A Visual tool for Data Quality Checking of Geomagnetic Data

Juan Carlos Méndez

jc.mendez (at) uniandes.edu.co

https://beta.observablehq.com/d/e7cc7a2094faa56c

Description

- **Dataset Title:** Geomagnetic Components of the Fúquene Observatory between 1955 and 2014
- **Number of Rows:** 525.960
- **Data Types :** Quantitative, temporal
- **The Problem:** The dataset has a huge number of rows (525.960). The owner of data doesn't have a visual tool that let him easily check data quality
- **The Task:** Allow the user to explore the data set in a visual and interactive way in order to detect possible quality problems in the data.

The Solution

- An interactive visualization was created using **d3** and **observablehq** that allows the user to explore the horizontal, vertical, declination and total intensity geomagnetic components in a summarized way by aggregation per year, month or day, as well as the exploration of the detail of the time series for a specific day.
- **Technologies:**
  - **Visualization:**
    - d3, observablehq, html5, css
  - **Data Preprocessing:** *(Summarization and Exploration)*
    - Python, Pandas, Jupyter, JupyterLab

Demo

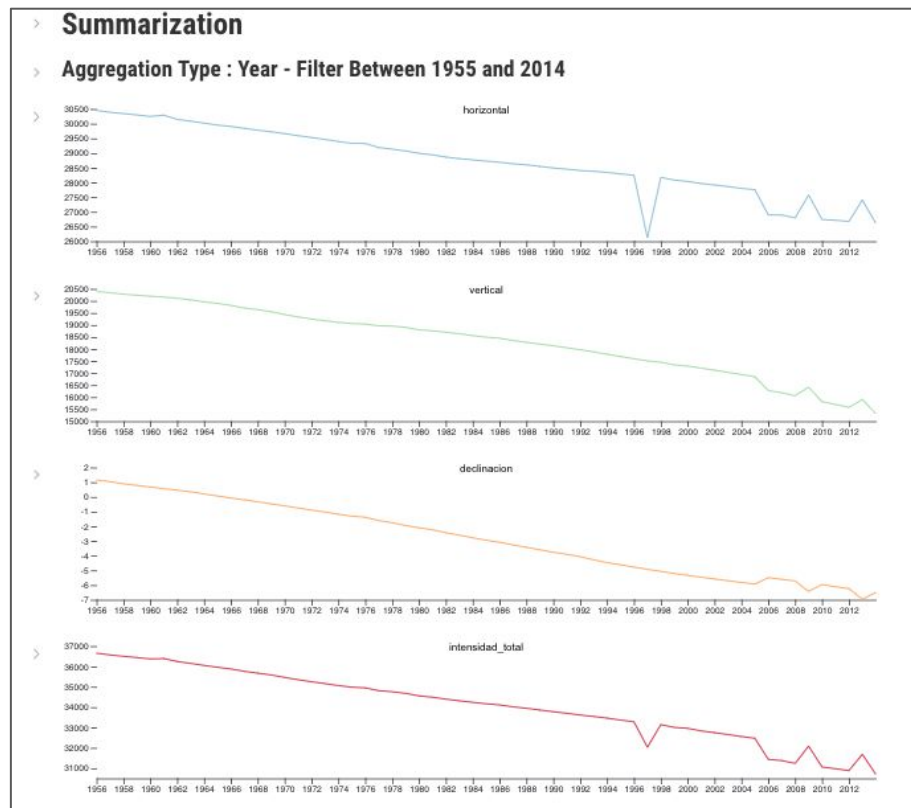[A Visual tool for Data Quality Checking of Geomagnetic Data](#)

Demo
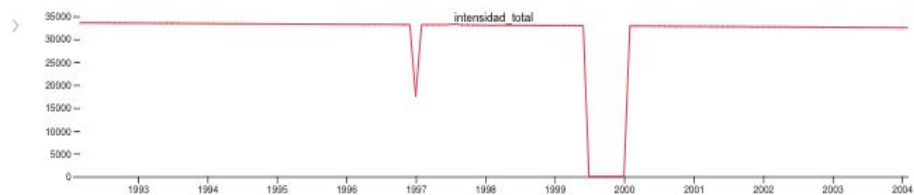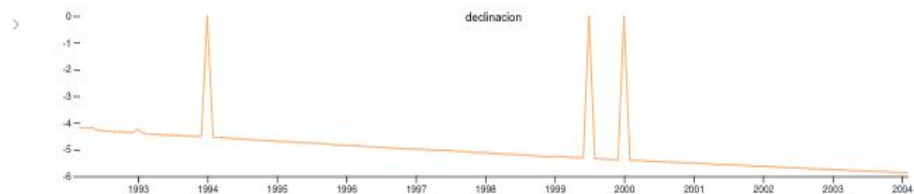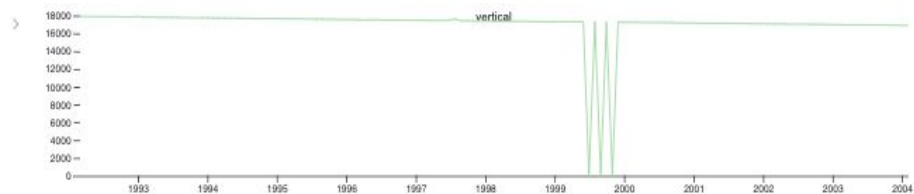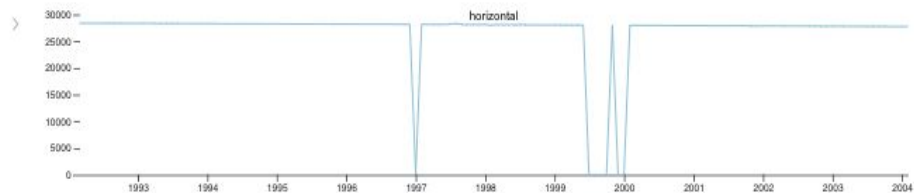
Demo
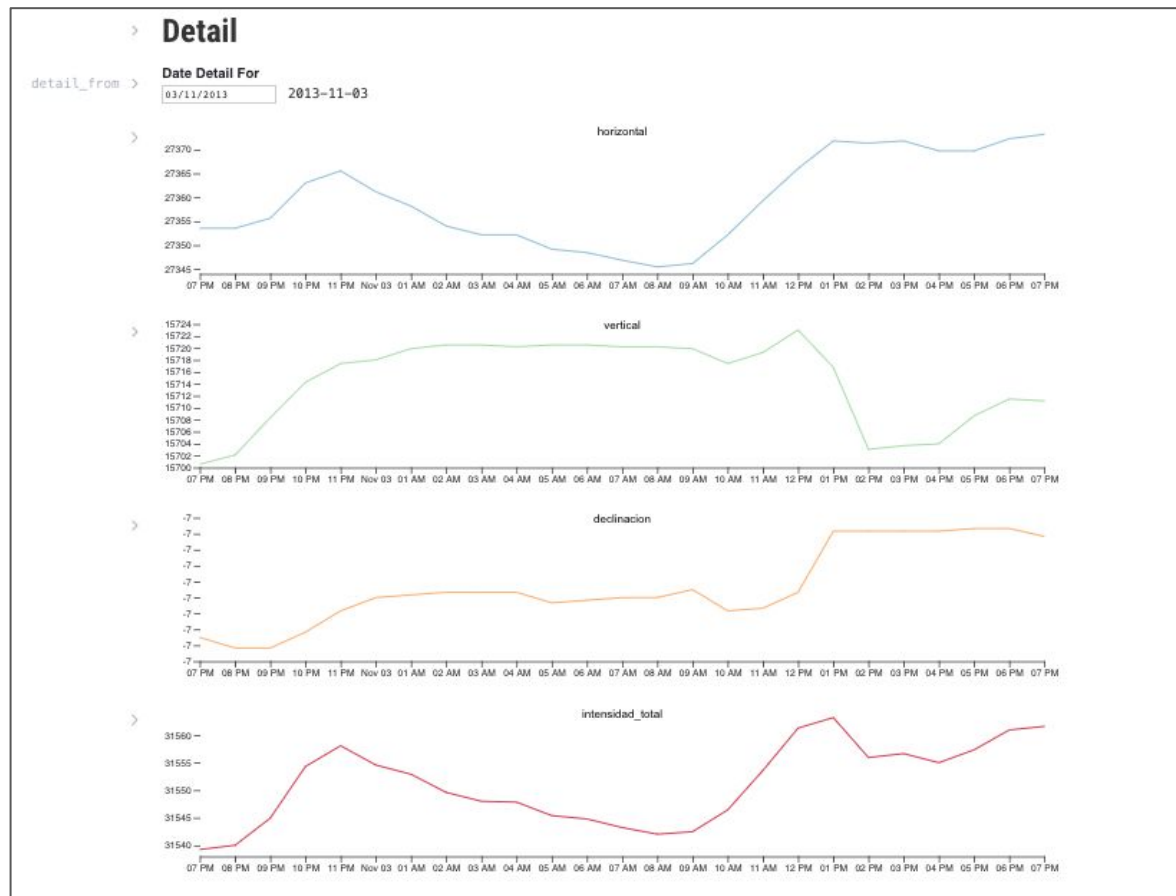
Insights

1. The user discovered some anomalies in the data using the interactive visualization.

2. According to the owner of the data, those anomalies were caused by errors during the process of the extraction and transformation of the original data.

3. The owner of the data will publish in the following days a new corrected version of the dataset

4. The owner of data is considering to include visual tools like this in his daily work.

Difficulties

- D3's Learning Curve (i.e. data parsing ,)
- It is not easy to *"debug"* using observablehq. *console.log("...blah...");*
- Non-Reproducible examples (i.e. old versions of D3)
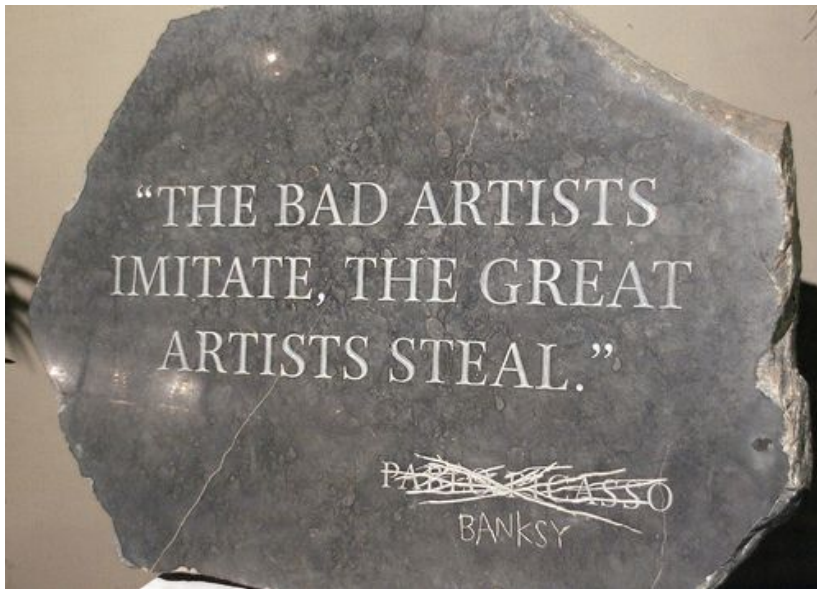
Achievements

- The user actually got insights!
- My first work using *observablehq* and it was possible to create a "*reusable*" chart.
- The user got interested about using interactive visualization tools