Term Project 1 - Peaks

A. BOGNAR

Introduction, File Structure

The Tidy Tuesdays is a great practice project that was designed mainly for R. However, it contains several relational databases, one of which is the **Himalayan Climbing Expeditions**. The data was collected until from 1905 to Spring 2019. To see further details check the original source.

In my project, I focused on building a data workflow similar to the one leared during the course.

To run the project all one has to do is:

1. Download the 3 **Clean** csv files and save them in your MySQL Upload Folder.
2. Run the SQL files in the following order:

   - Loading. Note that this is the only file where codes have to be changed.
   - Cleaning
   - Analytical Layer and Triggers
   - Views

Data and Analytics Questions

Clean files, Loading and Stored Procedures

Changes made from raw to clean:

- Date format to SQL readable
- Commas removed to allow separation.
- Duplicate Primary keys found: Expeditions 1 KANG10101, members KANG10101 #

Analytical Layer, ETL, Trigger

Data Mart, Views