# How to clean and compile the SportsBenefitsStatementsData.xlsx and the NGBHealthDataOutputExtract.csv for easier and better Tableau integration.

Before we can use Tableau to create our visualizations and run our analysis, we must review the provided data and make sure they’re clean and in an easy format for Tableau to view and pull from. Before walking through this resource please download the provided two files: NGBHealthDataOutputExtract.csv & SportsBenefitsStatementsData.xlsx into a folder you can easily access later.

Contents

[How to clean and compile the SportsBenefitsStatementsData.xlsx and the NGBHealthDataOutputExtract.csv for easier and better Tableau integration. 1](#_Toc196379295)

[Converting NGBHealthDataOutputExtract.csv file to Excel 1](#_Toc196379296)

[Cleaning the Excel Data 4](#_Toc196379297)

[Cleaning NGBHealthDataOutputExtract 4](#_Toc196379298)

[Cleaning SportsBenefitsStatementsData 6](#_Toc196379299)

## Converting NGBHealthDataOutputExtract.csv file to Excel

Tableau allows users to import data from both a CSV file and Excel file but before importing the data, the data needs to be cleaned and the easiest way to do that is in Excel. Therefore, to convert the CSV file to an Excel file follow these steps:

1. Open a new Excel file
2. Go to the “Data” tab at the top of the screen
3. At the left of the ribbon, click “Get Data” -> “From File” -> “From Text/CSV” A screenshot of a computer

   AI-generated content may be incorrect.
4. Find the downloaded NGBHealthDataOutputExtract.csv file and click on it
5. A new window will appear, click “Load” at the bottom right. A screenshot of a computer

   AI-generated content may be incorrect.
6. After processing for a little, a table with the CSV file should appear. Save this Excel file to indicate it has NGBHealthDataOutputExtract data. This concludes the conversion from CSV to Excel.

## Cleaning the Excel Data

Before importing into Tableau, clean and reorganize both Excel files to ensure the data has the correct data types and that there are no inconsistencies or duplicates values. The best practice is to make sure all the data is reviewed and cleaned even if you might not use it later. This provides flexibility later when making visualizations.

### Cleaning NGBHealthDataOutputExtract

1. Open the NGBHealthDataOutputExtract excel file
2. Remove Duplicates: Data -> Remove Duplicates -> Select All -> Ok
3. Ensure values in each column align with column headings.
   1. Column F (Staff Bucket Def) appears to have “30-Oct” as a value but the rest of the column has ranges as values. It looks like Excel misread this value when importing, since it should be “30-100” instead.
      1. To fix this:
         1. Filter column F by clicking the arrow down button next to the “Staff Bucket Def” heading.
         2. Uncheck all the boxes except for “30-Oct”
         3. Click “Ok”
         4. A screenshot of a computer

            AI-generated content may be incorrect.
         5. This filters the table to only show rows with “30-Oct” as the value in column F
         6. Click on the cell directly under the “Staff Bucket Def” header
         7. Type “30-100” and click enter
         8. To carry this change down the column, at the bottom right of the cell, there should be a small dot, double click the dot.

A screenshot of a calendar

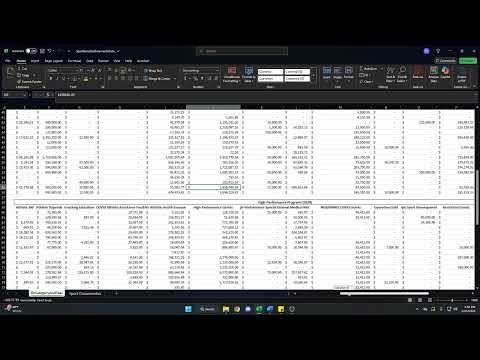
AI-generated content may be incorrect.

1. Ensure each column’s data type matches its values
   1. Columns G, M, N should all be currencies and should be shown as so.
      1. Hold Control and left click the letters above columns G, M, N, this should select them.
      2. A screenshot of a computer

         AI-generated content may be incorrect.Release control, go to the Number section in the ribbon and click the arrow down button right after “General”, select “Currency” from the drop down
2. Control + S to Save and this file’s data is all cleaned and ready to be imported into Tableau

### Cleaning SportsBenefitsStatementsData

Cleaning and reorganizing the SportsBenefitsStatementsData excel file is more complicated and complex than the previous one. Please refer to the video linked below for a step-by-step walkthrough of the cleaning process.

[](https://www.youtube.com/embed/kQd1KUVB4Ww?feature=oembed)

Once all the data is cleaned and formatted correctly, you can import them into Tableau for your analysis or to create your visualizations.