Project Assignment - Part 2

Anna Monreale, Cristiano Landi

December 18, 2024

Introduction

In Part 2 of the project, you are required to answer business questions using a datacube that you will create based on the database prepared earlier. Document the process of building your datacube in your report, and solve the business questions using MultiDimensional eXpressions (MDX) within SQL Management Studio. Some assignments are marked with a * symbol. This indicates that **you are only required to complete 4** of the marked assignments, which you can choose as you please. The last three assignments require you to create a dashboard. We recommend using the software demonstrated during the lectures. However, you may also use alternative software to generate the dashboard (e.g., MicroStrategy), provided that (i) the dashboard is interactive and (ii) the software can connect to our database. In any case, you must include a screenshot and a brief description/motivation as to why the dashboard is relevant to the business.

Delivery Instruction

When you want to deliver your project, compress all the files, including the PDF report, and create a .zip file named LDS_GroupID_part2.zip. Then, upload the file using the following Google form using your *studenti.unipi.it* email: https://forms.gle/y8tLrkjjebMuCCcW9. If you have problems using the form, email ALL teachers with the subject: **[LDS] PART2 Group_Id**. The zip file should contain all the files you created/used except the original dataset.

The project must be submitted by **January 5th** (inclusive). Projects uploaded after the deadline will not be accepted. Only one member of the group is required to upload the project.

If you discover any bugs or errors in the code from the first part during this part of the project, you can resubmit the first part using the Google form linked in this pdf. The resubmission must include a detailed list of all the changes made compared to the version of the code previously submitted.

Assignment 1

Build a datacube from the data of the tables in your database, defining the appropriate hierarchies. Create the needed measures based on the queries you need to answer.

Assignment 2*

For each month, show the total damage costs for each location and the grand total with respect to the location.

Assignment 3*

Compute the average yearly damage costs as follows: for each crash, calculate the total damage to the user divided by the number of distinct people involved in the crash. Then, compute the average of these values across all crashes in a year.

Assignment 4*

For each location, show the damage costs increase or decrease, in percentage, with respect to the previous year.

Assignment 5*

For each quarter, show all the locations where the number of vehicles involved exceeds the average number of vehicles involved in the corresponding quarter of the previous year. Also, report the increase as percentage.

Assignment 6*

For each vehicle type and each year, show the information and the (total) damage costs of the person with the highest reported damage.

Assignment 7*

Propose and solve a query showing some interesting and **non-trivial** facts you discover during the first part of the project.

Assignment 8

Complete one of the following assignments:

1. For each year, show the most frequent cause of crashes and the corresponding total damage costs. The primary crash contributing factor is given twice the weight of the secondary factor in the analysis. Additionally, show the overall most frequent crash cause across all years.

2. For each year, show the most risky crash type and its total damage costs. To measure how risky a crash type is, you should assign a weight to each type of injury you encounter in the data (for example, a fatal injury weighs 5 times an incapacitating one, which weighs twice a non-incapacitating injury).

Assignment 9

Create a dashboard that shows the geographical distribution of the total damage costs for each vehicle category.

Assignment 10

Create a plot/dashboard that you deem interesting w.r.t. the data available in your cube, focussing on data about the street.

Assignment 11

Create a plot/dashboard that you deem interesting w.r.t. the data available in your cube, focussing on data about the people involved in a crash