**Data preparation**

1.

* Open the workbook called 1\_1\_data\_preparation.xlsx from the Workbooks folder.
* Open the [**metadata sheet**](https://assets.datacamp.com/production/repositories/6386/datasets/0d84b751e28911f4a2c51b1a38c0100a55d8037e/Metadata%20Sheet%20-%20Customer%20Churn.pdf) in a separate page as this may come in handy throughout the course.

**Hint**

* The *File* menu can be found in the ribbon at the top of the workbook.
* To open a workbook, click on *Open* in the *File* menu. If you click on *Browse*, you can navigate to the Workbooks folder and open the 1\_1\_data\_preparation.xlsx file.

2.

* Navigate to Databel - Aggregate and convert the range of data ($A$1:$U$6670) into a tabular format.
* Rename this table "Aggregate".

**Hint**

* Highlight all the data to be included in the table: $A$1:$U$6670.
* Under the *Insert* tab in the ribbon, select *Table* .
* Ensure "My table has headers" is checked and click *OK*.
* Type in "Aggregate" for the *Table Name* under the *Table Design* ribbon.

3.

* Navigate to Databel - Customer and convert the range of data ($A$1:$AC$6688) into a tabular format.
* Rename this table "Customers".

**Hint**

* Highlight all the data to be included in the table: $A$1:$AC$6688.
* Under the *Insert* tab in the ribbon, select *Table*.
* Ensure "My table has headers" is checked and click *OK*.
* Type in "Customers" for the *Table Name* under the *Table Design* ribbon.

4.

Identify whether there are any duplicate values in the Customers table, there are two approaches that you can use:

* Remove duplicates feature
* Conditional formatting

**Hint**

**Remove duplicates feature**

* Highlight Customer ID in the Customers table.
* Click on *Data* in the menu at the top, then click on *Remove Duplicates*.
* Make sure all columns are selected in the pop-up window. Ensure that the *My list has headers* option is also selected. Click *OK*.

**Conditional formatting**

* With the Customer ID column selected, click on *Conditional Formatting > Highlight Cells Rules > Duplicate Values…*.
* In the *New Formatting Rule* window, leave all options as per default settings, with the formatting type being *Light Red Fill with Dark Red Text*.

5.

**Does our dataset contain any duplicate values?**

* Yes
* No

**Hint**

There are multiple ways to check for duplicate values in Excel including the **remove duplicates feature** and **applying conditional formatting**.

**Remove duplicates feature**

* Highlight Customer ID in the Customers table.
* Click on *Data* in the menu at the top, then click on *Remove Duplicates*.
* Make sure all columns are selected in the pop-up window. Ensure that the *My list has headers* option is also selected. Click *OK*.

**Conditional formatting**

* With the Customer ID column selected, click on *Conditional Formatting > Highlight Cells Rules > Duplicate Values…*.
* In the *New Formatting Rule* window, leave all options as per default settings, with the formatting type being *Light Red Fill with Dark Red Text*.

*If you're still stuck, review the solution in 1\_2\_calculating\_churn.xlsx from the Workbooks folder.*

**Calculating churn**

1.

Create a new column "Churned" in our Customers table that uses an IF() to convert the values in Churn Label based on the following criteria:

* "Yes" then 1
* "No" then O

**Hint**

* In column AD you should have a header "Churned" with the following formula: =IF([@[Churn Label]]="Yes", 1, 0)

2.

Create a blank *PivotTable* of the Customers table and place it in a new *Worksheet*. Rename this worksheet "Customer Pivots".

**Hint**

* Select any cell in the Customers table and click *Insert* then *PivotTable*.
* Ensure that the *Table/Range* input is Customers.
* Check *New Worksheet* and click *OK*.
* To rename a worksheet, double-click on it and type in the updated name.

3.

* In the *PivotTable*, display the total count of customers and number of churned customers.
* Update the column headers to user-friendly names such as "Total Customers" and "Churned Customers".

**Hint**

* Drag Customer ID and Churned to the *Values* section.
* Customer ID should be aggregated as a *Count*.
* Churned should be aggregated as a *Sum*.

4.

Great! Now we can easily identify how many customers have churned, but what if we want to find out what our churn rate is?

* Next to your *PivotTable*, create a new calculation with the header "Churn Rate" that divides churned customers by total customers.
* Format this as a % to two decimal places.

**Hint**

Your formula should look something like this:

* =GETPIVOTDATA("Churned Customers",$A$3)/GETPIVOTDATA("Total Customers",$A$3)
* B4/A4

To format a cell with decimal places:

* Under *Home* > *Number* there is a dropdown where you can change the format. Select *Percentage*.
* To change the decimal places, there are two icons under the dropdown to increase and decrease decimal. Click *Increase decimal* twice.

5.

**What's the total churn rate for "Databel"? (Answer format: XX.XX%)**

**26.86%**

**Hint**

Your formulas should look like the below:

* Churned: =IF([@[Churn Label]]="Yes", 1, 0)
* Churn Rate:
  + =GETPIVOTDATA("Churned Customers",$A$3)/GETPIVOTDATA("Total Customers",$A$3)
  + B4/A4

In your PivotTable, you should have:

* Columns: Values
* Values: Total Customer, Churned Customers

*If you're still stuck, review the solution in 1\_3\_investigating\_churn.xlsx from the Workbooks folder.*

**Investigating churn reasons**

1.

Create a blank *PivotTable* of the Customers table in the Customer Pivots worksheet.

**Hint**

* Select any cell in the Customers table and click *Insert* then *PivotTable*.
* Ensure that the *Table/Range* input is Customers.
* Check *Existing Worksheet* then navigate to Customer Pivots and select a cell (i.e Customer Pivots!$A$8) and click *OK*.

2.

* Analyze the total number of churned customers by Churn Reason.
* Rename the row header to "Churn Reason" and the column header to "Churned Customers".

**Hint**

* Drag Churn Reason into the *Rows* section.
* Drag Churned into the *Values* section. This should be aggregated as a *Sum*.

3.

* Order the churn reasons ascending, to the most popular churn reason appears at the bottom .
* Show the Churned Customers as a "% of Grand Total".

**Hint**

* To sort a *PivotTable*, right-click any value and navigate to *Sort* > *More Sort Options*.
  + Sort *Ascending (A to Z) by* Churned Customers.
* To convert *Values* in a *PivotTable* to "% of Grand Total", right-click on a cell and navigate the *Show Value As* > *% of Grand Total*.

4.

* Visualize your analysis with a *2D Bar Chart* and title it "Churn Reasons".
* Hide all field buttons on chart and delete the *Legend* .

**Hint**

* To create a *2D Bar Chart*, click anywhere in the *PivotTable* and navigate to *Insert* > *Charts* menu and click on *Insert Column or Bar Chart* button and select the *Clustered Bar* variant.
* To rename the chart, double-click on the title above the chart and type in a new name.
* To hide field buttons, right-click on the gray chart buttons and select *Hide all Field Buttons on Chart*.
* To delete a legend, right-click on the *Legend* and select *Delete*.

5.

**Which of the following is part of the top 3 churn reasons?**

* Competitor offered more data
* Price too high
* Poor expertise of phone support
* Competitor had better devices

**Hint**

In your PivotTable, you should have:

* Rows: Churn Reason
* Values: Churned Customers (as % of grand total)

*If you're still stuck, review the solution in 1\_4\_churn\_categories.xlsx from the Workbooks folder.*

**Digging deeper into churn categories**

1.

Create a blank *PivotTable* of the Customers table in the Customer Pivots worksheet.

**Hint**

* Select any cell in the Customers table and click *Insert* then *PivotTable*.
* Ensure that the *Table/Range* input is Customers.
* Check *Existing Worksheet* then navigate to Customer Pivots and select a cell (i.e Customer Pivots!$A$36) and click *OK*.

2.

* Analyze the total number of churned customers as % of grand total by Churn Category and Churn Reason.
* Rename the row header to "Churn Reason" and the column header to "Churned Customers".

**Hint**

* Drag Churn Category and Churn Reason into the *Rows* section.
* Drag Churned into the *Values* section. This should be aggregated as a *Sum*.
* To convert *Values* in a *PivotTable* to "% of grand total", right-click on a cell and navigate the *Show Value As* > *% of Grand Total*.

3.

We can see that category driving the highest % of churn is **Competitor**.

* Filter the *PivotTable* to only include this churn category.

**Hint**

* Drag Churn Category from *Rows* to *Filters* .
* Open *Filter* and de-select all categories except for **Competitor**.

4.

* Create a visualization of your choice to display the Churned Customers PivotTable you created and name it "Competitor Churn Analysis".
* Clean this visual up by removing unnecessary components such as field buttons and apply any style customizations.

**Hint**

* You can use either a bar chart, column chart, pie chart or donut chart to visualize the data.
* To hide field buttons: right-click on the gray chart buttons and select *Hide all Field Buttons on Chart*.

5.

**What % of customers churned due to "Competitor made better offer"? Rounded to two decimal places.**

**37.64%**

Hint

In your PivotTable, you should have:

* Filters: Churn Category
* Rows: Churn Reason
* Values: Churned Customers (as % of grand total)

*If you're still stuck, review the solution in 2\_1\_analyzing\_demographics.xlsx from the Workbooks folder.*