Data Cleaning Project

Cleaning the Wild Blueberry Dataset

Author: Ashish Patel

Date: 2020-12-12

Dataset: https://www.kaggle.com/lucasb/wild-blueberry-yield-prediction?select=WildBlueberryPollinationSimulationData.csv

```
In [1]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
        %matplotlib inline
        from sklearn.preprocessing import MinMaxScaler
        from sklearn.preprocessing import StandardScaler
        from sklearn.preprocessing import Normalizer
        from sklearn.preprocessing import LabelEncoder
        from sklearn.preprocessing import OrdinalEncoder
        from sklearn.preprocessing import OneHotEncoder
        from sklearn.preprocessing import Binarizer
        from sklearn.preprocessing import KBinsDiscretizer
        from sklearn.preprocessing import PolynomialFeatures
        from sklearn.preprocessing import PowerTransformer
        from sklearn.preprocessing import QuantileTransformer
        from sklearn.preprocessing import RobustScaler
        from sklearn.preprocessing import FunctionTransformer
        from sklearn.preprocessing import KernelCenterer
        from sklearn.preprocessing import MaxAbsScaler
        from sklearn.preprocessing import Imputer
        from sklearn.preprocessing import LabelBinarizer
        from sklearn.preprocessing import MultiLabelBinarizer
        from sklearn.preprocessing import scale
        from sklearn.preprocessing import minmax_scale
        from sklearn.preprocessing import maxabs_scale
        from sklearn.preprocessing import robust_scale
        from sklearn.preprocessing import normalize
        from sklearn.preprocessing import binarize
        from sklearn.preprocessing import add_dummy_feature
        from sklearn.preprocessing import PadaBoxer
        from sklearn.preprocessing import StandardScalerModel
        from sklearn.preprocessing import LabelEncoderModel
        from sklearn.preprocessing import OrdinalEncoderModel
        from sklearn.preprocessing import OneHotEncoderModel
        from sklearn.preprocessing import BinarizerModel
        from sklearn.preprocessing import KBinsDiscretizerModel
        from sklearn.preprocessing import PolynomialFeaturesModel
        from sklearn.preprocessing import PowerTransformerModel
        from sklearn.preprocessing import QuantileTransformerModel
        from sklearn.preprocessing import RobustScalerModel
        from sklearn.preprocessing import FunctionTransformerModel
        from sklearn.preprocessing import KernelCenterer
        from sklearn.preprocessing import MaxAbsScalerModel
        from sklearn.preprocessing import ImputerModel
        from sklearn.preprocessing import LabelBinarizerModel
        from sklearn.preprocessing import MultiLabelBinarizerModel
```