

FACULTY OF ENGINEERING AND TECHNOLOGY BACHELOR OF TECHNOLOGY

Data Mining and Business Intelligence

(DMBI) (203105454)

VI SEMESTER

Computer Science & Engineering Department





**CERTIFICATE**

*This is to certify that*

*Mr.*  **LUCKY PATHAN** *with Enrollment No.* **210303105790** has *successfully completed his laboratory experiments in the subject (with Code)* **Data Mining and Business Intelligence (203105454)** *from the department of* **Computer Science and Engineering** *during the academic year* ***2022-2023.***



**Date of Submission …..…………… Staff In charge …..……………**

**Head of Department …..……………**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **S**  **r. N**  **o** | **Experiment Title** | **Page No** | | **Date of Start** | **Date of Completing** | **Sign** | **Marks**  **(Out of 10)** |
| **From** | **To** |
| 1 | Design and Create cube by identifying measures and dimensions for Star Schema, Snowflake schema and fact Constellation Schema. | 1 | 3 |  |  |  |  |
| 2 | Make an OLAP cube and perform Roll Up and Drill Down operations on it. Show the Apex and Base cuboid for the same. Draw Star-net query model for the cube. | 4 | 5 |  |  |  |  |
| 3 | Create calculated member using arithmetic operators and member property of dimension member. | 6 |  |  |  |  |  |
| 4 | Design and Create cube by identifying measures and dimensions for Design storage using storage mode MOLAP, ROLAP and HOLAP. |  |  |  |  |  |  |
| 5 | Perform Pre-processing on a dataset. Apply various Filters and discuss the effect of each filter applied.   1. Handle Missing Values 2. Handle Infrequent Nominal Values 3. Derive an attribute from the existing attribute 4. Sampling 5. Discretization 6. Use Weka Too 7. Use XL Miner Tool |  |  |  |  |  |  |
| 6 | Perform different binning techniques to smooth out the noise in the dataset. Make sure that the user should have the choice to apply all the possible techniques. Show the output of different bins. Use histogram to partition the dataset into groups. |  |  |  |  |  |  |
| 7 | Perform regression on the data set using R programming. |  |  |  |  |  |  |
| 8 | Perform Association rule mining using WEKA tool. |  |  |  |  |  |  |
| 9 | Perform classification with WEKA tool   1. using Decision Tree Classifier 2. using Naïve Bayes Classifier 3. using Multilayer Perceptron. |  |  |  |  |  |  |
| 10 | Perform Clustering using WEKA too. |  |  |  |  |  |  |

# Practical 1

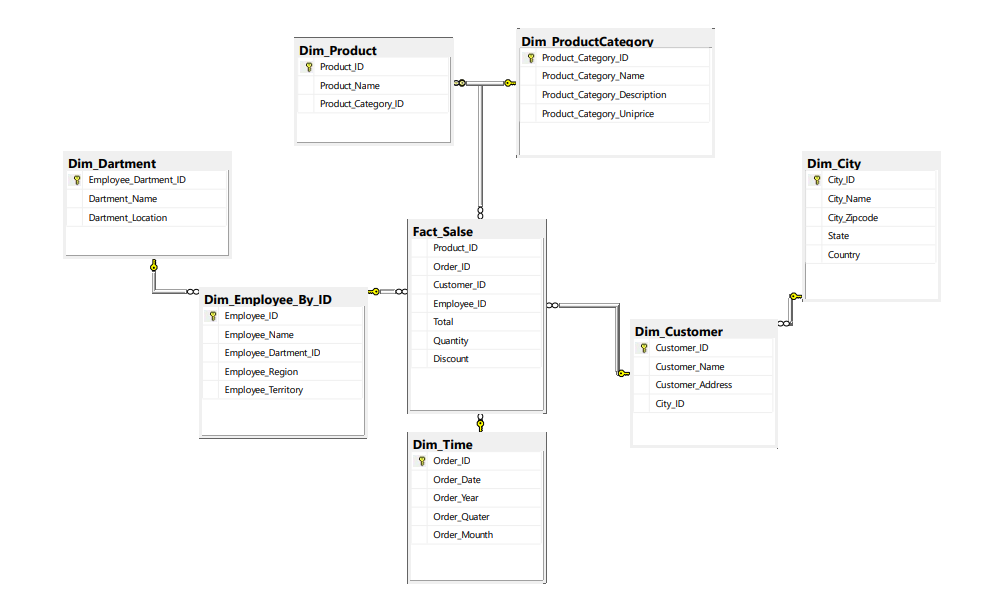
**Aim** : Design and Create cube by identifying measures and dimensions for Star Schema, Snowflake schema and fact Constellation Schema.

**OUTPUT:-**

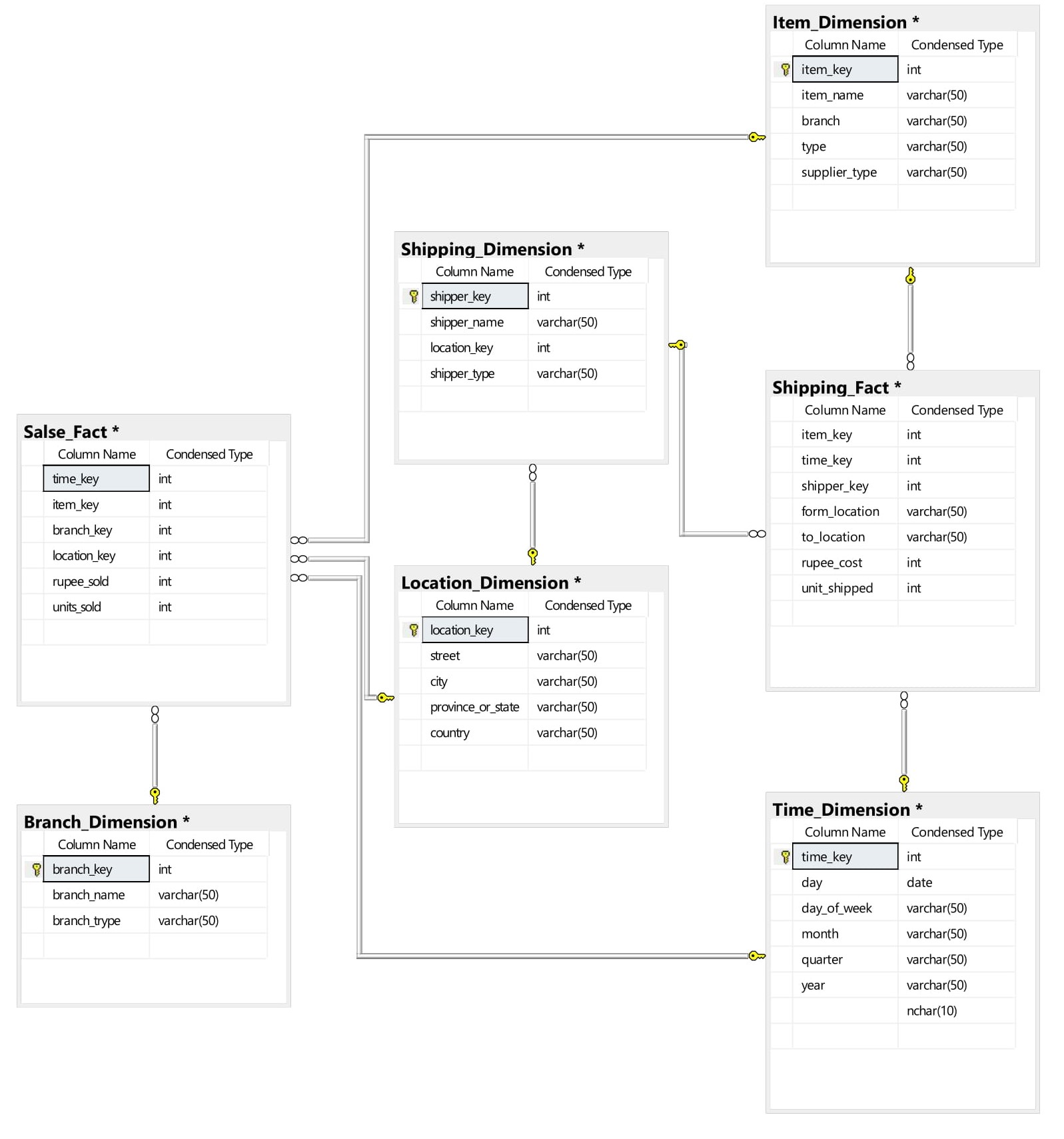
Star Schema -



SnowFlack Schema –



Constellation Schema -



# Practical 2

**Aim** : Make an OLAP cube and perform Roll Up and Drill Down operations on it. Show the Apex and Base cuboid for the same. Draw Star-net query model for the cube.

**CODE:-**

CREATE TABLE STUDENT(ID INT,NAME VARCHAR(50),CITY VARCHAR(30));

INSERT INTO STUDENT VALUES(1,'Lucky','Goghamba');

INSERT INTO STUDENT VALUES(2,'Swet','virpur');

INSERT INTO STUDENT VALUES(3,'Dhaval','Rampura');

INSERT INTO STUDENT VALUES(4,'Vasu','Rajcort');

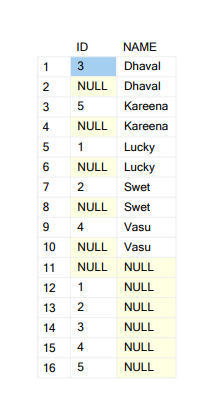
INSERT INTO STUDENT VALUES(5,'Kareena','Mumbai');

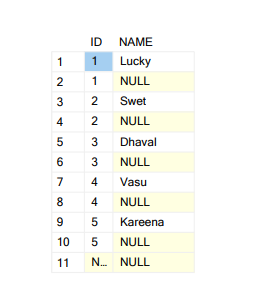
select ID, NAME from STUDENT group by cube(ID,NAME);

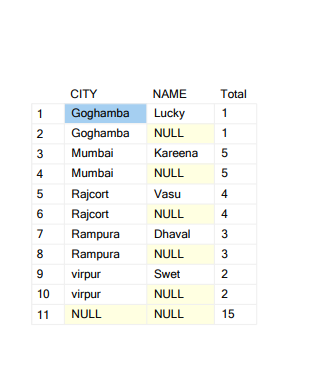
select ID, NAME from STUDENT group by rollup(ID,NAME);

select STUDENT.CITY,STUDENT.NAME ,SUM(STUDENT.ID) Total from STUDENT GROUP BY GROUPING SETS( ([CITY], [NAME]), ([CITY]), ());

**OUTPUT:-**

CUBE:-

ROLLUP:-

DRILL DOWN :-

# Practical 3

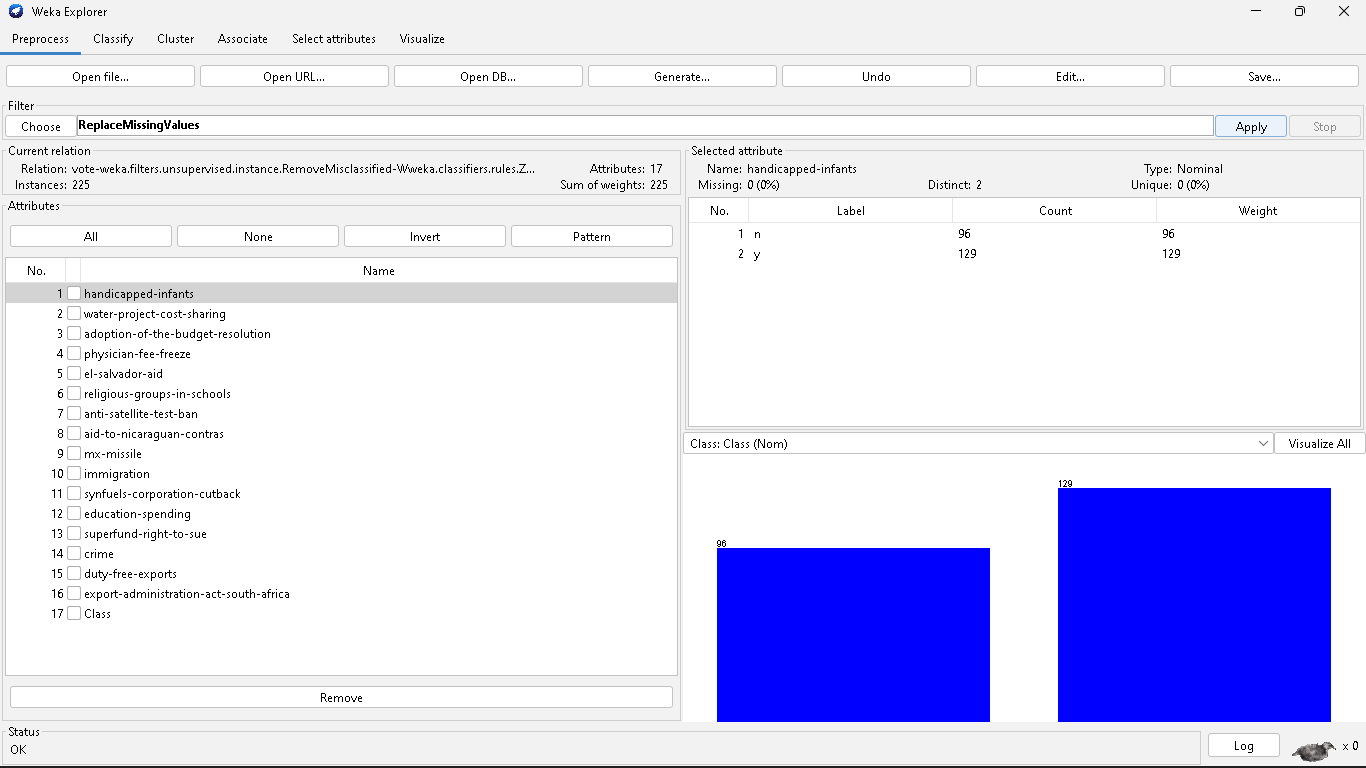
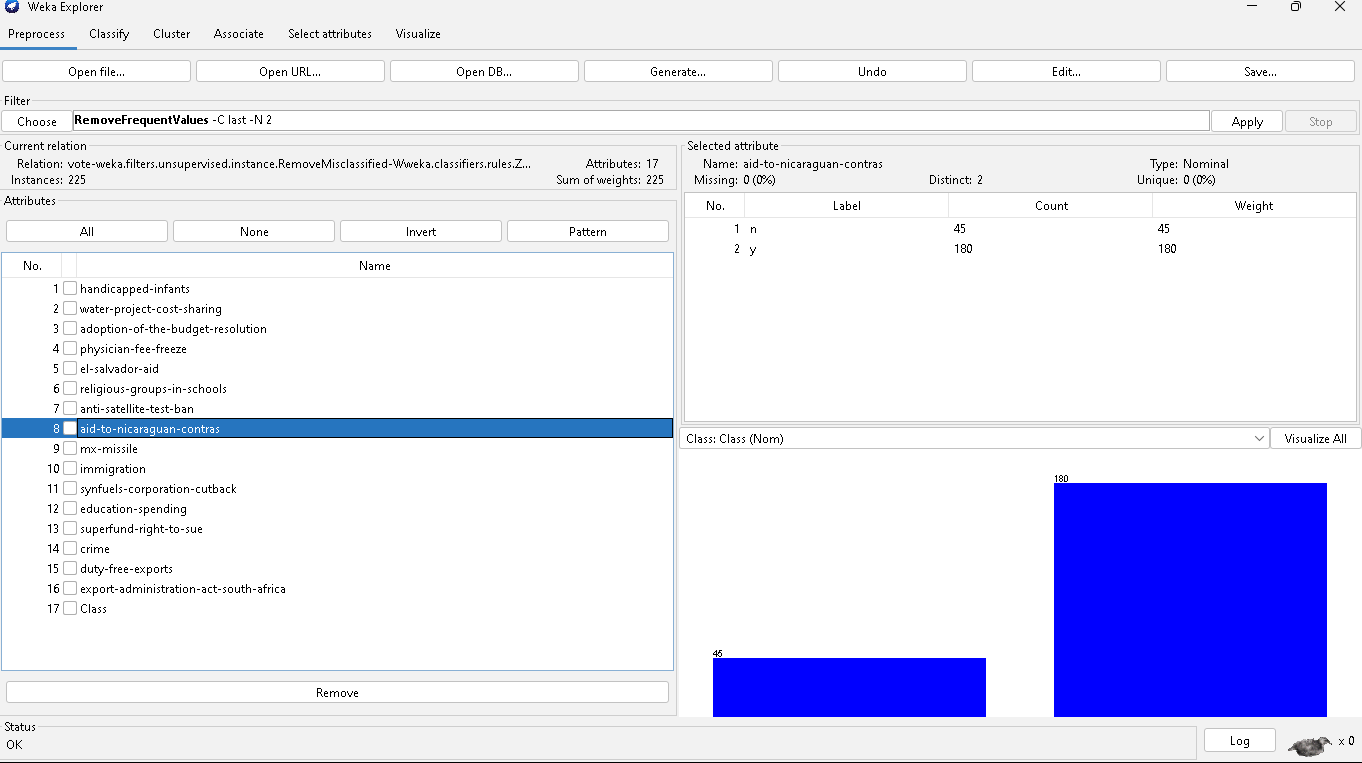
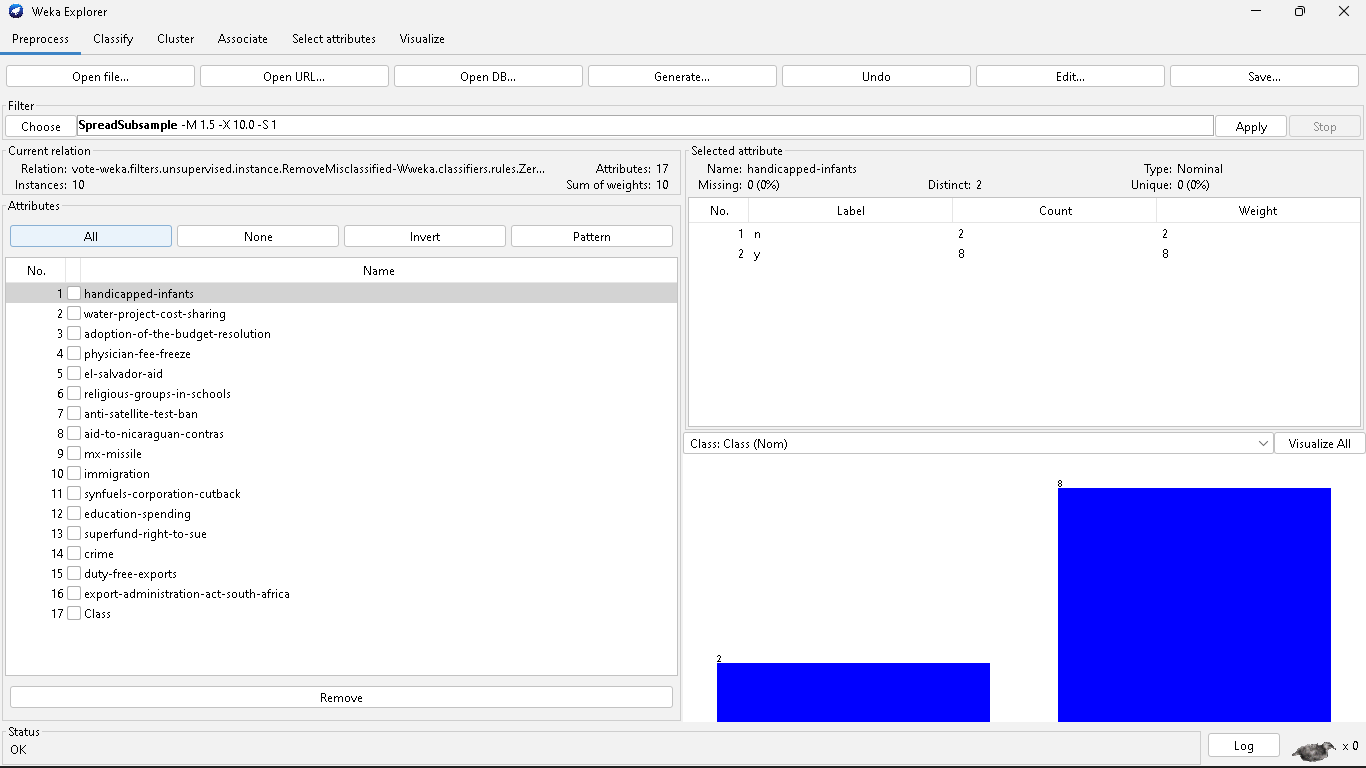
**Aim** : Create calculated member using arithmetic operators and member property of dimension member.

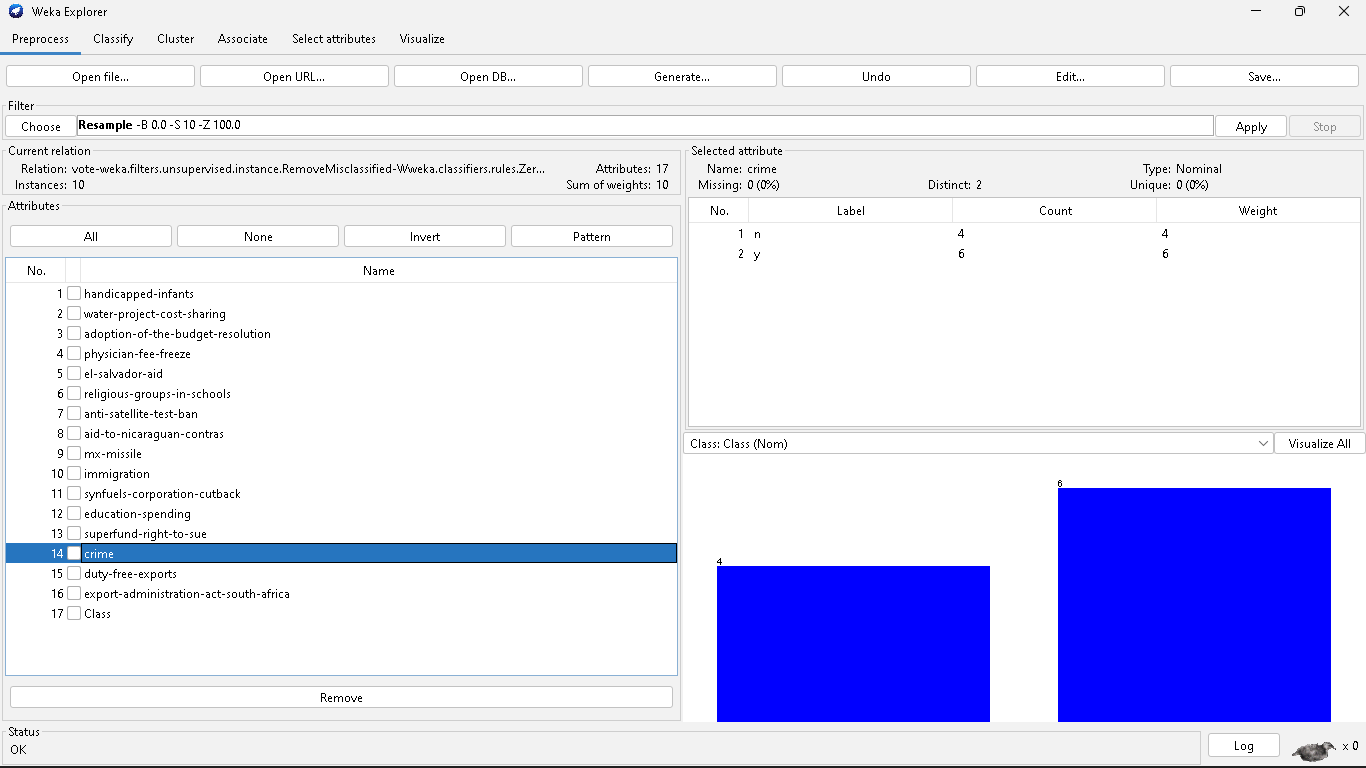
# Practical 4

**Aim** : Design and Create cube by identifying measures and dimensions for Design storage using storage mode MOLAP, ROLAP and HOLAP.

# Practical 5

**Aim** : Perform Pre-processing on a dataset. Apply various Filters and discuss the effect of each filter applied.

1. Handle Missing Values
2. Handle Infrequent Nominal Values
3. Derive an attribute from the existing attribute
4. Sampling



1. Discretization

