



En este proyecto trabajaremos con preguntas sobre el lenguaje Python. Los datos están extraídos de stackoverflow, y se corresponden con una colección de preguntas desde 2008 hasta 2016 relacionadas con Python. La colección completa está disponible en [Kaggle datastes](#). Los datos con los que trabajaremos incluyen distintas informaciones sobre las preguntas y, a partir de ellos, generaremos una serie de informes y gráficas que resumirán aspectos relevantes de las temáticas más consultadas.

El formato de entrada es CSV. Cada registro del fichero de entrada ocupa una línea y contiene cuatro informaciones sobre las preguntas (puntuación, título, año y etiqueta principal). Estas son las primeras líneas de un fichero de entrada:

	A	B	C	D
1	score	title	year	tag
2	21	How can I find the full path to a font from its display name on a Mac?	2008	photoshop
3	27	Get a preview JPEG of a PDF on Windows?	2008	pdf
4	40	Continuous Integration System for a Python Codebase	2008	extreme-programming
5	25	cx_Oracle: How do I iterate over a result set?	2008	cx-oracle
6	28	Using 'in' to match an attribute of Python objects in an array	2008	iteration
7	30	Class views in Django	2008	oop
8	20	Python and MySQL	2008	bpgsql
9	256	How do I use Python's itertools.groupby()?	2008	iteration
10	364	Adding a Method to an Existing Object Instance	2008	monkeypatching
11	251	How do you express binary literals in Python?	2008	literals

Además de distintos indicadores, generaremos un par de salidas gráficas que mostrarán, respectivamente, la evolución y la distribución de uso de ciertas etiquetas:

