

test.R

vikas

2021-09-29

```
#title: "Homewok 1"
#author: "Vikas Gunnam"
#date: "9/29/2021"
#output: pdf_document
#Gnumber: G01337418

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr    1.0.7
## v tidyr   1.1.4     v stringr  1.4.0
## v readr   2.0.1     vforcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()

library(tidyr)
library(dplyr)
library(ggplot2)

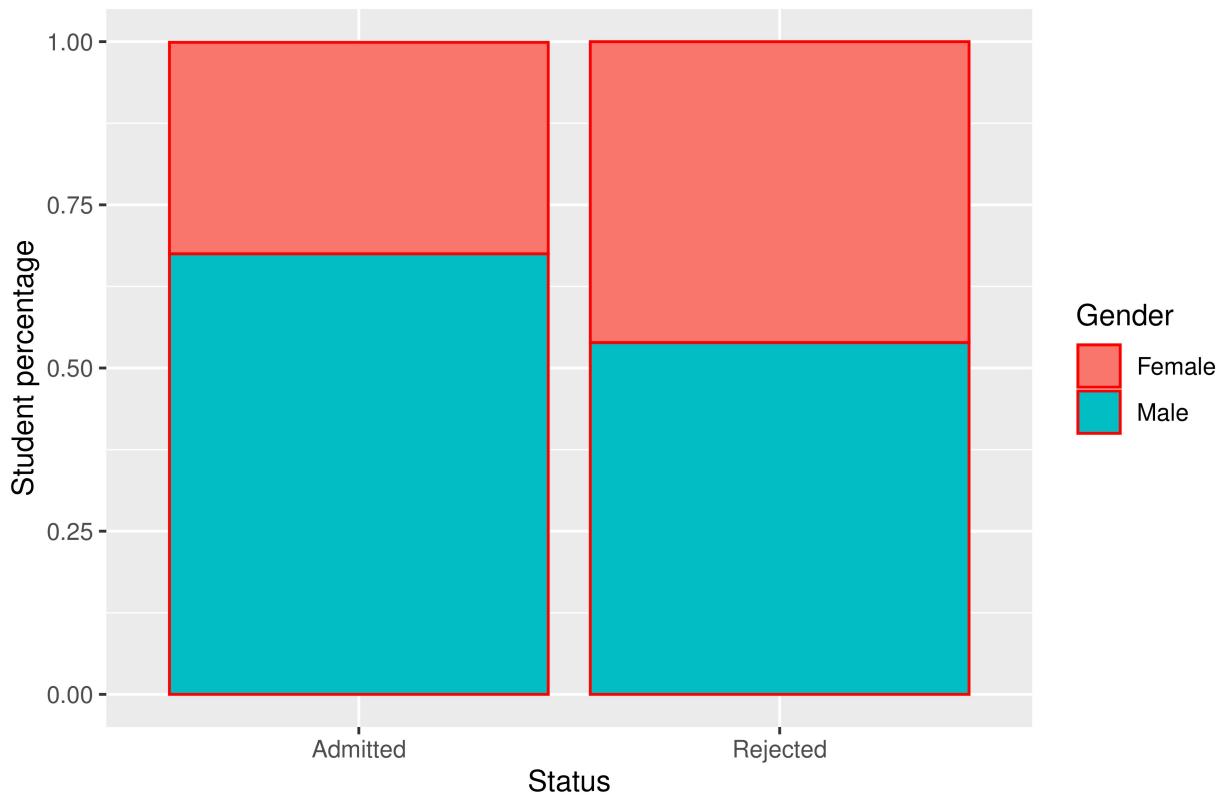
## 1) Create an appropriate graphic to show the university-level Admissions (Use Admission.csv).
q1 <- read.csv("C:/Users/vikas/assignment/Admission.csv")

## Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
## incomplete final line found by readTableHeader on 'C:/Users/vikas/assignment/
## Admission.csv'
q1

##   Gender   Admit Freq Prop
## 1 Male    Admitted 1158 0.675
## 2 Female  Admitted  557 0.324
## 3 Male    Rejected 1493 0.539
## 4 Female  Rejected 1278 0.461

admissions <- ggplot(q1 , aes(x=Admit,y=Prop, fill = Gender)) +
  geom_bar(stat = "identity", col="red") +
  labs(title = "University Level Admits and Rejects",
       y = "Student percentage",
       x = "Status")
admissions
```

University Level Admits and Rejects



```
##2) Are admissions gender biased? What other factors could influence admission rates? Discuss.
#According to the visualization we can see that the number of males admitted are slightly more than
#females On the other hand the rejection percentage remains almost same for both males and females.
#Overall it can be seen that the admits of males is slightly high compared to females.The number of
#applicants appliedfor the course is the major factor.Individually they may maintain equal male to
#female ratio based on courses but overall in some courses there may be a chance that females are
#not taking particular course.
```

```
## 3) Assume admissions are conducted at the department level.
```

```
##Create an appropriate graphic to show the department level Admissions. (use Admission_by_dept.csv).
q3 <- read.csv("C:/Users/vikas/assignment/Admission_by_dept.csv")
q3
```

```
##   Gender   Admit Dept Freq  Prop
## 1   Male Admitted    A  512 0.852
## 2 Female Admitted    A   89 0.148
## 3   Male Rejected    A  313 0.943
## 4 Female Rejected    A   19 0.057
## 5   Male Admitted    B  313 0.948
## 6 Female Admitted    B   17 0.052
## 7   Male Rejected    B  207 0.963
## 8 Female Rejected    B    8 0.037
## 9   Male Admitted    C  120 0.373
## 10  Female Admitted  C  202 0.627
```

```

## 11  Male Rejected    C  205 0.344
## 12 Female Rejected  C  391 0.656
## 13  Male Admitted   D  138 0.513
## 14 Female Admitted  D  131 0.487
## 15  Male Rejected   D  279 0.533
## 16 Female Rejected  D  244 0.467
## 17  Male Admitted   E   53 0.361
## 18 Female Admitted  E   94 0.639
## 19  Male Rejected   E  138 0.316
## 20 Female Rejected  E  299 0.684
## 21  Male Admitted   F   22 0.478
## 22 Female Admitted  F   24 0.522
## 23  Male Rejected   F  351 0.525
## 24 Female Rejected  F  317 0.475

data <- gather(q3, key = "Gender_Admitted", value ="Freq", 5:5)
data

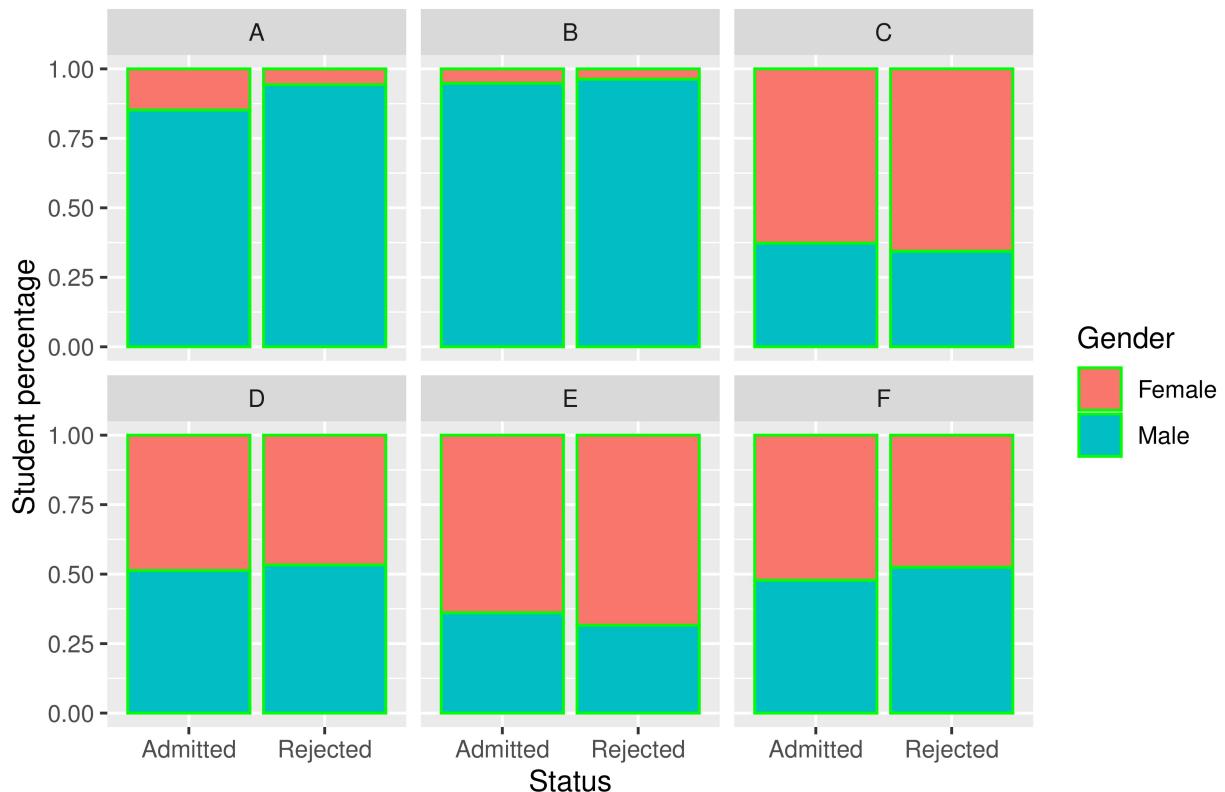
##   Gender     Admit Dept Gender_Admitted   Freq
## 1 Male Admitted A      Prop 0.852
## 2 Female Admitted A      Prop 0.148
## 3 Male Rejected  A      Prop 0.943
## 4 Female Rejected A      Prop 0.057
## 5 Male Admitted B      Prop 0.948
## 6 Female Admitted B      Prop 0.052
## 7 Male Rejected  B      Prop 0.963
## 8 Female Rejected B      Prop 0.037
## 9 Male Admitted C      Prop 0.373
## 10 Female Admitted C     Prop 0.627
## 11 Male Rejected  C      Prop 0.344
## 12 Female Rejected C      Prop 0.656
## 13 Male Admitted D      Prop 0.513
## 14 Female Admitted D      Prop 0.487
## 15 Male Rejected  D      Prop 0.533
## 16 Female Rejected D      Prop 0.467
## 17 Male Admitted E      Prop 0.361
## 18 Female Admitted E      Prop 0.639
## 19 Male Rejected  E      Prop 0.316
## 20 Female Rejected E      Prop 0.684
## 21 Male Admitted F      Prop 0.478
## 22 Female Admitted F      Prop 0.522
## 23 Male Rejected  F      Prop 0.525
## 24 Female Rejected F      Prop 0.475

sep_data_plot <- ggplot(data=data, aes(x=Admit, y = Freq, fill= Gender))+ 
  geom_bar(stat = "identity", col="green") +labs(title = "University Level Admits and Rejects",
                                                y = " Student percentage",
                                                x = "Status")+facet_wrap(~Dept)

sep_data_plot

```

University Level Admits and Rejects



```
##4) Do your conclusions differ? What other factors are at play?
##According to different departments the ratio is changing. For instance in department C
#the admit and reject ratio is same and we can see that males are dominant in both admits and
#rejects whereas in department E females are dominant in both admits and rejects.
```

##bonus question

```
qb <- read.csv("C:/Users/vikas/assignment/Original_Admissions_Data.csv")
qb
```

```
##   Dept Male_Admitted Male_Rejected Female_Admitted Female_Rejected
## 1   A        512       313         89          19
## 2   B        313       207         17          8
## 3   C        120       205        202         391
## 4   D        138       279        131         244
## 5   E        53        138         94         299
## 6   F        22        351         24         317
```

```
data <- gather(qb, key = "Gender_Admitted", value ="Freq", 5:5)
data
```

```
##   Gender Admit Dept Gender_Admitted   Freq
## 1   Male Admitted     A             Prop 0.852
## 2 Female Admitted     A             Prop 0.148
## 3   Male Rejected     A             Prop 0.943
```

```

## 4 Female Rejected A Prop 0.057
## 5 Male Admitted B Prop 0.948
## 6 Female Admitted B Prop 0.052
## 7 Male Rejected B Prop 0.963
## 8 Female Rejected B Prop 0.037
## 9 Male Admitted C Prop 0.373
## 10 Female Admitted C Prop 0.627
## 11 Male Rejected C Prop 0.344
## 12 Female Rejected C Prop 0.656
## 13 Male Admitted D Prop 0.513
## 14 Female Admitted D Prop 0.487
## 15 Male Rejected D Prop 0.533
## 16 Female Rejected D Prop 0.467
## 17 Male Admitted E Prop 0.361
## 18 Female Admitted E Prop 0.639
## 19 Male Rejected E Prop 0.316
## 20 Female Rejected E Prop 0.684
## 21 Male Admitted F Prop 0.478
## 22 Female Admitted F Prop 0.522
## 23 Male Rejected F Prop 0.525
## 24 Female Rejected F Prop 0.475

columnseparated_data <- separate(data,Gender_Admitted, into = c("Gender", "Admit") )

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 24 rows [1, 2, 3,
## 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].
#view(columnseparated_data)

sep_data_plot <- ggplot(data=data, aes(x=Admit, y = Freq, fill= Gender))+ 
  geom_bar(stat = "identity", col="green") +labs(title = "University Level Admits and Rejects",
                                                y = "Student percentage",
                                                x = "Status")+facet_wrap(~Dept)
sep_data_plot

```

University Level Admits and Rejects

