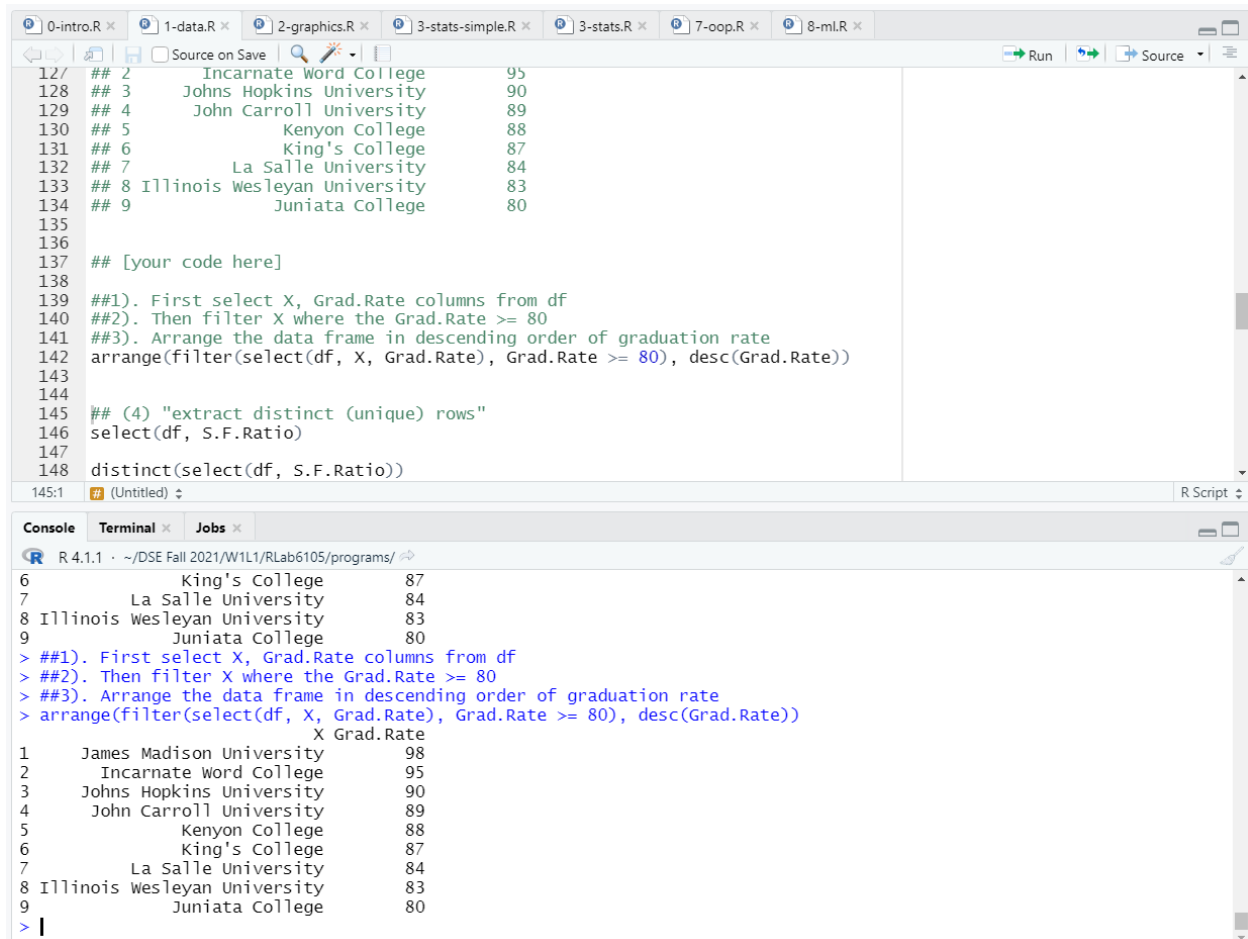


Q1)

- First select X, Grad.Rate columns from df
- Then filter X where the Grad.Rate  $\geq 80$
- Arrange the data frame in descending order of graduation rate.



The screenshot shows the RStudio environment with several open scripts. The active script, '3-stats.R', contains the following R code:

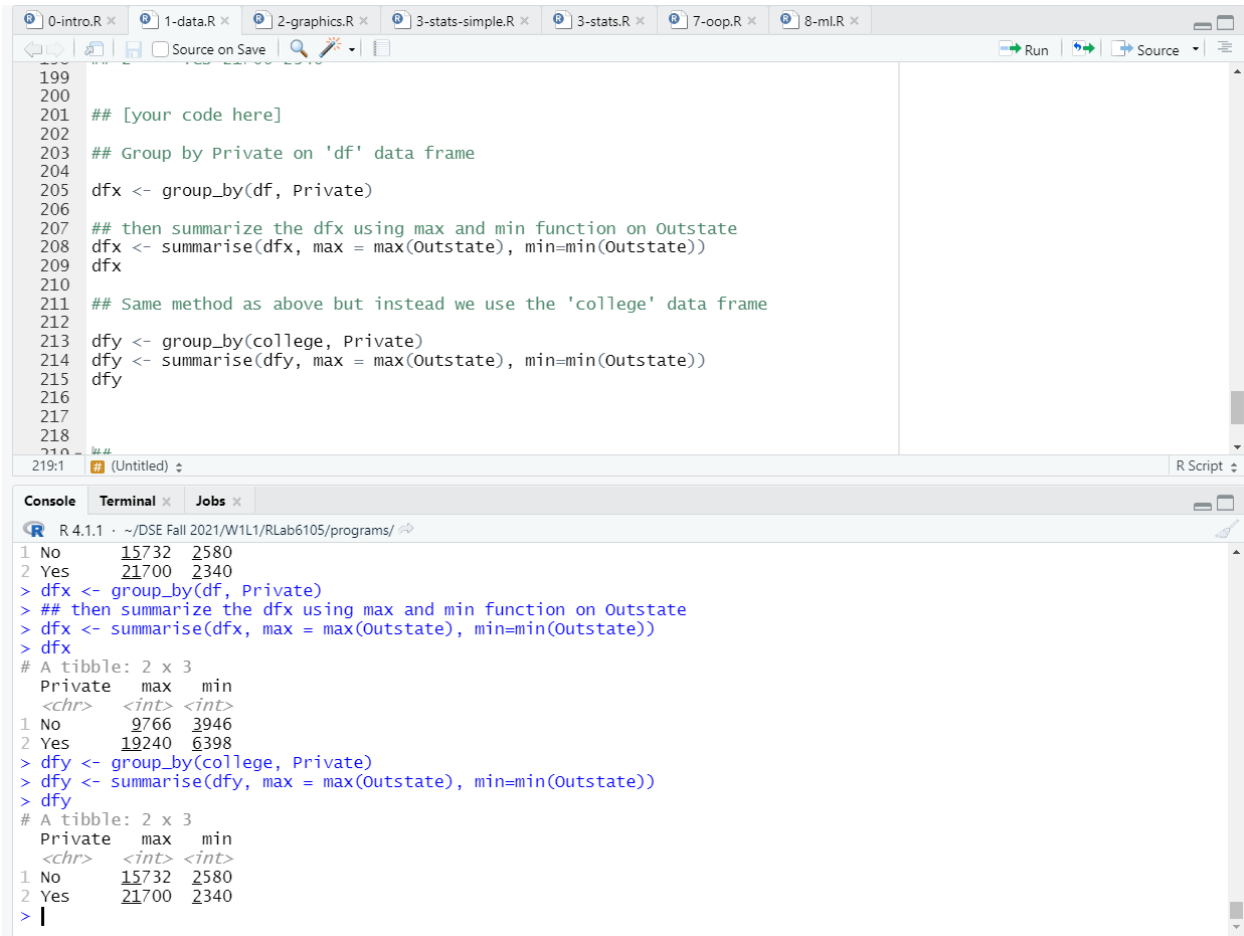
```
127 ## 2 Incarnate Word College 95
128 ## 3 Johns Hopkins University 90
129 ## 4 John Carroll University 89
130 ## 5 Kenyon College 88
131 ## 6 King's College 87
132 ## 7 La Salle University 84
133 ## 8 Illinois Wesleyan University 83
134 ## 9 Juniata College 80
135
136
137 ## [your code here]
138
139 ##1). First select X, Grad.Rate columns from df
140 ##2). Then filter X where the Grad.Rate  $\geq 80$ 
141 ##3). Arrange the data frame in descending order of graduation rate
142 arrange(filter(select(df, X, Grad.Rate), Grad.Rate  $\geq 80$ ), desc(Grad.Rate))
143
144
145 ## (4) "extract distinct (unique) rows"
146 select(df, S.F.Ratio)
147
148 distinct(select(df, S.F.Ratio))
```

The console output shows the result of the R code execution:

```
R 4.1.1 ~ /DSE Fall 2021/W1L1/RLab6105/programs/
6 King's College 87
7 La Salle University 84
8 Illinois Wesleyan University 83
9 Juniata College 80
> ##1). First select X, Grad.Rate columns from df
> ##2). Then filter X where the Grad.Rate  $\geq 80$ 
> ##3). Arrange the data frame in descending order of graduation rate
> arrange(filter(select(df, X, Grad.Rate), Grad.Rate  $\geq 80$ ), desc(Grad.Rate))
      X Grad.Rate
1 James Madison University 98
2 Incarnate Word College 95
3 Johns Hopkins University 90
4 John Carroll University 89
5 Kenyon College 88
6 King's College 87
7 La Salle University 84
8 Illinois Wesleyan University 83
9 Juniata College 80
> |
```

Q2)

- Group by Private on 'df' data frame to obtain dfx
- Then summarize the dfx using max and min function on Outstate
- For the next one, same method as above but instead we use the 'college' data frame



The screenshot shows the RStudio IDE with several tabs open at the top: 0-intro.R, 1-data.R, 2-graphics.R, 3-stats-simple.R, 3-stats.R, 7-oop.R, and 8-ml.R. The main editor window contains the following R code:

```
199  
200  
201 ## [your code here]  
202  
203 ## Group by Private on 'df' data frame  
204  
205 dfx <- group_by(df, Private)  
206  
207 ## then summarize the dfx using max and min function on Outstate  
208 dfx <- summarise(dfx, max = max(Outstate), min=min(Outstate))  
209 dfx  
210  
211 ## Same method as above but instead we use the 'college' data frame  
212  
213 dfy <- group_by(college, Private)  
214 dfy <- summarise(dfy, max = max(Outstate), min=min(Outstate))  
215 dfy  
216  
217  
218  
219 ##
```

The console window at the bottom shows the execution of the code:

```
R 4.1.1 · ~/DSE Fall 2021/W1L1/RLab6105/programs/  
1 No      15732 2580  
2 Yes      21700 2340  
> dfx <- group_by(df, Private)  
> ## then summarize the dfx using max and min function on Outstate  
> dfx <- summarise(dfx, max = max(Outstate), min=min(Outstate))  
> dfx  
# A tibble: 2 x 3  
  Private    max    min  
  <chr>    <int> <int>  
1 No       9766  3946  
2 Yes     19240  6398  
> dfy <- group_by(college, Private)  
> dfy <- summarise(dfy, max = max(Outstate), min=min(Outstate))  
> dfy  
# A tibble: 2 x 3  
  Private    max    min  
  <chr>    <int> <int>  
1 No      15732 2580  
2 Yes     21700 2340  
> |
```