

# Stat 140 - R Quiz – Sample

What's Your Name? \_\_\_\_\_

Which section are you in? \_\_\_\_\_

Please answer the questions below. If you wish, you may use a copy of the “Data Visualization with ggplot2” cheat sheet and a copy of the “Data Transformation with dplyr” cheat sheet.

1. Below are the first few rows of a data frame named `HELPrct`, which contains records for people enrolled in a substance abuse treatment program. Assuming the data set has already been loaded into R, what command could you use to produce this output?

```
##   age homeless substance mental_score physical_score
## 1  37   housed   cocaine    25.111990      58.41369
## 2  37 homeless  alcohol    26.670307      36.03694
## 3  26   housed   heroin     6.762923      74.80633
## 4  39   housed   heroin    43.967880      61.93168
## 5  32 homeless  cocaine    21.675755      37.34558
## 6  47   housed   cocaine    55.508991      46.47521
```

SOLUTION:

```
head(HELPrct)
```

2. Below is a more detailed look at the structure of the data frame. What command would you use to produce this output?

```
## 'data.frame':   453 obs. of  5 variables:
##  $ age          : int  37 37 26 39 32 47 49 28 50 39 ...
##  $ homeless     : Factor w/ 2 levels "homeless","housed": 2 1 2 2 1 2 2 1 1 1 ...
##  $ substance    : Factor w/ 3 levels "alcohol","cocaine",...: 2 1 3 3 2 2 2 1 1 3 ...
##  $ mental_score : num  25.11 26.67 6.76 43.97 21.68 ...
##  $ physical_score: num  58.4 36 74.8 61.9 37.3 ...
```

SOLUTION:

```
str(HELPrct)
```

3. What R command would you use to calculate the mean age of the study participants? We've discussed a couple of ways of doing this; pick your favorite.

SOLUTION:

```
mean(HELPrct$age)
```

```
summarize(HELPrct, mean_age = mean(age))
```

4. What R command would you use to calculate the standard deviation of the study participants' ages? We've discussed a couple of ways of doing this; pick your favorite.

SOLUTION:

```
sd(HELPrct$age)
```

```
summarize(HELPrct, sd_age = sd(age))
```

5. I want to calculate an overall measure of physical and mental health, as the sum of the `mental_score` and the `physical_score` for each study participant. I'd like to store this as a new variable in the data frame, named `combined_score`. What R command would you use to do this?

SOLUTION:

```
HELPrct <- mutate(HELPrct, combined_score = mental_score + physical_score)
```

There will be one question like 6a. or 6b: sort in ascending or descending order

6a. I want to sort the `HELPrct` in ascending order according to the `physical_score` variable, and store the sorted results in a new data frame called `HELPrct_sorted`. What R command would you use to do this?

SOLUTION:

```
HELPrct_sorted <- arrange(HELPrct, physical_score)
```

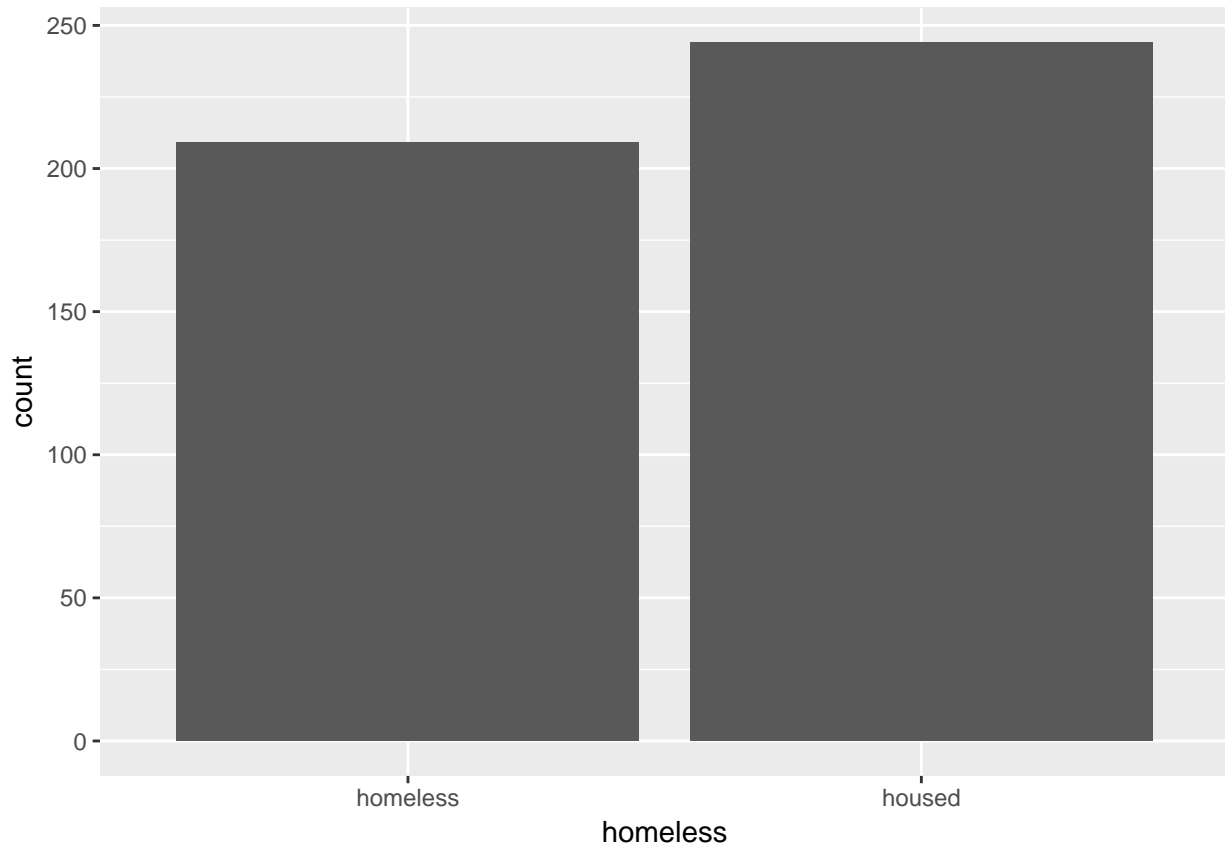
6b. I want to sort the `HELPrct` in descending order according to the `physical_score` variable, and store the sorted results in a new data frame called `HELPrct_sorted`. What R command would you use to do this?

SOLUTION:

```
HELPrct_sorted <- arrange(HELPrct, desc(physical_score))
```

There will be one question like 7a. through 7c.: write down the code to recreate a plot that is given. This plot will involve only one variable, so it will be either a bar plot for a categorical variable, or a histogram or density plot for a quantitative variable.

7a. What code would you use to make this plot?



Recall that the general template for making a ggplot is as below. Please fill in the blanks for the geometry type, the aesthetic mapping, and the data argument.

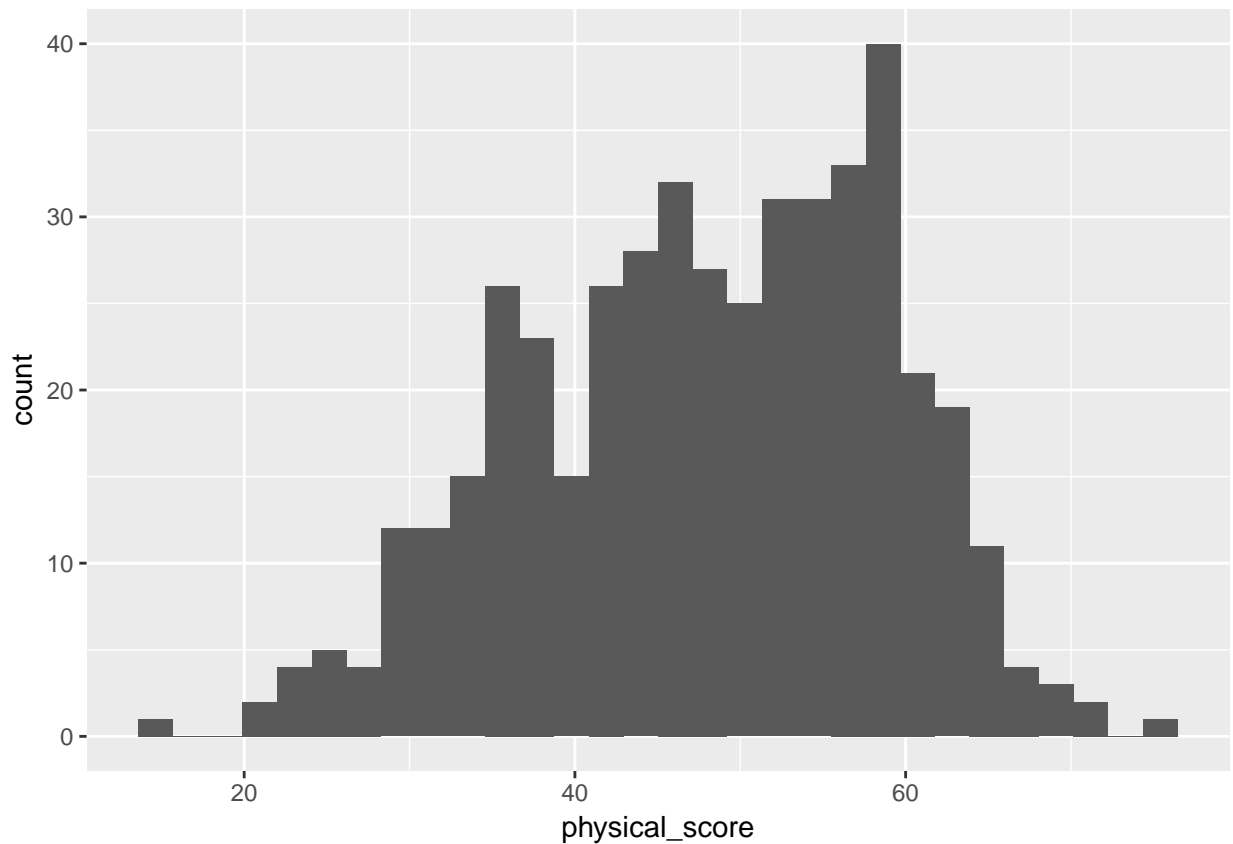
SOLUTION:

```
ggplot() +  
  geom_           (mapping = aes(           ),  
  data =         )
```

SOLUTION: (you would hand write in “bar”, “x = homeless”, and “HELPrct” in the blanks above)

```
ggplot() +  
  geom_bar(mapping = aes(x = homeless),  
  data = HELPrct)
```

7b. What code would you use to make this plot?



Recall that the general template for making a ggplot is as below. Please fill in the blanks for the geometry type, the aesthetic mapping, and the data argument.

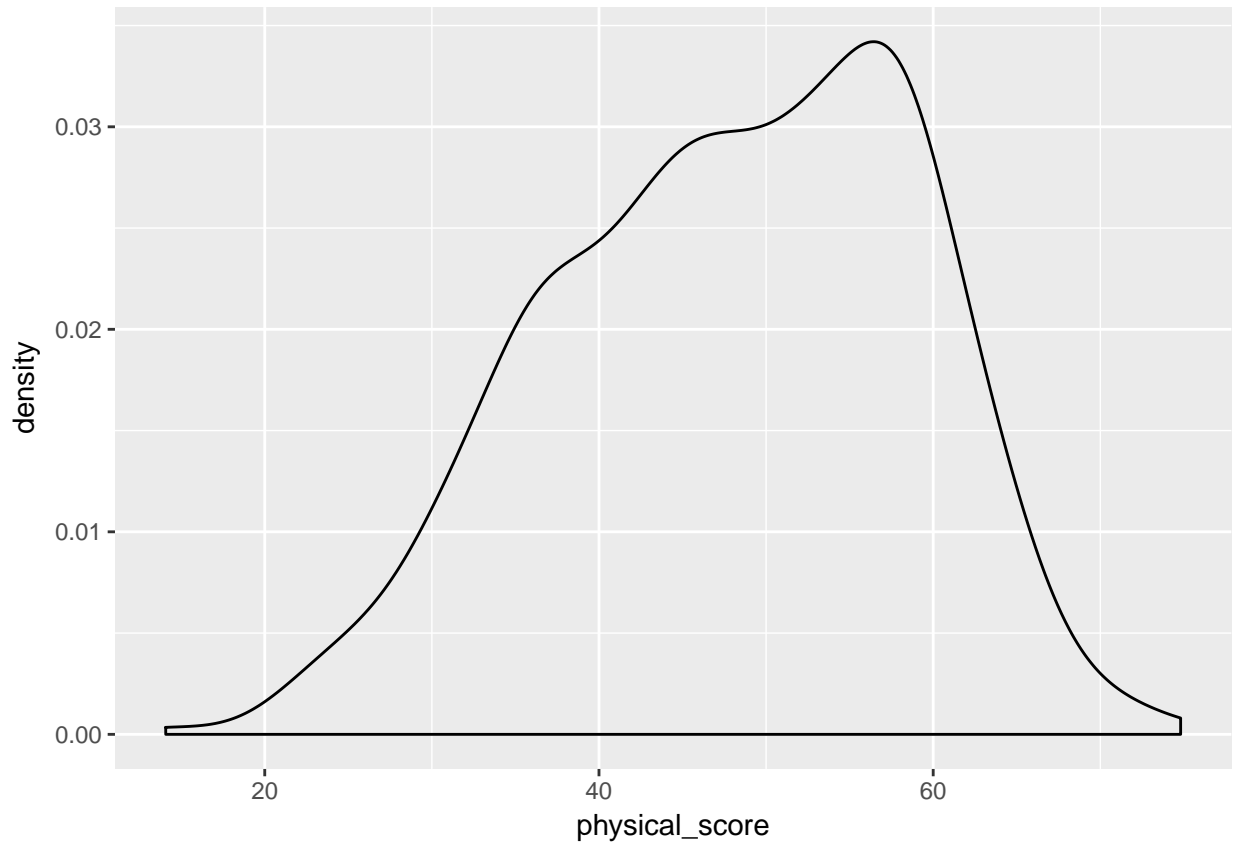
SOLUTION:

```
ggplot() +  
  geom_          (mapping = aes(  
    data =          )  
),
```

SOLUTION: (you would hand write in “histogram”, “x = physical\_score”, and “HELPrct” in the blanks above)

```
ggplot() +  
  geom_histogram(mapping = aes(x = physical_score),  
    data = HELPrct)
```

7c. What code would you use to make this plot?



Recall that the general template for making a ggplot is as below. Please fill in the blanks for the geometry type, the aesthetic mapping, and the data argument.

SOLUTION:

```
ggplot() +  
  geom_           (mapping = aes(  
    data =         )
```

SOLUTION: (you would hand write in “density”, “x = physical\_score”, and “HELPrct” in the blanks above)

```
ggplot() +  
  geom_density(mapping = aes(x = physical_score),  
    data = HELPrct)
```

8. For each of the following pairs of variables, circle the type of plot you would make, and write down the type of geometry you would use to make that plot. (more than one answer may be correct – if so, choose one)

8a. age and physical\_score

SOLUTION:

Bar Plot

Box Plot

Density Plot with groups shown in different colors

Scatter Plot

Geometry type: \_\_\_\_\_

SOLUTION: (you would circle the plot type above and write in the corresponding geometry)

Scatter Plot, geom\_point

8b. homeless and substance

SOLUTION:

Bar Plot

Box Plot

Density Plot with groups shown in different colors

Scatter Plot

Geometry type: \_\_\_\_\_

SOLUTION: (you would circle the plot type above and write in the corresponding geometry)

Bar Plot, geom\_bar

8c. homeless and physical\_score

SOLUTION:

Bar Plot

Box Plot

Density Plot with groups shown in different colors

Scatter Plot

Geometry type: \_\_\_\_\_

SOLUTION: (you would circle the plot type above and write in the corresponding geometry)

You can pick ONE of the following:

Box Plot, geom\_boxplot

OR

Density plot with groups shown in different colors, geom\_density

**8d. mental\_score and physical\_score**

SOLUTION:

Bar Plot

Box Plot

Density Plot with groups shown in different colors

Scatter Plot

Geometry type: \_\_\_\_\_

SOLUTION: (you would circle the plot type above and write in the corresponding geometry)

Scatter Plot, geom\_point