Introduction to databases

Do you know exactly what data is? Data is any kind of information around us! Information is facts or details about somebody or something and which you know or can learn about from different sources.

Let's explore some different types of data:

- In the fall, most trees have yellow leaves.
- Kim's eyes are blue.
- A car has 4 wheels.
- Today is Sunday.
- Anna's favorite food is spaghetti.

When a set of data is gathered, it can provide organized information about something specific! For example, the data a school gathers about its students makes up an organized set of information. Before computers, this type of data was usually kept in file cabinets.

People who use computers often use the words "information" and "data" in the same way. In the 1970s, computer scientists gave a new meaning to these words. "Data" means information that has not been checked. "Information" means data that has been checked and therefore it can be trusted.

What is a database?

A collection of data on a computer is called a database. Today, school data is no longer stored in huge file cabinets; instead, every school has an electronic database with all its students' "cards." The database is stored on a computer and is only a few megabytes in size! Also, each card looks the same as it did in the past, but now it is on the screen of a computer.

A database is a system for organizing data. It is a collection of raw data that can be changed, sorted, and quickly searched to show detailed information about something in particular. You can use database programs to manage electronic databases. A very simple example of a database is an electronic address book, which can include information about thousands of people.

For example, a school database might have one table including its students' information and another table including its teachers' information.

a database table is a small database of similar items. A database is organized into one or more tables.

A record in a database table is an item of information with some characteristics. A database table is a collection of records.

Every characteristic, or piece of information, is called a field. A field has a name and some data.

Create a database

Now it's time to create your own database. The first thing you have to do is decide what kind of information you want to include. Think of the characteristics you want to collect for each record and give them a title or a field name.

Although there are various programs for database management, you can use Microsoft Excel to create your database table in a very simple way!

Let's create an address book database of your friends' information. The field names are:

Name, Telephone, Home Address, Email Address, Birthday, Hobby.

To create a table:

- Open an empty spreadsheet.
- Type the database field names horizontally in different cells (from Al to Fl).
- Select Al to Fl.
- Click the Bold button Oto make the titles stand out.
- Now add one record for each of our friends. Every record must have information about the six different fields shown here.

HISTORY

Edgar F. "Ted" Codd was an English computer scientist who invented the relational model for database management in 1970 while working for IBM. His theory is the basis for relational databases and data management.

Format as Table

Now that your database information is ready, format it with a style of your choice to make Microsoft Excel understands it's a data table.

To apply a table style:

- Select your table cells from Al to F7.
- On the Home tab, in the Styles group, click Format as Table.
- Select a style you like! Let's select one in the Medium group.
- In the Format As Table window, select My table has headers.
- Click0K.
- Voila! Your table has a new style and the program knows that the titles are field headers.

If you want to add another friend later, you can simply start typing the new information on the first empty row under the data and Microsoft Excel will recognize it as a new record and include it in the table!

All the information in a database must relate to the same topic. For example, you cannot have information about sports in an animal database!

Filter and Sort

Apply a filter

To see a specific set of records (data), apply a filter. This is useful especially if you have really big tables and you only want to see a part of your data.

To apply a filter:

- Click the arrow button next to a field header, for example Name.
- Select only the names you want to display.
- This is called filtering. You have just applied a filter to your database table based on the content of one field.

You can Select All records by using the same steps and everything will be displayed as It was before.

Filter and Sort

If you have a lot of records, it's a good idea to sort them because this way it will be easier to find what you are looking for. You can put the data in alphabetical order for text fields or start from the smallest to the largest (and vice versa) for numbers.

To sort your data:

- Click the down arrow on the Name header.
- Click Sort A to Z to sort the table records alphabetically.
- All records will automatically change position in the table and will now be sorted based on the Name field. The arrow button of the field header will also change to show that the table is displayed in a specific order.

Multi-level sorting

Sometimes it's useful to sort your database data in multiple fields instead of just one. This is called multi-level sorting and it works like this:

Let's say you want to sort your data alphabetically by hobby and then by name. This just means that all your records will be sorted by hobby and if some of those records happen to have the same hobby, these will be sorted again by the name field.

To apply a multi-level sorting:

- If you have already sorted any field, such as the name field we described in the previous example, click the Undo button to work on the unsorted data.
- Select your table cells from Al to F7.
- On the Home tab, in the Editing group, click Sort & Filter and then Custom Sort.
- In the Sort by list, click Hobby.
- Click Add Level to add a second level of sorting to your data. A new row will appear.
- In the Then by list, click Name.
- Click OK.
- All records will be sorted based on the Hobby field and then based on the Name field.

Sorting Is a very good method to organize your data In any program that supports It. It's much easier to find information if it is in a certain order, rather than appearing randomly.

Custom Filters

Now let's see some more advanced filters. Let's say we want to display the records of kids whose birthday is in April.

To apply a filter:

- Click the arrow button next to a field header, in this example Birthday.
- Click Text Filters and then click Custom Filter.
- In the Birthday list click contains and in the text box next to it, type April.
- Click0K.
- As a result, only those records whose Birthday field contains the word April are displayed.

Notice the options in the drop-down list. The filter names are self-explanatory.

You can try out different filters to see what effect they have. You can also apply another filter in another field of your database right after the last one, to filter your records even more.

Keys and relationships

When working with databases, we always want things that are unique in real life to be represented uniquely in our database as well. For example, we know that each person is unique. Even when two people have the same name, they are still two different people. So, in a database with information about people, for example our friends, we want each record to describe a unique person.

Take a look at the database below:

Name	Phone Number	Home Address	E-mail Address	Birthday
Tom	212 500 2020	36 Cambridge Court	tom@digital-kids.com	2 April
Tom	212 500 4040	22 Alfred Drive	tomas@digital-kids.com	23 May

Here, two friends happen to have the same name. How do we understand that they are two different people? Of course, we look at their other details to tell them apart.

Now take a look at the database below:

Name	Phone Number	Home Address
Tom	212 500 2020	36 Cambridge Court
Tom	212 500 2020	36 Cambridge Court

What is going on here? When you see two identical records like these in a database, you may think that there is mistake and a record got duplicated. But how can, you be sure? For all you know, these two records could be for two different people, two roommates who happen to have the same name and address. So, now that we have seen how important it is to be able to have unique records in our databases, let's see how we can fix it so we don't come across any confusing records like the roommates in the previous example.

Primary key

To distinguish between records in a database table, so that *every* record is unique and uniquely identified, we use what is known in database lingo as a **primary key.**
A primary key is one or more fields that can uniquely identify each re cord rn-our database table. For example, if we know that there can't be two people with the same email in our friends' database, then we can use the field **Email Address** as the primary key. Similarly, if we are absolutely certain that the combination of name and email address is unique for our friends, then the fields **Name** and **Email Address** can be the primary key.

The choice of the primary key depends on the data and the various restrictions we know to be true Later on, we will see how the primary key is used to correlate records between different tables of our database.

ID Field
As you saw in the friends example at the beginning of the task, there are times when you can't find a good primary key even if you include all the fields. Sometimes, you must include so many fields in your primary key for it to be unique that it gets too big and you end up wasting lots of space. So what do you do in such cases? Well, if you can't find a primary key from your existing fields, you can create a new field for this purpose! We usually refer to that new field as the ID field. ID comes from the word "identity," a very fitting name, since this field identifies each record.

The ID field can contain any data we want but it must be unique for each record. The most common practice is to use incremental numbers.

When choosing a primary key, it's smart to always choose the smallest set of fields that can uniquely identify our records. We don't want to use extra fields if we don't have to!

Table relationships

Okay, let's sum up: you can choose a field from a table and call it the primary key and sometimes you create a new field called "ID" and use that as a primary key instead. But why do we go to all this trouble? Well, the primary key is actually very useful when you want to relate two tables. For example, let's say that your database consists of two tables. One table has information about the students of a school and the other has information about the books in the school library.

Student_ID	Name	Phone Number	Home Address	E-mail Address	Birthday
1	Tom	212 500 2020	36 Cambridge Court	tom@digital-kids.com	2 April
2	Tom	212 500 2020	36 Cambridge Court	tomas@digital-kids.com	2 April
3	Alex	212 500 5162	202 Newport Lane	alex@digital-kids.com	2 April
4	Stella	212 500 1234	2048 Central Avenue	stella@digital-kids.com	23 June

Book_ID	Title	Author	Date Published
1	A Tale of Two Cities	Charles Dickens	1859
2	The Little Prince	Antoine de Saint-Exupéry	1943

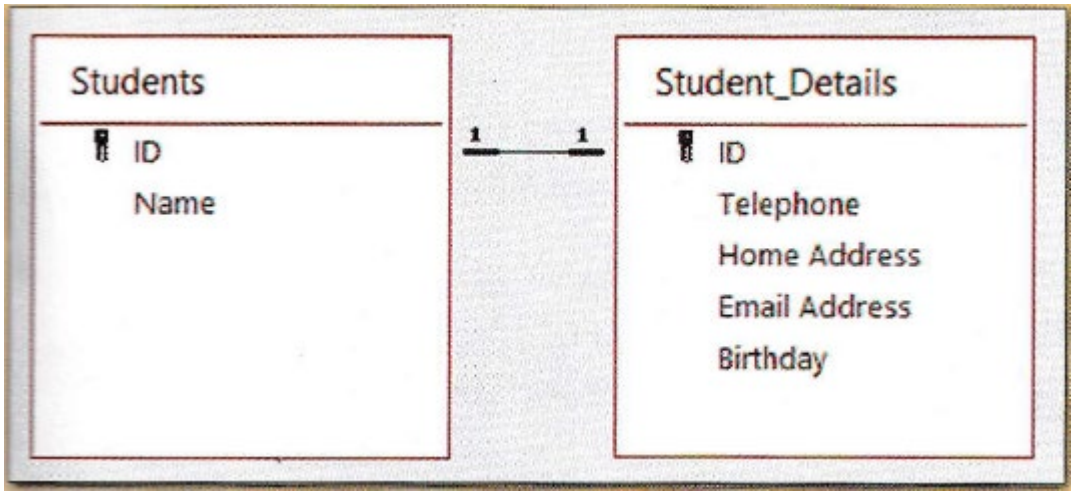
Imagine you want to create a third table with information about students who have borrowed books from the school library. To relate a student with a book, we simply need to use their respective primary keys and copy them to a new table along with any other information that might be relevant, as in the example table below.

Student_ID	Book_ID	Date Borrowed	Date Returned
1	1	04/22/2023	04/28/2023
4	2	04/23/2023	04/29/2023
3	1	05/13/2023	

From the table above we can see, for example, that Stella borrowed The Little Prince on 04/23/2023 and returned it on 04/29/2023. From this example, you can also see why it is so important for the primary key to be as small as possible. Since you copy the primary keys from each table to a new table in order to relate them, imagine what it would be like if, instead of using ID fields as primary keys, you copied all the information of the student and the book into a new record every time a student borrowed a book. It would be a complete waste of space.

The one to-one and one-to-many relationships

In a one-to-one relationship, each record in a database table is linked to only one other record in another table. In a one-to-one relationship. between Table A and Table B, each record in Table A is linked to another record in Table B. The number of records in Table A must equal the number of records in Table B.



The easiest way to create a one-to-one relationship is to divide an existing table of our database into two tables.

Student_ID	Name	Phone Number	Home Address	Email Address	Birthday
1	Tom	212 500 2020	36 Cambridge Court	tom@digital-kids.com	2 April
2	Tom	212 500 2020	36 Cambridge Court	tomas@digital-kids.com	2 April
3	Alex	212 500 5162	202 Newport Lane	alex@digital-kids.com	2 April
4	Stella	212 500 1234	2048 Central Avenue	stella@digital-kids.com	23 June

Student_ID	Name
1	Tom
2	Tom
3	Alex
4	Stella

This type of relationship is rarely used in real-life databases. However it can be used in certain cases to improve a process. For example, if we know that we are often going to search for a student's name but rarely for their other information, it makes sense to separate the data into two tables to do the search for names on a smaller table

In Microsoft access you don't need to specify the type of relationship you need. The program will do it for you, but it’s very important to understand how the tables are related to one another.
The solution of the lD field is so simple and elegant that It is very common to use it even when you can actually choose a primary key from the existing fields, because it ls always only one numerical field. Also, most real databases work faster with numbers than with text.

In a one-to-many relationship, each record oi a table can be related to many records of another table. This allows frequently used information to be saved only once in a table and referenced many times in all other tables. In a one-to-many relationship between Table A and Table B, each record in Table A is linked to 0, 1 or more records in Table B. The number of records in Table A is almost always smaller than the number of records in Table B.

Examples of one-to-many relationships can be:
> A mother with her children. A child will have only 1 (biological) mother. A mother will have zero, one, or many children.
> A car and its parts. A car has many parts. Each part can only belong to one car.
> Houses on a street. A street will have zero, one, or many houses. Each house will be on one street.

Lastly, there is also the many-to-many relationship, as in the example of the students and books borrowed from the library. This kind of relationship must use an extra table to map the relations between the other tables. In the example below a student may have borrowed zero, one or many books. A book may have been borrowed by zero, one or many students (at different times of course, as the book copies are unique).

You cannot directly create a many to many relationships. You have to odd o table and make two relationships of many to one and one to many.

Contact management

Now let's see how we can export our contacts from Outlook.com. Of course, one way to do it would be to navigate through the contacts, copy the information of each one and paste it into Microsoft Excel one by one. However, this method takes a long time, especially when the contact list is large. Ortlook.com offers the option to export all contacts at the same time, in .csv file format

CSV files (Comma Separated Values) are text files used mainly to transfer data between applications.

The easiest way to do this is to export our contacts to a single CSV file. Let's see how:

To export your contacts:

- Start Outlook.com.
- Click People.
- Click Manage and then click Export contacts.
- In the Export contacts window, click Export.
- The file with your contacts will be saved on your computer automatically.
- Click Options and then click Show in folder.
- In the Downloads folder, you can view the CSV file with your contacts.

Export only useful information. Otherwise, when you import your data into a database table, you will end up with a large collection of data you don't even need. This will slow down and complicate things.

This is a good way to back up your contacts. If something happens to your email account or your devices, you won't lose your contacts of your friends.

To import your contacts:

- Start Microsoft Excel
- On the Data tab, in the Get & Transform Data group, click From Text/CSV
- Find and select the CSV file that you created.
- The window of the file you -have selected will appear with all the data it contains in the Delimiter list, you can select the delimiter that, was used in your file. For example, choose Comma
- If you want to change some information, you can choose Transform Data or if the Data appear correct you, can choose Load.
- In the Power Query Editor, window, you can modify the data. For example, click the arow next to table icon and then click Use First Row as Headers
- Also, you can delete the columns you don't need. To do this, right-click a column and click Remove.
- When you're done, click Close & Load.
- Your contact records have been successfully imported into Microsoft Excel.

You can insert your table into an existing or new worksheet

Lab Data Collection

A data logger can take repeated measurements and then transfer the data to your computer for analysis. Just connect your data logger to the USB port of the computer and download the recorded data.

With a data logger, you can capture and display data as it happens, collect data continuously or at regular intervals over long periods of time. One kind of data logger is a weather station, which records everything about the current weather conditions. Scientists can then use this data to predict changes in the weather.

Using data logging software, you can display the data as a chart.

What can you do with a data logger

- Accurately measure sound and light as well as temperature.
- Take measurements that are too 'slow or too fast to record by traditional means.
- Collect data anywhere and then download the data to a computer for analysis.
- Get visual feedback of your measurements.