Jessica Curley, Guillermo Shick,

Dylan Smith, Jeffrey Winny

Steven Ramsier

February 22, 2019

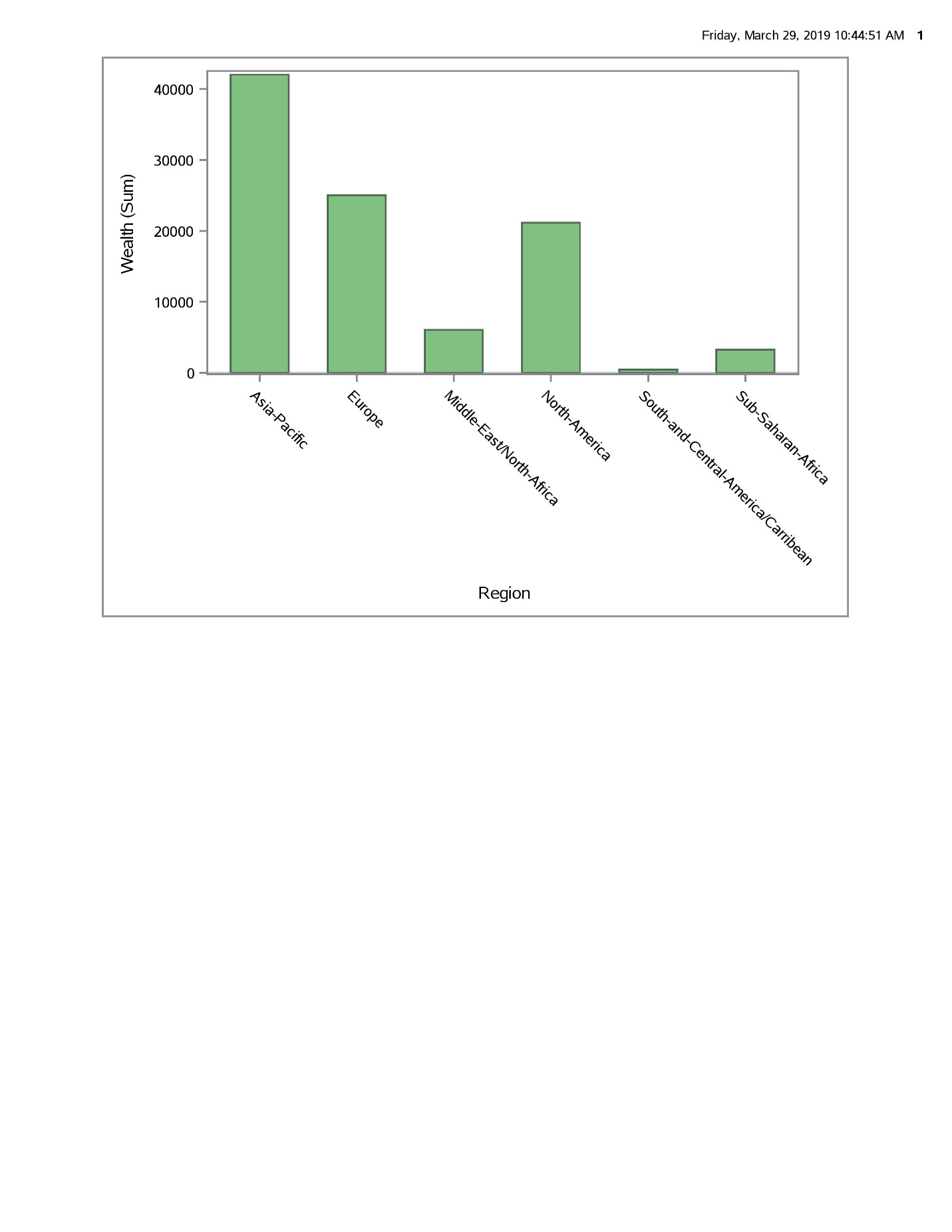
Team 12 Project B

**Introduction**

Our team was very interested in the power the internet has on a civilization, and the opportunities presented to individuals of internet-based societies. Again, our variables are internet penetration rate (Penetration); how many people based on the population use the internet, Region; the geographical location of the country, economic freedom index (freedomrank); an index of a country’s economic and social freedoms relative to each other, each nations’ GDP (Wealth) measured in billions of US Dollars; and population measured in millions. Our main goal is to learn the effect of internet penetration on a region’s economic freedom and GDP.

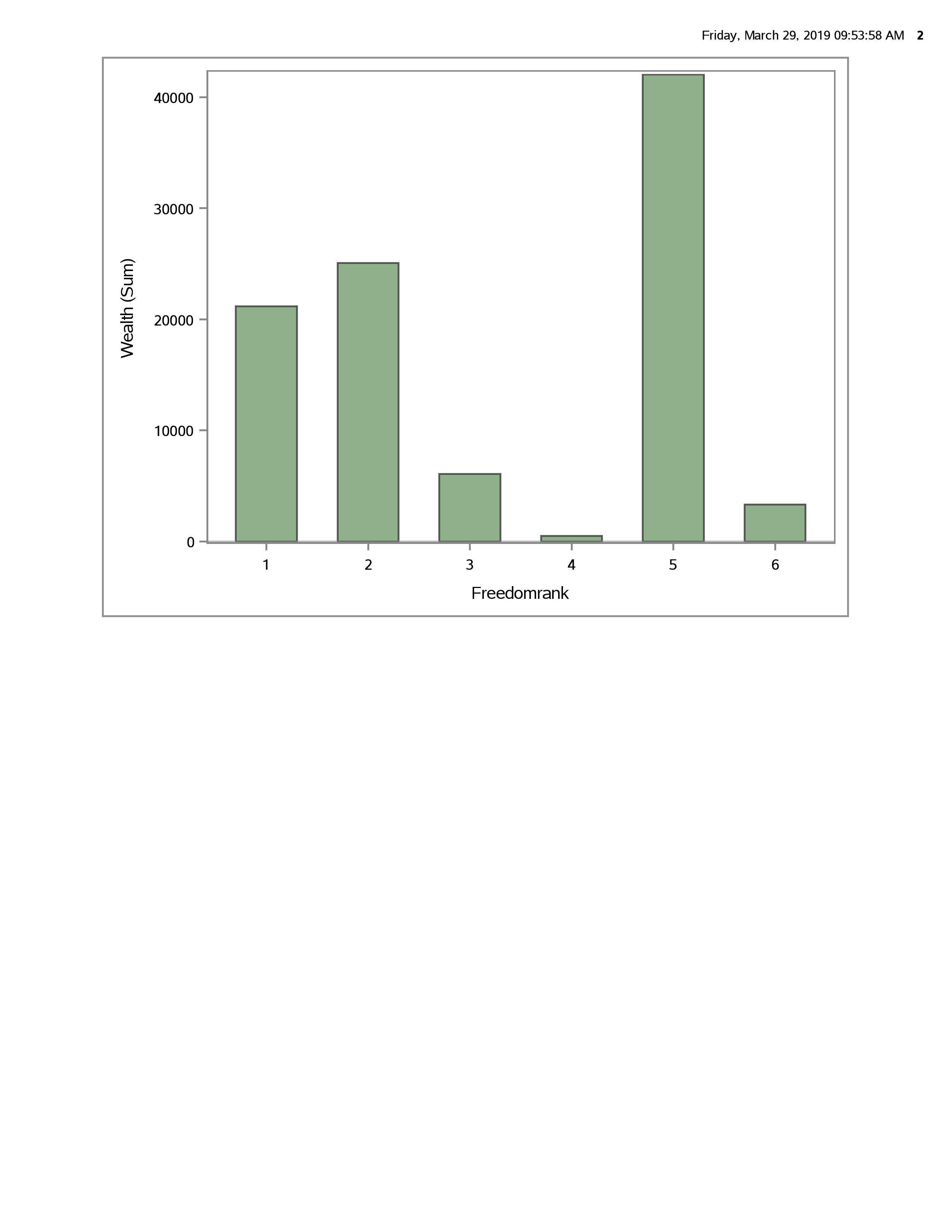
**Data Exploration via Graphical Displays**

Bar chart #1



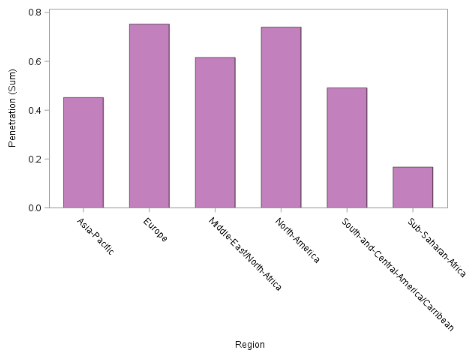
This chart displays the combined wealth of all the countries within each region. In general, there is a trend that liberal, democratic regions, such as the North America and Europe, are wealthier than their less free counterparts. The major exception is Asia, which is due to the fact that despite having possessing some of the most restrictive, totalitarian regimes in the world, the region is also home to some of the countries most free economies and societies. China is also a major outlier, being the world’s second largest economy and a heavily censored society, although they have made more recent strides towards increased economic freedoms lately.

Bar Chart #2



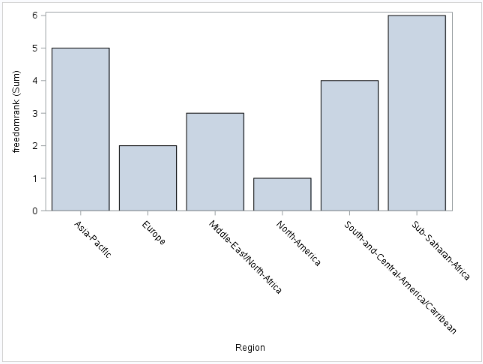
Here we have compared the combined wealth of the regions in billions of US Dollars and their freedom rank. We wanted to see if more wealth meant more economic freedom. Interestingly, in the Asia-Pacific region there is a spike, which is most likely due to the region possessing an almost even mix of wealthy and poor nations as well as free and totalitarian nations and mixes of both categories. Due to wealth being the combined GDP of all Asian nations, rich and poor alike as well as the existence of China, a very economically powerful nation with heavy societal restrictions the freedom rank for Asia is not a good overall estimator of the region’s overall freedom. The data overall though shows that wealthier nations are often more economically and socially free.

Bar Chart #3



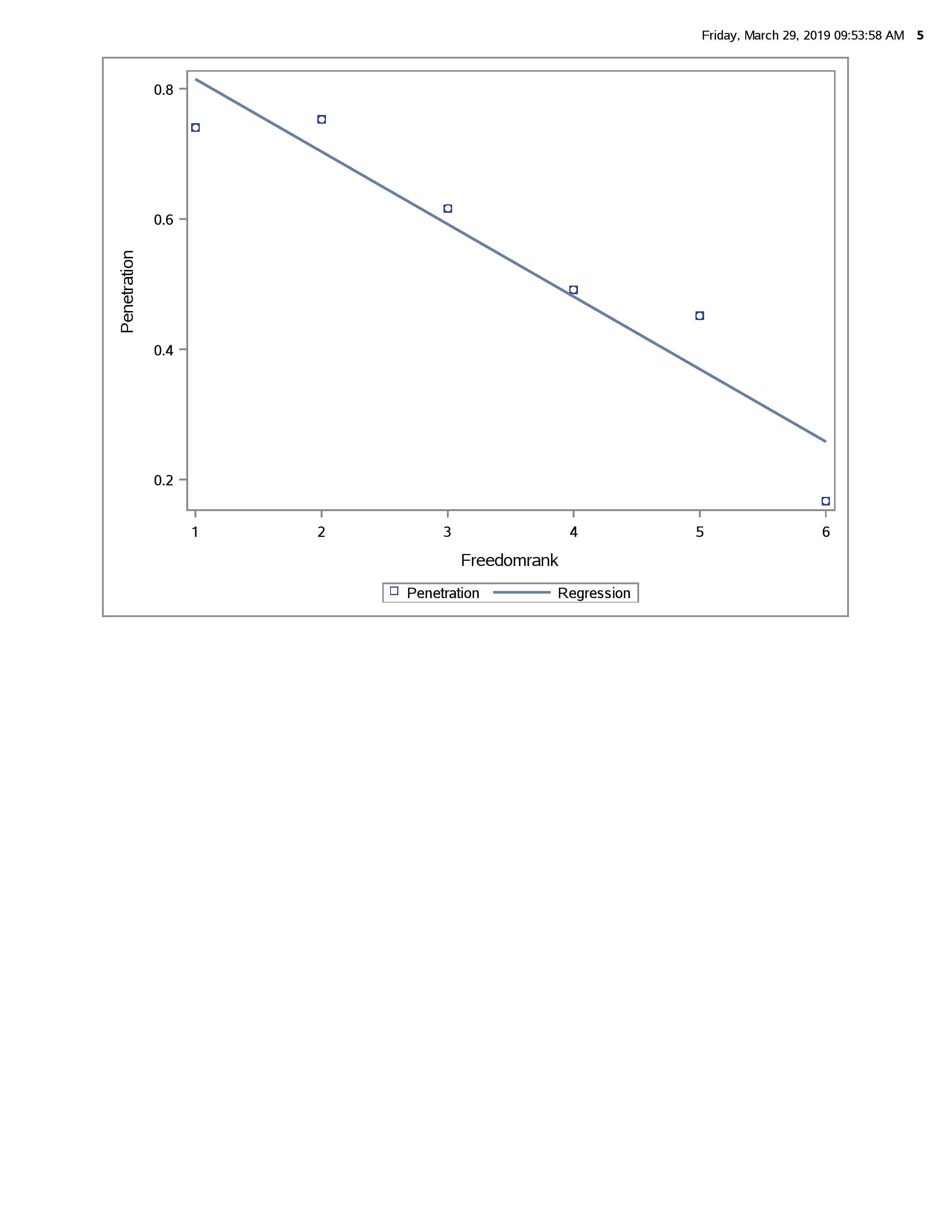
Here we wanted to see if the penetration percentage, based on population changed based on geographic location. Region does appear to play a role in internet access; North America and Europe have penetration rates of almost 80% while less than half of the populations of Asia and Central/South America have access to the internet and it is uncommon in Sub-Saharan Africa. This makes sense as access to the internet depends entirely on a nation’s ability to have enough resources to build the infrastructure necessary for the internet to work and its desire to do so.

Histogram #3



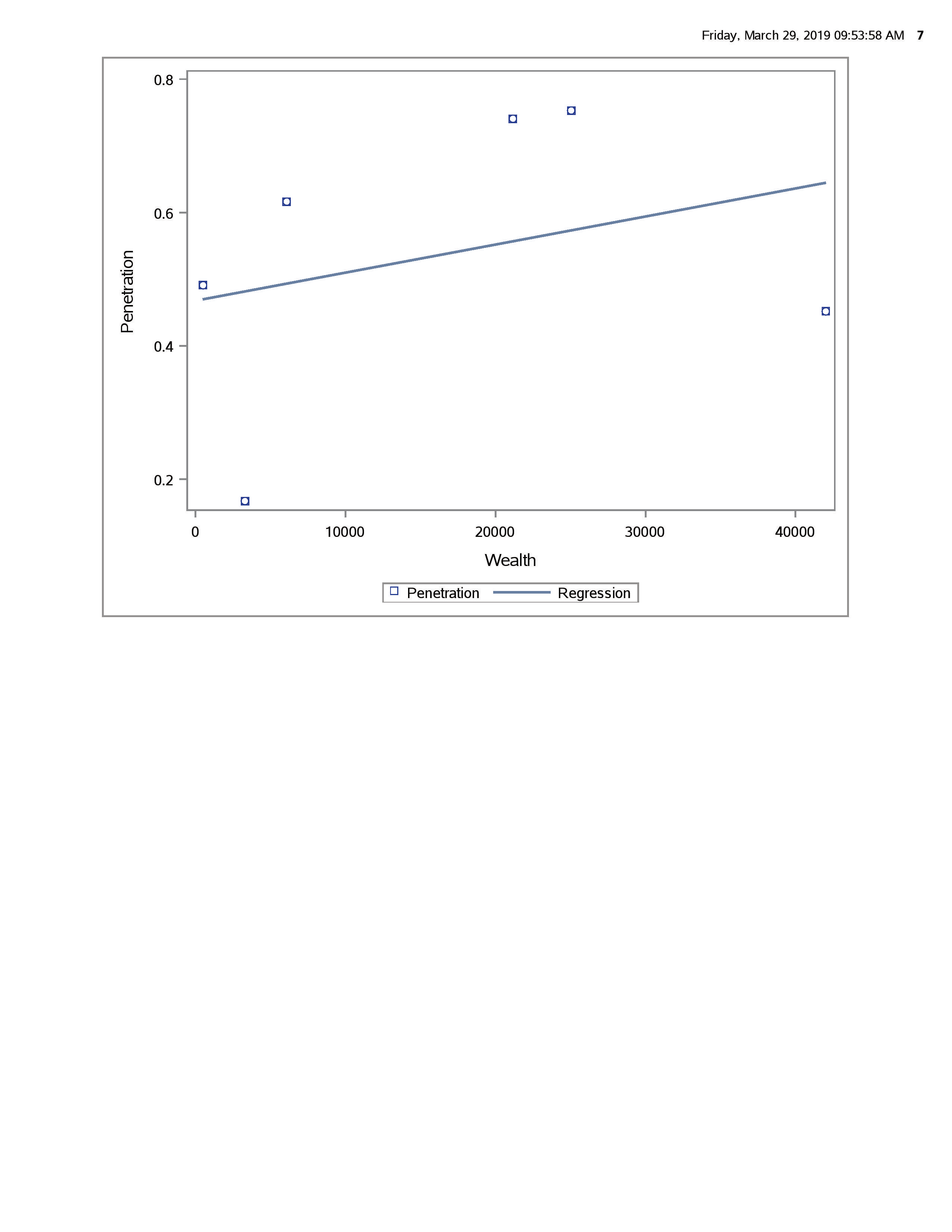
This histogram shows each region’s relative freedom rank when compared to each other. There are six regions and they are ranked in ascending order from most free to least free. Asia’s rank is the least representative as the region possessed a mix of both the most free and repressive nations in the world. Due to possessing as overall greater number of repressed nations, the region’s rank was second to lowest overall.

Scatterplot #1



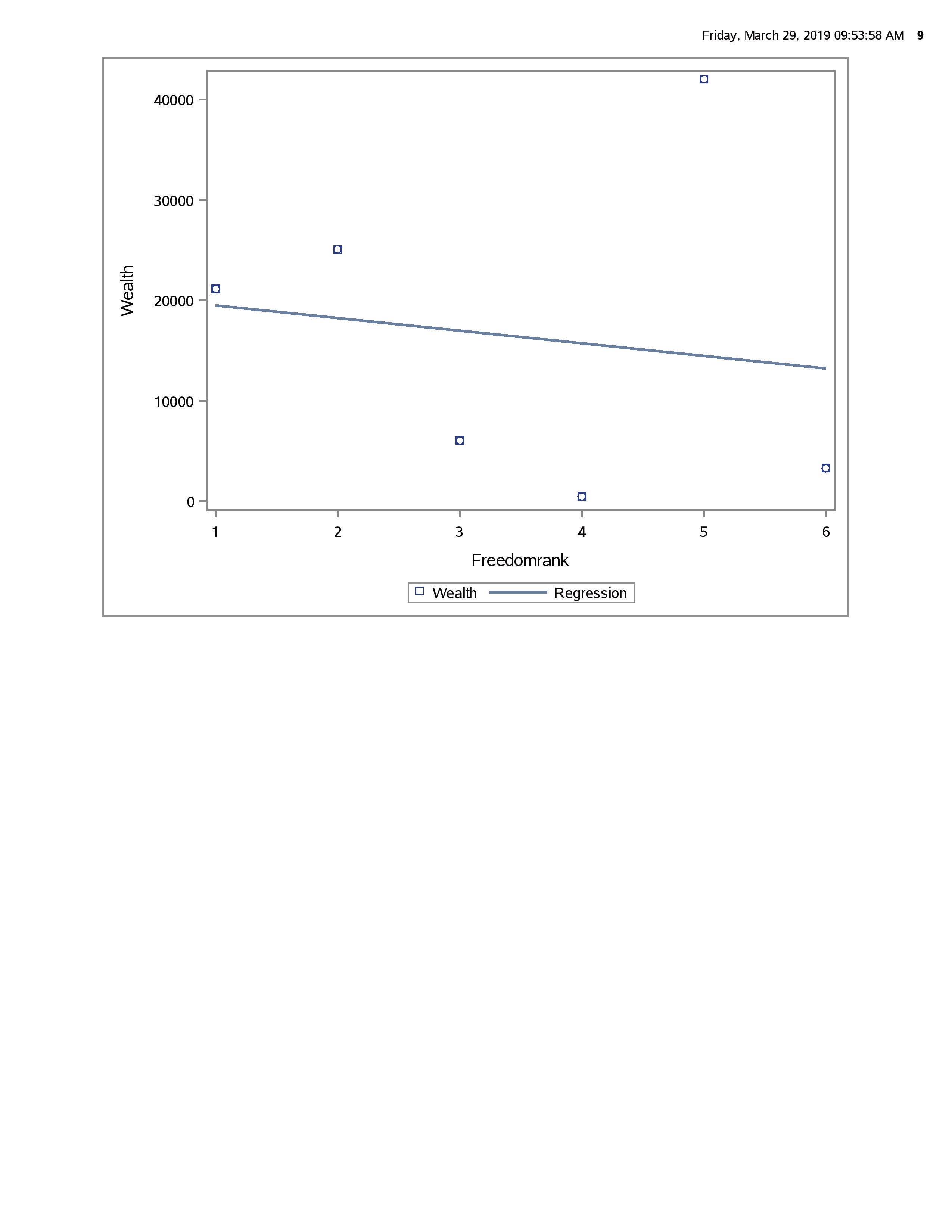
The following chart shows there is a moderately strong linear correlation between a region’s freedom rank and its populations’ access to the internet. Although the graph appears to present a negatively correlated relationship, remember that since a lower freedom rank indicates a freer society this actually shows the opposite that the more free a region is the higher percentage of its people have internet access.

Scatterplot #2



The results of this graph are interesting as it shows that there is a linear relationship between a region’s internet penetration and its wealth but that this is a very weak relationship. This is surprising as one would expect that greater wealth would allow for a better online infrastructure and greater opportunities for internet access but their does not appear to be a very strong relationship.

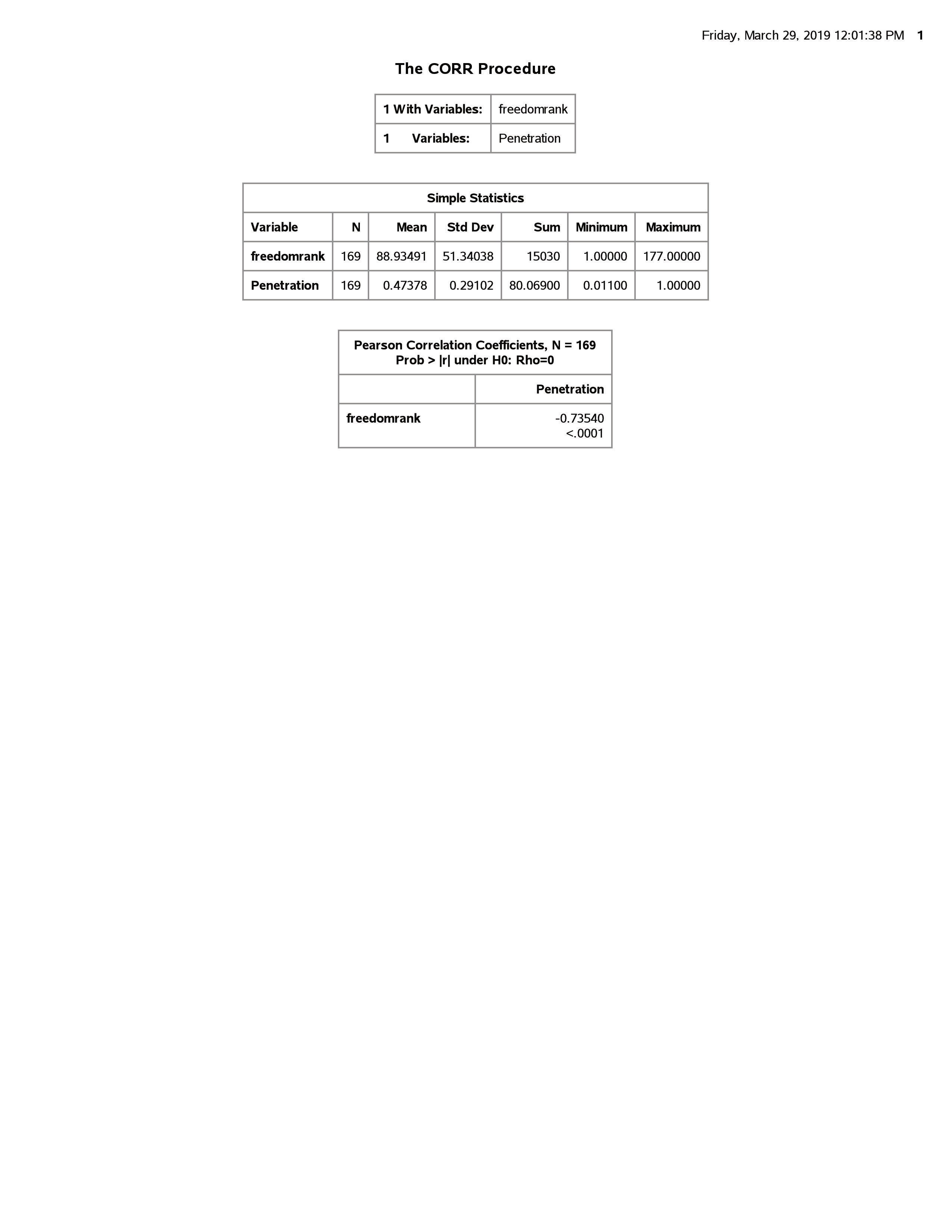
Scatterplot #3



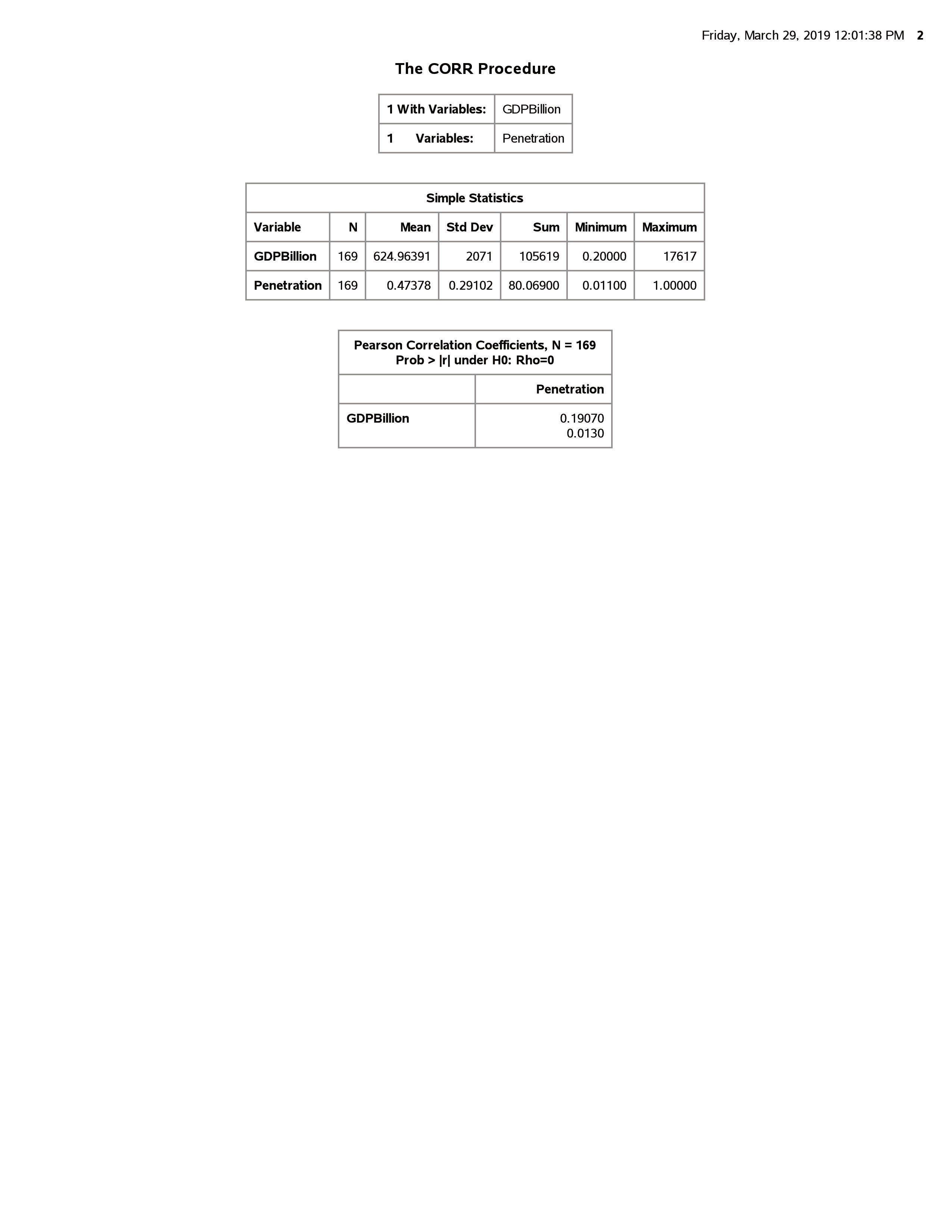
This graph comparing wealth to freedom is actually a little deceptive. If one were to only look at the regression line, there does not appear to be only a very weak correlation between the two. However, this is mostly due to the outlier of Asia, discussed in previous sections. If Asia were removed there would be a much stronger relationship between the variables.

**Data Exploration via Numerical Summaries**

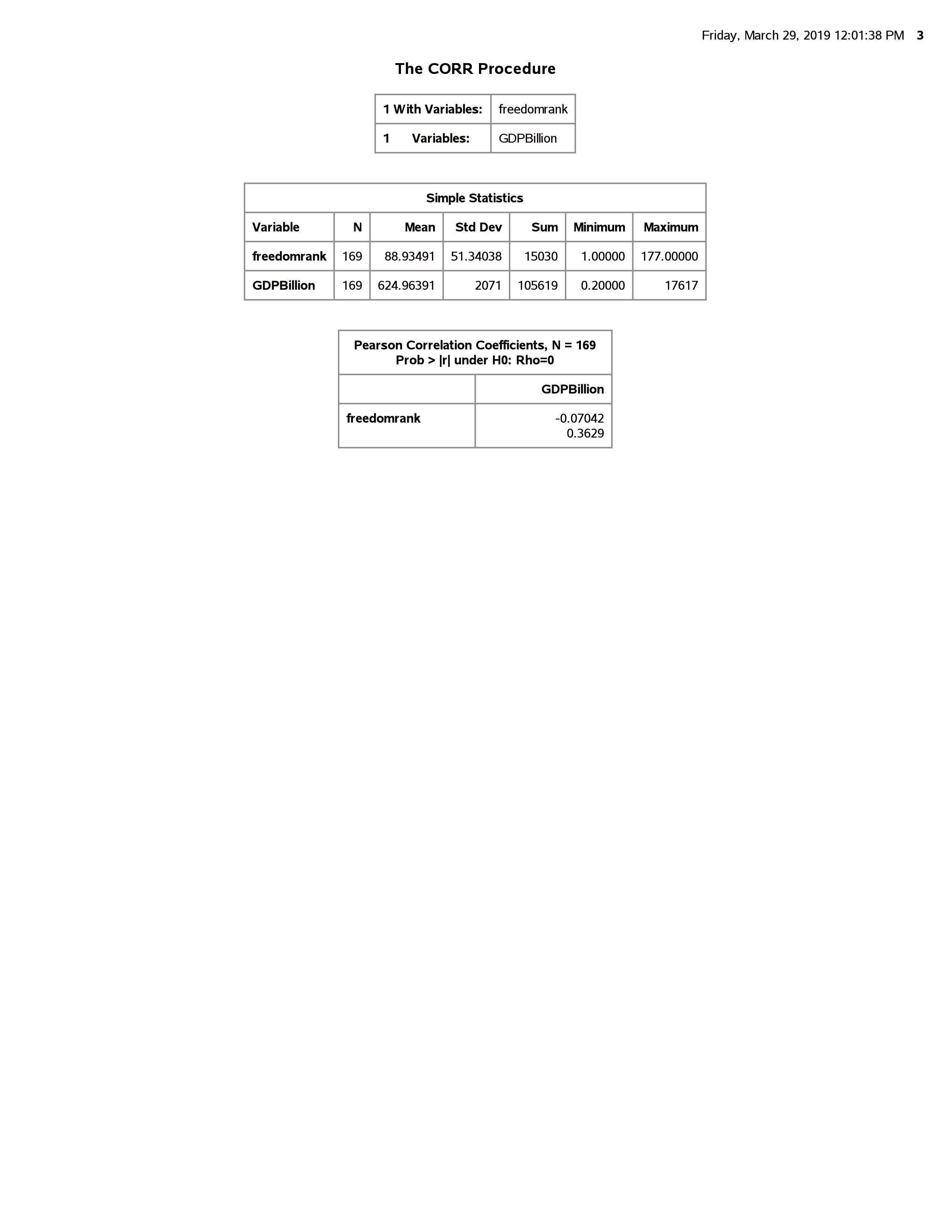
Note that for this section we actually used are original, untransformed data of comparing individual countries instead of regions. We felt this method would produce more representative results, and the primary reason for reducing all the data to six regions, making the graphical displays easier to follow, no longer applies. This also had the added bonus of not lumping all Asian countries in a single lump sum which had a tendency to skew the data in some cases.



A Correlation Procedure was performed upon the variables of a country’s freedom rank and its internet penetration. Like in graphical display earlier, there is shown to be a fairly strong linear relationship between freedom and internet access. The p-value for the null hypothesis that there is no correlation between the two variables is incredibly low, and we would reject the null hypothesis of no relationships under most significance levels. This is consistent with our hypothesis, that citizens of freer nations would have increased access to the internet as their governments are less concerned with their online activities and what information they may find.



A Correlation Procedure was performed upon the variables of a country’s GDP and its internet penetration. There appears to be a very weak linear relationship between the two variables although the p-value is small enough to reject the null hypothesis that there is no relationship between the variables under the standard 5% significance level. A country’s wealth does appear to have an affect on its populations access to the internet, which makes sense as internet infrastructure costs money to set up and maintain but this appears to be significantly less important a factor than how free a country is. Perhaps richer, more restrictive nations are not willingly to spend the money to create and maintain a potentially subversive system for their citizens.



A very surprising result is there does not appear to be any correlation between a nation’s wealth and how free a society they are. The calculated correlation coefficient is very weak and the p-value is so high to be unable to reject the null hypothesis of no relationship between the variables under almost any significance level. This relationship could be explored later and it is possible that there is a relationship between the two variables, just not a linear one, which is what the correlation coefficient measures.

**Conclusion**

After completing both graphical and numerical analysis on our data that some of our initial premises of our hypothesis were true and some false. There does appear to be a strong relationship between the freedom of a nation’s citizens and access to the internet as we had predicted. There is a similar relationship between wealth and internet access but it is not as strong as we originally had assumed. There appears to be no (linear) relationship between a nation’s wealth and the freedom of its citizens. We plan to explore the relationships between these three variables in depth and perhaps look and see if there is a non-linear relationship between freedom and wealth.

***We, the project team members, certify that the percentage of the effort listed by each of our names below is an accurate account of the original effort contributed by each team member in the producing of this project and report.***

Name : Percent of Total Work: Statistics Major:

Jeffrey Winny 25% No

Guillermo Shick 25% No

Jessica Curley 25% No

Dylan Smith 25% No