DATA 101 Exam 1

Kamila Palys

Due: Monday 10/26 at 11:59pm

Academic Honesty Statement (fill in your name)

I, Kamila Palys, hereby affirm that I have not communicated with or gained information in any way from my classmates or anyone other than the Professor during this exam, that I have not assisted anyone else with this exam, and that all work is my own.

Load packages and data

Questions

Question 1

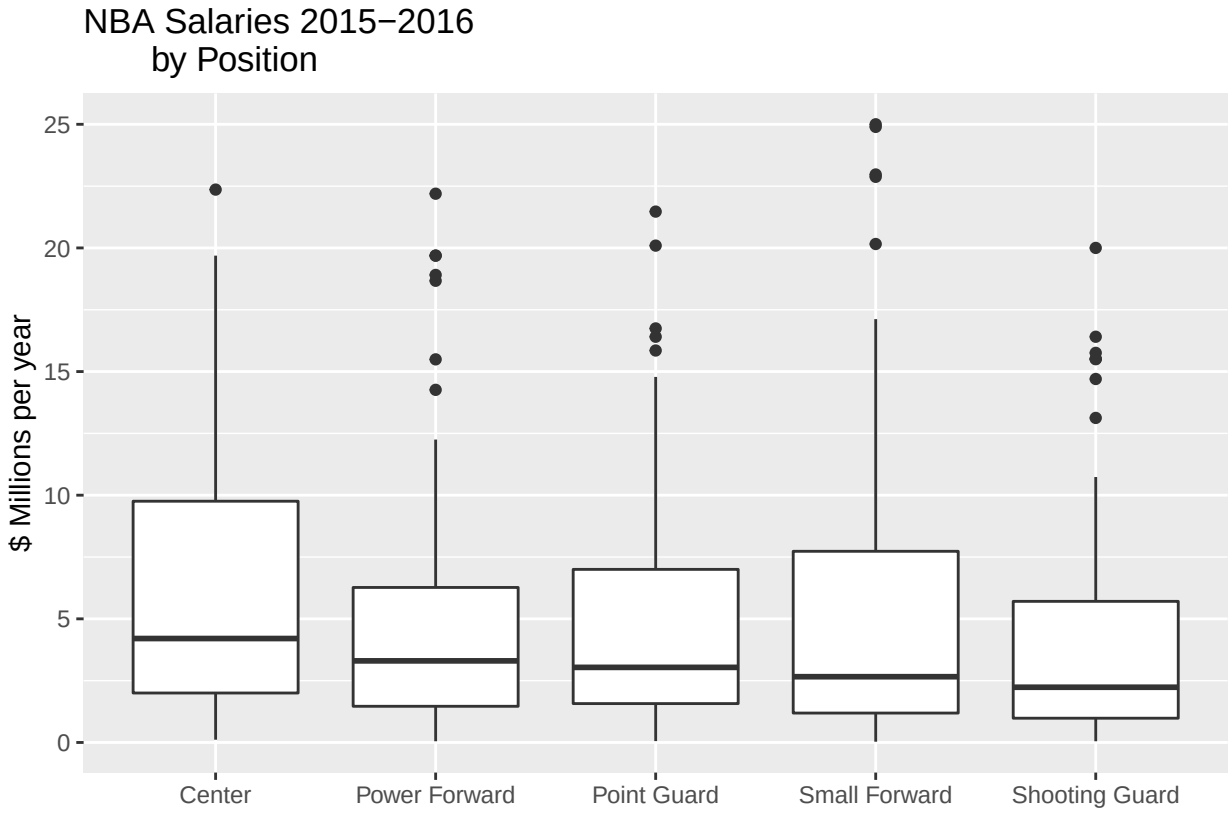
First, we will make a table to view the salaries of the NBA players in descending order by using the "select" and "arrange" functions to display the variables we want, in the order we want.

```
## # A tibble: 417 x 2
##    player          salary
##    <chr>            <dbl>
##  1 Kobe Bryant       25
##  2 Joe Johnson       24.9
##  3 LeBron James      23.0
##  4 Carmelo Anthony   22.9
##  5 Dwight Howard     22.4
##  6 Chris Bosh        22.2
##  7 Chris Paul        21.5
##  8 Kevin Durant      20.2
##  9 Derrick Rose      20.1
## 10 Dwyane Wade       20
## # ... with 407 more rows
```

From the table it is visible that Kobe Bryant has the highest salary of all the NBA players at $25 million.

Question 2

Now we will create a boxplot that compares the distribution of the players' salaries by position.

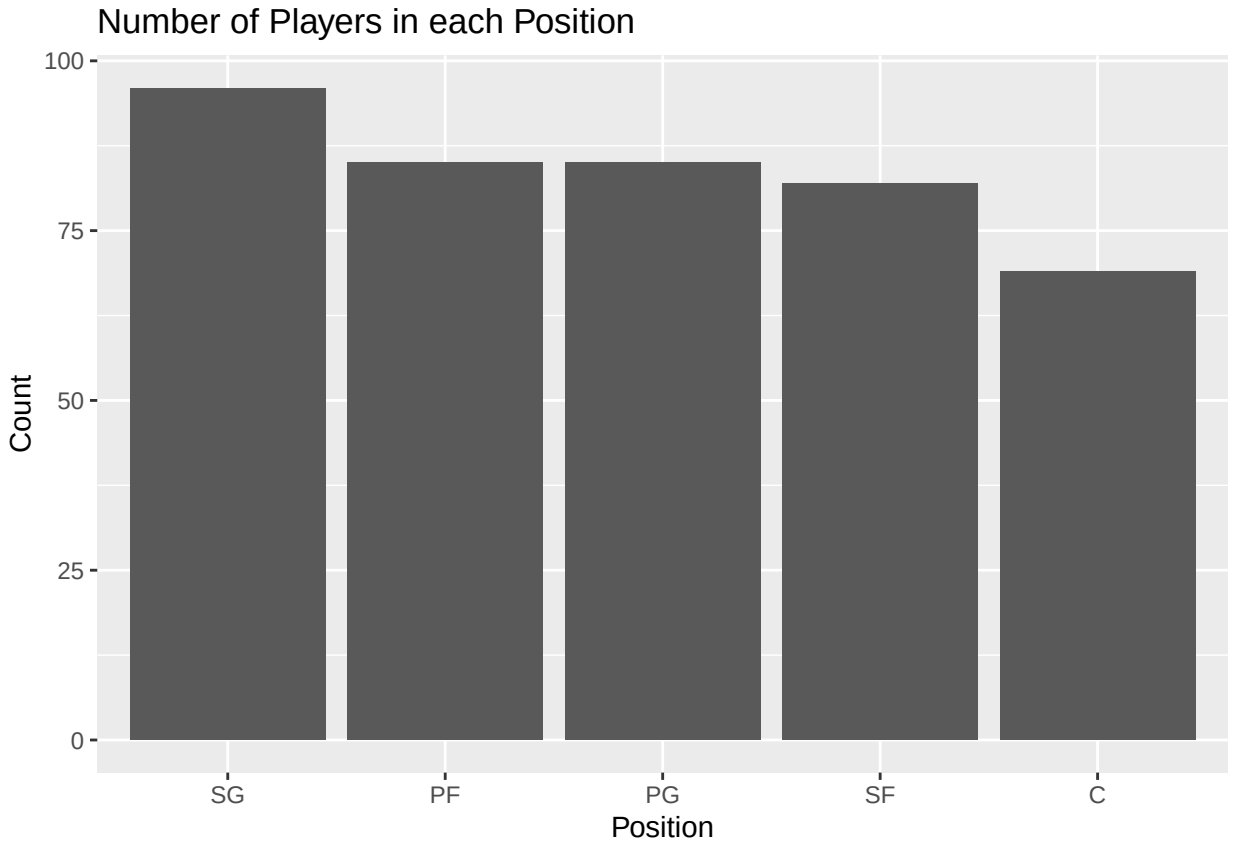


From this graph, we are able to see that the highest median salary is earned by players with the center position, while the lowest median comes from the shooting guard players. The ones who play center, however, also have a more even spread of their salaries, seeing as their interquartile range, or the middle 50%, is the largest of all the positions. There are also very little outliers in the center players' salaries, and there are several high outliers in all four of the other positions, so that is not to say that players playing a position other than center cannot earn a high salary.

Question 3

Now we will take a look at how many players there are in each position with "group_by" and the "count" function.

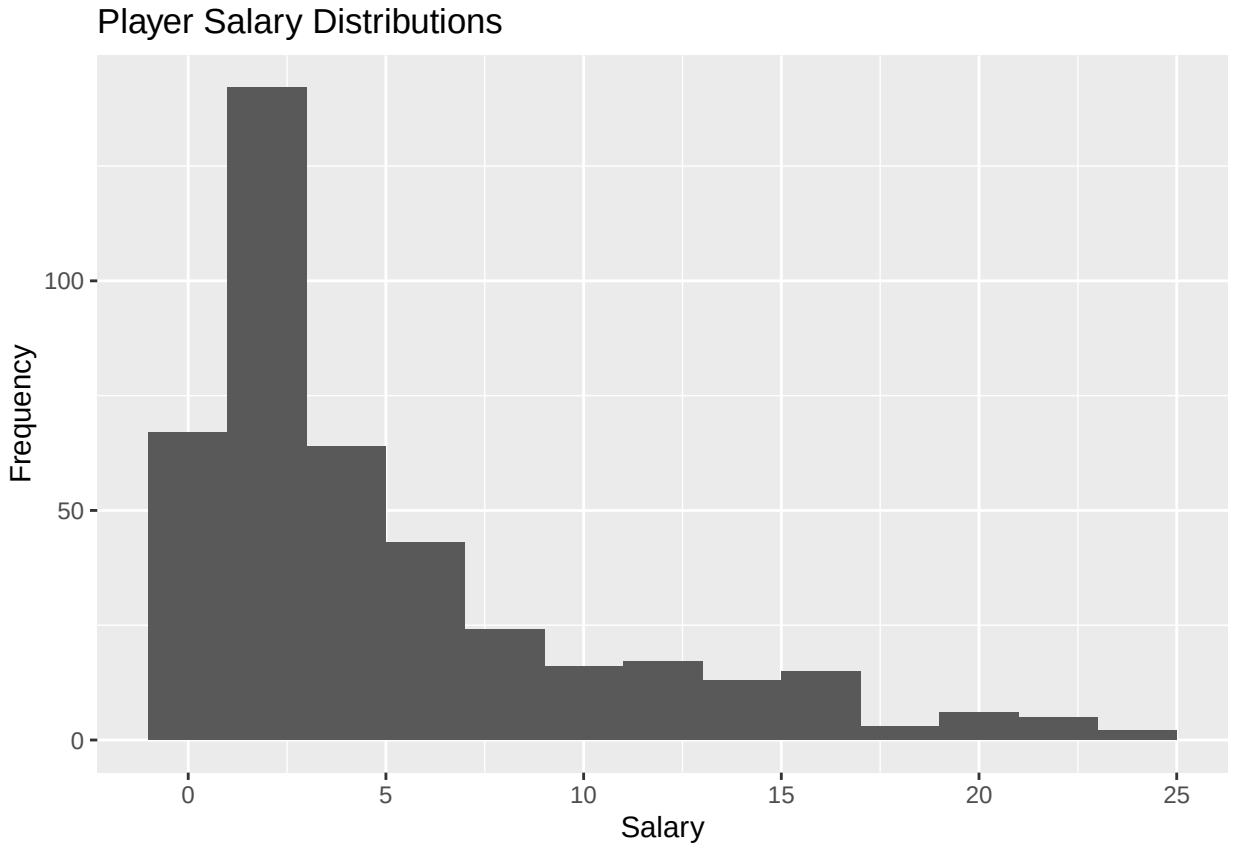
```
## # A tibble: 5 x 2
## # Groups:   position [5]
##   position     n
##   <chr>    <int>
## 1 SG          96
## 2 SF          82
## 3 PG          85
## 4 PF          85
## 5 C           69
```



With this table and bar graph, it is visible that there are the most players in shooting guard at 96, while the center position has only 69, which is the least amount of players playing a certain position. Additionally, there are 82 playing small forward, 85 playing point guard, and 85 playing power forward.

Question 4

Here a display is created in the form of a histogram to show the distribution of the salaries of the NBA players.



It is clear from this graph that a large number of players earned a salary between about $1 million to $3 million. It appears as though a large amount of players received a salary of up to $10 million, but a minority of them receive more than that and very few reach $20 million or more.

Question 5

Here, the average salaries per player in each team will be displayed.

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## Selecting by avg_salary
```

```
## # A tibble: 10 x 2
##    team                  avg_salary
##    <chr>                      <dbl>
##  1 Cleveland Cavaliers        10.2
##  2 Houston Rockets             7.11
##  3 Miami Heat                  6.79
##  4 Golden State Warriors       6.72
##  5 Chicago Bulls               6.57
##  6 San Antonio Spurs           6.51
##  7 Los Angeles Lakers          6.24
##  8 Sacramento Kings            6.22
##  9 Oklahoma City Thunder       6.05
## 10 Dallas Mavericks            5.98
```

As seen, the three top earning teams are the Cleveland Caveliers earning $10.2 million on average, Houston Rockets earning $7.11 million on average, and Miami heat with $6.79 million on average. The second and

third highest earning teams do not have an extreme difference in average salaries per player, but the Cleveland Caveliers have a high jump, with each player earning almost $3 million more on average than the second highest earning team. It is possible that this may be due to the Cleveland Caveliers having more famous basketball players that earn a salary that is considered to be an outlier.

Question 6

[Enter code and narrative here.]

Question 7

[Enter code and narrative here.]

Question 8

[Enter code and narrative here.]

Question 9

[Enter code and narrative here.]

Question 10

[Enter code and narrative here.]