R\_FinalTest\_ChaytanyaKumar

Chaytanya

12/16/2021

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.6 ✓ dplyr 1.0.7  
## ✓ tidyr 1.1.3 ✓ stringr 1.4.0  
## ✓ readr 2.0.1 ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(stringr)   
library(dplyr)  
library(ggplot2)  
library(tidyr)   
library(reshape2)

##   
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':  
##   
## smiths

library(readr)  
library(forcats)  
  
#1)Using appropriate R code, read in the emailed excel spread sheet  
#Collegedata. Use read\_csv(“Collegedata.csv”)  
  
collegedata <- read\_csv("Collegedata.csv")

## Rows: 7175 Columns: 12

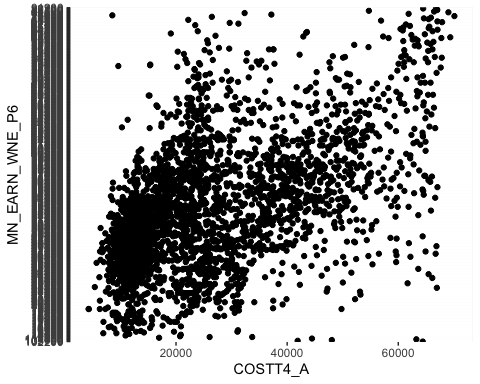
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (4): MN\_EARN\_WNE\_P6, INSTNM, GRAD\_DEBT\_MDN, AGE\_ENTRY  
## dbl (8): UNITID, OPEID, SAT\_AVG, ADM\_RATE, UGDS, COSTT4\_A, AVGFACSAL, ICLEVEL

##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

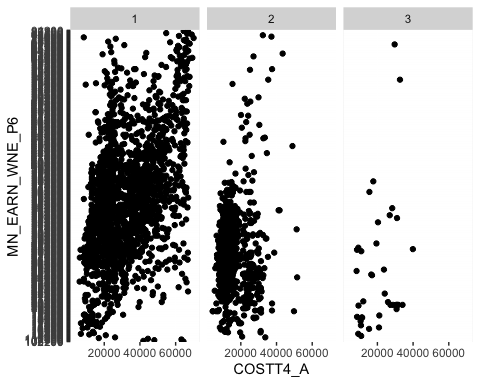
#view(collegedata)  
dim(collegedata)

## [1] 7175 12

#2)  
#Given the level of the institution, does there appear to be an association between  
#the average cost of attendance(x variable) and the mean earnings of students six years  
#after graduation(y variable)? Make appropriate plots to justify your response.   
#(Hint: Generate two plots to make your decision, first a standard scatter plot involving   
#the two continuous variables mentioned and then a facet plot over the appropriate categorical variable)  
  
college = collegedata%>%  
 filter(COSTT4\_A !="PrivacySuppressed", MN\_EARN\_WNE\_P6 !="PrivacySuppressed")  
  
ggplot(data = college)+  
 geom\_point(mapping = aes(x=COSTT4\_A, y=MN\_EARN\_WNE\_P6))



ggplot(data = college)+  
 geom\_point(mapping = aes(x=COSTT4\_A, y=MN\_EARN\_WNE\_P6))+  
 facet\_wrap(~ICLEVEL)



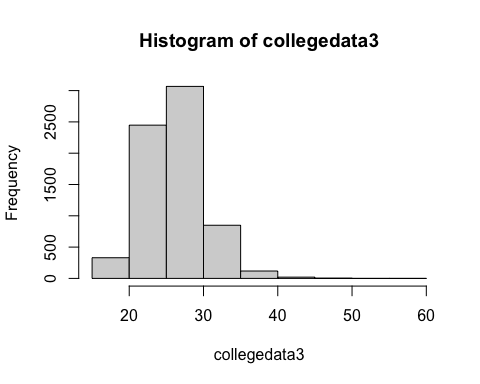
#From the scatter plot we can say that there is moderate positive association,   
#between the variables-average cost of attendance (x variable) and   
#the mean earnings of students six years.The facetplot also shows that the graph  
#has upward trend.  
  
#3)Use r code to produce a histogram of the average age of entry.   
#Comment on the distribution of this variable.  
  
  
collegedata2 = collegedata%>%  
select(AGE\_ENTRY)%>%  
filter(AGE\_ENTRY !="PrivacySuppressed" )  
collegedata2

## # A tibble: 6,839 × 1  
## AGE\_ENTRY   
## <chr>   
## 1 20.28374137  
## 2 23.60797466  
## 3 33.6722973   
## 4 22.72791963  
## 5 20.13099042  
## 6 21.1377014   
## 7 23.07102804  
## 8 30.38476563  
## 9 22.19980411  
## 10 20.51679587  
## # … with 6,829 more rows

str(collegedata2)

## tibble [6,839 × 1] (S3: tbl\_df/tbl/data.frame)  
## $ AGE\_ENTRY: chr [1:6839] "20.28374137" "23.60797466" "33.6722973" "22.72791963" ...

collegedata3<- as.numeric(collegedata2$AGE\_ENTRY)  
hist(collegedata3)



# From the graph we can see that the distribution is slightly skewed to the left.  
#We can interpret that the majority of the cases for the average age of entry  
#are between 20 to 30 years old.  
  
#4)Use r code that will produce output that shows the 10 institutions that  
#have the highest average age of entry?  
  
  
x4<-collegedata%>%  
filter(AGE\_ENTRY !="PrivacySuppressed" )%>%  
arrange(desc(AGE\_ENTRY))%>%  
print(n=10)

## # A tibble: 6,839 × 12  
## UNITID OPEID MN\_EARN\_WNE\_P6 INSTNM SAT\_AVG ADM\_RATE UGDS COSTT4\_A AVGFACSAL  
## <dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 383172 2.29e6 PrivacySuppre… Advan… NA NA 42 NA NA  
## 2 401223 3.87e6 PrivacySuppre… World… NA NA 124 20965 3550  
## 3 199971 3.57e6 PrivacySuppre… Carol… NA NA 32 NA NA  
## 4 481058 4.16e6 <NA> Grace… NA 0.727 99 17794 NA  
## 5 481368 4.21e6 <NA> Prest… NA NA 46 NA NA  
## 6 441511 3.51e6 PrivacySuppre… Apex … NA NA 755 11550 3136  
## 7 461236 4.16e6 <NA> Georg… NA NA 261 26225 3021  
## 8 139153 3.08e6 30200 Beula… NA NA 351 18230 3995  
## 9 454689 4.10e6 <NA> Taft … NA NA 8 NA 2546  
## 10 461449 4.17e6 <NA> Cosmo… NA NA 107 NA NA  
## # … with 6,829 more rows, and 3 more variables: GRAD\_DEBT\_MDN <chr>,  
## # AGE\_ENTRY <chr>, ICLEVEL <dbl>

#view(x4)  
  
  
#5)There are many universities with "American University" in the name. E.g.  
#"American University of Puerto Rico" and "National American  
 #University-Ellsworth AFB Extension". Use R code to create a data frame, called   
#`americandf`, that contains just the data from universities with  
#"American University" in the name.  
  
x1<-str\_subset(collegedata$INSTNM, "American University")  
#x1  
americandf<- collegedata%>%  
filter(INSTNM %in% x1)  
americandf

## # A tibble: 47 × 12  
## UNITID OPEID MN\_EARN\_WNE\_P6 INSTNM SAT\_AVG ADM\_RATE UGDS COSTT4\_A AVGFACSAL  
## <dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 127680 4.06e5 35200 Natio… NA NA 194 NA 3326  
## 2 131159 1.43e5 46800 Ameri… 1262 0.259 7276 60501 11386  
## 3 174385 4.06e5 35200 Natio… NA NA 152 18707 NA  
## 4 219204 4.06e5 35200 Natio… NA NA 1537 20069 4665  
## 5 219213 4.06e5 35200 Natio… NA NA 365 15870 6751  
## 6 241100 1.19e6 20700 Ameri… NA NA 583 16265 2937  
## 7 241128 1.19e6 20700 Ameri… NA NA 742 16393 2838  
## 8 242617 4.25e6 <NA> Inter… NA 0.691 3972 12881 4873  
## 9 242626 3.94e5 <NA> Inter… NA 0.466 3912 11158 4380  
## 10 242635 5.03e5 <NA> Inter… NA 0.465 3841 12672 4527  
## # … with 37 more rows, and 3 more variables: GRAD\_DEBT\_MDN <chr>,  
## # AGE\_ENTRY <chr>, ICLEVEL <dbl>

#view(americandf)  
  
  
#6)  
#Provide r code that will produce the number of colleges from the College Score data frame  
#that have average SAT scores that are above 1000.   
#( Do not produce the data frame. Your code should only yield the number)  
  
  
collegedata%>%  
 select(INSTNM,SAT\_AVG)%>%  
 filter(SAT\_AVG > 1000)%>%  
 count()

## # A tibble: 1 × 1  
## n  
## <int>  
## 1 849

#7)Provide r code that will show a data frame that lists the 10 highest Average SAT scores in decreasing order.   
#A partial data frame is given below.  
  
collegedata%>%  
arrange(desc(SAT\_AVG))%>%  
print(n=10)

## # A tibble: 7,175 × 12  
## UNITID OPEID MN\_EARN\_WNE\_P6 INSTNM SAT\_AVG ADM\_RATE UGDS COSTT4\_A AVGFACSAL  
## <dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 110404 113100 59800 Calif… 1555 0.0807 979 63471 19190  
## 2 166683 217800 107300 Massa… 1519 0.0794 4489 63250 17545  
## 3 144050 177400 73700 Unive… 1508 0.0794 5978 70100 17399  
## 4 166027 215500 102200 Harva… 1506 0.054 7447 64400 19152  
## 5 130794 142600 83500 Yale … 1502 0.0632 5471 66445 18333  
## 6 115409 117100 61600 Harve… 1496 0.129 829 69355 13165  
## 7 190150 270700 82600 Colum… 1496 0.0683 8124 69021 18089  
## 8 221999 353500 57700 Vande… 1495 0.108 6844 63532 12747  
## 9 186131 262700 78000 Princ… 1493 0.0652 5236 61860 17132  
## 10 227757 360400 61100 Rice … 1490 0.153 3879 58253 15378  
## # … with 7,165 more rows, and 3 more variables: GRAD\_DEBT\_MDN <chr>,  
## # AGE\_ENTRY <chr>, ICLEVEL <dbl>

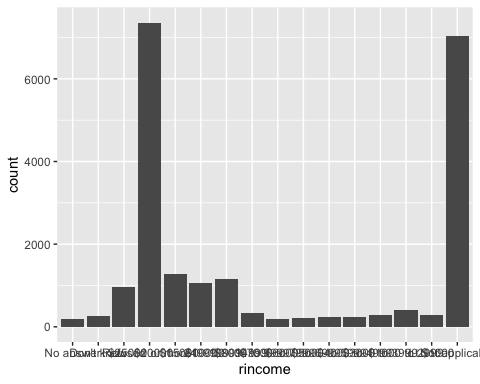
#)8Using the gss\_cat data frame, write r code that will produce the bar graph below.   
#And explain in one or two sentences why the bar graph is difficult to interpret.  
  
gss\_cat

## # A tibble: 21,483 × 9  
## year marital age race rincome partyid relig denom tvhours  
## <int> <fct> <int> <fct> <fct> <fct> <fct> <fct> <int>  
## 1 2000 Never married 26 White $8000 to 9999 Ind,nea… Prote… South… 12  
## 2 2000 Divorced 48 White $8000 to 9999 Not str… Prote… Bapti… NA  
## 3 2000 Widowed 67 White Not applicable Indepen… Prote… No de… 2  
## 4 2000 Never married 39 White Not applicable Ind,nea… Ortho… Not a… 4  
## 5 2000 Divorced 25 White Not applicable Not str… None Not a… 1  
## 6 2000 Married 25 White $20000 - 24999 Strong … Prote… South… NA  
## 7 2000 Never married 36 White $25000 or more Not str… Chris… Not a… 3  
## 8 2000 Divorced 44 White $7000 to 7999 Ind,nea… Prote… Luthe… NA  
## 9 2000 Married 44 White $25000 or more Not str… Prote… Other 0  
## 10 2000 Married 47 White $25000 or more Strong … Prote… South… 3  
## # … with 21,473 more rows

gss\_cat

## # A tibble: 21,483 × 9  
## year marital age race rincome partyid relig denom tvhours  
## <int> <fct> <int> <fct> <fct> <fct> <fct> <fct> <int>  
## 1 2000 Never married 26 White $8000 to 9999 Ind,nea… Prote… South… 12  
## 2 2000 Divorced 48 White $8000 to 9999 Not str… Prote… Bapti… NA  
## 3 2000 Widowed 67 White Not applicable Indepen… Prote… No de… 2  
## 4 2000 Never married 39 White Not applicable Ind,nea… Ortho… Not a… 4  
## 5 2000 Divorced 25 White Not applicable Not str… None Not a… 1  
## 6 2000 Married 25 White $20000 - 24999 Strong … Prote… South… NA  
## 7 2000 Never married 36 White $25000 or more Not str… Chris… Not a… 3  
## 8 2000 Divorced 44 White $7000 to 7999 Ind,nea… Prote… Luthe… NA  
## 9 2000 Married 44 White $25000 or more Not str… Prote… Other 0  
## 10 2000 Married 47 White $25000 or more Strong … Prote… South… 3  
## # … with 21,473 more rows

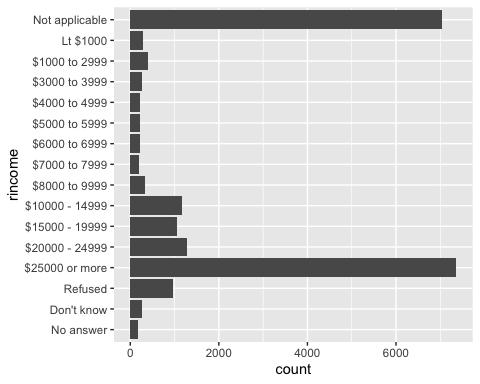
ggplot(data = gss\_cat)+  
geom\_bar(mapping = aes(x=rincome))



# as you can see, We are unable to read the labels of the levels of the   
#salary (tightly coupled due to lack of space)   
#and therefore it is difficult to interpret the Bar plot.   
  
#9)a) Now write r code from the same data set that produce the transformed bar graph   
#and comment on why it is an improvement.  
  
gss\_cat

## # A tibble: 21,483 × 9  
## year marital age race rincome partyid relig denom tvhours  
## <int> <fct> <int> <fct> <fct> <fct> <fct> <fct> <int>  
## 1 2000 Never married 26 White $8000 to 9999 Ind,nea… Prote… South… 12  
## 2 2000 Divorced 48 White $8000 to 9999 Not str… Prote… Bapti… NA  
## 3 2000 Widowed 67 White Not applicable Indepen… Prote… No de… 2  
## 4 2000 Never married 39 White Not applicable Ind,nea… Ortho… Not a… 4  
## 5 2000 Divorced 25 White Not applicable Not str… None Not a… 1  
## 6 2000 Married 25 White $20000 - 24999 Strong … Prote… South… NA  
## 7 2000 Never married 36 White $25000 or more Not str… Chris… Not a… 3  
## 8 2000 Divorced 44 White $7000 to 7999 Ind,nea… Prote… Luthe… NA  
## 9 2000 Married 44 White $25000 or more Not str… Prote… Other 0  
## 10 2000 Married 47 White $25000 or more Strong … Prote… South… 3  
## # … with 21,473 more rows

ggplot(data = gss\_cat)+  
geom\_bar(mapping = aes(x=rincome))+  
coord\_flip()



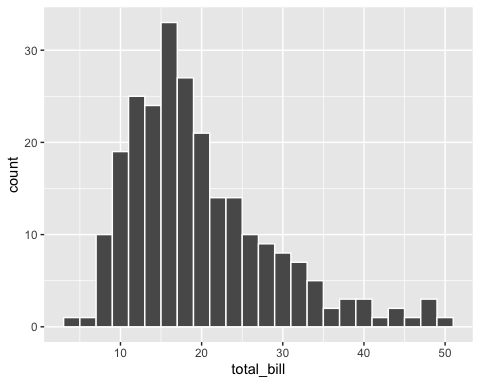
# We can evidently see the difference between the two. The former was difficult to interpret due to  
#congested labels of the levels of salary while later one shows levels of the income more defined and   
#easily to interpret after the co-oridinate flip.   
  
  
  
#b)Use r code to produce the tips data frame from the reshape2 package.   
#Name three categorical variables in the data frame.  
  
tips

## total\_bill tip sex smoker day time size  
## 1 16.99 1.01 Female No Sun Dinner 2  
## 2 10.34 1.66 Male No Sun Dinner 3  
## 3 21.01 3.50 Male No Sun Dinner 3  
## 4 23.68 3.31 Male No Sun Dinner 2  
## 5 24.59 3.61 Female No Sun Dinner 4  
## 6 25.29 4.71 Male No Sun Dinner 4  
## 7 8.77 2.00 Male No Sun Dinner 2  
## 8 26.88 3.12 Male No Sun Dinner 4  
## 9 15.04 1.96 Male No Sun Dinner 2  
## 10 14.78 3.23 Male No Sun Dinner 2  
## 11 10.27 1.71 Male No Sun Dinner 2  
## 12 35.26 5.00 Female No Sun Dinner 4  
## 13 15.42 1.57 Male No Sun Dinner 2  
## 14 18.43 3.00 Male No Sun Dinner 4  
## 15 14.83 3.02 Female No Sun Dinner 2  
## 16 21.58 3.92 Male No Sun Dinner 2  
## 17 10.33 1.67 Female No Sun Dinner 3  
## 18 16.29 3.71 Male No Sun Dinner 3  
## 19 16.97 3.50 Female No Sun Dinner 3  
## 20 20.65 3.35 Male No Sat Dinner 3  
## 21 17.92 4.08 Male No Sat Dinner 2  
## 22 20.29 2.75 Female No Sat Dinner 2  
## 23 15.77 2.23 Female No Sat Dinner 2  
## 24 39.42 7.58 Male No Sat Dinner 4  
## 25 19.82 3.18 Male No Sat Dinner 2  
## 26 17.81 2.34 Male No Sat Dinner 4  
## 27 13.37 2.00 Male No Sat Dinner 2  
## 28 12.69 2.00 Male No Sat Dinner 2  
## 29 21.70 4.30 Male No Sat Dinner 2  
## 30 19.65 3.00 Female No Sat Dinner 2  
## 31 9.55 1.45 Male No Sat Dinner 2  
## 32 18.35 2.50 Male No Sat Dinner 4  
## 33 15.06 3.00 Female No Sat Dinner 2  
## 34 20.69 2.45 Female No Sat Dinner 4  
## 35 17.78 3.27 Male No Sat Dinner 2  
## 36 24.06 3.60 Male No Sat Dinner 3  
## 37 16.31 2.00 Male No Sat Dinner 3  
## 38 16.93 3.07 Female No Sat Dinner 3  
## 39 18.69 2.31 Male No Sat Dinner 3  
## 40 31.27 5.00 Male No Sat Dinner 3  
## 41 16.04 2.24 Male No Sat Dinner 3  
## 42 17.46 2.54 Male No Sun Dinner 2  
## 43 13.94 3.06 Male No Sun Dinner 2  
## 44 9.68 1.32 Male No Sun Dinner 2  
## 45 30.40 5.60 Male No Sun Dinner 4  
## 46 18.29 3.00 Male No Sun Dinner 2  
## 47 22.23 5.00 Male No Sun Dinner 2  
## 48 32.40 6.00 Male No Sun Dinner 4  
## 49 28.55 2.05 Male No Sun Dinner 3  
## 50 18.04 3.00 Male No Sun Dinner 2  
## 51 12.54 2.50 Male No Sun Dinner 2  
## 52 10.29 2.60 Female No Sun Dinner 2  
## 53 34.81 5.20 Female No Sun Dinner 4  
## 54 9.94 1.56 Male No Sun Dinner 2  
## 55 25.56 4.34 Male No Sun Dinner 4  
## 56 19.49 3.51 Male No Sun Dinner 2  
## 57 38.01 3.00 Male Yes Sat Dinner 4  
## 58 26.41 1.50 Female No Sat Dinner 2  
## 59 11.24 1.76 Male Yes Sat Dinner 2  
## 60 48.27 6.73 Male No Sat Dinner 4  
## 61 20.29 3.21 Male Yes Sat Dinner 2  
## 62 13.81 2.00 Male Yes Sat Dinner 2  
## 63 11.02 1.98 Male Yes Sat Dinner 2  
## 64 18.29 3.76 Male Yes Sat Dinner 4  
## 65 17.59 2.64 Male No Sat Dinner 3  
## 66 20.08 3.15 Male No Sat Dinner 3  
## 67 16.45 2.47 Female No Sat Dinner 2  
## 68 3.07 1.00 Female Yes Sat Dinner 1  
## 69 20.23 2.01 Male No Sat Dinner 2  
## 70 15.01 2.09 Male Yes Sat Dinner 2  
## 71 12.02 1.97 Male No Sat Dinner 2  
## 72 17.07 3.00 Female No Sat Dinner 3  
## 73 26.86 3.14 Female Yes Sat Dinner 2  
## 74 25.28 5.00 Female Yes Sat Dinner 2  
## 75 14.73 2.20 Female No Sat Dinner 2  
## 76 10.51 1.25 Male No Sat Dinner 2  
## 77 17.92 3.08 Male Yes Sat Dinner 2  
## 78 27.20 4.00 Male No Thur Lunch 4  
## 79 22.76 3.00 Male No Thur Lunch 2  
## 80 17.29 2.71 Male No Thur Lunch 2  
## 81 19.44 3.00 Male Yes Thur Lunch 2  
## 82 16.66 3.40 Male No Thur Lunch 2  
## 83 10.07 1.83 Female No Thur Lunch 1  
## 84 32.68 5.00 Male Yes Thur Lunch 2  
## 85 15.98 2.03 Male No Thur Lunch 2  
## 86 34.83 5.17 Female No Thur Lunch 4  
## 87 13.03 2.00 Male No Thur Lunch 2  
## 88 18.28 4.00 Male No Thur Lunch 2  
## 89 24.71 5.85 Male No Thur Lunch 2  
## 90 21.16 3.00 Male No Thur Lunch 2  
## 91 28.97 3.00 Male Yes Fri Dinner 2  
## 92 22.49 3.50 Male No Fri Dinner 2  
## 93 5.75 1.00 Female Yes Fri Dinner 2  
## 94 16.32 4.30 Female Yes Fri Dinner 2  
## 95 22.75 3.25 Female No Fri Dinner 2  
## 96 40.17 4.73 Male Yes Fri Dinner 4  
## 97 27.28 4.00 Male Yes Fri Dinner 2  
## 98 12.03 1.50 Male Yes Fri Dinner 2  
## 99 21.01 3.00 Male Yes Fri Dinner 2  
## 100 12.46 1.50 Male No Fri Dinner 2  
## 101 11.35 2.50 Female Yes Fri Dinner 2  
## 102 15.38 3.00 Female Yes Fri Dinner 2  
## 103 44.30 2.50 Female Yes Sat Dinner 3  
## 104 22.42 3.48 Female Yes Sat Dinner 2  
## 105 20.92 4.08 Female No Sat Dinner 2  
## 106 15.36 1.64 Male Yes Sat Dinner 2  
## 107 20.49 4.06 Male Yes Sat Dinner 2  
## 108 25.21 4.29 Male Yes Sat Dinner 2  
## 109 18.24 3.76 Male No Sat Dinner 2  
## 110 14.31 4.00 Female Yes Sat Dinner 2  
## 111 14.00 3.00 Male No Sat Dinner 2  
## 112 7.25 1.00 Female No Sat Dinner 1  
## 113 38.07 4.00 Male No Sun Dinner 3  
## 114 23.95 2.55 Male No Sun Dinner 2  
## 115 25.71 4.00 Female No Sun Dinner 3  
## 116 17.31 3.50 Female No Sun Dinner 2  
## 117 29.93 5.07 Male No Sun Dinner 4  
## 118 10.65 1.50 Female No Thur Lunch 2  
## 119 12.43 1.80 Female No Thur Lunch 2  
## 120 24.08 2.92 Female No Thur Lunch 4  
## 121 11.69 2.31 Male No Thur Lunch 2  
## 122 13.42 1.68 Female No Thur Lunch 2  
## 123 14.26 2.50 Male No Thur Lunch 2  
## 124 15.95 2.00 Male No Thur Lunch 2  
## 125 12.48 2.52 Female No Thur Lunch 2  
## 126 29.80 4.20 Female No Thur Lunch 6  
## 127 8.52 1.48 Male No Thur Lunch 2  
## 128 14.52 2.00 Female No Thur Lunch 2  
## 129 11.38 2.00 Female No Thur Lunch 2  
## 130 22.82 2.18 Male No Thur Lunch 3  
## 131 19.08 1.50 Male No Thur Lunch 2  
## 132 20.27 2.83 Female No Thur Lunch 2  
## 133 11.17 1.50 Female No Thur Lunch 2  
## 134 12.26 2.00 Female No Thur Lunch 2  
## 135 18.26 3.25 Female No Thur Lunch 2  
## 136 8.51 1.25 Female No Thur Lunch 2  
## 137 10.33 2.00 Female No Thur Lunch 2  
## 138 14.15 2.00 Female No Thur Lunch 2  
## 139 16.00 2.00 Male Yes Thur Lunch 2  
## 140 13.16 2.75 Female No Thur Lunch 2  
## 141 17.47 3.50 Female No Thur Lunch 2  
## 142 34.30 6.70 Male No Thur Lunch 6  
## 143 41.19 5.00 Male No Thur Lunch 5  
## 144 27.05 5.00 Female No Thur Lunch 6  
## 145 16.43 2.30 Female No Thur Lunch 2  
## 146 8.35 1.50 Female No Thur Lunch 2  
## 147 18.64 1.36 Female No Thur Lunch 3  
## 148 11.87 1.63 Female No Thur Lunch 2  
## 149 9.78 1.73 Male No Thur Lunch 2  
## 150 7.51 2.00 Male No Thur Lunch 2  
## 151 14.07 2.50 Male No Sun Dinner 2  
## 152 13.13 2.00 Male No Sun Dinner 2  
## 153 17.26 2.74 Male No Sun Dinner 3  
## 154 24.55 2.00 Male No Sun Dinner 4  
## 155 19.77 2.00 Male No Sun Dinner 4  
## 156 29.85 5.14 Female No Sun Dinner 5  
## 157 48.17 5.00 Male No Sun Dinner 6  
## 158 25.00 3.75 Female No Sun Dinner 4  
## 159 13.39 2.61 Female No Sun Dinner 2  
## 160 16.49 2.00 Male No Sun Dinner 4  
## 161 21.50 3.50 Male No Sun Dinner 4  
## 162 12.66 2.50 Male No Sun Dinner 2  
## 163 16.21 2.00 Female No Sun Dinner 3  
## 164 13.81 2.00 Male No Sun Dinner 2  
## 165 17.51 3.00 Female Yes Sun Dinner 2  
## 166 24.52 3.48 Male No Sun Dinner 3  
## 167 20.76 2.24 Male No Sun Dinner 2  
## 168 31.71 4.50 Male No Sun Dinner 4  
## 169 10.59 1.61 Female Yes Sat Dinner 2  
## 170 10.63 2.00 Female Yes Sat Dinner 2  
## 171 50.81 10.00 Male Yes Sat Dinner 3  
## 172 15.81 3.16 Male Yes Sat Dinner 2  
## 173 7.25 5.15 Male Yes Sun Dinner 2  
## 174 31.85 3.18 Male Yes Sun Dinner 2  
## 175 16.82 4.00 Male Yes Sun Dinner 2  
## 176 32.90 3.11 Male Yes Sun Dinner 2  
## 177 17.89 2.00 Male Yes Sun Dinner 2  
## 178 14.48 2.00 Male Yes Sun Dinner 2  
## 179 9.60 4.00 Female Yes Sun Dinner 2  
## 180 34.63 3.55 Male Yes Sun Dinner 2  
## 181 34.65 3.68 Male Yes Sun Dinner 4  
## 182 23.33 5.65 Male Yes Sun Dinner 2  
## 183 45.35 3.50 Male Yes Sun Dinner 3  
## 184 23.17 6.50 Male Yes Sun Dinner 4  
## 185 40.55 3.00 Male Yes Sun Dinner 2  
## 186 20.69 5.00 Male No Sun Dinner 5  
## 187 20.90 3.50 Female Yes Sun Dinner 3  
## 188 30.46 2.00 Male Yes Sun Dinner 5  
## 189 18.15 3.50 Female Yes Sun Dinner 3  
## 190 23.10 4.00 Male Yes Sun Dinner 3  
## 191 15.69 1.50 Male Yes Sun Dinner 2  
## 192 19.81 4.19 Female Yes Thur Lunch 2  
## 193 28.44 2.56 Male Yes Thur Lunch 2  
## 194 15.48 2.02 Male Yes Thur Lunch 2  
## 195 16.58 4.00 Male Yes Thur Lunch 2  
## 196 7.56 1.44 Male No Thur Lunch 2  
## 197 10.34 2.00 Male Yes Thur Lunch 2  
## 198 43.11 5.00 Female Yes Thur Lunch 4  
## 199 13.00 2.00 Female Yes Thur Lunch 2  
## 200 13.51 2.00 Male Yes Thur Lunch 2  
## 201 18.71 4.00 Male Yes Thur Lunch 3  
## 202 12.74 2.01 Female Yes Thur Lunch 2  
## 203 13.00 2.00 Female Yes Thur Lunch 2  
## 204 16.40 2.50 Female Yes Thur Lunch 2  
## 205 20.53 4.00 Male Yes Thur Lunch 4  
## 206 16.47 3.23 Female Yes Thur Lunch 3  
## 207 26.59 3.41 Male Yes Sat Dinner 3  
## 208 38.73 3.00 Male Yes Sat Dinner 4  
## 209 24.27 2.03 Male Yes Sat Dinner 2  
## 210 12.76 2.23 Female Yes Sat Dinner 2  
## 211 30.06 2.00 Male Yes Sat Dinner 3  
## 212 25.89 5.16 Male Yes Sat Dinner 4  
## 213 48.33 9.00 Male No Sat Dinner 4  
## 214 13.27 2.50 Female Yes Sat Dinner 2  
## 215 28.17 6.50 Female Yes Sat Dinner 3  
## 216 12.90 1.10 Female Yes Sat Dinner 2  
## 217 28.15 3.00 Male Yes Sat Dinner 5  
## 218 11.59 1.50 Male Yes Sat Dinner 2  
## 219 7.74 1.44 Male Yes Sat Dinner 2  
## 220 30.14 3.09 Female Yes Sat Dinner 4  
## 221 12.16 2.20 Male Yes Fri Lunch 2  
## 222 13.42 3.48 Female Yes Fri Lunch 2  
## 223 8.58 1.92 Male Yes Fri Lunch 1  
## 224 15.98 3.00 Female No Fri Lunch 3  
## 225 13.42 1.58 Male Yes Fri Lunch 2  
## 226 16.27 2.50 Female Yes Fri Lunch 2  
## 227 10.09 2.00 Female Yes Fri Lunch 2  
## 228 20.45 3.00 Male No Sat Dinner 4  
## 229 13.28 2.72 Male No Sat Dinner 2  
## 230 22.12 2.88 Female Yes Sat Dinner 2  
## 231 24.01 2.00 Male Yes Sat Dinner 4  
## 232 15.69 3.00 Male Yes Sat Dinner 3  
## 233 11.61 3.39 Male No Sat Dinner 2  
## 234 10.77 1.47 Male No Sat Dinner 2  
## 235 15.53 3.00 Male Yes Sat Dinner 2  
## 236 10.07 1.25 Male No Sat Dinner 2  
## 237 12.60 1.00 Male Yes Sat Dinner 2  
## 238 32.83 1.17 Male Yes Sat Dinner 2  
## 239 35.83 4.67 Female No Sat Dinner 3  
## 240 29.03 5.92 Male No Sat Dinner 3  
## 241 27.18 2.00 Female Yes Sat Dinner 2  
## 242 22.67 2.00 Male Yes Sat Dinner 2  
## 243 17.82 1.75 Male No Sat Dinner 2  
## 244 18.78 3.00 Female No Thur Dinner 2

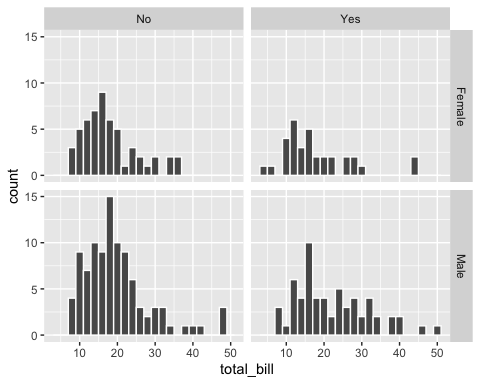
#view(tips)  
  
#Categorical variables are Fri,Sat, Sun,Thur from 'Day' column. Male and Female form 'Sex' ,  
#Lunch and Dinner from 'Time'.  
  
#10)Use r code to indicate how many levels exist for the factor day in the tips data frame  
#and determine the frequency of each level.  
  
tips%>%  
 group\_by(day)%>%  
 count()

## # A tibble: 4 × 2  
## # Groups: day [4]  
## day n  
## <fct> <int>  
## 1 Fri 19  
## 2 Sat 87  
## 3 Sun 76  
## 4 Thur 62

#11)Produce r code that will produce the following histogram from the tips data frame  
  
  
ggplot(data = tips)+  
geom\_histogram(mapping = aes(x=total\_bill),color='white',binwidth = 2)



#12)Write r code that will produce the following histograms from the tips data frame.  
  
ggplot(data = tips)+  
 geom\_histogram(mapping = aes(x=total\_bill), color = 'white',bins=2,binwidth = 2)+  
 facet\_grid(sex~smoker)



#13)Using the stringr::words data set along with str\_subset code,  
#produce R code that will show a 9 letter word that has the letter a in the middle.   
  
stringr::words

## [1] "a" "able" "about" "absolute" "accept"   
## [6] "account" "achieve" "across" "act" "active"   
## [11] "actual" "add" "address" "admit" "advertise"   
## [16] "affect" "afford" "after" "afternoon" "again"   
## [21] "against" "age" "agent" "ago" "agree"   
## [26] "air" "all" "allow" "almost" "along"   
## [31] "already" "alright" "also" "although" "always"   
## [36] "america" "amount" "and" "another" "answer"   
## [41] "any" "apart" "apparent" "appear" "apply"   
## [46] "appoint" "approach" "appropriate" "area" "argue"   
## [51] "arm" "around" "arrange" "art" "as"   
## [56] "ask" "associate" "assume" "at" "attend"   
## [61] "authority" "available" "aware" "away" "awful"   
## [66] "baby" "back" "bad" "bag" "balance"   
## [71] "ball" "bank" "bar" "base" "basis"   
## [76] "be" "bear" "beat" "beauty" "because"   
## [81] "become" "bed" "before" "begin" "behind"   
## [86] "believe" "benefit" "best" "bet" "between"   
## [91] "big" "bill" "birth" "bit" "black"   
## [96] "bloke" "blood" "blow" "blue" "board"   
## [101] "boat" "body" "book" "both" "bother"   
## [106] "bottle" "bottom" "box" "boy" "break"   
## [111] "brief" "brilliant" "bring" "britain" "brother"   
## [116] "budget" "build" "bus" "business" "busy"   
## [121] "but" "buy" "by" "cake" "call"   
## [126] "can" "car" "card" "care" "carry"   
## [131] "case" "cat" "catch" "cause" "cent"   
## [136] "centre" "certain" "chair" "chairman" "chance"   
## [141] "change" "chap" "character" "charge" "cheap"   
## [146] "check" "child" "choice" "choose" "Christ"   
## [151] "Christmas" "church" "city" "claim" "class"   
## [156] "clean" "clear" "client" "clock" "close"   
## [161] "closes" "clothe" "club" "coffee" "cold"   
## [166] "colleague" "collect" "college" "colour" "come"   
## [171] "comment" "commit" "committee" "common" "community"   
## [176] "company" "compare" "complete" "compute" "concern"   
## [181] "condition" "confer" "consider" "consult" "contact"   
## [186] "continue" "contract" "control" "converse" "cook"   
## [191] "copy" "corner" "correct" "cost" "could"   
## [196] "council" "count" "country" "county" "couple"   
## [201] "course" "court" "cover" "create" "cross"   
## [206] "cup" "current" "cut" "dad" "danger"   
## [211] "date" "day" "dead" "deal" "dear"   
## [216] "debate" "decide" "decision" "deep" "definite"   
## [221] "degree" "department" "depend" "describe" "design"   
## [226] "detail" "develop" "die" "difference" "difficult"   
## [231] "dinner" "direct" "discuss" "district" "divide"   
## [236] "do" "doctor" "document" "dog" "door"   
## [241] "double" "doubt" "down" "draw" "dress"   
## [246] "drink" "drive" "drop" "dry" "due"   
## [251] "during" "each" "early" "east" "easy"   
## [256] "eat" "economy" "educate" "effect" "egg"   
## [261] "eight" "either" "elect" "electric" "eleven"   
## [266] "else" "employ" "encourage" "end" "engine"   
## [271] "english" "enjoy" "enough" "enter" "environment"  
## [276] "equal" "especial" "europe" "even" "evening"   
## [281] "ever" "every" "evidence" "exact" "example"   
## [286] "except" "excuse" "exercise" "exist" "expect"   
## [291] "expense" "experience" "explain" "express" "extra"   
## [296] "eye" "face" "fact" "fair" "fall"   
## [301] "family" "far" "farm" "fast" "father"   
## [306] "favour" "feed" "feel" "few" "field"   
## [311] "fight" "figure" "file" "fill" "film"   
## [316] "final" "finance" "find" "fine" "finish"   
## [321] "fire" "first" "fish" "fit" "five"   
## [326] "flat" "floor" "fly" "follow" "food"   
## [331] "foot" "for" "force" "forget" "form"   
## [336] "fortune" "forward" "four" "france" "free"   
## [341] "friday" "friend" "from" "front" "full"   
## [346] "fun" "function" "fund" "further" "future"   
## [351] "game" "garden" "gas" "general" "germany"   
## [356] "get" "girl" "give" "glass" "go"   
## [361] "god" "good" "goodbye" "govern" "grand"   
## [366] "grant" "great" "green" "ground" "group"   
## [371] "grow" "guess" "guy" "hair" "half"   
## [376] "hall" "hand" "hang" "happen" "happy"   
## [381] "hard" "hate" "have" "he" "head"   
## [386] "health" "hear" "heart" "heat" "heavy"   
## [391] "hell" "help" "here" "high" "history"   
## [396] "hit" "hold" "holiday" "home" "honest"   
## [401] "hope" "horse" "hospital" "hot" "hour"   
## [406] "house" "how" "however" "hullo" "hundred"   
## [411] "husband" "idea" "identify" "if" "imagine"   
## [416] "important" "improve" "in" "include" "income"   
## [421] "increase" "indeed" "individual" "industry" "inform"   
## [426] "inside" "instead" "insure" "interest" "into"   
## [431] "introduce" "invest" "involve" "issue" "it"   
## [436] "item" "jesus" "job" "join" "judge"   
## [441] "jump" "just" "keep" "key" "kid"   
## [446] "kill" "kind" "king" "kitchen" "knock"   
## [451] "know" "labour" "lad" "lady" "land"   
## [456] "language" "large" "last" "late" "laugh"   
## [461] "law" "lay" "lead" "learn" "leave"   
## [466] "left" "leg" "less" "let" "letter"   
## [471] "level" "lie" "life" "light" "like"   
## [476] "likely" "limit" "line" "link" "list"   
## [481] "listen" "little" "live" "load" "local"   
## [486] "lock" "london" "long" "look" "lord"   
## [491] "lose" "lot" "love" "low" "luck"   
## [496] "lunch" "machine" "main" "major" "make"   
## [501] "man" "manage" "many" "mark" "market"   
## [506] "marry" "match" "matter" "may" "maybe"   
## [511] "mean" "meaning" "measure" "meet" "member"   
## [516] "mention" "middle" "might" "mile" "milk"   
## [521] "million" "mind" "minister" "minus" "minute"   
## [526] "miss" "mister" "moment" "monday" "money"   
## [531] "month" "more" "morning" "most" "mother"   
## [536] "motion" "move" "mrs" "much" "music"   
## [541] "must" "name" "nation" "nature" "near"   
## [546] "necessary" "need" "never" "new" "news"   
## [551] "next" "nice" "night" "nine" "no"   
## [556] "non" "none" "normal" "north" "not"   
## [561] "note" "notice" "now" "number" "obvious"   
## [566] "occasion" "odd" "of" "off" "offer"   
## [571] "office" "often" "okay" "old" "on"   
## [576] "once" "one" "only" "open" "operate"   
## [581] "opportunity" "oppose" "or" "order" "organize"   
## [586] "original" "other" "otherwise" "ought" "out"   
## [591] "over" "own" "pack" "page" "paint"   
## [596] "pair" "paper" "paragraph" "pardon" "parent"   
## [601] "park" "part" "particular" "party" "pass"   
## [606] "past" "pay" "pence" "pension" "people"   
## [611] "per" "percent" "perfect" "perhaps" "period"   
## [616] "person" "photograph" "pick" "picture" "piece"   
## [621] "place" "plan" "play" "please" "plus"   
## [626] "point" "police" "policy" "politic" "poor"   
## [631] "position" "positive" "possible" "post" "pound"   
## [636] "power" "practise" "prepare" "present" "press"   
## [641] "pressure" "presume" "pretty" "previous" "price"   
## [646] "print" "private" "probable" "problem" "proceed"   
## [651] "process" "produce" "product" "programme" "project"   
## [656] "proper" "propose" "protect" "provide" "public"   
## [661] "pull" "purpose" "push" "put" "quality"   
## [666] "quarter" "question" "quick" "quid" "quiet"   
## [671] "quite" "radio" "rail" "raise" "range"   
## [676] "rate" "rather" "read" "ready" "real"   
## [681] "realise" "really" "reason" "receive" "recent"   
## [686] "reckon" "recognize" "recommend" "record" "red"   
## [691] "reduce" "refer" "regard" "region" "relation"   
## [696] "remember" "report" "represent" "require" "research"   
## [701] "resource" "respect" "responsible" "rest" "result"   
## [706] "return" "rid" "right" "ring" "rise"   
## [711] "road" "role" "roll" "room" "round"   
## [716] "rule" "run" "safe" "sale" "same"   
## [721] "saturday" "save" "say" "scheme" "school"   
## [726] "science" "score" "scotland" "seat" "second"   
## [731] "secretary" "section" "secure" "see" "seem"   
## [736] "self" "sell" "send" "sense" "separate"   
## [741] "serious" "serve" "service" "set" "settle"   
## [746] "seven" "sex" "shall" "share" "she"   
## [751] "sheet" "shoe" "shoot" "shop" "short"   
## [756] "should" "show" "shut" "sick" "side"   
## [761] "sign" "similar" "simple" "since" "sing"   
## [766] "single" "sir" "sister" "sit" "site"   
## [771] "situate" "six" "size" "sleep" "slight"   
## [776] "slow" "small" "smoke" "so" "social"   
## [781] "society" "some" "son" "soon" "sorry"   
## [786] "sort" "sound" "south" "space" "speak"   
## [791] "special" "specific" "speed" "spell" "spend"   
## [796] "square" "staff" "stage" "stairs" "stand"   
## [801] "standard" "start" "state" "station" "stay"   
## [806] "step" "stick" "still" "stop" "story"   
## [811] "straight" "strategy" "street" "strike" "strong"   
## [816] "structure" "student" "study" "stuff" "stupid"   
## [821] "subject" "succeed" "such" "sudden" "suggest"   
## [826] "suit" "summer" "sun" "sunday" "supply"   
## [831] "support" "suppose" "sure" "surprise" "switch"   
## [836] "system" "table" "take" "talk" "tape"   
## [841] "tax" "tea" "teach" "team" "telephone"   
## [846] "television" "tell" "ten" "tend" "term"   
## [851] "terrible" "test" "than" "thank" "the"   
## [856] "then" "there" "therefore" "they" "thing"   
## [861] "think" "thirteen" "thirty" "this" "thou"   
## [866] "though" "thousand" "three" "through" "throw"   
## [871] "thursday" "tie" "time" "to" "today"   
## [876] "together" "tomorrow" "tonight" "too" "top"   
## [881] "total" "touch" "toward" "town" "trade"   
## [886] "traffic" "train" "transport" "travel" "treat"   
## [891] "tree" "trouble" "true" "trust" "try"   
## [896] "tuesday" "turn" "twelve" "twenty" "two"   
## [901] "type" "under" "understand" "union" "unit"   
## [906] "unite" "university" "unless" "until" "up"   
## [911] "upon" "use" "usual" "value" "various"   
## [916] "very" "video" "view" "village" "visit"   
## [921] "vote" "wage" "wait" "walk" "wall"   
## [926] "want" "war" "warm" "wash" "waste"   
## [931] "watch" "water" "way" "we" "wear"   
## [936] "wednesday" "wee" "week" "weigh" "welcome"   
## [941] "well" "west" "what" "when" "where"   
## [946] "whether" "which" "while" "white" "who"   
## [951] "whole" "why" "wide" "wife" "will"   
## [956] "win" "wind" "window" "wish" "with"   
## [961] "within" "without" "woman" "wonder" "wood"   
## [966] "word" "work" "world" "worry" "worse"   
## [971] "worth" "would" "write" "wrong" "year"   
## [976] "yes" "yesterday" "yet" "you" "young"

str\_subset(words, "....a....")

## [1] "character"

#14)Produce a string that will force a match for the regular expression \\””\   
#Use and show the R code command writelines to confirm your answer  
  
writeLines(" \\\\””\\\ ")

## \\””\

#15)Describe in words (two or three sentences) what the following regular expression   
#will match ^.\*e$.  
  
str\_subset(words, "^.\*e$")

## [1] "able" "absolute" "achieve" "active" "advertise"   
## [6] "age" "agree" "appropriate" "argue" "arrange"   
## [11] "associate" "assume" "available" "aware" "balance"   
## [16] "base" "be" "because" "become" "before"   
## [21] "believe" "bloke" "blue" "bottle" "cake"   
## [26] "care" "case" "cause" "centre" "chance"   
## [31] "change" "charge" "choice" "choose" "close"   
## [36] "clothe" "coffee" "colleague" "college" "come"   
## [41] "committee" "compare" "complete" "compute" "continue"   
## [46] "converse" "couple" "course" "create" "date"   
## [51] "debate" "decide" "definite" "degree" "describe"   
## [56] "die" "difference" "divide" "double" "drive"   
## [61] "due" "educate" "else" "encourage" "engine"   
## [66] "europe" "evidence" "example" "excuse" "exercise"   
## [71] "expense" "experience" "eye" "face" "figure"   
## [76] "file" "finance" "fine" "fire" "five"   
## [81] "force" "fortune" "france" "free" "future"   
## [86] "game" "give" "goodbye" "hate" "have"   
## [91] "he" "here" "home" "hope" "horse"   
## [96] "house" "imagine" "improve" "include" "income"   
## [101] "increase" "inside" "insure" "introduce" "involve"   
## [106] "issue" "judge" "language" "large" "late"   
## [111] "leave" "lie" "life" "like" "line"   
## [116] "little" "live" "lose" "love" "machine"   
## [121] "make" "manage" "maybe" "measure" "middle"   
## [126] "mile" "minute" "more" "move" "name"   
## [131] "nature" "nice" "nine" "none" "note"   
## [136] "notice" "office" "once" "one" "operate"   
## [141] "oppose" "organize" "otherwise" "page" "pence"   
## [146] "people" "picture" "piece" "place" "please"   
## [151] "police" "positive" "possible" "practise" "prepare"   
## [156] "pressure" "presume" "price" "private" "probable"   
## [161] "produce" "programme" "propose" "provide" "purpose"   
## [166] "quite" "raise" "range" "rate" "realise"   
## [171] "receive" "recognize" "reduce" "require" "resource"   
## [176] "responsible" "rise" "role" "rule" "safe"   
## [181] "sale" "same" "save" "scheme" "science"   
## [186] "score" "secure" "see" "sense" "separate"   
## [191] "serve" "service" "settle" "share" "she"   
## [196] "shoe" "side" "simple" "since" "single"   
## [201] "site" "situate" "size" "smoke" "some"   
## [206] "space" "square" "stage" "state" "strike"   
## [211] "structure" "suppose" "sure" "surprise" "table"   
## [216] "take" "tape" "telephone" "terrible" "the"   
## [221] "there" "therefore" "three" "tie" "time"   
## [226] "trade" "tree" "trouble" "true" "twelve"   
## [231] "type" "unite" "use" "value" "village"   
## [236] "vote" "wage" "waste" "we" "wee"   
## [241] "welcome" "where" "while" "white" "whole"   
## [246] "wide" "wife" "worse" "write"

#The following above regular expression is fetching all the words that ends with "e"   
#irrespective of the length of the string.   
  
#16)Using the methods demonstrated in class regarding Factors and Forcats,   
#use and show R code to create a factor that will enable you to sort the string vector   
#(“eight”, “four”, “ten”, “two”)  
#according to quantity, not alphabetical order.   
  
x <- c("eight", "four", "ten", "two")  
sort(x)

## [1] "eight" "four" "ten" "two"

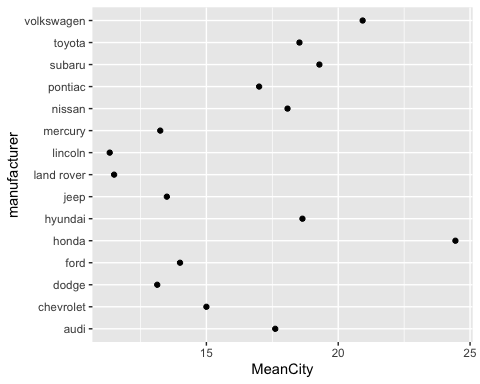
SortQunatity= c("two", "four", "eight", "ten")  
x1 = factor(x,levels = SortQunatity)  
sort(x1)

## [1] two four eight ten   
## Levels: two four eight ten

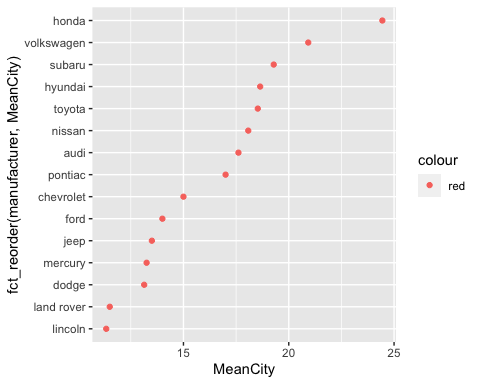
#17)Using the mpg data set, use and show R code that will produce a table that shows the  
#average city mileage (mean for cty) for each manufacturer.  
  
#view(mpg)  
  
mpgdata <- mpg %>%  
 group\_by(manufacturer) %>%  
 summarise(MeanCity = mean(cty))  
mpgdata

## # A tibble: 15 × 2  
## manufacturer MeanCity  
## <chr> <dbl>  
## 1 audi 17.6  
## 2 chevrolet 15   
## 3 dodge 13.1  
## 4 ford 14   
## 5 honda 24.4  
## 6 hyundai 18.6  
## 7 jeep 13.5  
## 8 land rover 11.5  
## 9 lincoln 11.3  
## 10 mercury 13.2  
## 11 nissan 18.1  
## 12 pontiac 17   
## 13 subaru 19.3  
## 14 toyota 18.5  
## 15 volkswagen 20.9

#18)Using the mpg data set and methods and code illustrated in the emailed Factors and Forcats R file:  
#use and show R code to generate a scatter plot that clearly shows how different manufacturers compare   
#with one another with regards to average or mean city mileage (cty)   
  
ggplot(data = mpgdata)+  
 geom\_point(mapping = aes(x=MeanCity, y = manufacturer))



ggplot(data = mpgdata) +  
 geom\_point(mapping = aes(x=MeanCity,  
 y = fct\_reorder(manufacturer,MeanCity),color='red'))



#19)To the mpg data table, apply a stringr function to print all observations of the  
#manufacturer variable in upper case letters. Thenuse R coding to produce rows 15 to 25.   
#The first five rows of the table are shown below.   
#Note that the row numbers 1 – 5 correspond to rows 15 – 20.  
  
mpg$manufacturer<-str\_to\_upper(mpg$manufacturer)  
mpg

## # A tibble: 234 × 11  
## manufacturer model displ year cyl trans drv cty hwy fl class  
## <chr> <chr> <dbl> <int> <int> <chr> <chr> <int> <int> <chr> <chr>  
## 1 AUDI a4 1.8 1999 4 auto… f 18 29 p comp…  
## 2 AUDI a4 1.8 1999 4 manu… f 21 29 p comp…  
## 3 AUDI a4 2 2008 4 manu… f 20 31 p comp…  
## 4 AUDI a4 2 2008 4 auto… f 21 30 p comp…  
## 5 AUDI a4 2.8 1999 6 auto… f 16 26 p comp…  
## 6 AUDI a4 2.8 1999 6 manu… f 18 26 p comp…  
## 7 AUDI a4 3.1 2008 6 auto… f 18 27 p comp…  
## 8 AUDI a4 quattro 1.8 1999 4 manu… 4 18 26 p comp…  
## 9 AUDI a4 quattro 1.8 1999 4 auto… 4 16 25 p comp…  
## 10 AUDI a4 quattro 2 2008 4 manu… 4 20 28 p comp…  
## # … with 224 more rows

mpgdata2<- mpg%>%  
 select(manufacturer, model, year)%>%  
 slice(15:25)  
  
mpgdata2

## # A tibble: 11 × 3  
## manufacturer model year  
## <chr> <chr> <int>  
## 1 AUDI a4 quattro 2008  
## 2 AUDI a6 quattro 1999  
## 3 AUDI a6 quattro 2008  
## 4 AUDI a6 quattro 2008  
## 5 CHEVROLET c1500 suburban 2wd 2008  
## 6 CHEVROLET c1500 suburban 2wd 2008  
## 7 CHEVROLET c1500 suburban 2wd 2008  
## 8 CHEVROLET c1500 suburban 2wd 1999  
## 9 CHEVROLET c1500 suburban 2wd 2008  
## 10 CHEVROLET corvette 1999  
## 11 CHEVROLET corvette 1999

#20)To the diamonds data table, apply R code to produce the table give below.  
#Note that the variable name color has been changed to Color  
  
diamonds

## # A tibble: 53,940 × 10  
## carat cut color clarity depth table price x y z  
## <dbl> <ord> <ord> <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>  
## 1 0.23 Ideal E SI2 61.5 55 326 3.95 3.98 2.43  
## 2 0.21 Premium E SI1 59.8 61 326 3.89 3.84 2.31  
## 3 0.23 Good E VS1 56.9 65 327 4.05 4.07 2.31  
## 4 0.29 Premium I VS2 62.4 58 334 4.2 4.23 2.63  
## 5 0.31 Good J SI2 63.3 58 335 4.34 4.35 2.75  
## 6 0.24 Very Good J VVS2 62.8 57 336 3.94 3.96 2.48  
## 7 0.24 Very Good I VVS1 62.3 57 336 3.95 3.98 2.47  
## 8 0.26 Very Good H SI1 61.9 55 337 4.07 4.11 2.53  
## 9 0.22 Fair E VS2 65.1 61 337 3.87 3.78 2.49  
## 10 0.23 Very Good H VS1 59.4 61 338 4 4.05 2.39  
## # … with 53,930 more rows

diamond1 <- diamonds%>%  
 group\_by(color)%>%  
 count()%>%  
 rename(Color = 'color', Frequency = 'n')%>%  
 arrange(desc(Frequency))  
  
diamond1

## # A tibble: 7 × 2  
## # Groups: Color [7]  
## Color Frequency  
## <ord> <int>  
## 1 G 11292  
## 2 E 9797  
## 3 F 9542  
## 4 H 8304  
## 5 D 6775  
## 6 I 5422  
## 7 J 2808

#21) Now using the table produced in problem 20 and ggplot coding, produce the bar graph shown below.  
  
  
ggplot(data=diamond1)+  
 geom\_bar(mapping = aes(y = Frequency,x=Color,fill=Color),stat="identity")

