**41.How to check whether residuals are normally distributed or not using python?**

**Objective:**

* To check distribution of residuals in python.

**Process:**

* Import the libraries.
* Read the sample data.
* Take model summary.
* Check residuals are follows normal distribution or not.
* Plot the residuals.

**Input:**

* Data set(csv file).

**Output:**

* Residuals plots
* qq plot for residuals.

**Source code:**

#import libraries

import scipy.stats as stats

import statsmodels.api as sm

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

#read the data set

data=pd.read\_csv('/home/soft27/soft27/Sathish/

Pythonfiles/Employee.csv')

#creating data frame

df=pd.DataFrame(data)

#assigning the independent variable

X = df[['rating','bonus']]

#assigning the dependent variable

Y = df['salary']

#Build multiple linear regression

X = sm.add\_constant(X)

#fit the variables in to the linear model

model = sm.OLS(Y, X, hasconst=True).fit()

#print the intercept and regression co-efficients

print\_model = model.summary()

print(print\_model)

#residuals visualization

mu = np.mean(model.resid)

sigma = np.std(model.resid)

pdf = stats.norm.pdf(sorted(model.resid), mu, sigma)

plt.hist(model.resid, bins=50, normed=True)

plt.plot(sorted(model.resid), pdf, color='r', linewidth=2)

plt.show()

#qq plot

fig, [ax1, ax2] = plt.subplots(1,2, figsize=(10,3))

sm.qqplot(model.resid, stats.t, fit=True, line='45',

ax = ax1)

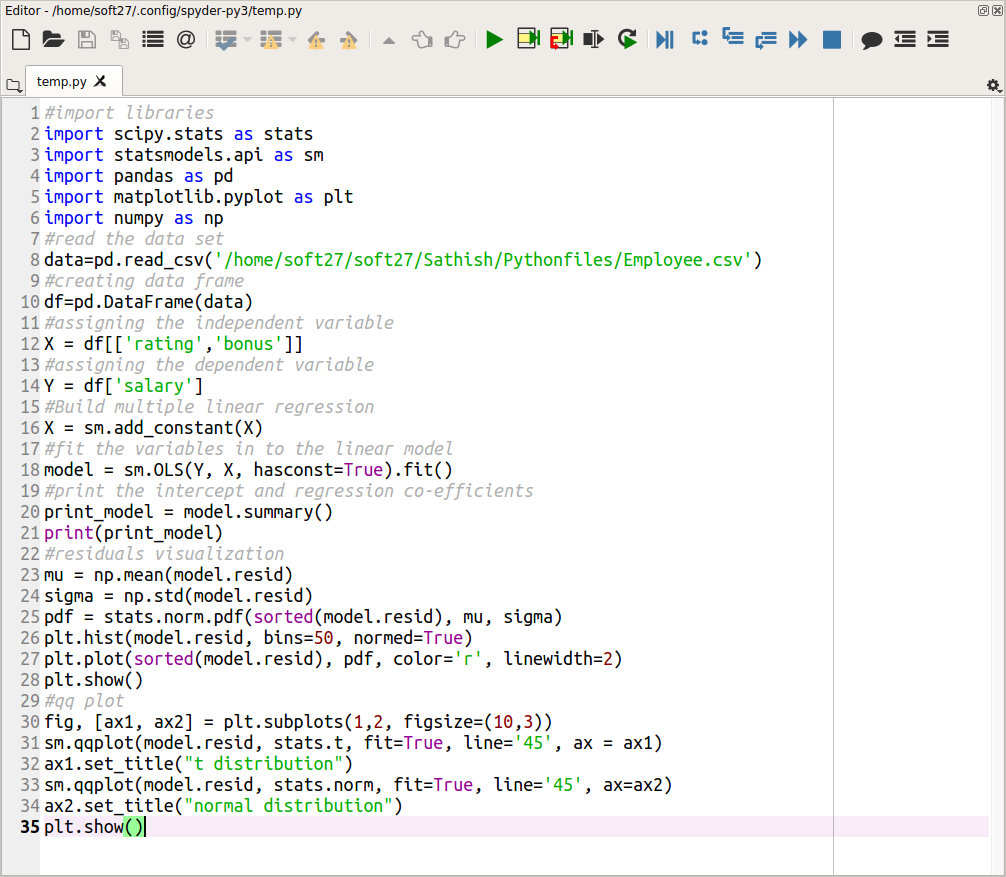
ax1.set\_title("t distribution")

sm.qqplot(model.resid, stats.norm, fit=True, line='45', ax=ax2)

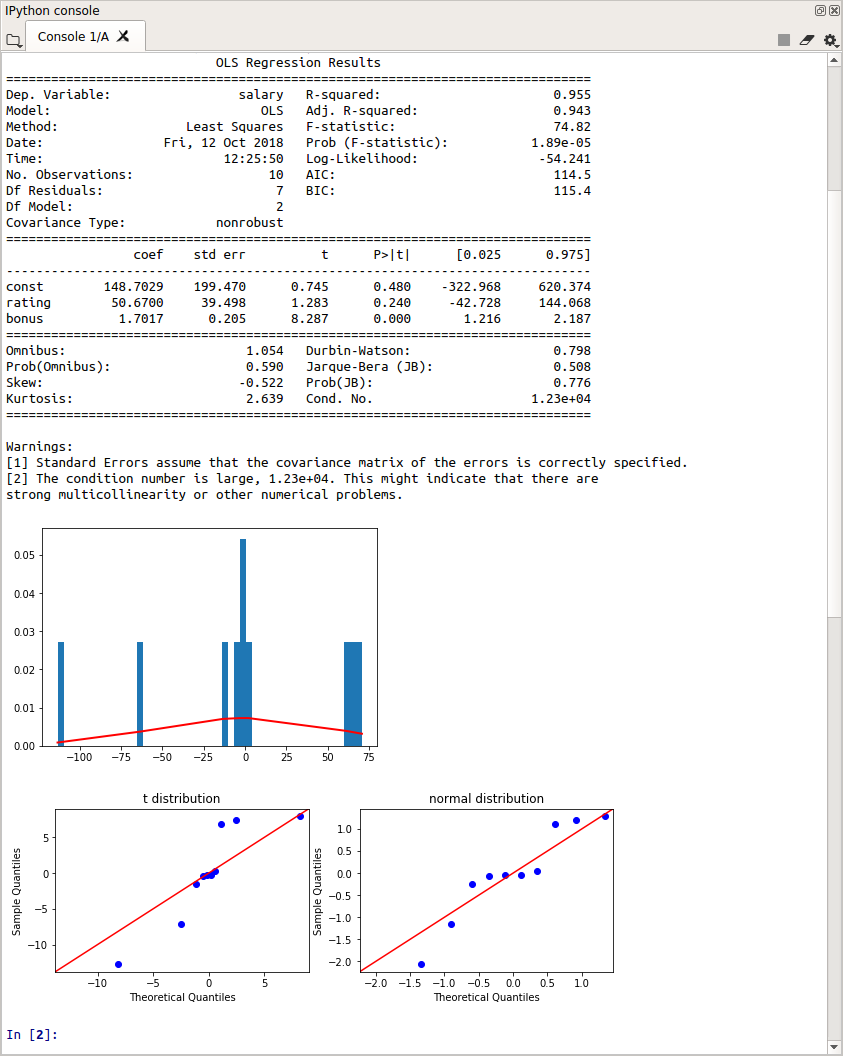
ax2.set\_title("normal distribution")

plt.show()

**Screen shots:**

****

**Output:**

****