Assignment #5
CSCI 3287
Brandon Glandt
4/5/2020

1) Question 1

```
1  use information_schema;
2  SELECT TABLE_NAME, TABLE_ROWS FROM TABLES WHERE TABLE_SCHEMA='aw';
3
```

100% 1:3

Result Grid | Filter Rows: Search | Export:

TABLE_NAME	TABLE_ROWS
DimAccount	99
DimCurrency	0
DimCustomer	18356
DimDepartmentGroup	7
DimEmployee	296
DimGeography	655
DimOrganization	14
DimProduct	158
DimProductCategory	4
DimProductSubcategory	37
DimPromotion	16
DimReseller	701
DimSalesReason	10
DimSalesTerritory	11
DimScenario	3
DimTime	1158
FactCurrencyRate	0
FactFinance	38480
FactInternetSales	59800

2) Question 2

```
1  use information_schema;
2  SELECT DISTINCT TABLE_NAME, COLUMN_NAME, column_key FROM COLUMNS WHERE COLUMN_KEY='PRI';
3
4
```

100% 88:2

Result Grid Filter Rows: Search Export:

TABLE_NAME	COLUMN_NAME	COLUMN_KEY
DimAccount	AccountKey	PRI
DimCustomer	CustomerKey	PRI
DimDepartmentGroup	DepartmentGroupKey	PRI
DimEmployee	EmployeeKey	PRI
DimGeography	GeographyKey	PRI
DimOrganization	OrganizationKey	PRI
DimProductSubcategory	ProductSubcategoryKey	PRI
DimReseller	ResellerKey	PRI
DimProduct	ProductKey	PRI
FactInternetSales	SalesOrderLineNumber	PRI
DimCurrency	CurrencyKey	PRI
DimProductCategory	ProductCategoryKey	PRI
DimPromotion	PromotionKey	PRI
DimSalesReason	SalesReasonKey	PRI
DimSalesTerritory	SalesTerritoryKey	PRI
DimScenario	ScenarioKey	PRI
DimTime	TimeKey	PRI
FactInternetSales	SalesOrderNumber	PRI

3) Question 3

The standard table naming convention that the AdventureWorksDW database designers use to differentiate dimension tables from fact tables in this star schema data warehouse are when they use dimension tables they start the name of the table with Dim and when they use fact tables they start the name of the table with Fact.

4) Question 4

I think the purpose of the recursive relationship on DimEmployee is because the primary key in that table is EmployeeKey and there are two multiples rows as well which are ParentEmployeeKey and SalesTerritoryKey. The two multiples of rows are the reason DimEmployee is recursive.

5) Question 5

The three types of models of bikes sold by AdventureWorks are Road, Mountain and Sport.

6) Question 6

2001:

```
use aw;

SELECT COUNT(product.Color) AS Amount, product.Color FROM (DimProduct product, DimTime time)
INNER JOIN FactInternetSales Sales ON (time.TimeKey = Sales.OrderDateKey)
WHERE time.TimeKey = Sales.OrderDateKey AND time.CalendarYear = '2001' AND Sales.ProductKey = product.ProductKey AND
product.EnglishDescription LIKE '%bike%'
GROUP BY product.Color ORDER BY Amount DESC;
```

100% 24:7

Result Grid Filter Rows: Search Export:

Amount	Color
775	Red
154	Black
84	Silver

The most popular color in 2001 was red.

2002:

```
use aw;

SELECT COUNT(product.Color) AS Amount, product.Color FROM (DimProduct product, DimTime time)
INNER JOIN FactInternetSales Sales ON (time.TimeKey = Sales.OrderDateKey)
WHERE time.TimeKey = Sales.OrderDateKey AND time.CalendarYear = '2002' AND Sales.ProductKey = product.ProductKey AND
product.EnglishDescription LIKE '%bike%'
GROUP BY product.Color ORDER BY Amount DESC;
```

100% 45:7

Result Grid Filter Rows: Search Export:

Amount	Color
1027	Red
320	Black
146	Yellow
101	Silver

The most popular color in 2002 was red.

2003:

```
use aw;

SELECT COUNT(product.Color) AS Amount, product.Color FROM (DimProduct product, DimTime time)
INNER JOIN FactInternetSales Sales ON (time.TimeKey = Sales.OrderDateKey)
WHERE time.TimeKey = Sales.OrderDateKey AND time.CalendarYear = '2003' AND Sales.ProductKey = product.ProductKey AND
product.EnglishDescription LIKE '%bike%'
GROUP BY product.Color ORDER BY Amount DESC;
```

100% 70:5

Result Grid Filter Rows: Search Export:

Amount	Color
1139	NA
1049	Yellow
770	Black
276	Blue
207	Silver
172	Red

The most popular color in 2003 was yellow.

2004:

```
use aw;

SELECT COUNT(product.Color) AS Amount, product.Color FROM (DimProduct product, DimTime time)
INNER JOIN FactInternetSales Sales ON (time.TimeKey = Sales.OrderDateKey)
WHERE time.TimeKey = Sales.OrderDateKey AND time.CalendarYear = '2004' AND Sales.ProductKey = product.ProductKey AND
product.EnglishDescription LIKE '%bike%'
GROUP BY product.Color ORDER BY Amount DESC;
```

100% 45:7

Result Grid Filter Rows: Search Export:

Amount	Color
1559	NA
1387	Yellow
839	Black
373	Blue
336	Silver

The most popular color in 2004 was yellow.

7) Question 7

2001:

```
use aw;

SELECT COUNT(Sales.UnitPrice) AS SalesVolume, geo.StateProvinceName
FROM DimGeography geo
INNER JOIN DimSalesTerritory territ ON geo.SalesTerritoryKey = territ.SalesTerritoryKey
INNER JOIN FactInternetSales Sales ON territ.SalesTerritoryKey = Sales.SalesTerritoryKey
INNER JOIN DimProduct product ON Sales.ProductKey = product.ProductKey
INNER JOIN DimTime time ON Sales.OrderDateKey = time.TimeKey
WHERE product.EnglishDescription LIKE '%bike%' AND time.CalendarYear = '2001' GROUP BY geo.StateProvinceName
ORDER BY SalesVolume DESC
LIMIT 4;
```

100% 43:7

Result Grid Filter Rows: Search Export: Fetch rows:

SalesVolume	StateProvinceNa...
19928	California
7486	New South Wales
5418	Washington
5088	England

The top four that showed the highest sales volume for 2001 are California, NSW, Washington and England.

2002:

```
1  use aw;
2
3  SELECT COUNT(Sales.UnitPrice) AS SalesVolume, geo.StateProvinceName
4  FROM DimGeography geo
5  INNER JOIN DimSalesTerritory territ ON geo.SalesTerritoryKey = territ.SalesTerritoryKey
6  INNER JOIN FactInternetSales Sales ON territ.SalesTerritoryKey = Sales.SalesTerritoryKey
7  INNER JOIN DimProduct product ON Sales.ProductKey = product.ProductKey
8  INNER JOIN DimTime time ON Sales.OrderDateKey = time.TimeKey
9  WHERE product.EnglishDescription LIKE '%bike%' AND time.CalendarYear = '2002' GROUP BY geo.StateProvinceName
10 ORDER BY SalesVolume DESC
11 LIMIT 4;
```

100% 77:9

Result Grid | Filter Rows: Search | Export: | Fetch rows:

SalesVolume	StateProvinceNa...
33182	California
9291	New South Wales
8610	Washington
8119	Texas

The top four that showed the highest sales volume for 2002 are California, NSW, Washington and Texas.

2003:

```
1  use aw;
2
3  SELECT COUNT(Sales.UnitPrice) AS SalesVolume, geo.StateProvinceName
4  FROM DimGeography geo
5  INNER JOIN DimSalesTerritory territ ON geo.SalesTerritoryKey = territ.SalesTerritoryKey
6  INNER JOIN FactInternetSales Sales ON territ.SalesTerritoryKey = Sales.SalesTerritoryKey
7  INNER JOIN DimProduct product ON Sales.ProductKey = product.ProductKey
8  INNER JOIN DimTime time ON Sales.OrderDateKey = time.TimeKey
9  WHERE product.EnglishDescription LIKE '%bike%' AND time.CalendarYear = '2003' GROUP BY geo.StateProvinceName
10 ORDER BY SalesVolume DESC
11 LIMIT 4;
```

100% 77:9

Result Grid | Filter Rows: Search | Export: | Fetch rows:

SalesVolume	StateProvinceNa...
72850	California
23940	Washington
22896	England
17825	Texas

The top four that showed the highest sales volume for 2003 are California, Washington, England and Texas.

2004:

```sql
use aw;

SELECT COUNT(Sales.UnitPrice) AS SalesVolume, geo.StateProvinceName
FROM DimGeography geo
INNER JOIN DimSalesTerritory territ ON geo.SalesTerritoryKey = territ.SalesTerritoryKey
INNER JOIN FactInternetSales Sales ON territ.SalesTerritoryKey = Sales.SalesTerritoryKey
INNER JOIN DimProduct product ON Sales.ProductKey = product.ProductKey
INNER JOIN DimTime time ON Sales.OrderDateKey = time.TimeKey
WHERE product.EnglishDescription LIKE '%bike%' AND time.CalendarYear = '2004' GROUP BY geo.StateProvinceName
ORDER BY SalesVolume DESC
LIMIT 4;
```

100% 77:9

Result Grid Filter Rows: Search Export: Fetch rows:

SalesVolume	StateProvinceNa...
95786	California
28014	Washington
25175	England
23437	Texas

The top four that showed the highest sales volume for 2004 are California, Washington, England and Texas.

8) Question 8

```
1  use aw;
2
3  SELECT DISTINCT (Sales.UnitPrice - Sales.ProductStandardCost) AS Margin,
4  product.ModelName AS Model, product.EnglishDescription
5  FROM FactInternetSales Sales
6  INNER JOIN DimTime time ON Sales.OrderDateKey = time.TimeKey
7  INNER JOIN DimProduct product ON product.ProductKey = Sales.ProductKey
8  WHERE time.CalendarYear = '2002'
9  ORDER BY Margin;
```

100% 1:8

Result Grid Filter Rows: Search Export:

Margin	Model	EnglishDescription
286.00	Road-650	Value-priced bike with many features of our top-…
296.00	Road-650	Value-priced bike with many features of our top-…
394.00	Road-550-W	Same technology as all of our Road series bike…
861.00	Road-250	Alluminum-alloy frame provides a light, stiff ride,…
924.00	Road-250	Alluminum-alloy frame provides a light, stiff ride,…
943.00	Mountain-200	Serious back-country riding. Perfect for all level…
953.00	Mountain-200	Serious back-country riding. Perfect for all level…
1407.00	Road-150	This bike is ridden by race winners. Developed…
1477.00	Mountain-100	Top-of-the-line competition mountain bike. Perfo…
1488.00	Mountain-100	Top-of-the-line competition mountain bike. Perfo…

The model of bike that yielded the highest margin for AdventureWorks is at margin 1488 as the model Mountain-100.