

Research 201

AAPA
American Academy of PAs

Using Advanced Excel Tools for Managing Your Data

Sean Kolhoff, PhD
Senior Research Analyst
American Academy of PAs

1

AAPA
American Academy of PAs

Excel Functions

2

Navigating the Excel Platform

- This presentation assumes you have a basic understanding of the Excel platform.
- If you are a true Excel novice, I recommend going to YouTube, Coursera, or LinkedIn Learning for intro information.

© American Academy of PAs. All rights reserved. These materials may not be reproduced without the express written permission of AAPA.

3

Excel Tips and Tricks

Using built in tools for quick summary stats

- Functions auto-update when numbers change
- Need correct cell references (be careful when adding/removing columns)

Employee ID	Hours Worked	Hourly Wage	Wages Owed
30399	45	60 \$	2,700.00
30941	44	66 \$	2,904.00
56967	51	56 \$	2,856.00
48298	42	62 \$	2,604.00
47392	44	59 \$	2,596.00
10928	53	56 \$	2,968.00
69382	55	55 \$	3,025.00
78292	46	63 \$	2,898.00
69382	57	65 \$	3,705.00
Total			
Mean			

4

Excel Tips and Tricks

Add conditional formatting to make cells stand out

Employee ID	Hours Worked	Hourly Wage	Wages Owed
30399	45	60 \$	2,700.00
30941	44	66 \$	2,904.00
56967	51	56 \$	2,856.00
48298	42	62 \$	2,604.00
47392	44	59 \$	2,596.00
10928	53	56 \$	2,968.00
69382	55	55 \$	3,025.00
78292	46	63 \$	2,898.00
69382	57	65 \$	3,705.00
Total	416.00	597.00	24749.00
Mean	41.60	59.70	2474.90

5

Manipulating/Rotating Data

The "Transpose" special paste option can let you turn rows into columns (and vice versa!)

Copy -> Right Click -> Paste Special -> Transpose

Be aware if you are copying cells that have formulas/reference other cells

NOTE: you cannot cut and transpose

Employee ID	Hours Worked	Hourly Wage	Bi Weekly Wages Owed
30399	45	60 \$	2,700.00
30941	44	66 \$	2,904.00
56967	51	56 \$	2,856.00

Employee ID	30399	30941	56967
Hours Worked	45	44	51
Hourly Wage	60	66	56
Bi Weekly Wages Owed	2,700.00	2,904.00	2,856.00

6

Relative, Absolute, and Mixed Cell References

Relative Cell Reference (ex. A1)

- Changes if the formula is moved or copied to a new cell
- Useful when: carrying out formula when the values do not matter
 - I.e. Hours Worked (B1) x Hourly Wage (B2) = Compensation (B1 * B2)

Absolute Cell Reference (ex. \$A\$1)

- If you move or copy the formula to a new cell, it will **always** reference the absolute cell reference
- Useful when: carrying out a formula that includes a value that is fixed
 - I.e. Hours Worked x Hourly Wage x Overtime = Compensation

Mixed Cell Reference

- A\$1** relative column, absolute row
 - If moved or copied, the **column changes**, but the row does not
- \$A1** absolute column, relative row
 - If moved or copied, the **row changes**, but the column does not

7

So, what does this look like?

Relative Cell References

A4	B4	C4
A5	B5	C5
A6	B6	C6

SA1 Locks Column
A\$1 Locks Row

Where do we go from here?

8

Excel Tips and Tricks

Using VLOOKUP to lookup a value in a table array and retrieve a matched value

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

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

Exact Match

Employee ID	On Call Hours	Department	Phone	RVUs
30339	Monday 8am-11am	Critical Access	203-555-0001	
30941	Saturday 8m - 7pm	Emergency Dept	203-555-0002	
14238	Tuesday 8am - 12pm	Inpatient Unit (n)	203-555-0003	
22546	Thursday 8am - 11pm	Intensive Care	203-555-0004	
29149	Sunday 9:30pm - 5:30pm	Operating room	203-555-0005	
14821	Wednesday 8am - 11am	Outpatient Unit	203-555-0006	
56987	Saturday 7pm - 10pm	Procedural suite	203-555-0007	
20125	Tuesday 8am-11pm	Emergency Dept	203-555-0008	
58053	Thursday 2am - 3am	Intensive Care	203-555-0009	

Employee ID: 22546
Phone: 203-555-0004
On Call Hours: 203-555-0004

9

Excel Tips and Tricks

Using VLOOKUP approximate match to lookup a value in a table array and retrieve a matched value

- [range_lookup] =TRUE (Data must be sorted ascending)

The screenshot shows an Excel spreadsheet with the following data tables:

Employee ID	Rate	Bonus
30399	2790	
30941	4987	
14238	5743	
22546	5148	
25149	5974	
14821	6646	
56967	5284	
97125	5086	
58053	3141	

Rate	Bonus %
2000	3%
3000	4%
4000	5%
5000	6%
6000	7%
7000	8%

The VLOOKUP formula is shown as: `=VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])` with the note "Approximate Match".

10

VLOOKUP: Pros and Cons

<p>Pros</p> <ul style="list-style-type: none"> • Easy to use • Effective on small data sets 	<p>Cons</p> <ul style="list-style-type: none"> • Easy to make mistakes <ul style="list-style-type: none"> • Requires mixed and absolute references • Adding/removing columns leads to invalid references • Limited functionality and flexibility <ul style="list-style-type: none"> • You can only look up values on the "right" of the reference value • Only works when the data is vertically arranged ("V" lookup) • Requires manual assignment, which can be cumbersome. <ul style="list-style-type: none"> • Manual work is slow, other functions (like INDEX) have more automation
--	---

11

Excel Tips and Tricks

Using INDEX vs. VLOOKUP

The screenshot shows an Excel spreadsheet with the following data table:

Employee ID	Department	Phone	Commission
30399	Critical Access	203-555-1000	
30941	Emergency Dept	203-555-1001	
14238	Department unit	203-555-1002	
22546	Intensive Care	203-555-1003	
25149	Operating room	203-555-1004	
14821	Outpatient Unit	203-555-1005	
56967	Procedural suite	203-555-1006	
97125	Emergency Dept	203-555-1007	
58053	Intensive Care	203-555-1008	

The screenshot also shows two summary tables:

INDEX MATCH	
Employee ID	30399
Department	
Phone	
Commission	

VLOOKUP	
Employee ID	30399
Department	Critical Access
Phone	

12

Excel Tips and Tricks

Pivot tables and charts

Employee ID	Hours Worked	Hourly Wage	Quarterly Wage	Bonus RVUs	Department
303991	37	\$	23,088.00		102 Critical Access
309411	42	\$	36,360.00		77 Emergency Dept/Observation unit
162328	46	\$	31,464.00		106 Inpatient unit (not ICU/CCU)
225441	35	\$	29,820.00		133 Intensive Care
251450	39	\$	27,612.00		96 Operating room
303991	37	\$	26,752.80		102 Critical Access
784851	42	\$	24,192.00		77 Emergency Dept/Observation unit
334959	41	\$	27,652.00		106 Inpatient unit (not ICU/CCU)
443991	56	\$	29,400.00		133 Intensive Care
884121	38	\$	27,360.00		96 Operating room

13

Data Visualization

14

Visualization on Excel: What charts to use

Clustered Column

- 2-4 data groups (series)
- Compared on the same data type
- Compensation differences between men and women

Bar vs. Column Chart

- Can better display multiple data series
- Bar charts allow you to have longer labels

15

Visualization on Excel: What charts to use

Line Charts

- Easier to view data trends over time
- Recommend to use over column charts if order of categories is important

RVUs by Department
Order IS Important

Department	February	March	April	May
Critical Access	4000	4500	5000	5500
Emergency Dept/Observation unit	4500	5000	5500	6000
Intensive Care	5000	5500	6000	6500
Outpatient Unit	5500	6000	6500	7000
Operating room	6000	6500	7000	7500
Inpatient unit (not ICU/CCLU)	6500	7000	7500	8000
Procedural suite	7000	7500	8000	8500

RVUs by Department

Department	Jan	Feb	Mar	Apr	May
Procedural suite	7000	7500	8000	8500	9000
Inpatient unit (not ICU/CCLU)	6500	7000	7500	8000	8500
Operating room	6000	6500	7000	7500	8000
Intensive Care	5500	6000	6500	7000	7500
Outpatient Unit	5000	5500	6000	6500	7000
Emergency Dept/Observation unit	4500	5000	5500	6000	6500
Critical Access	4000	4500	5000	5500	6000

Collection Agency of PA, All rights reserved. These materials may not be reproduced without the express written permission of AACE.

16

Visualization on Excel: What charts to use

Stacked Column (values) vs. Stacked Column (%)

- Both compare the proportion of the series within each category
- Stacked values make the trend in the absolute difference easier to interpret
- Stacked % emphasizes proportion of categories

New Vs Established Patients

Month	New Patients	Established Patients
Jan	3000	4500
Feb	3400	4750
March	3640	4320
April	3789	4100

Payments by Provider Types

Month	Provider A (%)	Provider B (%)	Provider C (%)
Jan	45%	31%	24%
Feb	41%	37%	22%
March	43%	35%	22%
April	45%	34%	21%

Collection Agency of PA, All rights reserved. These materials may not be reproduced without the express written permission of AACE.

17

Visualization on Excel: What charts to use

Scatter Chart

- Charts relationship between two variables
- More data points, stronger chart
- Illustrates "why" factors are related

Y (dependent)

X (Independent)

$y = 15.366x + 4207$
 $R^2 = 0.8277$

Pie Chart

- Composition of a single variable
- Works best with less than 5 categories
- Pie slices should be ordered from largest to smallest

Payment by Provider Type

Provider Type	Percentage
Provider A	45%
Provider B	31%
Provider C	24%

Payment by Provider Type

Provider Type	Percentage
Provider A	45%
Provider B	31%
Provider C	24%

Collection Agency of PA, All rights reserved. These materials may not be reproduced without the express written permission of AACE.

18

Visualization on Excel: What charts to use

Combination Charts

- Best used when your data series are two different *types* of data
 - For example, measuring % class attendance and average test score
- Data would not be interpreted well on the same scale/axis

The image shows two charts side-by-side, both titled "Number of Patients Seen and Average Waiting Time (mins)". The left chart is a combination bar and line chart where the y-axis ranges from 0 to 12,000 for the number of patients (blue bars) and 0 to 60 for the average waiting time (orange line). A red 'X' is placed over this chart. The right chart is a bar chart where the y-axis ranges from 0 to 12,000 for the number of patients (blue bars) and 0 to 60 for the average waiting time (orange line). This chart is marked as a better visualization.

19

AAPA
American Academy of PAs

Communicating your Findings

20

Communicating your findings

The language you use to share your results is related to how well you can leverage your data.

You are not trying to sell the *numbers*; you are trying to tell a story.

- Communicate the big picture of the results through your tables and figures
- Why does this matter? What should be done?

21

Stress the importance of the data (numerically)!

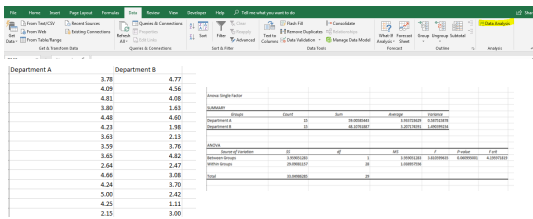
- Excel can be used to calculate statistical tests (which lets us say if the interactions between numbers matter)
- The Data Analysis button can be used to calculate:
 - ANOVA
 - Correlation
 - Moving averages
 - Regression
 - T-Test
 - Z-Test

Excel (especially scatter and line charts) can be great for forecasting and trend analysis

© Microsoft. All rights reserved. These materials may not be reproduced without the express written permission of MSFT.

22

Finding the Data Analysis Tool



© Microsoft. All rights reserved. These materials may not be reproduced without the express written permission of MSFT.

23

Communicating your findings

Here are some general tips for communicating your research

- Avoid jargon
Do not assume the audience all understands an industry-specific (or workplace specific) term/acronym. Plain language is easier to follow.
- Be concise
Keep your points short and include only what is necessary. Your audience can easily lose their train of thought.
- Know your audience
This is key. Relate your message/data to what your AUDIENCE cares about.
- Fact check yourself
Understand your data. Be prepared for questions. Some people in the audience may try to poke holes in your logic.
- Prepare for anything
You might lose power, the technology might fail, or you may run into someone interested in the half. Be ready to adapt your message to the situation.

© Microsoft. All rights reserved. These materials may not be reproduced without the express written permission of MSFT.

24

Communicating your findings

Looking for more? Try Grice's Maxims of Conversation

Maxim of Quantity: Be informative but say no more than is needed.
Maxim of Quality: Be truthful and say nothing that is not supported by your data/evidence.
Maxim of Relation: Be relevant to the discussion at hand.
Maxim of Manner: Be clear, brief, and orderly.

© American Academy of Pediatrics. All rights reserved. These materials may not be reproduced without the express written permission of AAPA.
