Q3#

```
import numpy as np
from sklearn import linear_model
file1 = open('train.txt', 'r')
Lines = file1.readlines()

 # [
 # "Density underwater weighing", x
 # "Percent body fat",
 # "Age",
 # "Weight",
 # "Neck circumference",
 # "Chest circumference",
 # "Abdomen 2 circumference",
 # "Hip circumference",
 # "Thigh circumference",
 # "Knee circumference",
 # "Ankle circumference",
 # "Biceps (extended) circumference",
 # "Forearm circumference",
 # "Wrist circumference"]
data = []
# Strips the newline character
for line in Lines:
        # slpit list of string and remove the newline tag
        values = line.strip().split("  ")
        # removing empty values from list
        values = list(filter(None, values))
        # converting string to float values
        for i in range(0, len(values)):
                values[i] = float(values[i])
        # list of data
        data.append(values)
# convert data to numpy array
data = np.array(data)

# y is Percent body fat
y = data[:,1]

# excluding the body density
X = data[:,2:15]
```

```
xo = []
for x in range(252):
        xo.append([1])
# add xo to features
X= np.append(X, xo, axis=1)

# calculate W0 to W13

X_transpose = np.transpose(X)
result_inverse =(np.linalg.inv((X_transpose.dot(X))))
xIntoY = X_transpose.dot(y)
Weights = (result_inverse.dot(xIntoY))
print (Weights)

# ###############
#   out put   #
# ###############

# [ 6.20786464e-02 -8.84446759e-02 -6.95904296e-02 -4.70600014e-01
#  -2.38641465e-02  9.54773458e-01 -2.07541123e-01  2.36099845e-01
#   1.52812146e-02  1.73995368e-01  1.81602416e-01  4.52024914e-01
#  -1.62063910e+00 -1.81884851e+01]


# print (X.dot(X_transpose))
# # y=y.astype('int')
# # clf = linear_model.LogisticRegression(C=1e40, solver='newton-cg')
# # fitted_model = clf.fit(X, y)

# fitted_model = linear_model.LinearRegression().fit(X, y)
# print ("reg score ",fitted_model.score(X, y))
# print ("coef ",fitted_model.coef_)
# print ("intercept_ ",fitted_model.intercept_)
# # example = [23,154.25,67.75,36.2,93.1,85.2,94.5,59.0,37.3,21.9,32.0,27.4,17.1]

# # testing on model
# example = [27,133.25,64.75,36.4,93.5,73.9,88.5,50.1,34.5,21.3,30.5,27.9,17.2]

# print ("predict : ",fitted_model.predict(np.array([example])))
```

Q4#

```
import pandas
import numpy as np
from sklearn import linear_model
#load the csv file
df = pandas.read_csv('heights_weights.csv')
# update the value with numbers
df['Gender'] = df['Gender'].replace(['Female'],0)
df['Gender'] = df['Gender'].replace(['Male'],1)

# convert to numpy array
data = df.to_numpy()

# taking the second and thrird row
X = data[:,1:3]
# take the first row
y = data[:,0]

# Fit (train) the Logistic Regression classifier
clf = linear_model.LogisticRegression(C=1e40, solver='newton-cg')
fitted_model = clf.fit(X, y)

prediction_result = clf.predict([(70,180)])

if prediction_result == 1:
        print ("prediction : Male")
elif prediction_result == 0:
        print ("prediction : Female")

print ("reg score ",fitted_model.score(X, y))
print ("coef ",fitted_model.coef_)
print ("intercept_ ",fitted_model.intercept_)

# Predict
# prediction : Male
# reg score  0.9194
# coef  [[-0.49261999  0.19834042]]
# intercept_  [0.69254178]
```