

Exercise on data warehouse design

Problem specifications

As part of the enhancement of the excellence of Made in Italy, the Italian Government wants to analyze the export of Italian wines abroad, to understand, for example, which are the most popular wines, from which areas of Italy they come, and which are the main foreign countries importers.

The Government has collected from Italian wine companies the list of purchase orders. Each purchase order is characterized by date of the order, amount in euros, quantity in liters, type of wine, type of packaging (e.g., bottle, carton, demijohn, etc.), foreign country of destination.

The analysts, starting from the purchase orders, want to create a data warehouse which synthesizes the collected data in order to answer a set of queries. Each order is also associated to its size: orders up to 100 liters are considered small, medium the ones up to 1000 liters, and large the ones above. Finally, for every winery, it is possible to find out the complete personal data (name, address, province, region, geographical area of the region).

The queries analysts are interested in the total quantity of wine exported (in liters), and the average price per liter, according to the variation of:

- month, 2 months period, 3 months period, semester, and year
- province, region and geographical area (North, Center, South) of the winery
- wine type
- packaging type
- foreign country of destination, and its continent
- order size (small, medium, large)