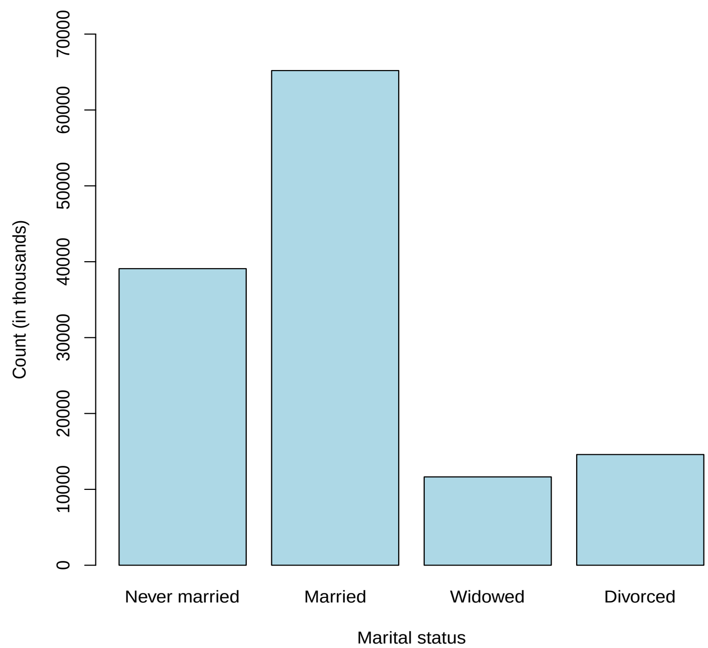
10.11. (a) To find how many women were **not** married, simply add the numbers for "Never married," "Widowed," and "Divorced." This is

$$39,087 + 11,642 + 14,591 = 65,320.$$

Because these counts are measured in 1000s, this means

$$65,320 \times 1000 = 65,320,000$$

women were not married in 2017.
(b) Here is the bar graph for these data:



This graph shows the distribution of marital status for American women in 2017. This is a categorical variable.
(c) It would only be correct to use a pie chart if "Never married," "Married," "Widowed," and "Divorced" were the only possible marital statuses. However, aren't there more statuses than these? What about "Separated?" We cannot use a pie chart unless we know the counts (or percentages) of the statuses that are not listed above. Pie chart percentages must add up to 100%.

Here is the R code I used to make the bar chart above:

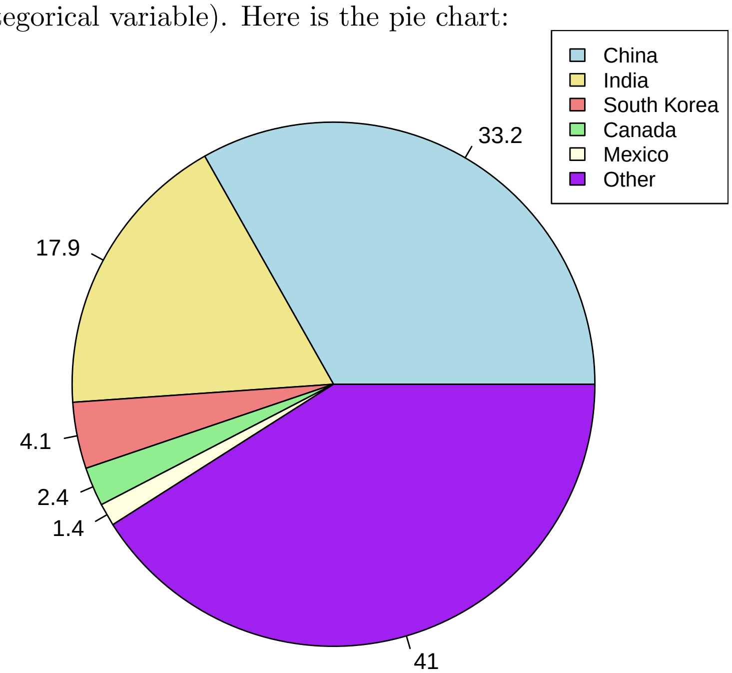
```
counts = c(39087,65193,11642,14591)
labels = c("Never married","Married","Widowed","Divorced")
barplot(counts, main="", xlab="Marital status", ylab="Count (in thousands)",
    names.arg=labels, ylim=c(0,70000), col="lightblue")
```

10.13. There is an increasing trend for each age group over time (from 2001-2015). The proportion who favor same-sex marriage is highest overall for the Millennial generation, followed by Generation X, followed by Baby Boomers, followed by the Silent Generation.

10.14. We do need an "Others" category so the percentages add up to 100%. Let's make a table first to display the percentages:

Category	China	India	South Korea	Canada	Mexico	Other	Total
Percentage	33.2%	17.9%	4.1%	2.4%	1.4%	**41.0%**	100%

We could make a bar graph or a pie chart to show the distribution of international student country of origin (a categorical variable). Here is the pie chart:

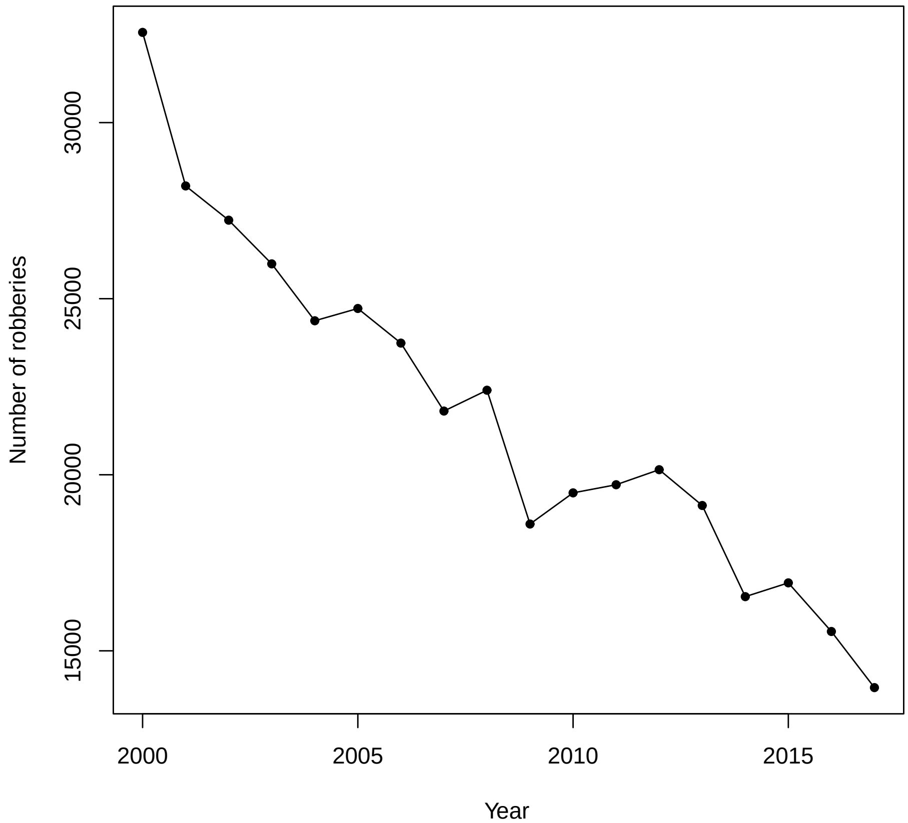


Here is the R code I used to produce the bar chart above:

```
percentages = c(33.2,17.9,4.1,2.4,1.4,41.0)
labels = c("China","India","South Korea","Canada","Mexico","Other")
pie(percentages,labels = percentages,main = "",
    col=c("lightblue","khaki","lightcoral","lightgreen","lightyellow","purple"))
legend("topright",labels,cex = 0.9,
    fill=c("lightblue","khaki","lightcoral","lightgreen","lightyellow","purple"))
```

10.19. The graph shows the percentage of educational attainment is close for males and females in all six categories. There is a slightly larger percentage for males in the "Less than high school" and "High school graduate" categories. The female percentages in the other categories are slightly larger than the male percentages.

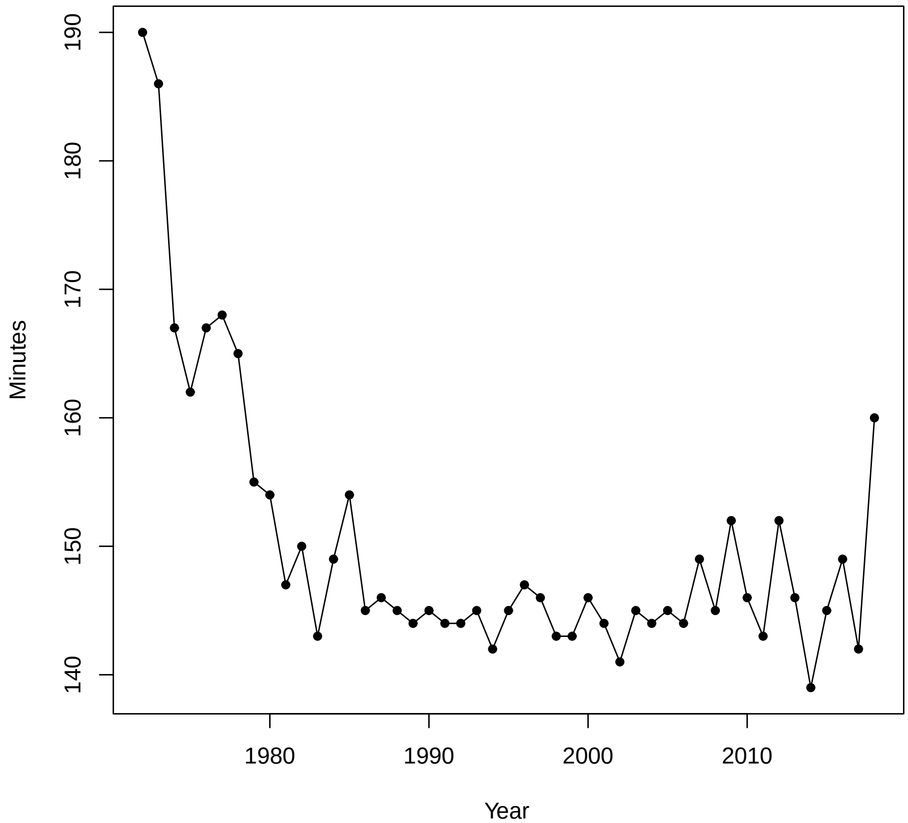
10.28. The number of robberies is a quantitative variable, and we are observing the values of this variable over time—each year between 2000 and 2017. A line graph (time series graph) is appropriate. From the graph (next page, top), we can see there is a clear decreasing trend in the number of robberies during 2000-2017. There are no sharp deviations from the overall decreasing trend. There is no seasonal variation because these are annual counts. We might see seasonal variation if we had monthly counts (e.g., if robberies consistently occurred more in some months than others and this pattern repeated itself over time).



Here is the R code I used to make this graph above:

```
year = seq(2000,2017,1)
robberies = c(32562,28202,27229,25989,24373,24722,23739,21809,22401,18601,
    19486,19717,20144,19128,16539,16931,15550,13956)
plot(year,robberies,ylab="Number of robberies",xlab="Year",type="o",pch=19,cex=0.75)
```

10.32. The winning time (in minutes) is a quantitative variable, and we are observing the values of this variable over time−each year between 1972 and 2018. A line graph (time series graph) is appropriate; see next page.



(b) The trend early on (during the 1970s) was decreasing over time. However, the times since 1980 have remained fairly stable. This is perhaps due to the natural "lower bound" on times that humans can reach. There does appear to be more variation in the winning times during the last 10 years or so. I'm not sure why this would happen.

Here is the R code I used to make this graph above:

```
winning.times = ts(read.table(file =
"C:\\Users\\tebbs\\OneDrive - University of South Carolina\\Documents\\
    texfiles\\Classes\\USC\\stat110\\s24\\data\\boston.txt",header=TRUE),start=1972)
plot(winning.times,ylab="Minutes",xlab="Year",type="o",pch=19,cex=0.75)
```