```
install.packages("tidyverse")
library(tidyverse)
install.packages("dplyr")
library("dplyr")
```

```
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.4     ✔ readr     2.1.4
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.4.4     ✔ tibble    3.2.1
✔ lubridate 1.9.3     ✔ tidyr     1.3.0
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ⓘ Use the conflicted package (<http://conflicted.r-lib.org/>) to
force all conflicts to become errors
Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)
```

```
mydata <-read.csv("StudentsPerformance.csv")
head(mydata)
summary(mydata)
glimpse(mydata)
```

```
  gender race.ethnicity parental.level.of.education lunch
1 female group B        bachelor's degree           standard
2 female group C        some college                standard
3 female group B        master's degree             standard
4 male   group A        associate's degree          free/reduced
5 male   group C        some college                standard
6 female group B        associate's degree          standard
  test.preparation.course math.score reading.score writing.score
1 none                      72            72            74
2 completed                 69            90            88
3 none                      90            95            93
4 none                      47            57            44
5 none                      76            78            75
6 none                      71            83            78
```

```
    gender           race.ethnicity     parental.level.of.education
 Length:1000        Length:1000        Length:1000
 Class :character   Class :character   Class :character
 Mode  :character   Mode  :character   Mode  :character
```

```
    lunch           test.preparation.course   math.score
reading.score
 Length:1000        Length:1000               Min.   :  0.00   Min.   :
17.00
 Class :character   Class :character          1st Qu.: 57.00   1st Qu.:
59.00
 Mode  :character   Mode  :character          Median : 66.00   Median :
70.00
                                              Mean   : 66.09   Mean   :
69.17
                                              3rd Qu.: 77.00   3rd Qu.:
79.00
                                              Max.   :100.00
Max.   :100.00
 writing.score
 Min.   : 10.00
 1st Qu.: 57.75
 Median : 69.00
 Mean   : 68.05
 3rd Qu.: 79.00
 Max.   :100.00

Rows: 1,000
Columns: 8
$ gender                     <chr> "female", "female", "female",
"male", "mal…
$ race.ethnicity             <chr> "group B", "group C", "group B",
"group A"…
$ parental.level.of.education <chr> "bachelor's degree", "some
college", "mast…
$ lunch                      <chr> "standard", "standard",
"standard", "free/…
$ test.preparation.course    <chr> "none", "completed", "none",
"none", "none…
$ math.score                 <int> 72, 69, 90, 47, 76, 71, 88, 40,
64, 38, 58…
$ reading.score              <int> 72, 90, 95, 57, 78, 83, 95, 43,
64, 60, 54…
$ writing.score              <int> 74, 88, 93, 44, 75, 78, 92, 39,
67, 50, 52…
```

```r
sum(is.na(mydata))
cdata <-na.omit(mydata)
summary(cdata)
```

```
[1] 0
```

```
   gender           race.ethnicity     parental.level.of.education
 Length:1000        Length:1000        Length:1000
 Class :character   Class :character   Class :character
 Mode  :character   Mode  :character   Mode  :character



   lunch            test.preparation.course   math.score
reading.score
 Length:1000        Length:1000               Min.   :  0.00   Min.   :
17.00
 Class :character   Class :character          1st Qu.: 57.00   1st Qu.:
59.00
 Mode  :character   Mode  :character          Median : 66.00   Median :
70.00
                                              Mean   : 66.09   Mean   :
69.17
                                              3rd Qu.: 77.00   3rd Qu.:
79.00
                                              Max.   :100.00
Max.   :100.00
 writing.score
 Min.   : 10.00
 1st Qu.: 57.75
 Median : 69.00
 Mean   : 68.05
 3rd Qu.: 79.00
 Max.   :100.00
```

```
x <-subset(cdata, select = c("parental.level.of.education"))
x
```

```
     parental.level.of.education
1    bachelor's degree
2    some college
3    master's degree
4    associate's degree
5    some college
6    associate's degree
7    some college
8    some college
9    high school
10   high school
11   associate's degree
12   associate's degree
13   high school
14   some college
15   master's degree
16   some high school
17   high school
```

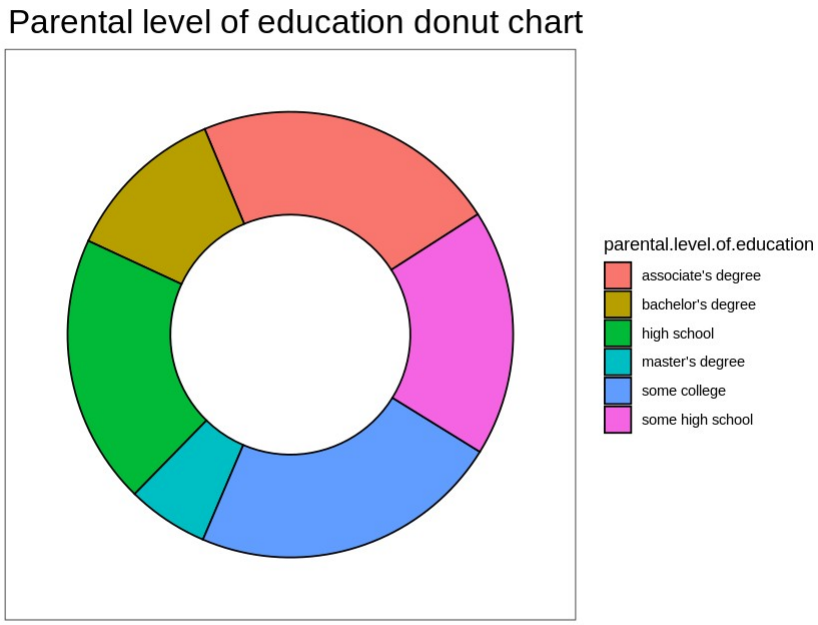
```
18   some high school
19   master's degree
20   associate's degree
21   high school
22   some college
23   some college
24   some high school
25   bachelor's degree
26   master's degree
27   some college
28   bachelor's degree
29   high school
30   master's degree
⋮    ⋮
971  bachelor's degree
972  some high school
973  high school
974  some college
975  some college
976  some college
977  some college
978  associate's degree
979  high school
980  associate's degree
981  high school
982  some high school
983  some high school
984  some college
985  some high school
986  high school
987  associate's degree
988  some high school
989  some high school
990  some college
991  high school
992  some high school
993  associate's degree
994  bachelor's degree
995  high school
996  master's degree
997  high school
998  high school
999  some college
1000 some college
```

```
count<-as.data.frame(table(x))
count
```

```
  parental.level.of.education Freq
1 associate's degree          222
```

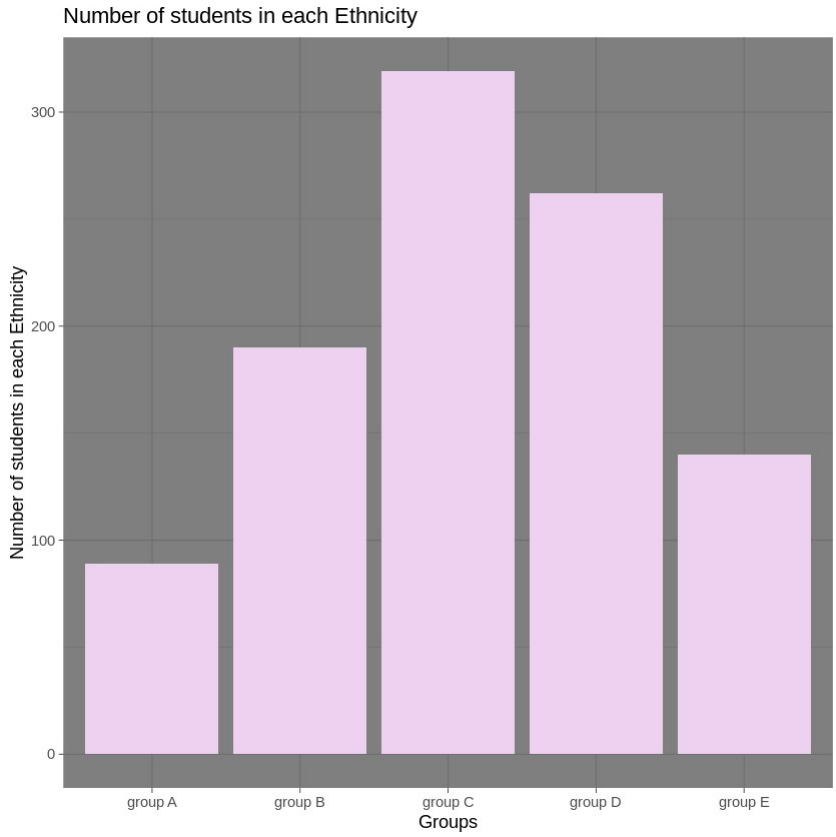
```
2 bachelor's degree           118
3 high school                 196
4 master's degree              59
5 some college                226
6 some high school            179
```

```
p = ggplot(count, aes(x=2, y=Freq, fill=parental.level.of.education))
+
  theme_bw()+
  geom_col(color="black")+
  coord_polar("y", start = 1) +
  theme(panel.background = element_blank(),
        axis.line = element_blank(),
        axis.title = element_blank(),
        axis.text = element_blank(),
        axis.ticks = element_blank(),
        panel.grid=element_blank(),
        plot.title = element_text(hjust = 0.5, size = 20)) +
  ggtitle(" Parental level of education donut chart") +
  xlim(0.5,2.5)
p
```

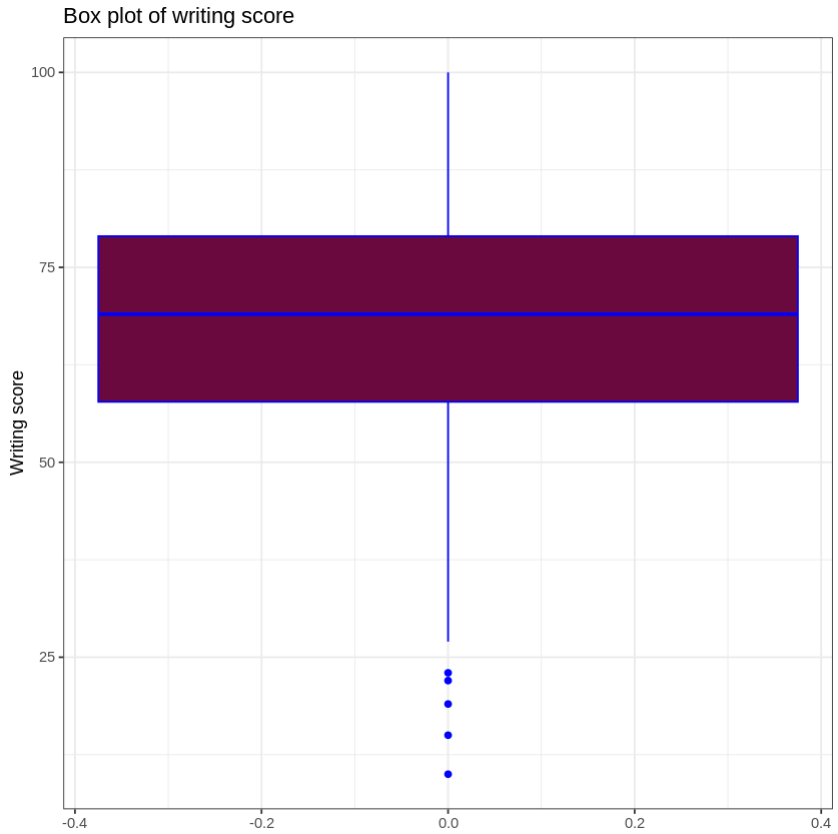



```
a <- ggplot(cdata, aes(x = race.ethnicity)) +
  geom_bar(fill="#eed1f0")+
```

```
  labs(x = "Groups", y = "Number of students in each Ethnicity", title
=
         "Number of students in each Ethnicity") +
  theme_dark()
a
```




```
ggplot(cdata,aes(y=writing.score))+
  geom_boxplot(color="blue",fill="#69093e")+
  theme_bw()+
  labs(y="Writing score",title="Box plot of writing score")
```




```
total <- cdata$math.score +cdata$reading.score + cdata$writing.score
total
```

```
  [1] 218 247 278 148 229 232 275 122 195 148 164 135 219 220 161 222
263  78
 [19] 134 173 198 210 151 215 225 219 178 211 205 207 217 189 193 120
266 241
 [37] 238 173 253 170 170 199 176 190 160 176 182 213 207 248 156 214
139 241
 [55] 242 117 253 156 175  27 225 112 178 220 176 192 119 206 175 160
194 181
 [73] 140 174 139 123  78 239 207 198 146 139 151 199 115 235 247 216
195 241
 [91] 211  97 221 138 257 241 193 205 187 194 213 216 265 151 274 197
287 193
[109] 198 206 264 166 154 163 299 235 242 239 175 189 260 272 274 176
236 268
[127] 207 216 238 151 255 130 231 240 228 161 146 181 192 201 219 185
151 147
[145] 239  94 247 207 224 293 198 212 195 150 190 247 218 180 234 192
234 248
[163] 154 154 264 296 155 212 229 212 219 260 246 200 148 259 158 225
184 297
[181] 205 176 147 204 146 197 221 196 145 252 200 238 192 201 238 168
```

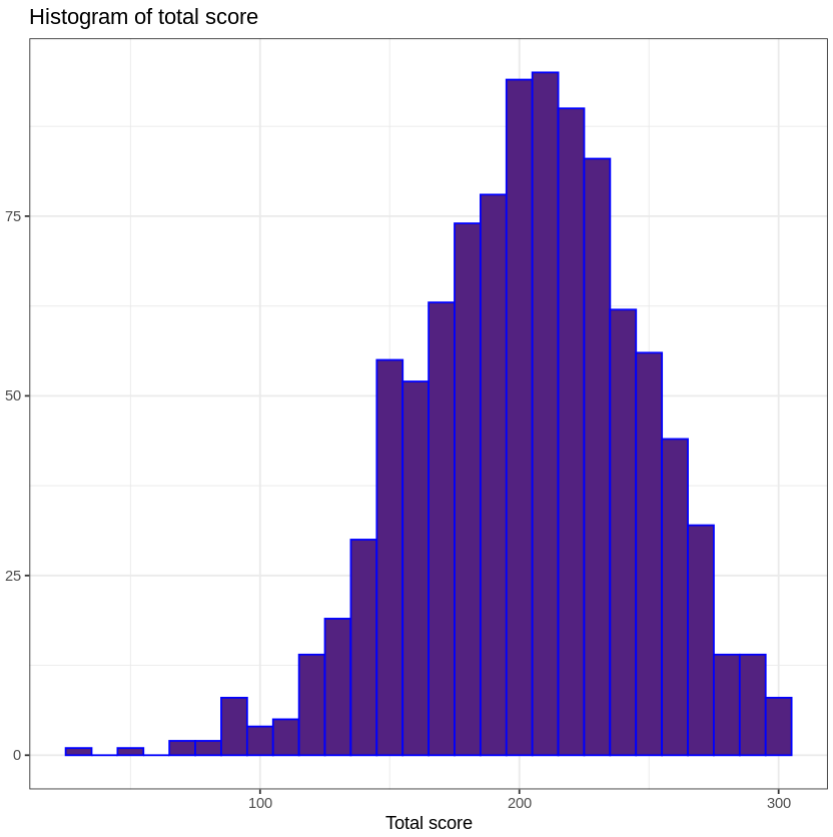
```
193 162
 [199] 153 233 237 223 215 194 142 223 206 237 231 185 238  90 164 167
259 242
 [217] 257 115 213 173 198 245 182 241 191 155 215 161 204 277 206 131
236 257
 [235] 249 232 196 198 157 244 207 246 163 153 218 232 239 196 191 181
145 237
 [253] 200 232 158 205 205 232 223 227 220 236 145 282 217 142 219 218
275 217
 [271] 193 149 158 206 249 234 270 197 231 186 147 139 236 226 129 245
267 238
 [289] 239 230 214 229 180 213 235 189 130 213 136 252 240 162 232 220
232 205
 [307] 267 147 191 158 229 183 196 204 198 202 250 233 203 184 230 240
232 149
 [325] 130 266 184  70 234 147 192 119 171 249 271 195 215 143  89 150
171 197
 [343] 219 206 219 227 180 266 215 237 186 175 221 161 200 203 185 177
182 259
 [361] 179 247 168  93 177 167 180 188 166 242 232 188 220 275 202 130
252 280
 [379] 231 192 249 282 231 132 124 218 207 202 190 218 199 222 210 167
216 134
 [397] 180 269 194 173 180 139 169 287 175 204 193 239 165 256 215 245
197 197
 [415] 202 226 213 214 189 167 268 179 172 224 114 191 255 205 177 185
189 203
 [433] 172 156 205 151 207 214 177 259 227 239 201 232 234 218 195 247
135 226
 [451] 234 289 216 172 176 140 257 155 300 205 166 153 217 258 195 262
95 204
 [469] 250 241 258 206 238 187 271 230 204 220 189 214 212 167 197 157
155 215
 [487] 164 209 174 243 226 198 256 259 186 198 154 213 203 219 235 273
178 276
 [505] 153 244 204 168 234 265 220 157 163 177 268 248 247 213 222 224
172 261
 [523] 184 163 154 187 177 132 117 206 208 154 181 261 210 232 169 179
222 275
 [541] 212 241 237 266 193 241 289 203 193 220 220 265 138 203 138 104
206 194
 [559] 193 201 228 211 278 223 145 136 292 210 158 203 216 279 171 180
226 148
 [577] 164 261 162 182 274 236 214 228 223 217 201 169 183 179 169 183
200 223
 [595] 291 207  69 155 219 228 177  88 234 168 263 209 251 152 187 186
192 183
 [613] 275 216 268 200 120 249 260 206 150 181 161 282 181 296 202 155
152 150
```

```
 [631] 201 204 221 248 248 209 228 260 236 223 144 252 232 233 195 227
194 194
 [649] 158 229 160 224 269 209 231 196 208 196 161 262 224 202 195 201
206 170
 [667] 207 249 208 230 210 159 223 233 224 180 227 228 234 184 213 188
171 113
 [685] 196 293 228 228 155 266 155 227 221 227 168 254 247 201 214 191
242 196
 [703] 258 194 186 199 116 177 255 161 267 250 297 246 204 257 223 291
226 244
 [721] 204 173 252 140 131 226 217 176 249 129 225 144 270 146 168 183
255 192
 [739] 225 176 225 150 247 204 157 225 215 201 169 257 208 205 231 252
166 271
 [757] 165 210 176 223 189 160 245 187 182 218 210 198 214 175 149 222
191 219
 [775] 191 162 208 122 233 261 159 251 243 173 251 127 228  89 194 178
170 196
 [793] 209 250 145 198 210 235 183 164 208 229 276 254 227 230 213 126
234 164
 [811] 118 141 185 247 235 267 168 207 200 256 267 261 152 240 158 187
194 210
 [829] 236 175 178 249 182 248 153 198 194 236 168 216 137 173 103 227
169 259
 [847] 261 168 199 224 205 192 265 210 181 290 208 225 147 232 168 254
119 222
 [865] 281 252 164 144 224 151 164 218 244 262 166 210 247 172 211 214
193 211
 [883] 208 145 156 218 288 185 208 143 262 262 190 237 190 104  92 220
184 225
 [901] 258 225 123 293 235 242 156 254 214 204 140 235 179 175 162 204
300 157
 [919] 214 278 206 134 204 184 213 198 172 197 133 155 211 181 205 219
275 191
 [937] 177 172 251 214 210 265 211 184 187 173 244 167 151 205 238 235
220 164
 [955] 200 191 252 292 167 227 178 155 300 216 187 204 198 186 211 239
289 219
 [973] 163 175 184 223 182 185 144 280  55 237 250 256 231 162 150 232
134 236
 [991] 242 225 207 208 188 282 172 195 223 249
```

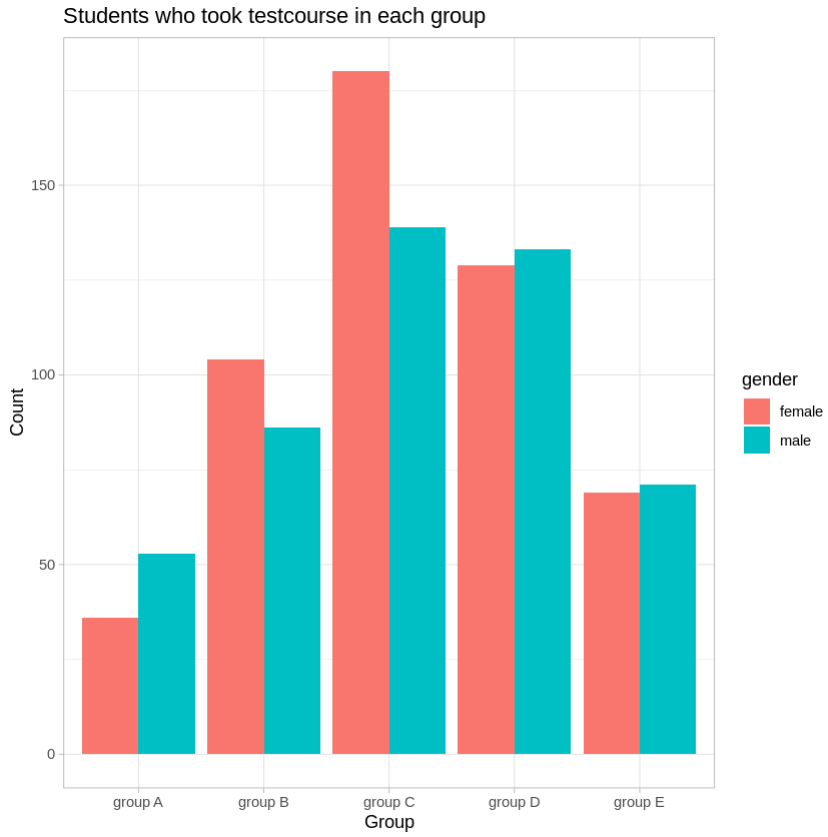
```
ggplot(cdata,aes(cdata$math.score +cdata$reading.score +
cdata$writing.score))+
  geom_histogram(binwidth=10,color="blue", fill="#532280")+
  theme_bw()+
  labs(x="Total score",y=NULL,title="Histogram of total score")
```

```
Warning message:
“Use of `cdata$math.score` is discouraged.
```

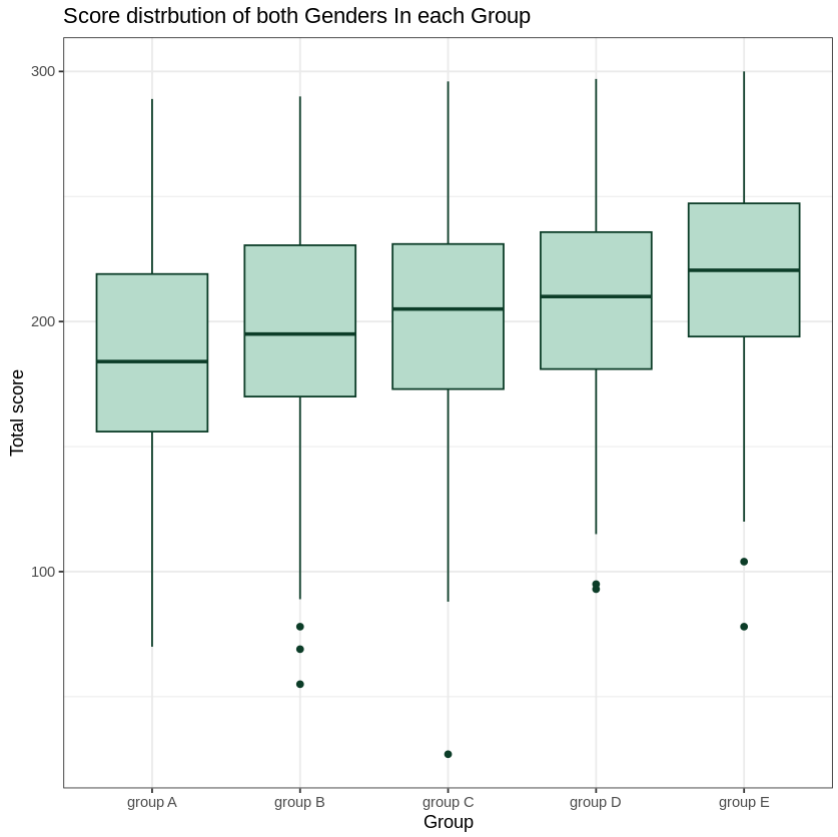
```
ⓘ Use `math.score` instead.”
Warning message:
“Use of `cdata$reading.score` is discouraged.
ⓘ Use `reading.score` instead.”
Warning message:
“Use of `cdata$writing.score` is discouraged.
ⓘ Use `writing.score` instead.”
```




```
ggplot(cdata,aes(race.ethnicity,fill=gender))+
  geom_bar(position="dodge")+
  theme_light()+
  labs(x="Group",y="Count",title="Students who took testcourse in each group")
```




```
ggplot(cdata,aes(race.ethnicity,total))+
  geom_boxplot(color="#0c3d28",fill="#b6dbcb")+
  theme_bw()+
  labs(x="Group",y="Total score",title="Score distrbution of both
Genders In each Group")
```




```
ndata <-cbind(cdata,total)
ndata
mat <-subset(ndata, select =

c("math.score","reading.score","writing.score","total"))
mat
```

	gender	race.ethnicity	parental.level.of.education	lunch
1	female	group B	bachelor's degree	standard
2	female	group C	some college	standard
3	female	group B	master's degree	standard
4	male	group A	associate's degree	free/reduced
5	male	group C	some college	standard
6	female	group B	associate's degree	standard
7	female	group B	some college	standard
8	male	group B	some college	free/reduced
9	male	group D	high school	free/reduced
10	female	group B	high school	free/reduced
11	male	group C	associate's degree	standard
12	male	group D	associate's degree	standard
13	female	group B	high school	standard
14	male	group A	some college	standard
15	female	group A	master's degree	standard
16	female	group C	some high school	standard

```
17   male   group C        high school                  standard
18   female group B        some high school             free/reduced
19   male   group C        master's degree              free/reduced
20   female group C        associate's degree           free/reduced
21   male   group D        high school                  standard
22   female group B        some college                 free/reduced
23   male   group D        some college                 standard
24   female group C        some high school             standard
25   male   group D        bachelor's degree            free/reduced
26   male   group A        master's degree              free/reduced
27   male   group B        some college                 standard
28   female group C        bachelor's degree            standard
29   male   group C        high school                  standard
30   female group D        master's degree              standard
⋮    ⋮      ⋮              ⋮                            ⋮
971  female group D        bachelor's degree            standard
972  male   group C        some high school             standard
973  female group A        high school                  free/reduced
974  female group D        some college                 free/reduced
975  female group A        some college                 standard
976  female group C        some college                 standard
977  male   group B        some college                 free/reduced
978  male   group C        associate's degree           standard
979  male   group D        high school                  standard
980  female group C        associate's degree           standard
981  female group B        high school                  free/reduced
982  male   group D        some high school             standard
983  male   group B        some high school             standard
984  female group A        some college                 standard
985  female group C        some high school             standard
986  male   group A        high school                  standard
987  female group C        associate's degree           standard
988  male   group E        some high school             standard
989  female group A        some high school             free/reduced
990  female group D        some college                 free/reduced
991  male   group E        high school                  free/reduced
992  female group B        some high school             standard
993  female group D        associate's degree           free/reduced
994  female group D        bachelor's degree            free/reduced
995  male   group A        high school                  standard
996  female group E        master's degree              standard
997  male   group C        high school                  free/reduced
998  female group C        high school                  free/reduced
999  female group D        some college                 standard
1000 female group D        some college                 free/reduced
     test.preparation.course math.score reading.score writing.score
total
1    none                     72         72            74
218
```

```
2    completed           69          90          88
247
3    none                90          95          93
278
4    none                47          57          44
148
5    none                76          78          75
229
6    none                71          83          78
232
7    completed           88          95          92
275
8    none                40          43          39
122
9    completed           64          64          67
195
10   none                38          60          50
148
11   none                58          54          52
164
12   none                40          52          43
135
13   none                65          81          73
219
14   completed           78          72          70
220
15   none                50          53          58
161
16   none                69          75          78
222
17   none                88          89          86
263
18   none                18          32          28
78
19   completed           46          42          46
134
20   none                54          58          61
173
21   none                66          69          63
198
22   completed           65          75          70
210
23   none                44          54          53
151
24   none                69          73          73
215
25   completed           74          71          80
225
26   none                73          74          72
```

219					
27	none	69	54	55	
178					
28	none	67	69	75	
211					
29	none	70	70	65	
205					
30	none	62	70	75	
207					
⋮	⋮	⋮	⋮	⋮	⋮
971	none	89	100	100	
289					
972	completed	78	72	69	
219					
973	completed	53	50	60	
163					
974	none	49	65	61	
175					
975	none	54	63	67	
184					
976	completed	64	82	77	
223					
977	completed	60	62	60	
182					
978	none	62	65	58	
185					
979	completed	55	41	48	
144					
980	none	91	95	94	
280					
981	none	8	24	23	
55					
982	none	81	78	78	
237					
983	completed	79	85	86	
250					
984	completed	78	87	91	
256					
985	none	74	75	82	
231					
986	none	57	51	54	
162					
987	none	40	59	51	
150					
988	completed	81	75	76	
232					
989	none	44	45	45	
134					
990	completed	67	86	83	

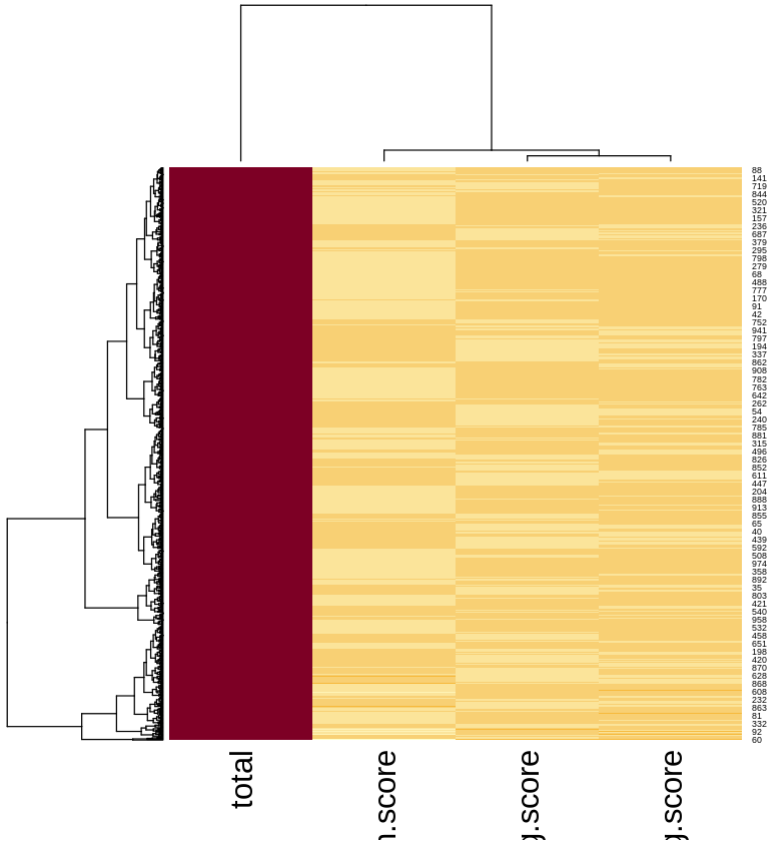
```
236
991  completed               86           81             75
242
992  completed               65           82             78
225
993  none                    55           76             76
207
994  none                    62           72             74
208
995  none                    63           63             62
188
996  completed               88           99             95
282
997  none                    62           55             55
172
998  completed               59           71             65
195
999  completed               68           78             77
223
1000 none                    77           86             86
249

     math.score reading.score writing.score total
1    72         72             74            218
2    69         90             88            247
3    90         95             93            278
4    47         57             44            148
5    76         78             75            229
6    71         83             78            232
7    88         95             92            275
8    40         43             39            122
9    64         64             67            195
10   38         60             50            148
11   58         54             52            164
12   40         52             43            135
13   65         81             73            219
14   78         72             70            220
15   50         53             58            161
16   69         75             78            222
17   88         89             86            263
18   18         32             28             78
19   46         42             46            134
20   54         58             61            173
21   66         69             63            198
22   65         75             70            210
23   44         54             53            151
24   69         73             73            215
25   74         71             80            225
26   73         74             72            219
27   69         54             55            178
```

```
28   67          69             75             211
29   70          70             65             205
30   62          70             75             207
⋮    ⋮           ⋮              ⋮              ⋮
971  89          100            100            289
972  78           72             69            219
973  53           50             60            163
974  49           65             61            175
975  54           63             67            184
976  64           82             77            223
977  60           62             60            182
978  62           65             58            185
979  55           41             48            144
980  91           95             94            280
981   8           24             23             55
982  81           78             78            237
983  79           85             86            250
984  78           87             91            256
985  74           75             82            231
986  57           51             54            162
987  40           59             51            150
988  81           75             76            232
989  44           45             45            134
990  67           86             83            236
991  86           81             75            242
992  65           82             78            225
993  55           76             76            207
994  62           72             74            208
995  63           63             62            188
996  88           99             95            282
997  62           55             55            172
998  59           71             65            195
999  68           78             77            223
1000 77           86             86            249
```

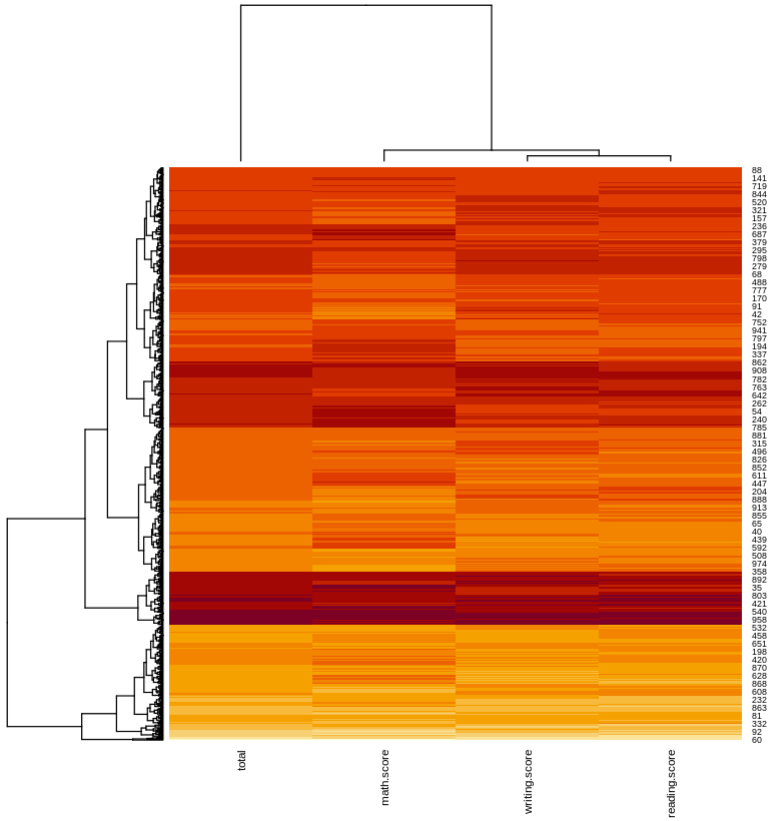
```
mdata=as.matrix(mat)
head(mdata)
```

```
  math.score reading.score writing.score total
1 72         72            74            218
2 69         90            88            247
3 90         95            93            278
4 47         57            44            148
5 76         78            75            229
6 71         83            78            232
```

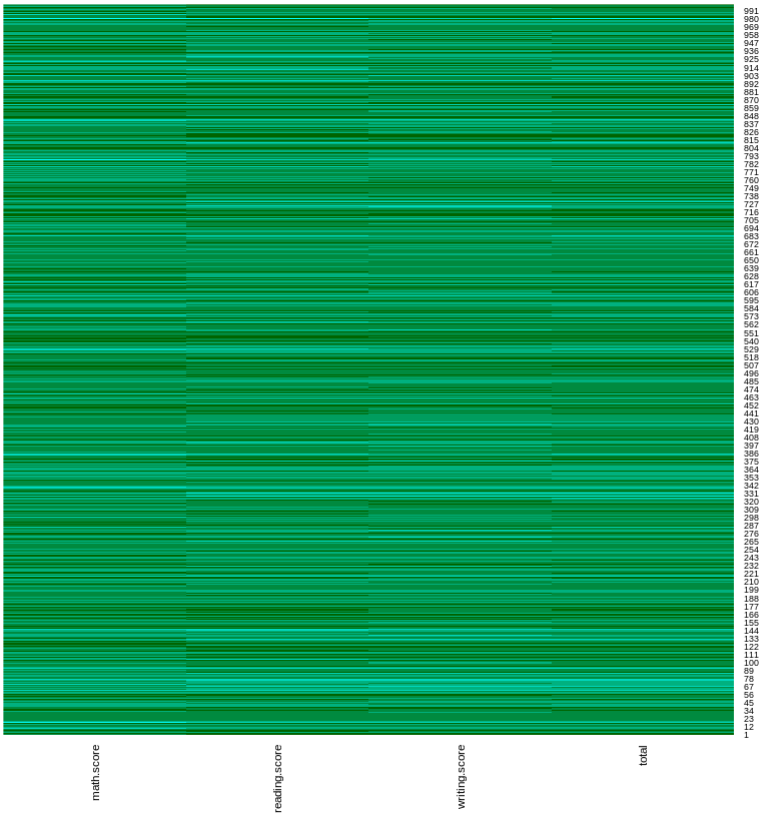
```
heatmap(mdata)
```




```
heatmap(mdata,cexCol = 0.7,scale="column")
```




```
new_colors <- colorRampPalette(c("cyan", "darkgreen"))
heatmap(mdata,cexCol = 0.7,scale="column",Colv = NA,Rowv =
          NA,col=new_colors(5))
```


991
980
969
958
947
936
925
914
903
892
881
870
859
848
837
826
815
804
793
782
771
760
749
738
727
716
705
694
683
672
661
650
639
628
617
606
595
584
573
562
551
540
529
518
507
496
485
474
463
452
441
430
419
408
397
386
375
364
353
342
331
320
309
298
287
276
265
254
243
232
221
210
199
188
177
166
155
144
133
122
111
100
89
78
67
56
45
34
23
12
1
math.score
reading.score
writing.score
total