3.1 Bar Graphs

Frequency, proportion, or percent of cases are typically presented graphically on bar graphs. If a bar graph is constructed for each level of a single categorical variable, the height of each bar is proportional to the value that it represents. If two categorical variables are involved, a **side-by-side bar graph** (also known as **grouped bar graph**) can be plotted that depicts all bars for one variable next to each other individually for each level of the other variable. Alternatively, a **stacked bar graph** shows levels of one variable as stacked segments on one bar for each level of the second variable. Below we consider an example where we produce the three types of bar graphs in SAS and R.

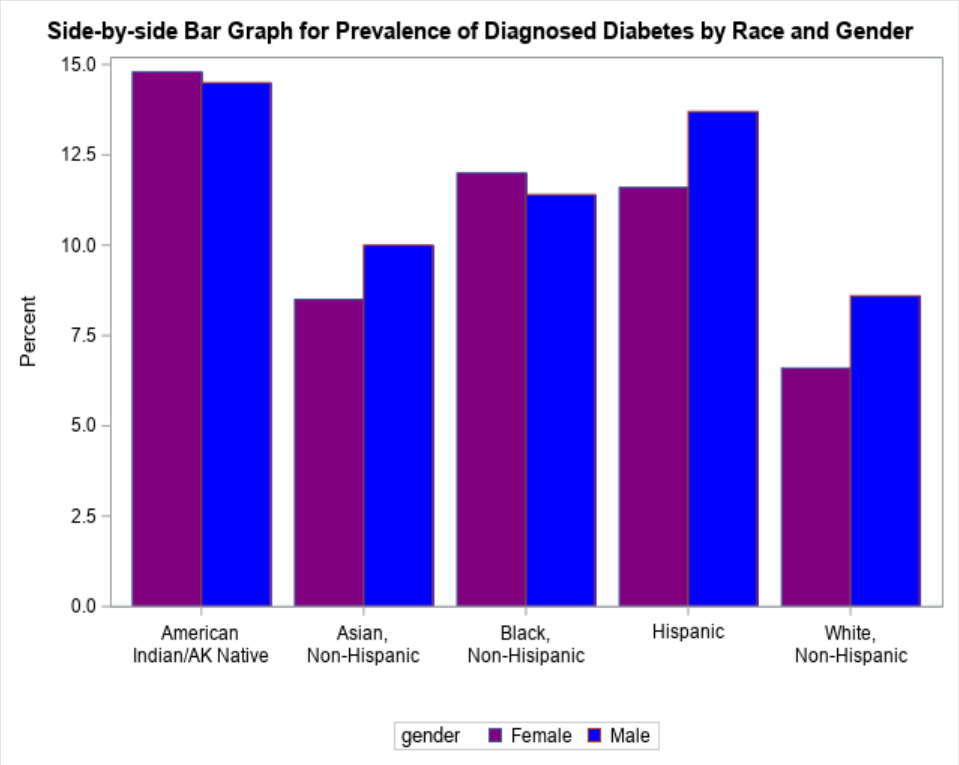
Example. The data given in the table below contain percent of diagnosed diabetes cases by race/ethnicity and gender among U.S. adults aged 18 years or older, in 2017–2018.

Race/ethnicity	Gender	
	Male	Female
American Indian/Alaska Native	14.5	14.8
Asian, Non-Hispanic	10.0	8.5
Black, Non-Hispanic	11.4	12.0
Hispanic	13.7	11.6
White, Non-Hispanic	8.6	6.6

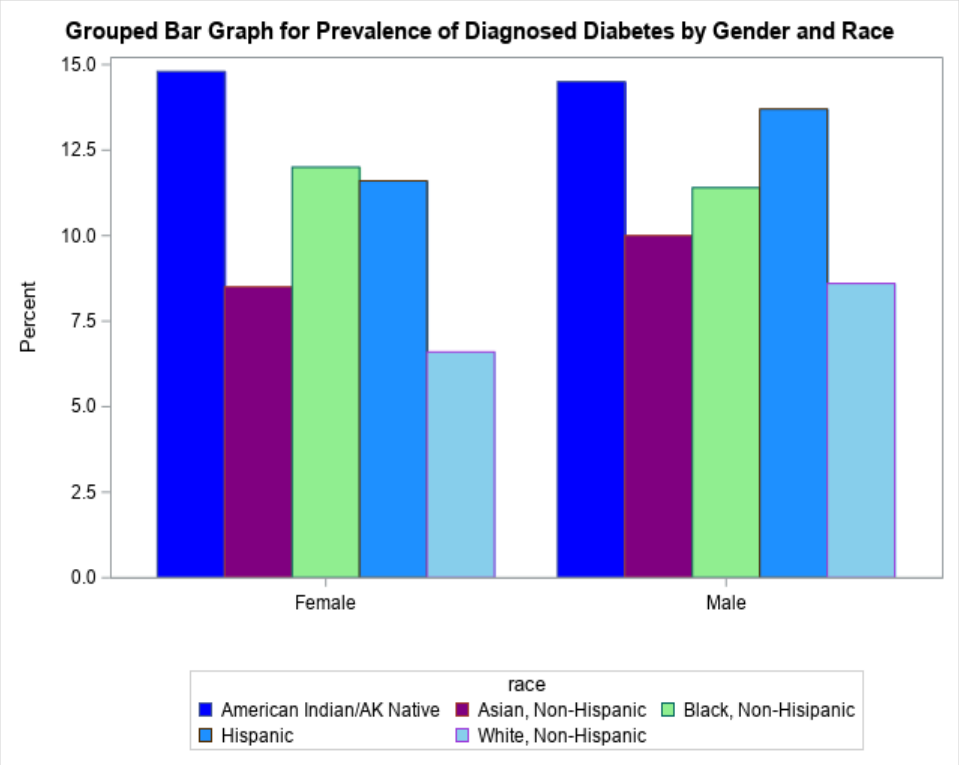
The following SAS and R codes produce bar graphs for these data.

In SAS:

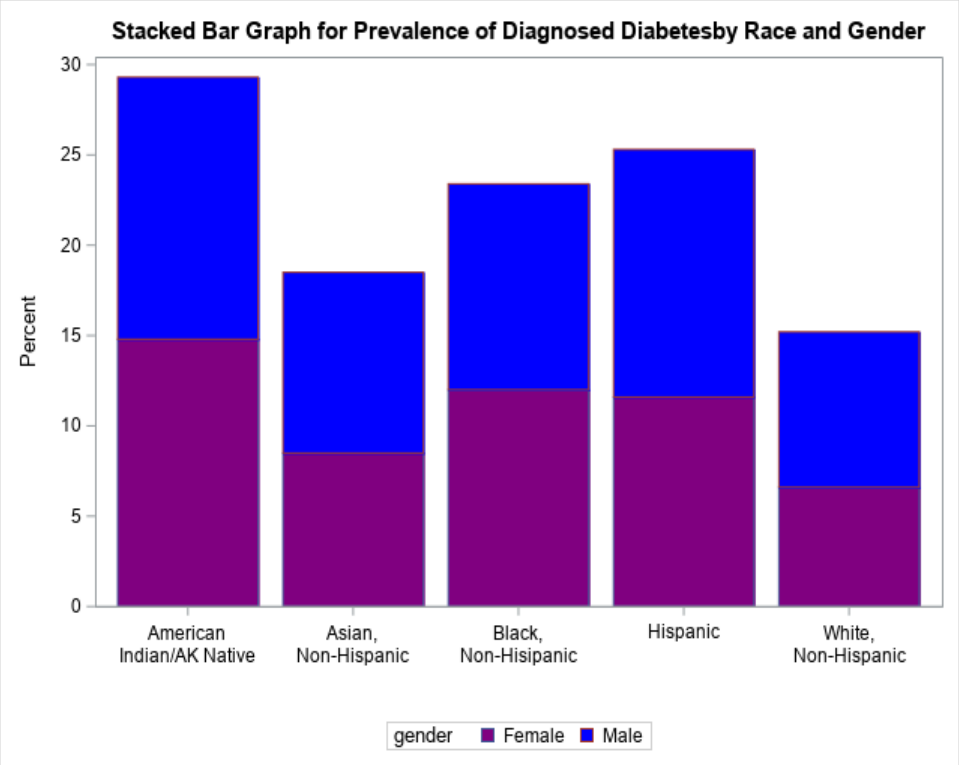
```
proc import ougt=diabetes1 datafile="./diabetes1.csv"
dbms=csv replace;
run;

title 'Side-by-side Bar Graph for Prevalence of Diagnosed Diabetes
by Race and Gender';
proc sgplot data=diabetes1; *SG=Statistical Graphics;
 styleattrs datacolors=(purple blue);
  vbar race /response=percent group=gender groupdisplay=cluster;
     xaxis label=' ';
      yaxis label='Percent';
run;
```



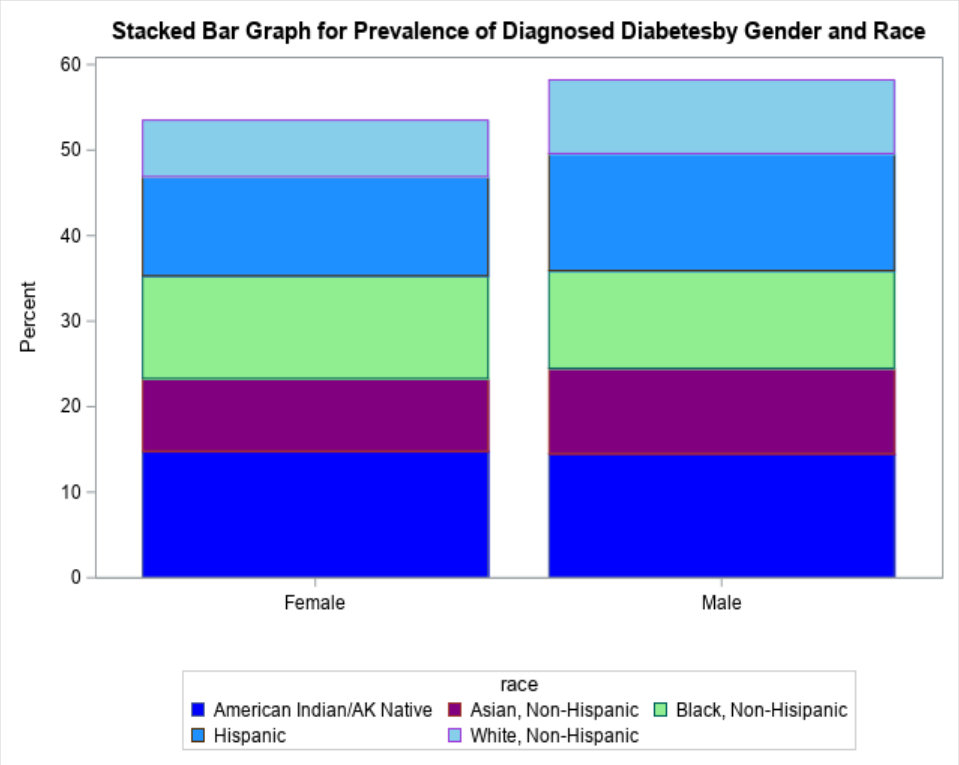

```
title 'Grouped Bar Graph for Prevalence of Diagnosed Diabetes
by Gender and Race';
proc sgplot data=diabetes1;
 styleattrs datacolors=(blue purple lightgreen dodgerblue skyblue);
  vbar gender /response=percent group=race groupdisplay=cluster;
     xaxis label=' ';
      yaxis label='Percent';
run;
```




```
title 'Stacked Bar Graph for Prevalence of Diagnosed Diabetes
by Race and Gender';
proc sgplot data=diabetes1;
 styleattrs datacolors=(purple blue);
  vbar race /response=percent group=gender groupdisplay=stack;
     xaxis label=' ';
      yaxis label='Percent';
run;
```



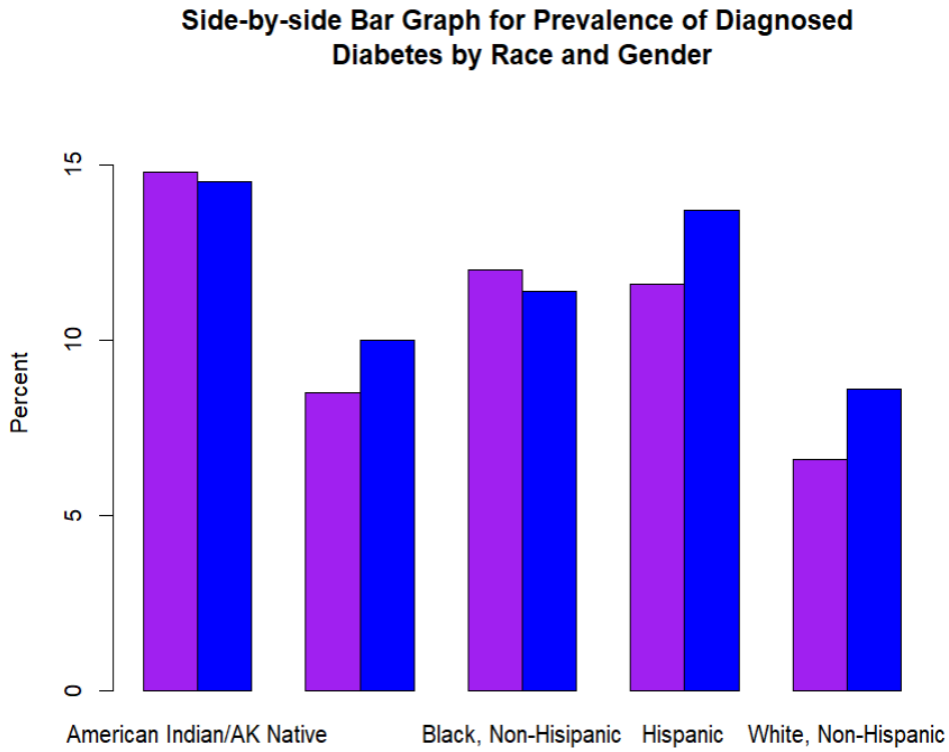

```
title 'Stacked Bar Graph for Prevalence of Diagnosed Diabetes
by Gender and Race';
proc sgplot data=diabetes1;
 styleattrs datacolors=(blue purple lightgreen dodgerblue skyblue);
  vbar gender /response=percent group=race groupdisplay=stack;
     xaxis label=' ';
      yaxis label='Percent';
run;
```



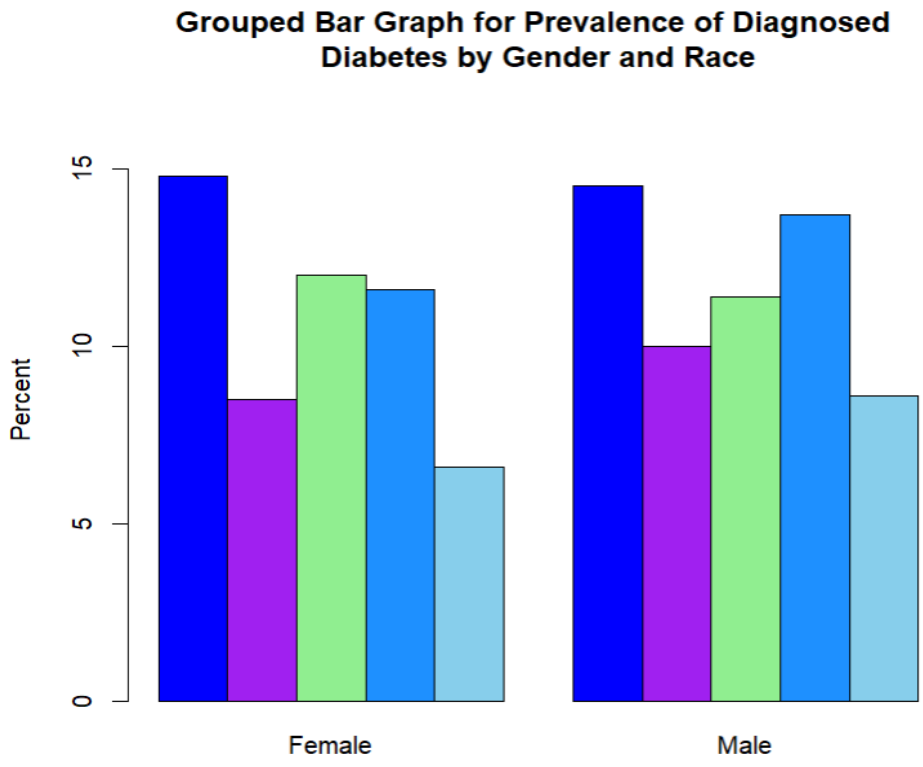
From the bar graphs, American Indians/Alaskan Natives have the highest percentages diagnosed with diabetes, for both men and women. Whites have the smallest such percentages. Fewer percentage of women than men are diagnosed for Asians, Hispanics, and Whites. For American Indians/Alaskan Natives and Blacks, higher percentage of women have diabetes than men.

In R:

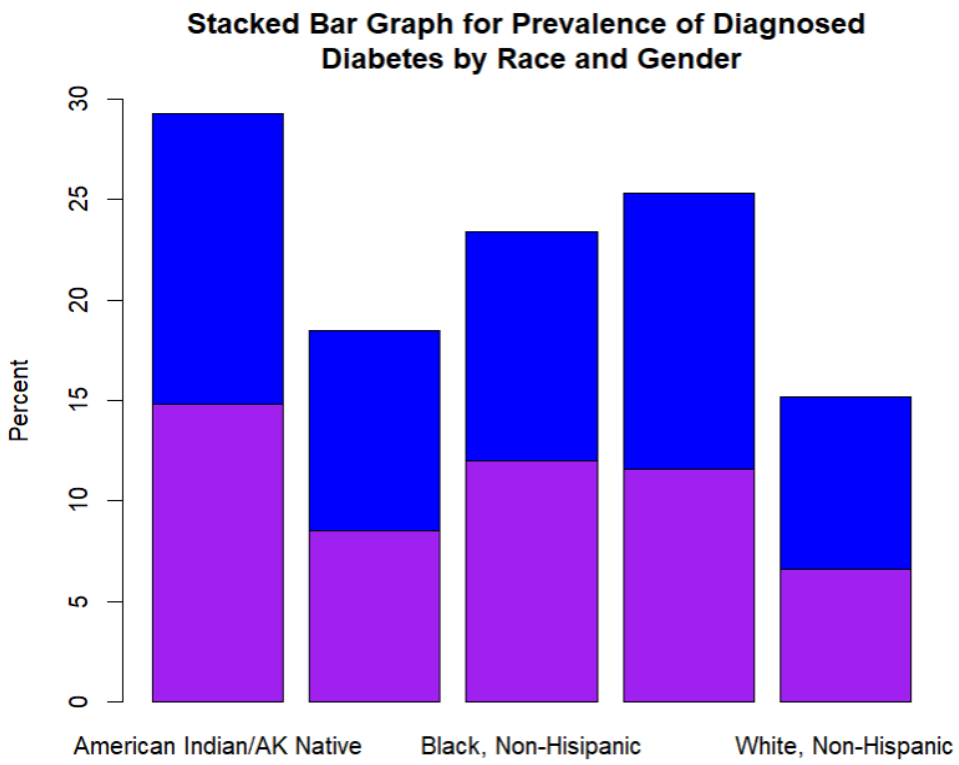
```
diabetes1<- read.csv(file="./diabetes1.csv", header=TRUE, sep=",")

table1 <- xtabs(percent ~ gender+race, data=diabetes1)
table2 <- xtabs(percent ~ race+gender, data=diabetes1)

barplot(table1, main='Side-by-side Bar Graph for Prevalence of Diagnosed
Diabetes by Race and Gender', ylim=c(0,17), xlab='', ylab='Percent',
col=c("purple","blue"), legend.text=rownames(table), beside=TRUE)
```



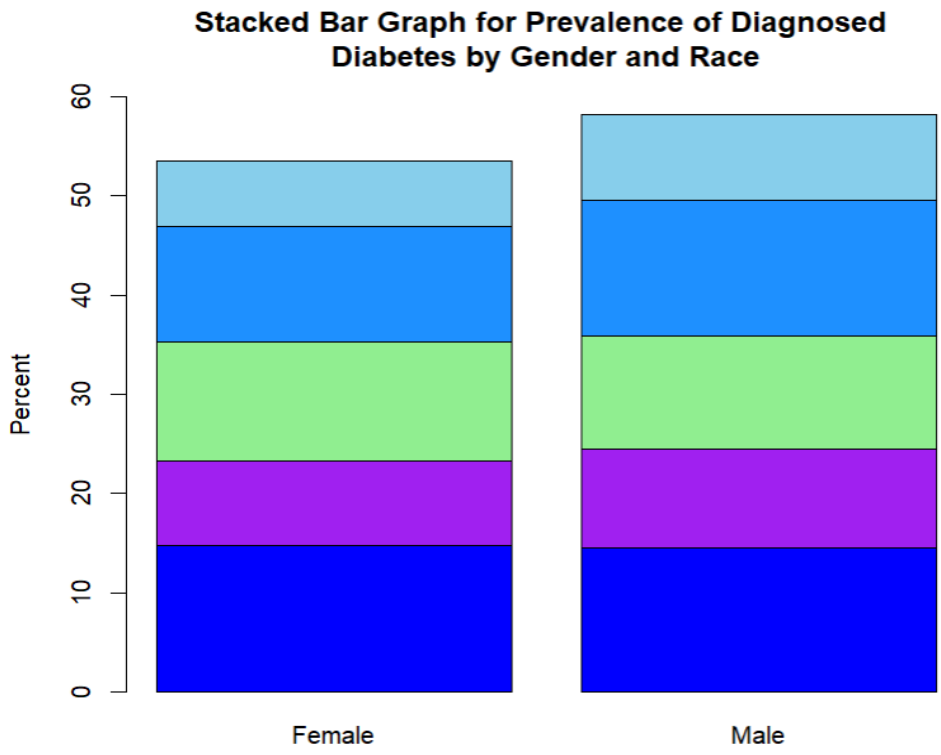

```
barplot(table2, main='Grouped Bar Graph for Prevalence of Diagnosed
Diabetes by Gender and Race', ylim=c(0,17), xlab='', ylab='Percent',
col=c("blue","purple","lightgreen","dodgerblue","skyblue"),
legend.text=rownames(table), beside=TRUE)
```




```
barplot(table1, main='Stacked Bar Graph for Prevalence of Diagnosed
Diabetes by Race and Gender', ylim=c(0,30), xlab='', ylab='Percent',
col=c("purple", "blue"), legend.text=rownames(table), beside=FALSE)
```




```
barplot(table2, main='Stacked Bar Graph for Prevalence of Diagnosed
Diabetes by Gender and Race', ylim=c(0,60), xlab='', ylab='Percent',
col=c("blue","purple","lightgreen","dodgerblue","skyblue"),
legend.text=rownames(table), beside=FALSE)
```


Stacked Bar Graph for Prevalence of Diagnosed Diabetes by Gender and Race
Percent
0
10
20
30
40
50
60
Female
Male

□