**MULTIVARIATE LINEAR REGRESSION**

**Using Technology: U.S. Economy Case Study**

U.S.. economic data 1976 to 1987

X1 = dollars/barrel crude oil

X2 = % interest on ten yr. U.S. treasury notes

X3 = foreign investments/billions of dollars

X4 = Dow Jones industrial average

X5 = GNP/billions of dollars

X6 = purchasing power U.S. dollar (1983 base)

X7 = consumer debt/billions of dollars

*Reference: Statistical Abstract of the United States 103rd and 109th ediiton*

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

df = pd.read\_csv('tech\_economy.csv')

df

df.describe

x = df.iloc[:, :-1].values

print(x)

y = df.iloc[:, -1].values

print(y)

from sklearn.linear\_model import LinearRegression

regressor = LinearRegression()

regressor.fit(x,y)

y\_pred = regressor.predict(x)

np.set\_printoptions(precision=2)

print(np.concatenate((y.reshape(len(y),1), y\_pred.reshape(len(y\_pred),1)) ,1))

print(regressor.predict([[20, 10, 300, 2500, 4700 ,0.80]]))

print(regressor.coef\_)

print(regressor.intercept\_)

**Equation for the model is:**

consumer debt = -854.387 - 0.804(dollars/barrel) - 4.89(% interest on 10 yrs) -1.67(foreign investments/billions) - 0.0033 (Dow Jones industrial average) + 38.5 (GNP/billions of dollars) + 299 (purchasing power US dollar)

from sklearn.metrics import mean\_squared\_error

mean\_squared\_error(y, y\_pred)