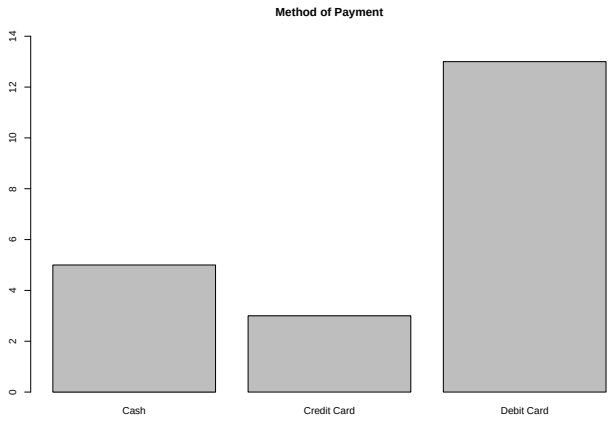
Stat 151.003 Fall 2006 (Kozdron)
Solutions to Assignment #1

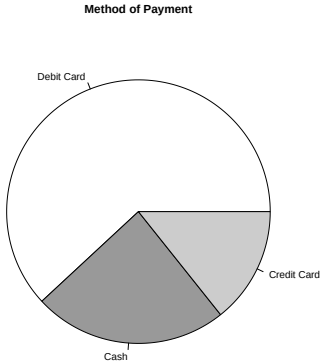
1. (a) This sample contains $n = 21$ observations. There are two observed variables, namely (i) amount spent, and (ii) method of payment. The observational unit is customer. (That is, for each customer in the sample, two variables were observed for that customer, namely amount spent and method of payment.) The amount spent variable is a quantitative variable which can be regarded as continuous. (The text indicates that discrete variables can be assigned the values $0, 1, 2, 3, \ldots$. For this reason, amount spent is classified as continuous.) The amount spent data can also be classified as ratio-level data. Method of payment, however, is categorical. Since there is no inherent ranking to the method of payment, this is a variable of the unordered type. That is, the method of payment data can be classified as ordinal-level data.

1. (b) One example of a tabular summary for the method of payment data is as follows:

Method of Payment	Frequency	Relative Frequency
Debit Card	13	$13/21 = 0.619$
Cash	5	$5/21 = 0.238$
Credit Card	3	$3/21 = 0.143$

Two possible graphical summaries are a bar graph and a pie chart.

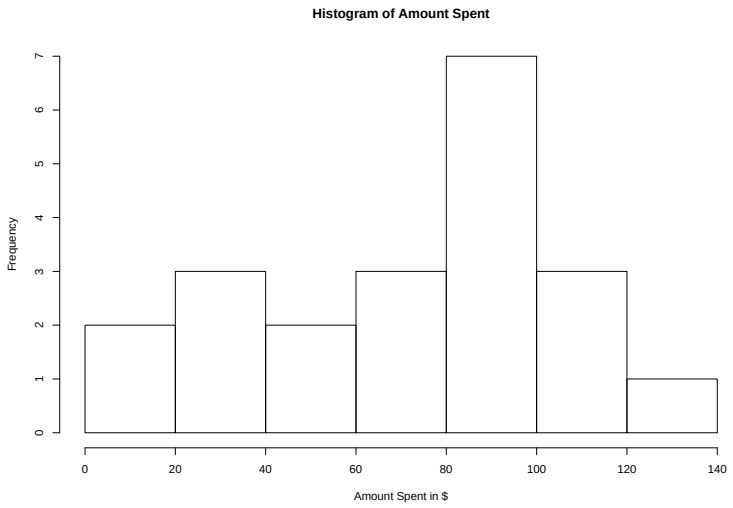




1. (c) One tabular summary that will be useful for describing the amount spent data is as follows:

Range – Amount spent ($)	Frequency	Relative Frequency
$0 \leq \$ < 20$	2	$2/21 = 0.095$
$20 \leq \$ < 40$	3	$3/21 = 0.143$
$40 \leq \$ < 60$	2	$2/21 = 0.095$
$60 \leq \$ < 80$	3	$3/21 = 0.143$
$80 \leq \$ < 100$	7	$7/21 = 0.333$
$100 \leq \$ < 120$	3	$3/21 = 0.143$
$120 \leq \$ < 140$	1	$1/21 = 0.048$

One possible graphical summary is a histogram.



1. (d) The mean amount spent is

$$\overline{x} = \frac{84.12 + 99.67 + \cdots + 112.67}{21} = \frac{1556.79}{21} = \$74.13.$$

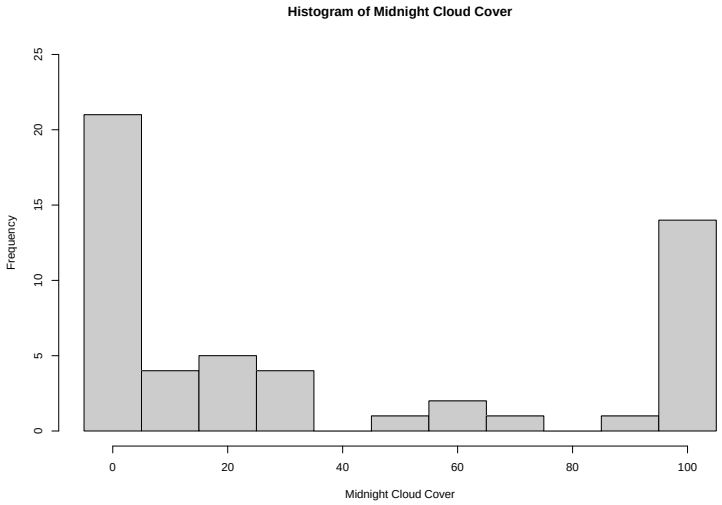
Since there are 21 data points, the median is found to be the 11th point when the data are arranged in increasing order. In this case, we find that the median amount spent is $MD = \$84.12$.

The median is the better measure of centre because the data is skewed left. (This is obvious from the histogram in **(c)**.) Remember that the median is a resistant measure of location and is preferred when data is skewed.

2. (a) The frequency table for the midnight cloud cover data is given by

Midnight Cloud Cover	Frequency	Relative Frequency
0	21	$21/53 = 0.396$
10	4	$4/53 = 0.075$
20	5	$5/53 = 0.094$
30	4	$4/53 = 0.075$
40	0	$0/53 = 0$
50	1	$1/53 = 0.019$
60	2	$2/53 = 0.038$
70	1	$1/53 = 0.019$
80	0	$0/53 = 0$
90	1	$1/53 = 0.019$
100	14	$14/53 = 0.264$

2. (b) The accompanying histogram is given by



2. (c) The mean midnight cloud cover is

$$\overline{x} = \frac{21 \times 0 + 4 \times 10 + \cdots + 14 \times 100}{53} = \frac{1990}{53} = 37.55.$$

Since there are 53 data points, the median is found to be the 27th point when the data are arranged in increasing order. In this case, we find that the median amount spent is $MD = 20$.

No, neither the mean nor the median is a good descriptive measure of this data. This is because the distribution (as seen from the histogram) is bimodal with peaks at both extremes of the data range (namely, at 0 and 100).

3. (a) We find that the net change in value per stock is given by the following table.

Company	Net change in value	Number of Stocks
Distiller I	-\$2.55	200
Distiller II	-\$2.50	200
Brewery A	-\$0.95	100
Brewery B	\$0.25	400
Vineyard X	-\$1.25	300
Vineyard Y	-\$1.00	300

Therefore, the net change in the value of the portfolio between its initial and final values is

$$(-2.55\times200)+(-2.50\times200)+(-0.95\times100)+(0.25\times400)+(-1.25\times300)+(-1.00\times300) = -\$1680.00.$$

3. (b) The mean change per stock is then given by

$$\overline{x} = \frac{-1680.00}{200+200+100+400+300+300} = \frac{-1680.00}{1500} = -\$1.12.$$