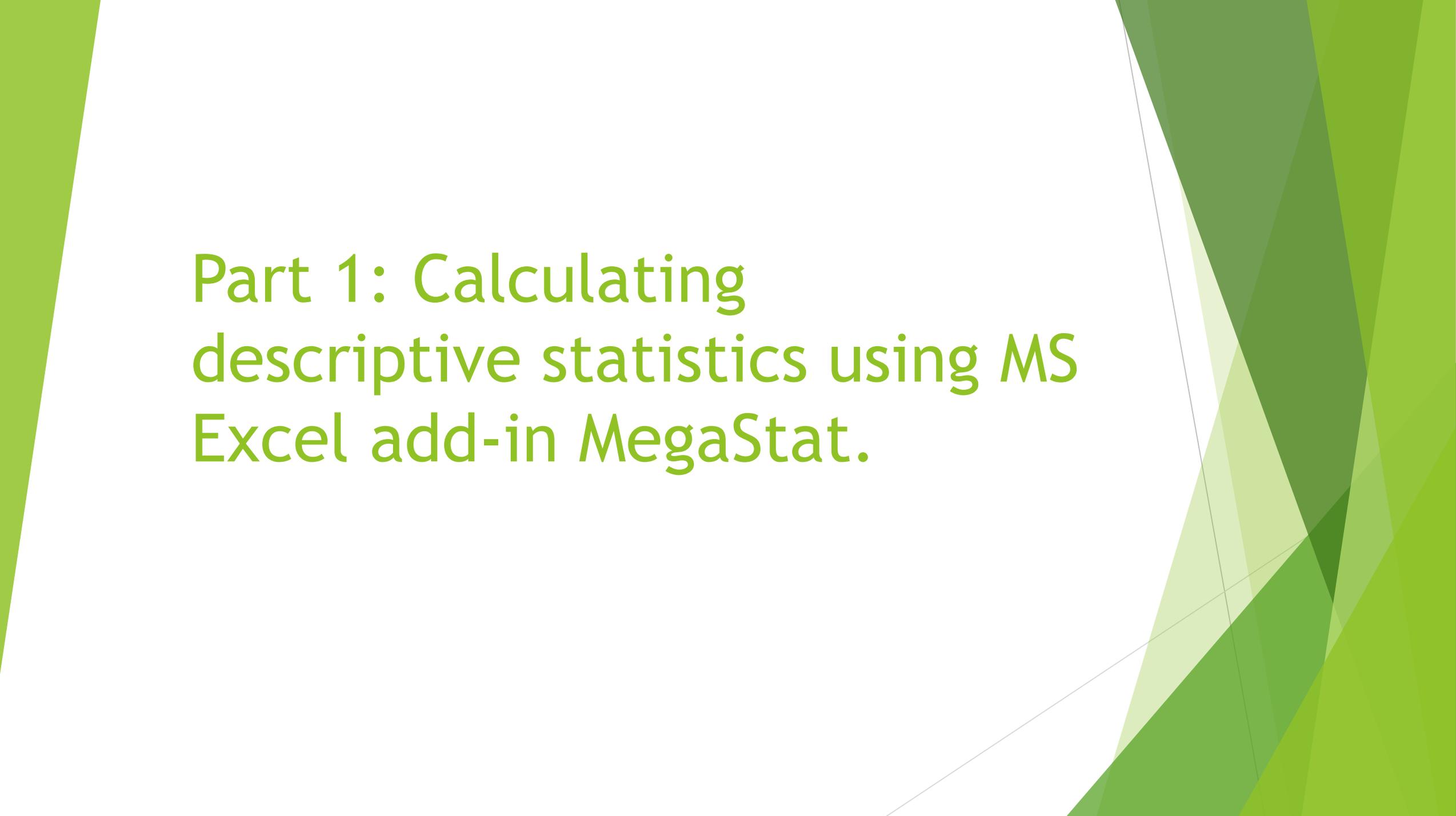


Chapter 3: Data Description

© STAT 110 Team 2018

The background features abstract, overlapping green geometric shapes in various shades, including light lime green, medium green, and dark forest green, creating a modern, layered effect.

Part 1: Calculating descriptive statistics using MS Excel add-in MegaStat.

3. Run MegaStat

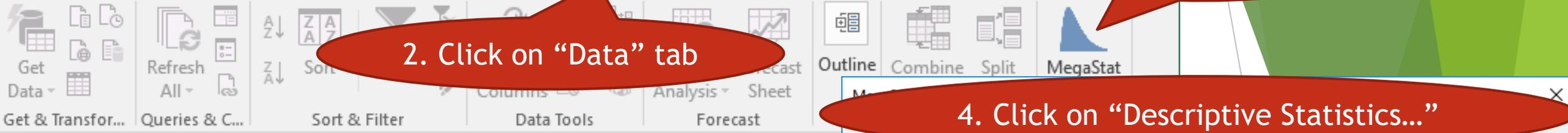
2. Click on "Data" tab

4. Click on "Descriptive Statistics..."

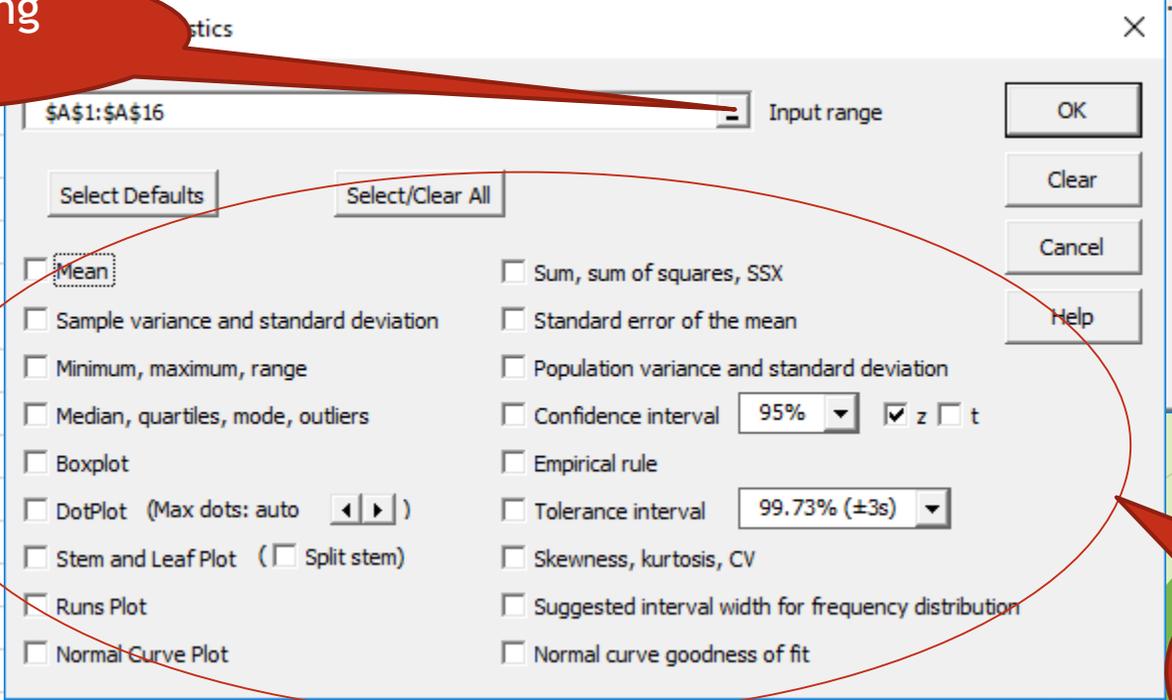
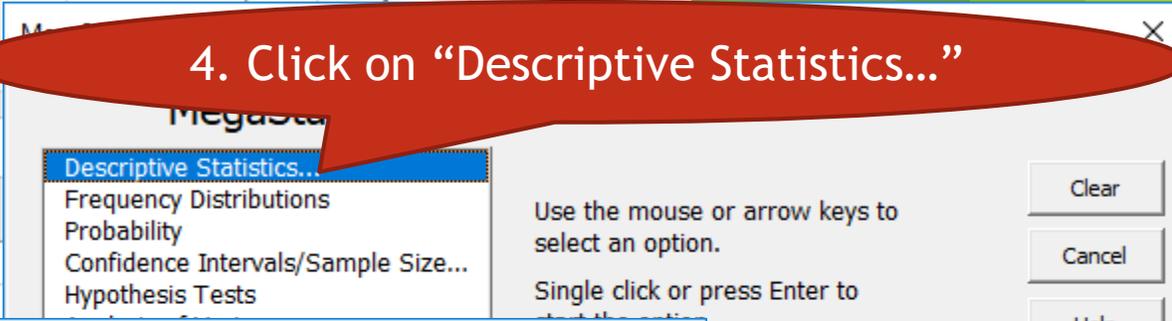
5. Locate data using "Input range"

1. Input data

6. Choose the required descriptive statistics



	A	B	C	D	E	F	G	H	I
1	parking								
2	750								
3	3400								
4	1962								
5	700								
6	203								
7	900								
8	8662								
9	260								
10	1479								
11	5905								
12	9239								
13	690								
14	9822								
15	1131								
16	2516								
17									
18									



3 - 1 Measures of Central Tendency

Choose to calculate the mean

Choose to calculate both the median and the mode

Descriptive statistics

\$A\$1:\$A\$16

Select Defaults Select/Clear All

<input checked="" type="checkbox"/> Mean	<input type="checkbox"/> Sum, sum of squares
<input type="checkbox"/> Sample variance and standard deviation	<input type="checkbox"/> Standard error of mean
<input type="checkbox"/> Minimum, maximum, range	<input type="checkbox"/> Population variance
<input checked="" type="checkbox"/> Median, quartiles, mode, outliers	<input type="checkbox"/> Confidence interval (95.00% 2 L
<input type="checkbox"/> Boxplot	<input type="checkbox"/> Empirical rule
<input type="checkbox"/> DotPlot (Max dots: auto ◀ ▶)	<input type="checkbox"/> Tolerance interval (99.73% (±3s) ▼
<input type="checkbox"/> Stem and Leaf Plot (<input type="checkbox"/> Split stem)	<input type="checkbox"/> Skewness, kurtosis, CV
<input type="checkbox"/> Runs Plot	<input type="checkbox"/> Suggested interval width for frequency distribution
<input type="checkbox"/> Normal Curve Plot	<input type="checkbox"/> Normal curve goodness of fit

IMPORTANT NOTES:

1. MegaStat calculates only one mode even if there are two modes or more.
2. If there is no mode, MegaStat will return the value #N/A

3 - 2 Measures of Variation

Choose to calculate the sample variance and standard deviation

Choose to compute the range

Choose to compute the coefficient of variation (CVar)

Descriptive statistics

\$A\$1:\$A\$16

Select Defaults Select/Clear All

Mean

Sample variance and standard deviation

Minimum, maximum, range

Median, quartiles, mode, outliers

Boxplot

DotPlot (Max dots: auto) ◀ ▶

Stem and Leaf Plot (Split stem)

Runs Plot

Normality test

Sum, sum of squares, SSX

Standard error of the mean

Population variance and standard deviation

Confidence interval 95% z t

Empirical rule

Tolerance interval 99.73% (±3s)

Skewness, kurtosis, CV

Suggested interval width for frequency distribution

Normal curve goodness of fit

Clear

Cancel

Help

IMPORTANT NOTE: Do not check this box unless you want to calculate the **POPULATION** variance and standard deviation.

3 - 3 Measures of Position

Descriptive statistics

Sample range: \$A\$1:\$A\$16

Defaults Select/Clear All

Sample variance and standard deviation

Minimum, maximum, range

Median, quartiles, mode, outliers

Boxplot

DotPlot (Max dots: auto)

Stem and Leaf Plot (Split stem)

Runs Plot

Normal Curve Plot

Sum, sum of squares, SSX

Standard error of the mean

Population variance and standard deviation

Confidence interval 95% z t

Empirical rule

Tolerance interval 99.73% ($\pm 3s$)

Skewness, kurtosis, CV

Suggested interval width for frequency distribution

Normal curve goodness of fit

Clear Cancel Help

Choose to calculate Q1 and Q3, and to detect outliers

IMPORTANT NOTE: MegaStat calculates Q1 and Q3 differently than those calculated according to the textbook.

Choose to get the boxplot

Part 2: Other descriptive
statistics calculated using *MS*
Excel only.

3 - 1 Measures of Central Tendency (Weighted Mean)

- ▶ From Example 3-14, we have:

Course	Credits (w)	Grade (X)
English	3	4
Psych.	3	2
Bio.	4	3
Physical	2	1

- ▶ To calculate the weighted mean, we should use the approach mentioned in the textbook. If the weights are integers, then another way to calculate the weighted mean is as follows.
- ▶ First, repeat each X according to the corresponding w, i.e. input 4, 4, 4, 2, 2, 2, 3, 3, 3, 3, 1, 1 in any column of Excel and then use the code “**=AVERAGE(<INPUT RANGE>)**” in any empty cell.

3 - 3 Measures of Position (Standard Score)

- ▶ To calculate the standard score, use:

=STANDARDIZE(<X>,<MEAN>,<STANDARD DEVIATION>)

- ▶ In Example 3-27, we have $X = 65$, Mean = 50, and standard deviation = 10. In any Excel cell, write “**=STANDARDIZE(65, 50, 10)**” to get 1.5.

3 - 3 Measures of Position (Percentile)

- ▶ To calculate percentile (P) given any value (X), use:

$$=(\text{COUNTIF}(\langle \text{DATA} \rangle, "< X") + 0.5) / n * 100$$

3 - 3 Measures of Position (Percentile)

- ▶ To calculate a value (X) given a percentile (P), do the following steps. Ascendingly order the data, by selecting the data, then going to the **DATA** tab, then clicking “Sort Smallest to Largest” on as shown below.

- ▶ Calculate the constant c as shown in book.

$$c = \frac{n \cdot p}{100}$$

- ▶ Count over as shown in the book.

