CIS 635 Data Mining Name \_\_\_Dhynasah Çakir\_\_

Homework 2

Hand in sheet

Write in or copy and paste the answers for the following:

**Part A**

1. age\_numeric\_\_\_\_\_ gender\_\_nominal\_\_\_ weather\_\_nominal\_\_\_\_ size\_\_\_\_ordinal\_\_\_
2. colleges dimensionality \_\_\_\_high\_\_\_\_ sparsity \_\_\_\_low\_\_\_\_\_  
     
   students dimensionality \_\_\_high\_\_\_\_\_ sparsity \_\_\_high\_\_\_\_\_  
     
   books dimensionality \_\_\_high\_\_\_\_\_\_ sparsity \_\_\_high\_\_\_\_\_
3. quotes :\_ordered data\_\_\_\_ customer: \_\_transaction data\_\_ facebook: graph\_\_\_\_ store transaction data\_\_\_
4. a. \_\_\_yes\_\_\_\_ b. \_\_\_no\_\_\_\_ c. \_\_\_both\_\_\_\_\_
5. leave \_blank if ordinal or nominal, or estimate value from its nearest neighbors by taking the average from those neighbors if its numeric and there is a strong regression.
6. Delete attribute
7. leave blank
8. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part B**

1. 8 9 -1 10 2. 1.6, 2, -1.3, 2.3 3. Not possible. Dimensions should match

4. \_\_2\_\_ \_\_4\_\_ \_\_6\_\_ 5. \_\_1\_\_ \_\_6\_\_ \_\_12\_\_ 6. \_49\_\_\_ \_\_45\_\_ \_\_42\_\_  
  
 \_\_12\_\_ \_\_\_10\_ \_\_8\_\_ \_\_2\_\_ \_\_5\_\_ \_\_11\_\_ \_\_84\_\_ \_\_81\_\_ \_\_78\_\_   
  
 \_\_24\_\_ \_\_22\_\_ \_\_20\_\_ \_\_3\_\_ \_\_4\_\_ \_\_10\_\_ \_198\_\_\_ \_\_189\_\_ \_\_180\_\_

**Part C**

Note that you do not to give answers for each number in the instructions – write in the **results** or the **command** as directed.

1. results for #3 4, 0, 0, 10, 5
2. results for #4 \_\_\_\_\_ [,1] [1,] 19
3. results for # 6 168
4. command for #8 \_\_\_\_\_x[1:10,2:3]\_\_\_\_
5. command for #10 \_\_\_\_\_\_\_\_\_\_\_\_\_A<- colMeans(x[,2:5],x$passed == "n")

print(A)

1. command for #12 \_\_\_\_\_\_\_\_\_\_\_ind1= sample(nrow(x)) %%3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. command for #13 \_\_\_\_\_\_\_\_\_\_\_\_mean(x[ind==1,4])\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_