a) $n = 51$

3 3 3 3 3 3 3 3 4 4 4 4 $\boxed{4}$ 5 5 5 5 5 6 6 6 7 7 7 7

Q1

$\boxed{8}$ 8 9 9 9 10 10 10 10 11 11 11 11 $\boxed{12}$ 13 15 15 15

m Q3

17 20 21 21 $\boxed{27 \quad 31 \quad 34 \quad 55}$

Outliers

Check for outliers

$Q1 - 1.5(IQR)$

$4 - 1.5(12-4)$

$4 - 1.5(8)$

$4 - 12$

-8

0

$Q3 + 1.5(IQR)$

$12 + 1.5(12-4)$

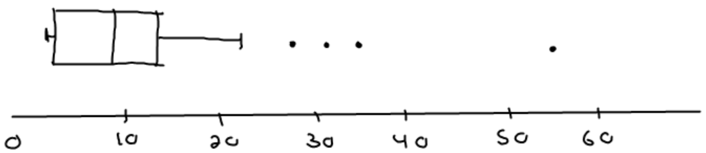
$12 + 1.5(8)$

$12 + 12$

24

3, 4, 8, 12, 21 (27, 31, 34, 55)

a)



b) Median and IQR. Right skewed distribution with outliers. Mean is not resistent to skewness and outlier. Mean will be inflated.

1/3

Standard deviation

$$\text{Variance} = S^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}$$

$$\text{Standard deviation} = \sqrt{\text{variance}} = \sqrt{S^2} = S$$

10 10 10

$s = 0$

5 10 15

$(5-10) + (10-10) + (15-10)$

$-5 + 0 + 5$

0

square so the + and − don't cancel out.

variance is square units

standard deviation is in units.

Properties of S

- s measures spread around mean only (not median or mode)
- $s = 0$ only occurs when there is no spread
  (all the data must be the same)
- s is not resistant.
  why? uses mean in its calculation

Ex marks $\bar{x} = 26.1$

8 13 14 16 23 26 28 33 39 61

$$\text{variance} = s^2 = \frac{(8-26.1)^2 + (13-26.1)^2 + (14-26.1)^2 + \ldots + (61-26.1)^2}{10-1}$$

$$\text{standard deviation} = \sqrt{\text{variance}}$$

$$s^2 = \frac{2263.6}{10-1}$$

$$s^2 = \frac{2263.6}{9}$$

$$\sqrt{s^2} = \sqrt{251.511}$$

$$\boxed{s = 15.8591}$$

Choosing a Summary

Median
Quartiles
↓
- better for skewed distributions

- resistant

Mean
Standard Deviation

- better for symmetrical distributions with no outliers

HW #5
1/1 - 1/3 In Review Packet