SoDA – Social Media Data Analysis

Jérôme Treboux, Fabian Cretton, Dominique Genoud, Anne Le Calvé, Florian Evéquoz,
Antonio J. Jara, Blaise Larpin[1]

[1]University of Applied Sciences Western Switzerland (HES–SO), Sierre, Switzerland;

Summary

1. Clients online behaviors analysis over multiple social media channels
2. Original methodology and tool that aims at formalizing the marketing need of the company
3. Two Swiss companies as use cases for the prototype development

Introduction

- **+12%** active social media accounts in **2015** compared to **2014 (+222 Million)**
- **2.078 Billion** active social media accounts in **2015 (29%** of the total population**).**
- **16 min/hour** is the time spent on social networks in the USA
- **23**% of Facebook users login at least **5** times a day

We explore how a combination of social data mining and information visualization can inform marketing decisions. We present the data mining tools and prototype visualizations that we developed for marketing specialists in two different use cases: a Swiss regional tourism promotion agency (Valais Wallis Promotion) and a Swiss online discount retailer (Qoqa.ch).

Fig. 1 – Work flow from the social network to the client

Methods
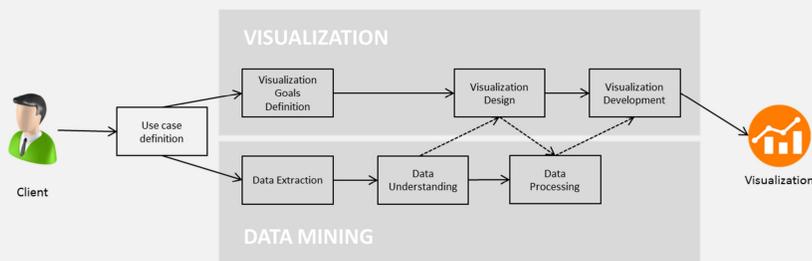
Fig. 2 – The methodology used to create the visualization in our use cases is based on CRISP-DM, with the addition of a visualization branch.

Results

- Use Case "Valais Wallis Promotion". Origin and destination of tourists and semantic visualization of tourists perception
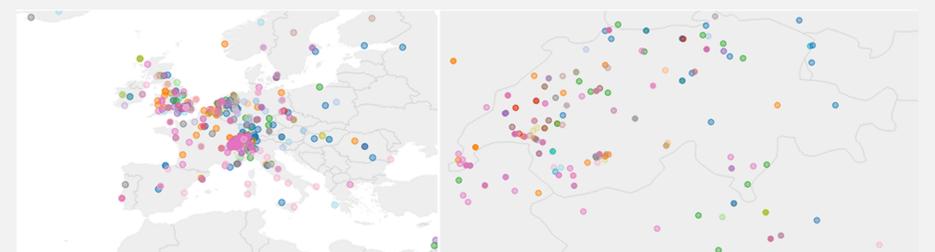
Fig. 3 – Origin of tourists visiting Valais according to Flickr. Position of dots corresponds to origin of tourist. Color corresponds to a destination. European map (left), Swiss map (right)
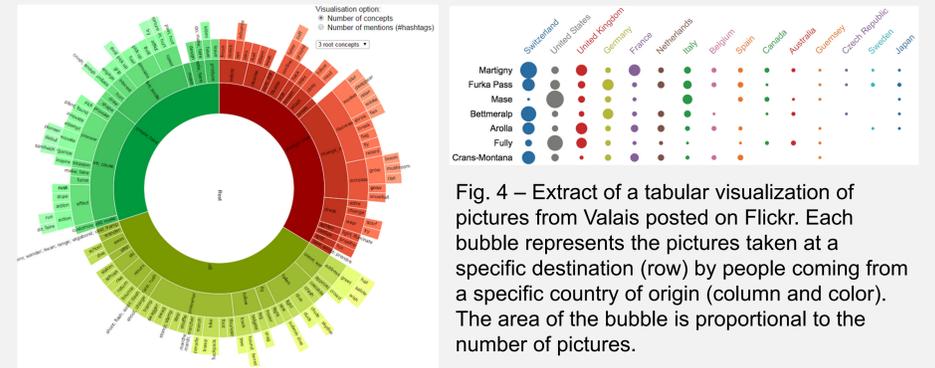
Fig. 4 – Extract of a tabular visualization of pictures from Valais posted on Flickr. Each bubble represents the pictures taken at a specific destination (row) by people coming from a specific country of origin (column and color). The area of the bubble is proportional to the number of pictures.

Fig. 5 – Sunburst visualization presenting the concepts most mentioned by tourists in the picture caption posted on their Instagram profile.

- Use case "Qoqa". Visualization of clients origin and gender for a better marketing strategy
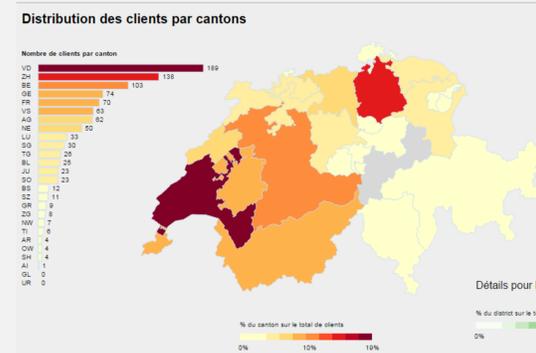
Fig. 6 – Extraction of Qoqa consumer distribution by canton. The colors represent the importance of a canton compared to others.

Fig. 7 – Example of client (left) and gender (right) distribution by district in the canton of Valais

Conclusion

1. Adapted CRISP-DM methodology to match the client vision
2. The variety of API to connect to social media vary often over time
3. Hashtags found in users' comments visualized through a novel way based on a multilingual network of concepts
4. Confirmation from the tourist expert that the information from social media is under used. They assessed the value of the tool