Take Home Quiz 2 - Solutions

Your Name

Due Wednesday March 26 at 10:00am

This quiz should take you approximately 25 minutes. Place your answers into this markdown document, knit it, and hand in the result as a PDF or Word document. You may use R, any reference material, and information already available on the internet. Do not work together and do not get help, except from Dr. Clair.

Problem 1 (10 points)

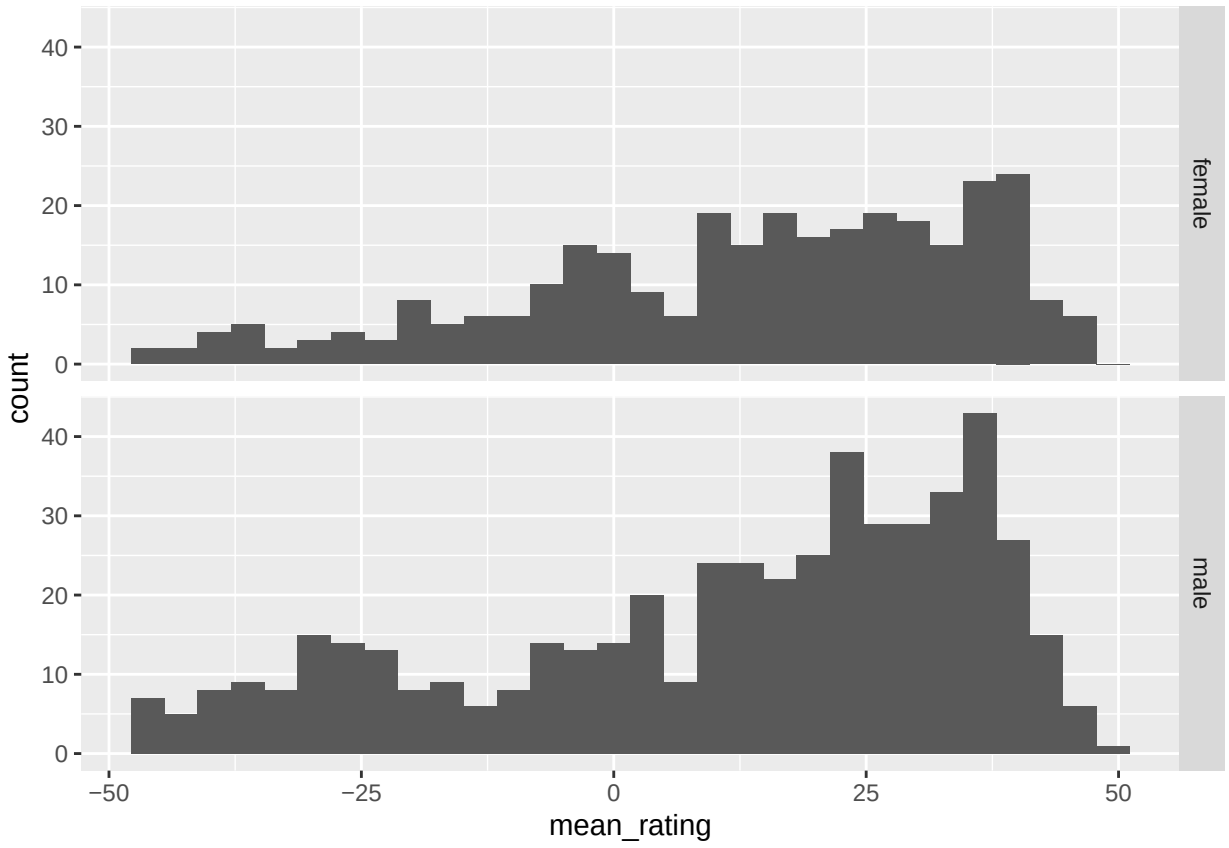
The data `personalities` is available on our course webpage at https://turtlegraphics.org/stat2300/data/personalities.csv

This has personality ratings for hundreds of fictional characters, across various spectrums. This problem only uses the BAP84 spectrum, which measures cruel vs. kind.

Make a histogram with facets for male and female that compares the personality ratings of male vs. female characters on the BAP84 spectrum.

Solution

```
personalities <- read.csv("https://turtlegraphics.org/stat2300/data/personalities.csv")
bap84 <- personalities |> filter(spectrum == 'BAP84')
bap84 |> ggplot(aes(x=mean_rating)) + geom_histogram() + facet_grid(gender~.)
```



Problem 2 (10 points)

Continue using the `personalities` data. Use a Wilcoxon rank sum test to check for a difference in personality ratings on the BAP84 spectrum for male vs. female characters.

Solution There is not a significant difference in BAP84 rating (p=0.4577)

```
wilcox.test(mean_rating ~ gender, data=bap84)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  mean_rating by gender
## W = 77495, p-value = 0.4577
## alternative hypothesis: true location shift is not equal to 0
```

Problem 3 (10 points)

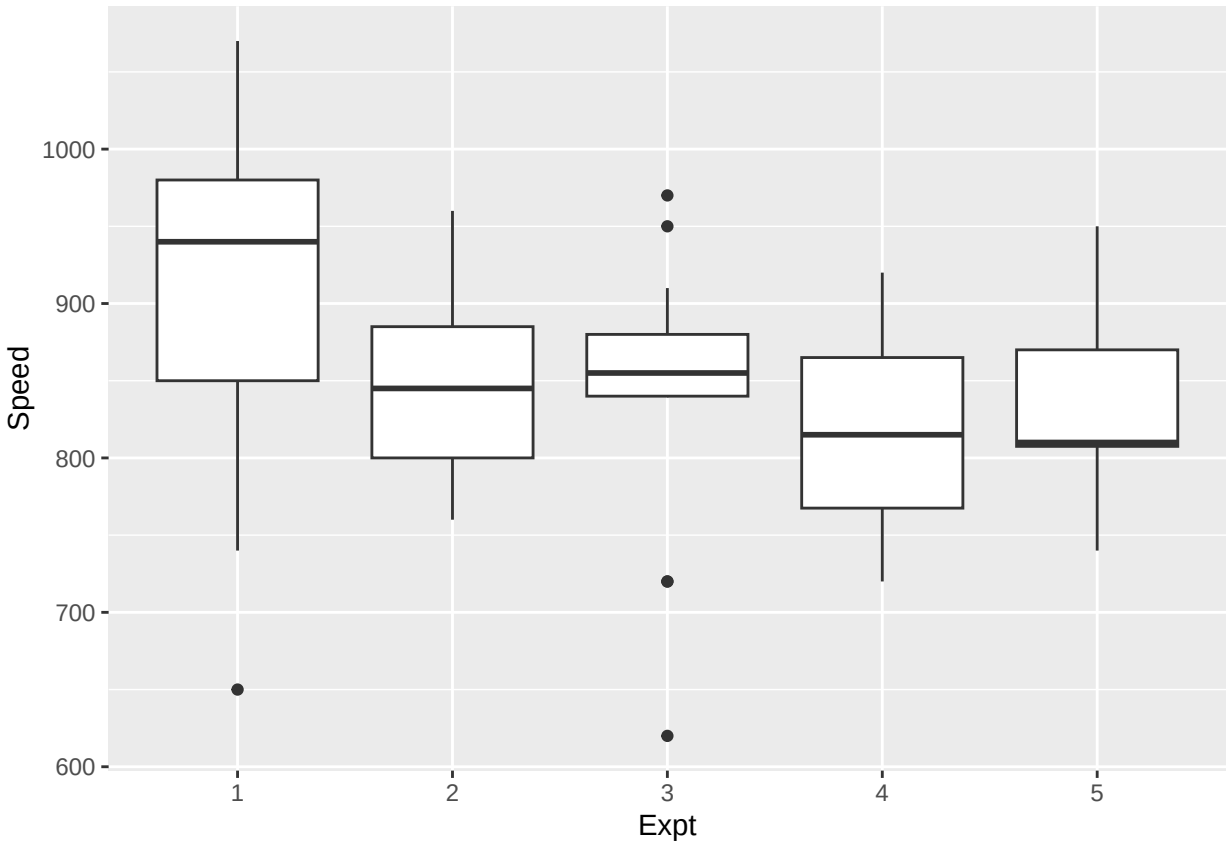
The built-in data set `morley` contains five experiments `Expt` measuring the speed of light. Each experiment ran 20 times.

a. Convert the `Expt` variable to a factor, you'll need this for the next few problems.

b. Use ggplot to boxplot the distribution of `Speed`, with one box per experiment.

Solution

```
morleyf <- morley |> mutate(Expt = factor(Expt))
morleyf |> ggplot(aes(x=Expt, y=Speed)) + geom_boxplot()
```



Problem 4 (10 points)

Use an analysis of variance to test for a difference in the mean speed of light for all five experiments. What do you conclude?

Solution

```
lm(Speed ~ Expt, data=morleyf) |> anova()
```

```
## Analysis of Variance Table
##
## Response: Speed
##           Df Sum Sq Mean Sq F value   Pr(>F)
## Expt       4  94514 23628.5  4.2878 0.003114 **
## Residuals 95 523510  5510.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Problem 5 (10 points)

Continue using the speed of light data.

Which one experiment differs significantly from the others? Remove it, then check with ANOVA that the other four have no difference in means.

Solution

```
pairwise.t.test(morleyf$Speed, morleyf$Expt)
```

```
##
```

```
##  Pairwise comparisons using t tests with pooled SD
##
## data:  morleyf$Speed and morleyf$Expt
##
##   1      2      3      4
## 2 0.1838 -      -      -
## 3 0.0610 1.0000 -      -
## 4 0.0028 0.8027 1.0000 -
## 5 0.0122 1.0000 1.0000 1.0000
##
## P value adjustment method: holm
```

```
morleyf %>% filter(Expt != 1) %>% lm(Speed ~ Expt, data=.) |> anova()
```

```
## Analysis of Variance Table
##
## Response: Speed
##           Df Sum Sq Mean Sq F value Pr(>F)
## Expt       3  14425  4808.3  1.1626 0.3296
## Residuals 76 314330  4135.9
```