

Data Warehouse Project

Education DataSet

Link: <https://www.kaggle.com/datasets/rkiattisak/student-performance-in-mathematics>

Brief description:

The Dataset contains information on the performance of high school students in mathematics, including their grades and demographic information. The data was collected from three high schools in the United States.

Columns:

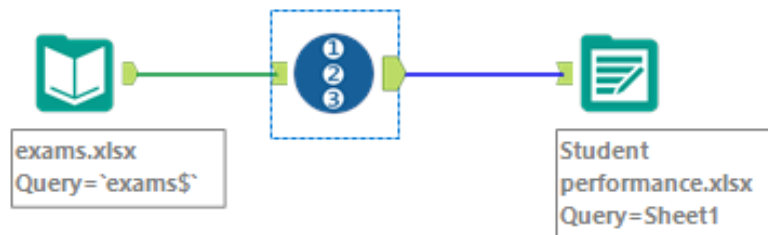
- **Gender:** The gender of the student (male/female)
- **Race/ethnicity:** The student's racial or ethnic background (Asian, African-American, Hispanic, etc.)
- **Parental level of education:** The highest level of education attained by the student's parent(s) or guardian(s)
- **Lunch:** Whether the student receives free or reduced-price lunch (yes/no)
- **Test preparation course:** Whether the student completed a test preparation course (yes/no)
- **Math score:** The student's score on a standardized mathematics test
- **Reading score:** The student's score on a standardized reading test
- **Writing score:** The student's score on a standardized writing test

Components used in Alteryx:

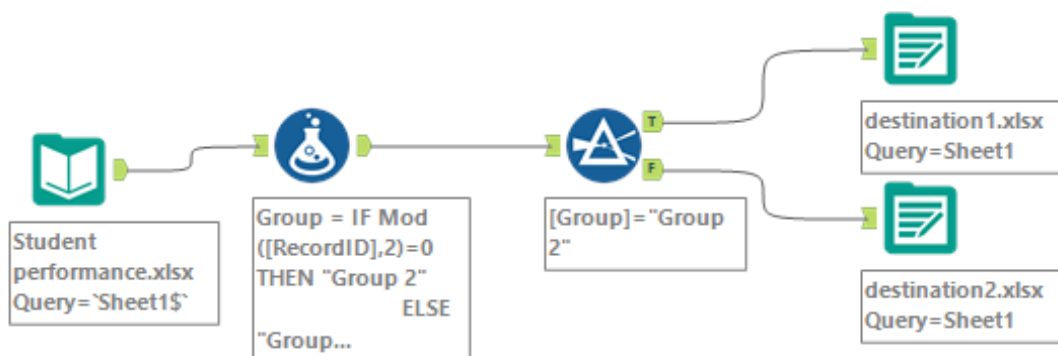
- 1- Input Data
- 2- Output Data
- 3- RecordID
- 4- Formula
- 5- Filter
- 6- Select
- 7- Summarize
- 8- Join
- 9- Transpose

Screenshots from Alteryx

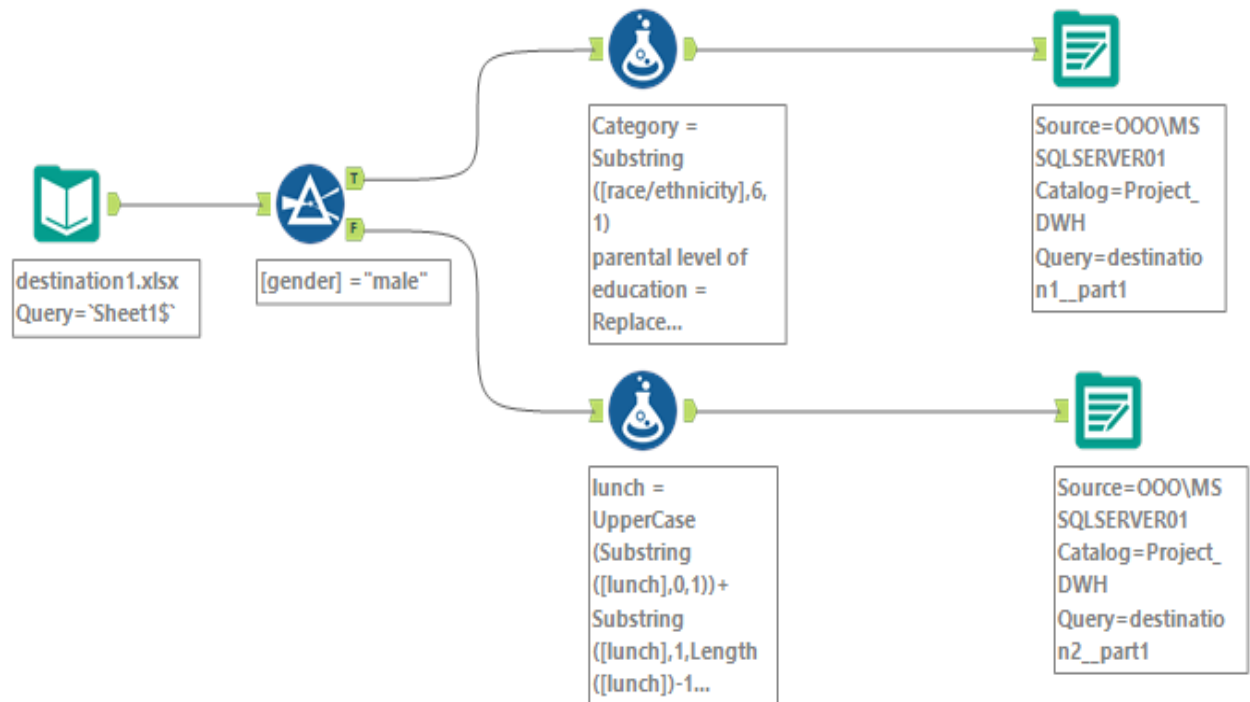
1- Add RecordID to data



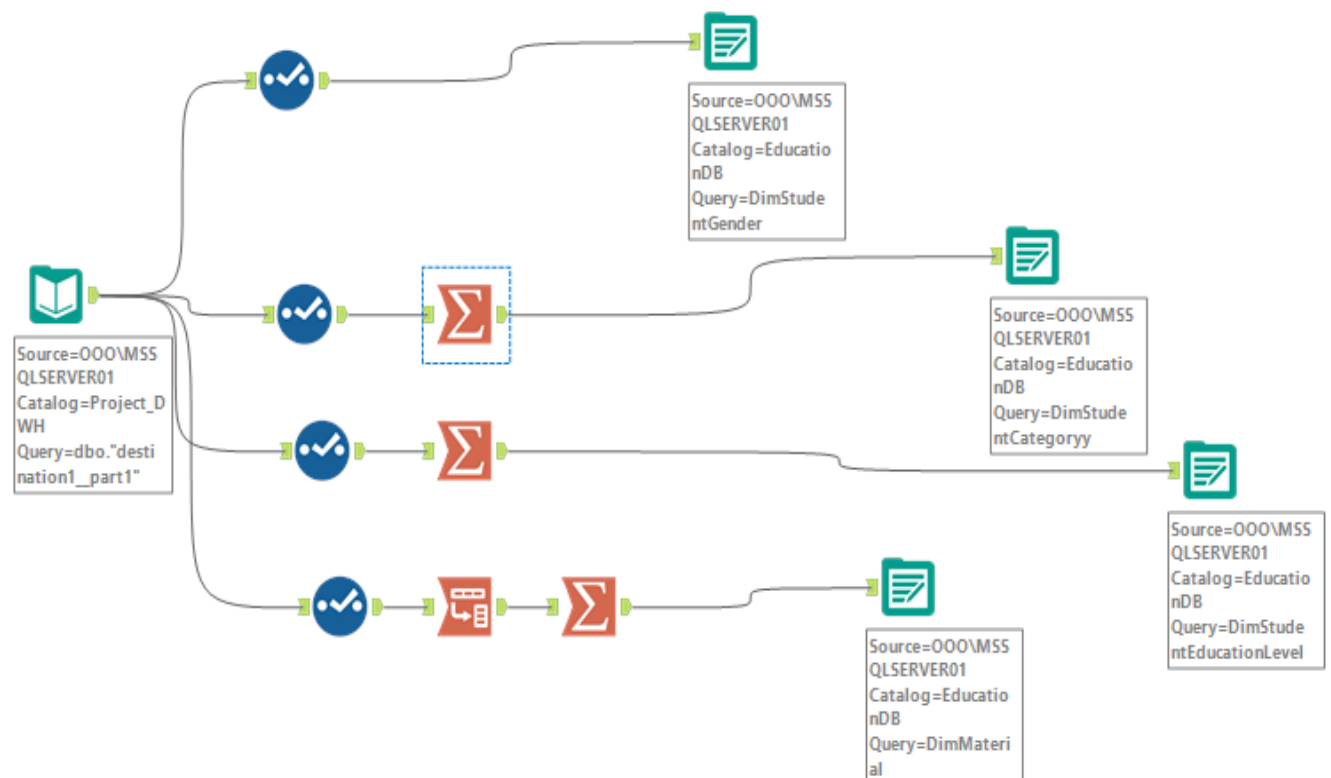
2- Split Dataset based on even or odd



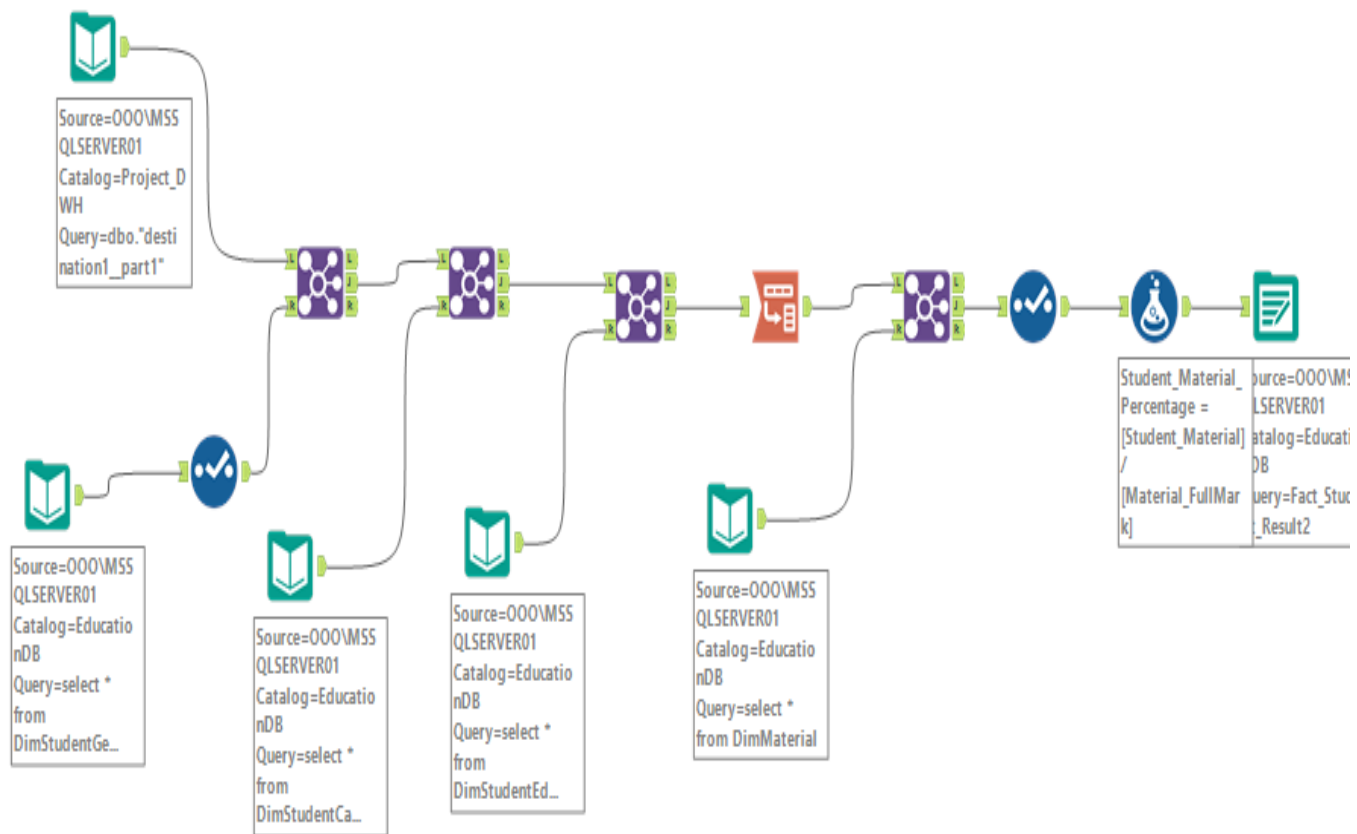
3- Transformations on data



4- Star Schema tables based on four dimensions



5- Fill fact_table with 4 dimensions



Companents used in SSIS

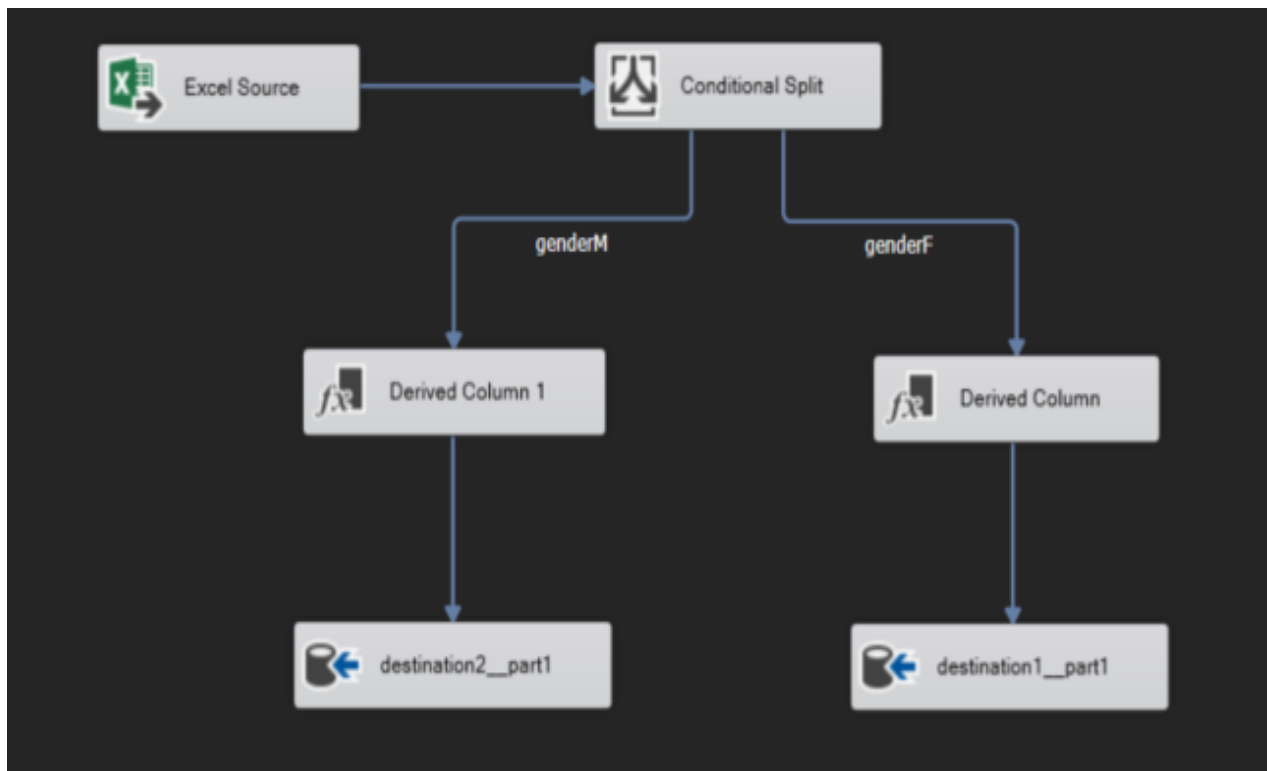
1-Excel Source

2-OLE DB Destination

3-Conditional Split

4-Derived Column

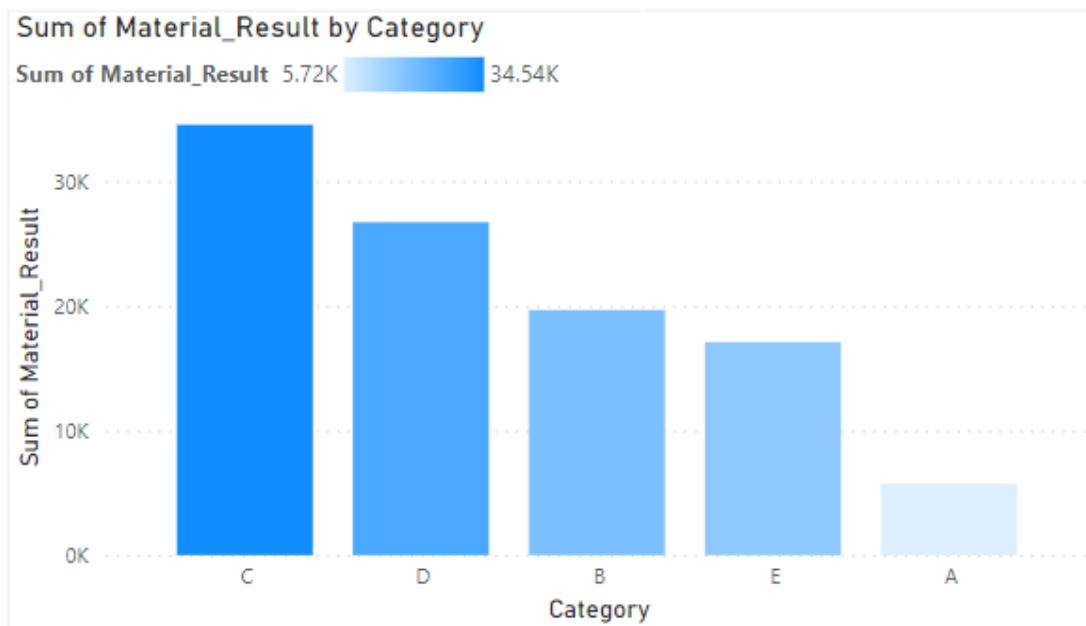
Screenshot from SSIS



Decisions based on charts:

1-bar chart

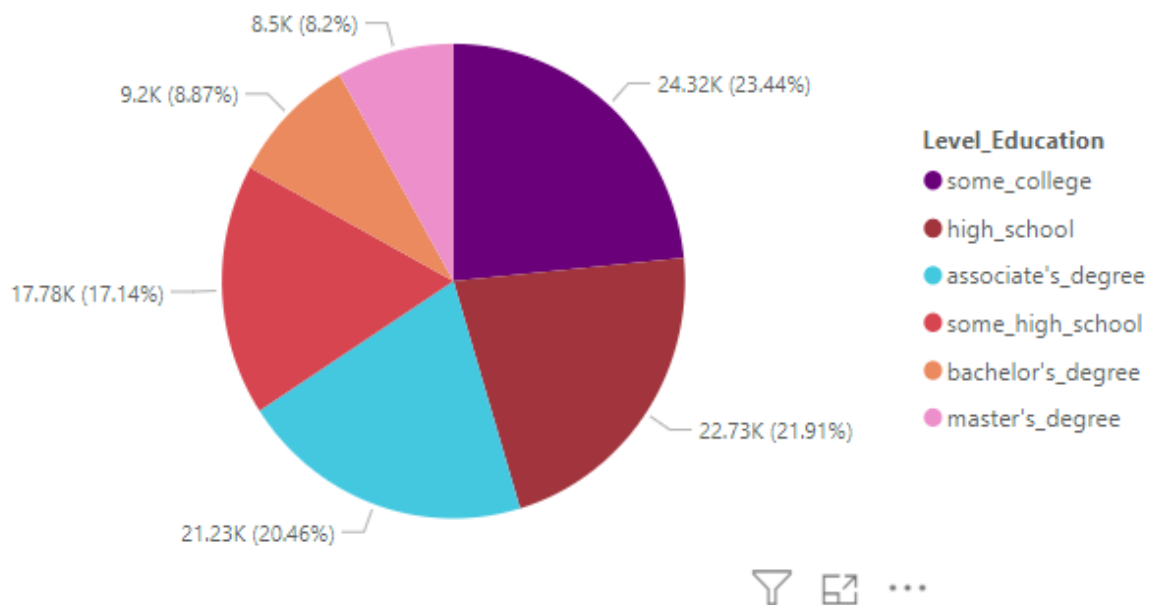
Decision: based on this chart the maximum values of materials (Exams) are in category C and the minimum values of materials(Exams) are in Category A



2-pie chart

Decision: in this chart the data is slices based on Level_Education and the Highest values in material exams are in some_college level and the lowest values in material exams are in master's_degree level.

Sum of Material_Result by Level_Education



3-bar chart

Decision: here the comparison is very convergent so here the compare between Female and male in average grades in material exam is considered the same

Average of Material_Result by Student_Gender

