EX NO: **DATA PRE-PROCESSING AND DATA CUBE : DATA PREPROCESSING METHODS ON STUDENT AND LABOR**
DATE: **DATASETS IMPLEMENT DATA CUBE FOR DATA WAREHOUSE ON 3-DIMENSIONAL DATA**

AIM:

BACKGROUND THEORY:

Data Preprocessing:

Data preprocessing involves cleaning, transforming, and organizing data to prepare it for analysis. We will use Orange's widgets to perform data preprocessing on student and labor datasets.

Implementing a Data Cube for Data Warehousing:

A data cube allows you to visualize and analyze data across multiple dimensions. In Orange, you can simulate a data cube using the "Pivot Table" widget for 3-dimensional data.

PROCEDURE:

1. Load Data:

o Drag the "File" widget to the canvas.

2. Handle Missing Values:

o Drag the "Edit Domain" widget to the canvas.

o Connect the "File" widget to the "Edit Domain" widget.

o Configure the "Edit Domain" widget to handle missing values (e.g., impute with mean/median).

3. Normalize Data:

o Drag the "Normalize" widget to the canvas.

o Connect the "Edit Domain" widget to the "Normalize" widget.

o Configure the "Normalize" widget to standardize the data.

4. Discretize Data (if necessary):

o Connect the "Normalize" widget to the "Discretize" widget.

o Configure the "Discretize" widget to bin continuous variables.

5. Feature Selection:

o Drag the "Select Columns" widget to the canvas.

o Connect the final preprocessing widget to the "Select Columns" widget & select relevant feature.
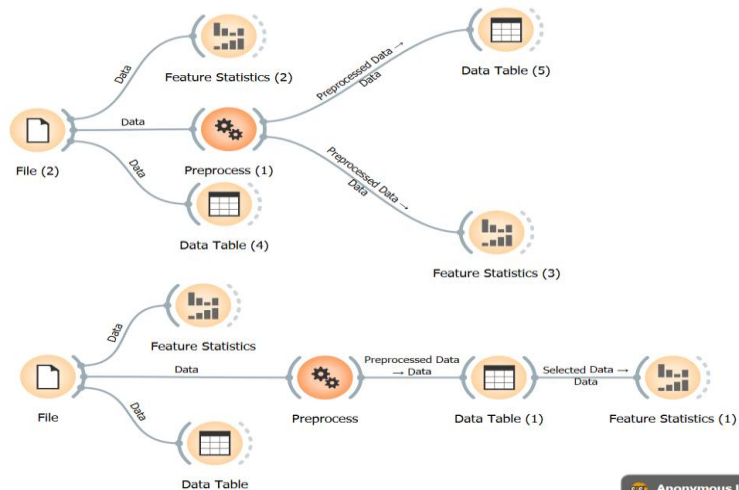
OUTPUT:

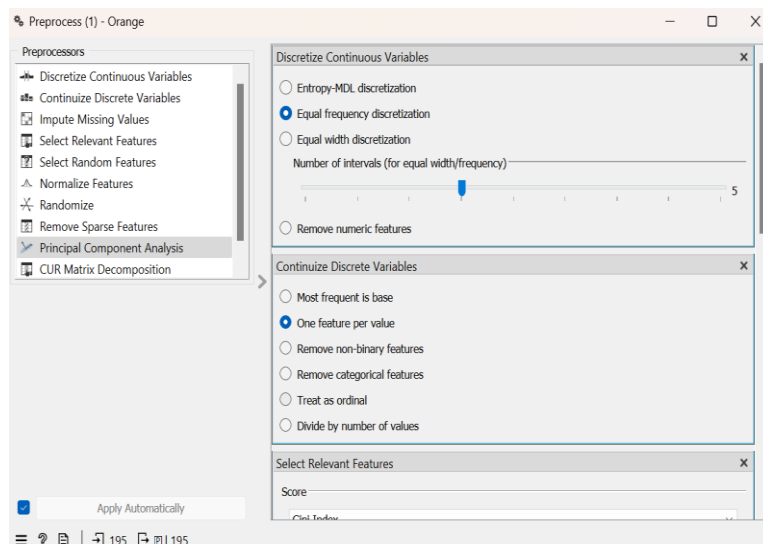

FIG: 1.1 – DATA PREPROCESSING AND FEATURE STATISTICS

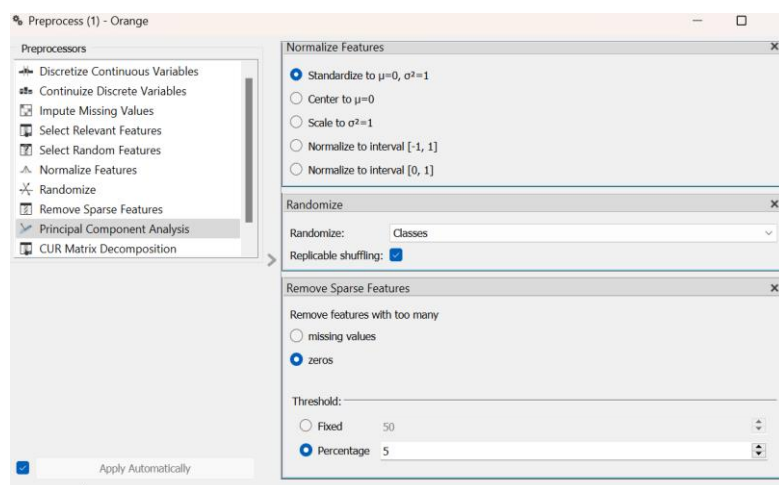FIG: 1.2 : PREPROCESSING DATA BY DISCRETIZING , CONTINOUS DISCRETE, RELEVANT FEATURES.

FIG: 1.2.1 : PREPROCESSING DATA BY NORMALIZINF, RANDOMIZING , REMOVE SAPRSE FEATURES.

RESULT: