Name: Amiabh Mahato (Amit Mahato)

1. Read in the CSV file “ML\_HW\_Data\_FisherIris.csv” into a matrix named as

“*Iris*”. Please do NOT output the whole matrix in our answer.

Ans:

Iris **=** pd.read\_csv('ML\_HW\_Data\_FisherIris.csv', header **= None**)

Iris.head()

0 1 2 3 4

0 2 14 33 50 0

1 24 56 31 67 1

2 23 51 31 69 1

3 2 10 36 46 0

4 20 52 30 65 1

2. Display **total number of rows** and total number of columns of the matrix “***Iris***”.

Ans: rows = Iris.shape[0]

cols = Iris.shape[1]

print("Total number of rows:" + str(rows))

print("Total number of columns:" + str(cols))

Total number of rows:150

Total number of columns:5

3. Display all the **row numbers** (i.e. record numbers) that have the 5th column less

than 0.

Ans: Iris.loc[Iris[4] < 0]

0 1 2 3 4

10 2 16 31 48 -10

23 15 45 29 60 -12

58 2 13 35 55 -10

89 12 47 28 61 -12

108 13 52 30 67 -11

136 18 63 29 73 -11

4. Remove the rows with the 5th column less than 0 from the “***Iris***” matrix. Please do

NOT output the whole resulting matrix in our answer.

Ans: Iris = Iris.loc[Iris[4] >= 0]

Iris.head(10)

0 1 2 3 4

0 2 14 33 50 0

1 24 56 31 67 1

2 23 51 31 69 1

3 2 10 36 46 0

4 20 52 30 65 1

5 19 51 27 58 1

6 13 45 28 57 2

7 16 47 33 63 2

8 17 45 25 49 1

9 14 47 32 70 2

5. Display **total number of rows** and **total number of columns** of the “***Iris***” matrix

again.

Ans: rows = Iris.shape[0]

cols = Iris.shape[1]

print("Total number of rows:" + str(rows))

print("Total number of columns:" + str(cols))

Total number of rows:144

Total number of columns:5

6. Copy the first 4 columns in the new “***Iris***” matrix into a new matrix “***X***”. Please

do NOT output the whole resulting matrix in our answer.

Ans: selected\_columns = Iris[[0, 1, 2, 3]]

X = selected\_columns.copy()

X.head()

0 1 2 3

0 2 14 33 50

1 24 56 31 67

2 23 51 31 69

3 2 10 36 46

4 20 52 30 65

7. Copy the 5th columns in the new “***Iris***” matrix into a new variable (or matrix) “***Y***”.

Please do NOT output the whole resulting matrix in our answer.

Ans: select = Iris[[4]]

Y = select.copy()

Y.head()

4

0 0

1 1

2 1

3 0

4 1

8. Display the **maximum value** and the **minimum value** of **EACH column** in “***X***”.

Ans: col0 = X[[0]]

col0\_max\_value = col0.max()

col0\_min\_value = col0.min()

print('Max value of column' + str(col0\_max\_value))

print('Min value of column' + str(col0\_min\_value))

Max value of column0 25

dtype: int64

Min value of column0 1

dtype: int64

Max value of column1 69

dtype: int64

Min value of column1 10

dtype: int64

Max value of column2 44

dtype: int64

Min value of column2 20

dtype: int64

Max value of column3 79

dtype: int64

Min value of column3 43

dtype: int64

9. Display **total number of elements** (i.e. items) in the third column of the matrix

“***X***” that are greater than 36.

Ans: column2 = X[2]

count = 0

for i in column2 :

if i > 36 :

count = count + 1

print ("Total number of elements in third column greater than 36 : " + str(count))

Total number of elements in third column greater than 36 : 15