***Analysing Company Performance with SQL***

**The Brief:**

The Northwind database contains the sales data for a fictitious company called Northwind Traders, which imports and exports specialty foods from around the world.

You are tasked to perform an analysis of the performance of this company using SQL.

**Description:**

In this individual assignment, you will have to load the provided dataset into a Postgres database and create SQL queries to answer some business questions.

You can download the dataset here: [link](https://drive.google.com/file/d/1xfUF79LxsYte9Wevce-Fhpw50uiS_0Jx/view?usp=sharing)

The dataset is composed of 13 tables:

|  |  |
| --- | --- |
| **Table** | **Rows** |
| Categories | 8 |
| Customers | 91 |
| Employees | 9 |
| EmployeeTerritories | 49 |
| Order Details | 2155 |
| Orders | 830 |
| Region | 4 |
| Products | 77 |
| Shippers | 6 |
| Suppliers | 29 |
| Territories | 53 |
| Usstates | 51 |

Here is the entity relationship diagram (ERD) for this dataset

**Business Questions:**

Question 1

For their annual review of the company pricing strategy, the Product Team wants to look at the products that are currently being offered for a specific price range ($20 to $50). In order to help them they asked you to provide them with a list of products with the following information:

1. their name
2. their unit price

Filtered on the following conditions:

1. their unit price is between 20 and 50 (greater or equal to 20 but less or equal than 50)
2. they are not discontinued

Finally order the results by unit price in a descending order (highest first).

Question 2

The Logistics Team wants to do a retrospection of their performances for the year 1998, in order to identify for which countries they didn’t perform well. They asked you to provide them a list of countries with the following information:

1. their average days between the order date and the shipping date (formatted to have only 2 decimals)
2. their total number of unique orders (based on the order id)

Filtered on the following conditions:

1. the year of order date is 1998
2. their average days between the order date and the shipping date is greater or equal 5 days
3. their total number of orders is greater than 10 orders

Finally order the results by country name in an ascending order (following alphabetical order).

Question 3

The HR Team wants to know for each employee what was their age on the date they joined the company and who they currently report to. Provide them with a list of every employees with the following information:

1. their full name (first name and last name combined in a single field)
2. their job title
3. their age at the time they were hired
4. their manager full name (first name and last name combined in a single field)
5. their manager job title

Finally order the results by employee age and employee full name in an ascending order (lowest first).

Question 4

The Logistics Team wants to do a retrospection of their global performances over 1997-1998, in order to identify for which month they perform well. They asked you to provide them a list with:

1. their year/month as single field in a date format (e.g. “1990-01-01” for January 1990)
2. their total number of orders
3. their total freight (formatted to have no decimals)

Filtered on the following conditions:

1. the order date is between 1997 and 1998
2. their total number of orders is greater than 35 orders

Finally order the results by total freight (descending order).

Question 5

The Pricing Team wants to know which products had an unit price increase and the percentage increase was not between 20% and 30%. In order to help them they asked you to provide them a list of products with:

1. their product name
2. their current unit price (formatted to have only 2 decimals)
3. their initial unit price (formatted to have only 2 decimals)
4. their percentage increase with the result formatted to an integer (e.g 50 for 50%)  using the following calculation:

*(Current Unit Price - Initial Unit Price) ÷ Initial Unit Price \* 100*

Filtered on the following conditions:

1. their percentage increase is not between 20% and 30%  (lower than 20 or greater than 30)

Finally order the results by percentage increase (ascending order).

Question 6

The Pricing Team wants to know how each category performs according to their price range. In order to help them they asked you to provide them a list of categories with:

1. their category name
2. their price range as:
   1. “1. Below $20”
   2. “2. $20 - $50”
   3. “3. Over $50”
3. their total amount (formatted to be integer) taking into account the offered discount (i.e. subtracting the discounted amount)
4. their volume of orders (number of orders in which the category was present)

Finally order the results by category name then price range (both ascending order).

Question 7

The Logistics Team wants to know what is the current state of our regional suppliers' stocks for each category of product. In order to help them they asked you to provide them a list of categories with:

1. their supplier region” as:
   1. “America”
   2. “Europe”
   3. “Asia-Pacific”
2. their category name
3. their total units in stock
4. their total units on order
5. their total reorder level

Finally order the results by category name, then supplier region and reorder level (each in ascending order).

Question 8

The Pricing Team wants to know for each currently offered product how their unit price compares against their categories average and median unit price. In order to help them they asked you to provide them a list of products with:

1. their category name
2. their product name
3. their unit price
4. their category average unit price (formatted to have only 2 decimals)
5. their category median unit price (formatted to have only 2 decimals)
6. their position against the category average unit price as:
   1. “Below Average”
   2. “Equal Average”
   3. “Over Average”
7. their position against the category median unit price as:
   1. “Below Median”
   2. “Equal Median”
   3. “Over Median”

Filtered on the following conditions:

1. They are not discontinued

Finally order the results by category name then product name (both ascending).

Question 9

The Sales Team wants to build a list of KPIs to measure employees' performances. In order to help them they asked you to provide them a list of employees with:

1. their full name (first name and last name combined in a single field)
2. their job title
3. their total sales amount excluding discount (formatted to have only 2 decimals)
4. their total number of unique orders
5. their total number of orders
6. their average product amount excluding discount (formatted to have only 2 decimals). This corresponds to the average amount of product sold (without taking into account any discount applied to it).
7. their average order amount excluding discount (formatted to have only 2 decimals). This corresponds to the ratio between the total amount of product sold (without taking into account any discount applied to it) against to the total number of unique orders.
8. their total discount amount (formatted to have only 2 decimals)
9. their total sales amount including discount (formatted to have only 2 decimals)
10. Their total discount percentage (formatted to have only 2 decimals)

Finally order the results by total sales amount including discount (descending).

Question 10

The Sales Team wants to build another list of KPIs to measure employees' performances across each category. In order to help them they asked you to provide them a list of categories and employees with:

1. their categories name
2. their full name (first name and last name combined in a single field)
3. their total sales amount including discount (formatted to have only 2 decimals)
4. their percentage of total sales amount including discount against his/her total sales amount across all categories (formatted to have only 5 decimals and maximum value up to 1)
5. their percentage of total sales amount including discount against the total sales amount across all employees (formatted to have only 5 decimals and maximum value up to 1)

Finally order the results by category name (ascending) then total sales amount (descending).

**Submission:**

You will submit the following **2 files** to Canvas:

* your final report (in Word document or PDF),
* a zip file containing your code:
  + SQL scripts (1 text file with .sql extension per business question) saved in a folder called ‘queries’
  + their resulting tables as CSV files saved in a folder called ‘csv’

Note: Each SQL script SHOULD NOT contain any comments, only the SQL statements.

The report will need to contain at least the following topics:

* Presentation of data
* Instructions to create and connect to database
* Answers to business questions with explanations, tables and/or graphs (as if you would have communicated your results back to the stakeholders)

**Data Dictionary:**

Categories Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | CategoryID | int |  |  |  | Categories of Northwind products. |
|  |  | IX | CategoryName | varchar | (15) |  |  | Name of food category. |
|  |  |  | Description | text |  | Yes |  | Full description of the category. Provide samples. |
|  |  |  | Picture | image |  | Yes |  | A picture representing the food category. |

CustomerCustomerDemo Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK | FK | IX | CustomerID | varchar | (5) |  |  |  |
| PK | FK | IX | CustomerTypeID | varchar | (10) |  |  |  |

CustomerDemographics Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | CustomerTypeID | varchar | (10) |  |  |  |
|  |  |  | CustomerDesc | ntext |  | Yes |  |  |

Customers Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | CustomerID | varchar | (5) |  |  | Unique five-character code based on customer name. |
|  |  | IX | CompanyName | varchar | (40) |  |  |  |
|  |  |  | ContactName | varchar | (30) | Yes |  |  |
|  |  |  | ContactTitle | varchar | (30) | Yes |  |  |
|  |  |  | Address | varchar | (60) | Yes |  | Street or post-office box. |
|  |  | IX | City | varchar | (15) | Yes |  |  |
|  |  | IX | Region | varchar | (15) | Yes |  | State or province. |
|  |  | IX | PostalCode | varchar | (10) | Yes |  |  |
|  |  |  | Country | varchar | (15) | Yes |  |  |
|  |  |  | Phone | varchar | (24) | Yes |  | Phone number includes country code or area code. |
|  |  |  | Fax | varchar | (24) | Yes |  | Phone number includes country code or area code. |

Employees Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | EmployeeID | int |  |  |  | Number automatically assigned to new employee. |
|  |  | IX | LastName | varchar | (20) |  |  |  |
|  |  |  | FirstName | varchar | (10) |  |  |  |
|  |  |  | Title | varchar | (30) | Yes |  | Employee's title. |
|  |  |  | TitleOfCourtesy | varchar | (25) | Yes |  |  |
|  |  |  | BirthDate | datetime |  | Yes |  |  |
|  |  |  | HireDate | datetime |  | Yes |  |  |
|  |  |  | Address | varchar | (60) | Yes |  | Street or post-office box. |
|  |  |  | City | varchar | (15) | Yes |  |  |
|  |  |  | Region | varchar | (15) | Yes |  | Street or post-office box. |
|  |  | IX | PostalCode | varchar | (10) | Yes |  |  |
|  |  |  | Country | varchar | (15) | Yes |  |  |
|  |  |  | HomePhone | varchar | (24) | Yes |  | Phone number includes country code or area code. |
|  |  |  | Extension | varchar | (4) | Yes |  | Internal telephone extension number. |
|  |  |  | Photo | image |  | Yes |  | Picture of employee. |
|  |  |  | Notes | text |  | Yes |  | General information about employee's background. |
|  | FK |  | ReportsTo | int |  | Yes |  | Employee's supervisor. |
|  |  |  | PhotoPath | varchar | (255) | Yes |  |  |

EmployeeTerritories Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK | FK | IX | EmployeeID | int |  |  |  |  |
| PK | FK | IX | TerritoryID | varchar | (20) |  |  |  |

OrderDetails Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK | FK | IX | OrderID | int |  |  |  | Same as Order ID in Orders table. |
| PK | FK | IX | ProductID | int |  |  |  | Same as Product ID in Products table. |
|  |  |  | UnitPrice | decimal |  |  | (0) |  |
|  |  |  | Quantity | smallint |  |  | (1) |  |
|  |  |  | Discount | decimal |  |  | (0) |  |

Orders Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | OrderID | int |  |  |  |  |
|  | FK | IX | CustomerID | varchar | (5) | Yes |  |  |
|  | FK | IX | EmployeeID | int |  | Yes |  | Same entry as in Employees table. |
|  |  | IX | OrderDate | datetime |  | Yes |  |  |
|  |  |  | RequiredDate | datetime |  | Yes |  |  |
|  |  | IX | ShippedDate | datetime |  | Yes |  |  |
|  | FK | IX | ShipVia | int |  | Yes |  | Same as Shipper ID in Shippers table. |
|  |  |  | Freight | decimal |  | Yes | (0) |  |
|  |  |  | ShipName | varchar | (40) | Yes |  | Name of person or company to receive the shipment. |
|  |  |  | ShipAddress | varchar | (60) | Yes |  | Street address only -- no post-office box allowed. |
|  |  |  | ShipCity | varchar | (15) | Yes |  |  |
|  |  |  | ShipRegion | varchar | (15) | Yes |  | State or province. |
|  |  | IX | ShipPostalCode | varchar | (10) | Yes |  |  |
|  |  |  | ShipCountry | varchar | (15) | Yes |  |  |

Products Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | ProductID | int |  |  |  | Number automatically assigned to new product. |
|  |  | IX | ProductName | nvarchar | (40) |  |  |  |
|  | FK | IX | SupplierID | int |  | Yes |  | Same entry as in Suppliers table. |
|  | FK | IX | CategoryID | int |  | Yes |  | Same entry as in Categories table. |
|  |  |  | QuantityPerUnit | nvarchar | (20) | Yes |  | (e.g., 24-count case, 1-liter bottle). |
|  |  |  | UnitPrice | money |  | Yes | (0) |  |
|  |  |  | UnitsInStock | smallint |  | Yes | (0) |  |
|  |  |  | UnitsOnOrder | smallint |  | Yes | (0) |  |
|  |  |  | ReorderLevel | smallint |  | Yes | (0) | Minimum units to maintain in stock. |
|  |  |  | Discontinued | bit |  |  | (0) | Yes means item is no longer available. |

Region Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | RegionID | int |  |  |  |  |
|  |  |  | RegionDescription | varchar | (50) |  |  |  |

Shippers Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | ShipperID | int |  |  |  | Number automatically assigned to new shipper. |
|  |  |  | CompanyName | varchar | (40) |  |  | Name of shipping company. |
|  |  |  | Phone | varchar | (24) | Yes |  | Phone number includes country code or area code. |

Suppliers Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | SupplierID | int |  |  |  | Number automatically assigned to new supplier. |
|  |  | IX | CompanyName | varchar | (40) |  |  |  |
|  |  |  | ContactName | varchar | (30) | Yes |  |  |
|  |  |  | ContactTitle | varchar | (30) | Yes |  |  |
|  |  |  | Address | varchar | (60) | Yes |  | Street or post-office box. |
|  |  |  | City | varchar | (15) | Yes |  |  |
|  |  |  | Region | varchar | (15) | Yes |  | State or province. |
|  |  | IX | PostalCode | varchar | (10) | Yes |  |  |
|  |  |  | Country | varchar | (15) | Yes |  |  |
|  |  |  | Phone | varchar | (24) | Yes |  | Phone number includes country code or area code. |
|  |  |  | Fax | varchar | (24) | Yes |  | Phone number includes country code or area code. |
|  |  |  | HomePage | ntext |  | Yes |  | Supplier's home page on World Wide Web. |

Territories Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Primary Key** | **Foreign Key** | **Index** | **Name** | **Data type** | **Length** | **Nullable** | **Default** | **Description** |
| PK |  | IX | TerritoryID | varchar | (20) |  |  |  |
|  |  |  | TerritoryDescription | varchar | (50) |  |  |  |
|  | FK |  | RegionID | int |  |  |  |  |