarters	Numerical	Number of quartered sections
sunstars	Numerical	Number of sun or star symbols
crescent	Numerical	1 if a crescent moon symbol present, else 0
triangle	Numerical	1 if any triangles present, 0 otherwise
icon	Factor (2 levels)	1 if an inanimate image present (e.g., a boat), otherwise 0
animate	Factor (2 levels)	1 if an animate image (e.g., an eagle, a tree, a human hand) present, 0 otherwise
text	Factor (2 levels)	1 if any letters or writing on the flag (e.g., a motto or slogan), 0 otherwise

Table 3 - Geometrical Characteristics

PART I – Classification

In this part of the project we will try to categorize the nation flags and identify their religion according to the characteristics of the flag. In the beginning of this analysis we do a descriptive analysis in order to understand better our data, something very important if we want to understand the behavior of the following models. We will try different classification algorithms (classifiers) and finally we compare our results and choose one of them. In more details, our classification analysis will include the following methods:

- Decision Trees
- Support Vector Machines

Descriptive Analysis

First, we import the data in **R studio**. At this stage we will take a quick view in the data in order to understand them better (Table 2 – Colour Characteristics and Table 3 - Geometrical Characteristics).
To begin, we create a bar graph with the number of the flags for each country (see Figure 1 - Number of flags for each Religion). Christians and Catholic are the most widespread religions.

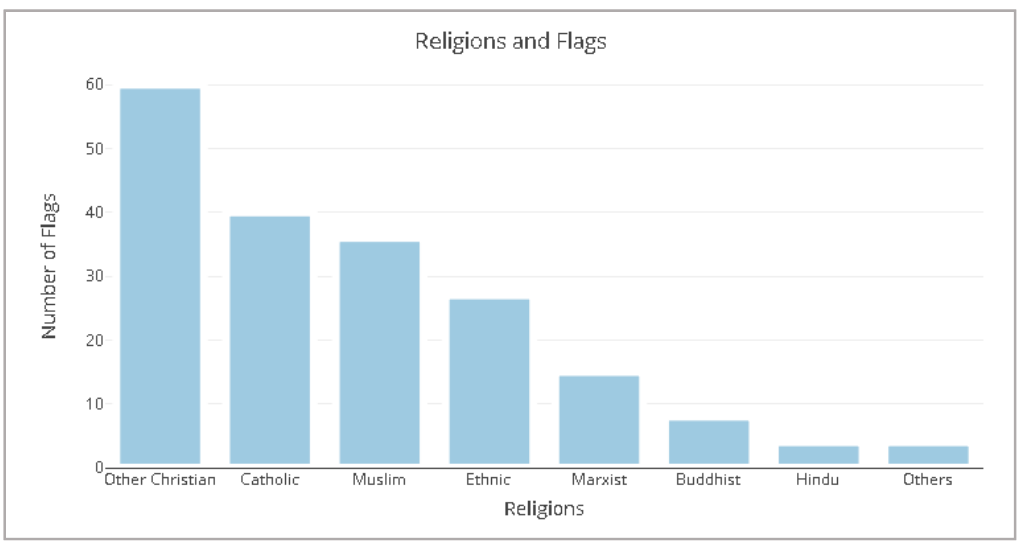

Figure 1 - Number of flags for each Religion

Figure 2 - Geometrical Characteristics

Respectively, we create bar plots for colour characteristics (see in Figure 3 –Number of Colours and Figure 4 - Main Hue Colour). We observe that the majority of flags are composed with 3 different colours (see in the left plot) and the most popular colour is Red.

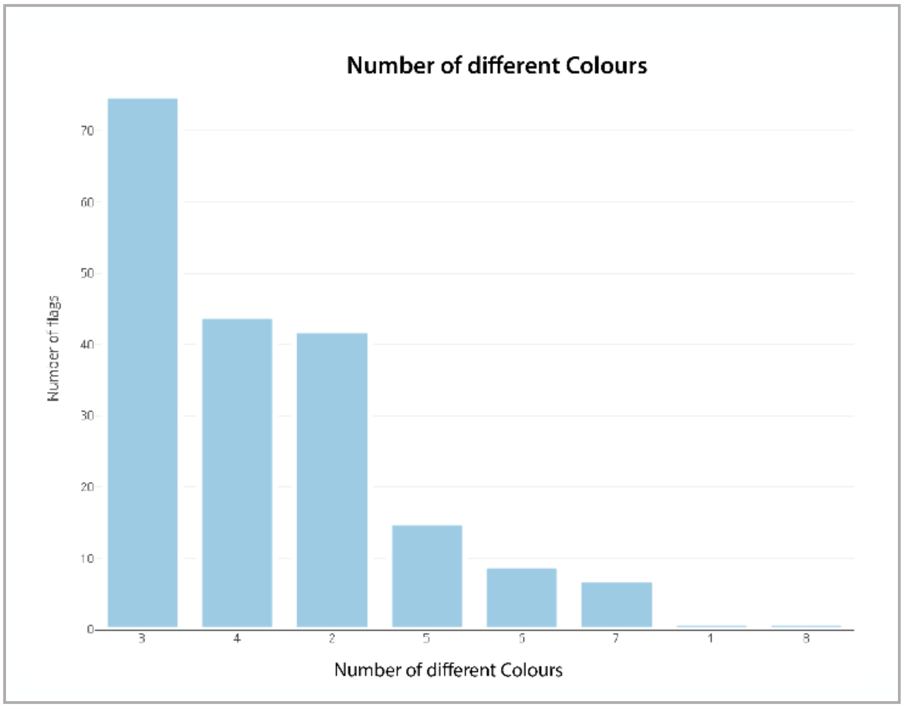

Figure 3 –Number of Colours

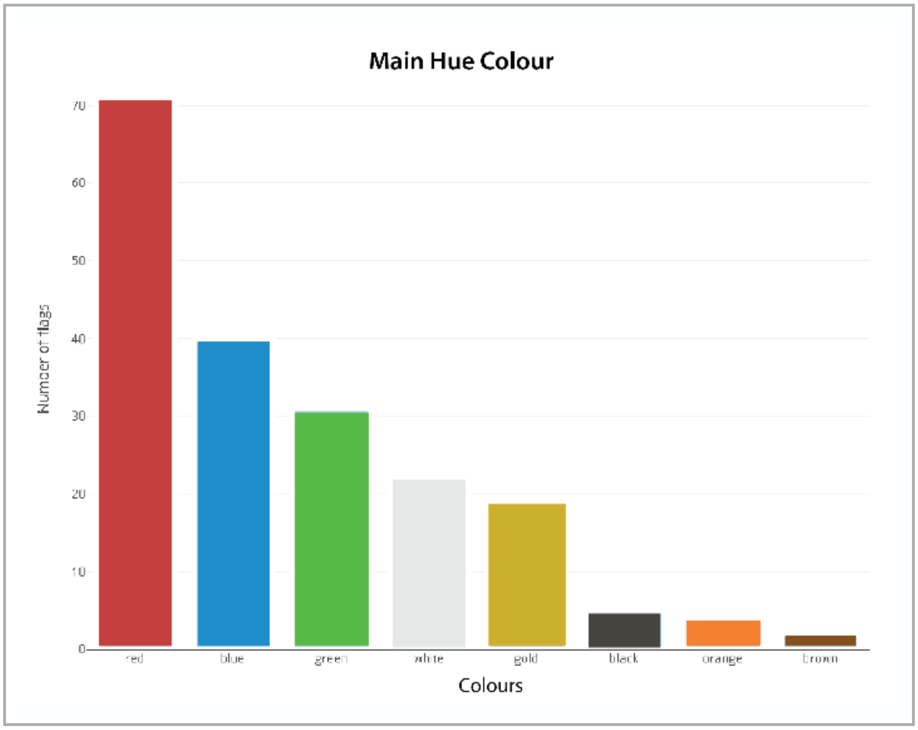

Figure 4 - Main Hue Colour

Classification Trees

In this part of the project we will create various classification trees (recursive partitioning). Firstly, we will create a maximal tree and after we will train other two different simpler models by using different methods. In more details, after the initial tree we will try to reduce the variables and grow again a new tree. Finally, we will use another approach by using a more pioneer method by using **Rpart**[1] library. Hence, at the end we have to combine the results of models and keep the model with highest prediction abilities.

1. CART Tree

Firstly we create our initial model, after we create other models by using different methods and finally we compare our results. For model evaluation we used an algorithm which do K-fold Cross validation. We preferred this technique because our dataset has only 194 observations, a very small number, which isn't

[1] For Rpart look in: https://cran.r-project.org/web/packages/rpart/rpart.pdf

enough for a simple split for training and test dataset. Our major criterion for model evaluation is the model accuracy. Accuracy is the number of good predictions divided with the total predictions.

The model after **tree** function (see in Figure 5 – Initial Classification Tree) is very complicated because we have a lot of mixed variables (numerical and categorical).

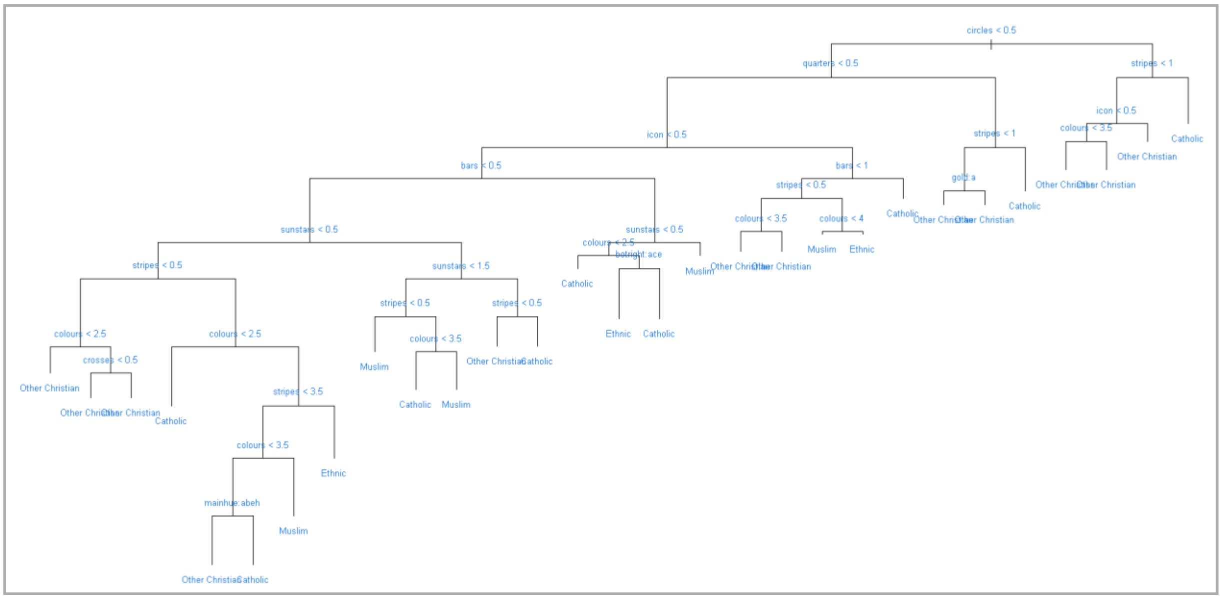

Figure 5 – Initial Classification Tree. The letters are too small to read them, we add this diagram for illustrator purposes and show the complexity of this model.

I would like to mention that we used “Gini” method for node splitting because we have categorical target variable, so we can’t use “variance” method. This tree is the largest tree that we could have with 30 terminal leaves. It’s very complicated; so it’s very difficult to interpret our results. We evaluate our model by using K-fold cross validation and finally to take the average of the accuracy score of each fold. The accuracy score of this model is 43%.

2. Classification Tree & Variable Selection with Random Forest

In this part we firstly do a variable selection before we train a tree model. For variable selection we used **Random Forest Method**. Random forest improves predictive accuracy by generating a large number of bootstrapped trees (based on random samples of variables) , classifying a case using each tree in this new "forest", and deciding a final predicted outcome by combining the results across all of the trees (an average in regression, a majority vote in classification). Breiman and Cutler's random forest approach is implemented via the Random Forest package.

With Random Forest Method we directly measure the impact of each feature on accuracy of the model. The general idea is to permute the values of each feature and measure how much the permutation decreases the accuracy of the model. Clearly, for unimportant variables, the permutation should have little to no effect on model accuracy, while permuting important variables should significantly decrease it. So, finally, we keep the 21 most important variables (see in Figure 6 - Variables Importance (Random Forest Method)).

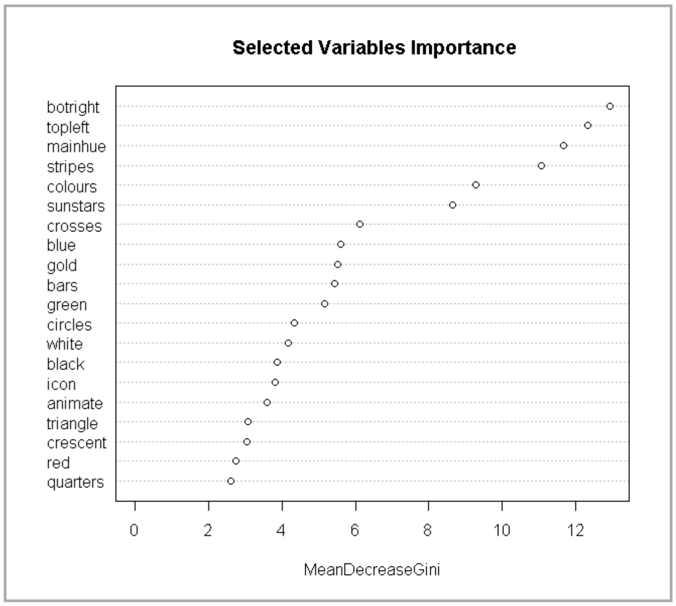

Figure 6 - Variables Importance (Random Forest Method)

After variable selection we train again the tree model (see in Figure 7 - Classification Tree (after variable selection). This tree is less complicated and has 22 terminal leaves.

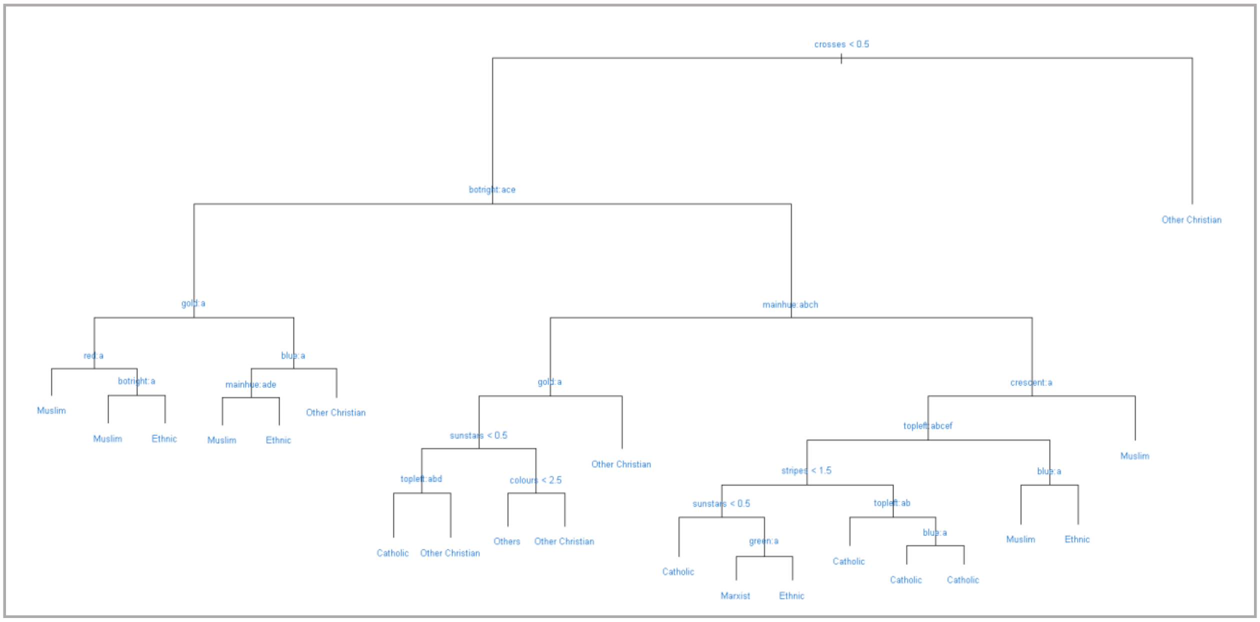

Figure 7 - Classification Tree (after variable selection). The letters are too small to read them, we add this diagram for illustrator purposes and show the complexity of this model.

Finally, this tree has 43% accuracy, which is the same with the previous model but it's less complicated.

3. Rpart Tree

Now, we create another tree by using **Rpart** library in R. The main difference from the **CART Tree** function is that Rpart handling of surrogate variables. Another difference is how pruning takes places. Specifically, Rpart treats differently ordinal and categorical variables. In a sense, our model from Rpart has 13 leaves (before pruning). After we prune our model in the point which minimizes the complexity parameter (cp). Finally, our pruned tree is much simpler, with 10 terminal leaves (see in Figure 8 - Pruned Classification Tree).

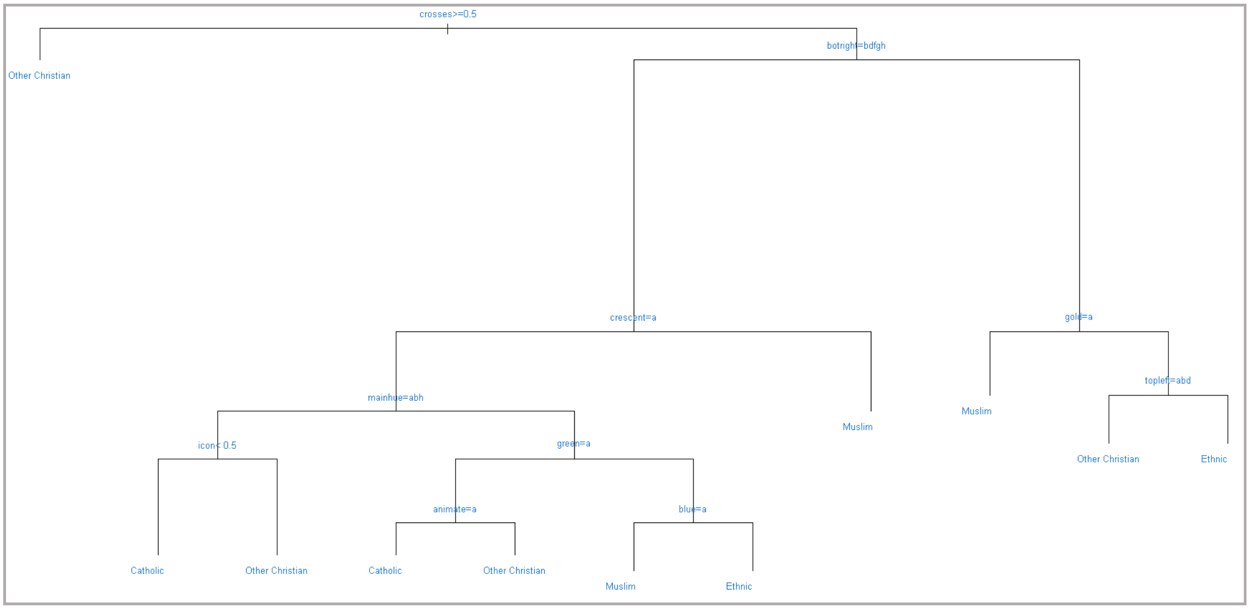

Figure 8 - Pruned Classification Tree. The letters are too small to read them, we add this diagram for illustrator purposes and show the complexity of this model.

Then we evaluate this model by using K-fold cross validation and finally we take the average of the accuracy score of each fold. The accuracy score of this model is extremely higher than the previous model 95%.

Support Vector Machines (SVM)

With Support Vector Machines (SVM), basically, we are looking for the optimal separating hyperplane between the two classes by maximizing the margin between the classes' closest point (see in Figure 9 - Decision Tree & SVM Comparison). In R we use "e1071" package. Firstly, we tune **cost** and **gamma** model parameters with a generic function (tune.svm()[2]) of statistical methods using a grid search over supplied parameter ranges (I would like to mention that this method is computational intensive and it takes several minutes to finish). After we have already the two parameters, we run the model and we calculate the accuracy of this method in order to compare it with the other classification methods.

[2] For more information about tune.svm look in : https://www.rdocumentation.org/packages/e1071/versions/1.6-8/topics/tune

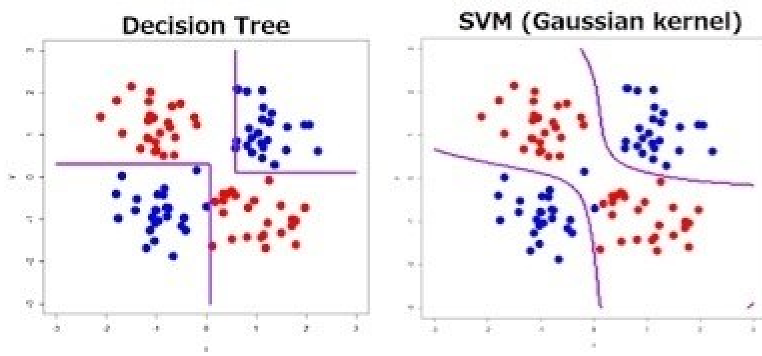

Figure 9 - Decision Tree & SVM Comparison. A decision tree will partition the feature space into half, using axis aligned linear decision boundaries. So, using a decision tree, you get a non -linear decision boundary, sometimes more than one; SVM method creates hyperplanes which maximizes the margin between the classes.

Models Comparison

At this stage we have to select one of the previous methods as the optimal for classification. Our selection criterion is the accuracy of the models and also the complexity of the classification tree models(see in Table 4 - Models Comparison). We calculate the accuracy for each model by using Cross Fold Validation because the volume of the data was not sufficient for a simple split to train and test partitions. As we can see, the model from Rpart library is the less complex model, with only 10 terminal leaves and the highest accuracy of all of them (95%). The support Vector Machines (SVM) method has a good accuracy value but not greater than the Rpart tree. Also, I would like to mention that the training of the SVM classifier it takes a lot of time.

Model Name	Accuracy	Complexity
Tree with the attributes	43%	30 terminal leaves
Tree with variable selection	43%	22 terminal leaves
Rpart tree	95%	10 terminal leaves
Support Vector Machines (SVM)	76%	----

Table 4 - Models Comparison

PART II – Clustering

In this part we have to do cluster analysis (clustering). In short, we have to create sets of flags in such a way that flags in the same cluster have common **geometrical** and **colour** characteristics to each other than those in the other clusters. It can be achieved by various algorithms that significantly differ in their notion of what constitutes a cluster and how to efficiently find them. We will use the following algorithms:

- PAM Clustering
- Hierarchical Clustering

Finally, we will compare our results in a sense of interpretation and clusters homogeneity.

Clusters Distance Metric

We need a linkage criterion in order to determine the distance between the sets of observations as a function of the pairwise distances between observations. In our dataset we have both numerical and factor variables which need different treatment. So, we prefer to use **Gower distance**[3].

The concept of Gower distance is actually quite simple. For each variable type, a particular distance metric works well and scales to fall between 0 and 1. Then, a linear combination, using user-specified weights (most simply an average), is calculated to create the final distance matrix. The metrics used for each data type are described below:

- **Quantitative (interval)**: range-normalized Manhattan distance
- **Ordinal**: variables are first ranked, then Manhattan distance is used with a special adjustment for ties
- **Nominal**: variables of *k* categories are first converted into *k* binary columns and then the Dice coefficient is used

[3] For Gower distance look in : https://www.rdocumentation.org/packages/StatMatch/versions/1.2.5/topics/gower.dist

Firstly, we calculate the Gower distance by using the **daisy**[4] function. If we want to check out the performance of the matrix we can do a quick search. We can search for the most similar pairs of flags, the pairs of flags which have the minimum distance; these flags are the flags of Syria and Iraq.

Figure 10 - Syria flag (on the left), Iraq flag (on the right)

Also, we can search to find the most dissimilar pairs in our Gower distance matrix. Our matrix shows that these flags are the flags of Haiti and Hong-Kong (see Figure 11 - Haiti flag (on the left), Hong Kong flag (on the right)).

Figure 11 - Haiti flag (on the left), Hong Kong flag (on the right)

Afterwards, we have to find the optimum number of the clusters by using Silhouettes plots (see in Figure 12 - Silhouette Plot). In Silhouette plot, the better Width is the highest. In our case the best number is nine clusters.

[4] For daisy function look in: http://stat.ethz.ch/R-manual/R-devel/library/cluster/html/daisy.html

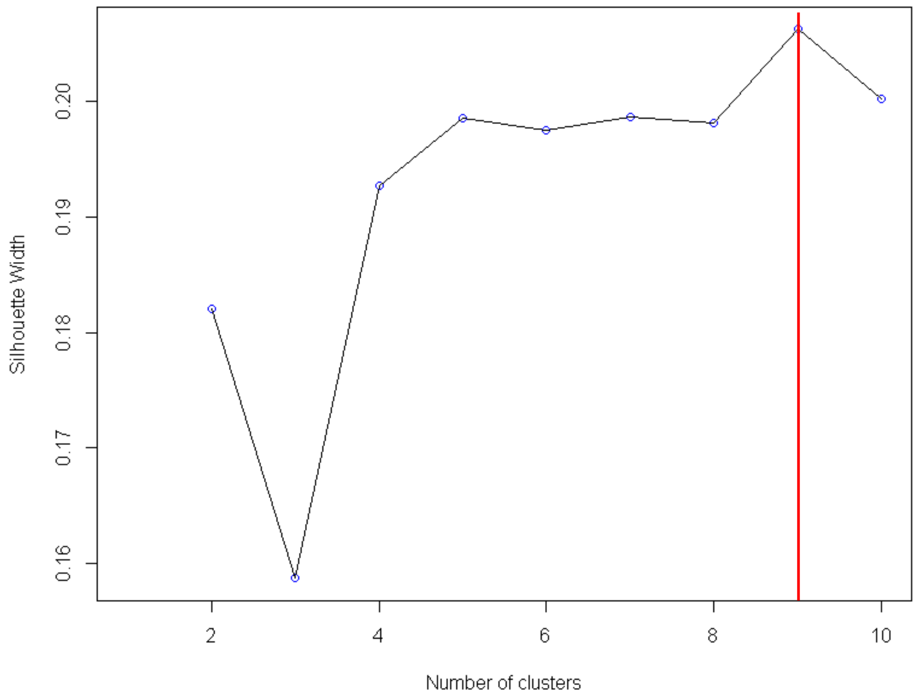

Figure 12 - Silhouette Plot

Although, in the following clustering methods we will set 9 in number of clusters.

Hierarchical Clustering

In this part we will use **Agglomerative Hierarchical Clustering**[5], this method is a "bottom up" approach; each observation starts in its own cluster, and pairs of clusters are merged as one moves up to the hierarchy. In order to decide which clusters should be combined, a measure of dissimilarity between sets of observations is required (a metric, in our occasion we have the Gower distance).

Firstly, we run the model and we create a dendogram with aim to have a first look to this clustering method. As we observe in Figure 13 - Hierarchical Clustering Dendrogram the 9 clusters can't be clear identified

[5] For Agglomerative Hierarchical Clustering function in R look in: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/hclust.html

with this method. But, we need more information if we want to have an all-around knowledge about this approach.

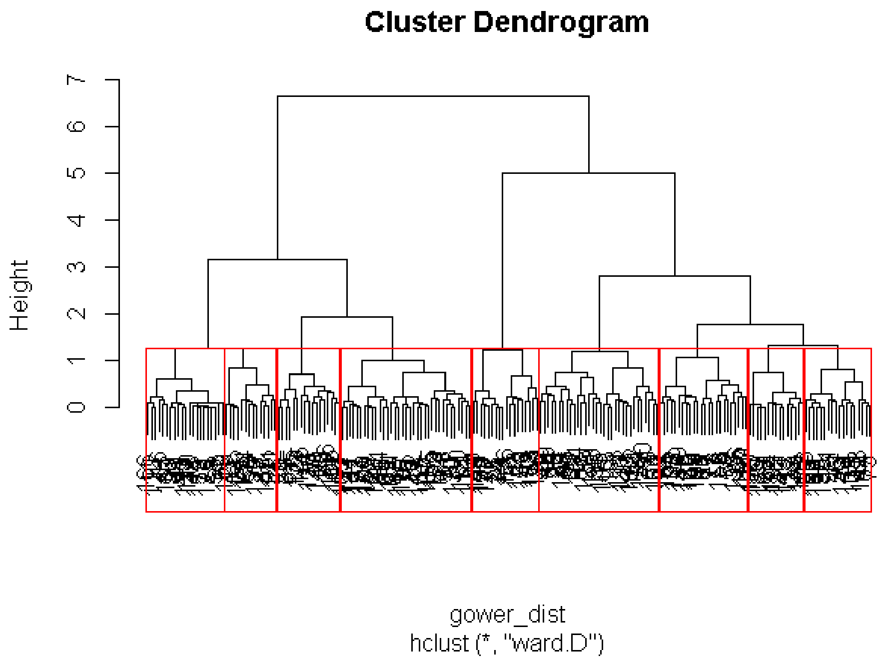

Figure 13 - Hierarchical Clustering Dendrogram

In the case of that, we create Silhouette plots with aim to study the separation distance between the resulting clusters. The silhouette plots displays a measure of how close each point in one cluster is to a point in the neighboring ones and thuds it provides a way to assess parameters, like the number of clusters visually (see in Figure 14 -Hierarchical Clustering - Silhouette Plot). As we can see, we don't have a very good clustering. A lot of observations look to belong to another cluster than they have assign. The red line in the plot is the limit of the lower Silhouette Width that the clusters should have. The majority of the clusters are below that line.

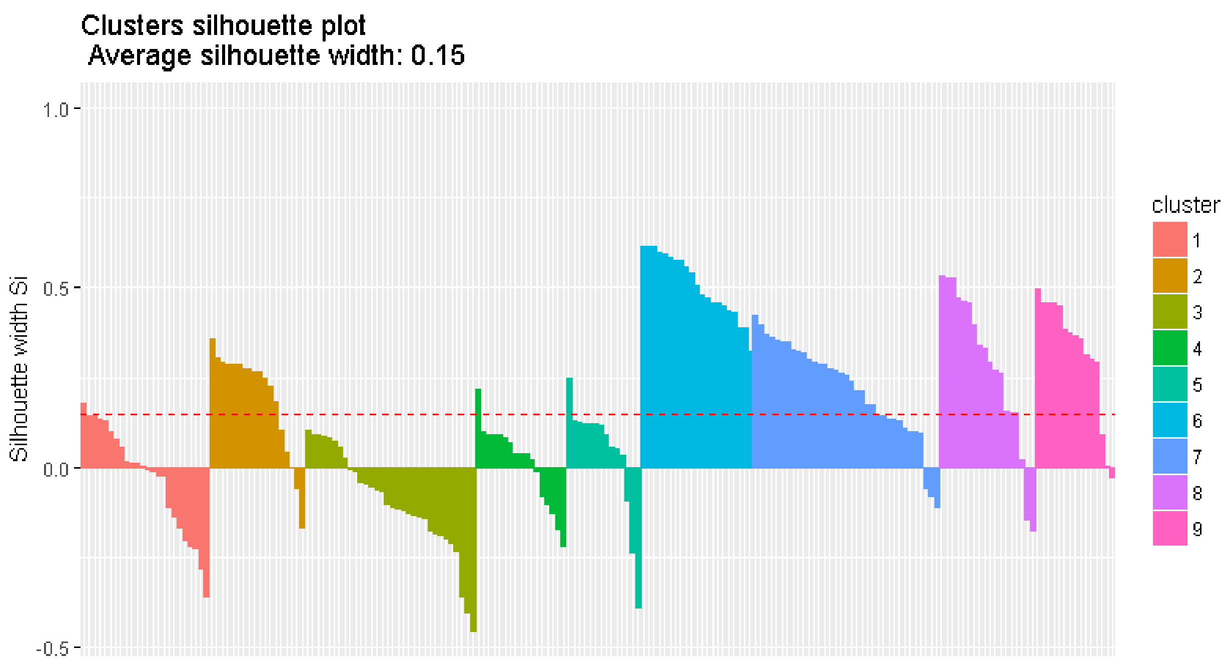

Figure 14 -Hierarchical Clustering - Silhouette Plot

PAM Method

The PAM method is **based on medoids** among the observations of the dataset. These observations should represent the structure of the data. We have to find a set of nine medoids; nine clusters are constructed by assigning each observation to the nearest medoid. The goal is to find nine representative objects which minimize the sum of the dissimilarities of the observations to their closest representative object. The algorithm first looks for a good initial set of medoids (this is called the BUILD phase). Then it finds a local minimum for the objective function; such a solution there is no single switch of an observation with a medoid that will decrease the objective (this is called the SWAP phase).

We create the nine clusters by using the **Pam function**[6] in R. Firstly, we run the silhouettes plot with aim to study the separation distance between the resulting clusters (see in Figure 15 – Pam Method - Silhouette Plot). As we can see we don't have a very well separated clusters, but we have better results than the Hierarchical Clustering method.

[6] For PAM function look in: https://stat.ethz.ch/R-manual/R-devel/library/cluster/html/pam.object.html

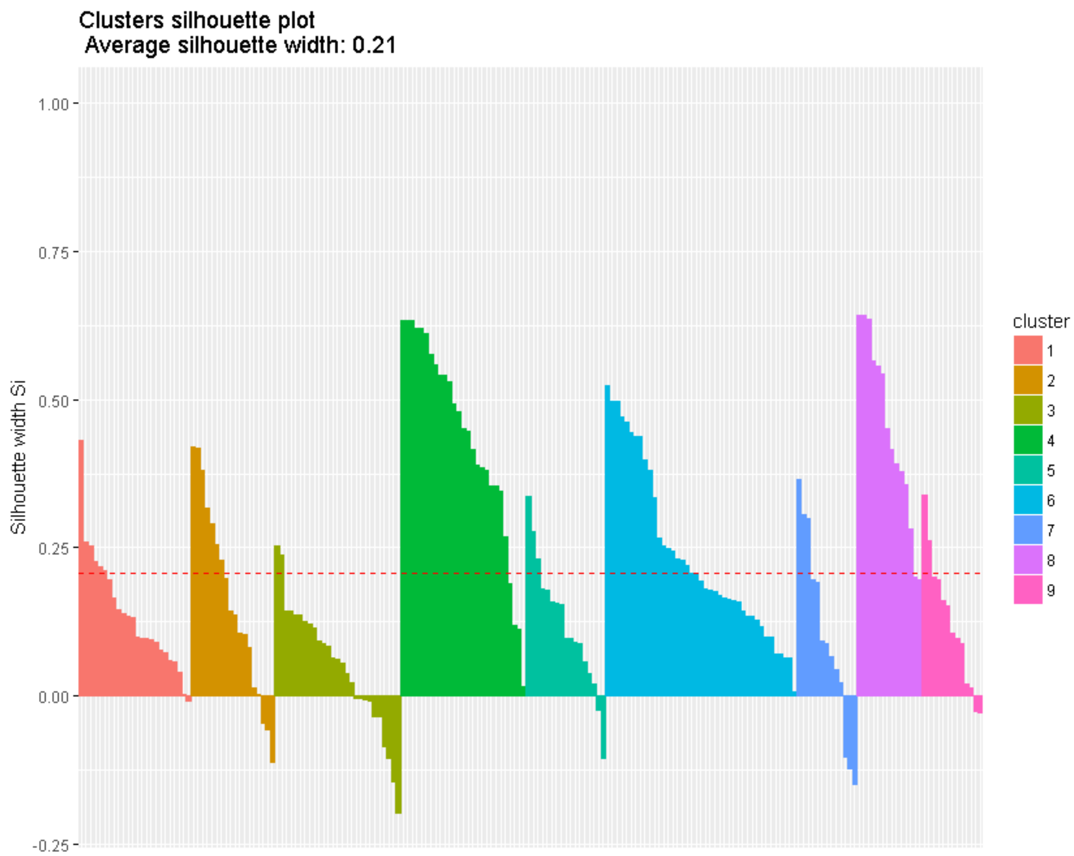

Figure 15 – Pam Method - Silhouette Plot

Cluster Analysis

Comparing the two clustering methods (Hierarchical Clustering and PAM Method) we observe that PAM methods create more clear clusters (based on Silhouette plots, look Figure 14 -Hierarchical Clustering - Silhouette Plot and Figure 15 – Pam Method - Silhouette Plot). So, we decide to select the PAM method. Now, we will do a descriptive analysis for each cluster in order to identify common characteristics of the clusters that maybe have sense. Firstly, we create pies with the percentage of the religions for each cluster in order to see if this attribute is important to distinguish the clusters (see in Figure 16 -Religions % per Cluster)

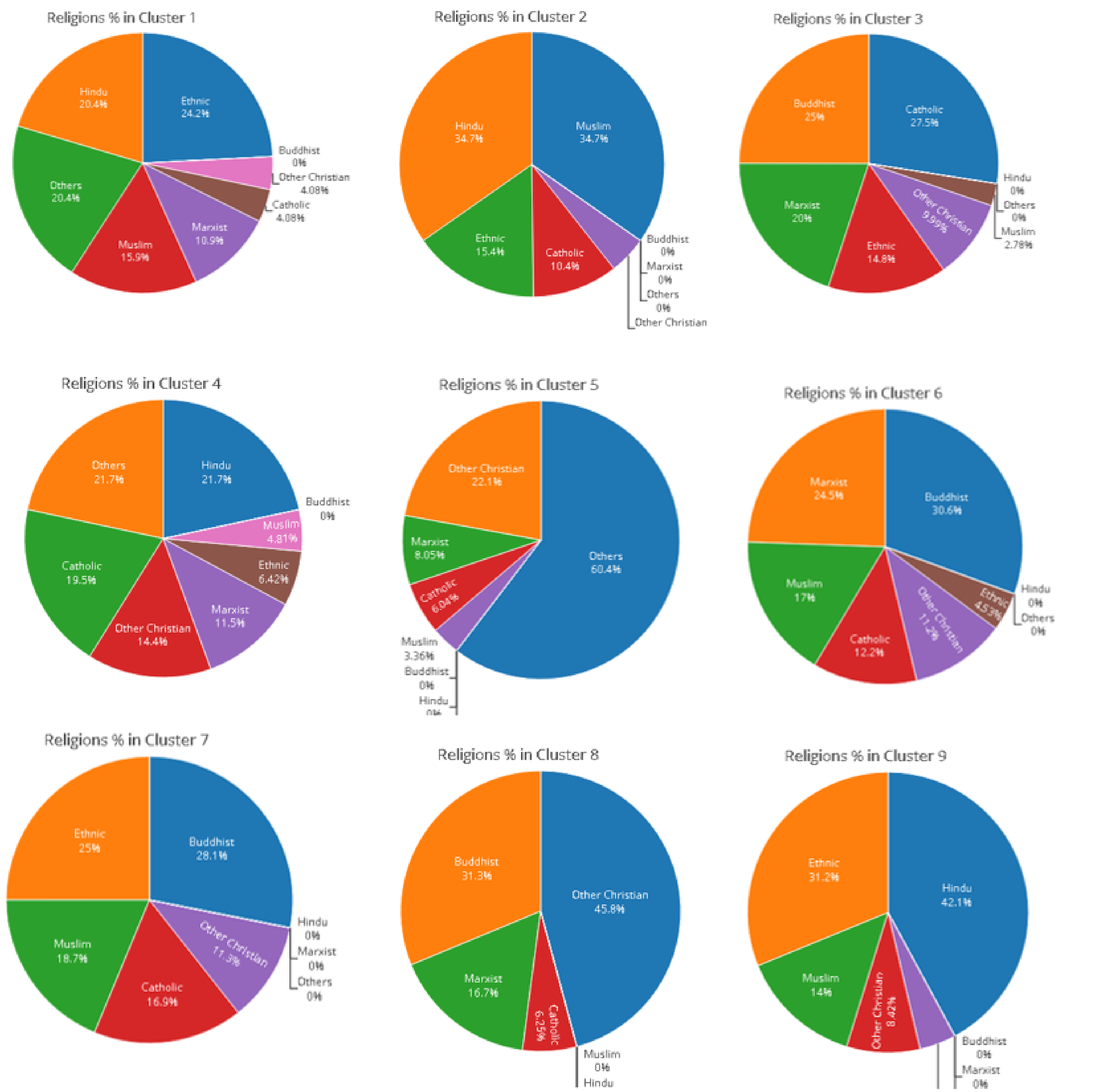

Figure 16 -Religions % per Cluster

As we can observe, in each cluster we have a different dominant religion (it's a good sign!). It's logical to have two times, two religions, because we have nine clusters and eight religions. After, we decide to investigate the percentage of the continents for each cluster (see in Figure 17 – Continents % per Cluster). I would like to mention that we didn't take into consideration the continents attributes in the clusters creation, because we base on flag characteristics.

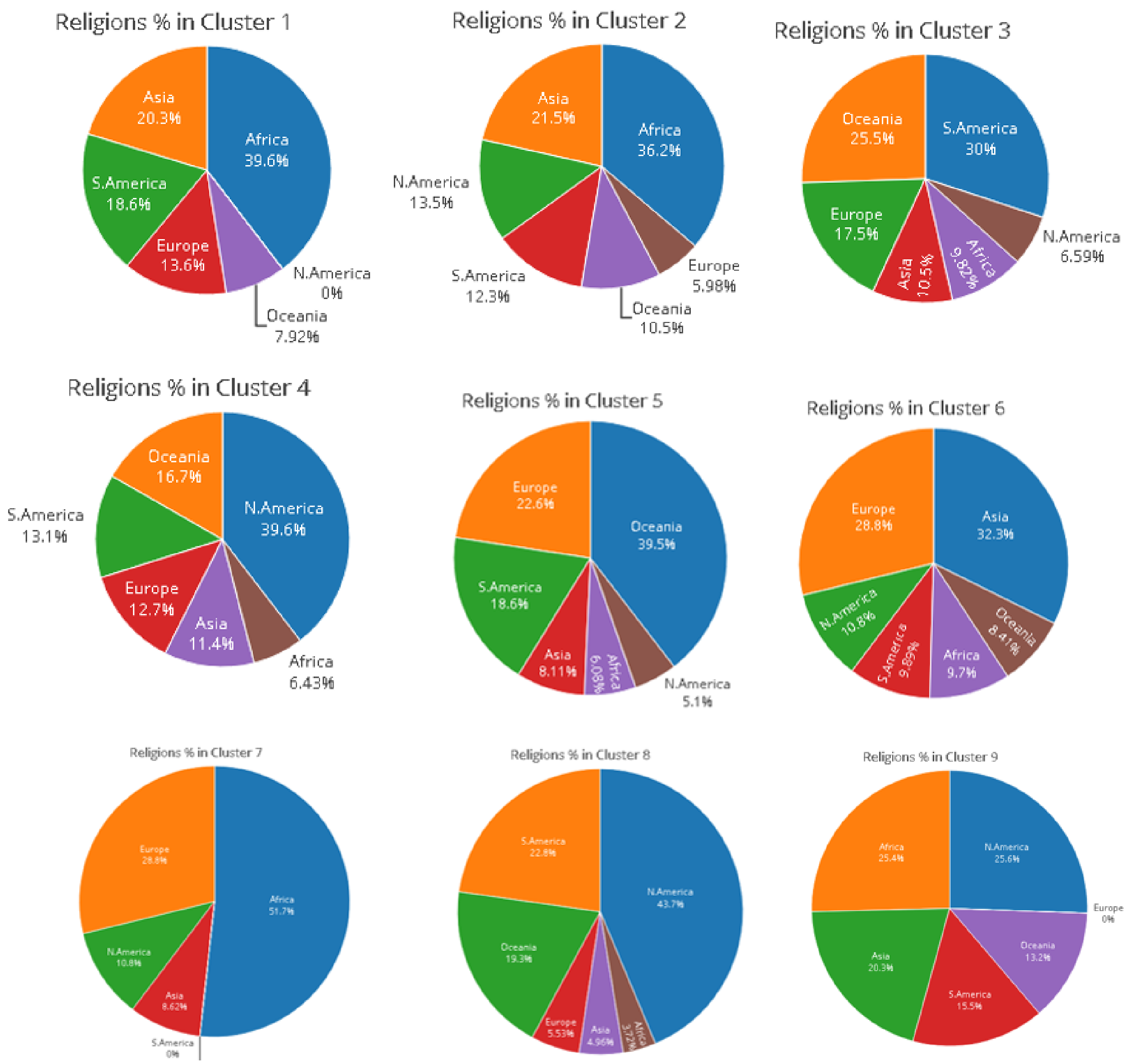

Figure 17 – Continents % per Cluster

If we summarize the above pies we have the following table (see in Table 5 -Clusters Summary Data)

Clusters ID	Dominant Continents	Dominant Religion
Cluster 1	Africa (39.6%), Asia (20.3%)	Ethnic (24.2%)
Cluster 2	Africa (36.2%), Asia (21.5%)	Muslim (34.7%)
Cluster 3	S.America (30%), Oceania (25.5%)	Catholic (27.5%)
Cluster 4	N.America (39.6%), Oceania (16,7%)	Hindu (21.7%)
Cluster 5	Oceania (39.5%), Europe (22.6%)	Others (60.4%)
Cluster 6	Asia (32.3%), Europe (28.8%)	Buddhist (30.6%)
Cluster 7	Africa (51.7%), Europe (28.8%)	Buddhist (28.1%)
Cluster 8	N.America (43.7%), S.America (22.8%)	Other Christian (45.8%)
Cluster 9	N.America (25.6%), Africa (25.4%)	Hindu (42.1%)

Table 5 -Clusters Summary Data

In the next table (see in Table 6 - Flags Sample for each Cluster) we have randomly sample 4 flags for each cluster in order to have a visual overview of the clusters (see in Appendix I, for an analytical overview of each cluster, Table 10 – Table 18). For example in the second cluster we have flags with 3 dominant colours which have the green as the most dominant.

Table 6 - Flags Sample for each Cluster

Finally, we have createD tables with all the flags of each cluster (see in Appendix I, the Table 7 - Countries of Clusters 1,2, 3 Table 8 - Countries of Clusters 4,5,6 Table 9 - Countries of Clusters 7,8,9)

Conclusions

In the first part of the project we had to find a method to classify the nation's flags in that way that we can identify the religion of the country based on their characteristics. We found that the best method was Classification Tree with Rpart library. Rpart package is a CART partitioning tree with different implementation. The Rpart programs build classification or regression models of a very general structure using a two stage procedure; the resulting models can be represented as binary trees. An example is some preliminary data gathered at Stanford on revival of cardiac arrest patients by paramedics.

In the second part of the project we had to create clusters of flags with common characteristics. At the beginning, in this method, we had to find the optimum number of clusters and then to assign each flag to one of them. We tried two methods (Hierarchical Clustering and PAM Method) and we evaluated the clusters by using Silhouette plots. Finally, we chose the PAM method as better for our occasion.

There are several distinctive characteristics by which we could **classify better the flags**. For example:

- Color separation types (e.g. two colours, tree colours)
- Flag shapes
- Predominant colour of the flag (Mainhue colour)
- Color symbolism (the same colour in different countries has different meaning)
- Symbols of the flags (for example, a cross, a star, a crescent)

So, we could have the following more clear clusters:

- **Cluster 1 -** One colour flags with a symbol in the center
- **Cluster 2 -** Two colour flags which divided into horizontal and vertical
- **Cluster 3 -** Three colour vertical stripes
- **Cluster 4 -** Three colour horizontal stripes
- **Cluster 5 -** Flags with diagonial division
- **Cluster 6 -** Flags which are relatively symmetrical
- **Cluster 7 -** Canton flags, these flags have one or two colours except the canton, e.g Greek flag
- **Cluster 8 -** Flags with vertical stripes at the flagpole
- **Cluster 9 -** Flags with the Scandinavian cross

- **Cluster 10 -** Flags with a separating triangle
- **Cluster 11 -** Multicolor flags
- **Cluster 12 -** Other flags, that are not in the above clusters

Appendix I

Figure 18 - Geometrical Characteristics

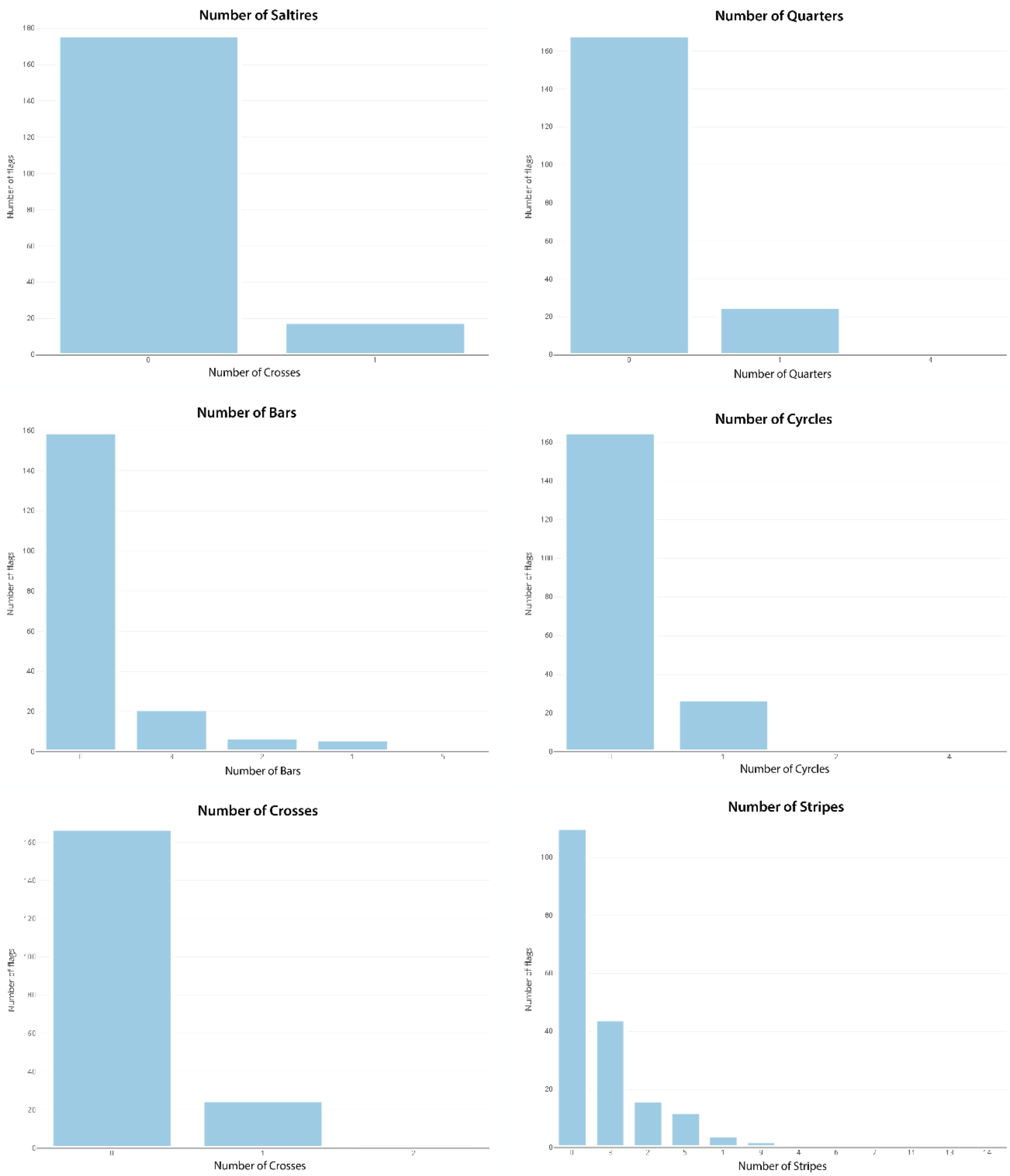

Figure 19 - Clusters Plot

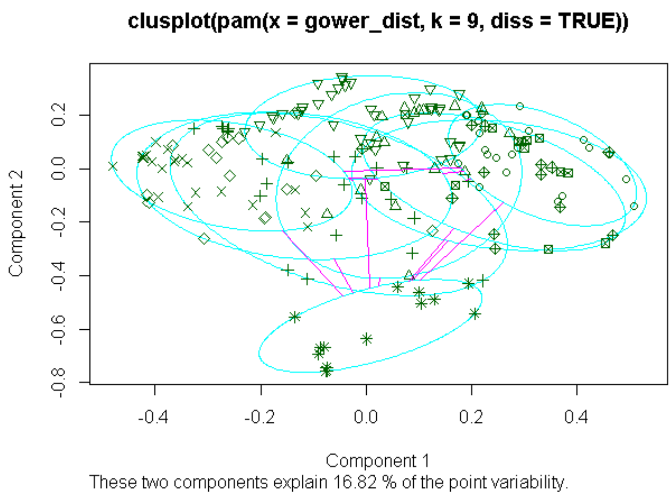

Table 7 - Countries of Clusters 1,2, 3

Cluster 1		Cluster 2		Cluster 3	
Afghanistan	Togo	Algeria	Benin	American-Samoa	Philippines
Albania	UAE	Bangladesh	Brazil	Andorra	Portugal
Angola	Vanuatu	Benin	Comorro-Islands	Antigua-Barbuda	Swaziland
Bolivia		Brazil	India	Bulgaria	Taiwan
Burkina		Comorro-Islands	Ireland	Burma	Uganda
Cape-Verde-Islands		India	Ivory-Coast	Central-African-Republic	USA
Congo		Ireland	Libya	Chad	Venezuela
Egypt		Ivory-Coast	Mauritania	Chile	Western-Samoa
Germany-DDR		Libya	Mexico	Colombia	Yugoslavia
Germany-FRG		Mauritania	Niger	Ecuador	
Ghana		Mexico	Nigeria	France	
Guinea-Bissau		Niger	Pakistan	French-Guiana	

Iraq	Nigeria	Saudi-Arabia	French-Polynesia
Kenya	Pakistan	Sierra-Leone	Kiribati
Malawi	Saudi-Arabia	Soloman-Islands	Liberia
Mauritius	Sierra-Leone	St-Vincent	Liechtenstein
North-Yemen	Soloman-Islands		Malaysia

Table 8 - Countries of Clusters 4,5,6

Cluster 4		**Cluster 5**	**Cluster 6**		
Anguilla	Micronesia	Australia	Austria	Luxembourg	Thailand
Argentina	Nauru	Cook-Islands	Bahrain	Malagasy	Tonga
Argentine	Nepal	Cyprus	Bhutan	Maldive-Islands	Trinidad-Tobago
Bahamas	Nicaragua	Czechoslovakia	Burundi	Malta	Tunisia
Barbados	North-Korea	Djibouti	Canada	Monaco	Turkey
Botswana	Qatar	Faeroes	China	Morocco	USSR
Costa-Rica	San-Marino	Finland	Denmark	Netherlands	Vietnam
Cuba	Somalia	Japan	Gambia	Norway	
Dominican-Republic	St-Lucia	Netherlands-Antilles	Gibraltar	Oman	
El-Salvador	Sweden	New-Zealand	Greenland	Peru	
Greece		Niue	Haiti	Poland	
Guatemala		Panama	Hungary	Puerto-Rico	
Honduras		South-Africa	Indonesia	Seychelles	
Iceland		South-Korea	Iran	Singapore	
Israel		Tuvalu	Kampuchea	South-Yemen	
Lesotho		UK	Laos	Spain	
Marianas		Uruguay	Lebanon	Switzerland	

Table 9 - Countries of Clusters 7,8,9

Cluster 7	Cluster 8	Cluster 9
Belgium	Belize	Brunei
Cameroon	Bermuda	Dominica
Ethiopia	British-Virgin-Isles	Equatorial-Guinea
Gabon	Cayman-Islands	Guyana
Grenada	Falklands-Malvinas	Jamaica
Guinea	Fiji	Jordan
Italy	Guam	Kuwait
Mali	Hong-Kong	Mozambique
Senegal	Montserrat	Papua-New-Guinea
Sri-Lanka	Parguay	Sao-Tome
Vatican-City	Romania	St-Kitts-Nevis
Zaire	St-Helena	Tanzania
Zambia	Turks-Cocos-Islands	Zimbabwe
	US-Virgin-Isles	

Table 10 - Cluster 1 Characteristics

Cluster 1					
Bars	0	**Black**	Yes	**Crescent**	0
Stripes	3	**Orange**	No	**Triangles**	0
Colours	4	**Mainhue**	Red	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	Yes	**Crosses**	0	**Text**	No
Blue	No	**Saltires**	0	**Topleft**	Red
Gold	Yes	**Quarters**	0	**Botrigth**	Green
White	No	**Sunstars**	1		

Table 11 - Cluster 2 Characteristics

Cluster 2					
Bars	0	**Black**	No	**Crescent**	0
Stripes	3	**Orange**	No	**Triangles**	0
Colours	3	**Mainhue**	Green	**Icon**	No
Red	No	**Circles**	0	**Animate**	No
Green	Yes	**Crosses**	0	**Text**	No
Blue	No	**Saltires**	0	**Topleft**	Green
Gold	No	**Quarters**	0	**Botrigth**	Green
White	Yes	**Sunstars**	0		

Table 12 - Cluster 3 Characteristics

Cluster 3					
Bars	0	**Black**	No	**Crescent**	0
Stripes	1	**Orange**	No	**Triangles**	0
Colours	3	**Mainhue**	Red	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	No	**Crosses**	0	**Text**	No
Blue	Yes	**Saltires**	0	**Topleft**	Blue
Gold	Yes	**Quarters**	0	**Botrigth**	Red
White	Yes	**Sunstars**	1		

Table 13 – Cluster 4 Characteristics

Cluster 4					
Bars	0	**Black**	No	**Crescent**	0
Stripes	1	**Orange**	No	**Triangles**	0
Colours	3	**Mainhue**	Blue	**Icon**	No
Red	No	**Circles**	0	**Animate**	No
Green	No	**Crosses**	0	**Text**	No
Blue	Yes	**Saltires**	0	**Topleft**	Blue
Gold	No	**Quarters**	0	**Botrigth**	Blue
White	Yes	**Sunstars**	0		

Table 14 - Cluster 5 Characteristics

Cluster 5					
Bars	0	**Black**	No	**Crescent**	0
Stripes	0	**Orange**	No	**Triangles**	0
Colours	3	**Mainhue**	White	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	No	**Crosses**	1	**Text**	No
Blue	Yes	**Saltires**	0	**Topleft**	White
Gold	No	**Quarters**	0	**Botrigth**	White
White	Yes	**Sunstars**	1		

Table 15 - Cluster 6 Characteristics

Cluster 6					
Bars	0	**Black**	No	**Crescent**	0
Stripes	0	**Orange**	No	**Triangles**	0
Colours	2	**Mainhue**	Red	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	No	**Crosses**	0	**Text**	No
Blue	No	**Saltires**	0	**Topleft**	Red
Gold	No	**Quarters**	0	**Botrigth**	Red
White	Yes	**Sunstars**	0		

Table 16 – Cluster 7 Characteristics

Cluster 7					
Bars	3	**Black**	No	**Crescent**	0
Stripes	0	**Orange**	No	**Triangles**	0
Colours	3	**Mainhue**	Gold	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	Yes	**Crosses**	0	**Text**	No
Blue	No	**Saltires**	0	**Topleft**	Green
Gold	Yes	**Quarters**	0	**Botrigth**	Red
White	No	**Sunstars**	0		

Table 17 - Cluster 8 Characteristics

Cluster 8					
Bars	0	**Black**	No	**Crescent**	0
Stripes	0	**Orange**	Yes	**Triangles**	0
Colours	6	**Mainhue**	Blue	**Icon**	Yes
Red	Yes	**Circles**	0	**Animate**	Yes
Green	Yes	**Crosses**	1	**Text**	Yes
Blue	Yes	**Saltires**	1	**Topleft**	White
Gold	Yes	**Quarters**	1	**Botrigth**	Blue
White	Yes	**Sunstars**	0		

Table 18 - Cluster 9 Characteristics

Cluster 9					
Bars	0	**Black**	Yes	**Crescent**	0
Stripes	0	**Orange**	Yes	**Triangles**	1
Colours	4	**Mainhue**	Green	**Icon**	No
Red	Yes	**Circles**	0	**Animate**	No
Green	Yes	**Crosses**	0	**Text**	No
Blue	No	**Saltires**	0	**Topleft**	Green
Gold	Yes	**Quarters**	1	**Botrigth**	Green
White	Yes	**Sunstars**	1		