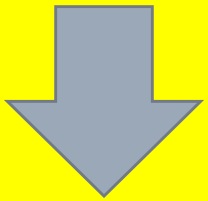
Excel & Data Analytics

Remember Data Wrangling?

The process of cleaning and unifying messy and complex data sets for easy access and analysis.



Organizing and processing data

Data Wrangling

Step #1 – Discovery

Step #2 – Structuring

Step #3 – Cleaning

Step #4 – Enriching

Step #5 – Validating

Step #6 – Publishing

What is Excel?

A program by Microsoft that is used for recording, analyzing and visualizing data in the form of a spreadsheet.

Why Excel?

- Performs various math functions on large data sets
- You can search, sort, filter; makes it easier to clean
- Beautify data and present with charts & tables
- Reporting, accounting & analysis is easier
- Provides security through locking cells, and passwords

Excel is HUGE

If you don't have Excel installed, and you don't have a Microsoft account, here is a site to signup for …

Free Microsoft Office Online | Word, Excel, PowerPoint

https://www.microsoft.com/en-za/microsoft-365/free-office-online-for-the-web

Lets play with Excel!

We will use the following spreadsheet together to learn more about Excel.

Excel_Practice_Student.xlsx

Note: This spreadsheet has 3 tabs at the bottom of the screen and is wide so scroll to the right to see all the content

Functions

AND

=AND(**[logical1], [logical2], ...**)

If a number is greater than or smaller than another number or is equal to something.

Produces “True” or “False”

AND USING IF

=IF(AND(**logical_test, logical_test**), **[“value_if_true”], [“value_if_false”]**)

Produces whatever outcome you need it to read if you don’t want “True” or “False” …

Like “Black” and “White”

AND			
Name	Type 1	Speed	Fire Type AND more than 70 Speed
Bulbasaur	Grass	45	FALSE
Ivysaur	Grass	60	
Venusaur	Grass	80	
Charmander	Fire	65	
Charmeleon	Fire	80	
Charizard	Fire	100	
Squirtle	Water	43	
Wartortle	Water	58	
Blastoise	Water	78	

Reading a function in Excel

=AND(B3="Fire",C3>70)

=AVERAGEIFS(C27:C38, B27:B38, F28, D27:D38,G28)

AVERAGEIF

Combine averages from different cells

=AVERAGEIF(**range, criteria, [average_range]**)

AVERAGEIFS

Calculates the average of a range based on one or more true/false conditions

=AVERAGEIFS(**average_range, criteria_range1, criteria1, ...**)

AVERAGEIF

Name	Type 1	Speed
Bulbasaur	Grass	45
Ivysaur	Grass	60
Venusaur	Grass	80
Charmander	Fire	65
Charmeleon	Fire	80
Charizard	Fire	100
Squirtle	Water	43
Wartortle	Water	58
Blastoise	Water	78

Type	Average Speed
Grass	61.66666667
Fire	
Water	

Reading a function in Excel

=AVERAGEIF(**B15:B23**,**D17**,**C15:C23**)

Functions

COUNT

Calculates number of cells used within a range that have numbers

=COUNT(value1:value2)

COUNTA

Calculates number of cells used within a range that have either numbers or letters

=COUNTA(value1:value2)

Functions

COUNTBLANK

Calculates number of cells used within a range that are blank

=COUNTBLANK(value1:value2)

COUNTIF

Calculates number of cells as specified

Have to use a $ for absolute values to be counted

=COUNTA($value$1:$value$2,criteria)

Functions

COUNTIFS

Counts cells in a range based on one or more true or false conditions

=COUNTIFS(**criteria_range1, criteria1, [criteria_range2, criteria2], ...**)

Functions

IF (EQUAL TO)

Returns values based on a true or false condition

`Click on D105`

=IF(B105="Grass", "Yes", "No")

Functions

IF (GREATER THAN)

Counts cells in a range based on one or more true or false conditions

I105

```
=IF(H105>500, "Yes", "No")
```

IFS

Returns values based on one or more true/false conditions

D117

=IFS(C117>90,"Fast", C117>50, "Normal", C117<=50,"Slow")

Functions

MEDIAN

Returns the middle value in the data

=MEDIAN(F117:K117)

MODE

Used to find the number seen most times.

=MODE.SNGL(B129:E134)

Functions

OR

Returns true/false based on two or more conditions.

```
=OR(B140="Water",C140>60)
```

OR WITH IF

Let's you check multiple conditions for the if function

```
=IF(OR(H140="water",C140>60),"Yes","No")
```

Functions

STDEV.P

Calculates the Standard Deviation for the entire **population**

=STDEV.P(D152:D173)

Measures how far a ‘typical’ observations is from the average of the data.

*Ignores cells with text and logic

STDEV.S

Calculates the Standard Deviation for a **sample**

=STDEV.S(L152:L171)

Functions

SUMIF

Calculates the sum of values in a range based on true/false conditions.

```
=SUMIF(C177:C185,F178,D177:D185)
```

SUMIFS

Calculate the sum of a range based on one or more true/false condition

```
=SUMIFS($D$189:$D$201,$C$189:$C$201,G189,$E$189:$E$201,H189)
```

Functions

VLOOKUP

Allows searches across columns

```
=VLOOKUP(G207,A205:E226,2,1)
```

lookup_value (required) in 1st column,

table_array (required) lookup range,

col_index_num (required) of lookup_value,

Approximate - 1/TRUE or Exact - 0/FALSE

XOR

Returns true/false based on two or more conditions

```
=XOR(B231="fire",C231<60)
```

Excel Part 2 – Charts & Pivot Tables

Windows Keyboard Shortcuts

CTRL + Z	undo
CTRL + W	close
CTRL + A	Select all
ALT + TAB	Switch apps
ALT + F4	Close apps
WIN + D	Show/Hide Desktop
CTRL + X	Cut
CTRL + C	Copy
CTRL + V	Paste

WIN + L/R Arrow	compare windows
WIN + up/down arrow	
WIN + double up/down	
ESC	
WIN + PrtScn	save screenshot
Shift + arrows	highlight text
CTRL + B/I/U	customize font

macOS Keyboard Shortcuts

Command + C = Copy

Command + X = cut

Command + V = paste

Control + Command + F = Fullscreen

Command + Mission Control = desktop

Inserting Charts

Visually Compare information inside of your data

Go to the Loans Tab on Excel Spreadsheet

Once your data is filled out…

- highlight the information use
- Highlight B1:C5
- Click Insert in the ribbon
- Bring up chart options
- Pick the ones you want

Pivot Tables

A summary of a large dataset that usually includes the total figures, average, minimum, maximum, etc.

Open Up Sales Spreadsheet

Once you bring up your spreadsheet

- Click on pivot table
- dataset should already be selected
- Select “new worksheet”
- Choose the values that you want
- Analyze your data

Sales Exercise

- Select Insert
- Select PivotChart
- Table/Range should be picked already
- Select: CompanyName, ProductName, UnitPrice, Quantity and SubTotal
- Select the Row Labels drop down, remove the “select all” tick, select Ana Trujiullo, press ok

Sales Exercise continued

- Go back to Insert and choose pivot table
- Make sure table is selected
- Choose FirstName
- Choose CompanyName
- Choose SubTotal and drag to values
- Select recommended Charts