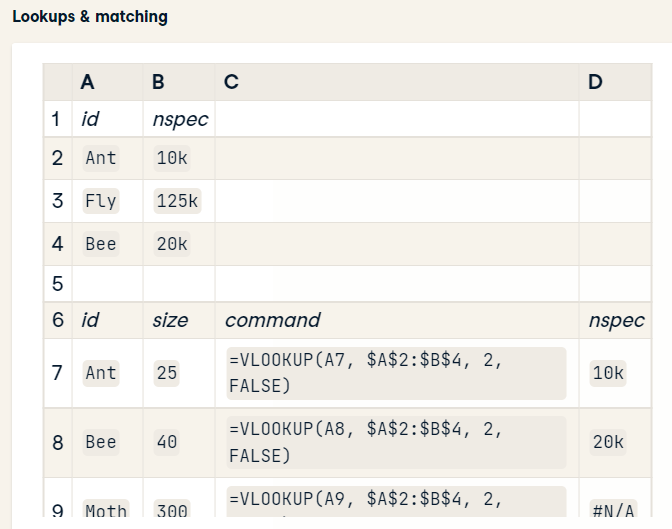
# Intermediate Spreadsheets

## 1. [1.] Lookups & matching [Aramalar ve eşleştirme]

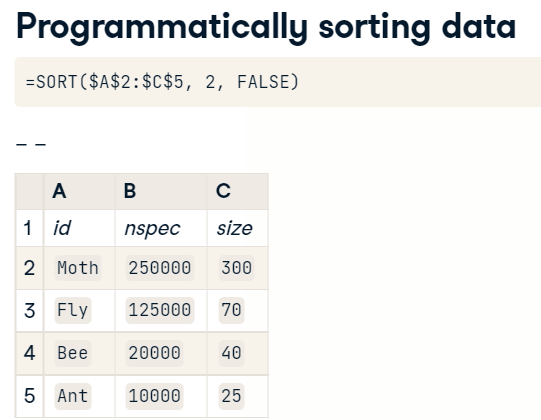
Now we're going to revisit and expand upon some of the data sorting and matching ideas from the Data Analysis with Spreadsheets course. [Şimdi, Elektronik Tablolarla Veri Analizi kursundaki bazı veri sıralama ve eşleştirme fikirlerini tekrar gözden geçirip genişleteceğiz.]

## 2. [2.] Left joins with VLOOKUP() [VLOOKUP() ile sola birleşimler]



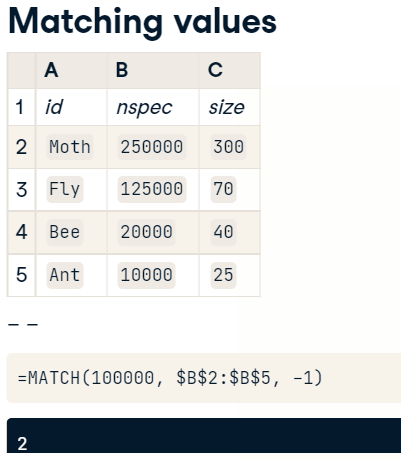
Throughout the DataCamp curriculum you'll repeatedly come across the idea of joining or merging two rectangular datasets together. [DataCamp müfredatı boyunca, iki dikdörtgen veri kümesini birleştirme veya birleştirme fikriyle tekrar tekrar karşılaşacaksınız.] In R this is often done with dplyr or data.table. [R'de bu genellikle dplyr veya data.table ile yapılır.] In Python you often use Pandas. [Python'da genellikle Pandalar kullanırsınız.] Spreadsheets have the ability to perform what's known as a left join using the VLOOKUP() function. [E-tablolar, VLOOKUP() işlevini kullanarak sol birleştirme olarak bilinen şeyi gerçekleştirme yeteneğine sahiptir.] This is a little complex, and the function takes four arguments. [Bu biraz karmaşıktır ve fonksiyon dört argüman alır.] Each of the datasets has to have an I-dot-D-dot. column with some common values. [Veri kümelerinin her birinin bir I-nokta-D-noktasına sahip olması gerekir. bazı ortak değerlere sahip sütun.] The first argument to VLOOKUP() is the value that you want to match. [DÜŞEYARA() işlevinin ilk argümanı, eşleştirmek istediğiniz değerdir.] Let's try matching Ant. [Karınca eşleştirmeyi deneyelim.] That's in cell A7 in the table we're merging into, so that's the first argument. [Bu, birleştirdiğimiz tabloda A7 hücresinde, yani ilk argüman bu.] The second argument is the range of the dataset that we want to find values in, given as absolute addresses. [İkinci argüman, mutlak adresler olarak verilen, değerlerini bulmak istediğimiz veri kümesinin aralığıdır.] In this case that dataset stretches from A2 to B4. [Bu durumda veri kümesi A2'den B4'e kadar uzanır.] The third argument is the number of the column that contains the values to be merged in. [Üçüncü argüman, birleştirilecek değerleri içeren sütunun numarasıdır.] In this case, we want the number of species from column B, the second column. [Bu durumda, ikinci sütun olan B sütunundaki türlerin sayısını istiyoruz.] The fourth argument is whether or not the lookup column is sorted. [Dördüncü argüman, arama sütununun sıralanıp sıralanmadığıdır.] Usually this will be FALSE. [Genellikle bu YANLIŞ olacaktır.]

## 3. [3.] Programmatically sorting data [Verileri programlı olarak sıralama]



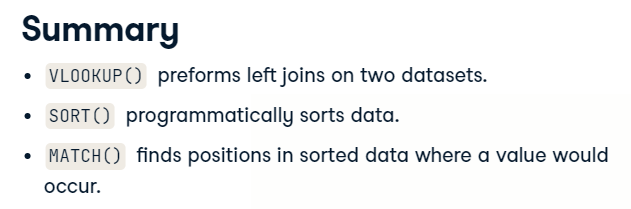
You can sort data by pointing and clicking in the menu, but if you want your analyses to be reproducible, you need to perform the sorting using code. [Menüde işaret ederek ve tıklayarak verileri sıralayabilirsiniz, ancak analizlerinizin tekrarlanabilir olmasını istiyorsanız, sıralamayı kod kullanarak yapmanız gerekir.] The SORT() function takes 3 arguments. [SORT() işlevi 3 bağımsız değişken alır.] First is the dataset, given as absolute addresses. [Birincisi, mutlak adresler olarak verilen veri kümesidir.] Second is the number of the column to sort by. [İkincisi, sıralanacak sütunun numarasıdır.] Finally, you pass TRUE to sort in ascending order and FALSE to sort in descending order. [Son olarak, artan düzende sıralamak için DOĞRU ve azalan düzende sıralamak için YANLIŞ'ı geçersiniz.] In the example you can see the insect data sorted in descending order of the number of species, in column two. [Örnekte, ikinci sütunda tür sayısına göre azalan düzende sıralanmış böcek verilerini görebilirsiniz.]

## 4. [4.] Matching values [eşleşen değerler]



MATCH() lets you find values in a sorted dataset. [MATCH(), sıralanmış bir veri kümesindeki değerleri bulmanızı sağlar.] It takes three arguments. [Üç argüman alır.] First is the value to find. [Birincisi, bulunacak değerdir.] In the example, you are looking for one hundred thousand. [Örnekte, yüz bin arıyorsunuz.] Second is the absolute address of the column of data, here from B2 to B5. [İkincisi, burada B2'den B5'e kadar olan veri sütununun mutlak adresidir.] Finally, you pass one if the data is sorted in ascending order or minus one if the data is sorted in descending order. [Son olarak, veriler artan düzende sıralanmışsa bir, azalan düzende sıralanmışsa eksi bir iletebilirsiniz.] In this case, one hundred thousand would occur after the 2nd value, so MATCH() returns two. [Bu durumda, 2. değerden sonra yüz bin gelir, dolayısıyla MATCH() iki döndürür.]

## 5. [5.] Summary [özet]



You saw three functions in this video. [Bu videoda üç fonksiyon gördünüz.] VLOOKUP() is the spreadsheet equivalent of a database left join. [DÜŞEYARA(), bırakılan bir veritabanının elektronik tablo eşdeğeridir.] The SORT() function does exactly what you might expect: it sorts datasets from smallest to largest or largest to smallest. [SORT() işlevi tam olarak beklediğiniz şeyi yapar: veri kümelerini en küçükten en büyüğe veya en büyükten en küçüğe sıralar.] Finally, MATCH() finds positions in sorted data where a value would occur. [Son olarak, MATCH(), sıralanmış verilerde bir değerin oluşacağı konumları bulur.]

## 6. [6.] Let's practice! [Hadi pratik yapalım!]

Time to find your perfect match. [Mükemmel eşinizi bulma zamanı.]

#### (1) A VLOOKUP refresher

In [**Data Analysis with Spreadsheets**](https://www.datacamp.com/courses/data-analysis-with-spreadsheets) you learned how to use [**VLOOKUP()**](https://support.google.com/docs/answer/3093318). They are important, so let's refresh your memory.

VLOOKUP() is like INDEX() in that it lets you look up values within a data block. It has the advantage that rather than you having to look and find the position of the cell to specify manually, you can specify the value you are looking for, and it automatically finds it. VLOOKUP() takes four arguments.

* First is the value that you are looking for, usually a string.
* Second is the data range, usually specified using absolute coordinates. The first column must contain the lookup values.
* Third is the column offset, the same as with INDEX().
* Fourth is whether or not the data is sorted by the lookup column. Usually you need to specify FALSE here.

##### Instructions

Find the same values from the INDEX() exercise, now using VLOOKUP(). Your data range is from $B$11 to $G$19.

* In cell I1, get the number of Blues in Sikkim. Call VLOOKUP(), passing "Sikkim", the data range, an offset of 5, and FALSE.
* In cell I2, get the number of White-Yellows in Mizoram Hills.

#### (2) Sorted!

It is often easier to make sense of datasets when the rows are ordered in some way, for example, when the values in a column go from smallest to largest (or largest to smallest). You can sort datasets using the [**SORT()**](https://support.google.com/docs/answer/3093150) function.

The first argument to SORT() is the range of the dataset, not including the header row. The second argument is the number of the column to sort on, starting with 1 as the left-most column. Thirdly, you set the sort direction: pass TRUE to sort in ascending order (low to high, or A to Z) and FALSE for descending order. If you want to break ties by sorting on further columns, you can pass another column index and another direction for each column that you want to sort with.

For example, to sort the dataset in range $D$2 to $H$100 by decreasing F values then increasing D values, you would write SORT($D$2:$H$100, 3, FALSE, 1, TRUE).

##### Instructions

In cell I2, write a command to sort the butterfly dataset by decreasing number of Skippers.

#### (3) Matching values

The [**MATCH()**](https://support.google.com/docs/answer/3093378) function let's you find the position of cells that match a particular criterion. It's a little tricky, so bear with this.

It works best when the data is already sorted (that's the case we'll consider here). The first argument is the limit value, the second argument is the data range, and the third argument is 1 if the column is sorted in ascending order and -1 for descending.

If a column of data, A2:A100 was sorted in ascending order, MATCH(1000, A2:A100, 1) would find the position of the largest value in A2:A100 that was less than or equal to 1000.

If B2:B100 was sorted in descending order, MATCH(1000, B2:B100, -1) would find the position of the smallest value in B2:B100 that was greater than or equal to 1000.

The dataset has been sorted by descending number of Skippers.

##### Instructions

* In cell I1, MATCH() the position of the smallest number of Skippers greater than or equal to 100. The data range is C2 to C45.
* In cell I2, get the ADDRESS() of that cell. The row is the match position plus one (for the header row), and it's the third column.
* In cell I3, get the value in that cell to find the smallest number of Skippers greater than 100.