

Project 1: Descriptive Statistics

Zack Garza

Math 142: Elementary Statistics
Effective Date of Report: June 22, 2014

CONTENTS

I. PART 1

The data set used for this analysis is “Data Set 8: Alcohol and Tobacco Use in Animated Children’s Movies” from Appendix B in *Elementary Statistics* by Triola. The data analyzed were the movie lengths in seconds.

Figure 1 shows the frequency distribution of these movie lengths.

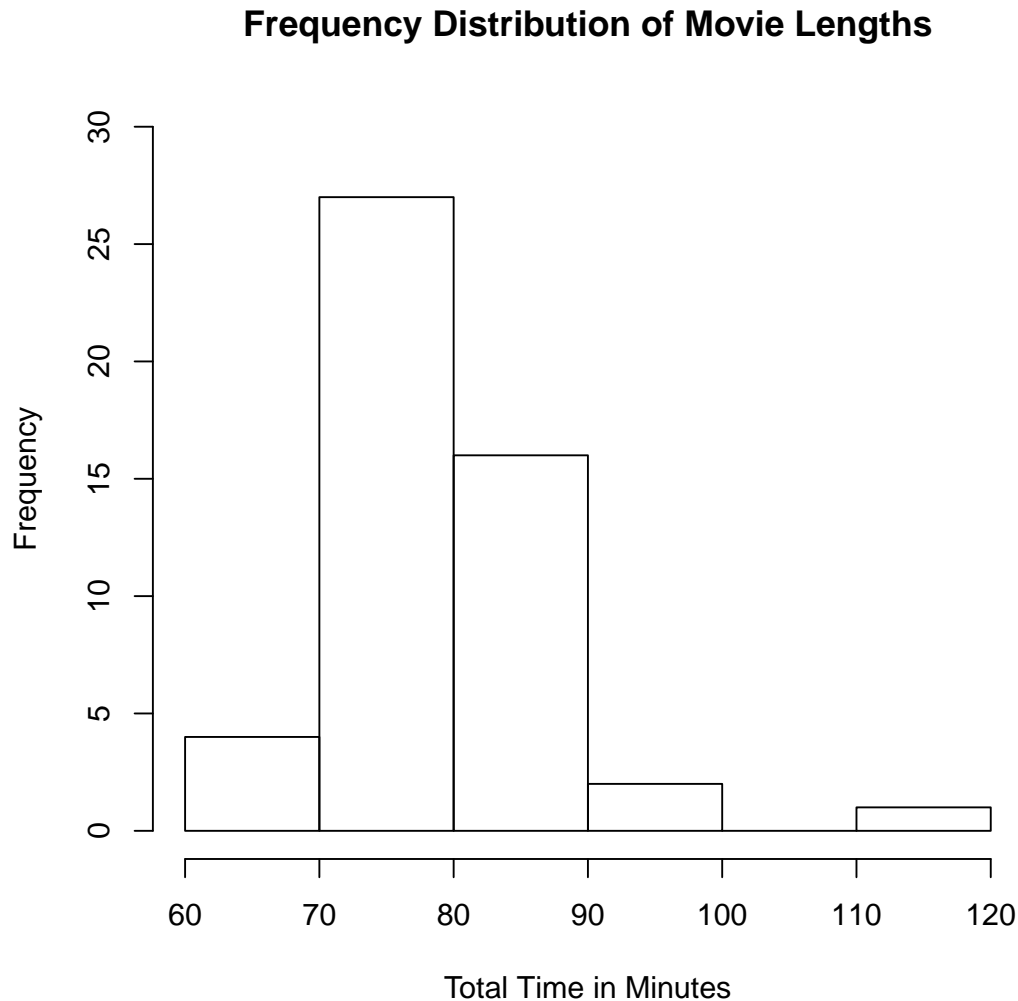


Figure 1. Running Times of a Sample of G-Rated movies.

II. MEASURES OF CENTER

Mean: 79.2

Median: 77.5

Mode: 75

Midrange: 92

Range: 56

Sample Standard Deviation: 8.96

5 Number Summary:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
64.00	74.00	77.50	79.20	82.75	120.00

III. MODIFIED BOX PLOT

In order to make a modified box plot, we define the outliers in the sample set to be those data that are outside of the interquartile range (or IQR). The IQR is equal to $Q_3 - Q_1$, so we first need to evaluate the quantiles. These are given by:

Quantiles:

0%	25%	50%	75%	100%
64.00	74.00	77.50	82.75	120.00

From this, we can identify the outliers as those that are above Q_3 or below Q_1 by $1.5 \times \text{IQR}$ and generate the boxplot accordingly.

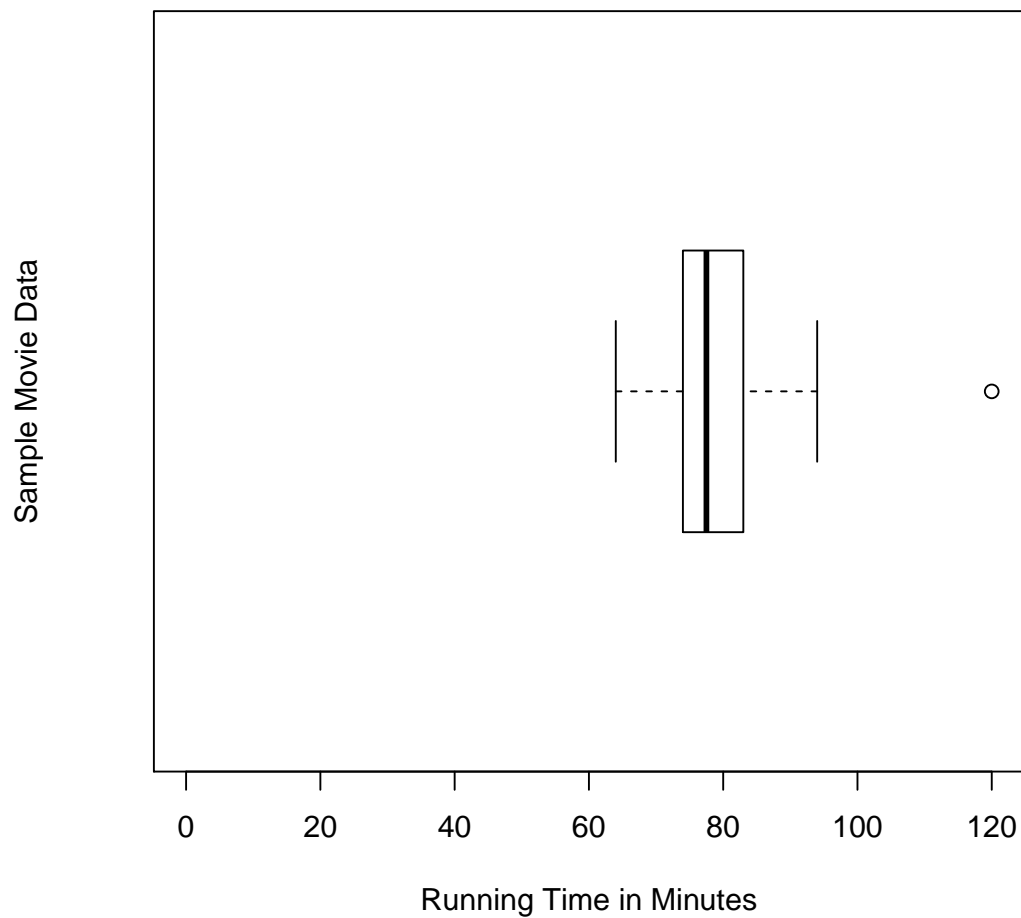


Figure 2. Modified box plot, using the 1.5 IQR method to identify outliers. The "whiskers" outside of the central box indicate the highest and lowest usual values.

IV. USUAL VALUES

The "usual values" all fall within ± 2 standard deviations of the mean – that is, they are in the range of 79.2 ± 8.96 .

Mean: 79.2

Lowest Usual Value: 61.3

Highest Usual Value: 97.1

V. SUMMARY OF FINDINGS