**What is AutoViz?**

AutoViz (Automatic Visualization) is a python library used for generating insightful visualizations, with just one line of code. AutoViz is relatively easy to use and is meant for beginners, intermediates, and experts! It is automatic (don't need to define the type of visualization) and efficient in handling large and complex datasets. AutoViz should be used for initial data exploration and quick overviews. Some of the output plots include: distribution plot, heat map, pairwise scatter plot, scatter plot, and violin plot.

**Getting started with AutoViz**

To begin using Autoviz, we must first install it and import AutoViz\_Class. We then instantiate a class. %matplotlib inline is needed to display charts inline when using Jupyter Notebooks. We cannot use '%matplotlib inline' when using PyCharm. We are then able to give AutoViz our desired dataset and it will automatically return visualizations.

**How to Install & First Example:**

pip install autoviz

from autoviz.AutoViz\_Class import AutoViz\_Class​

import numpy as np

import pandas as pd

AV = AutoViz\_Class()​

%matplotlib inline​

filename = "data/Housing.csv"

df = AV.AutoViz(filename)

We get a DQ issue output, pairwise scatter plot of all continuous variables, violin plot for continuous variables, and a heat map. Output is in svg format

**FixDQ() Function**

AutoViz provides data quality assessment by default and helps you fix data quality issues using the FixDQ() function​.

from autoviz import FixDQ​

fixdq = FixDQ()

fixdq.fit\_transform(filename)

We can use 'fixdq.fit\_transform(filename)' to provide a quick fix for our dataset.

**Second Example: AutoViz with a dependent variable, other arguments, and default output**

**Jupyter Notebooks**

target\_variable = "price"  
  
AV = AutoViz\_Class()  
  
dft = AV.AutoViz(  
 filename,

sep = “,”,

depVar= target\_variable

lowess = True

save\_plot\_dir = None  
)

We have used a target variable (price), a csv file for our data, and requested that there is a regression line shown (lowess=True). We get the same plots as above, but with each independent variable against our dependent variable. The output is in svg/png/jpg format (the default behavior) and shown in Jupyter Notebooks.

**Third Example: AutoViz with a dependent variable, pandas dataframe, and HTML output**

**PyCharm**

from autoviz import AutoViz\_Class  
import pandas as pd  
  
filename = pd.read\_csv(r"C:\Users\sgnew\Documents\OIM7502\_SP24\data\Housing.csv")  
target\_variable = "price"  
  
AV = AutoViz\_Class()  
  
dft = AV.AutoViz(  
 "",  
 sep="",  
 depVar=target\_variable,  
 dfte=filename,  
 chart\_format="html",  
 save\_plot\_dir=None  
)

We have used PyCharm to use HTML chart format. In this example, we also used a pandas dataframe. Notice how we did not use %matplot, like we did using Jupyter. Our output looks a little different than when we use svg/png/jpg chart format. We get an interactive distribution plot, heat map, an interactive pairwise scatter plot, an interactive scatter plot, and a violin plot.