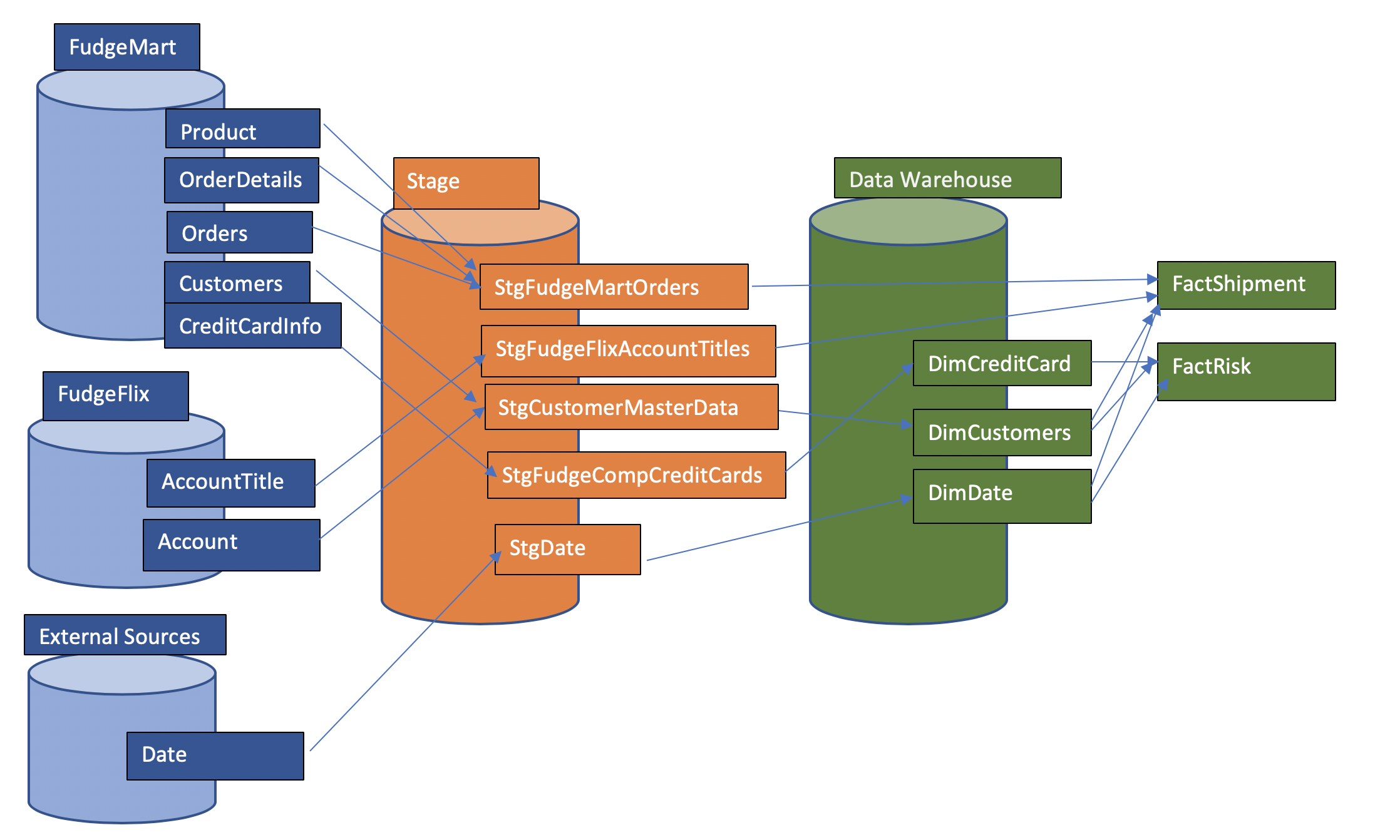
**IST722 Group Project Milestone#3 Document - Deliverable 5**

Team 1 - Xuehan Chen, Yue Wang, Yimin Xiao, Chiau Yin Yang

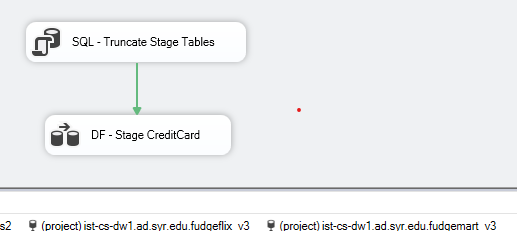
**Initial ETL done in SSIS**

1. Source to target map

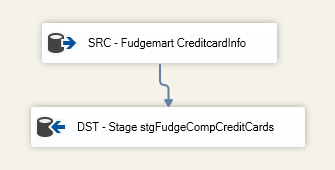


1. SSIS Package to stage your data

**Stage stgCreditCard -** Control flow

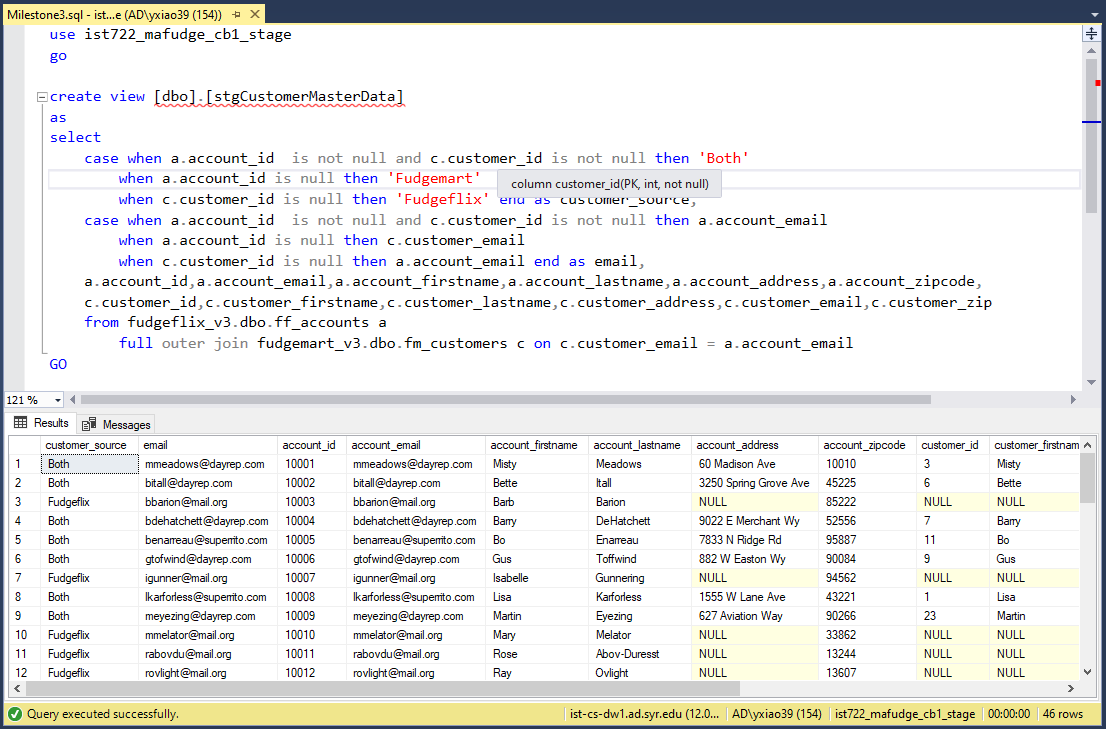


**Stage stgCreditCard -** Data flow

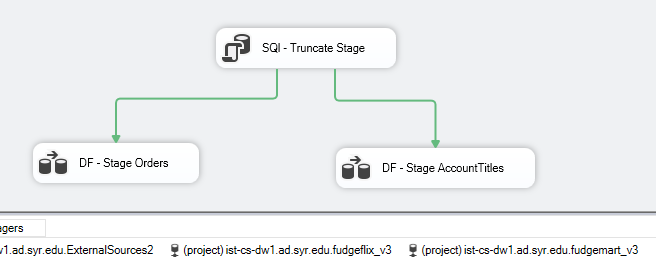


**Stage stgCustomer**

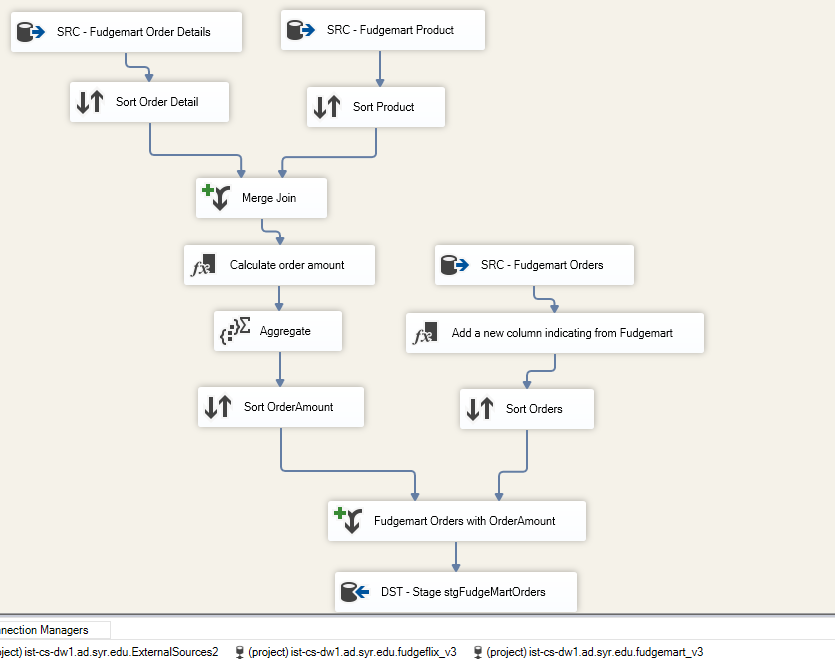
This is done by creating a view in SQL server. The view combines customers from both sources, with a source column indicating whether the data is from ‘Fudgemart’, ‘Fudgeflix’, or ‘Both’.



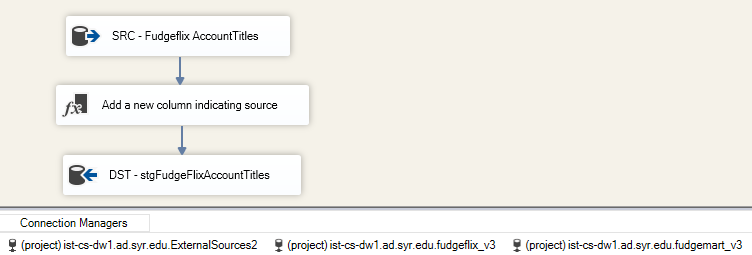
**Stage Shipment (Fudgemart Orders and Fudgeflix AccountTitles) -** Control flow



**Stage Shipment (Fudgemart Orders) -** Data flow

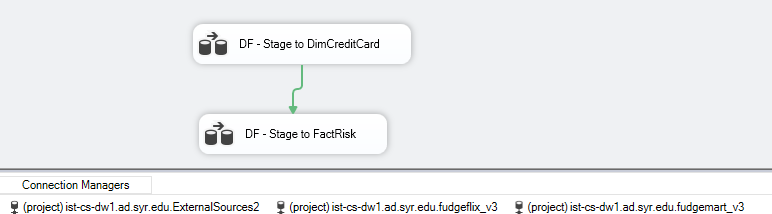


**Stage Shipment (Fudgeflix AccountTitles) -** Data flow

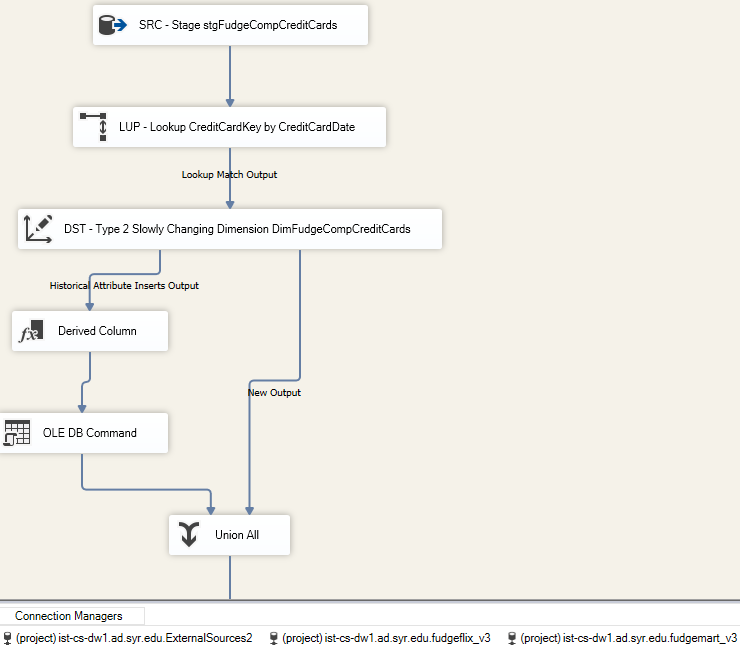


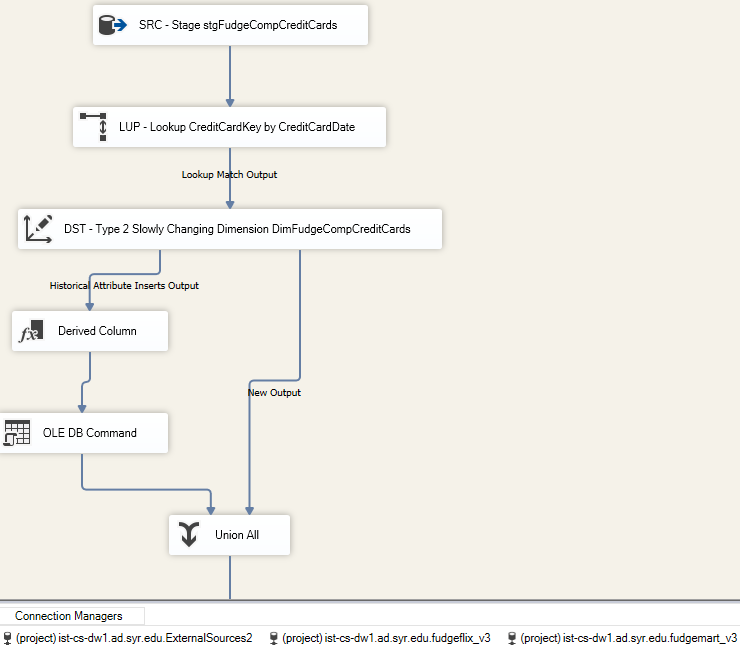
1. SSIS Package to load to DW from stage

**DW\_Risk -** Control flow

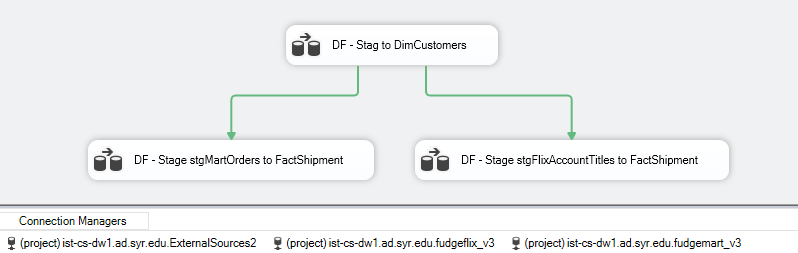


**DW\_Risk: stage to DimCredictCard** - Data flow

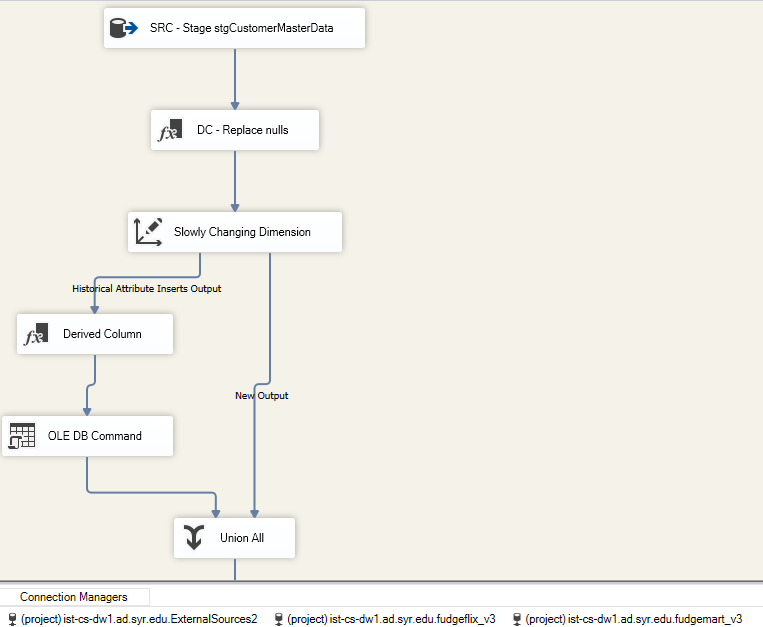


**DW\_Risk: FactRisk** - Data flow 

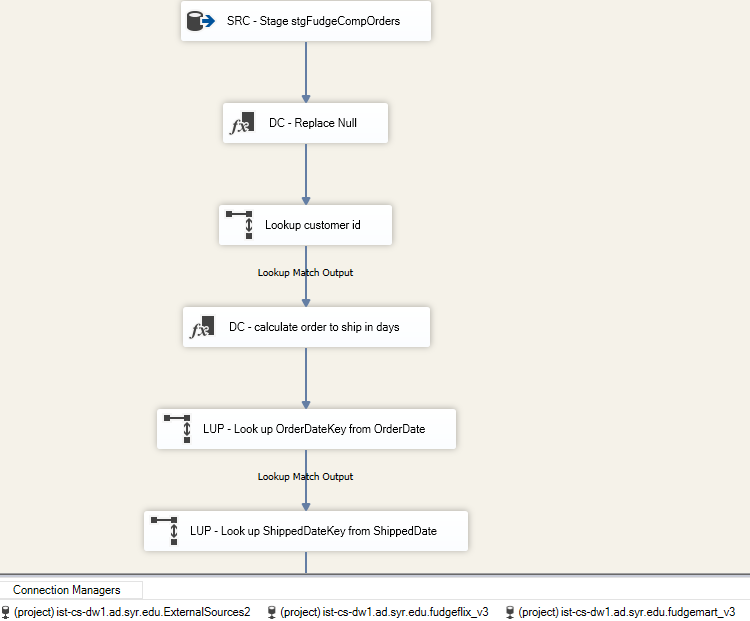
**DW\_Shipment -** Control flow



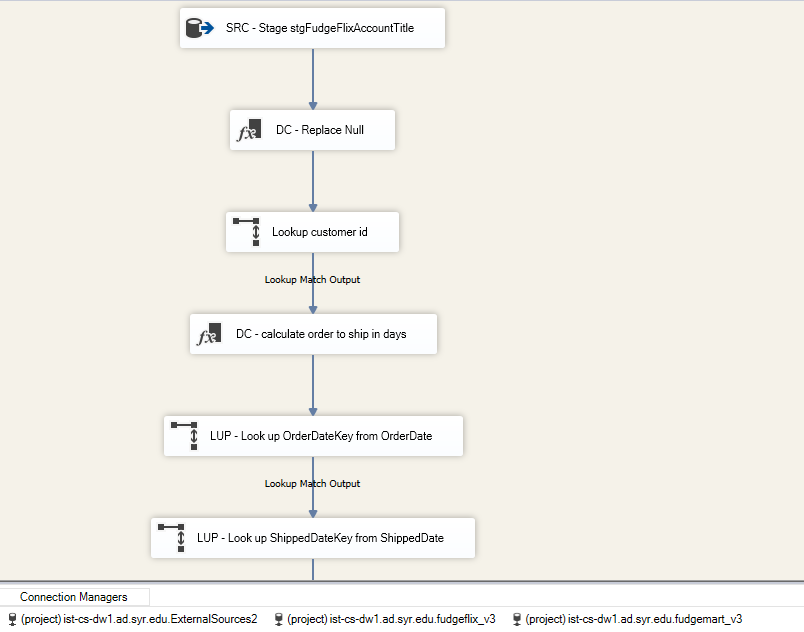
**DW\_Shipment: stgCustomers to DimCustomers** - Data flow



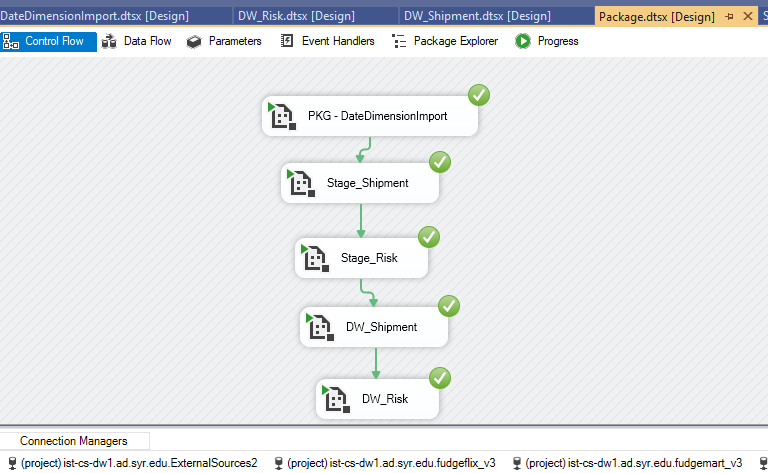
**DW\_Shipment: stgFudgeCompOrders to FactShipment** - Data flow



**DW\_Shipment: stage stgFlixAccountTitles to FactShipment** - Data Flow



**ETL process run successfully**

****

1. ETL Documentation: Source to target map, screenshots of data flows and explanation of ETL patterns used. Explain any data quality or survivorship rules you may have used.
   1. All the dimensions in the data warehouse are historical data. (DimCreditCard, DimCustomers, DimDate)
   2. All the facts in the data warehouse are slowly changing dimension. (FactShipment, FactRisk)
   3. All non-shipped product has a shipment data of the earliest data in DimDate.

