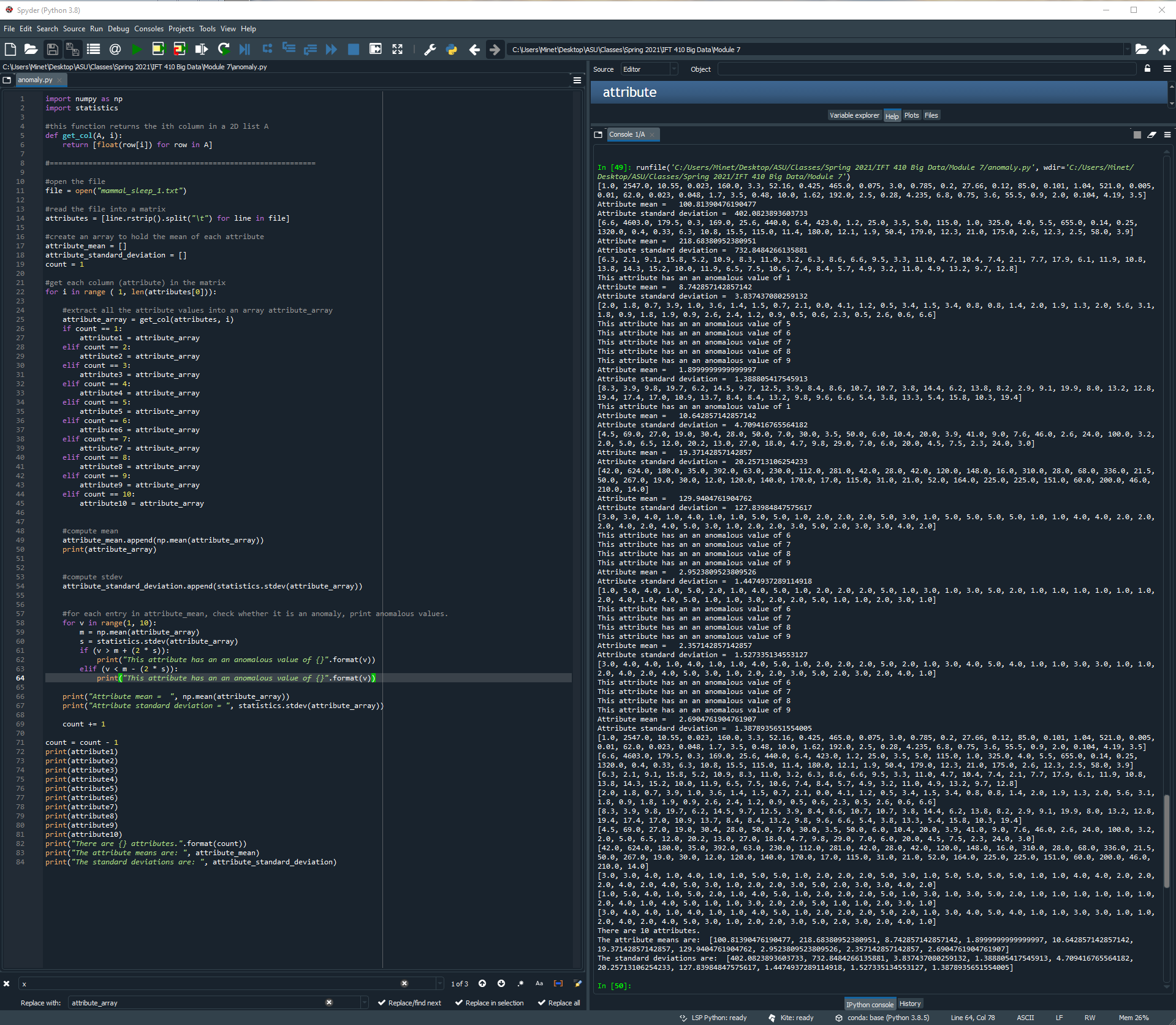
# Lab 9: Statistical Methods for Anomaly Detection



import numpy as np

import statistics

#this function returns the ith column in a 2D list A

def get\_col(A, i):

return [float(row[i]) for row in A]

#==============================================================

#open the file

file = open("mammal\_sleep\_1.txt")

#read the file into a matrix

attributes = [line.rstrip().split("\t") for line in file]

#create an array to hold the mean of each attribute

attribute\_mean = []

attribute\_standard\_deviation = []

count = 1

#get each column (attribute) in the matrix

for i in range ( 1, len(attributes[0])):

#extract all the attribute values into an array attribute\_array

attribute\_array = get\_col(attributes, i)

if count == 1:

attribute1 = attribute\_array

elif count == 2:

attribute2 = attribute\_array

elif count == 3:

attribute3 = attribute\_array

elif count == 4:

attribute4 = attribute\_array

elif count == 5:

attribute5 = attribute\_array

elif count == 6:

attribute6 = attribute\_array

elif count == 7:

attribute7 = attribute\_array

elif count == 8:

attribute8 = attribute\_array

elif count == 9:

attribute9 = attribute\_array

elif count == 10:

attribute10 = attribute\_array

#compute mean

attribute\_mean.append(np.mean(attribute\_array))

print(attribute\_array)

#compute stdev

attribute\_standard\_deviation.append(statistics.stdev(attribute\_array))

#for each entry in attribute\_mean, check whether it is an anomaly, print anomalous values.

for v in range(1, 10):

m = np.mean(attribute\_array)

s = statistics.stdev(attribute\_array)

if (v > m + (2 \* s)):

print("This attribute has an an anomalous value of {}".format(v))

elif (v < m - (2 \* s)):

print("This attribute has an an anomalous value of {}".format(v))

print("Attribute mean = ", np.mean(attribute\_array))

print("Attribute standard deviation = ", statistics.stdev(attribute\_array))

count += 1

count = count - 1

print(attribute1)

print(attribute2)

print(attribute3)

print(attribute4)

print(attribute5)

print(attribute6)

print(attribute7)

print(attribute8)

print(attribute9)

print(attribute10)

print("There are {} attributes.".format(count))

print("The attribute means are: ", attribute\_mean)

print("The standard deviations are: ", attribute\_standard\_deviation)