Measures of Location: Third Moment

John S Butler (TU Dublin)

Introduction

A different aspects of a distribution of data can be summarised by the measures of location:

1. The First Moment: Middle.
2. The Second Moment: Spread.
3. **The Third Moment: Symmetry.**

All that being said, I would always recommend plotting the data first before anything else.

A picture (histogram) is worth a thousand words.

Third Moment: Symmetry

Skewness

Definition 1:

Skewness is a measure of symmetry (or not symmetry) of a distribution. Pearson's Coefficient of Skewness number 1 uses the mode to calculate skewness, given by the formula is:

$$sk_1 = \frac{\bar{x} - Mode(x)}{\sigma}.$$

Definition 2:

Pearson's Coefficient of Skewness number 2 uses the median to calculate skewness, given by the formula is:

$$sk_2 = 3\frac{\bar{x} - Median(x)}{\sigma}$$

, where $\bar{x}$ is the average of the elements, σ is the standard #deviation, $Mode(x)$ is the mode of the elements and $Median(x)$ is the median of the elements.

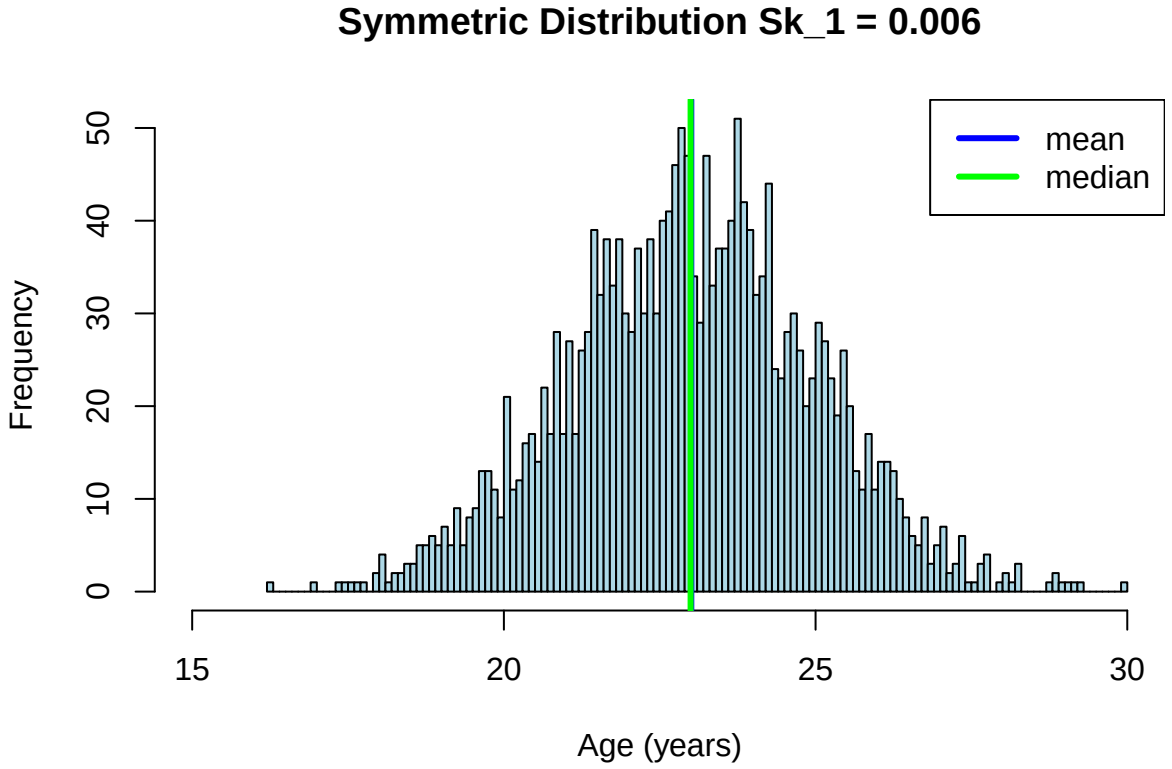
Interpretation of Skewness

How to interpret Skewness:

- A skewness near means the distribution is symmetric.
- A Negative skewness means the distribution is right skewed
- A Positive skewness means the distribution is left skewed

Symmetric Example (Age at a Good Concert)

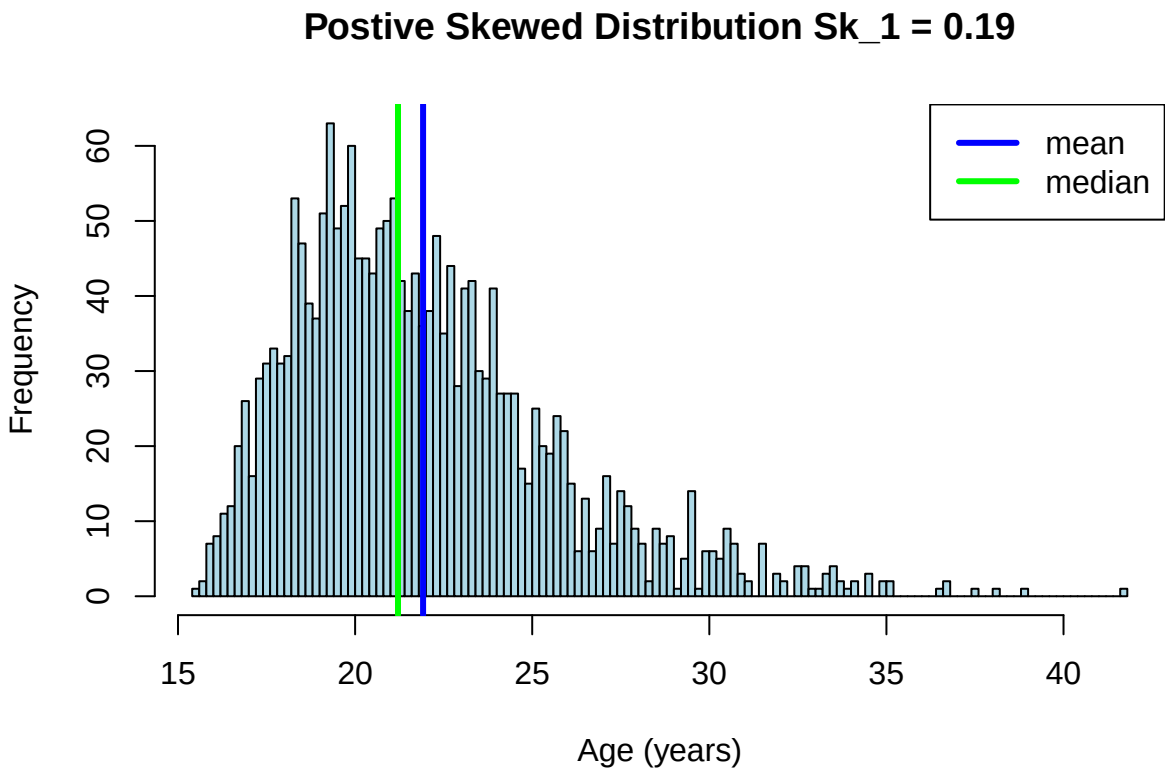
The figure below shows a symmetric histogram of 2000 concert attendees ages observations at a "good" concert:



The age distribution is equally distributed around the mean, hence the skewness is more or less 0.

Positive Skewness Example (Age at a Bieber Concert)

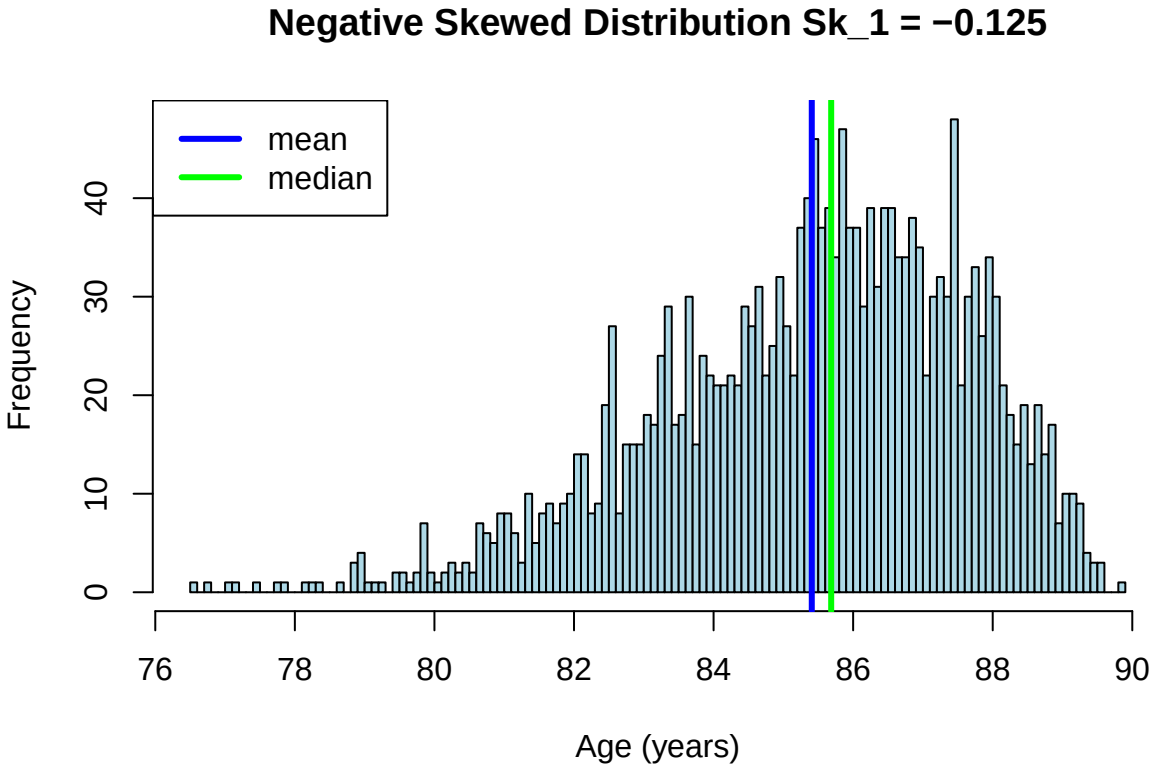
The figure below shows a positive skewness histogram of 2000 attendees age observations at a Justin Bieber concert:



The age distribution shows that the concert attendees are mostly young, hence the skewness is positive.

Negative Skewed Example (Age at an Andre Rieu Concert)

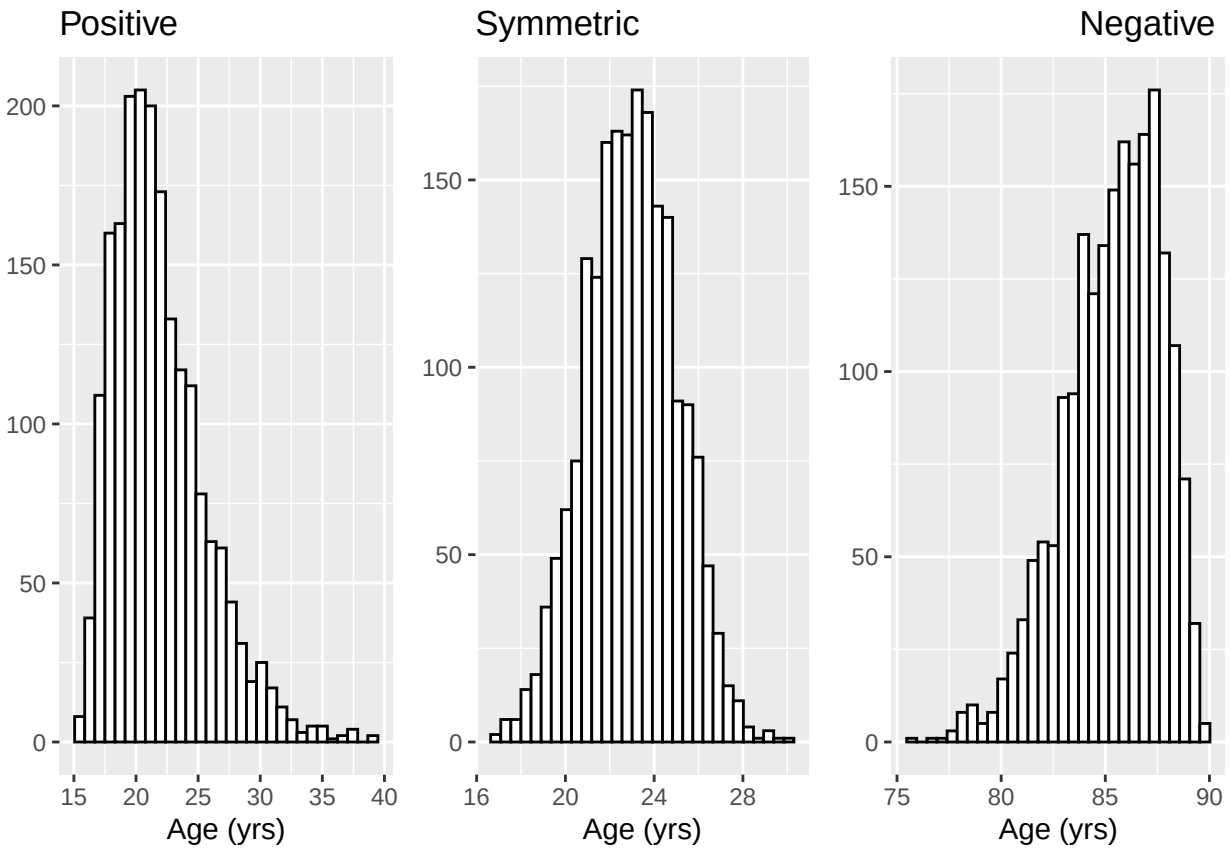
The figure below shows a negative skewness histogram of 2000 attendees age observations at a Andre Rieu concert:



The age distribution is skewed to the right side as the concert has mostly older adults, hence the skewness is negative.

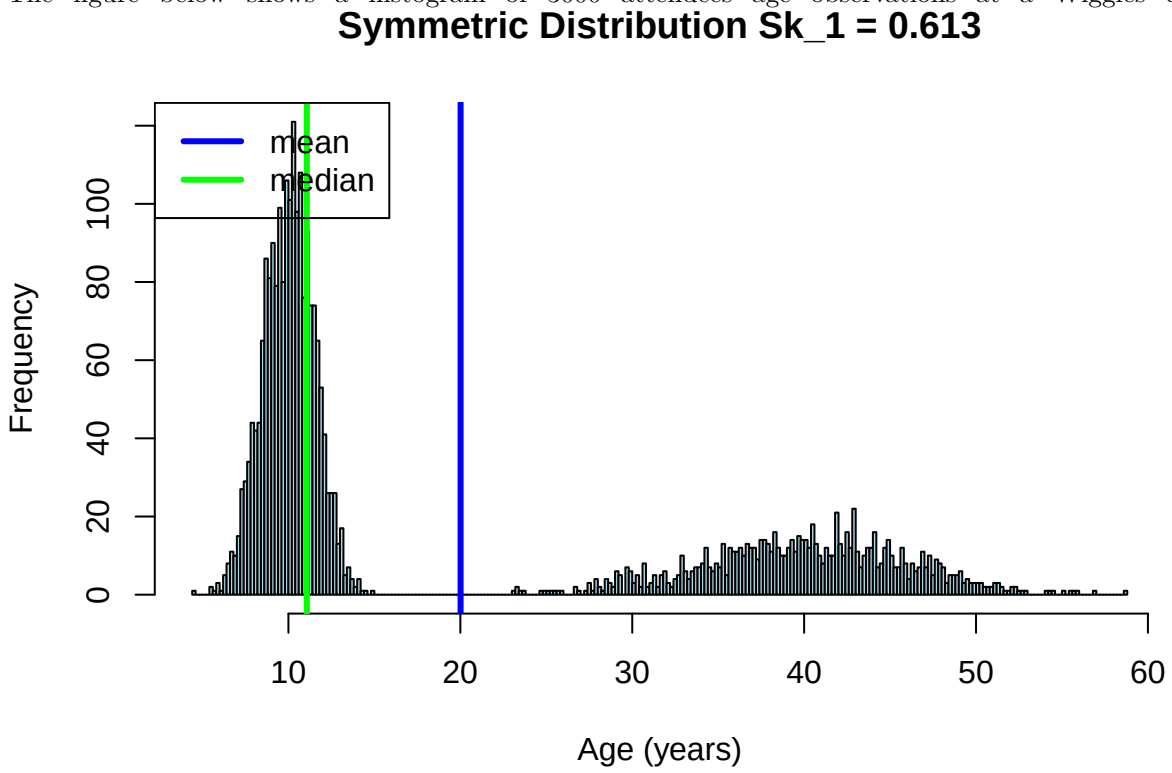
All in one plot

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```


Positive
Symmetric
Negative
200
150
100
50
0
15
20
25
30
35
40
Age (yrs)
150
100
50
0
16
20
24
28
Age (yrs)
150
100
50
0
75
80
85
90
Age (yrs)

Bi-modal distribution

The figure below shows a histogram of 3000 attendees age observations at a Wiggles concert:



The distribution is bimodal as there are 2000 excited children with 1000 parents wear earplugs, the skewness is positive but this is misleading.