How to create a find duplicates query

A **find duplicates query** allows you to search for and identify **duplicate records** within a table or tables. A duplicate record is a record that refers to the **same thing** or **person** as another record.

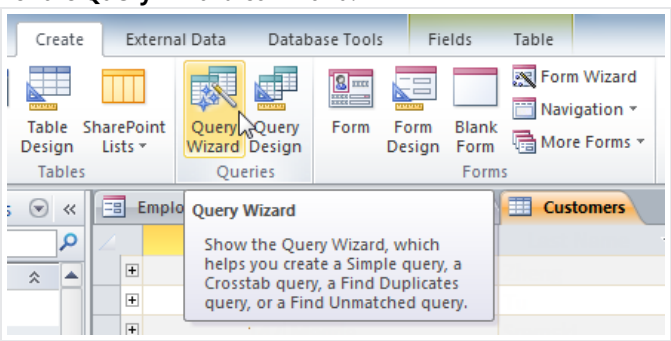
Not all records containing similar information are duplicates. For instance, records of two orders that were placed on different dates but that contained identical items would **not** be duplicate records. Likewise, not all duplicate records contain completely identical information. For example, two customer records could refer to the same person but include different addresses. The record with the out-of-date address would be the duplicate record.

Why is getting rid of duplicate records so important? Consider the example above. If we had multiple records for one customer, it would be difficult to view an order history for him since that information would be spread across many unlinked records. We might even deliver his order to the wrong address if the person entering the order information selects an outdated record. It's easy to see how having duplicate records can undermine the integrity and usefulness of your database.

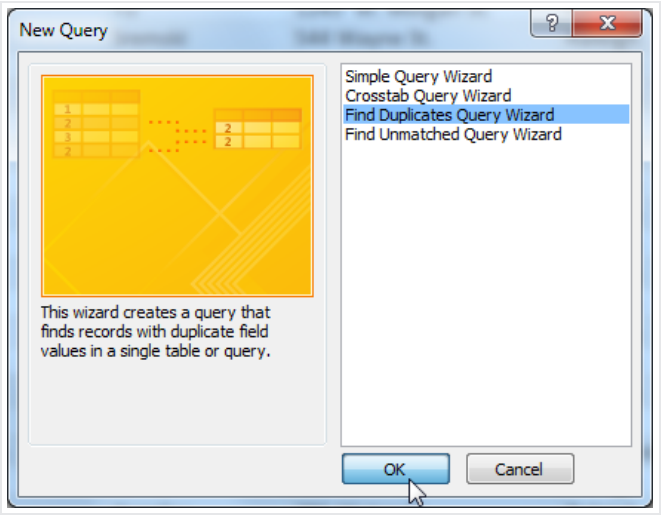
Fortunately, Access makes it easy to search for and locate potential duplicate records. Note that Access won't delete the records for you or help you figure out which one is current—you'll have to do these things for yourself. If you're familiar with the data in your database, though, getting rid of duplicate records will be a manageable task.

To create a find duplicates query:

1. Select the **Create** tab on the Ribbon, and locate the **Queries** group.
2. Click the **Query Wizard** command.

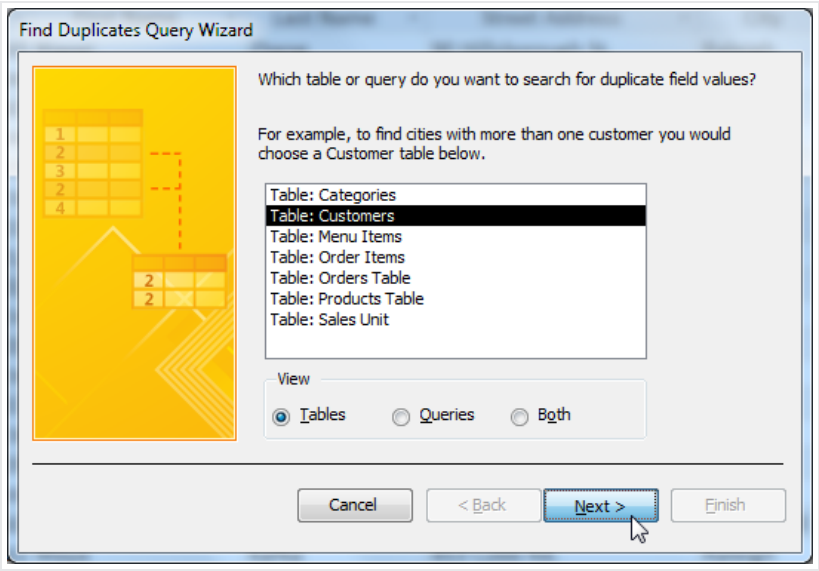


3. The **New Query** dialog box will appear. **Select** Find Duplicates Query from the list of queries, then click **OK**.

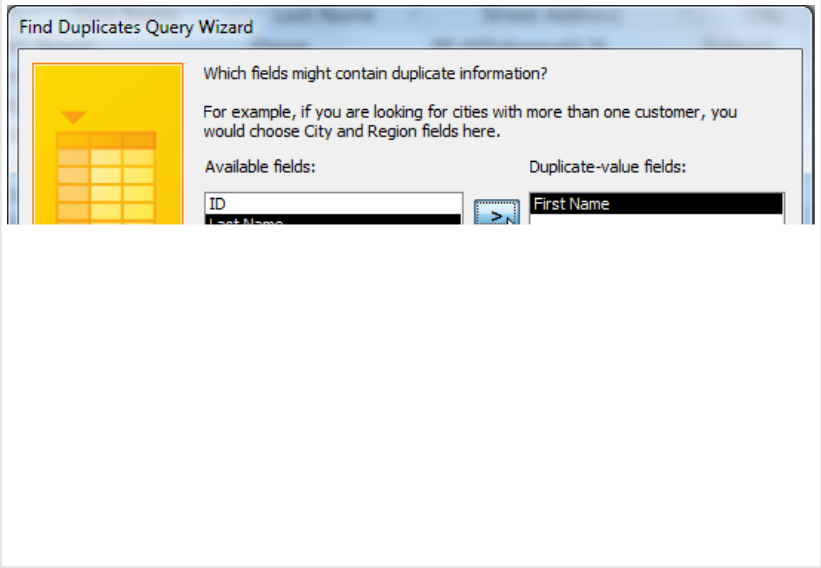


4. **Select** the table you want to search for duplicate records, then click **Next**. We're searching for duplicate customer records,

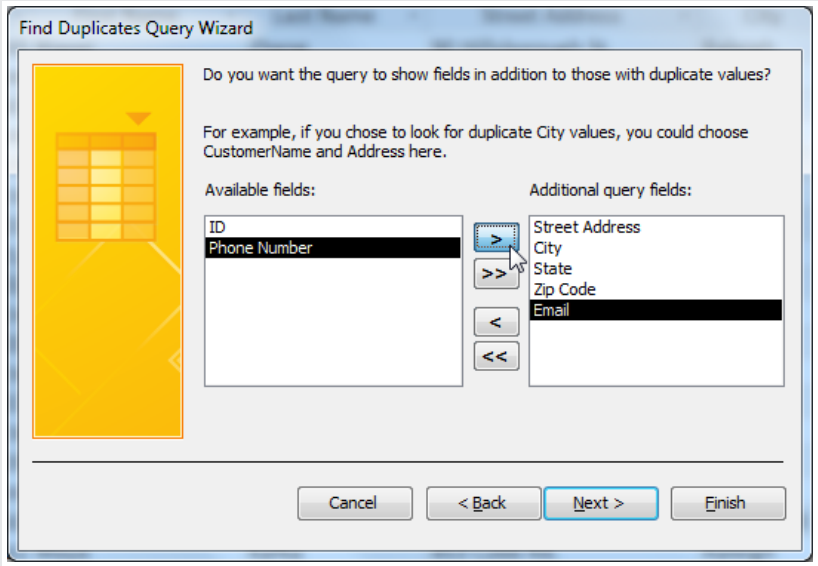
so we'll select the **Customers** table.



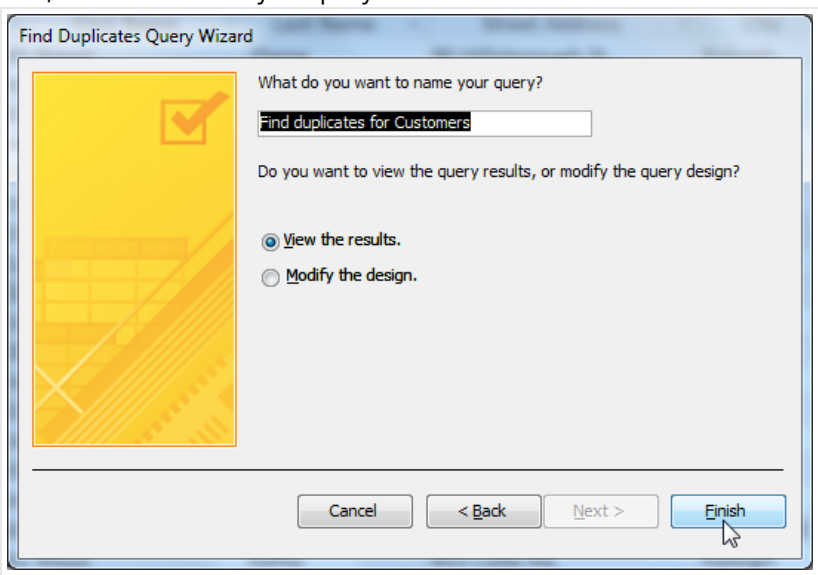
5. Choose the fields you wish to search for duplicate information by selecting them, then clicking the **right arrow button** > . Only select fields that should not be identical in nonduplicate records. For instance, since we're searching for duplicate customers we'll only select the **First Name** and **Last Name** fields, as it's unlikely that multiple people with the exact same first and last name would place orders at our bakery. When you've added the desired fields, click **Next**.



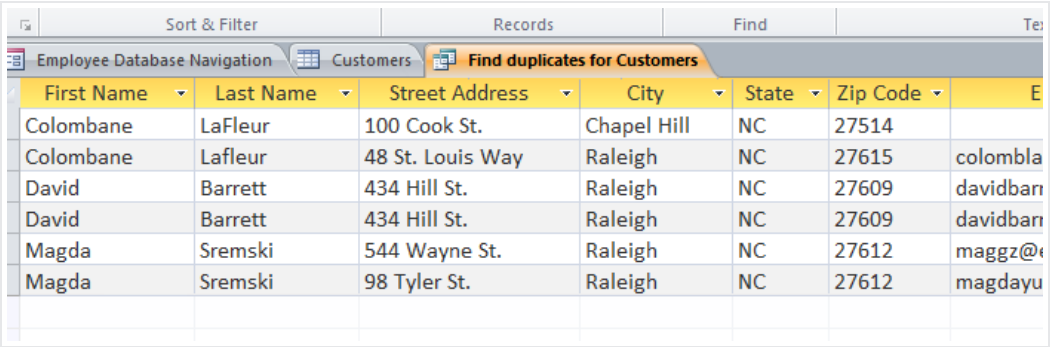
6. Select additional fields to view in your query results. Choose fields that will help you distinguish between the duplicate records and choose which one you want to keep. In our example, we'll add all of the fields relating to customer **addresses**—plus the **email address** and **phone number** fields—as records with identical customer names that might contain nonidentical information in these fields. When you're satisfied with the fields you've chosen, click **Next**.



7. Access will suggest a name for your query, but you can type in a different name. When you're satisfied with the query name, click **Finish** to run your query.



8. If Access found any duplicate records in your query, they will be displayed in the **query results**. **Review** the records, and **delete** any outdated or incorrect records as needed.



First Name	Last Name	Street Address	City	State	Zip Code	E
Colombane	LaFleur	100 Cook St.	Chapel Hill	NC	27514	
Colombane	Lafleur	48 St. Louis Way	Raleigh	NC	27615	colombla
David	Barrett	434 Hill St.	Raleigh	NC	27609	davidbar
David	Barrett	434 Hill St.	Raleigh	NC	27609	davidbar
Magda	Sremski	544 Wayne St.	Raleigh	NC	27612	maggz@
Magda	Sremski	98 Tyler St.	Raleigh	NC	27612	magdayu

Tips for resolving duplicate records

> **Save** your duplicate records queries, and run them often.

> **Investigate** potential duplicate records by looking at linked data in other tables. You can do this by searching for these records' **record ID numbers** in related tables. Is one record linked to mostly old orders while another contains recent

> ones? The latter is likely to be the current one.

> Once you decide which record to delete, make sure you won't be losing any information you might need. In our example, before we deleted our duplicate record, we found all of the orders linked to that record's **ID number** and replaced it with the ID number of the record we decided to keep.