B_c_2 : Autoviz

Utilisation de l'outil AutoViz pour une visualisation automatique des données.

Référence : https://github.com/AutoViML/AutoViz

In [1]:
```
%load_ext jupyter_black
import pandas as pd
import os
from autoviz import AutoViz_Class
```

```
Imported v0.1.720. After importing autoviz, execute '%matplotlib inline' to display ch
arts inline.
    AV = AutoViz_Class()
    dfte = AV.AutoViz(filename, sep=',', depVar='', dfte=None, header=0, verbose=1, lo
wess=False,
                      chart_format='svg',max_rows_analyzed=150000,max_cols_analyzed=30, save_
plot_dir=None)
```

In [2]:
```
path_to_interim_data = "../data/interim/"
demande_meteo_parquet = "demande_meteo.parquet"

df_import = pd.read_parquet(
    path=os.path.join(path_to_interim_data, demande_meteo_parquet),
    engine="pyarrow",
)
df = df_import["20190101":"20221231"]
```

In [3]:
```
AV = AutoViz_Class()
dft = AV.AutoViz(
    "",
    # sep=",",
    depVar="",
    dfte=df,
    header=0,
    verbose=1,
    lowess=False,
    chart_format="bokeh",
    max_rows_analyzed=150000,
    max_cols_analyzed=30,
    save_plot_dir=None,
)
```

```
Shape of your Data Set loaded: (35068, 2)
#############################################################################################
#
######################## C L A S S I F Y I N G  V A R I A B L E S  ###################
#
#############################################################################################
#
Classifying variables in data set...
    2 Predictors classified...
        No variables removed since no ID or low-information variables found in data se
t
No scatter plots with depVar when no depVar is given.
```

X-Axis

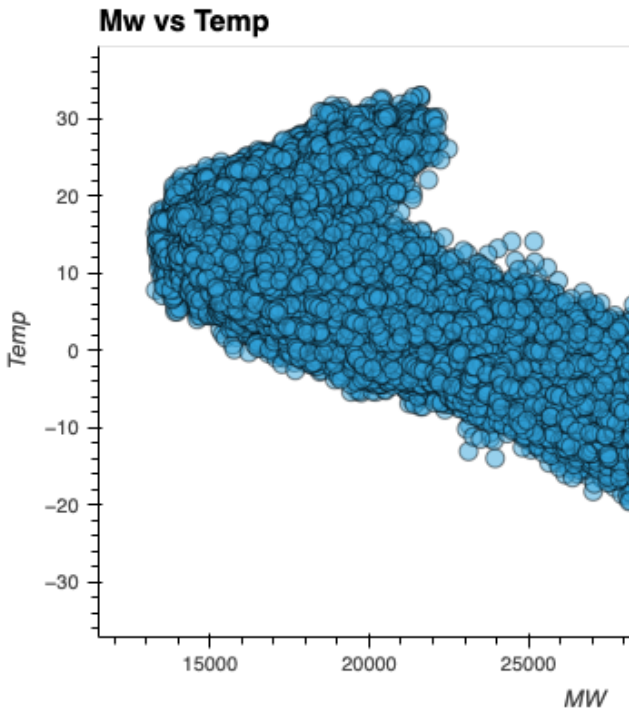
MW

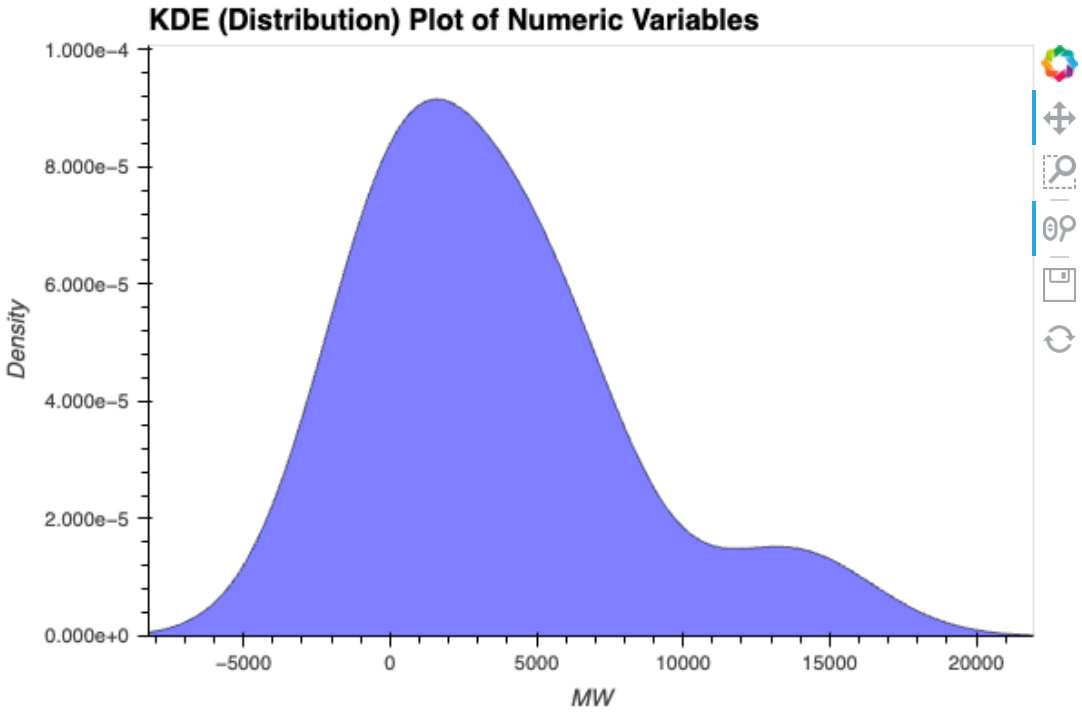
Y-Axis

Temp

Color

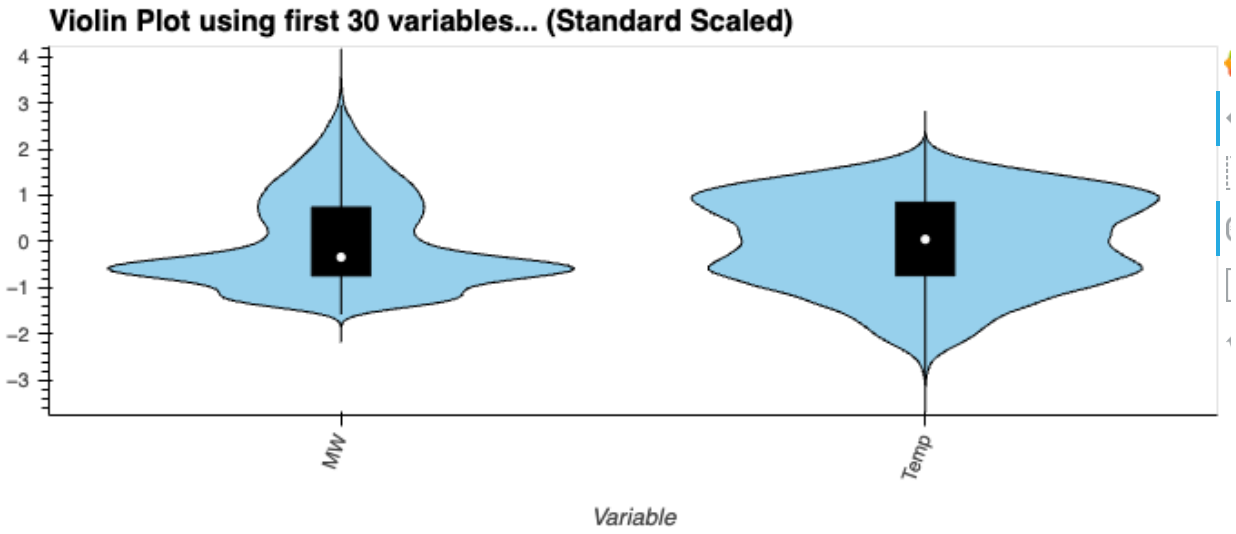
None

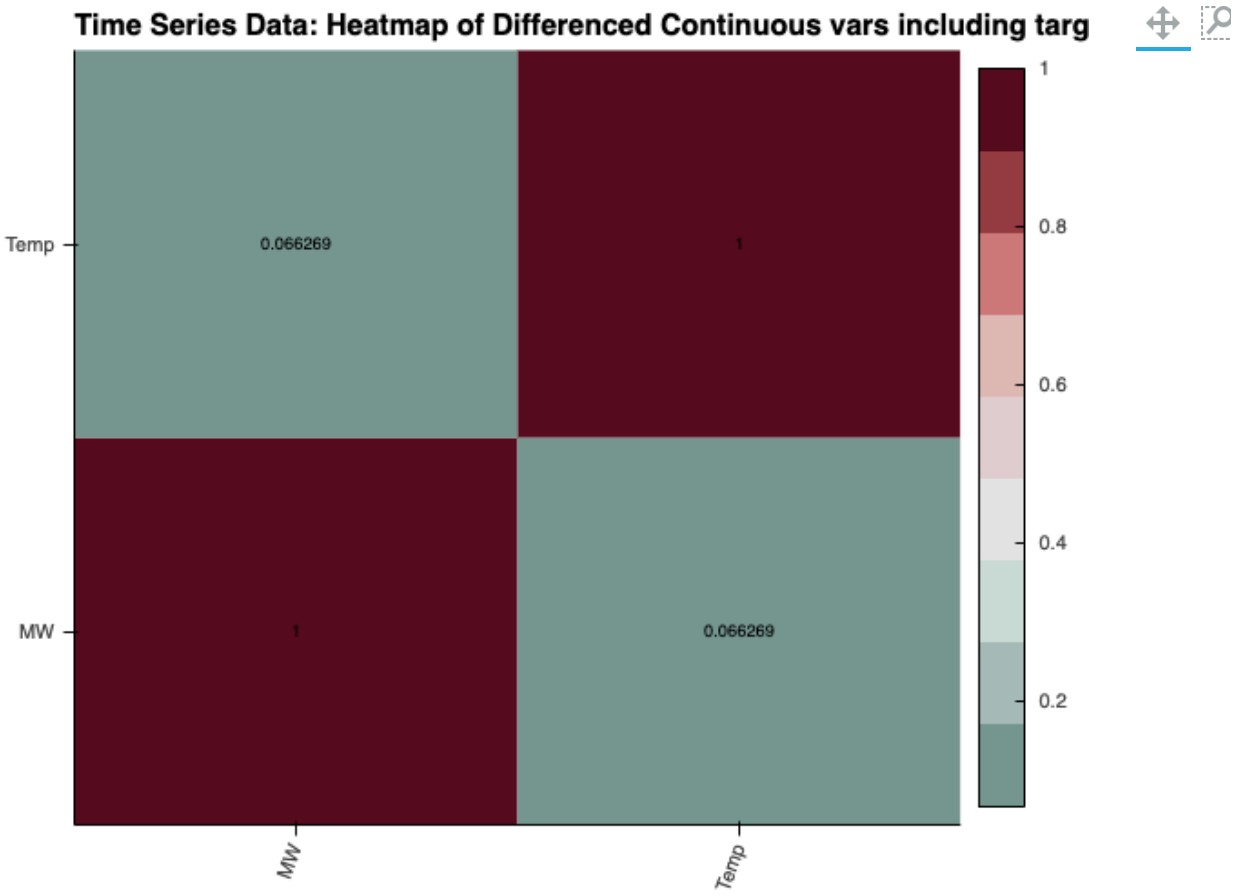




Select_

MW






```
Time to run AutoViz (in seconds) = 0
```

Nous pensions obtenir plus de visualisation avec cet outil, mais comme nous avons peu de colonne à analyser, le tout est limité. Nous pourrons possiblement y revenir après avoir réalisé les *features*.

En attendant, nous réutiliserons le *scatterplot* entre les MW et Temp, qui permettent de voir une pointe en forme de boomerang.