

# Final Project

*Evan Tiffany*

*11/14/2019*

```
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(ggplot2)
library(maps)
library(ggthemes)

popVWaste = filter(read.csv('coastal-population-vs-mismanaged-plastic.csv'), Coastal.population != "NA")
wasteVGDP = filter(read.csv('per-capita-mismanaged-plastic-waste-vs-gdp-per-capita.csv'), Per.capita.mismanaged.plastic.waste.kg.per.person != "NA")
waste = filter(read.csv('per-capita-plastic-waste-vs-gdp-per-capita.csv'), Per.capita.plastic.waste.kg.per.person != "NA")

dat = left_join(popVWaste, waste, by = c("Entity", "Code", "Year", "Total.population..Gapminder."))

## Warning: Column `Entity` joining factors with different levels, coercing to
## character vector

## Warning: Column `Code` joining factors with different levels, coercing to
## character vector

dat = left_join(dat, wasteVGDP, by = c("Entity", "Code", "Year", "Total.population..Gapminder."))

## Warning: Column `Entity` joining character vector and factor, coercing into
## character vector

## Warning: Column `Code` joining character vector and factor, coercing into
## character vector

dat <- dat[,0:9]
colnames(dat) <- c("Entity", "Code", "Year", "MismanagedWaste", "CoastalPopulation", "TotalPopulation", "GDP", "Population", "Year")

map.dat <- map_data("world")
map.dat$region <- gsub(paste0("\\b(",paste("USA", collapse="|"),")\\b"), "United States", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("UK", collapse="|"),")\\b"), "United Kingdom", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Virgin Islands", collapse="|"),")\\b"), "British Virgin Islands", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Ivory Coast", collapse="|"),")\\b"), "Cote d'Ivoire", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Faroe Islands", collapse="|"),")\\b"), "Faeroe Islands", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Republic of Congo", collapse="|"),")\\b"), "Congo", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Micronesia", collapse="|"),")\\b"), "Micronesia (country)", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Saint Martin", collapse="|"),")\\b"), "Sint Maarten (Dutch part of Saint Martin)", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Trinidad", collapse="|"),")\\b"), "Trinidad and Tobago", map.dat$region)
map.dat$region <- gsub(paste0("\\b(",paste("Tobago", collapse="|"),")\\b"), "Trinidad and Tobago", map.dat$region)
```

```

map.dat$region <- gsub(paste0("\\b(",paste("Antigua", collapse="|"),")\\b"), "Antigua and Barbuda", map
map.dat$region <- gsub(paste0("\\b(",paste("Barbuda", collapse="|"),")\\b"), "Antigua and Barbuda", map
map.dat$region <- gsub(paste0("\\b(",paste("Saint Kitts", collapse="|"),")\\b"), "Saint Kitts and Nevis
map.dat$region <- gsub(paste0("\\b(",paste("Nevis", collapse="|"),")\\b"), "Saint Kitts and Nevis", map
map.dat$region <- gsub(paste0("\\b(",paste("Saint Vincent", collapse="|"),")\\b"), "Saint Vincent and t
map.dat$region <- gsub(paste0("\\b(",paste("Grenadines", collapse="|"),")\\b"), "Saint Vincent and the
map.dat$region <- gsub(paste0("\\b(",paste("Democratic Republic of the Congo", collapse="|"),")\\b"), "

colnames(map.dat) <- c("long", "lat", "group", "order", "Entity", "subregion")

map.dat <- full_join(dat, map.dat, by = "Entity")
try <- filter(map.dat, is.na(lat))

```

## Overall Research Question:

Which countries are contributing most to plastic waste on Earth.

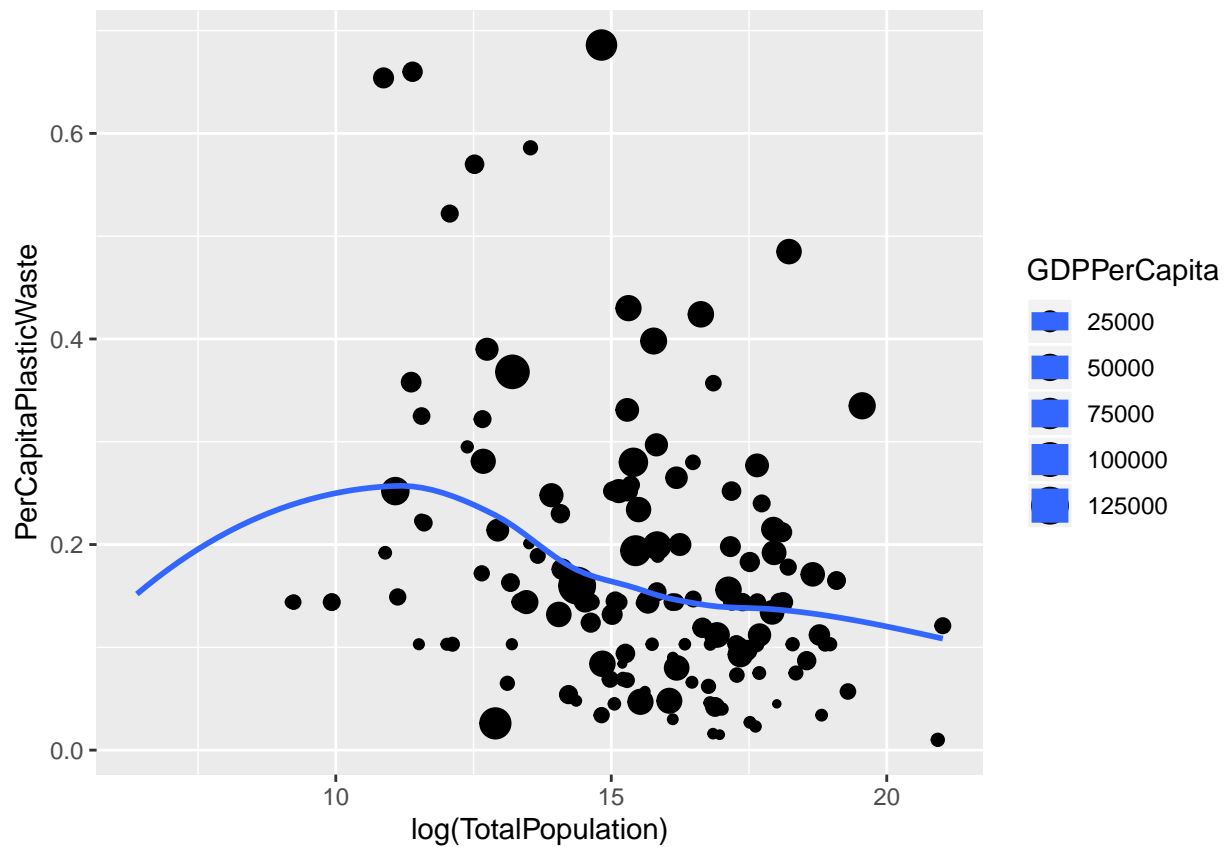
*Map of countries with the highest plastic waste compared to countries with the highest waste per capita*

```

dat2 <- filter(dat, PerCapitaPlasticWaste < 3)
ggplot(dat2, aes(x = log(TotalPopulation), y = PerCapitaPlasticWaste, size = GDPPerCapita)) + geom_point

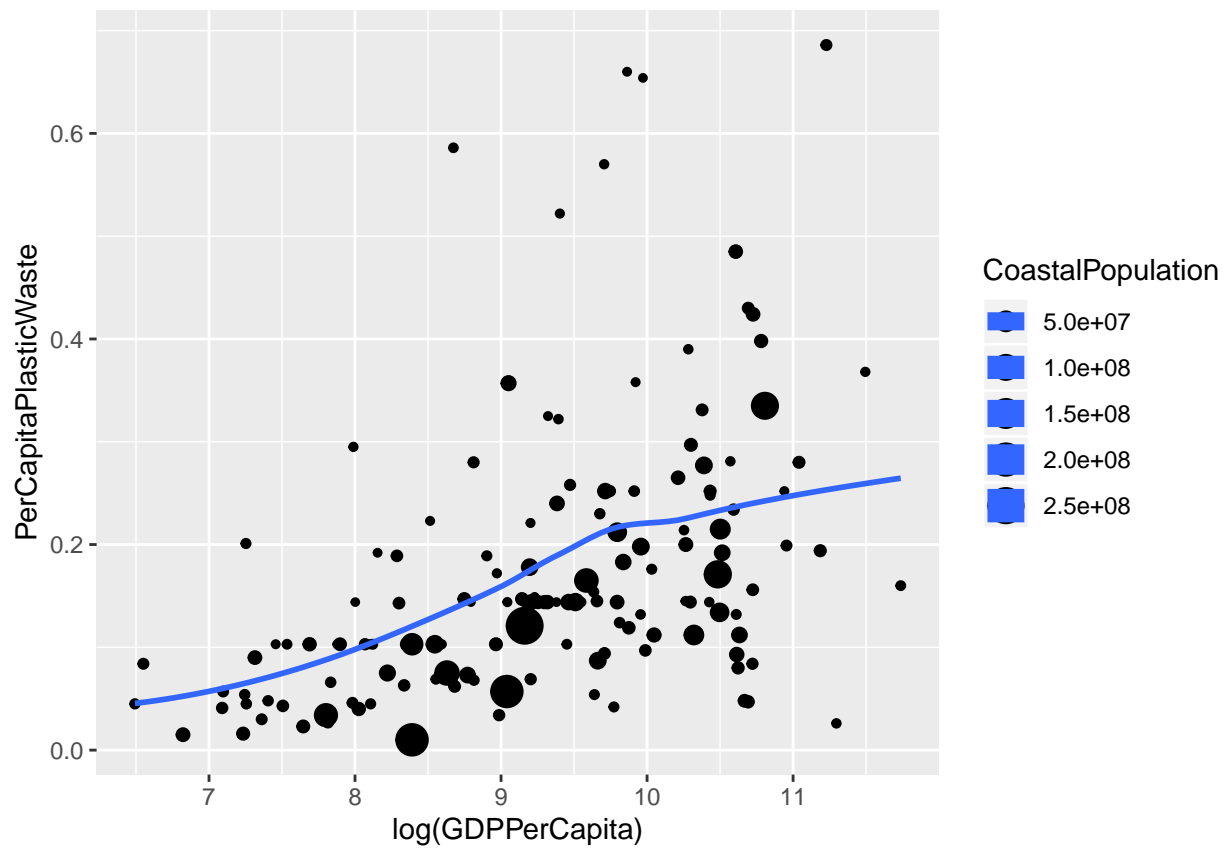
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## Warning: Removed 11 rows containing non-finite values (stat_smooth).
## Warning: Removed 41 rows containing missing values (geom_point).

```



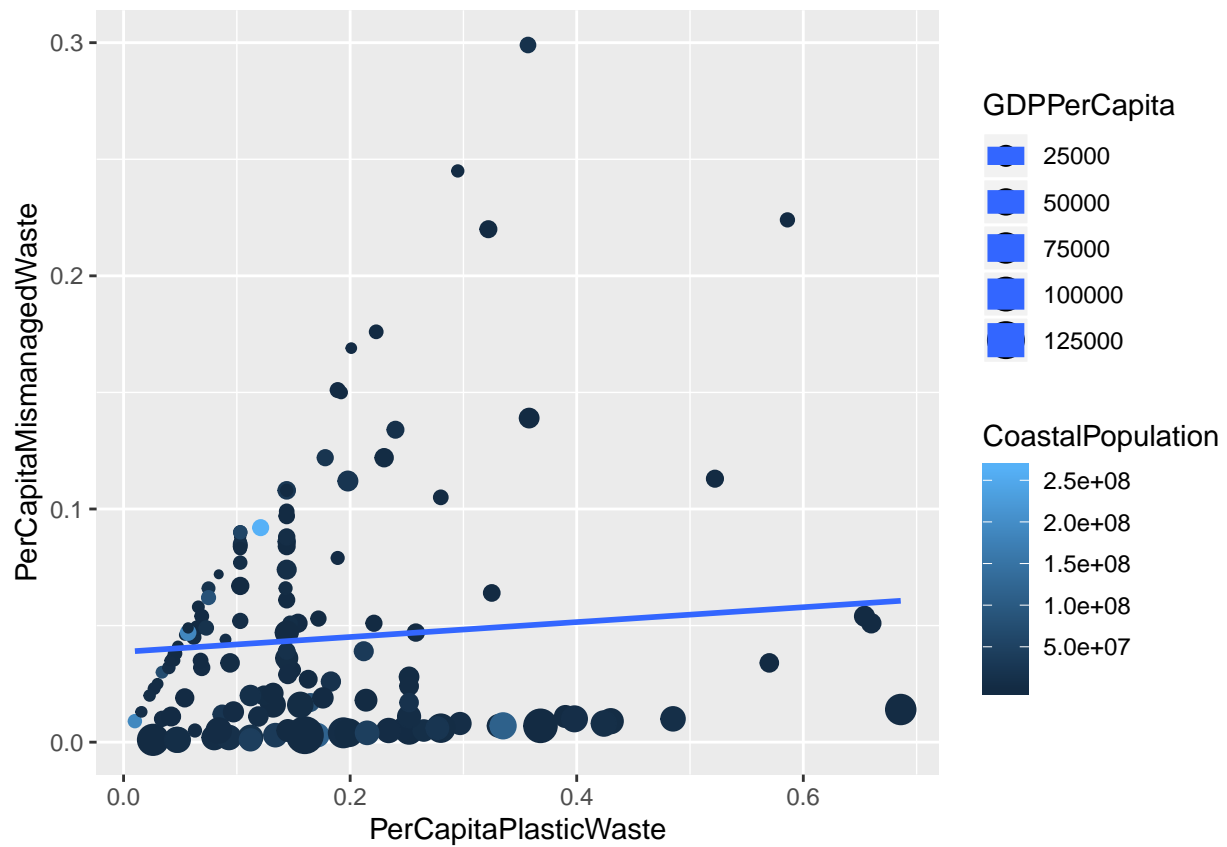
```
ggplot(dat2, aes(x = log(GDPPerCapita), y = PerCapitaPlasticWaste, size = CoastalPopulation)) + geom_point() +
  geom_smooth(method = 'loess', formula = 'y ~ x')

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## Warning: Removed 38 rows containing non-finite values (stat_smooth).
## Warning: Removed 38 rows containing missing values (geom_point).
```



```
ggplot(dat2, aes(x = PerCapitaPlasticWaste, y = PerCapitaMismanagedWaste, size = GDPPerCapita, color = 0
```

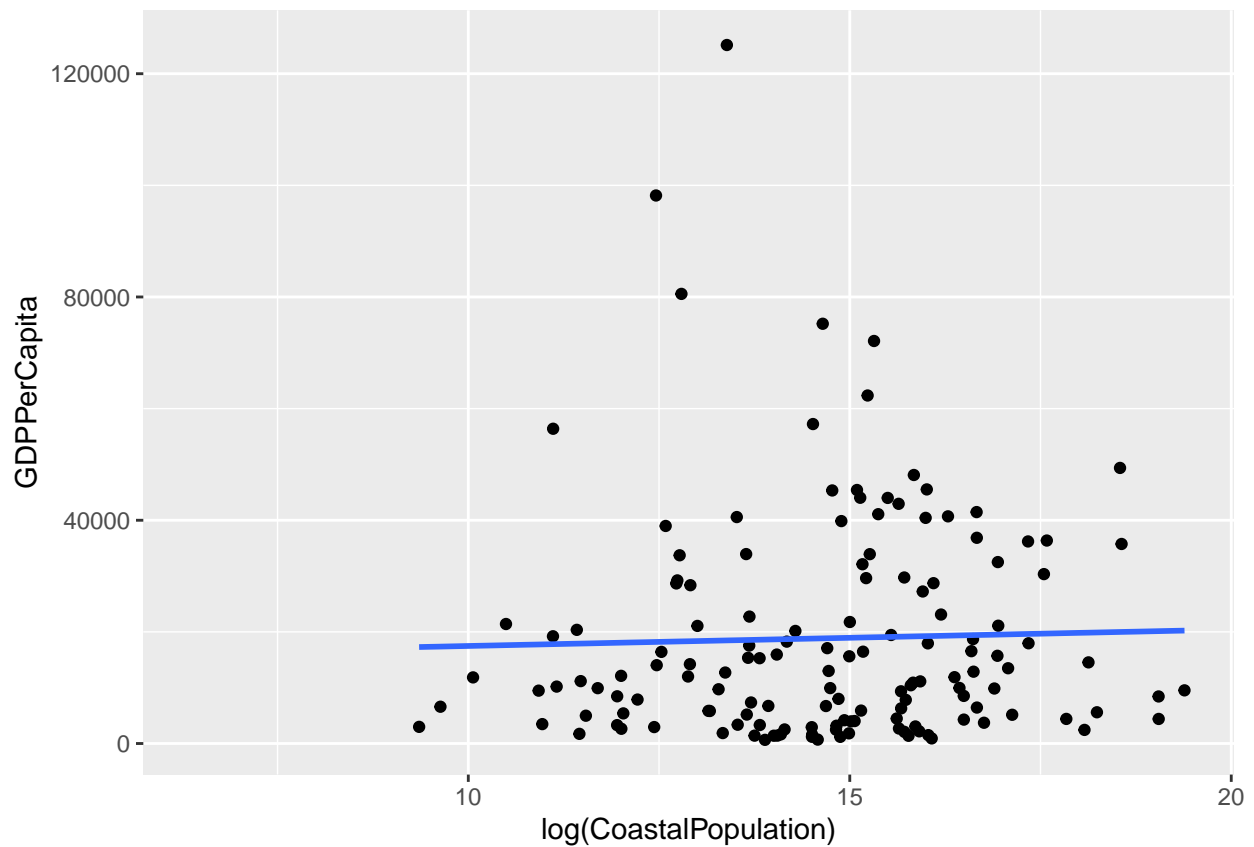
```
## Warning: Removed 38 rows containing missing values (geom_point).
```



```
ggplot(dat2, aes(x = log(CoastalPopulation), y = GDPPerCapita)) + geom_point() + geom_smooth(method = "lm")
```

```
## Warning: Removed 38 rows containing non-finite values (stat_smooth).
```

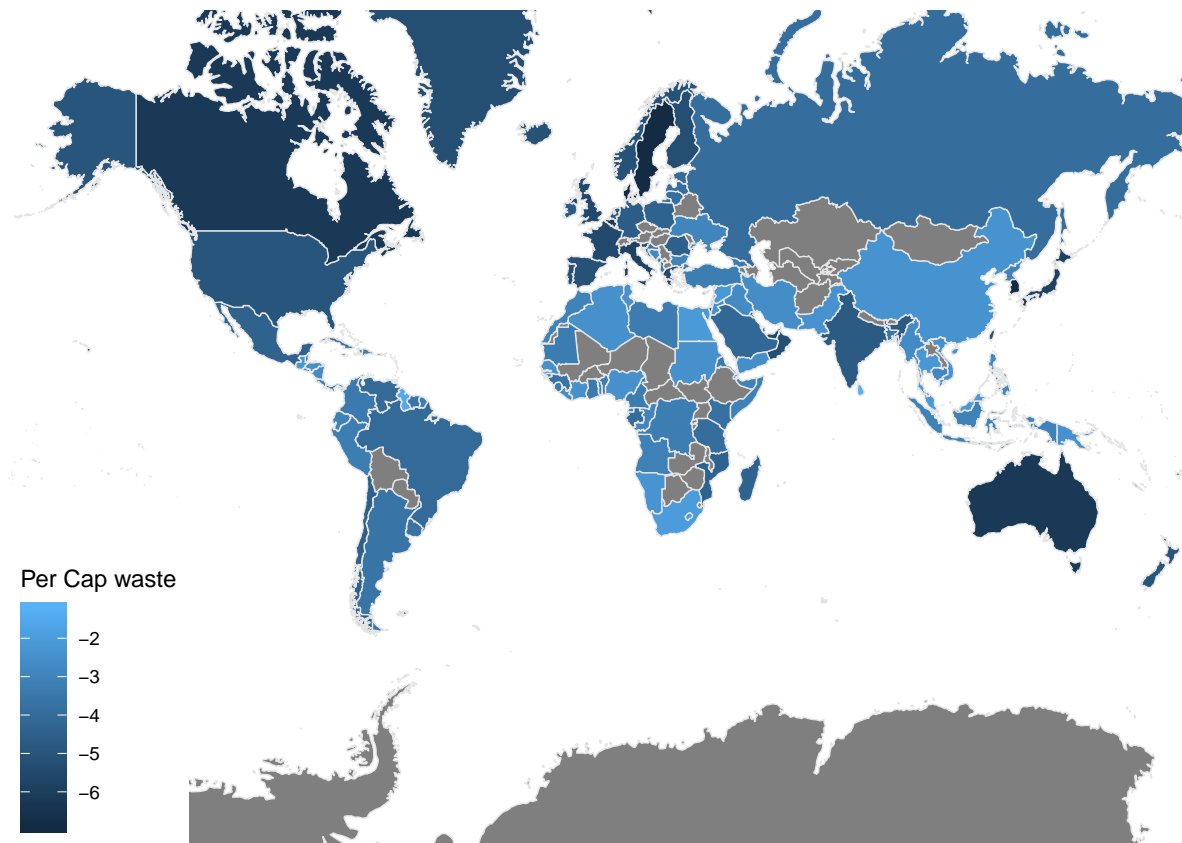
```
## Warning: Removed 38 rows containing missing values (geom_point).
```



```
who <- filter(dat2, MismanagedWaste > 7500000)

ggplot(map.dat, aes(long, lat, group = group, fill = log(PerCapitaMismanagedWaste))) + geom_polygon(col = "red", fill = "red")

## Coordinate system already present. Adding new coordinate system, which will replace the existing one
```



### **What predictors are important in determining plastic waste (other than population)**

Based on three(two of which we will probably remove from the final presentation/write-up) of the above models, we do not have evidence that coastal population is a meaningful predictor of per capita plastic waste or GDP per capita