SOC 301 - Social Statistics Name: _____
Quiz 3 - Individual October 24, 2016

---

Complete this quiz alone and no notes. Circle the letter of your choice and **write the letter of your choice** to the left of the question. Read carefully over the options and look over your quiz after you are done. ***You'll lose points for not following directions.***

1. If you want to plot two variables, where the explanatory is categorical and the response is numeric, which types of graphs would be appropriate to use?

   A. Faceted barplot & faceted histogram
   B. Faceted histogram & line graph
   C. Boxplot & faceted barplot
   D. Boxplot & faceted histogram

2. What code will make a plot showing the number of `Strongly Autocratic` and `Mildly Autocratic` countries in Africa over years?

   A.
   ```
   gap %>% filter(region == "Africa") %>%
     filter(dem_rank == "Strongly Autocratic" | "Mildly Autocratic") %>%
     ggplot(mapping = aes(x = year)) + geom_bar()
   ```
   B.
   ```
   gap %>% filter(region == "Africa") %>%
     filter(dem_rank == "Strongly Autocratic" | dem_rank == "Mildly Autocratic") %>%
     ggplot(mapping = aes(x = year)) + geom_bar()
   ```
   C.
   ```
   gap %>% filter(region == Africa) %>%
     filter(dem_rank == "Strongly Autocratic" | dem_rank == "Mildly Autocratic") %>%
     ggplot(mapping = aes(x = year)) + geom_bar()
   ```
   D.
   ```
   gap %>% filter(region == Africa) %>%
     filter(dem_rank == "Strongly Autocratic", "Mildly Autocratic") %>%
     ggplot(mapping = aes(x = year)) + geom_bar()
   ```

3. What is the difference between a faceted histogram and a faceted barplot?

   A. A faceted histogram is used for two categorical variables and a faceted barplot is best for one categorical and one numeric variable.
   B. A faceted barplot is used for two categorical variables and a faceted histogram is used for a numeric explanatory variable and a categorical response variable.
   C. A faceted barplot shows two categorical variables and a faceted histogram shows two numeric variables.
   D. A faceted histogram is for plotting a categorical explanatory variable and a numeric response variable while a faceted barplot is for plotting a categorical predictor and categorical response variable.

4. Using the `gap` data frame from Exam 1, find the mean `gdpPercap` of each `region` for each `year`.

   A. `gap %>% group_by(year) %>% summarize(mean_perCap = mean(gdpPercap))`
   B. `gap %>% summarize(region, year, mean_perCap = mean(gdpPercap))`
   C. `gap %>% group_by(region, year) %>% summarize(mean_perCap = mean(gdpPercap))`
   D. `gap %>% group_by(region) %>% summarize(mean_perCap = mean(gdpPercap))`

5. To produce a histogram with 4 bins, fill color green, and border color blue, what would the ending of the R chunk look like?

   A. `geom_histogram(bins = 4, color = blue, fill = green)`
   B. `geom_histogram(bins = 4, color = blue, fill = "green")`
   C. `geom_histogram(bins = 4, color = "blue", fill = "green")`
   D. `geom_histogram(bins = 4, aes(color = blue, fill = green))`

**More on back →**

SOC 301 - Social Statistics Name: _____
Quiz 3 - Individual October 24, 2016

---

6. Is the following data set in **tidy** format following Hadley Wickham's guidelines?

| | country | X1952 | X1957 | X1962 |
|---|---|---|---|---|
| 1 | Albania | -9 | -9 | -9 |
| 2 | Argentina | -9 | -1 | -1 |
| 3 | Armenia | -9 | -7 | -7 |
| 4 | Australia | 10 | 10 | 10 |

   A. Yes, it is neat and easy-to-read.
   B. Yes, it follows all of the requirements of a **tidy** data set as given in the textbook.
   C. No, it does not meet all of the requirements of a **tidy** data set.
   D. No, the values inside the data frame cannot be negative.

7. "`Error: could not function 'ggplot'`" appears after trying to knit an R Markdown document with the following code: `ggplot(data = alaska_flights, aes(x = origin) + geom_bar()`. You see a checkmark next to `ggplot2` in the **Packages** tab. What needs to be done to fix this error?
   A. You need to put `ggplot2` rather than `ggplot` in your code.
   B. `ggplot` is a function in `ggplot2`, so you need to make sure `ggplot2` is loaded via `library(ggplot2)` in a chunk above this chunk.
   C. The `g` in `ggplot` needs to be capitalized: (`Ggplot`)
   D. None of the above.

8. "`Error: object 'weather' not found`" appears after trying to knit an R Markdown document with the following code: `weather %>% filter(month == 1)`. What needs to be done to fix this error?
   A. You need to put `library(dplyr)` in a chunk above this code since the `weather` data frame is in the `dplyr` package.
   B. You need to put `library(nycflights13)` in a chunk above this code since the `weather` data frame is in the `nycflights13` package.
   C. Both A and B.
   D. None of the above.

9. What code is needed to choose only the data in `gap` for year 1952?
   A. `gap %>% filter(year = 1952)`
   B. `gap %>% select(year == 1952)`
   C. `gap + filter(year == 1952)`
   D. None of the above

10. Which of the following is correct about the five main verbs?
   A. `select` - Pick rows based on conditions about their values
   B. `arrange` - Sort the rows based on one or more variables
   C. `mutate` - Create summary measures of variables (or groups of observations on variables using `group_by`)
   D. `filter` - Chooses different variables from the data frame.

**REMINDER:** Circle the letter of your choice and **write the letter of your choice** to the left of the question. Read carefully over the options and look over your quiz after you are done.