Hands-On Activity: Using VLOOKUP

TOTAL POINTS 2

1.

Activity overview

Earlier, you learned about VLOOKUP, a function that uses vertical lookup to find specific values in a spreadsheet. In this activity, you will practice using VLOOKUP to consolidate information between two spreadsheets, clean data, and create a summary table from a query.

By the time you complete this activity, you will be able to use VLOOKUP to complete a variety of tasks in spreadsheets. This will enable you to clean and analyze data more efficiently, which is important for working with large datasets in your career as a data analyst.

What you will need

To get started, first access the VLOOKUP Practice Worksheet.

Click the link to the worksheet to create a copy. If you don't have a Google account, you may download the VLOOKUP Practice Worksheet directly from the attachments below.

Link to the worksheet: VLOOKUP Practice Worksheet

OR

Download VLOOKUP Practice Worksheet:

VLOOKUP Practice Sheet.xlsx

Search with VLOOKUP

Although you would usually clean your data prior to using VLOOKUP, this first step will illustrate why it's important to clean data first.

Imagine your research requires you to know how many hours an employee worked on a specific date. This is easy to do manually on a small spreadsheet and becomes harder as the amount of information grows or is spread across multiple spreadsheets. The VLOOKUP function provides a way to have the spreadsheet gather the information for you.

Assume you needed to figure out how many hours the employee Daniel Chan worked on January 3, 2020. In the spreadsheet you downloaded, it is easy to notice which number contains Daniel's name. But imagine if you had thousands of employees in your spreadsheet. It might not be easy to find his name without searching each cell. In this step, you are going to use Daniel Chan's name as the **lookup_value**, sometimes known as a search key, in VLOOKUP.

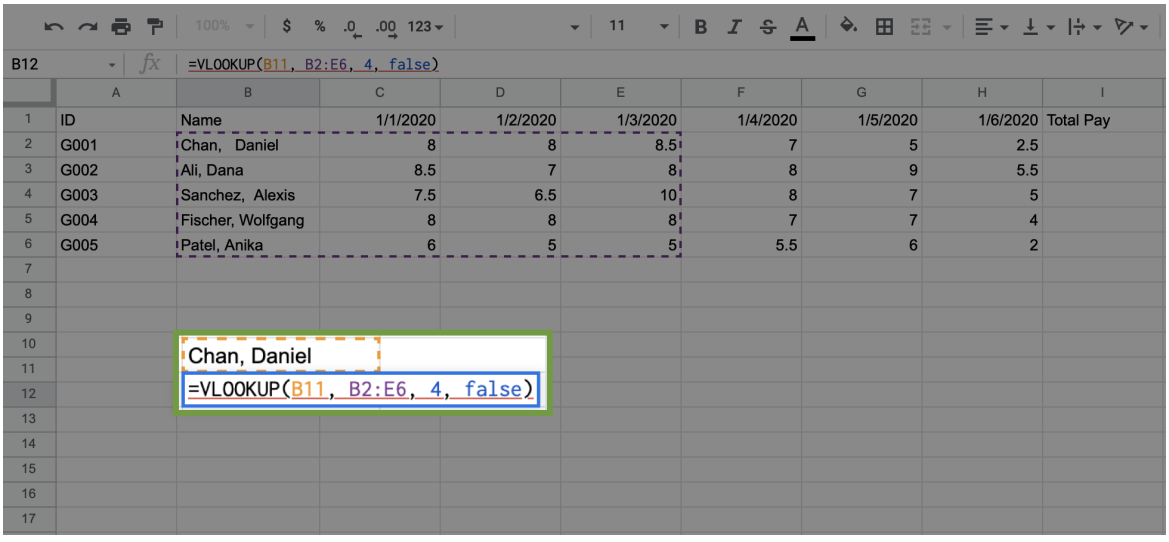
The syntax for the VLOOKUP function is **=vlookup(lookup_value, table_array, col_index_num, [range_lookup], true/false)**.

Search for the number of hours Daniel Chan worked on January 3, 2020.

1. In **B11** enter **Chan, Daniel.**

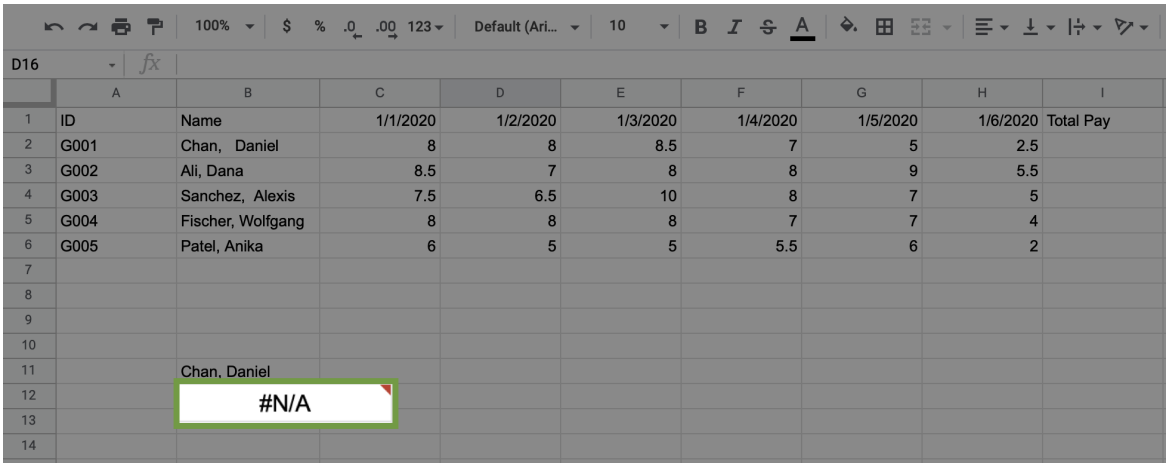
2. In **B12** enter **=VLOOKUP(B11, B2:E6, 4, false).**

As a refresher, this syntax means that the lookup value **is contained in cell B11**, the **table array contains cells B2 through E6**, you want to search in **column 4 of this array**, and you want an **exact match**. Remember that **column refers to the array column**, which represents the limits of your query.



	A	B	C	D	E	F	G	H	I
1	ID	Name	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Total Pay
2	G001	Chan, Daniel	8	8	8.5	7	5	2.5	
3	G002	Ali, Dana	8.5	7	8	8	9	5.5	
4	G003	Sanchez, Alexis	7.5	6.5	10	8	7	5	
5	G004	Fischer, Wolfgang	8	8	8	7	7	4	
6	G005	Patel, Anika	6	5	5	5.5	6	2	

3. Press **Enter** (Windows) or **Return** (Mac). The cell will now contain an error, **#N/A**.



	A	B	C	D	E	F	G	H	I
1	ID	Name	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Total Pay
2	G001	Chan, Daniel	8	8	8.5	7	5	2.5	
3	G002	Ali, Dana	8.5	7	8	8	9	5.5	
4	G003	Sanchez, Alexis	7.5	6.5	10	8	7	5	
5	G004	Fischer, Wolfgang	8	8	8	7	7	4	
6	G005	Patel, Anika	6	5	5	5.5	6	2	
7									
8									
9									
10									
11		Chan, Daniel							
12		#N/A							
13									
14									

Notice that the entry for Daniel Chan has extra spaces after the comma. Because **B11** does not contain those extra spaces, the search comes back with an error.

One option to fix this is to adjust the number of spaces until you get an exact match. However, this is not very efficient, and if you could identify the name, you would probably just use the cell number for your query.

The best way to handle this is to trim any extra spaces in the data. This is why it's important to clean your data prior to using VLOOKUP.

Prepare the data

Now you will prepare the data to help you more easily figure out how many hours employees worked. You first need to clean and label the data. Then, you can combine data from two spreadsheets using the trusty VLOOKUP function.

Clean and label the data

To trim the data, follow these steps:

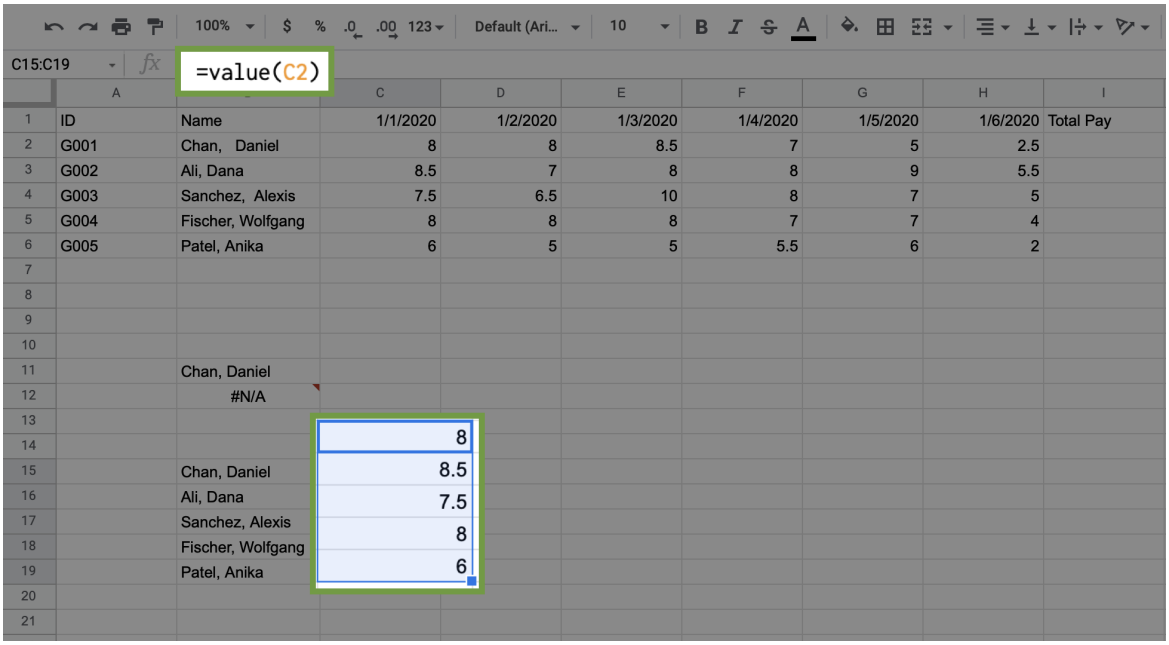
1. In cell **B15** type **=trim(B2).**

2. Click and drag down the bottom-right corner of the cell until you reach **B19**. The rest of the names will populate.

For this exercise, you are not replacing the trimmed data into the original table. There are many cases where you need to clean the data for your use, but you do not want to change data in the set with which you are working.

3. Scroll below the original data. In cell **C15** type **=value(C2).**

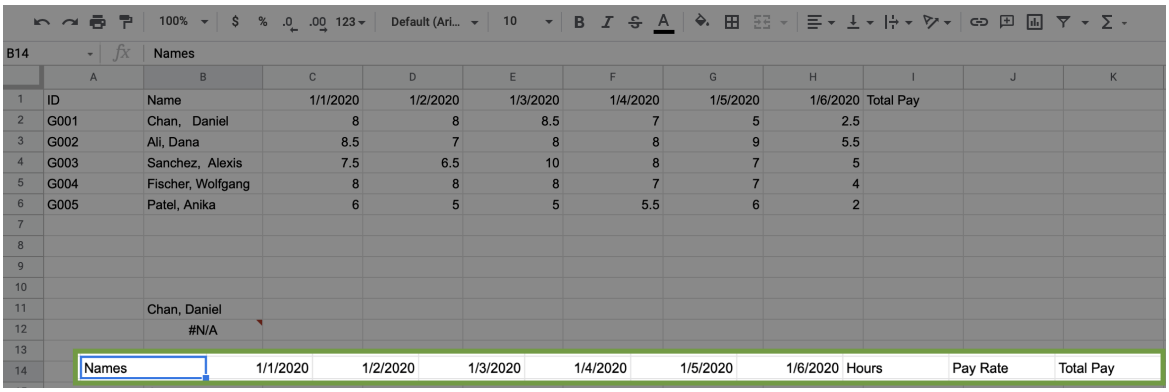
4. Click on the bottom-right corner of the cell and drag the cell down to populate the hours for the other employees.



It's also helpful to label the different columns for the data. Working with data gets messy quickly, and it is important to keep track of your value references.

Enter in the following labels:

- **B14:** Names
- **C14 – H14:** (Enter in the dates 1/1/2020 through 1/6/2020)
- **I14:** Hours
- **J14:** Pay Rate
- **K14:** Total Pay



Populate and sum the remaining hours

Use cells already populated in **C15** through **C19** to populate the remaining hours needed for each employee.

1. Click and drag the corner of **C15** to **H15** to populate the remaining hours for Daniel Chan.

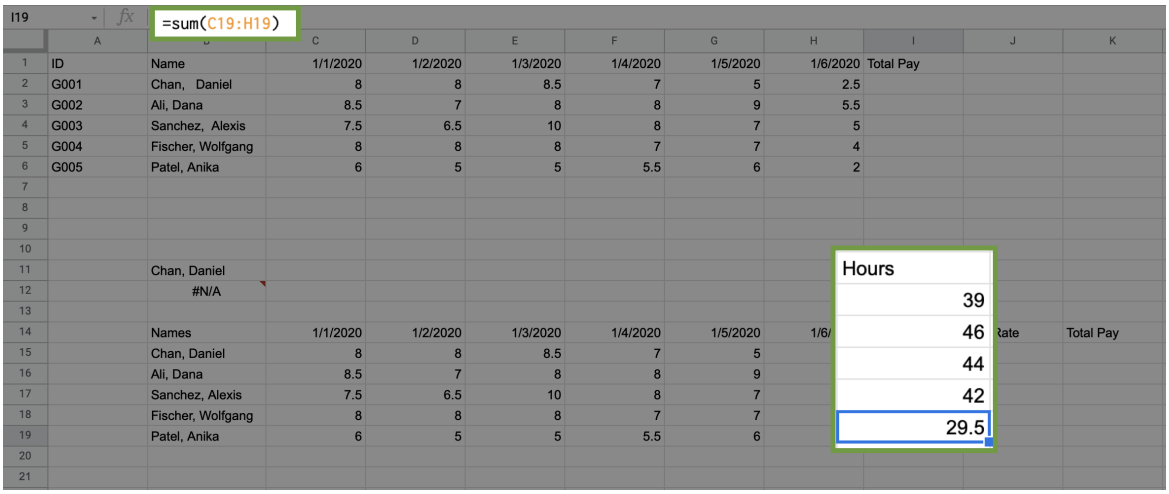
2. Repeat this process for the remaining employees.

	A	B	C	D	E	F	G	H	I	J	K
1	ID	Name	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Total Pay		
2	G001	Chan, Daniel	8	8	8.5	7	5	2.5			
3	G002	Ali, Dana	8.5	7	8	8	9	5.5			
4	G003	Sanchez, Alexis	7.5	6.5	10	8	7	5			
5	G004	Fischer, Wolfgang	8	8	8	7	7	4			
6	G005	Patel, Anika	6	5	5	5.5	6	2			
7											
8											
9											
10											
11		Chan, Daniel									
12		#N/A									
13											
14		Names	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Hours	Pay Rate	Total Pay
15		Chan, Daniel	8	8	8.5	7	5	2.5			
16		Ali, Dana	8.5	7	8	8	9	5.5			
17		Sanchez, Alexis	7.5	6.5	10	8	7	5			
18		Fischer, Wolfgang	8	8	8	7	7	4			
19		Patel, Anika	6	5	5	5.5	6	2			
20											

Now, fill in the **Hours** column for the employees.

3. In cell **I15** type **=sum(C15:H15).**

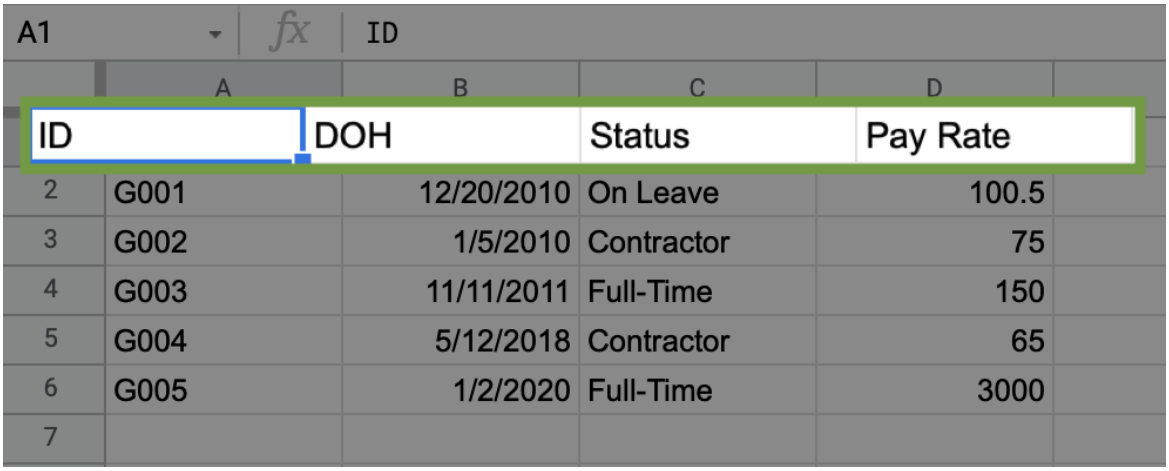
4. Click and drag down the bottom corner of cell **I15** to populate the sums for the remaining employees.



Import pay rate data

You might have noticed that pay information is missing from the spreadsheet. Data analysis often requires importing information from different data sources. In this case, the data required is on **Sheet2**. To import this data:

1. Click on **Sheet2**, which you can find at the bottom of the spreadsheet. **Employee ID, date of hire (DOH), status,** and **pay rate** are the data found on this sheet.



	A	B	C	D
	ID	DOH	Status	Pay Rate
2	G001	12/20/2010	On Leave	100.5
3	G002	1/5/2010	Contractor	75
4	G003	11/11/2011	Full-Time	150
5	G004	5/12/2018	Contractor	65
6	G005	1/2/2020	Full-Time	3000
7				

Now, use VLOOKUP to import pay rate data.

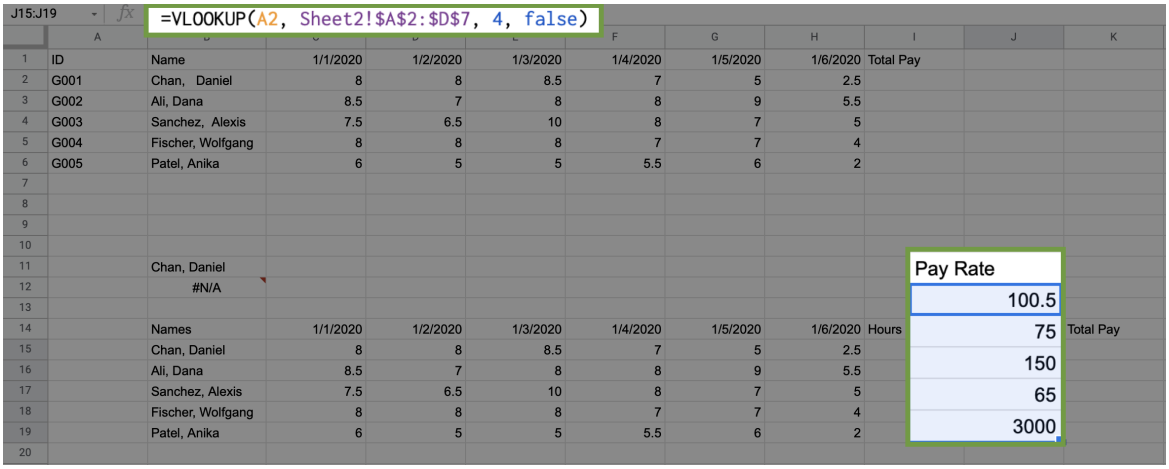
2. In **J15 (of sheet 1)** type: **=VLOOKUP(A2, Sheet2!$A$2:$D$6, 4, false).** Consider the syntax for this VLOOKUP function:

- **A2** refers to cell A2 in Sheet1.

Note: In Sheet2 the rate of pay, and related fields, are referenced by ID instead of employee name. You need to use employee ID to import the pay rate from Sheet2.

- **Sheet2!** refers to the sheet from which you want to access the data.
- **$A$2:$D$6** refers to the cells that make up the table array. The $ placed in front of the column tabs and cell numbers locks the formula so that it can be copied by dragging down the cell **J15** to import the pay rate for the other employees.
- **4** refers to the column from which the returned value will come. 4 means that the returned value will come from the 4th column in the selected array.
- **false** signifies that you want an exact, character-for-character match to the lookup value. If you put true instead, VLOOKUP would return an approximate match (or the closest match available) for the lookup value. This is not used very often in real-world situations.

3. Populate the pay rate for the remaining employees by dragging down the corner of the cell to copy the formula.



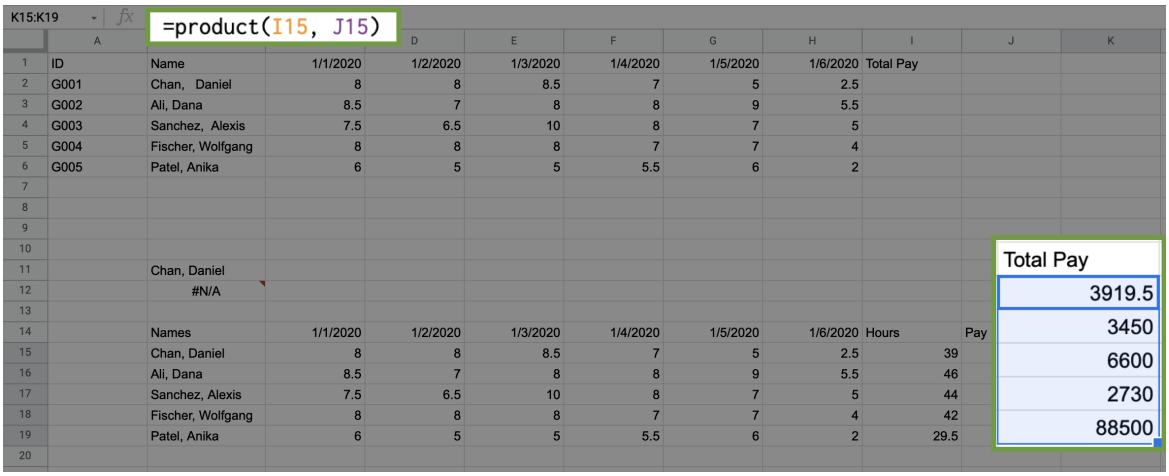
J15:J19 =VLOOKUP(A2, Sheet2!$A$2:$D$7, 4, false)

	A	B	C	D	E	F	G	H	I	J	K
1	ID	Name	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Total Pay		
2	G001	Chan, Daniel	8	8	8.5	7	5	2.5			
3	G002	Ali, Dana	8.5	7	8	8	9	5.5			
4	G003	Sanchez, Alexis	7.5	6.5	10	8	7	5			
5	G004	Fischer, Wolfgang	8	8	8	7	7	4			
6	G005	Patel, Anika	6	5	5	5.5	6	2			
7											
8											
9											
10											
11		Chan, Daniel								Pay Rate	
12		#N/A								100.5	
13											
14		Names	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Hours	75	Total Pay
15		Chan, Daniel	8	8	8.5	7	5	2.5		150	
16		Ali, Dana	8.5	7	8	8	9	5.5		65	
17		Sanchez, Alexis	7.5	6.5	10	8	7	5		3000	
18		Fischer, Wolfgang	8	8	8	7	7	4			
19		Patel, Anika	6	5	5	5.5	6	2			
20											

Now, calculate total pay.

4. In **K15** type **=product(I15, J15)**.

5. Drag cell **K15** down to populate the total pay for the remaining employees.



	A	Name	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Total Pay
1	ID								
2	G001	Chan, Daniel	8	8	8.5	7	5	2.5	
3	G002	Ali, Dana	8.5	7	8	8	9	5.5	
4	G003	Sanchez, Alexis	7.5	6.5	10	8	7	5	
5	G004	Fischer, Wolfgang	8	8	8	7	7	4	
6	G005	Patel, Anika	6	5	5	5.5	6	2	
7									
8									
9									
10									
11		Chan, Daniel							
12		#N/A							
13									
14		Names	1/1/2020	1/2/2020	1/3/2020	1/4/2020	1/5/2020	1/6/2020	Hours
15		Chan, Daniel	8	8	8.5	7	5	2.5	39
16		Ali, Dana	8.5	7	8	8	9	5.5	46
17		Sanchez, Alexis	7.5	6.5	10	8	7	5	44
18		Fischer, Wolfgang	8	8	8	7	7	4	42
19		Patel, Anika	6	5	5	5.5	6	2	29.5
20									

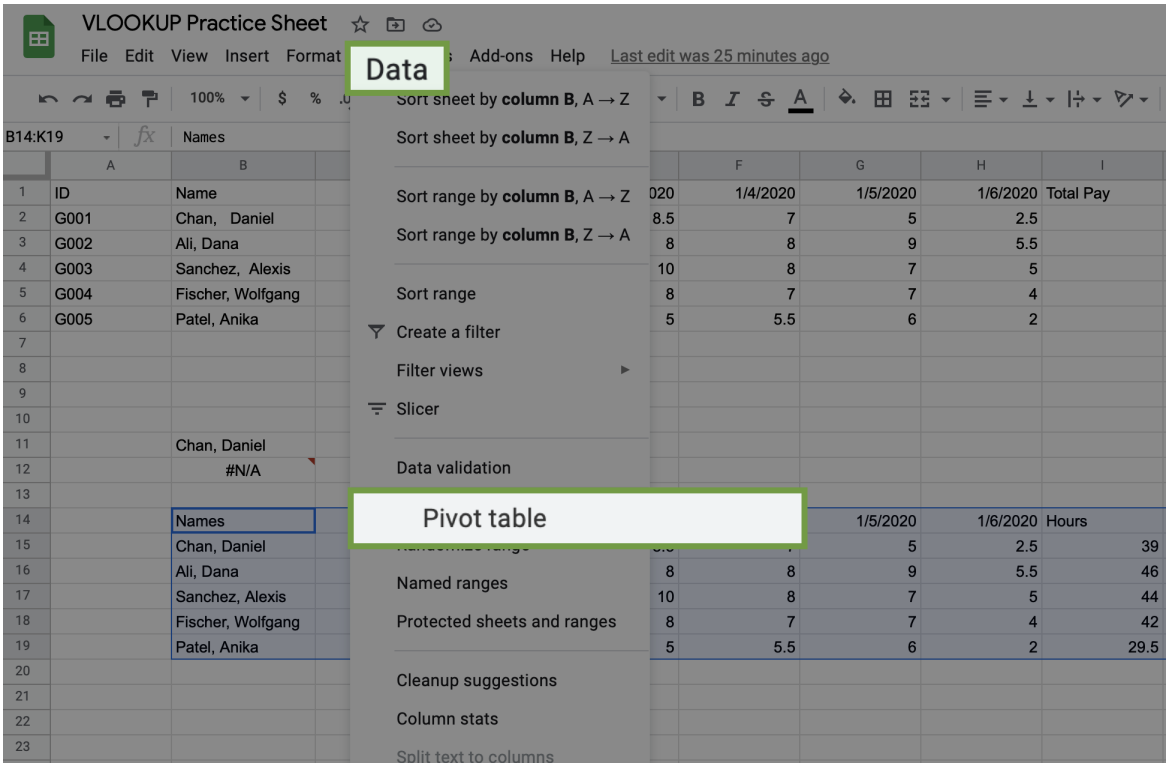
Create a summary table

Now that the data is clean and includes pay rate information, you can create a summary table, or pivot table. The following section demonstrates how to create a pivot table in Google Sheets. If you are using Excel, please follow the documentation for how to manually create a Pivot Table in Excel.

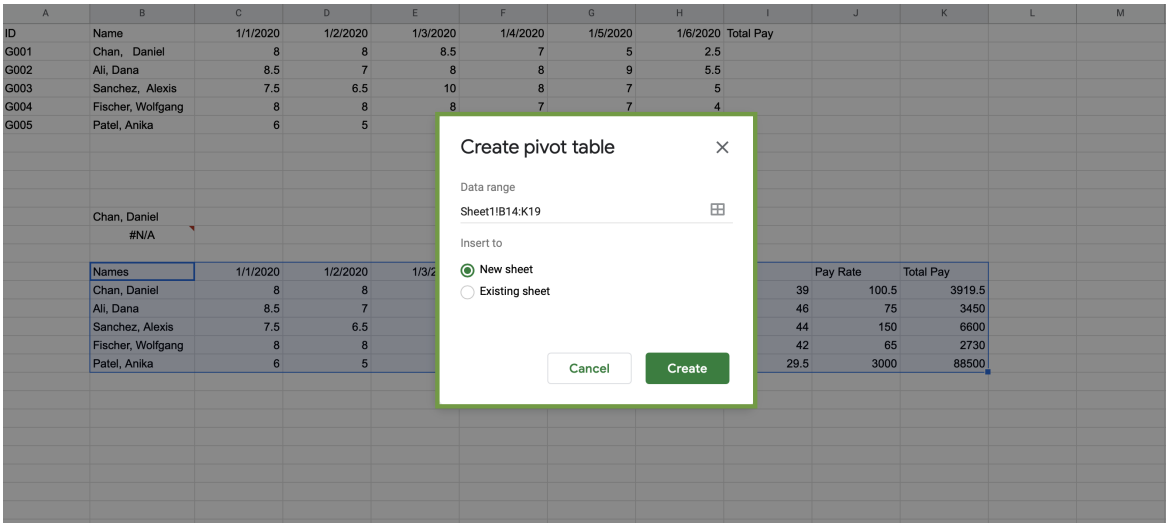
In Google Sheets, create a table for data in cells **(B14:K19)** using the following steps:

1. Select the data in cells **(B14:K19).**

2. Click on the **Data** tab, then select **Pivot Table.**



3. A pop-up window will display. Click on **New Sheet,** then click the **Create** button.



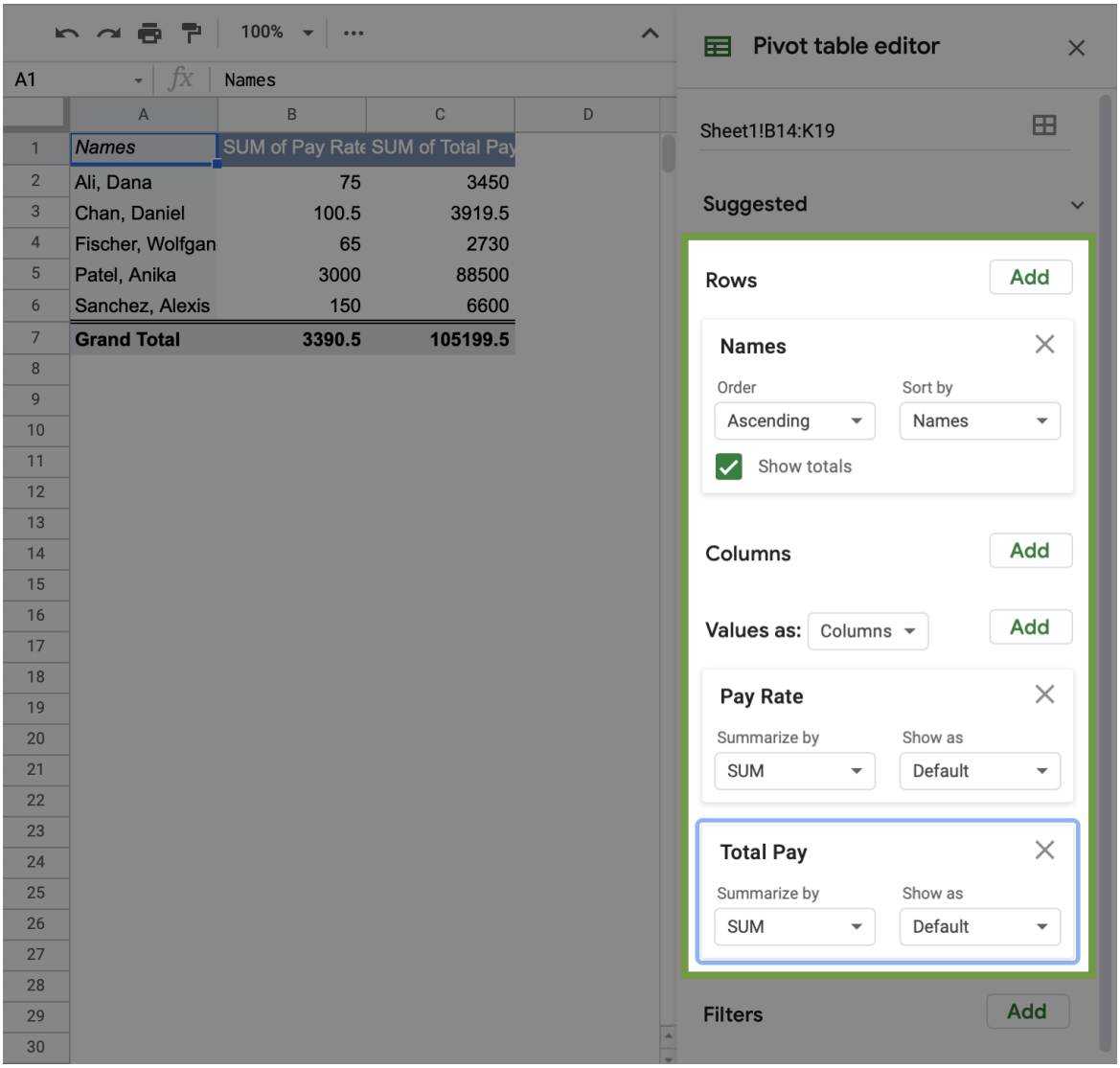
On the side of the new sheet, the **Pivot table editor** will display. The pivot table you are creating will contain each employee's name, pay rate, and total pay. Follow these steps to create the pivot table:

1. Click the **Add** button for **Rows.** Select **Names.**

2. Click the **Add** button for **Values.** Select **Pay Rate.**

3. Click the **Add** button for **Values** again. Select **Total Pay.**

The result should display like this:

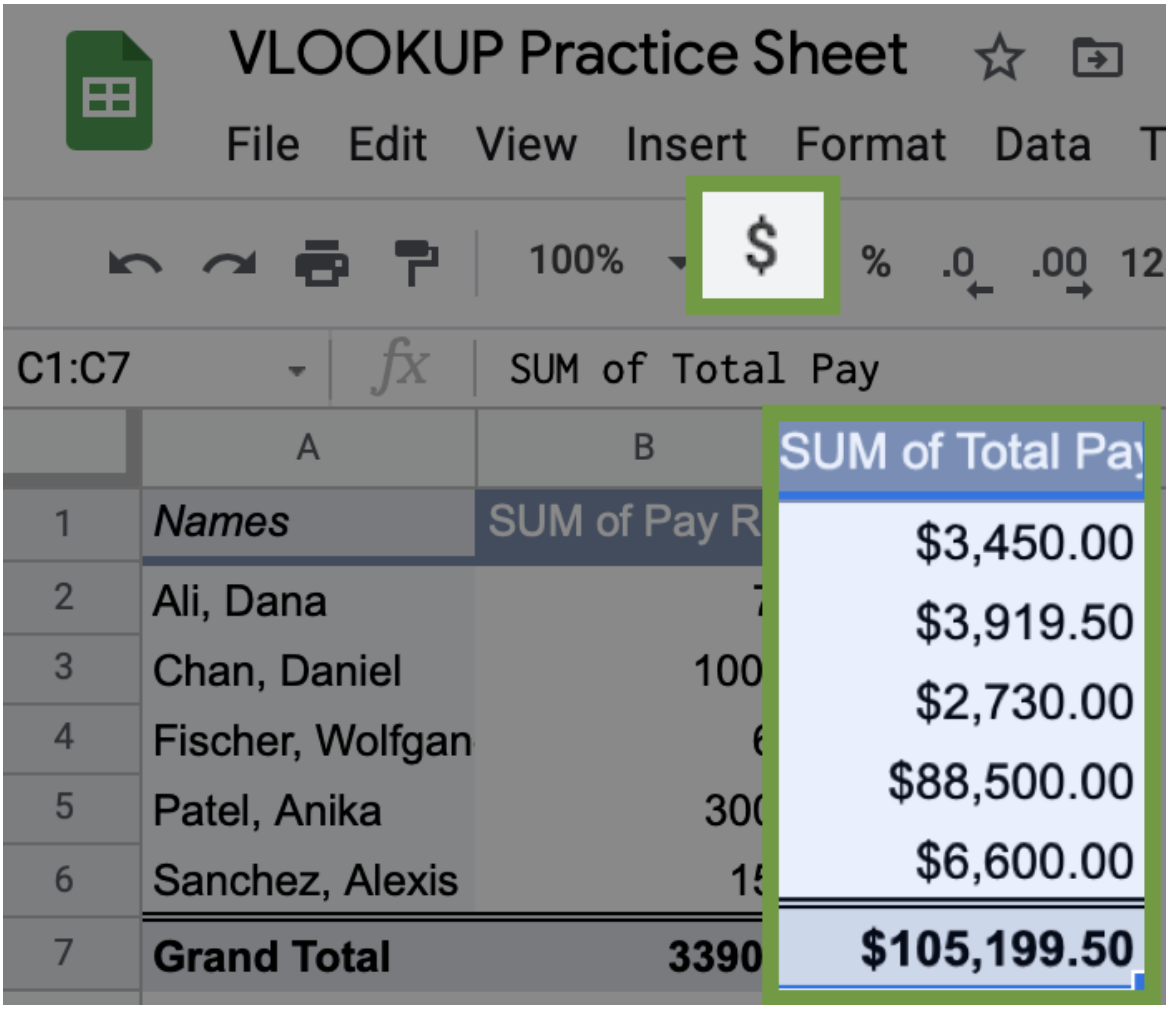


	A	B	C	D
1	*Names*	SUM of Pay Rate	SUM of Total Pay	
2	Ali, Dana	75	3450	
3	Chan, Daniel	100.5	3919.5	
4	Fischer, Wolfgan	65	2730	
5	Patel, Anika	3000	88500	
6	Sanchez, Alexis	150	6600	
7	**Grand Total**	**3390.5**	**105199.5**	

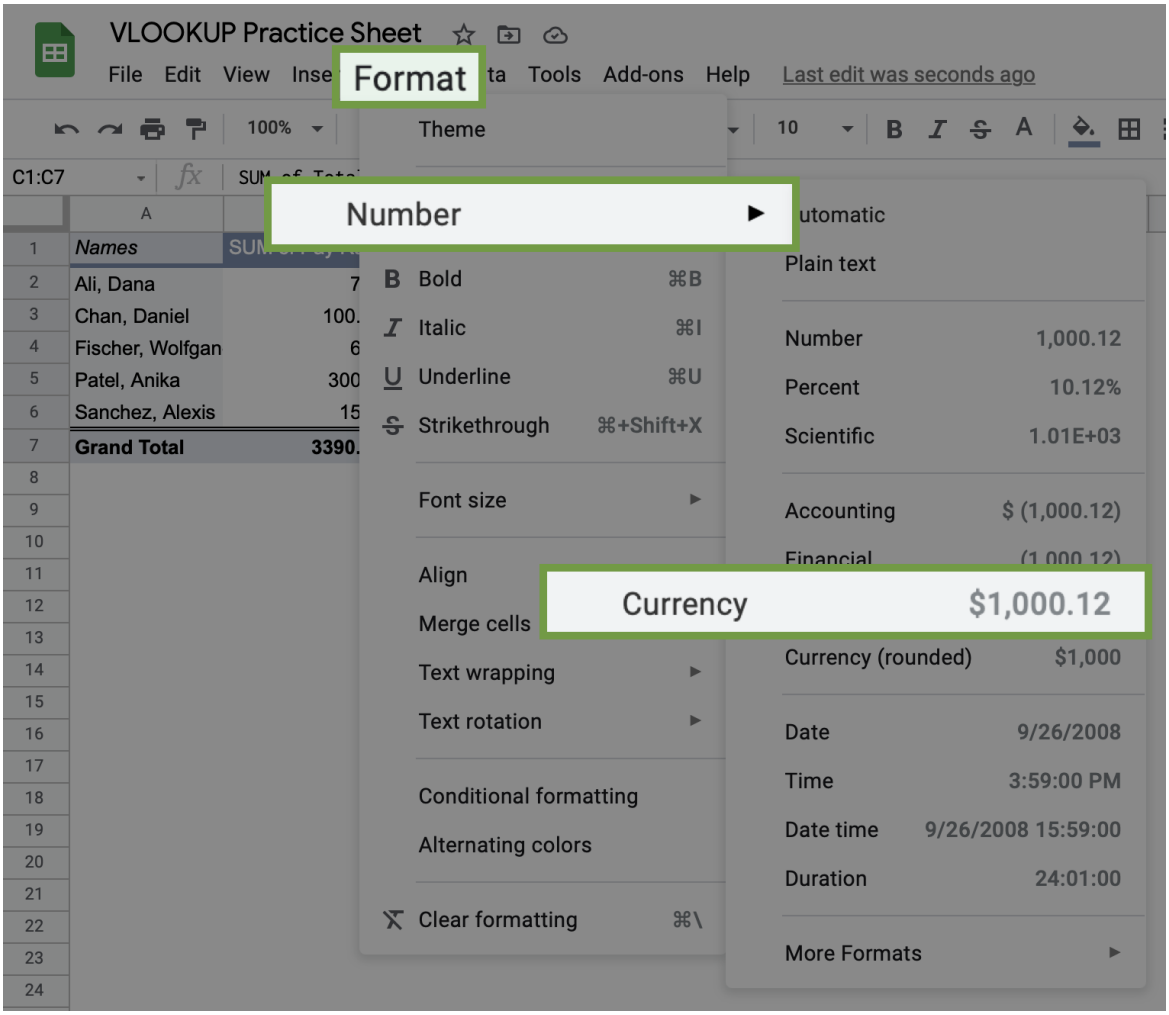
Next, convert the cells in the **Sum of Total Pay** column to currency.

4. Select the cells in the column **Sum of Total Pay.**

5. Click on the **$** symbol on the toolbar.



Alternatively, you can also click on the **Format** tab, select **Number**, then select **Currency.**



Congratulations! You have now used VLOOKUP and created a pivot table, two essential tools for analyzing data in spreadsheets.