Faculté des Sciences et Technologies
Département Informatique
Université Claude Bernard – Lyon 1

Data Visualization

Project specification

Facial expression emoji : visualization and analysis

Students	11217236 - Thi Ngoc Bich NGUYEN 11413751 – Cheng CHENG
Course	Data Visualization
Departement	Informatique
Tuteurs	Aurélien TABARD Romain VUILLEMOT Nicolas BONNEEL
Due date	12 Jan. 2017

1. Project Direction

1.1. Project Goals

Visualizing patterns over time:

- Frequency of using emoji
- Co-occurence between emoji
- Emoji ranking
- Popularity of emoji

Week	Tasks
Week 1	Choosing dataset (we excluded publication sin Facebook, LinkedIn,... but Twitter) Loading data from Twitter Storing data in mongoDB Learning D3js
Week 2	Processing data
Week 3	Writing algorithm to realize mapreduce job
Week 4	Visualizing data using D3js
Week 5	Testing and improving algorithm Visualizing data

1.2. Purpose

We want to visualize in realtime the frequency of using emoji, the emoji co-occurrence, emoji ranking over a specific period of time and then infer netizen's emotion by analyzing theirs publication in social network (in this case Twitter).

1.3. Project Background

Recent years, emoji or emoticon takes un important role in social networks by allowing users to express : their status of situation, sentiments… without using traditional texte.

2. Project Outcome

A website who give user a global view of tendanncy in using emoji. This website consists of 5 graphics:

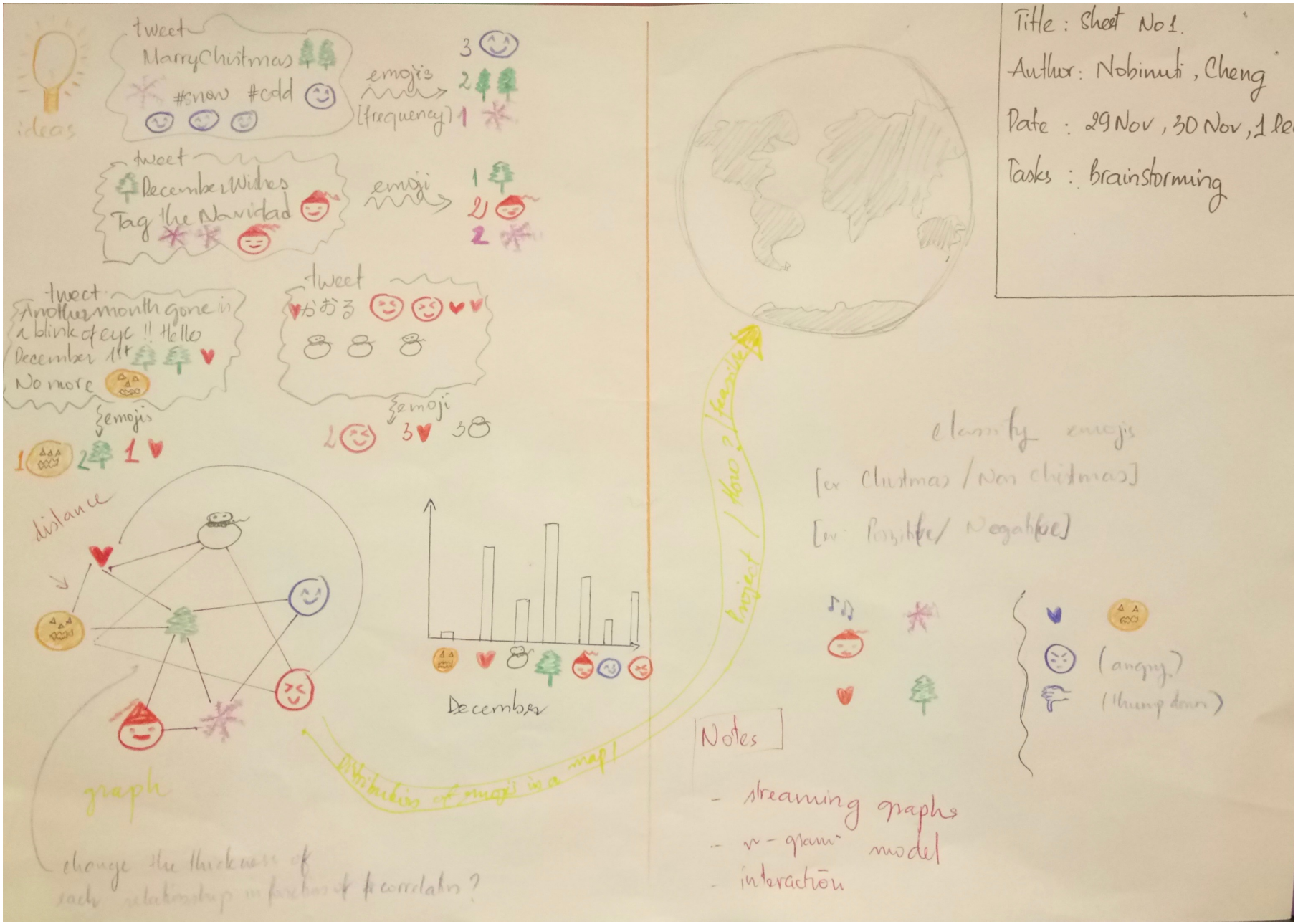
- Force-directed graph: relation between emoji, which are the emoji who connect to a specific one
- Bar chart: number of times each emoji appears in tweets
- Emoji-cloud: comparison of number of time each emoji appears in tweets in form of cloud
- Line graph (n-gram): the frequency of each emoji over the course of time
- Emoji co-occurrence matrix

3. Risks

The most important risks that have been identified:

- Visualize graphics in realtime
- Lack of having the possibility in storing all the tweets over a long period of time
- Overload data and interaction functionality

4. Design sketches


Title : Sheet No 1.
Author : Nobinuti , Cheng
Date : 29 Nov , 30 Nov , 1 De
Tasks : brainstorming
ideas
tweet
MarryChristmas
#snow #cold
emojis
[frequency]
3
2
1
tweet
December Wishes
Tag the Navidad
emoji
1
2
2
tweet
Another month gone in
a blink of eye !! Hello
December !!!
No more
emojis
1
2
1
tweet
おおる
emoji
2
3
3
distance
graph
December
Distribution of emoji in a map
change the thickness of
each relationship in function of the correlation ?
classify emojis
[ex Christmas / Non christmas]
[ex: Positive/ Negative]
(angry)
(thumb down)
Notes
- streaming graphs
- n-gram model
- interaction

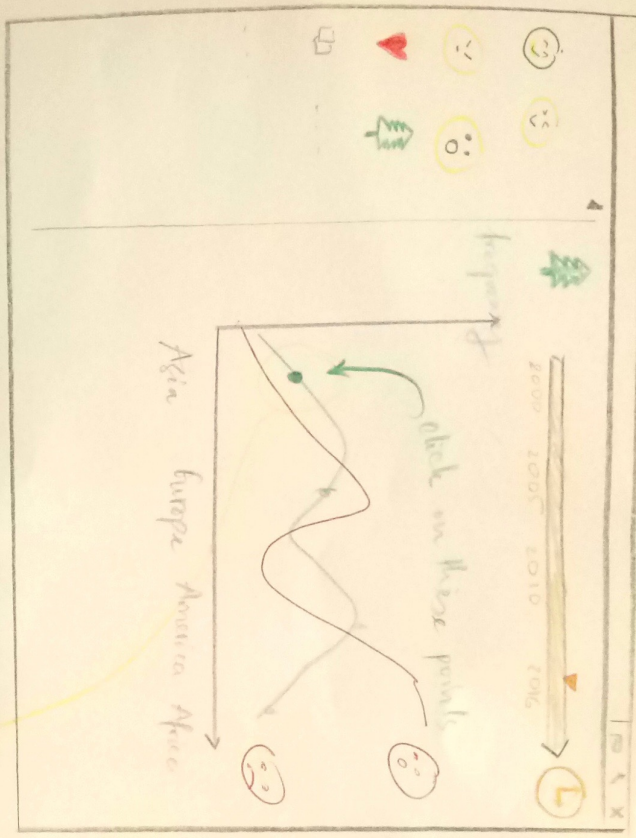
Title : Sheet no.2

Author: Nobinuti, Cheng

Date : 29 Nov, 30 Nov / 1 Dec

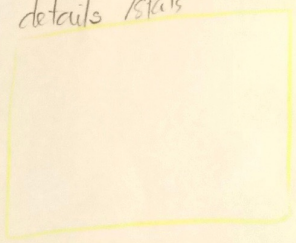
Tasks: . working with tweets
. filtering / cleaning tweets
.

Layout



Layout Operation

- Click on icon (emoji) in the smiley tab.
- Choose time (in timeline)



Focus Discussion

Pros	Cons
Plus spécifique que 2 graphe indépendant	Plus compliqué à réaliser
Plus précis	

Focus ~~Discussion~~ Focus

- les frequencies of each smiley

- How about visualize icon/smiley in a map of continents ?

Plus intuitive
Plus convivial

Il faut récupérer cordonnée géographique de chaque tweet (déjà fait)

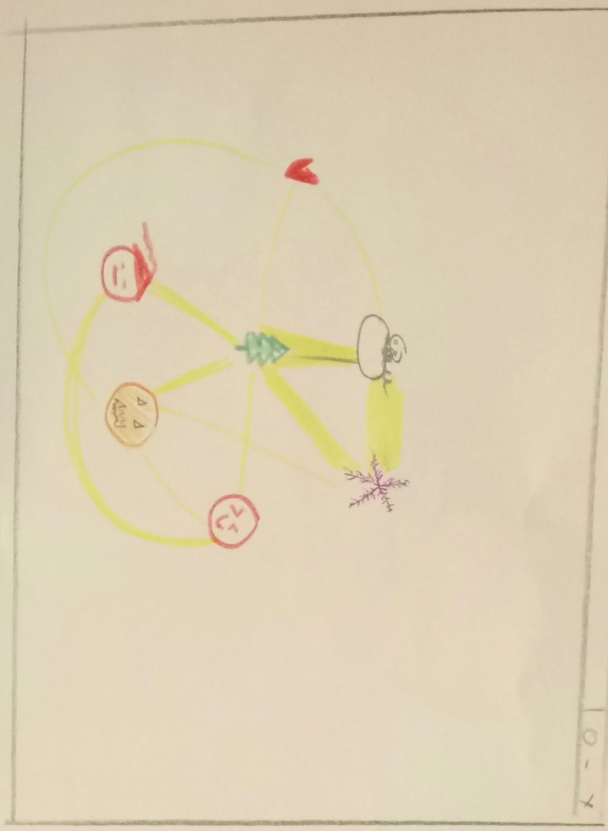
Title : Sheet no.3

Author : Nobinuti, Chung

Date : 29 Nov, 30 Nov, 1 Dec

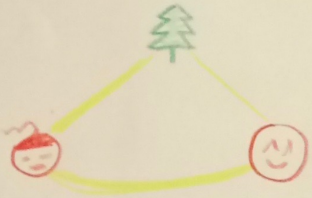
Tasks :

Layout



Operation

- When we click on one smiley, its relations will be highlighted



Focus

- focus on relations and correlation between smileys

Discussion

Pros	Cons
	- There're relations between smileys that is not coherent

- How about defining a particular event (ex: Christmas? New Year? Winter? Football?)

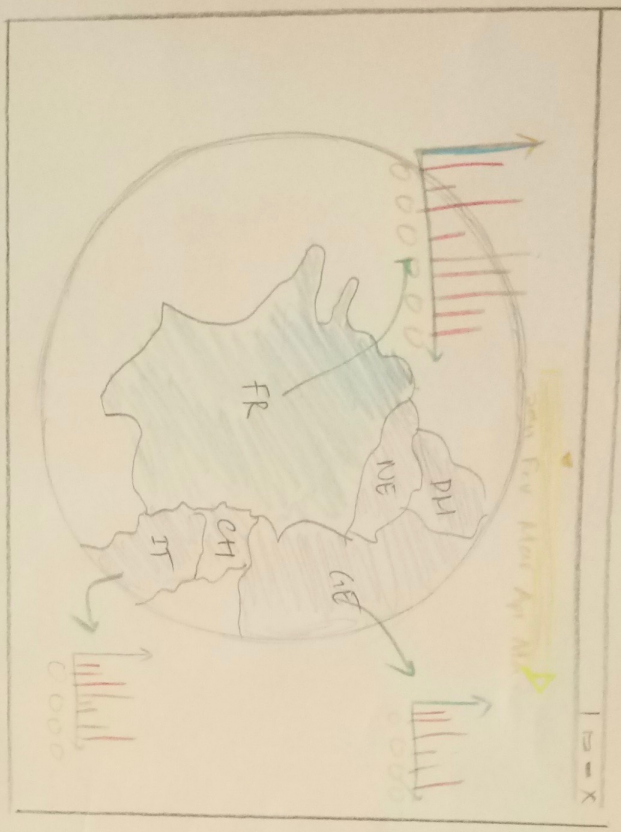
Title : Sheet no.4

Author : Nobinuti, Cheng

Date : 29 Nov, 30 Nov, 1 Dec

Tasks : fonctionalities

Layout



Operation

- Combine map and statistical result and timeline
- highlight selected country (specific play)

Discussion

Pros	Cons
- More information in on graphic	We cannot use realtime data for this visualizat

Focus

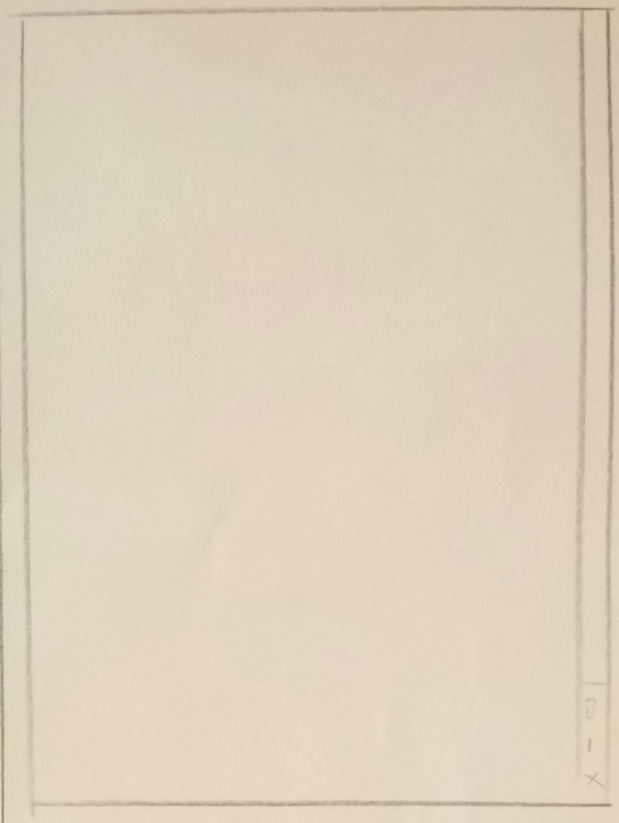
Title : Sheet No.5
Realization design

Author : Nobinulli, Cheng

Date : 29 Nov, 30 Nov, 1 Dec

Tasks :

Layout



Operation

Focus

Discussion / Detail

1. Algo : - [twitter API]
- [twitter Public Streaming API]

2. Dependencies : . MongoDB
. Node JS
. D3 JS

3. Timing :

4. Requirements: