TT TEXTBOOK TESTBANKS

Solutions Manual for Business Statistics Communicating with Numbers 4th Jaggia

SOLUTIONS MANUAL SAMPLE PREVIEW

PUBLISHER
McGraw-Hill

COPYRIGHT
2022

RESOURCE TYPE
Solutions Manual

ISBN
9781260716306

Chapter 2. Tabular and Graphical Methods

Solutions

1.

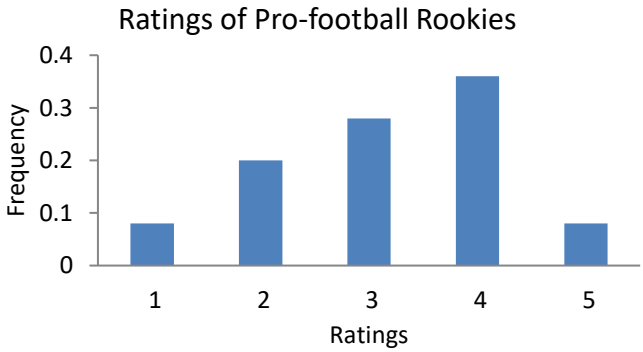
a. 22 (18 + 4) out of 50 rookies received a rating of 4 or better; 14 (10 + 4) out of 50 rookies received a rating of 2 or worse.

b.

Rating	Relative Frequency
1	4/50 = 0.08
2	10/50 = 0.20
3	14/50 = 0.28
4	18/50 = 0.36
5	4/50 = 0.08

8% of the rookies received a rating of 5.

c.



The rating of 4 has the highest frequency; the distribution is not symmetric, rather it is negatively skewed.

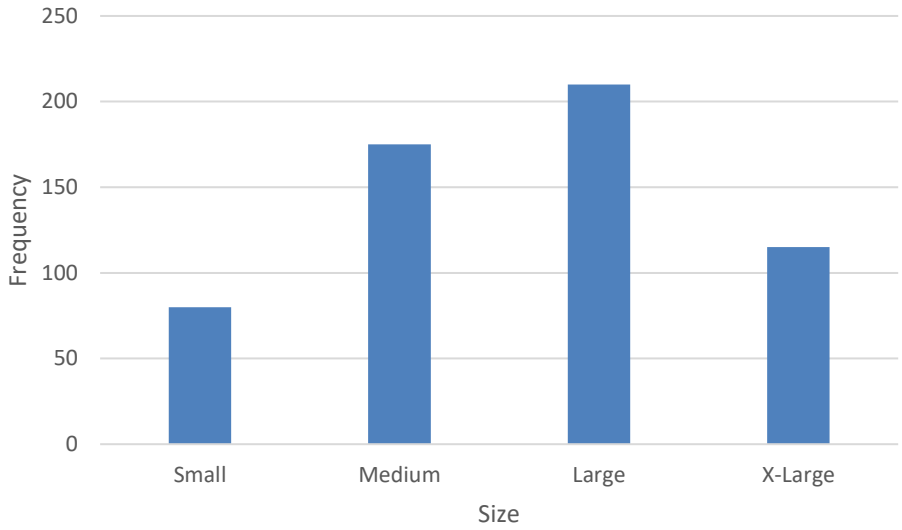
2.

a.

Size	Frequency	Relative Frequency
Small	80	80/580=0.138
Medium	175	175/580=0.302
Large	210	210/580=0.362
X-Large	115	115/580=0.198

The proportion of the sales for medium-sized shirts was 0.302.

b.



Sales of large-sized shirts had the highest frequency and sales of small-sized shirts had the lowest frequency.

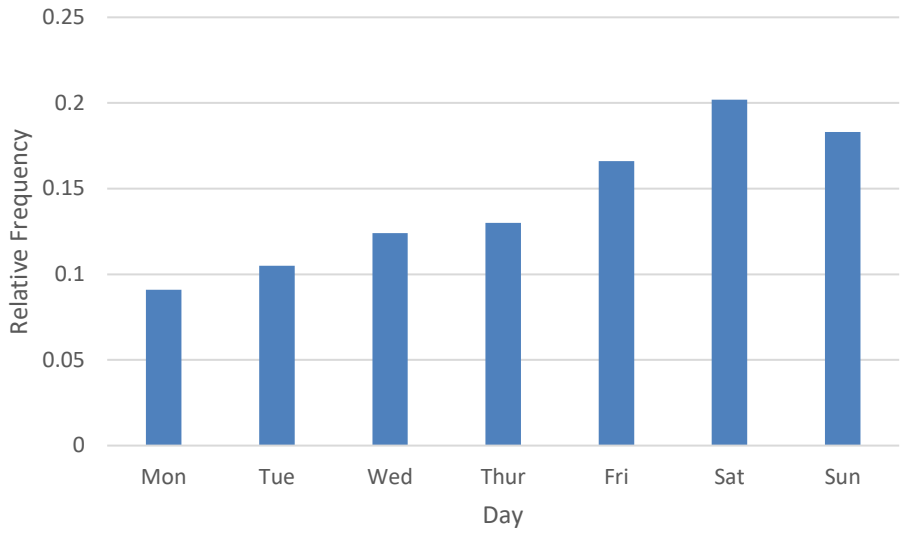
3.

a.

Day	Frequency	Relative Frequency
Mon	2504	2504/27500=0.091
Tue	2880	2880/27500=0.105
Wed	3402	3402/27500=0.124
Thur	3566	3566/27500=0.130
Fri	4576	4576/27500=0.166
Sat	5550	5550/27500=0.202
Sun	5022	5022/27500=0.183

The proportion of purchases that occurred on Wednesday was 0.124.

b.



The retailer saw the highest proportion of sales on the weekend. The lowest proportion of sales occurred on Monday.

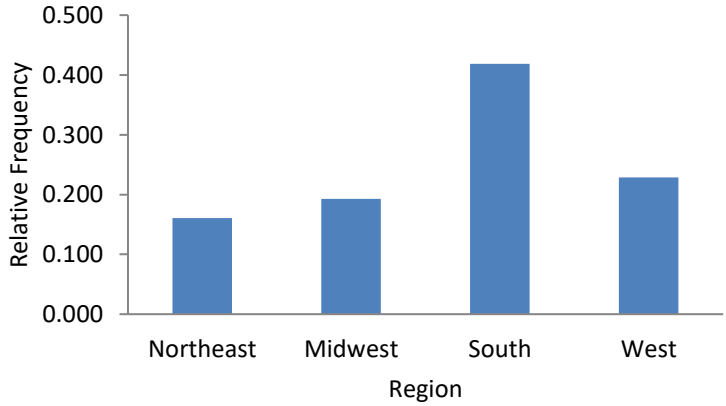
4.

a.

Region	Relative Frequency
Northeast	0.161
Midwest	0.193
South	0.418
West	0.228

19.3% of people in the Midwest are living below the poverty level.

b.



The South has the highest relative frequency as compared to the other three regions, which are roughly equal.

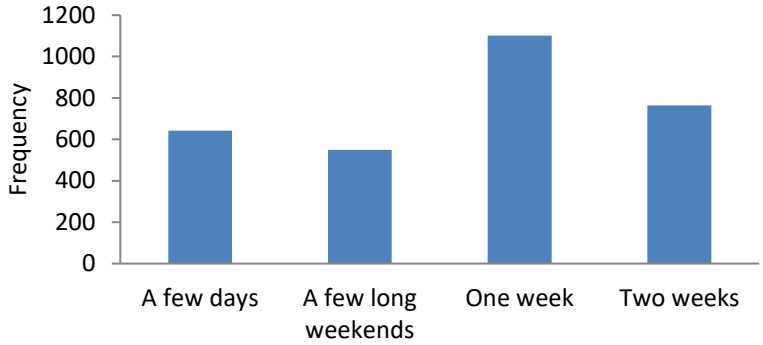
5.

a.

Response	Frequency
A few days	0.21(3057) = 642
A few long weekends	0.18(3057) = 550
One week	0.36(3057) = 1,101
Two weeks	0.25(3057) = 764

Approximately 1,101 people are going to take a one-week vacation.

b.



"One week" has the highest frequency as compared to the other categories, which are roughly equal.

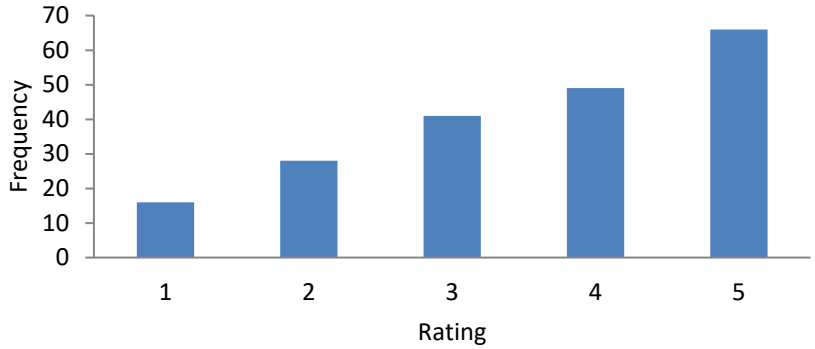
6.

a.

Rating	Frequency
1	16
2	28
3	41
4	49
5	66

A rating of 5 has the highest frequency

b.



The higher ratings have the higher frequencies.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
Frequency <- table(myData$Rating); Frequency
#Obtain bar chart
barplot(Frequency, col=("blue"))
abline(v=0)
```


7.

a.

Quality of Health	Frequency
Poor	18
Fair	26
Good	75
Excellent	31

Good is the most common response.

b.

Patients seem to be healthy, with the majority rated Good or Excellent.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
Frequency <- table(myData$Quality); Frequency
#Obtain bar chart
barplot(Frequency, col=("blue"))
abline(v=0)
```

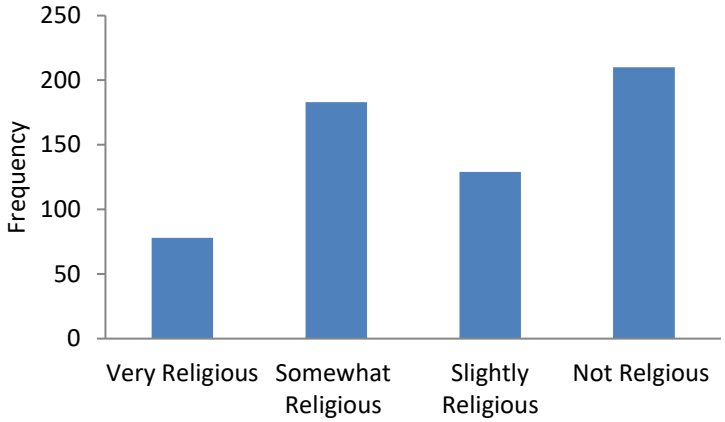

8.

a.

Faith	Frequency
Very Religious	78
Somewhat Religious	183
Slightly Religious	129
Not Religious	210

Not Religious is the most common response.

b.



210/600 = 0.35; about 35% responded "Not Religious" which is consistent with the earlier study.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution and relative frequency distribution
Frequency <- table(myData$Faith); Frequency
prop.table(Frequency)
#Obtain pie chart
pie(Frequency)
```

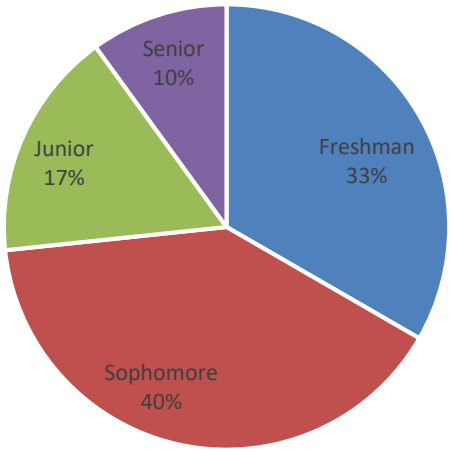

9.

a.

Status	Frequency	Relative Frequency
Freshman	20	0.33
Sophomore	24	0.40
Junior	10	0.17
Senior	6	0.10

There are 20 freshmen in her class; 40% of the class are sophomores.

b.



Freshmen and sophomores make up the majority of the class, 33% and 40%, respectively; 17% of the class are juniors and the remaining 10% are seniors.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution and relative frequency distribution
Frequency <- table(myData$Status)
prop.table(Frequency)
#Obtain pie chart
pie(Frequency)
```


10. This graph does not correctly depict what has happened to company's stock price over this period. By using a relatively high value as an upper limit on the vertical axis ($500), the rise in stock price appears dampened.

11. This graph does not correctly depict what has happened to sales over the most recent five-year period. The vertical axis has been stretched so that the increase in sales appears more pronounced than warranted.

12.

a. There were 378 total shipments.
b. 46 shipments were defective.
c. 80 shipments were from Vendor II.
d. 14 shipments were from Vendor I and were defective.

13.

a. 542 units passed inspection; 88 units failed inspection
b. 129 of the inspectors had high experience
c. 0.2667
d. 0.0730

14.

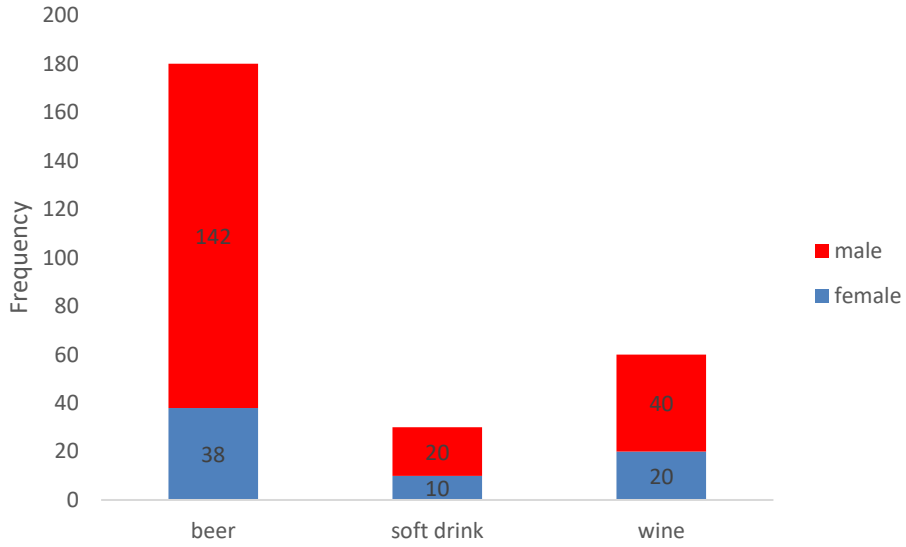
a.

	Sex		
Drink Choice	female	male	Total
beer	38	142	**180**
soft drink	10	20	**30**
wine	20	40	**60**
Total	**68**	**202**	**270**

202 of the customers were male; 60 of the customers drank wine.

b. 142/202 = 0.7030; 38/68 = 0.5588.

c.



The stacked column chart shows that beer is the popular drink at this bar, followed by wine and then soft drinks. Both men and women are more likely to choose beer over the other two options.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$Sex, myData$Drink); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable) , col=(c("blue","red"))
```

15.

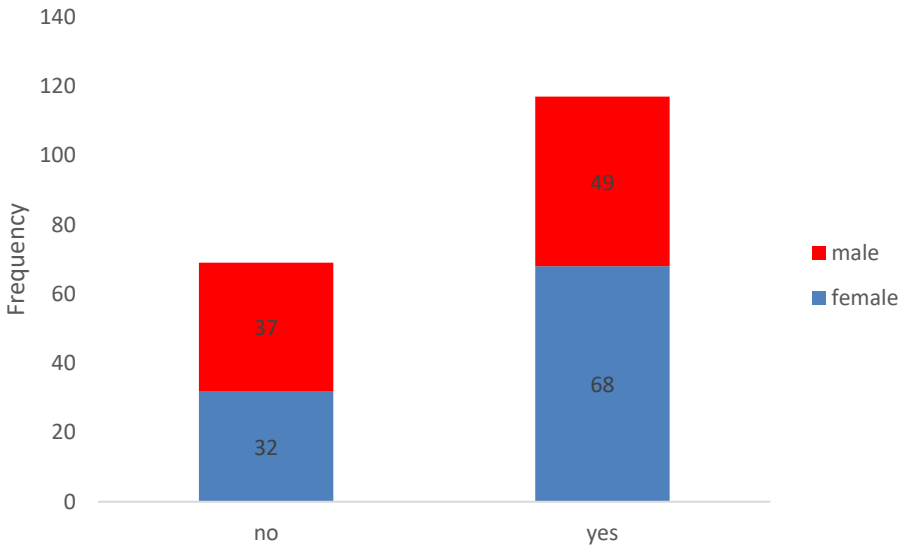
a.

	Gender		
Feasible	female	male	Total
no	32	37	**69**
yes	68	49	**117**
Total	**100**	**86**	**186**

100 of the students were female; 117 of the students thought it was feasible for men and women to be just friends.

b. 49/86=0.5698; 68/100 = 0.6800.

c.



The stacked column chart shows that yes is the most common answer. Both men and women are more likely to say that it is feasible for men and women to be friends.

However, women are more likely to say that it is feasible for men and women to be friends (0.68 > 0.57).

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$Sex, myData$Feasible); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable) , col=(c("blue","red"))
```

16.

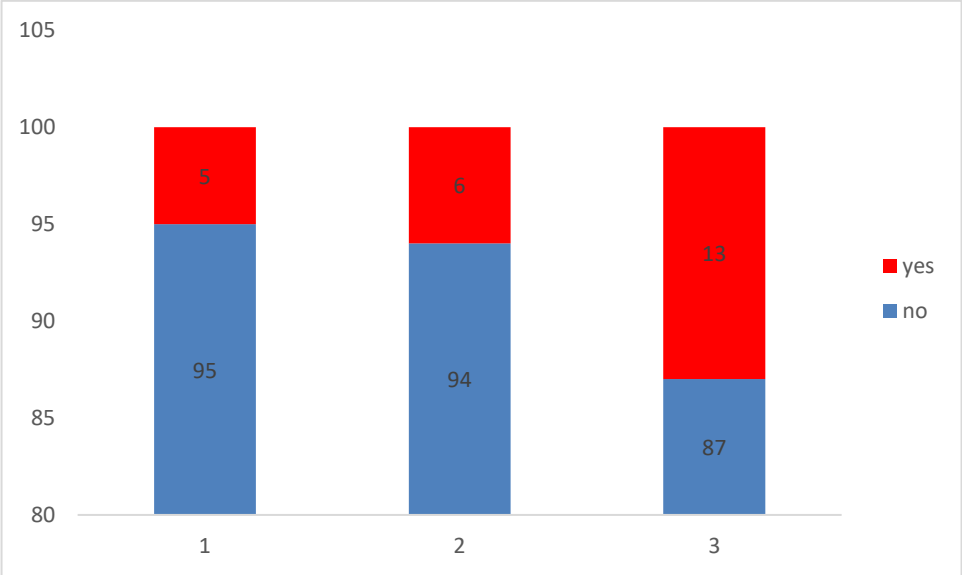
a.

	Shift			
Defective	1	2	3	**Total**
no	95	94	87	**276**
yes	5	6	13	**24**
Total	**100**	**100**	**100**	**300**

Five components from shift 1 were defective; 94 components from shift 2 were not defective.

b. 6/24=0.2500; 13/24 = 0.5417; components constructed during shift 3 seem to be defective at a higher rate.

c.



Defect rates seem to be highest in shift 3.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$Defective, myData$Shift); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable), col=(c("blue","red"))
```

17.

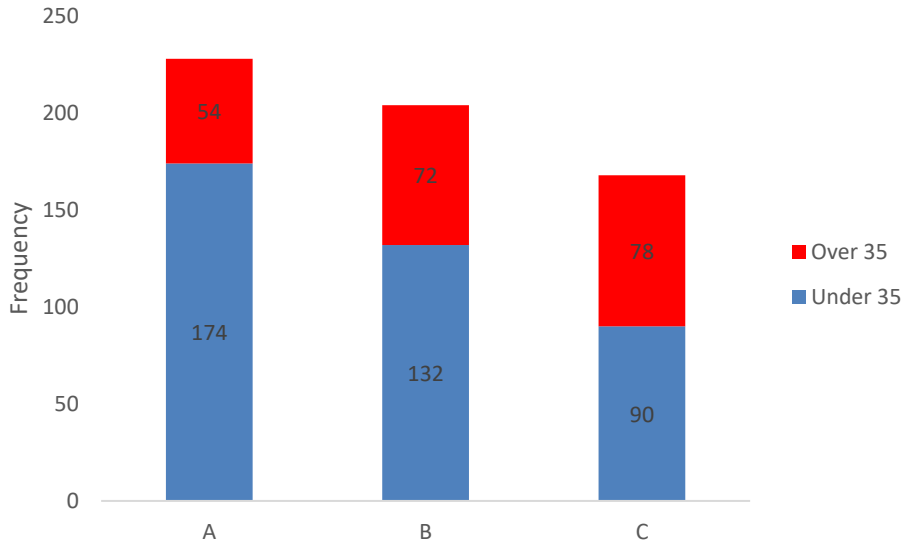
a.

	Under 35		
Brand	**no**	**yes**	**Total**
A	54	174	228
B	72	132	204
C	78	90	168
Total	**204**	**396**	**600**

228 purchases were from brand A; 396 purchases were from customers under 35.

b. 174/396=0.4394; 132/396=0.3333; 90/396=0.2273; The data suggests that customers under 35 prefer brand A.

c.



Younger customers seem to prefer A over B, and B over C, while older Customers seem to prefer C over B, and B over A

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$Age, myData$Brand); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable), col=(c("blue","red")))
```


18.

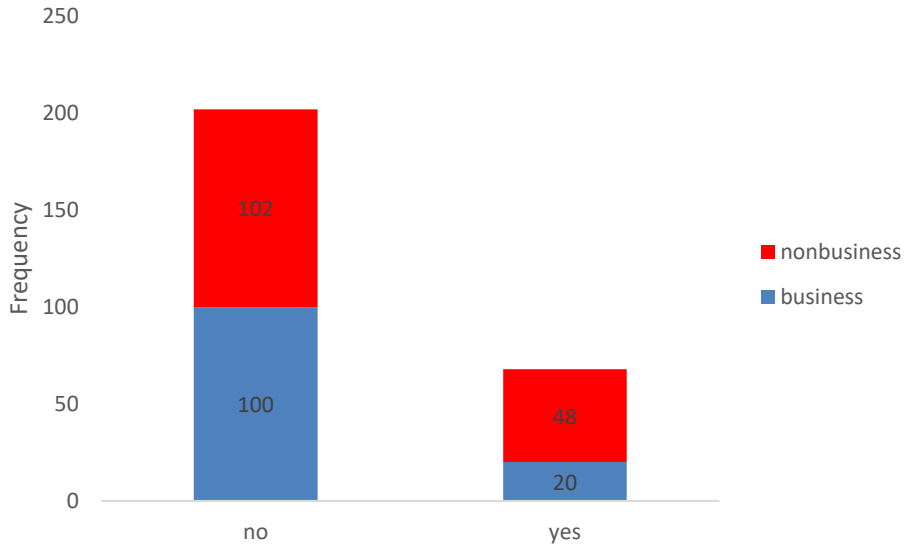
a.

	Major		
Study hard	**business**	**nonbusiness**	**Total**
no	100	102	202
yes	20	48	68
Total	120	150	270

120 students are business majors; 68 students study hard.

b. 20/120=0.1667; 48/150=0.3200; The data suggest that nonbusiness majors are more likely to study hard, which supports the report.

c.



The majority of both business and nonbusiness students do not study hard, but nonbusiness students are more likely to study hard.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$Major, myData$Study_Hard); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable), col=c("blue","red"))
```

19.

Interval	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
$10 < x \le 20$	12	12/56 = 0.214	12	12/56 = 0.214
$20 < x \le 30$	15	0.268	12 + 15 = 27	27/56 = 0.482
$30 < x \le 40$	25	0.446	52	0.928
$40 < x \le 50$	4	0.071	56	1.000

a. 0.268
b. 52
c. 0.482

20.

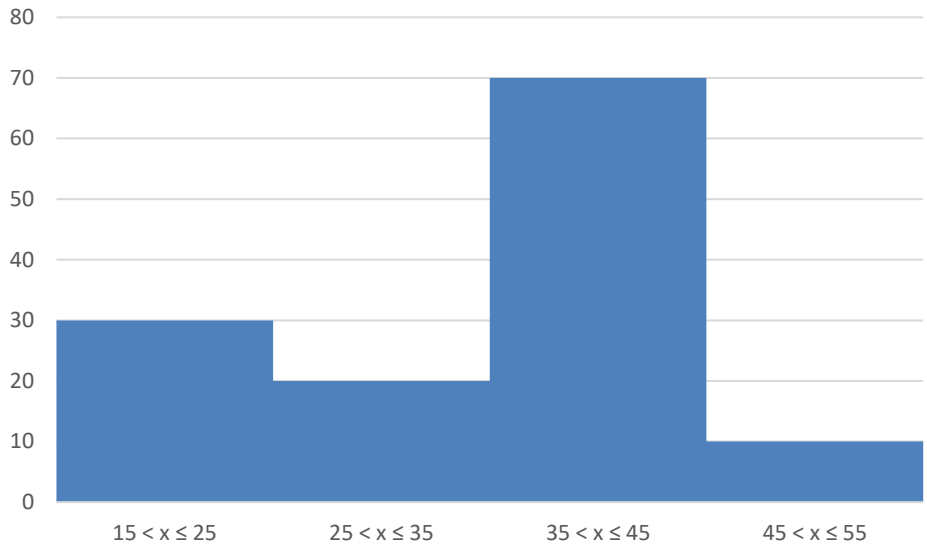
Interval	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1,000 < x ≤ 1,100	22	22/120 = 0.183	22	22/120 = 0.183
1,100 < x ≤ 1,200	38	0.317	22+38 = 60	60/120 = 0.500
1,200 < x ≤ 1,300	44	0.367	104	0.867
1,300 < x ≤ 1,400	16	0.133	120	1.000

a. 0.317
b. 104
c. 0.867; 0.133

21.

Interval	Cumulative Frequency	Frequency	Cumulative Relative Frequency
15 < x ≤ 25	30	30	30/130 = 0.231
25 < x ≤ 35	50	50 - 30 = 20	50/130 = 0.385
35 < x ≤ 45	120	120 - 50 = 70	0.923
45 < x ≤ 55	130	130 - 120 = 10	1.000

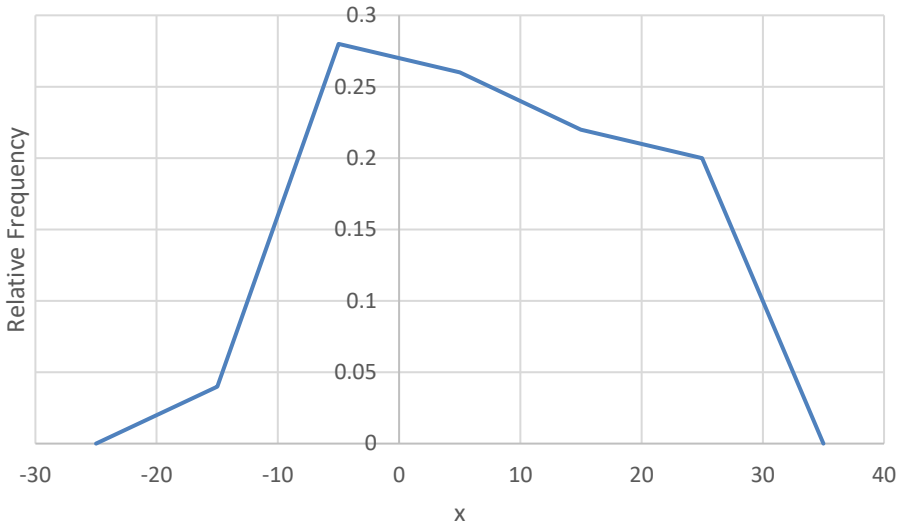
a. 70
b. 0.923
c. The distribution is not symmetric.



22.

Interval	Relative Frequency	Frequency	Cumulative Frequency	Cumulative Relative Frequency
-20 < *x* ≤ -10	0.04	0.04×50 = 2	2	12/56 = 0.214
-10 < *x* ≤ 0	0.28	14	2 + 14 = 16	27/56 = 0.482
0 < *x* ≤ 10	0.26	13	29	0.928
10 < *x* ≤ 20	0.22	11	40	1.000
20 < *x* ≤ 30	0.20	10	50	

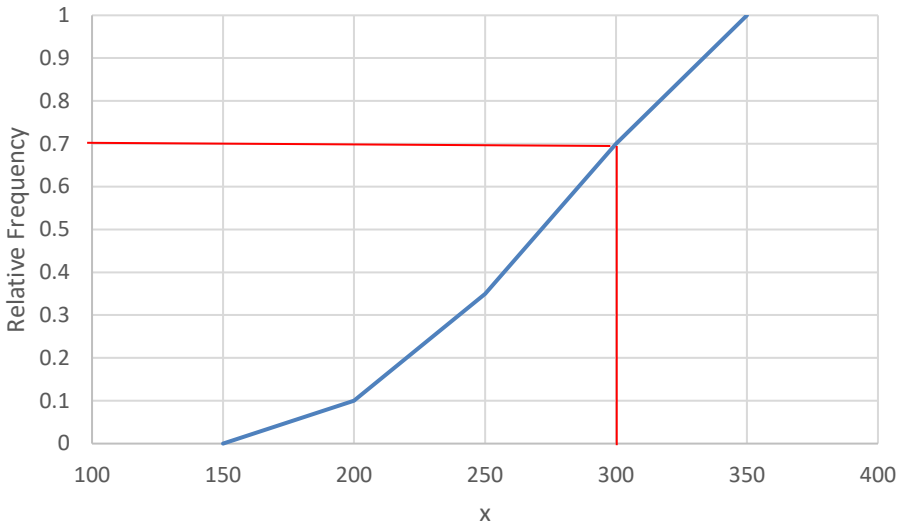
a. 14
b. 40
c. The distribution is not symmetric.



23.

Interval	Cumulative Relative Frequency	Relative Frequency
$150 < x \leq 200$	0.10	0.10
$200 < x \leq 250$	0.35	0.35 – 0.10 = 0.25
$250 < x \leq 300$	0.70	0.35
$300 < x \leq 350$	1.00	0.30

a. 0.35
b. 0.70 – see ogive below.



24.

a. No. The distribution is not symmetric. It is positively skewed.
b. 0.34.
c. 0.66(5,000)=3,300

25.

a. No. The distribution is not symmetric. It is positively skewed.
b. 0.43+0.43=0.86
c. (0.10+0.03)3,000=390

26.

a. Draw a vertical line through Variable X at 16. Once it intercepts with the ogive, draw a horizontal line to the y-axis. It intercepts at the point of about 0.40. Thus, about 0.40 of the observations are less than 16.

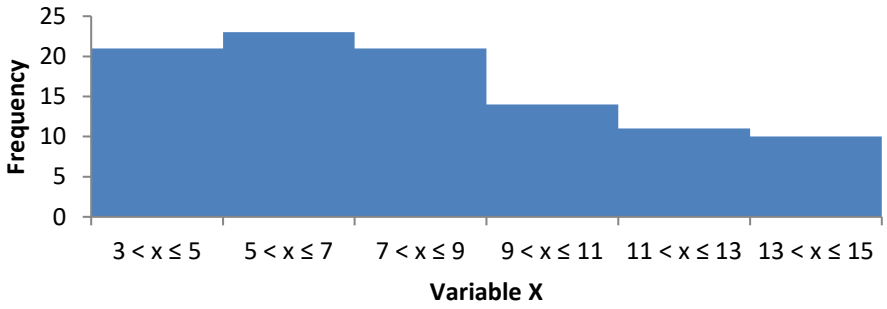
b. Draw a vertical line through Variable X at 20. Once it intercepts the ogive, draw a horizontal line to the y-axis. It intercepts at the point of about 0.80. Thus, about 0.80 of the observations are less than 20 implying that 0.20 are greater than 20. So 0.20(1,000)=200 observations are greater than 20.

27.

Interval	Frequency	Relative Frequency
$3 < x \leq 5$	21	0.21

5 < *x* ≤ 7	23	0.23
7 < *x* ≤ 9	21	0.21
9 < *x* ≤ 11	14	0.14
11< x ≤ 13	11	0.11
13 < x ≤ 15	10	0.10

a. 21
b. 0.23
c. The distribution is not symmetric.



The R Code

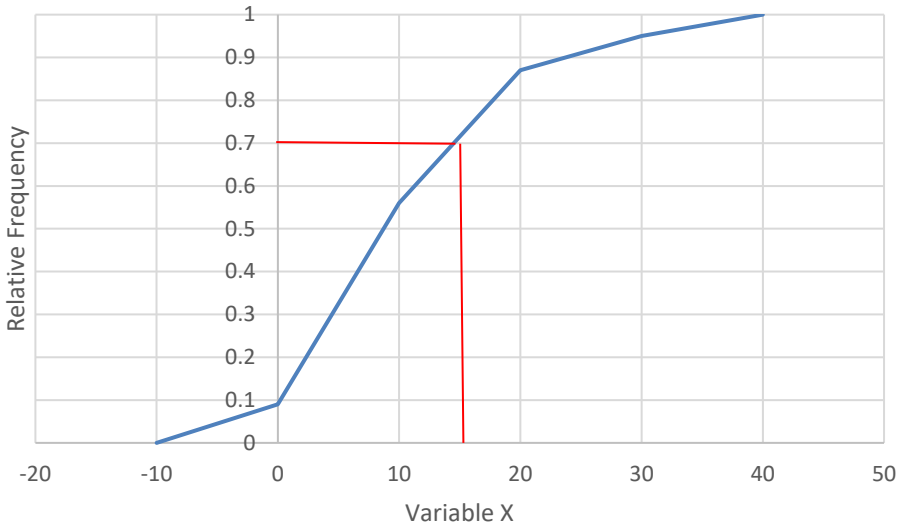
```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(3, 15, by=2)
intervals.cut <- cut(myData$X, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
#Obtain histogram
hist(myData$X, breaks=intervals, right=TRUE, col="blue")
```

28.

Interval	Frequency	Relative Frequency	Cumulative Relative Frequency
-10 < *x* ≤ 0	9	0.09	0.09
0 < *x* ≤ 10	47	0.47	0.56
10 < *x* ≤ 20	31	0.31	0.87
20 < *x* ≤ 30	8	0.08	0.95

$30 < x \leq 40$	5	0.05	1

a. 47
b. 0.31; 0.13
c. Approximately 0.70 of the observations are less than 15; see the ogive.



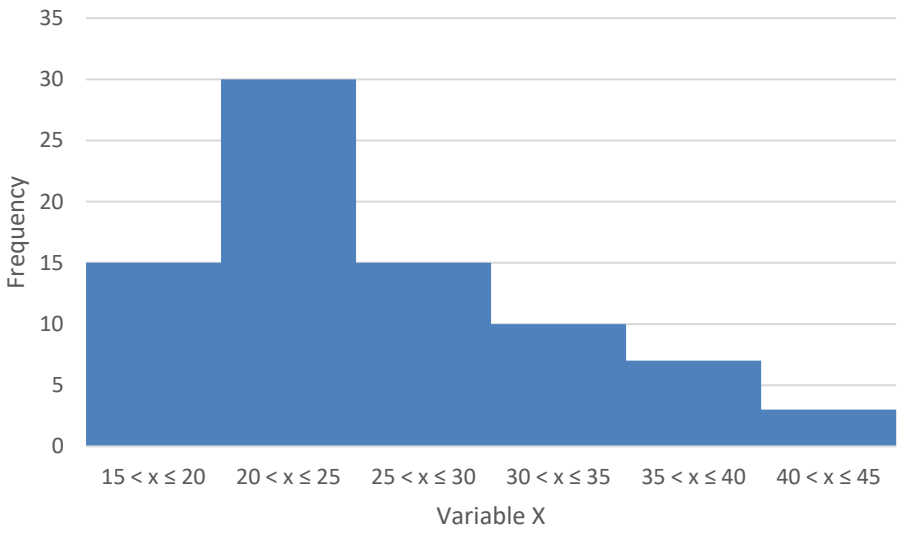
The R Code


```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(-10, 40, by=10)
intervals.cut <- cut(myData$X, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
```

29.

Average MPG	Frequency	Relative frequency	Cumulative relative frequency
$15 < x \leq 20$	15	15/80 = 0.1875	0.1875
$20 < x \leq 25$	30	0.3750	0.5625
$25 < x \leq 30$	15	0.1875	0.7500
$30 < x \leq 35$	10	0.1250	0.8750
$35 < x \leq 40$	7	0.0875	0.9625
$40 < x \leq 45$	3	0.0375	1.0000

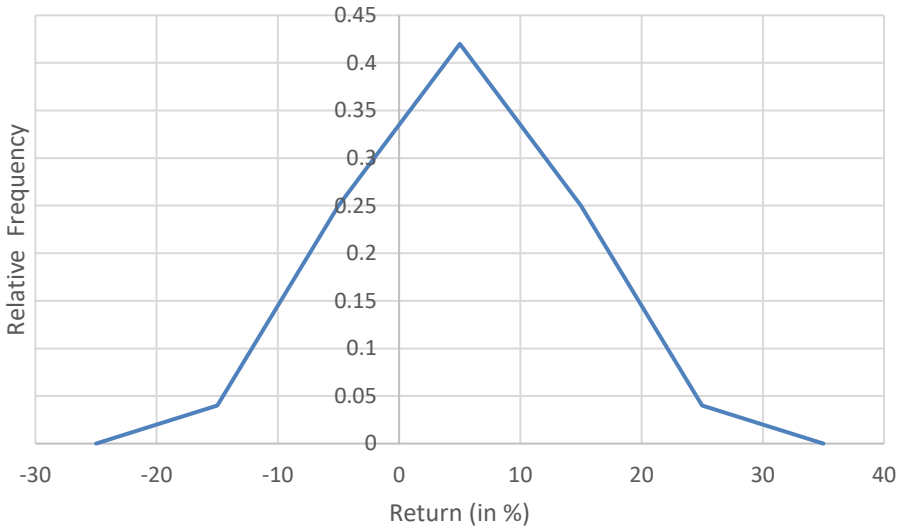
a. 0.375; 0.875; 0.125
b. It is not symmetric; it is positively skewed (see histogram).

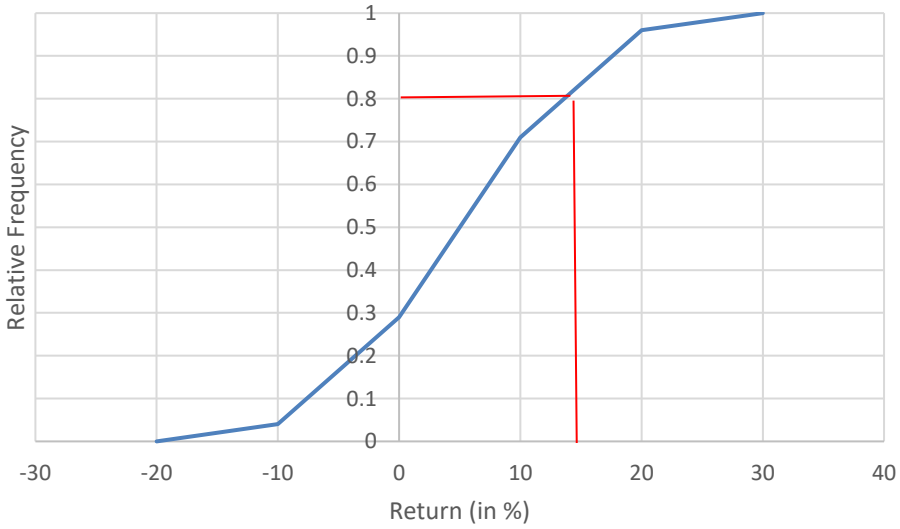


30.

Return (%)	Relative Frequency	Frequency	Cumulative Relative Frequency
$-20 < x \leq -10$	0.04	0.04(500) = 20	0.04
$-10 < x \leq 0$	0.25	125	0.04+0.25=0.29
$0 < x \leq 10$	0.42	210	0.71
$10 < x \leq 20$	0.25	125	0.96
$20 < x \leq 30$	0.04	20	1.00

a. 125; 0.71
b. The distribution is symmetric. (See polygon - it is mirror image of itself.)
c. 0.80 (See ogive.)



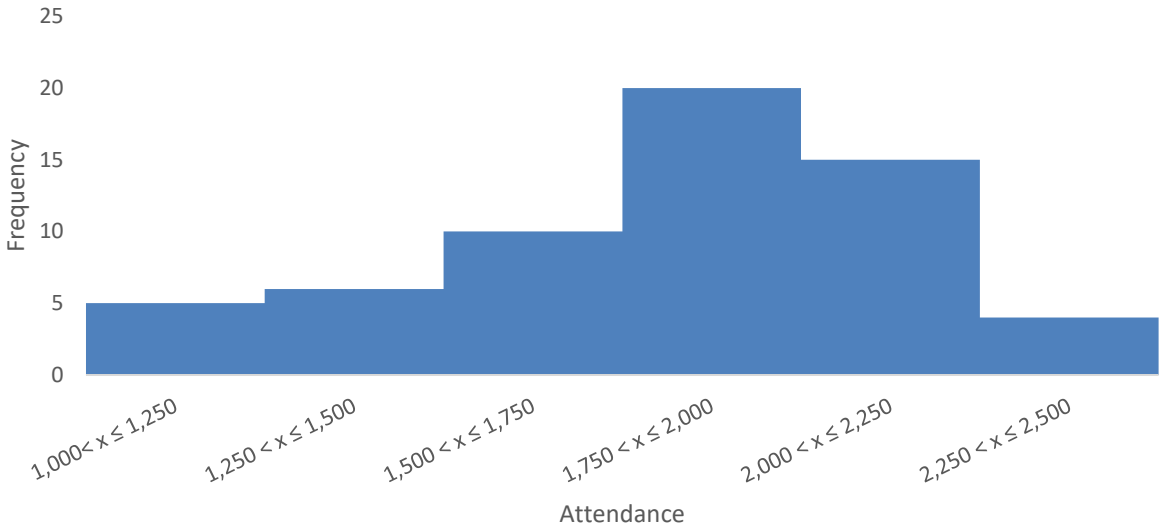


31.

Attendance	Frequency	Relative Frequency	Cumulative Relative Frequency
$1{,}000 < x \leq 1{,}250$	5	5/60 = 0.083	0.083
$1{,}250 < x \leq 1{,}500$	6	0.100	0.083+0.100 = 0.183
$1{,}500 < x \leq 1{,}750$	10	0.167	0.350
$1{,}750 < x \leq 2{,}000$	20	0.333	0.683
$2{,}000 < x \leq 2{,}250$	15	0.250	0.933
$2{,}250 < x \leq 2{,}500$	4	0.067	1.000

a. 0.33; 0.35; 0.65

b. The distribution is not symmetric; it is negatively skewed.

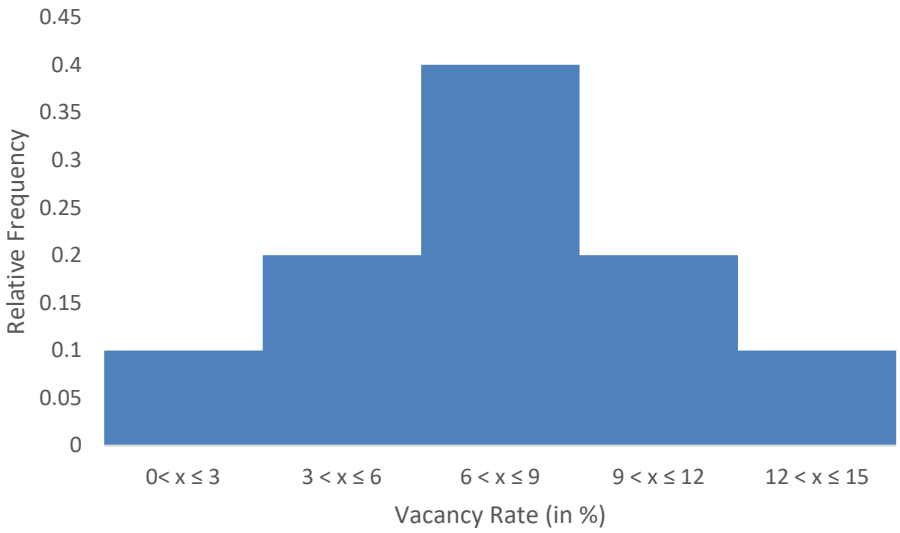


32.

Vacancy Rate (%)	Relative Frequency	Frequency	Cumulative Frequency	Cumulative Relative Frequency
$0 < x \leq 3$	0.10	0.10(50) = 5	5	0.10
$3 < x \leq 6$	0.20	10	5+10 = 15	0.10 +0.20 = 0.30
$6 < x \leq 9$	0.40	20	35	0.70
$9 < x \leq 12$	0.20	10	45	0.90
$12 < x \leq 15$	0.10	5	50	1.00

a. 20; 35; 0.30

b. The distribution is symmetric. (See histogram.)



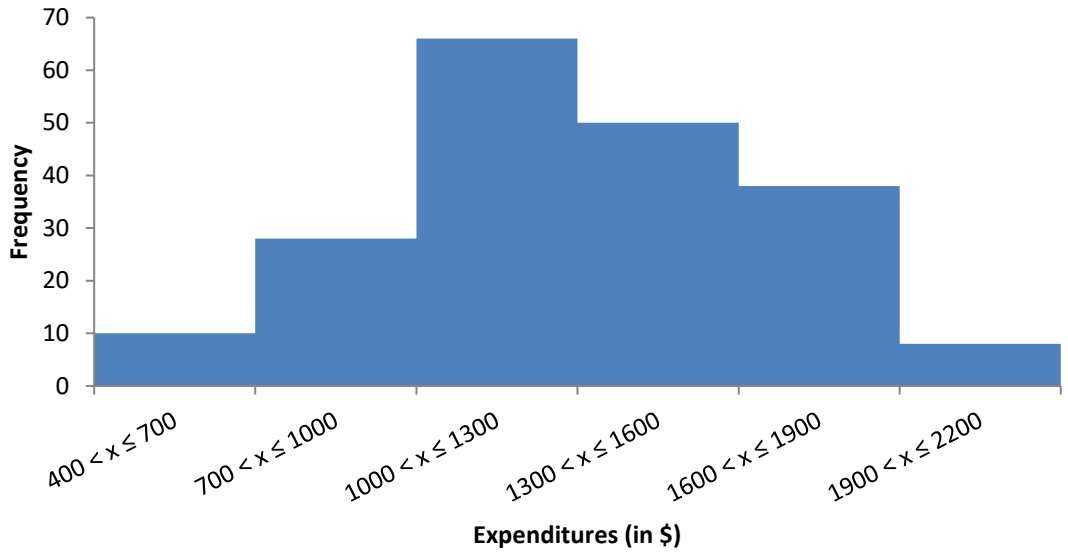
33.

a.

Expenditures	Frequency	Relative Frequency
400 < x ≤ 700	10	0.05
700 < x ≤ 1000	28	0.14
1000 < x ≤ 1300	66	0.33
1300 < x ≤ 1600	50	0.25
1600 < x ≤ 1900	38	0.19
1900 < x ≤ 2200	8	0.04

28 customers spent between $701 and $1,000; 0.48 of the customers spent more than $1,300.

b. The distribution is not symmetric.



The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(400, 2200, by=300)
intervals.cut <- cut(myData$Expenditures, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
#Obtain histogram
hist(myData$Expenditures, breaks=intervals, right=TRUE, col="blue")
```

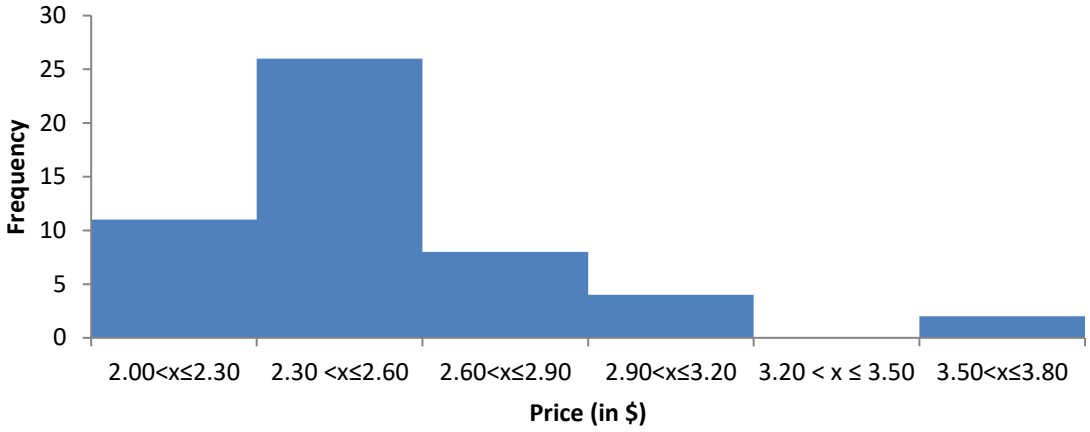
34.

a.

Gas Price	Frequency
$2.00 < x \leq 2.30$	11
$2.30 < x \leq 2.60$	26
$2.60 < x \leq 2.90$	8
$2.90 < x \leq 3.20$	4
$3.20 < x \leq 3.50$	0
$3.50 < x \leq 3.80$	2

The interval $2.30 < x \leq 2.60$ has the highest frequency. Six states (4+2=6) have average gas prices that are greater than $2.90.

b. No, the distribution is not symmetric. It is positively skewed.



The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(2.00, 3.80, by=0.30)
intervals.cut <- cut(myData$Price, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
#Obtain histogram
hist(myData$Price, breaks=intervals, right=TRUE, col="blue")
```

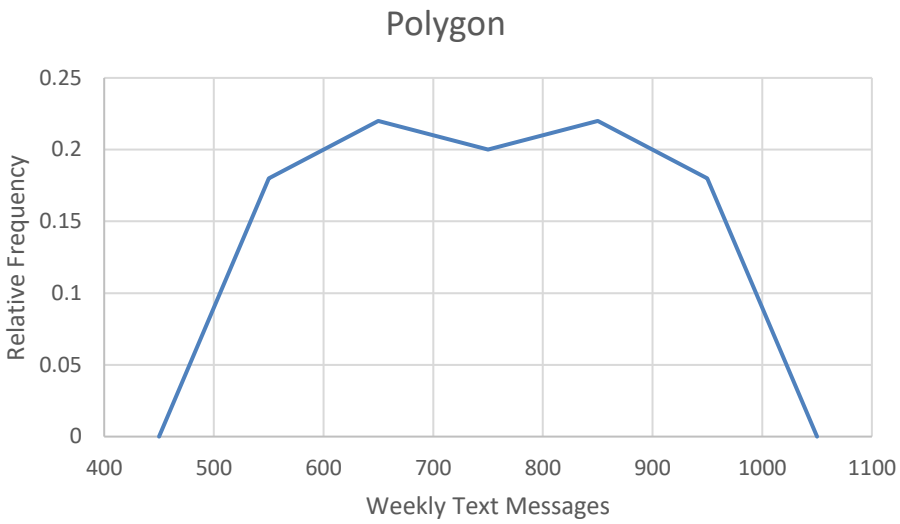

35.

a.

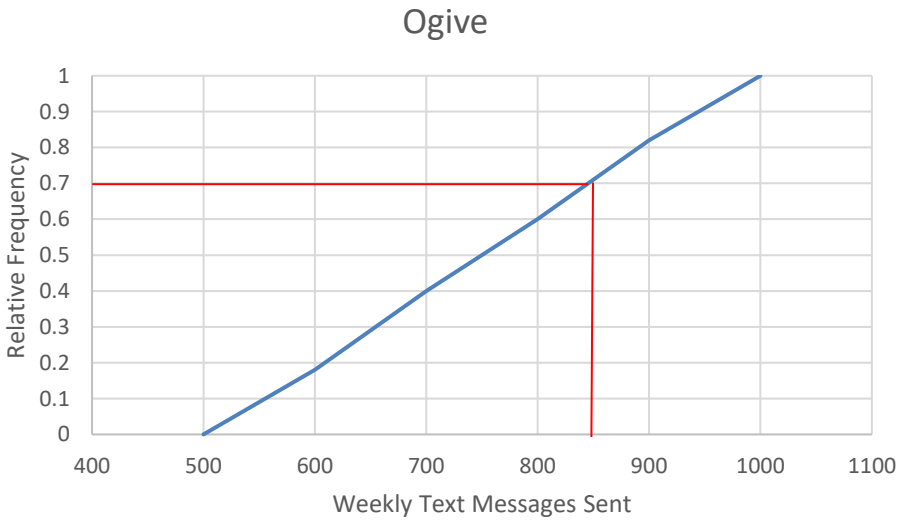
Text Messages	Frequency	Relative Frequency
500 < x ≤ 600	27	0.18
600 < x ≤ 700	33	0.22
700 < x ≤ 800	30	0.20
800 < x ≤ 900	33	0.22
900 < x ≤ 1,000	27	0.18

30 teens sent more than 700 text messages but no more than 800 text messages; 0.40 of the teens sent more than 800 text messages

b. The distribution is symmetric. (See polygon.)



c. Approximately 0.30 teens sent more than 850 text messages.



The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(500, 1000, by=100)
intervals.cut <- cut(myData$Texts, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
```

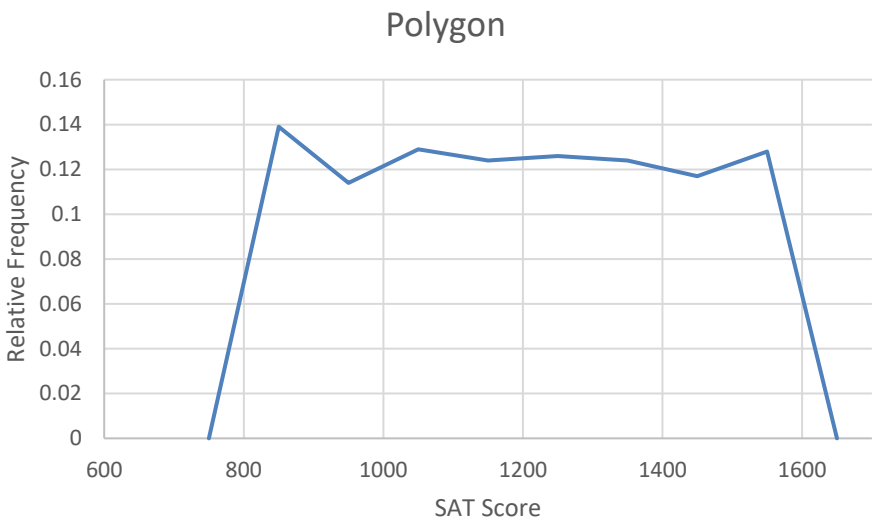
36.

a.

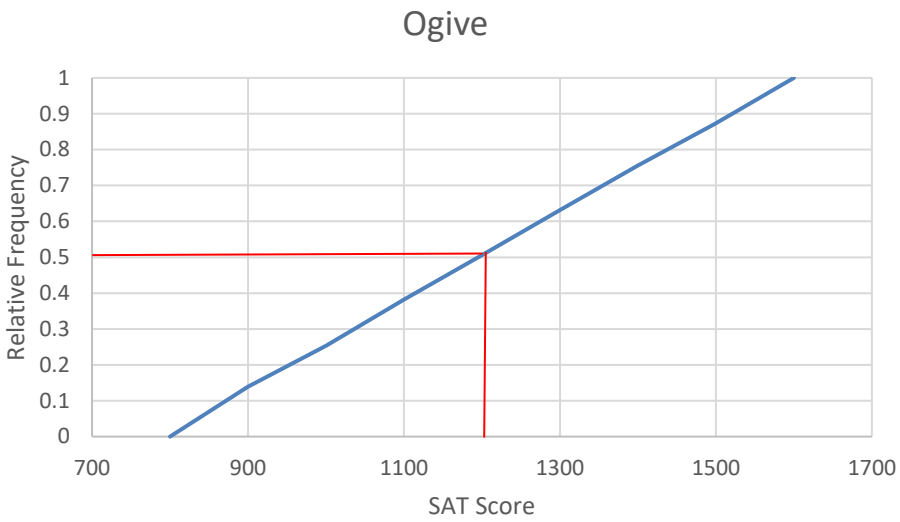
SAT Score	Frequency	Relative Frequency
$800 < x \leq 900$	171	0.139
$900 < x \leq 1000$	140	0.114
$1000 < x \leq 1100$	159	0.129
$1100 < x \leq 1200$	152	0.124
$1200 < x \leq 1300$	155	0.126
$1300 < x \leq 1400$	152	0.124
$1400 < x \leq 1500$	144	0.117
$1500 < x \leq 1600$	157	0.128

The interval with the highest frequency is $800 < x \leq 900$ (171 students); 157 students scored in the top interval; the proportion of students who scored in the interval $900 < x \leq 1000$ is 0.114.

b. The distribution is relatively symmetric. (See polygon)



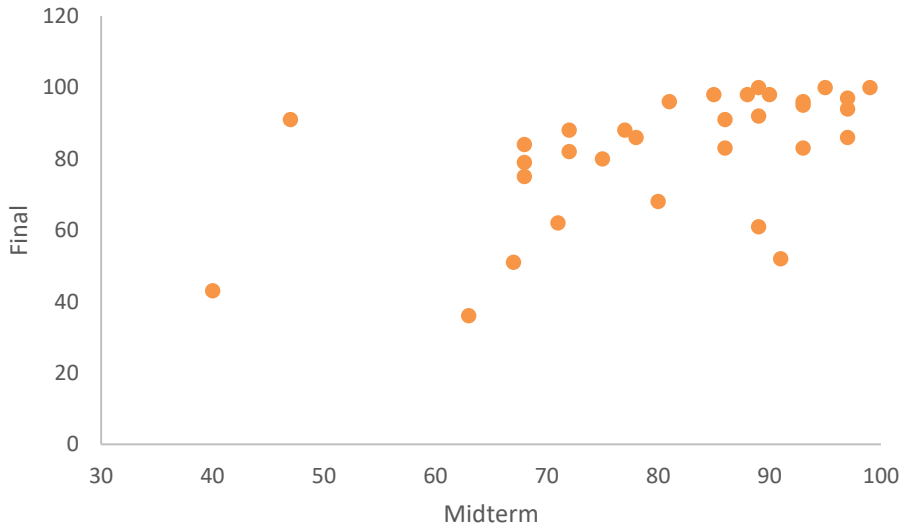
c. The proportion of students who scored below 1200 is approximately 0.50. (See ogive.)



The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(800, 1600, by=100)
intervals.cut <- cut(myData$SAT, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
```


37.

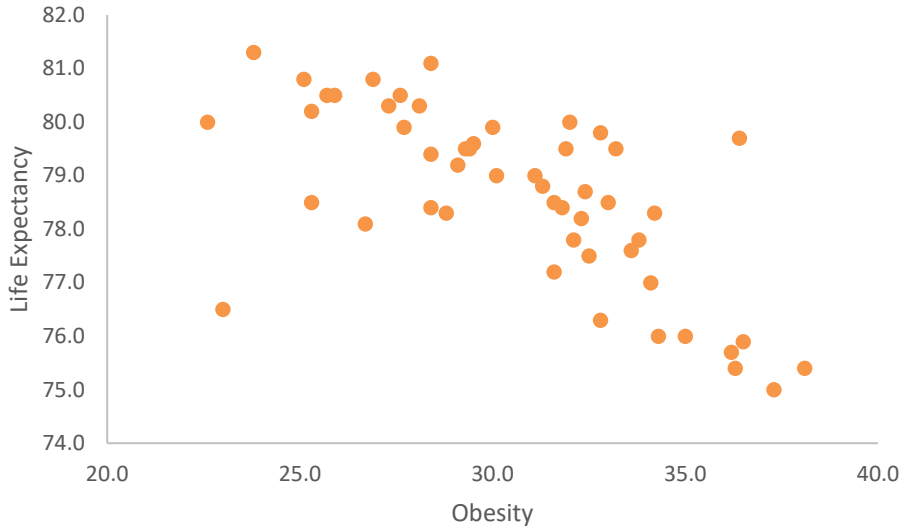


There appears to be a positive relationship between a student's midterm score and a student's final score.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$Final ~ myData$Midterm, col="chocolate", pch=16)
```

38.

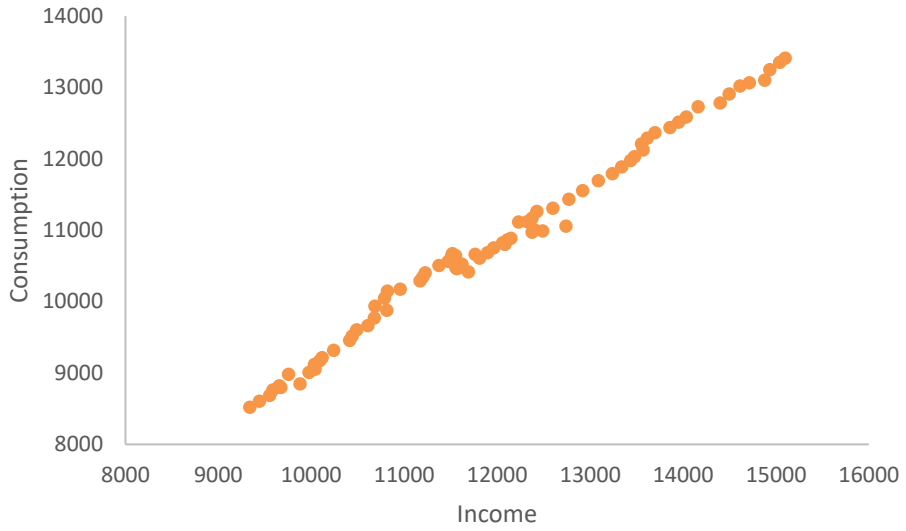


There appears to be a negative relationship between obesity and life expectancy.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$'Life Expectancy' ~ myData$Obesity, col="chocolate", pch=16)
```

39.

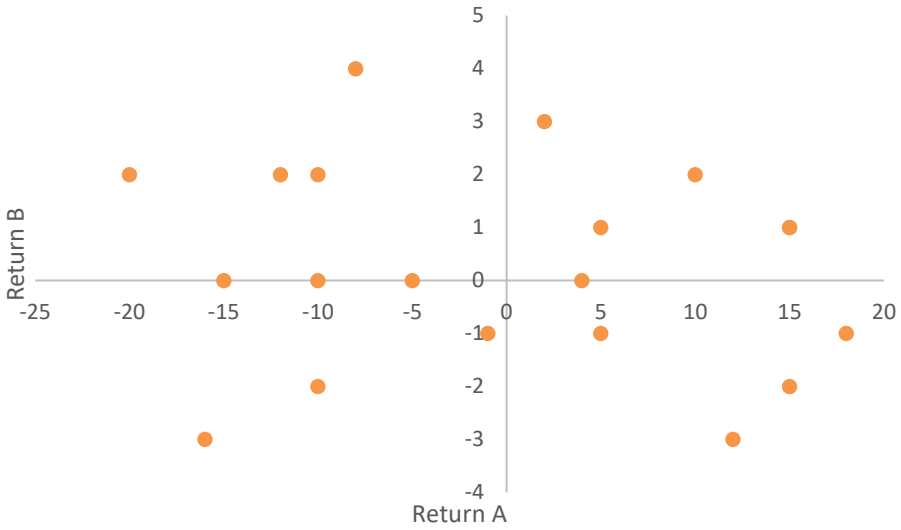


There appears to be a positive relationship between disposable income and consumption.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$Consumption~ myData$Income, col="chocolate", pch=16)
```

40.

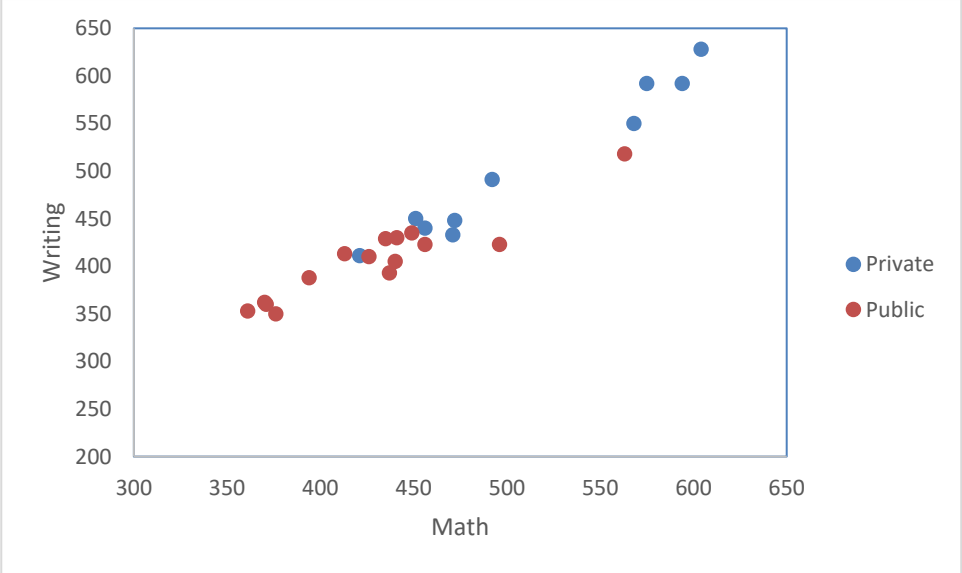


There appears to be no relationship between the returns of A and B, so investing in both would diversify risk.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$'Return B'~ myData$'Return A', col="chocolate", pch=16)
```


41.



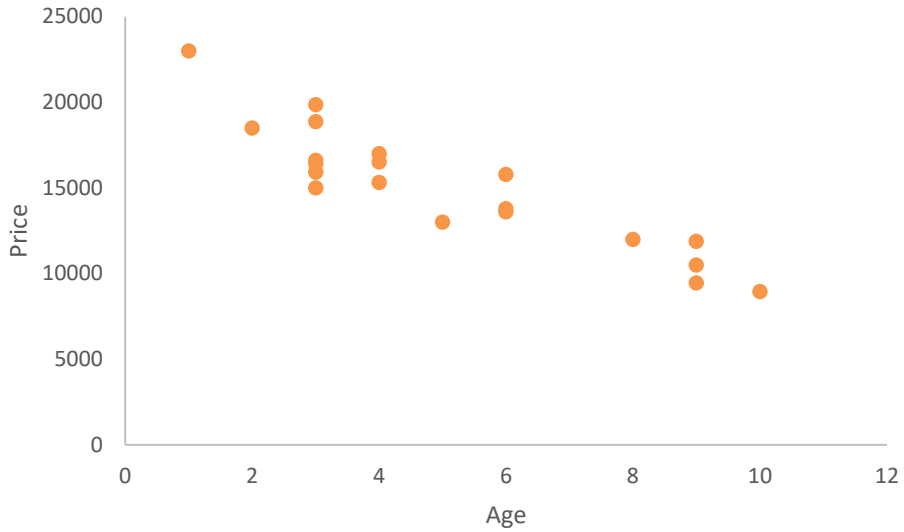
The scatterplot shows a positive linear relationship between the math and writing scores, and this relationship holds true for both public and private schools. The scatterplot suggests that private high schools tend to perform better than public high schools in terms of both math and writing scores.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot with categorical variable
plot(myData$Writing ~ myData$Math,pch=16, col=ifelse(myData$Type==”Private”, 20
  , 26))
legend(“bottomright”, legend=c(“Private”,”Public”), pch=16, col=c(20,26))
```

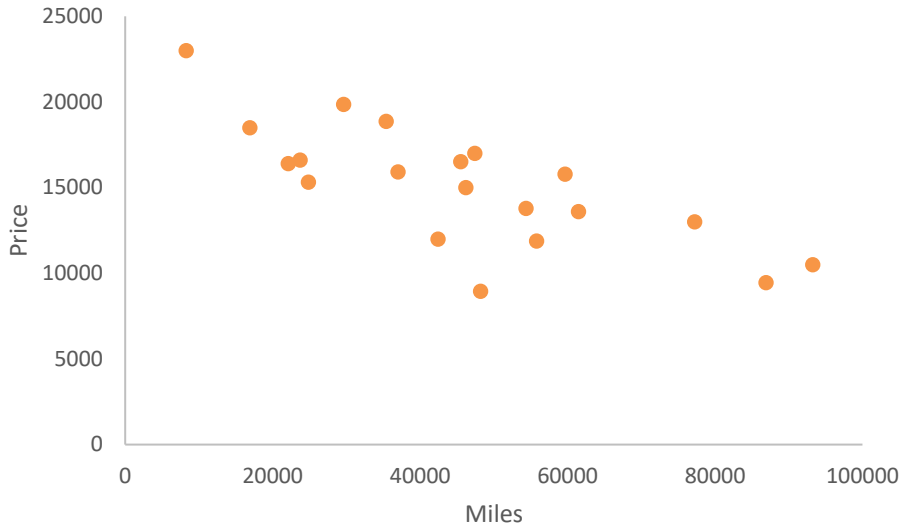
42.

a.



There appears to be a negative relationship between the price of a car and its age.

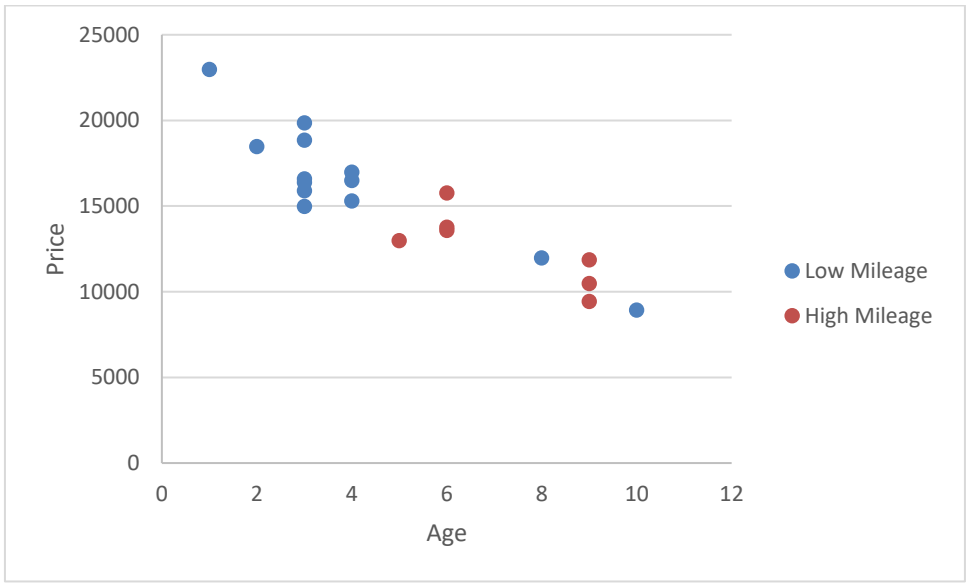
b.



There appears to be a negative relationship between the price of a car and its mileage.

c. Seven cars have mileage greater than 50,000.

d.

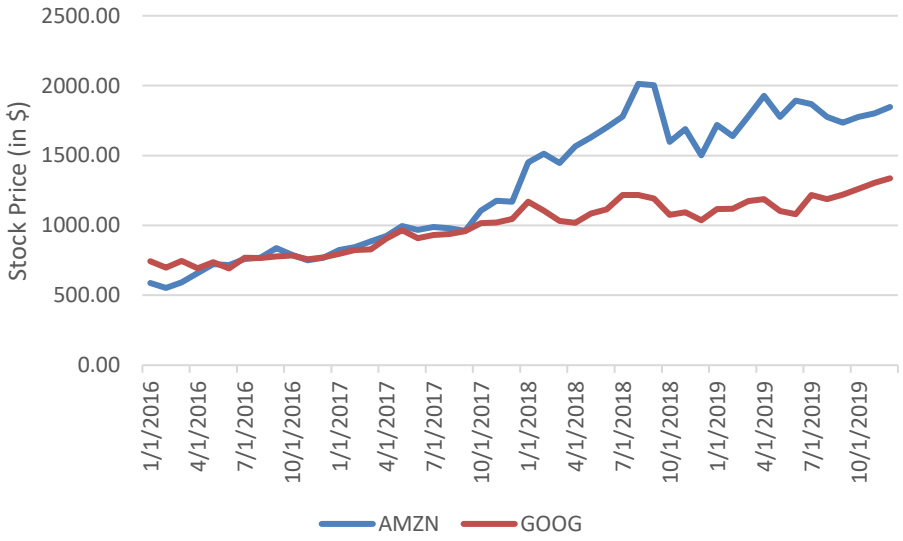


The negative relationship between age and price is consistent for cars of both mileage categories.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplots
plot(myData$Price~ myData$Age)
plot(myData$Price~ myData$Mileage)
#Converting Mileage to categorical variable
myData$Miles_Cat<-ifelse(myData$Miles<50000,"Low_Mileage","High_Mileage")
table(Miles_Cat)
# Obtain scatterplot with categorical variable
plot(myData$Price ~ myData$Age, pch=16,
     col=ifelse(myData$Miles_Cat==”Low_Mileage”, 20, 26))
legend(“bottomright”, legend=c(“Low_Mileage”,”High_Mileage”), pch=16,
     col=c(20,26))
```

43.

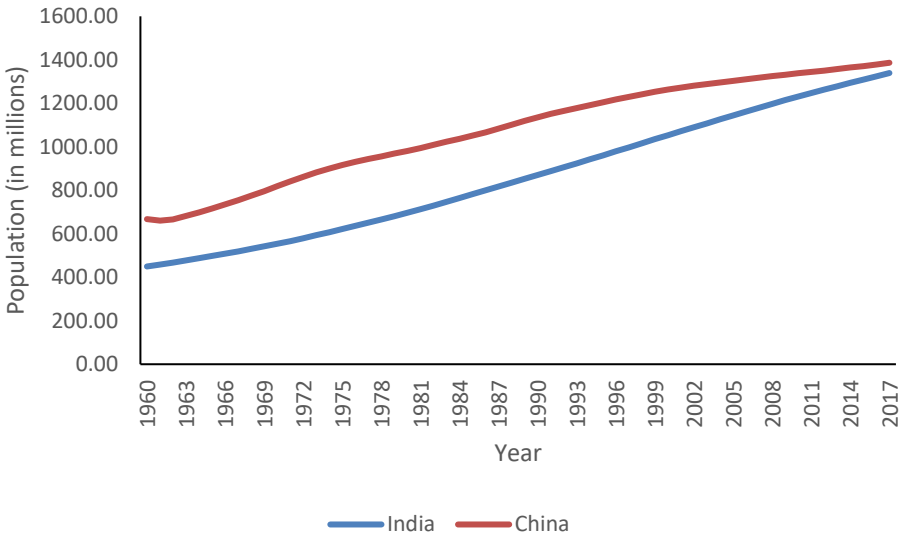


GOOG has an upward trend. Amazon seems to have a greater trajectory, even though it had a steep drop in the latter half of 2018.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain line chart for AMZN
plot(myData$AMZN~ myData$Date, col=”blue”, type=”l”)
#Incorporate red line for GOOG
lines(myData$GOOG~myData$Date, col=”red”, type=”l”)
#Incorporate legend
legend(“topleft”, legend=c(“AMZN”,”GOOG”), col=c(“blue”,”red”), lty=1)
```

44.



Both countries have a clear and consistent upward trend, but China's begins to stall slightly around 2000. India has slightly higher population growth over the past 40 years.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain line chart for India
plot(myData$India~ myData$Year, col="blue", type="l")
#Incorporate red line for China
lines(myData$China~myData$Year, col="red", type="l")
#Incorporate legend
legend("topleft", legend=c("India","China"), col=c("blue","red"), lty=1)
```


45.

Stem	Leaf
2	3 6 8
3	0 2 2 3 4 5 6
4	0 2 2 2 5 6 7
5	4 5 5

a. Observations ranged from a low of 2.3 to a high of 5.5.
b. The distribution is symmetric; it is a mirror image of itself.

46.

Stem	Leaf
-8	7 5 5 3 2 0 0 0
-7	9 7 5 3 3 2 1
-6	5 5 4
-5	2 0

(Keep in mind that these values are negative.)

a. Observations ranged from a low of -87 to a high of -50.
b. The distribution is not symmetric; it is positively skewed. Most of the numbers are in the lower stems of -8 and -7.

47.

Stem	Leaf
99	6 7 8
100	4 5
101	0 2 2 2 3 5 5 6
102	0 1 2 2 3 4 5

a. Temperatures ranged from a low of 99.6 to a high of 102.5.
b. The distribution is not symmetric; it has negative skew. The majority of patients recorded a temperature higher than 101.

48.

Stem	Leaf
7	3 4 6 7 8 8
8	0 1 2 3 4 4 4 4 7 8
9	0 0 0 1 1 2 2 2 3 3 4 4 4 4 4 5 6 6 6 8 8 9
10	6 7

Temperatures ranged from a low of 73 to a high of 107. The distribution is not symmetric; it has negative skew. Temperatures in 90s were the most frequent.

49.

Stem	Leaf
6	5 5 6 7 7
7	0 0 0 1 1 2 2 3 3 3 5 5 8 9
8	0 0 0 1 1 2

The officers' concerns are warranted. The data shows that the majority of cars exceed the 65 miles-per-hour limit.

50.

France

Stem	Leaf
1	9
2	1 2 2 2 2 2 3 3 4 4 5 5 5 5 7 7 9
3	1 1 1 2 3

Croatia

Stem	Leaf
2	1 2 3 3 4 4 5 6 6 6 8 8 9 9 9 9
3	0 0 0 2 2 2 3

France has a relatively younger team compared to Croatia. France's ages range from 19 to 33, while Croatia's ages range from 21 to 33. The majority of players on both teams are in their 20s. However, Croatia has a couple of more players in their 30s as compared to France.

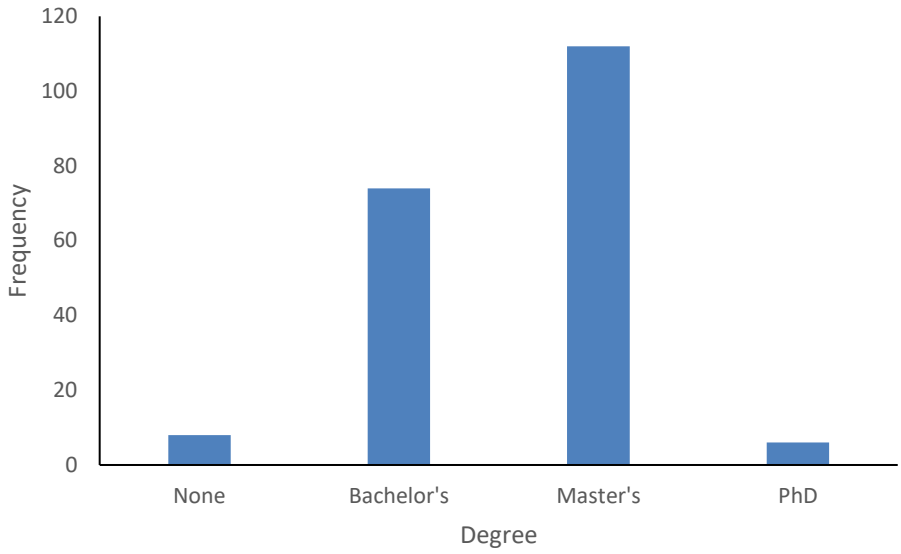
51.

a.

Degree	Frequency	Relative Frequency
None	8	0.04
Bachelor's	74	0.37
Master's	112	0.56
PhD	6	0.03

Yes, out of 200 CEOs more than half have advance degrees (112 + 6 =118); only 4% of CEOs do not have at least a Bachelor's degree.

b.



The majority of CEOs have a Bachelor's or a Master's degree. Very few CEOs have no degree or a PhD.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
Frequency <- table(myData$Degree); Frequency
prop.table(Frequency)
#Obtain bar chart
barplot(Frequency, col="blue")
abline(v=0)
```

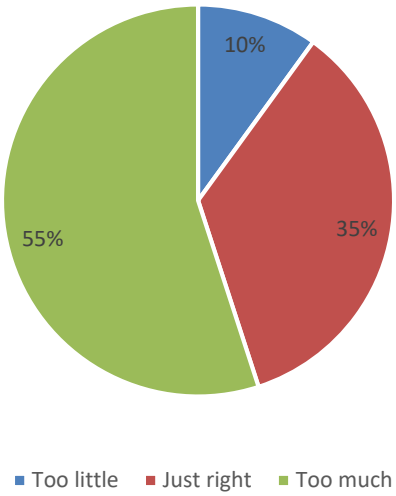

52.

a.

Response	Frequency	Relative Frequency
Too little	40	0.10
Just right	140	0.35
Too much	220	0.55

220 – number of respondents who felt that parents do too much for adult children; 0.10 – proportion of respondents who felt that parents do too little for adult children

b.



The majority of respondents, 55%, felt that parents do too much for their adult children; only 10% of respondents felt that parents do too little for their adult children.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
Frequency <- table(myData$Response); Frequency
prop.table(Frequency)
#Obtain pie chart
pie(Frequency)
```

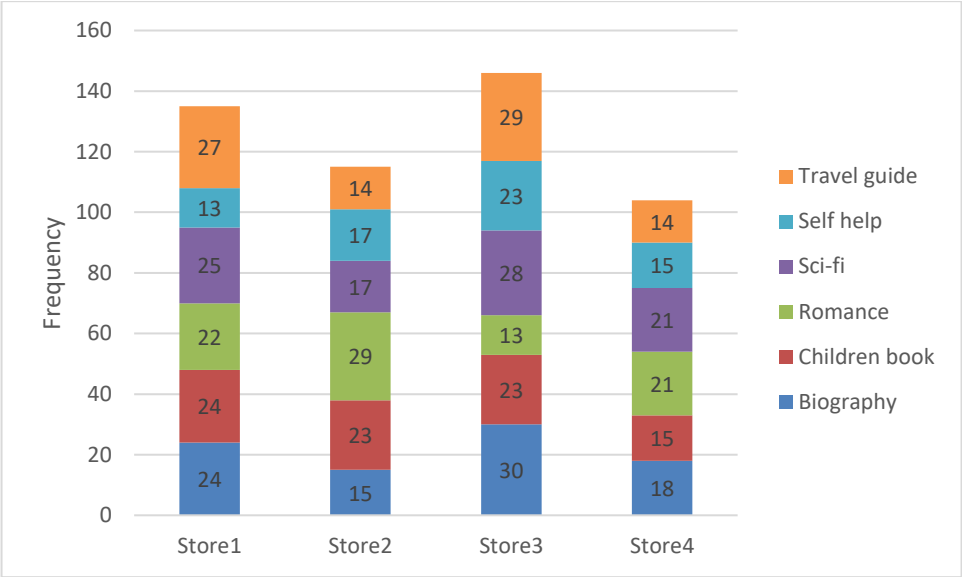

53.

a.

	BookType						
BookStore	Biography	Children book	Romance	Sci-fi	Self help	Travel guide	**Total**
Store1	24	24	22	25	13	27	**135**
Store2	15	23	29	17	17	14	**115**
Store3	30	23	13	28	23	29	**146**
Store4	18	15	21	21	15	14	**104**
Total	**87**	**85**	**85**	**91**	**68**	**84**	**500**

Of all stores, Store3 had the most transactions (146); of all book types sci-fi had the most transactions (91).

b.



Store2 has the most transactions for Romance books (29). Store3 has the most transactions for Sci-fi books.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain contingency table
myTable <- table(myData$BookType, myData$BookStore); myTable
prop.table(myTable)
#Obtain stacked column chart
barplot(myTable, legend=rownames(myTable), col=c(1,2,3,4,5,6))
```

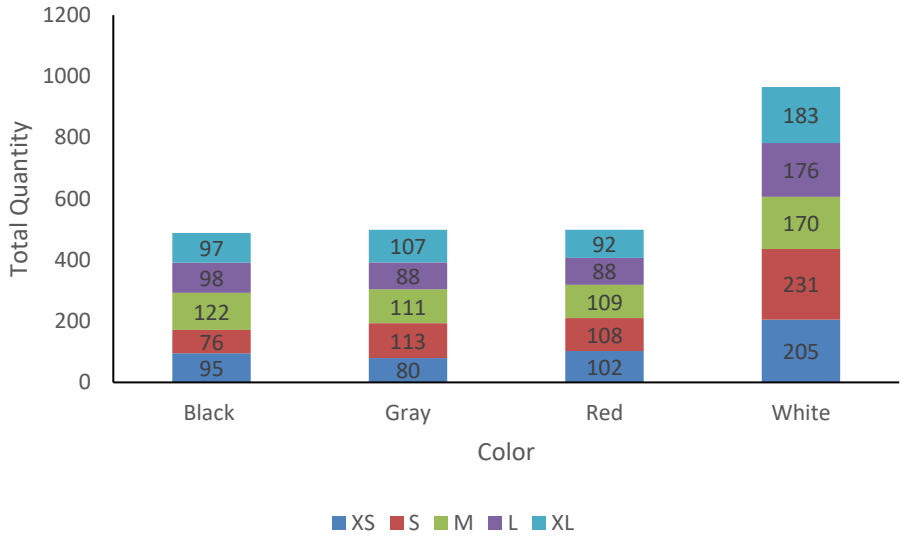
54.

a.

	Size					
Color	XS	S	M	L	XL	**Total**
Black	95	76	122	98	97	488
Gray	80	113	111	88	107	499
Red	102	108	109	88	92	499
White	205	231	170	176	183	965
Total	**482**	**528**	**512**	**450**	**479**	**2451**

109 M red t-shirts were sold; 183 XL white t-shirts were sold

b.



White S and White XS are the two most popular color/size combinations; Black S and Gray XS are the two least popular color/size combinations.

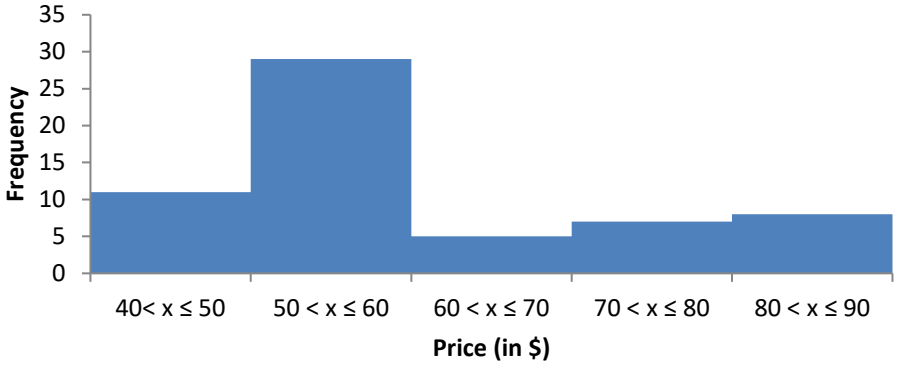
55.

Stock Price ($)	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
$40 < x \leq 50$	11	0.183	11	0.183
$50 < x \leq 60$	29	0.483	40	0.667
$60 < x \leq 70$	5	0.083	45	0.750
$70 < x \leq 80$	7	0.117	52	0.867
$80 < x \leq 90$	8	0.133	60	1

a. $50 < x \leq 60$; 11

b. 0.483; 45; 0.25

c. The distribution is not symmetric. It is positively skewed (See histogram.)



The R Code

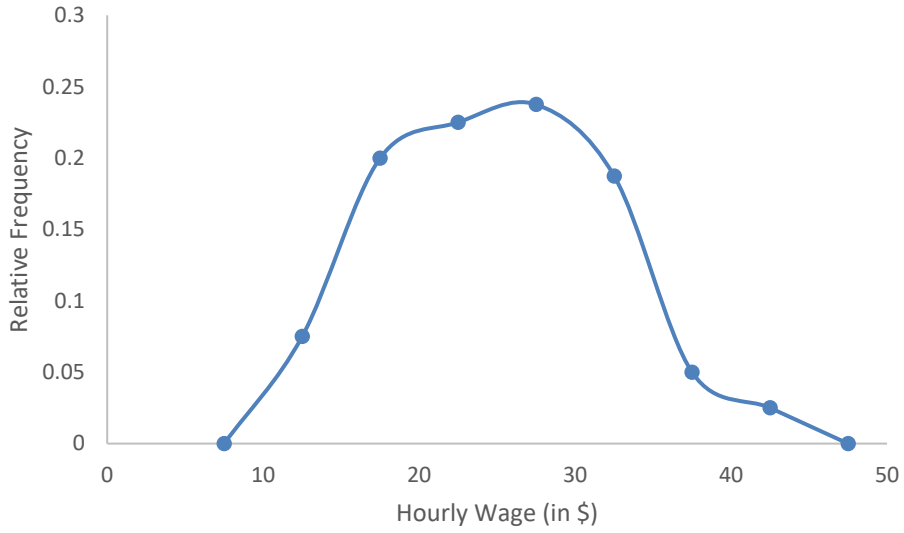
```
# Import the data into a data frame (table) and label it myData.
# Obtain frequency distribution
intervals <- seq(40, 90, by=10
intervals.cut <- cut(myData$Price, intervals, left=FALSE, right=TRUE)
myTable <- table(intervals.cut); myTable
View(myTable)
#Obtain histogram
hist(myData$Price, breaks=intervals, right=TRUE, col="blue")
```

56.

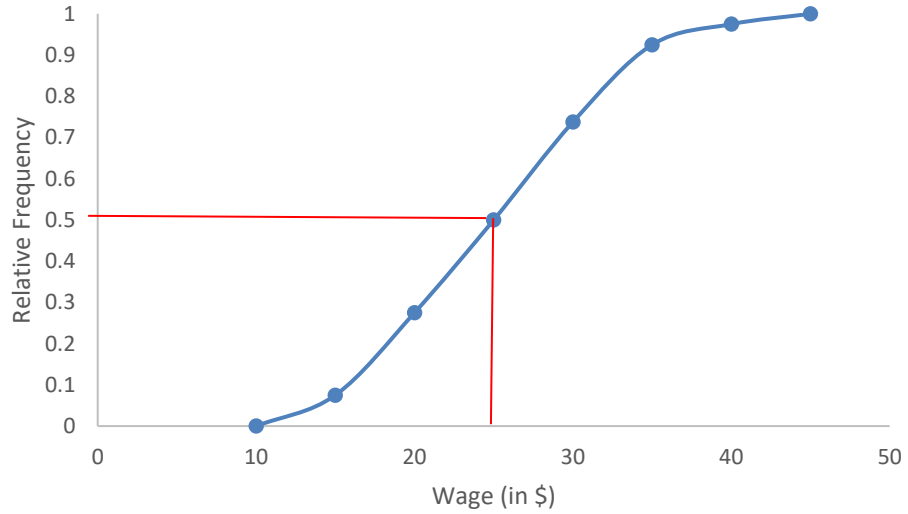
Hourly Wage ($)	Frequency	Relative Frequency	Cumulative Relative Frequency
10 < x ≤ 15	6	0.075	0.075
15 < x ≤ 20	16	0.200	0.275
20 < x ≤ 25	18	0.225	0.500
25 < x ≤ 30	19	0.238	0.738
30 < x ≤ 35	15	0.188	0.925
35 < x ≤ 40	4	0.050	0.975
40 < x ≤ 45	2	0.025	1.000

a. 0.851 (= 0.200 + 0.225 + 0.238 + 0.188); 0.075

b. The distribution is not approximately symmetric. (See polygon.)

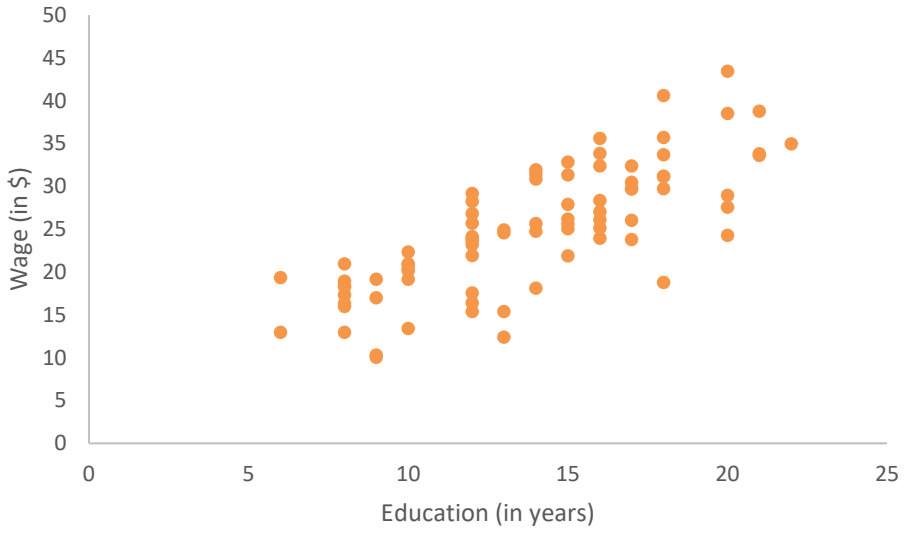


c. 0.50 (See the ogive.)



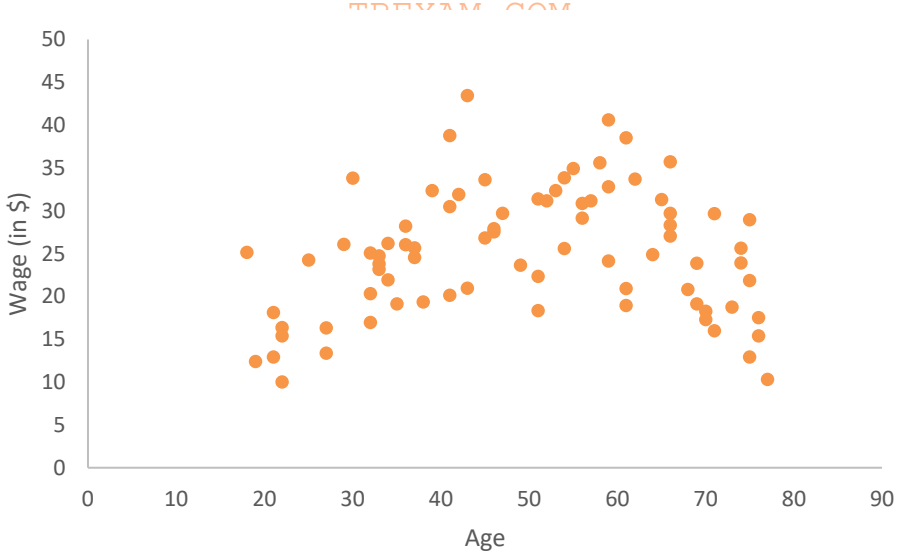
57.

a.



There appears to be a positive relationship between Wage and Education.

b.



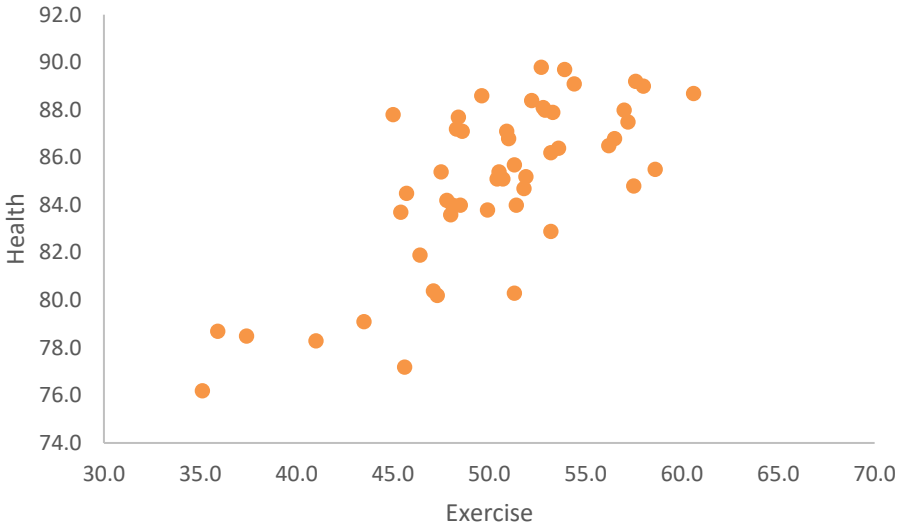
There appears to be a nonlinear relationship, an inverted U-shape relationship, between Wage and Age. The relationship between Wage and Age is initially positive and then after about age 50, seems to be negative.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplots
plot(myData$Wage~ myData$Education, col="chocolate", pch=16)
plot(myData$Wage~ myData$Age, col="chocolate", pch=16)
```

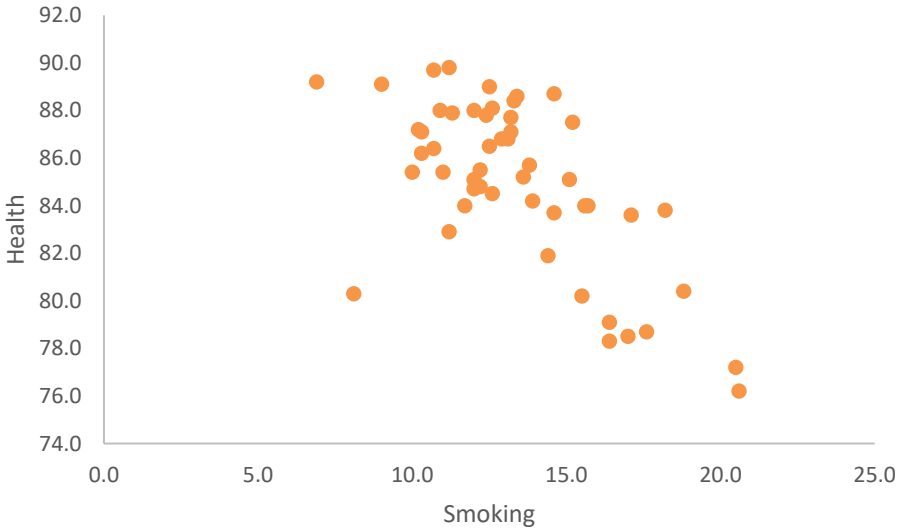
58.

a.



There appears to be a positive relationship between health and exercise.

b.



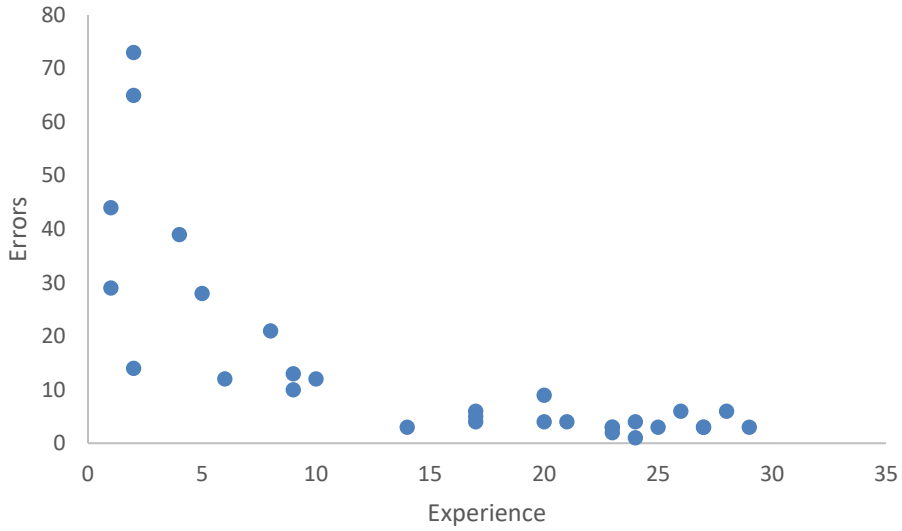
There appears to be a negative relationship between health and smoking.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplots
plot(myData$Health~ myData$Exercise, col="chocolate", pch=16)
plot(myData$Health~ myData$Smoking, col="chocolate", pch=16)
```

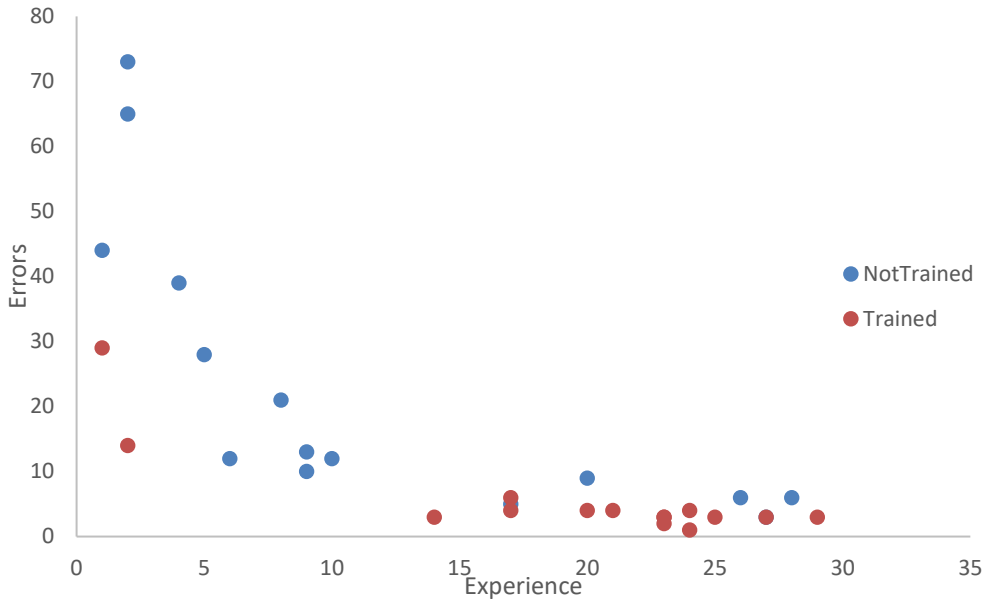
59.

a.



There appears to be a negative, somewhat nonlinear, relationship between pick errors and experience.

b.



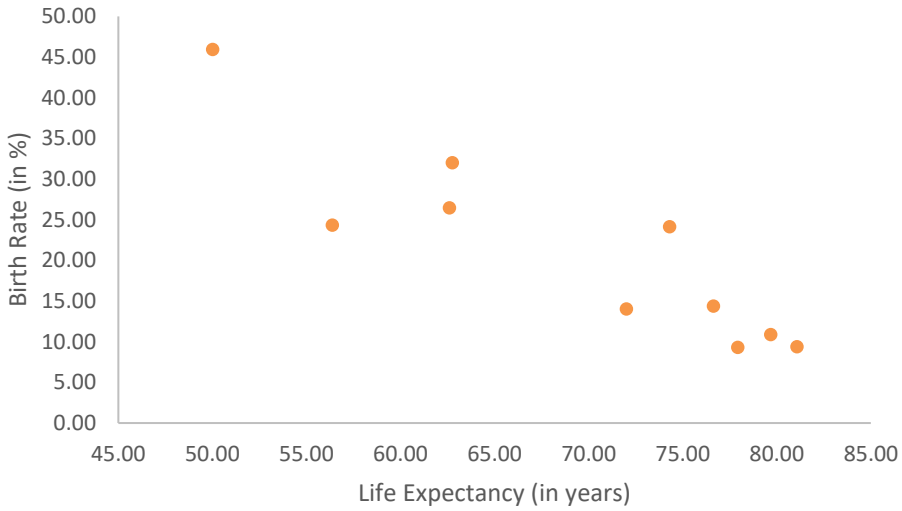
We still see a negative, somewhat nonlinear, relationship between pick errors and experience. We now also see that, in general, trained employees have relatively less pick errors and more experience compared to untrained employees.

The R Code


```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$Errors~ myData$Exper, col=”chocolate”, pch=16)
# Obtain scatterplot with categorical variable
plot(myData$Errors ~ myData$Exper, pch=16, col=ifelse(myData$Train==0, 20, 26))
legend(“right”, legend=c(“No Training”,”Training”), pch=16, col=c(20,26))
```

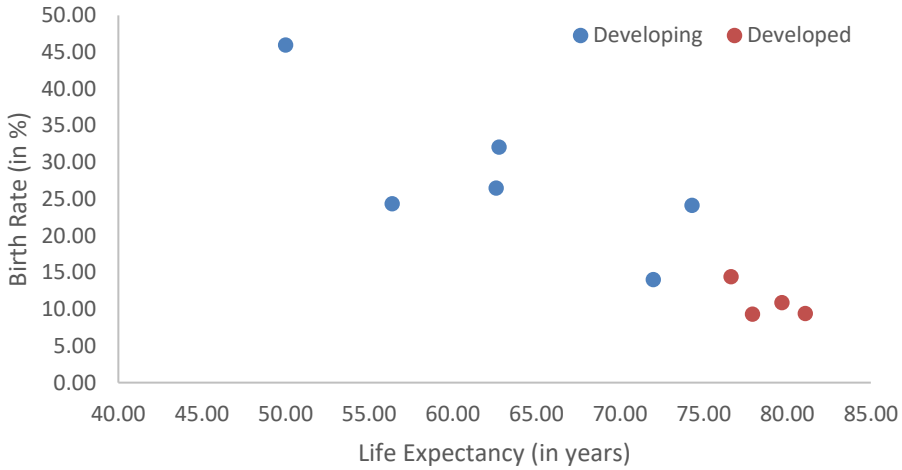
60.

a.



There appears to be a negative relationship between birth rate and life expectancy. That is countries with lower birth rates tend to have higher life expectancies.

b.

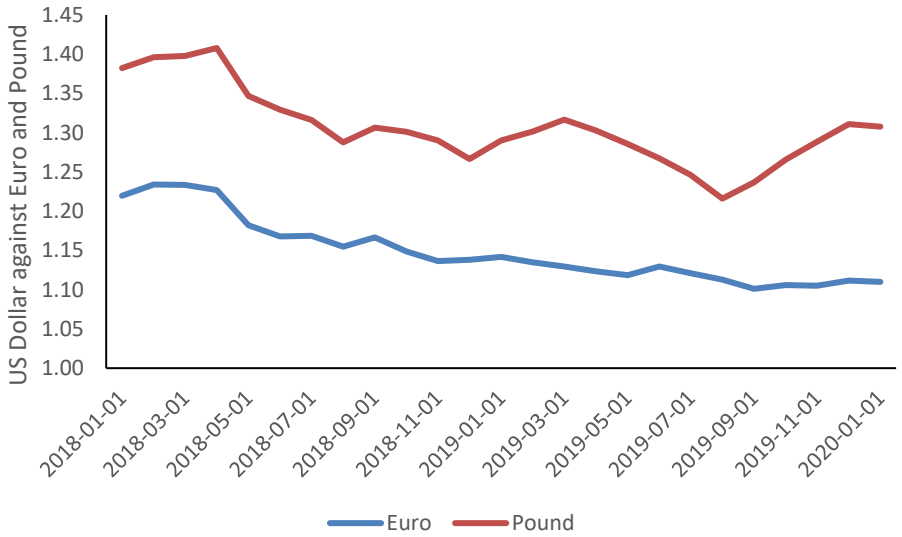


We still see a negative relationship between birth rate and life expectancy. We now also see that, in general, developed countries have lower birth rates and higher life expectancies compared to developing countries.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain scatterplot
plot(myData$Birth_Rate~ myData$Life_Exp, col="chocolate", pch=16)
# Obtain scatterplot with categorical variable
plot(myData$Birth_Rate ~ myData$Life_Exp, pch=16,
     col=ifelse(myData$Development=="Developing", 20, 26))
legend("right", legend=c("Developing","Developed"), pch=16, col=c(20,26))
```

61.



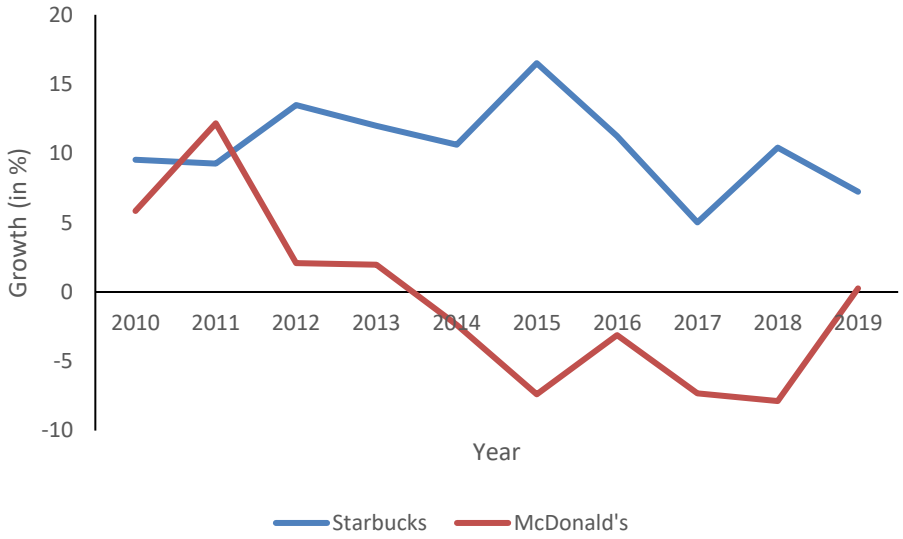
Over this time period, we see a steady decline in the USD against the Euro which means that the USD has been appreciating over this period. In January 2018, the USD against the Euro was 1.22 meaning that one euro could buy 22 cents more in goods and services than one dollar could; however, in January 2020, the USD against Euro was 1.11, meaning that one euro could by 11 cents more in goods and services than one dollar could.

Over this time period, we see a volatile decline in the USD against the Pound until around August 2019 (again implying appreciation of the USD). This low point corresponded to uncertainty surrounding Brexit. Since that time, the USD against the Pound has risen, almost reaching the same level as January 2018.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain line chart for Euro
plot(myData$Euro~ myData$Date, col="blue", type="l", ylim=c(1,1.5))
#Incorporate red line for Pound
lines(myData$Pound~myData$Date, col="red", type="l")
#Incorporate legend
legend("bottomleft", legend=c("Euro","Pound"), col=c("blue","red"), lty=1)
```

62.



From 2010 through 2019, Starbucks has seen positive growth rates in revenues in every year with a high of 16.51% in 2015 and a low of 5.02% in 2017. McDonald's, on the other hand, has seen quite a bit of volatility in its growth rates. It experienced five years of positive growth rates and five years of negative growth rates. Fortunately, for McDonald's, it appears that its growth rates might be trending back to positive.

The R Code

```
# Import the data into a data frame (table) and label it myData.
# Obtain line chart for Euro
plot(myData$Starbucks~ myData$Year, col="blue", type="l",ylim=c(-10,20))
#Incorporate red line for Pound
lines(myData$"McDonald's"~myData$Year, col="red", type="l")
#Incorporate legend
legend("bottomleft", legend=c("Starbucks","McDonald's"), col=c("blue","red"), lty=1)
```

63.

Stem	Leaf
3	5
4	6 6 8
5	4 6 6
6	2 3 4
7	0 0 1 5 5 5 7
8	0 1 2 3 4 6 8 9

a. Ages ranged from 35 to 89.
b. The distribution is not symmetric; it has negative skew. The majority of people are 70 years of age or older.

64.

Stem	Leaf
0	8 8 9 9
1	0 0 1 1 2 2 2 2 3 3 4 4 5 6 6 8 8 9 9 9
2	0 0 9 9
3	0 7

a. PEG ratios ranged from 0.8 to 3.7. The 1.0 to 1.9 range had the highest frequency.
b. The distribution is not symmetric; it is positively skewed. There are a few firms with relatively high PEG ratios.