STAT 479: PRACTICE MIDTERM 1

STAT 479

February 21, 2022

- This exam lasts from 1:20 - 2:10pm on February 28, 2022. There are 8 questions.
- This exam is closed note and closed computer.
- You may use a 1-page cheat sheet (8.5 x 11in or A4 size). You may use both sides, but the cheat sheet must be handwritten.
- If you need extra space, you may write on the back of the page. Please indicate somewhere that your answer continues.
- The instructors will only be able to answer clarifying questions during the exam. They will be sitting at the back of the room.

Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Total
Score									
Possible	2	2	3	3	5	3	6	6	30

Q1 [2 points]

Circle all the statements below that are true about small-multiple and compound figures. Assume that the `ggplot2`, `ggridges`, and `patchwork` packages have been loaded.

a. Each panel in a plot using `+ facet_grid(x ~ y)` will show a different subset of the data.

b. The number of vertical panels in a plot using `facet_wrap(~ x)` depends on the number of levels of `x`.

c. If `p1`, `p2`, and `p3` are each `ggplot2` objects, then using `p1 + p2 + p3` will combine those plots horizontally.

d. The ridge heights across `x` in a plot using `+ geom_ridge_density(aes(x, y))` encodes the variable `y`.

Q2 [2 points]

1. On the line, enter the letter that best matches the following definitions. A term may be used more than once or not at all.

   - ___ The use of the same visual encoding to several partitions of the dataset, shown together across adjacent panels.
   - ___ A detail view that appears when one or more items are selected in the main view.
   - ___ The use of different visual encodings to represent the same data, shown together across adjacent panels.
   - ___ A practice where items interactively selected in one overview are highlighted in all other views.

   A. Ridgeline plot B. Multiform display C. Details-on-demand D. Small multiples E. Linked brushing F. Tidy data

Q3 [3 points]

Discuss two common pitfalls in compound figures. What can be done to avoid them?

Q4 [3 points]

Sketch the code needed to transform `medals`,

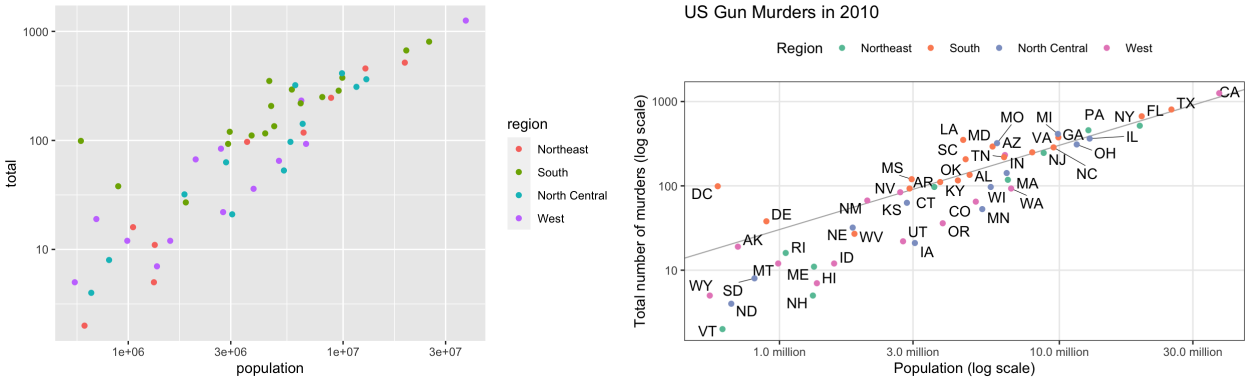
```
##   country gold silver
## 1     USA   10      7
## 2  Canada   20     26
```

into the form given below,

```
## # A tibble: 4 x 3
##   country medal  value
##   <chr>   <chr>  <dbl>
## 1 USA     gold      10
## 2 USA     silver     7
## 3 Canada  gold      20
## 4 Canada  silver    26
```

Q5 [5 points]

The questions below refer to the two figures below.



a. [2 points] Describe three distinct customizations used to transform the plot on the left into the plot on the right.

b. [3 points] Provide the `ggplot2` commands that were used for each transformation. Be as specific as possible.

Q6 [3 points]

This problem asks you to compare and contrast the `geom_ridge_density()` and `geom_rideline()` layers available in the `ggridges` package.

a. [3 points] Describe the differences in the visual encodings provided by these two layers. What kind of data do they assume, and how are they represented graphically?

b. [2 points] For each type of layer, provide one situation where one could be applied, but not the other.

Q7 [6 points]

Provide the reactivity graph associated with the shiny app below, which prints the summary and first few rows of one of three datasets, depending on the user's selection.

```
library(shiny)
ui <- fluidPage(
    textInput("caption", "Caption:", value = "Data Summary"),
    selectInput("dataset", "Dataset:", choices = c("rock", "pressure", "cars")),
    numericInput("obs", "Number of observations:", value = 10),
    verbatimTextOutput("summary"),
    dataTableOutput("view")
)
```

```
server <- function(input, output) {
  datasetInput <- reactive({
    if (input$dataset == "rock") {
      return (rock)
    } else if (input$dataset == "pressure") {
      return (pressure)
    }
    cars
  })

  output$caption <- renderText({ input$caption })
  output$summary <- renderPrint({
    dataset <- datasetInput()
    summary(dataset)
  })

  output$view <- renderDataTable({ head(datasetInput(), n = input$obs)})
}
```

Q8 [6 points]

Consider the dataset below, which contains the average `Asymmetry` across several groups of birds. Note that only the first few rows are shown,

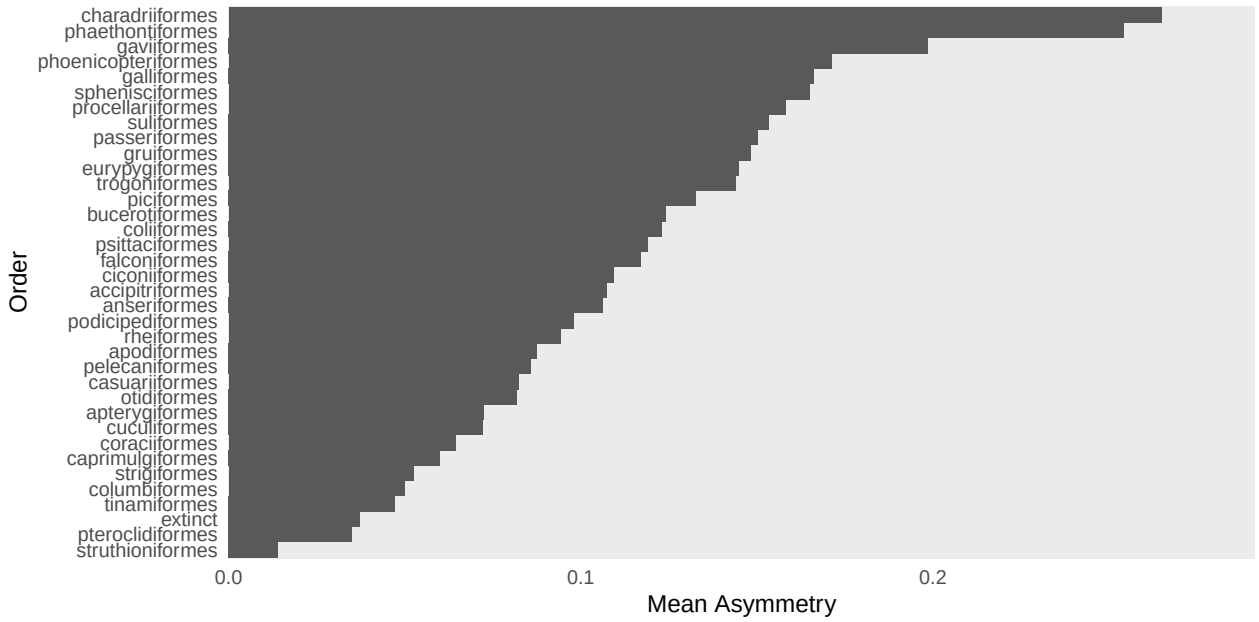
```
birds <- read_csv("https://raw.githubusercontent.com/krisrs1128/stat479_s22/main/exercises/data/birds
filter(!is.na(Order))

birds
```

```
## # A tibble: 1,400 x 9
##    Order           Family       MVZDatabase        Species             Asymmetry Ellipticity `AvgLeng
##    <chr>           <chr>        <chr>              <chr>                   <dbl>       <dbl>
##  1 ACCIPITRIFORMES Accipitridae Accipiter badius   Accipiter badius      0.138        0.344
##  2 ACCIPITRIFORMES Accipitridae Accipiter cooperii Accipiter cooperii    0.0937       0.272
##  3 ACCIPITRIFORMES Accipitridae Accipiter gentilis Accipiter gentilis    0.111        0.319
##  4 ACCIPITRIFORMES Accipitridae Accipiter nisus    Accipiter nisus       0.0808       0.239
##  5 ACCIPITRIFORMES Accipitridae Accipiter striatus Accipiter striatus    0.0749       0.254
##  6 ACCIPITRIFORMES Accipitridae Aegypius monachus  Aegypius monachus     0.07         0.348
##  7 ACCIPITRIFORMES Accipitridae Aquila chrysaetos  Aquila chrysaetos     0.119        0.306
##  8 ACCIPITRIFORMES Accipitridae Aquila rapax       Aquila rapax          0.125        0.352
##  9 ACCIPITRIFORMES Accipitridae Buteo albicaudatus Buteo albicaudatus    0.0818       0.284
## 10 ACCIPITRIFORMES Accipitridae Buteo brachyurus   Buteo brachyurus      0.140        0.237
## # ... with 1,390 more rows
```

a. [3 points] Provide code to convert the `Order` variable to lower case and to compute the mean `Asymmetry` within each `Order`.

b. [3 points] Provide code needed to generate the plot below. Be as specific as possible.


charadriiformes
phaethontiformes
gaviiformes
phoenicopteriformes
galliformes
sphenisciformes
procellariiformes
suliformes
passeriformes
gruiformes
eurypygiformes
trogoniformes
piciformes
bucerotiformes
coliiformes
psittaciformes
falconiformes
ciconiiformes
accipitriformes
anseriformes
podicipediformes
rheiformes
apodiformes
pelecaniformes
casuariiformes
otidiformes
apterygiformes
cuculiformes
coraciiformes
caprimulgiformes
strigiformes
columbiformes
tinamiformes
extinct
pteroclidiformes
struthioniformes
Order
0.0
0.1
0.2
Mean Asymmetry