

module2_lecture6

Mon, 12/27 3:51PM 17:34

SUMMARY KEYWORDS

provinces, numbers, calculate, excel, sum, inbuilt functions, suppose, crime rate, lowest crime rate, bmo, b4, inbuilt, crimes, population, highest crime rate, data, enter, total, cells, google sheets

SPEAKERS

Sumon Majumdar

So so far, I've introduced you to the notation of summation. And in this clip, I'm going to show you how you can use Excel, and also similarly Google Sheets, to actually do summations, if you wanted to add up a series of numbers. And I'm also going to tell you about a lot of other inbuilt functions, which are already there in Excel to help you easily do several sorts of computations. The particular data file that I'll be using is also posted on module. And you can download it to your computer, and perhaps follow along as I do the calculations here. Because as I've also mentioned before, learning software, there is no better way than to actually do these operations yourselves, right. That's how you can learn how to use a particular bit of software. Okay, let's get started.

So here is a data file here, where I first calculate, collected data on, on population across the various provinces in Canada. Right, so this is Newfoundland and Labrador, then Nova Scotia, PEI, New Brunswick, and so on. So these are the 2021, the second quarter estimates from a statistic scanner. So, and in this column are the actual population numbers. So now suppose I want to calculate the total population across Canada. So what I'm going to do is use a formula. And to use any formula in Excel, use the equal to side, and then the formula I'm going to use is the sum. So sum is an inbuilt formula in Excel. And what it does is, as the name suggests, it sums up over a set of numbers. So what's the set of numbers I want to add up? I want to start from here which is B4. And where do I want to go up to? I want to go up to Nunavut, which is B16. So I want to add up the set of numbers from B4 up to B16. So this is like here is like the index in the summation notation, the i runs from 4 to 16, because that indicates the set of cells over which we are summing. So now if I hit Enter, that gives you the total population of Canada, which is 38 million. Now, you need not want to add up all of the provinces, what say you wanted to only find out the total population of the Atlantic provinces. So this would be Newfoundland Labrador, Nova Scotia, PEI and New Brunswick. So if you want to look at only Atlantic provinces, say Atlantic provinces, again, I'm going to do equal to sum, right, and now I'm going to start from B4 and my index is going to run up to not all the provinces but only up to New Brunswick. So it's going to run from B4 through B7 and I hit Enter, so, this gives you the total population of the Atlantic provinces which is about, what is it, 2,044,000.

So, the sum is an inbuilt function in Excel. Excel has a lot of other such inbuilt functions. So for example, suppose you wanted to find the maximum so in a small data set like this, you can usually

eyeball and see that the maximum is at Ontario, right? But oftentimes you have larger data sets, maybe looking at sales, maybe looking at prices, right. And there it may not be as easy. So you can use the maximum sorry, the, the max that's the inbuilt function it returns, see, it even tells you what it does. It says it returns the largest value in a set of numbers. So again, max between B4 that's my first provinces population, and I want to go up to the last territories populations, that that's Nunavut. Close that parentheses, and hit Enter. So that tells you where that what's the maximum population among all these provinces.

Okay, you can calculate a lot of other things as well. For example, if you wanted to calculate the average, so that's already an inbuilt function in Excel, right? Suppose you wanted to calculate, let's say a minimum. Okay? There you have it. So, so there are a lot of inbuilt functions in Excel. And it's up to you to explore these inbuilt functions, and find out the one that's appropriate for your particular use. So let me do one other. A couple of other examples. So this is an example of crime. And so what I've done is here I've calculated, I have collected the crime statistics for different provinces. So for example, this here gives you the number of violent crimes that were in Newfoundland and Labrador in 2020. Then the next column here is the total non-violent crimes that took place in Newfoundland and Labrador in 2020. Okay, and same by for the other provinces. So for example, now, if you want to calculate the total violations, right, that means the total crimes, so these are total crimes, excluding any traffic violations, what I would do is add up the violent ones, plus the non-violent ones, right? So again, I used equal to denote that it's a formula, and I'm adding the violent ones plus the non-violent ones, hit Enter, that gives me the total number of crimes in Newfoundland and Labrador. Right, now, if I want to do the same for the other provinces, as I've shown you before, you can either just copy this cell and then paste it into the other cells, or there is an even simpler way that Excel gives you. So you look at the the bottom right hand corner of the cell for which you have already computed, see, it gives you a plus sign. Now, if you use that plus sign to expand, then it copies that formula into all of these other cells. So this then gives you the total number of crimes in Nova Scotia, PEI, New Brunswick, Quebec and so on.

Now, you may want to do other stuff, right? So you may be interested in the, say, the crime rate. So for example, suppose you're researching which is the safest province in the country, or which has the highest crime rate, or the lowest crime rate? So suppose you want to calculate the crime rate for a province, right? How is the crime, how do you calculate the crime rate? It's the total crimes as a percentage of the population. So let's compute that for Newfoundland and Labrador. So it's equal to the total number of violations. So the total crimes divided by the population. And to get it in a percentage form, we multiply it by 100. So what that tells you is that the crime rate in Newfoundland and Labrador is 6%. And now, suppose you want to compute this for all the other provinces, you just expand there, I forgot to expand it to the last one, which is Nunavut. And I just dragged that there, and that calculates you the crime rate for all the provinces. And now suppose you want to figure out, you want to organize these provinces, from the safest to the one with the highest crime rate. How do we do that? Again, what I'm going to do is I'm going to select all the data, all, the whole area in which there is this data on population and crime. And I want to solve this data from the lowest crime rate to the highest crime rate. So what I'm going to do is I'm going to make use of the sort and filter. This is one of the buttons here on the, on the top of Excel. And it gives you different options. So I'm going to choose Custom sort.

So what custom sort does is that it allows you to sort by any of the columns. See, you could either custom sort it by population, or by the number of violent crimes, the non-violent crimes. But what we

are interested in here is sorting by the crime rate. And this is in column G. So I want to sort by Column G. And it gives you the order, do you want to sorted from smallest to largest, or largest to smallest, so when the safest ones first, so I want to get it from smallest to largest. And I hit OK. And see, it has now sorted by provinces, according from the lowest crime rate to the highest, right. So if you now look across, the province with the lowest crime rate is Quebec, then comes Ontario, then PEI, Nova Scotia, and so on. So this is an easy way of doing data analysis using Excel. Now, you can, again, do lots of other stuff. So for example, if you want to look at the average crime rate, you may want to calculate the average of all of this again, if you look, if you use the average function from Excel, and that will tell you, okay, the if you want to look at from G4 to G16. And you want to find the average of these, that's, that gives you 15.02.

So let me now show you, let me show you another example, very much on the same lines, but this instead is someone's stock portfolio, right? So this person has stocks in lots of different companies, BMO, TD, Enbridge, and so on, right? And here we have the number of stocks that he has of BMO, and the price of it. And what I'm trying to calculate is the total valuation of his portfolio. So let's first start by calculating the value of his BMO shares. So he has 10 of them, right, that's in K4, times the price of each of them, which is in L4, right. And if I hit Enter, so that's the value of his BMO shares. And if I again drag this down here, this just copies the formula into all of the other cells. And if I now want to calculate the total value of his portfolio, it'll just be the sum of all of these values starting from BMO, and going up to, up to Twitter, close parenthesis. So that's the value of his, of his portfolio.

It's 7142.19. But now, if you're looking across, you see that one of these numbers here, the one for CCL, that's negative four. So do you wondering is this right? Or is this maybe a wrong entry? Maybe he, maybe it was a different number, or maybe his son just pressed a negative button while he was inputting there? We don't know what's happening, right? So maybe you want to ignore this. So you want to only calculate the total of, of the ones that you know are correct. So what Excel allows you to do is allows you to do summation with conditions, right. And this is the sum if, so what it does is that sum if, it adds the cell specified by a given condition or criteria. So what I'm going to do here is I'm going to sum up, right, again, from starting from here, from M4, and up to here, which is M17. But I want to only sum up the ones that I know to be correct. That's the ones which are greater than zero. Okay, so let me use greater than zero. So that's my criteria. So what it does is that it's summing up numbers, but only ones which satisfy this criteria greater than zero. And I hit Enter, and it calculates it to be 7234.07. So what it is doing is that it's not adding this, the negative number, it's adding up all the ones which are greater than zero.

You can do lots of other sort of criteria, right. So for example, if you want to sum up, maybe again, from M4 to M17. But you want a different criteria, that's, you put the criterias within inverted commas. So suppose you want to maybe add up only those of the bigger of the, those which you're really doing well in your portfolio. Maybe the ones which have value greater than 200, okay. So I've just arbitrarily found the criteria. So this is greater than 200. Again, I close the inverted commas close the parenthesis hit this, right? So this adds up the numbers which are all greater than 200. Right, so you can do sums over a subset of numbers which satisfy a particular criteria. Similarly, you can do averages over sets of numbers which satisfy a particular criteria, which is given by average if. So, what I'm going to stop this clip here, but what I hope you have got an introduction to is the various functions that are already inbuilt into Excel. And many of these functions are also inbuilt into Google Sheets. So what I would like you to do is to explore these inbuilt functions because these are really

handy and offer you easy ways to do a lot of data analysis without really having to sweat it out very much just by using all of these inbuilt functions. So I hope you will explore them, and I hope you will do some other practice problems that I have outlined for you on the module using this sort of data analysis using Excel and Google Sheets.