

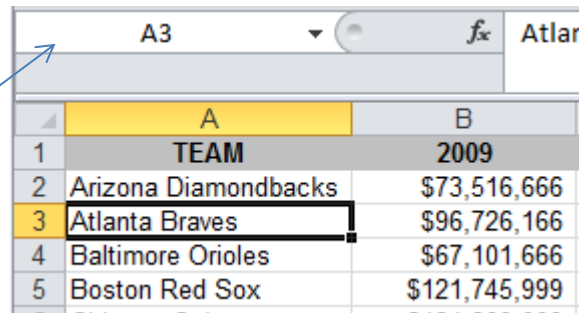
# Introduction to Excel

Use the spreadsheet called “MLBPayrolls2009\_2011” to follow along.

This tipsheet is specific to Microsoft Excel (2007 and newer) for Windows, however other spreadsheet software – such as Google Sheets – have the same functionality. Of course, the buttons and specific ways to do things might be different. But I often find answers by doing a web search using the Excel terms (i.e. “how to format cells in Google Sheets). Excel for Mac has all the same functions, but looks slightly different – especially in terms of where buttons and various tools are stored.

First, a few things about spreadsheets:

Each sheet is made up of columns (labeled by letters) and rows (labeled by numbers). Each cell is identified by an “cell address”, consisting of the letter and number. For example, in this picture, the value “Atlanta Braves” is stored in cell A3. Note that when the cursor is on a cell, the cell is outlined in black. Note the small black square in the corner – we’ll be using that in a couple minutes. Also note that the letter and number are highlighted in yellow and the cell address is displayed in the upper left corner of the sheet.



	A	B
1	TEAM	2009
2	Arizona Diamondbacks	\$73,516,666
3	Atlanta Braves	\$96,726,166
4	Baltimore Orioles	\$67,101,666
5	Boston Red Sox	\$121,745,999

If you’ve ever played the board game, “Battleship”, you’ll find this similar.

You’ll also notice that your cursor will change as you move it around. Most of the time, it will look like a big fat white plus sign. But if you hover over the top of the black box (around “Atlanta Braves”) you’ll get a tool that allows you to move that cell and if you hover in the lower right corner – where that little black square is – you’ll get a thin black cross. That thin black cross is used for copying formulas down or across the page. We’ll be using that a lot.



Used for selecting cells



The I-beam - indicates that you may type text in this area.



The fill handle - used for copying formula or extending a data series.



Used to select a whole row/column when positioned on the row number or column letter.



Appears at the border of the column letters. Drag to widen or narrow the width of a column.



Appears at the border between the row numbers. Drag to increase or decrease the height of a row.

## TERMS:

- **Columns** contain “categories” of data and are vertical
- **Rows** contain “individual records” and are horizontal
- A **cell address** consists of a letter followed by a number, i.e. “B4”
- The **active cell** is the cell with the darker border. Click on a cell to make it active.
- Each Excel file is described as a **workbook**, which can contain multiple **worksheets**.
- The **formula bar**, which is located in the menu bar, is similar to the address bar in a web browser. Here you can view and edit formulas that you’ve created in the worksheet.

Excel's original purpose was for accountants, so it is really useful for doing math on small or large datasets. But you'll find that Excel can do a lot of things – more than you'll probably ever need – including summarizing large datasets. This tipsheet is going to cover some of the main mathematical functions you will likely use. After learning these, the next thing you should learn is how to use Excel's Pivot Tables, which allows you to summarize your data (i.e. who is the largest donor in the governor's race? What's the average pay for each position in Major League Baseball? How many times has each restaurant been inspected?)

Any time you have a list of numbers that require some sort of calculation like *percent change from one year to the next, percent of total, average* or just a *simple total*, it's easier to put the figures into Excel and let the computer do the tough stuff for you. Excel is also extremely helpful when *comparing a list of numbers* to find out which one is the biggest, which is the smallest, etc.

This tipsheet will use total Major League Baseball payrolls to show how to calculate percent change from year to year, percent of total, average payroll and how to sort the list to compare the teams.

Here's a sample of what the data looks like when we start out.

	A	B	C	D
1	TEAM	2009	2010	2011
2	Arizona Diamondbacks	\$73,516,666	\$60,718,166	\$53,639,833
3	Atlanta Braves	\$96,726,166	\$84,423,666	\$87,002,692
4	Baltimore Orioles	\$67,101,666	\$81,612,500	\$85,304,038
5	Boston Red Sox	\$121,745,999	\$162,447,333	\$161,762,475
6	Chicago Cubs	\$134,809,000	\$146,609,000	\$125,047,329
7	Chicago White Sox	\$96,068,500	\$105,530,000	\$127,789,000
8	Cincinnati Reds	\$73,558,500	\$71,761,542	\$75,947,134
9	Cleveland Indians	\$81,579,166	\$61,203,966	\$49,190,566
10	Colorado Rockies	\$75,201,000	\$84,227,000	\$88,148,071
11	Detroit Tigers	\$115,085,145	\$122,864,928	\$105,700,231
12	Florida Marlins	\$36,834,000	\$57,029,719	\$56,944,000

Before doing any work on your data, here are a few other tips to keep in mind:

- 1) Make sure all your columns are labeled and that the column label only takes up one row. I frequently get datasets from government agencies that have two rows worth of labels – you'll have trouble sorting this type of spreadsheet, so you'll want to change the header/label row before you start working. Note that in the picture above, my column labels are highlighted in grey (I did this by highlighting the cells I wanted to change to grey and right-mouse click to choose "Format Cells", then changed the "Fill" color to grey. You don't need to color them like this...I just find it's useful to have the labels jump out at me.
- 2) Make sure that any stray information that came with your dataset – such as a source tag at the bottom or notes or titles or anything above the data table – are separated
- 3) Make sure your data table doesn't have any empty rows or empty columns. All the data should be neatly set up in consecutive columns and rows. In this case, we have one row for each team and each row has the team name on it. Then there are columns for the total each team spent on payroll in 2009, 2010 and 2011. This is a really important point because government agencies that send you data in Excel (or post it on their website) are horrible about following this. I often get datasets with blank columns jammed in the middle – seemingly to just make it look better.

### Percent Change:

To do this properly, it's crucial to have last year's total and this year's total lined up side by side (as shown above). It's possible to do this if the numbers are in other configurations, but it's much more difficult.

In the first blank column on the right (in this case Column E), create a new column header called "PctChange". Then in the cell just below the header (in this case, cell E3) we're going to create a formula that we can use to generate a percent change for all the teams — and it only needs to be typed once!

The formula is straight out of middle school math:

New value minus Old Value, divided by Old Value.

A way to remember this formula is this: N-0/0 (Think of the word "No!")

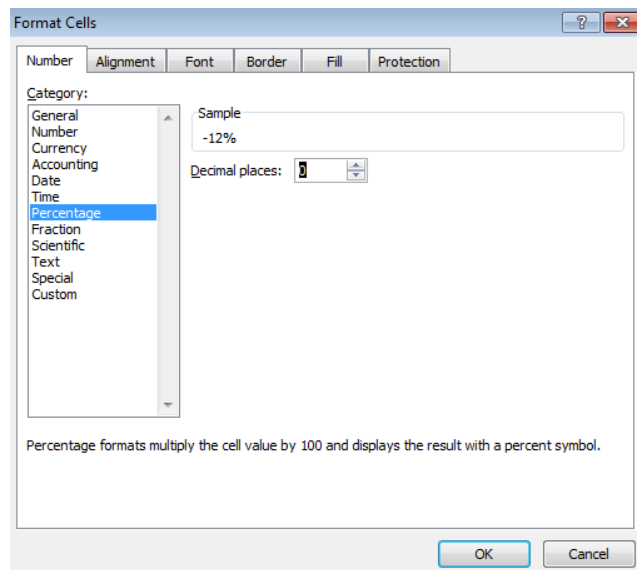
But with Excel, you set up a formula using the cell addresses – not the actual numbers. So in this case, the Arizona Diamondbacks 2010 payroll is in cell **C2** and their 2011 payroll is in **D2**.

So here's the formula:  $= (D2-C2)/C2$

	C	D	E	F
	2010	2011	PctChange	
6	\$60,718,166	\$53,639,833	$= (D2-C2)/C2$	
6	\$84,423,666	\$87,002,692		
6	\$81,612,500	\$85,304,038		
9	\$162,447,333	\$161,762,475		

When you are finished typing the formula, press the **Enter** key and you will see the result. It will probably be shown as decimals or maybe even as

currency . If so, put your cursor back on the value and right-mouse click and choose "Format Cells." There you choose the "Number" tab and set it to percentage and choose how many decimals you want to display.



Now to apply the formula to the other teams, you need to put your cursor back on the Diamondback's result and put your cursor over the lower right corner of the cell – where you will find a very small black square known as the "fill handle". (**Note:** when you move your cursor to different positions around cell E3 you will find your cursor changes shape.)

PctChange
0.28

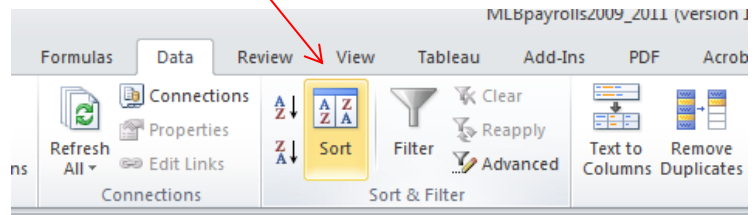
\*\*\*Notice the black square in the lower right corner (a.k.a. "fill handle")

When the cursor changes to a thin black cross, push down on your left mouse button and drag or "copy" the cell down through the subsequent cells until you reach the end of the list.

### Sorting the List:

To easily determine which team had the largest payroll increase and largest decrease, we need to sort the list. Put your cursor somewhere in the middle of your data area or text (you'll see a black box around whatever cell you landed in — this is the "active cell".) You don't need to highlight anything! Many beginners make the mistake of highlighting the column they want to sort by — don't do this! Just make sure your cursor is inside your data table — not out in the blank area of your worksheet — and then go to **Data** on the menu bar and choose **Sort...**

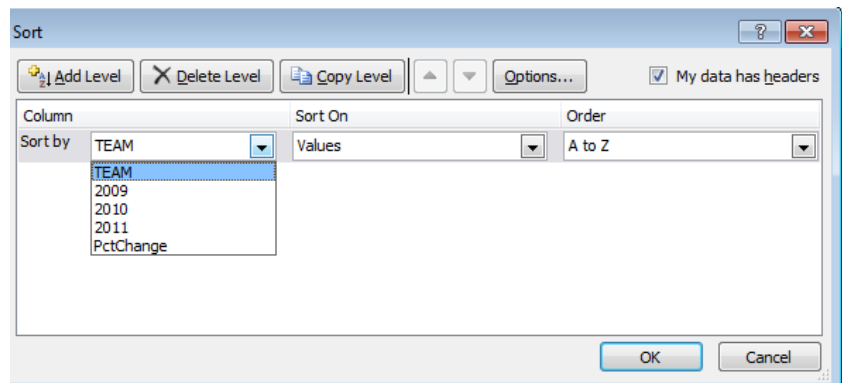
Excel will automatically highlight your data table and then you will get the chance to choose which field you want to sort by.



25047329

	C	D	E	F	G	H
	2010	2011	PctChange			
66	\$60,718,166	\$53,639,833	-12%			
66	\$84,423,666	\$87,002,692	3%			
66	\$81,612,500	\$85,304,038	5%			
99	\$162,447,333	\$161,762,475	0%			
00	\$146,609,000	\$125,047,329	-15%			
00	\$105,530,000	\$127,789,000	21%			

This will bring up a dialog box where you can choose which column you want to sort by. In this case, we want the list to go from highest percent change to lowest. So we would choose (from the pull-down menu in the left box called "sort by") the "PctChange" column and set it to go descending under the Order pull-down menu. (Note: If the pull-down menu doesn't show actual field names, but instead says "Column A", etc, check the checkbox in upper right that says "My data has headers.")



**Percent of Total:**

We may want to know what percent of all MLB money spent on payroll is made up by the Yankees’ monstrous payroll. Here’s how to do that.

First, we need to calculate a total payroll for all teams combined. And while we’re at it, let’s calculate the total for each of the 3 years. To do that, go to the bottom of column B (where the 2009 data is stored) and leave one blank line, then start a new formula on the next line to sum the figures. Instead of typing in all of the numbers in the list (like you would on a calculator), we’re going to add up all the numbers within a particular “range”. We use what’s called a **Sum** function.

Start the formula with an Equal sign, then within parentheses, indicate the range by typing the cell where the first value is located (in this case the list starts in B3) and then type a **colon** and the cell address where the last value is located. After you type this formula and press the **Enter** key you will see a total.

27	St. Louis Cardinals	\$88,528,409	\$93,540,751	\$105,433,572	13%
28	Tampa Bay Rays	\$63,313,034	\$71,923,471	\$41,053,571	-43%
29	Texas Rangers	\$68,178,798	\$55,250,544	\$92,299,264	67%
30	Toronto Blue Jays	\$80,538,300	\$62,234,000	\$62,567,800	1%
31	Washington Nationals	\$60,328,000	\$61,400,000	\$63,856,928	4%
32					
33		=sum(b2:b31)			
34					
35					
36					
37					

Then put your cursor back on your answer and get the thin black cross cursor and pull the formula to the right, so that it fills in the totals for 2010 and 2011.

ngton Nationals	\$60,328,000	\$61,400,000	\$63,856,928	4%
	\$2,666,728,494	\$2,717,759,865	\$2,786,161,291	

NOTE: If your answers show up as #####...that just means your column is not wide enough to display the whole answer. Go up to the top of the sheet and click between two of the letters (the column labels) and drag to the right in order to widen the column out.

Now to do the percent of total for each team, go to the first blank column on the right (in this case, the F column) and create a new header called “PctTotal”. In the cell just below your new header, we’ll do the formula for percent total.

Again it’s middle school math:  
 Team Payroll divided by Total League Payroll.

For this one, let’s just do the 2011 payroll.

Our first team’s payroll for 2011 is located in cell **D2**. Our total league payroll is in **D33**.

This time we need to add something special to the formula to tell it to always use cell D33 for the total number – even when we copy down the formula to other teams. **Note:** If you don't do this, Excel will use cell D34 for the second team, D35 for the third team, etc., and you won't get the correct answers!

To “lock” the formula on a particular cell, you put dollar signs (\$) around the column letter. Here's what it looks like:

	D	E	F	G	H
	2011	PctChange	PctTotal2011		
166	\$53,639,833	-12%	=D2/\$D\$33		
666	\$87,002,692	3%			
500	\$85,304,038	5%			
333	\$161,762,475	0%			

Now copy down the formula (get your thin black cross tool) like we did before and highlight the answers so that the formatting menu comes back at the top. Change the results to Percentage.

You can also re-sort the list so that you can see which team accounts for the biggest chunk of the total money spent on payroll in the MLB.

### Averages:

Calculating averages is similar to how we totaled the payroll list earlier. This time we're going to use the **Average** function. We'll calculate the average payroll for 2009 (later we'll do the other years).

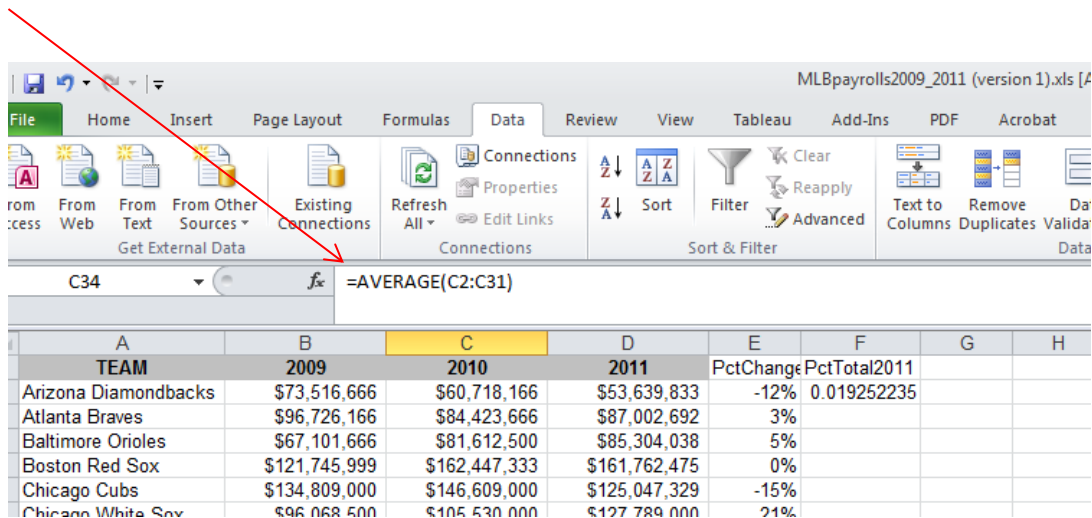
On the row just below your total MLB payroll, create a line for averages and start a new

30	Toronto Blue Jays	\$80,538,300	\$62,234,000	
31	Washington Nationals	\$60,328,000	\$61,400,000	
32				
33	TOTALS	\$2,666,728,494	\$2,717,759,865	\$
34	AVERAGE	=average(b2:b31)		
35				
36				
37				

formula in the B column on that line. Here's what the formula looks like:

After pushing enter, put your cursor in the lower right corner of the answer and get the thin black cross. This time copy the formula to the RIGHT so that you get an answer in the C column too. If you put your cursor on the new answer in the C column, you'll see that Excel accurately guessed that you wanted to do an average on C2:C33 instead.

To see, put your cursor on your new answer, then look at the top of the Excel page in the white bar (a.k.a. “formula bar”). Here’s what the original formula looks like in the formula bar:



## Some other functions, you’ll find useful:

### Median:

You can calculate a median value – the middle value – the same way we just did the averages. Except use the word “median” instead of the word average:

`=median(b2:b31)`

### Rank:

Instead of simply sorting your results to put them in “order”, this is a more sophisticated way to rank your records and to account for ties.

`=RANK(This Number, $Start Range:$End Range$, Order)`

This Number should be the cell where your data starts.

Start Range should be the cell where your data starts. Anchor with dollar

signs. End Range should be the last cell of your data. Anchor with dollar signs.

Order is either a 1 (smallest value will get assigned #1) or a 0 (largest value will get assigned #1).

Example: `=RANK(B2,$B$2:$B$31,1)`

### Per Capita or other Rates:

Per person or per capita rates are used to level the playing field. They’re often used to compare dissimilar places or events: crimes in cities with different populations, deaths from various diseases, etc. Also used for very big or very small numbers to make them easier to understand.

In Excel, if you wanted to calculate the per capita (or per person) for a list of cities with the number of murders in each city, you need a separate column/field that lists the population of each city.

Then in a new field you would use this formula:

`=murders/population`

Most likely, though, this will yield a very small number like 0.03. So you need to bring it up to something that readers will understand. To do that you multiply by a set number such as 1,000 or 10,000 or 100,000, like this:

`=(murders/population)*10000`

## Shortcuts and other useful tricks

**To highlight your block of data** (to sort or copy): Push Control-Shift-Asterisk (\*) at the same time. The cursor must be somewhere inside the chunk of data. (only works in Windows)

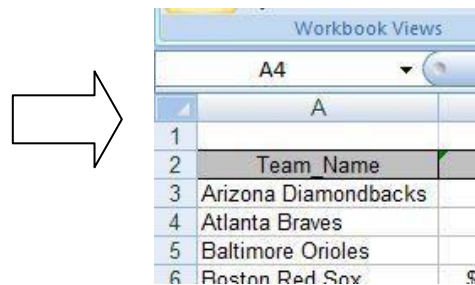
**To check the four corners of your highlighted data chunk:** While the data is highlighted, push Control-Period (.) at the same time, and repeat four times. Each time you push the keys it will go to a different corner of the data. This is useful for making sure you have highlighted the full chunk (and nothing extraneous) before sorting or copying.

**Freeze Panes:** To lock your field names in place so that you can always see them when you scroll down the page. Place your cursor in the cell just below the row that you want to lock into place, and all the way to the left of the page. Then go to View menu and select "Freeze Panes." It gives you several options, such as freezing the top row or freezing the far left column.

**To return to the top of your data:** Push Control and Home keys at the same time.

**Paste Special:** Use this function when you want to get rid of the formulas behind a column or row of data (until you get rid of the formulas, that row will be dependent on the other rows/columns used to calculate the formulas). Highlight the row(s) or column(s) that have the formulas. Copy the data using Ctrl-C or the copy button. Then put your cursor where you want to paste the data (best choice is to put it in a new column or row) and right-mouse click, then choose "Paste Special." A little box will come up: Under the Paste section at the top, choose "Values". Then say OK.

**To set all of your columns to the appropriate width:** Highlight all of your data by putting your cursor in the empty grey box in the left corner between the A & 1.



Then put your cursor in the grey, directly on the line between the A & B (or any two columns) and your cursor will change to more of a line. Then, double-click with your mouse and all of the columns on the page will widen or shrink to the maximum necessary width.

The same trick works for widening the rows (1, 2, 3, etc). When the data is highlighted, just put your cursor between the 1 & 2 (or any two rows) and double-click.



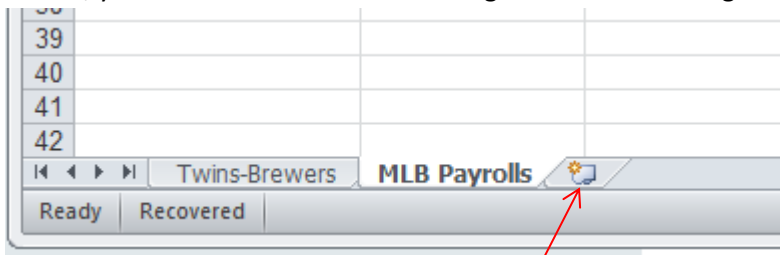
**Hide columns:** You can hide columns to get them out of your way or to avoid printing them by highlighting the columns you want (click on the letter at the top of the column) and right-mouse clicking to choose “Hide columns”. The columns will disappear. To get them back, highlight the two columns on either side of the ones that are missing and right-mouse click and choose “Unhide columns.”

**Add columns or rows:**

To add a new column in the middle of your data, click on the letter of the column immediately to the RIGHT of where you want the new column. So if you want a new column between columns B and C, you should click on C. Make sure the whole column is highlighted (if you clicked on “C”, Excel will automatically highlight the whole column) Then right-mouse click and choose “insert”.

To add a new row, click on the number of the row just BELOW where you want a new one. Make sure the whole row is highlighted, then right-mouse click and choose “insert”

**Worksheets:** You can toggle between different worksheets in an Excel workbook using the tabs in the lower left corner that have default names of “Sheet 1”, “Sheet 2”, etc. To change the name, double-click on the “Sheet 1” and it will turn black. Then you can start typing to give it a new name. To add a new worksheet, you’ll see a little button to the right of the tabs that gives you the option to add a new worksheet. You can move worksheets around so they appear in a different order, by clicking on the named tab with your mouse, holding down and dragging it in whatever direction you would like.



Click here to create a new worksheet

Created by  
@MaryJoWebster  
[Mjwebster71@gmail.com](mailto:Mjwebster71@gmail.com)  
Updated: October 2015