Predicting Growth of Gross Domestic Product per Capita

A machine learning approach

Klaas Fiete Krutein: kfkru@uw.edu

Mason Gionet: mgionet@uw.edu

Michael Shieh: mikesh13@uw.edu

Submitted on: December 12th, 2018

Abstract

Gross Domestic Product(GDP) per capita is a common metric for measuring economic performance of a metric, as such its growth is closely tracked to evaluate the overall economic conditions of a given country. As with many other economic factors, GDP per capita is dependent on a variety of factors focused on the country itself, and also in the context of global trade and politics. We began our research into GDP prediction with a dataset from the World Bank, with data collected from 1960 - 2016 on over 1600 features. We applied machine learning techniques to perform feature selection and build predictive models on the dataset over 3 iterations, successfully improving our results each time. Ultimately we can conclude that GDP per capita is widely complex and it is inefficient to build a global model over such a long time period. With a dataset this large, varied and sparse, we can achieve improved prediction accuracy by building more localized models focused on clustering countries that are of similar economic makeup and subject to the same macroeconomic trends.

Contents

Introduction

Productivity and economic strength are a key components to enable high quality of life in a country and in general, to support human development. As outlined by the United Nations the often referred Human Development Index (HDI) comprises of three main segments: the life expectancy index, the education index and the gross national income index (GNI)[1]. The GNI is closely related to the frequently reported gross domestic product (GDP) and differs only in how it is accounting for incomes of foreigners in the country and residents out of the country[2]. Thus, the GDP can also be used as an indicator for human development. In times of humanitarian crises of international impact, caused by armed conflicts, prosecution of minorities, corruption, nutrition supply bottlenecks and climate change, being able to predict human development is a particularly important asset to counteract the root causes and efficiently target areas that are facing issues with improving. Models that can simulate how the human development will change given the current situation are therefore of value.

While there is an entire organization landscape that aims at identifying regions with support needs, this project approaches the topic from a data analytics perspective and aims to provide a tool that can give a first overview which countries of the world will face significant issues in increasing their quality of life. However, this project is not designed to be used as a sole source of improvement possibilities, and should only be used as a framework to targeting this issue from a data perspective.

Problem Statement & Data Sources

Choosing an appropriate parameter to contribute to predicting human development growth is critical, as the HDI is a relatively recent index that did not emerge until the 1990s[1]. Thus, a more trustworthy indicator that strongly influences the HDI was identified as the GDP. To account for different population sizes, the GDP per capita is of particular interest. Furthermore, since the highest interest lies in the potential of growth, the GDP growth per capita is chosen as the dependent variable to predict. Hence, the objective of this project is to predict the GDP growth per capita for countries worldwide.

In general, predicting any kind of human development based on data is challenging as data quality is of particular importance. However, on many parameters, there are no international standards how they have to be calculated and whether they have to be recorded. As a result, data is often biased and sparse. While searching for data that can help to build prediction models, a dataset on world development indices by the World Bank [2] was assessed to be the most trustable and complete dataset that could be found, as it includes data from 1960 to 2016. While the number of non-available data is still large, the database should account for differences in calculations in a sufficient manner. The database reports 1600 unique indicators from 264 countries and geographic regions that can be used to predict the GDP growth per capita. However, the above mentioned sparsity of the data set requires pre-processing.

First, some basic assumptions have to be formulated:

1. The resulting model should be independent from time series effects and thus the year should not be used as a predictor

2. Using the country as a predictor might lead to a high prediction accuracy as the history of a country plays a role. However, including this as a predictor will make the models biased in a way that similar effects in other countries are given lower priority even though they could apply in the same way
3. Based on the sparsity it can be assumed that a large number of the 1600 indicators in the dataset will not lead to a sufficiently large sample size that can be used to build prediction models with reliable performance
4. Collinearity between response and predictor variables in the dataset could be existing

Since the data set is initially structured by year and then contains all the indicators sorted by country code and indicator in the style illustrated in Table 1.

Table 1: Schematic example illustration of raw data structure

Country name	**Country code**	**Indicator name**	**Indicator code**	**1960**	**1961**	**19...**
Saudi Arabia	ARB	Gross Domestic Product	GDP	value	value	...
Saudi Arabia	ARB	Gross Domestic Product per capita	GDP.PC	value	value	...
...	...	...	...	...	...	...

When reviewing this table from a data processing perspective, it is obvious that isolating values of each parameter and use them as predictors can be difficult. Considering the first and the second assumption, some data wrangling that transposes the right part of the dataset was performed. To execute this, the "dplyr" package from the scripting language R was used to restructure the data frame based on the unique identifiers of the indicator codes and re-assign them to the correct values. Table 2 illustrates the resulting table. The related code can be found in Appendix I.

Country name	**Country code**	**Year**	**GDP**	**GDP.PC**	**...**
Saudi Arabia	ARB	1960	value	value	...
Saudi Arabia	ARB	1961	value	value	...
...	...	...	...	...	...

Table 2: Schematic example illustration of pre-processed data structure

Thirdly, considering the third assumption, the data set gets reduced by dropping all variables, which contain more than 30% non-available values. This way, it can be ensured that the sample size does not get reduced by variables that show a low data completeness.

Lastly, the issue with potential collinearity regarding the response variable is addressed through dropping all other variables that are related to the GDP.

Proposed Methodology

Throughout this report, the modeling process shall follow an iterative approach. Thus, the methodology section starts with model selection, which is followed by tuning and optimization of these models for the dataset. Thirdly, the results of these models trained on the training dataset shall be validated on the testing dataset. Lastly, an evaluation step shall check, whether the results are satisfactory, or whether another iteration is needed to achieve satisfactory results. This section will begin describing the reasoning for the initialization of the first iteration. Further iterations will follow, if necessary in the results section. Figure 1 illustrates this iterative process.

Figure 1: Systematic approach to model selection

Since our dependent variable is continuous, there are a significant amount of methods that can be used. Due to the high dimensionality of the data set, we need to consider multiple methods that can effectively reduce the complexity while maintaining high prediction accuracy. The approach of this section is therefore to apply multiple methods to the dataset, perform a cross-validation and compare the prediction errors and correlations between the true response and the predicted response in the testing dataset. Of particular importance is that the project will aim at global methods as it can be assumed that human development is a problem of global nature with similar patterns around the world. Therefore, local models like Kernel and KNN methods and also models that are based on separating the dataset using these methods like Support Vector Machines (SVMs) are not reviewed in the scope of this project. Despite these limitations, directly applying the dataset to a linear regression function can be risky as the large number of covariates might lead to unintended use of highly confounded factors. In addition, the high dimensionality increases the chances of just picking covariates that are significant just by coincidence.

Therefore, ways to reduce the dimensionalities can be applied. One approach is to use methods that reduce the size of the dataset itself and only keep variables that show a significant correlation with the response variable. This could be achieved with a LASSO regression model to select predictors, which can then be fitted to a linear regression model afterwards. Considering that some variables in the dataset are assumed to be highly correlated or irrelevant for the prediction of the GDP growth, this appears to be a valuable approach and should be the first method that should be tested.

Secondly, the high dimensionality of the dataset could be traced back to the assumption that it is actually of a lower dimensionality and includes underlying patterns that can be filtered out. The

statistical tool most suitable to this, is a principal component analysis (PCA). Performing a PCA followed by a linear regression model appears to be as another valuable approach.

Lastly, methods can be considered that are actually able to handle the complexity of a large data set and do not try to force all included parameters into one equation. Since decision trees are able to solve highly complex data structures, a random forest that accounts for the data size is selected as the third methods. While this might be a valuable prediction tool, it shows issues with interpretability compared to the simpler regression models mentioned above.

While the selected methods only are a subset of methods that could be used to approach this topic, they appear to be a good starting point and provide a decent spread of analysis techniques to find a good fit, as outlined in the introduction section. Based on the results given by the three individual approaches, the best model can be picked and selected as a proposed method to predict the GDP growth per capita. To ensure that the models processed in the following section are comparable in their results, the same split between training and testing data was maintained through a fixed list of row indexes that is used for the definition of the training data on all three approaches.

Analysis and Results

Initial iteration

We begin our analysis by dividing our now dense (or 'non-sparse') dataset into 80% training and 20% testing. This gives us the ability to create robust models, while still maintaining a set to check for unwarranted bias. We then focus on approaching the prediction via feature reduction & regression (LASSO and PCA) as well as a full scale tree model (random forest).

LASSO and Linear Regression

LASSO (Least Absolute Shrinkage and Selection Operator) is a statistics model which penalizes the beta coefficients in any given regression. Increasing the penalty (lambda) on beta causes certain coefficients to go to zero, thus leaving them out of any further model. As shown below in Figure 1, we can see many of our variables go to zero rather quickly with others lasting longer despite the increasing lambda value

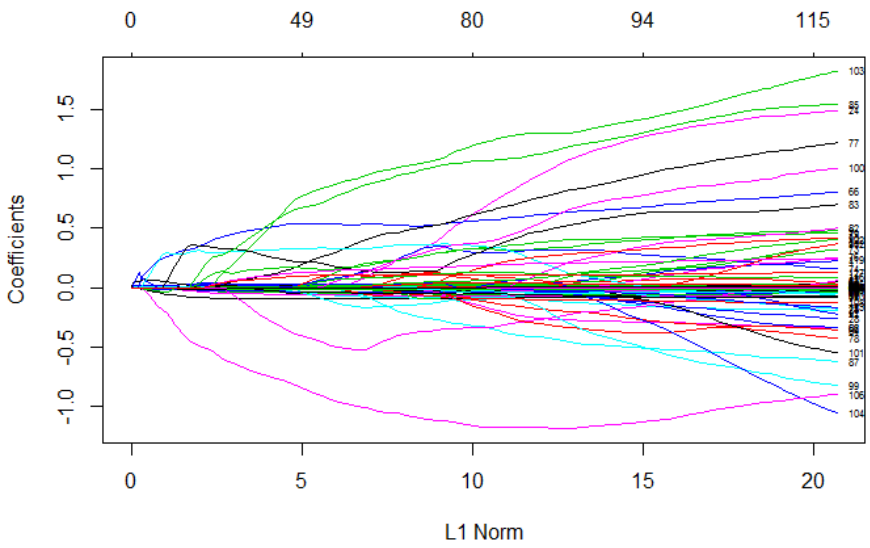


Figure 2: Plot of path solution trajectory

By using this approach, we reduce our modeling features from 1603 to 44, a reduction of 97%. This means our resulting model will be less computationally intensive to build, and contains only highly significant features. We then built a standard linear regression model with these variables. Testing the model against the testing set results in a mean squared error of 23.12, and a correlation of 0.323 between our predicted and actual GDP values. A scatterplot of actual vs. predicted GDP values showing the correlation can be seen below.

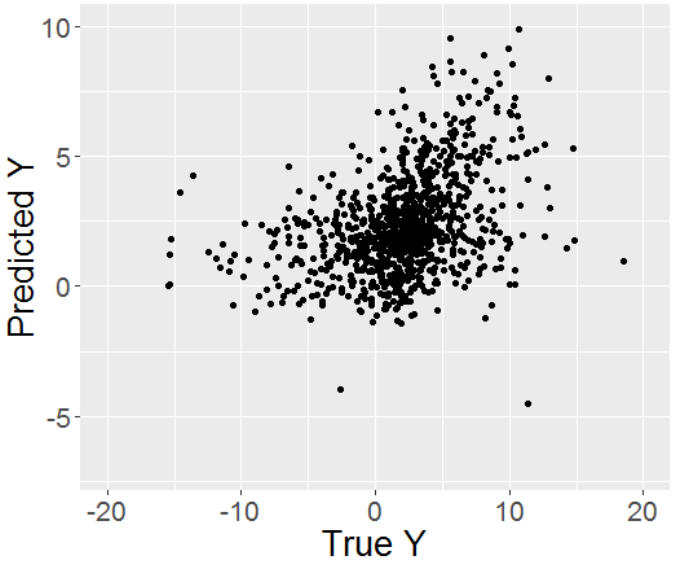


Figure 3: Comparison between predicted and true values of testing dataset

PCA and Linear Regression

Principal Component Analysis (PCA) is a method built on the assumption that, for a multivariate dataset that has many variables, the dimensionality of the dataset is smaller than it appears to be. Normally, we assume that in a model with n variables, those variables are n independent sources of variation that infuse uncertainty into the data. PCA, in contract, is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (entities each of which takes on various numerical values) into a set of values of linearly uncorrelated variables called principal components. In our case, there are thousands of variables that are recorded and potentially have related to the outcome, GDP growth index. We create 10 principal component to try to find lower dimensionalities and underlying patterns. Following are three plots that show the patterns of principal components.

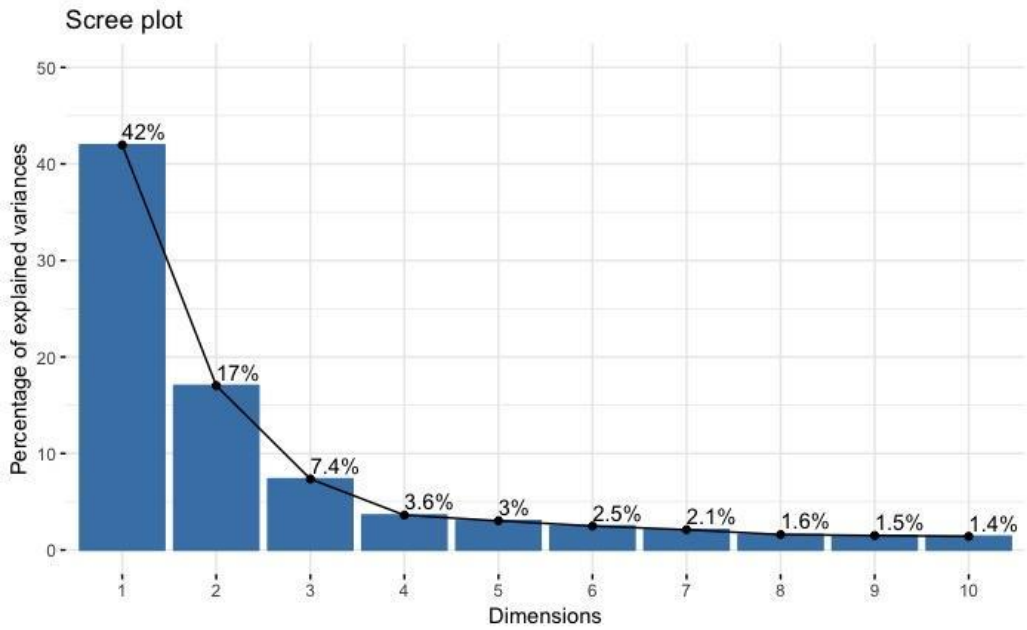


Figure 4: The percentage of explained variances from 10 principal components

As shown in Figure 3, There are three principal components that contribute most of the explained variances. Below are plots that shown the contribution of variables of the first principal component which explains a significant part of the variances.

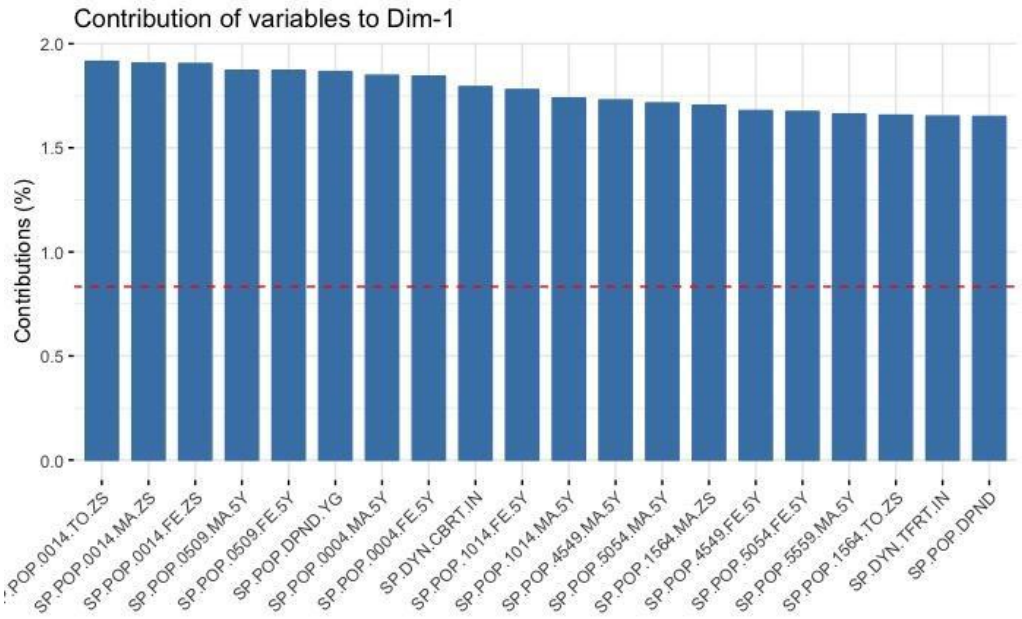


Figure 5: The contributions of variables for principal component 1

Afterwards, we put the 10 principle components into a linear regression model trained with training data, and summarized it as shown in Figure 5.

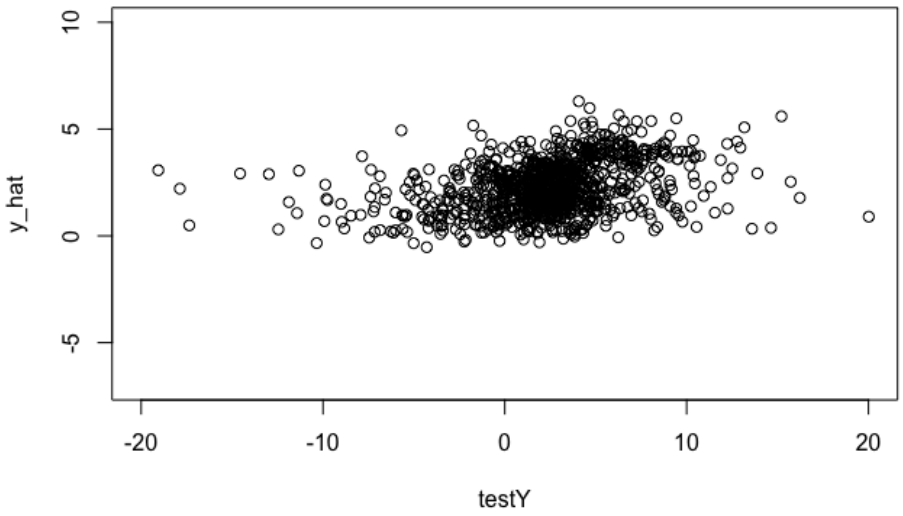


Figure 6: Comparison between predicted and true values of testing dataset

We use the testing data to predict GDP growths and find the correlations y_hat and testing data, and the mean MSE value of both. The correlation value is 0.303 and the mean MSE value is 22.862. From the principal components generated by the PCA model we can assume that there are some underlying patterns regarding the population data, since variables categorized in SP.POP are the main contributors for the principal components.

Random Forest

Random forest is a method that allows the prediction of a value, even if the dataset is complex and sparse. As a high number of random trees can be used to predict the response, it is more suitable than a simple decision tree. As already mentioned, the limitations are given by the difficulties to explain the underlying patterns and high computational effort. However, in our case we focus on prediction quality. To achieve a high prediction quality, tuning of the random forest parameters is necessary to obtain the optimal set that leads to the lowest error rate in the prediction of the test data. It was decided to tune for the number of trees, the number of features in the random forest and for the minimum bucket size of each node. At the same time, we want to avoid overfitting on the testing dataset itself. Thus, we combine our tuning process with a 10-fold cross validation on the training data. Based on this algorithm, the random forest with the lowest mean-squared error (MSE) is selected as the optimal model. The tuning values are defined as follows:

Number_of_trees = [50, 100, 150, 200, 250, 300]
Number_of_features = [5, 6, 7, 8, 9, 10, 11, 12] (interval between log(number of features in dataset)and $\sqrt{number\ of\ features\ in\ dataset}$)
Minimum_bucket_size = [20, 25, 30, 35, 40, 45, 50, 55]

In pseudo code, we can express the core part of the algorithm as follows:

```
For t in number_of_trees:
        For n in number_of_features:
                For b in minimum_bucket_size:
                        Select random value between 1 and 10
                        Perform k=10 fold cross validation on:
                                random_forest(n_of_trees=t, n_of_features=n, min_bucket=b)
                        Calculate MSE of resulting RF and save to data_frame
                End for
        End for
End for
Select minimum MSE from data_frame
Fit minimum MSE to test data
```

The full code in R is available in the appendix of this report. If we plot the produced models to visualize their MSE values, as can be seen in Figure 6, we can clearly identify the best model with properties of 100 trees, 11 features and a minimum bucket size of 55. This has been confirmed numerically by a data frame optimum search. The resulting MSE of this model is 23.90.

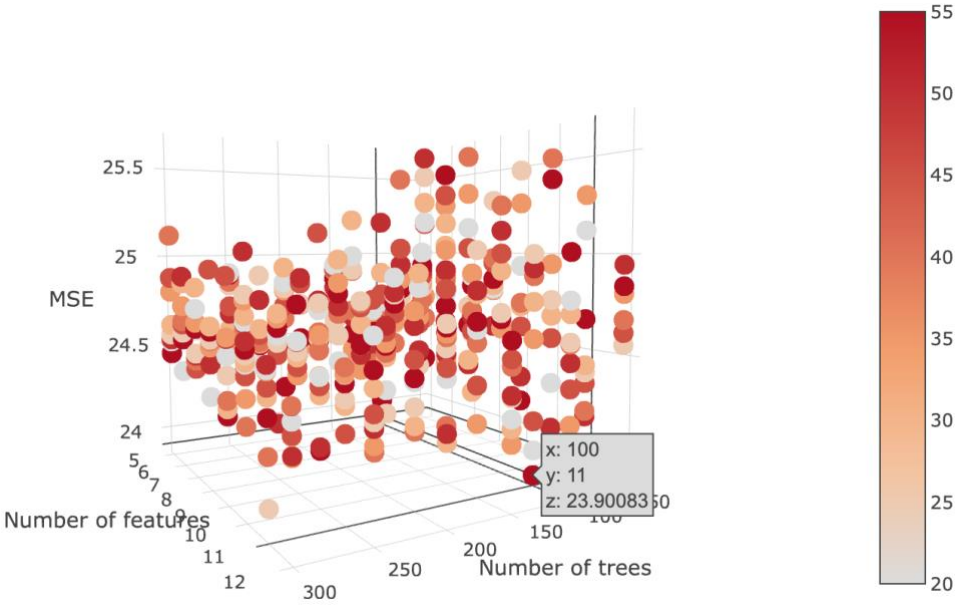


Figure 7: 3D-plot of generated random forest models based on number of trees, number of features and bucket size over MSE

When this model is fitted to the testing dataset, a clear correlation between predicted and true values can be found, as illustrated in Figure 7.

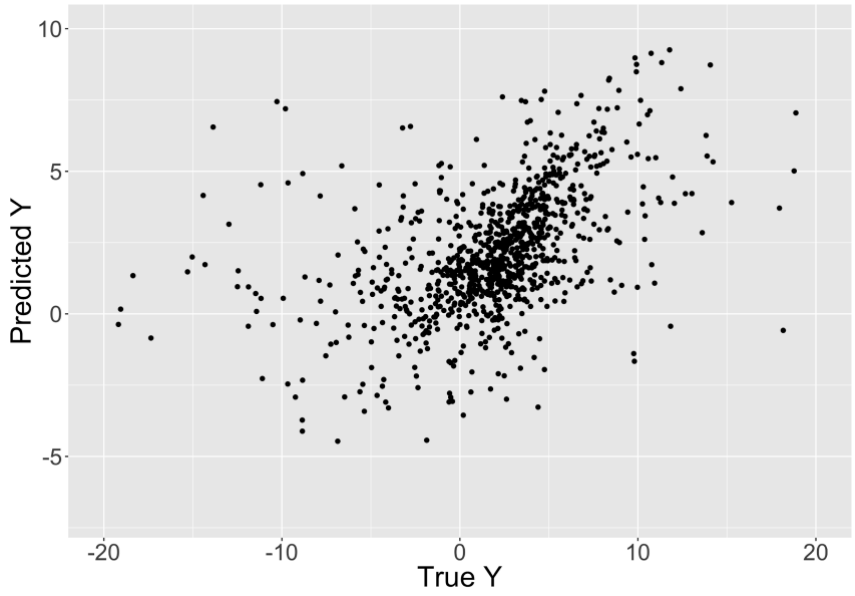


Figure 8: Comparison between predicted and true values of testing dataset

The correlation value between the two datasets is calculated as 0.399. While this is the best value that was achieved so far with the models, Figure 2 shows that the variance spread is high and that in some cases the prediction varied significantly from the tue growth values. This indicates that the dataset as it is might not describe the all influencing factors on GDP growth. This will be further discussed in the following sections.

Evaluation and Comparison of Initial Modeling Results

After approaching the dataset through three modeling approaches, we find the performance results in Table 3.

Modeling approach	**LASSO & Linear Regression**	**Random Forest**	**PCA & Linear Regression**
Correlation(y_pred, y_actual)	0.3229182	0.3996486	0.3035547
Mean Squared Error (MSE)	23.117314	23.90083	22.86172

Table 3: Performance comparison of initial modeling approaches

Given the performance measures, it can be easily identified that the three models all deliver similar results in prediction accuracy and correlations.

Insert discussion of the first two models here

For PCA & Linear Regression model, the performance of MSE is slightly better than other two models, but the correlation is not. This may be that although the principal components we create explain some underlying patterns, it is actually not an ideal tool for predicting our goals alone. If we also classify our dataset into different clusters, it may help our principal components to predict for higher correlation.

While the random forest showed a similar correlation of the model with the testing data set, it needs to be taken into account that the random forest modeling approach required a high amount of computational effort. Given the high dimensional cross-validation and the large number of tuning parameter values, the algorithm took about 15 hours to run to produce results for all tested configurations. This is an important disadvantage of this approach and should not be neglected when selecting this modeling strategy.

For all models, when reviewing the charts in the previous sections, the prediction quality is indeed, relatively low. While there are many reasons that could cause this prediction inaccuracy, one of the most likely ones is that the models are too exhaustive in the variety of countries with very different circumstances during different time periods they aim to cover. Thus, for the second iteration, methods are necessary that can account for different groups of models. Most practically, data clustering is a promising approach. Thus, the second iteration will re-evaluate the model prediction quality based on individual models that are fitted to identified clusters of the dataset. However, due to the high computational effort, the random forest approach is not being considered for a second iteration. While this would require parameter optimization for four different data subsets, the other two modeling approaches are more practical to use, while likely delivering a similar prediction quality.

Second Iteration

Cluster Modeling

Our next approach is to use mathematical clustering on the model to break down the dataset into smaller, more manageable subsets. Our goal of the clustering is to create more localized models, one

each cluster, to improve our accuracy. To cluster, we have to start again with our dataset. After removing all rows with too many missing values, we groupe all of the data by country name. We take the average values for each column. Then use the Mclust package in R to divide the cluster in 4 clusters which applies the Expectation-Maximization(EM) algorithm. Initially, we set G=1:9 which causes mclust to test the Bayesian Information Criterion for each different amount of clusters. We find that the algorithm recommended the use of just one cluster, which is essentially no cluster. Despite this result, we choose to test at G=4 to experiment with the effects of clustering on the dataset.

LASSO and Linear Regression on clusters

We then use LASSO regression as before to reduce the dimensionality individually for each cluster, and build linear regression models on each set on unique features. The mean squared error results for each cluster are shown below:

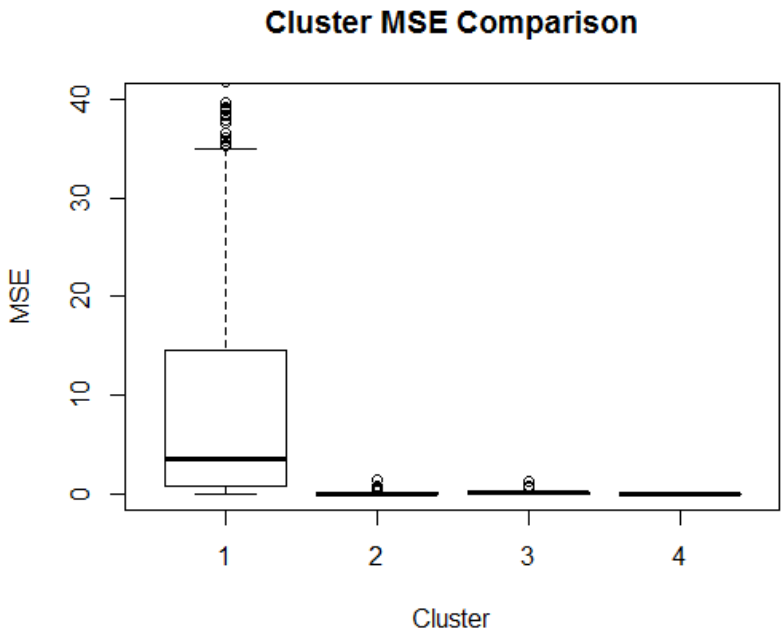


Figure 9: The responses per cluster for each clusters

Overall the results of this iteration are significantly better than those without the mathematical clustering. Across all clusters we have a mean squared error of 18.19, an improvement of 20%, a correlation between actual and predicted GDP values was 0.4349, improving 35%. Although there is a high discrepancy between our different models, largely given the change in sample size. As shown in the scatterplots for the models below, cluster 1 is trained on sample size of 3611 with mean error of 20.09 compared with cluster 4 trained on 9 data points resulting in a mean error of nearly 0 (5.981938e-29). This imbalance in sample size causes very high accuracy in clusters 2-4 as we have a very small sample size and likely overfitting in those models.

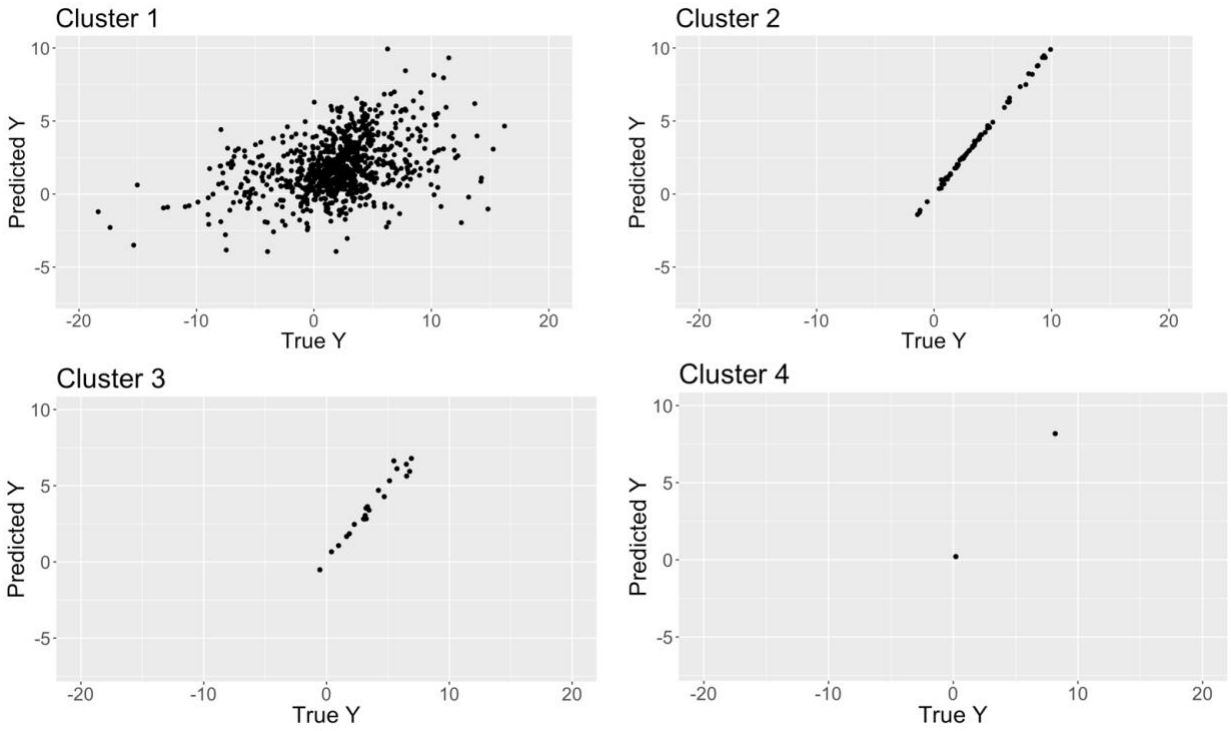


Figure 10: Scatterplots of results from each cluster.

PCA and Linear Regression on clusters

In the case of principal component analysis, the performance we obtain is also not satisfactory. Therefore, we transform our variables into several principal components using PCA. The number of responses per cluster are shown below, the x-axis represents the different clusters, the y-axis represents the actual GDP growth for each data point within the dataset. As we can see, most of the data are classified into cluster 1 and cluster 2.

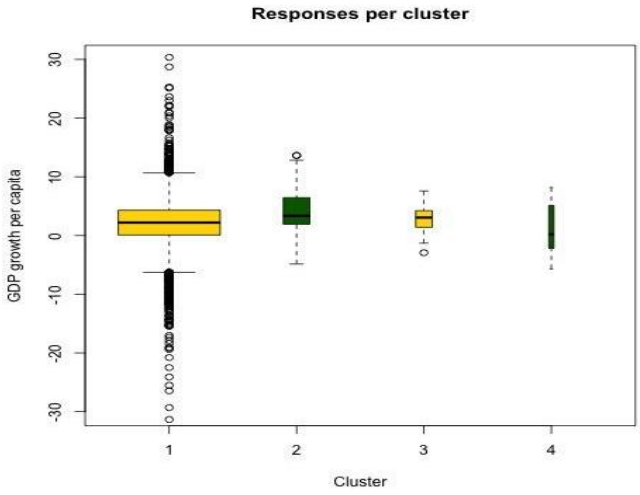


Figure 11: The responses per cluster for each clusters

Next, we cut the dataset by its classification into four different sub-dataset to perform principal component analysis. The scree plot below shows the explained variances by the principal components we created.

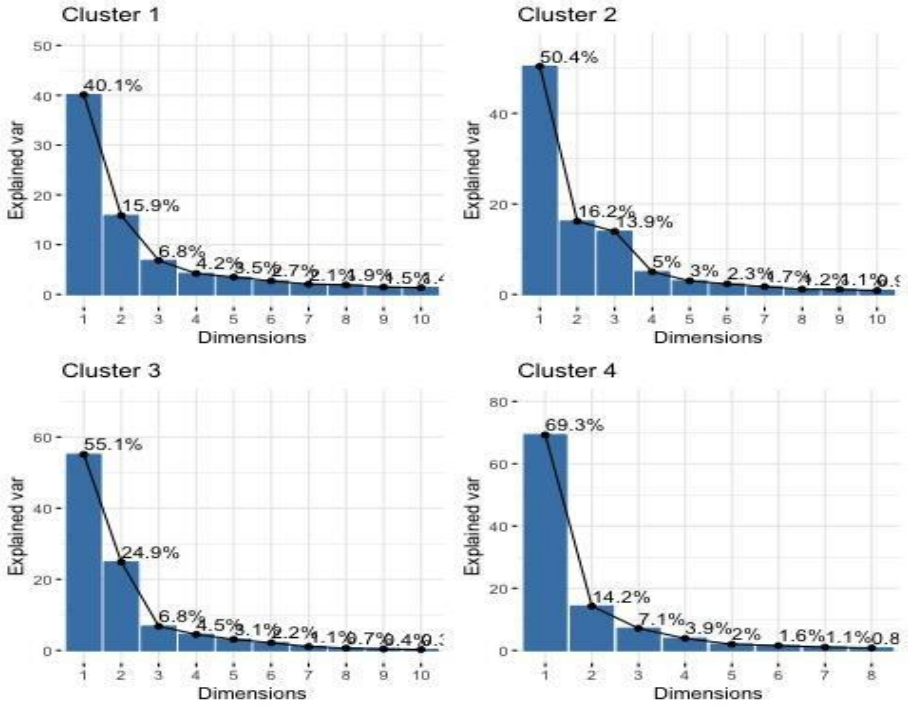


Figure 12: The explained variances for each cluster

After creating principal components, we again put them into a linear regression model to try to predict with those principal components.

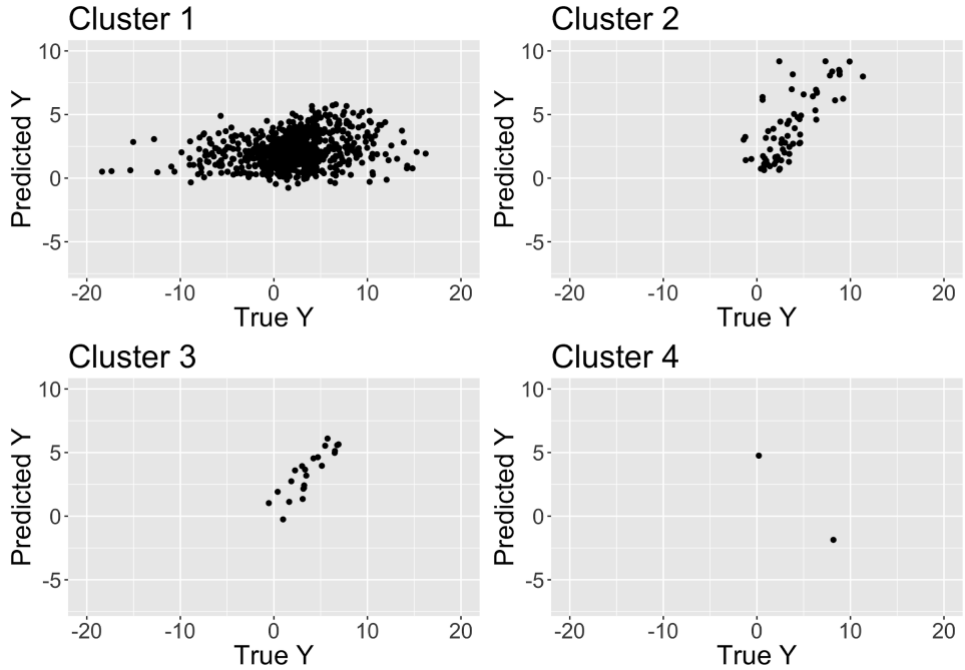


Figure 13: The scatterplot between True Y and Predicted Y for four clusters

The scatterplot above in Figure 13 shows that while the MSE for all cluster has improvements, the correlation for cluster 1 does not improve dramatically and the correlation for other three clusters improve significantly. However, since cluster 4 only has two data points, the significance may be misleading and the same can be suggested for cluster 3. For cluster 2, the performances of MSE and the correlation are really well. Below shows the scatterplot for four clusters that clearly shows cluster 2 and cluster 3 has a clear positive relationship. The result clearly shows that there are a lot of outliers for cluster 1, it could be the reason why for cluster 1 has such low correlation and a higher MSE compare to cluster 2 and 3. If we clear out the outliers, or classify our dataset into different kinds of clusters, the performance of this model can definitely improve.

Evaluation and Comparison of Second Iteration Results

From both of these methods, we saw an improvement in our mean squared error and an increase in correlation values. Across each cluster we saw that LASSO offered better results in both categories. Despite these improvements it is hard to find a heuristic meaning from the clustering since it was only based on the results of the EM algorithm. For example, the 4th cluster is composed only of data points from Qatar. Unfortunately the algorithm does not give us a clear why to distinguish why Qatar's annual GDP growth per capita differs so drastically from all other countries in our data.

Cluster	**1**		**2**		**3**		**4**	
Modeling approach	**LASSO**	**PCA**	**LASSO**	**PCA**	**LASSO**	**PCA**	**LASSO**	**PCA**
Correlation (y_pred, y_actual)	0.385	0.208	0.9939	0.7727	0.9786	0.869	1	-1
Mean-Squared Error (MSE)	20.195	22.79	0.1059	3.97	0.1809	1.12	1.5386e-28	60.766

Table 4: Performance comparison of secondary modeling approaches

Third Iteration

From the results of our second modeling iteration, we come away with two key takeaways. First, that clustering effectively decreases our error and increases the correlation between actual and predicted growth values. Second, despite the positive results, this clustering method is not very helpful for understanding which factors have an influence on driving changes in GDP. Therefore we choose to look back at the data source to find a useful metric for classification.

Clustering via Income Classification

The world bank have actually already broken our dataset into 4 income levels (high, upper medium, lower medium, and low) based on current income of the inhabitants. We choose to use this label to cluster our dataset and to answer a new question: How is GDP growth affected influenced by current income levels?

By adding this new classification method to our data, we break the overall set again into 4 clusters. Using this new method, we achieve a much more even spread in sample size, with the distribution in sample size becoming much smoother, with almost 25% of the total in each category.

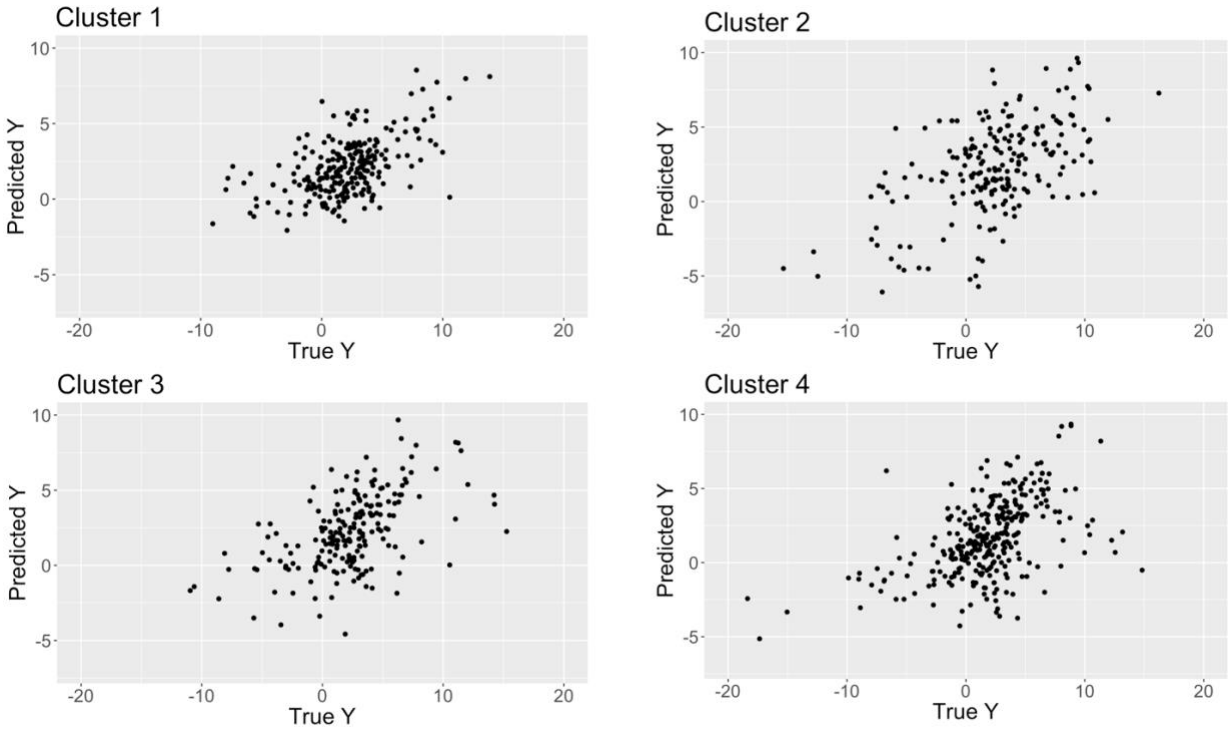


Figure 14: Correlation comparison across clusters.

We then run individual LASSO regression on each cluster and built new linear regression models for each. As shown in the two graphics below, this significantly improves our results. Our mean squared error across all clusters drops to 15.53895 and the correlation rises to 0.5544548. Again in these results we do not remove any outliers, which if done would improve our results even further.

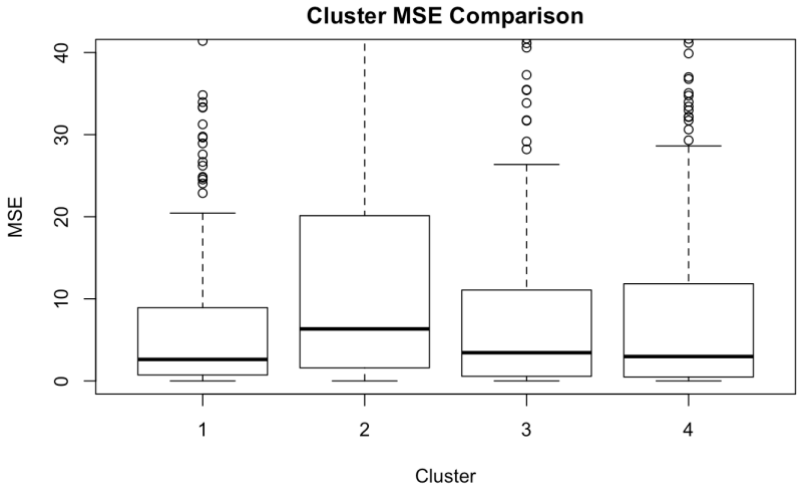


Figure 15: Performance comparison across clusters

Discussion and Conclusions

Evaluation of Final Results

In general, it can be discussed whether the low significance of the correlation between the predicted values and the true values in GDP growth should not raise other concerns. This indicates, that despite its large size, the dataset as used might not contain sufficient information to reliably describe GDP growth. Ways to address this prediction error would be testing for the effects of country and year, that were excluded in the beginning. These factors can account for time and regional effects. However, this might make it difficult to predict the values of future years, as obviously, no data is existing for future years yet, that could be used to make predictions.

Depending on the outcome of these alternative approaches, additional modeling techniques, such as local models can be explored. The reason why this could be of value is that certain factors might only support GDP growth up to a certain level. To illustrate that further, an example could be literacy rate of the public. A higher value in literacy rate might support GDP growth up to a certain threshold, while only larger higher education rates might actually correlate to higher GDP growth values. Hence, it could be of use to explore KNN or Kernel methods.

While using the same dataset, future work could also limit the size to more recent years, excluding parts with relatively sparse data. That way, more parameters could be used to predict GDP growth. Another angle of approach could be to interpret GDP growth as a result of a summary of events that happened over a longer time span than just one year. Long-term political decisions, previous incentives and local circumstances can influence the dynamics of a country over decades. That way, time-series analysis might be more suitable to predict this growth.

Based on what can be achieved using these alternative modeling approaches, it can be discussed whether the dataset itself might not be suitable in general, or whether there are factors that influence GDP growth that are not easily quantifiable and more of a qualitative nature. An example for this could be a seasonal trait like fear of certain political changes. Political instabilities and fear of e.g. mass migration, nuclear conflicts, food supply shortages, etc. that are not visible in reality but change the market behavior of the public could be data that is difficult to quantify and use as a predictor.

In general, it can be discussed, whether it is even desirable to be able to predict GDP growth of certain countries with high confidence. This could open the door to market speculations that could have even worse effects than activities in the 1990s where investor groups put market bets against currencies that let entire countries struggle and thus these tools could result in damage to society at large.

Lessons Learned

This project provided an opportunity to explore the opportunities machine learning methods offer in a very effective way. In particular using a large dataset showed quickly that there is always a trade-off between model complexity and computational efforts that need to be taken to achieve a high prediction accuracy, which becomes a time constraint given fixed computing power. While this dataset is still of

limited dimensions, it becomes clear what effect it can have to run models like random forests on datasets of a size of multiple terabytes in terms of computational effort and how powerful computers need to be to handle this high model complexity in a time frame that is useful in practice.

Furthermore, the iterative approach of this study illustrated the basic data wrangling and modeling approach. While to some extent, initial tuning of parameters for model building is valuable, the logical consequence out of the interpretation of the results is to execute further modeling iterations that aim at addressing the issues raised by the previous iteration. This is a very important lesson in analytics, as there are many tools for building statistical models, with so set selection criteria of what model to use for what problem. By testing multiple different models you gain more knowledge about your dataset, and how that model reacts to the nuances of the problem. The modeling approach presented in this paper therefore showed to be justified, as we continuously improved our results.

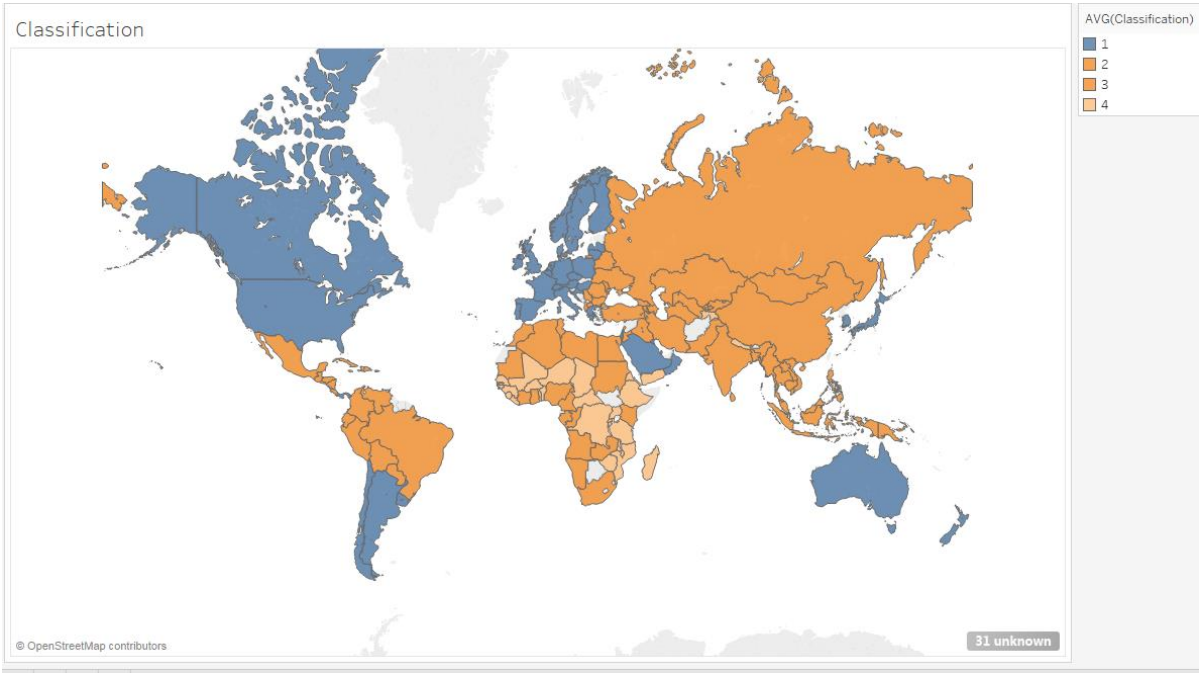
For the authors, a rather disappointing learning was that modeling quality can be optimized while still not leading to a satisfactory prediction quality, as all modeling outcome depends inherently on the ability of the data source to describe and explain the research question. This leads to the learning that preliminary data source selection requires significant attention when designing a research study.

It also was important to learn that when asking a broad questions, one can expect a broad answer. As shown in our third iteration, when we attempted to refine our prediction to subsets of the population we achieved much better prediction accuracy. This is a key takeaway to apply to future research questions: when building a predictive model, it is key to refine your questions on the dataset before building your models. This can be applied to the dataset as well. If we would have started by doing more exploratory data analysis, and found that with much less variability in the developed nations, and greater growth rates overall in developing nations we could have started building more localized models sooner, and would've had more time to refine that approach. For future analytics problems, this will be an important place to start, especially through leveraging visuals and background research into the dataset.

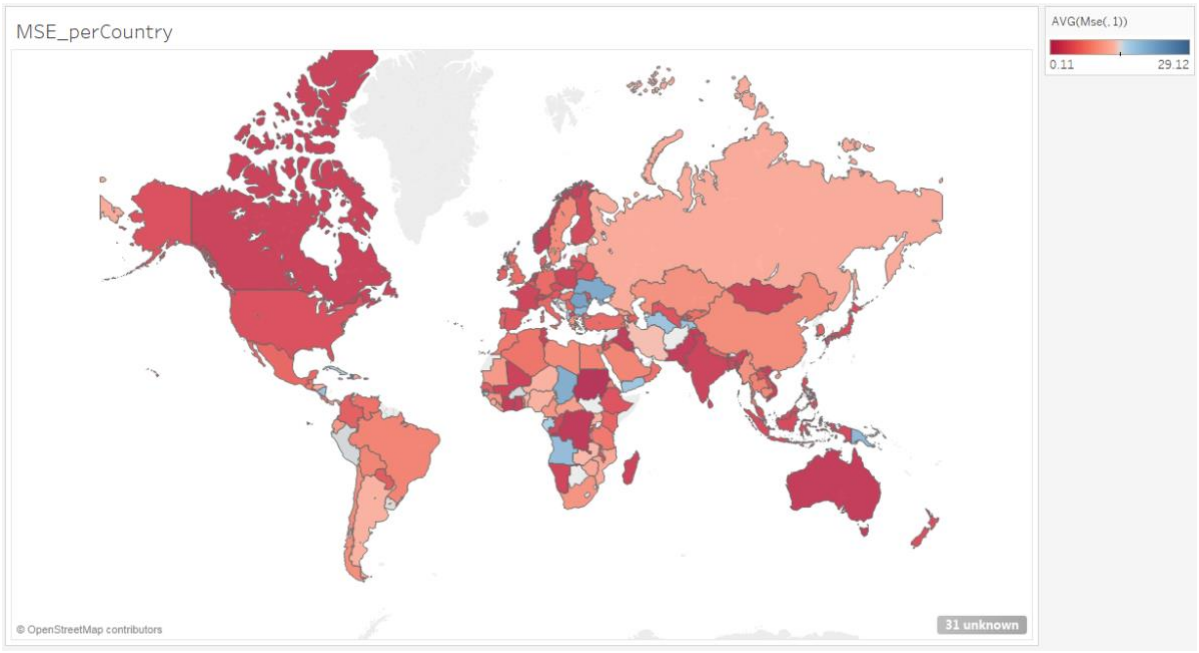
Overall this was a fantastic class project that leveraged our theoretical classroom learning to real data science problems, exposing us to many common issues faced by professionals in industry. By dealing with sparse data, ambiguity and high model complexity we are more prepared for entering into professional applications of machine learning, regardless of the industry.

Appendix

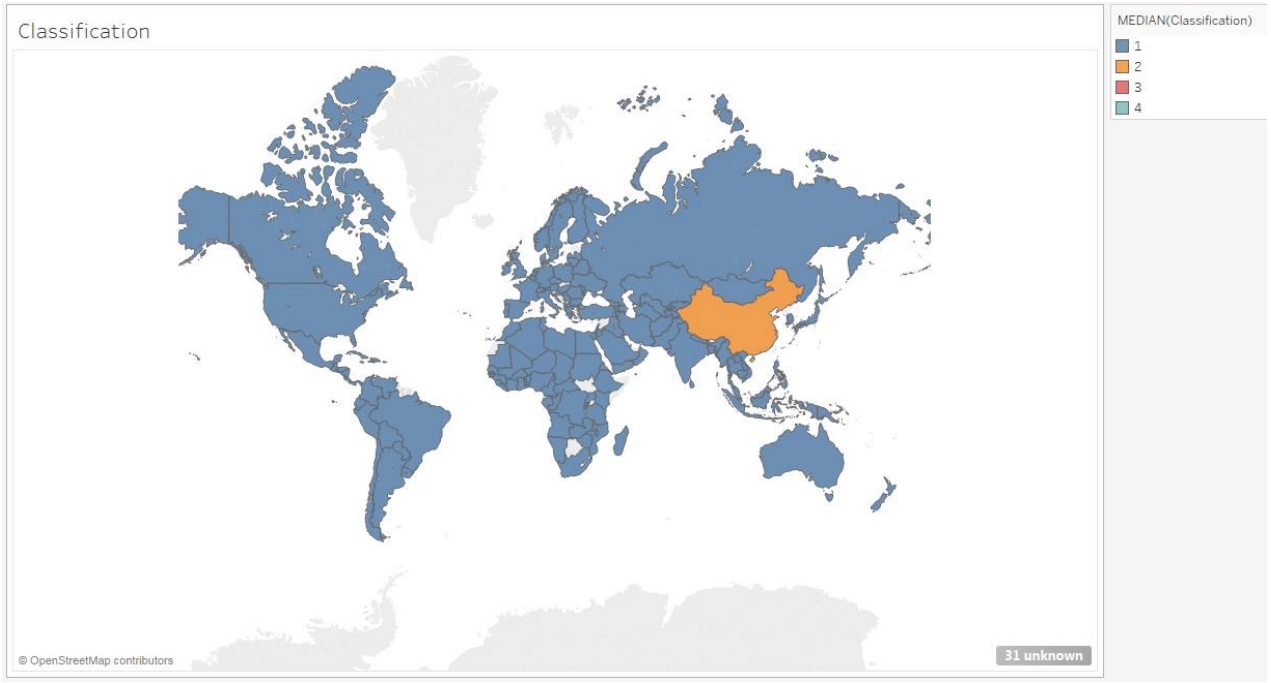
Clustering 1 - Classification



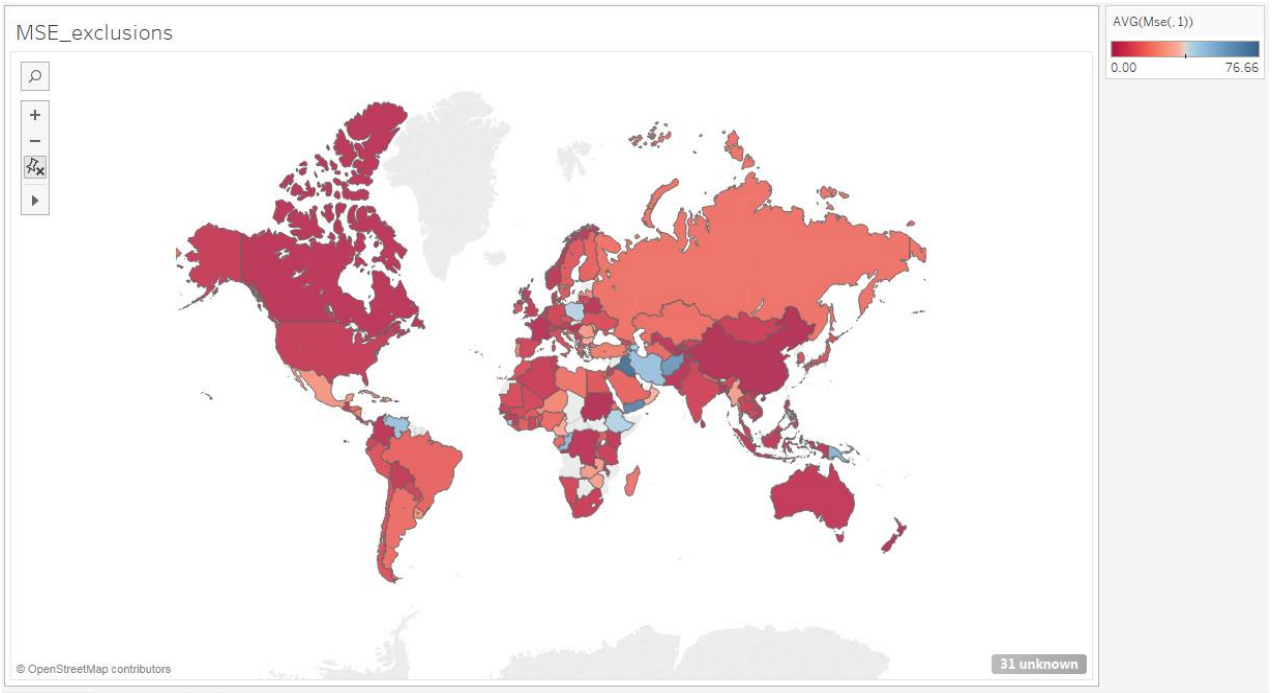
Clustering 1 - MSE (LASSO)



Clustering 2 - Classification



Clustering 2 - MSE (LASSO)



Bibliography

[1] United Nations Development Programme, “About Human Development,” *Human Development Reports*. Available: http://hdr.undp.org/en/humandev. [Accessed: Dec 11 2018]

[2] The World Bank, “World Development Indicators,” World Development Indicators datasheet (CSV), June 2010 [Revised Nov. 2018]. Available: https://datacatalog.worldbank.org/dataset/world-development-indicators. [Accessed: Dec 11 2018]

[3] K. Amadeo, “Gross National Income,” The Balance, Dec 11, 2018. Available: https://www.thebalance.com/gross-national-income-4020738. [Accessed: Dec 12 2018]