-- The data set 'drinks_cogs3.csv' consists of four columns named:
Country,
beer_servings,
spirit_servings,
wine_servings and
total_liters_of_pure_alcohol

-- Your task is to clean the data - (make sure that the numeric values don't have spaces, or extraneous characters, o's instead of 0's, etc.) and that they are numeric
-- You want to format the names of the countries so that they are capitalized properly - ie. United States v UNITED STATES
-- Find the blanks and determine if you should replace the blanks with a 0
-- Keep track of the missing values by using conditional highlighting
-- compute the total alcohol in liters:

- ○ The ABV (Alcohol by volume) for each drink (which is assumed to be constant for all countries)
- ○ The serving size for each type of drink (again assumed to be constant for all countries)
- ○ And The conversion from oz to liter for each drink, assuming that the servings are given in oz

-- for example - to calculate beer:
    Alcohol_beer for a country =
        (beer_servings for that country * beer_serving_size_in_liters * abv_beer)
-- you will need to compute the total for Alcohol_spirit and Alcohol_wine
-- Sum for each country - total_liters is the sum of the computed values

-- Answer the following questions:
1. What country has the highest consumption
2. What is the average consumption
3. How many countries are dry (with 0 consumption)

Excel functions to help you clean your data:

`=SUBSTITUTE(`

`SUBSTITUTE(text, old_text, new_text, [instance_num])`

SUBSTITUTE function

This article describes the formula syntax and usage of the **SUBSTITUTE** function in Microsoft Excel.

Description

Substitutes new_text for old_text in a text string. Use SUBSTITUTE when you want to replace specific text in a text string; use REPLACE when you want to replace any text that occurs in a specific location in a text string.

Syntax

SUBSTITUTE(text, old_text, new_text, [instance_num])

The SUBSTITUTE function syntax has the following arguments:

- **Text** Required. The text or the reference to a cell containing text for which you want to substitute characters.

- **Old_text** Required. The text you want to replace.

- **New_text** Required. The text you want to replace old_text with.

- **Instance_num** Optional. Specifies which occurrence of old_text you want to replace with new_text. If you specify instance_num, only that instance of old_text is replaced. Otherwise, every occurrence of old_text in text is changed to new_text.

TRIM function

This article describes the formula syntax and usage of the **TRIM** function in Microsoft Excel.

Description

Removes all spaces from text except for single spaces between words. Use TRIM on text that you have received from another application that may have irregular spacing.

`=trim`

`TRIM(text)`

If you want to see the code number that hides behind a character, type a character in a cell (*like the capital letter "A" in cell A1*), then select another cell and write a formula like the following

=CODE(A1)

The **CODE** function displays the **ASCII** code for the first character in a string detected in the target cell.

If we were to highlight and copy the trailing space from the value in cell **A9** and paste it into our test cell, the **CODE** function sees it as **ASCII** code **160**.

This non-breaking space is common amongst data that is copied from sources such as email readers, word processors, and web browsers.

Seeing Things from the Opposite Side

Another Excel text function is the **CHAR** function. This function takes an **ASCII** code number and represents it as its "human-friendly" version.
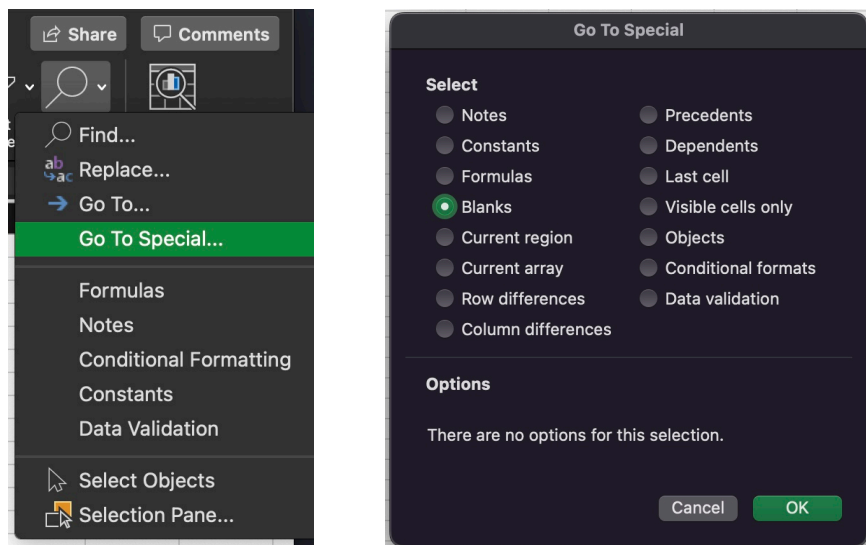
=CHAR(32) or =CHAR(160)

The above formulas will generate a traditional "space" and a non-breaking "space".

To our eyes, these both look the same, but Excel sees them as completely different characters, no different than seeing "A" and "B" as different characters.

The **TRIM** function can remove **ASCII 32** characters but not **ASCII 160** characters.

To find blanks in your dataset:

The below image uses a series of **ISTEXT** functions to test the "Product A" cells to see if the data is perceived as text by Excel. As you can see, Excel returns **TRUE** responses for all the tests.

	A	B	C	D	E	F	G
8	Plant code	Description	Product A	Product B			
9	AT309	Mueller & Co	1829	4321		TRUE	=ISTEXT(C9)
10	DE809	Bert Fond GmbH	21567	138		TRUE	=ISTEXT(C10)
11	AT430	Kind & Co	94865	423		TRUE	=ISTEXT(C11)
12	AT213	Wilfred Jay	487	556		TRUE	=ISTEXT(C12)
13			0				
14							

Another Excel text function called **VALUE** can read a number stored as text as a 'normal' number.

> **=VALUE(C9)**

Although the **VALUE** function ignores "space" characters, it fails to ignore non-breaking spaces.

If we nested the **VALUE** function within a **TRIM** function, it wouldn't solve our problem.

> **=TRIM(VALUE(C9))**

We need to use the **SUBSTITUTE** function.

> **=SUBSTITUTE(C9, CHAR(160),)**

But the result is still not a usable number because the original text is still perceived as text. We need to convert the text to a real number using the **VALUE** function.

> **=VALUE(SUBSTITUTE(C9, CHAR(160),))**

When to Use

Whenever you encounter data where...

- your numbers are not being perceived as numbers,
- you use the **TRIM** function and it fails,
- you use the **VALUE** function and it fails,

...use the **SUBSTITUTE** function to search for character code number **160** and replace it with "nothing".

=UPPER(

UPPER(**text**)

UPPER function

This article describes the formula syntax and usage of the **UPPER** function in Microsoft Excel.

Description

Converts text to uppercase.

Syntax

UPPER(text)

The UPPER function syntax has the following arguments:

- **Text**   Required. The text you want converted to uppercase. Text can be a reference or text string.

=LOWER(

LOWER(**text**)

LOWER function

This article describes the formula syntax and usage of the **LOWER** function in Microsoft Excel.

Description

Converts all uppercase letters in a text string to lowercase.

Syntax

LOWER(text)

The LOWER function syntax has the following arguments:

- **Text**   Required. The text you want to convert to lowercase. LOWER does not change characters in text that are not letters.

`=PROPER(`

PROPER(text)

PROPER function

This article describes the formula syntax and usage of the **PROPER** function in Microsoft Excel.

Description

Capitalizes the first letter in a text string and any other letters in text that follow any character other than a letter. Converts all other letters to lowercase letters.

Syntax

PROPER(text)

The PROPER function syntax has the following arguments:

- **Text**   Required. Text enclosed in quotation marks, a formula that returns text, or a reference to a cell containing the text you want to partially capitalize.