**Phần 1: GROUP BY**

**Câu 1:** Tổng annual revenue trên từng businesstype và ProductLine trong bảng DimReseller, chỉ trả về những bản ghi có tổng annual revenue lớn hơn 10 triệu

SELECT

BusinessType,

ProductLine,

SUM(AnnualRevenue) AS TotalAnnualRevenue

FROM DimReseller

GROUP BY BusinessType, ProductLine

HAVING SUM(AnnualRevenue) > 10000000

**Câu 2:** Dựa vào bảng FactInternetSales hãy trả về bảng có 3 cột:

* Mã khách hàng
* Ngày gần nhất mà khách hàng mua
* Số lượng đơn khách mua (Tìm hiểu COUNT VÀ COUNT DISTINCT)
* Chỉ về các khách hàng có ngày mua gần nhất từ năm 2012 đổ về sau
* SELECT
* CustomerKey,
* MAX(OrderDate) AS LatestOrderDate,
* COUNT(DISTINCT SalesOrderNumber) AS NumberOfOrders
* FROM FactInternetSales
* WHERE YEAR(OrderDate) >= '2012'
* GROUP BY CustomerKey

**Câu 3:** Doanh số bán hàng online theo từng năm ?

SELECT

YEAR(OrderDate) AS Year,

SUM(SalesAmount) AS TotalSales

FROM FactInternetSales

WHERE OrderDate IS NOT NULL

GROUP BY YEAR(OrderDate)

ORDER BY YEAR(OrderDate)

**Phần 2: Truy vấn lồng với IN**

**Câu 1:** Trả về tất cả các đơn hàng trong bảng InternetSales mà có khách hàng là Nam và sinh sau năm 1970 và mua các sản phẩm màu đen

SELECT

\*

FROM FactInternetSales

WHERE CustomerKey IN(

SELECT

CustomerKey

FROM DimCustomer

WHERE Gender = 'M'

AND YEAR(BirthDate) > '1970'

)

AND ProductKey IN(

SELECT

ProductKey

FROM DimProduct

WHERE Color = 'Black'

)

**Câu 2:** Trả về thông tin của top 5 khách hàng nữ chi tiêu nhiều nhất (Gợi ý: sử dụng truy vấn lồng kết hợp với GROUP BY vào bảng FactInternetSales để lọc các bản ghi trong bảng DimCustomer)

SELECT

\*

FROM DimCustomer

WHERE CustomerKey IN (

SELECT TOP 5

CustomerKey

FROM FactInternetSales

WHERE CustomerKey IN(

SELECT CustomerKey

FROM DimCustomer

WHERE Gender='F')

GROUP BY CustomerKey

ORDER BY SUM(SalesAmount) DESC

)

**Câu 3:** Trả về thông tin của top 5 sản phẩm bán chạy nhất màu đỏ có ListPrice > 500

SELECT

\*

FROM DimProduct

WHERE ProductKey IN (

SELECT TOP 5

ProductKey

FROM FactInternetSales

WHERE Color='Red' AND ListPrice > '500'

GROUP BY ProductKey

ORDER BY SUM(SalesAmount) DESC

)

**Phần 3: Sử dụng truy vấn lồng với CTE**

**Câu 1:** Hãy viết câu truy vấn cho biết các thông tin sau:

* Trung bình một reseller chi tiêu bao nhiêu tiền ?
* Trung bình một reseller mua bao nhiêu đơn hàng ?
* Trung bình một khách hàng online chi tiêu bao nhiêu tiền ?
* Trung bình một khách hàng online mua bao nhiêu đơn ?

**Nâng cao:** hãy thể hiện yêu câu trên thành một bảng duy nhất như dưới đây (tham khảo UNION và UNION ALL)

| **Phân loại** | **Trung bình chi tiêu** | **Trung bình số lượng đơn** |
| --- | --- | --- |
| Reseller |  |  |
| Online |  |  |

WITH ResellerSales AS (

SELECT

ResellerKey,

SUM(SalesAmount) AS TotalSalesRS,

COUNT(SalesOrderNumber) AS TotalOrdersRS

FROM FactResellerSales

WHERE ResellerKey IS NOT NULL

GROUP BY ResellerKey

),

OnlineSales AS (

SELECT

CustomerKey,

SUM(SalesAmount) AS TotalSalesOS,

COUNT(SalesOrderNumber) AS TotalOrdersOS

FROM FactInternetSales

GROUP BY CustomerKey

)

SELECT

'Reseller' AS PhânLoại,

AVG(TotalSalesRS) AS TrungBìnhChiTiêu,

AVG(TotalOrdersRS) AS TrungBìnhSốLượngĐơn

FROM ResellerSales

UNION

SELECT

'Online' AS N'PhânLoại',

AVG(TotalSalesOS) AS TrungBìnhChiTiêu,

AVG(TotalOrdersOS) AS TrungBìnhSốLượngĐơn

FROM OnlineSales

**Câu 2:** Trả về tên và tổng chi tiêu của top 5 khách hàng nữ sinh sau 1970, đồng thời thêm một cột Refund là bằng 10% tổng chi tiêu của các khách hàng đó

WITH TongMua AS(

SELECT

SUM(SalesAmount) 'Total',

CustomerKey

FROM FactInternetSales

GROUP BY CustomerKey)

SELECT TOP 5

FirstName,

MiddleName,

LastName,

Total

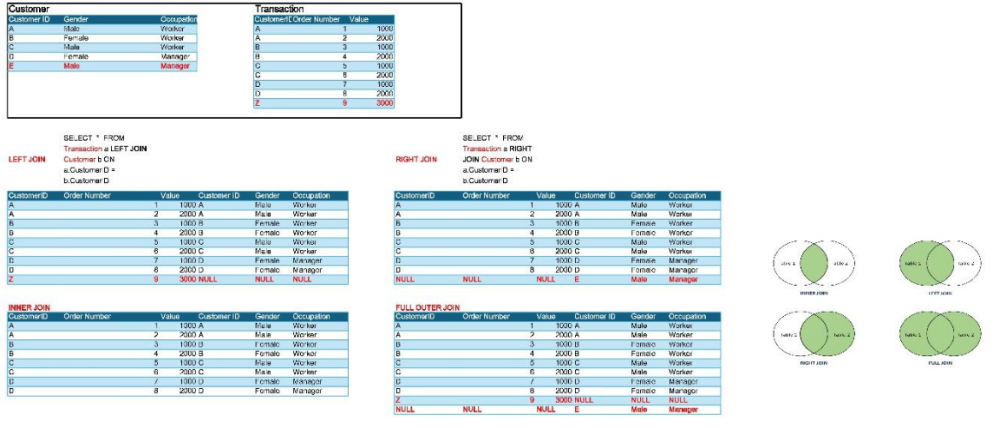
FROM DimCustomer INNER JOIN TongMua ON DimCustomer.CustomerKey = TongMua.CustomerKey

WHERE Gender = 'F' AND YEAR(BirthDate) >= '1970'

ORDER BY Total DESC

**Phần 4:** JOIN và các bài tập tổng hợp (không sử dụng IN trong các bài tập này)

* Tham khảo sơ đồ ERD để biết biết bảng nào join với bảng nào và với cột nào [*https://dataedo.com/samples/html/Data\_warehouse/doc/AdventureWorksDW\_4/modules/Internet\_Sales\_101/module.html*](https://dataedo.com/samples/html/Data_warehouse/doc/AdventureWorksDW_4/modules/Internet_Sales_101/module.html)



**Câu 1:** Trả về bảng kết quả chứa số đơn hàng của reseller, mã reseller, tên chương trình khuyễn mãi được hưởng. Lưu ý các dòng dữ liệu khi trả ra không được trùng nhau (Tham khảo SELECT DISTINCT)

* Thông tin ở bảng FactResellerSales và DimPromotion
* SELECT DISTINCT
* COUNT(SalesAmount) AS NumberOfOrders,
* R.ResellerKey,
* P.PromotionKey
* FROM FactResellerSales AS S
* INNER JOIN DimReseller AS R ON S.ResellerKey = R.ResellerKey
* INNER JOIN DimPromotion AS P ON S.PromotionKey = P.PromotionKey
* GROUP BY R.ResellerKey, P.PromotionKey

**Câu 2:** Trả về bảng kết quả chứa số đơn hàng, orderlinenumber, mã sản phẩm, số lượng mua, màu sắc, size, đơn giá, tên khách hàng, địa chỉ, giới tính

* Thông tin ở các bảng FactInternetSales, DimCustomer, DimProduct

**Câu 3:** Trả về bảng kết quả số lượng đơn hàng online trên mỗi Size sản phẩm

| **Màu sắc sản phẩm** | **Số lượng đơn hàng online** |
| --- | --- |
| SELECT  Product.Size,  Product.Color,  COUNT(SalesAmount) AS NumberOfOrders  FROM FactInternetSales AS Sales  INNER JOIN DimProduct AS Product ON Sales.ProductKey = Product.ProductKey  GROUP BY Product.Size, Product.Color  ORDER BY Product.Size, Product.Color |  |
|  |  |

**Câu 4:** Trung bình số lượng mua trên mỗi Size của các khách hàng sinh sau 1980

| **Size** | **Trung bình số lượng mua** |
| --- | --- |
| SELECT  Product.Size,  AVG(OrderQuantity) AS AverageQuantity  FROM FactInternetSales AS Sales  INNER JOIN DimProduct AS Product ON Sales.ProductKey = Product.ProductKey  INNER JOIN DimCustomer AS Customer ON Sales.CustomerKey = Customer.CustomerKey  WHERE Customer.BirthDate > ‘1980-12-31’  GROUP BY Product.Size  ORDER BY Product.Size |  |
|  |  |

**Câu 5:** Trả về bảng kết quả thể hiện doanh số từ năm 2012 – 2014 theo từng năm và theo từng quốc gia đồng thời thêm một cột Bonus vùng tương đương 10% doanh số

| **Quốc gia** | **Năm** | **Doanh số** | **Bonus vùng** |
| --- | --- | --- | --- |
| SELECT  Country.SalesTerritoryCountry,  YEAR(OrderDate) AS Year,  SUM(SalesAmount) AS SalesAmount,  SUM(SalesAmount) \* 0.1 AS BonusRegion  FROM FactInternetSales AS Sales  INNER JOIN DimSalesTerritory AS Country ON Sales.SalesTerritoryKey = Country.SalesTerritoryKey  WHERE YEAR(OrderDate) BETWEEN 2012 AND 2014  GROUP BY Country.SalesTerritoryCountry, YEAR(OrderDate)  ORDER BY Country.SalesTerritoryCountry, YEAR(OrderDate) |  |  |  |
|  |  |  |  |

**Câu 6:** Trả về mã đơn hàng online năm 2011 bởi khách hàng nữ sinh sau 1980, số orderline, mã khách hàng, mã sản phẩm, số lượng mua, và cột phân loại theo logic sau:

* Màu Black, Blue, Grey, Red được phân loại là DarkColor
* Silver, Silver/Black, White, Yellow được phân loại là BrightColor
* Các màu khác giữ nguyên

SELECT

SalesOrderNumber,

SalesOrderLineNumber,

Customer.CustomerKey,

Product.ProductKey,

OrderQuantity,

CASE

WHEN Product.Color IN ('Black', 'Blue', 'Grey', 'Red') THEN 'DarkColor'

WHEN Product.Color IN ('Silver', 'Silver/Black', 'White', 'Yellow') THEN 'BrightColor'

ELSE Product.Color

END AS ColorCategory

FROM FactInternetSales AS Sales

INNER JOIN DimCustomer AS Customer ON Sales.CustomerKey = Customer.CustomerKey

INNER JOIN DimProduct AS Product ON Sales.ProductKey = Product.ProductKey

WHERE YEAR(OrderDate) = 2011

AND Customer.Gender = 'F'

AND Customer.BirthDate > '1980-12-31'

* ORDER BY SalesOrderNumber

**Câu 7:** Giống câu 6 nhưng hãy tính tổng số lượng mua theo từng phân loại màu sắc

WITH C6 AS (

SELECT

SalesOrderNumber,

SalesOrderLineNumber,

Customer.CustomerKey,

Product.ProductKey,

OrderQuantity,

CASE

WHEN Product.Color IN ('Black', 'Blue', 'Grey', 'Red') THEN 'DarkColor'

WHEN Product.Color IN ('Silver', 'Silver/Black', 'White', 'Yellow') THEN 'BrightColor'

ELSE Product.Color

END AS ColorCategory

FROM FactInternetSales AS Sales

INNER JOIN DimCustomer AS Customer ON Sales.CustomerKey = Customer.CustomerKey

INNER JOIN DimProduct AS Product ON Sales.ProductKey = Product.ProductKey

WHERE YEAR(OrderDate) = 2011

AND Customer.Gender = 'F'

AND Customer.BirthDate > '1980-12-31'

)

SELECT

SUM(OrderQuantity),

ColorCategory

FROM C6

GROUP BY ColorCategory