

18 de Septiembre de 2023 – Madrid

# THE BRIDGE

## Data Science Full-Time

*Promoción Septiembre 2023*

*Lead Instructor*

**Miguel Nievas**

*Teacher Assistant*

**Jared Rivas**

*Teacher Assistant*

**Santiago Valencia**

*Teacher Assistant*

**Yeray Expósito**

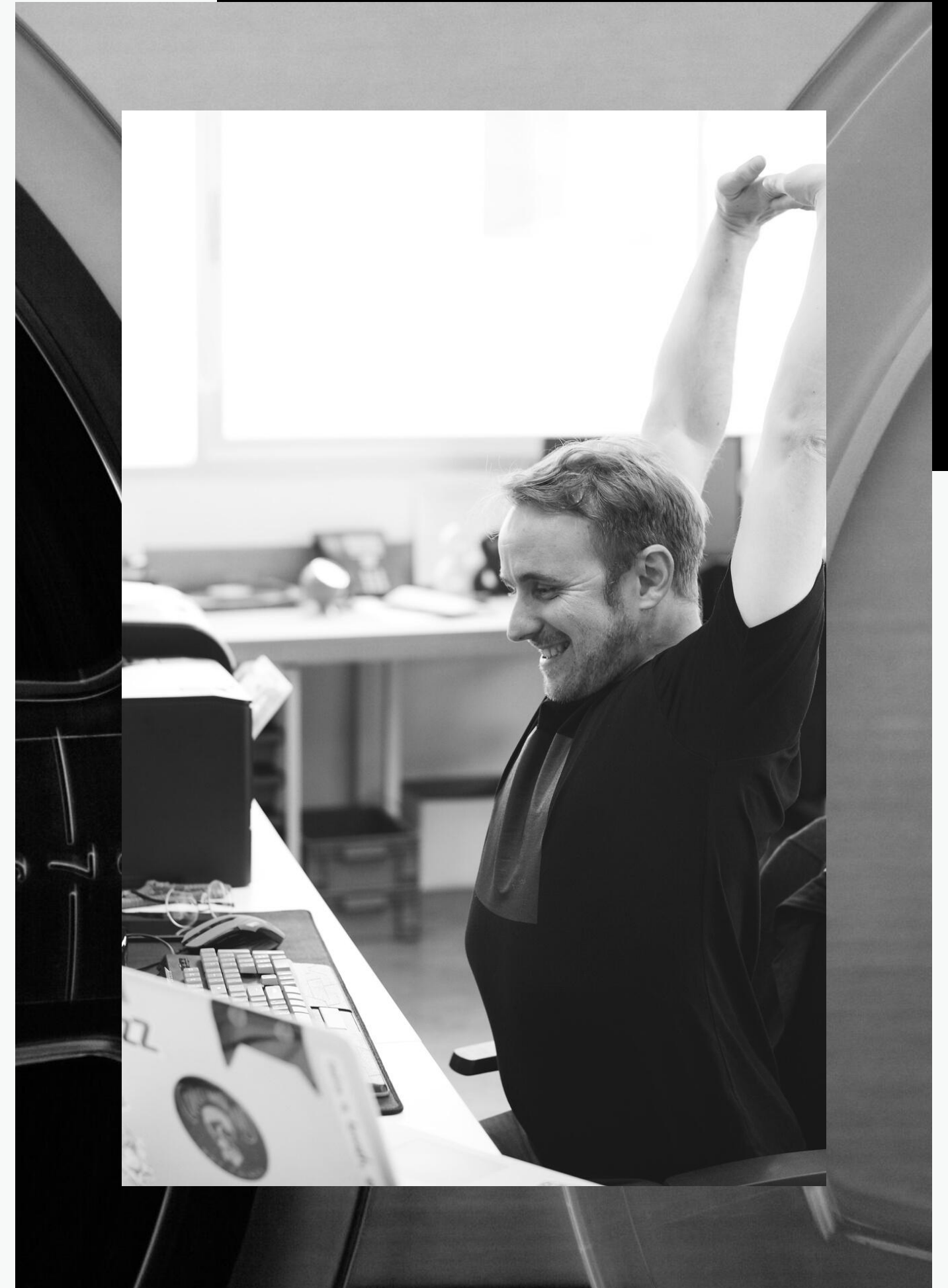
# Índice

- ❑ Hablemos de nosotros
- ❑ ¿Científico de datos?
- ❑ El Bootcamp

Calendario + Horario

Temario + Entregas + Herramientas

Guidelines



# Trayectoria



**Universidad de Zaragoza**  
*Ingeniero industrial*



**Karlsruher Institut für Technologie Universität**  
*Ingeniero IoT*



**Thyssenkrupp**  
*Analista de datos en logística*



**Decathlon**  
*Data Scientist en Making AI Accesible*



**Banco Santander**  
*Data Scientist en RRHH y Desarrollo de Negocio*



**The Bridge**  
*Data Science Lead Instructor*



## Hobbies



**Miguel Nievas**  
*Lead Instructor*

**Contacto**

[miguel.nievas@thebridgeschool.es](mailto:miguel.nievas@thebridgeschool.es)  
[LinkedIn](#)



# Os toca!

**Preséntate!**

**De dónde vienes**

**¿Porqué Data Science?**

**¿Qué esperas de este Bootcamp?**



# ¿Qué es un científico de datos?

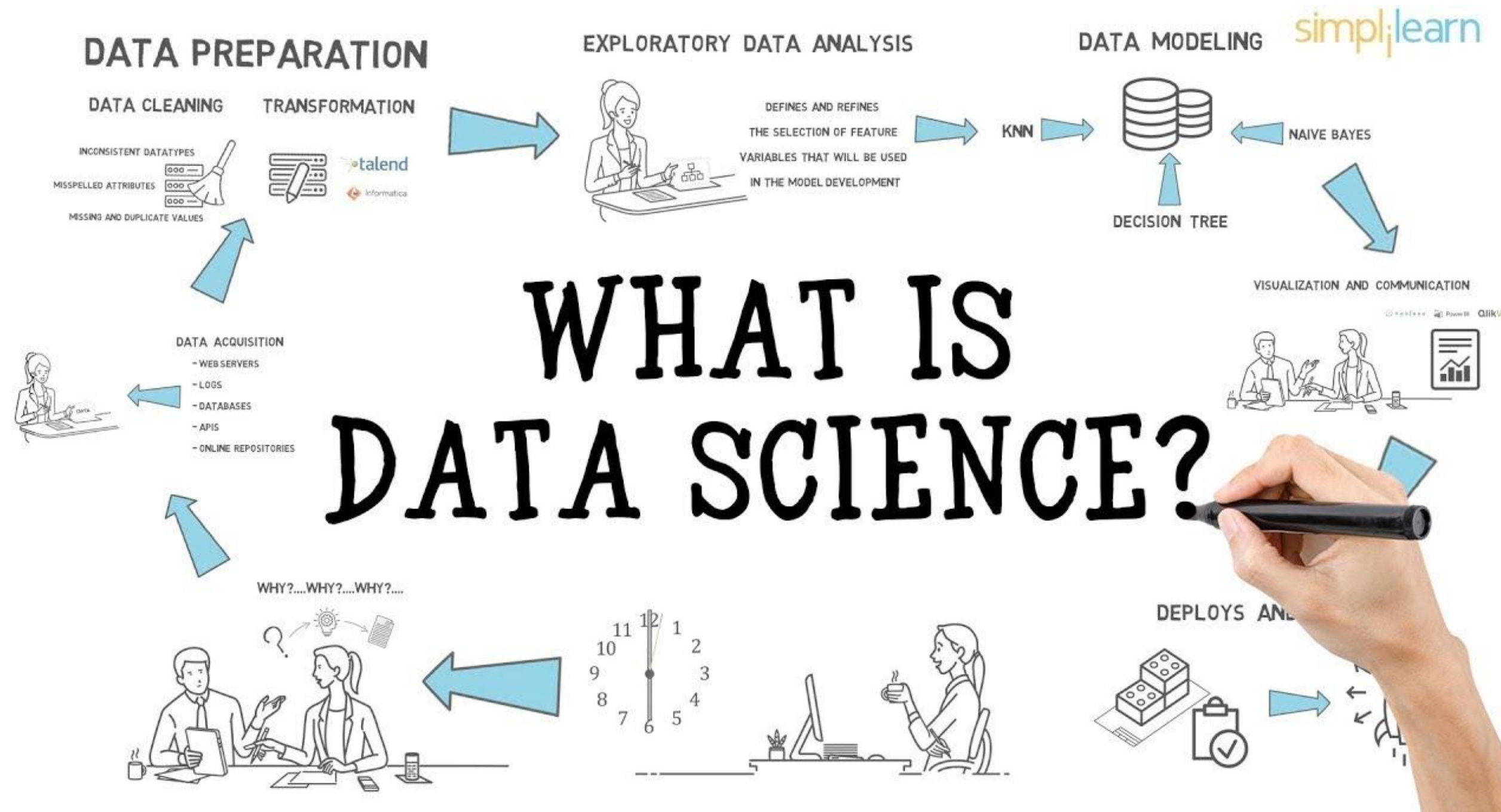


Josh Wills  
@josh\_wills

Follow

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

Reply Retweet Favorite More



# La profesión de Data Scientist

## ¿Qué trabajo realiza un Data Scientist?

Apoyo en la toma de decisiones basadas en datos

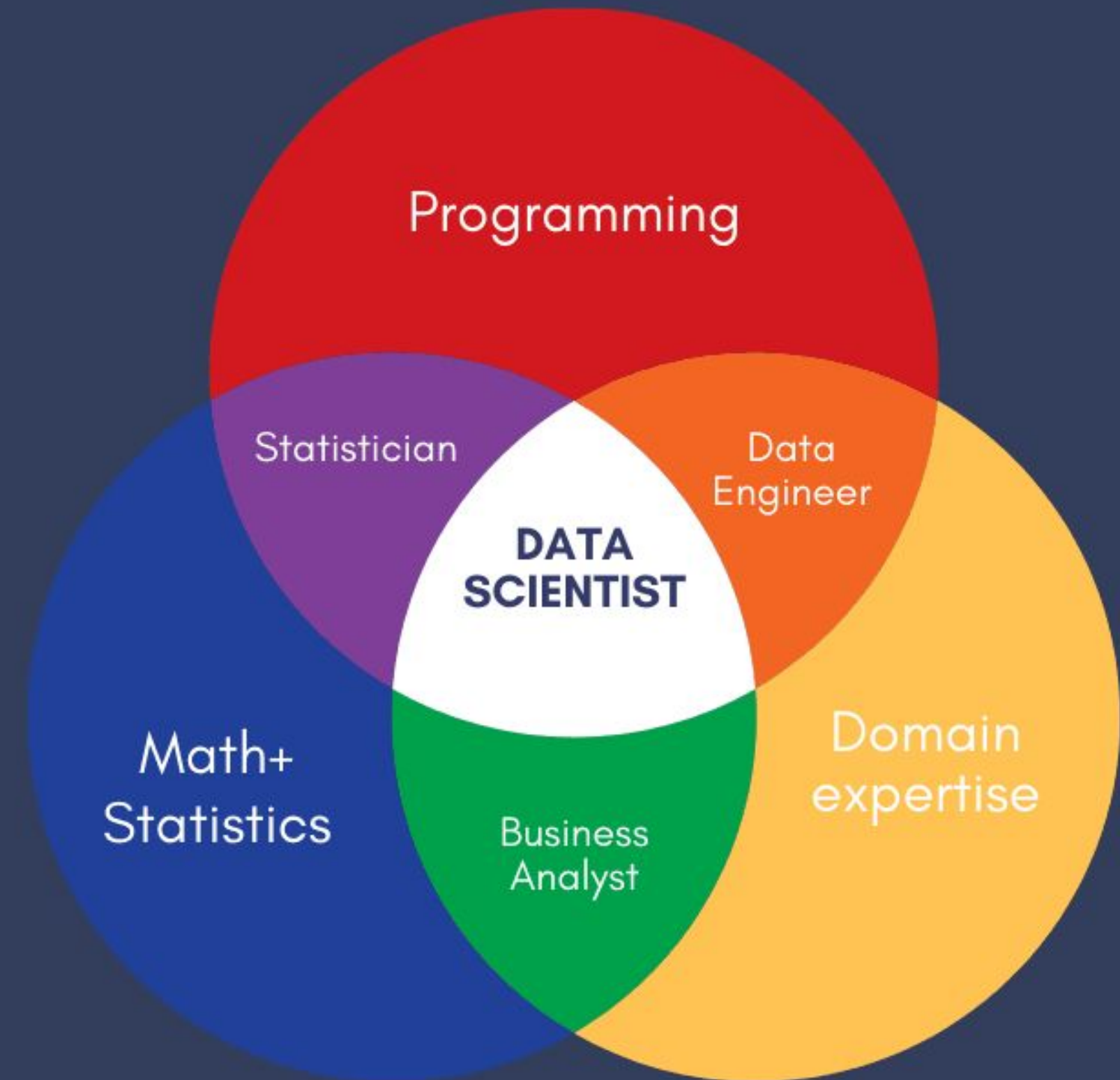
Resuelve problemas

Automatización de procesos

Algoritmos predictivos

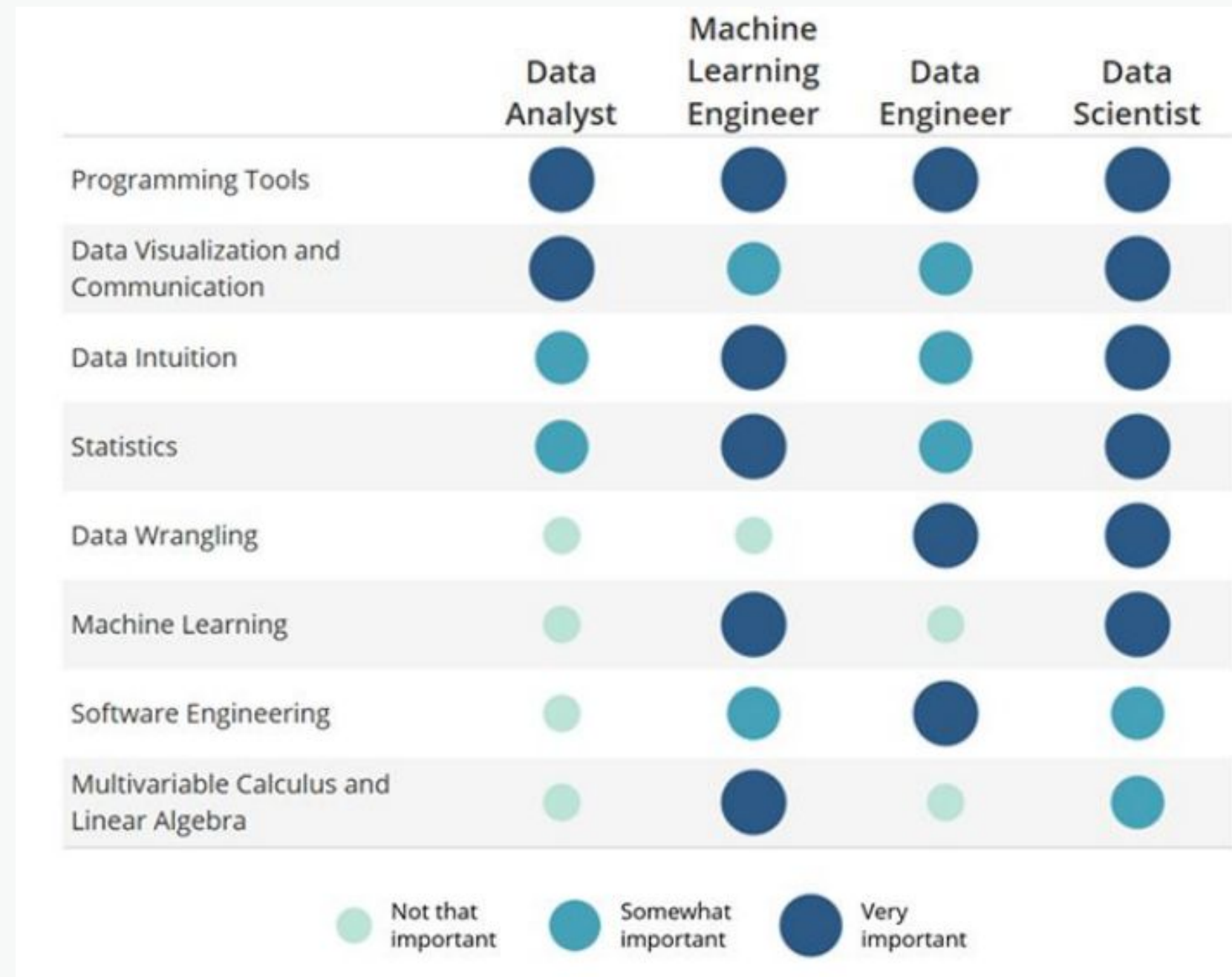
Visualización y storytelling

Servicio y gestión de bases de datos





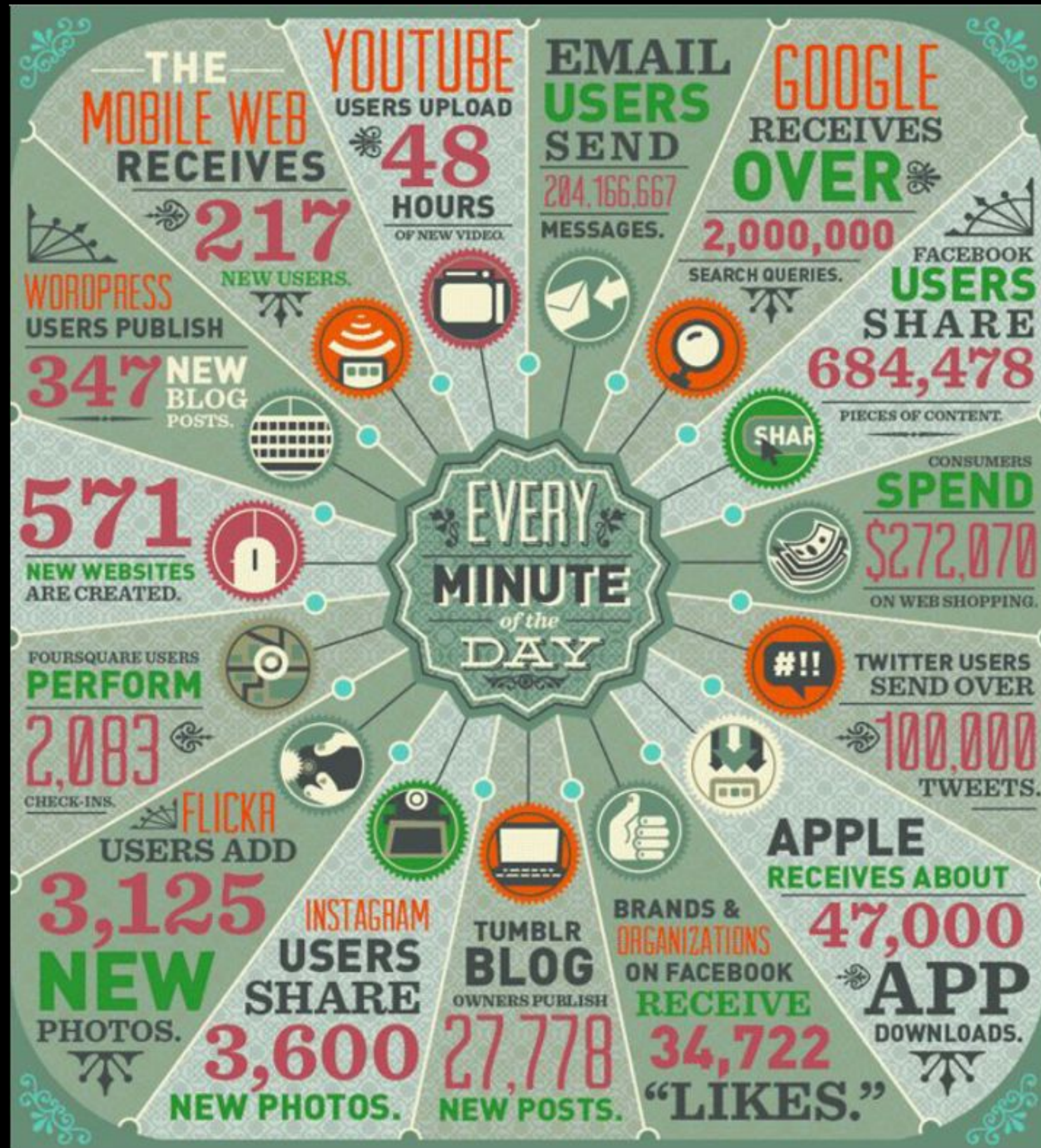
# ¿Qué significa serlo? II



[Fuente](#)

En qué dedicamos nuestro esfuerzo











# python

```
6 from watson.events import types
7 from watson.framework import events
8 from watson.http.messages import Response, Request
9 from watson.common.imports import get_qualified_name
10 from watson.common.contextmanagers import suppress
11
12
13 ACCEPTABLE_RETURN_TYPES = (str, int, float, bool)
14
15
16 class Base(ContainerAware, metaclass=abc.ABCMeta):
17     """The base class for all controllers.
18     Attributes:
19         __action__ (string): The last action that was called on the controller.
20     """
21
22     def execute(self, **kwargs):
23         method = self.get_execute_method(**kwargs)
24         self.__action__ = method
25         return method(**kwargs) or {}
26
27     @abc.abstractmethod
28     def get_execute_method(self, **kwargs):
29         raise NotImplementedError(
30             "You must implement get_execute_method")
```



# Hablemos de aplicaciones I

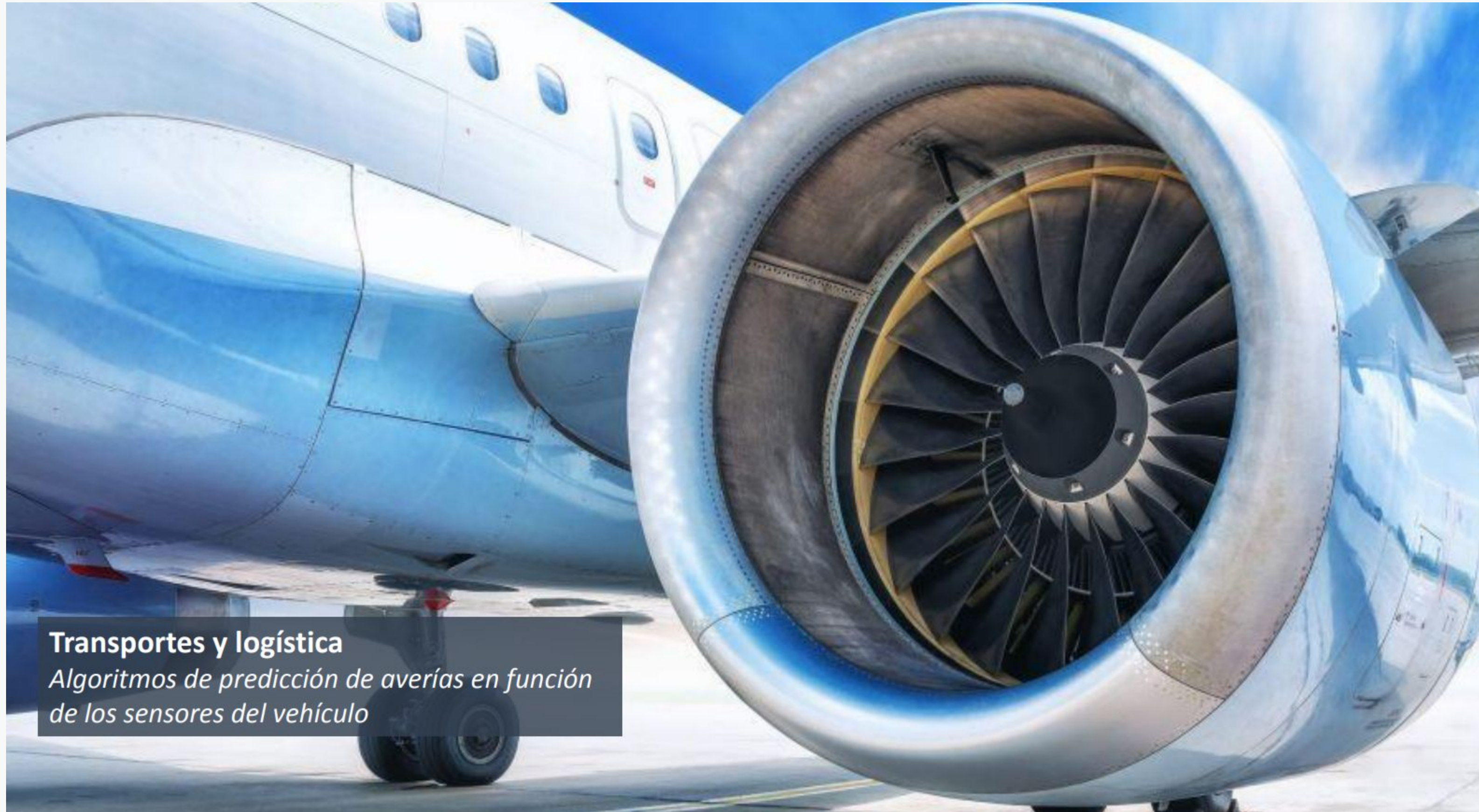


## Traders y banca

*Predicción de subidas y caídas del mercado, así como algoritmos de default*



# Hablemos de aplicaciones II

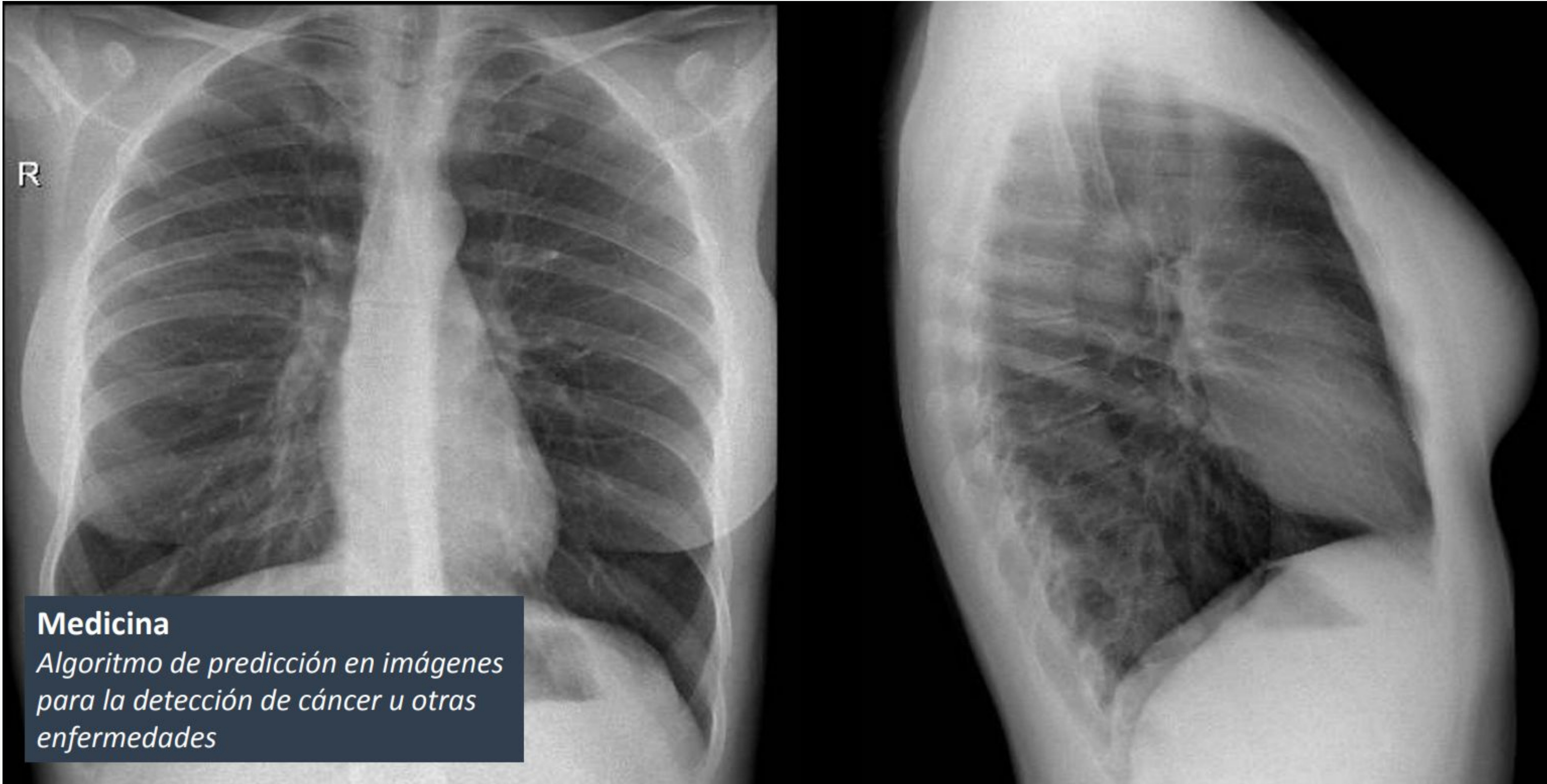


## Transportes y logística

*Algoritmos de predicción de averías en función de los sensores del vehículo*



# Hablemos de aplicaciones III



# Hablemos de aplicaciones IV





# Hablemos de aplicaciones V



Her (2013)



Ex-Machina (2014)

# Alerta Spoiler: Esto no es nuevo, pero... I

## Modelos Predictivos

### History [\[ edit \]](#)

Least squares linear regression, as a means of finding a good rough linear fit to a set of points was performed by [Legendre](#) (1805) and [Gauss](#) (1809) for the prediction of planetary movement. [Quetelet](#) was responsible for making the procedure well-known and for using it extensively in the social sciences.<sup>[26]</sup>

## Distribuciones de población

### History [\[ edit \]](#)

Bayes' theorem was named after [Thomas Bayes](#) (1701–1761), who studied how to compute a distribution for the probability parameter of a [binomial distribution](#) (in modern terminology). Bayes's unpublished manuscript was significantly edited by [Richard Price](#) before it was posthumously read at the Royal Society. Price edited<sup>[11]</sup> Bayes's major work "An Essay towards solving a Problem in the Doctrine of Chances" (1763), which appeared in *Philosophical Transactions*,<sup>[12]</sup> and

## Computación

### History [\[ edit \]](#)

See also: *Algorithm and Church–Turing thesis*

They were described in 1936 by [Alan Turing](#).

### Historical background: computational machinery [\[ edit \]](#)

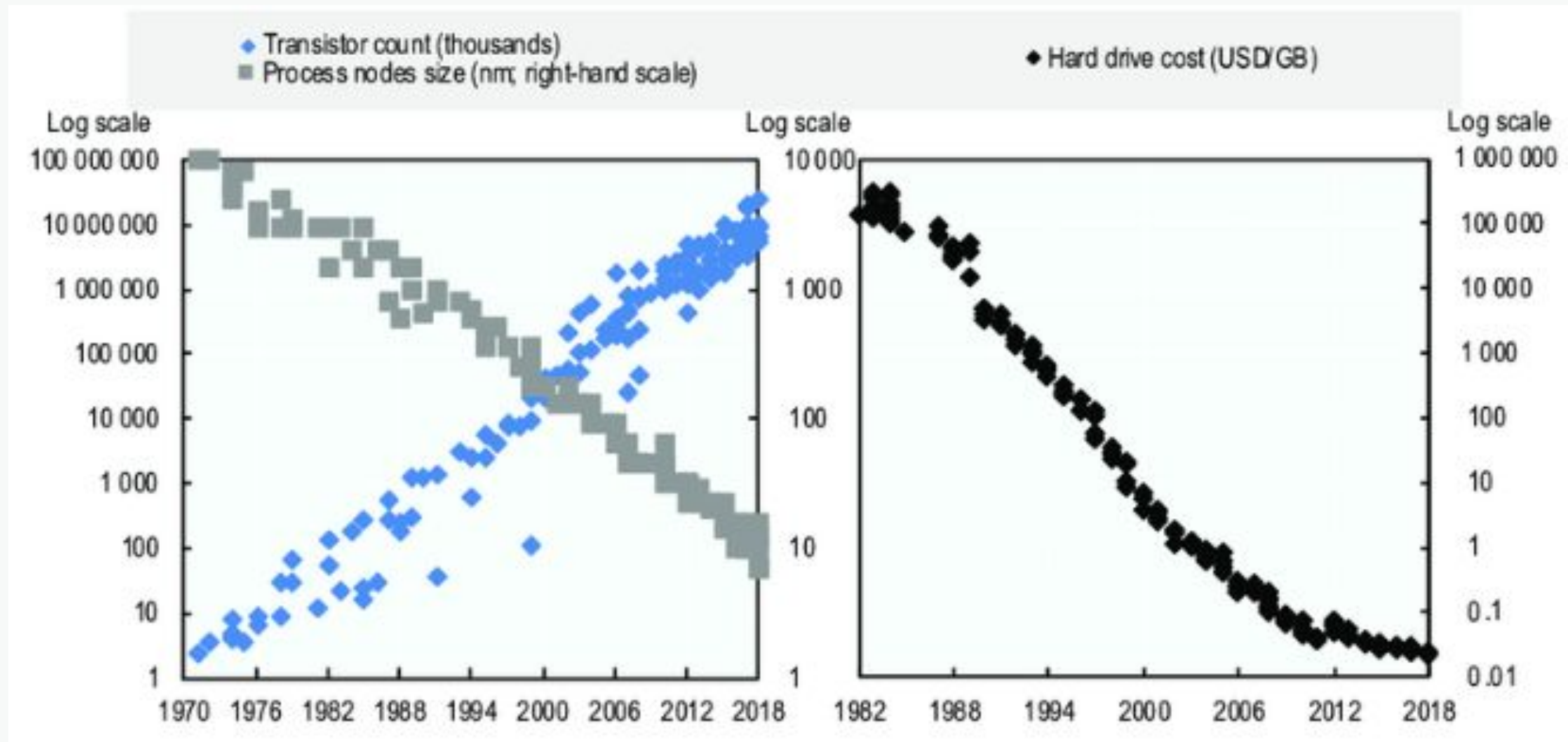
[Robin Gandy](#) (1919–1995)—a student of [Alan Turing](#) (1912–1954), and his lifelong friend—traces the lineage of the notion of "calculating machine" back to [Charles Babbage](#) (circa 1834) and actually proposes "Babbage's Thesis":





# Alerta Spoiler: Esto no es nuevo, pero... II

Procesado + Almacenamiento + Coste = Accesibilidad



[Fuente](#)

# Alerta Spoiler: Esto no es nuevo, pero... III

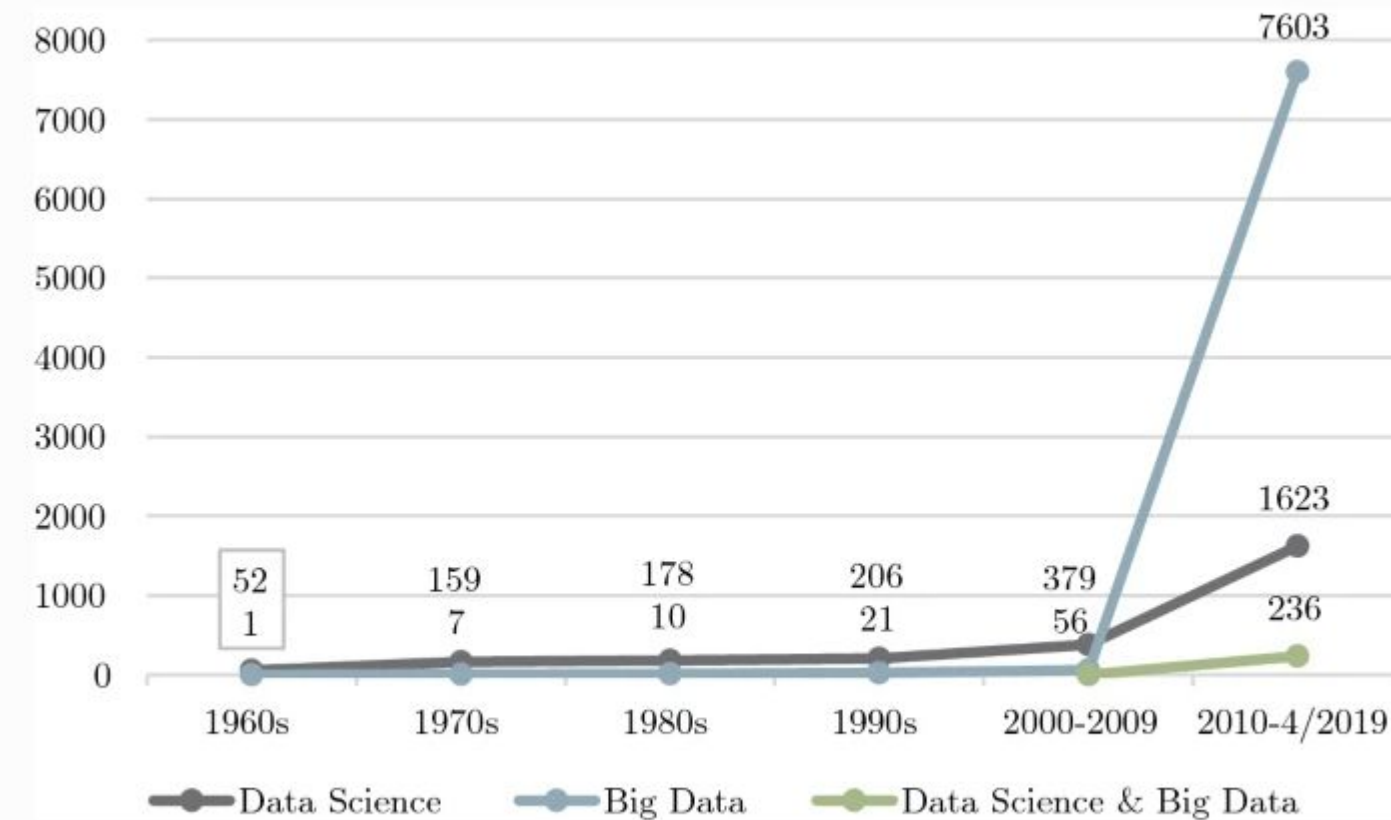
## Generación de datos

- **1.7MB of data** is created every second by every person during 2020.
- In the last two years alone, the astonishing **90%** of the world's data has been created.
- **2.5 quintillion bytes** of data are produced by humans every day.
- **463 exabytes** of data will be generated each day by humans as of 2025.
- **95 million** photos and videos are shared every day on Instagram.
- By the end of 2020, **44 zettabytes** will make up the entire digital universe.
- Every day, **306.4 billion emails** are sent, and **5 million Tweets** are made.



## Conocimiento e investigación

Fig. 1



Evolutionary trend in the number of publications covering data science and big data

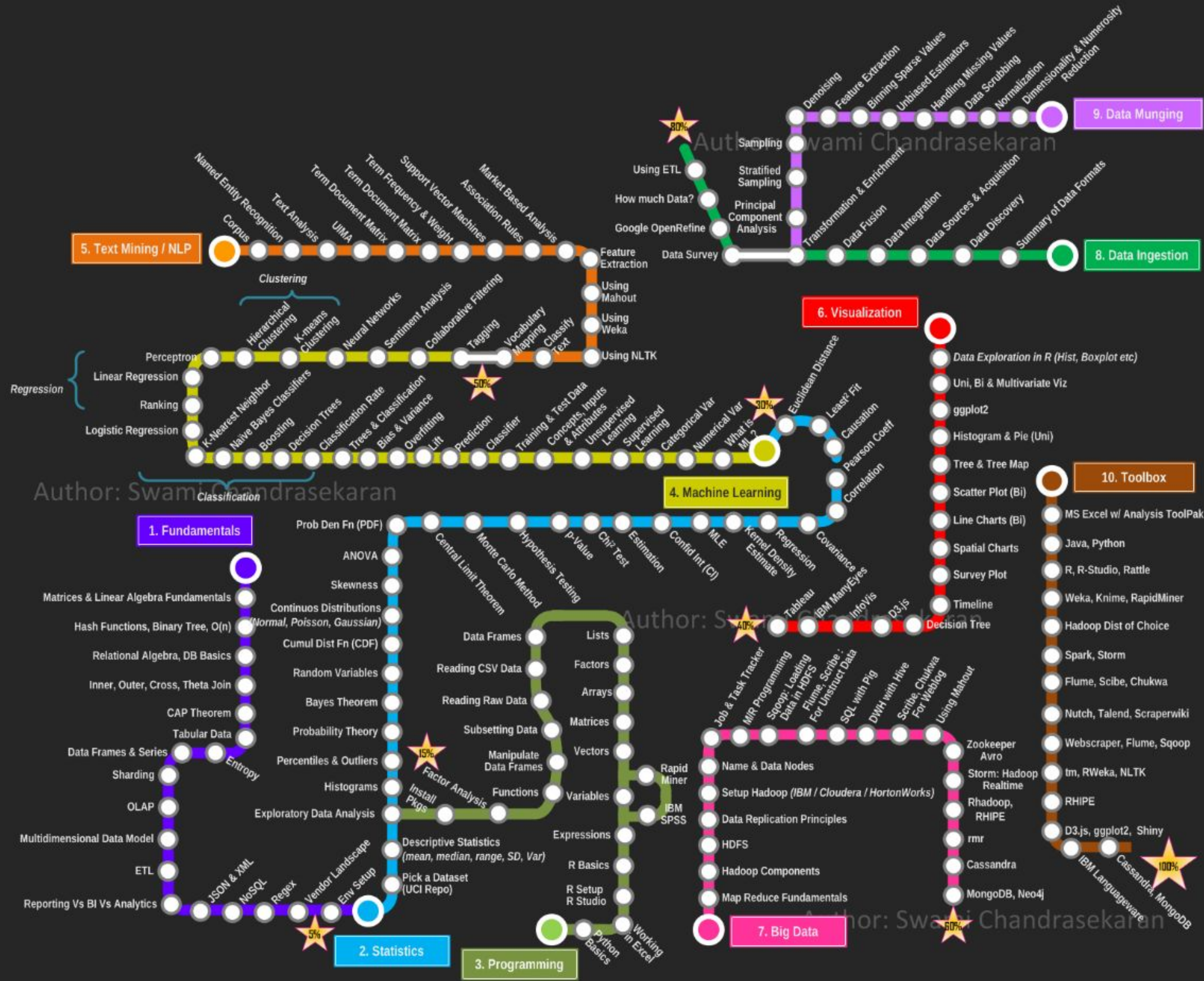
Fuente





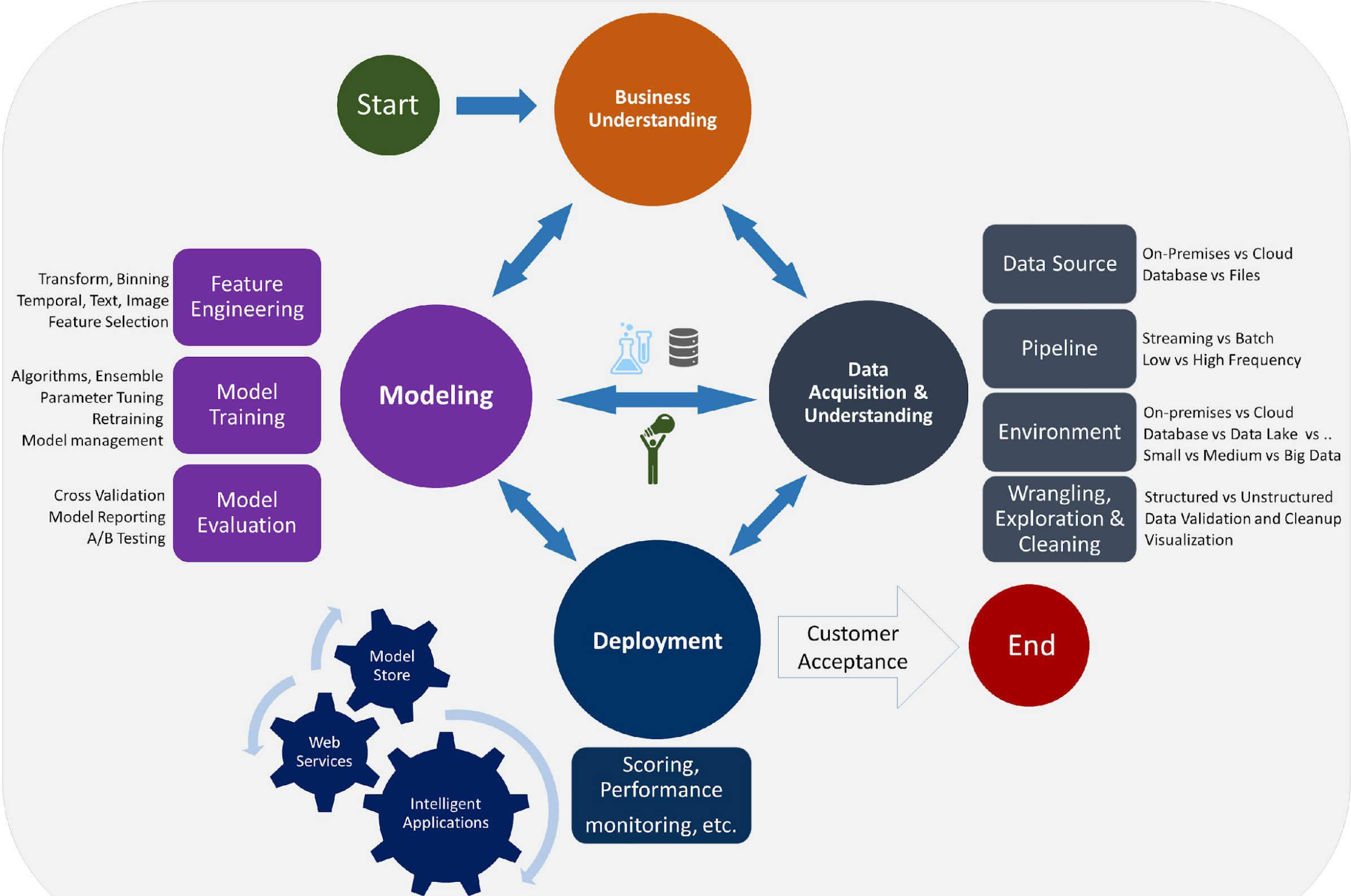
¿Qué piden las empresas?

Fuente



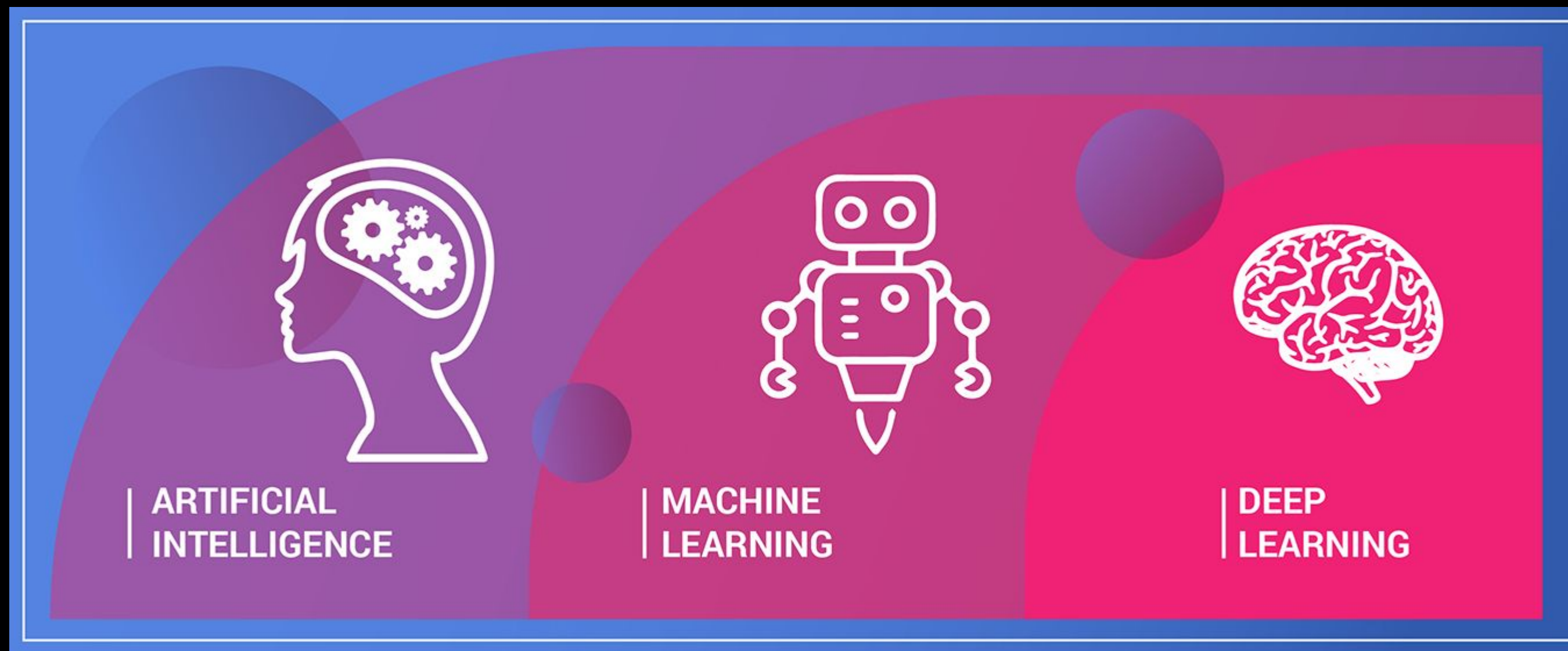


# Data Science Lifecycle



Ciclo de vida de  
un proyecto de  
Data Science





**IA** - Técnicas que posibilitan a máquinas **imitar la inteligencia humana**

**ML** - Algoritmos que dan a una máquina la **capacidad de aprender** a partir de datos.

**DL** - Rama de ML que utiliza **redes neuronales** con varias capas.

**NOT SURE IF HEAD IS ABOUT TO  
EXPLODE FROM CONFUSION**

**OR BRAIN IS GETTING BIGGER FROM  
KNOWLEDGE**

quickmeme.com



# Calendario

SEPTIEMBRE							OCTUBRE							NOVIEMBRE						
Lu	Ma	Mi	Ju	Vi	Sá	Do	Lu	Ma	Mi	Ju	Vi	Sá	Do	Lu	Ma	Mi	Ju	Vi	Sá	Do
				1	2	3							1			1	2	3	4	5
4	5	6	7	8	9	10	2	3	4	5	6	7	8	6	7	8	9	10	11	12
11	12	13	14	15	16	17	9	10	11	12	13	14	15	13	14	15	16	17	18	19
18	19	20	21	22	23	24	16	17	18	19	20	21	22	20	21	22	23	24	25	26
25	26	27	28	29	30		23	24	25	26	27	28	29	27	28	29	30			
							30	31												

DICIEMBRE							ENERO						
Lu	Ma	Mi	Ju	Vi	Sá	Do	Lu	Ma	Mi	Ju	Vi	Sá	Do
				1	2	3	1	2	3	4	5	6	7
4	5	6	7	8	9	10	8	9	10	11	12	13	14
11	12	13	14	15	16	17	15	16	17	18	19	20	21
18	19	20	21	22	23	24	22	23	24	25	26	27	28
25	26	27	28	29	30	31	29	30	31				

LEYENDA	
FESTIVO/NO LECTIVO	
RAMP UP	
CORE	
CR	Presentación Desafío / Periodo de Desafío
GRADUACIÓN	

# Contenido del Bootcamp II

## **Ramp Up**

**(15%)**

Aprenderás a manejarte con Python, sus principales herramientas y los fundamentos matemáticos de Data Science.

## **Data Analytics**

**(40%)**

Dominarás las técnicas de exploración y preparación de datos para su análisis y visualización.

## **Machine Learning**

**(30%)**

Construirás tus propios modelos de aprendizaje de datos.

## **Data Engineering**

**(15%)**

Conocerás cómo la ciencia de datos se convierte en un proceso de negocio y cómo desplegar nuestras soluciones de forma escalable



## Ramp Up

---

DS Toolkit

Git

Python

## Data Analysis

---

Exploratorio de  
datos

Visualización

Transformación  
de datos

SQL

## Machine Learning

---

Modelos  
supervisados

Modelos no  
supervisados

Time Series

Deep Learning

## Data engineering

---

Sistemas

Big Data

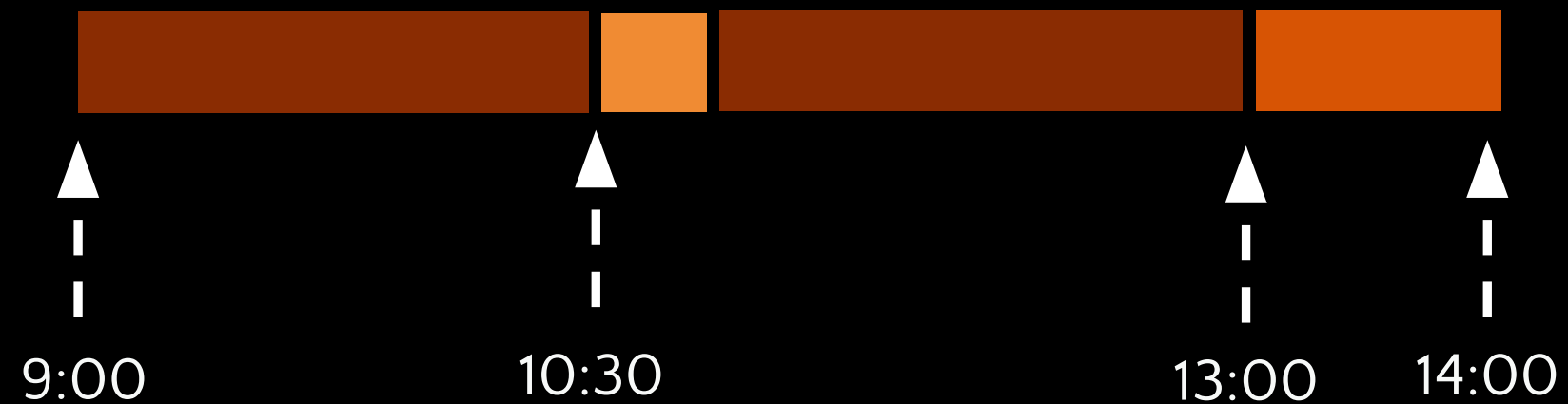
Apis

Cloud

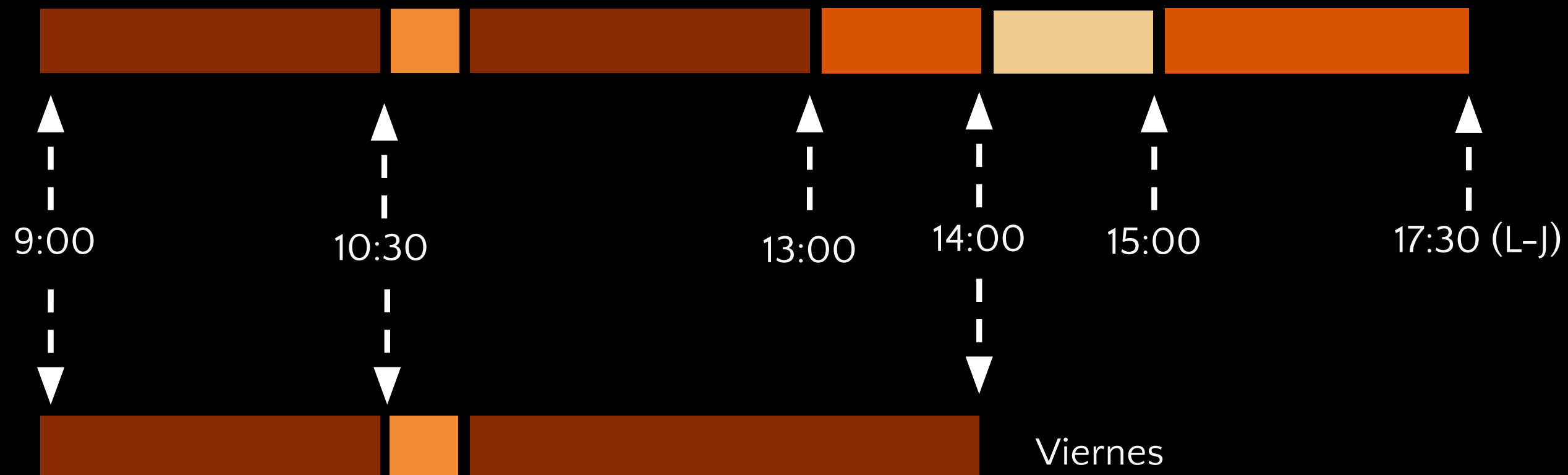


# Horario

## Ramp Up



## Bootcamp







Módulo	Evaluación	Formato
Ramp Up	Práctica	Entrega diaria
Ramp Up	Test	Test
Data Analysis	Proyecto individual	Presentación
Data Analysis	Proyecto individual	Presentación
Data Analysis	Proyecto grupal	Presentación
Machine Learning	Competición individual/grupal	Kaggle
Machine Learning	Proyecto individual	Presentación
Machine Learning	Proyecto grupal	Presentación
Data Engineering	Proyecto grupal	Presentación
Desafío Tripulaciones	Proyecto final	Presentación

# ¿Qué esperamos de ti?

- El ramp-up es **selectivo**. Habrá una prueba final.
- No entregar todos los **entregables**, faltar a clase o una mala actitud serán motivos para no aprobar el ramp-up. Los **proyectos** son obligatorios para graduarse en el bootcamp.
- Es obligatorio que l@s que estén en remoto tengan la **cámara** puesta todo el tiempo
- Las clases **serán grabadas** y se compartirá el enlace
- En clase se hablarán de **temas lectivos**.
- Habrá dos **delegados** (presencial + remoto)



# ¿Qué esperamos de ti?

- Anímate a **participar**, nos gusta que nos interrumpan.
- **Google es tu oráculo.** Todo lo que te puedas imaginar está ahí Lo que más vas a trabajar es el buscar respuestas de forma independiente.
- Tener una buena actitud de **respeto** a profesores y compañeros
- Atreverse a leer contenido en **inglés**
- Dormir bien, venimos a ofuscarnos, **tenemos que venir con energía.**

# Herramientas de clase

## SLACK

Canal del campus, de clase y canales privados. Utilizado en comunicaciones oficiales.

## GOOGLE CLASSROOM

Material de clase. Aquí se suben también las clases grabadas. **¡Necesitas una cuenta Google!**

## O'REILLY

Documentación, vídeos, libros de Data Science

## CALENDLY

Aplicación para reservar tutorías

## DISCORD

Aplicación para comunicación del desarrollo de la clase





**¡BIENVENIDOS!**

**Miguel - LI**

[miguel.nievas@thebridgeschool.es](mailto:miguel.nievas@thebridgeschool.es)

**Jared - TA**

[jared.rivas@thebridgeschool.es](mailto:jared.rivas@thebridgeschool.es)

**Santiago - TA**

[santiago.valencia@thebridgeschool.es](mailto:santiago.valencia@thebridgeschool.es)

**Yeray - TA**

[yeray.exposito@thebridgeschool.es](mailto:yeray.exposito@thebridgeschool.es)

**Preguntas?**