Data Analysis & Business Intelligence Made Easy with Excel Power Tools Excel Data Analysis Basics = E-DAB

Notes for Video:

E-DAB 07: Excel Data Analysis & BI Basics: Data Modeling: Excel Formulas, Power Query, Power Pivot?

Outcomes for Video:

1.	Data Modeling	2
2.	VLOOKUP Function	6
3.	Multiple Tables: Fact Tables and Dimension Tables (Lookup Tables)	10
4.	Power Pivot is just one of Many Tools in Excel	11
5.	Excel Power Pivot and Data Model PivotTables	12
2) Show Power Pivot Ribbon Tab in Excel	12
Z) Excel Power Pivot provides 3 Data Tools	13
5) Why the name Power Pivot?	13
8	3) Basic Advantages of Excel Power Pivot	14
g) Relationship feature works in versions of Excel 2013 or later	14
1	.0) DAX Formulas	14
6.	Implicit vs. Explicit DAX Measures :	15
7.	Power Query Merge feature	17
8.	Data Types in Power Query	18
9.	Overview of Three Examples in Video	19
10.	VLOOKUP Video Example	20
11.	Power Query Video Example	21
12.	Power Pivot Relationships feature & Implicit Measure feature	22

1. Data Modeling

- 1) Define Data Modeling:
 - 1. Configuring Raw Data into Proper Data Sets that can be used for creating information easily with tools like PivotTables, Power Pivot, Power BI Desktop and other tools.
- 2) Tools we use for Data Modeling:
 - 1. Data Modeling can be accomplished with many different tools such as Excel Spreadsheet Formulas, Excel features such as Text To Columns, Flash Fill, DAX Formulas and more, but the main tool we use in Excel and in Power BI is Power Query.
- 3) So far in this class, we have performed Data Modeling to convert Raw Data into a single Proper Data Set, such as:
 - 1. In Video #6, we used Power Query to Split by Delimiter to create a single Proper Data Set that we used as the source data for a PivotTable Report, as seen here:

	А	В	С	D	E	F	G
11	Description	Amount		Product	Date	Region	Sales
12	Carlota / 05/05/2016 / West	485		Carlota	5/5/2016	West	485
13	Aspen / 04/02/2016 / South	804		Aspen	4/2/2016	South	804
14	Yanaki / 05/02/2016 / South	790	==>>	Yanaki	5/2/2016	South	790
15	FlatTop / 04/11/2016 / South	965		FlatTop	4/11/2016	South	965
16	Carlota / 05/21/2016 / South	620		Carlota	5/21/2016	South	620
47		205			E /4 E /204 C	- ·	205

2. In Video #6, we used Power Query to append multiple Text Files into a single Proper Data Set:

TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT		CONTRACTOR DE LA CONTRACT			Product	Date	Region	Sales
-	-	-	-		Carlota	5/5/2016	West	485
				==>>	Aspen	4/2/2016	South	804
					Yanaki	5/2/2016	South	790
Aprilitit	July.tst	June.bit	May.bd		FlatTop	4/11/2016	South	965
					Carlota	5/21/2016	South	620
					A 1.1	Electore C		20

4. In example #1 this video, Video #7, we will use Spreadsheet Functions to gather the raw data from three different tables and converge it into a single Proper Data Set that we can use to build a requested report, as seen on the next page:

Start Tables:

in the

act Table = fTransactionsV			Dimension	Dimension Table = dProductV					Dimension Table = dWebSiteV		
ate 💌 Web Site ID 💽	Product ID 💌 Un	iits 💌	Product II	Product	▼ Flight Range (m)	Product Category Price	*	Cost 💌	Web Site ID 💌	Web Site Name	
12/8/19 cb69	S40DIS	1	Q20FRE	Quad		20 Freestyle	43.95	23.5	gb43	Gel Boomerangs	
12/18/20 am11	M50DIS	5	A22BEG	Aspen		22 Beginner	26.95	16.33	cb69	Colorado Boomeran	
12/9/19 gb43	A22BEG	3	\$40D15	Sunset		40 Distance	29.95	12.64	am11	Amazon	
4/10/20 gb43	A22BEG	1	C20FRE	Carlota	3	20 Freestyle	27.95	16.06	eb92	E-Bay	
2/25/19 cb69	A22BEG	1	8258EG	Bellen	3	25 Beginner	24.95	15.31			
12/23/20 gb43	A228EG	1	Y28BEG	Yanaki		28 Beginner	22.95	13.19			
12/1/20 cb69	Q20FRE	1	T20FRE	TriFly	3	20 Freestyle	19.95	8.53			
12/10/19 gb43	Q20FRE	2	F45DIS	Flattop		45 Distance	45.95	24.06			
12/23/19 am11	Y288EG	2	MSODIS	Majestic Bea	aut	50 Distance	49.95	26.29			
5/8/20 gb43	8258EG	2									

Excel Spreadsheet Formulas to create Data Model:

4	A	8	С	D	E	F	G	н	1	1	K	L	M
2				+ 1 1									
3		Data Mo	deling with	Excel Sprea	adsheet	Formu	las:						
4		Formula i	n cell F12: ='	VLOOKUP([@[Prod	uct ID]]	dProductV	,5,0)					
5		Formula i	n cell G12: =	[@Price]*[@Units]								
6		Formula i	n cell H12: =	VLOOKUP(@[Prod	uct ID]]	,dProduct\	/,4,0)					
7		Formula i	n cell 12: =\	LOOKUP([@[Web	Site ID]	,dWebSite	≥V,2,0)					
8													
9		Fact Table =	fTransactionsV								Dimension Table =	ProductV	
10													
11		Date 💌	Web Site ID 🎽	Product ID	Units *	Price 👕	Revenue 💌	Product Category	🕈 Web Site Name 🛛 🗶		Product ID	 Product 	 Flight Range (m)
12		12/8/19	cb69	S40DIS	1	29.95	29.95	Distance	Colorado Boomerangs		Q20FRE	Quad	20
13		12/18/20	am11	M50DIS	5	49.95	249.75	Distance	Amazon		A22BEG	Aspen	22
14		12/9/19	gb43	A22BEG	3	26.95	80.85	Beginner	Gel Boomerangs		S40DIS	Sunset	40
15		4/10/20	gb43	A22BEG	1	26.95	26.95	Beginner	Gel Boomerangs		C20FRE	Carlota	20
16		2/25/19	cb69	A22BEG	(1	26.95	26.95	Beginner	Colorado Boomerangs		B25BEG	Bellen	25
17		12/23/20	gb43	A22BEG	1	26.95	26.95	Beginner	Gel Boomerangs		Y28BEG	Yanaki	28
18		12/1/20	cb69	Q20FRE	1	43.95	43.95	Freestyle	Colorado Boomerangs		T20FRE	TriFly	20
19		12/10/19	gb43	Q20FRE	2	43.95	87.9	Freestyle	Gel Boomerangs		F45DIS	Flattop	45
20		12/23/19	am11	Y288EG	2	22.95	45.9	Beginner	Amazon		M50DIS	Majestic Beaut	50
21		5/8/20	gb43	B25BEG	2	24.95	49.9	Beginner	Gel Boomerangs				

Product Category	▼ Amazon	Colorado Boomerangs	E-Bay	Gel Boomerangs	Grand Total
Beginner	174,570	251,028	84,294	341,246	851,138
Distance	152,268	215,191	76,386	303,163	747,00
Freestyle	251,203	419,389	135,485	531,404	1,337,48
Grand Total	578,041	885,607	296,164	1,175,814	2,935,626
Sum of Revenue	Web Site Name				
Product Category	Amazon	Colorado Boomerangs	E-Bay	Gel Boomerangs	Grand Total
Beginner	5.95%	8.55%	2.87%	11.62%	28.999
Distance	5.19%	7.33%	2.60%	10.33%	25.459
Freestyle	8.56%	14.29%	4.62%	18.10%	45.569
Grand Total	19.69%	30.17%	10.09%	40.05%	100.009

5. In example #2 this video, Video #7, we will use Power Query to gather the raw data from two different tables in an Access database abd one table from an Excel Sheet and converge it into a single Proper Data Set that we can use to build a requested report, as seen below:



6. In example #3 this video, Video #7, we will use Power Pivot's Relationship feature and Implicit Measure feature to show three tables (from an Excel Sheet) in the PivotTable Field List and create our desired report, as seen here:

12/18/20 am11 12/9/19 gb43 4/10/20 gb43 2/25/19 cb69 12/23/20 gb43 12/1/20 cb69 12/10/19 gb43	Product ID Units S40DIS 1 M50DIS 5 A228EG 3 A228EG 1 A228EG 1 A228EG 1 A228EG 1 Q20FRE 1 Q20FRE 1 Q20FRE 2 Y28BEG 2	S40DIS Su C20FRE Ca B25BEG Be Y28BEG Ya T20FRE Tri F45DIS F12		tit Range (m) ♥ Prod 20 Free 22 Beg 40 Dist 20 Free 25 Beg 28 Beg 20 Free 45 Dist 20 Free 28 Beg 20 Free	estyle inner ance estyle inner inner	rice 43.95 26.95 29.95 27.95 24.95 24.95 22.95	Cost 23.5 16.33 12.64 16.06 15.31	Web Site ID gb43 cb69 am11 eb92	Web Site Name Gel Boomerangs Colorado Boomerangs Amazon E-Bay
12/18/20 am11 12/9/19 gb43 4/10/20 gb43 2/25/19 cb69 12/13/20 gb43 12/12/0 cb69 12/10/19 gb43 12/23/19 am11	M50DIS 5 A22BEG 3 A22BEG 1 A22BEG 1 A22BEG 1 Q20FRE 1 Q20FRE 2	A228EG As S40DIS Su C20FRE Ca B258EG Be Y288EG Ya T20FRE Tri F45DIS F12	pen nset rlota Ilen naki Fly sttop	22 Beg 40 Dist 20 Free 25 Beg 28 Beg 20 Free 45 Dist	inner ance estyle inner inner	26.95 29.95 27.95 24.95	16.33 12.64 16.06 15.31	cb69 am11	Colorado Boomerang: Amazon
12/9/19 gb43 4/10/20 gb43 2/25/19 cb69 12/23/20 gb43 12/1/20 cb69 12/10/19 gb43 12/23/19 am11	A22BEG 3 A22BEG 1 A22BEG 1 A22BEG 1 Q20FRE 1 Q20FRE 2	S40DIS SU C20FRE Ca B25BEG Be Y28BEG Ya T20FRE Tri F45DIS F12	nset rlota Ilen naki Fly sttop	40 Dist 20 Free 25 Beg 28 Beg 20 Free 45 Dist	ance estyle inner inner	29.95 27.95 24.95	12.64 16.06 15.31	am11	Amazon
4/10/20 gb43 2/25/19 cb69 12/23/20 gb43 12/12/0 cb69 12/20/20 gb43 12/20/19 gb43 12/23/19 am11	A228EG 1 A228EG 1 A228EG 1 Q20FRE 1 Q20FRE 2	C20FRE Ca B25BEG Be Y28BEG Ya T20FRE Tri F45DIS FIz	rlota Ilen naki Fly sttop	20 Free 25 Beg 28 Beg 20 Free 45 Dist	istyle inner inner	27.95 24.95	16.06 15.31	para la companya de la	
2/25/19 cb69 12/23/20 gb43 12/1/20 cb69 12/10/19 gb43 12/23/19 am11	A228EG 1 A228EG 1 Q20FRE 1 Q20FRE 2	B25BEG Be Y28BEG Ya T20FRE Tri F45DIS FIE	ilen naki Fly sttop	25 Beg 28 Beg 20 Free 45 Dist	inner inner	24.95	15.31	eb92	E-Bay
12/23/20 gb43 12/1/20 cb69 12/10/19 gb43 12/23/19 am11	A22BEG 1 Q20FRE 1 Q20FRE 2	Y28BEG Ya T20FRE Tri F45DIS Fla	naki Fly httop	28 Beg 20 Free 45 Dist	inner		the second se		
12/1/20 cb69 12/10/19 gb43 12/23/19 am11	Q20FRE 1 Q20FRE 2	T20FRE Tri F45DIS FIa	Fly attop	20 Free 45 Dist		22.95			
12/10/19 gb43 12/23/19 am11	Q20FRE 2	F45DIS FI	attop	45 Dist	style	10.05	13.19		
12/23/19 am11		and the second se				19.95	8.53 24.06		
				50 Dist	ance	49.95	26.29		
Text to Flash Rem Columns Fill Dupli Ianage Relationships tatus A Table tive ffransactions tive ffransactions	oove Data Conse cates Validation ~ Data Tools R (Web Site ID)	dWebSite	ps Manage Data Model			Navdshen) Web Sile ID) Web Sile Nave 3 -	Contraction	Re ID	dProduct8 Product8 Product00 Product Product Product Product Congrey Price Cont
Sum of Units Product Category	Web Site Name 🕞	lorado Boomeran	gs E-Bay 941 3,345	Gel Boomerangs 13,501	Grand Total 33,686	PivotTable Fi		• ×]
Distance	3,912		546 1,968	(1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.	19,204	Search		A	
Freestyle Grand Total	7,950	13,			42,313	dProductR			
		28, Iorado Boomeran		38,070 Gel Boomerangs	95,203 Grand Total	Product ID Product Flight Range Product Cat Price Cost			
Beginner	7.25%	10.4	4% 3.51%	14.18%	35.38%	_] ∰ dWebSiteR			
Distance	4.11%	5.8	3% 2.07%	8.17%	20.17%	Web Site ID			
Distance	8.35%	13.9	4.52%	17.64%	44.45%	记 Web Site N	ime		
Freestyle	19.71%	30.2	0% 10.10%	39.99%	100.00%	J III fTransactionsR			1

2. VLOOKUP Function

- 1) Looking things up in Lookup Tables is a common task in business, accounting and other professions.
- 2) Almost all Lookup Tables are Vertical because the first column contains the item that we try to match, and items are listed vertically.
 - i. Examples of Looking up items in a Vertical Lookup Table:
 - 1. This is a Price Lookup Table:

What is the price for Quad?

Product	Supplier	Price
Flying Eagle	Channel Craft	\$19.95
V Range	Colordao	\$18.95
Quad	Gel Boom	\$43.95
Bellen	Gel Boom	\$26.95
Carlota	Gel Boom	\$27.95

2. This is a Commission Bonus % Lookup Table:

If your sales are \$36,500, what is the Commission Bonus you earned?

		Bonus
Sales Lower Limit	Category	Commission %
	\$0 or more, all the way up to, but	
0	not including \$10,000	0
	\$10,000 or more, all the way up to,	
10000	but not including \$20,000	0.01
	\$20,000 or more, all the way up to,	
20000	but not including \$30,000	0.02
	\$30,000 or more, all the way up to,	
30000	but not including \$40,000	0.04
	\$40,000 or more, all the way up to,	
40000	but not including \$50,000	0.05
	\$50,000 or more, all the way up to,	
50000	but not including \$60,000	0.06
60000	Bigger than or equal to \$60,000	0.075

3. This is an Employee Lookup Table:

What is "Chukes, Hal" Zip Code?

Address	City	State Zip	
3557 1st St.	Seattle	WA	98114
6281 173rd St.	Tacoma	WA	98131
7149 1st Ave.	Seattle	WA	98133
7726 66th Ave.	Tacoma	WA	98111
316 66th Blvd. 151st St.	Kent	WA	98124
	3557 1st St. 6281 173rd St. 7149 1st Ave. 7726 66th Ave.	3557 1st St.Seattle6281 173rd St.Tacoma7149 1st Ave.Seattle7726 66th Ave.Tacoma	3557 1st St.SeattleWA6281 173rd St.TacomaWA7149 1st Ave.SeattleWA7726 66th Ave.TacomaWA

4. This is a Tax Lookup Table:

If your Income is \$3000, what is your tax?

Income	Tax	
\$0	\$	0.00
\$1,000	\$2	5.00
\$2,000	\$6	0.00
\$5,000	\$12	0.00
\$10,000	\$17	5.00

5. This is a Region Lookup Table:

What Region does "Miles, Josefina" represent?

Sales Rep	Region
Anderson, Robin	Midwest
Bryan, Viola	West
Gonzalez, Danny	East
Miles, Josefina	Southwest
Nguyen, Sheldon	East
Nichols, Claudia	Northwest
Pittman, Otis	West
Richardson, Kristina	Northwest
Rodgers, Bob	Midwest
Stokes, Taylor	Southwest

6. This is a Commission Bonus \$ Lookup Table:

If your sales are \$6000, what is your commission?

Your Sales	Commission
\$0	\$0
\$5,000	\$200
\$12,500	\$625
\$15,000	\$975
\$20,000	\$1,700

7. This is a Sales Category Table:

If your sales are \$375, what is the category?

Sales Lower Limit	Upper & Lower Limits	Category
	\$0 or more, all the way up to, but	
0	not including \$100	Below Par
	\$100 or more, all the way up to,	
100	but not including \$250	Par
	\$250 or more, all the way up to,	
250	but not including \$500	Average
	\$500 or more, all the way up to,	
500	but not including \$600	Above Average
600	Bigger than or equal to \$600	Excellent

- 3) What does VLOOKUP Function do?
 - i. VLOOKUP tries to find a match of an item in the first column of the Lookup Table and then retrieves (goes and gets) something from one of the other columns in the table and bring it back to the cell or formula.
 - ii. In VLOOKUP the V means Vertical.
 - Example: VLOOKUP can find a match for the Sales Number 17,382 in the sorted first column of the Lookup Table and retrieve the correct Bonus Commission %, 1.00%, from the 2nd column and bring it back to the cell C30, like in this picture:

1	A	В	C	D	E	F	G
29	Employee	Sales	Bonus Commission %?				
30	Rod Barbaria	17,382.00	=VLOOKUP(B30,\$E	\$32:\$F\$38,2)			
31	Elinore Cromedy	19,504.00	1.00%		Sales	Bonus Comm. %	
32	Twanda Spruce	29,999.99	2.00%		0	0	
33	Darius Raffety	30,000.01	4.00%		10000	0.01	
34	Frida Vickerson	16,081.00	1.00%		20000	0.02	
35	Sioux Radcoolinato	39,999.99	4.00%		30000	0.04	
36	Abdi Han	40,000.00	5.00%		40000	0.05	
37	Philly Johns	62,552.12	7.50%		50000	0.06	
38					60000	0.075	
39							
40							
41							

iv. Example: VLOOKUP can find a match for the Product "Quad" in the first column of the Lookup Table and retrieve the Quad's Price, 43.95, from the 3rd column and bring it back to the cell F23.

	A	В	С	D	E	F	G
46	Product	Price			Product	Supplier	Price
47	Quad	=VLOOKUP(A47,\$	E\$47:\$G\$51,3,0)		Flying Eagle	Channel Craft	\$19.95
48	Bellen	26.95			V Rang	Colorado	\$18.95
49	Flying Eagle	19.95			Quad	Gel Boom	\$43.95
50					Bellen	Gel Boom	\$26.95
51					Carlota	Gel Boom	\$27.95
52							

4) VLOOKUP Function arguments:

=VLOOKUP(lookup_value , table_array , col_index_num , [range_lookup])

- i. **lookup_value** = Item that you are trying to find in first column of lookup table.
- ii. **table_array** = Vertical table = Lookup table. First Column contains items you want to "match" with the lookup_value.
- iii. col_index_num = Which column in the lookup table has the items that you want to go and get and bring back to the cell? You have to count to determine which columns contains the items you want to retrieve: is it column 2, or column 3, or column 4, and so on.
- iv. **[range_lookup]** = Because there are two different types of lookup, we must tell VLOOKUP which of the two lookups we want it to do: either: Approximate Match Lookup or Exact Match Lookup. This argument tells VLOOKUP how to find a match in the first column of the Lookup Table.
 - 1. Approximate Match:
 - For "<u>Approximate Match</u>" we must put = TRUE or 1 or omitted.

TRUE - Approximate match FALSE - Exact match

- How Approximate Match works:
 - 1. For Approximate Match the VLOOKUP table MUST be sorted on the first column: Ascending, A to Z (Small to Big).
 - 2. This is how Approximate Match Lookup works:
 - i. It will look through the first column:
 - If the first value in the table is smaller than the lookup_value, VLOOKUP returns a Not Available Error: #N/A!
 - 2. Then it looks through the first column until it bumps into the first value bigger than it and then jump back one row. When it finds a match, it knows what row it should look in.
 - It actually does a "binary search", which is a technical computer term for "Approximate Match". "Binary Search" calculates quickly compared to "Exact Match".
 - 3. If the lookup_value is bigger than the last value, it stops at the last row.

2. Exact Match:

• For "Exact Match" we must put = FALSE or 0.

TRUE - Approximate match

- How Exact Match works:
 - 1. VLOOKUP will look though each item in the first column of the VLOOKUP table and try to find a match. When it finds a match, it knows what row it should look in.
 - 2. If VLOOKUP cannot find a match it will it will give an #N/A! error that tells you it did not find a match "it is not available".
- **Note about Exact Match**: If you have very large data sets, Exact Match Lookup may cause formula to calculate slowly because "Exact Match" Lookup must look through every item, one-by-one, until it finds a match.

3. Multiple Tables: Fact Tables and Dimension Tables (Lookup Tables)

- 1) In much of Data Analysis, Business Intelligence and Data Warehousing, we usually refer to our Proper Data Sets as either a Fact Table or a Dimension Table, as defined here:
 - 1. Fact Table (Also known as Transaction Table or Sales Table)
 - i. A Fact Table is a table that has numbers we need to summarize (like Sales or Units or dates or times). The word "Fact" equals a measurement of business activities (like amount of sales, or how many units sold, or clicks on web links)
 - ii. A Fact Table has Foreign Key columns that we will use in relationships (like Date, SalesRepKey, ProductKey)
 - 2. Dimension Table (also known as Lookup Table or Entity Table)
 - i. A Dimension Table is a table that has the first column as a Primary Key (Unique Identifier) for the Entity (Product, Sales Rep and so on) used in relationship with Fact Table.
 - ii. Remaining columns are attributes that we can use as:
 - Criteria / Filters / Categories / Report Labels for our Reports & Dashboards. These columsn are often referred to as "Filtering Column".
 - 2. Values we can lookup (like Price).
 - 3. Helper Columns (like Sort Helper Columns or Intermediate calculations).
 - iii. Synonyms for Filtering columns: Attribute, Criteria, Filters, Category, Categorical attributes, Constraints, Groupings, Report labels for the reports and analytics.
- 2) Fact Tables are usually large. Dimension Tables are comparatively small.
- 3) Example:

8) Fact Tables and Dimension Tables

Fact Tables have the Numbers we need to summarize (like Sales or Units)

This is called the Foreign Key (because there are duplicates). This is the Many Side in a Many-To One Relationship.

Date	ProductID	Units Sold	Sales (\$)
2/15/19	1043	6	287.7
2/15/19	1069	3	89.85
2/15/19	1069	8	239.6
2/15/19	2005	3	65.85
2/15/19	1069	11	329.45
2/15/19	1043	8	383.6
2/15/19	2005	6	131.7
2/15/19	2005	13	285.35
2/15/19	1043	5	239.75
2/15/19	1043	5	239.75

Dimension Tables, or Lookup Tables, have unique lists of attributes used as conditions to make the Fact Table calculations

This is called the Primary Key (because it is a Unique List). This is the One Side in a Many-To One Relationship.

ProductID	Product	Flight Rang	Retail Price
1043	Quad	20	47.95
1069	Aspen	25	29.95
2005	Sunspot	15	21.95

4. Power Pivot is just one of Many Tools in Excel

- 1) Excel is a program with many tools
- 2) Power Pivot is just one of the tools in Excel
- 3) The tool Power Pivot has two main parts:
 - i. Data Model
 - ii. Data Model PivotTables
- 4) The Data Model is made up of three main parts:
 - i. Columnar Database
 - ii. Relationships
 - iii. DAX Formulas
- 5) From the Data Model, we make Data Model PivotTables.
- 6) Picture:

Excel									
	Excel has many tools:								
Formulas	Excel Tables	Standard PivotTables	Power Query	Many More	Power Pivot				
				-12	1) Data Model				
					i) Columnar Database				
					ii) Relationships				
					iii) DAX Formulas				
					2) Data Model PivotTables				

5. Excel Power Pivot and Data Model PivotTables

- 1) Excel Power Pivot comes in Office 365.
- 2) Show Power Pivot Ribbon Tab in Excel
 - i. Click on the File Tab, then click on Options, then on the left, click on Add-ins, then in the Manage textbox dropdown, select "Com Add-ins", then check the check Box for Power Pivot.
 - ii. In Excel, the Power Ribbon Tab looks like this:

Click the "Manage Data Model" button to open Power Pivot Window to look at the Data Model

	\$ ₹										E-	DAB-07	7-DataModeling-Start.xlsx -	- Saved
File	[<u>fx</u>]	Insert	Page Layout		Data	Review	View	Developer	Add-ins	Help	Power Pivot	Q	Search	
Manage Data Mode	Measures ~ Calculat	KPIs ~ ions	Add to Data Model Tables	Set	ttings	Use t	his butt	on to add Exo	cel Tables f	rom an	Excel Worksh	eet to	the Data Model	

iii. When you open the Power Pivot Window, and click on Design View, you can see the Data Model that you created with the Relationships feature, as seen here:



- 3) Excel Power Pivot allows us to build "Data Model" PivotTables, as opposed to "Standard PivotTables"
 - i. Reminder from E-DAB video #4:
 - Standard PivotTables:
 - Use when you have about 50,000 rows of data or less.
 - Use when you have one Proper Data Set with all your Data.
 - You don't mind applying Number Formatting every time you make a PivotTable Calculation.
 - PivotTable Calculations are sufficient.

Data Model PivotTables:

- Big Data. Good for two reasons when you have large data sets:
 - i. File size is reduced when your Data is in the Data Model.
 - ii. You can easily build reports from millions of rows of data (Excel Spreadsheet only allows 1 million rows)
- Relationships for Multiple Tables. Great when you have more than one Proper Data Set as the source Data. Allows you to use the Relationship feature rather than the VLOOKUP function when you need to connect tables.
- DAX Formulas.
 - i. Has more options for calculations than a Standard PivotTable.
 - ii. Allows you to add Number Formatting to Formulas.
 - iii. Can use the same formula over and over.
- 4) Excel Power Pivot provides 3 Data Tools
 - i. Columnar Database = Behind the scenes in RAM Memory Efficient Big Data Analytics Database
 - ii. Relationships Between Tables = replace VLOOKUP and allow criteria and filters to affect reports and visualizations from one table to another.
 - iii. DAX Formulas:
 - 1. Efficiently Calculate Over Big Data.
 - 2. Many More Calculations than in Standard PivotTable
 - 3. Build One Formula that can work in many reports
 - 4. Add Number Formatting to Formulas
- 5) Why the name Power Pivot?
 - i. Because Microsoft wanted to use the same amazing PivotTable user interface to drag and drop fields to make reports but with more Power.
 - ii. The "Power" part of the name means:
 - 1. We can make PivotTables from "Big Data"
 - 2. We can make PivotTables from multiple Tables
 - 3. We can use DAX Formulas, which can process over big data efficiently and which allows us more varied calculations than in a Standard PivotTable.
 - iii. The "Pivot" part of the name means we can use a PivotTable user interface, that we all know and love!
- 6) Data Model = Name for Power Pivot's 3 Data Tools :
 - i. The Columnar Database, Relationships and DAX Formulas together are called the "Data Model".
- 7) From the Data Model we make Data Model PivotTables.
 - i. Synonyms for Data Model PivotTable:
 - 1. Power Pivot Report
 - 2. Power Pivot PivotTable
 - ii. Data Model PivotTables create summary reports with one or more calculations based on conditions / criteria / filters

8) Basic Advantages of Excel Power Pivot

- i. Can work on Millions of rows of data
- ii. Can Reduce file size on data sets with less than a million rows
- iii. Can use Relationships and Multiple Tables rather than VLOOKUP and a single Flat Table.
- iv. DAX formulas provide more variety that in a Standard PivotTable and can work efficiently on Big Data that is stored in the Columnar Database.
- 9) Relationship feature works in versions of Excel 2013 or later
 - i. This means that if you have Excel 2013 or 2016, but you do not have the correct version with the Power Pivot Com Add-in, you can still use the Relationships feature to add two or more tables to a PivotTable field list and then make a PivotTable based on multiple tables. However, if you do not have the correct version with the Com-Add-in, you will not be able to work in the Power Pivot Data Model Window.
- 10) DAX Formulas
 - i. DAX = Data Analysis Expressions
 - ii. Types of DAX Formulas:
 - 1. **Calculated Column** = New Columns add to tables in the Data Model. Video #8 will demonstrate Calculated Columns.
 - 2. **Measures** = Formulas used in Data Model PivotTables.
 - Two types of Measures:
 - i. Implicit Measures = formulas automatically created by Power Pivot.
 Video #7 will demonstrate Implicit Measures.
 - ii. **Explicit Measures** = formulas that Data Modeler creates. Video #8 will demonstrate Explicit Measures.
 - DAX Measures are different that the built-in calculation sin a Standard PivotTable, like "Summarize Values By" and "Show Values As".
 - When you create a DAX Measure you create a formula using DAX Functions like SUM, SUMX, AVERAGEX, CALCULATE, REALTED and others.
 - 3. **Table Formulas** = deliver a table of values. Video #9 will demonstrate Table Formulas.

6. Implicit vs. Explicit DAX Measures :

1)	To show Imp	licit Measures:				
	i. In Ex	cel, go to the Po	ower Pivot Rib	bbon Tab, then in the Data Mod	el group, clic	k Manage button.
		opens up the Po				
Show Implicit Measure button	Impl table	icit Measures Bu es in Power Pivo	utton. This wi t Data Model	go to the Advanced Ribbon Tab Il show the Implicit Measures in). leasure sin the Measure Grid in t	the Measure	e Grid (area below
					Power P	ivot Window
🌗 🚺 📕 숙 File Hor	52 (SS)S	ower Pivot for Exce Advanced	el - E-DAB-07-F	inished-Homework.xlsx	·	
Create and Manage	ect : <default> ▼ ectives</default>	Show Implicit Measures		Default Table Field Set Behavior Reporting Properties	Synonyms Language	
[Date]	•	fx		Reporting Properties	Lunguage	
	Sales 🐕 🔽	10 m	Product	Revenue		Add Column
1 10/7/1 1		Amazon.com	ILP Item		7448	
2 10/26/		Costco	XOU Item		4908	
	Sheliadawn	Home Depot	JHF Item		4427	
4 9/24/1	lon	Whole Foods	XOU Item		4344	
5 1/14/1 5	Stev <mark>en</mark>	McLendon's	MFB Item		5826	
	-1 1. 1			C		
				Sum of Revenue: 10834434 Count of Revenue: 1999	۲ ۲	
				Count of Revenue: 1999		
	\leq				Implicit	Measures
L Measur	re Grid 📃					

- 2) In general, it is okay to use Implicit Measures when you have a small data set (about 50,000 rows) and the built-in calculations in a Standard PivotTable are sufficient.
- 3) Compare and contrast Implicit and Explicit Measures picture is on next page.

Implicit DAX Measures

Drag and drop field into Data Model PivotTable/Power BI Values Area and the

Data Model makes a Read Only Measure for you

MUCH Less Control when you use Implicit

- 1) Number Field will use SUM Function
- 2) Text Field uses COUNT
- 3) You can't add Number Formatting to the Measure
- 4) You can't Reuse Formula
- 5) You can't change the name of the Measure
- 6) If you right-click an Implicit Measure in the Values Area of the PivotTable you can

change the "Summarize Values By", but then it adds yet another Implicit Measure

Explicit DAX Measure

You create the formula, choose the name, add the Number Formatting and use it over and over. You create formula in Measured Grid, or Measure dialog box

MUCH More Control when you use Explicit

- 1) You choose what functions go into your Measure
- 2) You can name the Measure
- 3) You can apply Number Formatting that will follow the Measure around when it is reused.
- 4) You can use the formula over and over.
- 5) You won't have a bunch of extra Measures created by right-click, "Summarize Values By".

7. <u>Power Query Merge feature</u>

1) Power Query Has Six Types of Merges / Joins. This picture summarizes pictorially the six types of merges / joins in Power Query:



- 2) What is a Merge / Join?
 - i. What does a Merge accomplish in Power Query?
 - The Left Outer Merge is a substitute for VLOOKUP. Merge can help us to add a column to a table by pulling matching values from a second table into the first table.
 - ii. Merge / Join Terminology:
 - Merge is the word that we use in Power Query.
 - Join is the word that is used in the SQL (Structured Query Language) and in other database languages.
 - Merge and Join will be synonyms for us.
- 3) Power Query Merges are similar to using VLOOKUP or Relationships.
 - i. VLOOKUP in Excel requires that you have two related columns if you want to lookup a value. We will see how to do this Merge using a Left Outer Merge in Power Query.
 - Relationships in the Excel Power Pivot or Power BI Desktop Data Model require that you have two Related Columns. Relationships in Data Models allow us to accomplish many tasks, one of which is like a Left Outer Merge.
- 4) Requirements for a Merge:
 - i. To Merge one or more queries, you must have the data imported into Power Query as a query.
 - ii. The Merge Feature is for Table Objects.
 - iii. Merges require Related Columns in one or more Table Objects.

Page 17 of 22

8. Data Types in Power Query.

1) Unlike Excel, we must properly Defined each Field with a Data Type. If we do not define the correct Data Type, for example a dollar amount as Currency, then some of the calculations in Power Query, Excel, Power Pivot and Power BI Desktop will not work correctly.

Data Types in Power Query	Short Definition
Decimal Number	Max 15 digits
Fixed Decimal Number	Max 4 decimals to right of decimal
Whole Number	No digits to right of decimal
Date/Time	Date and Time together
Date	Just Date
Time	Just Time
Date/Time/Timezone	Same as Date and Time
Duration	Length of Time
Text	Text - max length 268,435,456 Unicode characters
Ture/False	Ture/False

2) Here is a list of the Data Types in Power Query:

Data Types in Power Query - with Long Definition

Decimal Number – Represents a 64 bit (eight-byte) floating point number. It's the most common number type and corresponds to numbers as you usually think of them. Although designed to handle numbers with fractional values, it also handles whole numbers. The Decimal Number type can handle negative values from -1.79E +308 through -2.23E -308, 0, and positive values from 2.23E - 308 through 1.79E + 308. For example, numbers like 34, 34.01, and 34.000367063 are valid decimal numbers. The largest value that can be represented in a Decimal Number type is 15 digits long. The decimal separator can occur anywhere in the number. The Decimal Number type corresponds to how Excel stores its numbers.

Fixed Decimal Number – Has a fixed location for the decimal separator. The decimal separator always has four digits to its right and allows for 19 digits of significance. The largest value it can represent is 922,337,203,685,477.5807 (positive or negative). The Fixed Decimal Number type is useful in cases where rounding might introduce errors. When you work with many numbers that have small fractional values, they can sometimes accumulate and force a number to be slightly off. Since the values past the four digits to the right of decimal separator are truncated, the Fixed Decimal type can help you avoid these kinds of errors. If you're familiar with SQL Server, this data type corresponds to SQL Server's Decimal (19,4), or the Currency Data type in Power Pivot.

Whole Number – Represents a 64 bit (eight-byte) integer value. Because it's an integer, it has no digits to the right of the decimal place. It allows for 19 digits; positive or negative whole numbers between -9,223,372,036,854,775,808 (-2^63) and 9,223,372,036,854,775,807 (2^63-1). It can represent the largest possible number of the various numeric data types. As with the

Fixed Decimal type, the Whole Number type can be useful in cases where you need to control rounding.

Date/Time – Represents both a date and time value. Underneath the covers, the Date/Time value is stored as a Decimal Number Type. So you can actually convert between the two. The time portion of a date is stored as a fraction to whole multiples of 1/300 seconds (3.33 ms). Dates between years 1900 and 9999 are supported.

Date – Represents just a Date (no time portion). When converted into the model, a Date is the same as a Date/Time value with zero for the fractional value.

Time – Represents just Time (no Date portion). When converted into the model, a Time value is the same as a Date/Time value with no digits to the left of the decimal place.

Date/Time/Timezone – Represents a UTC Date/Time. Currently, it's converted into Date/Time when loaded into the model.

Duration – Represents a length of time. It's converted into a Decimal Number Type when loaded into the model. As a Decimal Number type it can be added or subtracted from a Date/Time field with correct results. As a Decimal Number type, you can easily use it in visualizations that show magnitude.

Text - A Unicode character data string. Can be strings, numbers, or dates represented in a text format. Maximum string length is 268,435,456 Unicode characters (256 mega characters) or 536,870,912 bytes.

True/False – A Boolean value of either a True or False.

9. Overview of Three Examples in Video

Problem: We have Multiple Tables that we need to use to make our PivotTable Reports **Answer:** 3 Solutions:

- 1) VLOOKUP Spreadsheet Function
- 2) Power Query Merge feature
- 3) Power Pivot (& Power BI) Relationship feature & Implicit Measure feature

VLOOKUP Spreadsheet Function

Use when you have about 50,000 rows or less of data Use when Tables are already in the Excel Spreadsheet Use when you want 1 table in the PivotTable Field List

Power Query Merge feature

Use when Tables are coming from an External data Source

Use when you want 1 table in the PivotTable Field List

* Power Query can also import multiple tables without using the Merge feature

Power Pivot Relationship feature & Implicit Measure feature

Use when you have more than about 50,000 records

Use when you want to have multiple tables in the PivotTable Field List

* Power Pivot & Power BI can be used on Big Data (millions of rows) with Explicit Measures





10.VLOOKUP Video Example

Business Situation: Boomerang Incorporated sells boomerang products on four different web sites: Gel Boomerangs, Colorado Boomerangs, Amazon, E-Bay.

Goal: Create a Total Revenue Report by Product Category and Web Site and a similar Report that shows the % Revenue each Total Revenue Number represents of the Grand Total.

Problem: They have three tables with the raw data we need to create the report.

Solution: VLOOKUP. The table is less than about 50,000 rows and we have the three tables in an Excel Sheet.

We will add extra columns to our Fact Table using VLOOKUP, and then we will make a PivotTable Report from 1 table in the PivotTable Field List.

Start Tables:

Fact Table = fTransactionsV			Dimension	Table = dProdu	ctV		Dimension Table = dWebSiteV		
Date 💌 Web Site ID	Product ID	Units 💌	Product II	* Product	Flight Range (m) 💌 Product Category 💌 Price	-	Cost 💌	Web Site ID	Web Site Name
12/8/19 cb69	\$40D15	1	Q20FRE	Quad	20 Freestyle	43.95	23.5	gb43	Gel Boomerangs
12/18/20 am11	M50DIS	5	A22BEG	Aspen	22 Beginner	26.95	16.33	cb69	Colorado Boomeran
12/9/19 gb43	A22BEG	3	\$40D15	Sunset	40 Distance	29.95	12.64	am11	Amazon
4/10/20 gb43	A228EG	1	C20FRE	Carlota	20 Freestyle	27.95	16.06	eb92	E-Bay
2/25/19 cb69	A22BEG	1	B25BEG	Bellen	25 Beginner	24.95	15.31		
12/23/20 gb43	A228EG	1	Y28BEG	Yanaki	28 Beginner	22.95	13.19		
12/1/20 cb69	Q20FRE	1	T20FRE	TriFly	20 Freestyle	19.95	8.53		
12/10/19 gb43	Q20FRE	2	F45DIS	Flattop	45 Distance	45.95	24.06		
12/23/19 am11	Y28BEG	2	M50DIS	Majestic Be	aut 50 Distance	49.95	26.29		
5/8/20 pb43	B25BEG	2							

Excel Spreadsheet Formulas to create Data Model:

A	В	c	D	E	F	G	н		1	К	L	M	N	0		P	Q R	S
	Data Mo	deling with	Excel Spre	adsheet	Form	ulas:												
	Formula	in cell F12: =	VLOOKUP(@Produ	uct ID]],dProdue	ctV,5,0)											
	Formula	in cell G12: =	[@Price]*	[@Units]														
	Formula	in cell H12: =	VLOOKUP	([@[Prod	uct ID]],dProdu	ictV,4,0)											
	Formula	in cell 112: =	VLOOKUP(@[Web 5	Site ID]],dWebS	SiteV,2,0)											
	Fact Table	fTransactionsV								Dimension Table =	dProductV						Dimension T	able = dWebSiteV
	Date 💌	Web Site ID 💌	Product ID	• Units •	Price	Revenue	Product Category	Web Site Name 💌		Product ID	· Product	 Flight Range (m) 	Product Category	· Price	✓ Cost	*	Web Site ID	• Web Site Name •
	12/8/19	eb69	S40DIS	1	29.9	5 29	.95 Distance	Colorado Boomerangs		Q20FRE	Quad		20 Freestyle		43.95	23.5	gb43	Gel Boomerangs
	12/18/20) am11	M50DIS	5	49.9	5 249	.75 Distance	Amazon		A22BEG	Aspen		22 Beginner		26.95	16.33	cb69	Colorado Boomerang
	12/9/19	gb43	A22BEG	3	26.9	5 80	0.85 Beginner	Gel Boomerangs		S40DIS	Sunset		40 Distance		29.95	12.64	am11	Amazon
	4/10/20	gb43	A22BEG	1	26.9	5 26	95 Beginner	Gel Boomerangs		C20FRE	Carlota		20 Freestyle		27.95	16.06	eb92	E-Bay
	2/25/19	9 cb69	A22BEG	1	26.9	5 26	6.95 Beginner	Colorado Boomerangs		B25BEG	Bellen		25 Beginner		24.95	15.31		
	12/23/20	gb43	A22BEG	1	26.9	5 26	5.95 Beginner	Gel Boomerangs		Y28BEG	Yanaki		28 Beginner		22.95	13.19		
	12/1/20	cb69	Q20FRE	1	43.9	5 43	.95 Freestyle	Colorado Boomerangs		T20FRE	TriFly		20 Freestyle		19.95	8.53		
	12/10/19	gb43	Q20FRE	2	43.9	5 8	7.9 Freestyle	Gel Boomerangs		F45DIS	Flattop		45 Distance		45.95	24.06		
	12/23/19	am11	Y288EG	2	22.9	5 4	5.9 Beginner	Amazon		MSODIS	Majestic Beaut		50 Distance		49.95	26.29		
	5/8/20	gb43	8258EG	2	24.9	5 4	9.9 Beginner	Gel Boomerangs										

Sum of Revenue Product Category	Web Site Name Amazon	Colorado Boomerangs	E-Bay	Gel Boomerangs	Grand Total
Beginner	174,570	251,028	84,294	341,246	851,138
Distance	152,268	215,191	76,386	303,163	747,008
Freestyle	251,203	419,389	135,485	531,404	1,337,480
Grand Total	578,041	885,607	296,164	1,175,814	2,935,626
Sum of Revenue	Web Site Name 🔻				
Product Category	▼ Amazon	Colorado Boomerangs	E-Bay	Gel Boomerangs	Grand Total
Beginner	5.95%	8.55%	2.87%	11.62%	28.99%
Distance	5.19%	7.33%	2.60%	10.33%	25.45%
Freestyle	8.56%	14.29%	4.62%	18.10%	45.56%
Grand Total	19.69%	30.17%	10.09%	40.05%	100.00%

11. Power Query Video Example

Business Situation: Boomerang Incorporated sells boomerang products on four different web sites: Gel Boomerangs, Colorado Boomerangs, Amazon, E-Bay. Goal: Create a Total Revenue Report by Product Category and Web Site and a similar Report that shows the % Revenue each Total Revenue Number represents of the Grand Total. Problem: They have three tables with the raw data we need to create the report. Two tables are in an Access Database and the third is below in the Excel Sheet. Solution: Power Query. We will use Power Query to import tables from Access and Excel into the Power Query Editor.

Then we will use the Merge feature and the Add Column feature. We will then load our data to the PivotTable Cache, and then we will make a PivotTable Report.



12. Power Pivot Relationships feature & Implicit Measure feature

Business Situation: Boomerang Incorporated sells boomerang products on four different web sites: Gel Boomerangs, Colorado Boomerangs, Amazon, E-Bay.

Goal: Create a Total Units Report by Product Category and Web Site and a similar Report that shows the % Revenue each Total Revenue Number represents of the Grand Total.

Problem: They have three tables with the raw data we need to create the report. The tables have about 50,000 records and are in the Excel Sheet and so we can choose to use VLOOKUP or Relationships. Solution: Power Pivot. We will use the Data Ribbon Tab, Data Tools, Relationship button and then an Implicit Measure in the Data Model PivotTable.

This allows us to create a One-To-Many Relationship between the tables, import them into the Power Pivot data Model, and then we will have three different table in our PivotTable Field List.

Start Tables:

act Table = fTransactionsR			Dimension Table = dProductR					Dimension Table = dWebSiteR		
Date 💌 Web Site ID 💌	Product ID 💌 Units	-	Product II	Product 💌 Flight Range (m)	Product Category Price	-	Cost 💌	Web Site ID	Web Site Name	
12/8/19 cb69	S40DIS	1	Q20FRE	Quad	20 Freestyle	43.95	23.5	gb43	Gel Boomerangs	
12/18/20 am11	M50DIS	5	A22BEG	Aspen	22 Beginner	26.95	16.33	cb69	Colorado Boomerang	
12/9/19 gb43	A228EG	3	S40DIS	Sunset	40 Distance	29.95	12.64	am11	Amazon	
4/10/20 gb43	A22BEG	1	C20FRE	Carlota	20 Freestyle	27.95	16.06	eb92	E-Bay	
2/25/19 cb69	A228EG	1	B25BEG	Bellen	25 Beginner	24.95	15.31			
12/23/20 gb43	A22BEG	1	Y28BEG	Yanaki	28 Beginner	22.95	13.19			
12/1/20 cb69	Q20FRE	1	T20FRE	TriFly	20 Freestyle	19.95	8.53			
12/10/19 gb43	Q20FRE	2	F45DIS	Flattop	45 Distance	45.95	24.06			
12/23/19 am11	Y28BEG	2	M50DIS	Majestic Beaut	50 Distance	49.95	26.29			

Create Relationships

Active

Text to	Flash Fill	Remove Duplicates	Data Validation ~		E E E Relationships	Manage Data Model
		3.000	Data T	ools		
Aanage Re	lationsh	iips				
Status 🔺	Table	1			Related Loo	kup Table
Active	C. C	actionsR (We	Cite 101		BALLAN CHART A	Web Site ID)



.

0

PivotTable Field List and Final Reports:

fTransactionsR (Product ID)

Sum of Units Product Category	Web Site Name - Amazon	Colorado Boomerangs	E-Bay	Gel Boomerangs	Grand Total	PivotTable Fields
Beginner	6,899	9,941	3,345	13,501	33,686	Choose fields to add to report
Distance	3,912	5,546	1,968	7,778	19,204	Search
Freestyle	7,950	13,265	4,307	16,791	42,313	J Ⅲ dProductR
Grand Total	18,761	28,752	9,620	38,070	95,203	Product ID
	Web Site Name - Amazon	Colorado Boomerangs	0.09/06/201	Gel Boomerangs	Report for the second second	Product Category Price Cost
Beginner	7.25%	10.44%	3.51%	14.18%	35.38%	
Distance	4.11%	5.83%	2.07%	8.17%	20.17%	J dWebSiteR
Freestyle	8.35%	13.93%	4.52%	17.64%	44.45%	🖉 Web Site Name
Grand Total	19.71%	30.20%	10.10%	39.99%	100.00%	J International International
						Web Site ID Product ID Units

dProductR (Product ID)