

StatCrunch Lab 3 for Statistics 301

Topics: summary statistics, boxplots, grouped data

Summary Statistics. Consider small *sample* of $n = 10$ temperatures:

0, 1, 1, 2, 2, 2, 3, 3, 5, 7.

Relabel var1 as temperature. Type temperatures into temperature column. Choose Stat, Summary Stats, Columns, select temperature, Next, select Sum, IQR, Unadj. Std. Dev., Unadj Variance, then Calculate.

- sample size: $n = 10$
- (population and sample) mean: $\bar{x} = 2.6$
- (sample) variance: $s^2 \approx 4.267$
- (sample) standard deviation: $s \approx 2.066$
- standard error: $\sigma_{\bar{x}} \approx 0.653$
- median: Median = 2
- range: Range = 7
- minimum: Min = 0
- maximum: Max = 7
- lower quartile: Q1 = 1
- upper quartile: Q3 = 3
- sum: $\sum x = 26$
- interquartile range: IQR = 2
- (population) variance (Unadj. Variance): $\sigma^2 = 3.84$
- (population) standard deviation (Unadj Std. Dev.): $\sigma \approx 1.96$

We will be interested in most of the statistics given here.

Boxplot. To create a boxplot from temperature data, Choose Graphics, Boxplot, select temperature, Next, check Use fences to identify outliers, check Draw boxes horizontally, Create Graph! Notice left whisker stops at largest value below upper fence $3 + 1.5(2) = 6$, that outlier 7 appears as a •.

Grouped Data. Consider discrete distribution table for sample of number of tablets used in high blood pressure experiment.

number of tablets	number of patients
1	5
2	10
3	4
4	1
total	20

1. *Grouped Average.* Click “Blank Data Table”. Relabel var1 as tablets, var2 as number. Type data into tablets and number columns. Choose Data, Computer expression, Expression: $\text{sum}(\text{tablets} * \text{number}) / 20$, New column name: grouped mean, Compute to find $\bar{x} = 2.05$.
2. *Grouped Standard Deviation.* Choose Data, Computer expression, Expression: $\text{sqrt}(\text{sum}(\text{number} * (\text{tablets} - 2.05)^2) / 19)$, New column name: grouped SD, Compute to find $s \approx 0.83$.