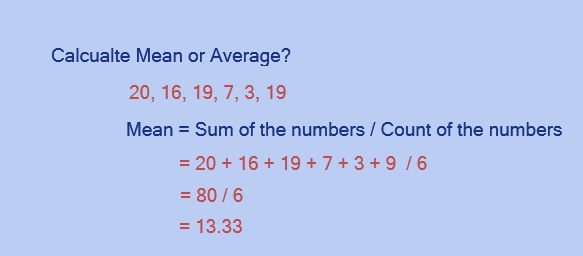
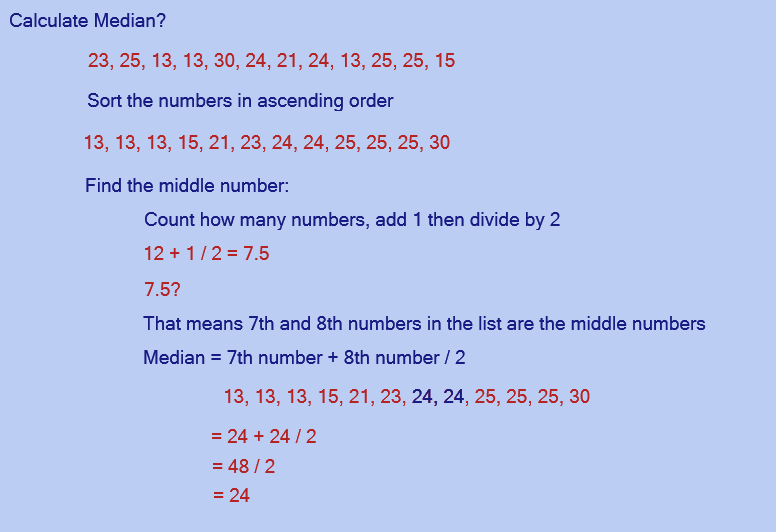
**Basic Statistics:**

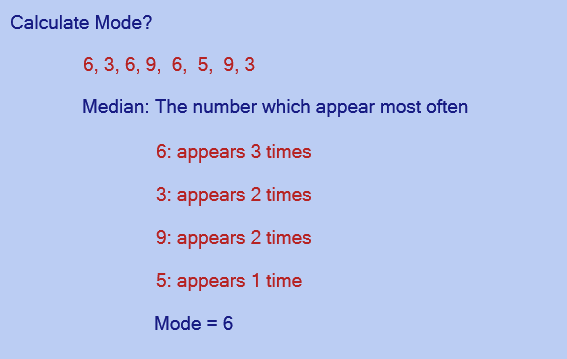
**Measure of Central Tendency:** Mean, Median, Mode

Mean or Average: 

Median: The Median is the ***"middle"*** of a sorted list of numbers.

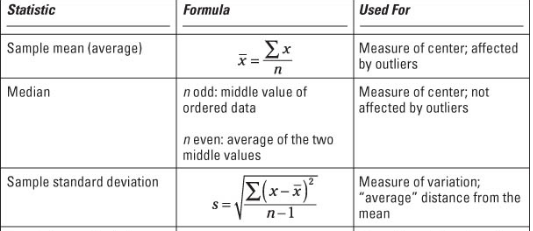


Mode: The number which appears most often in a set of numbers.



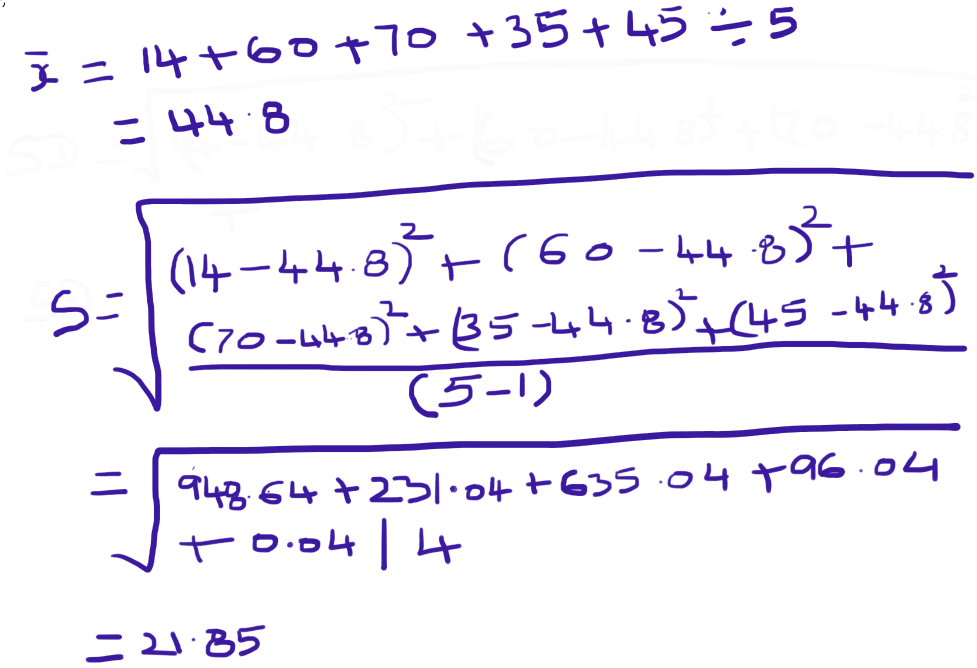
**Equations:**

Mean, Median and Standard Deviation



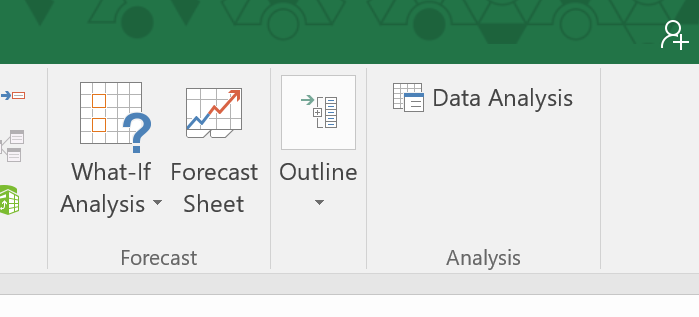
Standard Deviation:

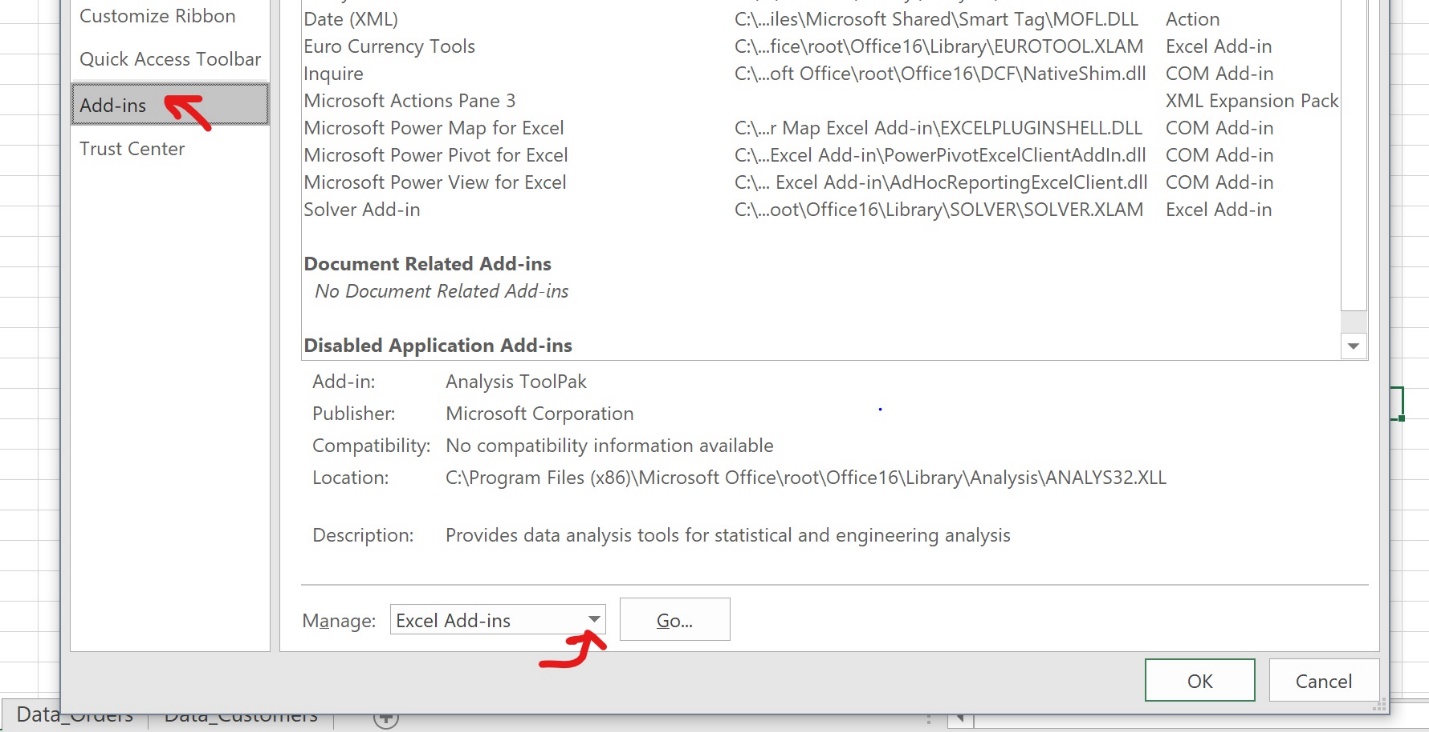
Example: 14, 60, 70, 35, 45



**Summary Statistics using Excel:**

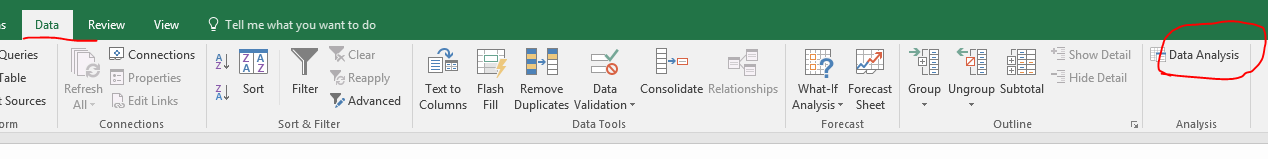
If the ‘Analysis’ box isn’t present in the ‘Data’ section, you will need to add it to Excel from the add-ins option

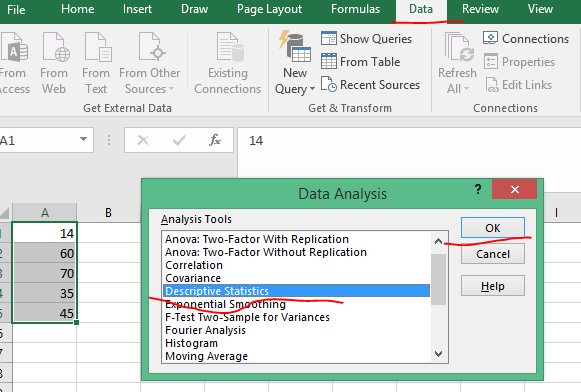


1. Click on File
2. Click on ‘options’
3. Click on the ‘add-ins’ section in the left toolbar
4. Click on the down-arrow under the ‘manage’ section
5. Select ‘Excel Add-ins’
6. Click ‘go’
7. Select the analysis ToolPak and click ‘OK’

Instead of calculating manually like above, we can do this in Excel. Here is a screen shot.

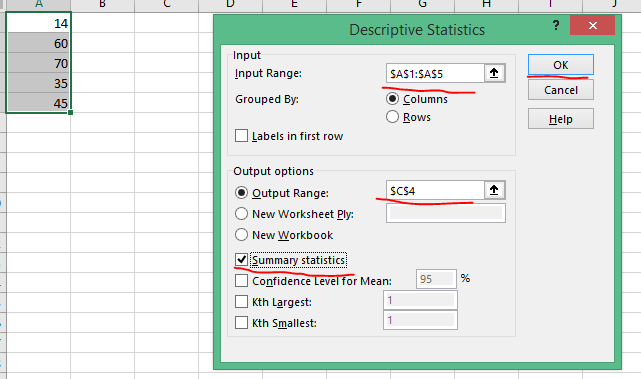
1. Open Excel
2. Select a single cell
3. Enter ‘=RANDBETWEEN(0,45)’ (you can use the autocomplete)
4. Drag the dot on the bottom right of the outlined cell down to a total of 5 cells
5. Click on ‘Data’ in the menu
6. Click on ‘Data Analysis’ in the menu (This option is all the way to the right)



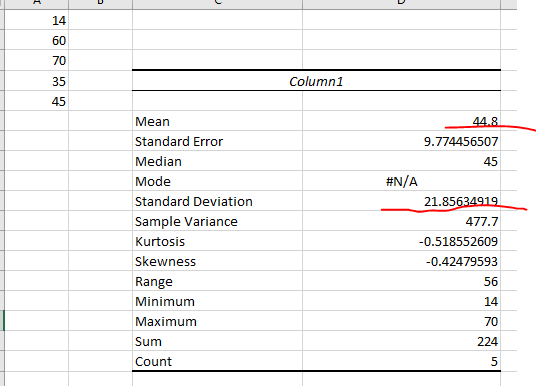


1. Click Ok
2. Select
   1. Input Range
   2. Output Range

Click Ok



7.



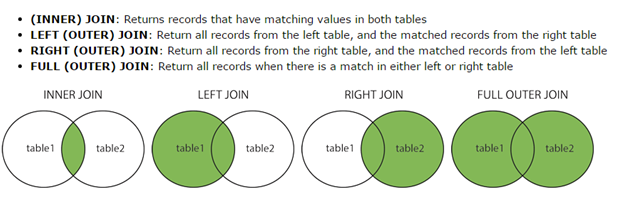
The Mean and Standard Deviation, we calculated early are matching.

**IMPORTANT**

Please try to use the Excel to display ‘Summary Statistics’ few times using different data sets, because we need this in the later lessons.

**Database Joins:**

* Inner Join
* Left Outer Join
* Right Outer Join
* Full Outer Join



To illustrate various joins, we created three tables - Customers, Employees and Orders. Please review the following and understand the concepts in joining the tables.

**IMPORTANT:**

You need this knowledge in Tableau to join the data from various sources

**Customers:**

**Employees:**

**Orders:**

**Inner Join:**

Display the data for Customers who have orders: CustomerID, Name, OrderID, Total

**Left Outer Join:**

Display the Orders whether any employees associated with the orders or not

(all orders)

**Right Outer Join:**

Display all the employees whether they are associated with any orders or not

**Full Outer Join:**

Display all the Orders and all the Employees – even when employees are not associated with any orders and any employee who has not taken any orders

**Discrete and Continuous Variables:**

A **discrete variable** is a **variable** whose value is obtained by counting.

A **continuous variable** is a **variable** whose value is obtained by measuring.

A **continuous variable** is one which can take on a value between any other two values

A **discrete variable** corresponds to a digital quantity, while a **continuous variable** corresponds to an analog quantity.

1. Discreet Variables

Examples:

* Gender
* Employment Status
* Number of Heads tossed
* Number of Cars you own
* Age in Years
* Countries in the World
* States in US
* Date
* Number of students present
* Number of red marbles in a jar

1. Continuous Variables

Examples

* Height of Students in a class
* Weight of Students in a class
* Annual salary (reported in Dollars and Cents)
* Drive time from your home to work
* Miles hiked in a month
* Date
* Temperature
* Time spent waiting
* Water consumed
* Color wavelength
* Direction of travel

NOTE: Date can be treated as Discreet or Continuous. We will see examples, when we work on Tableau

**Very Important:**

**You need to understand the concept of Discreet and Continuous variables. This concept is being used on every worksheet we develop in Tableau.**

Further reading:

https://www.youtube.com/watch?v=6IdJ1aPFDCs

<https://www.khanacademy.org/math/statistics-probability/random-variables-stats-library/discrete-and-continuous-random-variables/v/discrete-and-continuous-random-variables>