Relationship between country and revenue

First I identify the the tables that might contained the data that was relevant for revenue and countries.
The tables that I found to be relevant was **“SalesOrderHeader”**, **“SalesOrderDetail”** and **“SalesTerritory”**.
First I extracted the **TerritoryID** and the **LineTotal** from **SalesOrderHeader** and **SalesOrderDetail** by joining the table tables on **SalesOrderID** using the query

```
SELECT  TerritoryID, LineTotal
FROM Sales.SalesOrderHeader
JOIN Sales.SalesOrderDetail
ON Sales.SalesOrderHeader.SalesOrderID = Sales.SalesOrderDetail.SalesOrderID
```

	TerritoryID	LineTotal
1	5	2024.994000
2	5	6074.982000
3	5	2024.994000
4	5	2039.994000
5	5	2039.994000
6	5	4079.988000
7	5	2039.994000
8	5	86.521200

Next step I summerized the LineTotal by Territory ID
Using this query

```
SELECT TerritoryID, SUM(LineTotal) AS Revenue
FROM Sales.SalesOrderHeader
JOIN Sales.SalesOrderDetail
ON Sales.SalesOrderHeader.SalesOrderID = Sales.SalesOrderDetail.SalesOrderID
GROUP BY TerritoryID
```

	TerritoryID	Revenue
1	9	10655335.959317
2	3	7909009.005872
3	6	16355770.454862
4	7	7251555.646926
5	1	16084942.547585
6	10	7670721.035475
7	4	24184609.600810
8	5	7879655.072151
9	2	6939374.481005

The next thing step is to withdraw the **SaleLastYear, SalesYTD**, and **CountryRegionCode** from **SalesTerritory**

That was done by joining the result from the previous query to the **SalesTerritory** on **TerrtoryID**

```
SELECT CountryRegionCode, SalesYTD, SalesLastYear, Revenue
FROM Sales.SalesTerritory
JOIN
  (SELECT TerritoryID, SUM(LineTotal) AS Revenue
   FROM Sales.SalesOrderHeader
  JOIN Sales.SalesOrderDetail
  ON Sales.SalesOrderHeader.SalesOrderID = Sales.SalesOrderDetail.SalesOrderID
  GROUP BY TerritoryID) AS TerritoryRevenue
ON Sales.SalesTerritory.TerritoryID = TerritoryRevenue.TerritoryID
```

	CountryRegionCode	SalesYTD	SalesLastYear	Revenue
1	AU	5977814.9154	2278548.9776	10655335.959317
2	US	3072175.118	3205014.0767	7909009.005872
3	CA	6771829.1376	5693988.86	16355770.454862
4	FR	4772398.3078	2396539.7601	7251555.646926
5	US	7887186.7882	3298694.4938	16084942.547585
6	GB	5012905.3656	1635823.3967	7670721.035475
7	US	10510853.8739	5366575.7098	24184609.600810
8	US	2538667.2515	3925071.4318	7879655.072151

On the following step I grouped the totals by **CoutryRegionCode**

```
SELECT CountryRegionCode, SUM(SalesYTD) AS SalesYTD, SUM(SalesLastYear) AS
SalesLastYear, SUM(Revenue) AS Revenue
FROM Sales.SalesTerritory
JOIN
        (SELECT TerritoryID, SUM(LineTotal) AS Revenue
        FROM Sales.SalesOrderHeader
        JOIN Sales.SalesOrderDetail
        ON Sales.SalesOrderHeader.SalesOrderID = Sales.SalesOrderDetail.SalesOrderID
        GROUP BY TerritoryID) AS TerritoryRevenue
ON Sales.SalesTerritory.TerritoryID = TerritoryRevenue.TerritoryID
GROUP BY CountryRegionCode
ORDER BY Revenue DESC
```

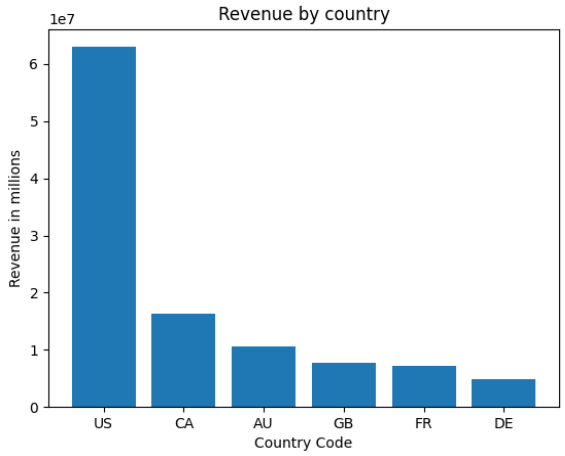
	CountryRegionCode	SalesYTD	SalesLastYear	Revenue
1	US	26411059.8792	19402504.6492	62997590.707423
2	CA	6771829.1376	5693988.86	16355770.454862
3	AU	5977814.9154	2278548.9776	10655335.959317
4	GB	5012905.3656	1635823.3967	7670721.035475
5	FR	4772398.3078	2396539.7601	7251555.646926
6	DE	3805202.3478	1307949.7917	4915407.595885

At last I query the table **Sales.vStoreWithAddresses** selecting **CountryRegionName** and counting it by the number of stores to find the relationship between number of stores and revenue..
I come to the conclusion that countries with high number of stores has a higher revenue.

```
SELECT CountryRegionName, COUNT(*)AS NumberOfStore
FROM Sales.vStoreWithAddresses
GROUP BY CountryRegionName
ORDER BY NumberOfStore DESC
```

	CountryRegionName	NumberOfStore
1	United States	437
2	Canada	115
3	Germany	40
4	France	40
5	United Kingdom	40
6	Australia	40

Once I had all the queries done I imported my cvs file into python.
So I could generation the charts.



Looking at the charts I come to the conclusion that countries with high number of stores has a higher revenue.