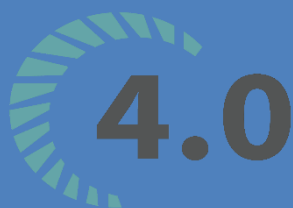


BỘ MÔN HỆ THỐNG THÔNG TIN – KHOA CÔNG NGHỆ THÔNG TIN
ĐẠI HỌC KHOA HỌC TỰ NHIÊN THÀNH PHỐ HỒ CHÍ MINH, ĐẠI HỌC QUỐC GIA TP HCM

HỆ THỐNG THÔNG TIN PHỤC VỤ TRÍ TUỆ KINH DOANH



Sinh viên thực hiện:

18127274 – Nguyễn Lê Đức Hoàng

18127241 – Nguyễn Đăng Triều

18127086 – Lê Thị Thùy Dương

ĐỒ ÁN/BÀI TẬP MÔN HỌC
HỌC KỲ I – NĂM HỌC 2021-2022



BẢNG THÔNG TIN CHI TIẾT CÁ NHÂN

STT	MSSV	Họ Tên	Mức độ hoàn thành	Đóng góp
1	18127274	Nguyễn Lê Đức Hoàng	100%	25%
2	18127241	Nguyễn Đăng Triều	100%	35%
3	18127086	Lê Thị Thùy Dương	100%	40%

Mục lục

Chương 1. Tổng quan về đề tài	4
1. Giới thiệu bộ dữ liệu	4
2. Mô tả dữ liệu.....	4
3. Dữ liệu liên quan yêu cầu phân tích.....	17
4. Phân tích profiling data	20
Chương 2. Quá trình SSIS	22
1. Source vào Stage.....	22
2. Stage vào NDS.....	25
3. NDS vào DDS	28
Chương 3: OLAP CUBE và MDX.....	56
1. Tạo ứng dụng SSAS	56
2. Tạo Data Source View.....	60
3. Tạo Cube.....	64
4. Sửa bảng chiều	70
5. Deploy cube lên server	71
6. Tạo mới Measure	72
7. Xóa dòng dữ liệu Unknow.....	73
Chương 4 Thực hiện ngôn ngữ MDX.....	74
1. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng (Fatal, Serious, Slight) ở các Địa phương (Local_Authority_(District)) trong tất cả các năm.....	74
2. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng ở các Địa Phương (Local_Authority_(District)) theo các Quý trong từng năm.....	75
3. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng và Thời Điểm Trong Ngày (Morning: 5am-12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) trong các năm.....	76

4. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng, Vùng (Urban_or_Rural_Area), và Kiểu Đường (Road Type) trong các năm.....	77
5. Tổng hợp số lượng tai nạn theo Mục Đích Hành Trình (Journey Purpose) và Loại Phương Tiện (Vehicle_Type).	78
6. Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng TNGT ở các vùng trong năm.	79
Chương 5: Trực quan hóa dữ liệu	81
1. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng (Fatal, Serious, Slight) ở các Địa phương (Local_Authority_(District)) trong tất cả các năm.....	81
2. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng ở các Địa Phương (Local_Authority_(District)) theo các Quý trong từng năm.....	81
3. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng và Thời Điểm Trong Ngày (Morning: 5am-12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) trong các năm.....	82
4. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng, Vùng (Urban_or_Rural_Area), và Kiểu Đường (Road Type) trong các năm.....	83
5. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng, Loại Nạn Nhân (Casualty Type) và Độ Tuổi trong các năm, Độ Tuổi được định nghĩa như sau:.....	84
6. Tổng hợp số lượng tai nạn theo Mục Đích Hành Trình (Journey Purpose) và Loại Phương Tiện (Vehicle_Type).	85
7. Thống kê số lượng tai nạn theo Mức Độ Nghiêm Trọng, Loại Phương Tiện (Vehicle Type), Built-up Road trong các năm.	86
8. Một vài biểu đồ cho tai nạn giao thông.....	86
9. Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng TNGT ở các vùng trong năm.	87
Chương 6: Data Mining.....	88

Chương 1. Tổng quan về đề tài

1. Giới thiệu bộ dữ liệu

Dữ liệu về các vụ tai nạn giao thông ở UK trong khoảng 3-4 năm.

- Dữ liệu UK Car Accidents 2005 - 2015 (SV chỉ lấy dữ liệu 3-4 năm, hoặc lấy dữ

liệu từ 2011-2014 được cung cấp sẵn):

<https://www.kaggle.com/silicon99/dft-accidentdata/discussion/28970?fbclid=IwAR1BvAiy8mEMy01XXAKtxLkX7Kx3kwPt3c3EYhwoxIWq5psikSAB2mVIF8A>

- Dữ liệu LSOA-Postcode mapping:

<https://geoportal.statistics.gov.uk/datasets/postcode-to-output-area-to-lower-layer-super-output-area-to-middle-layer-super-output-area-to-local-authority-district-august-2021-lookup-in-the-uk/about>

- Dữ liệu UK-Postcodes:

<https://github.com/academe/UK-Postcodes/blob/master/postcodes.csv>

2. Mô tả dữ liệu

2.1. Bảng Casualty

STT	Column	Description
1	Sex of Casualty	Giới tính xảy ra thương vong <ul style="list-style-type: none">• Giá trị 1 đại diện cho Nam• Giá trị 2 đại diện cho Nữ• Giá trị -1 dữ liệu bị thiếu hoặc ngoài phạm vi

2	Age Band of Casualty	<p>Các nhóm tuổi thương vong</p> <ul style="list-style-type: none"> • Giá trị 1 : nhóm tuổi từ 0 đến 5 tuổi • Giá trị 2 : nhóm tuổi từ 6 đến 10 tuổi • Giá trị 3 : nhóm tuổi từ 11 đến 15 tuổi • Giá trị 4 : nhóm tuổi từ 16 đến 20 tuổi • Giá trị 5 : nhóm tuổi từ 21 đến 25 tuổi • Giá trị 6 : nhóm tuổi từ 26 đến 35 tuổi • Giá trị 7 : nhóm tuổi từ 36 đến 45 tuổi • Giá trị 8 : nhóm tuổi từ 46 đến 55 tuổi • Giá trị 9 : nhóm tuổi từ 56 đến 65 tuổi • Giá trị 10 : nhóm tuổi từ 65 đến 75 tuổi • Giá trị 11 : nhóm tuổi từ 75 tuổi trở lên • Giá trị -1 : Mất dữ liệu hoặc nằm ngoài phạm vi
3	CasualtyType	<p>Các loại nạn nhân thương vong</p> <ul style="list-style-type: none"> • Giá trị 0 Người đi bộ • Giá trị 1 người đi xe đạp • Giá trị 2 Xe mô tô từ 50cc trở xuống cho người lái hoặc chở khách

		<ul style="list-style-type: none"> • Giá trị 3 Người lái hoặc chở khách từ 125cc trở xuống • Giá trị 4 Xe mô tô trên 125 phân khối đến người lái hoặc chở khách đến 500 phân khối • Giá trị 5 Xe mô tô trên 500cc chở hoặc người lái • Giá trị 8 Taxi / Thuê xe riêng • Giá trị 9 Người ngồi trên ô tô • Giá trị 10 xe buýt nhỏ (8-16 chỗ ngồi) hành khách • Giá trị 11 Người ngồi trên xe buýt hoặc xe khách (17 chỗ ngồi trở lên) • Giá trị 16 Người cưỡi ngựa • Giá trị 17 Người ngồi trên xe công nông • Giá trị 18 Xe điện • Giá trị 19 Xe Van / Xe chở hàng (3,5 tấn mgw trở xuống) • Giá trị 20 Xe chở hàng (người trên 3,5 tấn và dưới 7,5 tấn) • Giá trị 21 Xe chở hàng (7,5 tấn mgw trở lên) • Giá trị 22 Người lái xe tay ga cơ động • Giá trị 23 Người lái xe mô tô điện hoặc chở khách • Giá trị 90 Người ngồi trên xe khác • Giá trị 97 Xe mô tô - người lái hoặc hành khách phân khối không xác định
--	--	--

		<ul style="list-style-type: none"> Giá trị 98 Xe chở hàng (không rõ trọng lượng)
--	--	---

2.2. Bảng Accident

STT	Column	Description
1	Accident Severity	Đây là thuộc tính chỉ ra mức độ nghiêm trọng của 1 vụ tai nạn, gồm có 3 mức độ là(fatal, serious, slight).
2	Local Authority (District)	Thuộc tính chỉ ra tên của địa phương.
3	Time	Thời gian của vụ tai nạn.(giờ-phút)
4	Date	Ngày xảy ra tai nạn(DD/MM/YYYY).
5	Urban or Rural Area	Thuộc tính lưu thông tin khu vực xảy ra tai nạn. <ul style="list-style-type: none"> 1: Urban - khu đô thị 2: Rural - nông thôn 3: Unallocated - chưa phân bổ

6	Road Type	<p>Thuộc tính lưu thông tin loại đường nơi xảy ra tai nạn.</p> <ul style="list-style-type: none"> • 1: Roundabout - bùng binh • 2: One way street - đường một chiều • 3: Dual carriageway - đường hai chiều • 6: Single carriageway - đường đơn • 7: Slip road - đường trượt • 9: Unknown - không xác định • 12: One way street/Slip road - đường một chiều/ đường trượt • -1: Data missing or out of range - dữ liệu bị thiếu hoặc nằm ngoài phạm vi
7	Police Force	Chỉ ra lực lượng cảnh sát làm ở hiện trường tai nạn
8	Number of Vehicles	Chỉ ra số phương tiện trong vụ tai nạn
9	Number of Casualties	Chỉ ra số nạn nhân trong vụ tai nạn
10	Local_Authority_(Highway)	Chỉ ra thông tin của xa lộ.
11	Speed limit	Chỉ ra tốc độ giới hạn trên đường.

12	Light Conditions	Chỉ ra điều kiện ánh sáng trong lúc xảy ra tai nạn.
13	Weather Conditions	Chỉ ra điều kiện thời tiết trong lúc xảy ra tai nạn.
14	Road Surface Conditions	Chỉ ra điều kiện mặt đường trong lúc xảy ra tai nạn
15	LSOA	Chỉ ra vị trí tai nạn xảy ra

2.3. Bảng Vehicle

STT	Column	Description
1	Journey Purpose of Driver	<p>Chỉ ra mục đích của hành trình của người điều khiển phương tiện. và được chia thành các loại sau đây.</p> <ul style="list-style-type: none"> • 1: Journey as part of work • 2: Commuting to/from work - Đi làm đến / đi từ nơi làm việc • 3: Taking pupil to/from school - Đưa học sinh đến / từ trường • 4: Pupil riding to/from school - Cưỡi ngựa đến / từ trường • 5: Other - khác • 6: Not known - chưa biết

		<ul style="list-style-type: none"> • 15: Other/Not known (2005-10) - khác/ chưa biết • -1: Data missing or out of range - dữ liệu bị thiếu hoặc ngoài phạm vi
2	Vehicle Type	<p>Chỉ ra loại phương tiện sử dụng, và được chia thành các loại sau đây.</p> <ul style="list-style-type: none"> • 1: Pedal cycle - xe đạp • 2: Motorcycle 50cc and under - Xe mô tô 50cc trở xuống • 3: Motorcycle 125cc and under - Xe mô tô 125cc trở xuống • 4: Motorcycle over 125cc and up to 500cc - Xe mô tô trên 125cc đến 500cc • 5: Motorcycle over 500cc - Xe máy trên 500cc • 8: Taxi/Private hire car - Taxi / thuê xe riêng • 9: Car - Xe hơi • 10: Minibus (8 - 16 passenger seats) - Xe buýt nhỏ (8-16 chỗ ngồi) • 11: Bus or coach (17 or more pass seats) - Xe buýt hoặc xe khách (17 chỗ ngồi trở lên) • 16: Ridden horse - Ngựa hoang • 17: Agricultural vehicle - Xe công nông • 18: Tram - Xe điện

		<ul style="list-style-type: none"> • 19: Van / Goods 3.5 tonnes mgw or under - Van / Hàng hóa 3,5 tấn mgw trở xuống • 20: Goods over 3.5t. and under 7.5t - Hàng trên 3,5t. và dưới 7,5t • 21: Goods 7.5 tonnes mgw and over - Hàng hóa 7,5 tấn mgw trở lên • 22: Mobility scooter - Xe tay ga di động • 23: Electric motorcycle - Xe máy điện • 90: Other vehicle - Xe khác • 97: Motorcycle - unknown cc - Xe máy - phân khối không xác định • 98: Goods vehicle - unknown weight - Xe chở hàng - không rõ trọng lượng
3	Junction Location	<p>Chỉ ra vị trí giao lộ của phương tiện tai nạn, và được chia thành các loại sau đây</p> <ul style="list-style-type: none"> • 0: Not at or within 20 metres of junction - Không ở hoặc trong vòng 20 mét của đường giao nhau • 1:Approaching junction or waiting/parked at junction approach - Tiếp cận đường giao nhau hoặc chờ / đỗ ở đường giao nhau • 2:Cleared junction or waiting/parked at junction exit - Đường giao nhau đã thông thoáng

		<p>hoặc chờ / đỗ ở lối ra đường giao nhau</p> <ul style="list-style-type: none"> • 3: Leaving roundabout - Rời khỏi bùng binh • 4: Entering roundabout - Đi vào bùng binh • 5: Leaving main road - Ra khỏi đường chính • 6: Entering main road - Vào đường chính • 7: Entering from slip road - Đi vào từ đường trượt • 8: Mid Junction - on roundabout or on main road - Giao lộ Giữa - trên bùng binh hoặc trên đường chính • -1: Data missing or out of range - dữ liệu bị mất hoặc nằm ngoài phạm vi
4	Skidding Overturning and	<p>Chỉ ra phương tiện bị trượt hoặc lật, và được chia thành các loại sau đây.</p> <ul style="list-style-type: none"> • 0: None • 1: Skidded • 2: Skidded and overturned • 3: Jackknifed • 4: Jackknifed and overturned • 5: Overturned • -1: Data missing or out of range

5	Was Vehicle Left Hand Drive	<p>Chỉ ra người điều khiển phương tiện có lái xe bằng tay trái hay không.</p> <ul style="list-style-type: none"> • 1: No • 2: Yes • -1: Data missing or out of range - dữ liệu bị thiếu hoặc nằm ngoài phạm vi
6	Sex of Driver	<p>Chỉ ra giới tính của người điều khiển phương tiện.</p> <ul style="list-style-type: none"> • 1 : Male - Nam • 2 : Female - Nữ • 3:Not known - Chưa biết • -1: Data missing or out of range - dữ liệu bị thiếu hoặc nằm ngoài phạm vi
7	Age of Driver	Chỉ ra tuổi của người điều khiển phương tiện
8	Age Band of Driver	<p>Thuộc tính chỉ ra nhóm tuổi của nạn nhân, được lưu trữ theo quy luật code từ 1-11 (code – độ tuổi).</p> <ul style="list-style-type: none"> • 1: 0 - 5 • 2: 6 - 10 • 3: 11 – 15 • 4: 16 – 20

		<ul style="list-style-type: none"> • 5: 21 – 25 • 6: 26 – 35 • 7: 36 – 45 • 8: 46 – 55 • 9: 56 – 65 • 10: 66 – 75 • 11: Over 75 • -1: Data missing or out range
9	Age of Vehicle (manufacture)	Chỉ ra số tuổi của phương tiện.

2.4. Bảng Postcodes (UK Postcodes)

STT	Column	Description
1	city	Thị trấn(town) hoặc thành phố (city)
2	county	Quận(the county) hoặc khu vực đô thị (city conurbation) hoặc khu vực thành phố (municiple region).
3	country_code	Mã quốc gia gồm 3 chữ cái
4	country_name	Tên của quốc gia

5	region_code	Mã vùng (không chính thức)
6	region_name	Tên vùng
7	postcode	Mã bưu điện
8	easting	Giá trị hướng đông - tối đa 7 chữ số
9	northing	Giá trị hướng bắc - tối đa 7 chữ số
10	latitude	Góc vĩ độ tính bằng độ
11	longitude	Góc kinh độ tính bằng độ
12	iso3166-2	Mã quốc gia ISO

2.5. Bảng PCD_OA_LSOA_MSOA_LAD_AUG21_UK_LU

	Column	Description
1	pcd7	Mã postcode 7 ký tự
2	pcd8	Mã postcode 8 ký tự
3	pcds	Khoảng cách giữa quận và đơn vị ngành

4	dointr	Ngày postcode có hiệu lực
5	doterm	Ngày postcode hết hiệu lực
6	usertype	Loại người dùng
7	oa11cd	Mã Output Area dân số trung bình khu vực vào khoảng 310 người
8	lsoa11cd	Mã Lower layer super output area dân số vào khoảng 1500
9	msoa11cd	Mã Middle layer super output area dân số vào khoảng 7500
10	ladcd	Mã vùng đất
11	lsoa11nm	Tên của Lower layer super output area
12	msoa11nm	Tên của Middle layer super output area
13	ladnm	Tên vùng đất

3. Dữ liệu liên quan yêu cầu phân tích

3.1 Report the number of casualties by Severity (Fatal, Serious, Slight) in the Local Authority Districts over years

Các kiểu dữ liệu liên quan:

- Number of Casualties (Casualty table)
- Casualty Severity (Casualty table)
- Local Authority Districts (Accident table)
- Year (Accident table, column Date)

3.2 Report the number of casualties by Severity (Fatal, Serious, Slight) in the Local Authority Districts by Quarters in years.

Các kiểu dữ liệu liên quan:

- Number of Casualties (Casualty table)
- Casualty Severity (Casualty table)
- Local Authority Districts (Accident table)
- Quarters, Years (Accident table, column Date) - Dữ liệu chưa có, tự tạo thêm

3.3 Report the number of fatal casualties by Gender, Casualty Type and Age Band over years.

Các kiểu dữ liệu liên quan:

- Number of Casualties (Casualty table)
- Casualty Severity (Casualty table, Column Casualty Severity, value Fatal) - Dữ liệu bị thiếu sót
- Sex of Casualty (Casualty table)
- Casualty Type (Casualty table)
- Age_Band_of_Casualty (Casualty table)
- Years (Accident table, column Date) - Dữ liệu chưa có, tự tạo thêm

3.4 Report the number of accidents by Severity and Time of Day (Morning: 5am- 12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) over years.

Các kiểu dữ liệu liên quan:

- Number of Accidents(Accident table)
- Accident Severity (Accident table)
- Time of Day (Accident table) - Dữ liệu chưa có, tự tạo thêm từ cột Time
- Years (Accident table, column Date) - Dữ liệu chưa có, tự tạo thêm

3.5 Report the number of accidents by Severity, Urban or Rural Area and Road Type over years.

Các kiểu dữ liệu liên quan:

- Number of Accidents(Accident table)
- Accident Severity (Accident table)
- Urban_or_Rural_Area (Accident table)
- Road_Type (Accident table)
- Years (Accident table, column Date) - Dữ liệu chưa có, tự tạo thêm

3.6 Report the number of casualties by Severity, Casualty Type and Age Group over years, Age Group is defined as below:

Children: 0-15

Young adult: 0-17

Adult: 18-59

60 and over: 60-...

Các kiểu dữ liệu liên quan:

- Number of Casualties (Casualty table)
- Casualty Severity (Casualty table)
- Casualty Type (Casualty table)
- Age Group (Casualty table) - Dữ liệu chưa có, tự tạo thêm

3.7 Report the number of accidents by Journey Purpose and Vehicle Type.

Các kiểu dữ liệu liên quan:

- Number of Accidents(Accident table)
- Journey_Purpose_of_Driver (Vehicle table)
- Vehicle Type (Vehicle table)

3.8 Create a new attribute Built-up Road in Accidents table. Built-up Road may have 2 values:

Built-up road: if Speed Limit below 50 mph

Non Built-up road: if Speed Limit above 50 mph

Các kiểu dữ liệu liên quan:

- Built-up road (Accident table) - Dữ liệu chưa có, tự tạo thêm

3.9 Report the number of accidents by Severity, Vehicle Type, Built-up Road over years.

Các kiểu dữ liệu liên quan:

- Number of Accidents(Accident table)
- Accident Severity (Accident table)
- Vehicle Type (Vehicle table)
- Built-up road (Accident table)

3.10 Students design other reports about UK car accident.

Các kiểu dữ liệu liên quan:

- Number of Accidents
- Accident Severity
- Vehicle Type
- Year, Quater, Month, Session
- Urban or Rural Area
- Sex Of Driver
- Road Type

3.11 Use regional map to visually represent (by color) the number of car accidents in regions during a year.

Các kiểu dữ liệu liên quan:

- Number of Accidents(Accident table)
- Region
- Country

4. Phân tích profiling data

Tong quan bo data...4J6IM9Fjoker (54)

```

SELECT
s.Name AS Schema_Name,
t.Name AS Table_Name,
t.object_id,
t.type_desc,
p.rows AS Row_Count,
CAST(ROUND((SUM(a.used_pages) / 128.00), 2) AS NUMERIC(36, 2)) AS Used_MB,
CAST(ROUND((SUM(a.total_pages) - SUM(a.used_pages)) / 128.00, 2) AS NUMERIC(36, 2)) AS Unused_MB,
CAST(ROUND((SUM(a.total_pages) / 128.00), 2) AS NUMERIC(36, 2)) AS Total_MB
FROM sys.tables t
INNER JOIN sys.indexes i ON t.OBJECT_ID = i.object_id
INNER JOIN sys.partitions p ON i.object_id = p.OBJECT_ID AND i.index_id = p.index_id
INNER JOIN sys.allocation_units a ON p.partition_id = a.container_id
INNER JOIN sys.schemas s ON t.schema_id = s.schema_id
where p.index_id IN ( 0, 1 )
GROUP BY t.Name, t.object_id,t.type_desc,s.Name, p.Rows
  
```

100 %

Results Messages

	Schema_Name	Table_Name	object_id	type_desc	Row_Count	Used_MB	Unused_MB	Total_MB
1	dbo	Accidents1114	581577110	USER_TABLE	582027	103.45	1.43	104.88
2	dbo	Vehicles1114	597577167	USER_TABLE	1048575	99.41	3.29	102.70
3	dbo	Casualties1114	613577224	USER_TABLE	777820	52.28	5.98	58.26

File Accidents1114 có 582027 dòng

File Vehicles1114 có 1048575 dòng

File Casualties1114 có 777820 dòng

Profiling từng bảng...46IM9F\joker (54)

100 %

Results Messages

	TABLE_SCHEMA	TABLE_NAME	COLUMN_NAME	ColumnDataLength	Data Type	MinDataLength	MaxDataLength	AvgDataLength	MinDate	MaxDate	NoDistinct	NoNulls	NoZeroLength	PercentageNulls	PercentageZeroLength	NoDateWithHourminuteSecond
1	[dbo]	[Accidents1114]	[Accident_Index]	50	varchar	8	13	12	NULL	NULL	359457	0.0000	0.0000	0.0000	0.0000	NULL
2	[dbo]	[Accidents1114]	[Location_Easting_OSGR]	50	varchar	5	6	5	NULL	NULL	176334	0.0000	0.0000	0.0000	0.0000	NULL
3	[dbo]	[Accidents1114]	[Location_Northing_OSGR]	50	varchar	5	7	5	NULL	NULL	202366	0.0000	0.0000	0.0000	0.0000	NULL
4	[dbo]	[Accidents1114]	[Longitude]	50	varchar	4	9	8	NULL	NULL	501542	0.0000	0.0000	0.0000	0.0000	NULL
5	[dbo]	[Accidents1114]	[Latitude]	50	varchar	2	9	8	NULL	NULL	486247	0.0000	0.0000	0.0000	0.0000	NULL
6	[dbo]	[Accidents1114]	[Police_Force]	50	varchar	1	2	1	NULL	NULL	51	0.0000	0.0000	0.0000	0.0000	NULL
7	[dbo]	[Accidents1114]	[Accident_Severity]	50	varchar	1	1	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
8	[dbo]	[Accidents1114]	[Number_of_Vehicles]	50	varchar	1	2	1	NULL	NULL	20	0.0000	0.0000	0.0000	0.0000	NULL
9	[dbo]	[Accidents1114]	[Number_of_Casualties]	50	varchar	1	2	1	NULL	NULL	39	0.0000	0.0000	0.0000	0.0000	NULL
10	[dbo]	[Accidents1114]	[Date]	50	varchar	8	10	9	NULL	NULL	1461	0.0000	0.0000	0.0000	0.0000	NULL
11	[dbo]	[Accidents1114]	[Day_of_Week]	50	varchar	1	1	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
12	[dbo]	[Accidents1114]	[Time]	50	varchar	0	5	4	NULL	NULL	1440	0.0000	25.0000	0.0000	0.0043	NULL
13	[dbo]	[Accidents1114]	[Local_Authority_(District)]	50	varchar	1	3	2	NULL	NULL	380	0.0000	0.0000	0.0000	0.0000	NULL
14	[dbo]	[Accidents1114]	[Local_Authority_(Highway)]	50	varchar	9	9	9	NULL	NULL	207	0.0000	0.0000	0.0000	0.0000	NULL
15	[dbo]	[Accidents1114]	[1st_Road_Class]	50	varchar	1	1	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
16	[dbo]	[Accidents1114]	[1st_Road_Number]	50	varchar	1	4	2	NULL	NULL	5639	0.0000	0.0000	0.0000	0.0000	NULL
17	[dbo]	[Accidents1114]	[Road_Type]	50	varchar	1	1	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
18	[dbo]	[Accidents1114]	[Speed_Limit]	50	varchar	2	2	2	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
19	[dbo]	[Accidents1114]	[Junction_Detail]	50	varchar	1	2	1	NULL	NULL	10	0.0000	0.0000	0.0000	0.0000	NULL
20	[dbo]	[Accidents1114]	[Junction_Control]	50	varchar	1	2	1	NULL	NULL	5	0.0000	0.0000	0.0000	0.0000	NULL
21	[dbo]	[Accidents1114]	[2nd_Road_Class]	50	varchar	1	2	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
22	[dbo]	[Accidents1114]	[2nd_Road_Number]	50	varchar	1	4	1	NULL	NULL	5535	0.0000	0.0000	0.0000	0.0000	NULL
23	[dbo]	[Accidents1114]	[Pedestrian_Crossing-Human_Control]	50	varchar	1	1	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
24	[dbo]	[Accidents1114]	[Pedestrian_Crossing-Physical_Facilities]	50	varchar	1	1	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
25	[dbo]	[Accidents1114]	[Light_Conditions]	50	varchar	1	1	1	NULL	NULL	5	0.0000	0.0000	0.0000	0.0000	NULL
26	[dbo]	[Accidents1114]	[Weather_Conditions]	50	varchar	1	1	1	NULL	NULL	9	0.0000	0.0000	0.0000	0.0000	NULL
27	[dbo]	[Accidents1114]	[Road_Surface_Conditions]	50	varchar	1	2	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
28	[dbo]	[Accidents1114]	[Special_Conditions_at_Site]	50	varchar	1	2	1	NULL	NULL	9	0.0000	0.0000	0.0000	0.0000	NULL
29	[dbo]	[Accidents1114]	[Carriageway_Hazards]	50	varchar	1	2	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
30	[dbo]	[Accidents1114]	[Urban_or_Rural_Area]	50	varchar	1	1	1	NULL	NULL	2	0.0000	0.0000	0.0000	0.0000	NULL
31	[dbo]	[Accidents1114]	[Dist_Police_Officer_Atend_Scene_of_...]	50	varchar	1	2	1	NULL	NULL	4	0.0000	0.0000	0.0000	0.0000	NULL
32	[dbo]	[Accidents1114]	[LSOA_of_Accident_Location]	50	varchar	0	9	8	NULL	NULL	34866	0.0000	38891.0000	0.0000	6.6476	NULL

Query executed successfully.

DESKTOP-46IM9P\SVO (15.0 RTM) DESKTOP-46IM9P\joker... UK_Accidents 00:00:23 32 rows

Data Accidents1114 gồm 32 cột tất cả cột đều không bị missing value, ngoài ra còn có thông tin về min, max từng cột, độ dài cột, các giá trị duy nhất, ...

Profiling từng bảng...46IM9F\joker (54)*

100 %

Results Messages

	TABLE_SCHEMA	TABLE_NAME	COLUMN_NAME	ColumnDataLength	Data Type	MinDataLength	MaxDataLength	AvgDataLength	MinDate	MaxDate	NoDistinct	NoNulls	NoZeroLength	PercentageNulls	PercentageZeroLength	NoDateWithHourminuteSecond
1	[dbo]	[Vehicles1114]	[Accident_Index]	50	varchar	8	13	12	NULL	NULL	355421	0.0000	0.0000	0.0000	0.0000	NULL
2	[dbo]	[Vehicles1114]	[Vehicle_Reference]	50	varchar	1	2	1	NULL	NULL	68	0.0000	0.0000	0.0000	0.0000	NULL
3	[dbo]	[Vehicles1114]	[Vehicle_Type]	50	varchar	1	2	1	NULL	NULL	21	0.0000	0.0000	0.0000	0.0000	NULL
4	[dbo]	[Vehicles1114]	[Towing_and_Articulation]	50	varchar	1	2	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
5	[dbo]	[Vehicles1114]	[Vehicle_Manoeuvre]	50	varchar	1	2	1	NULL	NULL	19	0.0000	0.0000	0.0000	0.0000	NULL
6	[dbo]	[Vehicles1114]	[Vehicle_Location-Restricted_Lane]	50	varchar	1	2	1	NULL	NULL	11	0.0000	0.0000	0.0000	0.0000	NULL
7	[dbo]	[Vehicles1114]	[Junction_Location]	50	varchar	1	2	1	NULL	NULL	10	0.0000	0.0000	0.0000	0.0000	NULL
8	[dbo]	[Vehicles1114]	[Skidding_and_Overtaking]	50	varchar	1	2	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
9	[dbo]	[Vehicles1114]	[Hit_Object_in_Carriageway]	50	varchar	1	2	1	NULL	NULL	13	0.0000	0.0000	0.0000	0.0000	NULL
10	[dbo]	[Vehicles1114]	[Vehicle_Leaving_Carriageway]	50	varchar	1	2	1	NULL	NULL	10	0.0000	0.0000	0.0000	0.0000	NULL
11	[dbo]	[Vehicles1114]	[Hit_Object_of_Carriageway]	50	varchar	1	2	1	NULL	NULL	13	0.0000	0.0000	0.0000	0.0000	NULL
12	[dbo]	[Vehicles1114]	[1st_Point_of_Impact]	50	varchar	1	2	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
13	[dbo]	[Vehicles1114]	[Was_Vehicle_Left_Hand_Drive?]	50	varchar	1	2	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
14	[dbo]	[Vehicles1114]	[Journey_Purpose_of_Driver]	50	varchar	1	2	1	NULL	NULL	7	0.0000	0.0000	0.0000	0.0000	NULL
15	[dbo]	[Vehicles1114]	[Sex_of_Driver]	50	varchar	1	1	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
16	[dbo]	[Vehicles1114]	[Age_of_Driver]	50	varchar	1	3	1	NULL	NULL	101	0.0000	0.0000	0.0000	0.0000	NULL
17	[dbo]	[Vehicles1114]	[Age_Band_of_Driver]	50	varchar	1	2	1	NULL	NULL	12	0.0000	0.0000	0.0000	0.0000	NULL
18	[dbo]	[Vehicles1114]	[Engine_Capacity_[CC]]	50	varchar	1	5	3	NULL	NULL	1789	0.0000	0.0000	0.0000	0.0000	NULL
19	[dbo]	[Vehicles1114]	[Propulsion_Code]	50	varchar	1	2	1	NULL	NULL	13	0.0000	0.0000	0.0000	0.0000	NULL
20	[dbo]	[Vehicles1114]	[Age_of_Vehicle]	50	varchar	1	3	1	NULL	NULL	88	0.0000	0.0000	0.0000	0.0000	NULL
21	[dbo]	[Vehicles1114]	[Driver_IMD_Decile]	50	varchar	1	2	1	NULL	NULL	11	0.0000	0.0000	0.0000	0.0000	NULL
22	[dbo]	[Vehicles1114]	[Driver_Home_Area_Type]	50	varchar	1	2	1	NULL	NULL	4	0.0000	0.0000	0.0000	0.0000	NULL

Data Vehicles1114 gồm 22 cột tất cả cột đều không bị missing value, ngoài ra còn có thông tin về min, max từng cột, độ dài cột, các giá trị duy nhất, ...

Profiling từng bảng...46IM9F\joker (54)*

100 %

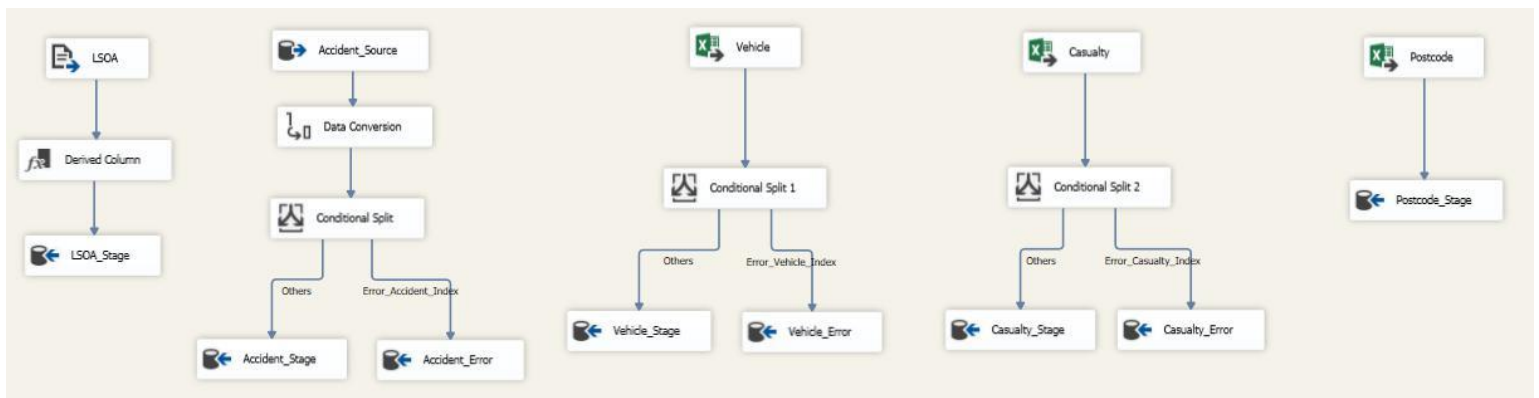
Results Messages

	TABLE_SCHEMA	TABLE_NAME	COLUMN_NAME	ColumnDataLength	DataType	MinDataLength	MaxDataLength	AvgDataLength	MinDate	MaxDate	NoDistinct	NoNulls	NoZeroLength	PercentageNulls	PercentageZeroLength	NoDateWithHourMinuteS
1	[dbo]	[Casualties1114]	[Accident_Index]	50	varchar	8	13	11	NULL	NULL	399339	0.0000	0.0000	0.0000	0.0000	NULL
2	[dbo]	[Casualties1114]	[Vehicle_Reference]	50	varchar	1	2	1	NULL	NULL	67	0.0000	0.0000	0.0000	0.0000	NULL
3	[dbo]	[Casualties1114]	[Casualty_Reference]	50	varchar	1	3	1	NULL	NULL	97	0.0000	0.0000	0.0000	0.0000	NULL
4	[dbo]	[Casualties1114]	[Casualty_Class]	50	varchar	1	1	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
5	[dbo]	[Casualties1114]	[Sex_of_Casualty]	50	varchar	1	2	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
6	[dbo]	[Casualties1114]	[Age_of_Casualty]	50	varchar	1	3	1	NULL	NULL	105	0.0000	0.0000	0.0000	0.0000	NULL
7	[dbo]	[Casualties1114]	[Age_Band_of_Casualty]	50	varchar	1	2	1	NULL	NULL	12	0.0000	0.0000	0.0000	0.0000	NULL
8	[dbo]	[Casualties1114]	[Casualty_Severity]	50	varchar	1	1	1	NULL	NULL	3	0.0000	0.0000	0.0000	0.0000	NULL
9	[dbo]	[Casualties1114]	[Pedestrian_Location]	50	varchar	1	2	1	NULL	NULL	12	0.0000	0.0000	0.0000	0.0000	NULL
10	[dbo]	[Casualties1114]	[Pedestrian_Movement]	50	varchar	1	2	1	NULL	NULL	11	0.0000	0.0000	0.0000	0.0000	NULL
11	[dbo]	[Casualties1114]	[Car_Passenger]	50	varchar	1	2	1	NULL	NULL	4	0.0000	0.0000	0.0000	0.0000	NULL
12	[dbo]	[Casualties1114]	[Bus_or_Coach_Passenger]	50	varchar	1	2	1	NULL	NULL	6	0.0000	0.0000	0.0000	0.0000	NULL
13	[dbo]	[Casualties1114]	[Pedestrian_Road_Maintenance_Worker]	50	varchar	1	2	1	NULL	NULL	4	0.0000	0.0000	0.0000	0.0000	NULL
14	[dbo]	[Casualties1114]	[Casualty_Type]	50	varchar	1	2	1	NULL	NULL	21	0.0000	0.0000	0.0000	0.0000	NULL
15	[dbo]	[Casualties1114]	[Casualty_Home_Area_Type]	50	varchar	1	2	1	NULL	NULL	4	0.0000	0.0000	0.0000	0.0000	NULL

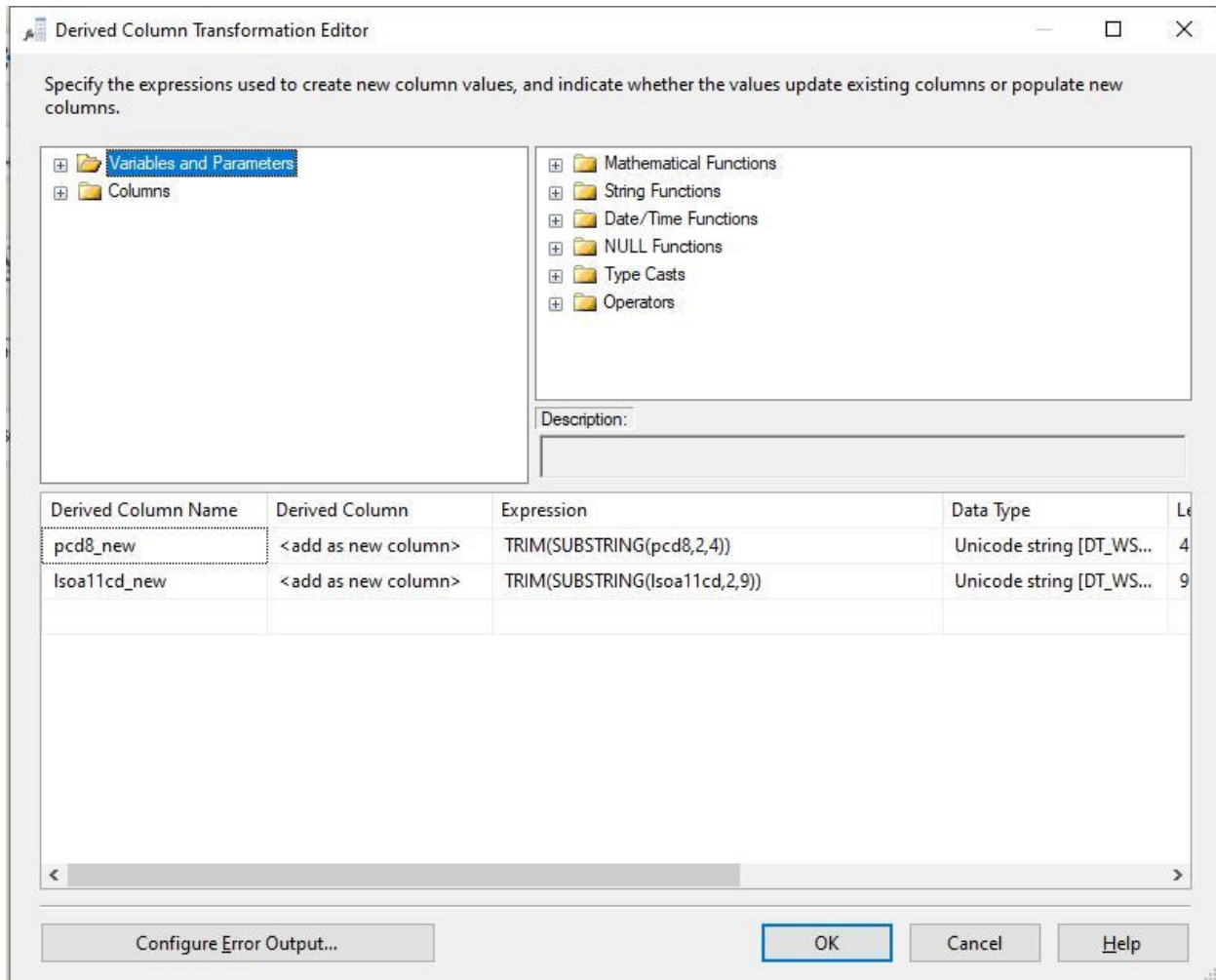
Data Vehicles1114 gồm 15 cột tất cả cột đều không bị missing value, ngoài ra còn có thông tin về min, max từng cột, độ dài cột, các giá trị duy nhất, ...

Chương 2. Quá trình SSIS

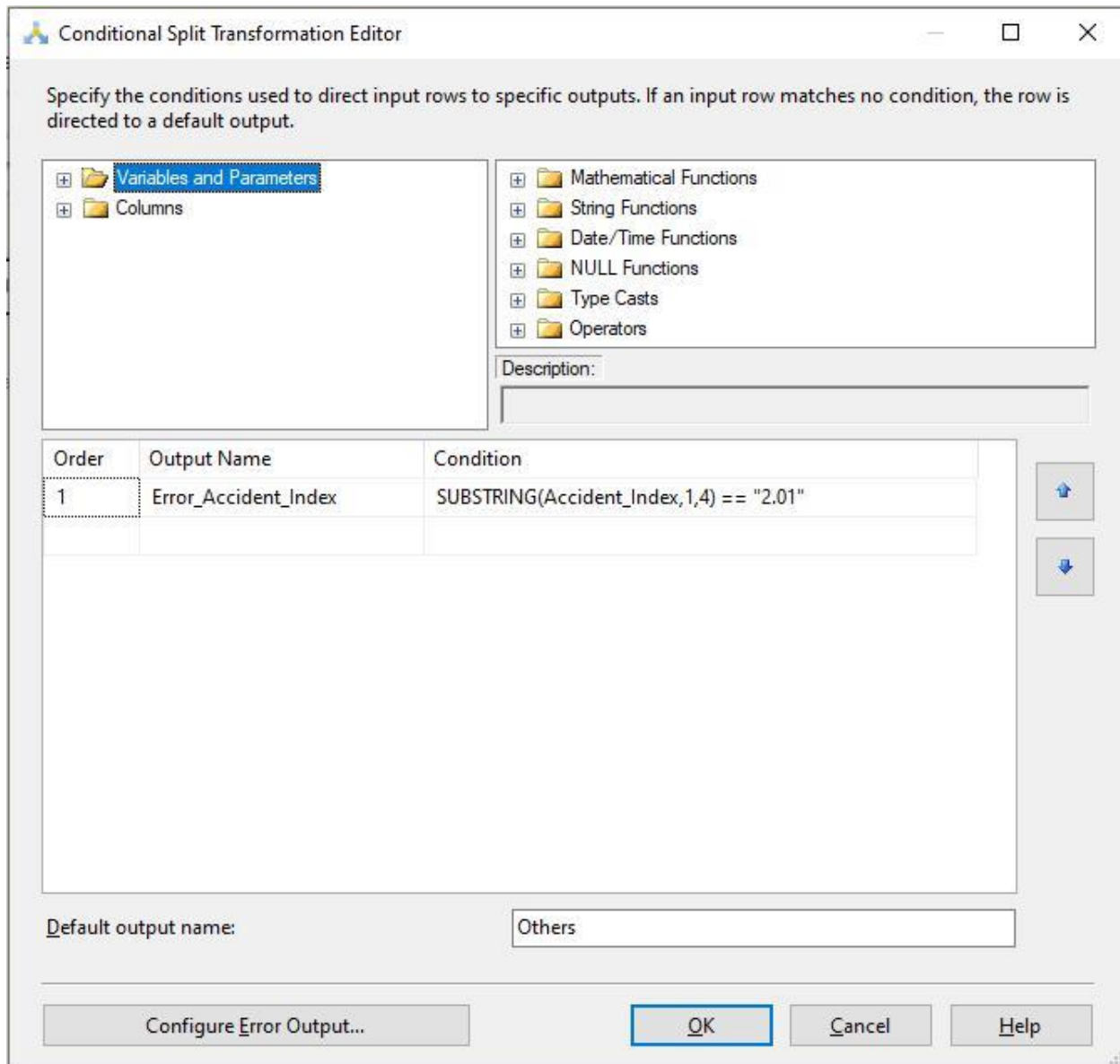
1. Source vào Stage



- Ở bảng LSOA khi đổ từ Source vào Stage chúng em dùng derive column để xử lý 2 cột Isoa11cd và pcd8 vì khi đổ LSOA bằng flat file data xuất hiện ký tự " " ở 2 đầu dữ liệu. Vì vậy đã chúng em dùng Derived Column để làm sạch dữ liệu
 - + Cột pcd8: Dùng substring từ vị trí thứ 2 để tránh dấu " và lấy đi 4 ký tự đầu của chuỗi để so khớp với data từ bảng Postcode và dùng trim để xóa khoảng cách do các chuỗi postcode có độ dài không cố định từ 2-4 ký tự.
 - + Cột Isoa11cd: Tương tự xử lý 9 ký tự của LSOA.



- Ở 3 bảng tiếp theo gồm Accident, Casualty và Vehicle thì chúng em dùng conditional split để loại bỏ các cột có thuộc tính Accident_Index bị lỗi định dạng từ source và đưa vào database riêng có tên Error. Dấu hiệu nhận biết là 4 ký tự đầu có dạng "2.01"



Quá trình conditional split

- Đặc biệt riêng bảng Accident chúng em không dùng file excel để làm source dữ liệu do data bị lỗi cột Date. Vì vậy team đã chọn cách import file excel vào thẳng sql server và dùng data từ đó để đổ vào Stage.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

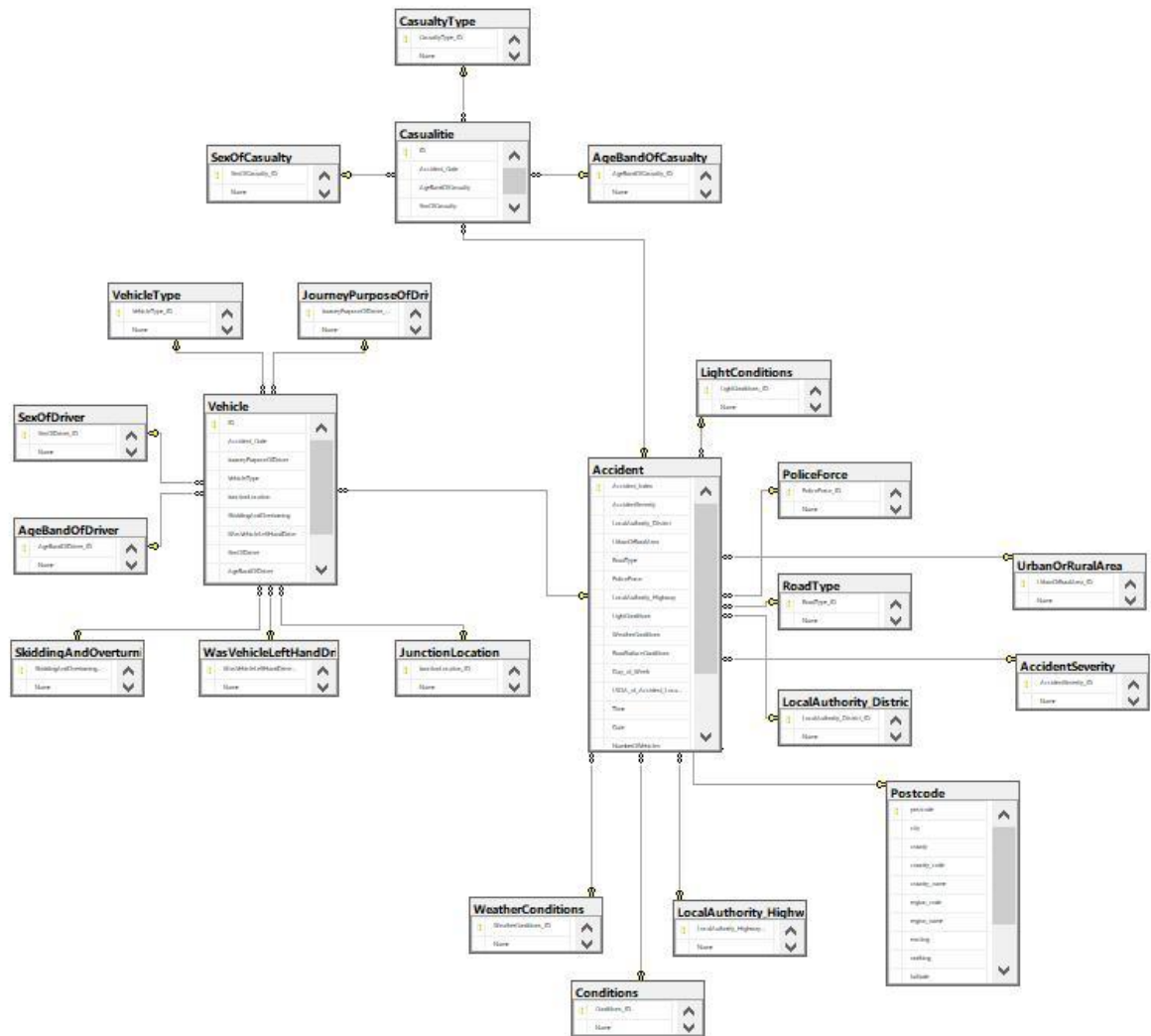
Available Input Columns

- ☒ Name
- ☒ Accident_Index
- ☒ Police_Force
- ☒ Accident_Severity
- ☒ Number_of_Vehicles
- ☒ Number_of_Casualties

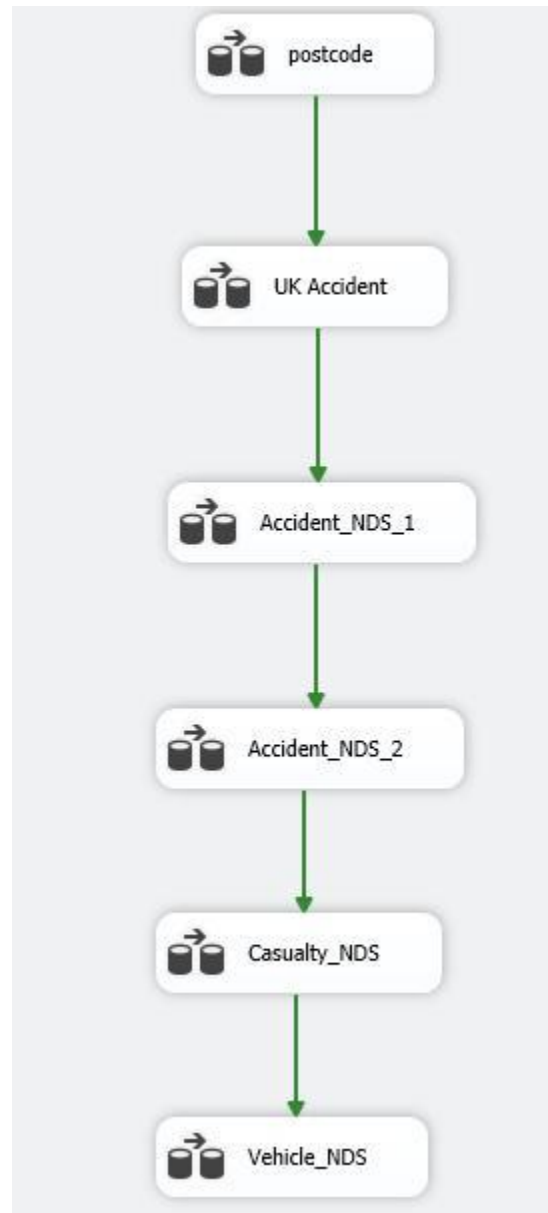
Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
Accident_Index	Copy of Accident_Index	Unicode string [DT_WSTR]	50			
Police_Force	Copy of Police_Force	Unicode string [DT_WSTR]	50			
Accident_Severity	Copy of Accident_Sev...	Unicode string [DT_WSTR]	50			
Number_of_Vehicles	Copy of Number_of_V...	Unicode string [DT_WSTR]	50			
Number_of_Casualties	Copy of Number_of_C...	Unicode string [DT_WSTR]	50			
Date	Copy of Date	Unicode string [DT_WSTR]	50			
Day_of_Week	Copy of Day_of_Week	Unicode string [DT_WSTR]	50			
Time	Copy of Time	Unicode string [DT_WSTR]	50			
Local_Authority_(High...	Copy of Local_Authori...	Unicode string [DT_WSTR]	50			

Configure Error Output... OK Cancel Help

2. Stage vào NDS



Thiết kế diagram dành cho NDS



Control flow của NDS

- 2 Data flow đầu tiên là để đổ data từ 2 file postcode và UK Accident dùng để tham chiếu khóa ngoại khi đổ các bảng chính phía sau.
- Accident_NDS:
 - + Kiểm tra hợp lệ các khóa ngoại của bảng Accident

- + Gộp bảng Accident và LSOA thông qua khóa ngoại LSOA để lấy ra cột pcd8 trong bảng LSOA
- + Tiếp tục dùng cột pcd8 để kiểm tra khóa ngoại postcode trong bảng Postcode. Sau đó chia dữ liệu về 2 bảng gồm những dòng có pcd8 hợp lệ và không hợp lệ
- + Với những dòng pcd8 không tồn tại thì sẽ chuyển về null và gộp lại với những dòng hợp lệ
- + Sau khi thỏa hết điều kiện thì sẽ được đổ vào NDS
- Casualty_NDS và Vehicle_NDS:
 - + Chỉ đơn giản kiểm tra hợp lệ của các khóa ngoại trước khi đổ vào NDS

3. NDS vào DDS

3.1. Lướt đồ sao DDS

- **Nhu cầu phân tích**

1 Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng (Fatal, Serious, Slight) ở các Địa phương (Local_Authority_(District)) trong tất cả các năm. Bảng

2 Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng ở các Địa Phương (Local_Authority_(District)) theo các Quý trong từng năm.

3 Thống kê số lượng người tử vong theo Giới Tính, Loại Nạn Nhân (Casualty Type) và Nhóm Tuổi (Age_Band_of_Casualty) theo các năm.

4 Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng và Thời Điểm Trong Ngày (Morning: 5am-12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) trong các năm.

5 Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng, Vùng (Urban_or_Rural_Area), và Kiểu Đường (Road Type) trong các năm.

6 Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng, Loại Nạn Nhân (Casualty Type) và Độ Tuổi trong các năm, Độ Tuổi được định nghĩa như sau:

Children: 0-15

Young adult: 0-17

Adult: 18-59

60 and over: 60-...

7 Tổng hợp số lượng tai nạn theo Mục Đích Hành Trình (Journey Purpose) và Loại Phương Tiện (Vehicle_Type).

8 Tạo thêm thuộc tính Built-up Road trong table Accidents. Built-up Road có 2

giá trị: Built-up road: Nếu tốc độ giới hạn (Speed Limit) dưới 50 mph Non Built-up road: Nếu tốc độ giới hạn từ 50 mph

9 Thống kê số lượng tai nạn theo Mức Độ Nghiêm Trọng, Loại Phương Tiện (Vehicle Type), Built-up Road trong các năm.

10 Thống kê số lượng tai nạn theo Mức Độ Nghiêm Trọng, Loại Phương Tiện (Vehicle Type), Built-up Road trong các năm.

11 Định nghĩa fact Variance để tính mức độ tăng giảm của TNGT theo đơn vị phần trăm qua các năm.

12 Xây dựng đồ thị/ biểu đồ cho các bảng thống kê ở trên.

13 Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng TNGT ở các vùng trong năm.

• **Mô hình hóa thiết kế chiều**

Các chiều liên quan sự kiện phân tích

- DimAccidentServity
- DimCasualtyType
- DimAgeBandOfCasualty
- DimAgeBandOfDriver
- DimDate
- DimGeography
- DimJourneyPurposeOfDriver
- DimJunctionLocation
- DimLocalAuthority_District

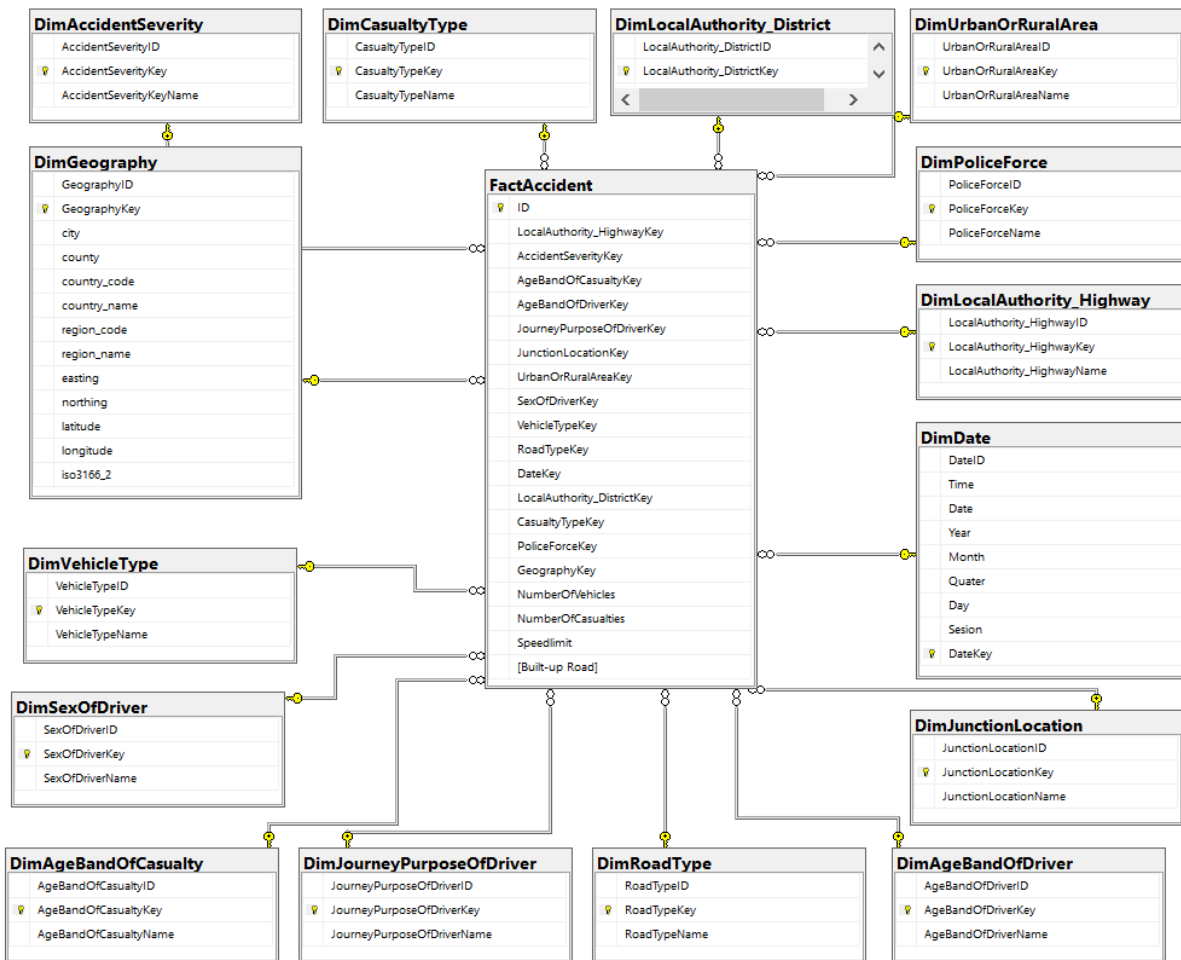
- DimLocalAuthority_Highway
- DimPoliceForce
- DimRoadType
- DimSexOfDriver
- DimUrbanOrRuralArea
- DimVehicleType
- FactAccident

- **Mô hình hóa thiết kế Fact**

Các giá trị có sẵn: NumberOfVehicles, NumberOfCasualties, Speedlimit

Các giá trị tính toán thêm: [Built-up Road]

- **Lượt đồ sao**



3.2. Tạo DataBase và đổ bảng vào

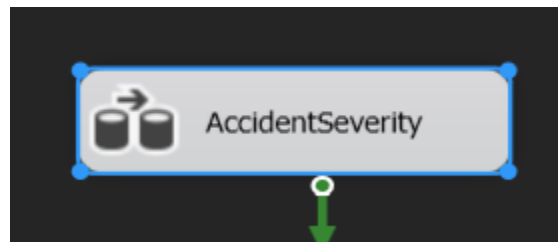

```
testco1thussql - DE...-4J6IM9F\joker (59)  DESKTOP-4J6IM9F\...DATH - Diagram_0*
USE [DDS_DATH]
GO
/***** Object: Database [DDS_DATH]    Script Date: 1/4/2022 9:40:29 AM *****/

/***** Object: Table [dbo].[DimAccidentSeverity]    Script Date: 1/4/2022 9:40:29 AM *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DimAccidentSeverity](
    [AccidentSeverityID] [int] IDENTITY(1,1) NOT NULL,
    [AccidentSeverityKey] [nvarchar](255) NOT NULL,
    [AccidentSeverityKeyName] [nvarchar](255) NULL,
    CONSTRAINT [PK_AS] PRIMARY KEY CLUSTERED
    (
        [AccidentSeverityKey] ASC
    )WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) GO
/***** Object: Table [dbo].[DimAgeBandOfCasualty]    Script Date: 1/4/2022 9:40:29 AM *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DimAgeBandOfCasualty](
    [AgeBandOfCasualtyID] [int] IDENTITY(1,1) NOT NULL,
    [AgeBandOfCasualtyKey] [nvarchar](255) NOT NULL,
    [AgeBandOfCasualtyName] [nvarchar](255) NULL,
    CONSTRAINT [PK_ABOC] PRIMARY KEY CLUSTERED
    (
        [AgeBandOfCasualtyKey] ASC
    )WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) GO
/***** Object: Table [dbo].[DimAgeBandOfDriver]    Script Date: 1/4/2022 9:40:29 AM *****/
SET ANSI_NULLS ON
```

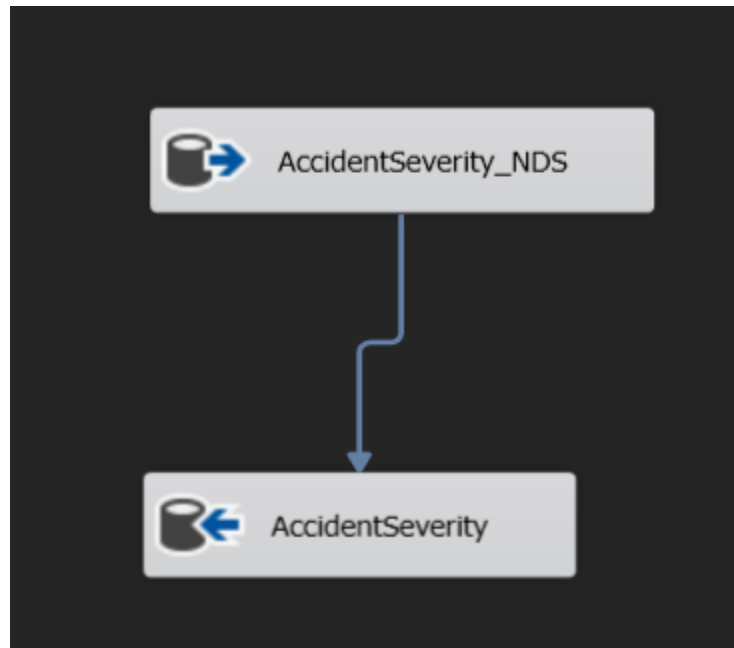
3.3. Tạo bảng Dim

- **DimAccidentSeverity**

Kéo thả Data Flow Task. Sau đó đổi tên Data Flow thành AccidentSeverity.



Chọn Flat File Source, OLE DB Destination. Đổi tên Flat File Source thành AccidentSeverity_NDS. - Đổi tên OLE DB Destination thành AccidentSeverity.



Nhấp chuột vào AccidentSeverity_NDS chọn DB Server và Table sau đó nhấn OK.

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
DESKTOP-4J6IM9F\SYO.NDS_DATH New...

Data access mode:
Table or view

Name of the table or the view:
[dbo].[AccidentSeverity]

Preview...

OK Cancel Help

Nhấp chuột vào AccidentSeverity chọn DB Server và Table

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:
DESKTOP-4J6IM9F\SYO.DDS_DATH New...

Data access mode:
Table or view - fast load

Name of the table or the view:
[dbo].[DimAccidentSeverity] New...

☐ Keep identity ☒ Table lock
☐ Keep nulls ☒ Check constraints

Rows per batch:

Maximum insert commit size:

View Existing

OK Cancel Help

Sang mục Mapping tiến hành mapping nguồn dữ liệu NDS với bảng Dim sau đó chọn OK.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Available Input Columns

Name
AccidentSeverity_ID
Name

Available Destination Columns

Name
AccidentSeverityKey
AccidentSeverityKeyName
AccidentSeverityID

Input Column	Destination Column
AccidentSeverity_ID	AccidentSeverityKey
Name	AccidentSeverityKeyName
<ignore>	AccidentSeverityID

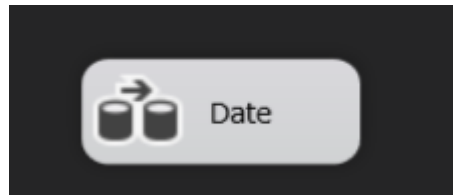
OK Cancel Help

- Tương tự cho các bảng còn lại: DimAgeBandOfCasualty, DimAgeBandOfDriver, DimCasualtyType, DimGeography, DimJourneyPurposeOfDriver, DimJunctionLocation, DimLocalAuthority_District, DimLocalAuthority_Highway,

DimPoliceForce, DimRoadType, DimSexOfDriver, DimUrbanOrRuralArea,
DimVehicleType.

- **DimDate**

Kéo thả Data Flow Task. Sau đó đổi tên Data Flow thành Date.



Chọn Flat File Source, Data Conversion, Derived Column, Sort, OLE DB Destination. Đổi tên Flat File Source thành Accident_NDS. - Đổi tên OLE DB Destination thành Date.

Nhấp chuột vào Accident_NDS chọn DB Server và Table

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
DESKTOP-4J6IM9F\SYO.NDS_DATH New...

Data access mode:
Table or view

Name of the table or the view:
[dbo].[Accident]

Preview...

OK Cancel Help

Tại tab Column chọn các cột cần sử dụng sau đó nhấn OK.

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager

Columns

Error Output

Available External Columns

- ☐ Name
- ☒ Accident_Index
- ☐ AccidentSeverity
- ☐ LocalAuthority_District
- ☐ UrbanOrRuralArea
- ☐ RoadType
- ☐ PoliceForce
- ☐ LocalAuthority_Highway
- ☐ LightConditions
- ☐ WeatherConditions

External Column	Output Column
Date	Date
Time	Time
Day_of_Week	Day_of_Week
Accident_Index	Accident_Index

OK Cancel Help

Mở Data Conversion để tiến hành biến đổi lại dữ liệu, chuyển đổi Date từ dạng nvarchar sang kiểu date và lưu là ConvertDate, chuyển Date từ dạng nvarchar sang dạng varchar và lưu là DateKey.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

- ☐ Name
- ☒ Date
- ☐ Time
- ☐ Day_of_Week
- ☐ Accident_Index

Input Column	Output Alias	Data Type	Length	Precision	Scale
Date	ConvertDate	database date [DT_DBDATE]			
Date	DateKey	string [DT_STR]	255		

Configure Error Output... OK Cancel Help

Mở Derived Column để tiến hành tính toán một số giá trị, lấy ra Year, Quarter, Month, Day từ ConvertDate đã chuyển ở bước Data Conversion bằng các hàm Date/Time Functions. Lấy ra Session dùng để lưu các buổi trong một ngày.

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters
Columns

Mathematical Functions
String Functions
Date/Time Functions
NULL Functions
Type Casts
Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length	Precision
Year	<add as new column>	YEAR(ConvertDate)	four-byte signed inte...		
Month	<add as new column>	MONTH(ConvertDate)	four-byte signed inte...		
Quarter	<add as new column>	DATEPART("QQ",ConvertDate)	four-byte signed inte...		
Day	<add as new column>	DAY(ConvertDate)	four-byte signed inte...		
sessionoftheday	<add as new column>	(DT_I4)(LEN(Time) == 0 ? "0" : (LEN(Time) == 5 ? SUBSTRING(Time,1,2) : SUBSTRING(Time,1,1)))	four-byte signed inte...		

Configure Error Output... OK Cancel Help

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters
Columns

Mathematical Functions
String Functions
Date/Time Functions
NULL Functions
Type Casts
Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length
Session	<add as new column>	(DT_STR,10,1252)(sessionoftheday < 5 ? "Night" : (sessionoftheday < 13 ? "Morning" : (sessionoftheday < 18 ? "Afternoon" : "Night")))	string [DT_STR]	10

Configure Error Output... OK Cancel Help

Sau khi biến đổi tính toán thì tiến hành Sort để nhằm xóa các giá trị bị trùng lặp.

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass Thr...
<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Time	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Day_of_Week	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Accident_Index	<input checked="" type="checkbox"/>
<input type="checkbox"/>	ConvertDate	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	DateKey	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	Con
DateKey	DateKey	ascending	1	

☒ Remove rows with duplicate sort values

OK Cancel Help

Nhấp chuột vào OLE DB Destination để đưa dữ liệu vào DimDate, chọn DB và Table

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:
DESKTOP-4J6IM9F\SYO.DDS_DATH New...

Data access mode:
Table or view - fast load

Name of the table or the view:
[dbo].[DimDate] New...

☐ Keep identity ☒ Table lock
☐ Keep nulls ☒ Check constraints

Rows per batch:

Maximum insert commit size:

< > View Existing

OK Cancel Help

Sang mục Mapping tiến hành mapping nguồn dữ liệu NDS với bảng Dim sau đó chọn OK.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager

Mappings

Error Output

Available Input Co...

Name

Date

Time

Day_of_Week

Accident_Index

ConvertDate

DateKey

Year

Month

Quater

Available ...

Name

DateID

Time

Date

Year

Month

Quater

Day

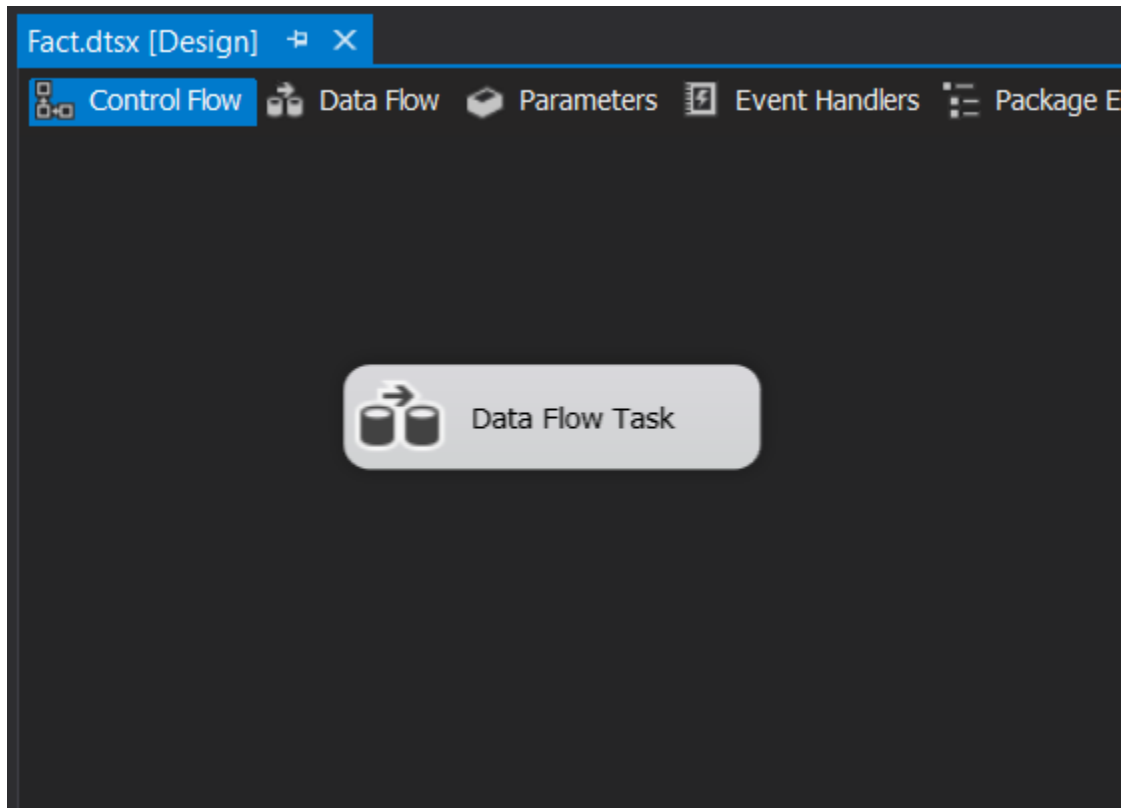
Date...

Input Column	Destination Column
<ignore>	DateID
Time	Time
ConvertDate	Date
Year	Year
Month	Month
Quater	Quater
Day	Day
DateKey	DateKey
Session	Sesion

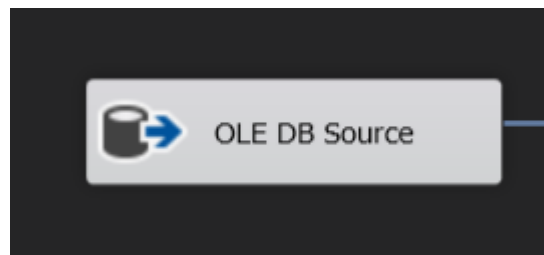
OK Cancel Help

• FactAccident

Kéo thả Data Flow Task để tiến hành xây Data Flow cho bảng Fact.



Kéo task OLE DB Source để tiến hành nạp source từ NDS vào



Chọn OLE DB Source sau đó tiến hành kết nối DB và viết truy vấn để lấy các dữ liệu cần thiết từ các bảng của NDS

.... Code

```
SELECT a.Accident_Index, a.AccidentSeverity, a.LocalAuthority_District,  
a.UrbanOrRuralArea, a.RoadType, a.PoliceForce, a.LocalAuthority_Highway, a.[Date],
```

a.NumberOfVehicles, a.NumberOfCasualties, a.Speedlimit, c.AgeBandOfCasualty,
c.CasualtyType,

c.SexOfCasualty, v.AgeBandOfDriver, v.AgeOfDriver,
v.AgeOfVehicle_Manufacture, a.pcd8, v.JourneyPurposeOfDriver, v.JunctionLocation,
v.SexOfDriver, v.VehicleType

FROM Accident AS a INNER JOIN

Casualtie AS c ON a.Accident_Index = c.Accident_Code INNER JOIN

Vehicle AS v ON a.Accident_Index = v.Accident_Code

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:

DESKTOP-4J6IM9F\SYO.NDS_DATH

New...

Data access mode:

SQL command

SQL command text:

```
SELECT a.Accident_Index, a.AccidentSeverity,
a.LocalAuthority_District, a.UrbanOrRuralArea, a.RoadType,
a.PoliceForce, a.LocalAuthority_Highway, a.[Date],
a.NumberOfVehicles, a.NumberOfCasualties, a.Speedlimit,
c.AgeBandOfCasualty, c.CasualtyType,
c.SexOfCasualty, v.AgeBandOfDriver, v.AgeOfDriver,
v.AgeOfVehicle_Manufacture, a.pcd8, v.JourneyPurposeOfDriver,
v.JunctionLocation, v.SexOfDriver, v.VehicleType
FROM Accident AS a INNER JOIN
Casualtie AS c ON a.Accident_Index =
c.Accident_Code INNER JOIN
Vehicle AS v ON a.Accident_Index = v.Accident_Code
```

Parameters...

Build Query...

Browse...

Parse Query

Preview...

OK

Cancel

Help

Sau khi đưa dữ liệu từ NDS vào tiến hành Lookup cho bảng DimAccidentServity

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Cache mode

☒ Full cache
☐ Partial cache
☐ No cache

Connection type

☐ Cache connection manager
☒ OLE DB connection manager

Specify how to handle rows with no matching entries

Redirect rows to no match output

OK Cancel Help

Connect đến bảng

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Specify a data source to use. You can select a table in a data source view, a table in a database connection, or the results of an SQL query.

OLE DB connection manager:
DESKTOP-4J6IM9F\SYO.DDS_DATH New...

☒ Use a table or a view:
[dbo].[DimAccidentSeverity] New...

☐ Use results of an SQL query:
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Tại mục Column chuột phải vào khoảng trống các Table chọn Edit Mapping

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Available Input Columns

Name
AccidentSeverity
LocalAuthority_District
UrbanOrRuralArea
RoadType
PoliceForce
LocalAuthority_High...
Date
NumberOfVehicles
NumberOfCasualties

Available Lookup Columns

<input type="checkbox"/> Name		I...
<input type="checkbox"/> AccidentSeverityID		
<input type="checkbox"/> AccidentSeverityKey		
<input type="checkbox"/> AccidentSeverityKeyName		

Lookup Column	Lookup Operation	Output Alias
---------------	------------------	--------------

OK

Cancel

Help

Tiến hành mapping

Create Relationships

Input Column	Lookup Column
AccidentSeverity	AccidentSeverityKey

OK Cancel

Tương tự cho các bảng còn lại

Sau khi Lookup các bảng tiến hành Conversion lại dữ liệu.

Biến đổi Number Of Vehicle và Number Of Casualty từ dạng nvarchar sang dạng int.

Biến đổi Date từ dạng nvarchar sang dạng varchar.

Biến đổi Speedlimit từ dạng nvarchar sang dạng int.

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input Columns

- ☒ Name
- ☐ AccidentSeverity
- ☐ LocalAuthority_District
- ☐ UrbanOrRuralArea
- ☐ RoadType
- ☐ PoliceForce
- ☐ LocalAuthority_Highway

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
NumberOfCasualties	Copy of NumberOfC...	four-byte signed integ...				
NumberOfVehicles	Copy of NumberOfV...	four-byte signed integ...				
Date	ConvertDate	string [DT_STR]	255			1252 (ANS
Speedlimit	ConvertSpeedlimit	four-byte signed integ...				

Configure Error Output... OK Cancel Help

Sử dụng Derived Column để chuyển thuộc tính Build Up Road có giá trị Speedlimit < 50 và > = 50

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

- Variables and Parameters
- Columns

- Mathematical Functions
- String Functions
- Date/Time Functions
- NULL Functions
- Type Casts
- Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length	Precision	Scale	Code Page
Built-up Road v	<add as new column>	(DT_STR,20,1252)(ConvertSpeedlimit < 50 ? "Built-up Road" : "No...)	string [DT_STR]	20			1252 (ANSI - Latin I)

Configure Error Output... OK Cancel Help

Cuối cùng đổ dữ liệu vào Bảng Fact. Tạo một OLE DB Destination. Chọn DB và Table

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:

Data access mode:

Name of the table or the view:

☐ Keep identity ☒ Table lock
☐ Keep nulls ☒ Check constraints

Rows per batch:

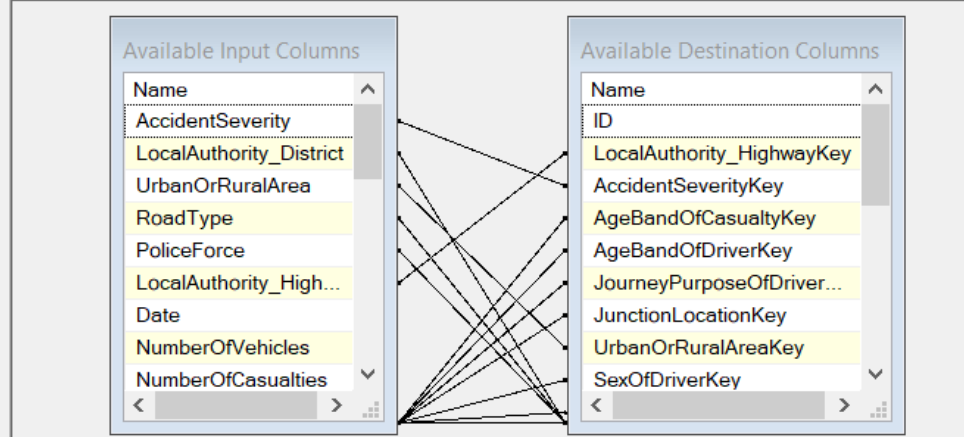
Maximum insert commit size:

Mapping lại dữ liệu

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output



Input Column	Destination Column
<ignore>	ID
LocalAuthority_Highway	LocalAuthority_HighwayKey
AccidentSeverity	AccidentSeverityKey
AgeBandOfCasualty	AgeBandOfCasualtyKey
AgeBandOfDriver	AgeBandOfDriverKey
JourneyPurposeOfDriver	JourneyPurposeOfDriverKey
JunctionLocation	JunctionLocationKey
UrbanOrRuralArea	UrbanOrRuralAreaKey
SexOfDriver	SexOfDriverKey
VehicleType	VehicleTypeKey
RoadType	RoadTypeKey
ConvertDate	DateKey
LocalAuthority_District	LocalAuthority_DistrictKey
CasualtyType	CasualtyTypeKey
PoliceForce	PoliceForceKey
pcd8	GeographyKey
Copy of NumberOfVehicles	NumberOfVehicles

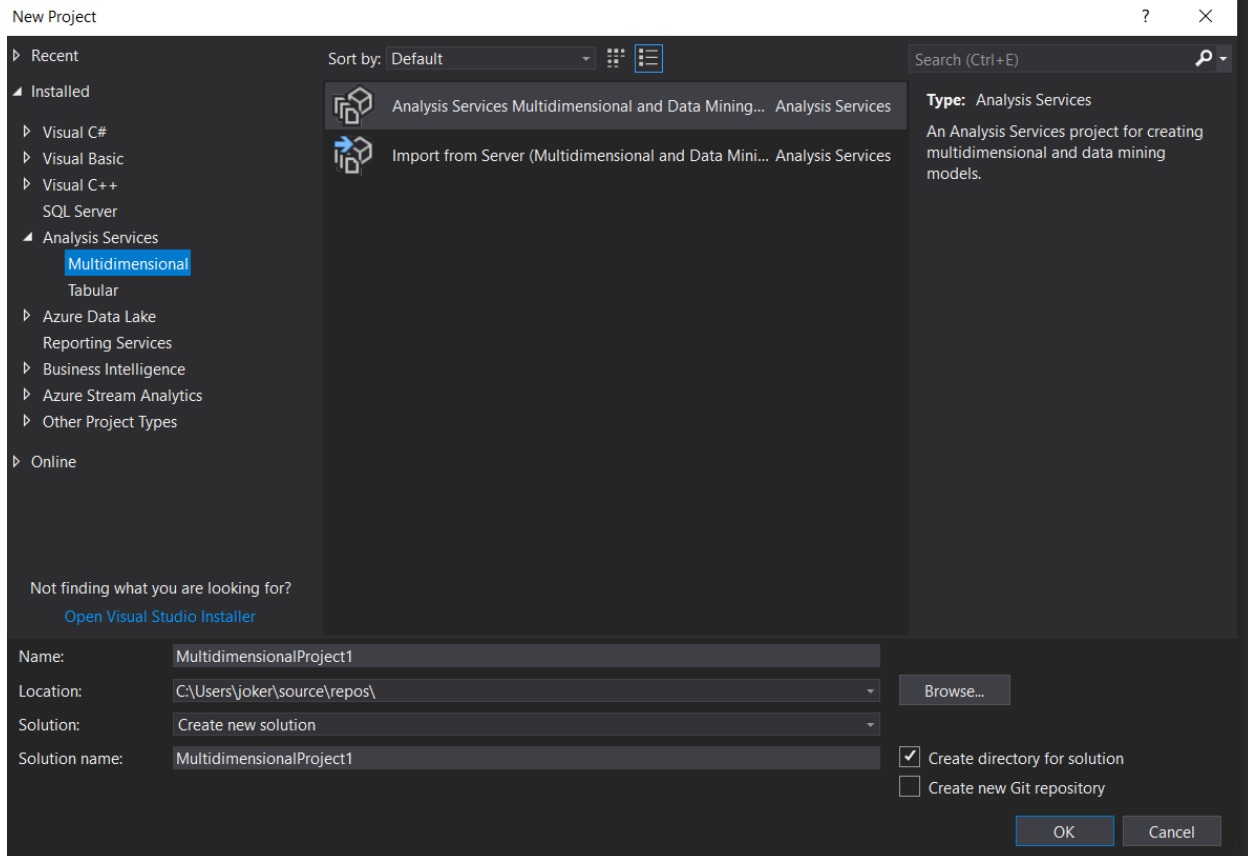
OK

Cancel

Help

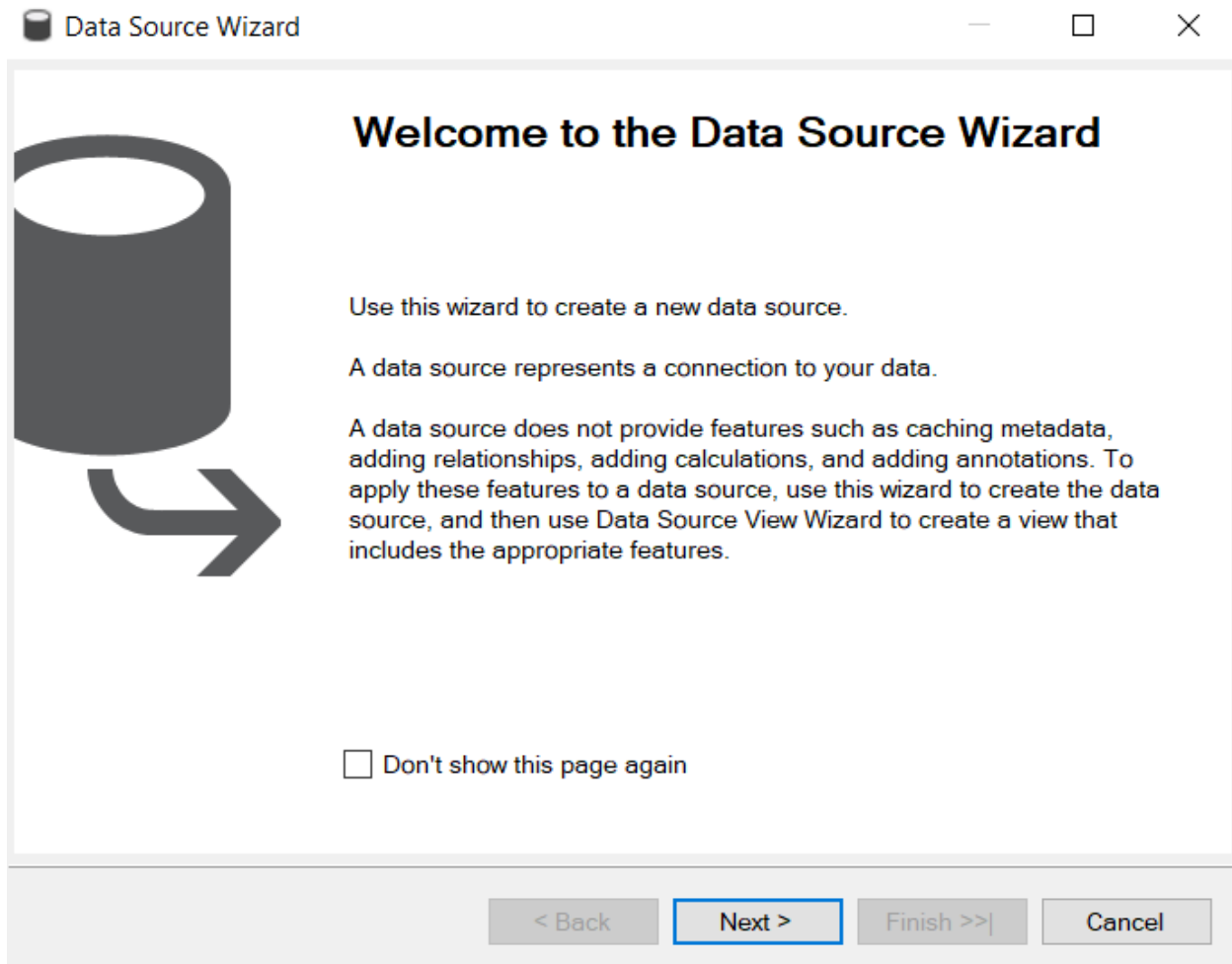
Chương 3: OLAP CUBE và MDX

1. Tạo ứng dụng SSAS



1. Tạo Data Source

Chuột phải vào DataSource và chọn New Data Source.Chọn Next để tiếp tục.



Connect đến Data có sẵn hoặc tạo connect mới. Connect đến DDS đã tạo trước đó.

Data Source Wizard

Select how to define the connection

You can select from a number of ways in which your data source will define its connection string.

☐ Create a data source based on another object

☒ Create a data source based on an existing or new connection

Data connections:

DESKTOP-4J6IM9F\SYO
DESKTOP-4J6IM9F\SYO.AdventureWorksDW2012
DESKTOP-4J6IM9F\SYO.DDS_DATH
DESKTOP-4J6IM9F\SYO.NDS_DATH

Data connection properties:

Property	Value
Data Source	DESKTOP-4J6IM9F\
Integrated Sec...	SSPI
Provider	MSOLAP.8

New...

Delete

< Back

Next >

Finish >>|

Cancel

Nhập User Name và pass của Window

58

Data Source Wizard

Impersonation Information

You can define what Windows credentials Analysis Services will use to connect to the data source.

☒ Use a specific Windows user name and password

User name:

Password:

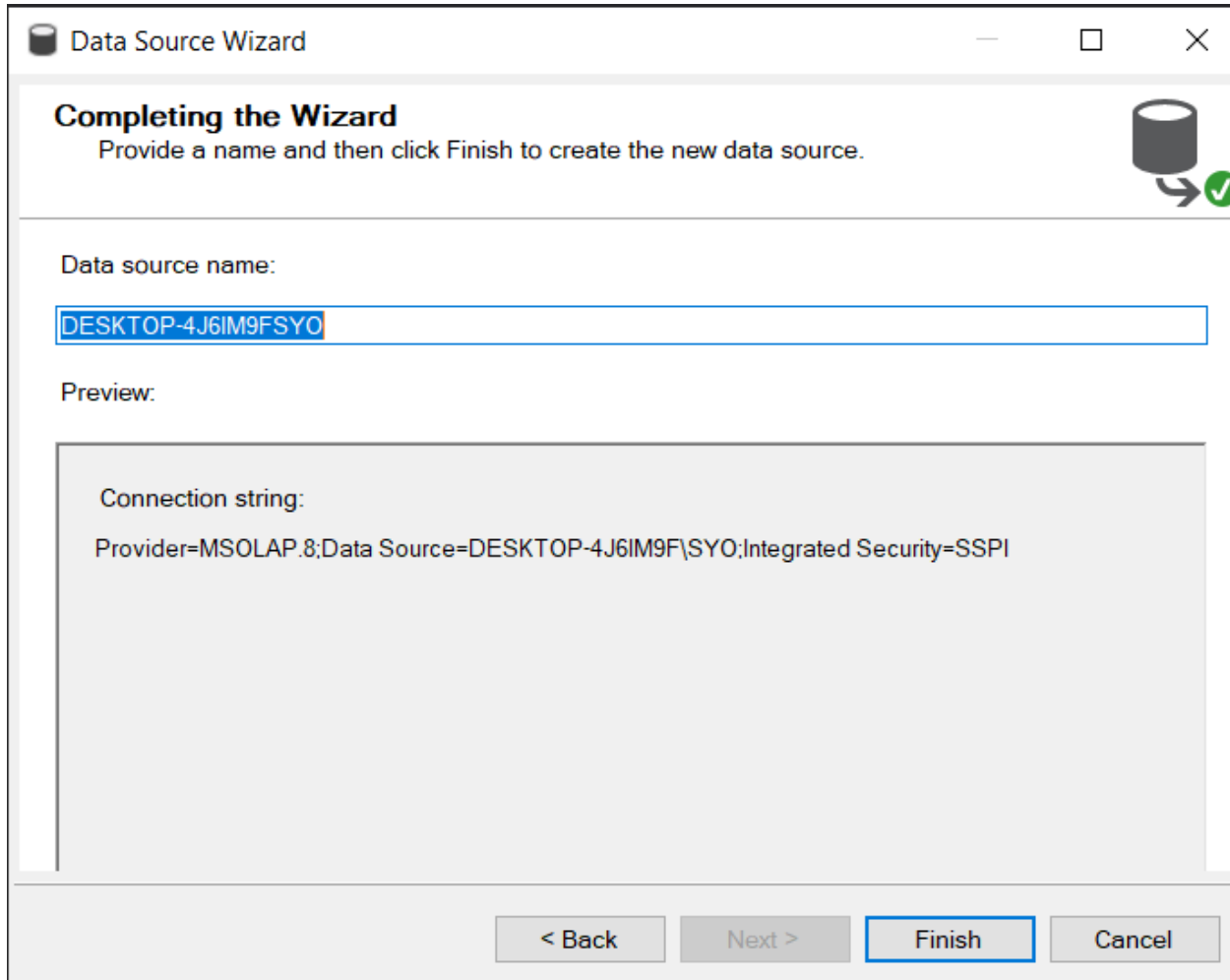
☐ Use the service account

☐ Use the credentials of the current user

☐ Inherit

< Back **Next >** Finish >>| Cancel

Sau đó chọn Finish.



The image shows a screenshot of the 'Data Source Wizard' window. The title bar reads 'Data Source Wizard'. The main heading is 'Completing the Wizard', followed by the instruction 'Provide a name and then click Finish to create the new data source.' To the right of this text is an icon of a database cylinder with a green checkmark. Below the instruction, there is a text box labeled 'Data source name:' containing the text 'DESKTOP-4J6IM9FSYO'. Underneath this is a 'Preview:' section with a light gray background, displaying the 'Connection string:' as 'Provider=MSOLAP.8;Data Source=DESKTOP-4J6IM9F\SYO;Integrated Security=SSPI'. At the bottom of the window are four buttons: '< Back', 'Next >', 'Finish' (which is highlighted with a blue border), and 'Cancel'.

2. Tạo Data Source View

Chọn chuột phải vào Data Source Views và New Data Source View để tiến hành tạo data mới. Sau đó chọn Next.

Data Source Wizard

Completing the Wizard
Provide a name and then click Finish to create the new data source.

Data source name:

Preview:

Connection string:
Provider=MSOLAP.8;Data Source=DESKTOP-4J6IM9F\SYO;Integrated Security=SSPI

< Back Next > **Finish** Cancel

Chọn data source đã tạo từ bước Data Source trước đó hoặc new một data mới

Data Source View Wizard

Select a Data Source
Select an existing relational data source or create a new one.

Relational data sources:

DDS DATH

Data source properties:

Property	Value
Data Source	DESKTOP-4J6IM9F...
Initial Catal...	DDS_DATH
Integrated ...	SSPI
Provider	SQLNCLI11.1

New Data Source... Advanced...

< Back Next > Finish >>| Cancel

Chọn ra các bảng cần sử dụng và next (Lấy hết các bảng từ DDS)

Data Source View Wizard

Select Tables and Views

Select objects from the relational database to be included in the data source view.

Available objects:

Name	Type
DimAccidentSeverity (dbo)	Table
DimAgeBandOfCasualty (dbo)	Table
DimAgeBandOfDriver (dbo)	Table
DimCasualtyType (dbo)	Table
DimDate (dbo)	Table
DimGeography (dbo)	Table
DimJourneyPurposeOfDriver ...	Table
DimJunctionLocation (dbo)	Table
DimLocalAuthority_District (d...	Table
DimLocalAuthority_Highway (Table

Filter:

☐ Show system objects

Included objects:

Name	Type
------	------

Add Related Tables

< Back **Next >** Finish >>| Cancel

Sau đó chọn Finish

Data Source View Wizard

Completing the Wizard
Provide a name, and then click Finish to create the new data source view.

Name:

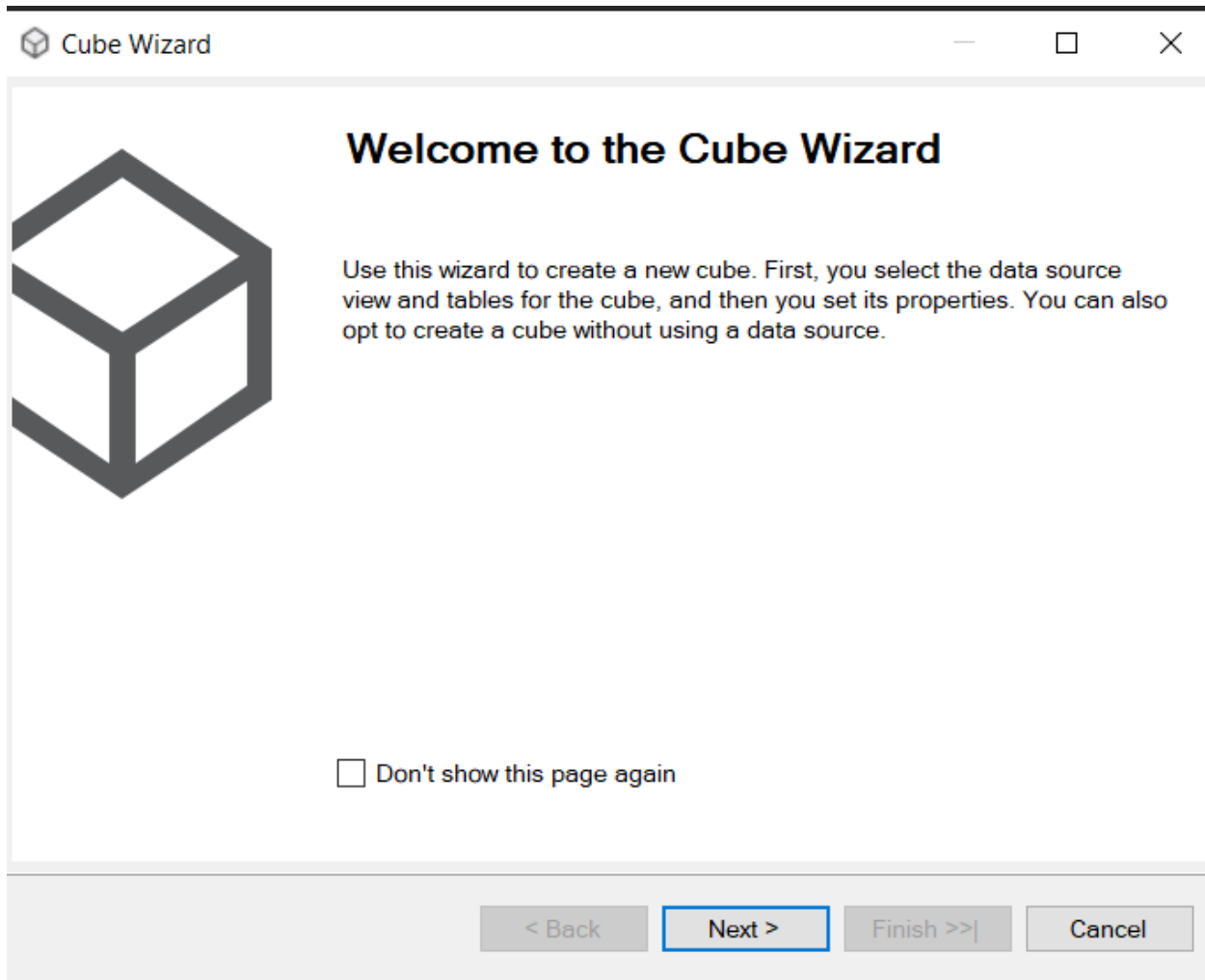
Preview:

DDS DATH 1

< Back Next > **Finish** Cancel

3. Tạo Cube

Chuột phải vào Cubes và chọn New Cube sau đó Next



Sử dụng table đã tồn tại và chọn Next

Cube Wizard

Select Creation Method

Cubes can be created by using existing tables, creating an empty cube, or generating tables in the data source.

How would you like to create the cube?

☒ Use existing tables

☐ Create an empty cube

☐ Generate tables in the data source

Template:

(None)

Description:

Create a cube based on one or more tables in a data source.

< Back **Next >** Finish >> | Cancel

Chọn nút Suggest và chọn bảng Fact

Cube Wizard

Select Measure Group Tables

Select a data source view or diagram and then select the tables that will be used for measure groups.

Data source view:
DDS DATH


Measure group tables:

Suggest

<input type="checkbox"/>	DimDate
<input type="checkbox"/>	DimGeography
<input type="checkbox"/>	DimJourneyPurposeOfDriver
<input type="checkbox"/>	DimJunctionLocation
<input type="checkbox"/>	DimLocalAuthority_District
<input type="checkbox"/>	DimLocalAuthority_Highway
<input type="checkbox"/>	DimPoliceForce
<input type="checkbox"/>	DimRoadType
<input type="checkbox"/>	DimSexOfDriver
<input type="checkbox"/>	DimUrbanOrRuralArea
<input type="checkbox"/>	DimVehicleType
<input checked="" type="checkbox"/>	FactAccident
<input type="checkbox"/>	sysdiagrams

< Back Next > Finish >> Cancel

Chọn các Measure cho Fact






 Cube Wizard

Select Measures

Select measures that you want to include in the cube.



☒ Measure

- ☒  Fact Accident
- ☒  Number Of Vehicles
- ☒  Number Of Casualties
- ☒  Speedlimit
- ☒  Fact Accident Count

< Back

Next >

Finish >>|

Cancel

Chọn các bảng Dim

Cube Wizard


Select Existing Dimensions
Select existing dimensions to include in the cube.




☒ Dimension

- ☒ Dim Age Band Of Casualty
- ☒ Dim Urban Or Rural Area
- ☒ Dim Junction Location
- ☒ Dim Casualty Type
- ☒ Dim Date
- ☒ Dim Police Force
- ☒ Dim Local Authority Highway
- ☒ Dim Sex Of Driver
- ☒ Dim Vehicle Type
- ☒ Dim Local Authority District
- ☒ Dim Journey Purpose Of Driver
- ☒ Dim Road Type
- ☒ Dim Age Band Of Driver
- ☒ Dim Accident Severity
- ☒ Dim Geography

< Back Next > Finish >>| Cancel


Sau đó chọn Finish


Cube Wizard


Completing the Wizard


Name the cube, review its structure, and then click Finish to save the cube.





Cube name:


Preview:



Measure groups



Fact Accident



Number Of Vehicles



Number Of Casualties



Speedlimit



Fact Accident Count



Dimensions



Dim Age Band Of Casualty



Dim Urban Or Rural Area



Dim Junction Location


Dim Casualty Type


Dim Date


Dim Police Force


Dim Local Authority Highway


Dim Sex Of Driver

< Back

Next >

Finish

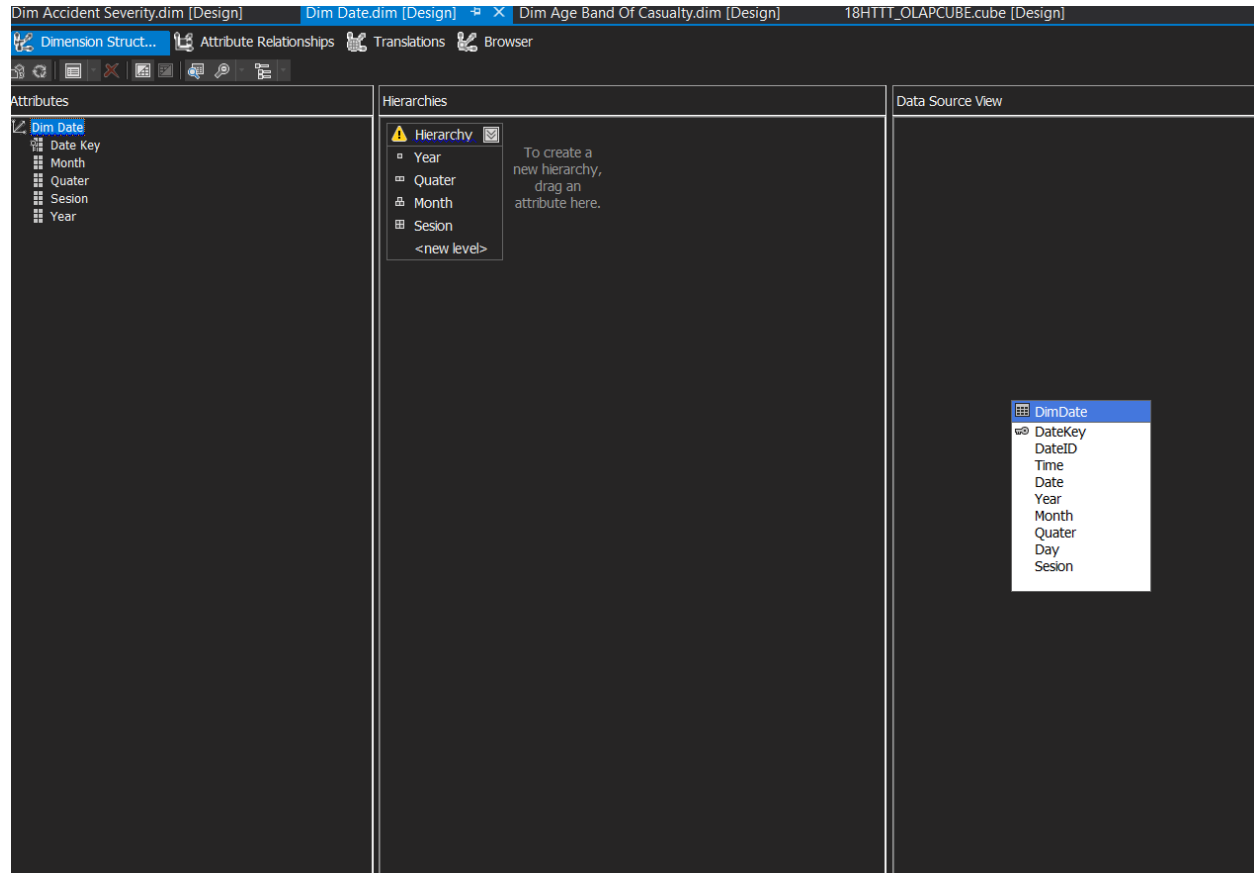
Cancel

4. Sửa bảng chiều

Với bảng Dim_Date

Trong Folder Dimensions. Nhấp đúp chuột vào Dimdate - Trong Data Source View, di chuyển các trường trong Dimdate vào Dim date trong Attributes.

70



Tương tự cho các bảng còn lại

5. Deploy cube lên server

Trong pane Solution Explore , chuột phải vào tên project và chọn Properties

18HTTT_07_OLAPCUBE Property Pages

Configuration: Active(Development) Platform: N/A Configuration Manager...

Configuration Properties

- Build
- Debugging
- Deployment

Deployment Server

Deployment Server Edition: Developer

Deployment Server Version: 15.0

Outputs

Output Path: bin\

Remove Passwords: True

Deployment Server Edition

The edition of the Analysis Services instance to which the project will be depl...

OK Cancel Apply

Tại mục Deployment đổi server thành server name của SQL Server sau đó nhấn Start để tiến hành deploy.

18HTTT_07_OLAPCUBE Property Pages

Configuration: Active(Development) Platform: N/A Configuration Manager...

Configuration Properties

- Build
- Debugging
- Deployment

Options

Processing Option: Default

Transactional Deployment: False

Server Mode: Deploy Changes Only

Target

Server: DESKTOP-4J6IM9F\SYO

Database: 18HTTT_07_OLAPCUBE

Server

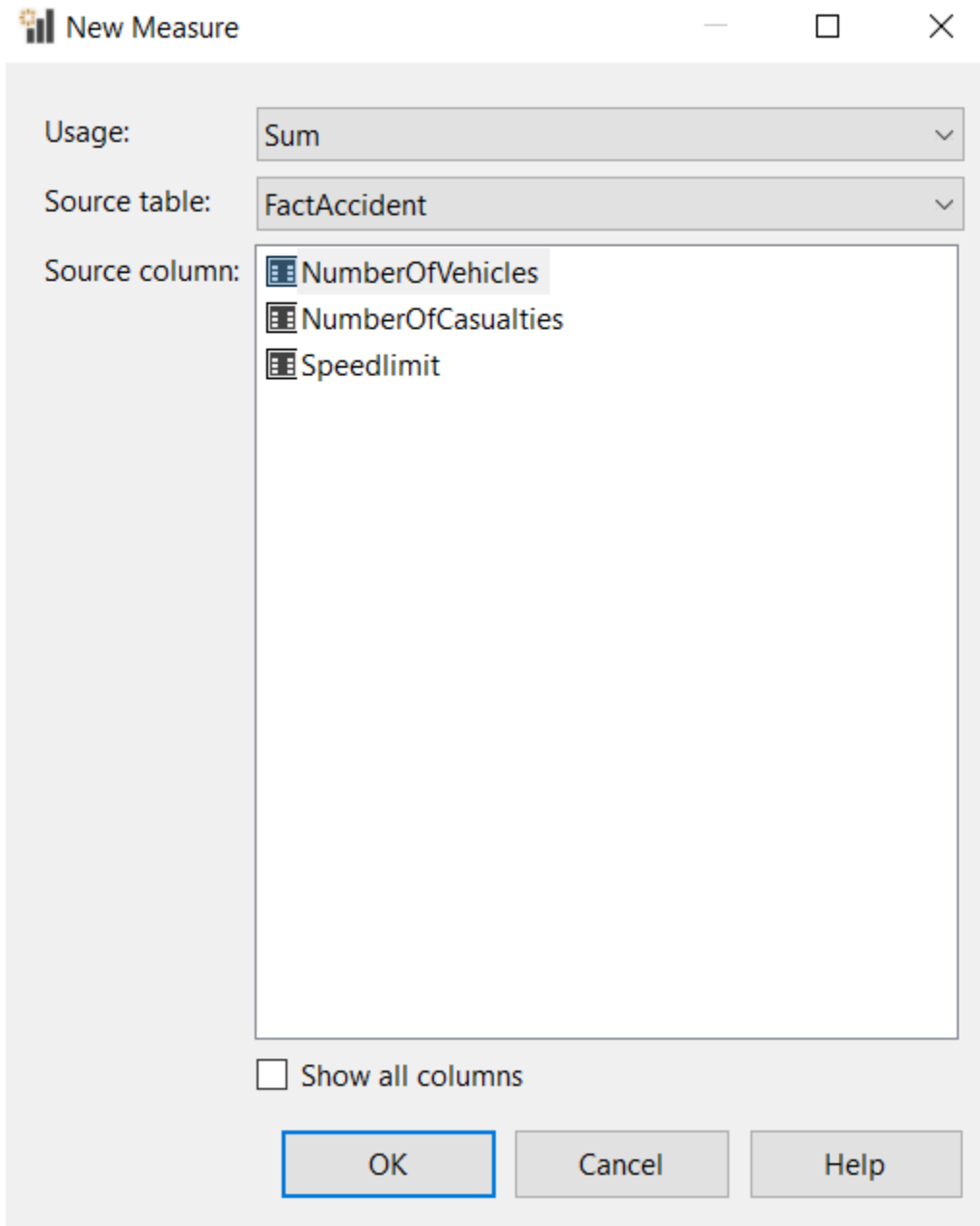
The Analysis Services instance to which the project will be deployed.

OK Cancel Apply

6. Tạo mới Measure

Chuyển sang Cube Designer bằng cách đúp chuột vào 18HTTT_OLAPCUBE.cube.

Trong tab Cube Structure ở ô Measure, nhấp phải vào Fact và chọn New Measure



New Measure

Usage: Sum

Source table: FactAccident

Source column:

- NumberOfVehicles
- NumberOfCasualties
- Speedlimit

☐ Show all columns

OK Cancel Help

7. Xóa dòng dữ liệu Unknow

Chọn bảng DimAccidentServity, chuột phải vào tên bảng ở mục Attribute và chọn Properties. Chỉnh sửa ở dòng UnknowMemberName, đổi Visible thành Hidden.

Properties	
Dim Accident Severity Dimension	
Advanced	
AttributeAllMemberName	
CurrentStorageMode	Molap
CurrentStringStoresCompatibilityLevel	1050
DependsOnDimension	
ErrorConfiguration	(custom)
Language	
MiningModelID	
ProcessingGroup	ByAttribute
ProcessingMode	Regular
ProcessingPriority	0
ProcessingRecommendation	None
ProcessingState	Unprocessed
Source	DDS DATH (Data source view)
StringStoresCompatibilityLevel	1050
UnknownMember	Hidden
UnknownMemberName	
WriteEnabled	False
Basic	
Description	
ID	Dim Accident Severity
Name	Dim Accident Severity
Type	Regular
Misc	
Collation	
MdxMissingMemberMode	Default
Storage	
ProactiveCaching	(none)
StorageMode	Molap
UnknownMember	
Specifies the existence of an unknown member, and whether that member is visible or hidden. Fact data not associated with a member can be associated with the unknown member.	

Tương tự cho các bảng còn lại.

Chương 4 Thực hiện ngôn ngữ MDX

1. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng (Fatal, Serious, Slight) ở các Địa phương (Local_Authority_(District)) trong tất cả các năm.

select

non empty ([Dim Date].[Hierarchy].[Year],[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns,

non empty ([Dim Local Authority District].[Local Authority District Name].[Local Authority District Name]) on rows

from [18HTTT_OLAPCUBE]

where [Measures].[Number Of Casualties]

```

select
non empty ([Dim Date].[Hierarchy].[Year],[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name])
on columns,
non empty ([Dim Local Authority District].[Local Authority District Name].[Local Authority District Name]) on rows
from [18HTTT_OLAPCUBE]
where [Measures].[Number Of Casualties]

```

100 %

Messages Results

	2011	2011	2011	2012	2012	2012	2013	2013	2013	2014	2014	2014
	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight
Amber Valley	19	113	1220	35	267	1070	(null)	106	197	(null)	(null)	(null)
Ashfield	10	216	1243	38	218	1161	21	157	1249	115	229	1169
Aylesbury Vale	61	344	1429	106	639	1415	28	471	1556	322	224	1520
Babergh	(null)	(null)	(null)	(null)	(null)	2	31	139	994	16	143	1669
Barking and Dagenham	21	139	1824	11	466	1618	64	78	1267	6	66	1988
Barnet	8	487	3915	8	580	3715	30	340	3493	7	356	4666
Barnsley	16	479	1758	43	343	1580	11	324	1478	5	596	3166
Basildon	3	118	2290	19	196	1763	2	217	1390	18	370	2177
Bassetlaw	69	374	1120	49	443	1202	80	319	1319	45	200	959
Bedford	7	124	1778	1	199	1489	11	139	1183	13	203	1650
Bexley	15	160	1683	8	131	1482	1	130	1074	1	77	1573
Birmingham	310	1540	10605	41	1770	8977	355	1467	10659	19	1723	11446
Blackburn with Darwen	30	228	1858	19	327	2451	(null)	356	1699	34	286	1622
Blackpool	49	199	1932	2	180	2032	31	277	1817	(null)	173	2119
Bolsover	(null)	103	1344	52	135	1623	(null)	35	487	(null)	(null)	(null)

2. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng ở các Địa Phương (Local_Authority_(District)) theo các Quý trong từng năm.

select

non empty ([Dim Date].[Year].[Year],[Dim Date].[Quater].[Quater],

[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns,

`non empty` ([Dim Local Authority District].[Local Authority District Name].[Local Authority District Name]) on rows

`from` [18HTTT_OLAPCUBE]

`where` [Measures].[Number Of Casualties]

```
select
non empty ([Dim Date].[Year].[Year],[Dim Date].[Quater].[Quater],
[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns,
non empty ( [Dim Local Authority District].[Local Authority District Name].[Local Authority District Name]) on rows
from [18HTTT_OLAPCUBE]
where [Measures].[Number Of Casualties]
```

	2011	2011	2011	2011	2011	2011	2011	2011	2011	2011	2011	2011	2011	2012	2012	2012	2012	2012	2012
	1	1	1	2	2	2	3	3	3	4	4	4	4	1	1	1	2	2	2
	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal
Amber Valley	1	28	240	(null)	18	360	(null)	26	300	18	41	320	(null)	24	242	(null)	149		
Ashfield	(null)	105	415	(null)	66	233	10	27	311	(null)	18	284	36	48	199	2	67		
Aylesbury Vale	1	37	410	(null)	49	332	9	25	282	51	233	405	16	121	230	54	280		
Babergh	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)	(null)		
Barking and Dagenham	(null)	45	449	11	24	384	8	30	347	2	40	644	(null)	324	380	(null)	19		
Barnet	2	132	1031	4	96	1007	1	110	1075	1	149	802	1	80	869	1	203		
Barnsley	10	42	502	1	148	298	3	70	484	2	219	474	6	132	249	3	113		
Basildon	(null)	43	491	(null)	31	252	3	30	1212	(null)	14	335	1	58	394	18	17		
Bassetlaw	6	143	325	34	164	238	1	34	280	28	33	277	4	76	256	22	79		
Bedford	2	43	386	(null)	33	504	2	20	477	3	28	411	1	71	416	(null)	20		
Bexley	(null)	49	360	14	42	411	1	31	277	(null)	38	635	1	30	382	7	28		
Birmingham	21	277	2196	12	429	2877	50	455	2884	227	379	2648	16	356	2432	8	718		
Blackburn with Darwen	12	15	314	1	38	508	16	81	623	1	94	413	(null)	28	665	1	25		
Blackburn	(null)	67	413	(null)	57	211	40	21	758	(null)	54	150	(null)	26	576	1	50		

3. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng và Thời Điểm Trong Ngày (Morning: 5am-12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) trong các năm.

`select`

`non empty` ([Dim Date].[Sesion].[Sesion]) on rows,

`non empty` ([Dim Date].[Hierarchy].[Year],

[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns

from [18HTTT_OLAPCUBE]

where [Measures].[Fact Accident Count]

```

select
non empty ([Dim Date].[Session].[Session]) on rows,
non empty ([Dim Date].[Hierarchy].[Year],
[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns
from [18HTTT_OLAPCUBE]
where [Measures].[Fact Accident Count]

```

100 %

	2011	2011	2011	2012	2012	2012	2013	2013	2013	2014	2014	2014
	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight
Afternoon	1022	10494	70068	1002	10320	65183	1091	8688	54130	1002	10127	63877
Morning	1093	10464	65146	1035	10326	65967	861	9461	60925	1072	10585	68758
Night	1000	8392	52021	632	7450	43473	787	8183	49001	848	7936	53599

4. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng, Vùng (Urban_or_Rural_Area), và Kiểu Đường (Road Type) trong các năm.

select

non empty ([Dim Urban Or Rural Area].[Urban Or Rural Area Name].[Urban Or Rural Area Name],[Dim Road Type].[Road Type Name].[Road Type Name]) on rows,

non empty ([Dim Date].[Hierarchy].[Year],[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]) on columns

from [18HTTT_OLAPCUBE]

where [Measures].[Fact Accident Count]

```
select
non empty ([Dim Urban Or Rural Area].[Urban Or Rural Area Name].[Urban Or Rural Area Name],[Dim Road Type].[Road Type Name]
non empty ([Dim Date].[Hierarchy].[Year],[Dim Accident Severity].[Accident Severity Key Name].[Accident Severity Key Name]
from [18HTTT_OLAPCUBE]
where [Measures].[Fact Accident Count]
```

100 %

Messages Results

		2011	2011	2011	2012	2012	2012	2013	2013	2013	2014	2014	2014
		Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight	Fatal	Serious	Slight
Rural	Dual carriageway	573	3893	15914	525	2347	14212	487	2337	12683	488	2602	14268
Rural	One way street	20	76	223	13	25	153	(null)	42	207	1	49	155
Rural	Roundabout	18	377	4703	13	452	4650	12	446	3688	14	451	4261
Rural	Single carriageway	1504	9141	38266	1304	9281	35161	1317	8794	32834	1488	9663	34982
Rural	Slip road	8	111	1382	6	147	1437	14	95	1049	2	96	1197
Rural	Unknown	16	16	323	(null)	24	187	34	31	158	3	16	113
Urban	Dual carriageway	222	1940	17343	108	1730	16526	190	1890	16506	155	2077	18207
Urban	One way street	14	285	2348	21	253	2187	8	318	2516	16	375	2833
Urban	Roundabout	18	686	8372	13	662	7851	21	655	7135	27	691	8463
Urban	Single carriageway	721	12707	97097	651	13072	90987	654	11614	86196	725	12540	100526
Urban	Slip road	1	83	772	15	59	875	2	82	859	(null)	54	1058
Urban	Unknown	(null)	35	492	(null)	44	397	(null)	28	225	3	34	171

5. Tổng hợp số lượng tai nạn theo Mục Đích Hành Trình (Journey Purpose) và Loại Phương Tiện (Vehicle_Type).

select

non empty ([Dim Vehicle Type].[Vehicle Type Name].[Vehicle Type Name]) on rows,

non empty ([Dim Date].[Hierarchy].[Year],[Dim Journey Purpose Of Driver].[Journey Purpose Of Driver Name].[Journey Purpose Of Driver Name]) on columns

from [18HTTT_OLAPCUBE]

where [Measures].[Fact Accident Count]

```

select
non empty ([Dim Vehicle Type].[Vehicle Type Name].[Vehicle Type Name]) on rows,
non empty ([Dim Date].[Hierarchy].[Year],[Dim Journey Purpose Of Driver].[Journey Purpose Of Driver Name].[Journey Purpose Of Driver Name]) on columns,
from [18HTTT_OLAPCUBE]
where [Measures].[Fact Accident Count]

```

	2011	2011	2011	2011	2011	2011	2011
	Commuting to/from work	Journey as part of work	Not known	Other	Pupil riding to/from school	Taking pupil to/from school	Commuting to/from work
Agricultural vehicle	3	232	66	(null)	(null)	(null)	
Bus or coach (17 or more pass seats)	18	4824	639	4	2	102	
Car	13965	15412	131760	1328	201	2251	
Data missing or out of range	(null)	(null)	1	(null)	(null)	(null)	
Electric motorcycle	(null)	(null)	(null)	(null)	(null)	(null)	
Goods 7.5 tonnes mgw and over	34	3323	484	(null)	(null)	1	
Goods over 3.5t and under 7.5t	33	903	400	1	(null)	1	
Goods vehicle - unknown weight	(null)	(null)	(null)	(null)	(null)	(null)	
Minibus (8 - 16 passenger seats)	31	418	252	1	(null)	43	
Mobility scooter	(null)	(null)	5	5	(null)	(null)	
Motorcycle - unknown cc	(null)	(null)	2	(null)	(null)	(null)	
Motorcycle 125cc and under	749	599	3698	31	26	5	
Motorcycle 50cc and under	213	202	1585	26	61	7	
Motorcycle over 125cc and up to 500cc	210	168	1065	4	6	(null)	
Motorcycle over 500cc	768	480	4160	51	2	6	

6. Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng TNGT ở các vùng trong năm.

select

non empty ([Dim Date].[Year].[Year]) on columns,

non empty ([Dim Geography].[Region Name].[Region Name]) on rows

from [18HTTT_OLAPCUBE]

where [Measures].[Fact Accident Count]


```
select
non empty ([Dim Date].[Year].[Year]) on columns,
non empty ([Dim Geography].[Region Name].[Region Name]) on rows
from [18HTTT_OLAPCUBE]
where [Measures].[Fact Accident Count]
```

100 %

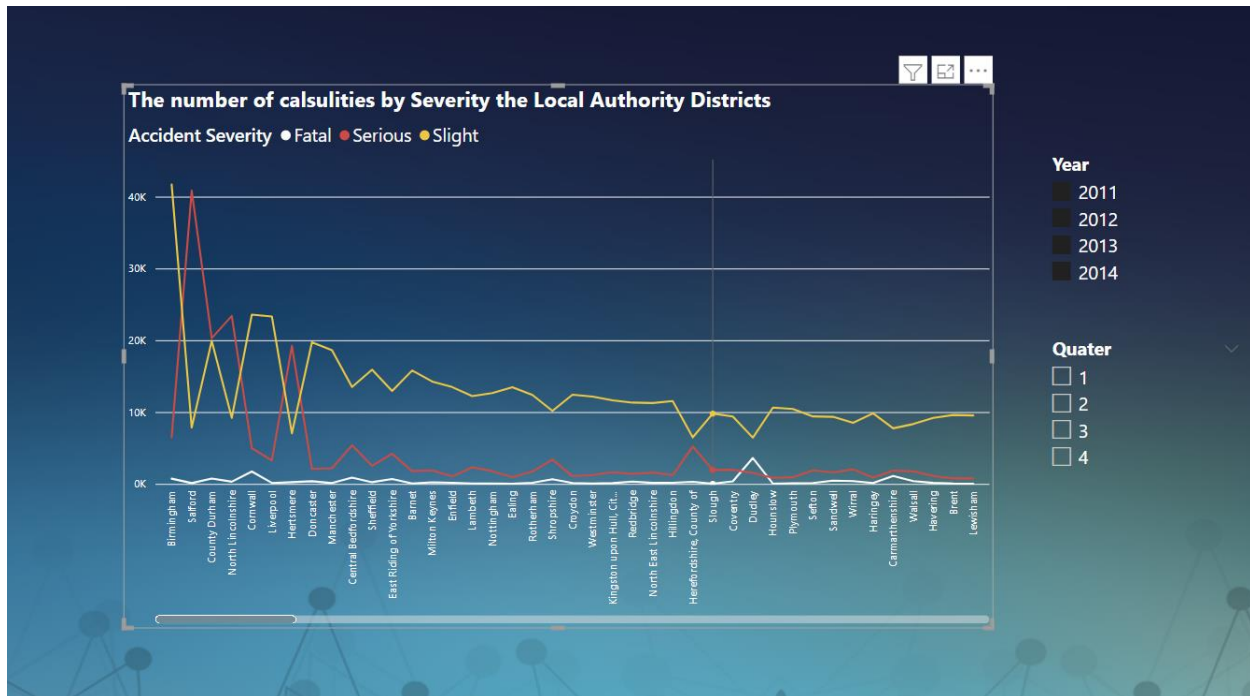
Messages

Results

	2011	2012	2013	2014
#N/A	1115	1012	1076	902
East Midlands	24138	23042	19155	15096
East of England	21617	20874	23666	27096
London	50608	50015	49294	56049
North East	5760	5937	5609	6116
North West	33380	29668	27988	28831
South East	25432	22314	13511	14023
South West	17330	15370	15572	16156
Wales	6863	6843	6162	6391
West Midlands	19010	17212	17942	23112
Yorkshire and the Humber	14447	13101	13152	24032

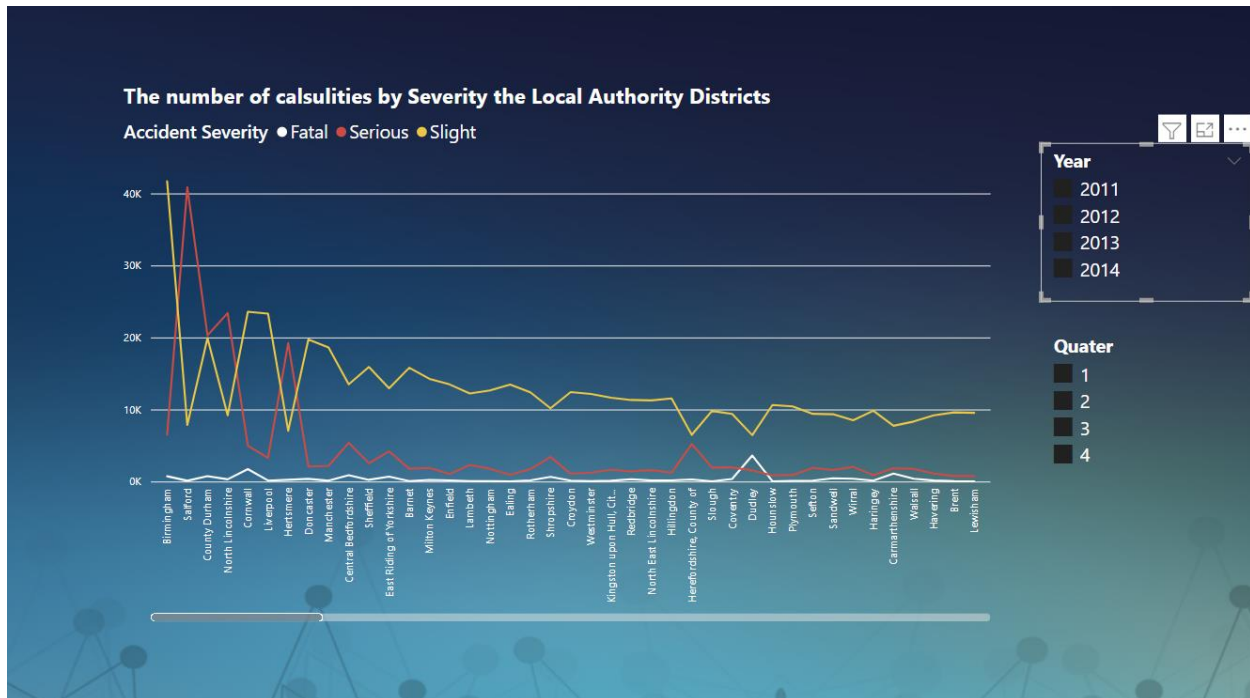
Chương 5: Trực quan hóa dữ liệu

1. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng (Fatal, Serious, Slight) ở các Địa phương (Local_Authority_(District)) trong tất cả các năm.



Từ năm 2012 đến 2014 Mức độ gây ra tử vong ở East Dorset là cao nhất, mức độ nhẹ ở Birmingham là cao nhất, mức độ nghiêm trọng ở Central Bedfordshire

2. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng ở các Địa Phương (Local_Authority_(District)) theo các Quý trong từng năm.



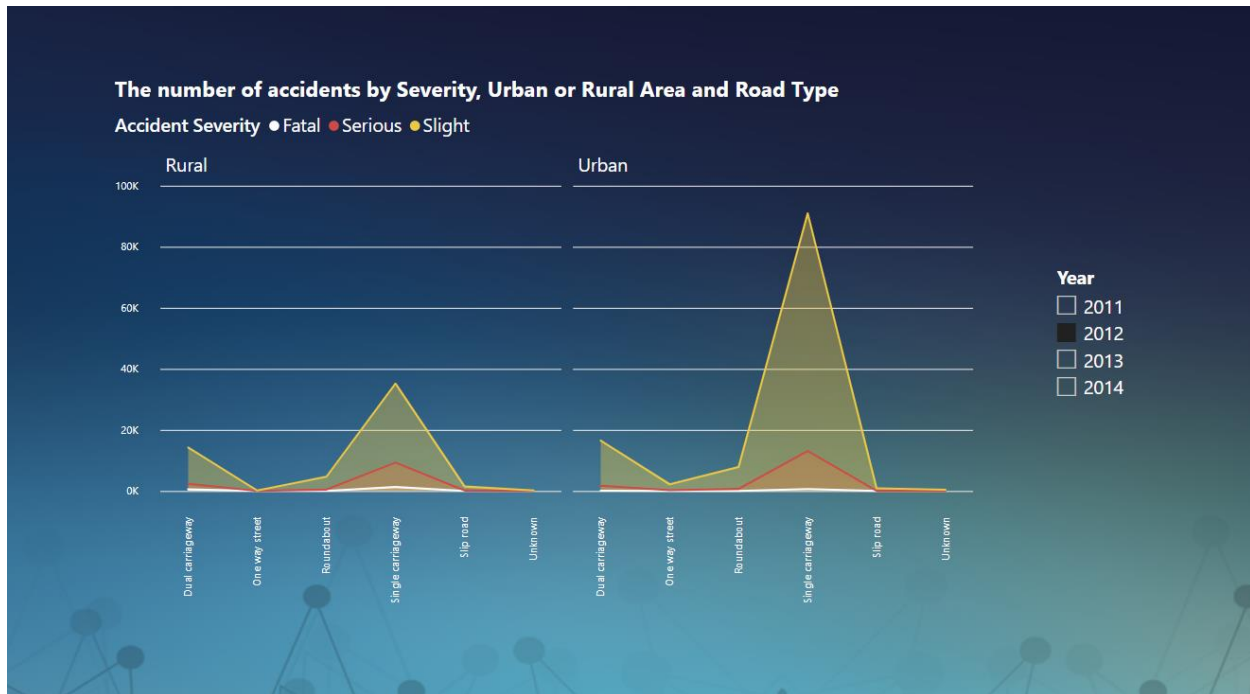
Từ năm 2012 đến 2014 qua các quý, Mức độ gây ra tử vong ở Uttlestford là cao nhất, mức độ nhẹ ở Birmingham là cao nhất, mức độ nghiêm trọng ở Salfard

3. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng và Thời Điểm Trong Ngày (Morning: 5am-12pm, Afternoon: 12pm-5pm, Evening: 5pm-9pm, Night: 9pm-5am) trong các năm.



Năm 2012 Số lượng tai nạn gây ra tử vong, nhẹ hoặc nghiêm trọng đều vào khoảng buổi sáng

4. Thống kê số lượng TNGT theo Mức Độ Nghiêm Trọng, Vùng (Urban_or_Rural_Area), và Kiểu Đường (Road Type) trong các năm.



Năm 2012 đa phần tai nạn loại nhẹ và nghiêm trọng diễn ra ở thành phố còn mức độ gây tử vong lại diễn ra ở vùng nông thôn nhiều hơn

5. Thống kê số lượng nạn nhân theo Mức Độ Nghiêm Trọng, Loại Nạn Nhân (Casualty Type) và Độ Tuổi trong các năm, Độ Tuổi được định nghĩa như sau:

Children: 0-15

Young adult: 0-17

Adult: 18-59

60 nd over: 60-...

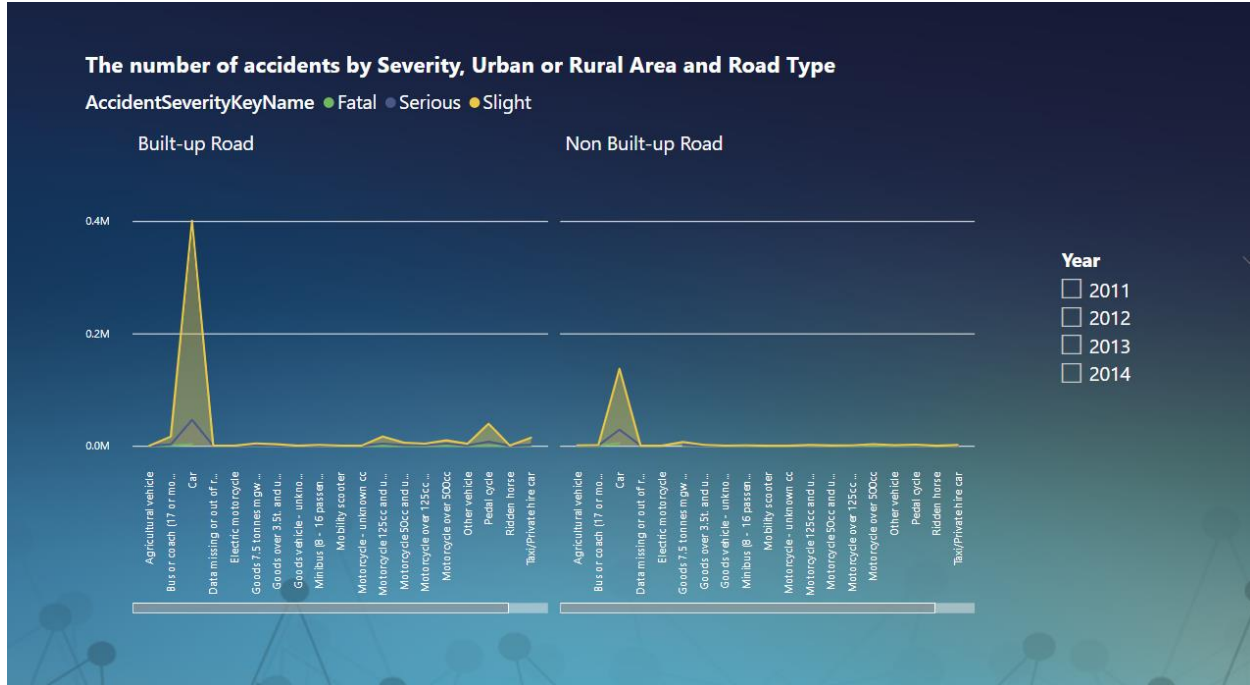


6. Tổng hợp số lượng tai nạn theo Mục Đích Hành Trình (Journey Purpose) và Loại Phương Tiện (Vehicle_Type).

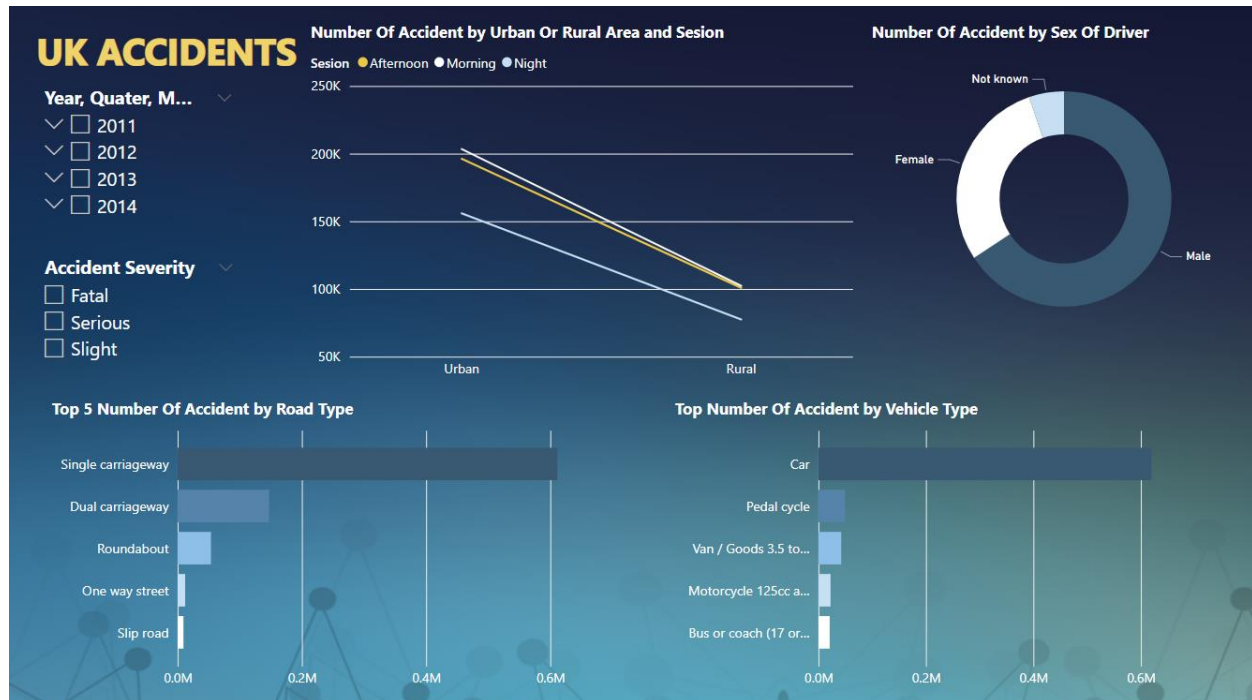
The number of accidents by Journey Purpose and Vehicle Type.

Vehicle Type	Commuting to/from work	Data missing or out of range	Journey as part of work	Not known	Other	Pupil riding to/from school	Taking pupil to/from school
Agricultural vehicle	18		1108	273	7		
Bus or coach (17 or more pass seats)	89		18046	2201	26	184	385
Car	54433	2	56832	488984	9606	1115	8753
Data missing or out of range				20	1		
Electric motorcycle				12			
Goods 7.5 tonnes mgw and over	137		12462	1690	12	6	4
Goods over 3.5t. and under 7.5t	145		3634	1268	33		2
Goods vehicle - unknown weight	4		149	146	1		3
Minibus (8 - 16 passenger seats)	93		1538	866	20	9	141
Mobility scooter	1		7	199	20	2	1
Motorcycle - unknown cc	14		5	127	4	1	
Motorcycle 125cc and under	3351		2680	16143	192	120	16
Motorcycle 50cc and under	770		742	5242	131	215	24
Motorcycle over 125cc and up to 500cc	879		639	4477	56	18	4
Motorcycle over 500cc	2821		1652	14999	237	15	19
Other vehicle	188		2406	2895	29	7	12
Pedal cycle	7397		3357	37055	580	1008	116
Ridden horse			11	292	3		
Taxi/Private hire car	388		12992	4291	58	13	152
Tram			86	4			
Van / Goods 3.5 tonnes mgw or under	3288		20625	18008	353	18	106
Total	74016	2	138971	599192	11369	2731	9738

7. Thống kê số lượng tai nạn theo Mức Độ Nghiêm Trọng, Loại Phương Tiện (Vehicle Type), Built-up Road trong các năm.



8. Một vài biểu đồ cho tai nạn giao thông



9. Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng TNGT ở các vùng trong năm.



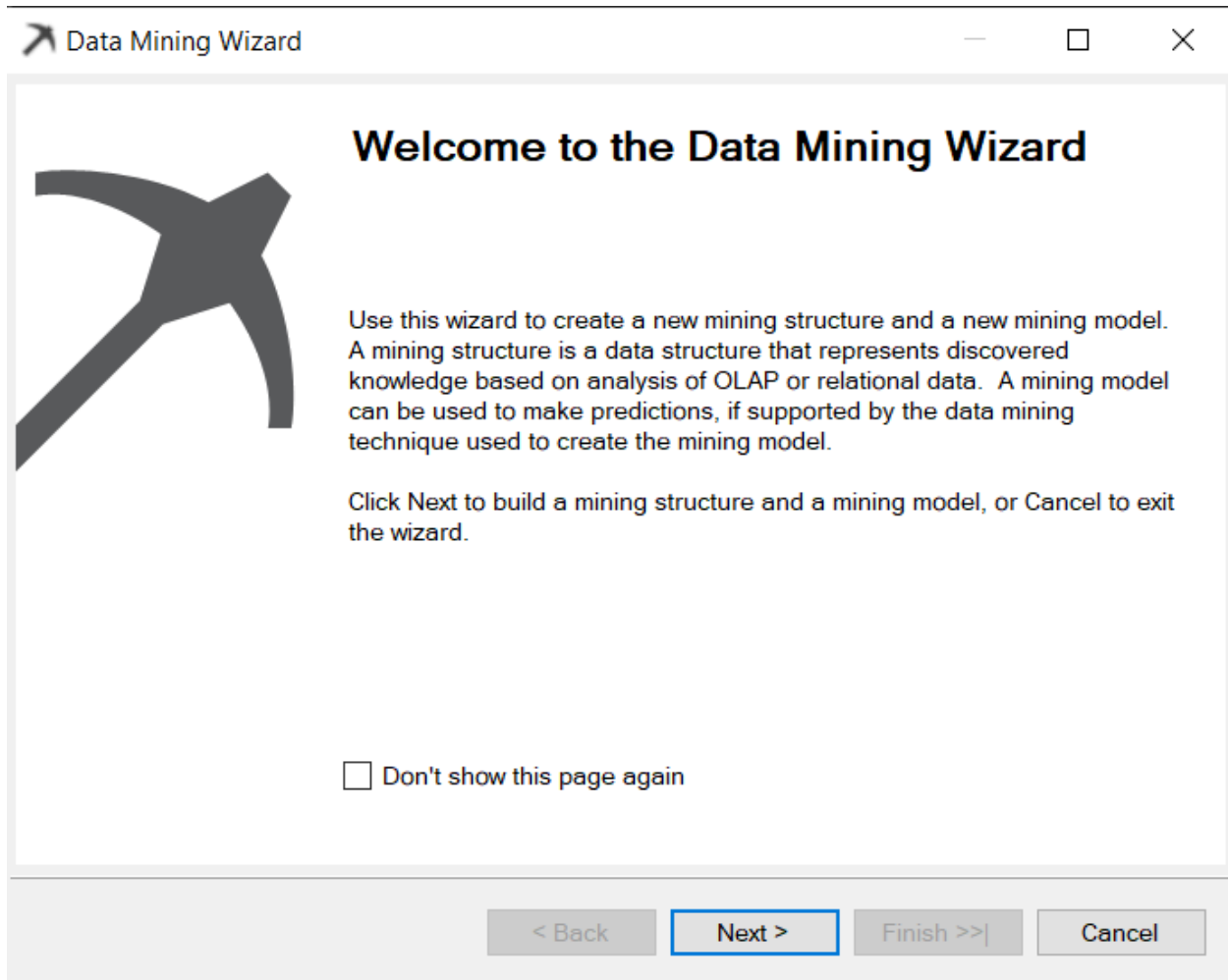
Chương 6: Data Mining

Phân tích cho yêu cầu tỉ lệ xảy ra tai nạn ở vùng thành thị và nông thôn đồng thời dự đoán tỉ lệ tai nạn một người xảy ra tai nạn giao thông bằng Decision Tree.

1. Thực hiện quá trình Data mining

- Tạo Data Source: Tương tự bước tạo data source ở mục Tạo Data Source của Olap Cube (Lấy NDS làm dữ liệu nguồn)
- Tạo Data Source View: Tương tự bước tạo data source ở mục Tạo Data Source của Olap Cube
- Tạo Mining Structure

Chuột phải vào mining structures chọn New Data Structures sau đó Next



Chọn data có sẵn

Data Mining Wizard

Select the Definition Method
Select the method to be used while creating the mining structure definition.

Which method do you use to define the mining structure?

☒ From existing relational database or data warehouse

☐ From existing cube

Description:

This method defines a mining structure based on tables and columns from an existing relational database.

< Back Next > Finish >>| Cancel

Chọn loại Data mining structure

Data Mining Wizard

Create the Data Mining Structure
Specify if mining model should be created and select the most applicable technique.

☒ Create mining structure with a mining model

Which data mining technique do you want to use?

Microsoft Decision Trees

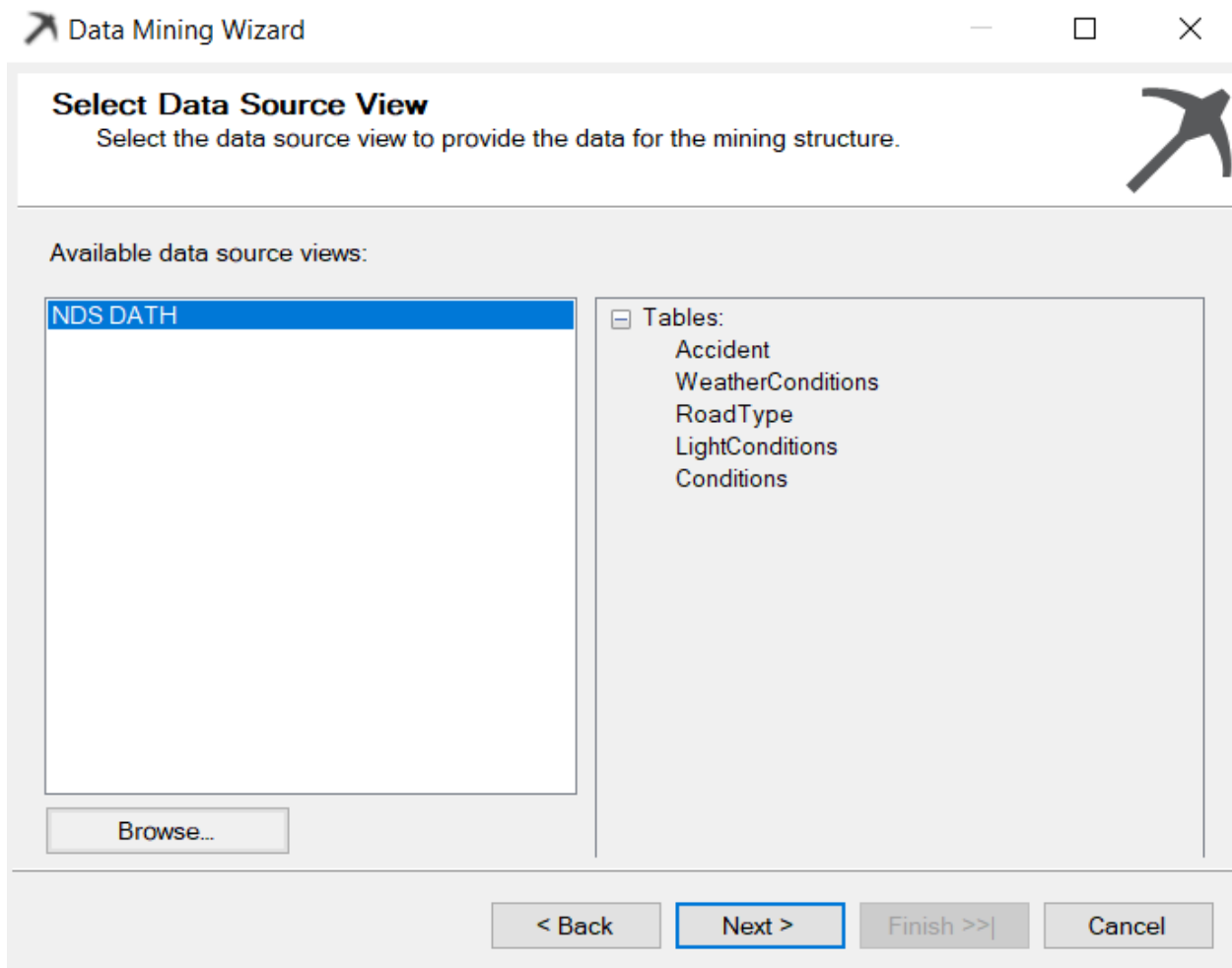
☐ Create mining structure with no models

Description:

The Microsoft Decision Trees algorithm is a classification algorithm that works well for predictive modeling. The algorithm supports the prediction of both discrete and continuous attributes.

< Back Next > Finish >>| Cancel

Connect đến CSDL tạo ở Data Source View




Tiếp theo chọn khóa chính

➤ Data Mining Wizard
— □ ×


Specify Table Types

Specify the type of tables to use for your analysis.



Input tables:

Tables	Case	Nested
Accident	<input type="checkbox"/>	<input type="checkbox"/>
Conditions	<input type="checkbox"/>	<input type="checkbox"/>
LightConditions	<input type="checkbox"/>	<input type="checkbox"/>
RoadType	<input type="checkbox"/>	<input type="checkbox"/>
WeatherConditions	<input type="checkbox"/>	<input type="checkbox"/>

 Select a case table.

< Back
Next >
Finish >>|
Cancel

Chọn các Input đầu vào và biến muốn dự đoán

Data Mining Wizard

Specify the Training Data
Specify the columns used in your analysis.

Mining model structure:


Tables/Columns	Key	Input	Predi...
<input checked="" type="checkbox"/> Accident_Index	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> AccidentSeverity	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Date	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Day_of_Week	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> LightConditions	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> LocalAuthority_District	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> LocalAuthority_Highway	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> LSOA_of_Accident_Location	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> NumberOfCasualties	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> NumberOfVehicles	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> podS	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> PoliceForce	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> RoadSurfaceConditions	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> RoadType	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> Speedlimit	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> Time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> UrbanOrRuralArea	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/> WeatherConditions	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Recommend inputs for currently selected predictable:

Suggest


< Back Next > Finish >> Cancel

Thay đổi kiểu dữ liệu








 Data Mining Wizard

Specify Columns' Content and Data Type

Specify mining structure columns' content and data type.



Mining model structure:

Columns	Content Type	Data Type
 Accident Index	Key	Text
 Light Conditions	Discrete	Text
 Road Surface Conditions	Discrete	Text
 Road Type	Discrete	Text
 Speedlimit	Discrete	Long
 Urban Or Rural Area	Discrete	Text
 Weather Conditions	Discrete	Text

Detect continuous or discrete for numeric columns:

Detect


< Back


Next >

Finish >>|


Cancel


Chia tập train test và số trường hợp lớn nhất cho mẫu testing là 100000

 Data Mining Wizard — □ ×

Create Testing Set 

Specify the number of cases to be reserved for model testing.



Percentage of data for testing: % 

Maximum number of cases in testing data set: 


Description:

Input data will be randomly split into two sets, a training set and a testing set, based on the percentage of data for number of cases in testing data set you provide. The training set is used to create the mining model. The testing model accuracy.

[Percentage of data for testing] specifies percentages of cases reserved for testing set.
[Maximum number of cases in testing data set] limits total number of cases in the testing set.
If both values are specified, both limits are enforced.


 

Đổi tên cho mining structure

 Data Mining Wizard

Completing the Wizard

Completing the Data Mining Wizard by providing a name for the mining structure.





Mining structure name:








Mining model name:

☐ Allow drill through

Preview:

 DecisionFree

 Columns

-  Accident Index
-  Light Conditions
-  Road Surface Conditions
-  Road Type
-  Speedlimit
-  Urban Or Rural Area
-  Weather Conditions

< Back

Next >

Finish

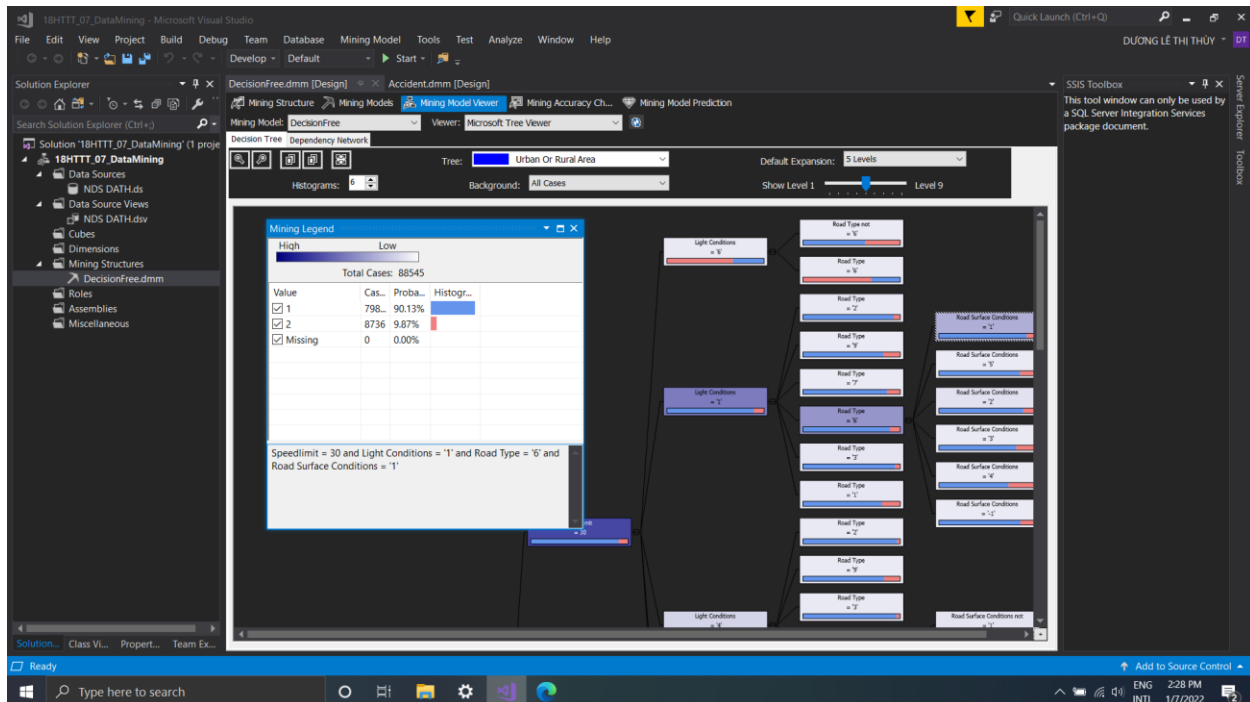
Cancel

Sau khi Finish tiến hành Deploy để chạy Model

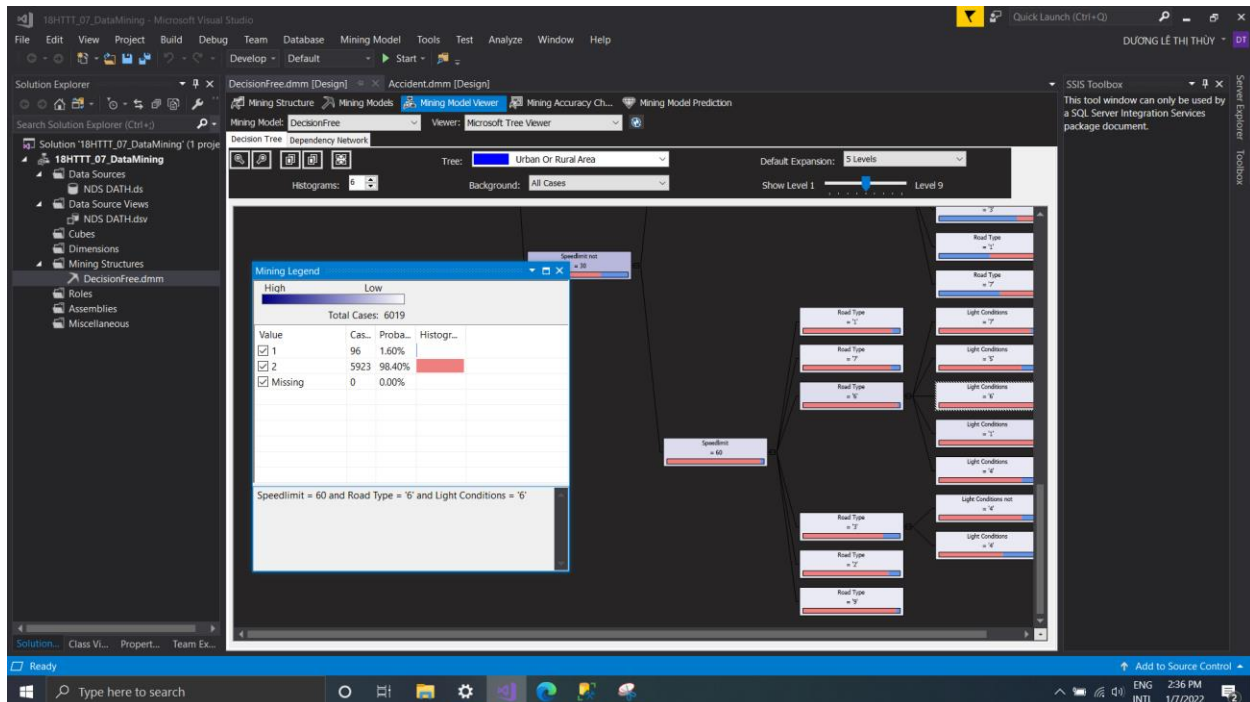
2. Desicion Free

Xem thông tin của 5 lớp

Trong điều kiện tốc độ giới hạn là 30, điều kiện ánh sáng là 1, Road Type là 6, Road Surface là 1 thì tỉ lệ xảy ra tai nạn giao thông của thành thị là 90.13%



Trong điều kiện tốc độ giới hạn là 60, Road Type là 6, Light Condition là 6 tỉ lệ xảy ra tai nạn ở nông thôn là 98.40%



Dependency Network của thuật toán Decision Free cho thấy những thuộc tính ảnh hưởng đến việc xảy ra tai nạn giao thông.

Những thuộc tính có ảnh hưởng giảm dần: Speedlimit -> RoadType -> Light Conditions
-> Road Surface Condition -> Weather Condition

