Lab-1

Aim: Recording types of data and various file formats. Identifying data sources. Handling traditionally to start with at small scale.

Rahil Acharya

CE004

20CEUOD004

Data Types

1. Student Information : Structured
2. Book Information   : Structured
3. Feedback Form   : Structured
4. Issuing and returning book information : Structured
5. E-Learning content : unstructured
6. Web Pages   : Semi Structured

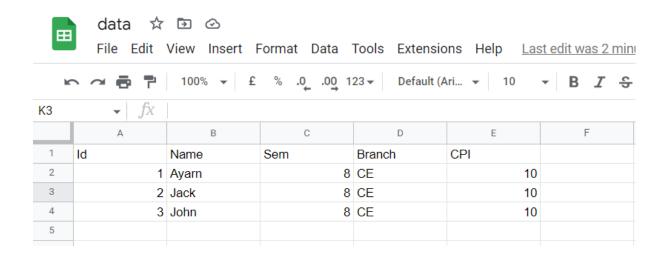**Database** vs. **Data Warehouse**

- **Database**

- Used in OLTP

- Optimised for write operation

- Current Data

- Application Oriented
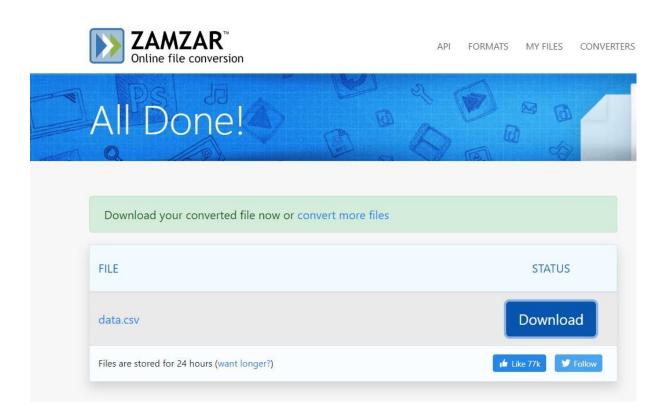
- Normalised Data

- **Data Warehouse**

- Used in OLAP

- Optimised for read operation

- Historical Data

- Subject Oriented

- Denormalized Data
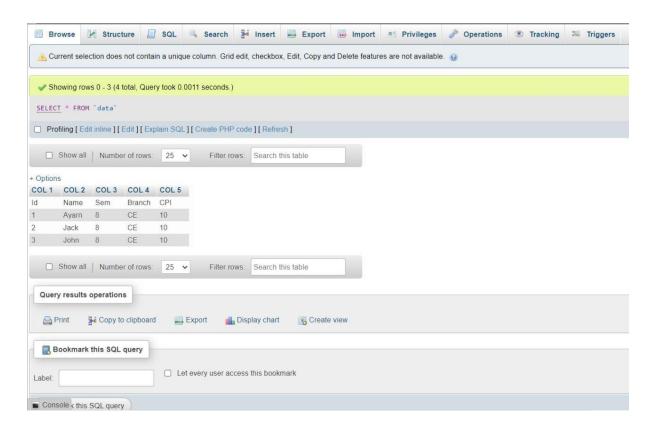
- **Given the spreadsheet file convert it into a csv**

https://www.zamzar.com/convert/xls-to-csv/


```
Id,Name,Sem,Branch,CPI
1,Ayarn,8,CE,10
2,Jack,8,CE,10
3,John,8,CE,10
```

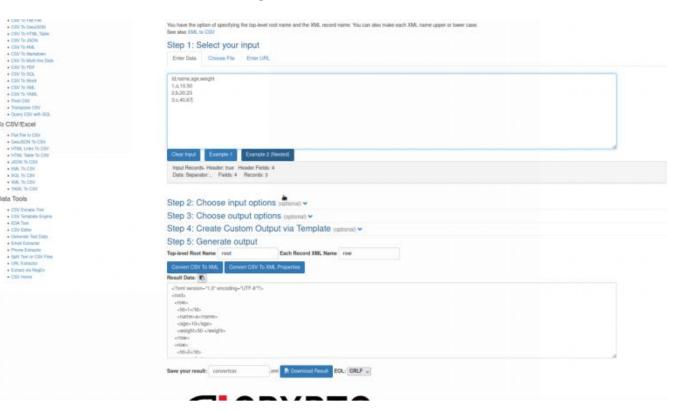- Import a csv into MySQL database table

- Write a computer program to read records from database and generate data file.
  - XML
  - JSON

Convert CSV to Database file into sqlite.

```
[user1@hadoop-clone Personal]$ sqlite3
SQLite version 3.7.17 2013-05-20 00:56:22
Enter ".help" for instructions
Enter SQL statements terminated with a ";"
sqlite> .mode csv
sqlite> .import Student.csv Student
Error: no such table: Student
sqlite> CREATE TABLE Student(id integer PRIMARY KEY, name text, age integer, weight integer);
sqlite> .import Student.csv Student
Error: datatype mismatch
sqlite> INSERT INTO Student (7, "abc", 20, 40);
Error: near "7": syntax error
sqlite> INSERT INTO Student VALUES (7, "abc", 20, 40);
sqlite> SELECT * FROM Student;
7,abc,20,40
sqlite> .import Student.csv Student
sqlite> SELECT * FROM Student;
1,a,10,50
2,b,20,23
3,c,40,67
7,abc,20,40
sqlite> .tables
Student
sqlite> .output Student.sql
sqlite> .dump
sqlite> .exit
```

CSV To XML Using 'convertcsv.com'

CSV To JSON Using 'csvjson.com'