

Skewedness

- **Skewed Data:** A distribution with which extreme values (values that are much smaller or much larger than the rest)
 - In a **Nonskewed** distribution, the Mean and Median are usually very similar if not identical.
 - In a **Skewed** distribution, the extreme value(s) create(s) a gap between the mean and the median.
- Skewed data set can be detected visually when in a stem and leaf notation:

a) Positively Skewed Data:

```

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

```

- When there is a "tail" on the bottom end of the stem-and-leaf diagram, the data is positively skewed.

b) Negatively Skewed Data:

```

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

```

- When there is a "tail" at the top end of the stem-and leaf diagram, the data is negatively skewed

c) Nonskewed Data:

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

- When the steam-and-leaf diagram is relatively symmetrical with the "peak" located toward the centre, the data is nonskewed.
- **IMPORTANT → 2 rules that are always true:**
 - 1) When Mean > Median, the distribution is positively skewed
 - 2) When Mean < Median, the distribution is negatively skewed

Example 6

Is the following distribution [29, 33, 46, 49, 62, 71] negatively or positively skewed?

- To solve this, simply calculate the Mean and Median and compare:
 - Mean = 48.33, Median = 47.5
 - Mean > Median → The data is positively skewed.
- It's that simple :)