

Desarrollo de carrera en la Ciencia de datos

Ángela Devia

Líder de desarrollo e innovación en Científicas de Datos

MUJERES CON SUPERPODERES 4.0

LA MUJER EN LA ERA DIGITAL

01

BUSINESS UNDERSTANDING

Ask relevant questions and define objectives for the problem that needs to be tackled.

07

DATA VISUALIZATION

Communicate the findings with key stakeholders using plots and interactive visualizations.

02

DATA MINING

Gather and scrape the data necessary for the project.

DATA SCIENCE

sudeen.c

03

DATA CLEANING

Fix the inconsistencies within the data and handle the missing values.

06

PREDICTIVE MODELING

Train machine learning models, evaluate their performance, and use them to make predictions.

05

FEATURE ENGINEERING

Select important features an construct more meaningful ones using the raw data that you have.

04

DATA EXPLORATION

Form hypotheses about your defined problem by visually analyzing the data.

Data

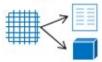


Traditional data warehouse, transactions, descriptive



Distributed datamarts, spreadsheets

Organize



Model





DECISIONS

"consumer oriented agile insight"





Big Data



variety

velocity

volume

Unstructui notes, logs



Social Med...
pulse, emerging issues



Survey Research attitudes, opinions



Sense-making



Sensors

on-premise - cloud-managed services

Data Science





Simplify

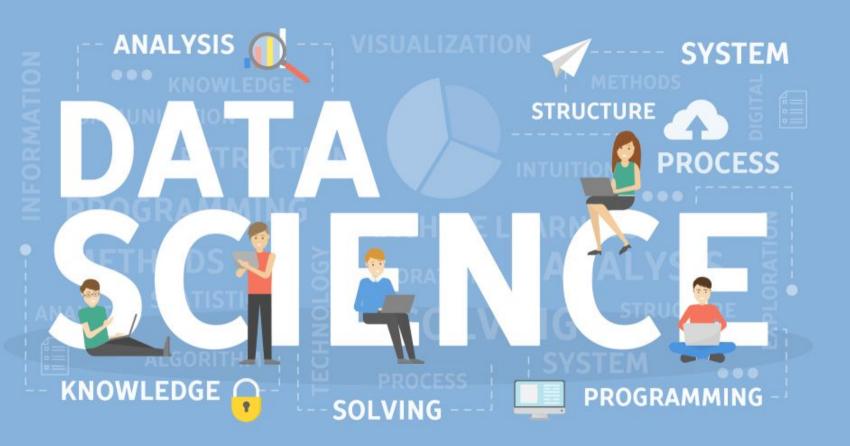














¿Quién puede ser científic@ de datos?

Carreras frecuentes:

Estadística - Ingeniería Industrial - Matemáticas - Ciencias de la computación

Carreras que hemos conocido en Ciencias de Datos:

Periodismo - Administración de empresa - Derecho - Economía -

Ciencias Políticas y Sociales - Pedagogía - Ciencias básicas

Realmente, ¡no hay límite o restricción alguna!



Roles en las Ciencias de Datos e IA

- Ingeniera de Datos (Data engineer)
- Tomadora de decisiones (Decision maker)
- Analista de datos (Data analyst)
- Analista experto (Expert analyst)
- Estadística (Statistician)
- Ingeniera de aprendizaje de máquina aplicado (Applied Machine Learning Engineer)
- Científica de datos (data scientist)
- Gestora de analítica (Analytics Manager / Data Science Leader)
- Experta en datos cualitativos (Qualitative Expert / Social Scientist)
- Investigadora (Researcher)

Ejemplo de aplicación de Ciencia de Datos

Ibagué e Islandia

Ibagué / Población

553,526

2015



También se buscó



8.081 millones



Las fuentes incluyen: Eurostat, Organización de las Naciones Unidas

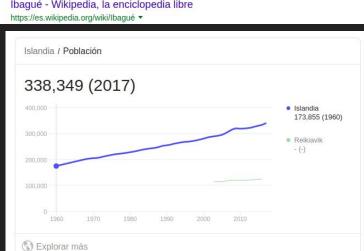
Bucaramanga 521,857



Comentarios

Comentarios

Ibagué - Wikipedia, la enciclopedia libre





Ibagué

Municipio en Colombia

Ibagué es una ciudad del oeste de Colombia conocida por su patrimonio musical. En la Plaza de Bolívar, bordeada de árboles, está la Catedral de Ibagué, que tiene un campanario con una cúpula. Cerca está el Conservatorio de Tolima, que presenta conciertos. El Museo de Arte de Tolima tiene una colección variada. En el noroeste, se encuentra el verde Cañón del Combeima, hogar de osos de anteojos, cóndores y zorros. Más al noroeste, están las cumbres nevadas del Parque Nacional Los Nevados.





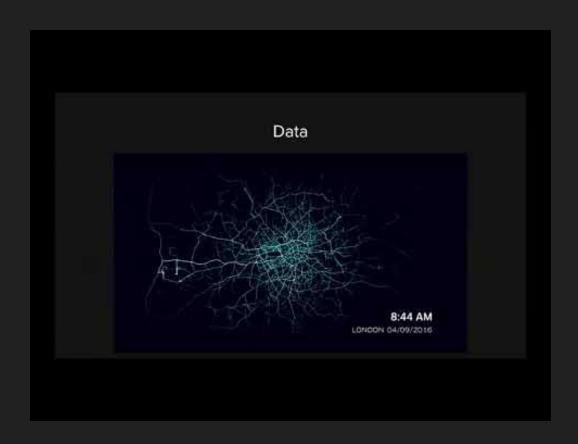


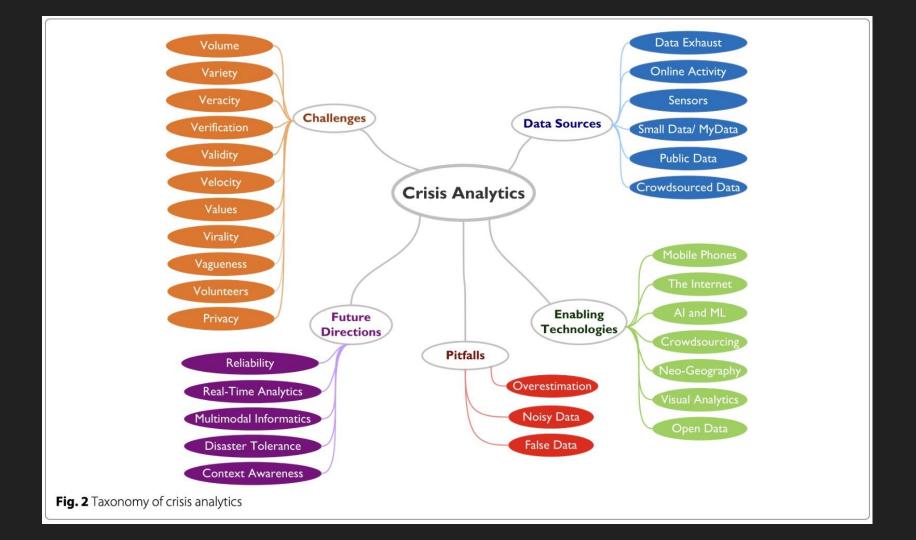
Ejemplos de aplicación de Ciencia de Datos

Inteligencia Artificial: Tráfico en China



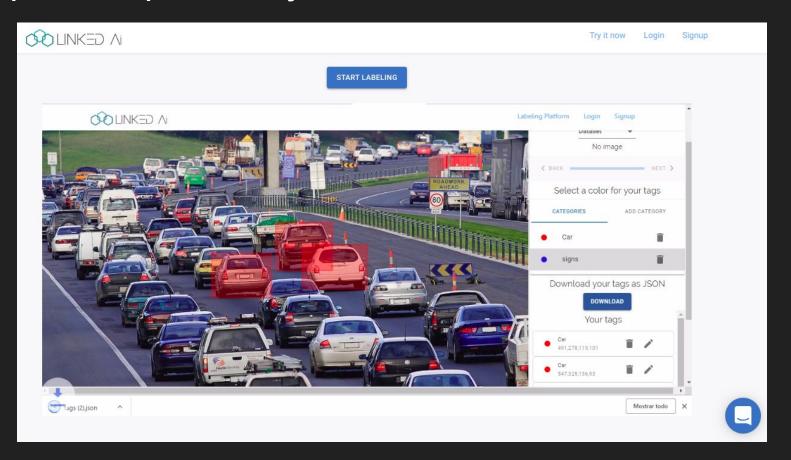
Uber Eats





Empresas que trabajan con el aprovechamiento de Datos

Ejemplos en Colombia y el mundo





ABOU

NEWS

EVENTS

KNOWLEDGE

OUR WORK

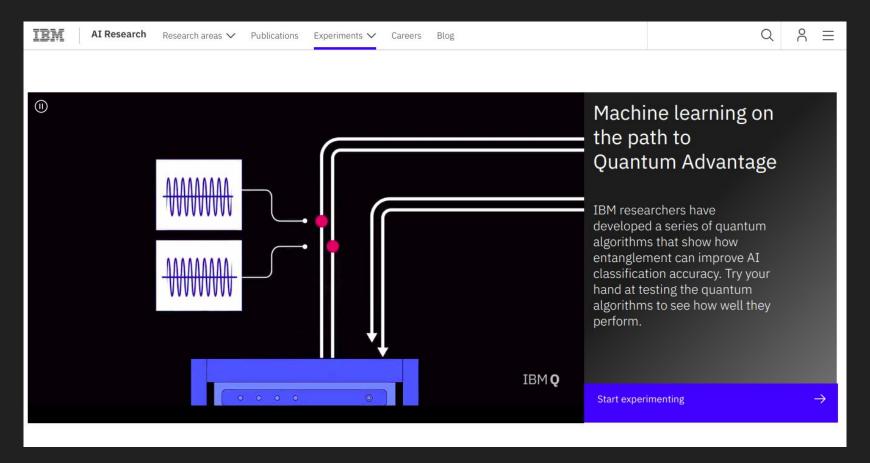
GET INVOLVED

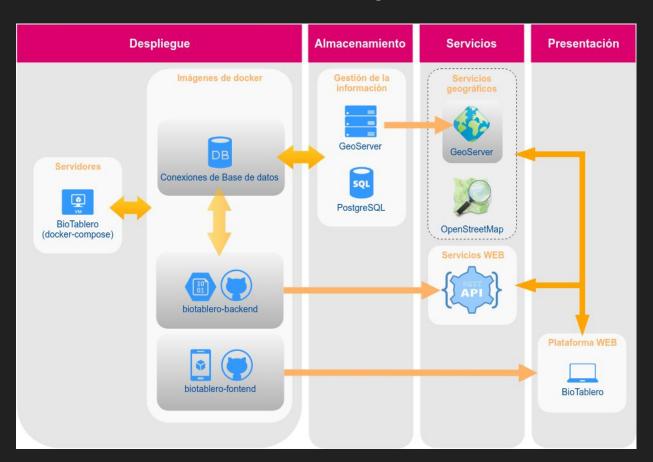
Q

KNOWLEDGE HUB

Curated knowledge, data, case studies and learning resources for those with an interest in Artificial Intelligence in New Zealand.

Welcome to the AI Forum NZ Knowledge Hub. This is a curated collection of key knowledge, data, case studies and learning resources for those with an interest in Artificial Intelligence and its benefits for New Zealand.

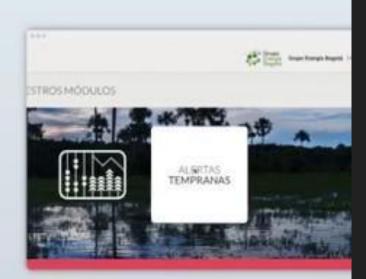




Proyecto en el cual trabajo actualmente:

Biotablero - Instituto Humboldt

http://biotablero.humboldt .org.co Herramientas web para consultar cifras e indicadores y facilitar la toma de decisiones sobre biodiversidad



Herramientas para trabajar los datos y generar valor

Código abierto

Con licencia (periodo de prueba)



MysQL

loins are what



Transform



SQL

If you change the way you look at things, the things you look at change.

Model



Visualize



Other Tools











Al usually natural stupidity







you see, salt looks like sugar

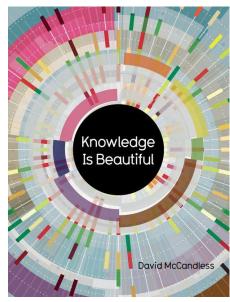


Dont trust everything I have enough tools, said no data scientist ever



iMás información!

Existen demasiados recursos en visualización de datos, pero uno de mis súper recomendados:



TEDTalk

¿Cómo participamos en la industria 4.0?

Aprendiendo, aplicando y probando nuevas tecnologías

Compartiendo nuestras habilidades y conocimientos

Explorando información



Hoy, ¿qué puedo hacer para comenzar a comprender sobre la ciencia de datos?

Recomendaciones:

- Buscar en Google
- Usar YouTube
 - Luis Serrano & ML
- Aprender en DataCamp
- Explorar Coursera y
 EdX

- ...

Excellent Quotes by Warren Buffet



On Earning: "Never depend on single income. Make investment to create a second source".

On Spending: "If you buy things you do not need, soon you will have to sell things you need".

On Savings: "Do not save what is left after spending, but spend what is left after saving".

On Taking Risk: "Never test the depth of river with both the feet".

On Investment: "Do not put all eggs in one basket".

On Expectations: "Honesty is very expensive gift. Do not expect it from cheap people".

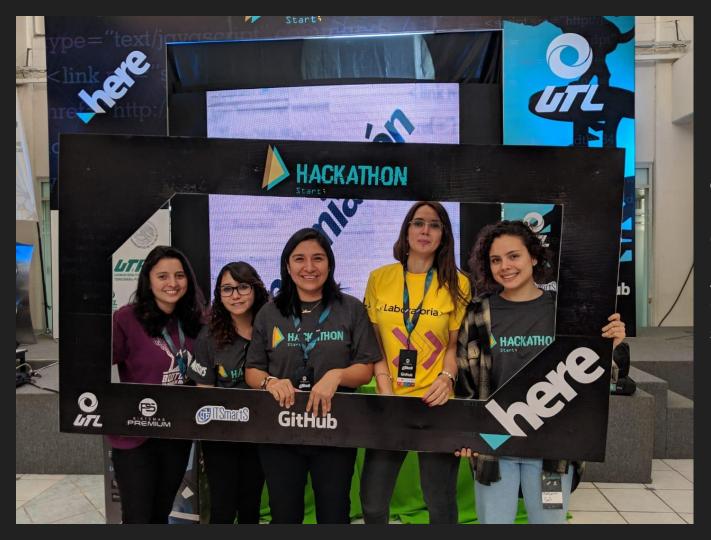
El idioma no tiene por qué ser barrera

La ciencia de datos ha sido usada para facilitarnos la vida y aprender de contenidos que están en otros idiomas es más fácil. Usemos herramientas como Traductores en línea y correctores de estilo para apoyar nuestro trabajo.



KEEP CALM HABLA **ESPANGLISH**

biculturalmom.com



Científicas de datos +

Fernanda de HERE Technologies @herehacklatam

Lucero - Laboratoria Perú Loreto - Laboratoria Chile (@laboratoriala)

Sol - ComIT de Argentina



Gracias por su atención y bienvenidas todas las personas a la Industria 4.0

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21th century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- Optimization: gradient descent and variants



PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative



Gracias por su atención y participación

Cualquier inquietud, contáctenos:

contacto@cientificasdedatos.com

