**Technologies used:**

|  |  |
| --- | --- |
| Database | SQL Server |
| ETL Tool | SSIS (SQL Server Integration Services) |
| Analysis tool | SSAS Cube (SQL Server Analysis Services) – Multidimensional cube. |
| Reporting Tool | Excel |

**Steps performed for the case study:**

1. Downloaded the JSON file from location <https://s3-eu-west-1.amazonaws.com/csparkdata/ol_cdump.json>
2. Imported JSON file “ol\_cdump.json” into below SQL Server

Server: Localhost

Database: demo

Schema: casestudy

Table: o1\_cdump.

1. Performed the cleaning and loading of data into below table using SSIS package.

SSIS Package: Data Loading.dtsx

Server: Localhost

Database: demo

Schema: casestudy

Table: mv\_o1\_cdump

1. Split the data and loaded into Layer 1 tables using SSIS package.

SSIS Package: Data Loading.dtsx

Server: Localhost

Database: demo

Schema: casestudy

Table: author,work,edition

1. Data modeling and creation of Facts and Dims using step 4 data.

Server: Localhost

Database: datawarehouse

Schema: cs

Procedures: usp\_td\_author, usp\_td\_edition, usp\_td\_work, usp\_tf\_FactlessFact

Dimension tables: td\_author, td\_edition, td\_genres, td\_languages, td\_publishers, td\_subjects, td\_work

Fact Table: tf\_FactlessFact

1. Developed the SSAS Multidimensional cube for slicing and dicing using above dimension and fact tables.
2. Profiled the data and created different pivots with measures. Pivots available in the separate excel document.