

Chapter 4 #13, 21, 24, 32, 40

**13. Summaries.**

a) (1 points) The mean price of the electric smooth top ranges is \$1001.50.

b) (2 points) In order to find the median and the quartiles, the list must be ordered.

565 750 850 900 1000 1050 1050 1200 1250 1400

The median price of the electric smooth top ranges is \$1025.

Quartile 1 = \$850 and Quartile 3 = \$1200.

c) (1 points) The range of the distribution of prices is  $\text{Max} - \text{Min} = \$1400 - \$565 = \$835$ .

The IQR =  $Q3 - Q1 = \$1200 - \$850 = \$350$ .

**21. The Great One.**

a) (3 points) The stem and leaf plot is shown below.

0 4

0 66

0 99

1 0

1 233

1 44

1 666

1 89

2 0011

points scored

(2|0 means 200 points)

b) (1 points) The distribution looks slightly left skewed. The scores range from about 40–210 (Note that these are truncated values). The median (the average of the 10th and the 11th values) is about 140. There are no outliers.

c) (3 points) If we consider this as a left-skewed distribution, then the five-number summary is an appropriate summary and is given below.

**Descriptive Statistics: points scored**

Variable	Minimum	Q1	Median	Q3	Maximum
points scored	48.0	98.3	145.5	192.8	215.0

**24. Family income.**

a) (3 points) This is because the distribution of income is right skewed because of the relatively smaller proportion of very rich families. Shown below is a histogram of a simulated data set having mean and median close values given.

**Descriptive Statistics: Income05**

Variable	N	Mean	Q1	Median	Q3
Income05	3468	70348	27747	57887	102918

b) (3 points) In 2005, the mean is  $(70300-57700)/57700 = 21.8\%$  higher compared with the median. In 1980, the mean is  $(62300-58300)/58300 = 6.9\%$  higher compared with the median. This shows that the distribution has become more skewed. This means the relatively small proportion of the rich has gotten richer compared to the larger proportion of the poor. Shown below is a histogram of simulated data sets having mean and median close values given.

**Descriptive Statistics: Income80, Income05**

Variable	N	Mean	Q1	Median	Q3
Income80	3468	62283	27362	58320	86471
Income05	3468	70348	27747	57887	102918

**32. (1 points) Acid rain.** The distribution of the pH readings of water samples in Allegheny County, Pennsylvania, is bimodal. A roughly uniform cluster is centred on a pH of 4.4. This cluster ranges from pH of 4.1 to 4.9. Another smaller, tightly packed cluster is centred on a pH of 5.6. Two readings in the middle seem to belong to neither cluster.

**40. Incarceration rates.**

a) (3 points) The histogram of incarceration rates is given below. A stem and leaf plot is also an appropriate display.

```

0 2222222222233333333333444
0 55555555555666666666677777777777888888888888888899999999999
1 0000000000111111122222222333333333444444
1 55555555555666667777778889999999
2 001122334444
2 55666778899
3 00112224444
3 556667799
4 00122
4 6678
5 34
5 8
6 3
6
7
7 5

```

Incarceration rate

(6|3 means 630 prisoners per 100 000 of nation's population)

b) (3 points) For skewed distributions the five-number summary is more appropriate.

**Descriptive Statistics: Incarceration rate**

Variable	N	Minimum	Q1	Median	Q3	Maximum
Incarceration rate	217	22.00	75.50	123.00	213.00	751.00

c) (3 points) The distribution of incarceration rates is right skewed with median 123 and interquartile range  $213.0 - 75.5 = 137.5$ . The maximum value (751.0 in US) is an outlier. (In fact, the  $1.5 \times \text{IQR}$  rule shows many outliers: any value greater than  $Q3 + 1.5 \times \text{IQR} = 419.25$  is an outlier). Canada's incarceration rate is 108 and this is below the median but greater than the first quartile.