

```
#####
```

```
### Data Visualization Assignment 4.1 ###
```

```
# Connor Harrison, Mar 10, 2019 #
```

```
#####
```

```
# Load Packages
```

```
library(tidyverse)
```

```
library(readr)
```

```
library(readxl)
```

```
# Load Data
```

```
ForeignAssistance_data <- read_csv("~/Georgetown Docs/Data/ForeignAssistance-  
FullDataSet/ForeignAssistance-FullDataSet-2017-and-Later.csv")
```

```
region <- read_excel("~/Georgetown Docs/Data/countries of the world.xls")
```

```
gdp_full <- read_excel("~/Georgetown Docs/Data/gapdata_gdp_ppp_v14.xlsx")
```

```
population <- read_csv("~/Georgetown  
Docs/Data/UNdata_Export_20190210_173214793/UNdata_Export_20190210_173214793.csv")
```

```
# Clean GDP Data
```

```
gdp <- gdp_full[, 1:3]
```

```
gdp_2018 <- filter(gdp, Year=='2018')
```

```
gdp_2018 <- select(gdp_2018, Country='Area', gdp='GDP per capita - with interpolations')
```

```
# Keep Data for 2018
```

```
FA_2018 <- filter(ForeignAssistance_data, Award_Transaction_Fiscal_Year==2018)
```

```
# Keep and Rename Relevant Variables
```

```
FA_2018 <- select(FA_2018, Year = 'Award_Transaction_Fiscal_Year', Category =  
'Award_Transaction_US_Foreign_Assistance_Category',
```

```
Amount = 'Award_Transaction_Value', Country = 'Recipient_Location')
```

```
# Order by Country
```

```
FA_2018 <- arrange(FA_2018, desc(Country, Category))
```

```
# Create Variable: Sum of Total FA for each country
```

```
FA_2018_sum_total <- FA_2018 %>% group_by(Country) %>%  
  mutate(Sum_Amount = sum(Amount))
```

```
# Drop Duplicate Observations
```

```
FA_2018_sum_total <- select(FA_2018_sum_total, Year, -Category, Country, Sum_Amount, -Amount)  
FA_2018_unique_total <- unique(FA_2018_sum_total)
```

```
# Clean Region and Country Data to Merge
```

```
region <- select(region, Country='Data is public domain from US government.', region='..2')  
FA_2018_regions <- left_join(FA_2018_unique_total, region, by="Country")
```

```
# Check Missing Values on Merge
```

```
region_missing <- filter(FA_2018_regions, is.na(region))  
region_missing <- arrange(region_missing, desc(Country))
```

```
# Join GDP Data
```

```
FA_2018_complete <- left_join(FA_2018_regions_total, gdp_2018, by="Country")
```

```
# Drop observations with missing gdp data
```

```
FA_2018_complete <- filter(FA_2018_complete, !is.na(gdp))
```

```
# Add Missing Region Observations
```

```
FA_2018_complete <- FA_2018_complete %>%  
  mutate(region=replace(region, Country=="Trinidad and Tobago", "LATIN AMER. & CARIB"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
  mutate(region=replace(region, Country=="Timor-Leste", "OCEANIA"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
  mutate(region=replace(region, Country=="South Sudan", "SUB-SAHARAN AFRICA"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
  mutate(region=replace(region, Country=="Sao Tome and Principe", "SUB-SAHARAN AFRICA"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
  mutate(region=replace(region, Country=="Bosnia and Herzegovina", "EASTERN EUROPE"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
  mutate(region=replace(region, Country=="Montenegro", "EASTERN EUROPE"))
```

```
# Base Plot
```

```
ggplot(data = FA_2018_complete,
```

```
  aes(area = Sum_Amount, fill = gdp, label = Country)) +
```

```
  geom_treemap() +
```

```
  geom_treemap_text(colour = "white", place = "centre") +
```

```
  scale_fill_gradient(limits = c(0, 15000)) +
```

```
  facet_wrap(~region, ncol = 3)
```

```
# Too many regions w/ facet
```

```
# Drop N. America and NA regions
```

```
# Combine Remaining Regions
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="NEAR EAST", "NEAR EAST & NORTH AFRICA"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="NORTHERN AFRICA", "NEAR EAST & NORTH AFRICA"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="BALTICS", "EUROPE"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="WESTERN EUROPE", "EUROPE"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="EASTERN EUROPE", "EUROPE"))
```

```
FA_2018_complete <- FA_2018_complete %>%
```

```
mutate(region=replace(region, region=="C.W. OF IND. STATES", "CENTRAL ASIA"))
```

```
FA_2018_complete <- filter(FA_2018_complete, region != "NORTHERN AMERICA" & region != "NA")
```

```
# Just SSA Data Set
```

```
FA_2018_ssa <- filter(FA_2018_complete, region == "SUB-SAHARAN AFRICA")
```

```
# Revised Plot 1.0
```

```
ggplot(data = FA_2018_complete,
```

```
  aes(area = Sum_Amount, fill = gdp, label = Country)) +
```

```
geom_treemap() +
```

```
geom_treemap_text(colour = "white", place = "centre") +
```

```
scale_fill_gradient2(limits = c(0, 15000), low = "white", mid = "skyblue", high = "blue4", name="GDP
per Capita") +
facet_wrap(~region, ncol = 3) +
labs(title = paste("Strategic Funding"),
      subtitle = "US Regional Foreign Assistance is targeted to Strategic Partners \nrather than countries
with the greatest financial need",
      caption = "Source: ForeignAssistance.gov; World Bank; UN Data\n*Tile Area is Propotional to Share
of Regional Aid")
```