Data_Assignment_3

Marko Suchy

Introduction

This project focuses on 3 networks reaching across Europe: cultural similarity, trading partnership, and militaristic alliances. Data for each of these relationship types, as well as node and edge attributes, were pulled from public online sources and cleaned and prepared using Python.

Cultural similarity data was scraped from a study done by Dr. Jochen Roose, of the Freie Universität Berlin, based on responses to the European Social Survey. The study constructed an index from 0 to 1 for the cultural similarity between most European countries. Over 500 data points exist from this study, and are used as edges in the cultural similarity network below. The countries with similarity indices are used as the base set of nodes for our networks.

Economic trading data was scraped from the World Integrated Trade Solution website. Data was pulled from the top 5 'export & import partners' for each country in our base set of nodes. An edge between partner and country was created if both were included in out base set of nodes. Information about the dollar amount of trade between countries (in millions of USD) and partner share percent were also scraped from the WITS website, and included as edge attributes.

Alliance data was taken from the 'Correltes of War' database and prepared in python. This dataset held Alliances across hundreds of countries throughout history (starting before World War 1.) All alliances with end dates were removed. The data includes many types of agreements: defense, neutrality, nonaggression, entente, and asymmetric. One big disadvantage to the data found is that it is missing data from Austria, Ireland, Israel, Slovenia, Sweden, Switzerland, Cyprus.

Node attributes: democracy index, press freedom index, population, GDP, and region, were all pulled from various Wikipedia tables. More reading on the measurement of each of these attributes can be found in the list data sources.

The project is interesting because it may offer some questions for further study regarding relationships in Europe, which is currently in a volitile state due to the war in Ukraine. It also provided an excellent opportunity for me personally to practice my data scraping and wrangling skills in both Python and R.

Exploratory Visualizations

We import and join cultural similarity network data, and create a subgraph including only ties with a weight greater than 0.78. This threshold is selected because it seems to suggest a significant level of connection, without creating too sparse a network.

```r
cultural_exchange_edgelist <- read.csv("cultural_exchange_edgelist")
cultural_strength <- read.csv("cultural_exchange_strength")
cultural_node_attributes <- read_csv("node_attributes_cultural_exchange")
cultural_exchange_network <- as.network(cultural_exchange_edgelist, directed = FALSE)

set.edge.attribute(cultural_exchange_network, "weight", cultural_strength[['strength']])
```

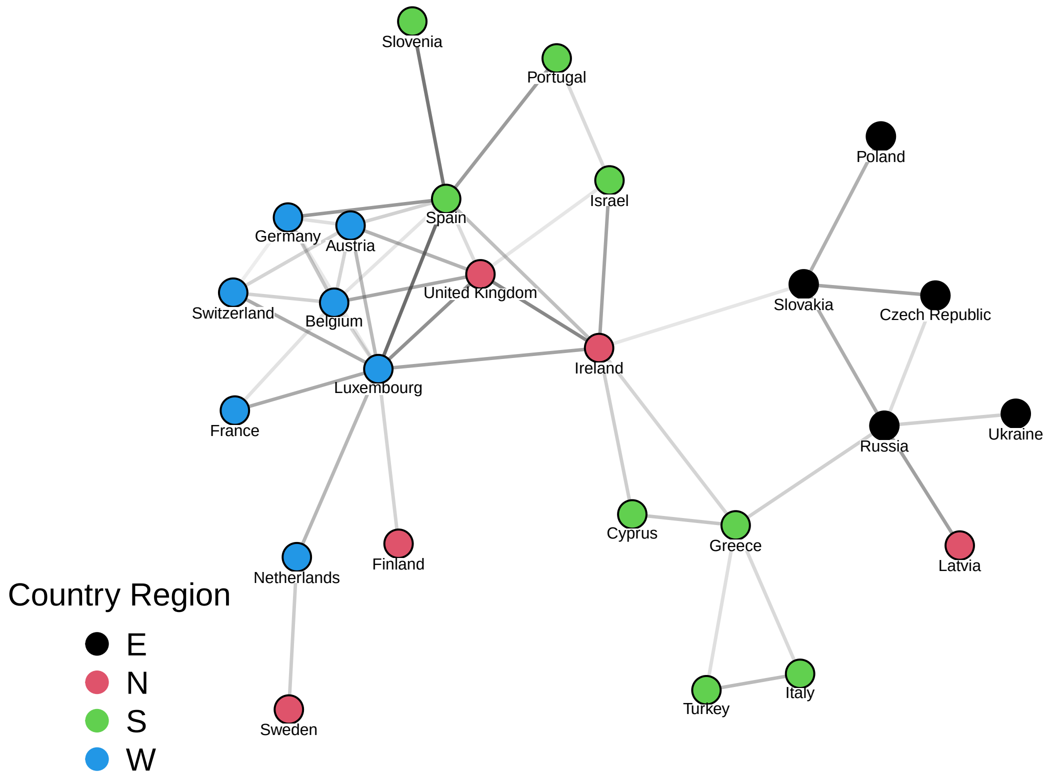
```r
set.vertex.attribute(cultural_exchange_network, "democracy_index",
                     cultural_node_attributes$democracy_index)
set.vertex.attribute(cultural_exchange_network, "press_freedoom",
                     cultural_node_attributes$press_freedom)
set.vertex.attribute(cultural_exchange_network, "population",
                     cultural_node_attributes$population)
set.vertex.attribute(cultural_exchange_network, "GDP",
                     cultural_node_attributes$GDP)
set.vertex.attribute(cultural_exchange_network, "region",
                     cultural_node_attributes$region)
cultural_similarity_subgraph <- get.inducedSubgraph(
  cultural_exchange_network, eid=which(cultural_exchange_network %e% "weight" > .78))
```

Now we will set edge opacity based on weight, and the vertex color based on the node attribute region. Then we plot the network, including a legend.

```r
cultural_exchange_network.ecol <- gray(1 - (cultural_exchange_network %e% "weight"),
                                       alpha = (cultural_exchange_network %e% "weight")^4)
cultural_exchange_network.vcol <- as.factor(cultural_similarity_subgraph %v% "region")

op <- par(mar=c(0,0,0,0))

gplot(cultural_similarity_subgraph, gmode = "graph", displaylabels = TRUE, label.pos = 1,
      boxed.labels = T, label.border = 0, label.cex = .5, label.pad = 0,
      edge.col = cultural_exchange_network.ecol,
      vertex.col = cultural_exchange_network.vcol)
legend("bottomleft", legend = c("E", "N", "S", "W"),
       col = as.factor(levels(cultural_exchange_network.vcol)), pch=19,pt.cex=1.5,bty="n",
       title="Country Region")
```




```
par(op)
```

This graph reveals that region appears an important factor in determining connection. However it seems that for Northern countries, and possibly southern countries, the importance of region is less. Also, Ireland is the only node connected to all other regions, and is positioned as a node with high bridging capital.

Now we will import and join our data on the economic relationships between countries.

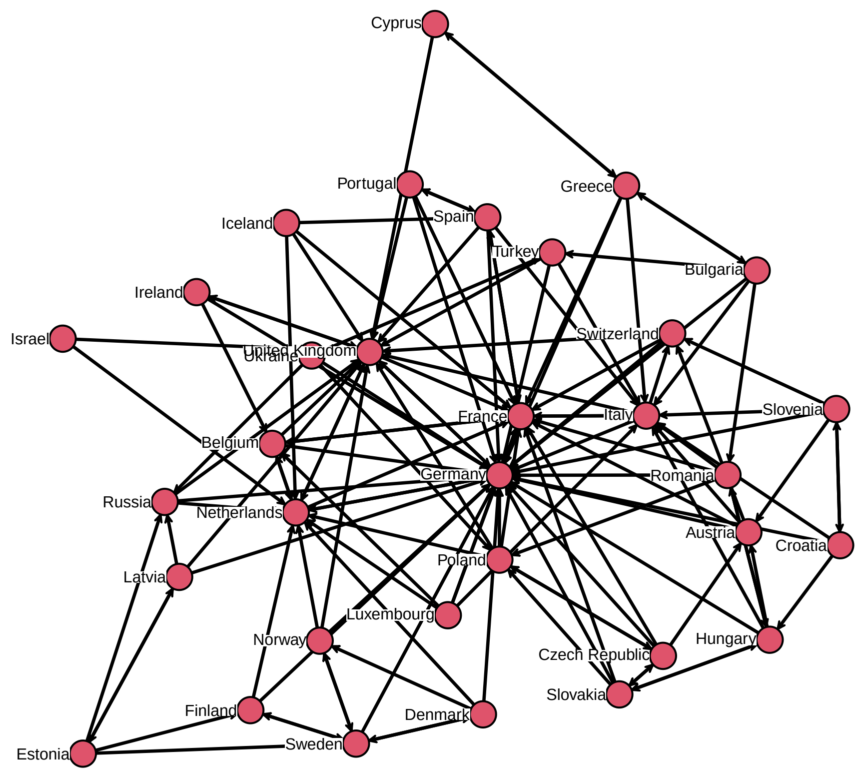
```
economic_exchange_edgelist <- read.csv("trade_edge_list")
economic_exchange_edge_attributes <- read_csv("trade_edge_attributes")
economic_node_attributes <- read_csv("node_attributes_economic_exchange")
economic_exchange_network <- as.network(economic_exchange_edgelist)

set.edge.attribute(economic_exchange_network, "trade_in_USD",
                   economic_exchange_edge_attributes$Trade)
set.edge.attribute(economic_exchange_network, "share_percent",
                   economic_exchange_edge_attributes$`Share Percent`)
set.vertex.attribute(economic_exchange_network, "democracy_index",
                     economic_node_attributes$democracy_index)
set.vertex.attribute(economic_exchange_network, "press_freedom",
                     economic_node_attributes$press_freedom)
set.vertex.attribute(economic_exchange_network, "population",
                     economic_node_attributes$population)
set.vertex.attribute(economic_exchange_network, "GDP",
                     economic_node_attributes$GDP)
```

```
set.vertex.attribute(economic_exchange_network, "region",
                     economic_node_attributes$region)
```

When we plot the graph, we see that France, Germany, Italy, and the United Kingdom are highly central. France and Germany's spots as number two and number two suggest that maybe region has something to do with degree centrality. Western countries, being closer to the Americas, might be more likely to be a profitable trading partner. This theory will be tested later via exponential random graph models (ERGMS.)

```
op <- par(mar=c(0,1,0,0))
gplot(economic_exchange_network, gmode = "digraph", displaylabels = TRUE, label.pos = 2,
      boxed.labels = T, label.border = 0, label.cex = .5, label.pad = 0,
      arrowhead.cex = .3)
```




```
par(op)
```

Finally, we will import alliance data, and plot it.

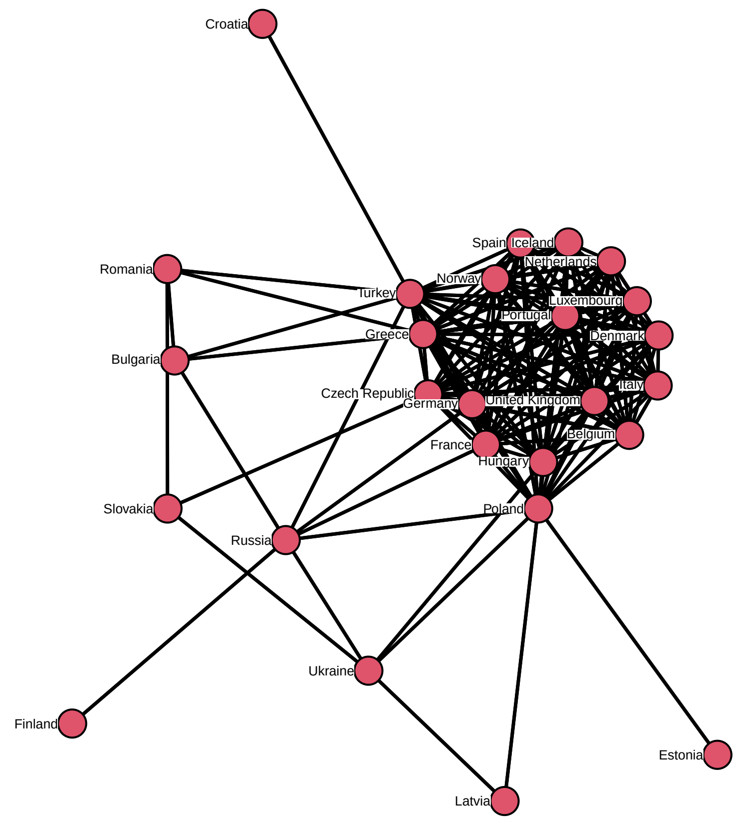
```
alliance_edgelist <- read.csv("alliance_edge_list")
alliance_edge_attributes <- read.csv("alliance_edge_attributes")
alliance_node_attributes <- read.csv("node_attributes_alliance_network")

alliance_network <- as.network(alliance_edgelist, multiple = T, directed = F)

set.edge.attribute(alliance_network, "defense", alliance_edge_attributes$defense)
```

```r
set.edge.attribute(alliance_network, "neutrality", alliance_edge_attributes$neutrality)
set.edge.attribute(alliance_network, "nonaggression",
                   alliance_edge_attributes$nonaggression)
set.edge.attribute(alliance_network, "entente", alliance_edge_attributes$entente)
set.edge.attribute(alliance_network, "asymmetric", alliance_edge_attributes$asymmetric)
set.vertex.attribute(alliance_network, "region", alliance_node_attributes$region)

op <- par(mar=c(0,0,0,0))
gplot(alliance_network, gmode = "graph", displaylabels = TRUE, label.pos = 2,
      boxed.labels = T, label.border = 0, label.cex = .4, label.pad = 0)
```




```r
par(op)
```

This plot reveals that most of Europe, especially those countries in Nato, are heavily connected via alliances. It is possible to view specif alliance types, however we will not do this as it really does not seem to reveal much thart's interesting.

Actor Prominence

Because network detection algorithms did not work very well on our networks, we will focus on interesting findings regarding actor prominence.

We will start by examining the centrality data by node in the cultural similarity subgraph.

```r
vertex <- cultural_similarity_subgraph %v% "vertex.names"
degree_centrality <- degree(cultural_similarity_subgraph, gmode="graph")
closeness_centrality <- closeness(cultural_similarity_subgraph, gmode="graph")
betweenness_centrality <- betweenness(cultural_similarity_subgraph, gmode="graph")

centrality_df <- data.frame(vertex, degree_centrality, closeness_centrality,
                            betweenness_centrality)
centrality_df[nrow(centrality_df) + 1,] =
  c("Network", centralization(cultural_similarity_subgraph,degree),
               centralization(cultural_similarity_subgraph,closeness),
               centralization(cultural_similarity_subgraph,betweenness))

print(centrality_df)
```

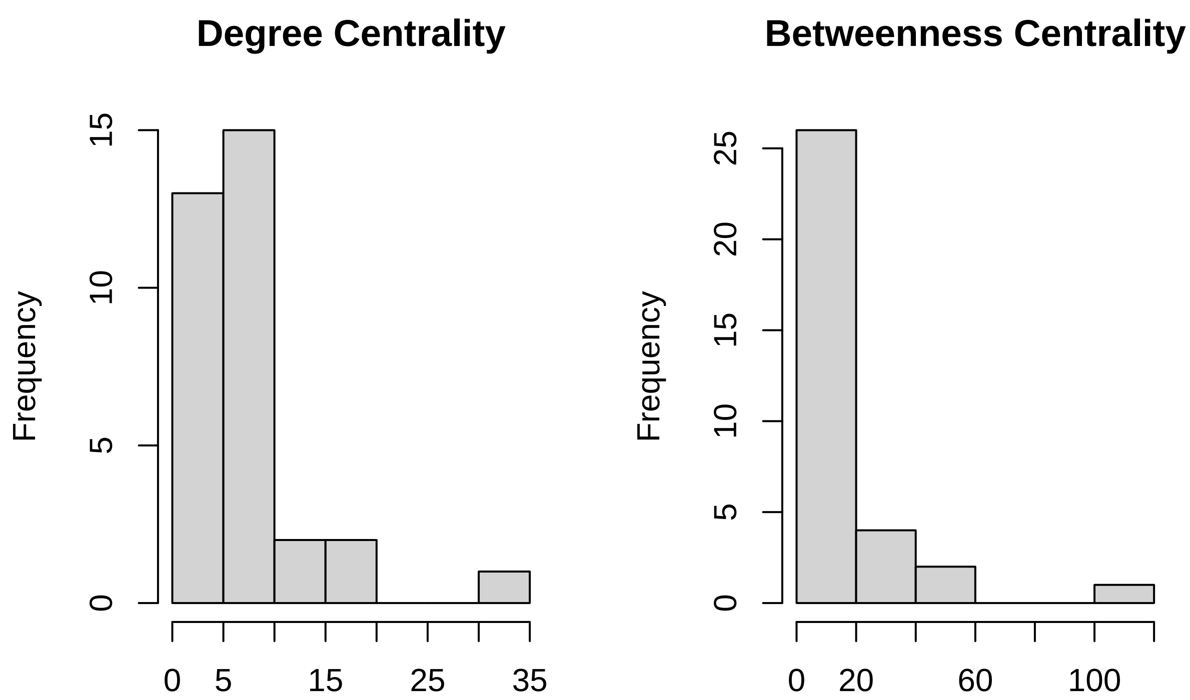
```
##            vertex degree_centrality closeness_centrality betweenness_centrality
## 1         Austria                 6    0.387096774193548       1.72619047619048
## 2         Belgium                 7     0.39344262295082       5.55952380952381
## 3  Czech Republic                 2    0.324324324324324                      0
## 4         Finland                 1    0.342857142857143                      0
## 5          France                 2    0.347826086956522                      0
## 6         Germany                 5                0.375                   0.75
## 7          Greece                 5    0.436363636363636                     70
## 8         Ireland                 7    0.545454545454545       149.035714285714
## 9          Israel                 3    0.387096774193548                      6
## 10          Italy                 2    0.311688311688312                      0
## 11         Latvia                 1    0.263736263736264                      0
## 12     Luxembourg                10     0.51063829787234       106.119047619048
## 13    Netherlands                 2    0.352941176470588                     23
## 14         Poland                 1                  0.3                      0
## 15       Portugal                 2    0.338028169014085      0.809523809523809
## 16         Russia                 5    0.352941176470588                   51.5
## 17       Slovakia                 4    0.421052631578947                     61
## 18       Slovenia                 1    0.328767123287671                      0
## 19          Spain                 8                 0.48       51.5119047619048
## 20         Sweden                 1    0.263736263736264                      0
## 21    Switzerland                 4    0.358208955223881                      0
## 22         Turkey                 2    0.311688311688312                      0
## 23 United Kingdom                 6    0.461538461538462       13.9880952380952
## 24        Ukraine                 1    0.263736263736264                      0
## 25         Cyprus                 2     0.39344262295082                      0
## 26        Network 0.289855072463768    0.190310664065831       0.48081111973775
```

Regarding degree centrality in the cultural similarity subgraph, we found that generally smaller countries held many strong connections, while bigger countries held fewer. This seems to make sense to the extent that the people of different neighboring smaller countries live within closer proximity to one another, while most people of big countries are only near their fellow countrymen.

From these finding, one wonders weather the prominence of an actor in the cultural similarity subgraph drives the greater culture of Europe, or weather it is the result of some other driving force. Surprisingly, Luxembourg emerged as Europe most prominent cultural actor. It wasthe actor with the highest degree centrality, and high closeness centrality and betweenness centrality. It is hard to believe that such a small country would drive the culture of Europe, so it seems out actor prominence reflects the effects of cultural similarity in Europe, rather than reveals the source.

We will now turn to centrality and actor prominence in the economic exchange network. We will examine histograms of degree centrality and betweeness centrality.

```
par(mfrow=c(1, 2))
hist(degree(economic_exchange_network, gmode="digraph"), labels = F, breaks = "Sturges",
     main = "Degree Centrality", xlab = "")
hist(betweenness(economic_exchange_network, gmode="digraph"), labels = F,
     breaks = "Sturges", main = "Betweenness Centrality", xlab = "")
```



These reveal that there seem to be a very small group of extremely prominent actors in this network.

To no one's surprise we see that Germany has the highest degree centrality within the network at 31 ties. This is consistent with Germany as Europe's largest national economy. Germany is followed by France and the UK, both at 20 ties, and Italy and the Netherlands, both at 15 ties. Other than the Netherlands, this follows the pattern of European economy by size exactly.

Interestingly, France is the country with betweeness centrality above 100. This finding, while hard to understand on its face, may open an intersting area for further exlploration.

There is little overlap between the prominent countries of our cultural similarity network and the prominent countries in our economic trading partner network. This seems to have to do with a country's size. Small countries seem culturally similar to many other countries but don't have a lot of trading partners, while large countries have have a lot of trading partners but a lot of cultural individuality.

IV Exponential Random Graph Models

We will now turn our attention to analysis of our networks through ERGMs. We start with a null model built on our cultural similarity subgraph.

```
cultural_mod0 <- ergm(cultural_similarity_subgraph ~ edges)
summary(cultural_mod0)
```

```
## Call:
## ergm(formula = cultural_similarity_subgraph ~ edges)
##
## Maximum Likelihood Results:
##
##       Estimate Std. Error MCMC % z value Pr(>|z|)
## edges  -1.7346     0.1617      0  -10.73   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 415.9  on 300  degrees of freedom
##  Residual Deviance: 253.6  on 299  degrees of freedom
##
## AIC: 255.6  BIC: 259.3  (Smaller is better. MC Std. Err. = 0)
```

We will make a model with considers node level attributes population, press freedom, democracy index, and region.

```
cultural_mod1 <- ergm(cultural_similarity_subgraph ~ edges + nodecov("population") +
                        nodecov("press_freedoom") + nodecov("democracy_index") + nodefactor("region"))
summary(cultural_mod1)
```

```
## Call:
## ergm(formula = cultural_similarity_subgraph ~ edges + nodecov("population") +
##     nodecov("press_freedoom") + nodecov("democracy_index") +
##     nodefactor("region"))
##
## Maximum Likelihood Results:
##
##                         Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                 -9.695e-01  3.194e+00      0  -0.304   0.7615
## nodecov.population     1.331e-09  4.700e-09      0   0.283   0.7771
## nodecov.press_freedoom -2.577e-02  3.044e-02      0  -0.847   0.3973
## nodecov.democracy_index  1.521e-01  2.223e-01      0   0.684   0.4939
## nodefactor.region.N     3.434e-01  5.207e-01      0   0.660   0.5095
## nodefactor.region.S     1.832e-01  4.120e-01      0   0.445   0.6565
## nodefactor.region.W     9.745e-01  4.874e-01      0   1.999   0.0456 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 415.9  on 300  degrees of freedom
##  Residual Deviance: 244.4  on 293  degrees of freedom
##
## AIC: 258.4  BIC: 284.3  (Smaller is better. MC Std. Err. = 0)
```

We see that of these characteristics, the one that makes a statistically significant difference in tie formation is being from the Western region of Europe. Being from the West increases a countries likelihood to have more strong ties. This seems to indicate that Western Europe may have more power in influencing the overall European culture than countries of other regions.

In our next model, we will include a dyadic term evaluating if regions match. It is expected that there is a positive correlation between being in the same region and having a connection, as cultures are tyoically regional.

```
cultural_mod2 <- ergm(cultural_similarity_subgraph ~ edges + nodecov("population")
                      + nodecov("press_freedoom") + nodecov("democracy_index")
                      + nodefactor("region") + nodematch("region"))
summary(cultural_mod2)
```

```
## Call:
## ergm(formula = cultural_similarity_subgraph ~ edges + nodecov("population") +
##     nodecov("press_freedoom") + nodecov("democracy_index") +
##     nodefactor("region") + nodematch("region"))
##
## Maximum Likelihood Results:
##
##                            Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                    -1.092e+00  3.455e+00      0  -0.316    0.752
## nodecov.population        1.719e-09  5.199e-09      0   0.331    0.741
## nodecov.press_freedoom   -3.014e-02  3.301e-02      0  -0.913    0.361
## nodecov.democracy_index   1.838e-01  2.424e-01      0   0.758    0.448
## nodefactor.region.N       3.105e-01  5.106e-01      0   0.608    0.543
## nodefactor.region.S      -9.457e-02  3.836e-01      0  -0.247    0.805
## nodefactor.region.W       6.941e-01  4.857e-01      0   1.429    0.153
## nodematch.region          1.931e+00  3.589e-01      0   5.380   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 415.9  on 300  degrees of freedom
##  Residual Deviance: 215.3  on 292  degrees of freedom
##
## AIC: 231.3  BIC: 261  (Smaller is better. MC Std. Err. = 0)
```

Indeed we find a significant positive correlation in our dyadic term. Interestingly, it seems that this was previously confounding the significance of a country being Western, as that has become not significant. Therefore it seems the local effects of culture greatly outweigh those from Western Europe.

Our next model will consider a structural term, geometrically weighted degree distribution. Unfortunately the model would not compile with a gwesp term, which would be easier to interpret.

```
set.seed=42
cultural_mod3 <- ergm(cultural_similarity_subgraph ~ edges + nodecov("population") +
                      nodecov("press_freedoom") + nodecov("democracy_index") +
                      nodematch("region") + gwdegree(decay = 2, fixed = FALSE, cutoff = 20),
                      control = control.ergm(MCMLE.maxit = 30))
summary(cultural_mod3)
```

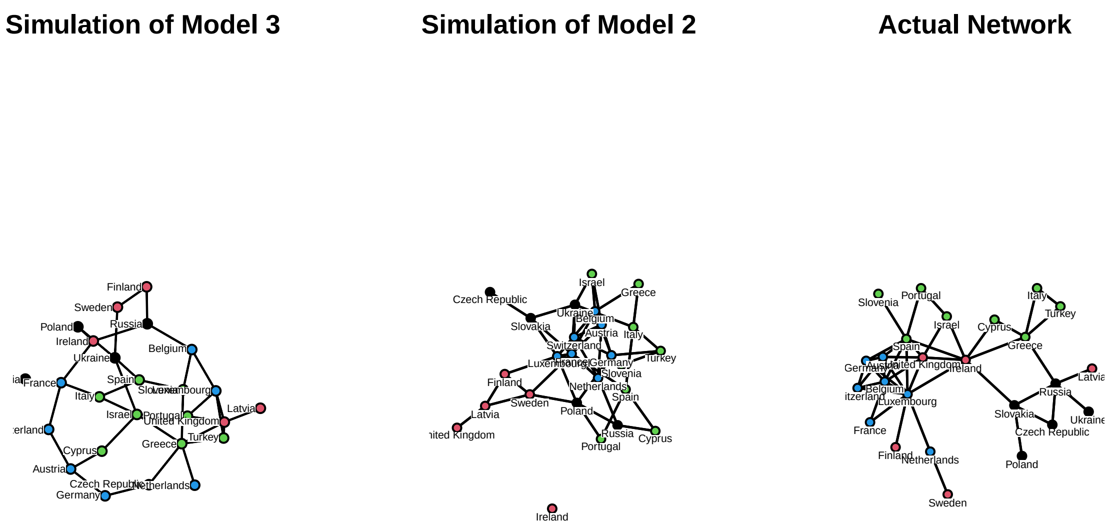
```
## Call:
## ergm(formula = cultural_similarity_subgraph ~ edges + nodecov("population") +
```

```
##     nodecov("press_freedoom") + nodecov("democracy_index") +
##     nodematch("region") + gwdegree(decay = 2, fixed = FALSE,
##     cutoff = 20), control = control.ergm(MCMLE.maxit = 30))
##
## Monte Carlo Maximum Likelihood Results:
##
##                          Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                  -2.673e-01  2.311e+00      0  -0.116   0.9079
## nodecov.population      1.448e-09  4.548e-09      0   0.318   0.7502
## nodecov.press_freedoom -5.725e-02  2.810e-02      0  -2.037   0.0416 *
## nodecov.democracy_index 4.573e-01  2.261e-01      0   2.022   0.0432 *
## nodematch.region        1.379e+00  3.398e-01      0   4.057   <1e-04 ***
## gwdegree                3.830e-01  1.729e+00      0   0.222   0.8247
## gwdegree.decay          3.489e-28  2.167e+00      0   0.000   1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 415.9  on 300  degrees of freedom
##  Residual Deviance: 228.0  on 293  degrees of freedom
##
## AIC: 242  BIC: 267.9  (Smaller is better. MC Std. Err. = 0.1778)
```

We find the structural component to have no significant result. Likely this is due to the selection of only strong ties with a weight of .78 or above.

We now will simulate our second and third models as well as the original network, to compare.

```r
sim1 <- simulate(cultural_mod3, nsim = 1, seed = 42)
sim2 <- simulate(cultural_mod2, nsim = 1, seed = 42)
cultural_exchange_network.vcol <- as.factor(cultural_similarity_subgraph %v% "region")
par(mfrow=c(1, 3))
plot(sim1, vertex.cex = 2, displaylabels = TRUE, label.pos = 2, boxed.labels = T,
      label.border = 0, label.cex = .5, label.pad = 0, arrowhead.cex = .3,
     vertex.col = cultural_exchange_network.vcol,
     main = "Simulation of Model 3")
gplot(sim2, gmode = "graph", displaylabels = TRUE, label.pos = 1,
      boxed.labels = T, label.border = 0, label.cex = .5, label.pad = 0,
      vertex.col = cultural_exchange_network.vcol,
      main = "Simulation of Model 2")
gplot(cultural_similarity_subgraph, gmode = "graph", displaylabels = TRUE, label.pos = 1,
      boxed.labels = T, label.border = 0, label.cex = .5, label.pad = 0,
      vertex.col = cultural_exchange_network.vcol,
      main = "Actual Network")
```



In these simulations we see decent fit. Both model 2 and model 3 simulations captured the grouping by region quite well, which makes sense as this is the models ony significant term.

To further examine fit, we will use the gof function.

```
gof2 <- gof(cultural_mod2)
print(gof2)
```

```
##
## Goodness-of-fit for degree
##
##          obs min mean max MC p-value
## degree0    0   0 0.71   4       1.00
## degree1    6   0 2.56   6       0.08
## degree2    7   0 4.61  10       0.32
## degree3    1   1 5.02  10       0.08
## degree4    2   0 4.68  10       0.30
## degree5    3   0 3.35   8       1.00
## degree6    2   0 2.21   6       1.00
## degree7    2   0 1.17   5       0.66
## degree8    1   0 0.48   3       0.72
## degree9    0   0 0.16   1       1.00
## degree10   1   0 0.05   2       0.08
##
## Goodness-of-fit for edgewise shared partner
##
##      obs min  mean max MC p-value
```

```
## esp0  11  11 21.68  32       0.02
## esp1  13   2 13.82  30       0.98
## esp2   2   0  6.19  15       0.28
## esp3   8   0  2.15   9       0.02
## esp4   7   0  0.74  11       0.04
## esp5   3   0  0.18   4       0.06
## esp6   1   0  0.01   1       0.02
##
## Goodness-of-fit for minimum geodesic distance
##
##     obs min  mean max MC p-value
## 1    45  30 44.77  59       1.00
## 2    82  48 95.13 138       0.62
## 3    87  51 89.75 119       0.76
## 4    61   9 37.84  75       0.08
## 5    23   0 10.50  47       0.26
## 6     2   0  2.11  21       0.64
## 7     0   0  0.32   9       1.00
## 8     0   0  0.09   4       1.00
## Inf   0   0 19.49 109       0.98
##
## Goodness-of-fit for model statistics
##
##                                 obs          min         mean          max
## edges                   4.50000e+01 3.000000e+01 4.477000e+01 5.900000e+01
## nodecov.population      2.84519e+09 1.814266e+09 2.811237e+09 3.837651e+09
## nodecov.press_freedoom  7.13640e+03 4.760400e+03 7.103692e+03 9.457500e+03
## nodecov.democracy_index 6.97420e+02 4.703500e+02 6.939812e+02 9.225400e+02
## nodefactor.region.N     1.60000e+01 4.000000e+00 1.587000e+01 2.400000e+01
## nodefactor.region.S     2.50000e+01 1.000000e+01 2.489000e+01 3.900000e+01
## nodefactor.region.W     3.60000e+01 2.300000e+01 3.546000e+01 5.100000e+01
## nodematch.region        2.60000e+01 1.800000e+01 2.608000e+01 3.500000e+01
##                         MC p-value
## edges                         1.00
## nodecov.population            0.94
## nodecov.press_freedoom        0.96
## nodecov.democracy_index       1.00
## nodefactor.region.N           1.00
## nodefactor.region.S           1.00
## nodefactor.region.W           1.00
## nodematch.region              1.00
```

Because there are only few P-values close to one, we see that our model may not be a great fit, even though it generated alright simulations.

We will now turn to our second set of ERGMs, associated with our economic exchange model. Once again we will start with a null model.

```
economic_mod0 <- ergm(economic_exchange_network ~ edges)
summary(economic_mod0)
```

```
## Call:
## ergm(formula = economic_exchange_network ~ edges)
##
```

```
## Maximum Likelihood Results:
##
##       Estimate Std. Error MCMC % z value Pr(>|z|)
## edges -1.94591    0.09305      0  -20.91   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 1463.9  on 1056  degrees of freedom
##  Residual Deviance:  795.7  on 1055  degrees of freedom
##
## AIC: 797.7  BIC: 802.7  (Smaller is better. MC Std. Err. = 0)
```

Followed by a model that considers the node level attributes population, press freedom index, democracy index, and GDP.

```
economic_mod1 <- ergm(economic_exchange_network ~ edges + nodecov("population") +
                      nodecov("press_freedom") + nodecov("democracy_index") + nodecov("GDP"))
summary(economic_mod1)
```

```
## Call:
## ergm(formula = economic_exchange_network ~ edges + nodecov("population") +
##     nodecov("press_freedom") + nodecov("democracy_index") + nodecov("GDP"))
##
## Maximum Likelihood Results:
##
##                           Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                   -1.935e+00  1.993e+00      0  -0.971   0.3317
## nodecov.population      -1.290e-08  6.637e-09      0  -1.944   0.0519 .
## nodecov.press_freedom    1.350e-02  1.667e-02      0   0.810   0.4181
## nodecov.democracy_index -2.088e-01  1.579e-01      0  -1.322   0.1861
## nodecov.GDP              9.277e-07  1.668e-07      0   5.562   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 1463.9  on 1056  degrees of freedom
##  Residual Deviance:  677.1  on 1051  degrees of freedom
##
## AIC: 687.1  BIC: 711.9  (Smaller is better. MC Std. Err. = 0)
```

In this model we see that only GDP is significant. This makes a lot of sense, as naturally countries with larger economies have a higher amount of countries trading with them. Population is also significant at the .1 level. Interestingly the coefficient on populaton is negative, meaning that as population increases the average number of trading partner decreases. This is quite surprising considering the countries with the highest GDPs typically have high populations.

we will now include a dyadic term, considering if two nodes belong to the same region. We expect the result to be positive, as two countries in the same region are geographically close, which facilitates convinient trading.

```
economic_mod2 <- ergm(economic_exchange_network ~ edges + nodecov("population") +
                        nodecov("press_freedom") + nodecov("democracy_index") +
                        nodecov("GDP") + nodematch("region"))
summary(economic_mod2)
```

```
## Call:
## ergm(formula = economic_exchange_network ~ edges + nodecov("population") +
##     nodecov("press_freedom") + nodecov("democracy_index") + nodecov("GDP") +
##     nodematch("region"))
##
## Maximum Likelihood Results:
##
##                           Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                   -1.862e+00  1.969e+00      0  -0.945   0.3444
## nodecov.population      -1.518e-08  6.814e-09      0  -2.227   0.0259 *
## nodecov.press_freedom    1.509e-02  1.710e-02      0   0.883   0.3775
## nodecov.democracy_index -2.690e-01  1.630e-01      0  -1.650   0.0990 .
## nodecov.GDP              1.050e-06  1.749e-07      0   6.004   <1e-04 ***
## nodematch.region         1.536e+00  2.239e-01      0   6.859   <1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 1463.9  on 1056  degrees of freedom
##  Residual Deviance:  630.1  on 1050  degrees of freedom
##
## AIC: 642.1  BIC: 671.8  (Smaller is better. MC Std. Err. = 0)
```

This dyad term produces a highly significant positive coefficient. Therefore our theory of trading convince seems to prove correct. Also in this model, population becomes significant at the .05 level and democracy index at the .1 level. These two coefficients are both negative as well, which is quite counterintuitive.

Our final economic model includes a structural term, geometrically weighted dyadwise shared partner distribution, which is the sum of geometrically weighted non-edgewise shared partner distribution and geometrically weighted edgewise shared partner distribution.

```r
set.seed(42)
economic_mod3 <- ergm(economic_exchange_network ~ edges + nodecov("population") +
                        nodecov("press_freedom") + nodecov("democracy_index") +
                        nodecov("GDP") + nodematch("region") +
                        dgwdsp(fixed = F), control = control.ergm(MCMLE.maxit = 25))
summary(economic_mod3)
```

```
## Call:
## ergm(formula = economic_exchange_network ~ edges + nodecov("population") +
##     nodecov("press_freedom") + nodecov("democracy_index") + nodecov("GDP") +
##     nodematch("region") + dgwdsp(fixed = F), control = control.ergm(MCMLE.maxit = 25))
##
## Monte Carlo Maximum Likelihood Results:
##
##                           Estimate Std. Error MCMC % z value Pr(>|z|)
## edges                   -1.001e+00  2.184e+00      0  -0.458  0.64671
## nodecov.population      -2.094e-08  7.772e-09      0  -2.695  0.00704 **
## nodecov.press_freedom    2.377e-02  1.964e-02      0   1.210  0.22617
## nodecov.democracy_index -3.727e-01  1.868e-01      0  -1.995  0.04601 *
## nodecov.GDP              1.460e-06  2.223e-07      0   6.570  < 1e-04 ***
## nodematch.region         1.747e+00  2.419e-01      0   7.222  < 1e-04 ***
## gwdsp.OTP               -3.231e-01  4.706e-02      0  -6.867  < 1e-04 ***
## gwdsp.OTP.decay          2.607e-31  2.302e-01      0   0.000  1.00000
## ---
```

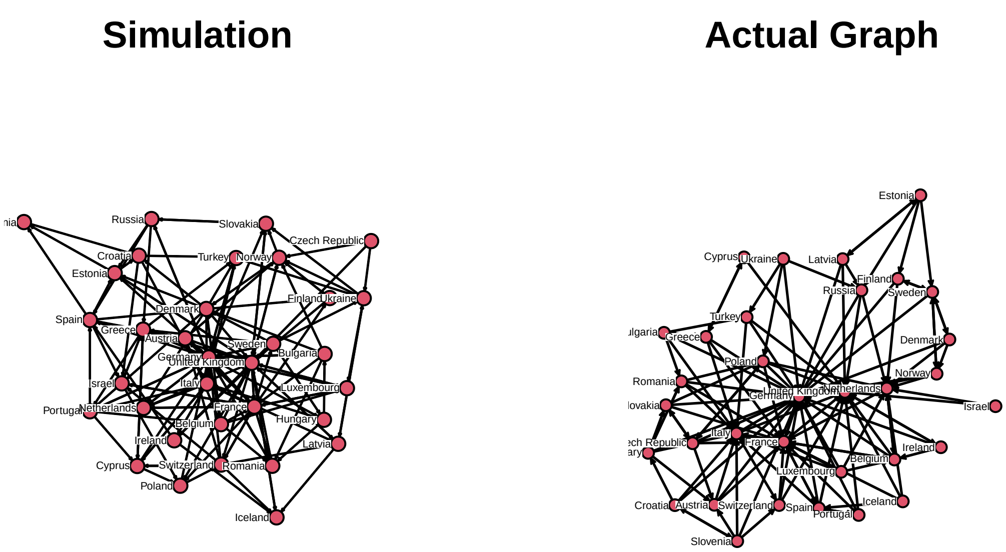
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      Null Deviance: 1463.9  on 1056  degrees of freedom
##  Residual Deviance:  525.1  on 1048  degrees of freedom
##
## AIC: 541.1  BIC: 580.8  (Smaller is better. MC Std. Err. = 0.6276)
```

The addition of this term yields a very significant negative coefficient, and also drops the AIC of our model. Our model's fit is greatly improved, but interpreting the results of our negative coefficient is challenging. Possibly this is due to each leg of a triad, as measured by the dgwdsp term, having to have both directions in it. But in our economic network, most ties are uni-directional, toward the large economic players. Thus this coefficient reflects that the largest players top five outgoing trade partners are quite limited.

In this model significance of democracy index and population also jump up, while the coefficients remain negative.

Finally we will compare a simulated network to out actual network.

```r
sim2 <- simulate(economic_mod2, nsim = 1, seed = 42)
par(mfrow=c(1, 2))
plot(sim2,vertex.cex = 2, displaylabels = TRUE, label.pos = 2, boxed.labels = T,
      label.border = 0, label.cex = .3, label.pad = 0, arrowhead.cex = .5,
     main = "Simulation")
gplot(economic_exchange_network, gmode = "digraph", displaylabels = TRUE, label.pos = 2,
      boxed.labels = T, label.border = 0, label.cex = .3, label.pad = 0,
      arrowhead.cex = .5, main = "Actual Graph")
```



From these graphs it appears are model isn't a bad fit, as Germany, the UK, other central actors remain in the center of the graph.

Conclusion

In this project we learned a lot about the prominent actors. While examining prominence we developed a theory that some networks favor small countries for prominence, while others favor large countries. Specifically small networks were tied strongly via culture to many other countries, while large countries were tied economically to many other countries. We then applied ERGMs and found that the leading cause of countries cultural similarity was actually matching regions at the node level. Our ERGMs looking at the economic connection indicate that countries with a large population are actually less likely to form economic ties. This flies in the face of our previous theory, but also seems quite dubious. Issues with our ERGMs may have been caused by limiting factors in our data collection, which only include each country's top five trading partners. In future, it would be good to collect data on more trading partners, and consider information about the weights of each trading partner.

Data Sources https://en.wikipedia.org/wiki/Democracy_Index https://en.wikipedia.org/wiki/Press_Freedom_Index https://en.wikipedia.org/wiki/List_of_countries_and_dependencies_by_population https://en.wikipedia.org/wiki/United_Nations_geoscheme_for_Europe https://correlatesofwar.org/data-sets/formal-alliances/ http://userpage.fu-berlin.de/~jroose/index_en/main_indexvaluesaz.htm https://wits.worldbank.org/countrysnapshot/en/RUShttps://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal)