Data Management

Excel Tips and Tricks to Summarize Data

Danny Limoges, PharmD
Vaccine & 340B Coordinator
St. Bernard Hospital and Health Care Center

Please download and open the Excel file:

https://bit.ly/2VLMdUk

(case sensitive)

Data Management

The presenter has no conflicts of interest to declare.

Please download and open the Excel file:

https://bit.ly/2VLMdUk

(case sensitive)

Learning Objectives

- Demonstrate ways to format and arrange data using Microsoft Excel©
- Create a pivot table to summarize data.
- Use VLOOKUP to retrieve information from a table.

Please download and open the Excel file:

https://bit.ly/2VLMdUk

(case sensitive)

What is Excel?

More than a spreadsheet…

- Conditional Formatting
- Filtering / Sorting
- Relational Data
- External Data Sources

	A	B	C
1	ID	Date Written	Last Fill Static
2	A8211	8/26/2020	10/2/2020
3	A4801	9/14/2020	12/1/2020
4	A5633	2/16/2021	4/22/2021
5	A5627	2/6/2021	3/1/2021
6	A1199	6/29/2021	9/16/2021

F2 =VLOOKUP(A2,Fills,3,FALSE)

	A	B	C	D	E	F
1	PT ID	DOB	ObservationID	A1c	Primary Insurance	Last Insulin Fill
2	A5633	3/30/1996	6255118	5.8	Humana PPO	4/22/2021
3	A8211	6/8/1991	8952402	6.2	BCBS Medicaid	10/2/2020
4	A4185	9/10/1966	1254798	6.1	BCBS Medicaid	6/16/2021
5	A4801	2/4/1952	9612090	8.4	Medicaid HFS	12/1/2020
6	A5627	5/10/1988	1114137	5.1	Self Pay	3/1/2021
7	A1199	9/22/1968	1991924	7.2	BCBS PPO	9/16/2021
8	A3311	6/17/1965	2341813	5.9	Humana PPO	2/6/2021
9	A3014	10/5/1990	7890682	8.0	BCBS Medicaid	10/7/2021
10	A3352	4/14/1960	3003214	6.2	Medicaid HFS	1/27/2021
11	A5285	8/24/1955	5378398	7.9	Self Pay	7/24/2021
12	A1866	8/8/1957	3861641	5.8	BCBS Medicaid	4/16/2021
13	A7168	10/17/1983	1388769	6.1	BCBS Medicaid	10/8/2021
14	A1198	8/21/1973	3003251	7.3	Medicaid HFS	6/7/2021
15	A9342	8/14/1984	1645981	6.3	Self Pay	3/8/2021
16	A3186	4/6/1968	3638689	5.4	BCBS PPO	2/22/2021
17	A9195	6/13/1977	4922070	6.0	Humana PPO	12/20/2020

What is Excel?

More than a spreadsheet…

- Computational Engine
- Machine Learning AI
- Pivot Tables
- Graphing

Insurance	Insulin Fills
BCBS Medicaid	5
BCBS PPO	2
Humana PPO	3
Medicaid HFS	3
Self Pay	3
Grand Total	**16**

I25

	A	B	C	D	E	F	G
1	PT ID	DOB	ObservationID	A1c	Primary Insurance	Last Insulin Fill	Month of Last Fill
2	A5633	3/30/1996	6255118	5.8	Humana PPO	4/22/2021	4
3	A8211	6/8/1991	8952402	6.2	BCBS Medicaid	10/2/2020	10
4	A4185	9/10/1966	1254798	6.1	BCBS Medicaid	6/16/2021	6
5	A4801	2/4/1952	9612090	8.4	Medicaid HFS	12/1/2020	12
6	A5627	5/10/1988	1114137	5.1	Self Pay	3/1/2021	3
7	A1199	9/22/1968	1991924	7.2	BCBS PPO	9/16/2021	9
8	A3311	6/17/1965	2341813	5.9	Humana PPO	2/6/2021	2
9	A3014	10/5/1990	7890682	8.0	BCBS Medicaid	10/7/2021	10
10	A3352	4/14/1960	3003214	6.2	Medicaid HFS	1/27/2021	1
11	A5285	8/24/1955	5378398	7.9	Self Pay	7/24/2021	7
12	A1866	8/8/1957	3861641	5.8	BCBS Medicaid	4/16/2021	4
13	A7168	10/17/1983	1388769	6.1	BCBS Medicaid	10/8/2021	10
14	A1198	8/21/1973	3003251	7.3	Medicaid HFS	6/7/2021	6
15	A9342	8/14/1984	1645981	6.3	Self Pay	3/8/2021	3
16	A3186	4/6/1968	3638689	5.4	BCBS PPO	2/22/2021	2
17	A9195	6/13/1977	4922070	6.0	Humana PPO	12/20/2020	12

Formatting and Arranging Data

Introduction

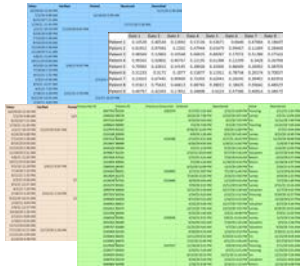
Challenges of arranging data in Excel:

- User must define and maintain structure
- Data types are ambiguous, not enforced
- Slow with large data sets
- Cell contents are hidden by default

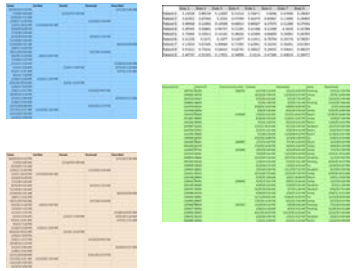
Formatting and Arranging Data

Introduction

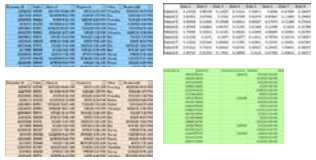
Typical Workflow



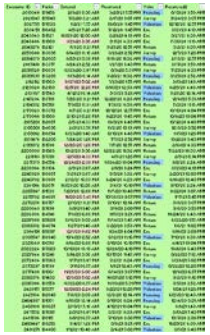
Assemble



Arrange



Filter,
Format Data



Relate,
Compute

Visual
Formatting

Formatting and Arranging Data

Microsoft Guidelines for Organizing Data:

- Put similar items in the same column
- Keep ranges of data separate
- Position critical data above or below the range
- Avoid blank rows and columns
- Display all rows and columns in a range

	A	B	C	D	E	F
1	Excel Data Formatting and Arranging					
2	Extracted:	9/16/2021				
3						
4	Patient ID	DOB	ObservationID	A1c	Primary Insurance	
5	A7041	6/26/1954	6489896	6.9	Humana PPO	
6	A7517	9/5/1967	9154796	6.9	BCBS Medicaid	
7	A3495	11/11/1989	7996850	7.8	BCBS Medicaid	
8	A9403	10/23/1969	9112538	8.2	Medicaid HFS	
9	A5212	2/17/1971	8560585	5.7	Self Pay	
10	A9854	6/14/1964	4125455	6.6	BCBS PPO	
11	A2277	7/10/1957	3702324	5.6	Humana PPO	
12	A1520	3/16/1993	2943682	6.2	BCBS Medicaid	
13	A1212	2/12/1998	8996416	8.4	Medicaid HFS	
14	A8272	4/23/1972	6785043	5.2	Self Pay	
15	A8203	2/8/1961	1894787	8.2	BCBS Medicaid	
16	A4370	10/26/1950	6191482	5.7	BCBS Medicaid	
17	A7303	7/18/1960	8117153	5.2	Medicaid HFS	

Guidelines for organizing and formatting data on a worksheet. Available at: https://support.microsoft.com/en-us/office/guidelines-for-organizing-and-formatting-data-on-a-worksheet-90895cad-6c85-4e02-90d3-8798660166e3. Accessed July 16, 2021.

Formatting and Arranging Data

Introduction

Best Practices for Arrangement:

- One observation / fact per row
- Define what each row represents
  - One set of demographics / settings per row?
  - Log of observations?
- Label all columns meaningfully
- Format columns as correct data types

Formatting and Arranging Data

Best Practices

One observation / fact per row

Patient MRN	Time 1	Glucose 1	Potassium 1	Time 2	Glucose 2	Potassium 2
A9084	5:00 AM	166	3.7	1:00 PM	164	4.1
A2573	5:30 AM	231	4.9	1:30 PM	218	5.5

Patient MRN	Time	Glucose	Potassium
A9084	5:00 AM	166	3.7
A2573	5:30 AM	231	4.9
A9084	1:00 PM	164	4.1
A2573	1:30 PM	218	5.5

Formatting and Arranging Data

Best Practices

Define what each row represents

Order ID	Transaction Sent	Acq	Units
1	8/12/21 9:30 AM	$ 5.55	90
2	8/12/21 2:36 PM	$ 204.63	15
3	8/12/21 4:44 PM	$ 151.96	30

One row per order

Order ID	Transaction Sent	Acq	Units	Transaction Type
1	8/12/21 9:30 AM	$ 5.55	90	Sent
1	8/12/21 9:39 AM	$ (5.55)	90	Reversed
1	8/12/21 9:40 AM	$ 6.58	90	Sent
2	8/12/21 2:36 PM	$ 204.63	15	Sent
2	8/13/21 10:42 PM	$ (204.63)	15	Reversed

One row ***per transaction*** per order

Formatting and Arranging Data

Best Practices

Label all columns meaningfully

Guidelines for a good column name…

1. Accurate
2. Differentiates it from other columns
3. “Reads” at a glance

Formatting and Arranging Data

Best Practices

Format columns as correct data types

	Input	Correct Type
Date	"5/6"	5/6/2021
Number	1234	1234
Different Decimals	44.567	44.57
	97.1	97.10
	71	71.00
Finances	10	$ 10.00
	5.25	$ 5.25
	7.75	$ 7.75

Formatting and Arranging Data

Practice

Vaccines Administered					
Week	Monday	Tuesday	Wednesday	Thursday	Friday
7/18/2021	5	5	4	5	4
7/25/2021	5	5	4	5	5
8/1/2021	6	3	6	5	5

How can this table improve?

A: One observation per row

Vaccines Administered	
Day	Vaccines
7/18/2021	5
7/19/2021	5
7/20/2021	4
7/21/2021	5
7/22/2021	4
7/25/2021	5
7/26/2021	5
7/27/2021	4
7/28/2021	5
7/29/2021	5

Formatting and Arranging Data

Practice

How can this table improve? One observation per row

Vaccines Administered	
Day	Vaccines
7/18/2021	10 Moderna
7/19/2021	9 Moderna
7/20/2021	10 Moderna, 6 Pfizer
7/21/2021	9 Moderna
7/22/2021	10 Moderna
7/25/2021	8 Moderna, 5 Pfizer
7/26/2021	10 Moderna
7/27/2021	10 Moderna
7/28/2021	9 Moderna
7/29/2021	10 Moderna

Vaccines Administered		
Day	Vaccine	Given
7/18/2021	Moderna	10
7/19/2021	Moderna	9
7/20/2021	Moderna	10
7/20/2021	Pfizer	6
7/21/2021	Moderna	9
7/22/2021	Moderna	10
7/25/2021	Moderna	8
7/25/2021	Pfizer	5
7/26/2021	Moderna	10
7/27/2021	Moderna	10
7/28/2021	Moderna	9
7/29/2021	Moderna	10

Formatting and Arranging Data

Practice

How is the "one observation per row" structure helpful?

- Can create a Table
- Can create Pivot Tables
- Demo…
  - Go to Tab "Practice 1 – Table"

Vaccines Administered		
Day	**Vaccine**	**Given**
7/18/2021	Moderna (use up this lot)	10
7/19/2021	Moderna (use up this lot)	9
7/20/2021	Moderna	10
7/20/2021	Pfizer (use up this lot)	6
7/21/2021	Moderna	9
7/22/2021	Moderna	10
7/25/2021	Moderna	8
7/25/2021	Pfizer	5
7/26/2021	Moderna	10
7/27/2021	Moderna	10
7/28/2021	Moderna	9
7/29/2021	Moderna	10

Formatting and Arranging Data

Formatting Data

Types of Formatting

- Data Type
  - Number / Text / Date
- Visual
  - Presentation of Number / Text / Date
  - Conditional Formatting

Formatting and Arranging Data

Common Formats

General		
General Cell with Formula	1234	
General Cell with Text	Text	
Text		
Plain Text:	Text	
Number formatted as text:	1234	
Formula formatted as text:	=1000 + 234	

Formatting and Arranging Data

Common Formats

Number		
Number	1234.00	-1234.00
Number with Commas, Red Negative*	1,234.00	1,234.00
Accounting		
Dollar Value	$ 1,234.00	
Dollar Value (4 decimals)	$ 1,234.0000	

* = Must use "Format Cells" menu

Formatting and Arranging Data

Common Formats

Date	
Short Date	8/10/2021
Long Date	Tuesday, August 10, 2021
Month/Day*	8/10
Month/Day (3-letter month)*	Aug-21
Only the day*	10
The real value (days since 1/0/1900):	44418

* = Must use "Format Cells" menu

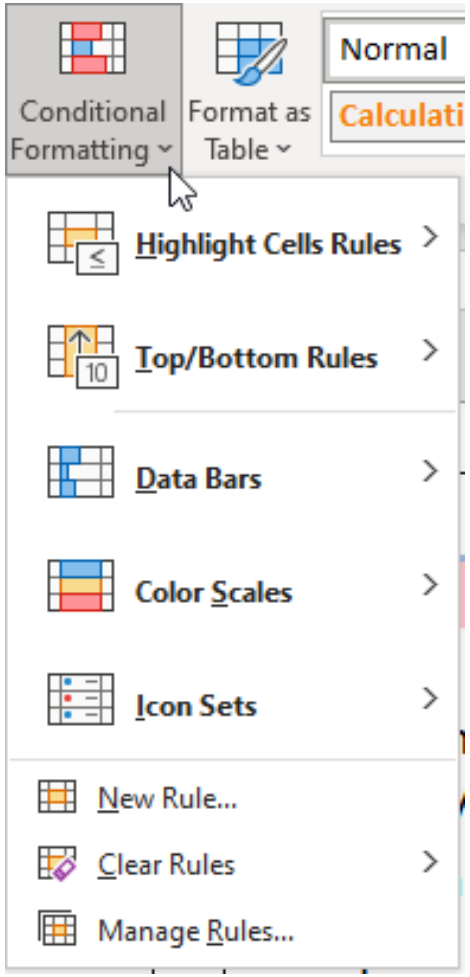
Live demo of formatting…

Formatting and Arranging Data

Conditional Formatting

To apply Conditional Formatting:

1. Select range of data to format
2. Click “Conditional Formatting”, select type
3. To customize, click Conditional Formatting again
   A. Click “Manage Rules…” and change settings



Formatting and Arranging Data

Conditional Formatting

Uses for Conditional Formatting

- Make large datasets comprehensible
- Create a color-based filter
- Highlight noteworthy values and alerts
- Find duplicate values

Formatting and Arranging Data

Conditional Formatting

Text

Equals "Ertapenem"
cefazolin
levofloxacin
ertapenem
imipenem/cilastatin
cephalexin
Cellcept
meropenem

Contains "penem"
cefazolin
levofloxacin
ertapenem
imipenem/cilastatin
cephalexin
Cellcept
meropenem

Contains "ce*in"
cefazolin
levofloxacin
ertapenem
imipenem/cilastatin
cephalexin
Cellcept
meropenem

** = Zero to many wildcards*

Contains "ce?in"
cefazolin
levofloxacin
ertapenem
imipenem/cilastatin
cephalexin
Cellcept
meropenem

? = One wildcard

Formatting and Arranging Data

Conditional Formatting

Numbers

Data Bars	Color Scales	Icon Sets	Duplicate values
7	7	▲ 100%	1
6	6	▲ 75%	1
5	5	▲ 50%	2
4	4	▲ 25%	3
3	3	▬ 0%	4
2	2	▼ -25%	5
1	1	▼ -50%	6

Pivot Tables

Uses for Pivot Tables

- Summarize data
- Reformat data
- Find outliers
- Target subsets of data

Pivot Tables

Anatomy of a Pivot Table

Filter

Columns

Rows

Values

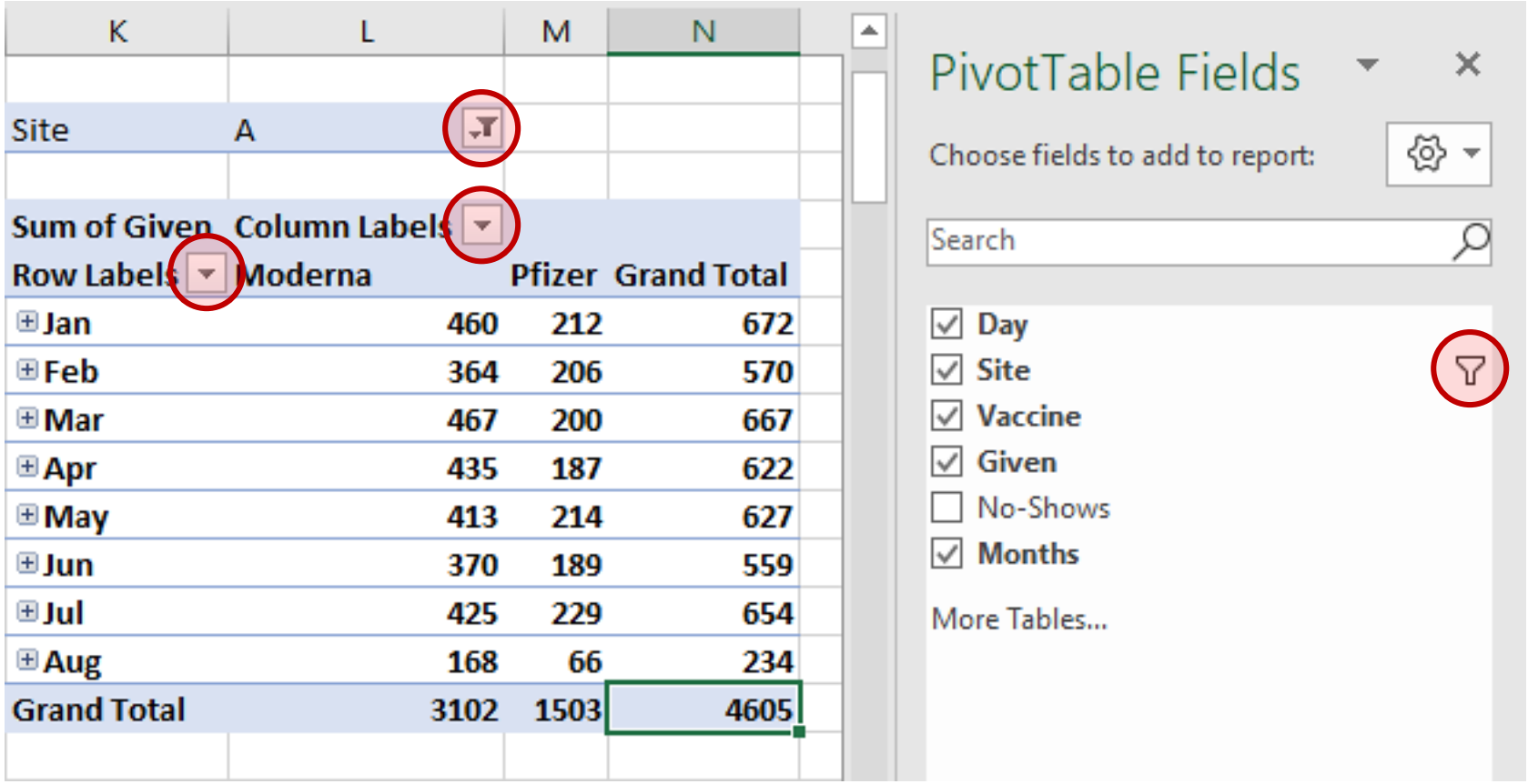
Site	A

Sum of Given	Column Labels		
Row Labels	Moderna	Pfizer	Grand Total
Jan	460	212	672
Feb	364	206	570
Mar	467	200	667
Apr	435	187	622
May	413	214	627
Jun	370	189	559
Jul	425	229	654
Aug	168	66	234
Grand Total	3102	1503	4605

Pivot Tables

Note that filters may be in multiple locations.

(O365 Version Pictured)



Pivot Tables

Live demo:

- Basic Setup
- “Rule of Two” for beginners
- Sum vs Count
- Totals
- Changing Layout

Pivot Tables

Practice

Open Excel tab “Practice 2 – Pivot Table”

Solve: How many Moderna vaccines were given in July?

Discussion and live solution in 3 minutes.

VLOOKUP

Used to extract ***Relational Data***

Units Ordered	Fruit Ordered	Unit Price	Order Price		Fruit	Price Per Unit
1	Oranges	VLOOKUP			Apples	$5
3	Oranges				Oranges	$10
2	Bananas				Bananas	$20
3	Apples					
2	Oranges					
2	Bananas					
3	Bananas					
3	Bananas					
1	Bananas					

VLOOKUP

Used to extract ***Relational Data***

Units Ordered	Fruit Ordered	Unit Price	Order Price		Fruit	Price Per Unit
1	Oranges	$ 5	$ 5		Apples	$5
3	Oranges	$ 5	$ 15		Oranges	$10
2	Bananas	$ 20	$ 40		Bananas	$20
3	Apples					
2	Oranges					
2	Bananas					
3	Bananas					
3	Bananas					
1	Bananas					

VLOOKUP

Equation

	B				F	G
Units Ordered	Fruit Ordered	Unit Price	Order Price		Fruit	Price Per Unit
1	Oranges	=VLOOKUP(B2,F:G,2,FALSE			Apples	$5
3	Oranges	VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])			Oranges	$10
2	Bananas				Bananas	$20
3	Apples					

lookup_value: cell holding the ***key*** value

table_array: range of data with ***key*** on left side

col_index_num: column containing data to pull. 1 = First Column.

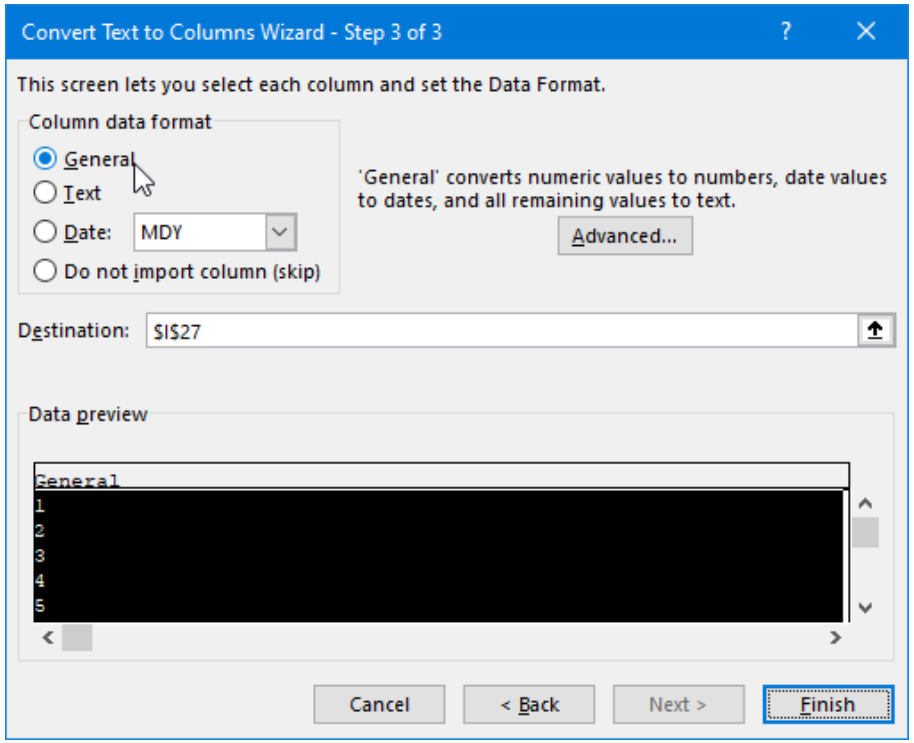
[range_lookup]: optional method to use to search. Use FALSE.

VLOOKUP

Common Pitfall: Different data types for keys

Numbers as Text	VLOOKUP Name
1	#N/A
2	#N/A
3	#N/A
4	#N/A
5	#N/A
6	#N/A
7	#N/A

Number	Name
1	One
2	Two
3	Three
4	Four
5	Five
6	Six
7	Seven



Solution: Convert "Numbers as Text" to general data using **Text to Columns**.

VLOOKUP

Practice

Open Excel tab “Practice 3 – VLOOKUP”

- Work on the three problems.
- Collaborate with neighbors.
- Group Discussion in 5 minutes.

Assessment Question #1

A user is copying and pasting a new and different type of data into a spreadsheet in Excel. Which of the following describes the best location to place the new data?

A. In a new tab or more than one blank row and column away from existing data.
B. At the end of the existing data, with extra columns added to the original table if needed.
C. Filter the existing data so that the keys match that of the new data then paste immediately next to the filtered table.
D. Rather than pasting the data, the user should link between the documents

Assessment Question #2

The pharmacy is investigating the cost and effectiveness of a new initiative to counsel patients on a medication before discharge. Assuming that each attempt to find the patient in-room and counsel them is tracked on an Excel sheet, what is the best way to record this data?

A. Each patient should have one row of data and each attempt will be a new set of columns.
B. Each attempt should be its own row, with the patient, time recorded and results for each entry.
C. Each patient should have one row that shows the time of the latest attempt and its results.
D. Record the successful attempts in new rows in one table and unsuccessful attempts in another table with the same column names

Assessment Question #3

If a user wants to summarize only one group of values out of the larger dataset using a pivot table, what is the best part of the pivot table to adjust?

A. Values
B. Rows
C. Columns
D. Filter

Assessment Question #4

A VLOOKUP equation is not working - it is returning a value of #N/A. The user has verified that the equation is pointing to the correct “lookup_value” and “table_array” and that the key being looked up is in both. What is the most likely reason why the VLOOKUP is not working?

A. The key may be formatted as a text value in one area and a number value in the other.
B. The available memory is low therefore VLOOKUP will not calculate correctly at this time.
C. VLOOKUP is not the correct formula to use in this situation.
D. VLOOKUP can only search a limited number of cells in a table.

Assessment Question #5

In the following VLOOKUP equation, which column holds the value that will be returned by the equation?

VLOOKUP(A1, B:C, 2, FALSE)

a. A

b. B

c. C

d. D

References

- McFedries P. *Microsoft Excel 2019 Formulas and Functions*. Microsoft Press; 2019.
- Harvey G. *Excel 2019 All-in-One For Dummies*. John Wiley & Sons; 2018.
- Guidelines for organizing and formatting data on a worksheet. Available at: https://support.microsoft.com/en-us/office/guidelines-for-organizing-and-formatting-data-on-a-worksheet-90895cad-6c85-4e02-90d3-8798660166e3. Accessed July 16, 2021.

Data Management: Excel Tips and Tricks to Summarize Data

CE Code: X5131

You will need this code to claim credit on CESally.com.

See the directions in the Syllabus.