

Table of Top/Bottom 5 countries in Change in some Variable

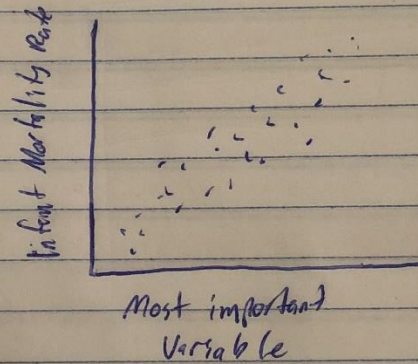
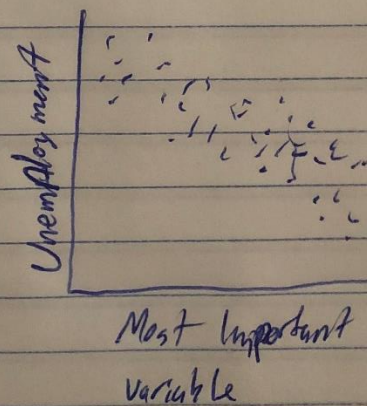
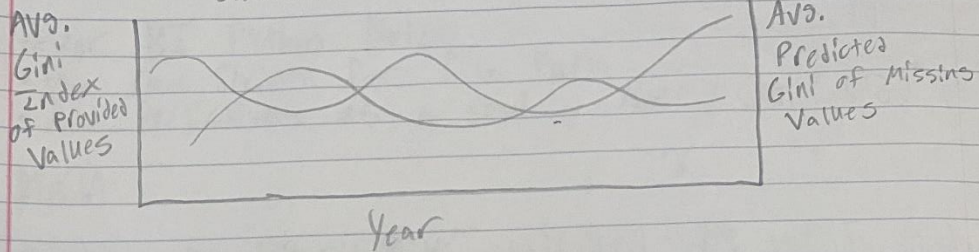


Table of Top 10 countries in difference between actual/Expected values

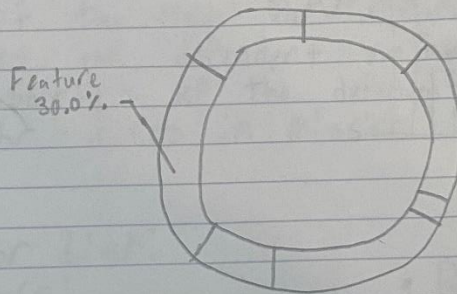


## Comparison of Provided and Predicted Gini's

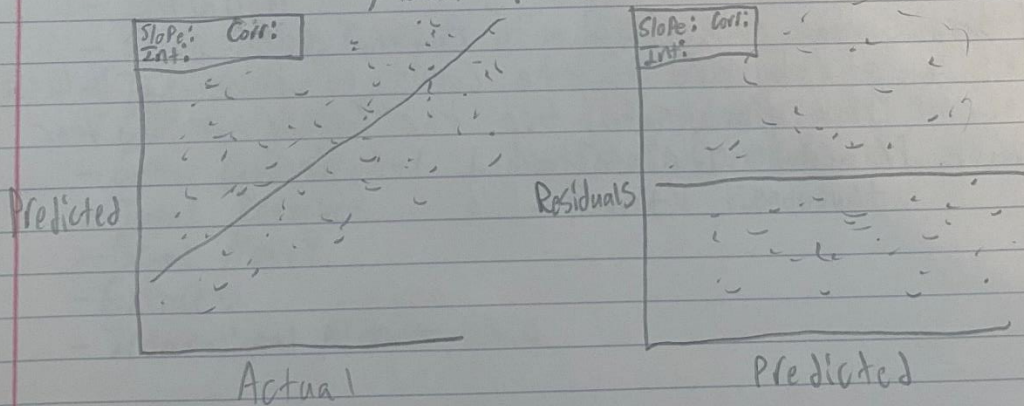


Shows trends of missing data vs. trends of provided data

## Feature Importances



## Model Performance





DF =

$(std * imp) + mean$

### Gini Index ML Predictions Visualizations



### Map of Missing Values

"World Map"

- Bubbles for each spot with missing values
- Size of bubble is the count of missing values
- Scroll over bubble to see the country and year with predicted Gini attached to year

## Gini Index

1990 x x x | [ ] | x x x

2020 x x x | [ ] | x x x

## Infant mortality

1990 [ ] | x x x

2020 [ ] | x

## Life Expectancy

same mock graphic  
as above

## Unemployment

same mock graphic  
as above

## Documenting and Responding to Visualization Napkin Drawing Feedback

Here is the feedback that we received about our original napkin drawings for visualizations:

[Yesterday 5:27 PM] Caitlin Ryan

It looks like you have multiple versions of showing the most important features and are probably choosing just one. I think one of the bar graphs will do a better job showing the information than the donut chart since it is easier to see the difference between size of the bars especially when they are sorted in descending/ascending order.

I don't know what exactly you are predicting with your model but it looks like there were a few scatterplots that were actual v. predicted, I know one was life expectancy. But you also had different things regarding the Gini coefficient so I'm not sure if that was also an actual v. predicted, I know there is one that has a line chart of the average of provided values v. predicted Gini for missing values. And there's also a map for the missing values, so I am guessing you are predicting the Gini value based on different known factors, I do like the idea of being able to see where factors are missing and what a predicted Gini value would be.

I do like the idea of a table showing the top 10 countries with the biggest gap between predicted and actual Gini values.

The last bit I noticed, the box and whiskers plot and scatter plots of most important variable and some value, they look like they are going to start adding too much to the dashboard. I don't know if you want to pick one thing to look at both a box and whiskers and a scatter plot of it since I think even choosing one style and including 4 versions will still add clutter.

[10:44 AM] Dan Angelina

Like caitlin said i noticed many visuals for features and agree that the bar chart is the best option for that topic. For the life expectancy with the "star", is that information all predicted or is it just one point on the visual that you are using the ml model to predict. I like the idea for showing trends of missing values and showing differences between countries as context for the information your sharing.

[10:44 AM] Dan Angelina

I did notice that you had a lot of visuals and you might want to consider narrowing down what you are looking at before putting too much on the dashboard.

[10:46 AM] Dan Angelina

I like the use of different visual models but dont let that cause you too make unnecessary visuals or share info that may not be useful in answering the questions that you are looking to answer.

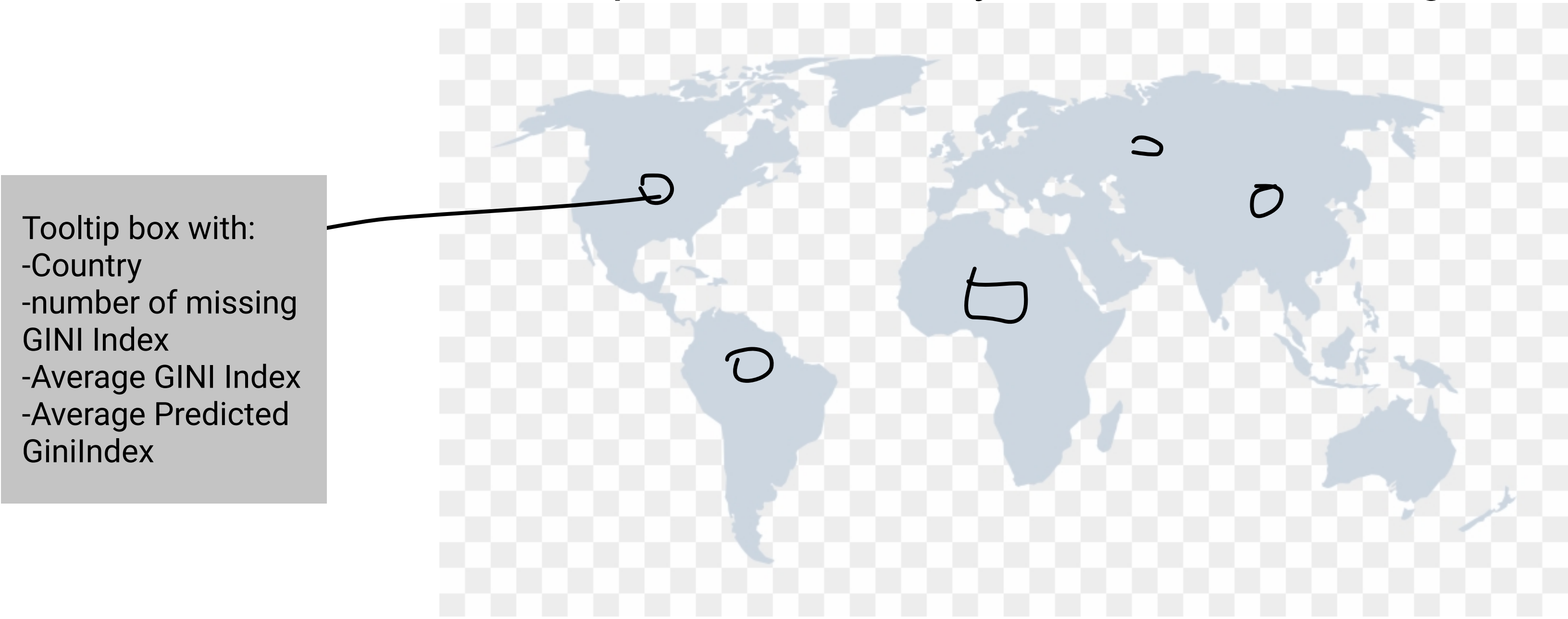
## **Responding to Feedback**

Based on this feedback, we made a couple of major choices about our visualizations:

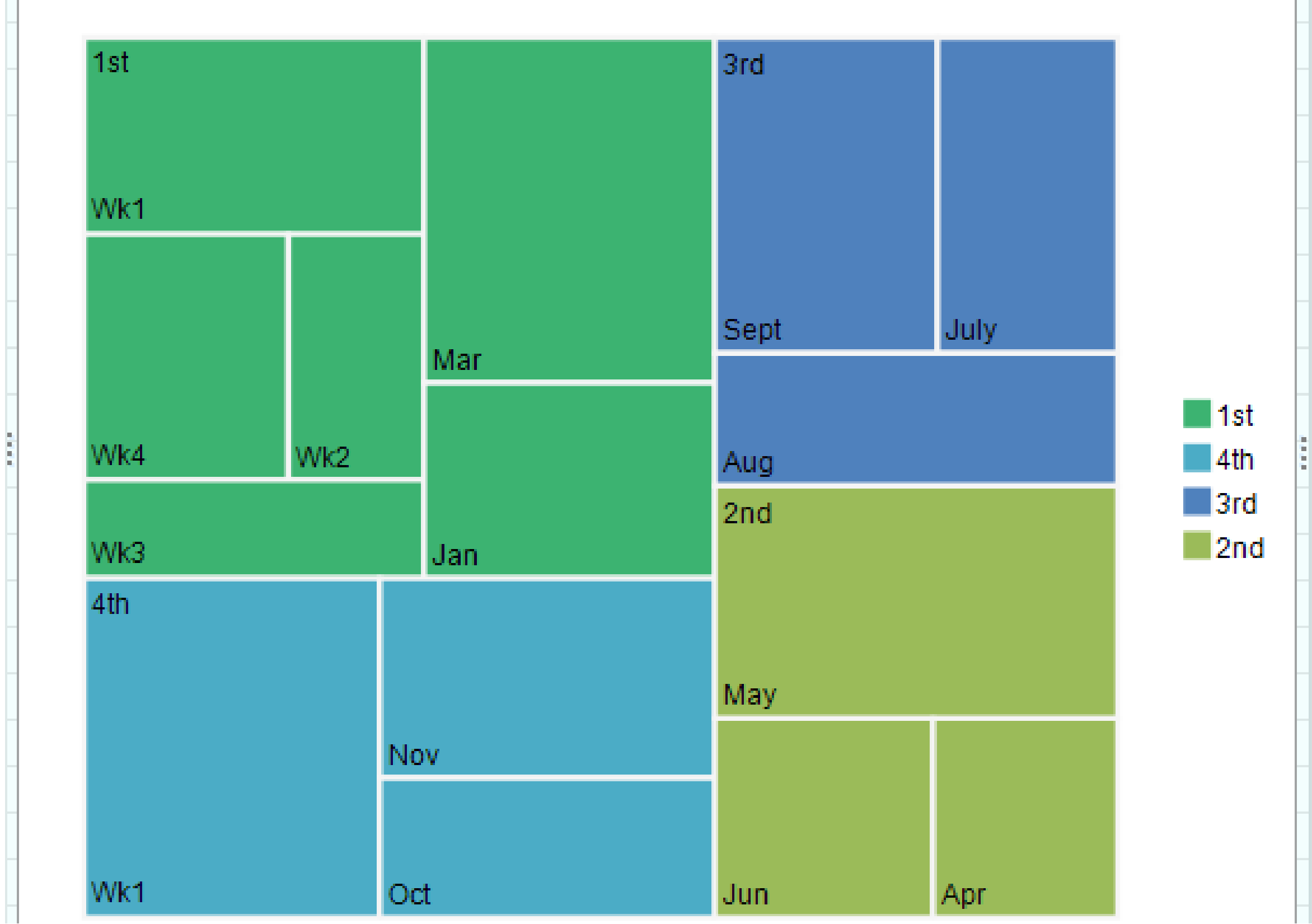
- 1) We need to reduce the number of dependent variables that we're working on. There are too many different graphs right now, and that actually makes our story harder to understand. So, we've decided to look at only Gini Index and Unemployment Rate, rather than looking at all four dependent variables.
- 2) In general, we need to make our labels clearer and our graphs easier to interpret. Where possible, we've added labels, converted donut charts to tree maps, and tried to order our data in an intuitive way.



World Map of Countries by Number of Missing Gini Index



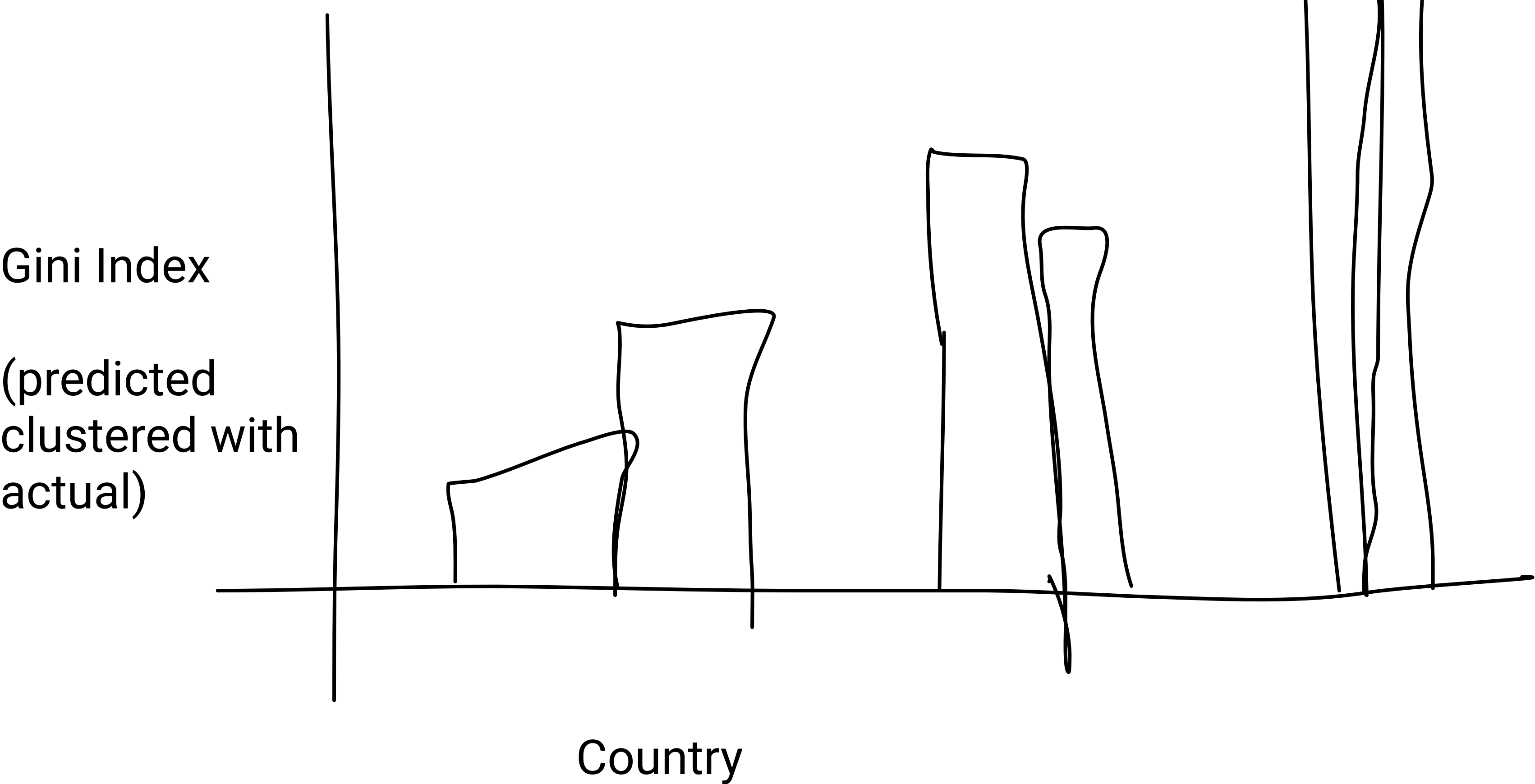
Tree map of GINI Index Model Features by Feature Importance



Actual GINI Index vs. Predicted GINI Index

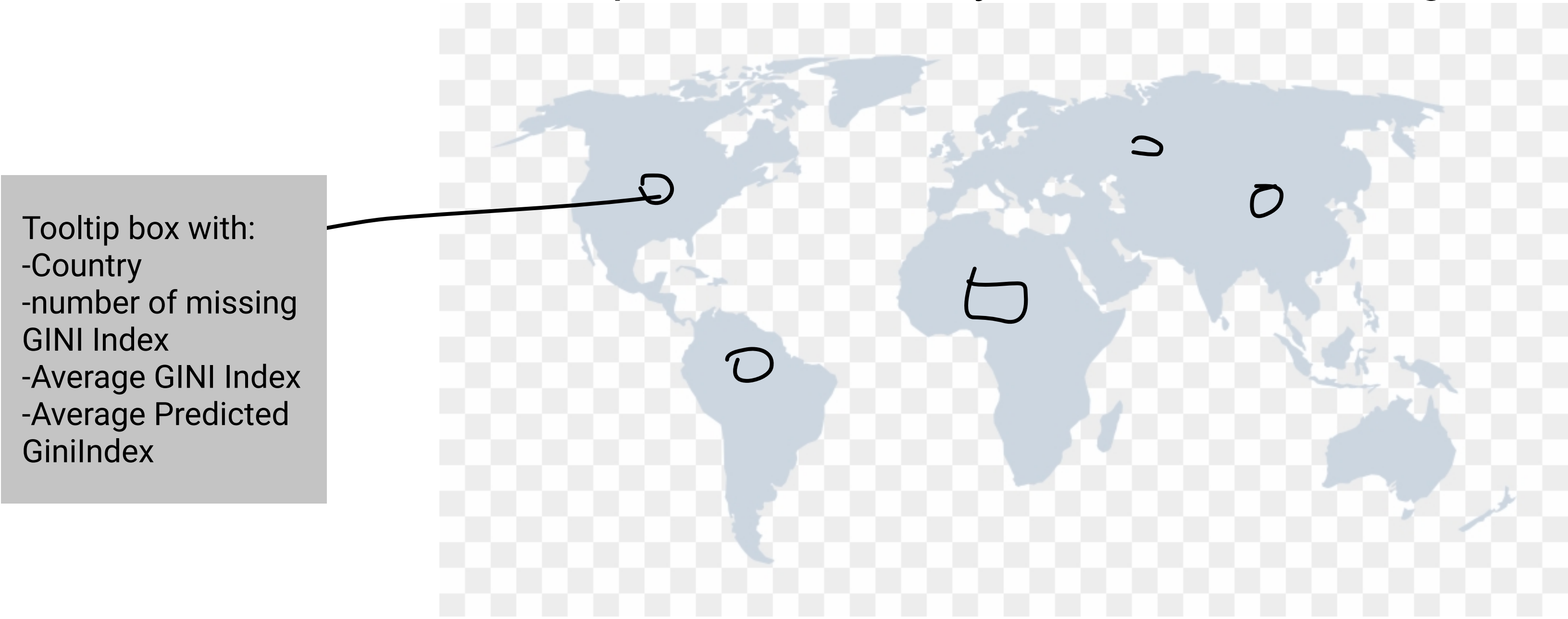


Clustered Bar Char of Predicted and Actual Gini Index

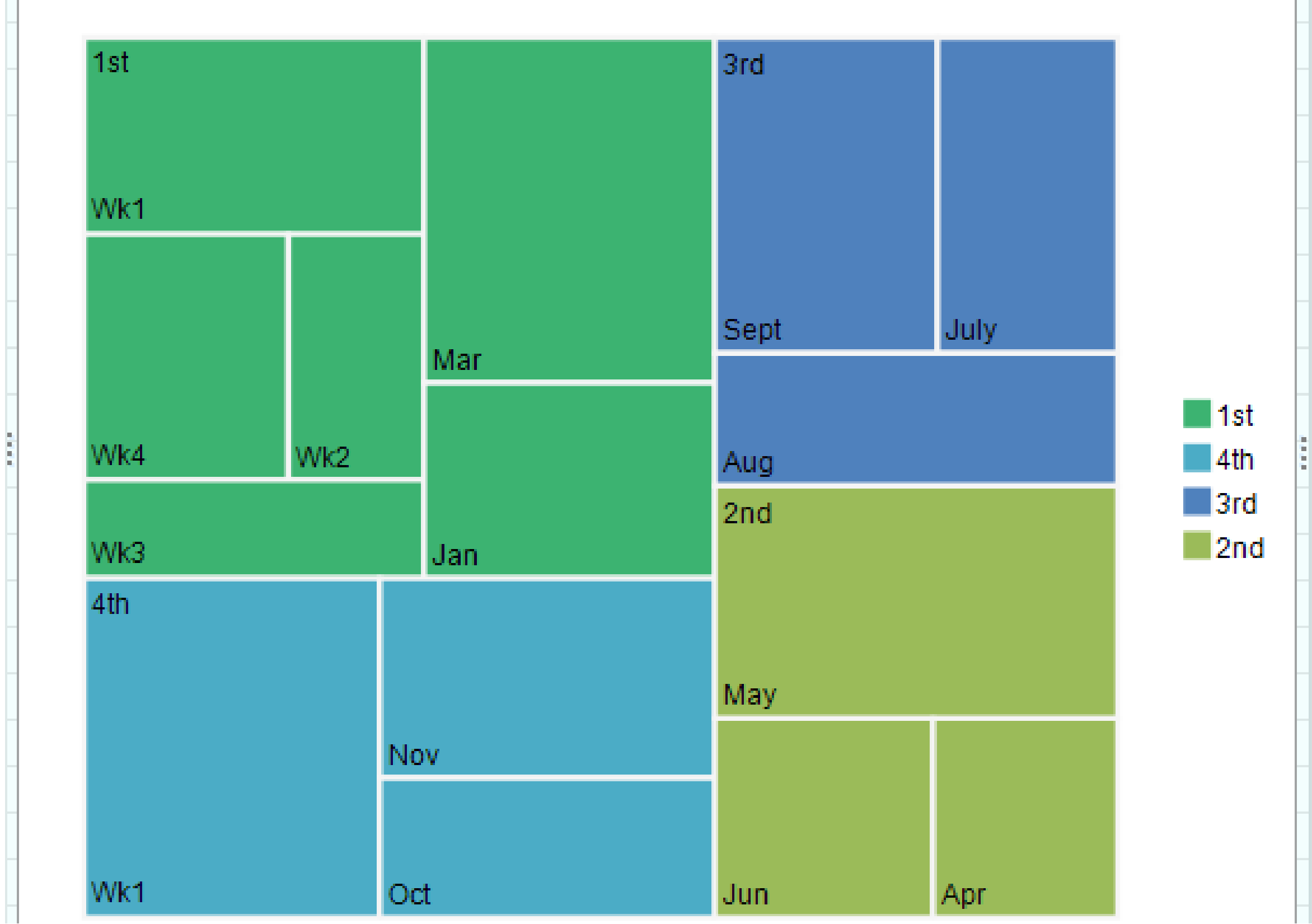




World Map of Countries by Number of Missing Gini Index



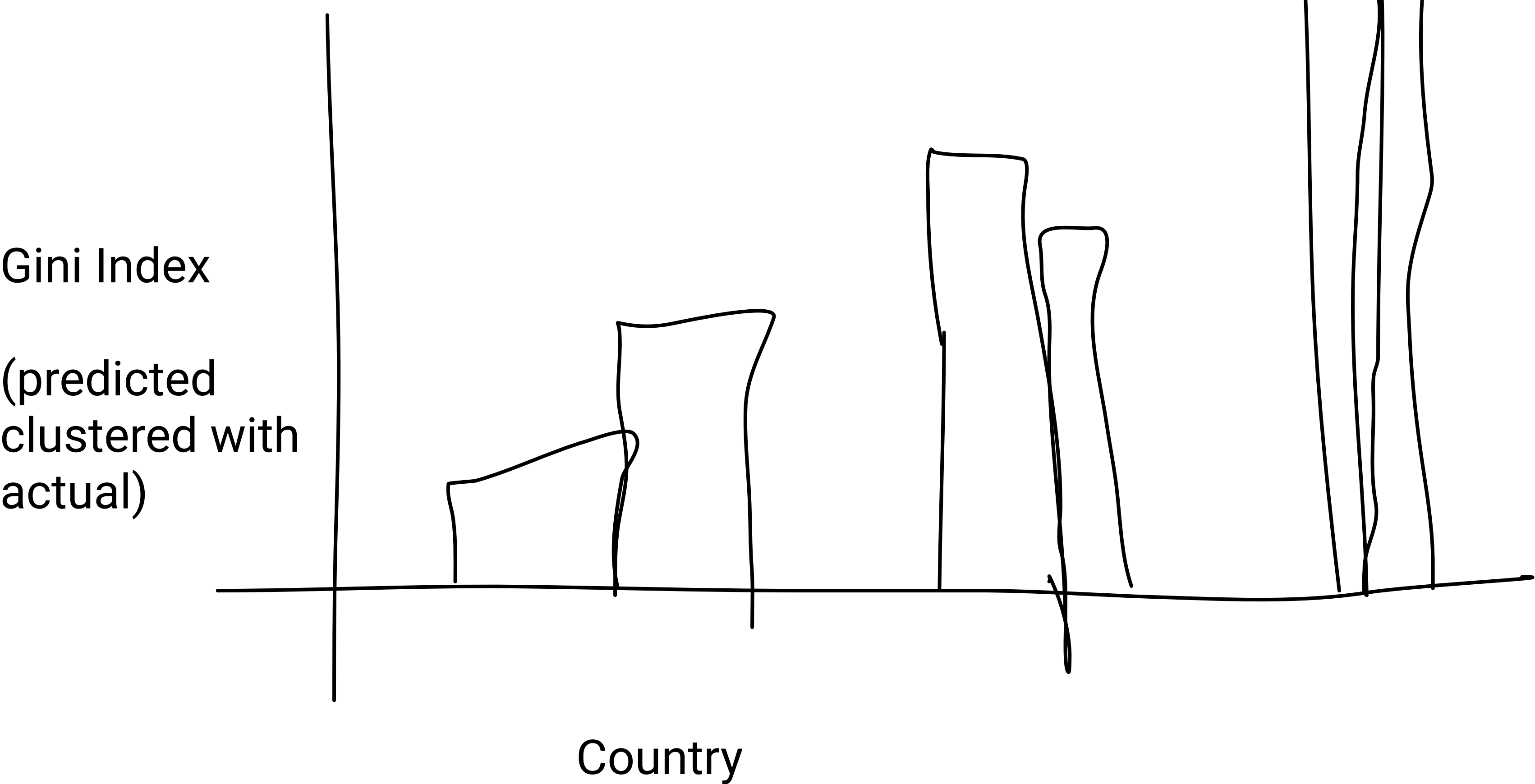
Tree map of GINI Index Model Features by Feature Importance



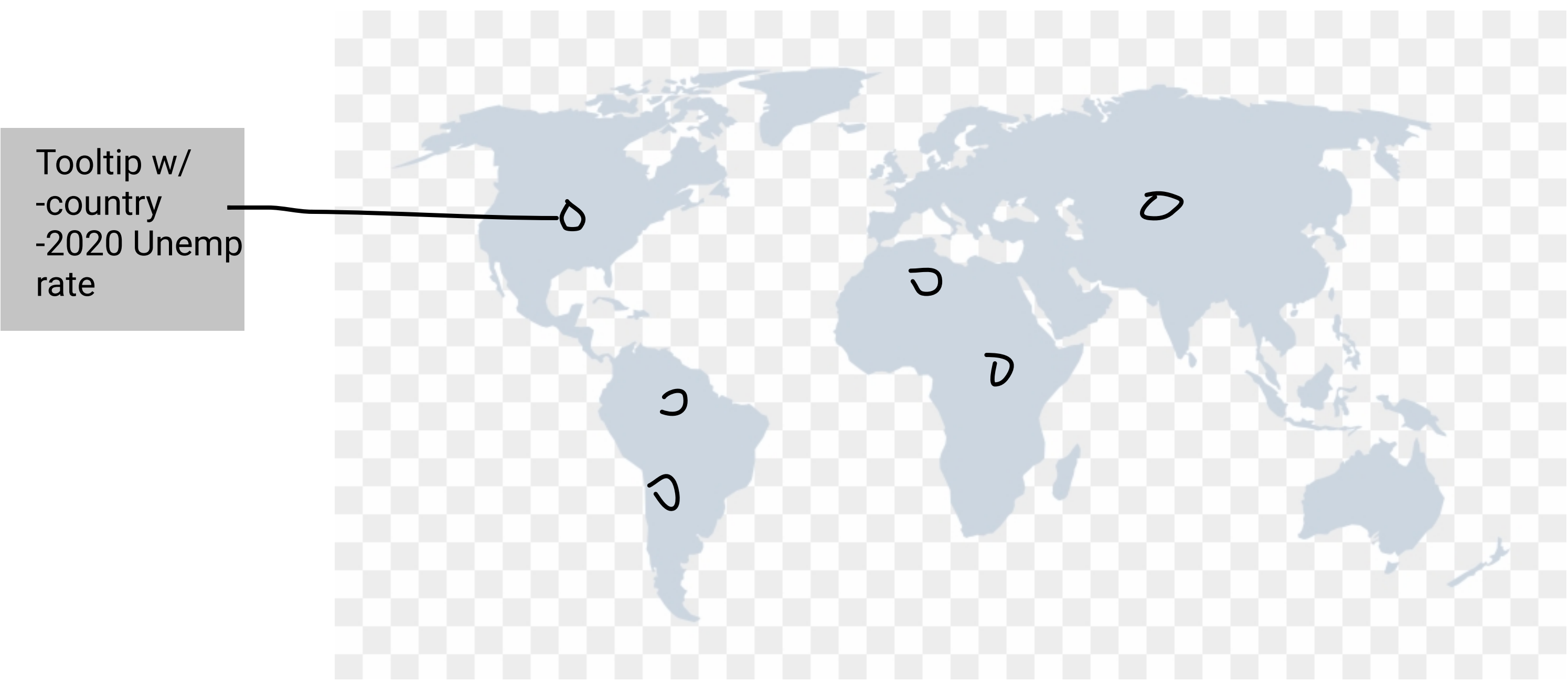
Actual GINI Index vs. Predicted GINI Index



Clustered Bar Char of Predicted and Actual Gini Index



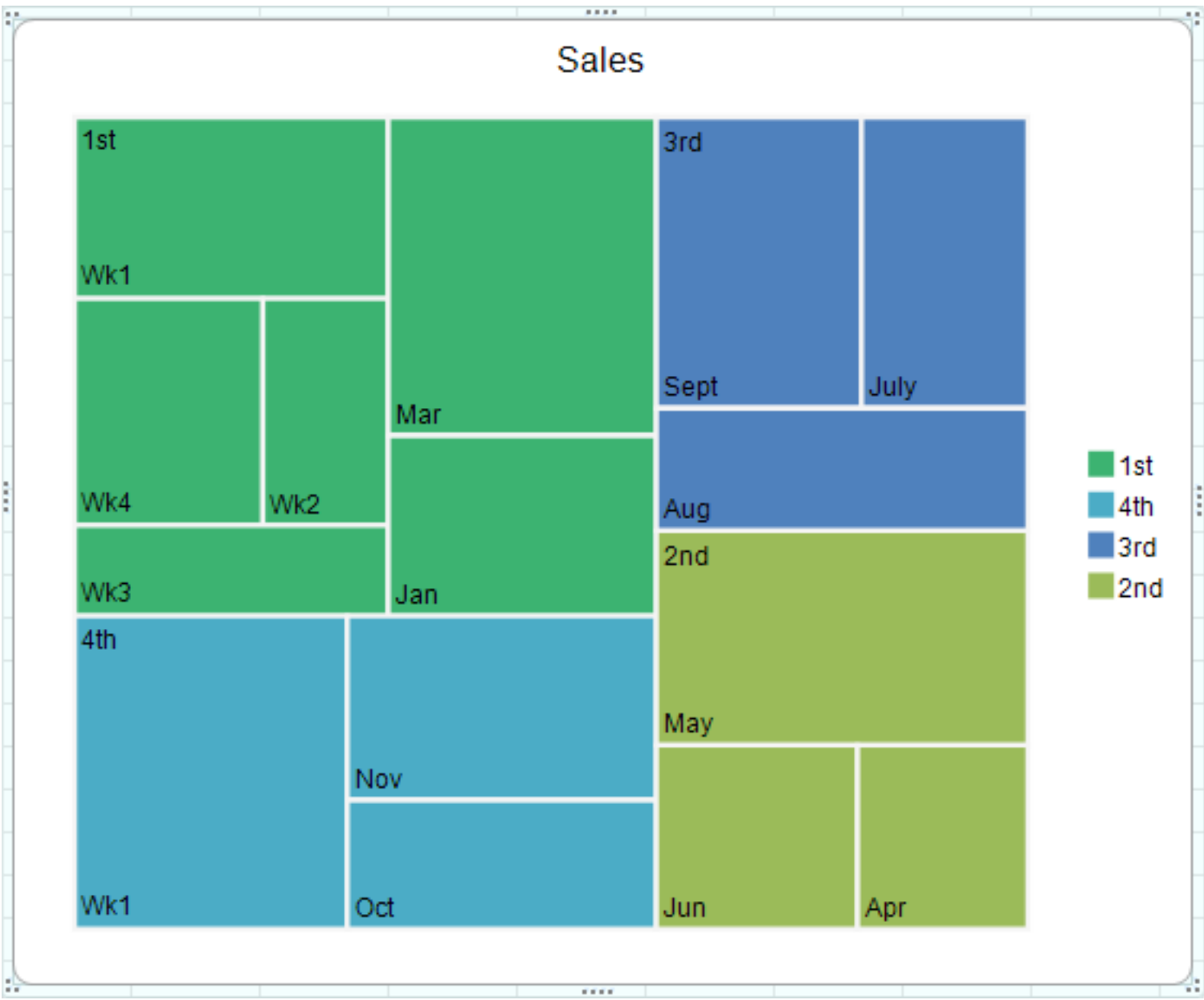
Heat Map of Countries by 2020 Unemployment Rate



Heat Map by Predicted Unemp Rate in 2021



Feature Importance Tree Map for Unemp Rate Model



Actual vs. Predicted Unemp Rate Scatterplot

