

Universidad de Buenos Aires

Facultad De Ingeniería

Año 2020 - 1er Cuatrimestre

**ORGANIZACIóN DE DATOS (75.06)**

TRABAJO PRÁCTICO N.º 1

TEMA: Análisis exploratorio de datos

GRUPO: 23 – Barbijo De Uranio

INTEGRANTES:



Repositorio en GitHub:

<https://github.com/JulioCastillo1/TP1-Datos>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

75.06 – Organización de datos Curso 01 – Luis Argerich

INTRODUCCION

El objetivo de este trabajo práctico es realizar un análisis exploratorio del set de datos de la competencia de Kaggle: https://www.kaggle.com/c/nlp-getting-started​ .

El archivo train.csv cuenta con los siguientes campos:

* Id: identificador unico para cada tweet
* text: el texto del tweet
* location: Ubicación desde donde fue enviado (podría no estar)
* keyword: Un keyword para el tweet (podría faltar)
* target: En train.csv, indica si se trata de un desastre real (1) o no (0)

Se pretende descubrir cosas sobre los datos que puedan resultar interesantes.