Sistemas Inteligentes para la Gestión de la Empresa

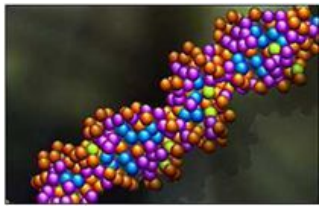
2016 - 2017

- **Tema 1. Introducción a la Ciencia de Datos**
- **Tema 2. Depuración y Calidad de Datos**
- **Tema 3. Análisis Predictivo para la Empresa**
- **Tema 4. Análisis de Transacciones y Mercados**
- **Tema 5. Modelos avanzados de Analítica de Empresa**
- **Tema 6. Big Data**
- **Tema 7. Aplicaciones de la Ciencia de Datos en la Empresa**

Ciencia de Datos, Minería de Datos, Big Data

Nuestro mundo gira en torno a los datos

- Ciencia
  - Bases de datos de astronomía, genómica, datos medio-ambientales, datos de transporte, …



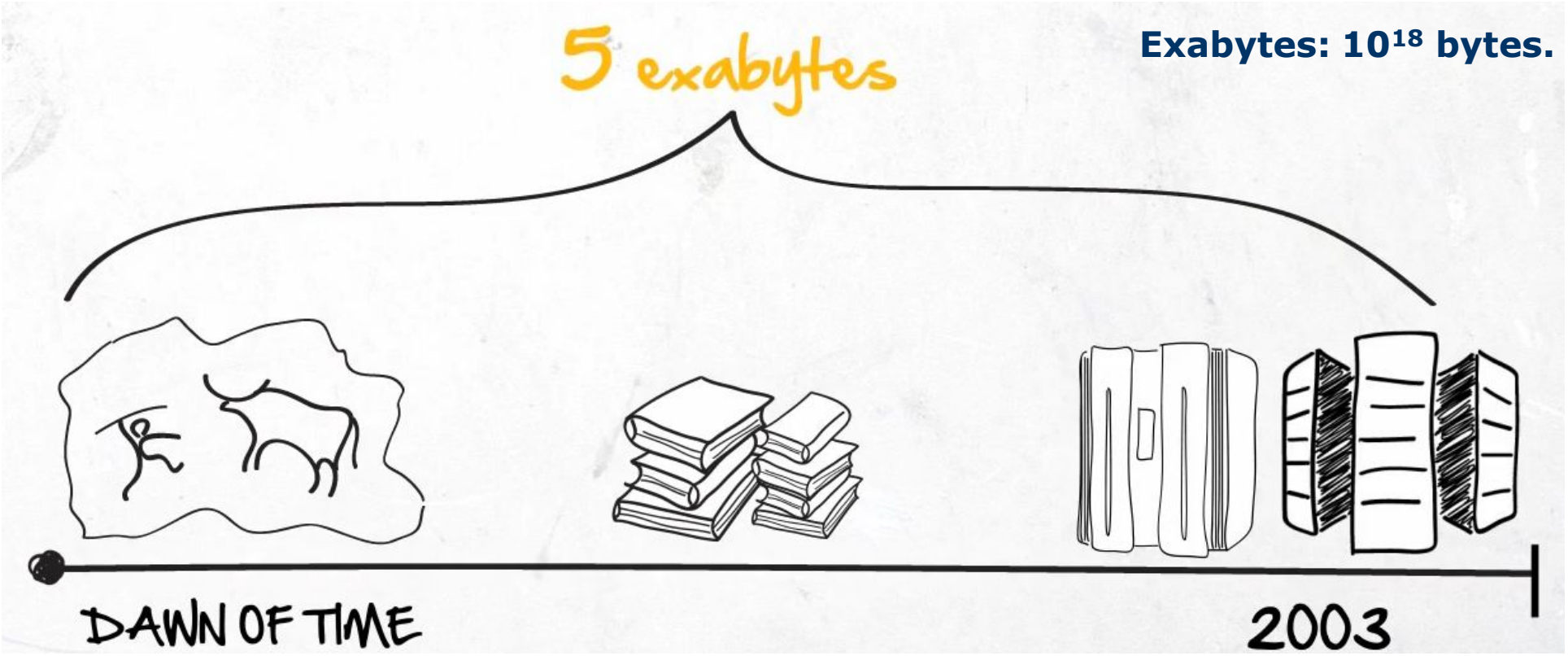
- Ciencias Sociales y Humanidades
  - Libros escaneados, documentos históricos, datos sociales, …

- Negocio y Comercio
  - Ventas de corporaciones, transacciones de mercados, censos, tráfico de aerolíneas, …

- Entretenimiento y Ocio
  - Imágenes en internet, películas, ficheros MP3, …
- Medicina
  - Datos de pacientes, datos de escaner, radiografías …

- Industria, Energía, …
  - Sensores, …

La explosión de los datos

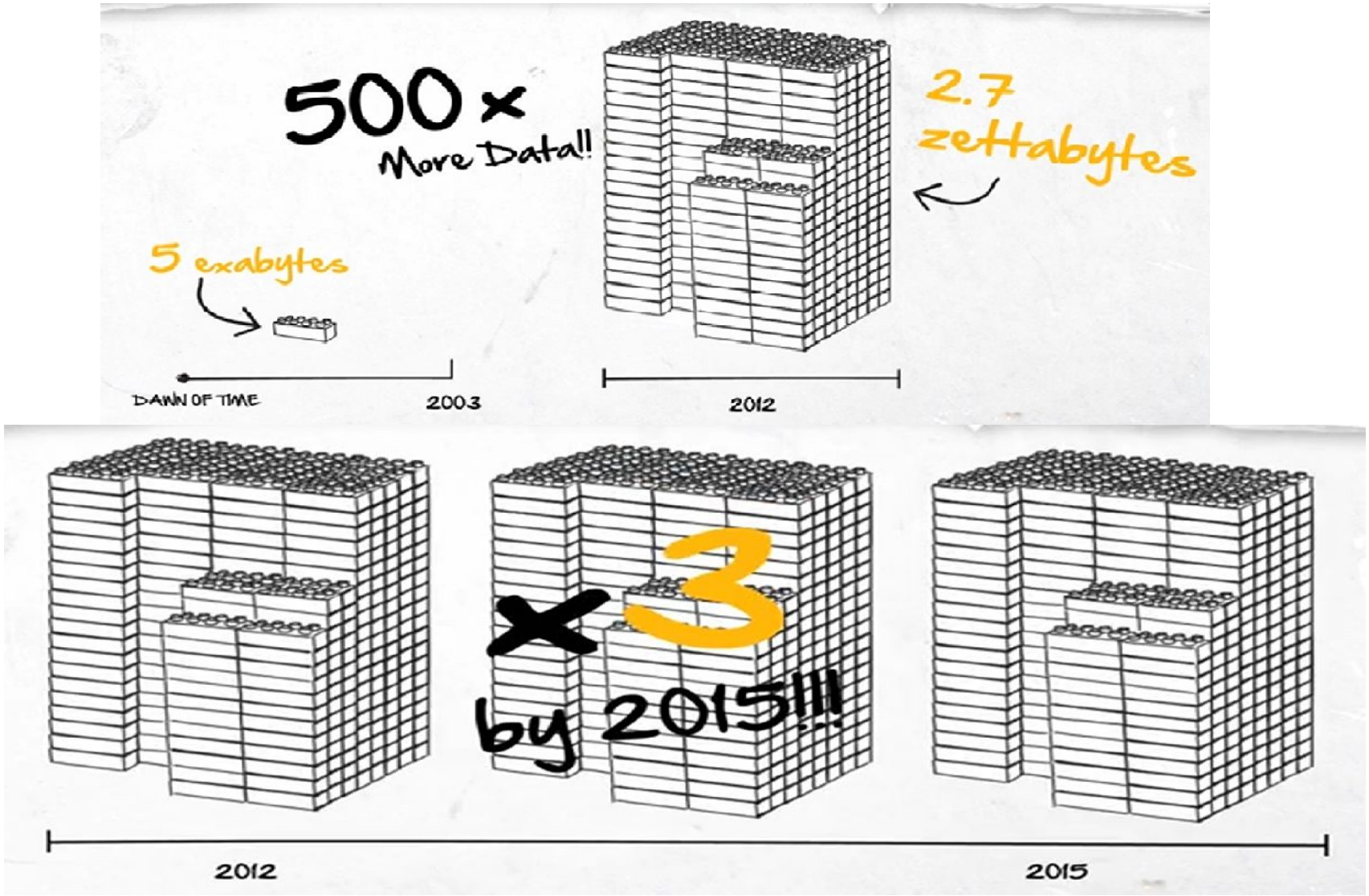


Exabytes: 10^{18} bytes.

1 EB = 10^3 PB = 10^6 TB = 10^9 GB = 10^{12} MB = 10^{15} KB = 10^{18} bytes.

La explosión de los datos

Zettabyes: 10^{21} bytes.



El problema de la explosión de información:

- existencia de herramientas para la recolección de información
- madurez de la tecnología de bases de datos
- bajo precio del hardware

➔ cantidades gigantescas de datos almacenados en bases de datos, *data warehouses* y otros tipos de almacenes de información

Somos ricos en datos pero pobres en conocimiento

El progreso y la innovación ya no se ven obstaculizados por la capacidad de recopilar datos, sino por la capacidad de gestionar, analizar, sintetizar, visualizar, y descubrir el conocimiento de los datos recopilados de manera oportuna y en una forma escalable

Ciencia de Datos, Minería de Datos, Big Data

ELMUNDO.es
Líder mundial en español | Miércoles 04/09/2013. Actualizado 16:27h.

Alex ' Sandy' Pentland, director del programa de emprendedores del 'Media Lab' del Massachusetts Institute of Technology (MIT)

INTERNET | Campus Party Europa 2013

'Es la década de los datos y de ahí vendrá la revolución'

Alex 'Sandy' Pentland, durante su ponencia. | M. Sáinz

Considerado por 'Forbes' como uno de los siete científicos de datos más poderosos del mundo

Ben Chams - Fotolia

http://www.elmundo.es/elmundo/2013/09/03/navegante/1378

Índice

- ¿Qué es la Ciencia de Datos?
- El poder de los datos y su impacto en nuestra sociedad
- Herramientas y Lenguajes en Ciencia de Datos. Repositorio de Kaggle
- Comentarios Finales

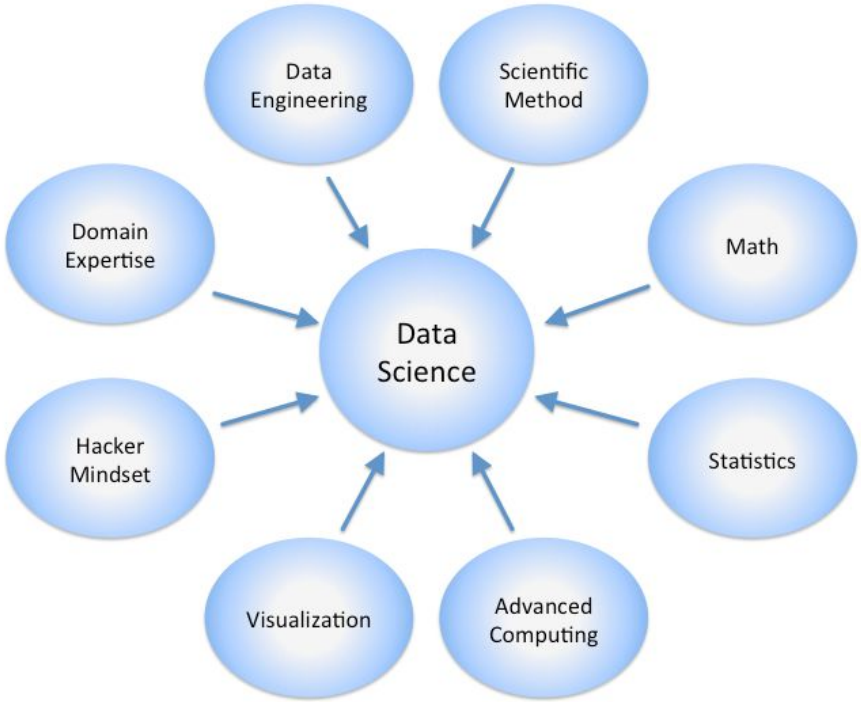
Índice

- **¿Qué es la Ciencia de Datos?**
- El poder de los datos y su impacto en nuestra sociedad
- Herramientas y Lenguajes en Ciencia de Datos. Repositorio de Kaggle
- Comentarios Finales

Ciencia de Datos

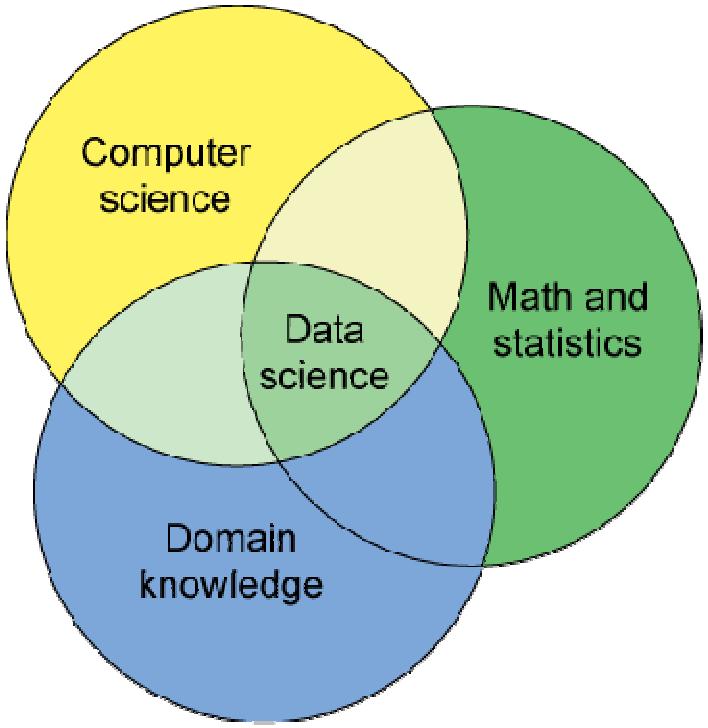
Data Science

Ciencia de Datos es el ámbito de conocimiento que engloba las habilidades asociadas a la extracción de conocimiento de datos, incluyendo Big Data

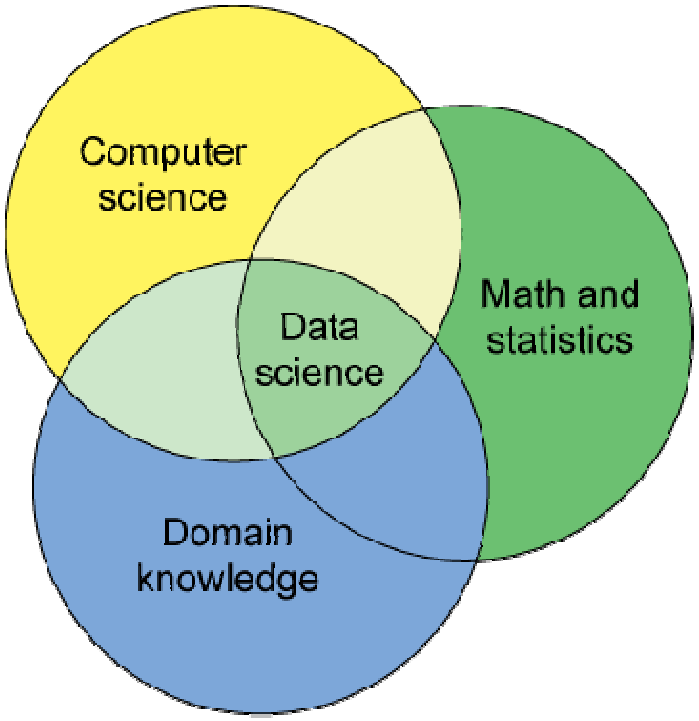


Ciencia de Datos

Data Science o la Ciencia de Datos incorpora diferentes elementos y se basa en las técnicas y teorías de muchos campos, incluyendo las matemáticas, estadística, ingeniería de datos, reconocimiento de patrones y aprendizaje, computación avanzada, visualización, modelado de la incertidumbre, almacenamiento de datos y la informática de alto rendimiento con el objetivo de extraer el significado de datos y la creación de productos de datos.



Ciencia de Datos

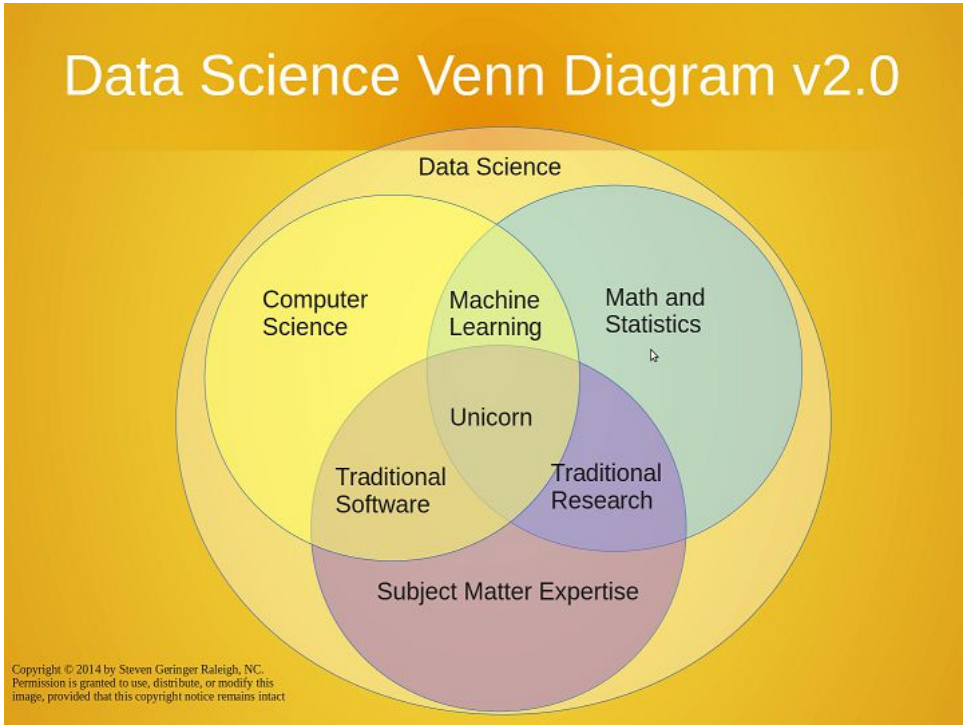


Es un término relativamente nuevo que se utiliza a menudo de manera intercambiable con analítica de negocio. La ciencia de datos busca utilizar todos los datos disponibles y relevantes para "extraer conocimiento" que pueda ser fácilmente comprendido por los expertos en el área de aplicación. Un experto de la ciencia de datos se denomina un científico de datos.

Ciencia de Datos

¿Qué es un Científico de Datos?

Un científico de datos es un profesional que debe dominar las ciencias matemáticas y la estadística, acabados conocimientos de programación (y sus múltiples lenguajes), ciencias de la computación y analítica.



Ciencia de Datos

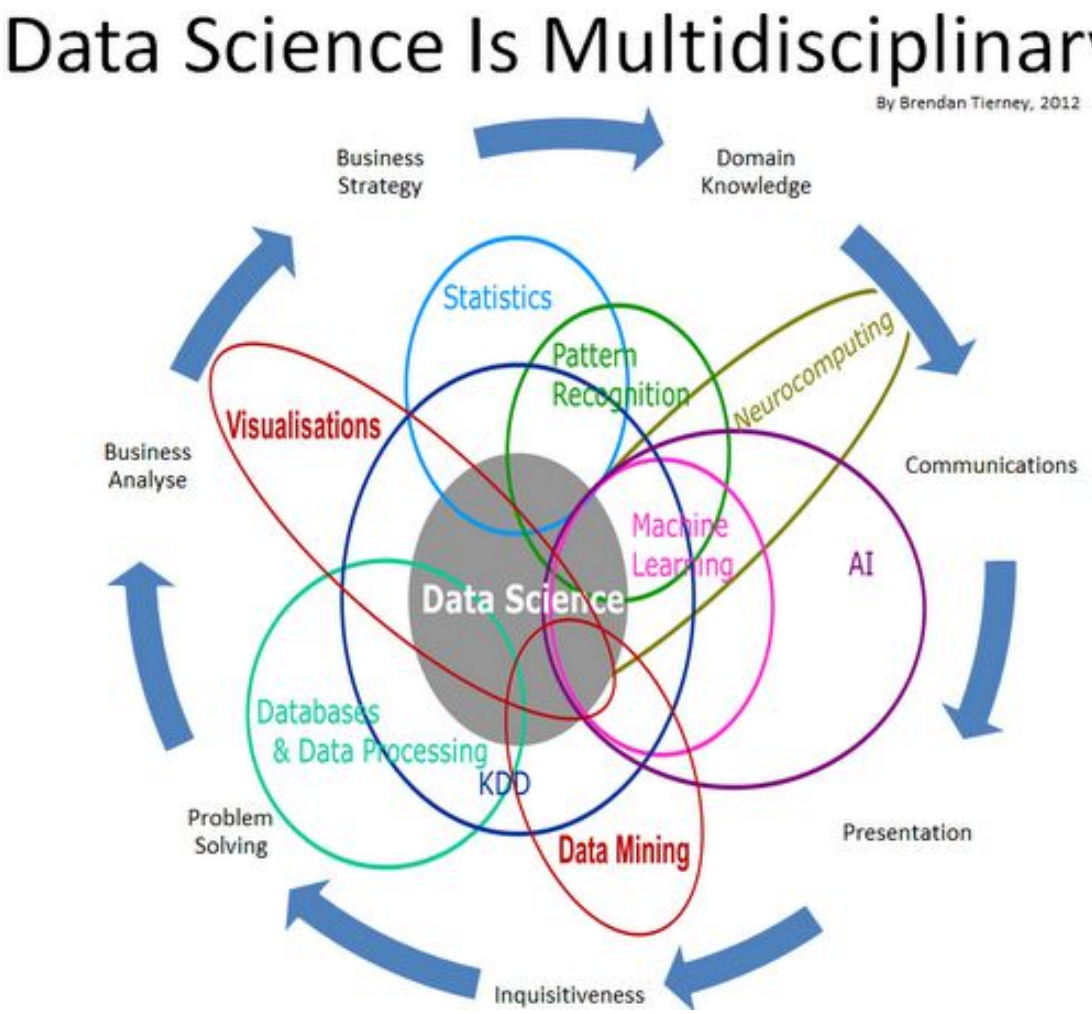
José Antonio Guerrero: uno de los mejores científicos de datos del mundo (Plataforma Kaggle)

¿Qué es un científico de datos?

"Es una persona con fundamentos en matemáticas, estadística y métodos de optimización, con conocimientos en lenguajes de programación y que además tiene una experiencia práctica en el análisis de datos reales y la elaboración de modelos predictivos.
De las tres características quizás la más difícil es la tercera; no en vano la modelización de los datos se ha definido en ocasiones como un arte. Aquí no hay reglas de oro, y cada conjunto de datos es un lienzo en blanco."

Leer más: http://www.elconfidencial.com/tecnologia/2013-12-19/un-matematico-andaluz-desconocido-es-el-mejor-cientifico-de-datos-del-mundo_67675/

Ciencia de Datos

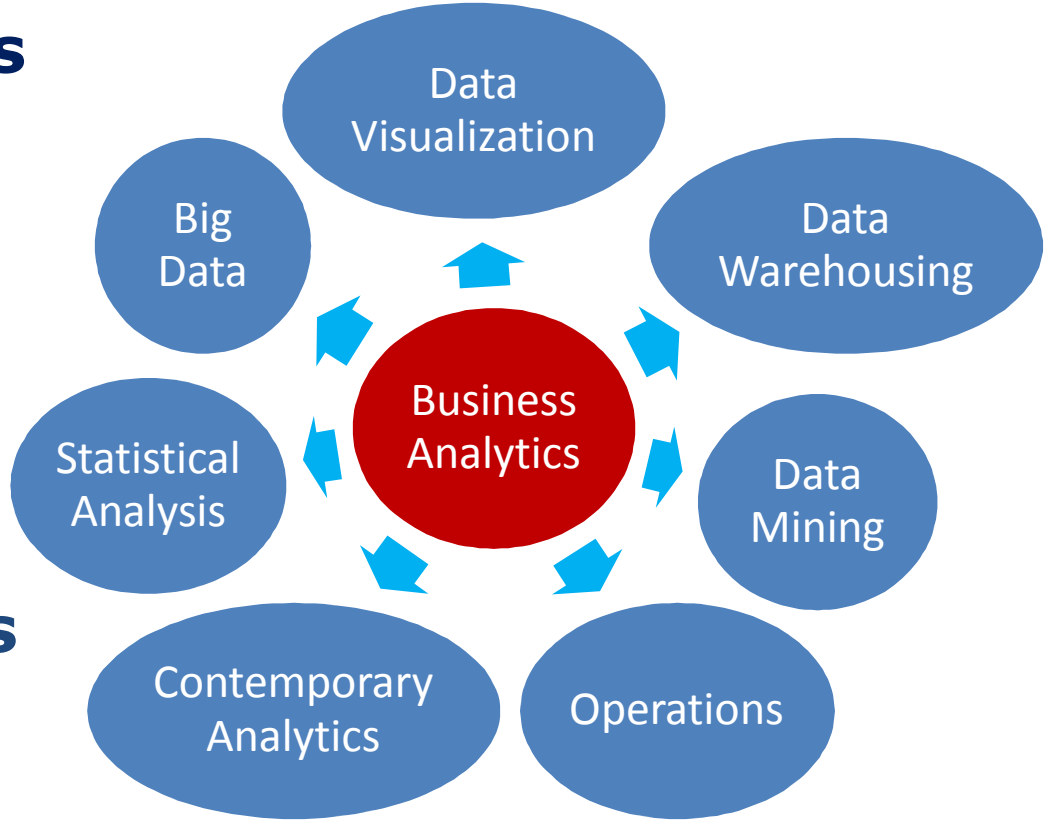


Business Analytics

Data Science

=

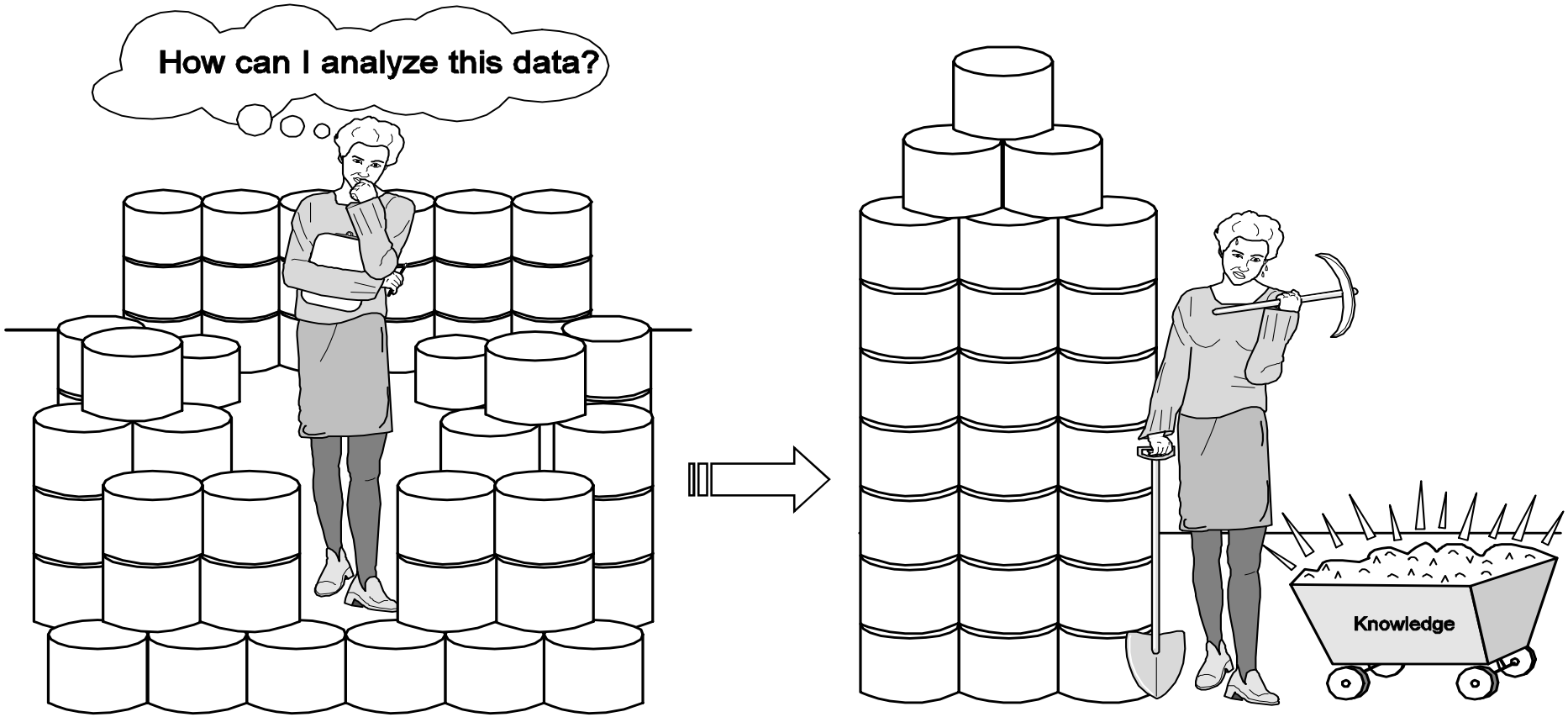
Business Analytics



Data mining: Data Preprocessing, Supervised learning, unsupervised learning, forecasting
Contemporary Analytics: text mining, network analytics, social analytics, customer analytics, web analytics, risk analytics, information retrieval and recommendations
Statistical Analysis: Estimation and inference; and regression models
Operations: Simulation and optimization

Ciencia de Datos

Minería de Datos



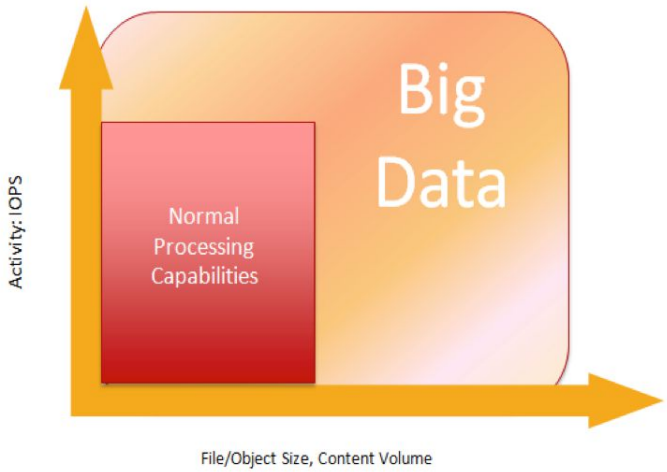
We have rich data,
but poor information

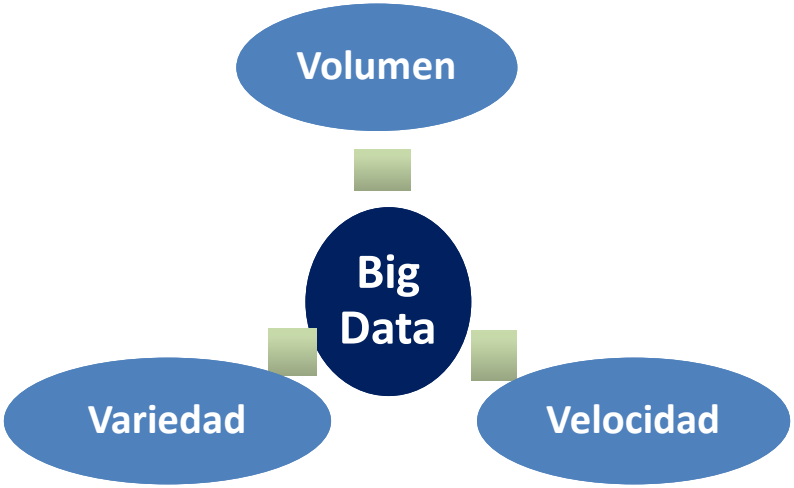
Data mining-searching for knowledge
(interesting patterns) in your data.

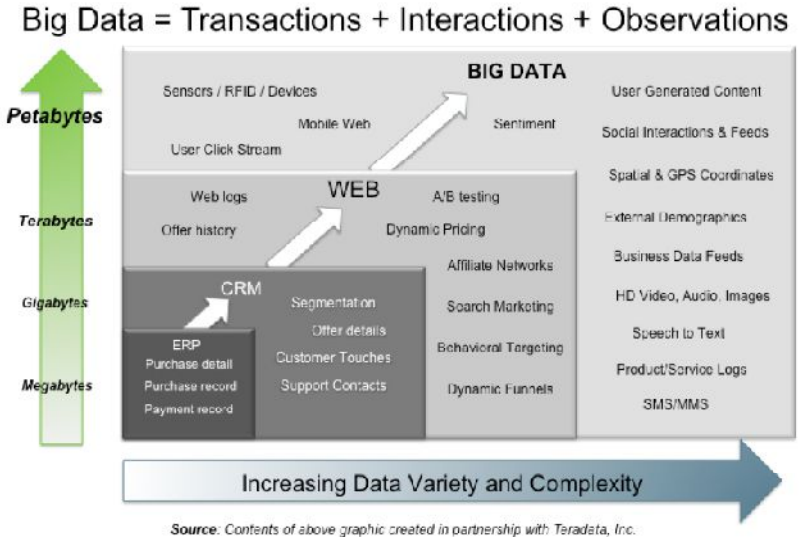
Ciencia de Datos

Big Data

"***Big Data***" son datos cuyo volumen, diversidad y complejidad **requieren nueva arquitectura, técnicas, algoritmos y análisis** para gestionar y extraer valor y conocimiento oculto en ellos ...







Índice

- ¿Qué es la Ciencia de Datos?
- **El poder de los datos y su impacto en nuestra sociedad**
- Herramientas y Lenguajes en Ciencia de Datos. Repositorio de Kaggle
- Comentarios Finales

Análisis de asociaciones

Pañales y cerveza. Ficción y leyenda para ilustrar el análisis de transacciones

Si compro cerveza, entonces compro pañales 60%

Si compro pañales, entonces compro cerveza 100%

El poder de los datos

Análisis de asociaciones en transacciones de tarjetas de crédito

Fuente: Big Data. La revolución de los datos masivos. Pag. 160

El poder de los datos

Análisis de transacciones

Fuente: Big Data. La revolución de los datos masivos. Pag. 77

El poder de los datos

Análisis de transacciones: Un chivo expiatorio

Target (cadena de grandes almacenes) que utiliza el análisis de transacciones y asociaciones.

Unos días después el director llamó al padre para disculparse.

Respuesta conciliadora del padre:

"He estado hablando con mi hija –dijo el padre– Resulta que en mi casa han tenido lugar ciertas actividades de las que yo no estaba del todo informado. Mi hija sale de cuentas en agosto. Soy yo el que les debe una disculpa".

Fuente: Big Data. La revolución de los datos masivos. Pag. 77.

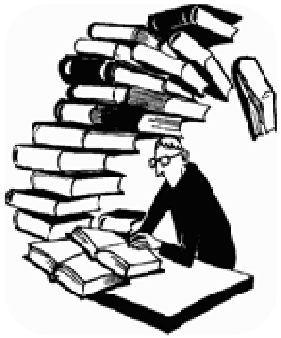
El poder de los datos

Análisis de transacciones

Amazon: Sistema de recomendación

Críticos y editores literarios

La voz de Amazon (1995)



Greg Linde (1997) propuso un sistema de recomendaciones, filtrado colaborativo

"artículo a artículo"

Dilema: ¿Lo que los clics decían o lo que opinaban los críticos?

Los datos incrementaron tremendamente las ventas

Ahora más de 1/3 de las ventas son gracias a las recomendaciones

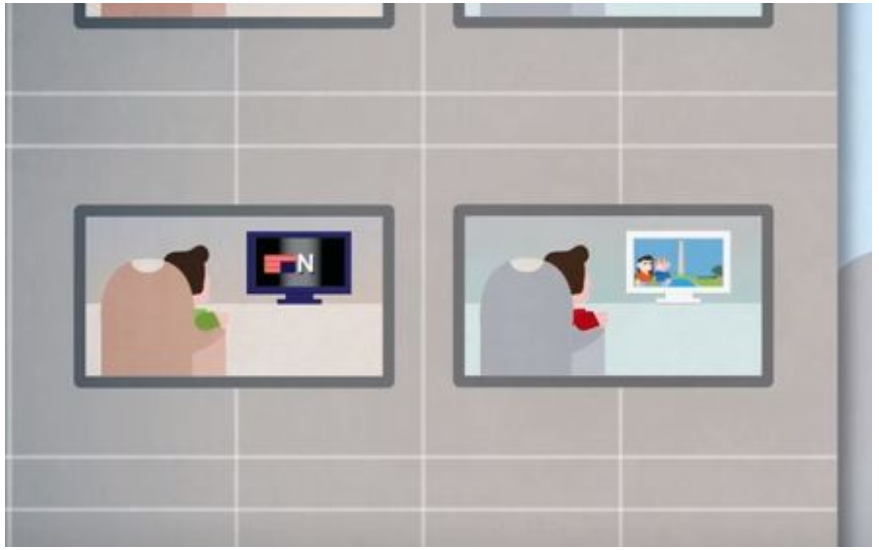
Fuente: Big Data. La revolución de los datos masivos. Pag. 69.

El poder de los datos

Netflix: Sistema de recomendación

Para Netflix, compañía de alquiler de películas online, las tres cuartas partes de los pedidos nuevos surgen de las recomendaciones.

Netflix y Amazon son dos empresas cuyo plan de negocio está basada en big data y sistemas de recomendación



Analizando Twitter para medir la Salud Pública

You Are What You Tweet

Un sistema de filtrado de datos de Twitter puede inferir aspectos de salud analizando 144M de tuits (2011-2013)

Se obtienen 13 grupos coherentes de mensajes correlacionados

- **Gripe estacional (r= 0.689) y alergias (r = 0.810)**
- **Ejercicio y obesidad relacionados con datos geográficos, ..**

Discovering Health Topics in Social Media Using Topic Models
Michael J. Paul, Mark Dredze, Johns Hopkins University, Plos One 9(8)
e103408, 2014
doi:10.1371/journal.pone.0103408

Banca: Identificación de personas con las compras de tarjetas de crédito

http://elpais.com/elpais/2015/01/29/ciencia/1422520042_066660.h

PRIVACIDAD EN INTERNET »

Cuatro compras con la tarjeta bastan para identificar a cualquier persona

- Los patrones de uso de las tarjetas permiten descubrir la identidad del 90% de una muestra de 1,1 millones de personas anónimas, según demuestra un estudio del MIT

Science

AAAS

Unique in the shopping mall: On the reidentifiability of credit card metadata

Yves-Alexandre de Montjoye,[1]* Laura Radaelli,[2] Vivek Kumar Singh,[1,3] Alex "Sandy" Pentland[1]

Large-scale data sets of human behavior have the potential to fundamentally transform the way we fight diseases, design cities, or perform research. Metadata, however, contain sensitive information. Understanding the privacy of these data sets is key to their broad use and, ultimately, their impact. We study 3 months of credit card records for 1.1 million people and show that four spatiotemporal points are enough to uniquely reidentify 90% of individuals. We show that knowing the price of a transaction increases the risk of reidentification by 22%, on average. Finally, we show that even data sets that provide coarse information at any or all of the dimensions provide little anonymity and that women are more reidentifiable than men in credit card metadata.

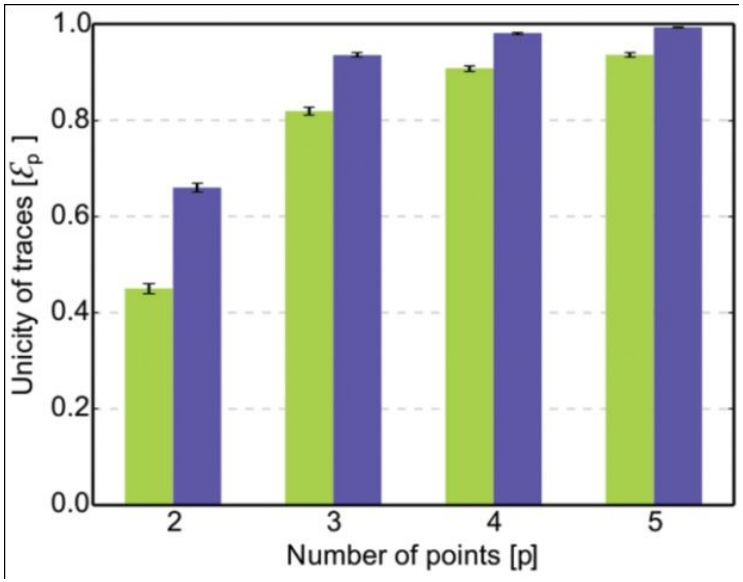
http://www.sciencemag.org/content/347/6221/536

Banca: Identificación de personas con las compras de tarjetas de crédito

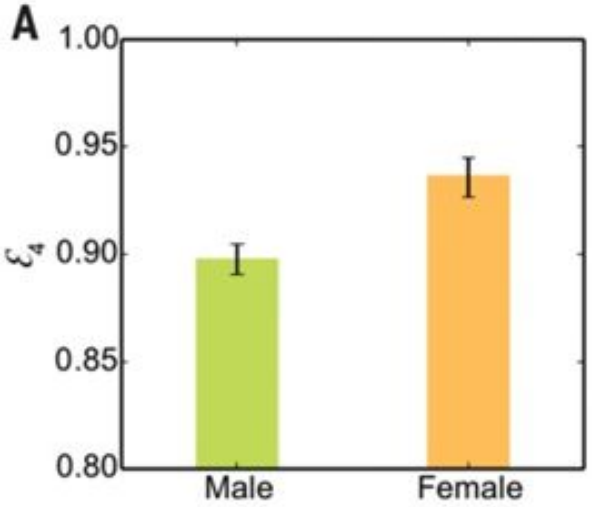
PRIVACIDAD EN INTERNET »

Cuatro compras con la tarjeta bastan para identificar a cualquier persona

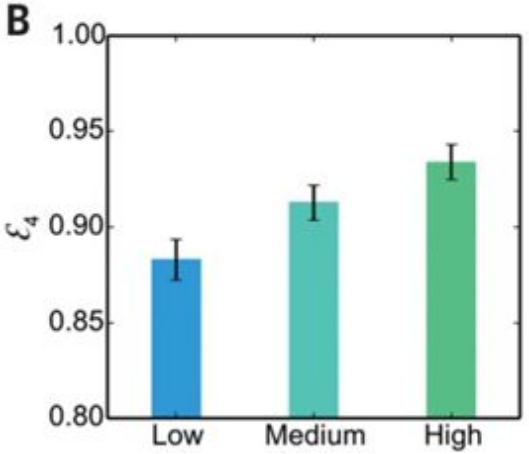
- Los patrones de uso de las tarjetas permiten descubrir la identidad del 90% de una muestra de 1,1 millones de personas anónimas, según demuestra un estudio del MIT



Identificación por el número de compras



Identificación por el género



Identificación por el poder adquisitivo

Impacto en la sociedad

Salud
Redes sociales como fuente de datos
Industria, comercio, banca, …
Ocio y cultural (Ej. Recomendaciones)
Política, Bien social (Social good)

Gran Impacto en la Sociedad y presencia en los medios de comunicación

EL MUNDO

Líder mundial en español | Miércoles 04/09/2013. Act

INTERNET | Campus Party Europa 201

'Es la década d
ahí vendrá la re

Alex 'Sandy' Pentland, durante su ponenci

http://www.elmundo.es/elm

EL PAÍS

ECONOMÍA

ECONOMÍA EMPRESAS MERCADOS BOL

El maná de los datos

- La conversión de datos en información útil par
millones de dólares en 2015. La herramienta '

SUSANA BLÁZQUEZ | Madrid | 29 SEP 2013 - 01:0

Archivado en: Citigroup Cap Gemini Sogeti S/
IBM Telefónica Aplicaciones informáticas Tecnol

EL PAIS

ECONOMÍA

ECONOMÍA EMPRESAS MERCADOS BOLSA MIS AHORROS VIVIENDA TECNOLOGÍA OF

EMPRENDEDORES »

El Big Data echa una mano al campo

• Una empresa española recoge miles de datos para predecir las cosechas

MARÍA FERNÁNDEZ | 30 NOV 2014 - 00:00 CET

Archivado en: Bases datos Emprendedores Aplicaciones informáticas Empresas
Programas informáticos Economía Informática Industria

Impacto Económico

La demanda de profesionales formados en Ciencia de Datos y *Big Data* es enorme.

Se estima que la conversión de datos en información útil generó un mercado de 132.000 millones de dólares en 2015 y que se crearán más de 4.4 millones de empleos.

España necesitaba para 2015 más de 60.000 profesionales con formación en Ciencia de Datos y *Big Data*.

España necesitará 60.000 profesionales de Big Data hasta 2015

22 octubre, 2013 Eventos 18

España necesitará 60.0

Toledo.

EL PAÍS PORTADA INTERNACIONAL POL

ECONOMÍA

ECONOMÍA EMPRESAS MERCADOS BOLSA FINANZAS PERSONALES VIVIENDA TECNOLOGÍ

ESTÁ PASANDO Multa a la banca Revuelo en Hacienda Eléctricas y renovables Paro

El maná de los datos

- La conversión de datos en información útil para las empresas generará un mercado de 132.000 millones de dólares en 2015. La herramienta 'big data' sacará del mercado a quien no la use

SUSANA BLÁZQUEZ | Madrid | 29 SEP 2013 - 01:00 CET 10

Archivado en: Citigroup Cap Gemini Sogeti SAP Oracle ING Bank BBVA Mapfre Bases datos IBM Telefónica Aplicaciones informáticas Tecnología Empresas Programas informáticos Economía

http://economia.elpais.com/economia/2013/09/27/actualidad/1380283725_938376.html

http://www.revistacloudcomputing.com/2013/10/espana-necesitara-60-000-profesionales-de-big-data-hasta-2015/?goback=.gde_4377072_member_5811011886832984067#!

Índice

- ¿Qué es la Ciencia de Datos?
- El poder de los datos y su impacto en nuestra sociedad
- **Herramientas y Lenguajes en Ciencia de Datos. Repositorio de Kaggle**
- Comentarios Finales

Herramientas, Lenguajes, Kaggle

Una web sobre el software libre para Ciencia de Datos …

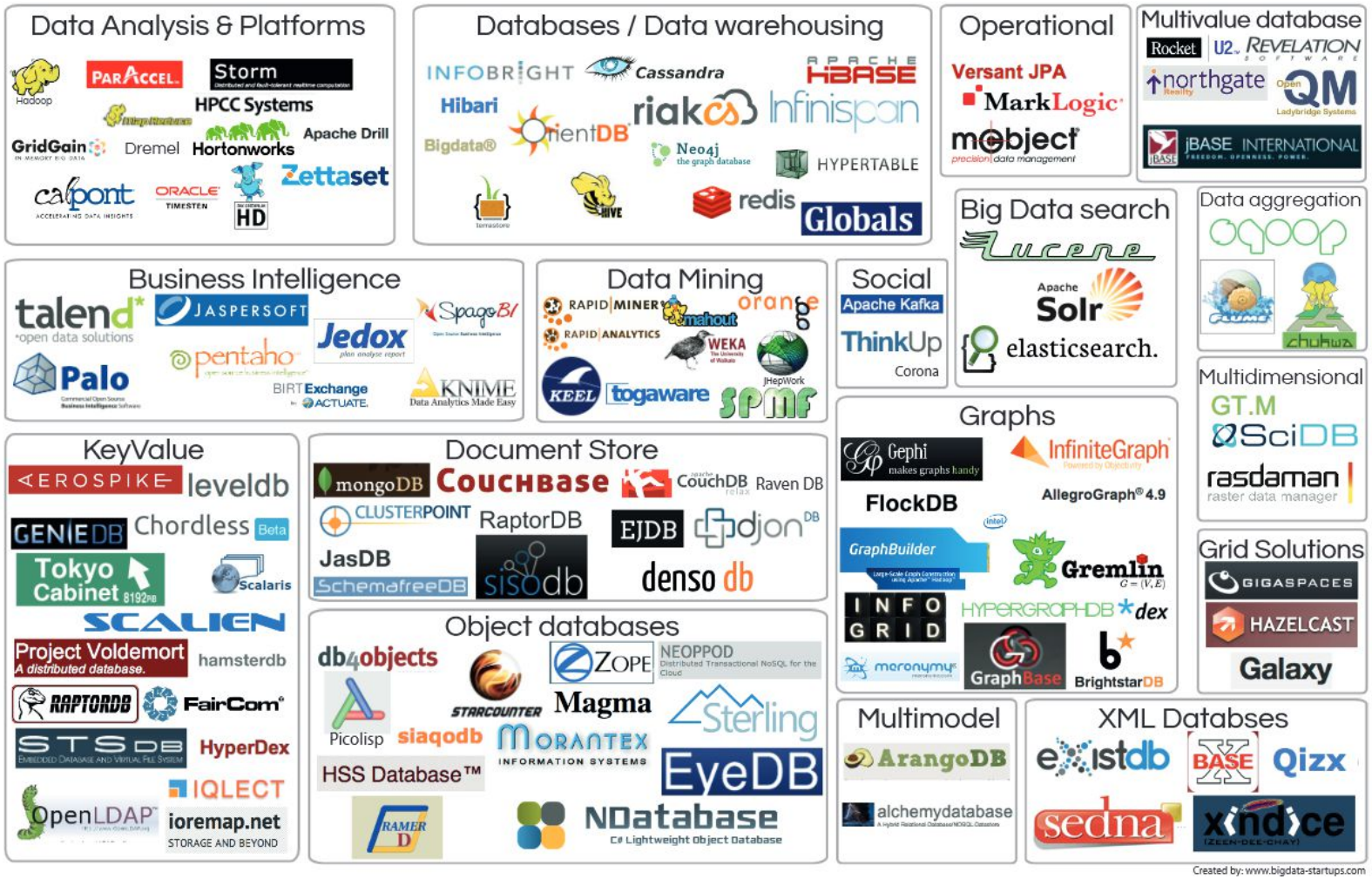
Software (open source tools)

http://www.bigdata-startups.com/open-source-tools/

Herramientas, Lenguajes, Kaggle

Una web sobre el software libre para Ciencia de Datos …

http://www.bigdata-startups.com/open-source-tools/



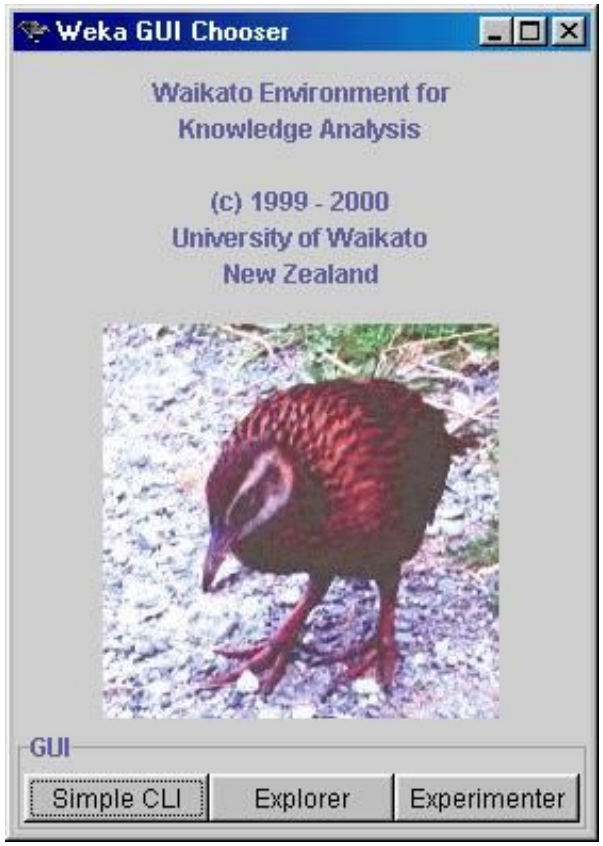
Herramientas, Lenguajes, Kaggle

KNIME (o Konstanz Information Miner) es una plataforma de minería de datos que permite el desarrollo de modelos en un entorno visual. KNIME está desarrollado sobre la plataforma Eclipse y programado, esencialmente, en java.

Fue desarrollado originalmente en el departamento de bioinformática y minería de datos de la Universidad de Constanza, Alemania, bajo la supervisión del profesor Michael Berthold. En la actualidad, la empresa KNIME.com GmbH, radicada en Zúrich, Suiza, continúa su desarrollo además de prestar servicios de formación y consultoría.

https://www.knime.org/

Herramientas, Lenguajes, Kaggle



Weka

- **The University of Waikato, New Zealand**
- **Machine learning software in Java implementation**

http://www.cs.waikato.ac.nz/ml/weka/

Herramientas, Lenguajes, Kaggle

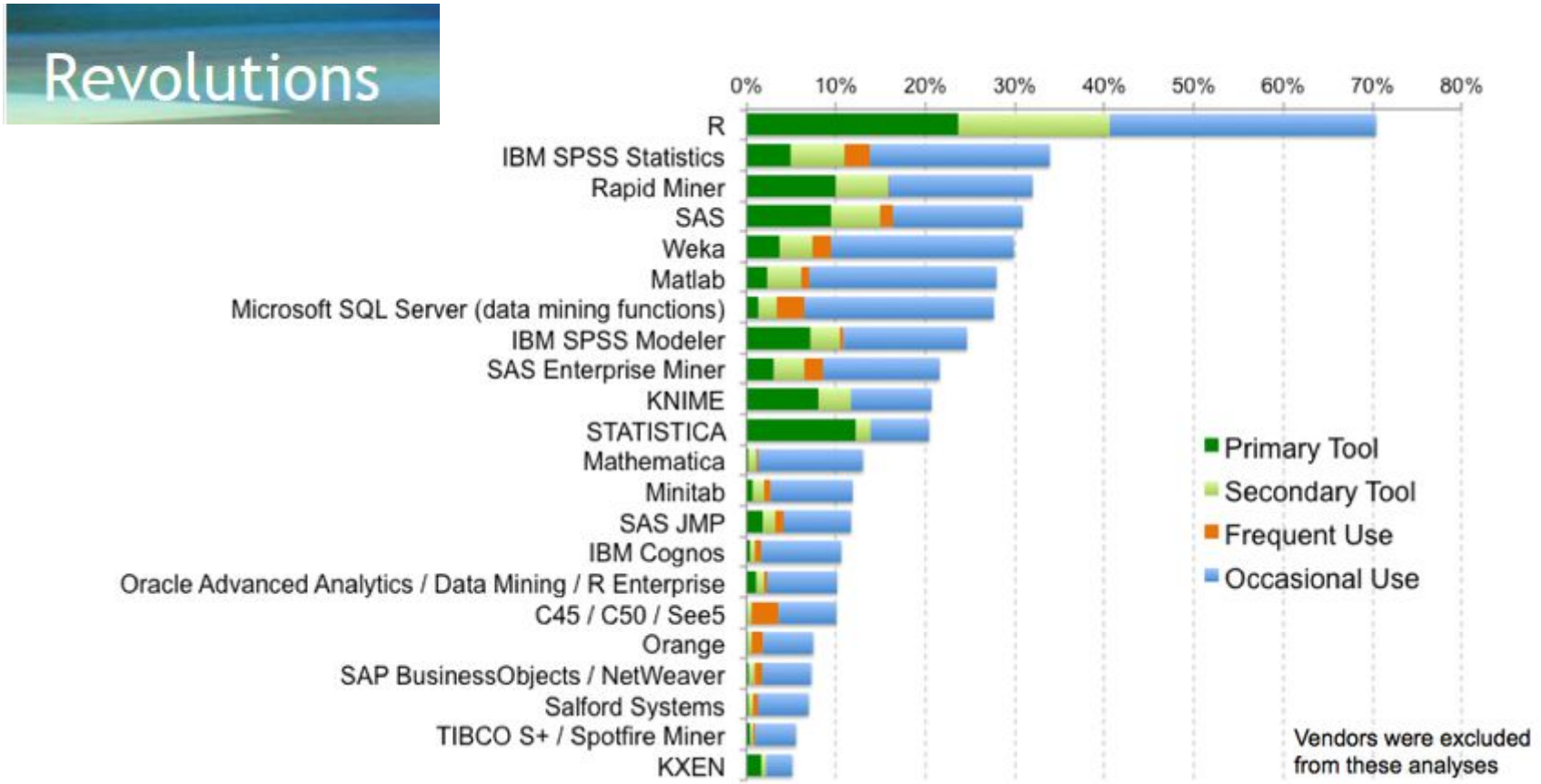
KEEL

- **University of Granada**
- **Machine learning software in Java implementation**

http://www.keel.es/

Herramientas, Lenguajes, Kaggle

Sobre herramientas de minería de datos

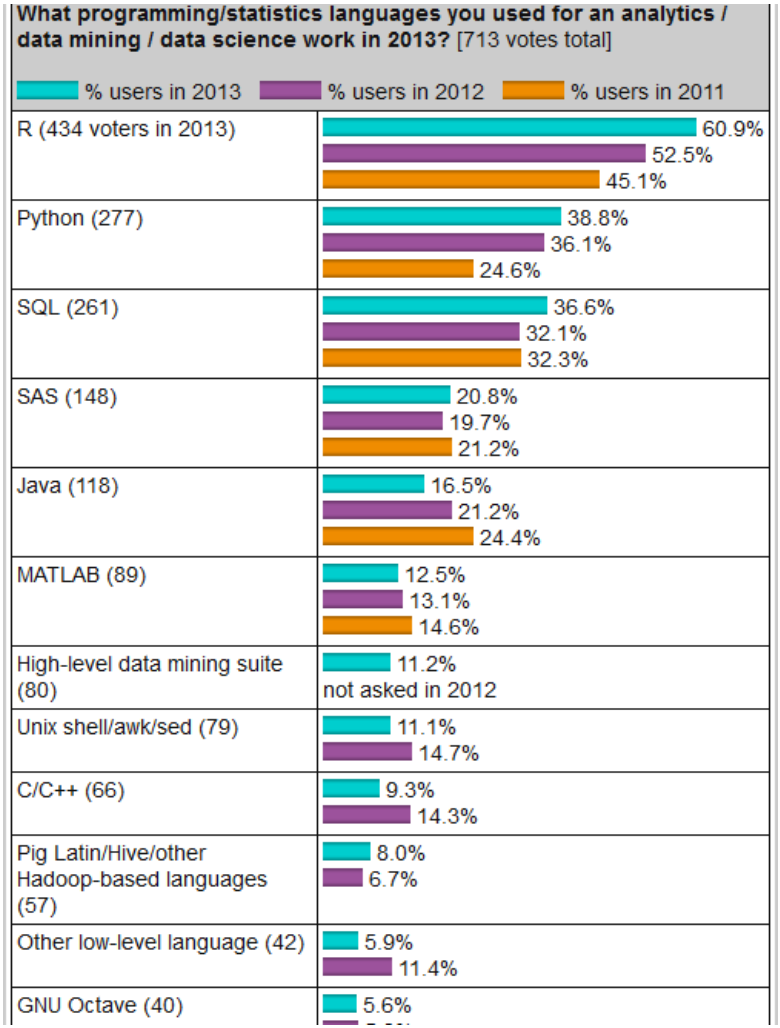


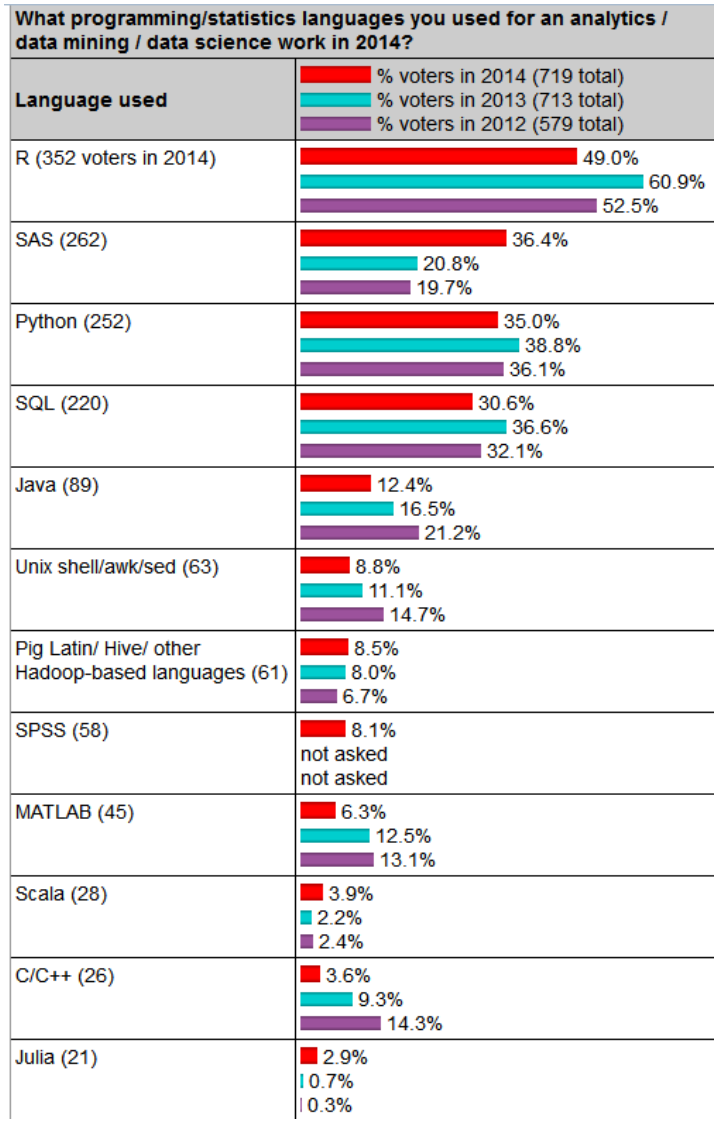
http://blog.revolutionanalytics.com/2013/10/r-usage-skyrocketing-rexer-poll.html

Herramientas, Lenguajes, Kaggle

Sobre los lenguajes de programación (R, Phyton, ...).

Lenguajes a usar para Data Science





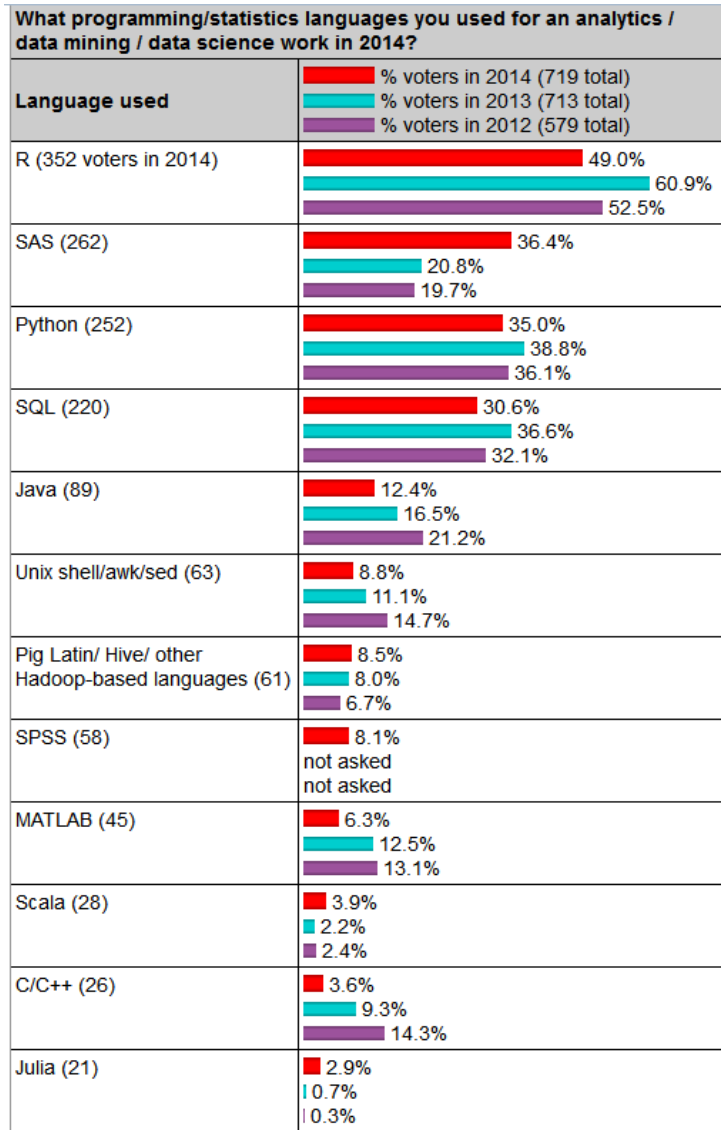
Herramientas, Lenguajes, Kaggle

Sobre los lenguajes de programación (R, Phyton, ...).

Consolidation among top 4 languages: R, SAS, Python, and SQL, and decline in usage of less popular languages for data mining: Java, Unix shell, MATLAB, C/C++, Perl, Octave, Ruby, Lisp, F.

By Gregory Piatetsky, Aug 18, 2014.

http://www.kdnuggets.com/polls/2014/languages-analytics-data-mining-data-science.html



What programming/statistics languages you used for an analytics / data mining / data science work in 2014?

Language used	% voters in 2014 (719 total)	% voters in 2013 (713 total)	% voters in 2012 (579 total)
R (352 voters in 2014)	49.0%	60.9%	52.5%
SAS (262)	36.4%	20.8%	19.7%
Python (252)	35.0%	38.8%	36.1%
SQL (220)	30.6%	36.6%	32.1%
Java (89)	12.4%	16.5%	21.2%
Unix shell/awk/sed (63)	8.8%	11.1%	14.7%
Pig Latin/ Hive/ other Hadoop-based languages (61)	8.5%	8.0%	6.7%
SPSS (58)	8.1%	not asked	not asked
MATLAB (45)	6.3%	12.5%	13.1%
Scala (28)	3.9%	2.2%	2.4%
C/C++ (26)	3.6%	9.3%	14.3%
Julia (21)	2.9%	0.7%	0.3%

Herramientas, Lenguajes, Kaggle

Sobre los lenguajes de programación (R, Phyton, …). El website CRAN

cran.r-project.org/

The Comprehensive *R* Archive Network

CRAN
Mirrors
What's new?

Contributed Packages

Available Packages

Currently, the CRAN package repository features 5799 available packages.

Table of available packages, sorted by date of publication

Table of available packages, sorted by name

http://cran.r-project.org/web/views/MachineLearning.html

Herramientas, Lenguajes, Kaggle

Sobre herramientas de minería de datos

http://scikit-learn.org/stable/

Classification

Identifying to which set of categories a new observation belong to.

Applications: Spam detection, Image recognition.
Algorithms: *SVM, nearest neighbors, random forest, ...* — *Examples*

Regression

Predicting a continuous value for a new example.

Applications: Drug response, Stock prices.
Algorithms: *SVR, ridge regression, Lasso, ...* — *Examples*

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes
Algorithms: *k-Means, spectral clustering, mean-shift, ...* — *Examples*

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency
Algorithms: *PCA, feature selection, non-negative matrix factorization.* — *Examples*

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter tuning
Modules: *grid search, cross validation, metrics.* — *Examples*

Preprocessing

Feature extraction and normalization.

Application: Transforming input data such as text for use with machine learning algorithms.
Modules: *preprocessing, feature extraction.* — *Examples*

Herramientas, Lenguajes, Kaggle

Sobre herramientas de minería de datos

Anaconda

http://docs.continuum.io/anaconda/

Anaconda is a free collection of powerful packages for Python that enables large-scale data management, analysis, and visualization for Business Intelligence, Scientific Analysis, Engineering, Machine Learning, and more.

****

Large learning problems

theano **Phyton library** **https://pypi.python.org/pypi/Theano**

Deep Learning

Pylearn is a Python library for machine learning, built on top of Theano, our library for defining, optimizing and evaluating mathematical expressions involving multi-dimensional arrays.

Herramientas, Lenguajes, Kaggle

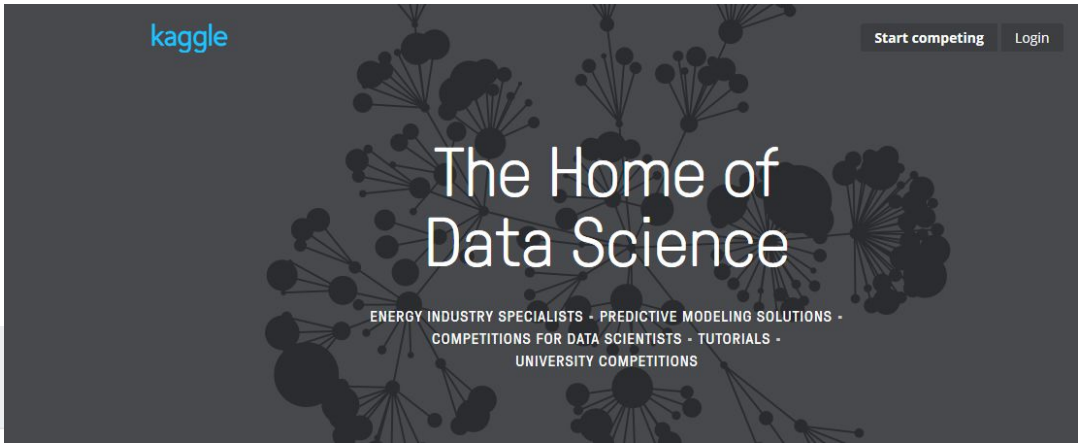
… y un buen enlace para comenzar a practicar, KAGGLE

Kaggle: The Home of Data Science

http://www.kaggle.com/

Es un portal web que ofrece competiciones, tutoriales, actividades académicas …

kaggle in Class

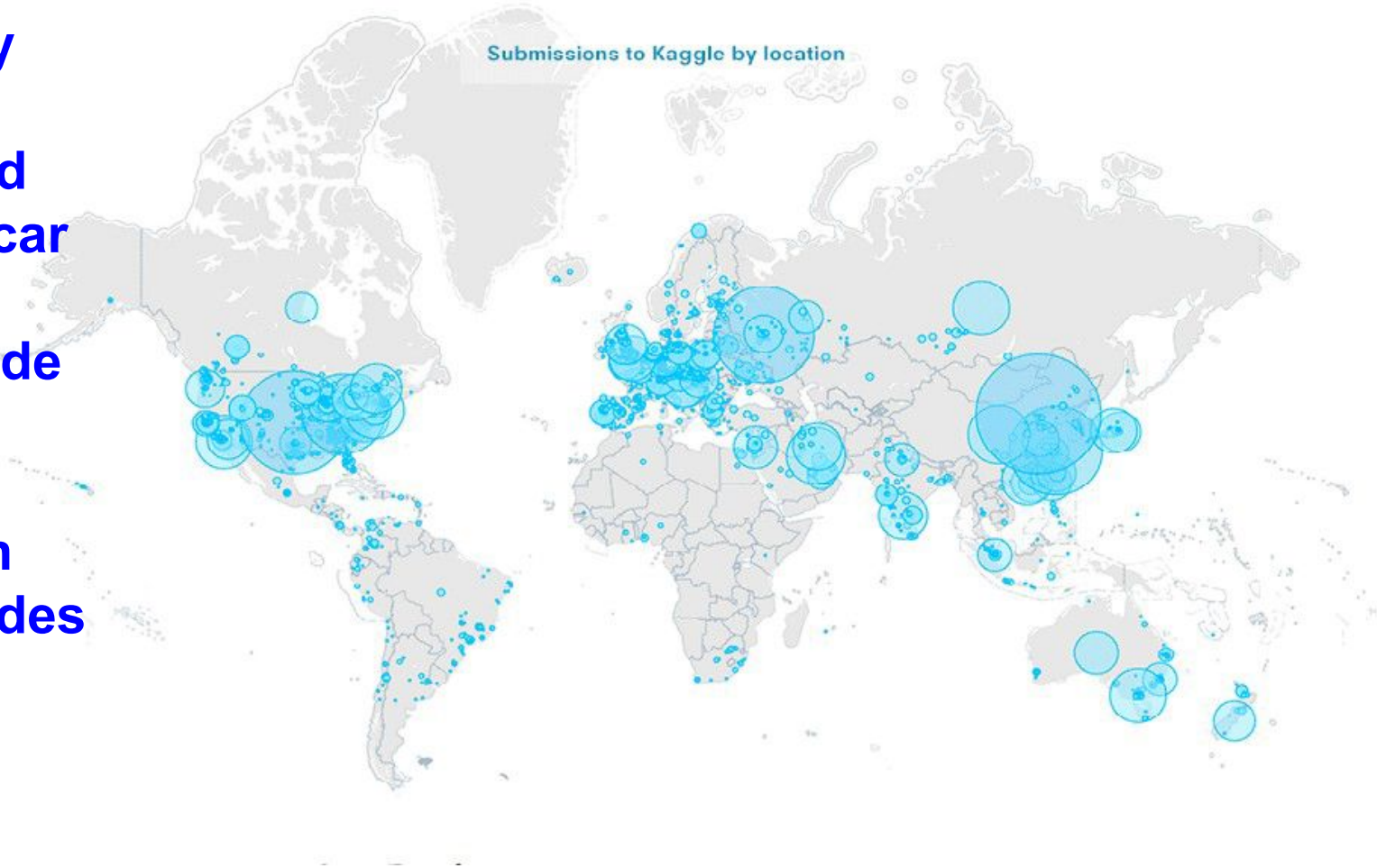


Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGLE

Kaggle: The Home of Data Science

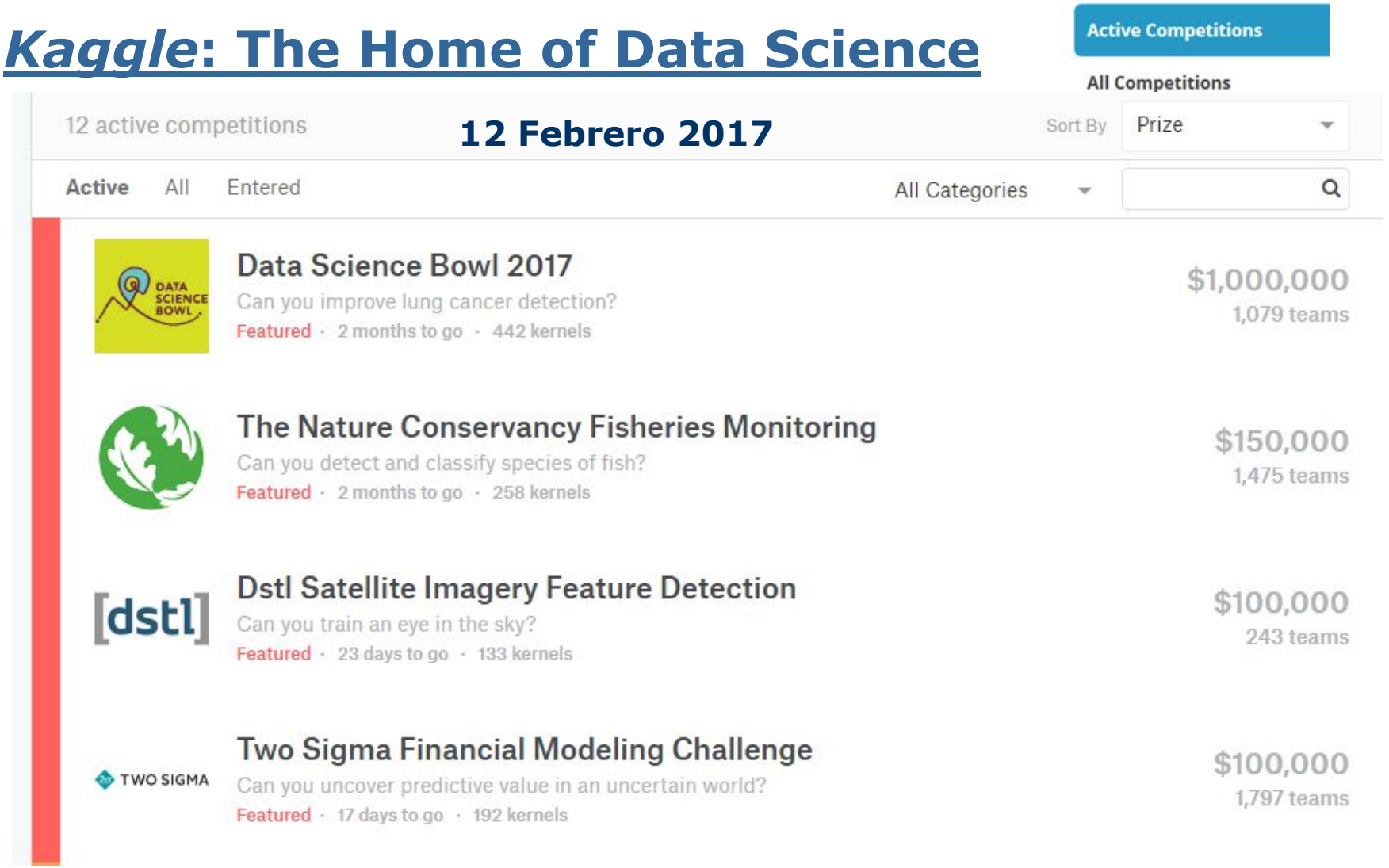
Es una muy buena oportunidad para practicar en la resolución de problemas reales y la adquisición de habilidades en Data Science.



Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGLE

Kaggle: The Home of Data Science

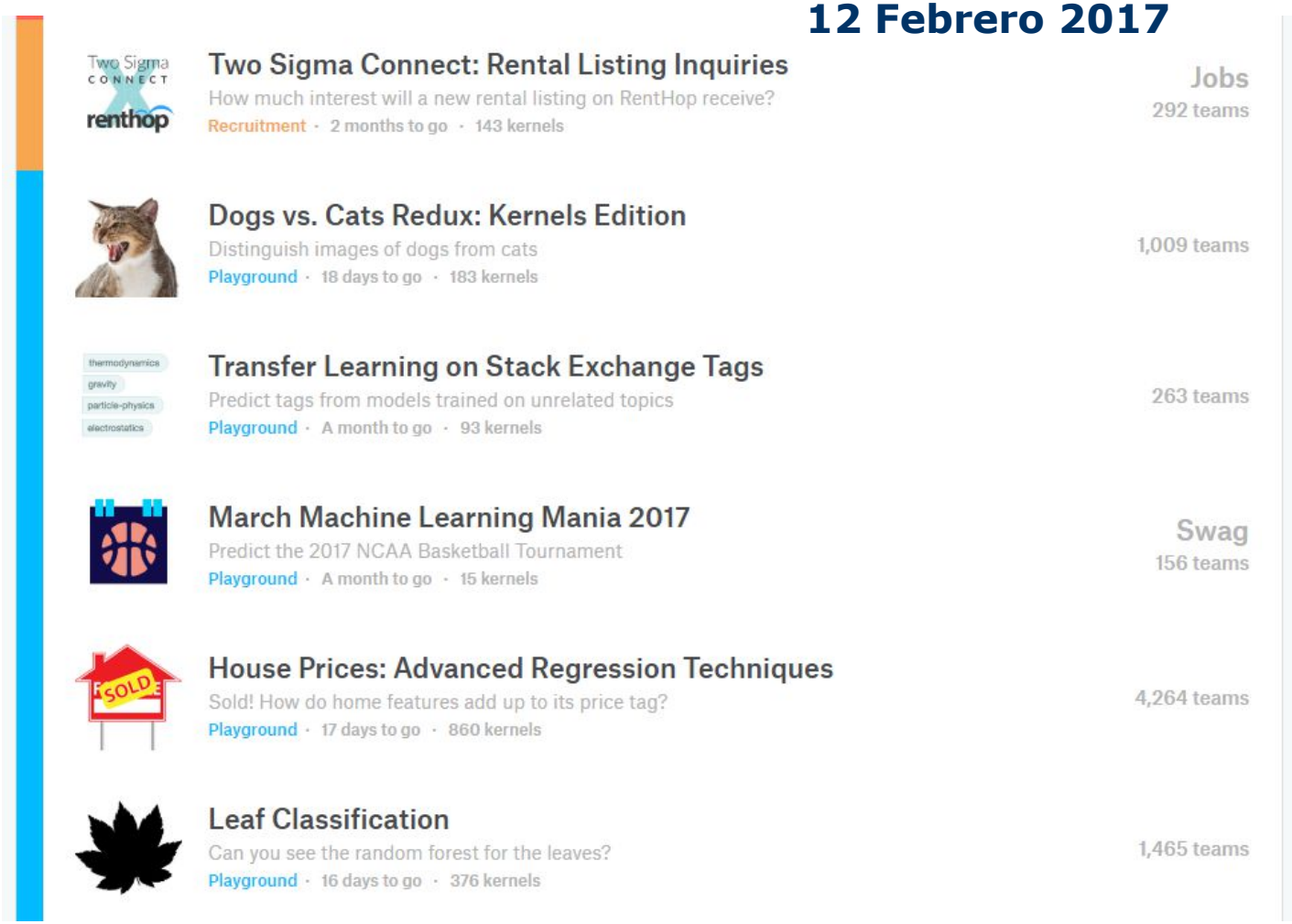


Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGEL

Kaggle: Go from Big Data to Big Analytics

12 Febrero 2017



Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGEL

Kaggle: Go from Big Data to Big Analytics

12 Febrero 2017

Digit Recognizer

Classify handwritten digits using the famous MNIST data

Getting Started · 3 years to go · 2,318 kernels

1,425 teams

Titanic: Machine Learning from Disaster

Predict survival on the Titanic using Excel, Python, R & Random Forests

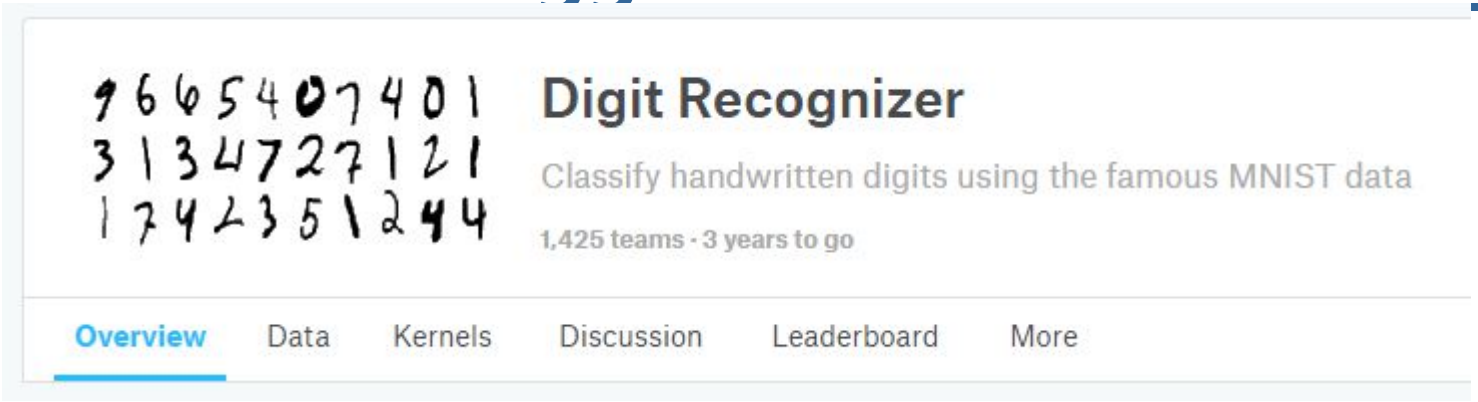
Getting Started · 3 years to go · 5,814 kernels

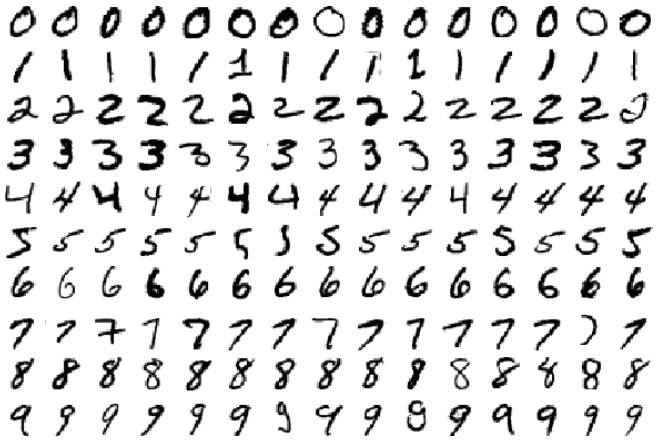
5,870 teams

Herramientas, Lenguajes, Kaggle

... y un buen enlace para comenzar a practicar, KAGGLE

Kaggle: The Home of Data Science





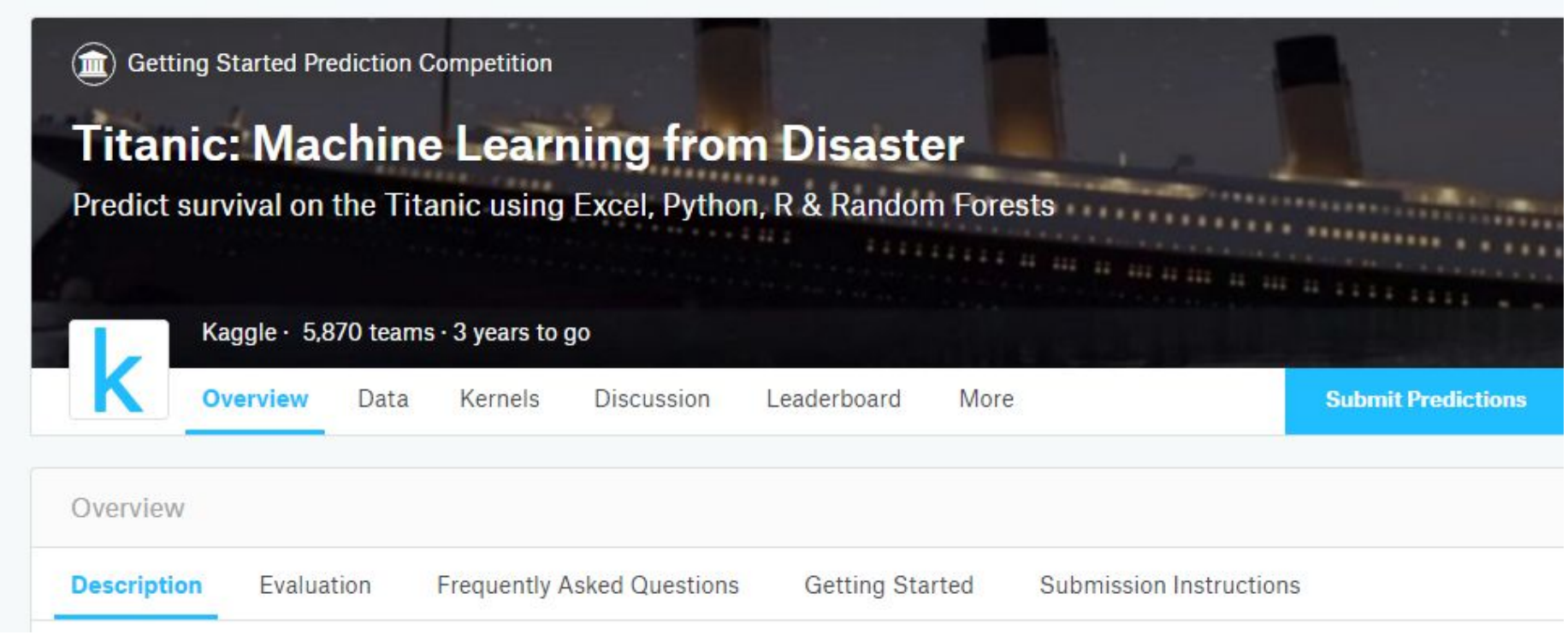
MNIST data

Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGEL

Kaggle: Go from Big Data to Big Analytics

https://www.kaggle.com/c/titanic



Herramientas, Lenguajes, Kaggle

… y un buen enlace para comenzar a practicar, KAGGLE

kaggle in Class

Sign Up

Predicting unemployment in the Great Recession

Monday, September 23, 2013 — Knowledge • 84 teams — Friday, December 6, 2013

Dashboard

Home
Data
Make a submission

Information
Description
Evaluation
Rules
Prizes

Forum

Leaderboard
Public
Private

Leaderboard

1. GobbleGobble
2. Bryan
3. Jim Monteleone
4. DMFK
5. mudkips
6. Steven Balough

Competition Details » Get the Data » Make a submission

This competition is private-entry. You can view but not participate.

Stats 202 Prediction Challenge.

Who can participate?

This challenge is restricted to students enrolled in Stats 202 at Stanford University in the fall of 2013.

The data

We will use the National Longitudinal Study of Youth (NLSY79). This landmark study published by the National Bureau of Labor Statistics has followed a cohort of baby boomers since 1979, recording various data on employment, education, poverty, family, attitudes, etc.

What is your goal?

Based on interviews spanning from 1979 to 2006, you will be tasked with predicting the number of weeks that each person was unemployed during 2010, at the peak of the Great Recession.

kaggle in Class — Sign up — Login

Academic Machine Learning Competitions

Theory, meet practice.

Kaggle hosts free projects for hundreds of universities around the globe. Engage students with an oppurtunity to apply machine learning to real problems.

Learn about hosting

Berkeley · University of California · UCL · Columbia · Cornell · Erasmus University · Harvard · The University of Melbourne · Michigan · University of Oxford · Stanford University · University of Toronto

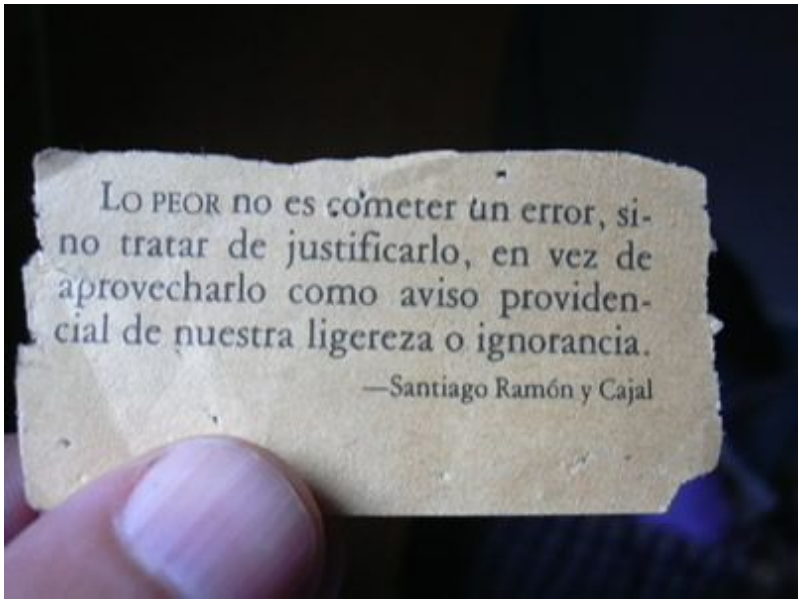
Índice

- ¿Qué es la Ciencia de Datos?
- El poder de los datos y su impacto en nuestra sociedad
- Herramientas y Lenguajes en Ciencia de Datos. Repositorio de Kaggle
- **Comentarios Finales**

Comentarios Finales

Hay que evitar los errores comunes

- Aprender de cosas que no son ciertas
  - Patrones que no representan ninguna regla subyacente
  - Datos que no reflejan lo relevante
  - Datos con un nivel de detalle erróneo
- Aprender cosas ciertas, pero inútiles
  - Aprender información ya conocida
  - Aprender cosas que no se pueden utilizar



Hay que obtener conocimiento útil

Comentarios Finales

Una demanda creciente de profesionales en "Big Data" y "Ciencia de Datos"

Oportunidades

La demanda de profesionales formados en Ciencia de Datos y *Big Data* es enorme.

Se estima que la conversión de datos en información útil generará un mercado de 132.000 millones de dólares en 2015 y que se crearán más de 4.4 millones de empleos.

España necesitará para 2015 más de 60.000 profesionales con formación en Ciencia de Datos y *Big Data*.


EL PAÍS

PORTADA INTERNACIONAL POL

ECONOMÍA

ECONOMÍA EMPRESAS MERCADOS BOLSA FINANZAS PERSONALES VIVIENDA TECNOLOG

ESTÁ PASANDO Multa a la banca Revuelo en Hacienda Eléctricas y renovables Paro

El maná de los datos

- La conversión de datos en información útil para las empresas generará un mercado de 132.000 millones de dólares en 2015. La herramienta 'big data' sacará del mercado a quien no la use

SUSANA BLÁZQUEZ | Madrid | 29 SEP 2013 - 01:00 CET 10

Archivado en: Citigroup Cap Gemini Sogeti SAP Oracle ING Bank BBVA Mapfre Bases datos IBM Telefónica Aplicaciones informáticas Tecnología Empresas Programas informáticos Economía

http://economia.elpais.com/economia/2013/09/27/actualidad/1380283725_938376.html

Comentarios Finales

Una demanda creciente de profesionales en "Big Data" y "Ciencia de Datos"

Oportunidades (en España)

http://www.revistacloudcomputing.com/2013/10/espana-necesitara-60-000-profesionales-de-big-data-hasta-2015/?goback=.gde_4377072_member_5811011886832984067#!

España necesitará 60.000 profesionales de Big Data hasta 2015

22 octubre, 2013 Eventos 18

España necesitará 60.000 profesionales de Big Data hasta 2015

"España va a necesitar alrededor de sesenta mil profesionales del Big Data de aquí a 2015", así lo ha asegurado Francisco Javier Antón, Subdirector General de Tecnologías del Ministerio de Educación, Cultura y Deportes en una mesa redonda sobre beneficio y aplicación de Big Data en pymes, moderada por Daniel Tapias de Sigma Technologies, celebrada durante el 4º Congreso Nacional de CENTAC de Toledo.

"Existe una demanda mundial para formar a 4,4 millones de profesionales de la gestión Big Data desde ingenieros, gestores y científicos de datos", comenta Antón. Sin embargo, "las empresas todavía no ven en el Big Data un modelo de negocio", lamenta. "Solo se extrae un 1% de los datos disponibles en la red", añade. "Hace falta formación y concienciación.

Sistemas Inteligentes para la Gestión de la Empresa

2016 - 2017

- Tema 1. Introducción a la Ciencia de Datos
- Tema 2. Depuración y Calidad de Datos. Preprocesamiento de datos
- Tema 3. Análisis Predictivo para la Empresa
- Tema 5. Análisis de Transacciones y Mercados
- Tema 4. Modelos avanzados de Analítica de Empresa
- Tema 6. Big Data
- Tema 7. Aplicaciones de la Ciencia de Datos en la Empresa