

EDA and SQL project

Dataset 'Online Retail'

Context - Transnational data set which contains all the transactions occurring between December 10th, 2010 (01/12/2010) and December 9th, 2011 (09/12/2011) for a UK based non-store online retail

Main customers of the company is wholesalers

What do I want as business insights

1. What are the total sales of the business? month over month evolution
2. Can I see the classification of the products and total sales?
3. What is the average unit price?
4. What is our most expensive product and their sales?
5. What is our least expensive product and their sales?
6. What is the average sale per month?
7. Who are our customers?
8. Who is our biggest buyer?
9. Which countries do we sell to?
10. Which countries based on sales are our top 3?

Data assumptions

- Filter the data to only analyze 6 months (01-01-2011 to 31-06-2011)
- Invoice transactions which the quantity is negative or zero or unit_price = 0 are considered returns
 - quantity number of returns