

## Tutorial 1

|   |    |
|---|----|
| A little about SAS and the libraries .....        | 1  |
| Before you start .....                            | 1  |
| How to start your work in SAS .....               | 2  |
| How to create a new project .....                 | 2  |
| How to run a libname statement.....               | 4  |
| How to import a dataset.....                      | 5  |
| How to create a line plot?.....                   | 11 |
| A) The Line Plot Wizard:.....                     | 11 |
| B) The Line plot .....                            | 15 |
| How to use “Modify task” for your advantage ..... | 18 |
| How to create a histogram .....                   | 20 |
| How to transform variables .....                  | 24 |

### A little about SAS and the libraries

SAS works with datasets in its format. You can import Excel files, but then SAS will convert them into its own format. SAS saves its files in libraries. There are two kinds of libraries:

- temporary, i.e. every time you finish your session with SAS and leave it, the file will be erased. It's called WORK library.
- permanent, i.e. every time you finish your session with SAS and leave it, the file will be accessible even after. You can start a new session and access the file; you can copy the file and send it by email etc. Usually, you want to save your files in the permanent library. That's why you will start your **EVERY SESSION** with a libname statement.

What is a library? It is basically a place on your computer (a folder of your choice) where you want SAS to store your files. It doesn't matter which folder you choose, but you should remember where that folder is so that you could always access the file.

### Before you start

Create a folder on your computer and place there all the datasets you are going to be using during this course. You can keep placing datasets there in any format – SAS or Excel. That will be your ***library***, which SAS will refer to every time you want to open or save a dataset (in its native format).

*This tutorial is based on a bike rentals dataset where we will be trying to predict a number of rides.*

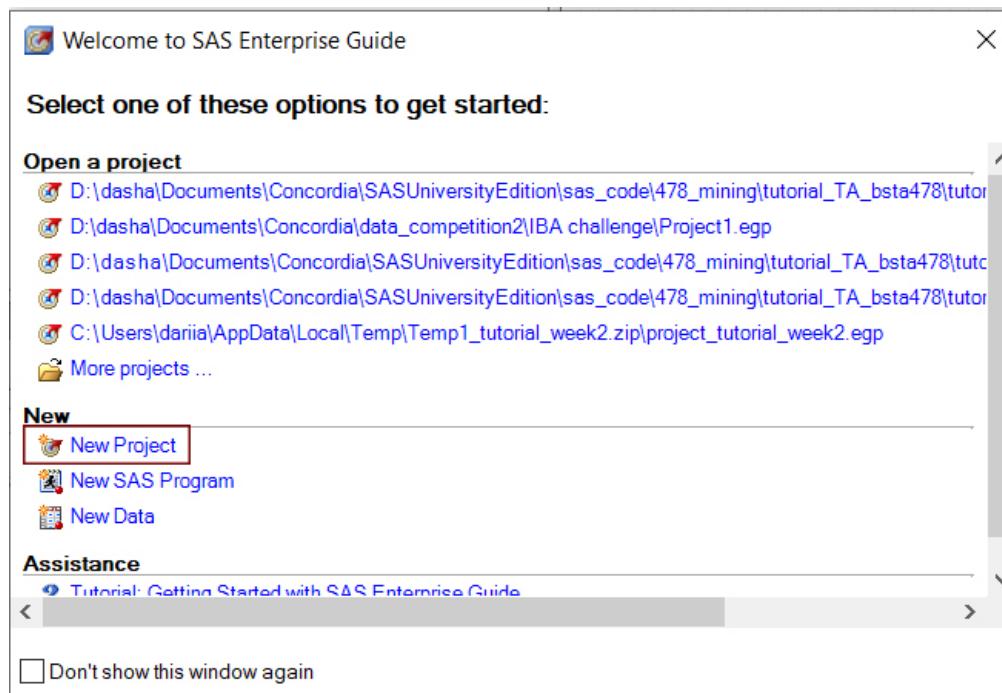
### SAS EG: CRASH COURSE ☺

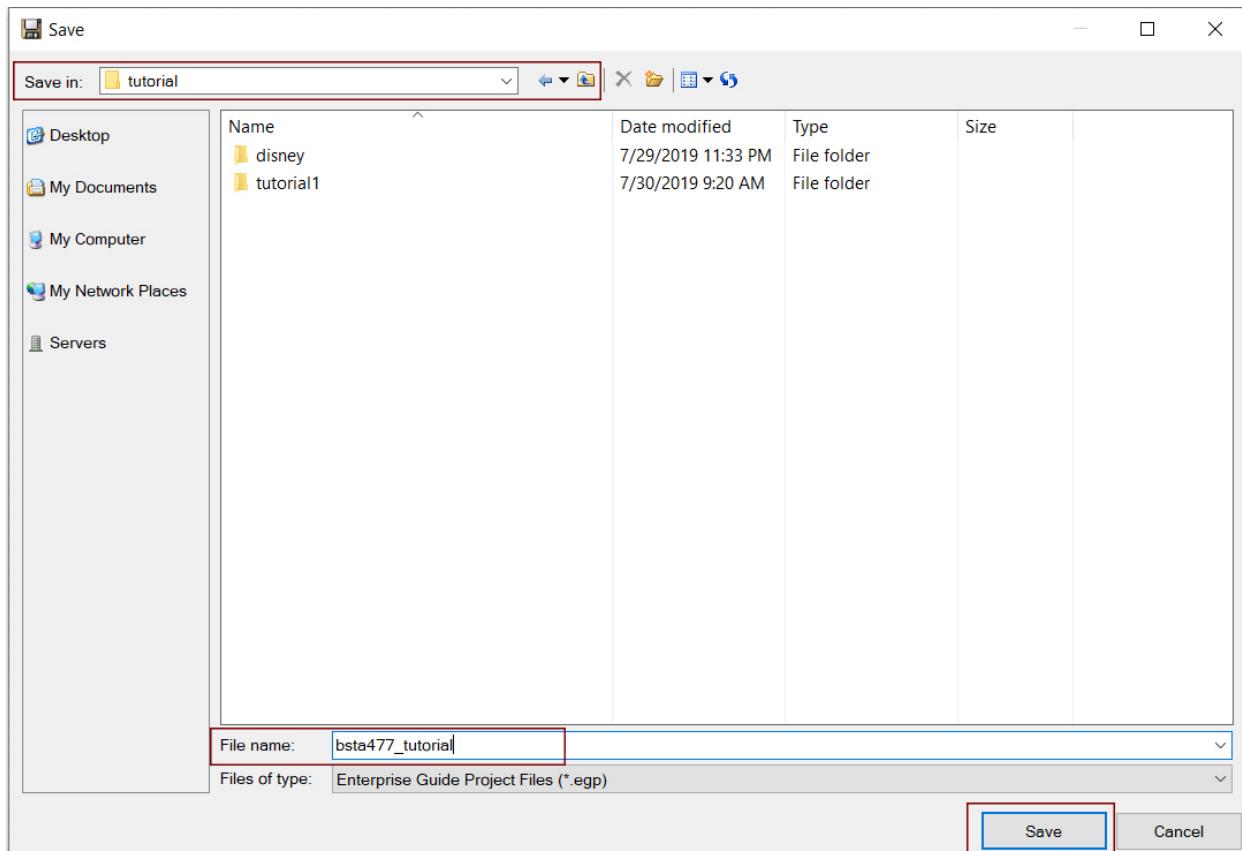
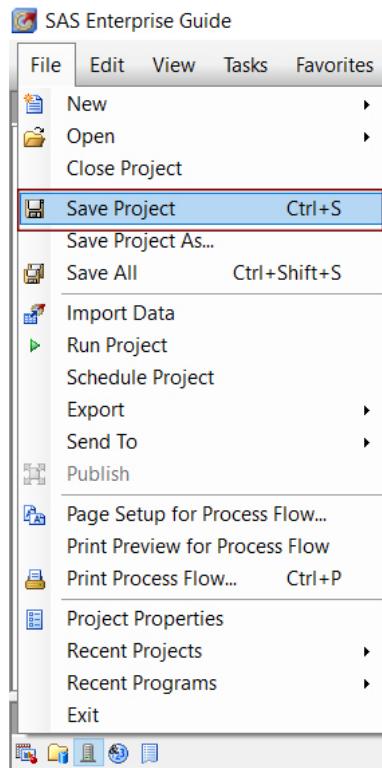
#### How to start your work in SAS

**Create a new project → Save it to a specific location on your computer** (never skip this step if you want to refer to this project again, and you will!!!) → **Run a libname statement** before you perform any operations in SAS (you will need SAS to know where to save files in case you want to save them permanently or open files SAS already has in the library) → **Import or open a dataset** to start working

#### How to create a new project

**Open SAS EG → New → New Project → Save project → Locate the folder where you want to save the file → Give it a name (i.e. bst477\_tutorials)**





## How to run a libname statement

**ALWAYS START A NEW SESSION WITH THIS STEP!!! NEVER SKIP IT!**

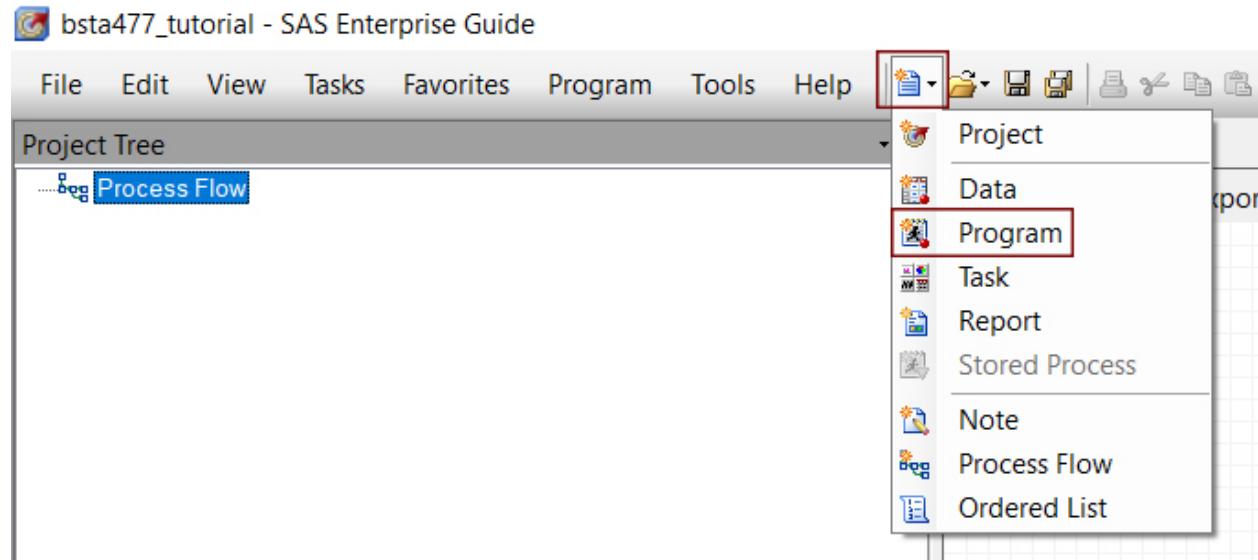
**Click on the icon that looks like a sheet of paper → Choose program → Type the libname statement (as shown below)**

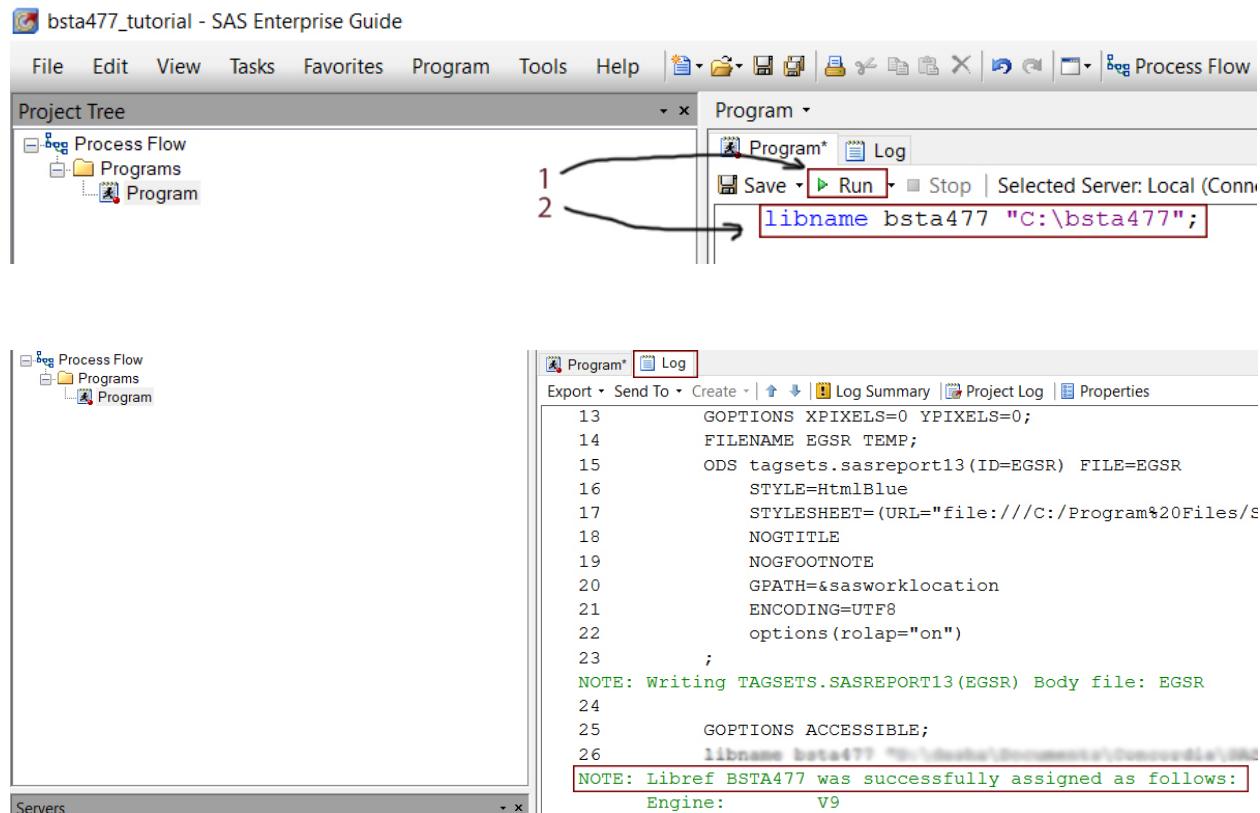
The anatomy of the libname statement. It has four parts:

- **a command** to SAS that right now you are creating a permanent library for your future SAS datasets (libname)
- **the name of the library** that you selected; it must have maximum 8 symbols and should never start with a digit (bst477)
- **the exact path** to the folder on your computer/flashdrive where the library should be physically located; just open the folder where you are planning to save all the files and copy the address from the address bar of that folder (remember to not click on any file in that folder!!!). Always! use quotation marks
- **a semicolon** to tell SAS that the command is complete, and it can start performing it!

```
libname bst477 "C:\bst477";
```

Remember to always check the log. It will show whether the library has been successfully assigned or not!



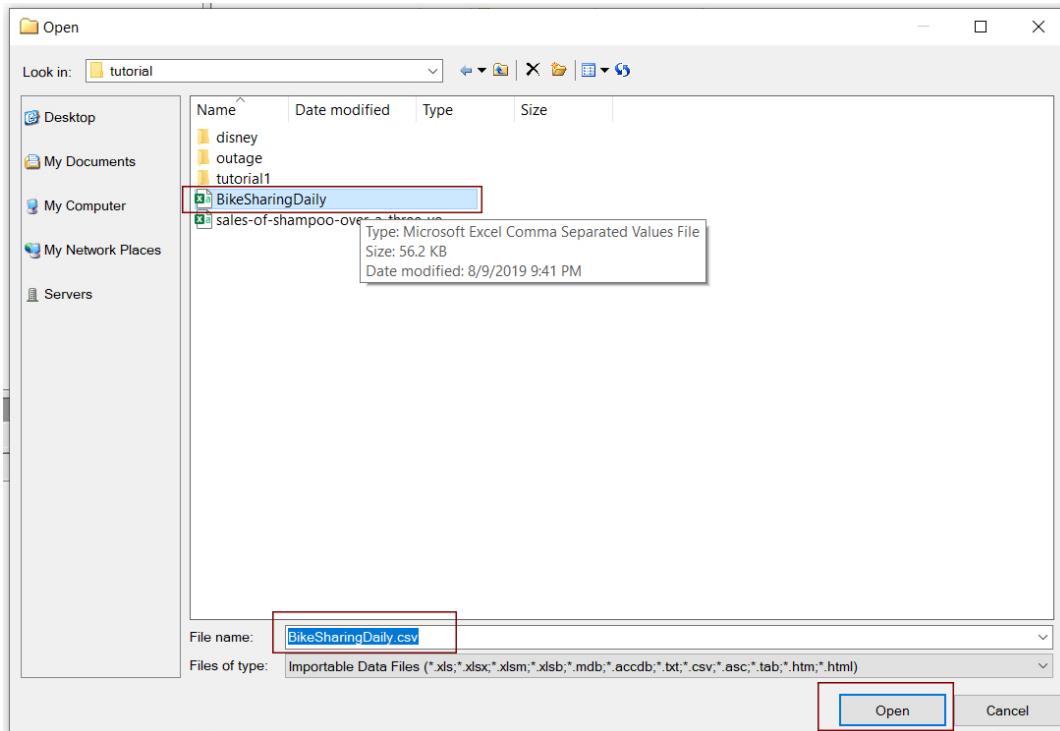
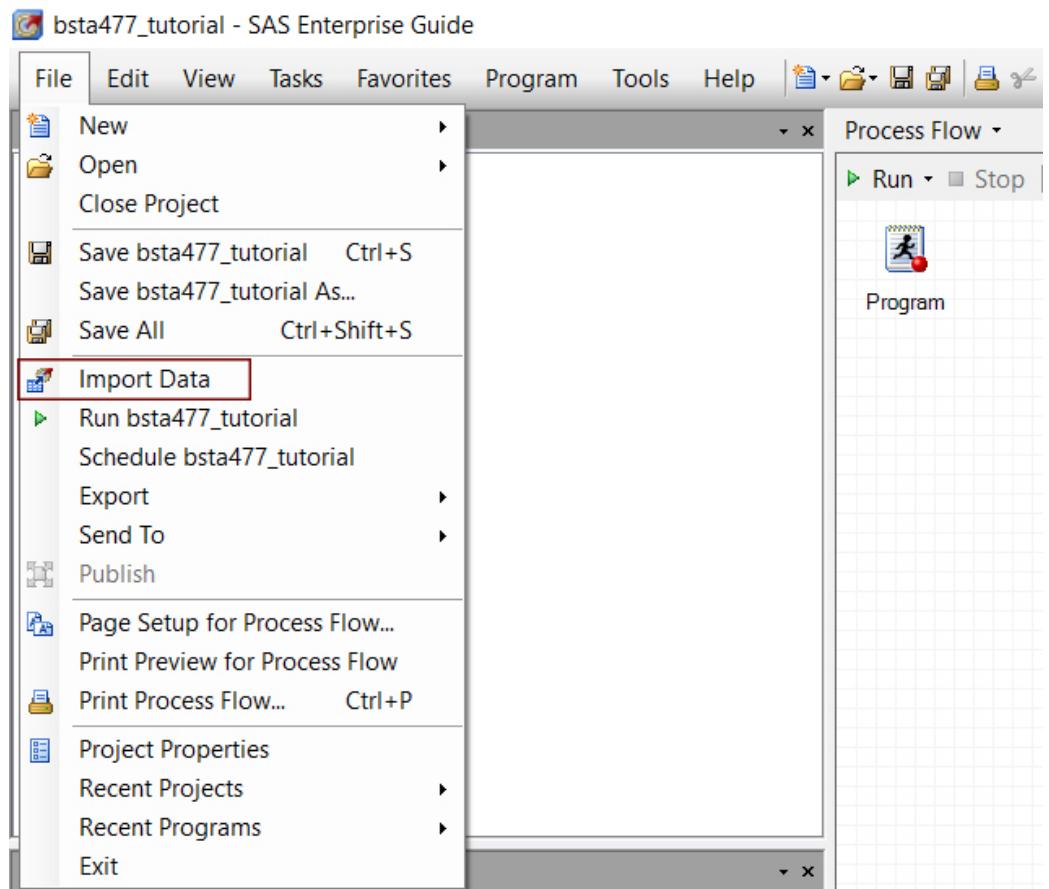


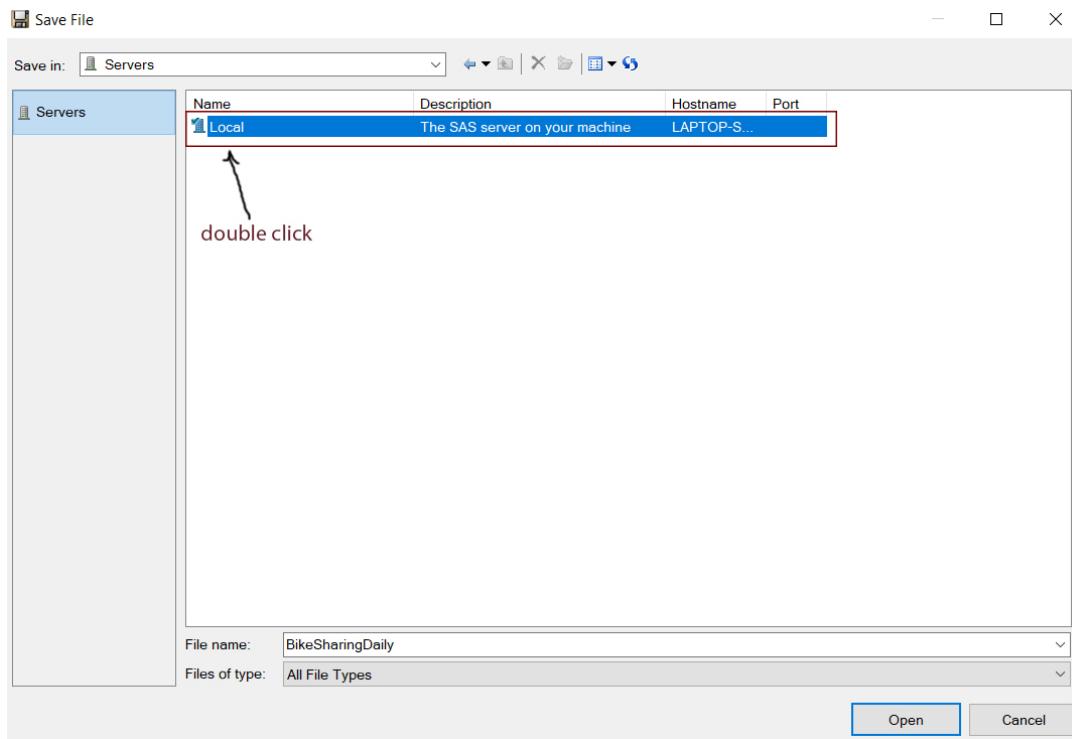
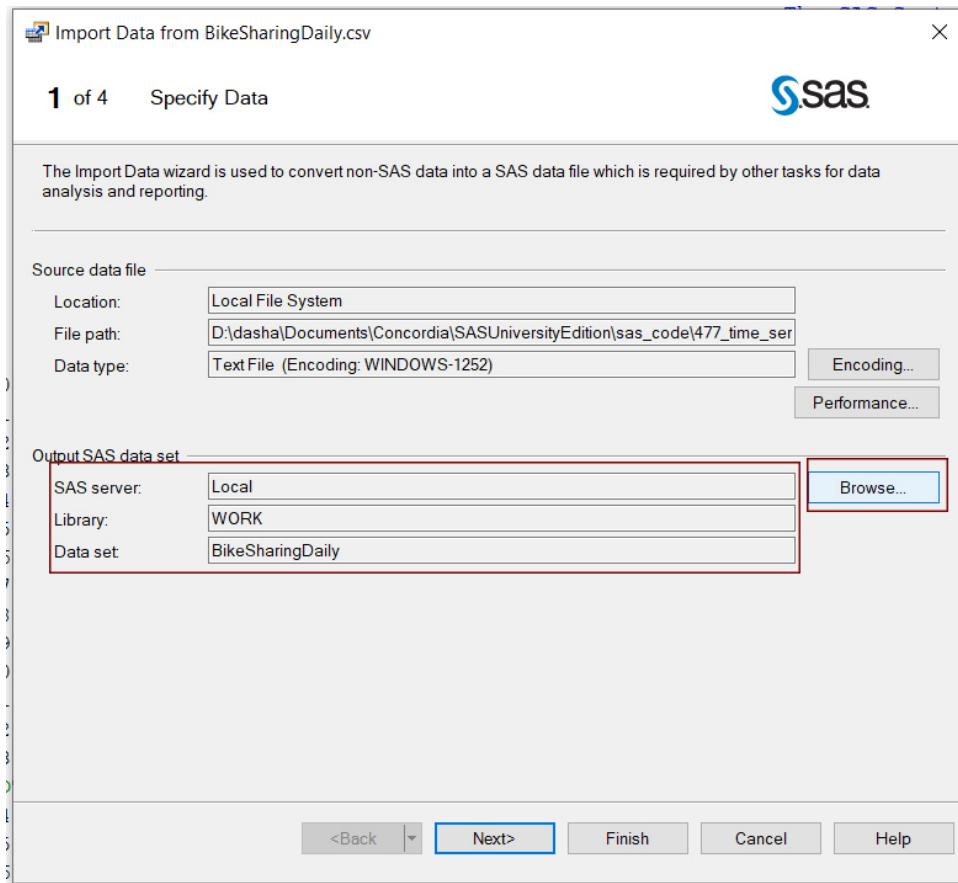
## How to import a dataset

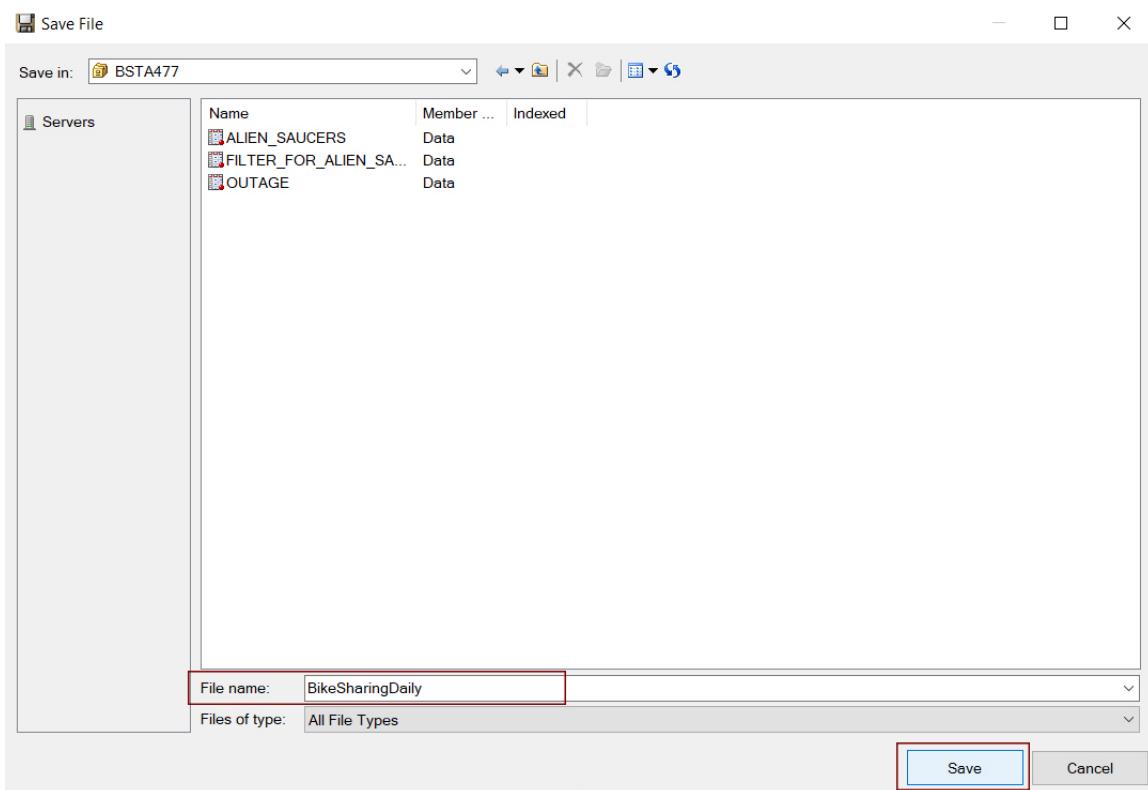
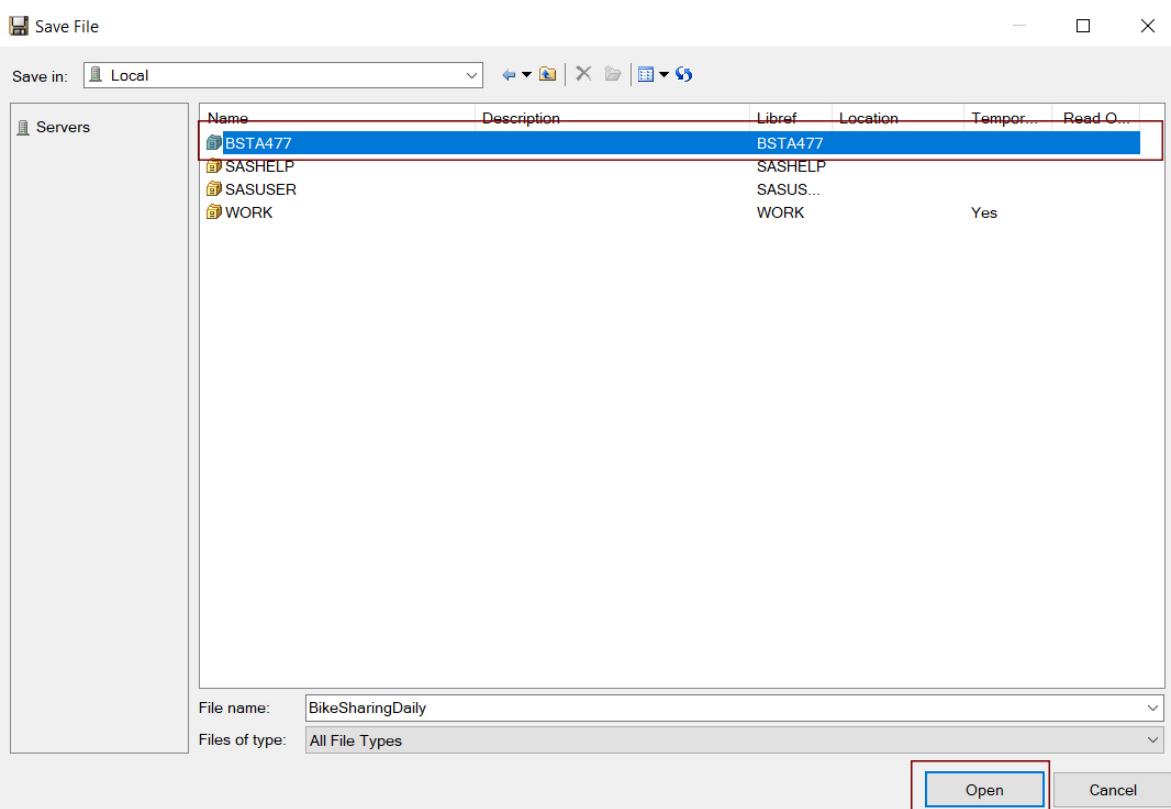
Now we are ready to start work! The idea of using SAS for time series business forecasting is to forecast sales, waiting times, temperature, stock prices etc. using various models, which we will be working on. What you need to start your work is a dataset! SAS cannot “see” datasets with non-SAS extension, so if you have a file in Excel or csv (they are the best!), you will have to import it.

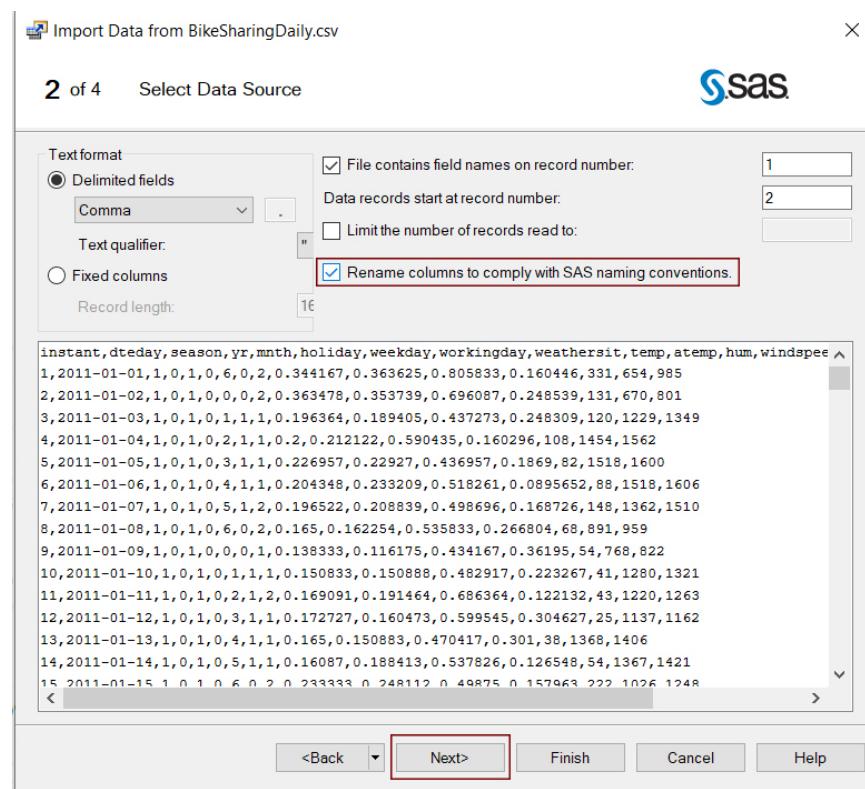
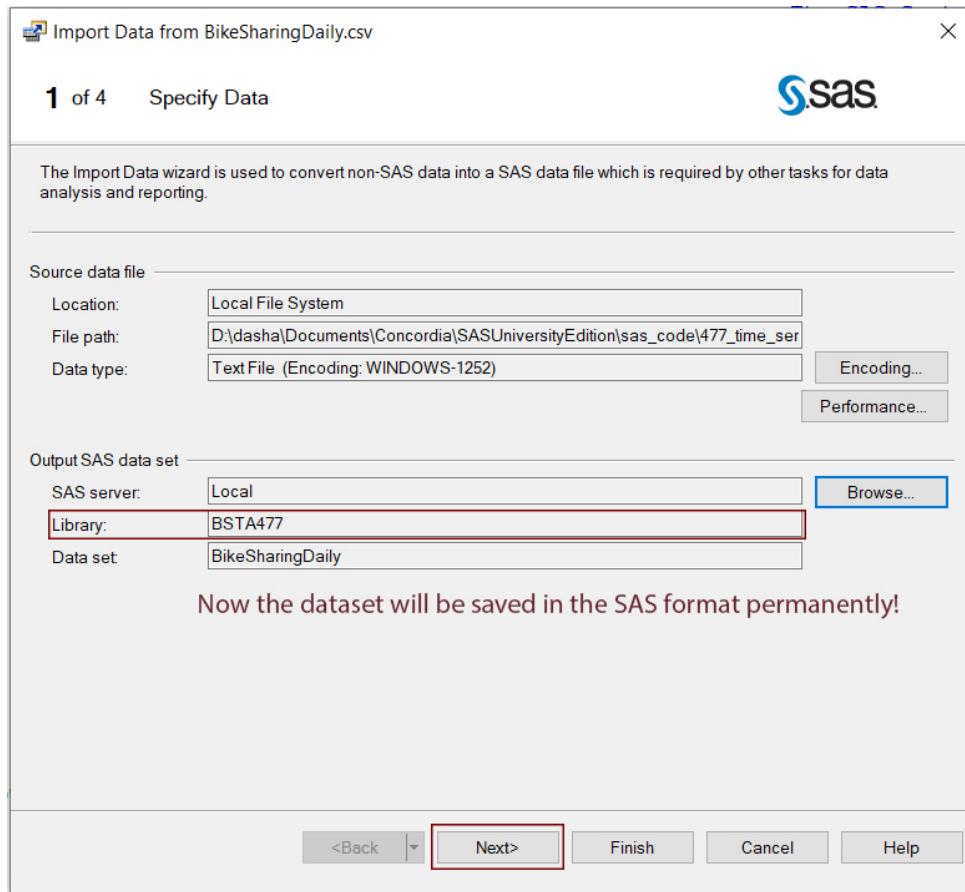
**File → Import Data → Locate the file on your computer/flashdrive → Open → To save the file permanently, change the library from WORK to the one you created in the LIBNAME statement: Browse → Servers → Local → BSTA477 → Open → Give your SAS dataset a name (or keep the original name if you want) → Save → Next → Check the “Rename columns to comply with SAS naming conventions” (it will be easier for you since SAS has its own way of naming variables) → Next → Next → Finish**

Note: You can skip steps 3 and 4, but you should never skip the first two steps because in the first one you will save the newly imported file permanently and you will have access to it after the session is over. In the second step you will change the variable names to the ones understood by SAS.









Import Data from BikeSharingDaily.csv

3 of 4 Define Field Attributes

Select columns and define attributes:

| Inc                                 | Source Name | Name       | Label      | Type   | Source Informat | Len. | Output Format | Output Informat |
|-------------------------------------|-------------|------------|------------|--------|-----------------|------|---------------|-----------------|
| <input checked="" type="checkbox"/> | instant     | instant    | instant    | Number | BEST3.          | 8    | BEST3.        | BEST3.          |
| <input checked="" type="checkbox"/> | dteday      | dteday     | dteday     | Date   | YYMMDD10.       | 8    | YYMMDD10.     | YYMMDD10.       |
| <input checked="" type="checkbox"/> | season      | season     | season     | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | yr          | yr         | yr         | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | mnth        | mnth       | mnth       | Number | BEST2.          | 8    | BEST2.        | BEST2.          |
| <input checked="" type="checkbox"/> | holiday     | holiday    | holiday    | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | weekday     | weekday    | weekday    | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | workingday  | workingday | workingday | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | weathersit  | weathersit | weathersit | Number | BEST1.          | 8    | BEST1.        | BEST1.          |
| <input checked="" type="checkbox"/> | temp        | temp       | temp       | Number | COMMA9.         | 8    | BEST9.        | BEST9.          |
| <input checked="" type="checkbox"/> | atemp       | atemp      | atemp      | Number | COMMA9.         | 8    | BEST9.        | BEST9.          |
| <input checked="" type="checkbox"/> | hum         | hum        | hum        | Number | COMMA8.         | 8    | BEST8.        | BEST8.          |
| <input checked="" type="checkbox"/> | windspeed   | windspeed  | windspeed  | Number | COMMA9.         | 8    | BEST9.        | BEST9.          |
| <input checked="" type="checkbox"/> | casual      | casual     | casual     | Number | BEST4.          | 8    | BEST4.        | BEST4.          |
| <input checked="" type="checkbox"/> | registered  | registered | registered | Number | BEST4.          | 8    | BEST4.        | BEST4.          |
| <input checked="" type="checkbox"/> | cnt         | cnt        | cnt        | Number | BEST4.          | 8    | BEST4.        | BEST4.          |

Select All   Clear All   Modify...

<Back   Next>   Finish   Cancel   Help

Import Data from alien\_saucers.csv

4 of 4 Advanced Options

Embed the data within the generated SAS code.

Import the data using SAS/ACCESS Interface to PC Files whenever possible.

Remove characters that can cause transmission errors from text-based data files.

Generalize import step to run outside SAS Enterprise Guide.

Maximum record length for input text file in bytes:

<Back   Next>   Finish   Cancel   Help

| instant | dteday     | season | yr | mnth | holiday | weekday | workingday | cnt |
|---------|------------|--------|----|------|---------|---------|------------|-----|
| 1       | 2011-01-01 | 1      | 0  | 1    | 0       | 6       | 0          | 0   |
| 2       | 2011-01-02 | 1      | 0  | 1    | 0       | 0       | 0          | 0   |
| 3       | 2011-01-03 | 1      | 0  | 1    | 0       | 1       | 0          | 1   |
| 4       | 2011-01-04 | 1      | 0  | 1    | 0       | 0       | 2          | 1   |
| 5       | 2011-01-05 | 1      | 0  | 1    | 0       | 3       | 1          | 1   |

## How to create a line plot?

When you work with time series, a nice way to assess what kind of data you have available is to plot it using a line plot. It will show you whether your data has a trend, seasonality or if it is stationary. The dates will be on the x-axis, and the data you are trying to predict are plotted on the y-axis. In SAS EG this plot is called a line plot and you have a few options:

- a) Line Plot Wizard or
- b) Line Plot

### A) The Line Plot Wizard:

Open the datafile you want to explore → Graph → Line Plot Wizard → Next → Horizontal (a date/datetime variable) → Vertical (the target, the variable you want to predict) → Next → Type: Line, Symbol: Point → Next → Appearance options: Tick marks → Next → Graph: Give your graph a name → Finish.

The screenshot shows the SAS Enterprise Guide interface with the 'Import Data (BikeSharingDaily.csv)' window open. The 'Graph' menu is highlighted, and a dropdown menu is displayed with several options: Bar Chart Wizard..., Bar Chart..., Pie Chart Wizard..., Pie Chart..., Line Plot Wizard... (which is highlighted with a blue selection bar), Line Plot..., Scatter Plot..., Scatter Plot Matrix..., and Area Plot....

**Line Plot for Local:BSTA477.BIKESHARINGDAILY**

**1 of 5 Verify Data**

The Line Plot Wizard helps you create two-dimensional line plots, with options for markers and line interpolations.

Data:

SAS server: Local  
Library: BSTA477  
Data set: BIKESHARINGDAILY

Task filter: None

Edit... <Back Next> Finish Cancel Help

---

**Line Plot for Local:BSTA477.BIKESHARINGDAILY**

**2 of 5 Assign variables to roles**

Required

Horizontal: dteday  
Vertical: cnt

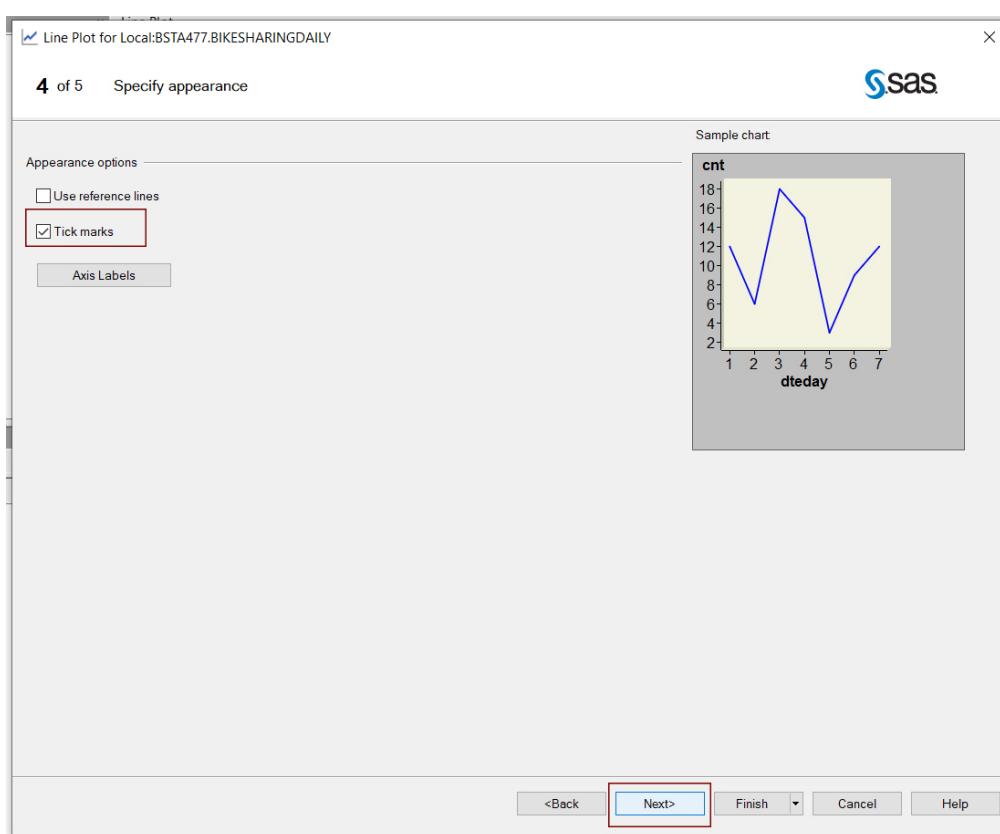
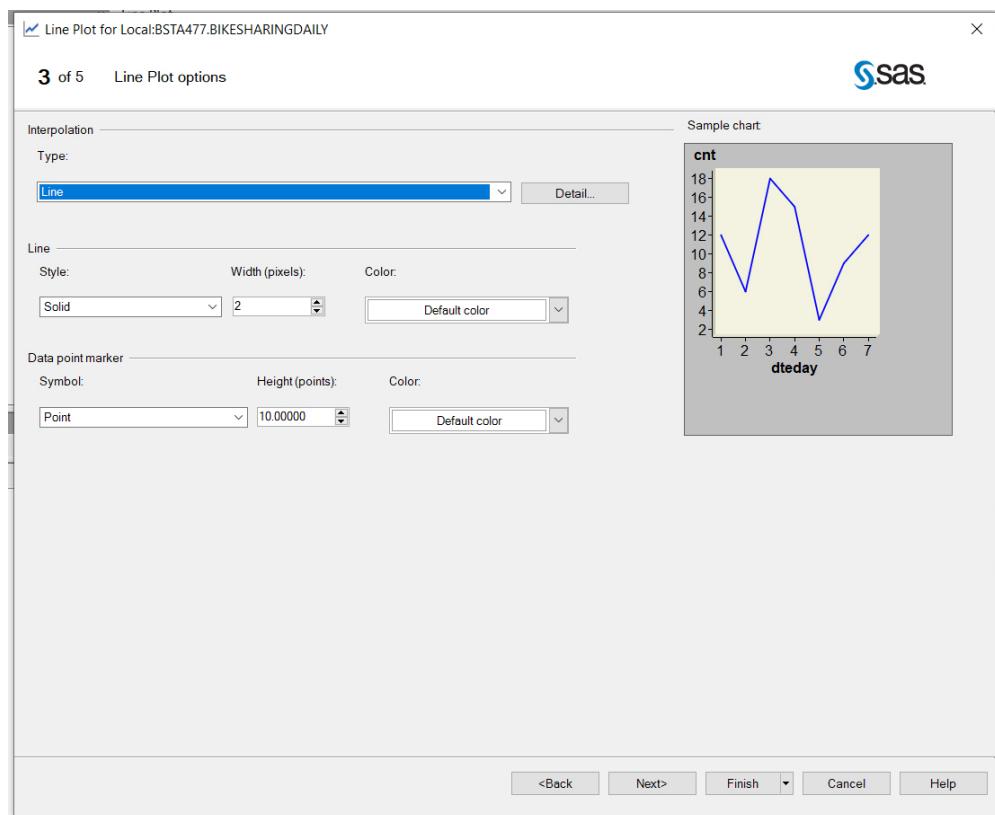
Optional

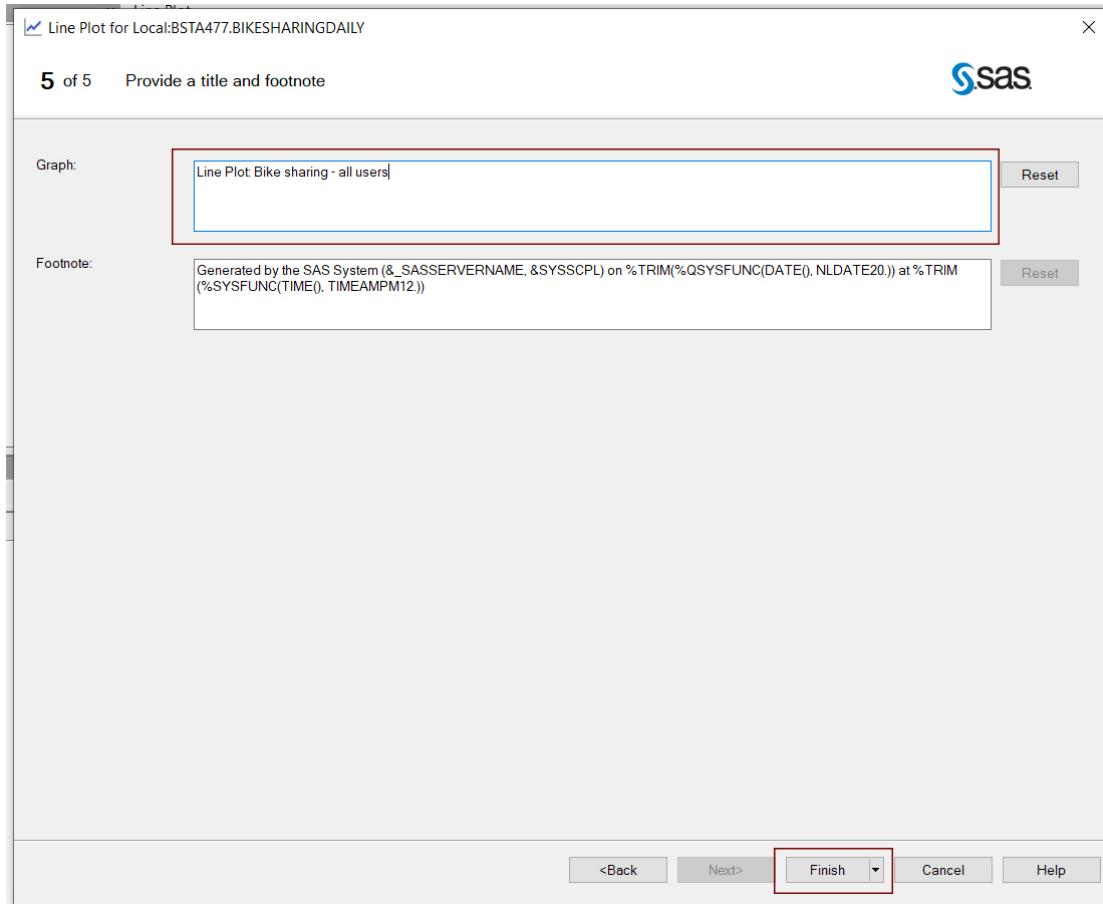
Chart by: instant  
 Summarize for each distinct horizontal value  
Function: Sum

Sample chart:

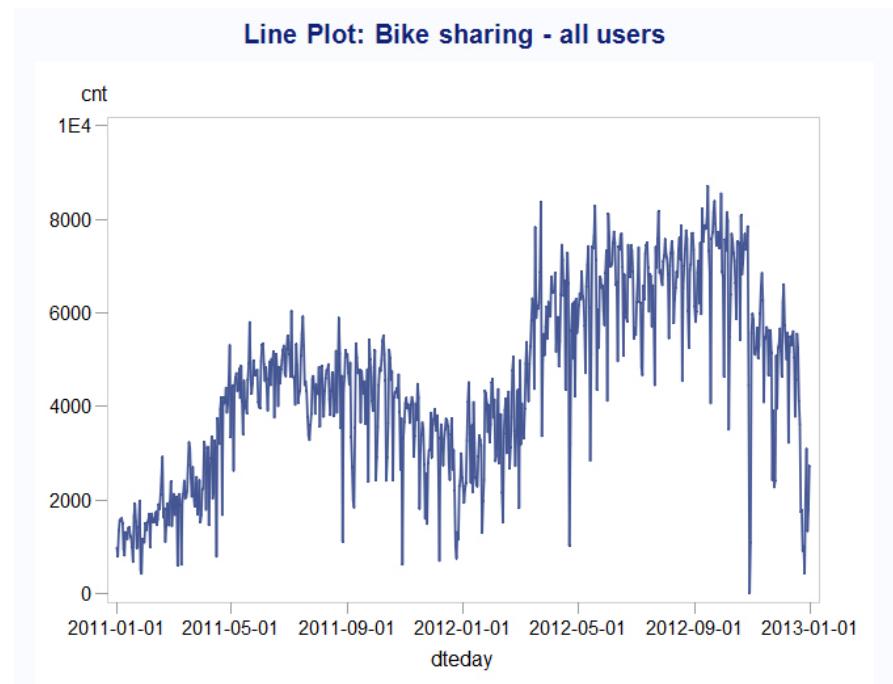
| dteday | cnt |
|--------|-----|
| 1      | 11  |
| 2      | 7   |
| 3      | 17  |
| 4      | 16  |
| 5      | 3   |
| 6      | 10  |
| 7      | 12  |

<Back Next> Finish Cancel Help





Here's the resulting plot.



## B) The Line plot

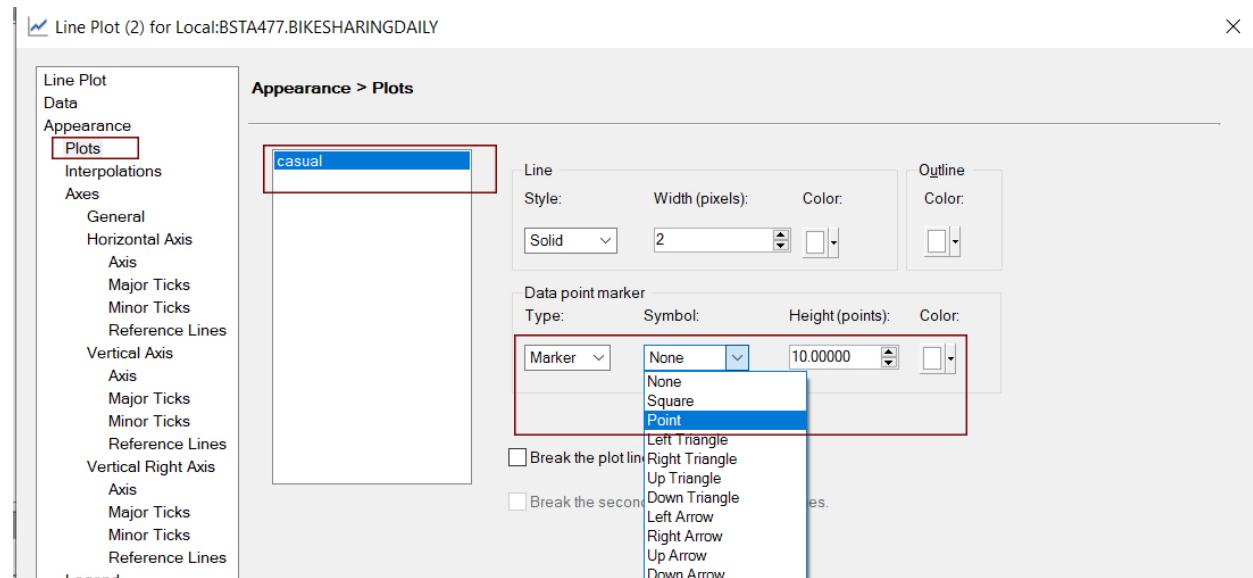
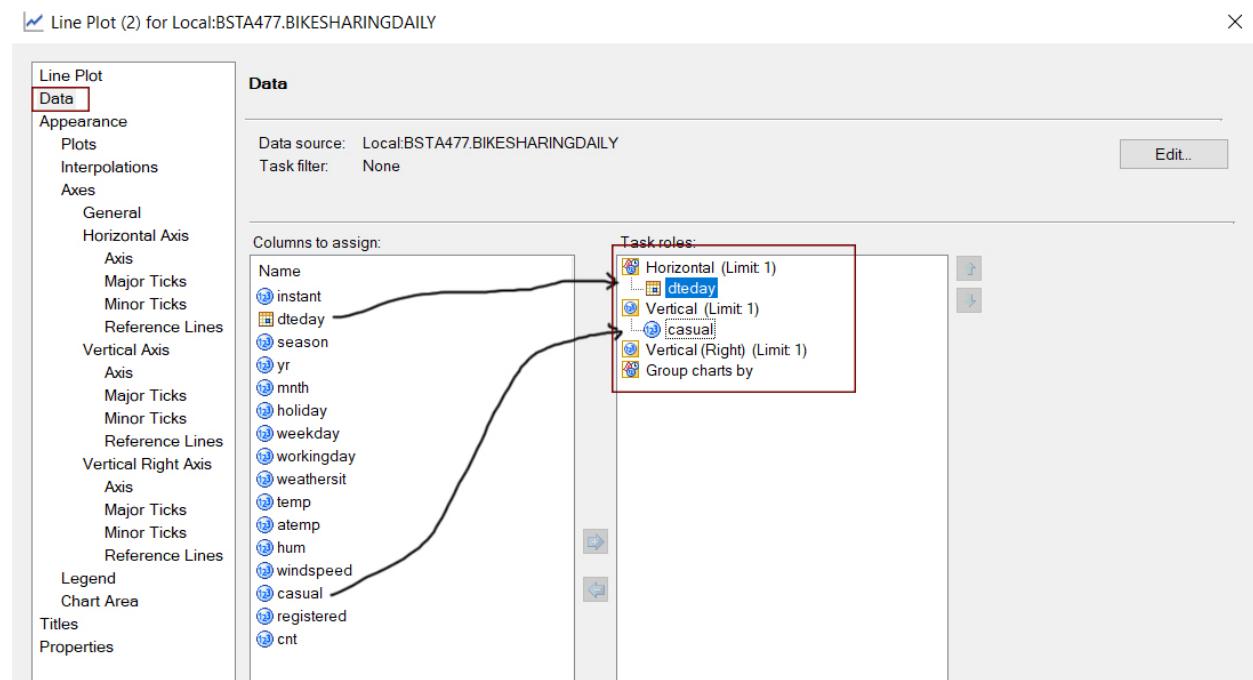
Open the datafile you want to explore → Graph → Line Plot → Choose line plot from the list of the graphs on the right → Data (where you will select the variables you want to plot): the datetime/date variable should be assigned to the Horizontal and the target should be assigned to the Vertical role → Plots → Marker: Point → Titles: Take off the check from the “Use default text” and give a title to your graph → Run

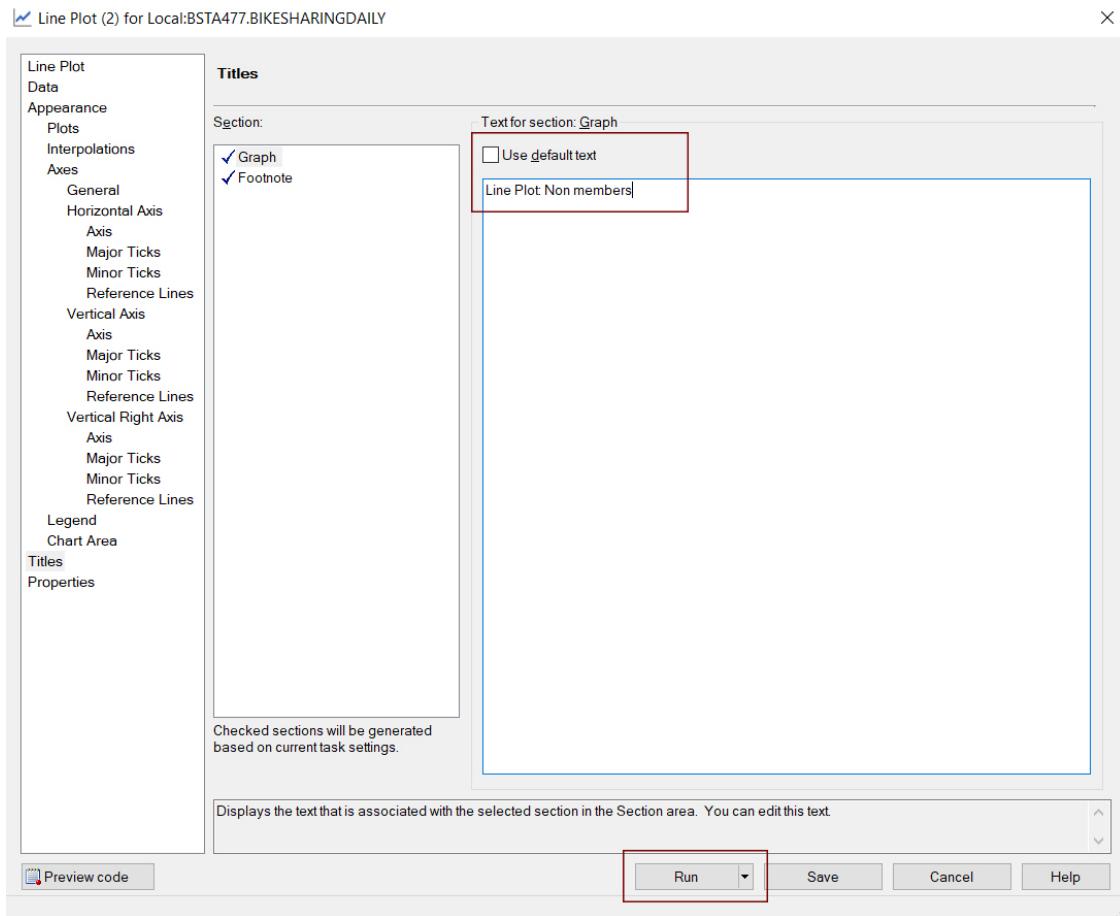
As you could notice, choosing the Line plot you get more control over your graph when building it. The Line plot wizard is a faster way of making line plots. Also Tableau is a great way for visually exploring your data.

The screenshot shows the RapidMiner interface with the 'Line Plot' tab selected. The 'Graph' menu is open, and the 'Line Plot...' option is highlighted with a red box. A preview window on the right shows a line plot for 'instant' vs 'weekday'.

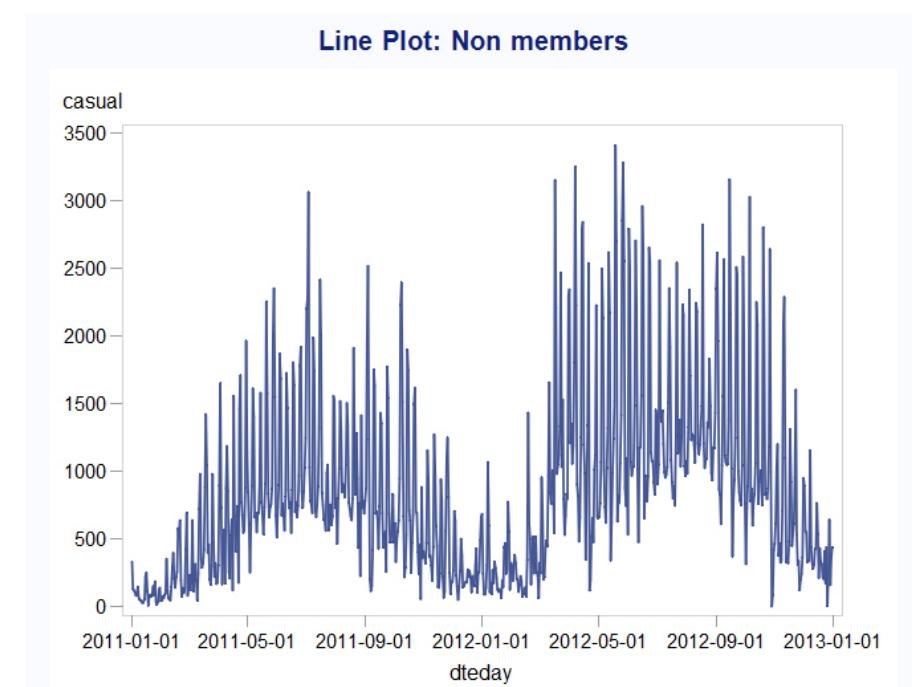
| instant | dteday     | season | yr | weekday |
|---------|------------|--------|----|---------|
| 1       | 2011-01-01 | 1      | 1  | 6       |
| 2       | 2011-01-02 | 1      | 1  | 0       |
| 3       | 2011-01-03 | 1      | 1  | 1       |
| 4       | 2011-01-04 | 1      | 1  | 2       |
| 5       | 2011-01-05 | 1      | 1  | 3       |
| 6       | 2011-01-06 | 1      | 1  | 4       |
| 7       | 2011-01-07 | 1      | 1  | 5       |
| 8       | 2011-01-08 | 1      | 1  | 6       |
| 9       | 2011-01-09 | 1      | 1  | 0       |

The screenshot shows the Tableau interface with the 'Line Plot (2)' window open. The left sidebar shows the configuration options: Line Plot, Data, Appearance, Plots, Interpolations, Axes, General, Horizontal Axis, Vertical Axis, and Major Ticks. The 'Plots' section is expanded, showing various plot types: Line Plot (selected), Spline Plot, Needle Plot, Step Plot, Scatter Plot with Regression Line, Smooth Plot, and Standard Deviation Plot. Below these are other options: Lagrange Interpolation Plot, Multiple vertical column line plot..., and Multiple line plots by group column. A 'Scatter Plot' icon is also visible.





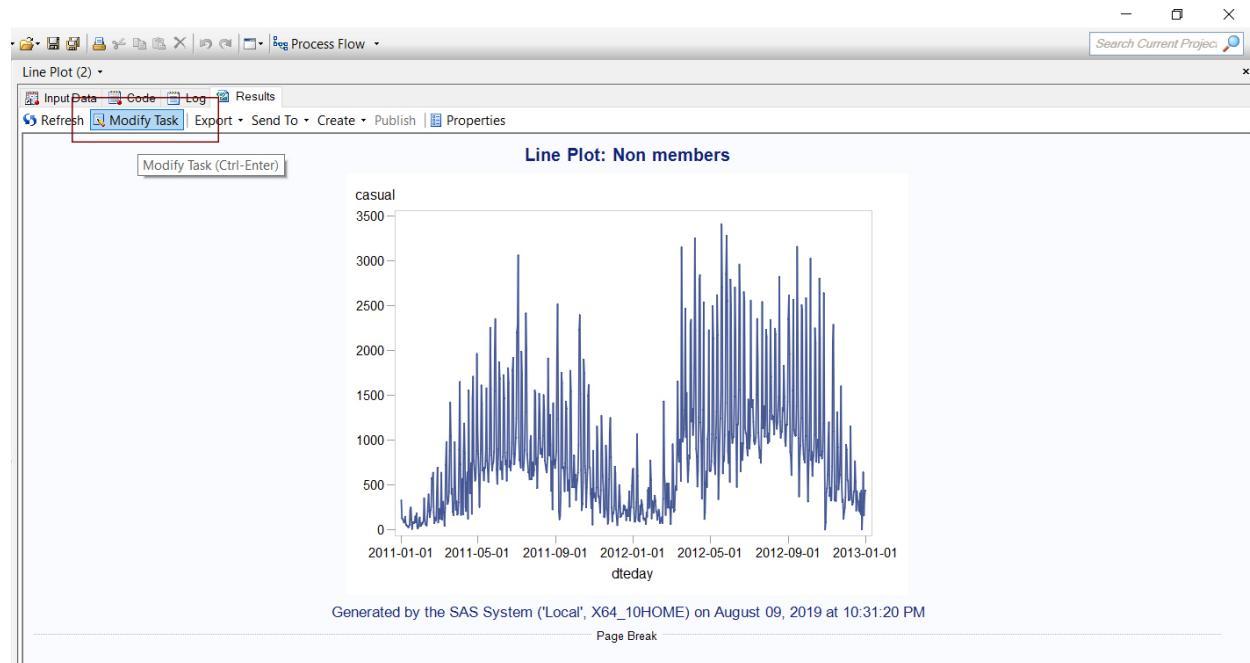
Here's the resulting plot.

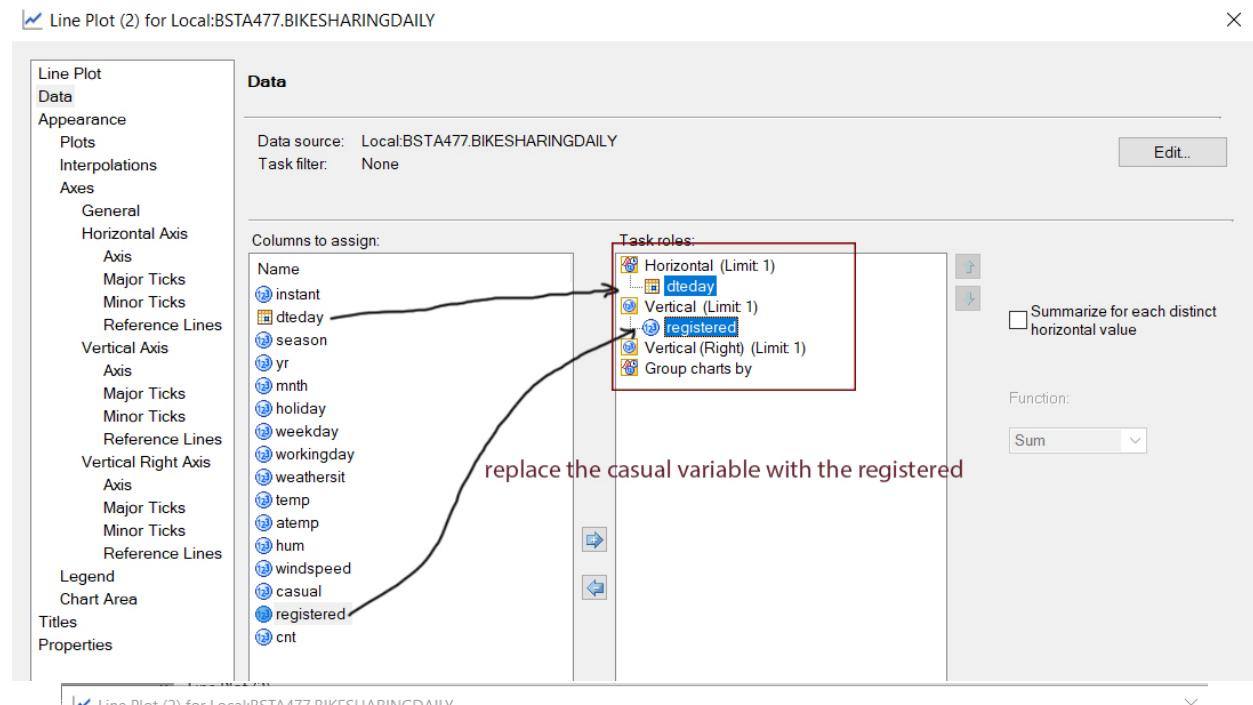


### How to use “Modify task” for your advantage

Sometimes you want to make a few line plots, which is our case. Using the Line plot wizard, we created a line plot for the total count of rides by both members and non-members. Then using the line plot, we created a line plot for the casual (non-member) rides. Now, we can use the previous line plot task to create a line plot for registered users. For that you should be at the page with the newly created line plot and follow these steps:

**Page with the created graph → Data: Replace the variable from the previous run (casual in this case) with another variable you want to plot this time (registered in our case) → Titles: Change the title of the graph → Run → No (choosing this option, you will create a new line plot and not overwrite the previously created plot; thus, you will end up having both time series plots)! Do you see how much time it saves you?**





Line Plot (2) for Local:BSTA477.BIKESHARINGDAILY

Data

Data source: Local:BSTA477.BIKESHARINGDAILY Task filter: None Edit...

Columns to assign:

- Name
- instant
- dteday
- season
- yr
- mnth
- holiday
- weekday
- workingday
- weathersit
- temp
- atemp
- hum
- windspeed
- casual
- registered
- cnt

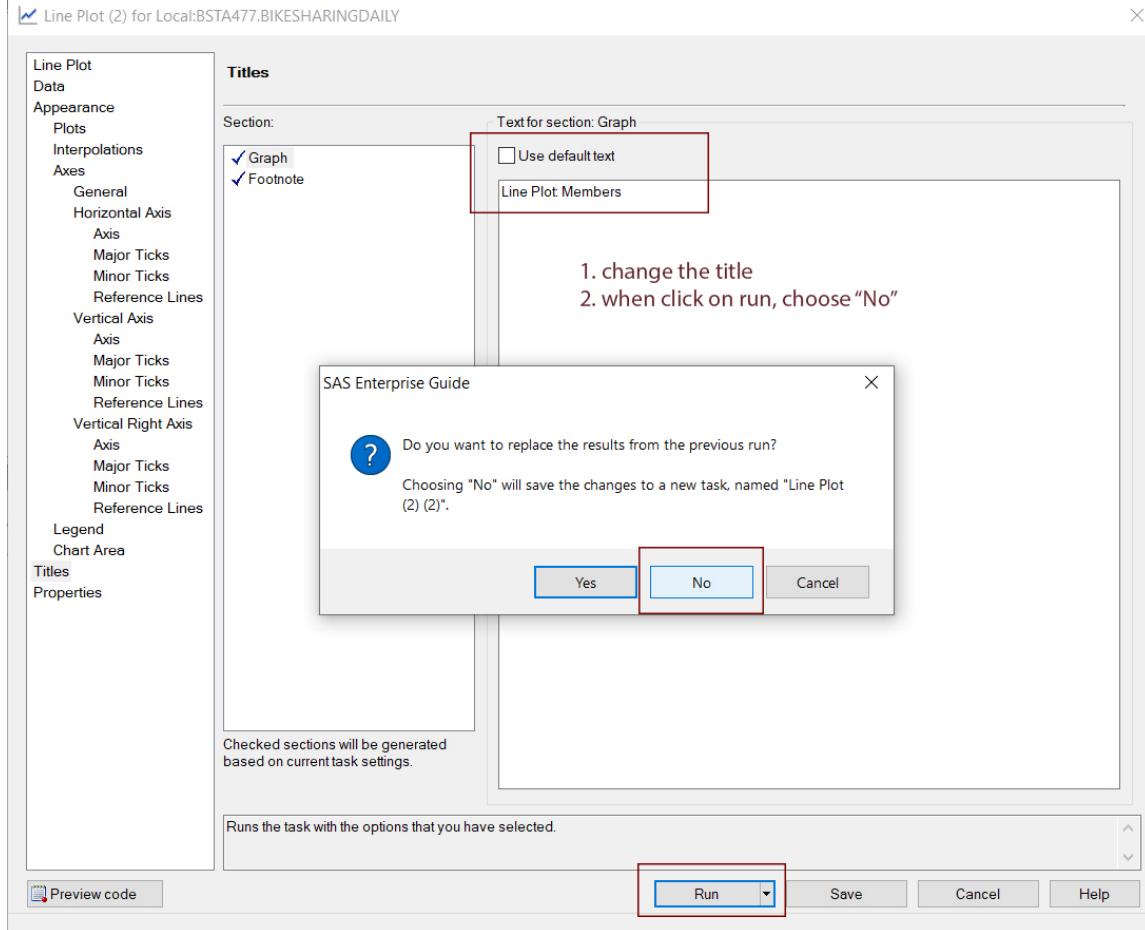
Task roles:

- Horizontal (Limit 1)
- dteday
- Vertical (Limit 1)
- registered
- Vertical (Right) (Limit 1)
- Group charts by

Summarize for each distinct horizontal value

Function: Sum

replace the casual variable with the registered



Line Plot (2) for Local:BSTA477.BIKESHARINGDAILY

Titles

Section: Graph

Text for section: Graph

Graph  
Footnote

Use default text

Line Plot Members

1. change the title  
2. when click on run, choose "No"

SAS Enterprise Guide

Do you want to replace the results from the previous run?  
Choosing "No" will save the changes to a new task, named "Line Plot (2) (2)".

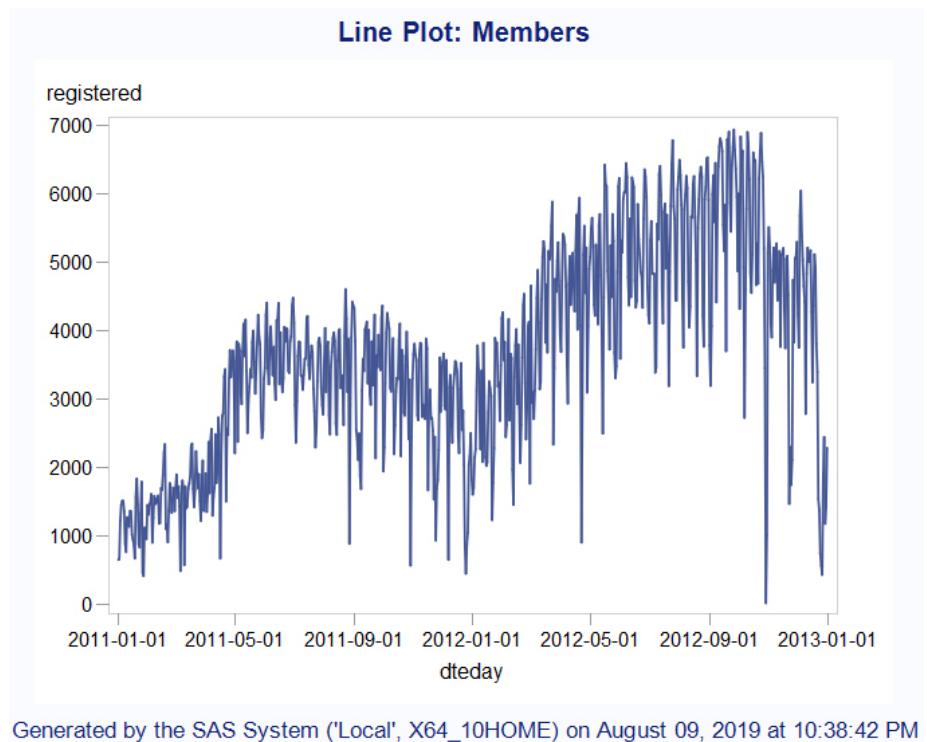
Yes No Cancel

Checked sections will be generated based on current task settings.

Runs the task with the options that you have selected.

Run Save Cancel Help

Preview code



## How to create a histogram

You also want to check whether the target variable is normally distribution. If it isn't, then you will need to transform it. Our target variable is going to be one of the three: casual (non-member rides), registered (member rides) or cnt (total count of rides).

To check the distribution of the continuous variable, you should create a histogram and explore the variable distribution. You can also include additional statistics of your choice from the list provided by SAS such as mean, median, skewness.

**With the dataset open: Describe → Distribution Analysis → Data: Drag the variables you want to explore to the right pane and place them under the Analysis variables → Plots: Appearance: Histogram Plot (put a check) and Background color: White → Plots: Inset, check “Include Inset” and check the statistics you are interested in → Run**

As we can see from the histogram below, the casual variable needs transformation as it is highly skewed to the right. Thus, we will need to take a log of it and instead of predicting the casual (non-member) rides, we will be predicting the log(casual) and then taking an exponential function on its predictions to obtain the needed results.

Data Imported from BikeSharingDaily.csv

|    | instant | dteday     | season | holiday | weekday |
|----|---------|------------|--------|---------|---------|
| 1  | 1       | 2011-01-01 | 1      | 0       | 6       |
| 2  | 2       | 2011-01-02 | 1      | 0       | 0       |
| 3  | 3       | 2011-01-03 | 1      | 0       | 1       |
| 4  | 4       | 2011-01-04 | 1      | 0       | 2       |
| 5  | 5       | 2011-01-05 | 1      | 0       | 3       |
| 6  | 6       | 2011-01-06 | 1      | 0       | 4       |
| 7  | 7       | 2011-01-07 | 1      | 0       | 5       |
| 8  | 8       | 2011-01-08 | 1      | 0       | 6       |
| 9  | 9       | 2011-01-09 | 1      | 0       | 0       |
| 10 | 10      | 2011-01-10 | 1      | 0       | 1       |
| 11 | 11      | 2011-01-11 | 1      | 0       | 2       |
| 12 | 12      | 2011-01-12 | 1      | 0       | 3       |

Distribution Analysis for Local:BSTA477.BIKESHARINGDAILY

**Data**

Data source: Local:BSTA477.BIKESHARINGDAILY  
Task filter: None

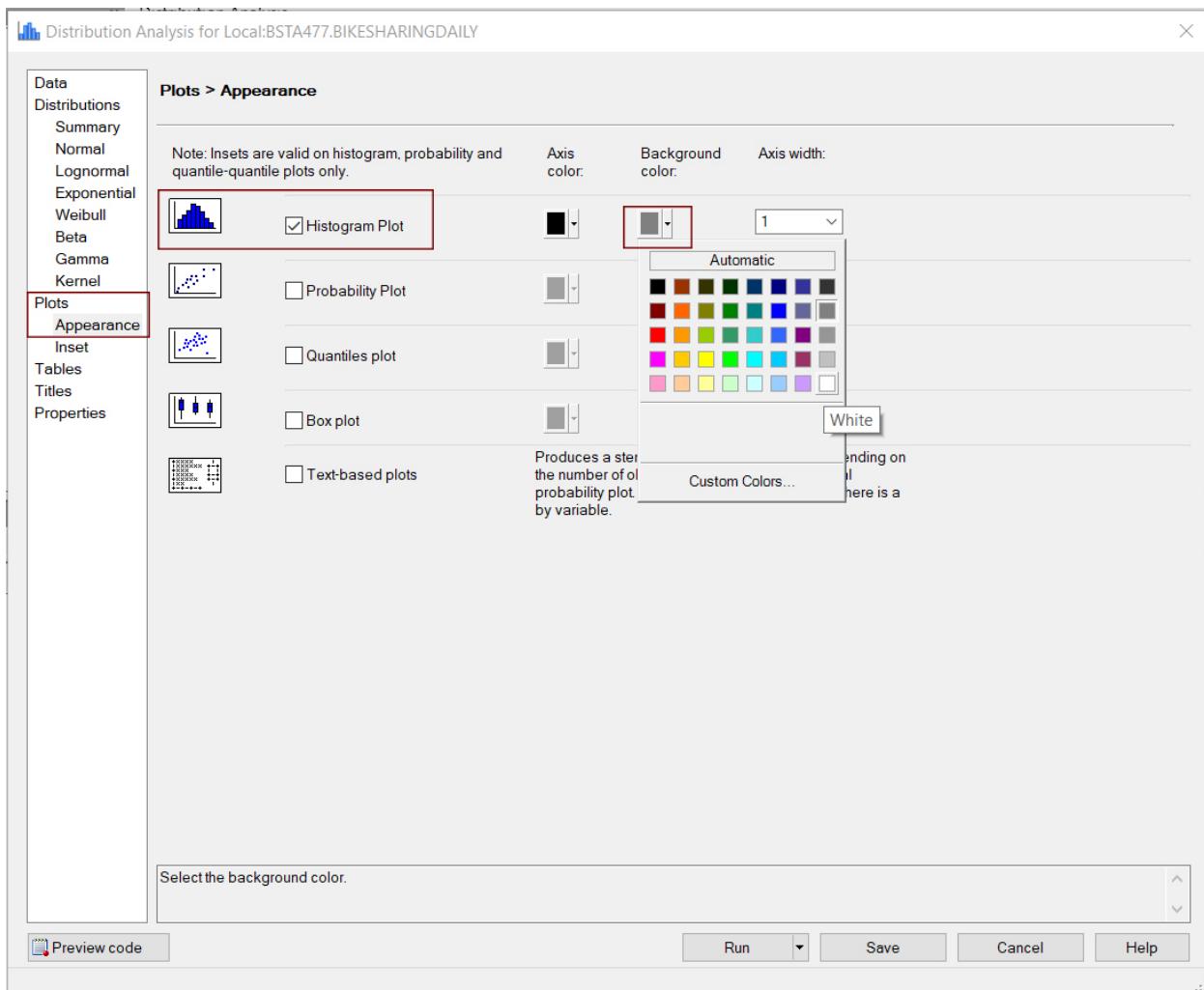
Variables to assign:

|      |            |
|------|------------|
| Name | instant    |
|      | dteday     |
|      | season     |
|      | yr         |
|      | mnth       |
|      | holiday    |
|      | workingday |
|      | weathersit |
|      | temp       |
|      | atemp      |
|      | hum        |
|      | windspeed  |
|      | casual     |
|      | registered |
|      | cnt        |

Task roles:

|                                    |  |
|------------------------------------|--|
| Analysis variables                 | <input checked="" type="checkbox"/> casual     |
|                                    | <input checked="" type="checkbox"/> registered |
|                                    | <input checked="" type="checkbox"/> cnt        |
| Group analysis by                  | <input type="checkbox"/>                       |
| Frequency count (Limit 1)          | <input type="checkbox"/>                       |
| Relative weight (Limit 1)          | <input type="checkbox"/>                       |
| Classification variables (Limit 2) | <input type="checkbox"/>                       |

Select a role to view the context help for that role.



**Distribution Analysis for Local:BSTA477.BIKESHARINGDAILY**

**Plots > Inset**

**Include inset**

**Inset statistics**

Sample size  
 Sum of the weights  
 Sample mean  
 Sum of the observations  
 Standard deviation  
 Variance  
 Skewness  
 Kurtosis  
 Largest value  
 Smallest value  
 Number of observations  
 Range  
 Most frequent value  
 Number of missing values  
 Uncorrected sum of squares  
 Corrected sum of squares  
 Coefficient of variation  
 Standard error of the mean  
 1st percentile  
 5th percentile  
 10th percentile  
 Lower quartile (25th percentile)  
 Median (50th percentile)  
 Upper quartile (75th percentile)  
 90th percentile  
 95th percentile  
 99th percentile

**Inset location:** Northwest

**Text:**

**Frame:**

**Background:**

**Inset label:**

**Color:**

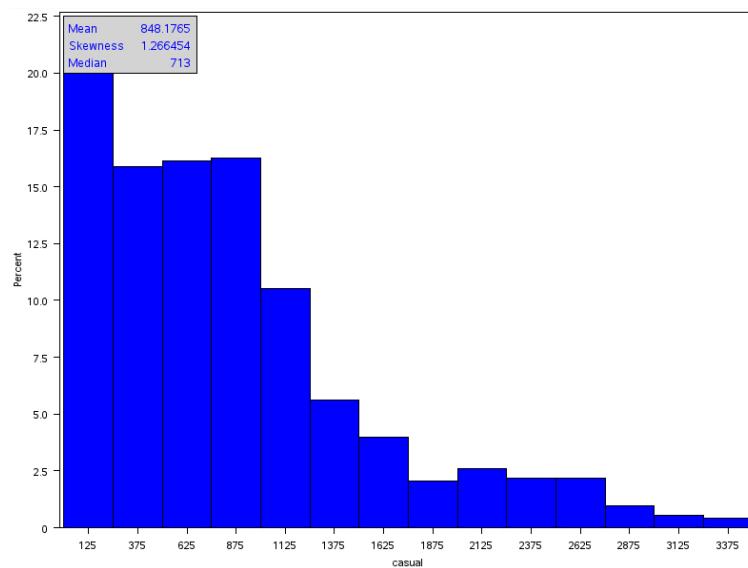
**Background:**

**Inset text height:** 2.00000

Select the statistics to include in the inset.  
 Calculates a value that exceeds half of the sample data values and is exceeded by half of the sample data values.

**Preview code**

**Run** **Save** **Cancel** **Help**



What about the other two variables? How are they distributed? Do they need transformation?

## How to transform variables

Basically, a variable transformation in SAS will lead to creating a new variable. For that we will need to run a Query, which will result not only in creating a new variable, but also in creating a new dataset. We will continue working with the new dataset afterwards since it will contain the newly created variable we will need.

With the dataset open:

**STEP1: Query Builder → Drag the variables you want to see in your new dataset into the right pane.**

If you want to include all the columns from the previous file, drag the t1.(BIKESHARINGDAILY) into the right pane.

**STEP2:** make sure you save the new dataset permanently. **In the top right corner click on the Change button next to the Output name → Servers → Local → BSTA477 → Open → Save**

Otherwise, you will create a temporary file that you will not be able to refer to in your next session and to access the file you will always have to repeat all these steps, which is not productive. If you look at the output file name, you will see that originally the file is saved in the WORK library!

**STEP3:** Create a new variable:

**Click on the icon located on the right that looks like a calculator (Add a New Computed Column) → Advanced expression → Next → Type in the field: log and an opening bracket and then drag the variable you want to take a log of and close the bracket to have: log(t1.casual) → Next → In the Column Name type the name of the new variable: Log\_casual; Format: Click on the Change button → Categories: Numeric, BESTD w.d; set the overall width to 8 → OK → Next → Finish → Run**

The screenshot shows the SAS interface with the 'One-Way Frequencies' window open. At the top, there are several tabs: Input Data, Code, Log, Results, Filter and Sort, and Query Builder. The 'Query Builder' tab is highlighted with a red box. Below the tabs is a menu bar with options like Where, Data, Describe, Graph, Analyze, Export, Send To, and a Process Flow button. Underneath the menu is a data grid showing two rows of data. The columns are labeled instant, dteday, season, yr, mnth, holiday, and weekday. Row 1 has values 1, 2011-01-01, 1, 0, 1, 0, 0. Row 2 has values 2, 2011-01-02, 1, 0, 1, 1, 0.

Query Builder for Local:BSTA477.BIKESHARINGDAILY

Query name: Query Builder Output name: WORKQUERY\_FOR\_BIKESHARINGDAILY Change...

Computed Columns Prompt Manager Preview Tools Options

Add Tables Delete Join Tables Select Data Filter Data Sort Data

| Column Name | Source Column | Sub... | Format | Details |
|-------------|---------------|--------|--------|---------|
| instant     | t1.instant    |        |        |         |
| dteday      | t1.dteday     |        |        |         |
| season      | t1.season     |        |        |         |
| yr          | t1.yr         |        |        |         |
| mnth        | t1.mnth       |        |        |         |
| holiday     | t1.holiday    |        |        |         |
| weekday     | t1.weekday    |        |        |         |
| workingday  | t1.workingday |        |        |         |
| weathersit  | t1.weathersit |        |        |         |
| temp        | t1.temp       |        |        |         |
| atemp       | t1.atemp      |        |        |         |
| hum         | t1.hum        |        |        |         |
| windspeed   | t1.windspeed  |        |        |         |
| casual      | t1.casual     |        |        |         |
| registered  | t1.registered |        |        |         |
| cnt         | t1.cnt        |        |        |         |

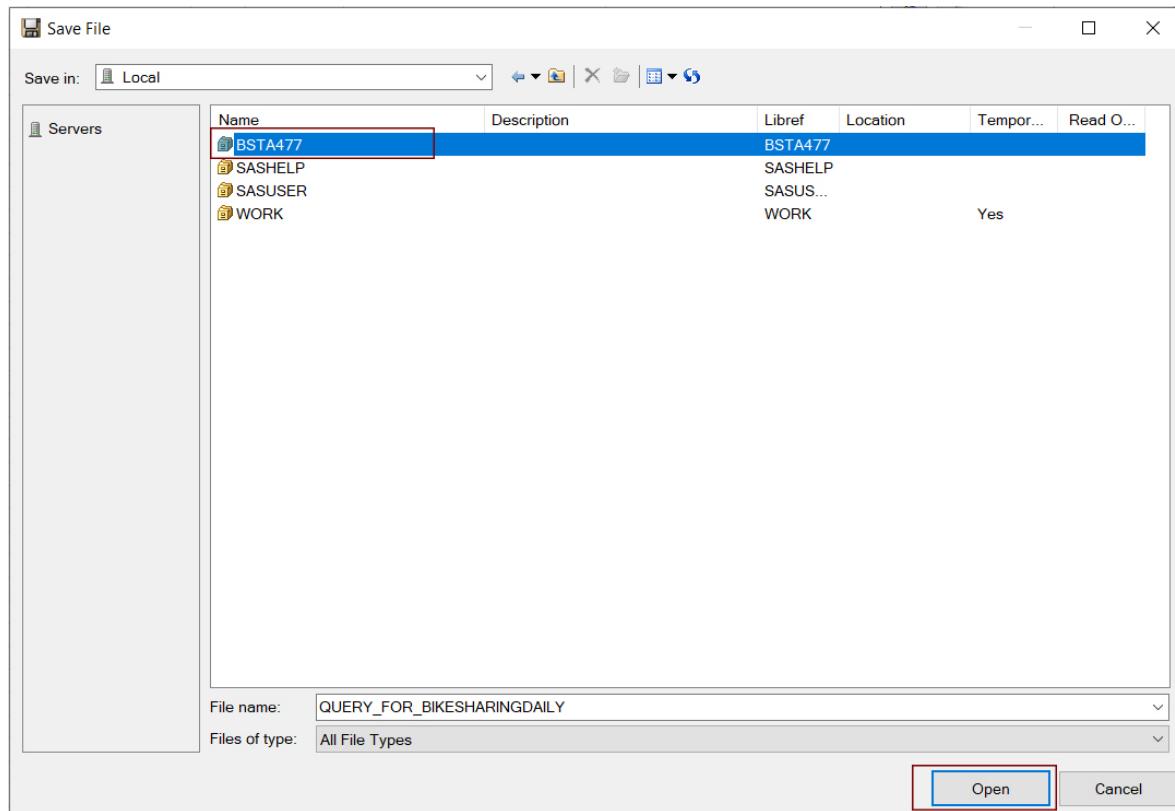
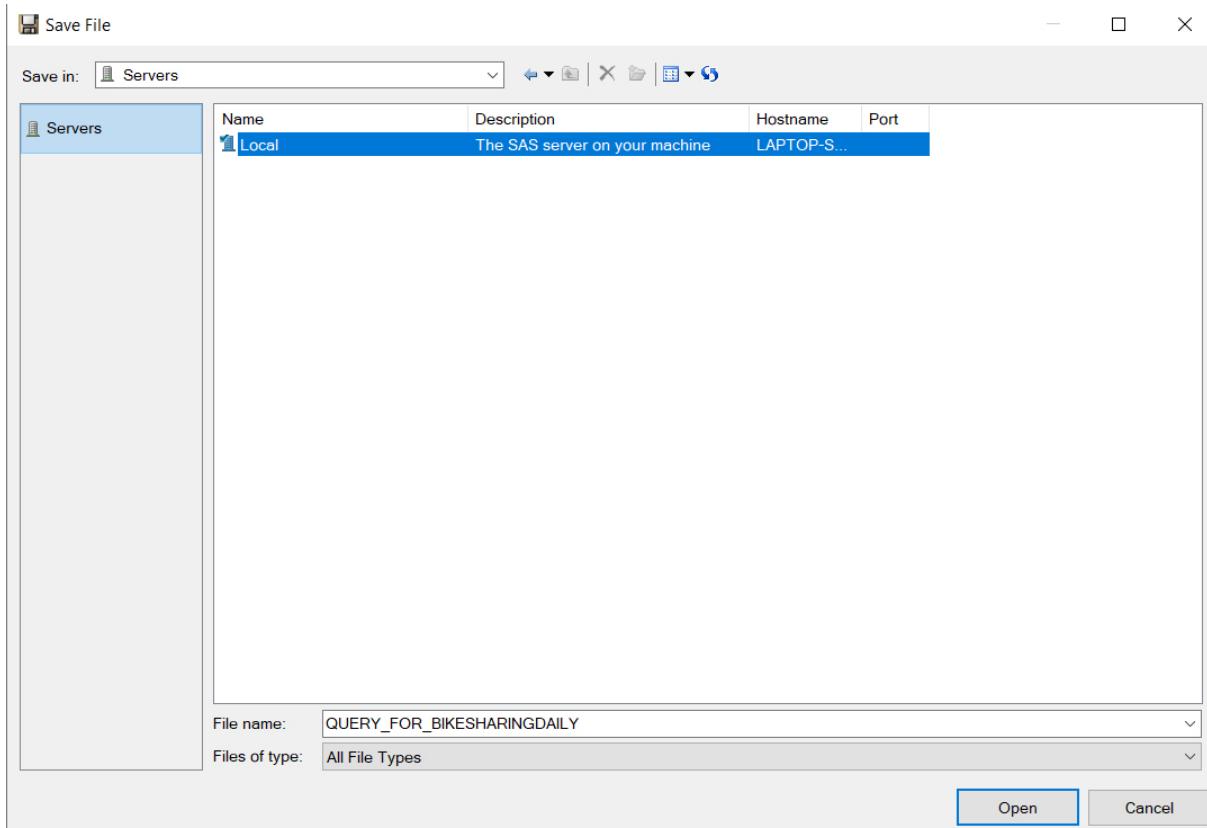
Query Builder for Local:BSTA477.BIKESHARINGDAILY

Query name: Query Builder Output name: WORKQUERY\_FOR\_BIKESHARINGDAILY Change...

Computed Columns Prompt Manager Preview Tools Options

Add Tables Delete Join Tables Select Data Filter Data Sort Data

| Column Name | Source Column | Sub... | Format | Details |
|-------------|---------------|--------|--------|---------|
| instant     | t1.instant    |        |        |         |
| dteday      | t1.dteday     |        |        |         |
| season      | t1.season     |        |        |         |
| yr          | t1.yr         |        |        |         |
| mnth        | t1.mnth       |        |        |         |
| holiday     | t1.holiday    |        |        |         |
| weekday     | t1.weekday    |        |        |         |
| workingday  | t1.workingday |        |        |         |
| weathersit  | t1.weathersit |        |        |         |
| temp        | t1.temp       |        |        |         |
| atemp       | t1.atemp      |        |        |         |
| hum         | t1.hum        |        |        |         |
| windspeed   | t1.windspeed  |        |        |         |
| casual      | t1.casual     |        |        |         |
| registered  | t1.registered |        |        |         |
| cnt         | t1.cnt        |        |        |         |



Save File

Save in: BSTA477

| Name                   | Member ... | Indexed |
|------------------------|------------|---------|
| ALIEN_SAUCERS          | Data       |         |
| BIKESHARINGDAILY       | Data       |         |
| FILTER_FOR_ALIEN_SA... | Data       |         |
| OUTAGE                 | Data       |         |

File name: QUERY\_FOR\_BIKESHARINGDAILY

Files of type: All File Types

Save Cancel

RINGDAILY

Output name: BSTA477.QUERY\_FOR\_BIKESHARINGDAILY Change...

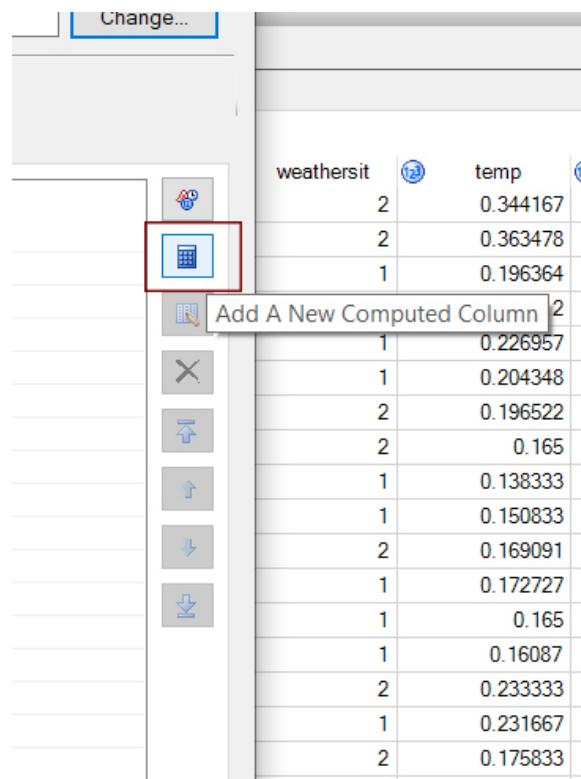
Preview Tools Options

Select Data Filter Data Sort Data

| Column Name | Source Column | Format | Details |
|-------------|---------------|--------|---------|
| instant     | t1.instant    |        |         |
| dteday      | t1.dteday     |        |         |
| season      | t1.season     |        |         |
| yr          | t1.yr         |        |         |
| mnth        | t1.mnth       |        |         |
| holiday     | t1.holiday    |        |         |
| weekday     | t1.weekday    |        |         |
| workingday  | t1.workingday |        |         |
| weathersit  | t1.weathersit |        |         |
| temp        | t1.temp       |        |         |
| atemp       | t1.atemp      |        |         |
| hum         | t1.hum        |        |         |
| windspeed   | t1.windspeed  |        |         |
| casual      | t1.casual     |        |         |
| registered  | t1.registered |        |         |
| cnt         | t1.cnt        |        |         |

Add A New Computed Column

| weathersit | temp     |
|------------|----------|
| 2          | 0.344167 |
| 2          | 0.363478 |
| 1          | 0.196364 |
| 1          | 0.226957 |
| 1          | 0.204348 |
| 2          | 0.196522 |
| 2          | 0.165    |
| 1          | 0.138333 |
| 1          | 0.150833 |
| 2          | 0.169091 |
| 1          | 0.172727 |
| 1          | 0.165    |
| 2          | 0.233333 |
| 1          | 0.231667 |
| 2          | 0.175833 |



| weathersit | temp     |
|------------|----------|
| 2          | 0.344167 |
| 2          | 0.363478 |
| 1          | 0.196364 |
| 1          | 0.226957 |
| 1          | 0.204348 |
| 2          | 0.196522 |
| 2          | 0.165    |
| 1          | 0.138333 |
| 1          | 0.150833 |
| 2          | 0.169091 |
| 1          | 0.172727 |
| 1          | 0.165    |
| 1          | 0.16087  |
| 2          | 0.233333 |
| 1          | 0.231667 |
| 2          | 0.175833 |

New Computed Column

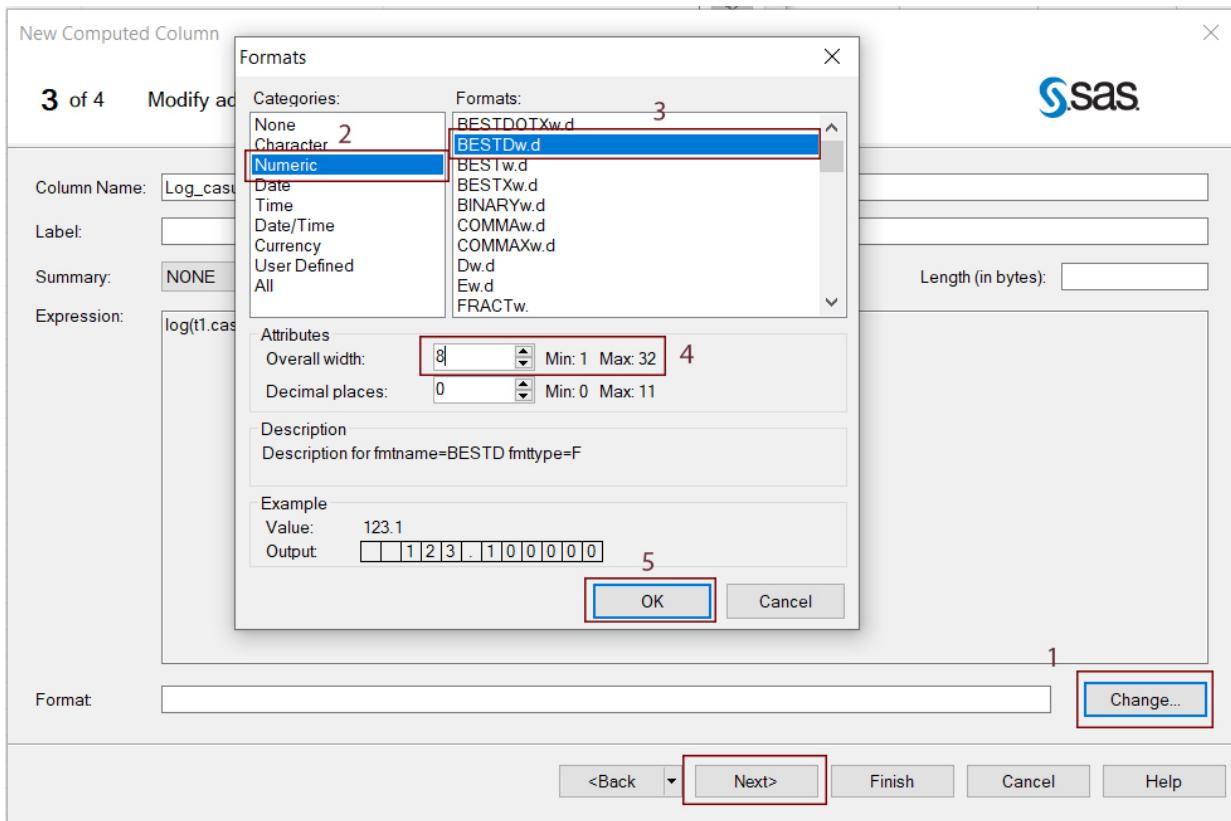
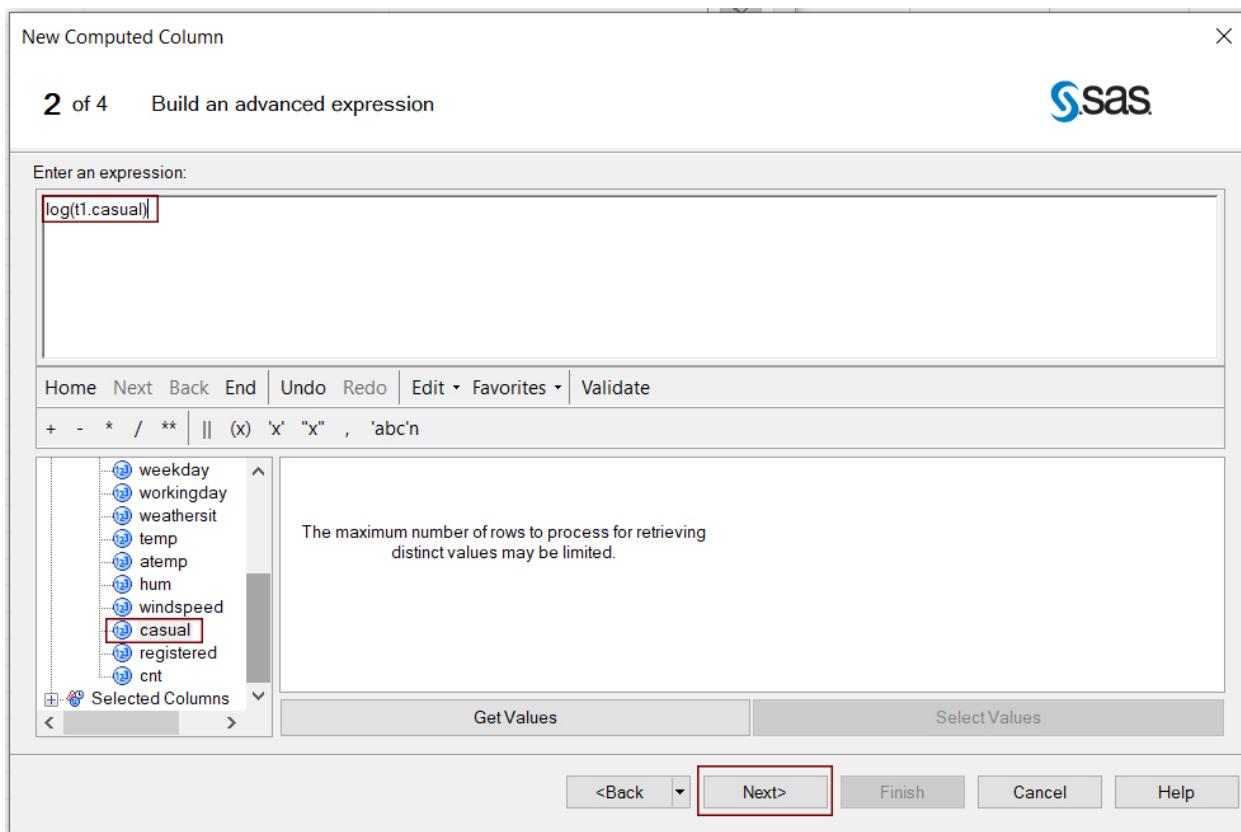
1 of 4 Select a type

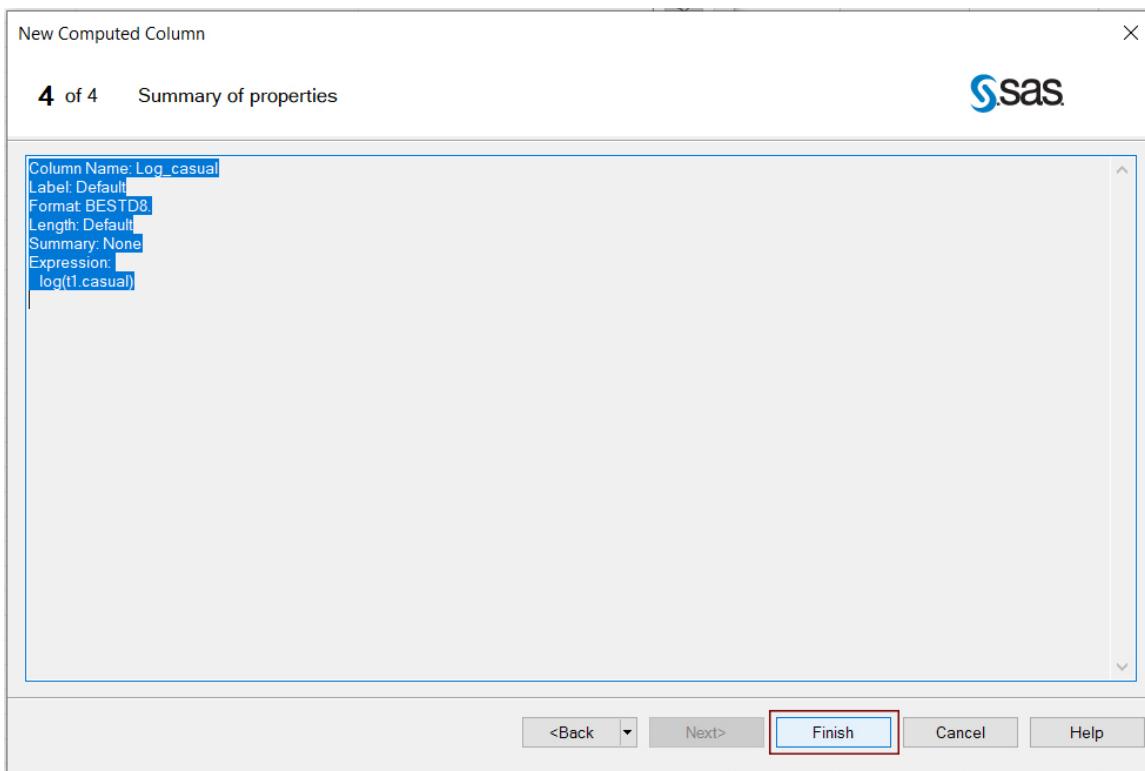


Summarized column  
 Recoded column  
 Advanced expression  
 From an existing computed column

Convert to an advanced expression

<Back | **Next>** | Finish | Cancel | Help





| Column Name       | Source Column | Su...   | Format         | Details |  |
|-------------------|---------------|---------|----------------|---------|--|
| t1.instant        | t1.instant    |         |                |         |  |
| dteday            | t1.dteday     |         |                |         |  |
| season            | t1.season     |         |                |         |  |
| yr                | t1.yr         |         |                |         |  |
| mnth              | t1.mnth       |         |                |         |  |
| holiday           | t1.holiday    |         |                |         |  |
| weekday           | t1.weekday    |         |                |         |  |
| workingday        | t1.workingday |         |                |         |  |
| weathersit        | t1.weathersit |         |                |         |  |
| temp              | t1.temp       |         |                |         |  |
| atemp             | t1.atemp      |         |                |         |  |
| hum               | t1.hum        |         |                |         |  |
| windspeed         | t1.windspeed  |         |                |         |  |
| casual            | t1.casual     |         |                |         |  |
| registered        | t1.registered |         |                |         |  |
| cnt               | t1.cnt        |         |                |         |  |
| <b>Log_casual</b> | Computed      | BESTD8. | log(t1.casual) |         |  |

a new field

Select distinct rows only

Run Save and Close Cancel Help

Now we are ready for forecasting.