بسم الله الرحمن الرحيم

* ML → produces knowledge
* BI → produces visualization.

* Data should be
1- Integrated 2-Historical 3-Time-stamped 4-Subjecte oriented
       5-cleansed

* ETL: extract - transform - load

* Unioning files into 1 source → to get single version of truth

* Connection Managers → هي connections بتاخد قائمة لل
project كله (و ده الجنب) تطبيق على كل package (و ده ؟)

* SSIS packages → To transform data.
* package parse → visual شايل في ال package
* Miscellaneous → لو عايز تحط related files للمشروع

* Control Flow → ال Flow بتاع ال package ده كله
* Data Flow → شايل ال باب ال data بس
* parameters → لو غيرت في ال package
* Event Handlers → error
* package Explorer → ده بيكون ال hierarchy ال all Folders في
                         كل package

(Task) عبارة عن file على القرص أو في DB

مبدئي بعمل source ? ، transform ? ، destination

& * ال Fact Table بتجمع بين جداول { X } ال dimension table و ال Fk
   measures

* الـ measure يعني حاجة بتتقاس بيس performance بتاع
شئ معين زي مثلا quantity & sales

* Difference between Table or View & Table of view - fast load *

في الـ Table or view العادية .. بيعمل insert للـ rows واحد
واحد .. في الـ load بيعمل كريم insert بجميع الـ insert واحد واحد
بس مشكلة في باب لو row في النص حصل فيه مشكلة.
معينش حاجة بتشتغل خالص.
← مميزه أسرع : Fast load
← مميزه أكثر : Table or View العادية

(Task Solution)

| Distributor |

↓

| OLE DB Destination |

لو الـ big int

* في الـ Flat File بيرجع أخذ أول الـ String لـ unicode string
* و الـ integer بتبقى integer Four-byte unsigned
لو int عادي → Four-byte signed integer
* لو عندي Row في مسكلة في أي حتى row يكون "
String يبقى كل column الى integer عندي ٣ حلول:
← أكبر programming
← أكبر column derived
← استبعاد دلوقتي

بعمل كده من خلال الـ configure error output button:
بحط مكان الحاجه اللي فيها مشكله NULL ① Ignore Failure
لو محتاج الـ Rows دي فيها غلط و اقدر اعالجها تانيه ② Redirect Row
يوقف كل حاجه تانيه ③ Fail component

Select sum (Sales) From [Company] [dbo] [salesFact
Session2]
where Substring ([Date] , 4, 2) = '06'

* Change (10/07/19) to (2019-07-10)

"20" +
SUBSTRING ([Date] , 7, 2) + " - " +
SUBSTRING ([Date] , 4, 2) + " - " +
SUBSTRING ([Date] , 1, 2)

[Distributor2] "Source"
↓
[Derived Column]
↓
[SalesFact] "Destination"
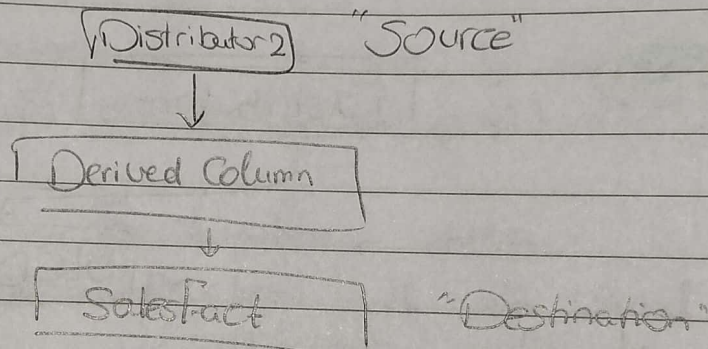
* To delete old data in a table → Truncate table

* في ال Mapping اقفل دايا تكت اسم ال Column الجديد
عشان تكلك ال data الجديد

نفس الكود بس العزء الي تاني سطر من تكتبه
الي هو بتاع الشهر و هدخط باله "06" +

* Detach → هنا هنخزن ال buffer معناها وكل ال loading
زي ما انت عايز

* لو عندي بقى في Flat File نفس الوقت 2Formats لل <date>
مرة بتكتب بطريقة 8062019 ومرة 08062019 ..
طالما في حالتين يبقى مسمح بكتر Component بين الاتنين
Conditional Split .

↳ المطلوب باب هي الاتنين يتحولوا لل Format :
2019 - 06 - 08

(If length = 7) 8062019 ,   (If length = 8) 08062019

SUBSTRING ([Date], 4,4) + "-" +   SUBSTRING([Date], 5,4) + "-" +
"   " , 2, 2) + "-0" +   "   " , 3, 2) + "-" +
"   " , 1, 1)   "   " , 1, 2)

Flow

Distributor 1

→ Data Conversion

↓

Conditional Split

Length = 7   Length = 8

Handle data of length 7   Handle data of length 8

→ Union All ←

↓

OLE DB Destination
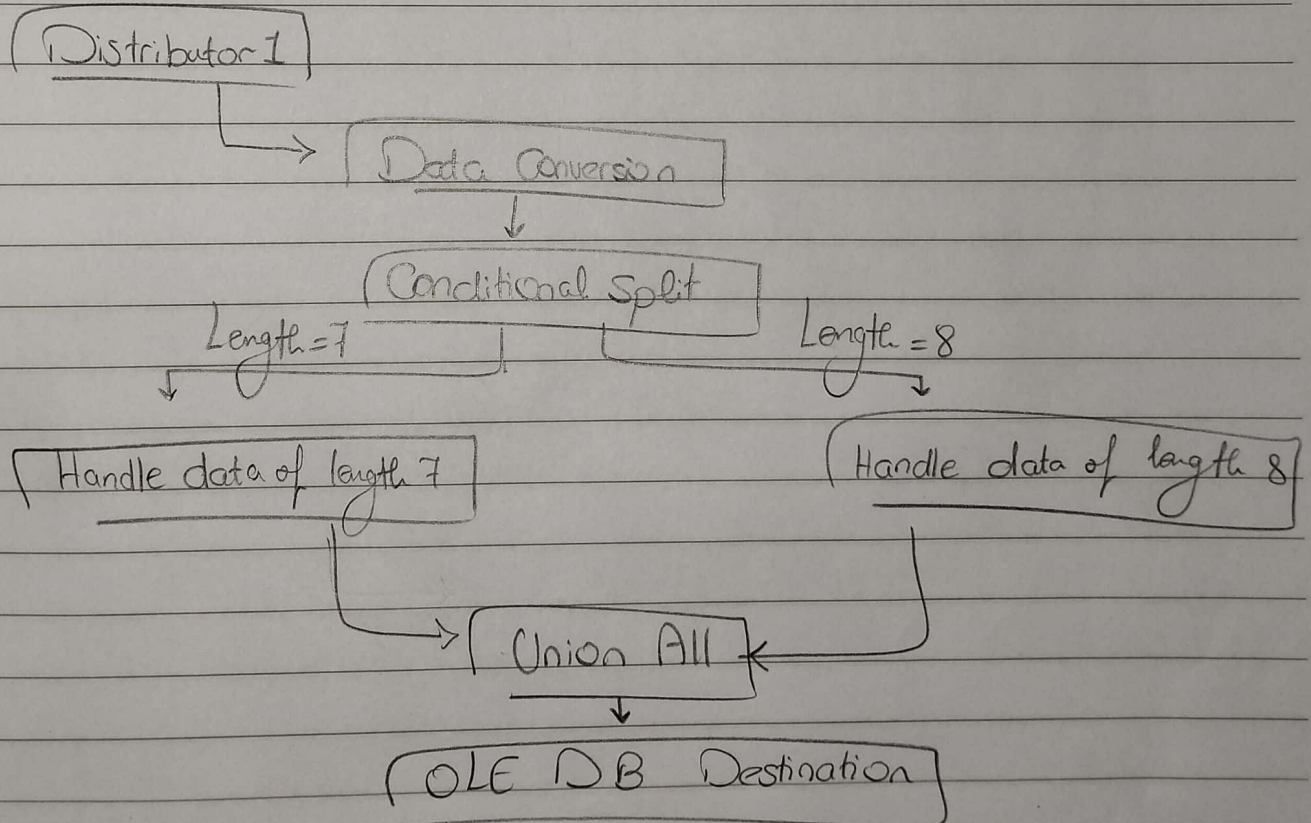
"To obtain single version of truth"

* أحنا استخدم ال union all عن أن نجمع results و في الاخر انا كان ايلي احجة ~ تجمع rows
أشوف بيبي في الاخر و انا بحرب بس كل شيء يتجمع (rows)
...احمد ال m