UN Peacekeeping and Defense Spending

Technical Report

Submitted by Drew Brooks

In fulfillment of the requirements of the capstone project for the University of Missouri Master's in Data Science and Analytics

January 2021

Code Repository:
https://github.com/Drew-Brooks/UN-Peacekeeping-and-Defense-Spending
Dashboard:
https://shiny.sgn.missouri.edu/students/ban.0878/DATA-SCI-8656/PSDS_capstone/

Section 1: Background and Research Approach

The general research question guiding this project is: which countries contribute the most troops, experts/observers, and police to UN peacekeeping operations and why?

The International Peacekeeping Institute Peacekeeping Database[1] presents month-by-month totals of police, observers, and troops broken out by contributor countries and mission. It also includes useful data such as the regional blocs or affiliations of both contributor and host countries, and the distance between contributor capitals and PKO mission headquarters. Starting with this database, I intend to explore the connection between factors such as proximity to the mission, history of contribution, and defense spending to countries' willingness to contribute personnel to peacekeeping missions. The first two of these factors were the subjects of previous IPI research, so my project primarily concentrates on the role of defense spending. The defense spending factor has not been previously explored by IPI researchers. My hypothesis is that countries that spend relatively less on defense use UN peacekeeping operations (for which they are reimbursed by common funds) as a way of supplementing their defense budgets and gaining deployment experience for their troops and police.

Background on IPI research: IPI, a nonprofit think tank based in New York City, compiled the IPI Peacekeeping Database in 2013 as part of their "Providing for Peacekeeping Project (PPP)." The goal of the PPP was to "analyze the factors that encourage or discourage states from contributing to UN peacekeeping operations."[2] Among the researchers' key findings, they note that although more countries are contributing to peacekeeping operations every year since 1990, proportionally more uniformed personnel are coming from fewer countries. The regional source of most contributions has shifted from Europe in the 1990s to South Asia and Africa today, driven by large increases from Bangladesh, India, Pakistan, Nigeria, Rwanda, and Ethiopia. In a follow-up analysis from 2018, IPI researchers also found that the proportion of peacekeepers coming from countries that border conflict regions has increased since

---

[1] IPI Peacekeeping Database, International Peace Institute, accessed 10/14/2020, available at www.providingforpeacekeeping.org

[2] Chris Perry and Adam C. Smith, "Trends in Uniformed Contributions to UN Peacekeeping: A New Dataset, 1991-2012," International Peace Institute, accessed 10/14/2020, http://www.ipinst.org/wp-content/uploads/publications/ipi_e_pub_trends_un_peacekeeping.pdf.

1990, from less than 3 percent to about 20 percent.[3] The increasing share of peacekeeping contributions from the "global south" is a potential indication that budgetary issues are influencing the decisions of whether and how much to contribute.

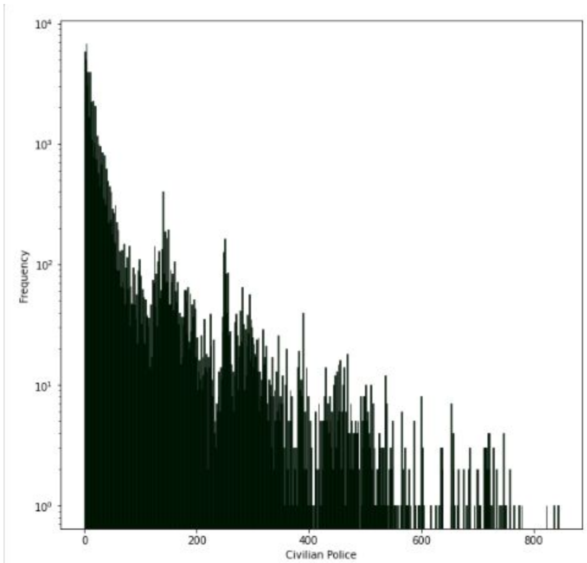
Section 2: Exploratory Data Analysis and Data Carpentry
The IPI dataset consists of 151,844 entries, 1 each for every contributor country for every mission, every month from November 1990 to November 2018. There are 73 columns: Date, Contributor, Mission, 30 geographic and political bloc characteristics for the contributor, the same 30 geographic and political bloc characteristics for the mission country, and the target data fields: Experts on Mission, Formed Police Units, Individual Police, Civilian Police, Troops, Military Observers (what Experts on Mission were called prior to November 2009)[4] and "Total," which is the sum of all contributions in the row.

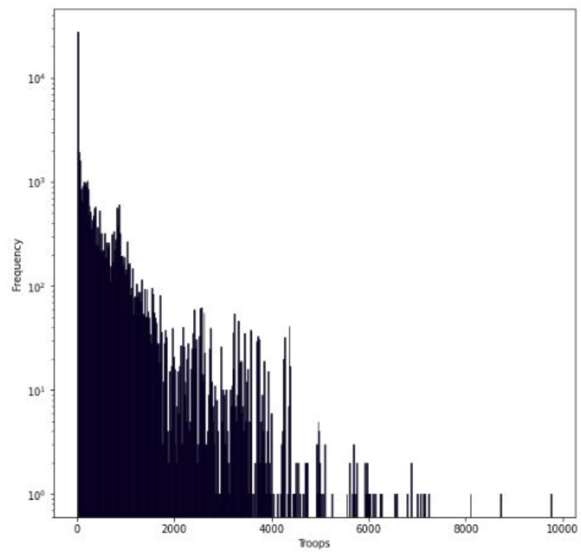
The statistics for the key target data fields are below:

	Experts_on_Mission	Formed_Police_Units	Inidividual_Police	Civilian_Police	Troops	Observers	Total
count	34667.000000	3893.000000	27412.000000	62919.000000	57429.000000	93570.000000	151844.000000
mean	5.881472	212.052915	19.712681	40.160937	322.605913	6.440814	142.623271
std	7.532372	106.214632	33.836267	78.653013	631.394785	8.719371	436.197382
min	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	2.000000	140.000000	4.000000	4.000000	3.000000	2.000000	3.000000
50%	3.000000	147.000000	10.000000	12.000000	29.000000	4.000000	8.000000
75%	7.000000	279.000000	22.000000	33.000000	372.000000	8.000000	39.000000
max	135.000000	574.000000	565.000000	845.000000	9769.000000	791.000000	9779.000000

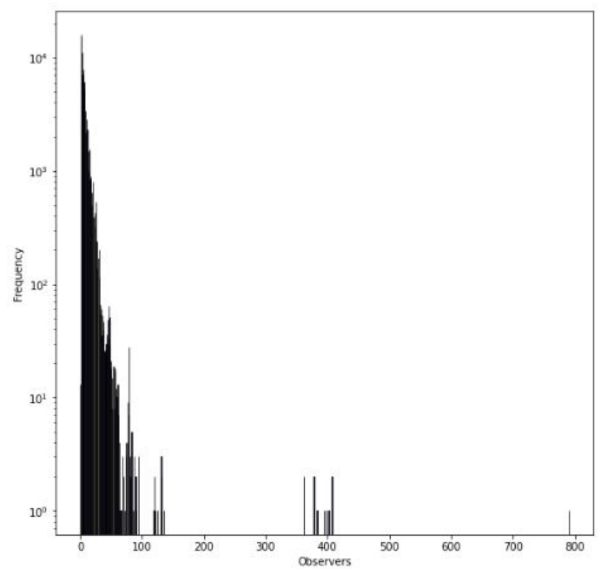
The spread between the 75th percentile contributions and the maximum contribution size is significant. There are lots of small contributions, and a few extremely large contributions. This makes me think it will be fruitful to segment my analysis to see if there are different factors that lead some countries to send only token contributions, some to send sizable contingents, and some to send very large numbers of troops. I am most interested in the Police and Troop totals, since these are the heart of peacekeeping operations:

[3] Thong Nguyen and Paul D. Williams, "Neighborhood Dynamics in UN Peacekeeping Operations, 1990-2017," International Peace Institute, accessed 11/4/2020, https://www.ipinst.org/2018/04/neighborhood-dynamics-in-un-peacekeeping-operations.
[4] IPI Peacekeeping Database: Coding Manual, https://s3.amazonaws.com/ipi-pdf-document-store/ppp-papers/ipi_pub_ppp_coding_manual.pdf, Accessed 10/14/2020/





The "Observers" graph is interesting. Since these personnel are unarmed and serve a reporting and accountability function, it makes sense that most contributions would be at numbers less than 100.[5] The cluster of observers around 400 and the one tiny spike by 800 are outliers, though. When we drill down to see what accounts for them, we find:




```
1993-01-31, Ukraine, UNPROFOR, 398
1993-02-28, Ukraine, UNPROFOR, 378
1993-03-31, Ukraine, UNPROFOR, 378
1993-04-30, Ukraine, UNPROFOR, 394
1993-05-31, Ukraine, UNPROFOR, 362
1993-06-30, Ukraine, UNPROFOR, 362
1993-07-31, Ukraine, UNPROFOR, 407
1993-08-31, Ukraine, UNPROFOR, 407
1993-09-30, Ukraine, UNPROFOR, 402
1993-10-31, Ukraine, UNPROFOR, 397
1993-11-30, Ukraine, UNPROFOR, 383
2000-02-28, Guinea, UNAMSIL, 791
```

UNPROFOR was the UN Protection Force in the former Yugoslavia.[6] It is not immediately clear why Ukraine had such a disproportionately large number of observers in that mission, but given the former USSR's interest in the parties of the conflict it is not unreasonable.

[5] "United Nations Military Observers," Government of Canada Department of National Defence, https://www.canada.ca/en/department-national-defence/services/operations/allies-partners/united-nations/military-observers.html, accessed 11/15/2020.

[6] UNPROFOR, https://peacekeeping.un.org/sites/default/files/past/unprof_b.htm, accessed 11/15/2020.

UNAMSIL was the UN Mission in Sierra Leone, and ran from October 1999 to December 2005. Since the official UN mission data sheet says the maximum total number of observers allowed was 260, we must conclude that the 791 number from Guinea is a bad piece of data. In fact for all the following months, the number of Guinean observers is 12 and the number of total personnel is around 790.

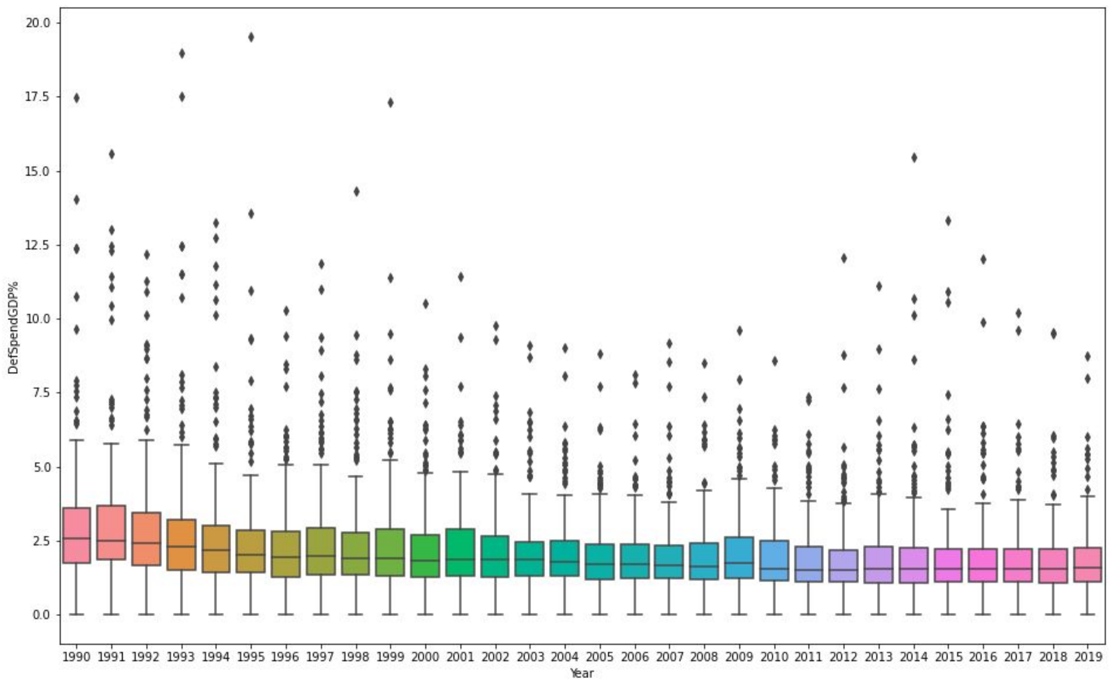
In order to examine the effect of defense spending levels on contributions, I have to merge data on defense spending with the IPI dataset. The World Bank has some good defense economic data.[7] I will use spending as a percent of Gross Domestic Product so that the measurements are more consistent despite inflation and exchange rates. Since my hypothesis relates to defense burden rather than absolute spending power, this measure makes sense. The World Bank data has defense spending as a percent of GDP for an increasing number of countries each year (82 in 1960, peaking at 1999 in 2005, settling at 193 in 2018). When I plot the summary statistics for the years 1990-2018, I noticed that one country (Kuwait) spent 117% of GDP on defense in 1991, the year of the Gulf War:

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	...	2009	2010
count	151.000000	163.000000	174.000000	188.000000	189.000000	185.000000	186.000000	187.000000	182.000000	184.000000	...	196.000000	195.000000
mean	3.516357	3.904546	3.147916	3.012078	2.809148	2.539348	2.453418	2.478946	2.616191	2.625296	...	2.084927	1.930156
std	4.490422	9.296615	3.093961	2.980125	2.956534	2.268414	2.151968	1.915472	2.956071	3.131128	...	1.478538	1.310584
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000
25%	1.728140	1.846909	1.674065	1.496312	1.440544	1.422541	1.285024	1.362277	1.364396	1.307838	...	1.226803	1.151639
50%	2.563538	2.507998	2.435879	2.308353	2.202313	2.035957	1.959158	1.981771	1.896522	1.895821	...	1.736615	1.558146
75%	3.624722	3.696459	3.452684	3.202397	3.024573	2.831627	2.821924	2.931261	2.792795	2.911898	...	2.608292	2.495444
max	48.517267	117.349823	31.786024	21.329114	29.727685	19.539112	20.344819	11.860415	32.497067	34.377768	...	9.617199	8.565677

8 rows × 29 columns

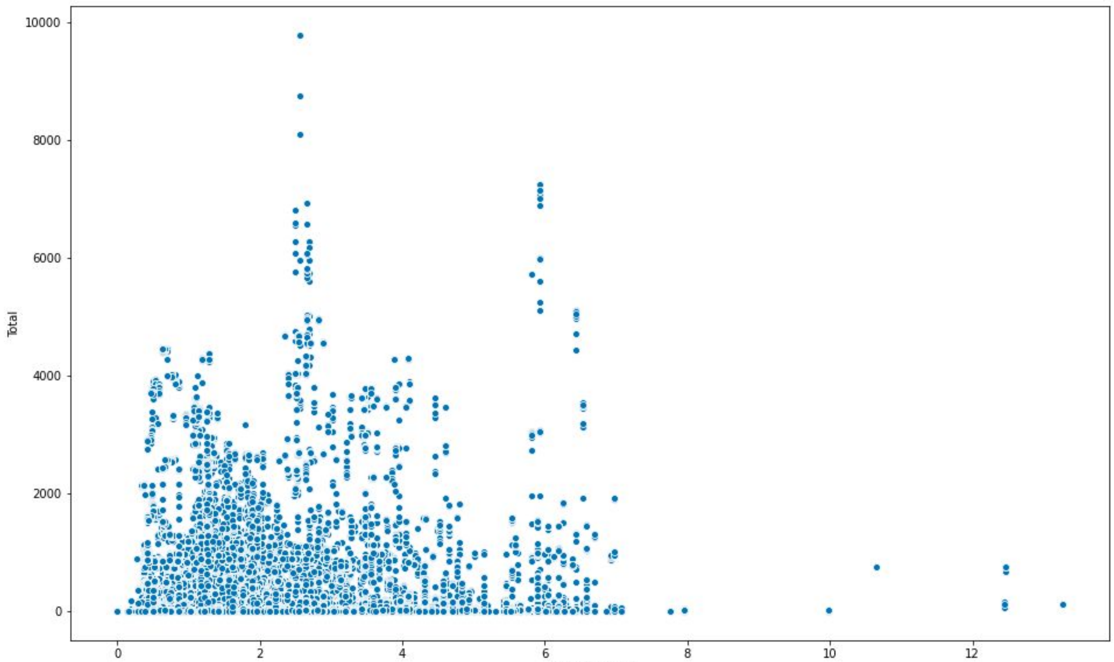
I checked, and Kuwait did not have any peacekeepers in other countries that year. When I throw out that anomaly, plus some other cases where countries had defense burdens over 20% but no peacekeepers deployed, I get the below box plot:

[7] Military Expenditure % GDP, World Bank, accessed 10/14/2020, https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS



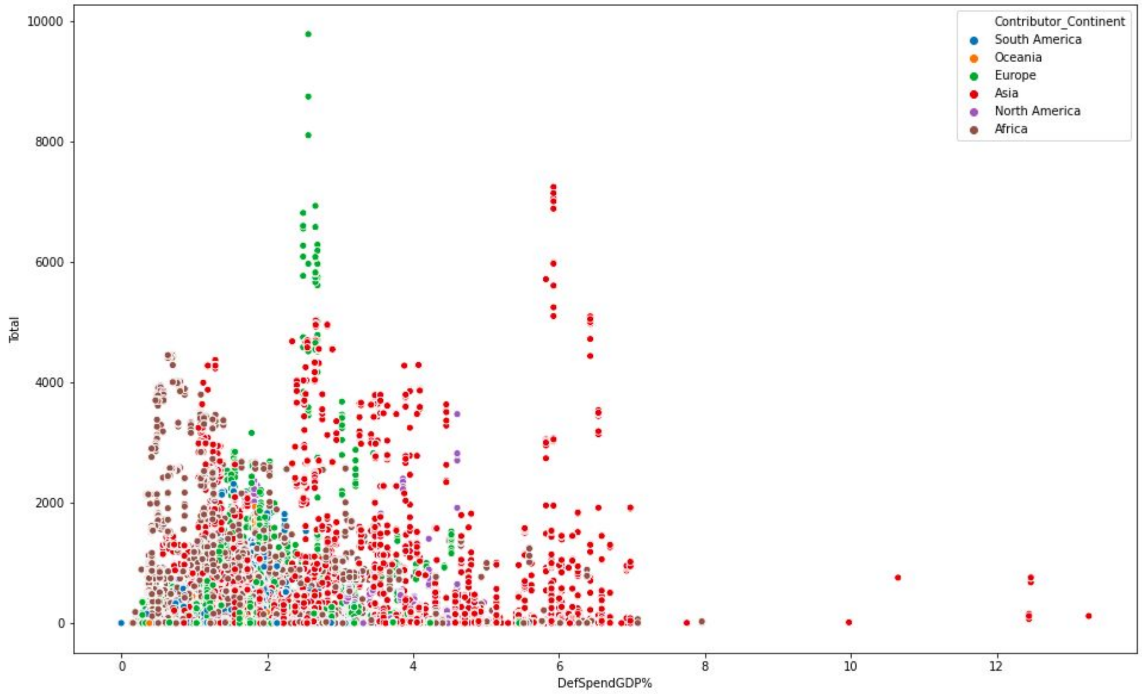
The average defense burden has shrunk since the end of the Cold War, from 3.52% in 1990 to 1.85% in 2018—a decline of almost 50%.

Once I melted and merged the defense spending data with the contributors dataframe, I could plot spending vs contributions:

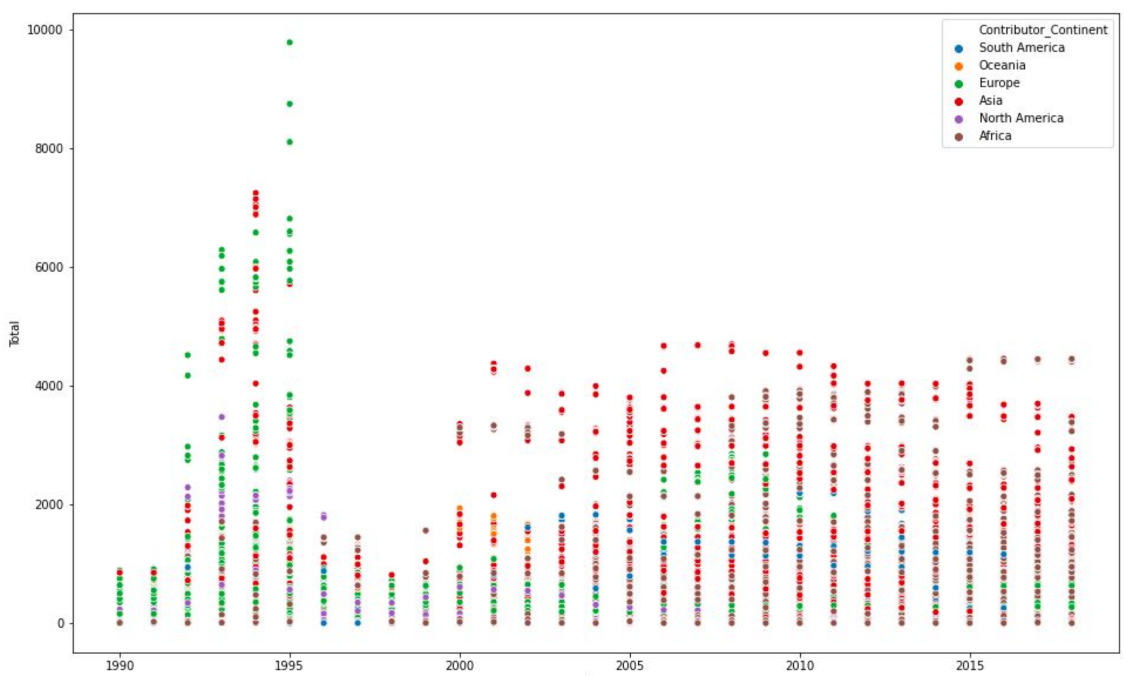


At first this doesn't tell much of a story, but when I change the color to show continents, it suggests some groupings that might be interesting to explore. Specifically, some

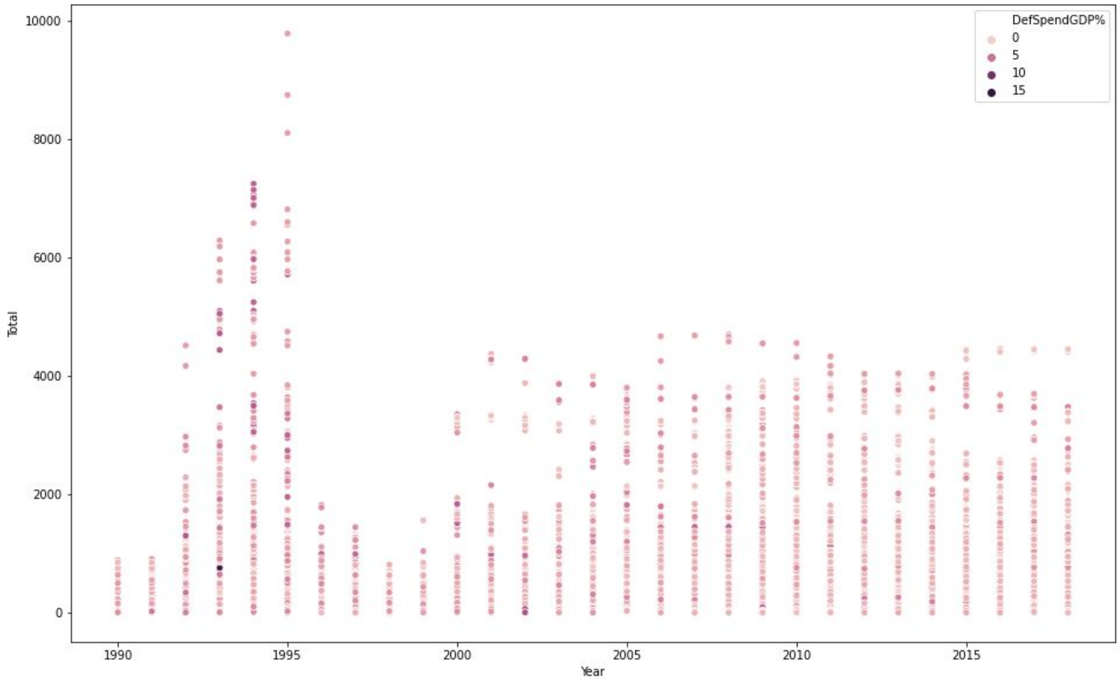
African, European, and Asian countries that appear to be punching above their weight:



When I plot contributors by year, you can definitely see the shift from Europe to Asia and the Africa as the primary contributors over the decades since the Cold War:



The pattern by defense spending is less obvious, but there might be something there, since the high spenders seem more concentrated higher on the y-axis in previous decades:



To complete my analysis of defense spending, I will use a linear regression to determine what relationship, if any, there is between Troop levels and spending, both overall and with the addition of Date as a time variable.

Section 3: Regression Analysis

Beginning with a .csv file saved down after the data carpentry work detailed above, I selected just the dependent and independent variables I was interested in, and also renamed a few to correct misspelling and get rid of problematic variables. I also changed all NA values to 0, because it's useful to the analysis to know when countries were sending zero troops:

```
vars <- c("Date", "Experts_on_Mission", "Formed_Police_Units",
"Inidividual_Police", "Civilian_Police", "Troops", "Observers", "Total",
"DefSpendGDP%")
contrib <- contrib[,(colnames(contrib) %in% vars)]
contrib$Date <- as_date(contrib$Date)
contrib$Date <- as.numeric(contrib$Date)
contrib[is.na(contrib)] = 0
names(contrib)[names(contrib) == "Inidividual_Police"] <-
"Individual_Police"
names(contrib)[names(contrib) == "DefSpendGDP%"] <- "DefSpend"
```

I then split the data into 70% training and 30% test, and examined the correlations in the training data:

```
cor(train_data)
```

	Date	Experts_on_Mission	Formed_Police_Units	Individual_Police	Civilian_Police	Troops	Observers	Total	DefSpend
Date	1.00000000	0.28646502	0.14327954	0.1929675851	0.08582623	0.02683658	-0.09091557	0.03472192	-0.1705215839
Experts_on_Mission	0.28646502	1.00000000	0.21997439	0.2053625588	0.17084971	0.33529559	0.46221204	0.35110962	-0.0100090933
ormed_Police_Units	0.14327954	0.21997439	1.00000000	0.2609786941	0.75018136	0.20667393	0.07556786	0.29208203	0.0455272277
Individual_Police	0.19296759	0.20536256	0.26097869	1.0000000000	0.45146189	0.09894434	0.04195991	0.15129567	0.0003110808
Civilian_Police	0.08582623	0.17084971	0.75018136	0.4514618852	1.00000000	0.22106110	0.07049268	0.33654846	0.0965606090
Troops	0.02683658	0.33529559	0.20667393	0.0989443450	0.22106110	1.00000000	0.30046829	0.99262007	0.0450727009
Observers	-0.09091557	0.46221204	0.07556786	0.0419599069	0.07049268	0.30046829	1.00000000	0.31482117	0.0696045275
Total	0.03472192	0.35110962	0.29208203	0.1512956735	0.33654846	0.99262007	0.31482117	1.00000000	0.0563869971
DefSpend	-0.17052158	-0.01000909	0.04552723	0.0003110808	0.09656061	0.04507270	0.06960453	0.05638700	1.0000000000

The strongest correlations seem to be that one kind of contribution begets others. This makes sense as nations will send "packages" of forces to missions that include multiple kinds of personnel. The specific hypothesis I want to test, though, is the idea that nations with low defense spending send extra peacekeepers for the training and because of the reimbursement money. The low correlation between defense spending percentages and troops indicates little support for this hypothesis, but I fit a regression just to see.

In [8]:
```
spend_reg <- lm(train_data$Troops ~ train_data$DefSpend)
summary(spend_reg)
```

```
Call:
lm(formula = train_data$Troops ~ train_data$DefSpend)

Residuals:
   Min     1Q Median     3Q    Max 
-256.6 -124.7 -114.4  -99.2 9633.0 

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)    
(Intercept)           92.205      2.481   37.16   <2e-16 ***
train_data$DefSpend   17.084      1.190   14.36   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 420.9 on 101228 degrees of freedom
Multiple R-squared:  0.002032,	Adjusted R-squared:  0.002022 
F-statistic: 206.1 on 1 and 101228 DF,  p-value: < 2.2e-16
```

Indeed, the R-squared of 0.002 shows very little support for the hypothesis, or rather does not allow me to reject the null hypothesis that defense spending levels do not impact troop contributions to UN peacekeeping operations.

Because of the shift over time in contributions from relatively wealthier to less wealthy countries, I supposed there might be more of an effect from defense spending levels if I added in a time variable:

In [9]:
```r
# So the immediate result is that there is very little predictive value in def
spend_reg2 <- lm(train_data$Troops ~ train_data$DefSpend + train_data$Date)
summary(spend_reg2)
```

```
Call:
lm(formula = train_data$Troops ~ train_data$DefSpend + train_data$Date)

Residuals:
   Min     1Q Median     3Q    Max 
-248.1 -128.5 -113.2  -91.8 9654.9 

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)    
(Intercept)         1.162e+01  7.631e+00   1.523    0.128    
train_data$DefSpend 1.938e+01  1.207e+00  16.057   <2e-16 ***
train_data$Date     5.629e-03  5.041e-04  11.165   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 420.6 on 101227 degrees of freedom
Multiple R-squared:  0.003259,	Adjusted R-squared:  0.003239 
F-statistic: 165.5 on 2 and 101227 DF,  p-value: < 2.2e-16
```

The R-squared value increased to 0.003, but that still is not enough of an effect to conclude there is any significant correlation, even over time. Finally, although my hypothesis about defense spending is most likely to impact Troops (since police are typically funded under different budget lines, i.e. interior ministries), I tested to see if there was an effect on Total contributions:

In [10]:
```r
# Maybe there's more correlation with total contributions vs just Troops
spend_reg3 <- lm(train_data$Total ~ train_data$DefSpend + train_data$Date)
summary(spend_reg3)
```

```
Call:
lm(formula = train_data$Total ~ train_data$DefSpend + train_data$Date)

Residuals:
   Min     1Q Median     3Q    Max 
-306.8 -143.4 -120.9  -82.6 9647.6 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)         -4.1483589  7.9405719  -0.522    0.601    
train_data$DefSpend 25.3351792  1.2560079  20.171   <2e-16 ***
train_data$Date      0.0075292  0.0005246  14.353   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 437.7 on 101227 degrees of freedom
Multiple R-squared:  0.005204,	Adjusted R-squared:  0.005184 
F-statistic: 264.8 on 2 and 101227 DF,  p-value: < 2.2e-16
```

Again the R-squared increased, but 0.005 is still so miniscule that I cannot reject the null hypothesis. There was no point in trying to predict values in the test data because the correlation effect was so small. This analysis does not show a significant effect on peacekeeping contributions stemming from defense spending levels.

Section 4: Clustering and Classification

To examine other factors that might impact contribution levels and perhaps create a policy-useful model to identify countries where peacekeeping contributions could be sourced, I took advantage of the geographic and regional bloc information compiled by the IPI to use a clustering algorithm to try and find useful groupings that could tell me something about national tendencies to contribute or not contribute to peacekeeping operations.

Starting with the same .csv as I did for the regressions, I performed the same cleaning steps, then dropped all columns that were non-numeric, as well as the longitude/latitude columns so those values would not throw off the clustering algorithm:

```
# Replace NAs with zero
contrib[is.na(contrib)] = 0
# Drop geo columns and factors
nonnum <- names(Filter(is.factor, contrib))
xcluded <- c(nonnum, 'Mission_HQ_Longitude',"Mission_HQ_Latitude","Contributor_Capital_Latitude","Contributor_
clustdata <- contrib[, -which(names(contrib) %in% xcluded)]
names(clustdata)

```

```
'' 'Contributor_NAM' 'Contributor_G77' 'Contributor_AU' 'Contributor_Arab_League' 'Contributor_OIC' 'Contributor_CIS' 'Contributor_G20'
'Contributor_EU' 'Contributor_NATO' 'Contributor_G8' 'Contributor_OECD' 'Contributor_ASEAN' 'Contributor_OAS' 'Contributor_Shanghai'
'Contributor_GCC' 'Contributor_UMA' 'Contributor_COMESA' 'Contributor_CENSAD' 'Contributor_EAC' 'Contributor_ECCAS'
'Contributor_ECOWAS' 'Contributor_IGAD' 'Contributor_SADC' 'Mission_NAM' 'Mission_G77' 'Mission_AU' 'Mission_Arab_League'
'Mission_OIC' 'Mission_CIS' 'Mission_G20' 'Mission_EU' 'Mission_NATO' 'Mission_G8' 'Mission_OECD' 'Mission_ASEAN' 'Mission_OAS'
'Mission_Shanghai' 'Mission_GCC' 'Mission_UMA' 'Mission_COMESA' 'Mission_CENSAD' 'Mission_EAC' 'Mission_ECCAS' 'Mission_ECOW
'Mission_IGAD' 'Mission_SADC' 'Experts_on_Mission' 'Formed_Police_Units' 'Individual_Police' 'Civilian_Police' 'Troops' 'Observers' 'Total'
'Year' 'DefSpend'
```

To make it more computationally feasible, I then limited the data to just those columns that pertained to the contributor, then sampled 10,000 rows:

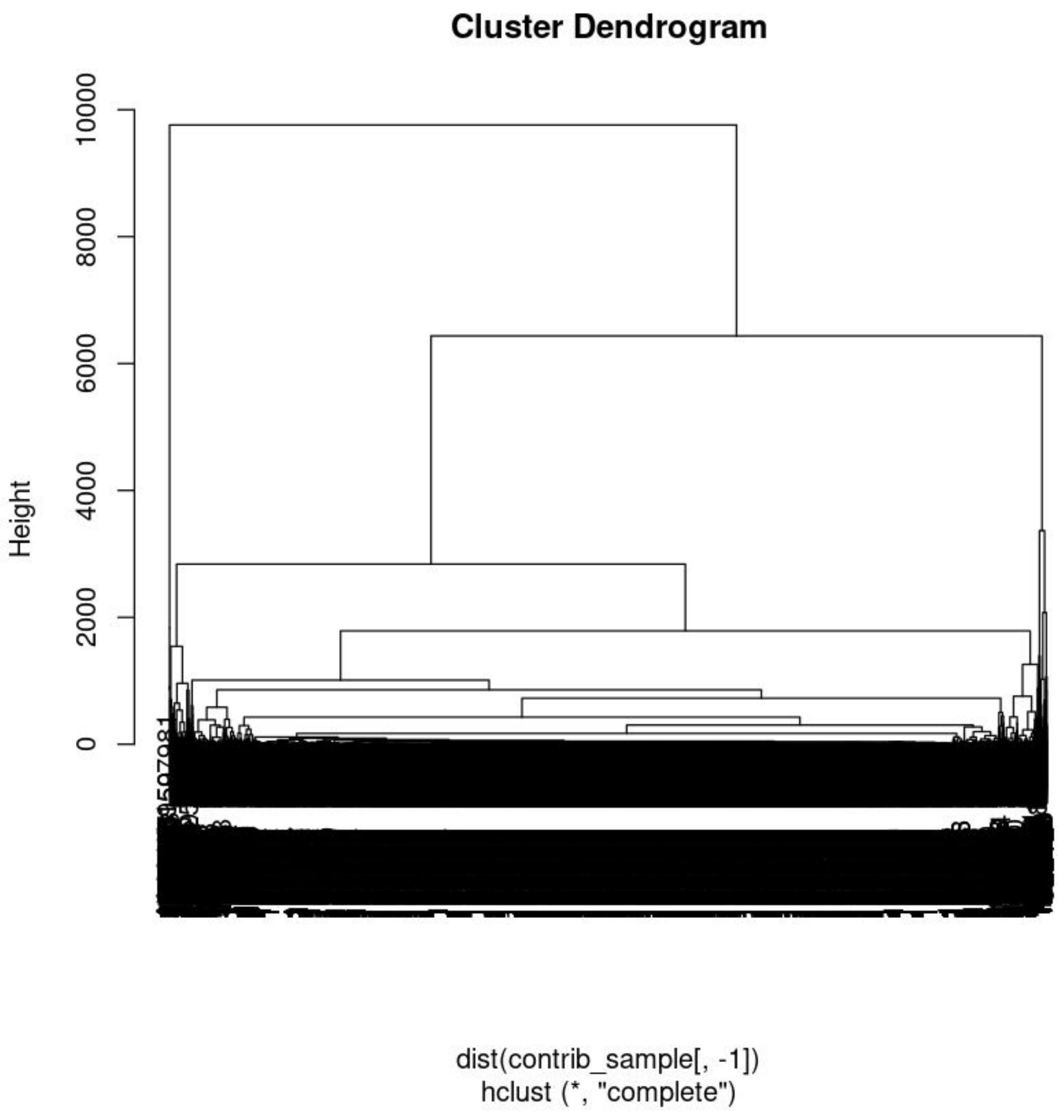
```
# Just cluster on characteristics of contributor
contrib_clustdata <- clustdata %>% dplyr::select(-contains("Mission"))
head(contrib_clustdata)
```

	Contributor_NAM	Contributor_G77	Contributor_AU	Contributor_Arab_League	Contributor_OIC	Contributor_CIS	Contr
0	0	1	0	0	0	0	
1	0	1	0	0	0	0	
2	0	1	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	
5	0	0	0	0	0	0	

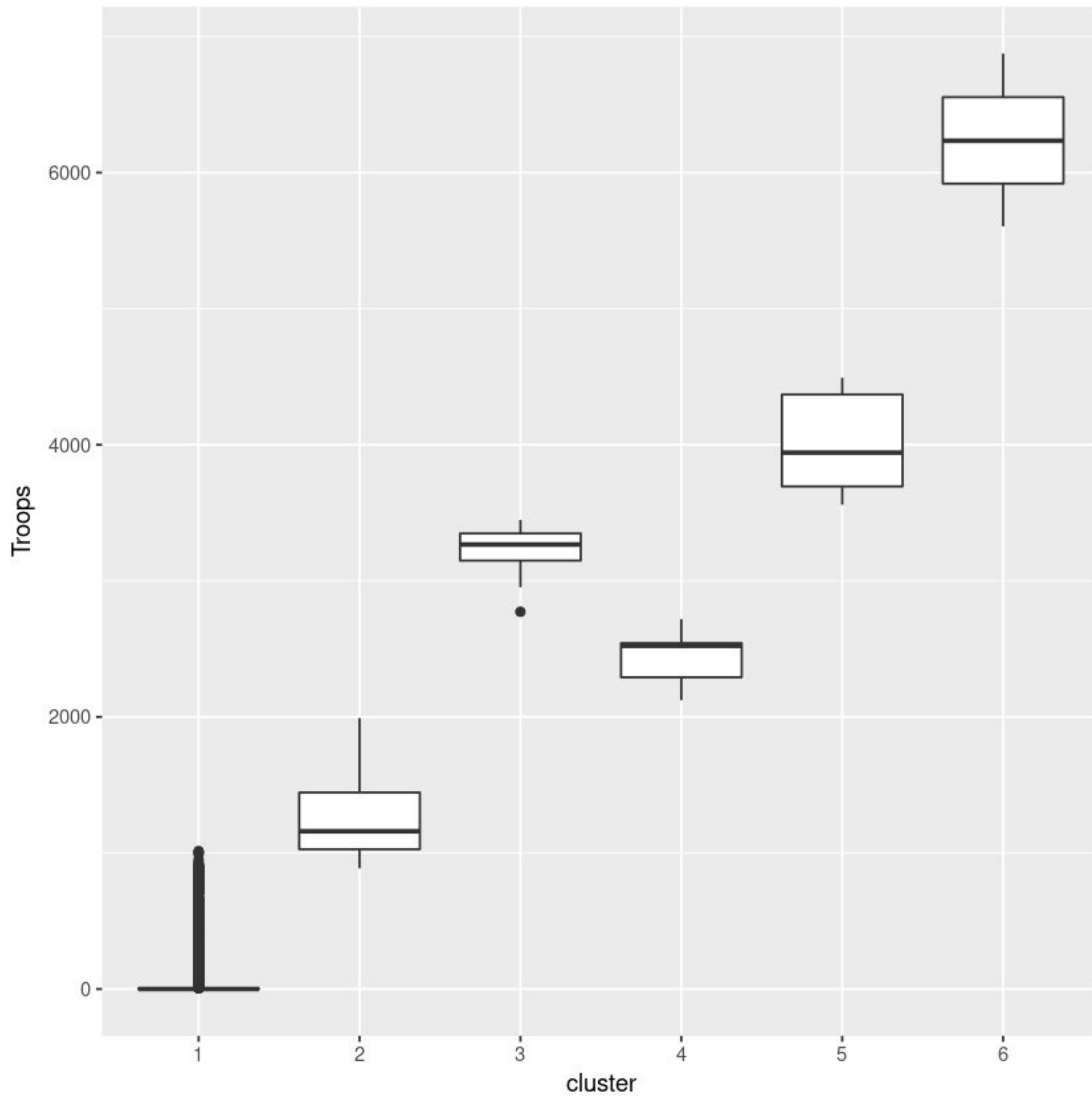
```
contrib_sample <- sample_n(contrib_clustdata,10000)
contrib_sample <- contrib_sample[,-31] # Drop the Year
str(contrib_sample)
```

Next I used the hierarchical clustering package (because I did not know how many clusters would be useful before I started) to find clusters in the data (excluding column 1, the :

```
set.seed(100)
clust <- hclust(dist(contrib_sample[,-1]), method="complete")
plot(clust)
```



Since the troop levels are the most relevant to my research question, I adjusted the cut line until I got a useful separation based on the variable Troop:



These clusters appear to be very useful from an analytic perspective, since they step up in terms of the numbers of troops contributed in each from token contributions to very substantial brigade and even division-sized forces. The skewness towards few or no troops is still very present, however, since cluster 1 includes 9654 out of the 10,000 observations, and cluster 2 contains 244.

I expected this skewness to impact the usefulness of my classification model to provide useful predictions of troop contributions, but went ahead anyway. I split into training and test groups, then used a Latent Dirichlet allocation to predict the cluster of each observation:

```r
# Splitting into train/test to see if an LDA model can predict the clusters
library(caTools)
set.seed(100) # set.seed() will help us to reproduce the results.
split = sample.split(contrib_sample$cluster, SplitRatio=0.7)

sample_train  = subset(contrib_sample, split==TRUE)

# Test data will have the rest 30% of data
sample_test  = subset(contrib_sample, split==FALSE)
```

```r
library(MASS)
ldafit=lda(cluster~ .,data=sample_train[,-1]) #All variable except the contributor name
summary(ldafit)
```

```
Warning message in lda.default(x, grouping, ...):
"variables are collinear"
```

```
        Length Class  Mode
prior     6    -none- numeric
counts    6    -none- numeric
means   180    -none- numeric
scaling 150    -none- numeric
lev       6    -none- character
svd       5    -none- numeric
N         1    -none- numeric
call      3    -none- call
terms     3    terms  call
xlevels   0    -none- list
```

The model achieved 97.8% accuracy, but the confusion table shows that the errors are primarily misclassifying contributors with higher troop levels into clusters of lower troop levels, which is not especially useful. Since the point would be to identify countries that should be contributing more but are not, I would rather see errors in the other direction--countries that appear as if they should be contributing more than they are.

```r
pred <- predict(ldafit, sample_test)
```

```r
# Create a confusion table.
table(pred$class, sample_test$cluster)
mean(pred$class==sample_test$cluster)

```

	1	2	3	4	5	6
1	2835	0	0	0	0	0
2	61	71	0	0	0	0
3	0	0	6	1	1	0
4	0	2	0	16	0	0
5	0	0	0	0	5	0
6	0	0	0	0	0	1

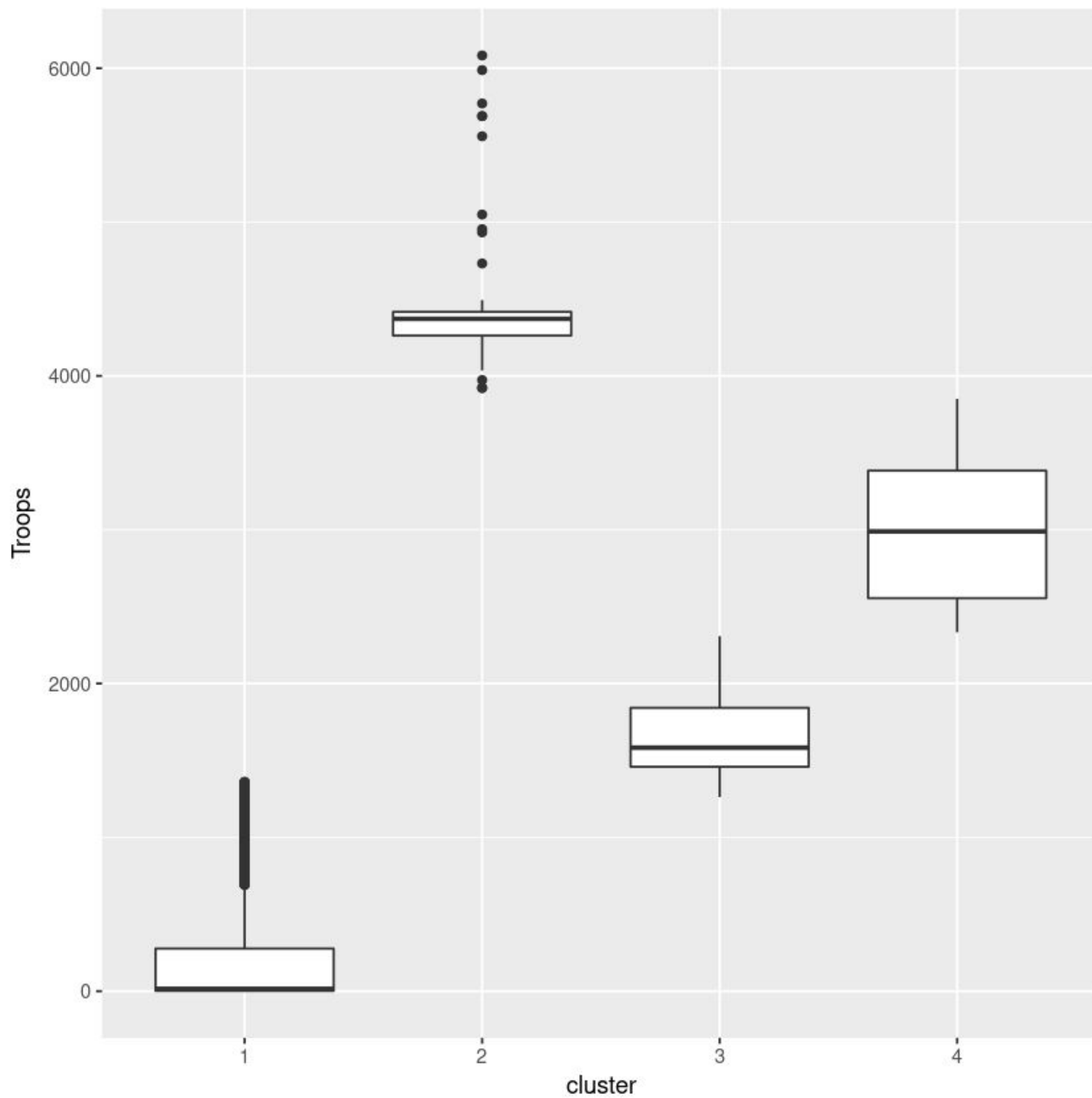
0.978326108702901

To counteract the left-skewness, I repeated the clustering and classification with only rows where the troop contribution was greater than zero:

```r
# re-do clustering and LDA with just nonzero contributions
contrib_clustdata2 <- subset(contrib_clustdata, contrib_clustdata$Troops>0)
contrib_sample2 <- sample_n(contrib_clustdata2,10000)
contrib_sample2 <- contrib_sample2[,-31] # Drop the Year
```

```r
set.seed(100)
clust2 <- hclust(dist(contrib_sample2[,-1]), method="complete")
#plot(clust2)
```

The clustering showed less skewness, with only 9446 observations in cluster 1 and 273 and 237 in cluster 3 and 4, respectively. I reduced the cut to four clusters since any more than that was returning zero observations above four. They are still analytically useful in terms of cut lines, with troop levels between 1-1000, 1000-2000, 2000-4000, and over 4000.



With this method, there are more errors classifying countries as contributing more than they actually do, which suggests it could be useful for policy-predictive purposes:

```r
pred2 <- predict(ldafit2, sample_test2)
```

```r
# Create a confusion table.
table(pred2$class, sample_test2$cluster)
mean(pred2$class==sample_test2$cluster)
```

	1	2	3	4
1	2806	0	0	0
2	0	9	0	2
3	28	0	80	5
4	0	4	2	64

0.986333333333333

Section 5: Dashboard

In order to maximize the usefulness of this data for analytic and policy consumers, I created a prototype data dashboard to display peacekeeping contributors by report month and type of contribution. Because the IPI data ends in 2018, I took data for the dashboard direct from UN because it is kept up to date, although it only goes back to 2002 instead of 1990. The dashboard uses a choropleth to display contributions by color saturation, and clicking on a country shows a bar graph below with the contributions by type.

The prototype dashboard is here:
https://shiny.sgn.missouri.edu/students/ban.0878/DATA-SCI-8656/PSDS_capstone/

I am still working with the analytic consumers of this data to perfect the dashboard--the next step will be to add another tab with a filterable data table showing which mission each contributor is participating in. If it is accepted in its final form, I will then move it to a permanent shiny server and set it up to auto ingest new data from the UN site.

Section 6: Conclusions

This analysis did not find evidence of a significant direct relationship between defense spending levels and troop contributions to UN missions. Further analyses of specific countries, or regressions limited to single years might show more direct relationships, but those would be less meaningful from a policy standpoint because they would note allow for useful prediction unless the results carried over multiple years or multiple countries.

The analysis did, however, suggest that using the continent and regional bloc alignment of the contributing country along with defense spending could allow an analyst to create a useful predictive model. To refine the model further, we would need to add in data for all countries of the world instead of just those with some contribution, and possibly segment the data more by time blocks to try to increase the effect since political circumstances change so much over the decades. A close examination of the misclassified countries might also reveal other variables that would improve the model. Ultimately, however, purely quantitative modeling such as this research used is only useful in conjunction with deep subject matter expertise and a holistic understanding of the factors that weigh on national decisions to participate in peacekeeping operations.