

## Introducción al Big Data

### Ejercicio U11\_E01

MasterD

## 1. TENDENCIAS EN BIG DATA

Como estamos en una unidad de introducción al Big Data, vamos a plantearte unos ejercicios sencillos consistentes simplemente en buscar información para ampliar algunos conceptos.

En este ejercicio en concreto vamos a tratar de estudiar cómo ha evolucionado el impacto del Big Data. A falta de otras fuentes o métricas más oficiales, vamos a basarnos en las estadísticas e indicadores de popularidad que publica Google a través de su portal Google Trends.

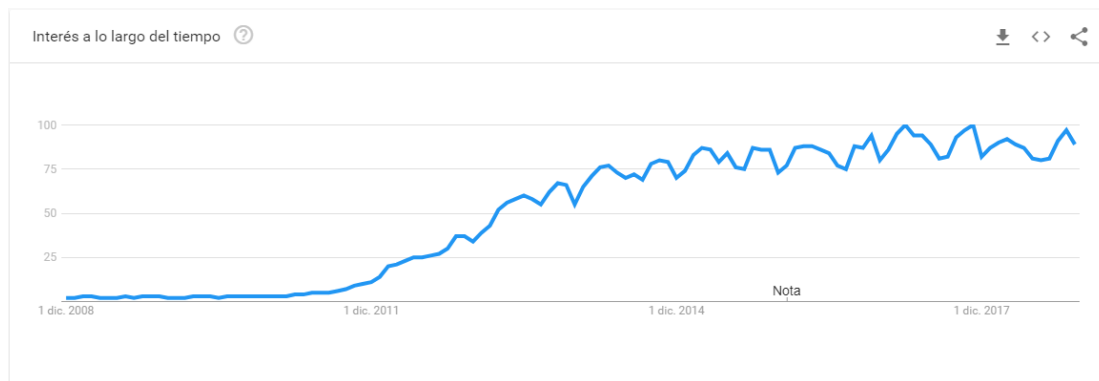
Accede a Google Trends y analiza cómo ha evolucionado el interés por el término 'big data'. Si te sale una lista de opciones, elige el tema 'Macrodatos' en español. Amplía el análisis a todo el mundo y el periodo a los últimos 10 años.

<https://trends.google.com/trends/explore?date=2008-11-20%202018-11-20&q=%2Fm%2F0bs2j8q>

Según la propia descripción de Google:

Los números reflejan el interés de búsqueda en relación con el valor máximo de un gráfico en una región y un periodo determinados. Un valor de 100 indica la popularidad máxima de un término, mientras que 50 y 0 indican que un término es la mitad de popular en relación con el valor máximo o que no había suficientes datos del término, respectivamente.

Es decir, al instante de mayor cantidad de consultas sobre el tema se le asigna un valor de 100. A partir de ahí, se escalan el resto de cantidades. De este modo, un periodo con un índice 20 representa que había un 80% menos de interés que en el periodo punta.

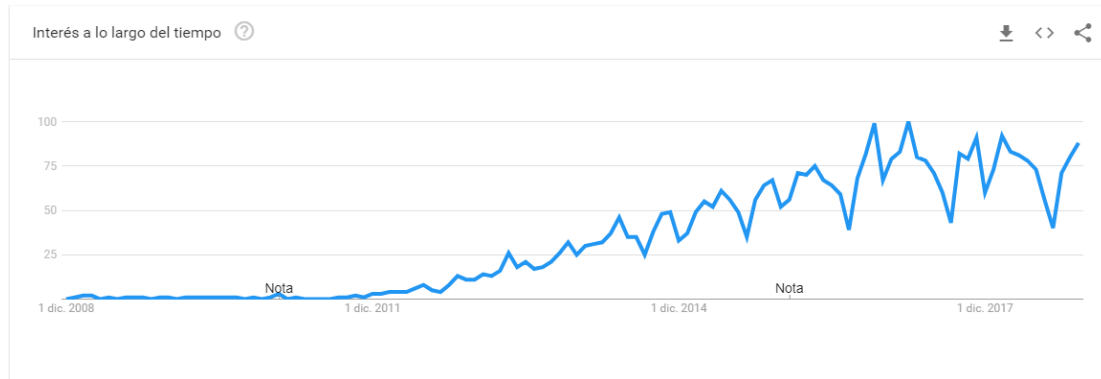


¿Identificas el periodo de mayor crecimiento? ¿Dirías que se ha estabilizado?

Según los datos, entre mitad de 2011 hasta inicios de 2015 se produjo el mayor crecimiento en el interés global por el Big Data. Desde entonces, la popularidad de este tema se mantiene más o menos estable.

Restringe el análisis ahora a España. ¿Crees que se ha alcanzado el mismo interés que el visto a nivel mundial?

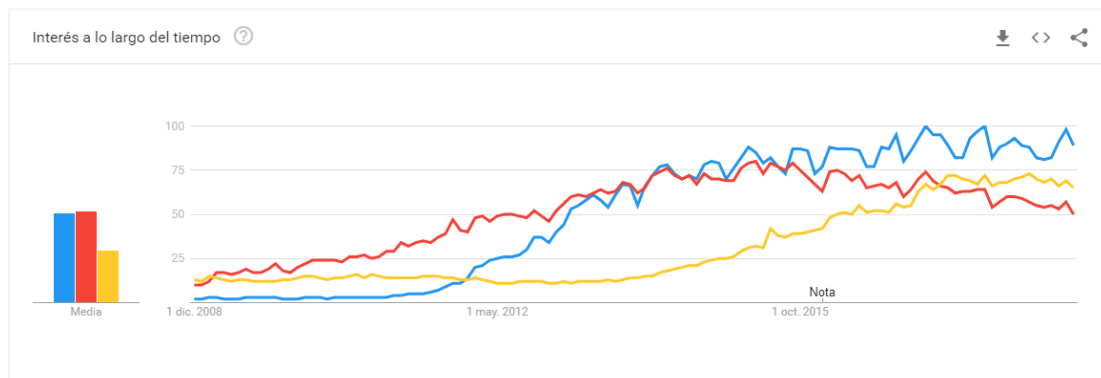
<https://trends.google.com/trends/explore?date=2008-11-20%202018-11-20&geo=ES&q=%2Fm%2F0bs2j8q>



En el caso de España, el despegue del interés por estas tecnologías es algo posterior y con un crecimiento menos pronunciado que a escala mundial. Podríamos decir que en España llegamos algo más tarde al fenómeno Big Data.

Vamos a añadir dos elementos más correspondientes a las tecnologías Big Data más extendidas (y que veremos en las siguientes unidades). Añade a la comparativa 'Apache Hadoop' y después 'Apache Spark'. Si te aparecen varias opciones, elige la correspondiente a la categoría de software.

<https://trends.google.com/trends/explore?date=2008-11-20%202018-11-20&q=%2Fm%2F0bs2j8q,%2Fm%2F0fdjtq,%2Fm%2F0ndhxqz>



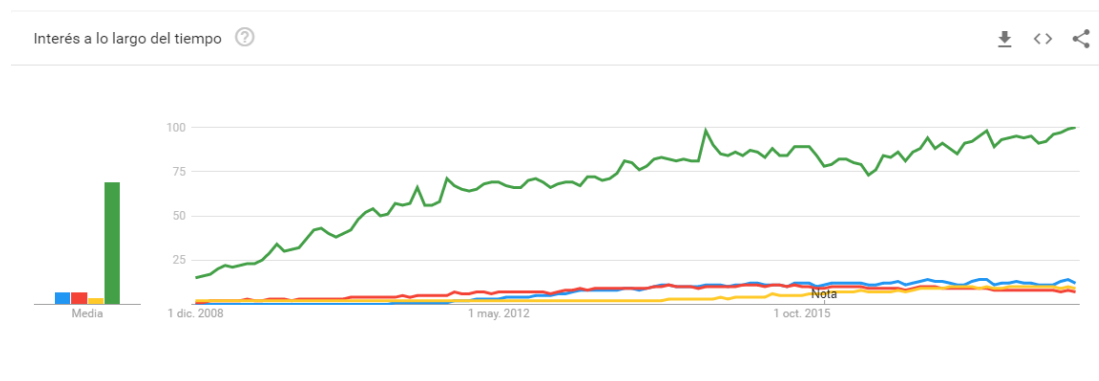
¿Qué te parece la evolución del interés por Hadoop? ¿Y la de Spark? ¿Cuál dirías que es la tecnología que va a liderar el campo Big Data próximamente?

A tenor de la gráfica, podemos señalar que el periodo de máximo interés por Hadoop se produjo entre 2013 y 2016 (aproximadamente). A partir de ahí, esta tecnología empezó a decaer un poco en favor de nuevas tecnologías, como es el

caso de Spark, que sobrepasó a Hadoop en relevancia en mayo de 2017. Es algo habitual en el mundo de las tecnologías de la información. Las tecnologías novedosas suelen experimentar una suerte de *boom*, diríamos que se ponen de moda y captan la atención de desarrolladores y empresas. Pero después la espuma baja. Pero no significa que Hadoop vaya a desaparecer, simplemente ocurre que el interés se queda en los ámbitos y usuarios que realmente se benefician y utilizan la tecnología. Hadoop goza de muy buena salud y todo indica que seguirá así por un buen tiempo. Ahora es Spark quien ha alcanzado gran popularidad, pero también terminará ajustándose. Y otra nueva tecnología se aupará a lo alto de la ola...

Para terminar, vamos a añadir el *cloud computing* a la comparativa. Incluye el tema 'Computación en la nube'.

[https://trends.google.com/trends/explore?date=2008-11-20%202018-11-20&q=%2Fm%2F0bs2j8q,%2Fm%2F0fdjtq,%2Fm%2F0ndhxqz,%2Fm%2F02y\\_9m3](https://trends.google.com/trends/explore?date=2008-11-20%202018-11-20&q=%2Fm%2F0bs2j8q,%2Fm%2F0fdjtq,%2Fm%2F0ndhxqz,%2Fm%2F02y_9m3)



Comparado con el resto de temas, ¿qué te parece el interés por el *cloud computing*? ¿Qué motivos se te ocurren para la diferencia con el resto de conceptos?

Como ves, la diferencia es aplastante. El *cloud computing* parece arrasar en interés, multiplicando casi por 10 en algunos momentos al resto de temas. La razón en realidad es que la computación en la nube es un concepto muy amplio, que abarca gran cantidad de servicios y tecnologías: cuentas de almacenamiento, correo, fotografías, redes sociales... Todo está en la nube. En la última unidad del curso te explicaremos qué hay exactamente detrás de la nube, y cómo se relaciona con las tecnologías Big Data.