**CS612 - Lab 3**

**Q1: Do the follow Questions**

1. Distinguish between noise and outliers. Be sure to consider the following questions.
2. Is noise ever interesting or desirable? Outliers? Noise is not interesting or

**Noise in attributes is undesirable by default, as it distorts the original attribute values. Outliers can potentially be legitimate data objects (or values), i.e. identifying them can be the main objective of some data mining tasks. Thus, outliers can potentially be interesting/desirable, but noise is not**

1. Can noise objects be outliers?

**Yes, noise objects sometimes can be considered as outliers., Noise in attribute values can make the data look more randomized or unusual. Thus, it is possible that some instances in noisy data will appear as outlier**

1. Are noise objects always outliers?

**No, noise objects are not always outliers. The noise points can be mixed together with the non-noise point(normal data). So noise objects are not always outliers.**

1. Are outliers always noise objects?

**No, outliers are not always noise objects because outliers can be objects of interest in some data mining applications, such as network intrusion detection. So these outliers can be legitimate data objects that appear to not belong in the data set. Those outliers would typically not classify as noise objects.**

1. Consider the following sets of values. For each set find the mean, measurement and Bias

****

**Bias=Mean - Known Value**

* 1. Set A = {1.2, 0.87, 1.013, 1.0001, and 0.986},

**Standard Deviation, s: 0.11864869152249**

|  |  |
| --- | --- |
| **Bias:** | **0.01382** |
| **Sum, Σx:** | **5.0691** |
| **Mean, x̄:** | **1.01382** |
| **Variance, s2:** | **0.01407751** |

* 1. Set B = {6.215, 6.23, 6.33, 6.001, and 6.123},

**Standard Deviation, s: 0.12401088661888**

|  |  |
| --- | --- |
| **Bias:** | **0.1798** |
| **Sum, Σx:** | **30.899** |
| **Mean, x̄:** | **6.1798** |
| **Variance, s2:** | **0.0153787** |

* 1. Set C = {8.222, 8.452, 8.0001, 8.201, and 8.111},

**Standard Deviation, s**: **0.16720598673493**

|  |  |
| --- | --- |
| **Bias:** | **0.197** |
| **Sum, Σx:** | **40.9861** |
| **Mean, x̄:** | **8.19722** |
| **Variance, s2:** | **0.02795784** |

* 1. Set D = {12.205, 12.099, 12.103, 12.101, and 11.896}

**Standard Deviation, s: 0.11270403719477**

|  |  |
| --- | --- |
| **Bias:** | **0.0808** |
| **Sum, Σx:** | **60.404** |
| **Mean, x̄:** | **12.0808** |
| **Variance, s2:** | **0.0127022** |

**Question 2 – Working with Square Matrix (Programming Question)**

* Write an object oriented program in python that does the following:
* Part a)
  + Interactively get a positive number, “n”, from a user
    - If “n” is less or equal to 3 report an error and quit the program.
  + Otherwise, try to get n\*n numbers from the first file.
  + For example, if n=5 you are required to get 25 numbers from the first file in order to make a 5 by 5 matrix.
    - Therefore, If there is less than 25 numbers in the file, report error and quit the program
    - If there are more than 25 numbers (for size 5) from the file, get the first 25 numbers, make a matrix of 5 by 5 and ignore the other numbers
* Part b)
  + try to get n\*n numbers from the second file.
  + If the value of “n” to create the first matrix, you must use the same value “n” to create the second matrix.
  + The two matrix should have the same size. So if the first size of the first matrix is 5, the size of the second matrix must be 5 too.
  + Again read the integer number from the second file.
    - For n=5, If there is less than 25 numbers (for size = 5) in the second file, report error and quit the program
    - If there are more than 25 numbers from the file, get the first 25 numbers, make the second matrix of 5 by 5 and ignore the other numbers
* Part c)
  + Do the product of the two matrices and print the resulting matrix with proper heading
  + Do the dot product of the two matrices and print the resulting matrix with proper heading
  + Transpose the first and the second matrices and multiply them with each other and print the resulting matrix with proper heading
  + Do the dot product of the transposed matrices and print the resulting matrix with proper heading
  + Divide the first matrix by the second one. In case of division by zero you need to show the word “undefined” in the proper matrix location. Print the resulting matrix with proper heading

As I mentioned above, you need to create a class with several function. Your main program should call the appropriate class function and print all your output in an output file with proper heading

Main program should consist of the following:

1. N = get a size for a square matrix
2. Create an object call it myMatrix
3. M1 = myMatrix. GetMatrix(N, file1)
4. M2 = myMatrix. GetMatrix(N, file2)
5. M1\_Multiply\_M2 = myMatrix.Product(M1, M2)
6. M1\_DotMultiply\_M2 = myMatrix.DotProduct(M1, M2)
7. M1\_Trans = myMatrix.Transpose(M1)
8. M2\_Trans = myMatrix.Transpose(M2)
9. M1Trans\_Multiply\_M2Trans = myMatrix.Product(M1\_Trans, M2\_Trans)
10. M1Trans\_DotMultiply\_M2Trans = myMatrix.DotProduct(M1\_Trans, M2\_Trans)
11. myMatrix.Print(M1)
12. myMatrix.Print(M2)
13. myMatrix.Print(M1\_Multiply\_M2)
14. myMatrix.Print(M1\_DotMultiply\_M2)
15. myMatrix.Print(M1\_Trans)
16. myMatrix.Print(M2\_Trans)
17. myMatrix.Print(M1Trans\_Multiply\_M2Trans)
18. myMatrix.Print(M1Trans\_DotMultiply\_M2Trans)

The following explains a direction that you need to follow to provide the output of your program

1. First, place the following in a file and call that file “file1.txt”
2. 2 3 4 5 6

7 8 9 10 11 12

13 14 15 16 17 18

19 20 21 22 23 24

25 26 27 28 29 30

31 32 33 34 35 36

1. First, place the following in a file and call that file “file2.txt”
2. 200 30 40 50 60

70 80 90 100 110 120

130 140 150 160 170 180

190 200 210 22 230 240

250 260 270 28 290 300

310 320 330 34 350 360

Now run your program and you should see the following on the screen. Note that anything that is shown in red is what you need to enter as input to your program. Anything shown in blue is what your program prints on the screen as soon as the program starts. Now run the program and provide the input

**>> Enter the dimension of your matrix:** **2**

**Error: \*\*\*\*\* This dimension is out of bound. The program stops in here.**

**(\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*End of the Program \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*)**

Now re-execute your code and you should see the following

**>> Enter the dimension of your matrix:** **7**

**Error: \*\*\*\*\* We can only handle up to 6 dimension at this time. The program stops in here.**

**(\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*End of the Program \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*)**

Again re-execute your code and you should see the following

**>> Enter the dimension of your matrix:** **5**

**====================================================================================**

**The content of the first matrix is:**

**<print the content – tab delimited>**

**====================================================================================**

**The content of the second matrix is:**

**<print the content – tab delimited>**

**====================================================================================**

**The product of the two matrices is:**

**<print the result – tab delimited>**

**====================================================================================**

**The dot-product of the two matrices is:**

**<print the result – tab delimited>**

**====================================================================================**

**The result of matrix1 divided by matrix2 is:**

**<print the result – tab delimited>**

**====================================================================================**

**The transpose of the first matrix is:**

**<print the result – tab delimited>**

**====================================================================================**

**The transpose of the second matrix is:**

**<print the result – tab delimited>**

**====================================================================================**

**The product of the transpose of the two matrices is:**

**<print the result – tab delimited>**

**====================================================================================**

**The dot product of the transpose of the two matrices is:**

**<print the result – tab delimited>**

**====================================================================================**

**The result of matrix1 divided by matrix2 is:**

**<print the result – tab delimited>**

**====================================================================================**

**(\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*End of the Program \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*)**

Now change the content of the second matrix as follows ***(the changes are shown in the highlighted area)*** and run the program again as directed above. The only change is setting up the dimension to be equal to 4:

1. **20 30 40 50 60**

**70 80 90 100 110 120**

**130 140 150 160 170 180**

**190 200 0 22 230 240**

**250 260 270 28 290 300**

**0 320 330 34 350 360**

**Test your program with a data set that is provided in this lab.**

* **Place the questions with the answers to question 1 in one file, call it “Lab3 -Q1”**
* **the code for the program in another file call it “Lab3-Q2-Code”**
* **the snapshot of the output of the program in the third file., call it “Lab3-Q2-Output”**
* **Place all the files in a folder and call the folder based on your name and your partner’s name (ex:Jack-and-Nancy-Lab1)**
* **Zip the folder and only one person in the team is required to submit the work.**