Numbered Figures begin here:

```
library(tidyverse)
library(readxl)
```

Figure 1: Packages

Group	Leniency
neutral	6
smile	3.5
smile	4.5
smile	6
smile	4
neutral	2.5
smile	7.5
smile	2.5
smile	3.5
neutral	4
neutral	2.5
neutral	4.5
smile	3.5
smile	9
neutral	3
smile	3
smile	5
neutral	4.5
smile	5.5
smile	5

Figure 2: Smiles leniency data

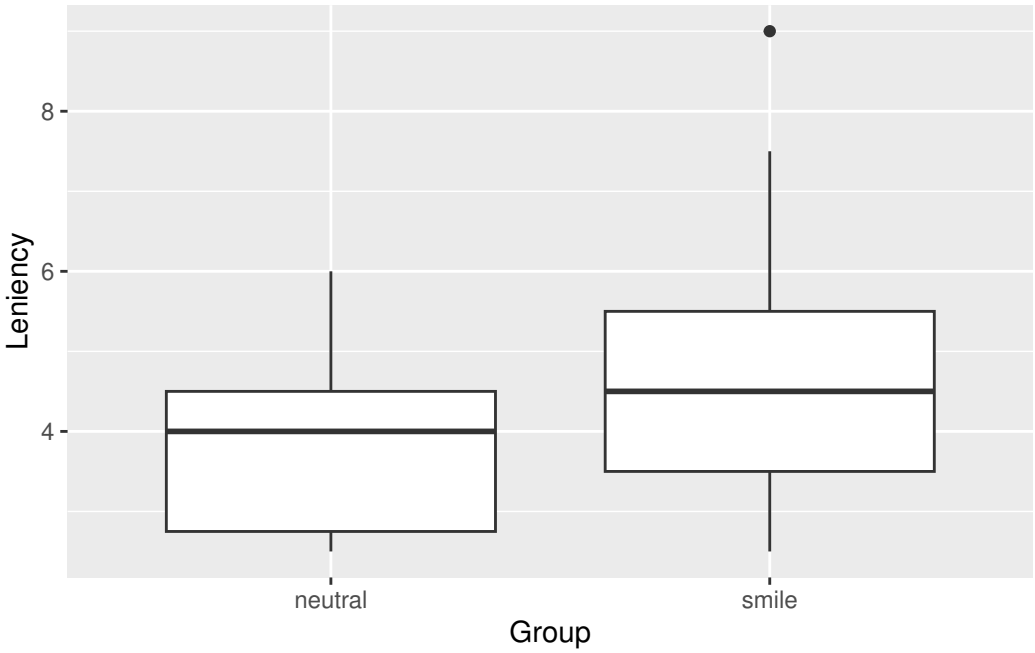

Figure 3: Smiles leniency plot

```
# A tibble: 30 x 3
   age_group fact_correct opinion_correct
   <chr>            <dbl>           <dbl>
 1 18-49                3               5
 2 18-49                5               5
 3 18-49                5               5
 4 50+                  4               1
 5 18-49                2               4
 6 50+                  5               5
 7 18-49                5               5
 8 50+                  4               2
 9 18-49                2               5
10 50+                  4               3
11 50+                  2               5
12 18-49                3               5
13 50+                  1               4
14 18-49                3               3
15 50+                  3               3
16 50+                  3               2
17 18-49                5               5
18 50+                  3               3
19 50+                  2               5
20 18-49                5               5
21 50+                  5               1
22 18-49                2               5
23 50+                  4               3
24 18-49                3               1
25 50+                  5               5
26 18-49                1               5
27 50+                  3               5
28 50+                  4               3
29 50+                  1               4
30 18-49                5               5
```

Figure 4: Fact and opinion survey data (30 randomly chosen rows)

```
fact_opinion %>% count(age_group) -> counted
ggplot(counted, aes(x = age_group)) + geom_bar()
```

Figure 5: Some code

```
ggplot(fact_opinion, aes(x = fact_correct, fill = age_group)) +
  geom_bar(position = "dodge")
```

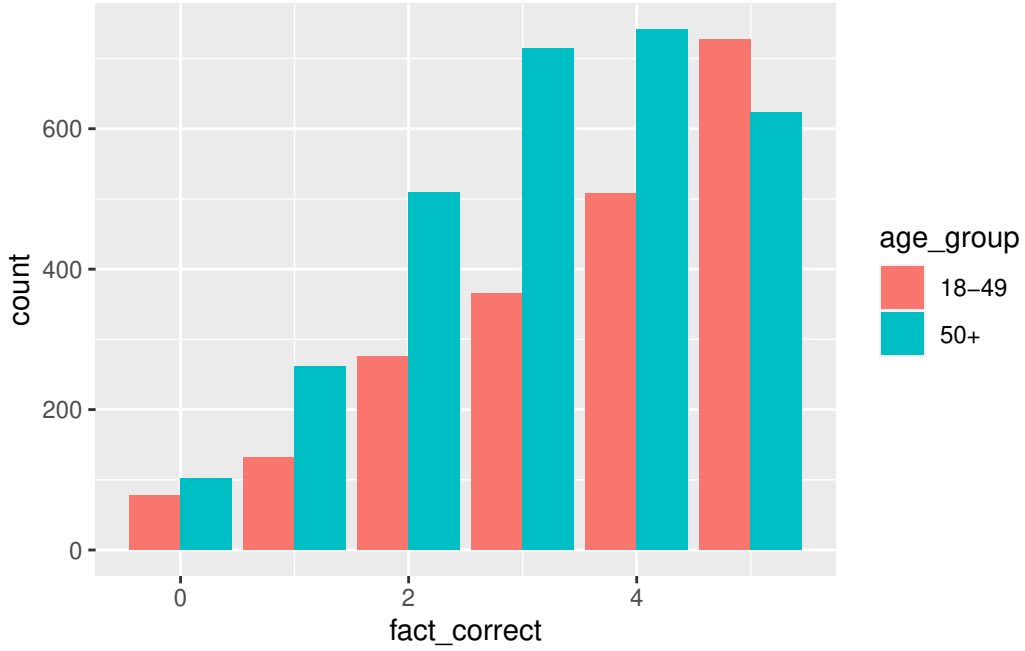

Figure 6: Fact and opinion survey plot

```
# A tibble: 30 x 6
   Region        Status Sex    Cause                     Rate    SE
   <chr>         <chr>  <chr>  <chr>                    <dbl> <dbl>
 1 HHS Region 08 Urban  Male   Unintentional injuries    55.3   0.7
 2 HHS Region 10 Urban  Male   Cancer                   191.    1.1
 3 HHS Region 10 Urban  Female Cerebrovascular diseases  35.2   0.4
 4 HHS Region 06 Urban  Male   Alzheimers                20.7   0.3
 5 HHS Region 10 Urban  Male   Unintentional injuries    49.8   0.6
 6 HHS Region 03 Rural  Female Cancer                   157.    1.4
 7 HHS Region 10 Rural  Male   Cerebrovascular diseases  37.1   1
 8 HHS Region 02 Rural  Male   Flu and pneumonia         19.6   0.9
 9 HHS Region 09 Rural  Male   Heart disease            206.    2.7
10 HHS Region 02 Urban  Female Flu and pneumonia         14.5   0.2
11 HHS Region 06 Urban  Male   Cancer                   202.    0.7
12 HHS Region 07 Rural  Male   Lower respiratory         65.9   0.9
13 HHS Region 01 Rural  Female Diabetes                  15     0.6
14 HHS Region 07 Urban  Female Diabetes                  16     0.3
15 HHS Region 09 Rural  Male   Unintentional injuries    79.1   1.8
16 HHS Region 04 Rural  Male   Unintentional injuries    79.1   0.7
17 HHS Region 01 Urban  Female Cancer                   140.    0.8
18 HHS Region 08 Urban  Female Nephritis                  8.3   0.3
19 HHS Region 07 Rural  Female Cancer                   150.    1.3
20 HHS Region 07 Rural  Male   Unintentional injuries    68.1   1
21 HHS Region 01 Rural  Male   Lower respiratory         51.7   1.3
22 HHS Region 09 Urban  Female Suicide                    5.2   0.1
23 HHS Region 05 Urban  Female Nephritis                 12.9   0.1
24 HHS Region 06 Rural  Male   Unintentional injuries    77.2   0.9
25 HHS Region 08 Rural  Male   Unintentional injuries    71     1.3
26 HHS Region 08 Urban  Female Alzheimers                30.9   0.5
27 HHS Region 10 Rural  Male   Alzheimers                22.9   0.8
28 HHS Region 06 Urban  Male   Heart disease            220.    0.8
29 HHS Region 05 Rural  Female Unintentional injuries    32.3   0.4
30 HHS Region 03 Rural  Male   Lower respiratory         62.1   1
```

Figure 7: US regional mortality rates data (randomly chosen rows)

```
  my_url <- "http://ritsokiguess.site/datafiles/shrimp.csv"
  shrimp <- read_csv(my_url)
```

```
Rows: 18 Columns: 1
-- Column specification --------------------------------------------------------
Delimiter: ","
dbl (1): percent

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
  shrimp
```

```
# A tibble: 18 x 1
   percent
     <dbl>
 1    32.2
 2    33
 3    30.8
 4    33.8
 5    32.2
 6    33.3
 7    31.7
 8    35.7
 9    32.4
10    31.2
11    26.6
12    30.7
13    32.5
14    30.7
15    31.2
16    30.3
17    32.3
18    31.7
```

Figure 8: Shrimp cocktail data

```r
with(shrimp, t.test(percent, mu = 34, alternative = "less"))
```

```
    One Sample t-test

data:  percent
t = -5.0761, df = 17, p-value = 4.674e-05
alternative hypothesis: true mean is less than 34
95 percent confidence interval:
    -Inf 32.5503
sample estimates:
mean of x
 31.79444
```

Figure 9: Code and output for an analysis on the shrimp data

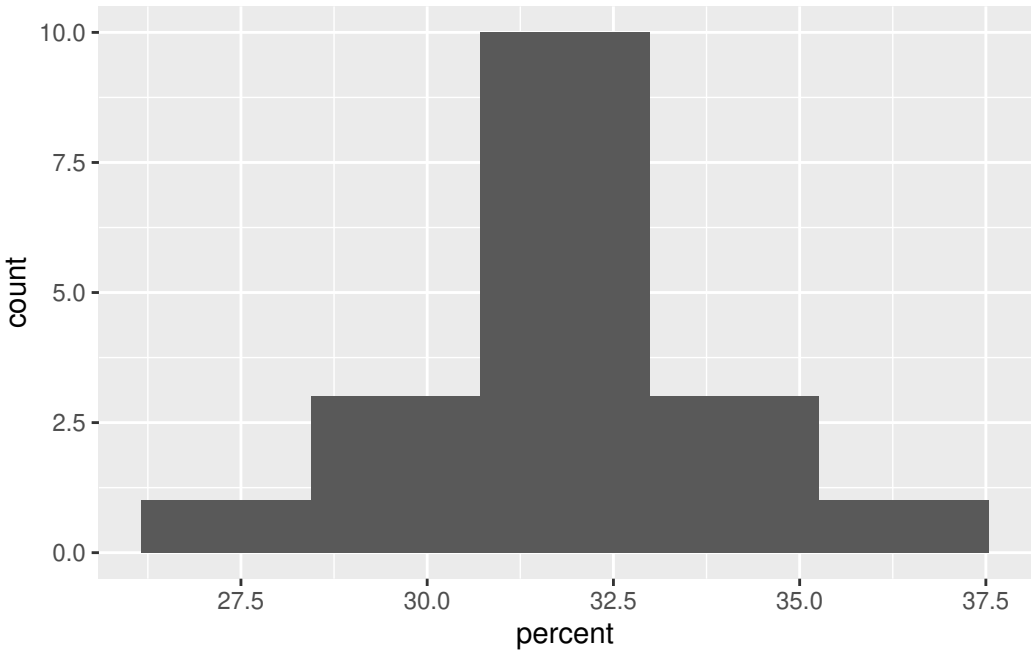

Figure 10: Histogram of shrimp data

```
# A tibble: 19 x 2
   company      emission
   <chr>           <dbl>
 1 manufacturer      2.7
 2 manufacturer      3.1
 3 manufacturer      3.1
 4 manufacturer      2.9
 5 manufacturer      2.5
 6 manufacturer      3.4
 7 manufacturer      3.4
 8 manufacturer      3.4
 9 manufacturer      2.4
10 competitor        3.7
11 competitor        3
12 competitor        3.5
13 competitor        3.8
14 competitor        2.8
15 competitor        3.5
16 competitor        3.4
17 competitor        3.6
18 competitor        2.7
19 competitor        3.7
```

Figure 11: Carbon monoxide data

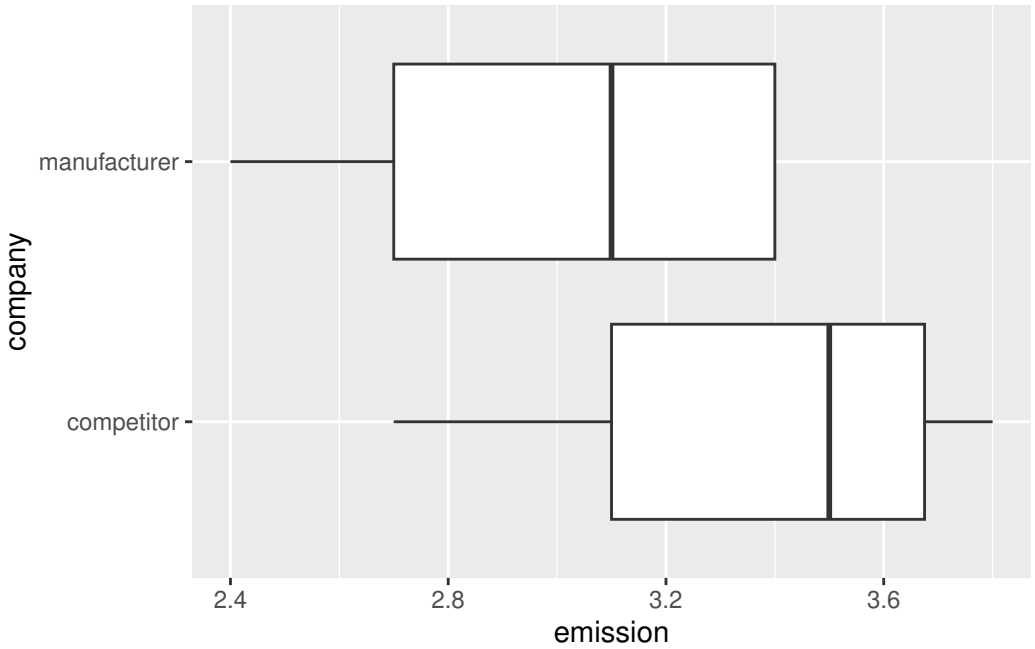

Figure 12: Plot for carbon monoxide data. Note that one of the whiskers for "manufacturer" is very short.

```
    Welch Two Sample t-test

data:  emission by company
t = 2.1187, df = 16.842, p-value = 0.02465
alternative hypothesis: true difference in means between group competitor and group manufacturer is gre
95 percent confidence interval:
 0.06802198        Inf
sample estimates:
  mean in group competitor mean in group manufacturer
                  3.370000                   2.988889
```

Figure 13: Test output for carbon monoxide data