Data Warehouse (TP3)

Fonction analytique_ Using SQL Server

ANNASSIRI Fatima Zahra

BISD(M2)2020

Exercice 1:

La création des tables Emp et Dept du compte Scott dans la base de données SQL :

```
□CREATE TABLE dept(deptno INT NOT NULL, dname VARCHAR(14), loc VARCHAR(13));

  INSERT INTO dept VALUES(10, 'ACCOUNTING', 'NEW YORK');
  INSERT INTO dept VALUES(20, 'RESEARCH', 'DALLAS');
INSERT INTO dept VALUES(30, 'SALES', 'CHICAGO');
INSERT INTO dept VALUES(40, 'OPERATIONS', 'BOSTON');
  SELECT * FROM dept;
```

Les résultats donnés par :



Création de la table Emp :

```
□CREATE TABLE emp (empno INT NOT NULL, ename VARCHAR(10),
             job VARCHAR(9), mgr INT, hiredate DATE, sal decimal, comm INT, deptno INT);
            INSERT INTO emp VALUES(7369, 'SMITH', 'CLERK', 7902, '1980-12-17', 800, NULL, 20);
INSERT INTO emp VALUES(7499, 'ALLEN', 'SALESMAN', 7698, '1981-02-20', 1600, 300, 30);
INSERT INTO emp VALUES(7521, 'WARD', 'SALESMAN', 7698, '1981-02-22', 1250, 500, 30);
INSERT INTO emp VALUES(7566, 'JONES', 'MANAGER', 7839, '1981-04-02', 2975, NULL, 20);
INSERT INTO emp VALUES(7654, 'MARTIN', 'SALESMAN', 7698, '1981-09-28', 1250, 1400, 30);
INSERT INTO emp VALUES(7698, 'BLAKE', 'MANAGER', 7839, '1981-09-28', 1250, 1400, 30);
INSERT INTO emp VALUES(7782, 'CLARK', 'MANAGER', 7839, '1981-06-09', 2450, NULL, 30);
INSERT INTO emp VALUES(7788, 'SCOTT', 'ANALYST', 7566, '1981-04-19', 3000, NULL, 20);
INSERT INTO emp VALUES(7849, 'KING', 'PRESIDENT', NULL, '1981-11-17', 5000, NULL, 10);
INSERT INTO emp VALUES(7849, 'TURNER', 'SALESMAN', 7698, '1981-09-08', 1500, NULL, 30);
INSERT INTO emp VALUES(7876, 'ADAMS', 'CLERK', 7788, '1987-05-23', 1100, NULL, 20);
INSERT INTO emp VALUES(7900, 'JAMES', 'CLERK', 77698, '1981-12-03', 950, NULL, 30);
100 % +

        ename
        job
        mgr
        hiredate
        sal
        comm
        dep

        SMITH
        CLERK
        7902
        1980-12-17
        800
        NULL
        20

                                                                                                                                                            comm deptno
               7369
                7499 ALLEN SALESMAN 7698 1981-02-20 1600 300 30
  3
               7521 WARD SALESMAN 7698 1981-02-22 1250 500
  4
                7566
                                   JONES
                                                          MANAGER 7839 1981-04-02 2975 NULL 20

        7654
        MARTIN
        SALESMAN
        7698
        1981-09-28
        1250

        7698
        BLAKE
        MANAGER
        7839
        1981-05-01
        2850

                                                                                                                                                               1400
  6
                                                                                                                                                               NULL
                7782 CLARK MANAGER 7839 1981-06-09 2450 NULL 10
```

Question 1:

10

11 12 13

14

7788 SCOTT ANALYST 7566 1981-04-19 3000 NULL 20 7839 KING PRESIDENT NULL 1981-11-17 5000 NULL 10 7844 TURNER SALESMAN 7698 1981-09-08 1500 NULL 30

 7876
 ADAMS
 CLERK
 7788
 1987-05-23
 1100
 NULL
 20

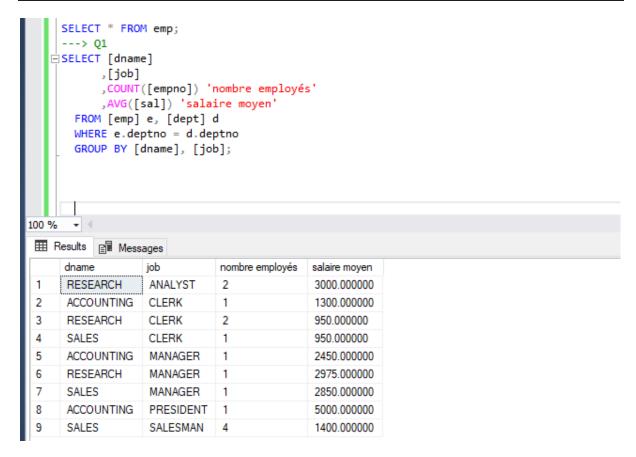
 7900
 JAMES
 CLERK
 7698
 1981-12-03
 950
 NULL
 30

 7902
 FORD
 ANALYST
 7566
 1981-12-03
 3000
 NULL
 20

7934 MILLER CLERK 7782 1982-01-23 1300 NULL 10

Donner le nombre d'employés et le salaire moyen par département et par emploi :

```
SELECT [dname]
    ,[job]
    ,COUNT([empno]) 'nombre employés'
    ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY [dname], [job];
```



Question 2:

Donner le nombre d'employés et le salaire moyen par département et par emploi :

- Proposition n°1: SQL du base

```
SELECT [dname]
    ,[job]
    ,COUNT([empno]) 'nombre employés'
    ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY [dname], [job]
UNION ALL
SELECT [dname]
    ,NULL
    ,COUNT([empno]) 'nombre employés'
    ,AVG([sal]) 'salaire moyen'
```

```
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
 WHERE e.deptno = d.deptno
 GROUP BY [dname];
   □SELECT [dname]
          ,[job]
           ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
      FROM [emp] e, [dept] d
      WHERE e.deptno = d.deptno
      GROUP BY [dname], [job]
    UNION ALL
      SELECT [dname]
          ,NULL
          ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
      FROM [emp] e, [dept] d
      WHERE e.deptno = d.deptno
      GROUP BY [dname];
100 %
      + 4
Results Messages
     dname
                            nombre employés salaire moyen
    RESEARCH
                ANALYST
                                           3000.000000
 2
     ACCOUNTING CLERK
                                           1300.000000
                             1
 3
     RESEARCH
                  CLERK
                             2
                                           950.000000
 4
     SALES
                  CLERK
                                           950.000000
 5
     ACCOUNTING MANAGER
                                           2450.000000
                             1
 6
     RESEARCH
                MANAGER
                                           2975.000000
 7
                  MANAGER
                                           2850.000000
 8
     ACCOUNTING PRESIDENT 1
                                           5000.000000
9
                  SALESMAN
     SALES
                                           1400.000000
 10
     ACCOUNTING NULL
                             3
                                           2916.666666
     RESEARCH
                             5
 11
                  NULL
                                           2175.000000
     SALES
                  NULL
                             6
                                           1566.666666
 12
```

- Proposition n°2: unsing Grouping SET

```
SELECT [dname]
    ,[job]
    ,COUNT([empno]) 'nombre employ�s'
    ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY
GROUPING SETS (
    ([dname], [job]),
    ([dname])
);
```

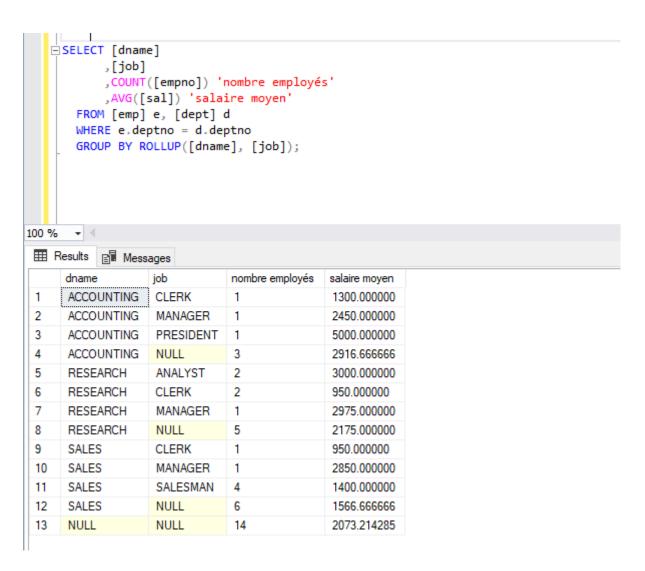
```
,[job]
           ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
       FROM [emp] e, [dept] d
       WHERE e.deptno = d.deptno
       GROUP BY
         GROUPING SETS (
             ([dname], [job]),
             ([dname])
         );
     -
100 %
Results Messages
     dname
                              nombre employés
                                             salaire moyen
     ACCOUNTING
                  CLERK
                                             1300.000000
2
     ACCOUNTING
                   MANAGER
                                             2450.000000
3
     ACCOUNTING
                   PRESIDENT
                                             5000.000000
                              1
4
     ACCOUNTING NULL
                              3
                                             2916.666666
5
     RESEARCH
                   ANALYST
                              2
                                             3000.000000
6
     RESEARCH
                   CLERK
                              2
                                             950.000000
 7
     RESEARCH
                   MANAGER
                               1
                                             2975.000000
     RESEARCH
                   NULL
                              5
                                             2175.000000
8
9
     SALES
                   CLERK
                                             950.000000
                               1
     SALES
 10
                   MANAGER
                               1
                                             2850.000000
 11
     SALES
                   SALESMAN
                              4
                                             1400.000000
                              6
     SALES
                   NULL
                                             1566.666666
 12
```

Proposition n°3: using Group BY ROLLUP OR ROLLUP

```
SELECT [dname]
    ,[job]
    ,COUNT([empno]) 'nombre employés'
    ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY ROLLUP([dname], [job]);
```

USING ROLLUP

```
SELECT [dname]
    ,[job]
    ,COUNT([empno]) 'nombre employ◆s'
    ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY [dname], ROLLUP ([job]);
```



Question 3:

Mêmes questions qu'en 2, mais utiliser les fonctions CASE WHEN...THEN... ELSE... END et GROUPING pour améliorer l'affichage :

```
SELECT GROUPING([dname])
    ,GROUPING([job])
    ,[dname]
    ,[job]
    ,COUNT([empno]) 'nombre employ�s'
     ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY [dname], ROLLUP ([job]);

SELECT CASE
    WHEN GROUPING([dname]) = 0 THEN [dname]
    ELSE 'ALL'
END 'dname'
    ,CASE
```

```
WHEN GROUPING([job]) = 0 THEN [job]

ELSE 'ALL'

END 'job'

,COUNT([empno]) 'nombre employés'

,AVG([sal]) 'salaire moyen'

FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d

WHERE e.deptno = d.deptno

GROUP BY [dname], ROLLUP ([job]);

SELECT [dname]

,CASE

WHEN GROUPING([job]) = 0 THEN [job]

ELSE 'ALL'

END 'job'

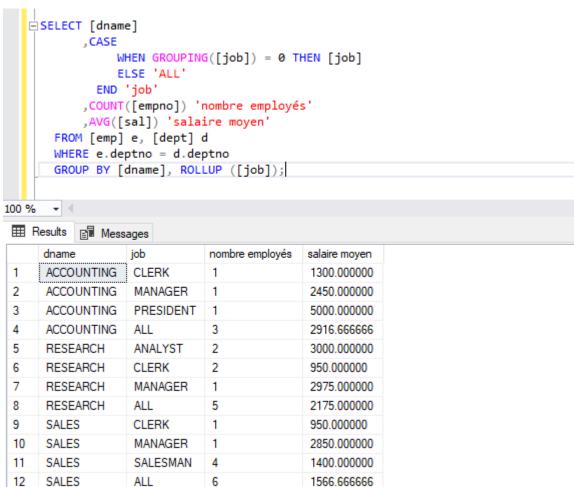
,COUNT([empno]) 'nombre employés'

,AVG([sal]) 'salaire moyen'

FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d

WHERE e.deptno = d.deptno

GROUP BY [dname], ROLLUP ([job]);
```



Question 4:

Mêmes questions qu'en 3 en ajoutant l'année d'embauche comme dimension :

```
SELECT [dname]
    ,CASE
      WHEN GROUPING([job]) = 0 THEN [job]
      ELSE 'ALL'
    END 'job'
      ,[hiredate]
    ,COUNT([empno]) 'nombre employ◆s'
      ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY [dname], ROLLUP ([job], [hiredate]);
```



Question 5:

Même question que 3 mais en ajoutant tous les sous-totaux. Proposer une solution en SQL de base et une solution utilisant la clause CUBE.

Solution de base SQL :

```
SELECT [dname]
      ,[job]
    ,COUNT([empno]) 'nombre employ♦s'
      ,AVG([sal]) 'salaire moyen'
  FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
 WHERE e.deptno = d.deptno
 GROUP BY [dname], [job]
 SELECT [dname]
      ,'NONE'
    ,COUNT([empno]) 'nombre employés'
      ,AVG([sal]) 'salaire moyen'
 FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
 WHERE e.deptno = d.deptno
 GROUP BY [dname]
  SELECT 'NONE'
      ,[job]
    ,COUNT([empno]) 'nombre employés'
      ,AVG([sal]) 'salaire moyen'
 FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
 WHERE e.deptno = d.deptno
 GROUP BY [job]
UNION ALL
 SELECT 'NONE'
      ,'NONE'
    ,COUNT([empno]) 'nombre employés'
      ,AVG([sal]) 'salaire moyen'
 FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
 WHERE e.deptno = d.deptno;
```

```
,[job]
           ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
       FROM [emp] e, [dept] d WHERE e.deptno = d.deptno
      GROUP BY [dname], [job]
     UNION ALL
       SELECT [dname]
          ,'NONE'
          ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
       FROM [emp] e, [dept] d WHERE e.deptno = d.deptno
       GROUP BY [dname]
     UNION ALL
       SELECT 'NONE'
           ,[job]
           ,COUNT([empno]) 'nombre employés'
          ,AVG([sal]) 'salaire moyen'
       FROM [emp] e, [dept] d WHERE e.deptno = d.deptno
      GROUP BY [job]
     UNION ALL
       SELECT 'NONE'
          ,'NONE'
          ,COUNT([empno]) 'nombre employés'
           ,AVG([sal]) 'salaire moyen'
      FROM [emp] e, [dept] d WHERE e.deptno = d.deptno;
100 % -
dname
                            nombre employés salaire moyen
     RESEARCH
                 ANALYST
 1
                                           3000.000000
     ACCOUNTING CLERK
                             1
                                           1300.000000
 3
     RESEARCH
                  CLERK
                             2
                                           950.000000
 4
     SALES
                  CLERK
                             1
                                           950.000000
 5
     ACCOUNTING MANAGER
                                           2450.000000
```

- Solution avec la clause CUBE :

```
SELECT CASE
    WHEN GROUPING([dname]) = 0 THEN [dname]
    ELSE 'NONE'

END 'dname'
    ,CASE
    WHEN GROUPING([job]) = 0 THEN [job]
    ELSE 'NONE'

END 'job'
    ,COUNT([empno]) 'nombre employés'
    ,AVG([sal]) 'salaire moyen'

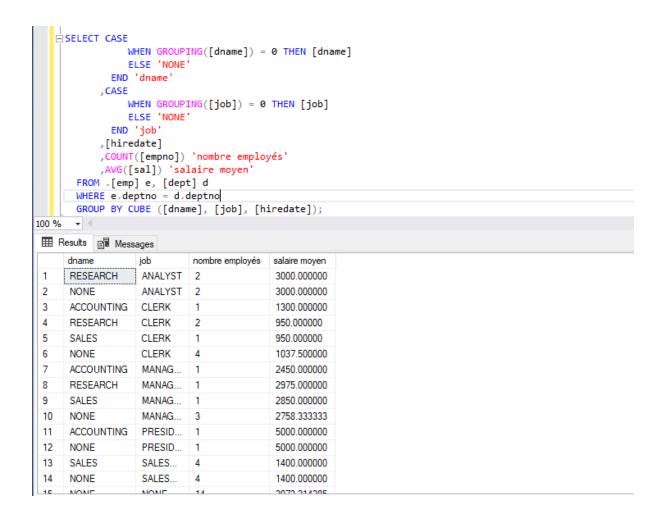
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY CUBE ([dname], [job]);
```



Question 6:

Même question que 5, mais en ajoutant l'année d'embauche comme dimension :

```
SELECT CASE
    WHEN GROUPING([dname]) = 0 THEN [dname]
    ELSE 'NONE'
END 'dname'
    ,CASE
    WHEN GROUPING([job]) = 0 THEN [job]
    ELSE 'NONE'
END 'job'
    ,[hiredate]
    ,COUNT([empno]) 'nombre employés'
     ,AVG([sal]) 'salaire moyen'
FROM [tp2-fa].[dbo].[emp] e, [tp2-fa].[dbo].[dept] d
WHERE e.deptno = d.deptno
GROUP BY CUBE ([dname], [job], [hiredate]);
```



Exercice 2:

La création et le remplissages des tables : Chicago, Toronto et Vancouver.

```
Le resultat donné par :
```

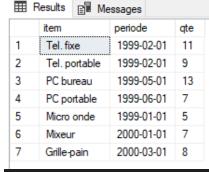
```
SELECT * FROM chicago;
```



```
CREATE TABLE toronto(item VARCHAR(15), periode DATE, qte INTEGER);
INSERT INTO toronto VALUES('Tel. fixe', '1999-02-01', 11);
INSERT INTO toronto VALUES('Tel. portable', '1999-02-01', 9);
INSERT INTO toronto VALUES('PC bureau', '1999-05-01', 13);
INSERT INTO toronto VALUES('PC portable', '1999-06-01', 7);
INSERT INTO toronto VALUES('Micro onde', '1999-01-01', 5);
INSERT INTO toronto VALUES('Mixeur', '2000-01-01', 7);
INSERT INTO toronto VALUES('Grille-pain', '2000-03-01', 8);
```

Le resultat donné par :

SELECT * FROM toronto;



```
CREATE TABLE vancouver(item VARCHAR(15), periode DATE, qte INTEGER);
INSERT INTO vancouver VALUES('Tel. fixe', '1999-10-01', 15);
INSERT INTO vancouver VALUES('Tel. portable', '1999-12-01', 5);
INSERT INTO vancouver VALUES('PC bureau', '1999-06-01', 2);
INSERT INTO vancouver VALUES('PC portable', '1999-07-01', 8);
INSERT INTO vancouver VALUES('Micro onde', '1999-02-01', 10);
INSERT INTO vancouver VALUES('Mixeur', '2000-02-01', 5);
INSERT INTO vancouver VALUES('Grille-pain', '2000-01-01', 5);
```

Le résultat est donné par :

```
SELECT * FROM vancouver;
```



Question 1:

Donner le nombre Faire une requête pour obtenir une table Ventes ayant pour attribut Ville, Période, Item, Qte, regroupant les informations des trois villes. et le salaire moyen par département et par emploi :

```
SELECT * INTO ventes FROM

(SELECT 'chicago' as 'ville', *

FROM [tp2-fa].[dbo].[chicago]

UNION ALL

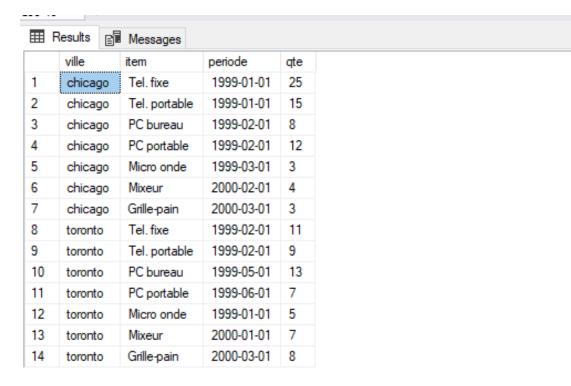
SELECT 'toronto' as 'ville', *

FROM [tp2-fa].[dbo].[toronto]

UNION ALL

SELECT 'vancouver' as 'ville', *

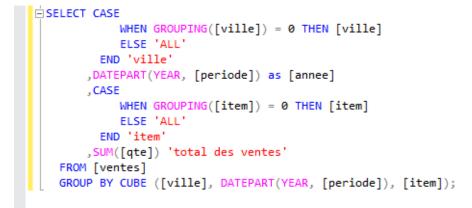
FROM [tp2-fa].[dbo].[vancouver]) AS ventes;
```

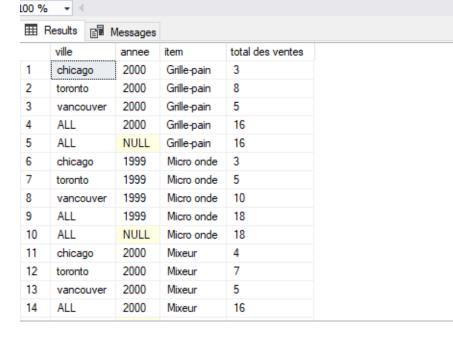


Question 2:

Construire le cube de données du total des ventes selon les dimensions ville, année et item :

```
SELECT CASE
     WHEN GROUPING([ville]) = 0 THEN [ville]
     ELSE 'ALL'
END 'ville'
, DATEPART(YEAR, [periode]) as [annee]
,CASE
     WHEN GROUPING([item]) = 0 THEN [item]
     ELSE 'ALL'
END 'item'
,SUM([qte]) 'total des ventes'
FROM [tp2-fa].[dbo].[ventes]
GROUP BY CUBE ([ville], DATEPART(YEAR, [periode]), [item]);
```

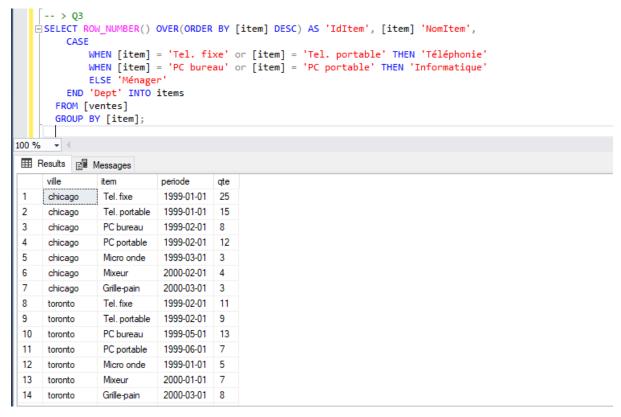




Question 3:

On souhaite maintenant regrouper les objets par département (Téléphonie, Informatique et Ménager). Créer la table Items ayant pour attributs IdItem, NomItem, Dept et la peupler à partir de Ventes :

```
SELECT ROW_NUMBER() OVER(ORDER BY [item] DESC) AS 'IdItem', [item] 'NomItem',
    CASE
        WHEN [item] = 'Tel. fixe' or [item] = 'Tel. portable' THEN 'Téléphonie'
        WHEN [item] = 'PC bureau' or [item] = 'PC portable' THEN 'Informatique'
        ELSE 'Ménager'
    END 'Dept' INTO items
FROM [tp2-fa].[dbo].[ventes]
GROUP BY [item];
```



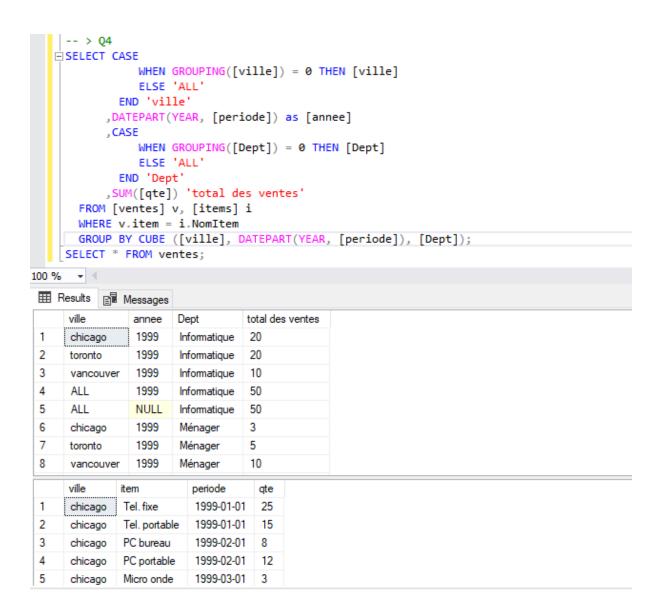
Question 4:

Construire le cube de données du total des ventes selon les dimensions ville, année et département.

```
SELECT CASE
    WHEN GROUPING([ville]) = 0 THEN [ville]
    ELSE 'ALL'

END 'ville'
,DATEPART(YEAR, [periode]) as [annee]
,CASE
    WHEN GROUPING([Dept]) = 0 THEN [Dept]
    ELSE 'ALL'
    END 'Dept'
,SUM([qte]) 'total des ventes'

FROM [tp2-fa].[dbo].[ventes] v, [tp2-fa].[dbo].[items] i
WHERE v.item = i.NomItem
GROUP BY CUBE ([ville], DATEPART(YEAR, [periode]), [Dept]);
```



Question 5:

Donner le total des ventes par produit et par département selon les dimensions ville et année :

Le résultat est donnés par :

```
□SELECT [item], [Dept],
              CASE
                 WHEN GROUPING([ville]) = 0 THEN [ville]
                 ELSE 'ALL'
              END 'ville'
            ,DATEPART(YEAR, [periode]) as [annee]
            ,SUM([qte]) 'total des ventes'
       FROM [ventes] v, [items] i
       WHERE v.item = i.NomItem
       GROUP BY [item], [Dept], CUBE ([ville], DATEPART(YEAR, [periode]));
     SELECT * FROM ventes;
100 % - 4
Results Messages
                 Dept
                         ville
                                           total des ventes
      item
                                    annee
     Grille-pain
                 Ménager chicago
                                    2000
                                            3
 2
     Grille-pain
                                    2000
                                           8
                 Ménager toronto
 3
                                           5
      Grille-pain
                                    2000
                 Ménager
                         vancouver
 4
                         ALL
                                    2000
                                           16
      Grille-pain
                 Ménager
 5
                          ALL
                                    NULL
                                           16
      Grille-pain
                 Ménager
 6
                                            3
      Micro onde
                 Ménager
                         chicago
                                    1999
 7
                                    1999
                                            5
      Micro onde
               Ménager
                         toronto
                         vancouver 1999
 8
                                           10
      Micro onde Ménager
              item
                          periode
                                     qte
 1
      chicago
              Tel. fixe
                          1999-01-01
                                     25
 2
      chicago Tel. portable 1999-01-01 15
 3
      chicago PC bureau
                          1999-02-01 8
 4
      chicago PC portable 1999-02-01 12
 5
      chicago Micro onde 1999-03-01 3
```

■ The end --