

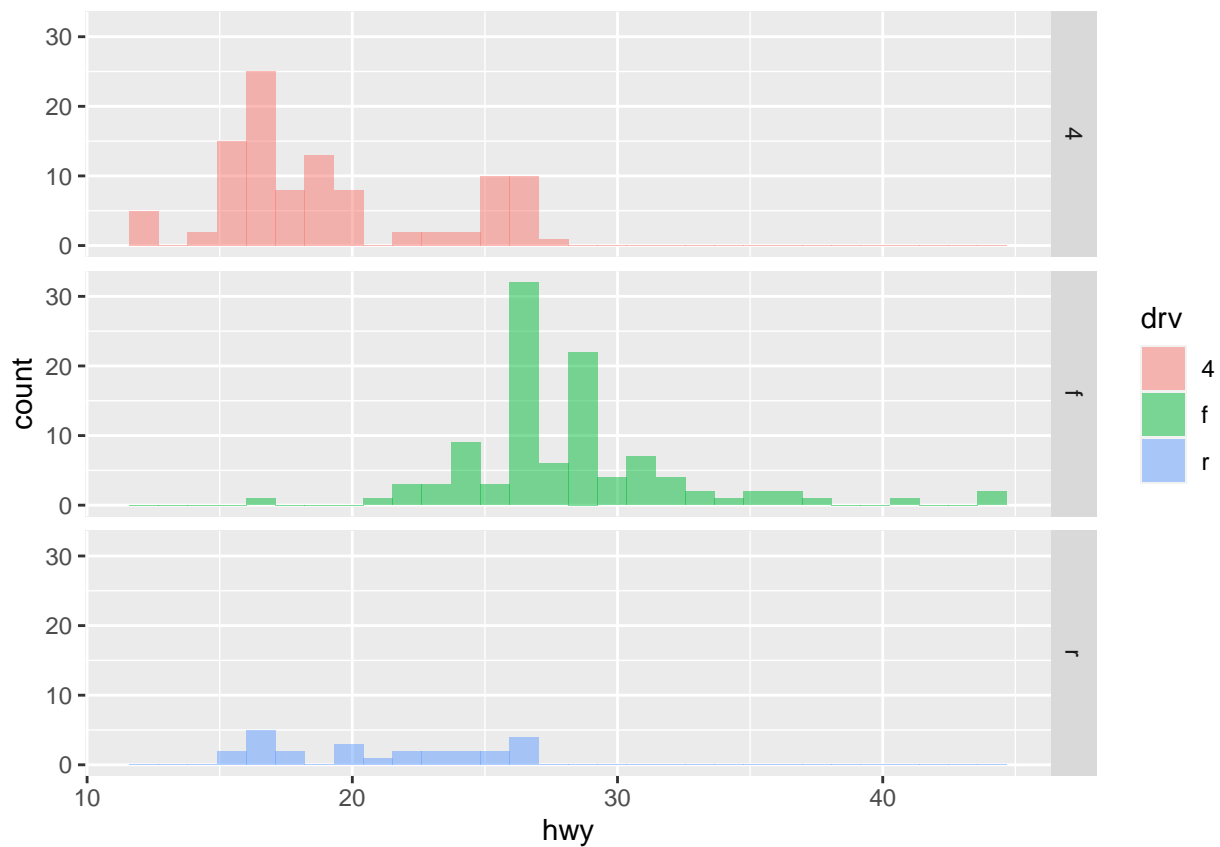
Internal Assessment 1

ECON03SEC1 (Backlog)
Department of Economics
Presidency University, Kolkata
Full Marks: 50
23 Nov 2023

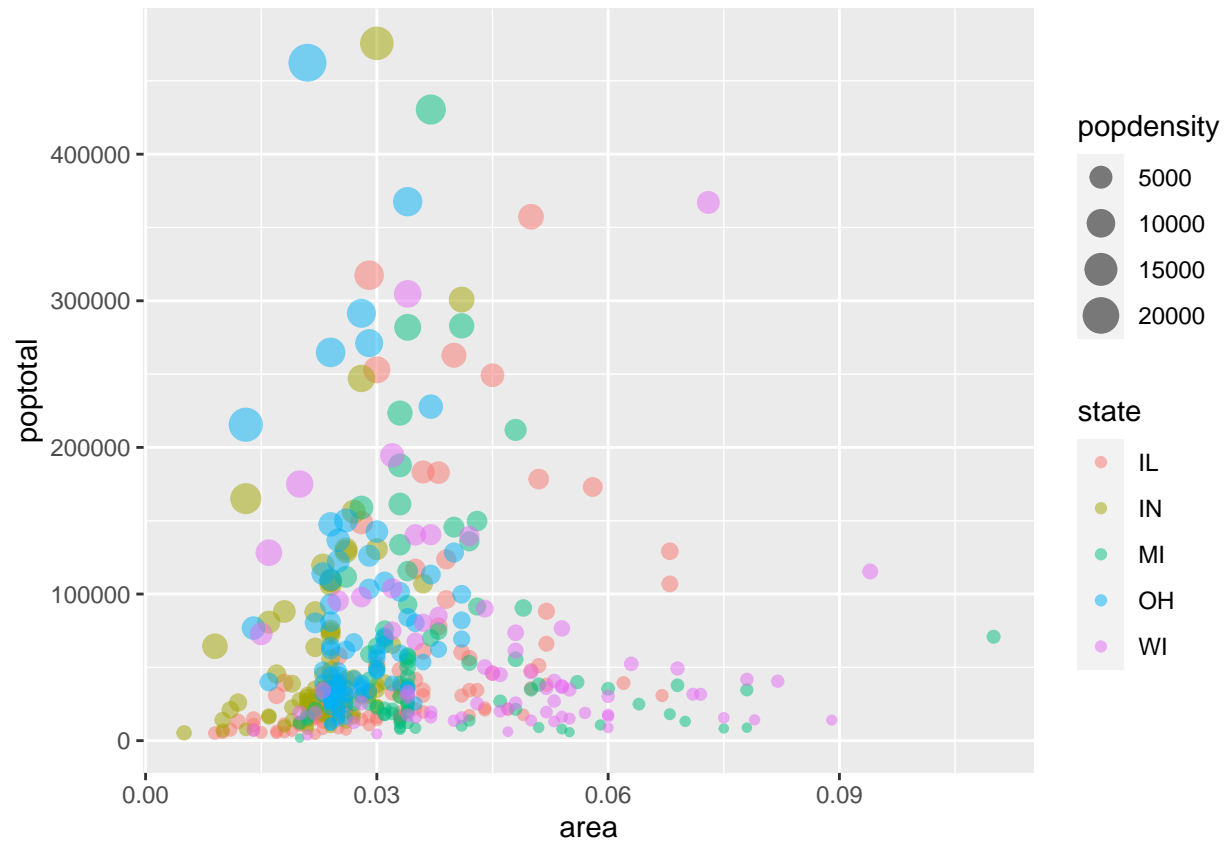
ggplot2

Answer any 2 questions. [2 x 2.5 = 5]

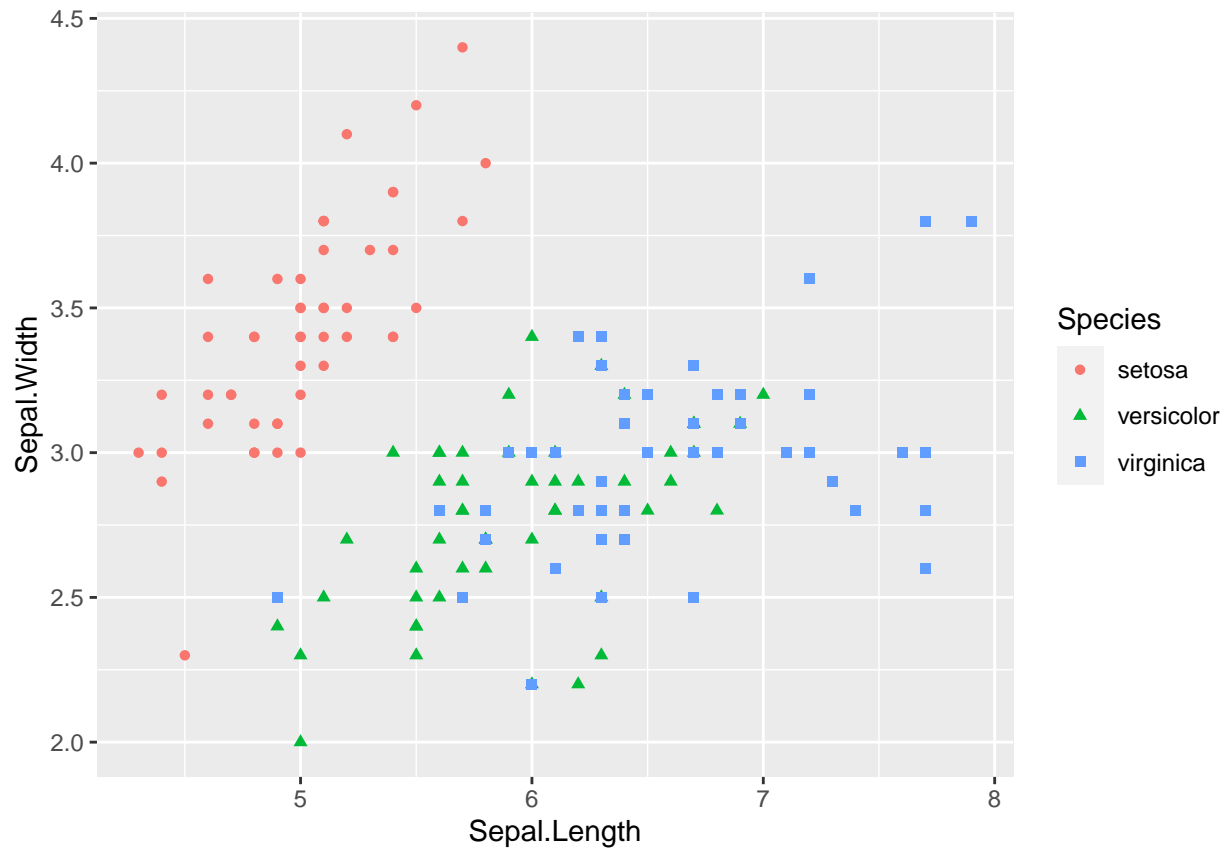
1. Using the `mpg` dataset in the `ggplot2` package, replicate the following plot.



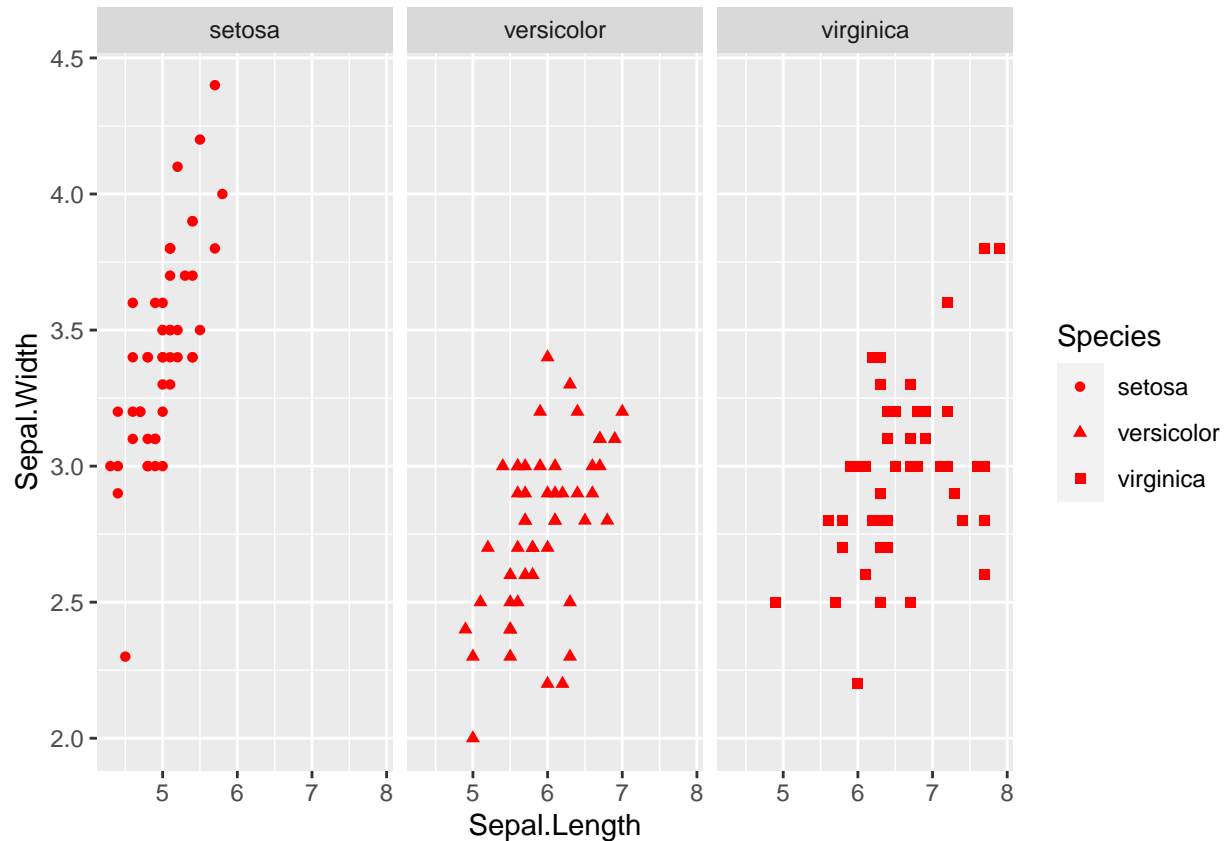
2. Using the `midwest` dataset in the `ggplot2` package, replicate the following plot.



3. Using the `iris` dataset in the base R `datasets` package, replicate the following plot.



4. Using the `iris` dataset in the base R `datasets` package, replicate the following plot.



dplyr

Answer any 3 questions. [$3 \times 5 = 15$]

- Which species have blue eyes in the `starwars` dataset in the `dplyr` package?
- How many female humans are there in the `starwars` dataset in the `dplyr` package ?
- In the `starwars` dataset in the `dplyr` package, what is the average mass of female Human species ?
- In the `starwars` dataset in the `dplyr` package, how many species have a fair skin colour?

tidyr

Answer any 3 questions. [$3 \times 5 = 15$]

- Tidy and replicate the `fish_encounters` dataset in the `tidyr` package as given below.

```
## # A tibble: 5 x 12
##   fish Release I80_1 Lisbon Rstr Base_TD BCE BCW BCE2 BCW2 MAE MAW
##   <fct>   <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>
## 1 4842         1     1     1     1     1     1     1     1     1     1     1
## 2 4843         1     1     1     1     1     1     1     1     1     1     1
## 3 4844         1     1     1     1     1     1     1     1     1     1     1
## 4 4858         1     1     1     1     1     1     1     1     1     1     1
```

```
## 5 4861      1      1      1      1      1      1      1      1      1      1      1
```

10. Tidy and replicate the `who` dataset in the `tidyr` package as given below.

```
## # A tibble: 56 x 6
##   country iso2 iso3 year Diagnosis Value
##   <chr>   <chr> <chr> <dbl> <chr>   <dbl>
## 1 India   IN     IND   2002 new_sp_m3544 55829
## 2 India   IN     IND   2002 new_sp_m2534 54719
## 3 India   IN     IND   2002 new_sp_m4554 44532
## 4 India   IN     IND   2002 new_sp_m1524 39923
## 5 India   IN     IND   2002 new_sp_f2534 31946
## 6 India   IN     IND   2002 new_sp_f1524 28573
## 7 India   IN     IND   2002 new_sp_m5564 28199
## 8 India   IN     IND   2002 new_sp_f3544 21378
## 9 India   IN     IND   2002 new_sp_m65   14960
## 10 India  IN     IND   2002 new_sp_f4554 13233
## # i 46 more rows
```

11. Tidy and replicate the `world_bank_pop` dataset in the `tidyr` package as given below.

```
## # A tibble: 4 x 4
##   country indicator Year Population
##   <chr>   <chr>      <chr>      <dbl>
## 1 IND     SP.POP.GROW 2000          1.82
## 2 IND     SP.URB.GROW 2000          2.60
## 3 IND     SP.URB.TOTL 2000 293168849
## 4 IND     SP.POP.TOTL 2000 1059633675
```

12. Tidy and replicate the `population` dataset in the `tidyr` package as given below.

```
## # A tibble: 2 x 20
##   country '1995' '1996' '1997' '1998' '1999' '2000' '2001' '2002' '2003' '2004'
##   <chr>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 India   9.56e8 9.73e8 9.90e8 1.01e9 1.03e9 1.04e9 1.06e9 1.08e9 1.09e9 1.11e9
## 2 China   1.24e9 1.25e9 1.26e9 1.27e9 1.27e9 1.28e9 1.29e9 1.30e9 1.30e9 1.31e9
## # i 9 more variables: '2005' <dbl>, '2006' <dbl>, '2007' <dbl>, '2008' <dbl>,
## #   '2009' <dbl>, '2010' <dbl>, '2011' <dbl>, '2012' <dbl>, '2013' <dbl>
```

base R

Answer any 3 questions. [$3 \times 5 = 15$]

13. Explain the following codes and their outputs.

```
a1 <- 12; class(a1); length(a1)
names(a1) <- 'Number'; names(a1)
```

14. Explain the following codes and their outputs.

```
a2 <- matrix(1:9, nrow = 3)
colnames(a2) <- c("A", "B", "C")

a2[c(TRUE, FALSE, TRUE), c("B", "A")]
```

15. Explain the following codes and their outputs.

```
month_levels <- c(
  "Jan", "Feb", "Mar", "Apr", "May", "Jun",
  "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"
)
a4 <- factor(c("Dec", "Apr", "Jan", "Mar"), levels = month_levels)
a4
```

16. Explain the following codes and their outputs.

```
a5 <- factor(c("high", "low", "medium", "medium", "high"),
             levels = c("low", "medium", "high"), ordered = TRUE)
a5
```