

Retrasos de Rodalies Barcelona en Twitter

Una denuncia del servicio público de trenes en la comarca de Barcelona mediante redes sociales por Roberto Lara Martín



Fuente: <https://commons.wikimedia.org/w/index.php?curid=31801516>

Contexto

Este *dataset* fue obtenido de Twitter a modo de ejercicio para una asignatura de un máster. El objetivo del ejercicio era practicar el *web scraping* y el análisis de datos.

Contenido

El archivo, en formato CSV, contiene la cantidad de veces que un usuario de Twitter ha usado uno de los siguientes *hashtags*: #ROD1, #ROD2, #ROD3, #ROD4, #ROD7, #ROD8, #FGCVALL y #FGCLLO, que son utilizados para denunciar un problema del servicio en las líneas R1, R2, R3, R4, R7, R8, red de Metro del Vallès y red de Metro Llobregat-Anoia respectivamente. Además de los *hashtags*, contamos también los comunicados de incidencia del organismo

oficial (Gencat). Los mensajes fueron recogidos del 23/10/2017 al 29/10/2017. Aunque una semana pueda parecer poco, los resultados obtenidos son suficientes para establecer conclusiones acerca del servicio.

Así, en el fichero encontramos tres columnas:

- La primera, “línea”, contiene el nombre de la línea o red de Metro de Rodalies Barcelona, tal y como se detallan arriba.
- La segunda, “numero_tweets”, contiene el nombre total de denuncias mediante *hashtag* y comunicados de incidencia oficiales.
- La tercera, “numero_tweets_oficial”, contiene el nombre total de los comunicados de incidencia oficiales.

Agradecimientos

El nacimiento de esta batería de *hashtags* es dado por el proyecto [rodalia.info](http://www.rodalia.info)¹ que, junto a la implicación altruista de algunos usuarios de Rodalies Barcelona, ha construido un sistema de alertas vía Twitter “a tiempo real” sobre el estado del servicio. Así, mis agradecimientos van a todos ellos, por hacer posible este conjunto de datos.

Inspiración

Cada día miles de personas utilizan la red ferroviaria de la comarca de Barcelona, ya sean habitantes que van a trabajar o turistas que acaban de llegar al aeropuerto. Este transporte se convierte así en una herramienta indispensable para el desplazamiento de personas. Parece evidente pues que la puntualidad y la comunicación del estado de la línea sean los principales objetivos a cumplir, cosa que no se da en la realidad. Este *dataset*, mediante la ayuda de los usuarios finales, responde a la cuestión de cuál es la línea más conflictiva, la que menos, la que más se avisa y la que menos, con un fin de denuncia a los responsables y de alerta a los consumidores.

¹ <http://www.rodalia.info/es>

Licencia

Con el objetivo de llegar al máximo público posible y para reconocer el esfuerzo que me ha supuesto, este *dataset* se encuentra bajo licencia CC BY-SA 4.0.