HW 10 Instructional Project 1

STAT 5400

Due: Nov 15, 2024 9:30 AM

Problems

1. ggplot2 (prepared by Connor Curtiss, Pradeep Maripala, Ariadna Orbe Vivero, Behrooz Khalil Loo, Bowen Su)

Complete the two questions on the slides: ggplot2.pptx, on ICON.

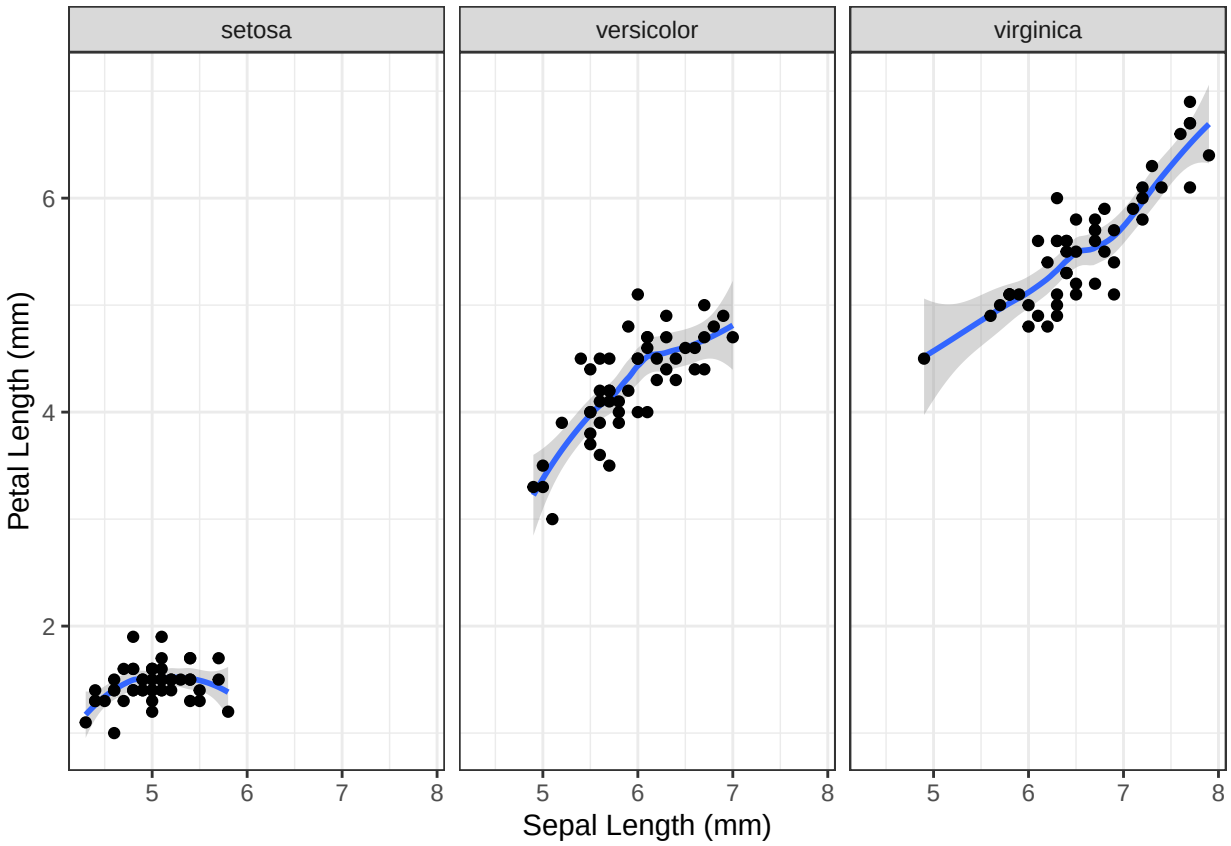
• Write code using ggplot2 that produces a plot facet wrapped by species that shows the relationship between sepal length and petal length. Use geom_smooth to draw lines of best fit through the points

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
ggplot(iris, mapping = aes(x = Sepal.Length, y=Petal.Length)) +
  geom_smooth() +
  geom_point() +
  facet_wrap(~Species) +
  labs(x="Sepal Length (mm)", y="Petal Length (mm)") +
  theme_bw()
```

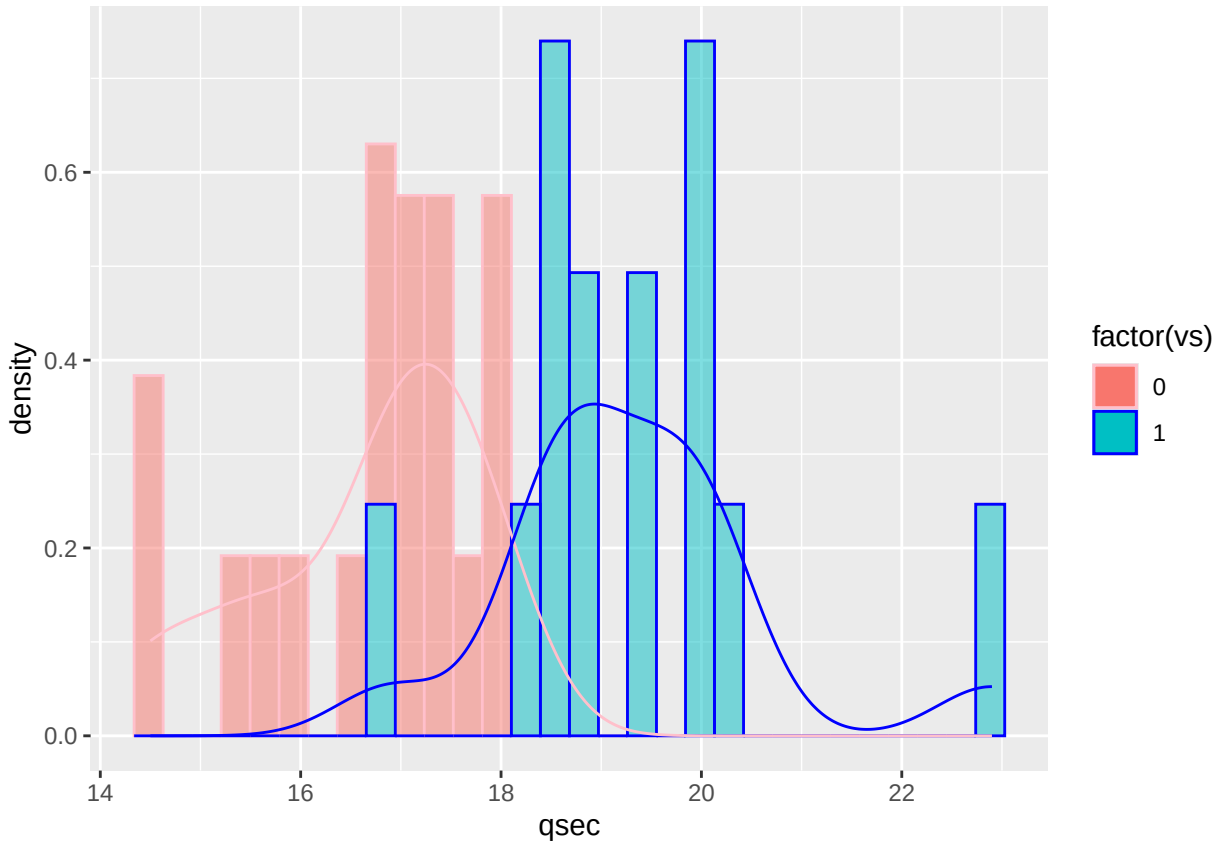
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Can you recreate the following graph using ggplot2?

```
ggplot(data=mtcars, mapping = aes(x = qsec, color=factor(vs)))+
     geom_histogram(aes(y = after_stat(density), fill = factor(vs)),alpha = 0.5) +
  geom_density() +
  scale_color_manual(values = c( "pink", 'blue'))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



2. Numpy and Pandas (prepared by Akhilesh Karra, Vaishnavi Soni, Nathan Munshower, Michael Sullivan, Dongwei Zhang)

Complete the questions on the slides: Numpy and Pandas.pptx, on ICON.

```
import numpy as np
arr = np.arange(1,101).reshape(10,10)
arr
array([[  1,   2,   3,   4,   5,   6,   7,   8,   9,  10],
       [ 11,  12,  13,  14,  15,  16,  17,  18,  19,  20],
       [ 21,  22,  23,  24,  25,  26,  27,  28,  29,  30],
       [ 31,  32,  33,  34,  35,  36,  37,  38,  39,  40],
       [ 41,  42,  43,  44,  45,  46,  47,  48,  49,  50],
       [ 51,  52,  53,  54,  55,  56,  57,  58,  59,  60],
       [ 61,  62,  63,  64,  65,  66,  67,  68,  69,  70],
       [ 71,  72,  73,  74,  75,  76,  77,  78,  79,  80],
       [ 81,  82,  83,  84,  85,  86,  87,  88,  89,  90],
       [ 91,  92,  93,  94,  95,  96,  97,  98,  99, 100]])

arr[-5:,-5:]
array([[ 56,  57,  58,  59,  60],
       [ 66,  67,  68,  69,  70],
       [ 76,  77,  78,  79,  80],
       [ 86,  87,  88,  89,  90],
       [ 96,  97,  98,  99, 100]])
```

```
arr[-5:,-5:].flatten() + 5
array([ 61,  62,  63,  64,  65,  71,  72,  73,  74,  75,  81,  82,  83,
        84,  85,  91,  92,  93,  94,  95, 101, 102, 103, 104, 105])

np.median(arr[-5:,-5:].flatten() + 5)
83.0
```

3. Practice of Julia (prepared by Phoebe Low and Ting-Hung Yu, STAT 5400 Fall 2020 alumni)

Watch the following videos and read Julia slides.pdf on the ICON site.

https://uicapture.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=f4b86ede-d182-4e87-9047-ac4800feeacf

https://uicapture.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=fd3ad007-deb3-4aa3-9f49-ac4800fee7b9

Finish the homework question assigned on the slides.

Use the online Julia platform https://repl.it/languages/julia or install Julia locally to write a function, say m(x), that finds mean and variance (no packages needed) simultaneously.

- Test on sequence of integers 1:10000.
- Paste the code below on the RMarkdown file, and attach a screenshot of results to the PDF.

```
put your code here

"Hello World"

function m(x)
  total = 0
  length = 0
  var = 0
  for i in x
    total = total + i
    length =length + 1
  end
  mn = total/length

  for i in x
    var = var + (i-mn)^2
  end
  var = var/(length-1)
  return (mn, var)
end

println(m(1:10000))


# Use the code below to insert the screenshot into pdf
```

4. Connect R to ChatGPT Watch the following videos, and explore using ChatGPT in R throught API. You do not need to submit anything for this question. https://www.youtube.com/watch?v=szPIuzQ-jco&ab_channel=AnalyticoHub

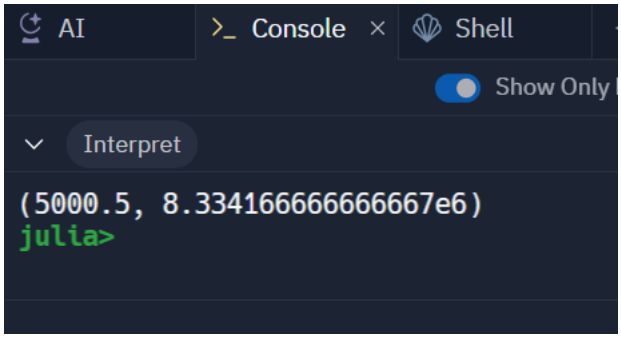


Figure 1: Caption for the picture.