

# **Analysis of Customer Orders in Retail Application**

## **From Oracle Database**

# Project Ideas



As the business bloom nowadays, the analysis of customer behavior becomes more and more important. By analysing the orders of customers, it will provide massive information for decision-makers.

- Understanding customer preferences
- Optimizing inventory management
- Personalizing marketing and sales efforts
- Predicting future demand

# Objectives & Deliverables

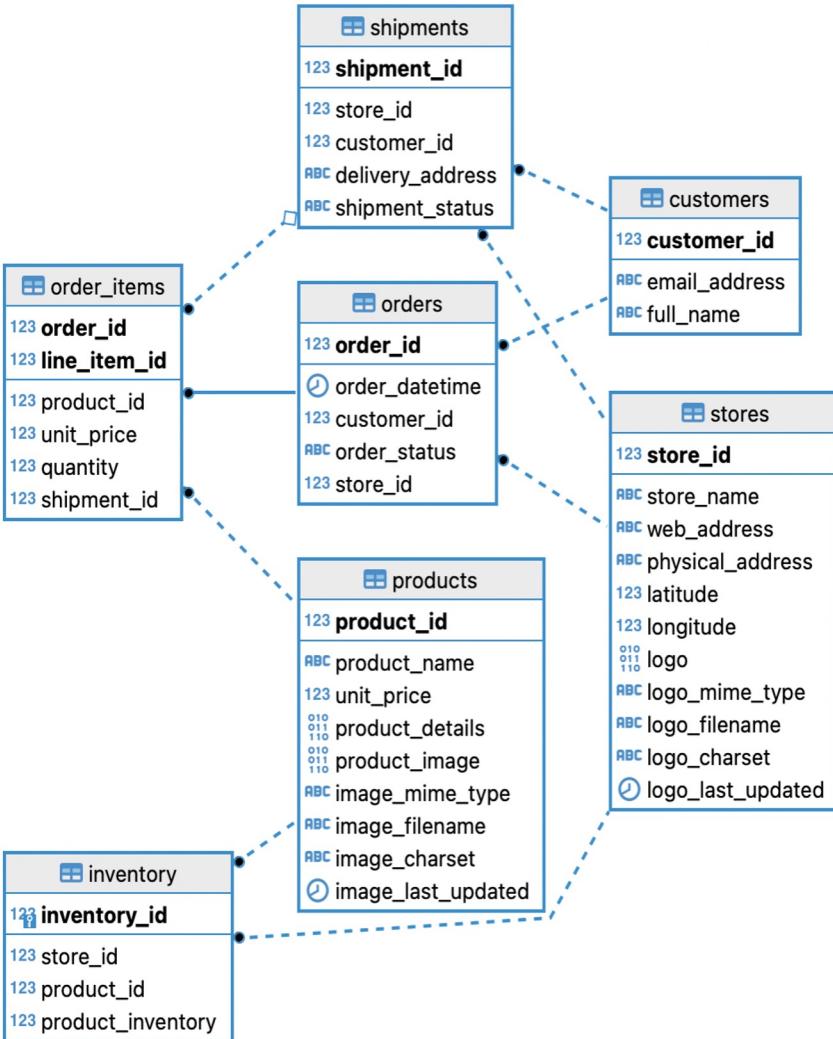
## Objectives:

- Create a database and data mart to track custom orders and relative information
- Improve inventory management and production efficiency by analyzing custom order data
- Increase customer satisfaction and track the order status specifically



## Deliverables:

- Operational database and data mart in SQL Server
- Automate the process of Extracting, Transforming, and Loading data from multiple sources by ETL and SSIS
- Multi-dimension Cube by using SSAS

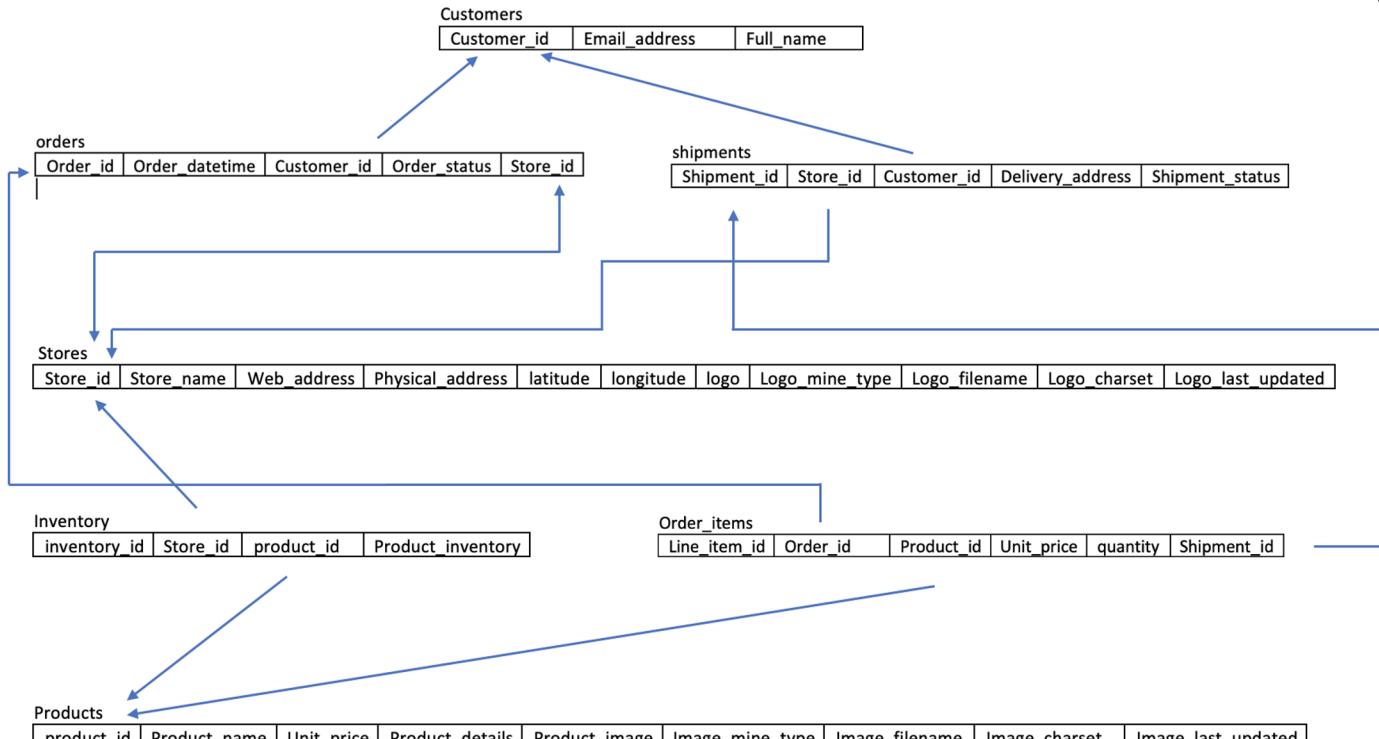


# Operational Database

## -ER Diagram

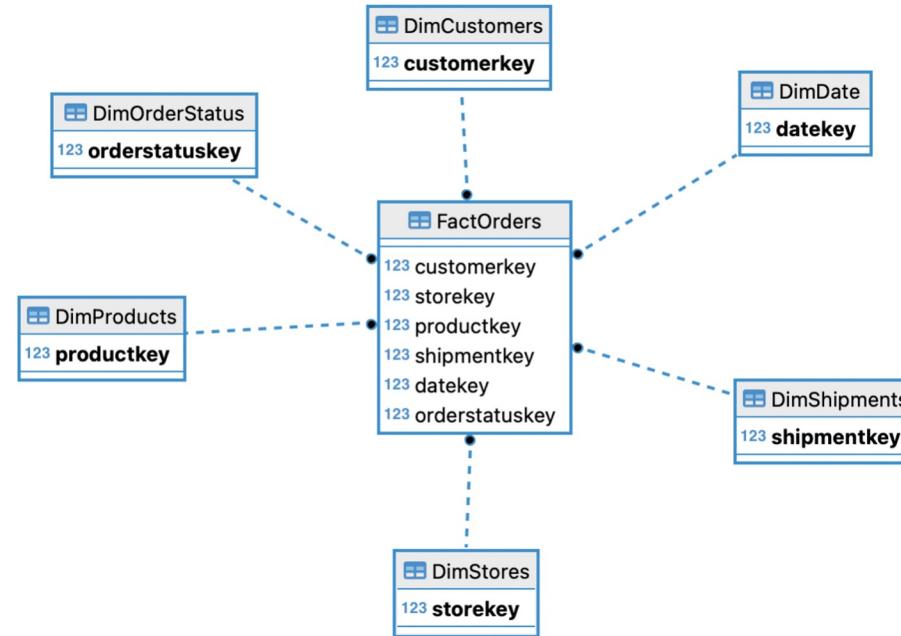
# Operational Database

-3rd NF



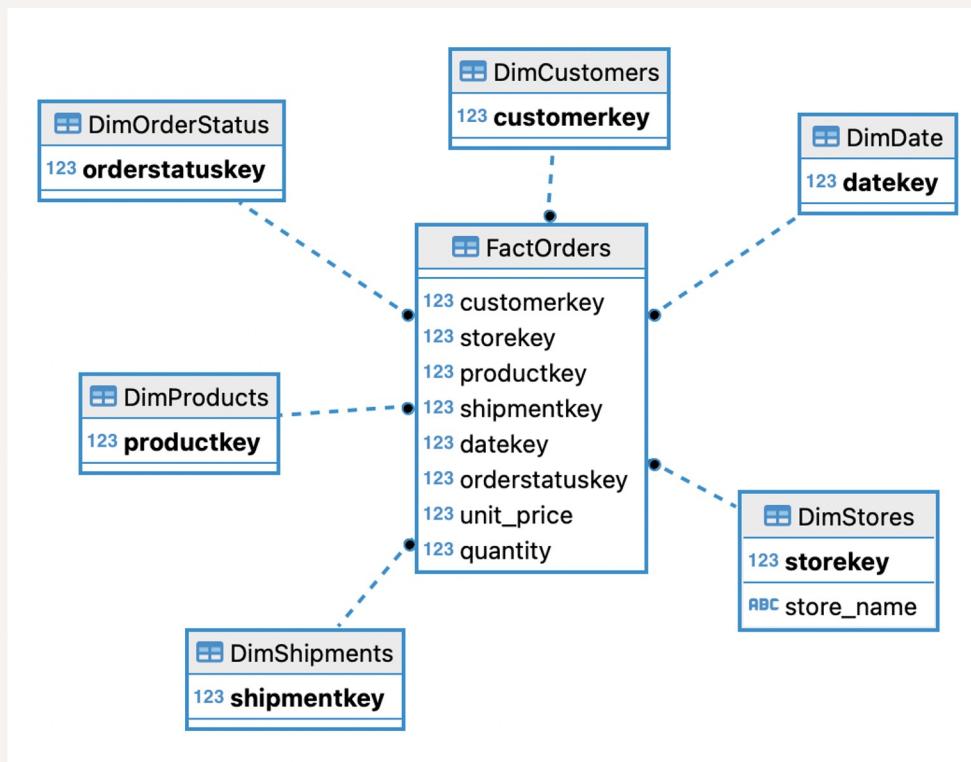
# Star schema

- Decide on the dimensions and the measures
- Adding primary key and foreign key



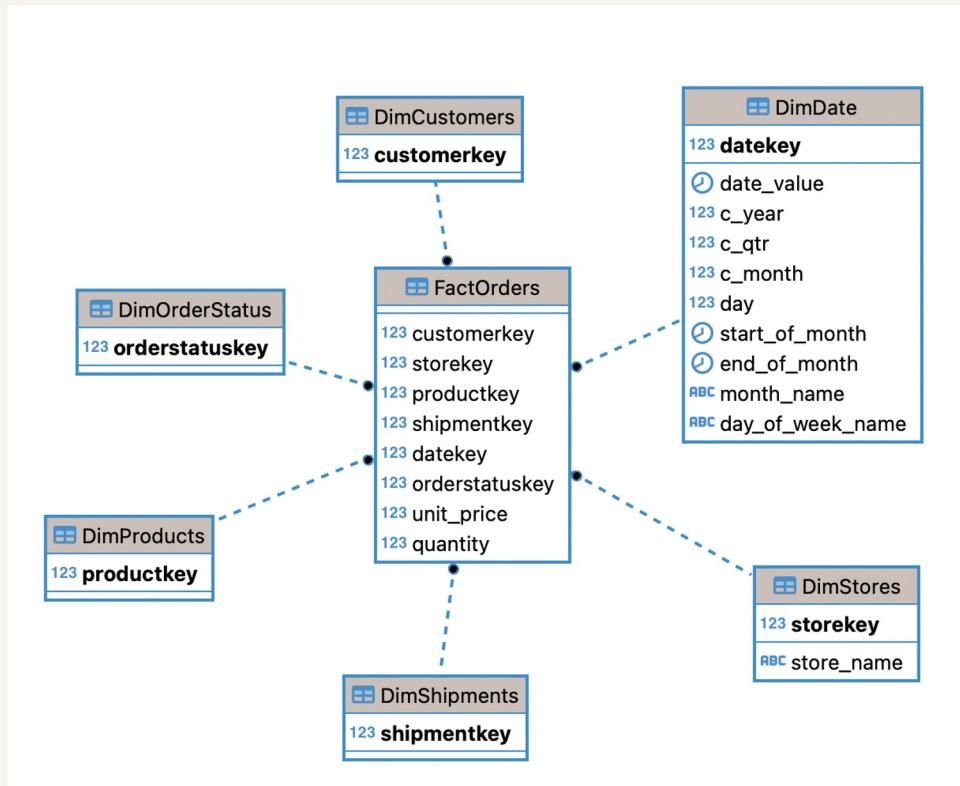
# Star schema

- Adding measurement into fact table

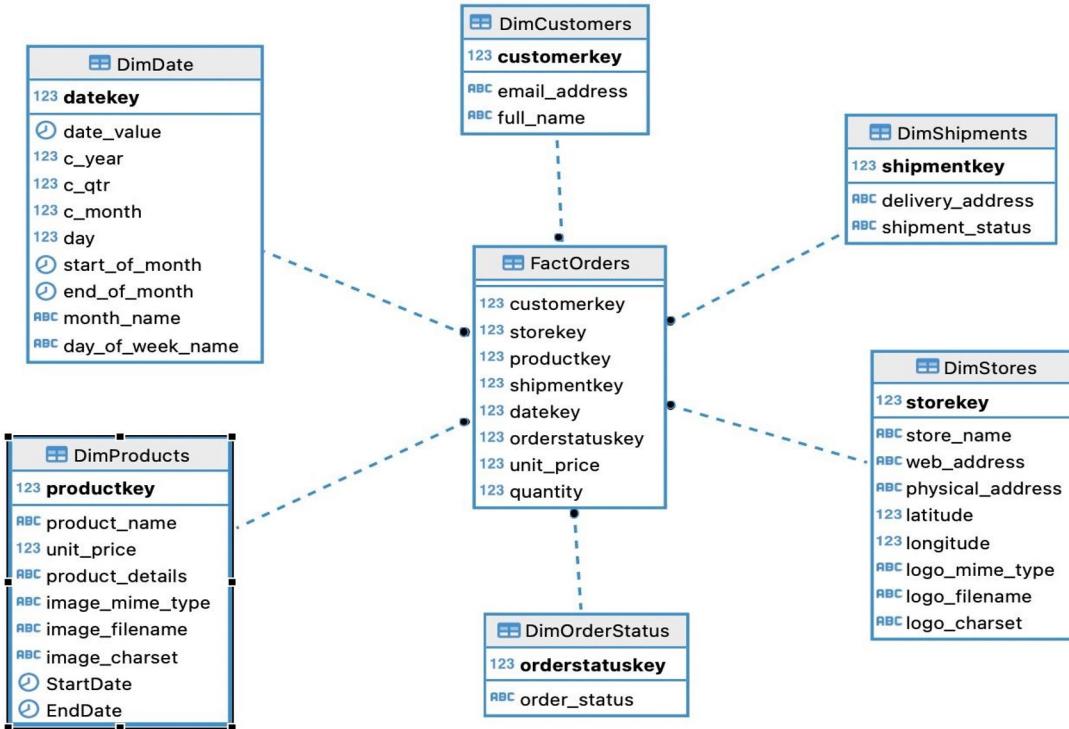


# Star schema

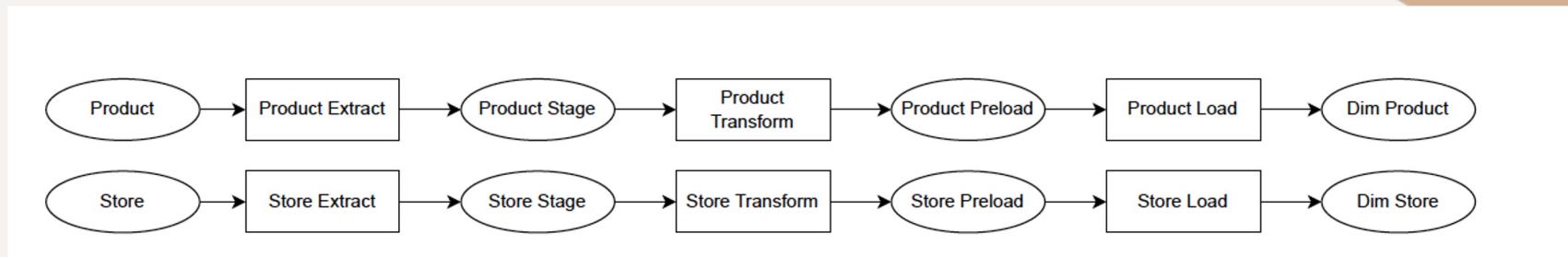
- Represent the hierarchy of date dimension table



# Star schema



# ETL Diagram



- > Products\_Extract
- > Products\_Load
- > Products\_Transform
- > Stores\_Extract
- > Stores\_Load
- > Stores\_Transform

- > Products\_Preload
- > Products\_stage
- > Stores\_Preload
- > Stores\_stage

# DimStores - SCD Type 1

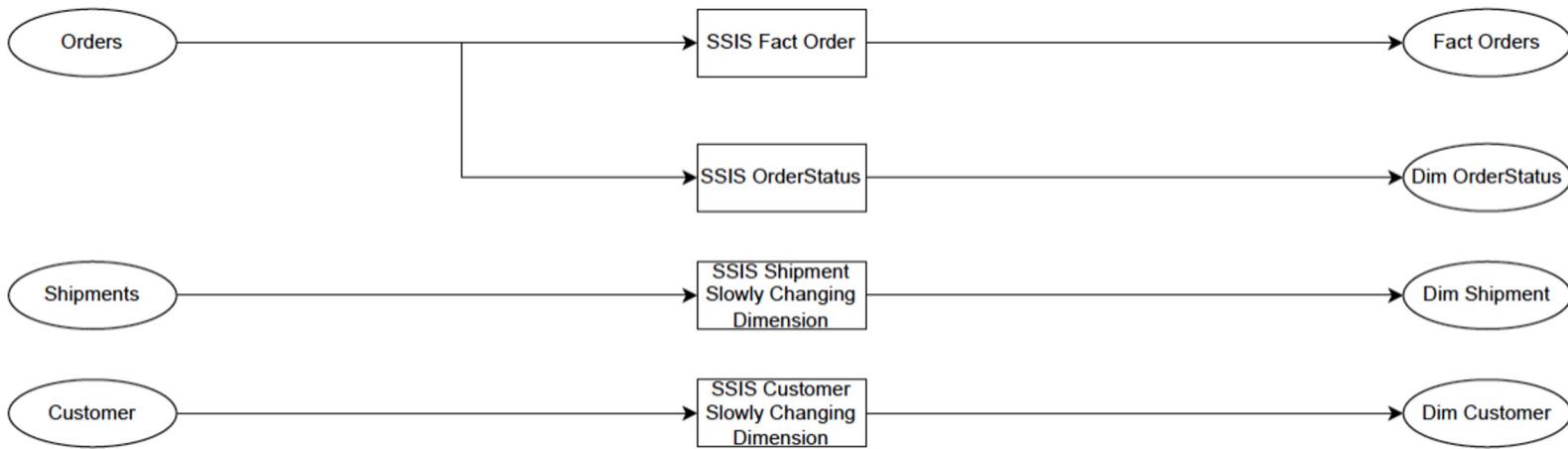
DimStores										Enter a SQL expression to filter results (use Ctrl+Space)			
Grid	123	ABC store_name	ABC web_address	ABC physical_address	123	123	latit	ABC logo_m	ABC logo_f	ABC logo_charset	123		
1	24	Online	<input checked="" type="checkbox"/> https://www.example.com	[NULL]			NULL]	[NULL]	[NULL]	[NULL]	[NULL]		
2	25	San Francisco	[NULL]	Redwood Shores¶ 500 Oracle Parkway¶ Rm -122			38	[NULL]	[NULL]	[NULL]	[NULL]		
3	26	Seattle	[NULL]	1501 Fourth Avenue¶ Suite 1800¶ Seattle, WA 98101			48	[NULL]	[NULL]	[NULL]	[NULL]		
4	27	New York City	[NULL]	205 Lexington Ave¶ 7th Floor¶ New York, NY 10017			-74	41	[NULL]	[NULL]	[NULL]		
5	28	Chicago	[NULL]	233 South Wacker Dr.¶ 45th Floor¶ Chicago, IL 60606			-88	42	[NULL]	[NULL]	[NULL]		
6	29	London	[NULL]	One South Place¶ London, EC2M 2RB			0	52	[NULL]	[NULL]	[NULL]		
7	30	Bucharest	[NULL]	Floreasca Park¶ 43 Soseaua Pipera, corp B, et. 26			26	44	[NULL]	[NULL]	[NULL]		
8	31	Berlin	<input checked="" type="checkbox"/> https://www.example.com	[NULL]			NULL]	[NULL]	[NULL]	[NULL]	[NULL]		
9	32	Utrecht	[NULL]	Hertogsweertweg 163-167, 3543 AS Utrecht, NL			5	52	[NULL]	[NULL]	[NULL]		
10	33	Madrid	[NULL]	C/ José Echegaray 6B¶ Las Rozas, Madrid 28230			-4	40	[NULL]	[NULL]	[NULL]		
11	34	Johannesburg	[NULL]	Woodmead North Office Park¶ 54 Maxwell Drive			28	-26	[NULL]	[NULL]	[NULL]		
12	35	Lagos	[NULL]	1391 Tiamiyu Savage St, Victoria Island, Lagos			3	6	[NULL]	[NULL]	[NULL]		
13	36	Bengaluru	[NULL]	193, Bannerghatta Main Rd, Industrial Area, Bengaluru			78	13	[NULL]	[NULL]	[NULL]		
14	37	Mumbai	[NULL]	First International Financial Center¶ Unit No. 73			73	19	[NULL]	[NULL]	[NULL]		
15	38	New Dehli	[NULL]	F-01/02, First Floor¶ Salcon Rasvillas, D-1, New Dehli			77	29	[NULL]	[NULL]	[NULL]		
16	39	Sydney	[NULL]	Riverside Corporate Park¶ 4 Julius Avenue			151	-34	[NULL]	[NULL]	[NULL]		
17	40	Perth	[NULL]	Level 9¶ 225 St Georges Terrace¶ Perth, WA 6000			116	-32	[NULL]	[NULL]	[NULL]		

# DimProducts - SCD Type2

DimProducts | Enter a SQL expression to filter results (use Ctrl+Space)

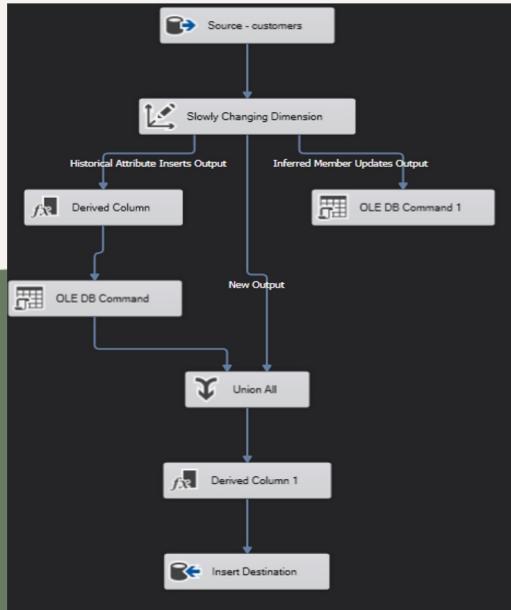
	123	ABC	product_name	123	ABC	ABC	ABC	ABC	StartDate	EndDate
Grid				ur	ur	image	image	image		
Text										
1	1	Boy's Shirt (White)	29.55	L	FLOWER	ENVIRONMENT	[NULL]	[NULL]	2013-01-01	[NULL]
2	2	Women's Shirt (Green)	16.67	XXL	SUN	ENVIRONMENT	[NULL]	[NULL]	2013-01-01	[NULL]
3	3	Boy's Sweater (Green)	44.17	S			[NULL]	[NULL]	2013-01-01	[NULL]
4	4	Boy's Trousers (White)	43.71	M	BEACH	ENVIRONMENT	[NULL]	[NULL]	2013-01-01	[NULL]
5	5	Girl's Shorts (Red)	38.28	XL	[NULL]	[NULL]	[NULL]	[NULL]	2013-01-01	[NULL]
6	6	Boy's Socks (Grey)	19.16	S	FLOWER	ENVIRONMENT	[NULL]	[NULL]	2013-01-01	[NULL]
7	7	Boy's Socks (Black)	19.58	M	SEA	ENVIRONMENT	[NULL]	[NULL]	2013-01-01	[NULL]
8	8	Boy's Coat (Brown)	21.16	XL	[NULL]	[NULL]	[NULL]	[NULL]	2013-01-01	[NULL]

# SSIS



# SSIS Slowly Changing Dimension

## Example(DimCustomer - SCD Type 2)



Select a Dimension Table and Keys

Select a dimension table to load and map columns in the transformation input to

Connection manager: DestinationConnectionOLEDB [New...](#)

Table or view: [dbo].[DimCustomers]

Input Columns	Dimension Columns	Key Type
customer_id	customerkey	Business key
email_address	email_address	Not a key column
	EndDate	
full_name	full_name	Not a key column
	StartDate	

# SSIS Slowly Changing Dimension

## Example(DimCustomer - SCD Type 2)

**Slowly Changing Dimension Columns**  
Manage the changes to column data in your slowly changing dimensions by setting the

**Fixed Attribute**  
Select this type when the value in a column should not change. Changes are treated as errors.

**Changing Attribute**  
Select this type when changed values should overwrite existing values. This is a Type 1 change.

**Historical Attribute**  
Select this type when changes in column values are saved in new

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
email_address	Historical a...
full_name	Historical a...

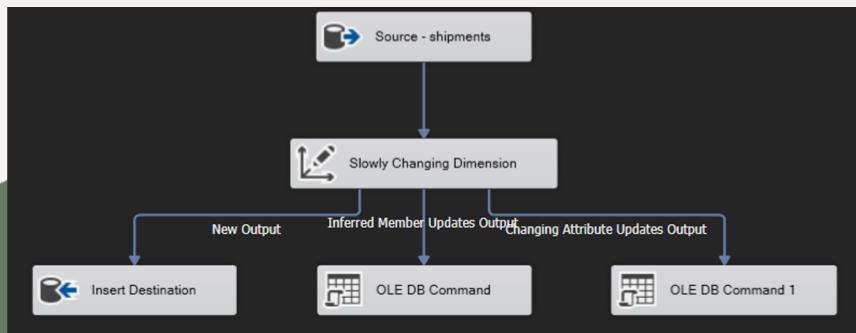
**Historical Attribute Options**  
You can record historical attributes using a single column or start and end date columns.

Use a single column to show current and expired records  
Column to indicate current record:  
Value when current: \_\_\_\_\_  
Expiration value: \_\_\_\_\_

Use start and end dates to identify current and expired records  
Start date column: StartDate  
End date column: EndDate  
Variable to set date values: System::ContainerStartTime

# SSIS Slowly Changing Dimension

## Example(DimShipment - SCD Type 1)



**Select a Dimension Table and Keys**  
Select a dimension table to load and map columns in the transformation input to

Connection manager: **SourceConnectionOLEDB** [New...](#)

Table or view: **[dbo].[DimShipments]**

Input Columns	Dimension Columns	Key Type
delivery_addr...	delivery_address	Not a key column
shipment_sta...	shipment_status	Not a key column
shipment_id	shipmentkey	Business key

# SSIS Slowly Changing Dimension

## Example(DimShipment - SCD Type 1)

**Slowly Changing Dimension Columns**

Manage the changes to column data in your slowly changing dimensions by setting the



**Fixed Attribute**

Select this type when the value in a column should not change. Changes are treated as errors.

**Changing Attribute**

Select this type when changed values should overwrite existing values. This is a Type 1 change.

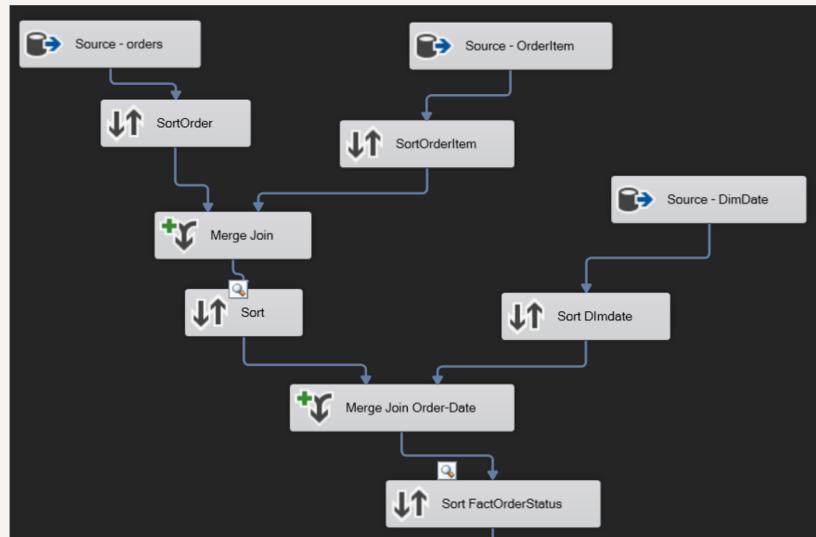
**Historical Attribute**

Select this type when changes in column values are saved in new

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
delivery_address	Changing a...
shipment_status	Changing a...

# SSIS Example (Fact Order)



# SSIS Example (Fact Order)

Sort :

The screenshot shows the 'Available Input Columns' list and a corresponding 'Input Column' mapping table.

**Available Input Columns:**

Name	Pass Thru...
<input checked="" type="checkbox"/> order_id	<input type="button" value="..."/>
<input type="checkbox"/> line_item_id	<input checked="" type="checkbox"/>
<input type="checkbox"/> product_id	<input checked="" type="checkbox"/>
<input type="checkbox"/> unit_price	<input checked="" type="checkbox"/>
<input type="checkbox"/> quantity	<input checked="" type="checkbox"/>

**Input Column Mapping:**

Input Column	Output Alias	Sort Type	Sort Order	Con...
order_id	order_id	ascending	1	

Join:

The screenshot shows two 'SortOrder' components and a 'SortOrderItem' component, with a connection line between the first two.

**SortOrder Components:**

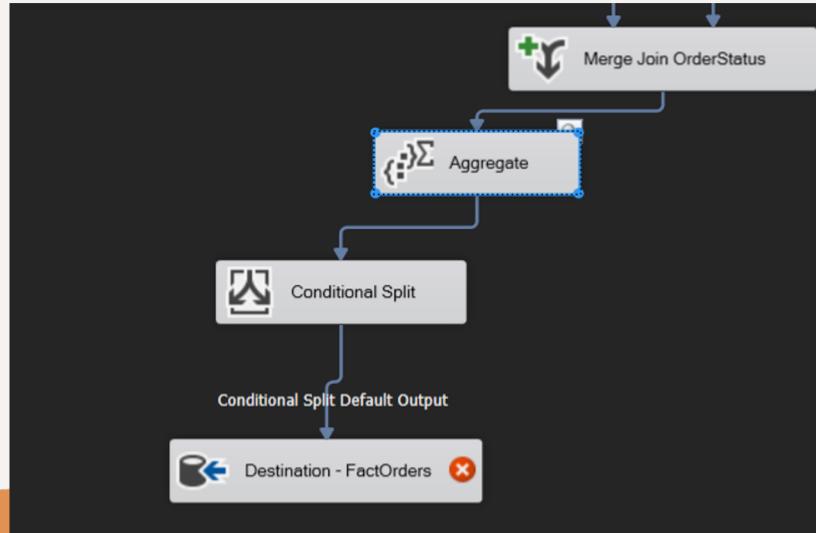
Name	Ord...	Join...
<input checked="" type="checkbox"/> order_id	1	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> order_datetime	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> customer_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> order_status	0	<input type="checkbox"/>

Name	Ord...	Join...
<input type="checkbox"/> order_id	1	<input checked="" type="checkbox"/>
<input type="checkbox"/> line_item_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> product_id	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> unit_price	0	<input type="checkbox"/>
<input checked="" type="checkbox"/> quantity	0	<input type="checkbox"/>

**Input Column Mapping:**

Input	Input Column	Output Alias
SortOrder	order_id	order_id
SortOrder	order_datetime	order_datetime
SortOrder	customer_id	customer_id
SortOrder	order_status	order_status
SortOrder	store_id	store_id
SortOrderItem	line_item_id	line_item_id
SortOrderItem	product_id	product_id
SortOrderItem	unit_price	unit_price
SortOrderItem	quantity	quantity

# SSIS Example (Fact Order)



# SSIS Example (Fact Order)

Aggregate :

Input Column	Output Alias	Operation
shipment_id	shipment_id	Group by
orderstatuskey	orderstatuskey	Group by
datekey	datekey	Group by
unit_price	unit_price	Average
quantity	quantity	Sum

Split :

Order	Output Name	Condition
1	RemoveNull	ISNULL(shipment_id)    ISNULL(order_id)    ISNULL(cu...

# SSIS Example (Fact Order)

- Result of Top 1000 rows

Results Messages

	customerkey	storekey	productkey	shipmentkey	datekey	orderstatuskey	unit_price	quantity
1	31	1	27	602	20180622	3	39.91	3
2	182	1	44	1091	20180916	3	39.32	2
3	193	1	41	331	20180504	3	8.66	4
4	305	1	40	883	20180814	3	34.06	2
5	355	1	40	1218	20181010	3	34.06	2
6	172	1	18	266	20180424	3	24.46	2
7	172	1	14	1657	20181228	3	26.14	5
8	365	1	6	368	20180510	3	38.28	4
9	43	1	28	1803	20190123	3	10.24	2
10	87	1	14	934	20180823	3	26.14	3
11	154	1	37	343	20180505	3	29.51	5
12	202	1	41	1735	20190113	3	8.66	3
13	32	1	9	440	20180524	3	21.16	2
14	121	1	27	72	20180314	3	39.91	4

Query executed successfully.

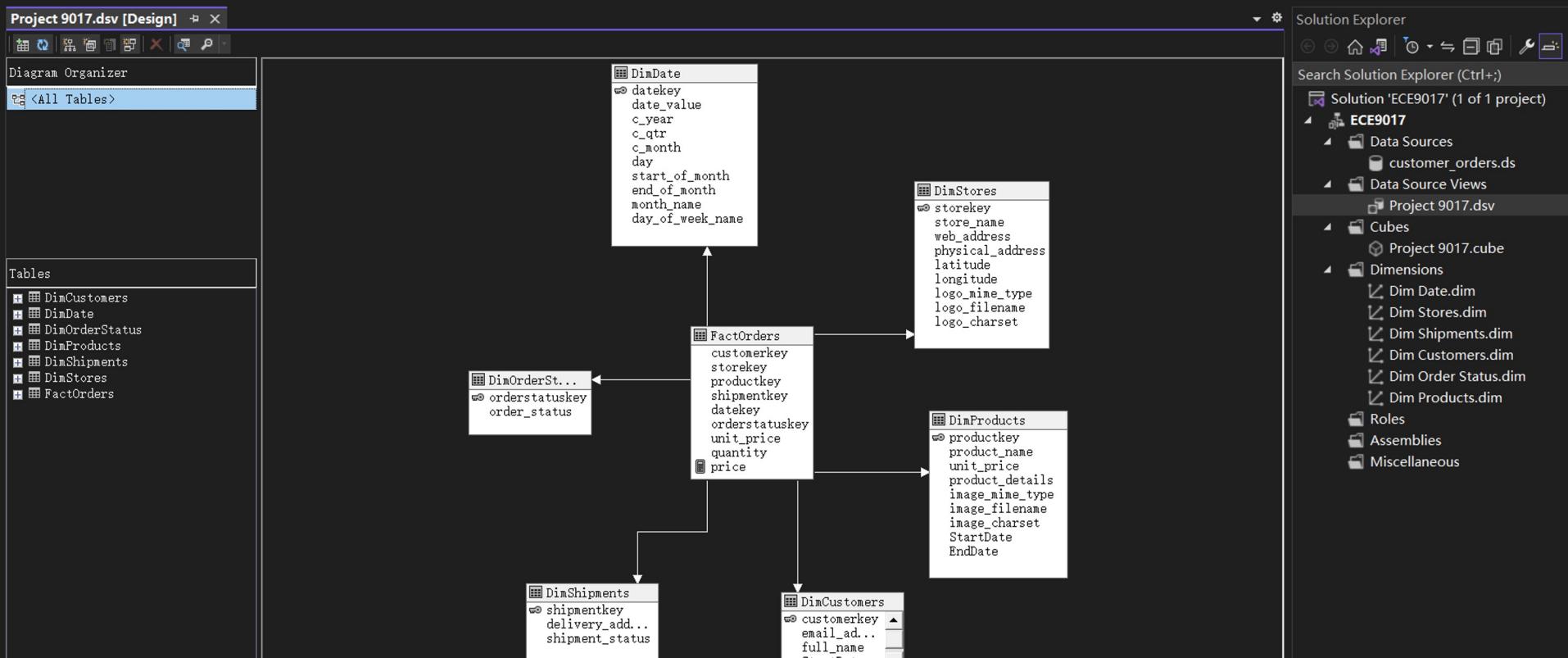
| LAPTOP-15UNHFNB\ECE9017 (15... | LAPTOP-15UNHFNB\zacky ... | CO | 00:00:00 | 1,000 rows

# SSAS

- Help businesses gain deeper insights into the data
- Advanced calculations and interactive reporting
- Providing powerful analysis

# Data Source View

Creating a data source view that defines the relationships between the data sources and the cube, and determines how the data is organized and presented in the cube.



# Dimension Hierarchies

Defining the hierarchies of dimensions in the cube, which can be used to drill down into the data and analyze it from different perspectives.

The screenshot shows a user interface for defining dimension hierarchies. The top navigation bar includes tabs for "Dimension Struct...", "Attribute Relations...", "Translations...", and "Browser". Below the navigation is a toolbar with various icons. The main area is divided into three panels:

- Attributes:** A list of attributes under the "Dim Date" dimension, including c Month, c Qtr, c Year, Date Value, Datekey, Day, Day Of Week Name, End Of Month, Month Name, and Start Of Month.
- Hierarchies:** A panel titled "层次结构" (Hierarchies) with a warning icon. It lists existing hierarchy levels: c Year, c Qtr, c Month, and Day, along with a placeholder "<new level>". A tooltip explains: "To create a new hierarchy, drag an attribute here."
- Data Source View:** A panel showing the data source view for the "Dim Date" dimension, listing "DimDate".

# Dimension Hierarchies

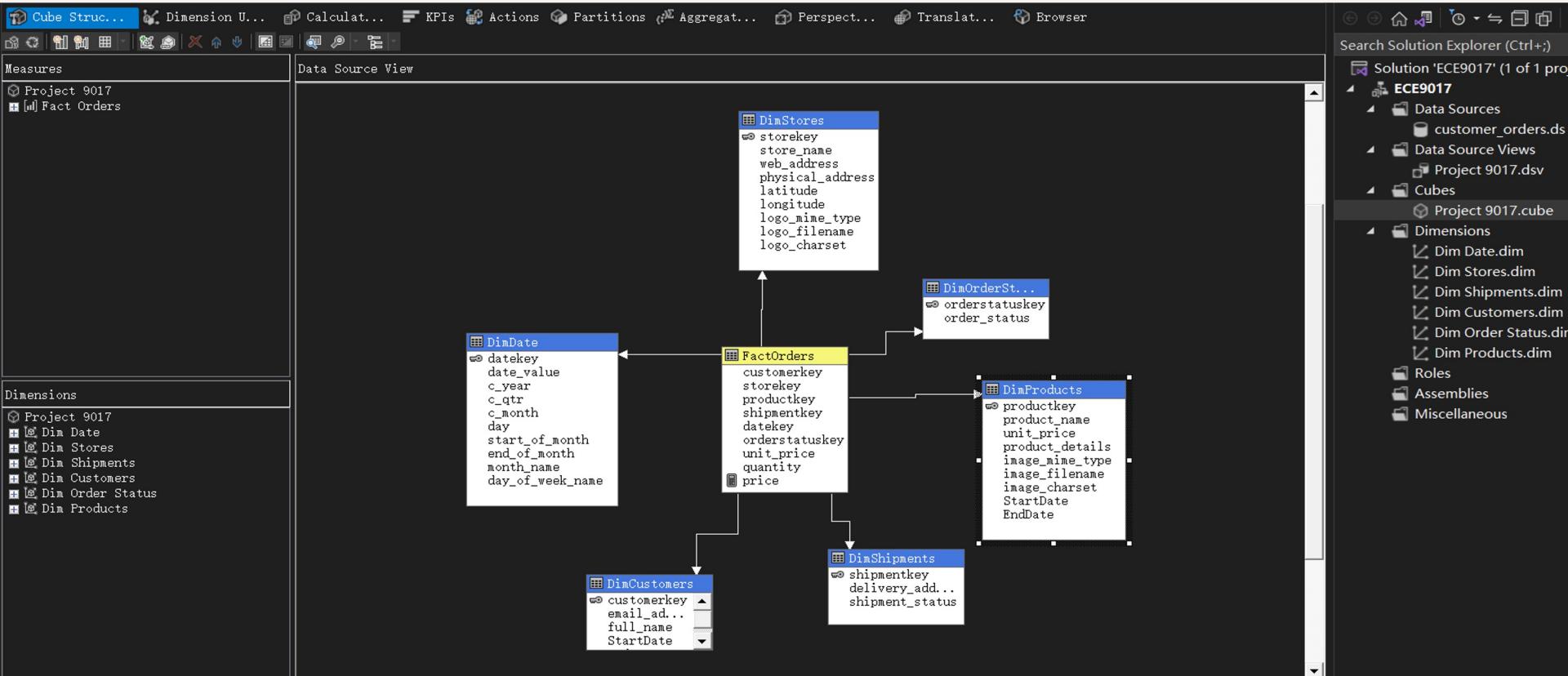
Screenshot of the Analysis Services Dimension Designer interface showing the configuration of a dimension hierarchy.

The interface includes:

- Top navigation bar with tabs: Project 9017 [Browse], SQLQuery3.sql, SQLQuery2.sql, and LAPTOP-31OCEV...o.DimCustomers.
- Language dropdown set to Default.
- Toolbars for Edit as Text, Import..., MDX, and various icons.
- Left sidebar with sections: Project 9017, Metadata, Search Model, Measure Group, and a tree view of dimensions and hierarchies.
- Tree view on the left:
  - Day Of Week Name
  - End Of Month
  - Month Name
  - Start Of Month
    - Members
    - Start Of Month
  - 层次结构
    - Members
    - c Year
    - c Qtr
    - c Month
    - Day
  - Dim Order Status
  - Order Status
  - Members
- Main workspace:
  - Table: Dimension, Hierarchy, Operator, Filter Expression.
  - Table: <Select dimension>, Hierarchy, Operator, Filter Expression.
  - Table: c Year, AVG PRICE (2018, 26.45798...).

# Cube Structure

Defining the structure of the cube, including the measures, dimensions, and hierarchies that will be included in the cube.



# Measures and Measure Group

Defining the measures that will be used to analyze the data, and grouping them into logical sets called measure groups.

The screenshot shows the 'Project 9017' metadata interface. At the top, there's a 'Search Model' bar and a 'Measure Group:' dropdown set to '<All>'. Below this is a tree view of the data model:

- Project 9017
  - Measures
    - Fact Orders
      - Avg Quantity
      - Avg Unit Price
      - Fact Orders Count
      - Max Unit Price
      - Min Unit Price
      - Price
      - Quantity
    - Avg Price
  - KPIs
    - SalesKPI
  - Dim Customers
  - Dim Date
    - c Month
    - c Qtr
    - c Year

# Dimension Usage

Defining how the dimensions in the cube will be used to slice and dice the data, and how they will relate to each other.

The screenshot shows the 'Dimension Usage' dialog box in Analysis Services Management Studio. The dialog is titled 'Define Relationship'. It lists dimensions on the left and their corresponding keys in the Fact Orders measure group on the right. A preview window on the left shows a blue icon for the dimension and a yellow icon for the fact table. The main area contains settings for the relationship:

- Select relationship type: Regular
- The dimension table is joined directly to the fact table.
- Granularity attribute: Datekey
- Dimension table: DimDate
- Measure group table: FactOrders
- Relationship:
  - Dimension Columns: datekey
  - Measure Group Columns: datekey

At the bottom right is an 'Advanced...' button.

# Calculation

Defining calculated measures and members in the cube, which can be used to perform advanced analysis and derive insights from the data.

The screenshot shows the Analysis Services Management Studio interface. On the left, the Script Organizer pane displays a tree structure with a node labeled 'CALCULATE' expanded, and a child node '2 [AVG PRICE]' selected. The main workspace shows the properties for this calculated member:

- Name:** [AVG PRICE]
- Parent Properties:** Parent hierarchy: Measures, Parent member: (empty)
- Expression:** [Price]/[Quantity]  
No issues found
- Additional Properties:** Format string: (empty), Visible: True, Non-empty behavior: (empty), Associated measure group: (Undefined), Display folder: (empty), Color Expressions: (empty), Font Expressions: (empty)

At the bottom right of the workspace, there are buttons for Ln: 1, Ch: 19, SPC, and CRLF.

# Calculation

The screenshot shows the SSAS Management Studio interface. On the left, the 'Project 9017' navigation pane is open, displaying the structure of the project. Under the 'Measures' node, there is a folder named 'Fact Orders' which contains a calculated member named 'AVG PRICE'. This member is highlighted with a blue background.

The main workspace displays a table with two rows. The columns are labeled 'Year' and 'AVG PRICE'. The first row shows the year 2018 with a value of 26.1099879299941. The second row shows the year 2019 with a value of 26.7100644699139. A red rectangular box highlights this table.

c Year	AVG PRICE
2018	26.1099879299941
2019	26.7100644699139

# KPIs

Defining key performance indicators (KPIs) in the cube, which can be used to monitor and measure the performance of the business based on predefined metrics.

The screenshot shows the KPI Organizer interface in Analysis Services Management Studio. The left sidebar contains navigation links like Cube Structure, Dimension Usage, Calculations, KPIs (which is selected), Actions, Partitions, Aggregates, Perspectives, Translations, and Browser. Below that are Calculation Tools, Metadata, Functions, Templates, and Search Model. A Measure Group dropdown is set to <All>. The main pane displays the configuration for a KPI named "SalesKPI".

**KPI**

- Name: SalesKPI
- Associated measure group: <All>

**Value Expression**

```
[Measures].[Price]
```

No issues found

Ln: 1 Ch: 19 SPC CRLF

**Goal Expression**

```
100000
```

No issues found

Ln: 1 Ch: 7 SPC CRLF

**Status**

- Status indicator: Gauge
- Status expression:

```
CASE
WHEN kpivalue('SalesKPI')<kpigoal('SalesKPI') THEN -1
WHEN kpivalue('SalesKPI')>kpigoal('SalesKPI') THEN 1
ELSE 0
END
```

No issues found

Ln: 5 Ch: 4 SPC CRLF

**Trend**

- Trend indicator: Standard arrow
- Trend expression:

# KPIs

Project 9017 [Browse] SQLQuery3.sql ...EVMN\dqsun (53)\* SQLQuery2.sql ...EVMN\dqsun (56) LAPTOP-31OCEV...o.DimCustomers

Language: Default

Edit as Text Import... MDX

Dimension Hierarchy Operator Filter Expression

<Select dimension>

Project 9017

Metadata

Search Model

Measure Group:

<All>

Project 9017

Measures

Fact Orders

- AVG Quantity
- AVG Unit Price
- Fact Orders 计数
- Price
- Quantity
- 最大的 Unit Price
- 最小的 Unit Price

Avg PRICE

KPIs

- SalesKPI

Dim Customers

- Customerkey
- Full Name

Dim Date

- c Month
- c Qtr
- Members

c Year	SalesKPI Value	SalesKPI Goal	SalesKPI Status
2018	346114.000...	100000	1
2019	74574.4999...	100000	-1

# MDX Query-1

Using the Multidimensional Expressions (MDX) query language to retrieve data from the cube, perform complex calculations, and analyze the data in various ways.



```
SELECT NON EMPTY
    { [Measures].[Price], [Measures].[AVG PRICE], [Measures].[Fact Orders Count]
} ON COLUMNS,
NON EMPTY
    { ([Dim Date].[c Year].[c Year].ALLMEMBERS *
[Dim Order Status].[Order Status].[Order Status].ALLMEMBERS )
} ON ROWS
FROM [Project 9017]
```

c Year	Order Status	Price	AVG PRICE	Fact Orders Count
2018	COMPLETE	343717.580000002	26.138219011407	4354
2018	REFUNDED	2396.42	22.6077358490566	38
2019	COMPLETE	73316.6399999997	26.7578978102189	900
2019	REFUNDED	1257.86	24.1896153846154	16

# MDX Query-2



```
SELECT
NON EMPTY
{
    [Measures].[Price], [Measures].[AVG PRICE]
} ON COLUMNS,

NON EMPTY
{ ([Dim Stores].[Store Name].[Store Name].ALLMEMBERS * [Dim Date].[c Year].[c Year].ALLMEMBERS )
}
ON ROWS

FROM [Project 9017]
```

Store Name	c Year	Price	AVG PRICE
Online	2018	346114.00000...	26.1099879299941
Online	2019	74574.499999...	26.7100644699139

# Excel Pivot

行标签	Price	AVG PRICE	Fact Orders Count
2018			
COMPLETE	343717.58	26.13821901	4354
REFUNDED	2396.42	22.60773585	38
2019			
COMPLETE	73316.64	26.75789781	900
REFUNDED	1257.86	24.18961538	16
总计	420688.5	26.21438809	5308

a. Data showed in Excel Pivot Table

c Year	Order Status	Price	AVG PRICE	Fact Orders Count
2018	COMPLETE	343717.5800000002	26.138219011407	4354
2018	REFUNDED	2396.42	22.6077358490566	38
2019	COMPLETE	73316.6399999997	26.7578978102189	900
2019	REFUNDED	1257.86	24.1896153846154	16

b. Previous data showed in browser

The screenshot shows a browser-based pivot table interface. At the top right is a search bar with a magnifying glass icon. Below it is a sidebar with several filter checkboxes:

- AVG PRICE
- AVG Quantity
- AVG Unit Price
- Fact Orders Count
- Max Unit Price
- Min Unit Price
- Price
- Quantity

Below the sidebar is a section labeled "KPI" with a small icon. To the right of the sidebar is a message: "在以下区域间拖动字段:" (Drag fields between these areas). On the far right is a vertical scroll bar.

At the bottom, there are two main sections: "筛选" (Filter) and "列" (Columns).

**筛选 (Filter):**

- Σ 数值 (Σ Value)

**行 (Rows):**

- 层次结构 (Hierarchical structure) dropdown menu
- Order Status dropdown menu

**值 (Values):**

- Price dropdown menu
- AVG PRICE dropdown menu
- Fact Orders Count dropdown menu



# Result & Conclusion

**By ETL process using Stored procedure and SSIS :**

- Stored Procedure
- SSIS
  - ETL for Type 1 & 2 SCD.
  - Transform data from multiple Dim tables to Fact Table

**By multi-dimension analysis in SSAS:**

- Customer data about the average price
  - Determine pricing strategies
  - Assess supplier performance
- Track progress towards our sales goals
  - Identify performance gaps
  - Determine areas for improvement
  - Monitor progress towards overall business goals

# Thank you!

