

# Assessment 08 - Basic Plots

*Gabriele Mineo - Harvard Data Science Professional*

## Scatterplots

We made a plot of total murders versus population and noted a strong relationship: not surprisingly states with larger populations had more murders. You can run the code in the console to get the plot.

```
library(dslabs)
data(murders)

population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total

plot(population_in_millions, total_gun_murders)
```

Note that many states have populations below 5 million and are bunched up in the plot. We may gain further insights from making this plot in the log scale.

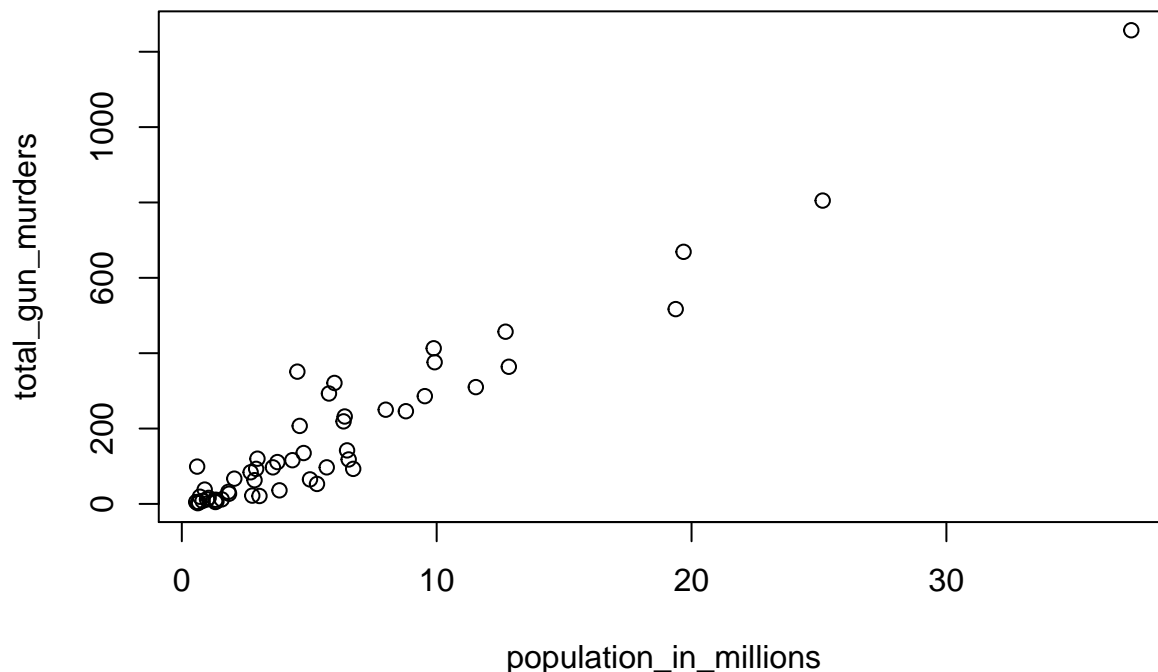
### \*Instructions

Transform the variables using the log, to the base 10, transformation Plot the log transformed total murders versus population

```
# Load the datasets and define some variables
library(dslabs)
data(murders)

population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total

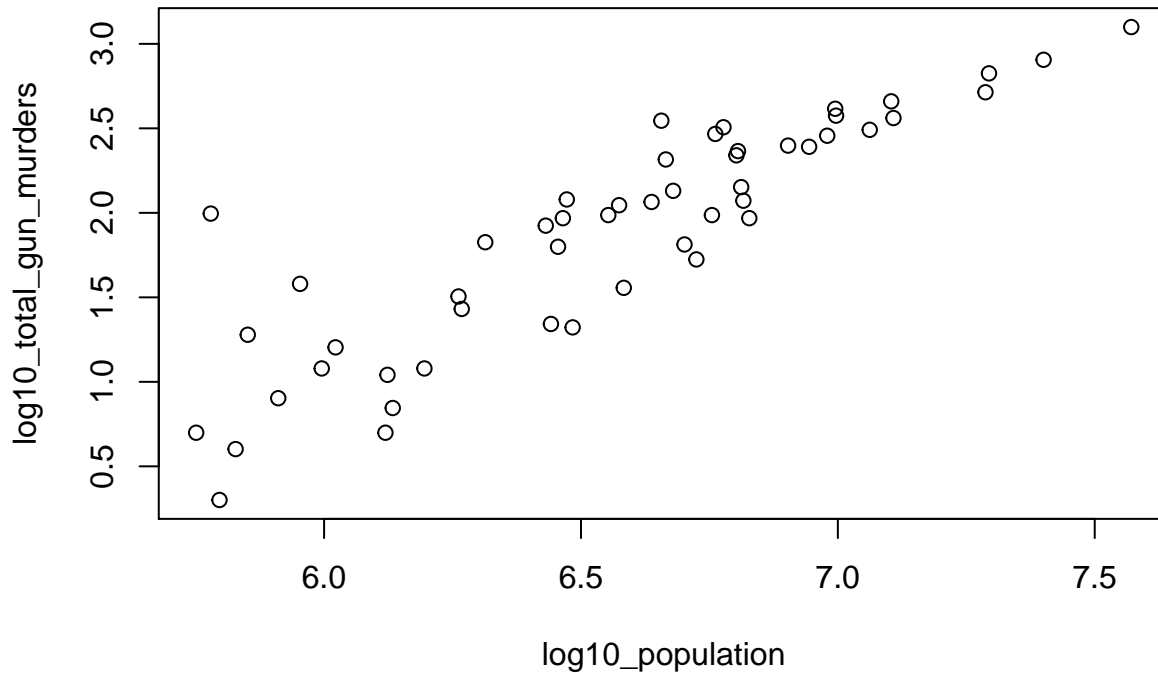
plot(population_in_millions, total_gun_murders)
```



```

# Transform population using the log10 transformation and save to object log10_population
log10_population <- log10(murders$population)
# Transform total gun murders using log10 transformation and save to object log10_total_gun_murders
log10_total_gun_murders <- log10(total_gun_murders)
# Create a scatterplot with the log scale transformed population and murders
plot(log10_population, log10_total_gun_murders)

```



## Histograms

Now we are going to make a histogram.

### Instructions

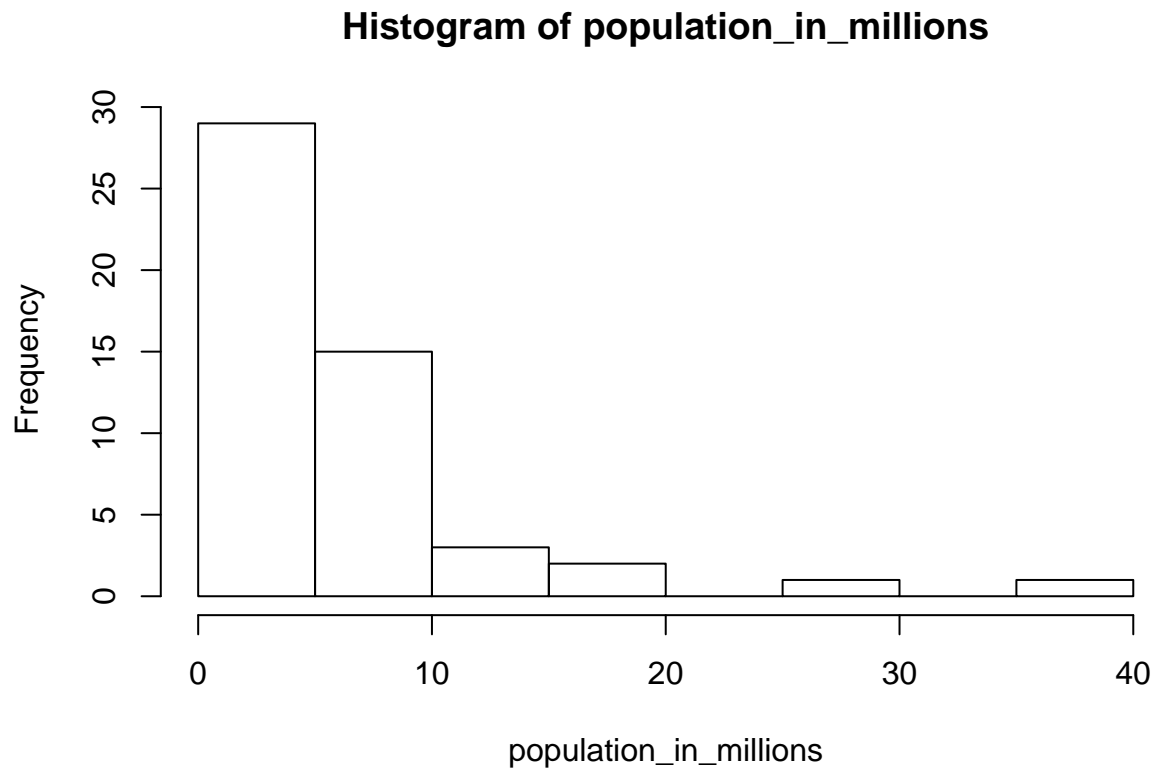
- Compute the population in millions and save it to the object `population_in_millions`
- Create a histogram of the state populations using the function `hist`

```

# Store the population in millions and save to population_in_millions
population_in_millions <- murders$population/10^6

# Create a histogram of this variable
hist(population_in_millions)

```



## Boxplots

Now we are going to make boxplots. Boxplots are useful when we want a summary of several variables or several strata of the same variables. Making too many histograms can become too cumbersome.

### Instructions

In one line of code:

- Stratify the state populations by region.
- Generate boxplots for the strata.

Note that you can achieve this using this `population~region` inside `boxplot` to generate the strata and specify the dataset with the `data` argument.

```
# Create a boxplot of state populations by region for the murders dataset  
boxplot(population~region,data=murders)
```

