UNIVERSITÀ DELLA CALABRIA

BUSINESS INTELLIGENCE AND ANALYTICS

DATA WAREHOUSE

GLOBAL TERRORISM ANALYSIS

Luis Angel Rodriguez Reiners

MAT: 214686

The current project is based on the analysis of global terrorism attacks since 1970 to 2018. The database used defines terrorist attacks as the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation. For the analysis made in this project were selected a subset of columns, there wasn't much problems with the original dataset, except for NULL values in some relations, which were substituted for fixed values.

The original database was in a csv. file, which in a first step was loaded to a relational database, Figure 1. shows its relational model.
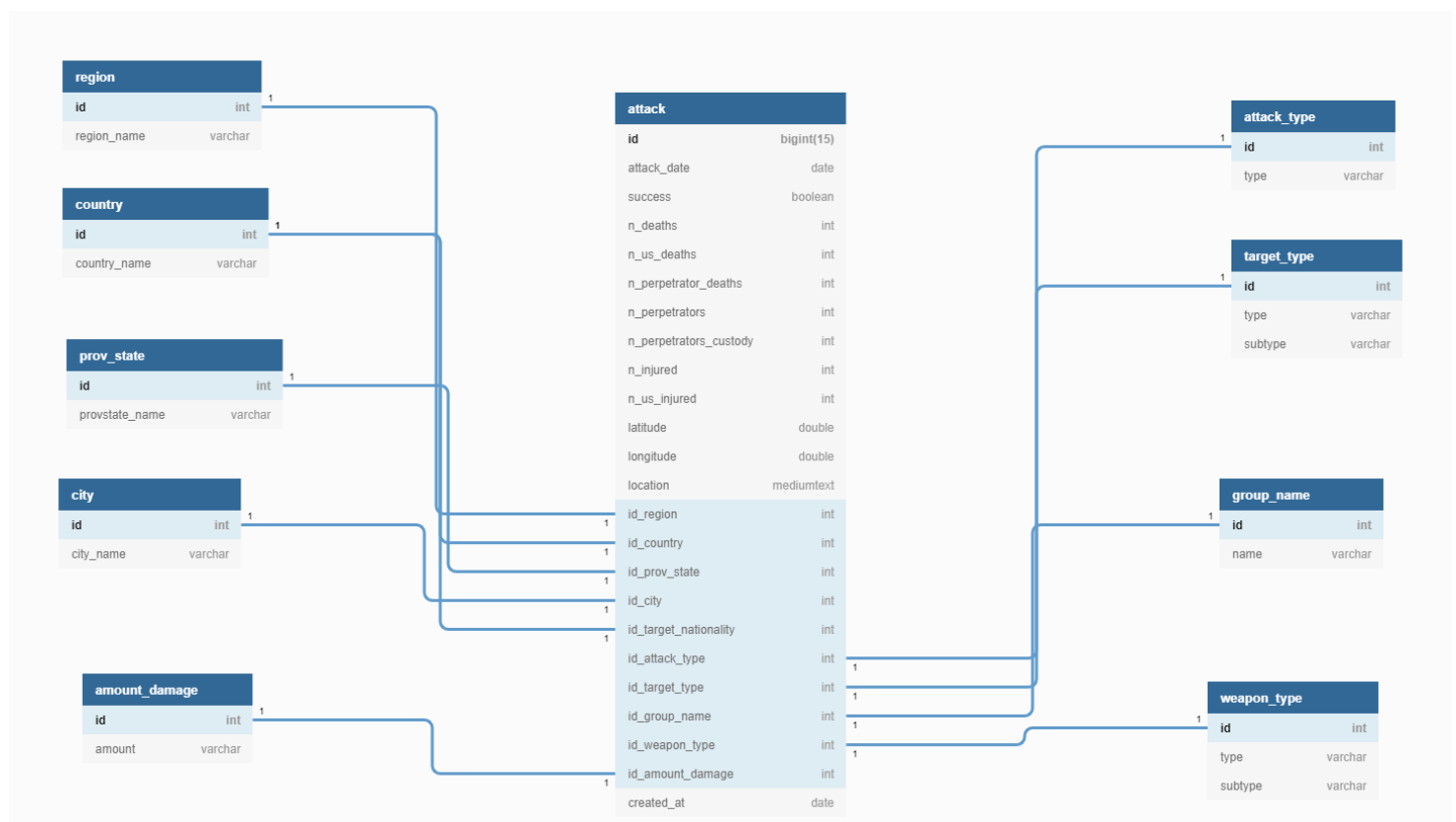

Figure 1. Relation Model Diagram

The extracted attribute tree schema is shown in Figure 2. and Figure 3. shows the fact schema.
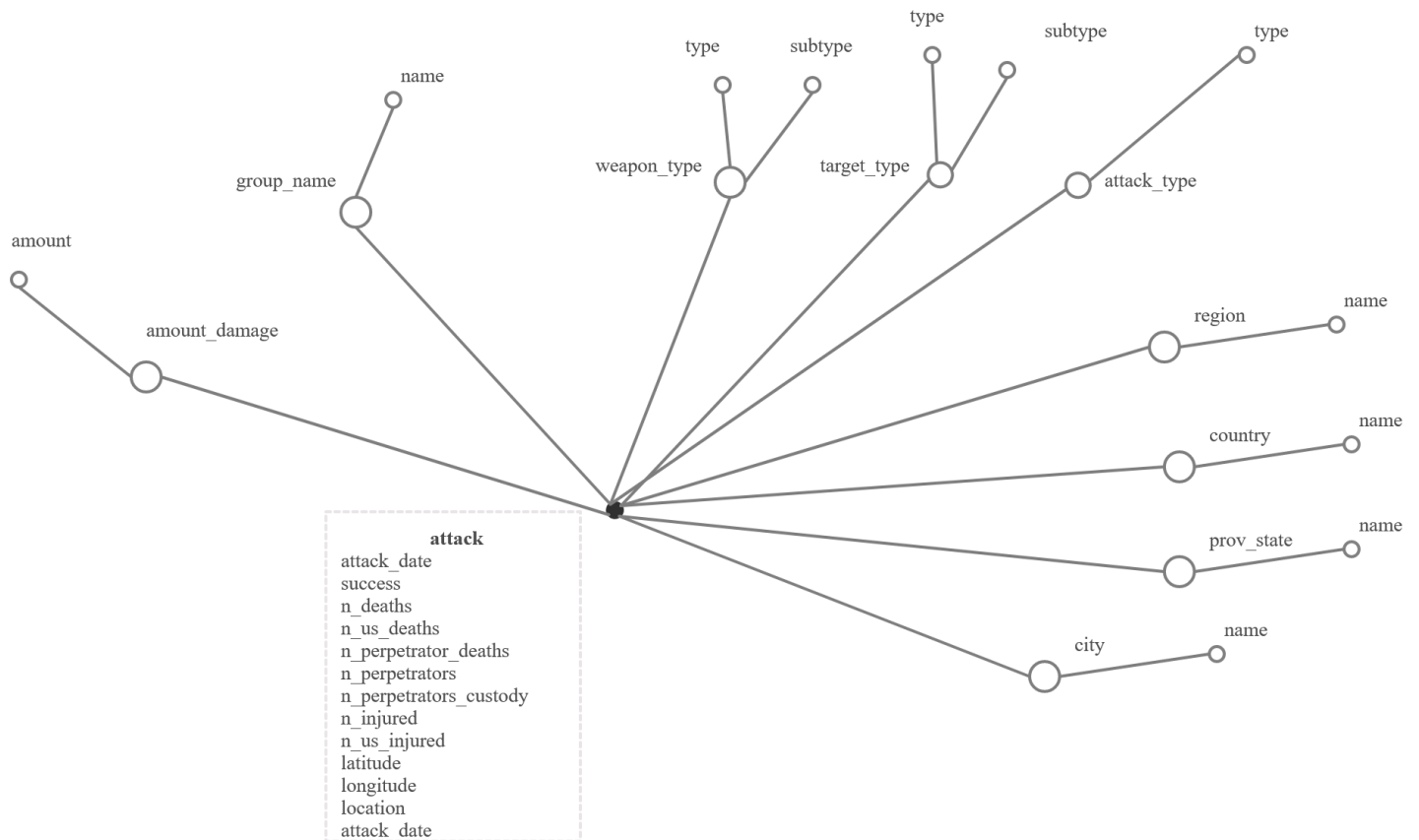

Figure 2. Attribute tree schema

Figure 3. Fact Schema

The selected logical model for working in the project was Relational Online Analytical Processing model (ROLAP). It assumes the relational model for representing multidimensional data; besides, ROLAP is supported by a lot of number of tools.

Arriving to this point, the star scheme has been generated. It is shown in Figure 4.
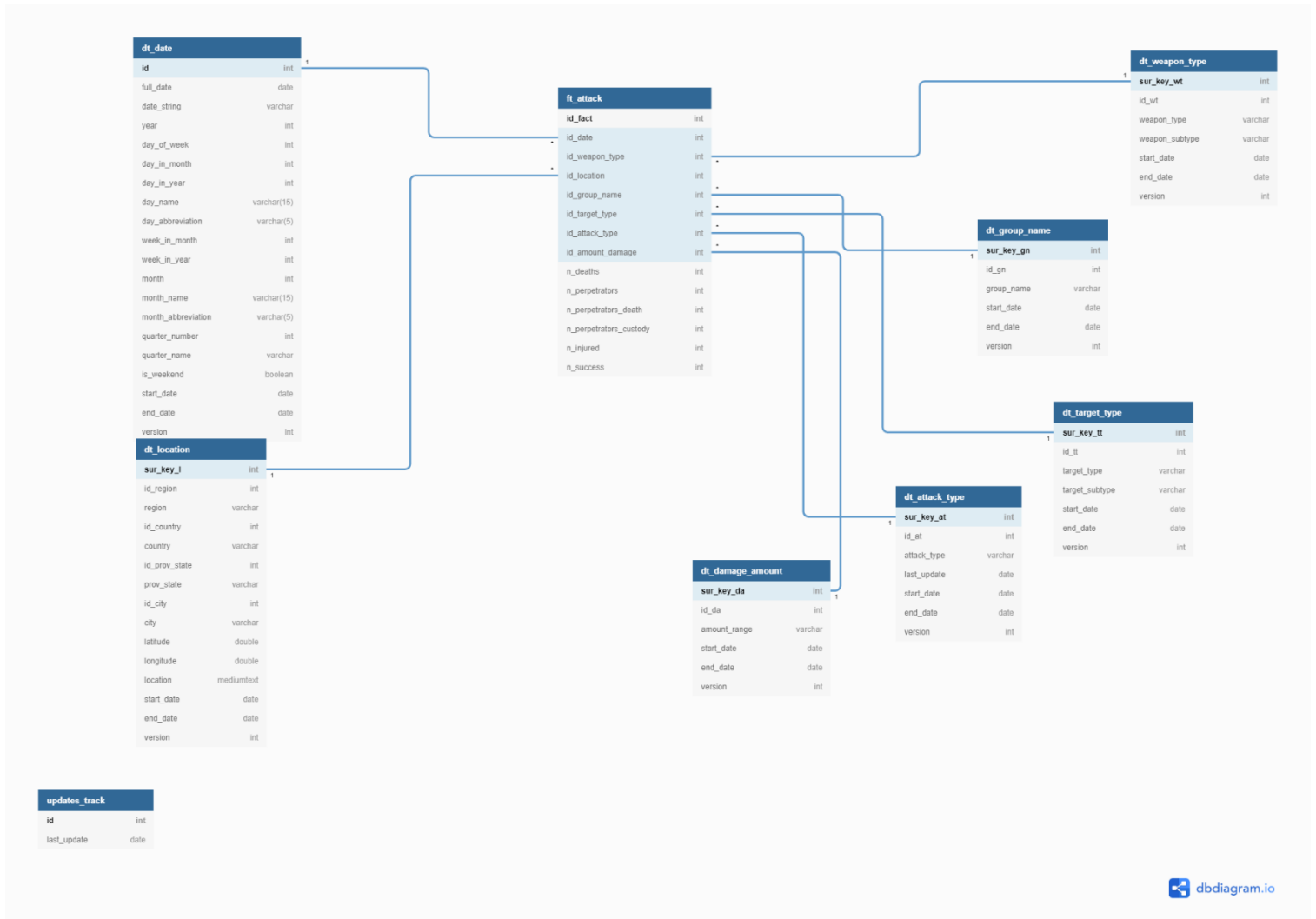

Figure 4. Star schema relational model.

In the project, for populating the data warehouse were used different jobs and transformations through Spoon tool in Pentaho Data Integration suite. Firstly, the data was extracted and transformed to a database (reconciled data), and then, from this point the data was loaded to the data warehouse. The principal jobs were:

- Load Database From CSV – JOB: the transformation extracts the data from the data source, transforms and loads it to the data stage area.
- Create-Update-DataWarehouse-Job: the transformation is used to calculated different metrics from the data stage area, and then to load them into the data warehouse.

BUSINESS INTELLIGENCE AND ANALYTICS

DATA WAREHOUSE

GLOBAL TERRORISM ANALYSIS

Luis Angel Rodriguez Reiners

MAT: 214686