

Project PEACH is a large scale open source community driven Data Science project by the Computer Science Department at University College London. The goal of Project PEACH is to provide a data science tool for medical professionals and researchers that can aid them in diagnostic and analytics processes through the use of big data, machine learning and data visualisation. <https://code-4-health.org/peach>

Project Supervisors: Dr. Dean Mohamedally, Dr. Atia Rafiq, Marios Constantinides, George Spithourakis

Team Members: Nikolaos Georgiadis, Narmin Huseynli, Efthymia Kazakou, Johann Kristthorsson, Mahalakshmi Sabapathi, Siri Vinay, Stephen Whelan

Project Contact: Stephen Whelan (Lead Architect, Template 1), stephen.whelan.15@ucl.ac.uk

The following document is a description of the file structure for the Social Media data set.

PEACH_hashtagLog.csv - (common delimited)

This file is a running output of user's tweets.

File Structure

| Field | Values | Description |
|--------------|---------------------------------|---|
| User Id | Integer | This field is a user's unique Id. Each user maintains their unique Id throughout all data sets in Project PEACH. |
| DateTime | DateTime (YYYYMMDD HH:MM:SS) | This field is the date and time the hashtag was tweeted by the user. |
| Hashtag | String | This value is the user's tweet. The tweets fall into the following broad categories {social/alcohol related, food, outdoors activities, pets, traveling}. |