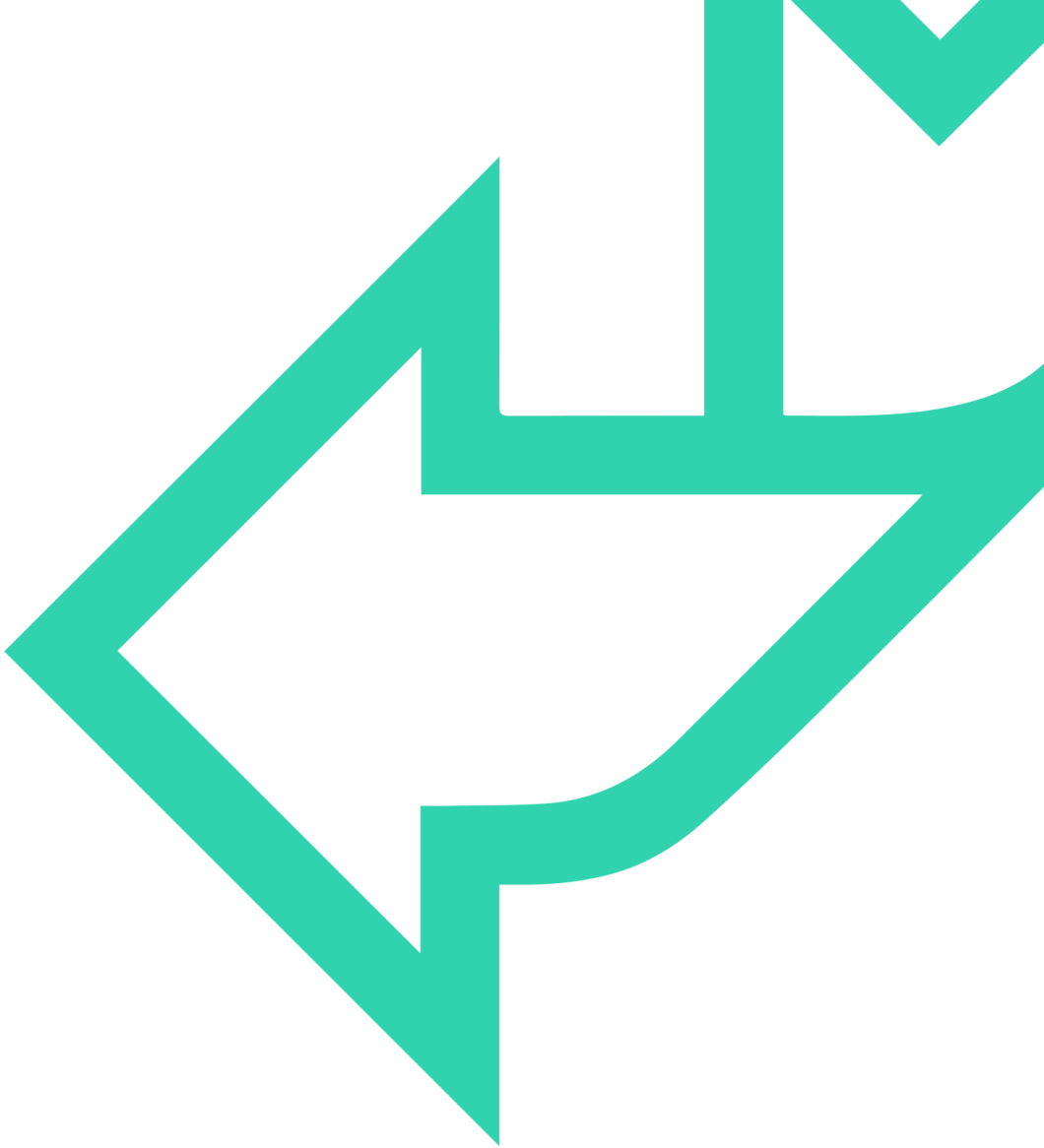
POWERING POTENTIAL

Data Essentials L3

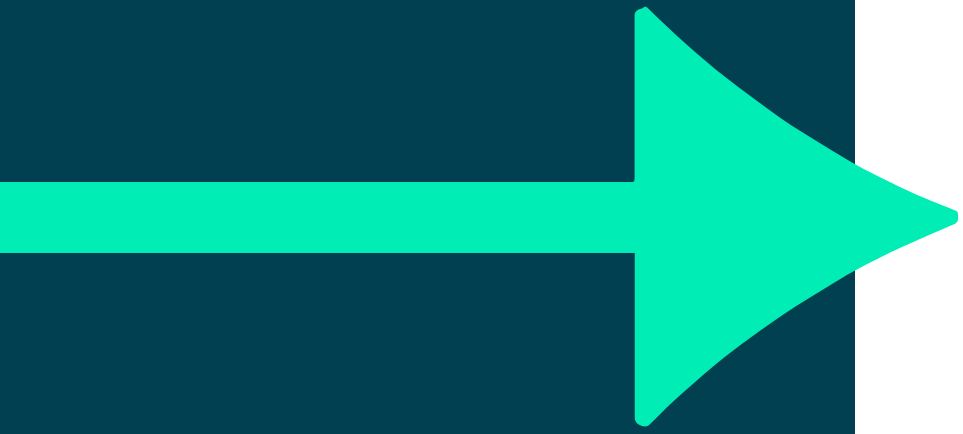
DAY 2



Module 2: From Data to Insight

POWER QUERY ACCESS

Check if you have the Power Query feature in your Excel.



Data blending

DATA BLENDING

What, why, how?

1. What does it mean to ‘blend data’ together?
2. Examples of data sources that can be used to blend data.
3. What are the benefits of data blending?
4. Methods of data blending.
5. Examples of data blending from your own role / organisation. If not applicable, general examples of data blending from any industry.

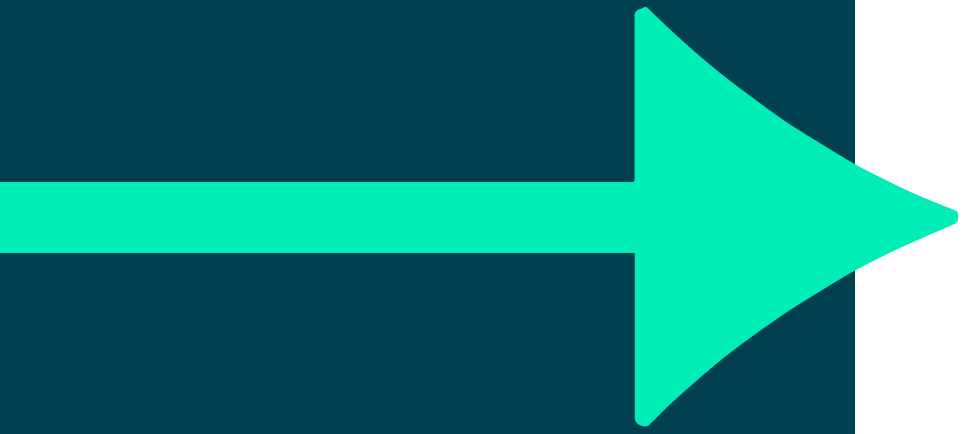
Think, research, take notes.

Share with class.

DATA BLENDING

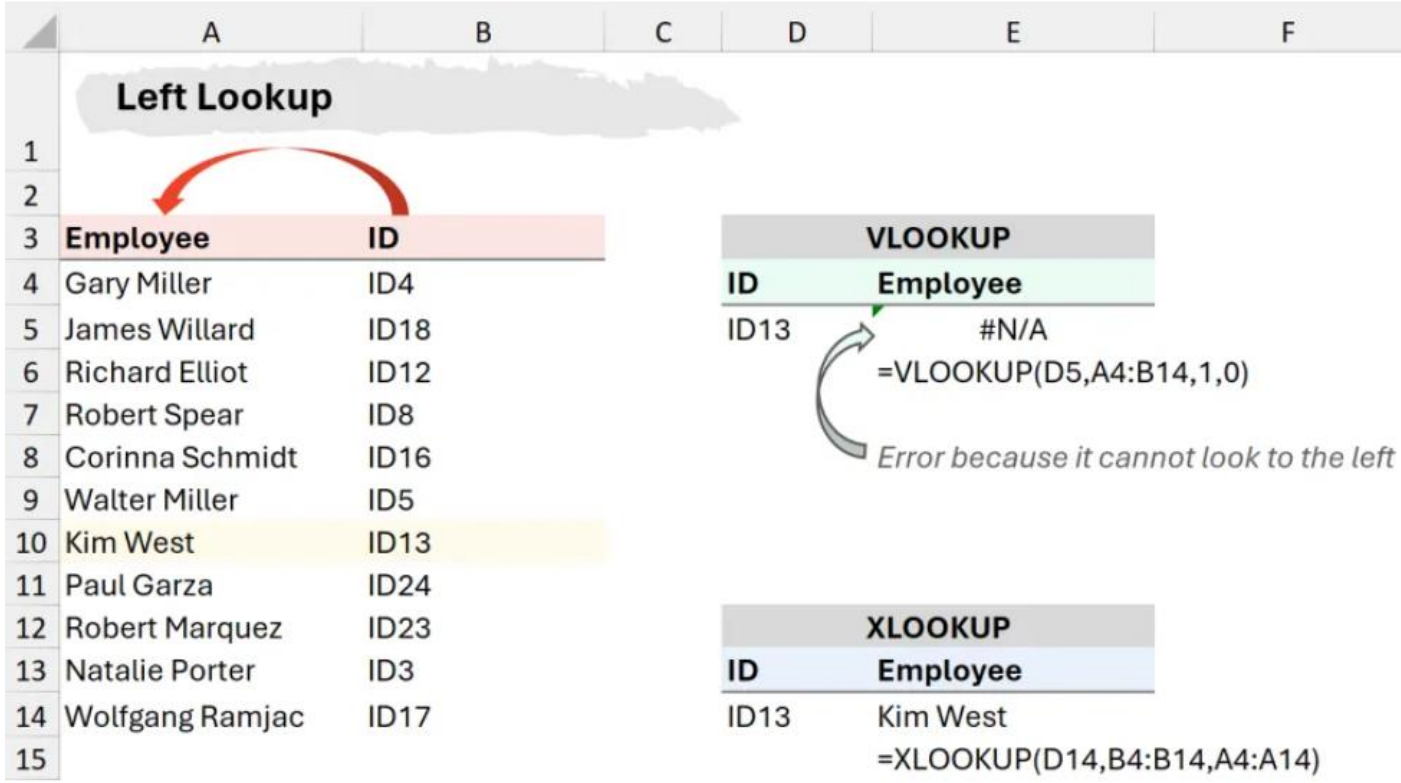
Methods

- Lookup functions (VLOOKUP, XLOOKUP, etc.)
- Joining data (horizontally)
- Merging data (vertically)



LOOKUP FUNCTIONS

- VLOOKUP
- XLOOKUP
- INDEX and MATCH



JOINING MERGING DATA

Order Table

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

Order Value Table

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

- New **columns** are added.
- Based on there being a matching unique field in both tables (needed to join).

APPENDING DATA

Purchase Data 2022

PurchaseID	Purchase Date	Purchase Amount	CustomerID
123	01/01/2022	£123.00	IU7
783	01/02/2022	£78.00	YT5
817	01/03/2022	£12.00	QT6

Purchase Data 2021

Purchase ID	Purchase Date	Purchase Amount	CustomerID
389	01/01/2021	£75.00	QT6
156	01/02/2021	£126.00	JR4
901	01/03/2021	£250.00	NM8

Purchase Data 2020

PurchaseID	Purchase Date	Purchase Amount	CustomerID
739	01/01/2020	£350.00	SD1
198	01/02/2020	£176.00	NM8
272	01/03/2020	£120.00	PW3

Purchase Data 2020-2022

PurchaseID	Purchase Date	Purchase Amount	CustomerID
123	01/01/2022	£123.00	IU7
783	01/02/2022	£78.00	YT5
817	01/03/2022	£12.00	QT6
389	01/01/2021	£75.00	QT6
156	01/02/2021	£126.00	JR4
901	01/03/2021	£250.00	NM8
739	01/01/2020	£350.00	SD1
198	01/02/2020	£176.00	NM8
272	01/03/2020	£120.00	PW3

- New **rows** are added.

Lookup functions

LOOKUP FUNCTIONS

KSBs

S5$_1$ Manipulate and link different data sets as required.

XLOOKUP()

	A	B	C	D	E
1	Turkey	**Ankara**		=XLOOKUP(A1, tblCapitals[Country], tblCapitals[Capital])	
2	*Input*	*Output*			
3					
4	**Country**	**Capital**		What to search for	
5	Costa Rica	San José			
6	Iceland	Reykjavík		Where to search for it	
7	Turkey	Ankara			
8	Wales	Cardiff		What to return	
9	Germany	Berlin			
10	Liberia	Monrovia			
11	Burundi	Bujumbura			
12	Croatia	Zagreb			
13	Nigeria	Abuja			
14					

Open file

‘Learner Guide - Practice with XLOOKUP.pdf’

Joining tables

JOINING TABLES

KSBs

K6$_{1+2}$ The value of data to the business. How to undertake blending of data from multiple sources.

K9$_1$ Basic statistical methods and simple data modelling to extract relevant data and normalise unstructured data.

S3$_1$ Summarise and explain gathered data.

S4$_1$ Blend data sets from multiple sources and present in format appropriate to the task.

JOINING TABLES

Order Table

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

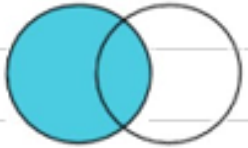
Order Value Table

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

JOINING TABLES

LEFT JOIN

Left Join



OrderCustomer
(Left Table)

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

Left Join

OrderValue
(Right Table)

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

OrderNum	TotalValue
4000	£58.97
	£9.32

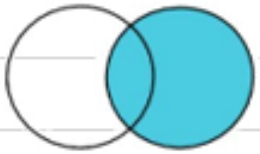
OutputLeftJoin

OrderID	Customer	OrderNum	TotalValue
1000	Alice	1000	31.14
1000	Alice	1000	15.92
2000	Bob	2000	6.53
2000	Cathy	2000	6.53
	Diego		
3000	Emily		

JOINING TABLES

RIGHT JOIN

Right Join



OrderCustomer
(Left Table)

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

Right Join

OrderValue
(Right Table)

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

OrderID	Customer
	Diego
3000	Emily

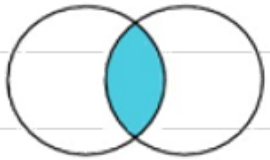
OutputRightJoin

OrderID	Customer	OrderNum	TotalValue
1000	Alice	1000	31.14
1000	Alice	1000	15.92
2000	Bob	2000	6.53
2000	Cathy	2000	6.53
		4000	58.97
			9.32

JOINING TABLES

INNER JOIN

Inner Join



OrderCustomer
(Left Table)

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

Inner Join

OrderValue
(Right Table)

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

OrderID	Customer
	Diego
3000	Emily

OrderNum	TotalValue
	£58.97
	£9.32

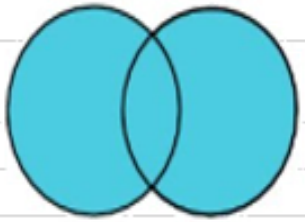
OutputInnerJoin

OrderID	Customer	OrderNum	TotalValue
1000	Alice	1000	31.14
1000	Alice	1000	15.92
2000	Bob	2000	6.53
2000	Cathy	2000	6.53

JOINING TABLES

FULL(OUTER) JOIN

Full (Outer) Join



OrderCustomer
(Left Table)

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

Full Join

OrderValue
(Right Table)

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

OutputFullJoin

OrderID	Customer	OrderNum	TotalValue
1000	Alice	1000	31.14
1000	Alice	1000	15.92
2000	Bob	2000	6.53
2000	Cathy	2000	6.53
	Diego		
3000	Emily		
		4000	58.97
			9.32

EXPERIMENTS WITH JOINS

OrderCustomer
(Left Table)

OrderID	Surname	Firstname
1	Smith	John
2	Johnson	Jane
3	Williams	Robert
4	Davis	Emily
5	Wilson	Sarah
6	Anderson	Michael
7	Taylor	Jessica
8	Harris	David
10	Mariya	Khan

OrderValue
(Right Table)

OrderNum	Surname	OrderTotal
1	Smith	150
2	Johnson	200.5
3	Williams	75.2
4	Davis	300
NULL	Wilson	120.75
6	Anderson	90.5
7	Taylor	NULL
9	Patel	250.8

Open file

'Joins Playground.pdf'

LOOKUP FORMULAS VS. JOINS

OrderCustomer

(Left Table)

OrderID	Customer
1000	Alice
2000	Bob
2000	Cathy
	Diego
3000	Emily

OrderValue

(Right Table)

OrderNum	TotalValue
1000	£31.14
1000	£15.92
2000	£6.53
4000	£58.97
	£9.32

Output

OrderID	Customer	OrderNum	TotalValue
1000	Alice	1000	31.14
1000	Alice	1000	15.92
2000	Bob	2000	6.53
2000	Cathy	2000	6.53
	Diego		
3000	Emily		

NULL

Question

Could this left join operation be achieved in Excel using **only** lookup functions such as XLOOKUP()? **(No copy and paste!)**

If so: prove it!

If not: explain why not.

KNOWLEDGE CHECK: JOINS

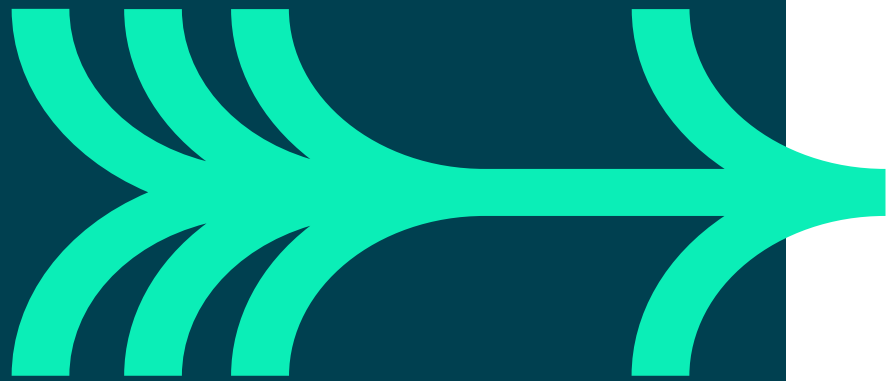
Joins

Question 1

What are the four basic types of join introduced earlier?

A: Middle, left, right, full
B: Up, left, right, down
C: Inner, left, right, full
D: Inner, left, right, outer

KNOWLEDGE CHECK: JOINS



Joins

Question 2

In a left join, which rows are **guaranteed** to make it into the output table?

A: All rows from the Left Table

B: All rows from the Right Table

C: All rows from both the Left and Right tables

D: No rows are guaranteed to make it into the output table

KNOWLEDGE CHECK: JOINS

Joins

Question 3

Which type of join do we need if the output table must contain only rows from the Left Table that have a match in the right Table **and vice versa**?

A: Inner join
B: Left join
C: Right join
D: Full join

JOIN (MERGE) DATA

POWER QUERY

Open file

‘Merge Data For PQ..xlsx’

APPEND DATA

POWER QUERY

Open file

‘Append Data For PQ.xlsx’