

A Network Tour of Data Science

GeoTweet - Summary

Christian ABBET, Maxime BONHENBLUST, Nicolas MASSEREY

November 28, 2016

This project use geolocated tweets in the Swiss area during the past years. The data set is provided by Mr Michele Catasta (Applied Data Analysis lecturer). Data are described by Table1

This project aim to reconstruct mobility flow of user. The project will focus on finding home and work location for each user and estimate the distance traveled to go to work. The project will then classify the results comparing cities/countryside, cantons and Swiss/foreigner workers.

An other aim of this project is to locate people going to huge manifestation such as music festival, Salon de l'Auto or Foire du Valais. Then estimate the distance traveled to go to the event depending on its location in Switzerland or the canton the user is living in.

Name	Type	Type precision
id	bigint(20)	UNSIGNED
userId	bigint(20)	UNSIGNED
createdAt	timestamp	-
text	text	UTF-8
longitude	float	-
latitude	float	-
placeId	varchar(250)	UTF8
inReplyTo	bigint(20)	UNSIGNED
source	int(10)	UNSIGNED
truncated	bit(1)	-
placeLatitude	float	-
placeLongitude	float	-
sourceName	varchar(255)	UTF-8
sourceUrl	varchar(255)	UTF-8
userName	varchar(250)	UTF-8
screenName	varchar(250)	UTF-8
followersCount	int(10)	UNSIGNED
friendsCount	int(10)	UNSIGNED
statusesCount	int(10)	UNSIGNED
userLocation	varchar(200)	UTF-8

Table 1: **Data set data details**