6a Apply and explore various plotting functions on UCI data sets.

Dataset: Iris

a. Normal curves

b. Density and contour plots

c. Correlation and scatter plots

d. Histograms

e. Three dimensional plotting

import numpy as np

import matplotlib.pyplot as plt

from scipy.stats import norm

import pandas as pd

import seaborn as sns

from mpl\_toolkits.mplot3d import Axes3D

df=pd.read\_csv("iris.data")

df.columns = ['SepalLengthCm','SepalWidthCm','PetalLengthCm','PetalWidthCm','class']

df1=df['SepalLengthCm']

df2=df['SepalWidthCm']

df3=df['PetalLengthCm']

print('HEAD =\n',df1.dtype,df1.head())

#normal curve

mu = df1.mean()

sigma = df1.std()

print('mean, std devia',mu,sigma)

x = np.linspace(mu - 3\*sigma, mu + 3\*sigma, 1000)

plt.plot(x, norm.pdf(x, mu, sigma), color='red', linewidth=3)

plt.show()

#Density plots allow to visualize the distribution of a numeric variable for one or several groups.

df1.plot(kind = 'density')

plt.show()

#histogram

plt.hist(df['SepalLengthCm'], bins = 20, color = "green")

plt.title("Sepal Length in cm")

plt.xlabel("Sepal\_Length\_cm")

plt.ylabel("Count")

#correlation

corr = df.corr()

# According to the correlation matrix results PetalLengthCm and

#PetalWidthCm have possitive correlation which is proved by the plot above

print('Correlation amoung columns in iris dataset\n',corr)

#Scatterplot

df.plot(kind ="scatter",

x ='SepalLengthCm',

y ='PetalLengthCm')

plt.grid()

#Contour plots also called level plots are a tool for doing multivariate analysis and visualizing 3-D plots in 2-D space

# X, Y: 2-D numpy arrays with same shape as Z or 1-D arrays such that len(X)==M and len(Y)==N (where M and N are rows and columns of Z)

#Z: The height values over which the contour is drawn. Shape is (M, N)

# Creating 2-D grid of features

[X, Y] = np.meshgrid(df1,df2)

fig, ax = plt.subplots(1, 1)

Z = X \*\* 2 \* Y \*\* 2

# plots filled contour plot

ax.contourf(X, Y, Z)

ax.set\_title('Filled Contour Plot')

ax.set\_xlabel('SepalLengthCm')

ax.set\_ylabel('SepalWidthCm')

plt.show()

#3D Plotting

fig = plt.figure()

ax = fig.gca(projection='3d')

ax.plot\_trisurf(df1, df2, df3, cmap=plt.cm.viridis, linewidth=0.2)

plt.show()

6b. Download adult dataset

Attributes (Age,Workclass.,Final Weight.,Education.,Education Number of Years.,Marital-,status.,Occupation.,,elationship.,Race.,Sex.,Capital-gain.,Capital-loss.,Hours-per-week.,Native-country.)

from UCI repository and construct 2 normal curve one for age and another for final weight

6c. construct 2 density plots one for age and the other for final weight. And also construct histogram for age and final weight.

6d. Find the correlation between all attributes in the dataset and also final the correlation coefficient between age and final weight. Then construct scatter plot for age Vs final weight.

6e. construct contour plot for age, final weight and relationship sin(age)^2+cos(final weight)^2 and also construct a 3D scatter plot between age, final weight and sin(age)^2+cos(final weight)^2