Lab 01 - Plastic waste

Aleksandra Medina October 8, 2020

Load packages and data

```
library(tidyverse)
```

```
plastic_waste <- read_csv("data/plastic-waste.csv")
```

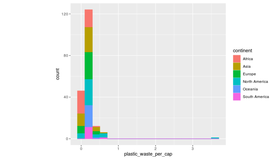
Exercise 1

After plotting a histogram representing each continent's plastic waste, we can make a few observations:

- there is an outlier (Trinidad and Tobago)
- the continent producing the least plastic waste is Africa
- the greatest variety within a continent is in North America
- most continents produce approximately 0.2-0.4 kg per capita of plastic waste
- South America produces surprisingly large amounts of plastic waste

```
# insert code here
ggplot(data = plastic_waste, aes(x = plastic_waste_per_cap, color = continent, fill = continent)) + geom
```

```
## Warning: Removed 51 rows containing non-finite values (stat_bin).
```

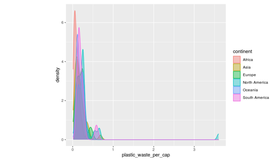


Exercise 2

```
# insert code here

ggplot(data = plastic_waste,
       mapping = aes(x = plastic_waste_per_cap,
                     color = continent,
                     fill = continent)) +
  geom_density(alpha = 0.4)
```

```
## Warning: Removed 51 rows containing non-finite values (stat_density).
```



Exercise 3

A geom is the entire geometric object, thus the alpha (opacity) applies to the entire graph, not just parts of it. Color and fill change for different graphs, while the geom stays untouched as a larger geometric object.

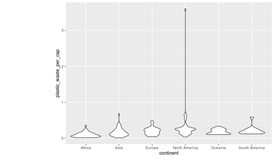
Exercise 4

The violin plot does a better job emphasizing how big of an outlier Trinidad and Tobago is. It also does a better job showing the overall distribution of the countries within each continent and the mode of each continent. This helps us see that the continents that have more of their countries producing higher amounts of plastic per capita are Europe and North America. This could suggest a hypothesis that more developed countries produce more plastic per capita and vice versa.

However, box plots show quartiles and help us identify additional outliers better.

```
# insert code here
ggplot(data = plastic_waste,
       mapping = aes(x = continent,
                     y = plastic_waste_per_cap)) +
  geom_violin()
```

```
## Warning: Removed 51 rows containing non-finite values (stat_ydensity).
```

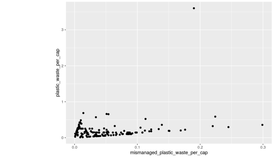


Exercise 5

You can see somewhat of a linear relationship suggesting that the percentage of mismanaged waste stays approximately the same as a percentage of the total plastic waste produced (with outliers, of course). Fortunately, the plot seems to suggest that most countries are at the beginning of the graph meaning that the proportion of mismanaged plastic waste is quite low in the majority of countries.

```
# insert code here
ggplot(data = plastic_waste,
       mapping = aes(x = mismanaged_plastic_waste_per_cap,
                     y = plastic_waste_per_cap)) +
  geom_point()
```

```
## Warning: Removed 53 rows containing missing values (geom_point).
```

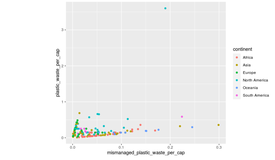


Exercise 6

There is a pattern that emerges here. Europe (while producing comparatively a lot of plastic waste), is the best at managing it and has the lowest proportion of mismanaged plastic waste. Africa seems to display the strongest linear correlation of the amount of plastic waste produced and the amount of plastic waste mismanaged. A couple of countries in Asia, Oceania and South America score exceptionally high on the plastic waste mismanaged scale.

```
# insert code here
ggplot(data = plastic_waste,
       mapping = aes(x = mismanaged_plastic_waste_per_cap,
                     y = plastic_waste_per_cap,
                     color = continent,
                     fill = continent)) +
  geom_point()
```

```
## Warning: Removed 51 rows containing missing values (geom_point).
```



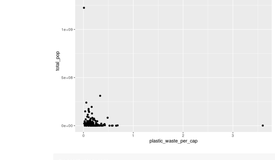
Exercise 7

I think, in order to better see a correlation, I should get rid of the three outliers - China and India for their population and Trinidad and Tobago for its plastic waste production. However, there still seems to be a stronger correlation between plastic waste per capita and total coastal population, compared to plastic waste per capita and total population.

```
# insert code here

ggplot(data = plastic_waste,
       mapping = aes(x = plastic_waste_per_cap,
                     y = total_pop)) +
  geom_point()
```

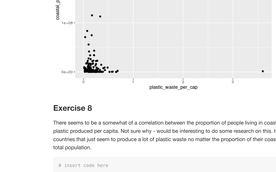
```
## Warning: Removed 61 rows containing missing values (geom_point).
```




```
# insert code here

ggplot(data= plastic_waste,
       mapping = aes(x = plastic_waste_per_cap,
                     y = coastal_pop)) +
  geom_point()
```

```
## Warning: Removed 51 rows containing missing values (geom_point).
```



Exercise 8

There seems to be a somewhat of a correlation between the proportion of people living in coastal areas and the amount of plastic produced per capita. Not sure why - would be interesting to do some research on this. However, there are some countries that just seem to produce a lot of plastic waste no matter the proportion of their coastal population compared to total population.

```
# insert code here

new_plastic_waste <- plastic_waste %>% filter (plastic_waste_per_cap < 3)

new_plastic_waste <- transform(new_plastic_waste,
                               coastal_pop_prop = coastal_pop/ total_pop
                               )

# scale_color_viridis
ggplot(data = new_plastic_waste,
       mapping = aes(x = coastal_pop_prop,
                     y = plastic_waste_per_cap,
                     ))+
  geom_point()+
  geom_smooth()+

  labs(title = "Plastic Waste vs Costal Population Proportion", subtitle = "by Continent" , x = "Coastal
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 10 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 10 rows containing missing values (geom_point).
```