Excel functions for Data Analytics

Text functions

Joining text:

CONCAT - Combines the text from multiple ranges and/or strings, but it doesn't provide the delimiter or IgnoreEmpty arguments.

```
=CONCAT(B2," ", C2)
```

CONCATENATE * deprecated, replaced by CONCAT

```
=CONCATENATE(B2, " ",C2)
```

TEXTJOIN - Combines the text from multiple ranges and/or strings, and includes a delimiter you specify between each text value that will be combined.

```
TEXTJOIN(delimiter, ignore_empty, text1, [text2], …)
```

Splitting text:

LEFT - Returns the leftmost characters from a text value

```
LEFT(text, [num_chars])
```

RIGHT - Returns the rightmost characters from a text value

```
RIGHT(text,[num_chars])
```

MID - Returns a specific number of characters from a text string starting at the position you specify

```
MID(text, start_num, num_chars)
```

Combining text:

- CLEAN - Removes all nonprintable characters from the text
- TRIM - Removes spaces from text
- UPPER - Converts text to uppercase
- LOWER - Converts text to lowercase
- PROPER – capitalize 1st letter

```
TRIM(CLEAN(B2))
```

Replacing and removing

- SUBSTITUTE(where, what, with what) - Substitutes new text for old text in a text string

```
SUBSTITUTE(text, old_text, new_text, [instance_num])
```

```
=SUBSTITUTE(A2, "Sales", "Cost")
```

Numbers and Date Functions

Converting data:

- VALUE - Converts a text argument to a number
- TEXT - Formats a number and converts it to text

* *Applying the VALUE() function to a non-numeric cell will result in an error. Instead:*

1. Copy column as a value
2. Find . and replace with **,**
3. Format as a currency

Date functions:

- TODAY – returns date
- NOW – returns date and time
- MONTH
- DAY
- YEAR
- DATE – for entering a valid date

Performing calculations with dates:

- DAYS – Returns the number of days between dates
- WORKDAY – returns serial number of dates before or after a specified number of workdays
- WORKDAY.INTL - function takes a date and returns the nearest working n days in the future or past, based on an offset value you provide. Unlike the WORKDAY function, WORKDAY.INTL allows you to customize which days are considered weekends (non-working days). You can use WORKDAY.INTL function to calculate completion dates, ship dates, etc. that take into account non-working days.
- NETWORKDAYS – returns workdays between two dates
- NETWORKDAYS.INTL – same with custom weekend parameters

Calculating dates from a given date

- EOMONTH – returns the last day of the month

=EOMONTH (start_date, months)

=EOMONTH(A2,1) - Date of the last day of the month, one month after the date in A2

=EOMONTH(A2,-3) - Date of the last day of the month, three months before the date in A2

- EDATE – function returns date on the same day of the month, n months in the past or future. You can use EDATE to calculate expiration dates, maturity dates, and other due dates. Use a positive value for months to get a date in the future, and a negative value for dates in the past.

=EDATE (start_date, months)

=EDATE(A2,-1) - The date, one month before the date above

***Calculate the number of days overdue:**

=DAY(TODAY()-A3)

=TODAY()-A3

***Calculate last day of a month**

=DATE(2020,MONTH(1&A2)+1,1)-1

Minus one day

Next month

1st day of the next month

Cell Referencing

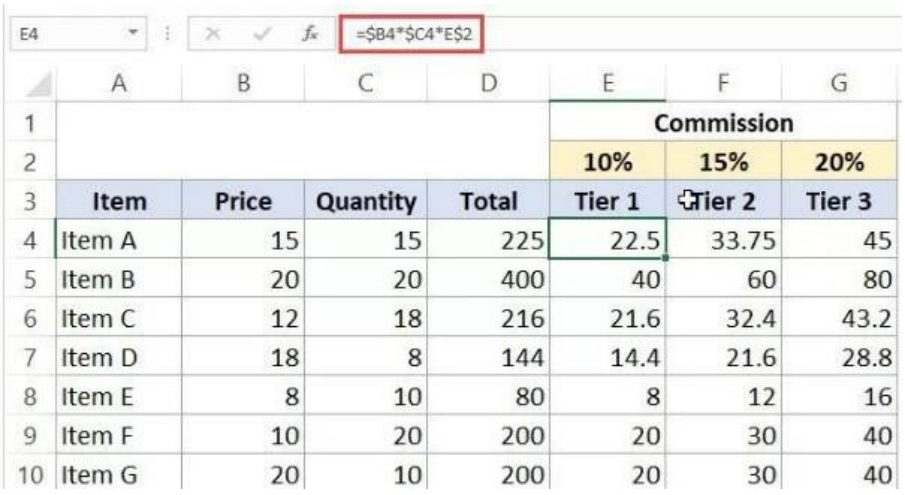
Relative: C1=A1*B1 – then C2=A2*B2

Absolute cell references are useful when you don't want the cell reference to change as you copy formulas. A dollar symbol, when added in front of the row and column number, makes it absolute (i.e., stops the row and column number from changing when copied to other cells).

E2=D2*G1 - then E3=D3*G1

Mixed cell references are a bit more tricky than the absolute and relative cell references.
There can be **two types of mixed cell references**:

- The row is locked while the column changes when the formula is copied.
- The column is locked while the row changes when the formula is copied.



	A	B	C	D	E	F	G
1					Commission		
2					10%	15%	20%
3	Item	Price	Quantity	Total	Tier 1	Tier 2	Tier 3
4	Item A	15	15	225	22.5	33.75	45
5	Item B	20	20	400	40	60	80
6	Item C	12	18	216	21.6	32.4	43.2
7	Item D	18	8	144	14.4	21.6	28.8
8	Item E	8	10	80	8	12	16
9	Item F	10	20	200	20	30	40
10	Item G	20	10	200	20	30	40

=$B4*$C4*E$2

$B4 (and $C4) – In this reference, the dollar sign is right before the Column notation but not before the Row number. This means that when you copy the formula to the cells on the right, the reference will remain the same as the column is fixed.

E$2 – In this reference, the dollar sign is right before the row number, and the Column notation has no dollar sign. This means that when you copy the formula down the cells, the reference will not change as the row number is locked.

Changing the cell reference

When you select a cell reference (in the formula bar or in the cell in edit mode) and press F4, it changes the reference.

Suppose you have the reference =A1 in a cell.

Here is what happens when you select the reference and **press the F4 key**.

- Presthe s **F4 key once**: The cell reference changes from A1 to **A1** (becomes 'absolute' from 'relative').
- Press the **F4 key two times**: The cell reference changes from A1 to **A$1** (changes to mixed reference where the row is locked).
- Press the **F4 key three times**: The cell reference changes from A1 to **$A1** (changes to mixed reference where the column is locked).
- Press the **F4 key four times**: The cell reference becomes **A1** again.

Logical and LOOKUP functions

Comperason operators:

= (equal sign) *Equal to* =A1=B1

> (greater than sign) *Greater than* =A1>B1

< (less than sign) *Less than* =A1<B1

>= (greater than or equal to sign) *Greater than or equal to* =A1>=B1

<= (less than or equal to sign*) Less than or equal to* =A1<=B1

<> (not equal to sign) *Not equal to* =A1<>B1

Function	Description	Formula Example	Formula Description
AND	Returns TRUE if all of the arguments evaluate to TRUE.	**=AND(A2>=10, B2<5)**	The formula returns TRUE if a value in cell A2 is greater than or equal to 10, and a value in B2 is less than 5, FALSE otherwise.
OR	Returns TRUE if any argument evaluates to TRUE.	**=OR(A2>=10, B2<5)**	The formula returns TRUE if A2 is greater than or equal to 10 or B2 is less than 5, or both conditions are met. If neither of the conditions it met, the formula returns FALSE.
XOR	Returns a logical Exclusive Or of all arguments.	**=XOR(A2>=10, B2<5)**	The formula returns TRUE if either A2 is greater than or equal to 10 or B2 is less than 5. If neither of the conditions is met or both conditions are met, the formula returns FALSE.
NOT	Returns the reversed logical value of its argument. I.e. If the argument is FALSE, then TRUE is returned and vice versa.	**=NOT(A2>=10)**	The formula returns FALSE if a value in cell A1 is greater than or equal to 10; TRUE otherwise.

IF function

`=IF(C2="Yes",1,2)` *- says IF(C2 = Yes, then return a 1, otherwise return a 2)*

IFS function

`=IFS([Something is True1, Value if True1,Something is True2,Value if True2,Something is True3,Value if True3)`

LOOKUP function

Use LOOKUP when you need to look in a single row or column and find a value from the same position in a second row or column.

- Use **VLOOKUP** to search one row or column, or to search multiple rows and columns (like a table). **It's a much-improved version of LOOKUP**. Watch this video about how to use VLOOKUP.
- In Microsoft 365, use **XLOOKUP** - it's not only faster, but it also lets you search in any direction (up, down, left, right).

```
LOOKUP(lookup_value, lookup_vector, [result_vector])
```

```
VLOOKUP(value, table_array, col_index, approx_match)
```

INDEX – returns a value in a list or a table based on coordinate location

MATCH – locates the position of a lookup value in a row, column, or a table

```
=INDEX(Population, MATCH(A4, Country, 0)
```