Assignment Data Storytelling

Data Scientist, EC Education
April 2023

Submission Deadline: 2023-05-30

1. General information

Your individual project submission will take the form of:

1) a report/dashboard in the Power BI Cloud service, plus 2) a supplementary file, in any format, where you **briefly** describe:

a) *What is the Data Storytelling (communication impact) you are attempting to deliver via this report/dashboard.*
b) *Explain how you have structured your report/dashboard in order to deliver that impact*
c) *Explain for each of your visualizations:*
    a. *Why you chose that way to visualise and communicate the information.*
    b. *How you planned for that visualisation to support and help deliver the dashboard's overall impact.*
d) *Why you believe your technical analysis is correct and what steps you took to confirm your results.*

As the focus of this assignment is on your ability to create and communicate impactful stories via visualisations, you will have a large degree of artistic freedom regarding the overall layout, designs, and forms of visualisations you choose to employ.

It is allowed to import additional information sources to the data, or even for you to add information manually, if you believe it will provide extra clarity and meaning to the task (for example, information about which continent a country belongs to).

Additionally, remember that if there is a lot of redundancy in the data, it is generally wise to break up the material.

Finally, the assessment criteria are provided for you in section 3, which specify the minimum demands and requirements which will need to be implemented and convincingly presented to receive either a **G, VG, or Incomplete** marking for this assignment.

2 Task description

2.1 Assignment

Your assignment is to study the data, to learn and explore the relationships between each of the countries and various exported shipments, and to create visualisations which present a cohesive understanding of what you have discovered and what you recommend as a result.

To help focus your work and to make it easier for you to craft an impactful, actionable, and results-based presentation, I strongly recommend for you to identify relatively early in your data analysis:

1) **a role for yourself** and the company/organisation for whom you are working, and
2) **the audience for whom you are presenting** the visualisations you have created.

2.2 Data

In the attached csv files, you will find records pertaining to meat transactions between several countries, one file for each country. In the columns therein you will find the following information for each export transaction:

- **Sender** - The country that exports the goods.
- **Receiver** - The country that imports the goods.
- **Sender Budget** - The annual budget the sending country has for imports.
- **Sender Expected Revenue** - The expected annual gross revenue from meat exports.
- **Receiver Budget** - The annual budget the recipient country has for imports.
- **Receiver Expected Revenue** - The expected annual gross revenue from meat exports.
- **Meat type** - Type of meat exported.
- **Meat price** - The price per kilo for the type of meat exported.
- **Toxicity** - Unfortunately, all meat in this assignment has some degree of toxicity in it(!)
- **Day Number** - Which day it is, measured from the start of the measurements.
- **Discount** - Different countries have different trade agreements between them. This factor therefore determines the amount of the discounting levels in each export transaction.

Good Luck!

3 Assessment criteria

3.1 For a G:

- Make at least one DIM-Table and one FACT-table and connect them.
- Make at least 6 different types of visualizations and at least 2 separate sheets.
- Implement at least 2 different DAX columns.
- Implement at least 2 different DAX measures.
- Answer (and visualize) which countries have exported and imported the lowest and highest volumes of meat throughout the period of time.
- Answer (and visualize) the day on which the least meat shipped corresponding to the least value of goods exported (take i.e. consideration of price per kilo and volume).

3.2 For a VG:

In addition to what is required for G:

- Answer (and visualize) which countries have exceeded their import budget for the year (i.e., taking into account shipping volumes, kilo price and discount)
- Answer (and visualize) which countries have fallen below their Expected Revenue for the year (also taking into account here volume, kilo price and discount)
- Implement Drill Downs/Hierarchies in at least one of your visualizations.
- Implement at least one DAX which uses RELATED in its expression.
- Include a Python or R script where it feels natural.
    - Suggestions for this could be, for example: to either use it as a visual tool or use it to load data from *https://github.com/evahegnar/SSBI* rather than through the csv files.
- Visualize all exports/imports for the entire time span with a Flow Map where all edge sizes are controlled by total exports.
- Make Discount into a DIM table which contains all pairs of countries.
    - Note that discounts for exporting from country A to country B does not have to be the same as the discount for exporting from country B to country A. You may need to make a new column to succeed getting indexes with composite keys to merge.

Note: using DAX with various relationships can make it easier to address these requirements