CSE 310 L Data Warehousing and Mining

Date: 29-08-2023

1. Write a Python program to **fill in the missing values**

Create a file named “Emplyee.txt” having “EmpName”, “EmpId”, “EmpDesig”, and “EmpSalary” as the attributes.

Check for empty cells and fill the missing values with “NA” for “EmpName”, “EmpId” and “EmpDesig”, with the average salary for “EmpSalary”.

**Input:**

EmpName EmpId EmpDesig EmpSalary

T. Ramesh 20001 Asst. Prof. 150000

M. Soumya 20002 Asst. Prof. 120000

R. Shyam 20003 140000

S. Raghu Asst. Prof.

**Output:**

EmpName EmpId EmpDesig EmpSalaray

T. Ramesh 20001 Asst. Prof. 150000

M. Soumya 20002 Asst. Prof. 120000

R. Shyam 20003 NA 140000

S. Raghu NA Asst. Prof. 136666

1. Write Python program to smoothen the data using Binning (Equal frequency Binning and Smoothing by bin mean)

**Input:**

Enter the number of products: 12

Enter the price of the products:

15 13 16 16 25 20 20 21 22 22 19 25

Enter the number of bins: 4

**Output:**

Bin 1: 14.67, 14.67, 14.67

Bin 2: 18.33, 18.33, 18.33

Bin 3: 21, 21, 21

Bin 4: 24, 24, 24