

# CURSO **CIENCIA DE DATOS**

---

4 DE JULIO DE 2022



**¿Quién  
soy?**

Ángeles Scetta



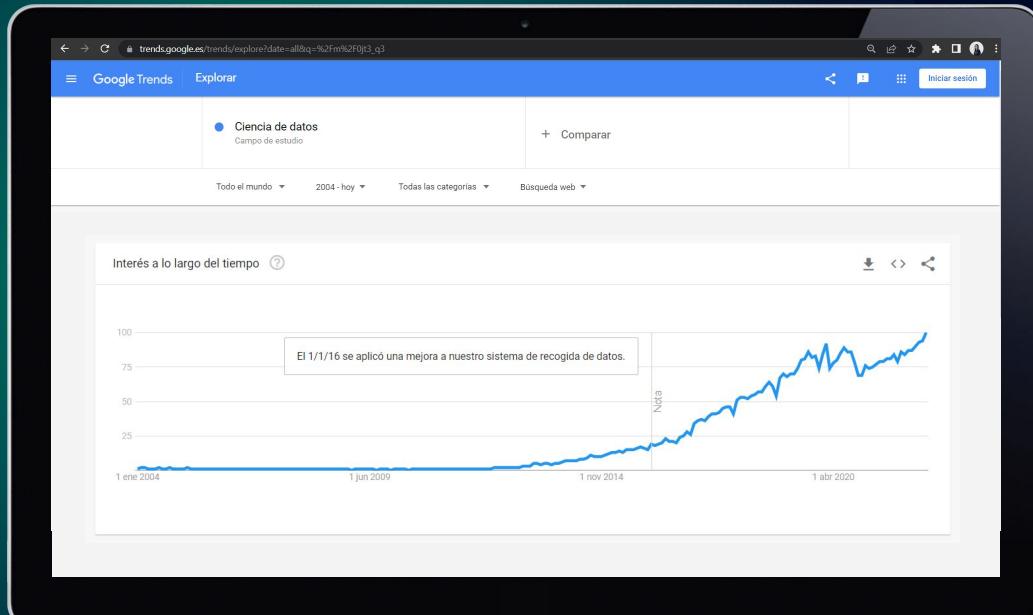
¿Quién  
soy?

Franco Galeano

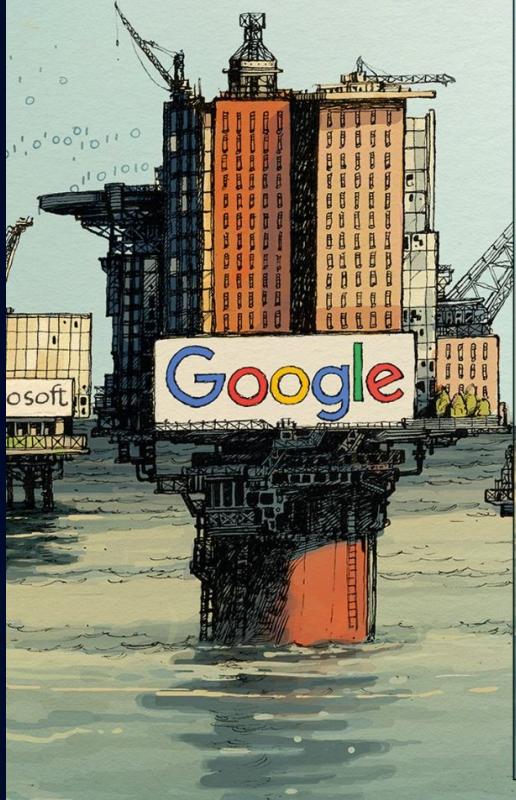
# ¿DE QUÉ VAMOS A HABLAR EN ESTE CURSO?

Datos **Políticas Públicas** Evidencia  
**Programación** Información Geográfica **Urbanismo**  
**Procesamiento** Visualización **Mapas** Gráficos Innovación  
**Inteligencia Artificial** Machine Learning Deep Learning  
Estadística **Datasets** Datos Abiertos **Transformación**  
**Digital** Cultura de Datos **Gobernanza de Datos** Científicos de Datos  
**Analistas de Datos** Equipos **Reutilización Código** Python  
**RStudio** Ciudades Planificación Urbana **Medición de Impacto**  
Decisiones **Arquitectura de Datos** **Ingeniería de Datos**  
Sensores IoT **Big Data** **Transformación Digital** Cultura de Datos  
**Gobernanza de Datos** Científicos de Datos **Analistas de Datos**  
Equipos **Reutilización Código** Python **RStudio**  
**Ciudades** Planificación Urbana **Medición de Impacto**  
Decisiones **Arquitectura de Datos** Ingeniería de Datos  
Sensores IoT **Big Data**

# ¿POR QUÉ HABLAR DE CIENCIA DE DATOS?



Fuente: <https://trends.google.es/trends/>



MAY 6TH-12TH 2017

The Economist

Crunch time in France  
Ten years on: banking after the crisis  
South Korea's unfinished revolution  
Biology, but without the cells

# The world's most valuable resource

Data and the new rules of competition

The cover of The Economist magazine from May 6th to 12th, 2017. The title 'The Economist' is at the top in a red box. Below it are four headlines: 'Crunch time in France', 'Ten years on: banking after the crisis', 'South Korea's unfinished revolution', and 'Biology, but without the cells'. The main headline 'The world's most valuable resource' is in large, bold black letters. Below it, the subtitle 'Data and the new rules of competition' is also in bold black letters. The background of the cover features a stylized illustration of several buildings under construction on stilts in the water, with binary code floating around them.



# El 'big data' revela qué elementos estimulan el dinamismo de una ciudad

LA REVISTA • EDICIÓN

## ¿Será posible reemplazar al Indec con big data y algoritmos?

El diluvio de datos que genera el avance de la tecnología instala el debate respecto de si se modificará la tradicional modalidad de encuestas para medir la realidad social; el catálogo de los desafíos que vienen de aquí en más

≡ EL PAÍS

## 'Big data' para evitar otra burbuja inmobiliaria

El desembarco de la tecnología aporta transparencia al sector y ayuda al particular a decidir si comprar o vender

NEGOCIOS

≡ EL PAÍS

ESPACIO ECO ›

## Tecnologías inteligentes al rescate de las ciudades

'Big data', 'machine learning' y nuevas metodologías: ejemplos de cómo la innovación y la ciencia ciudadana pueden lograr urbes más sostenibles

PLANETA  
FUTURO 

≡ EL PAÍS

## 10 maneras de mejorar nuestra calidad de vida usando 'big data'

La importancia de los datos para poder alcanzar los Objetivos del Milenio es grande. Algunos ejemplos que ilustran cómo podemos beneficiarnos de ellos

**infobae**  
AMÉRICA TELESHOW DEPORTES TENDENCIAS CULTURA MIX5411

TECNO

## El impacto del Big Data en las ciudades inteligentes

Cómo se emplea la información que se recolecta por medio de cámaras y sensores en la administración pública. Beneficios y desafíos de la sociedad de la información

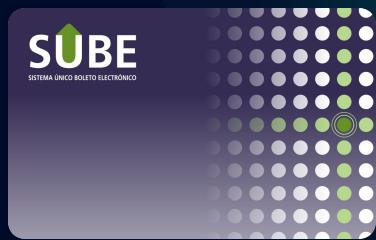
# ¿CÓMO SE GENERAN LOS DATOS?



- TELÉFONOS CELULARES
- APPS MÓVILES
- REDES SOCIALES
- SISTEMAS DE BICICLETAS PÚBLICOS
- SISTEMAS DE BOLETO ELECTRÓNICO
- TARJETAS BANCARIAS
- SENSORES Y TECNOLOGÍAS IOT EN DIFERENTES COMPONENTES DE LA CIUDAD (SMART CITIES)

## ¡DATOS URBANOS!

...para tomar mejores decisiones a la hora de desarrollar políticas públicas



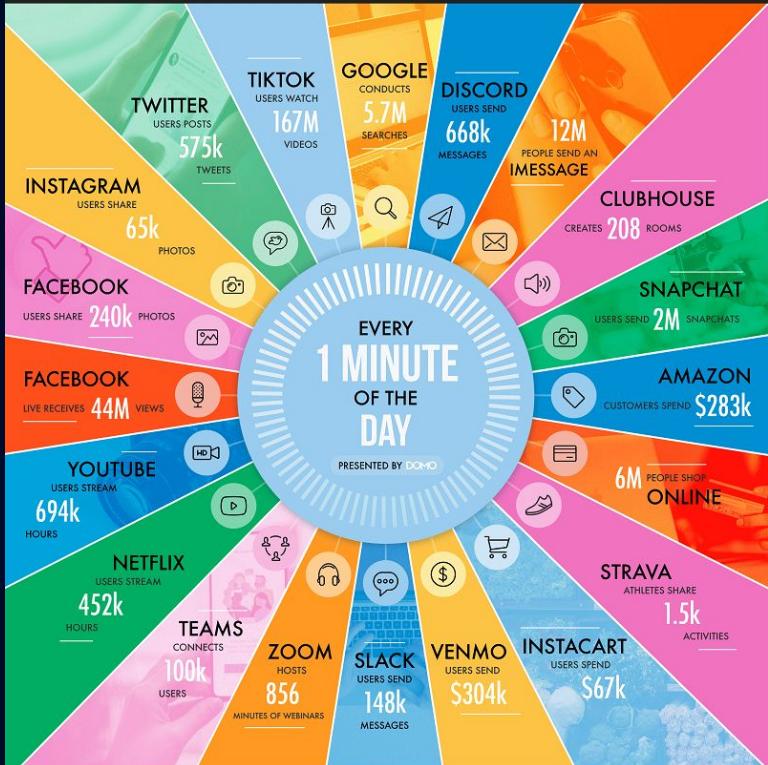
Y MUCHOS MÁS...



# Data Never Sleeps 9.0

## How much data is generated every minute?

The 2020 pandemic upended everything, from how we engage with each other to how we engage with brands and the digital world. At the same time, it transformed how we eat, how we work and how we entertain ourselves. Data never sleeps and it shows no signs of slowing down. In our 9th edition of the "Data Never Sleeps" infographic, we bring you a glimpse of how much data is created every digital minute in our increasingly data-driven world.



As of July 2021, the internet reaches 65% of the world's population and now represents 5.17 billion people—a 10% increase from January 2021. Of this total, 92.6 percent accessed the internet via mobile devices. According to Statista, the total amount of data consumed globally in 2021 was 79 zettabytes, an annual number projected to grow to over 180 zettabytes by 2025.

### Global Internet Population Growth (IN BILLIONS)



As the world changes, businesses need to change too—and that requires data. Domo gives you the power to make data-driven decisions at any moment, on any device, so that you can make smart choices in a rapidly changing world. Every click, swipe, share, or like tells you something about your customers and what they want, and Domo is here to help you and your business make sense of all of it.

Learn more at [domo.com](https://domo.com)



SOURCES: LOCAL IQ, BUSINESS OF APPS, DUSTIN STOUT, HOOTSUIT, EXPANDED RAMBLINGS, INTERNET WORLD STATS, STATISTA, CNBC, BRANDWATCH, KILL THE CABLE BILL, YOUTUBE, INSTA, THE VERGE, MANAGEMENT COMMUNICATION: A CASE ANALYSIS APPROACH, INTERNET LIVE STATS, SODA, STATISTA

Fuente: DOME

<https://www.domo.com/blog/what-data-never-sleeps-9-0-proves-about-the-pandemic/>



# ¡BIG DATA!

CONJUNTO DE DATOS QUE POR SU  
**VOLUMEN, VELOCIDAD Y**  
**VARIEDAD** NECESITAN SER  
PROCESADOS CON HERRAMIENTAS  
ESPECÍFICAS



## VOLUMEN

Gran cantidad de datos/información.



## VELOCIDAD

Datos que se actualizan con frecuencia alta.



## VARIEDAD

Datos que por sus diferentes orígenes, presentan diversos formatos:

- **Estructurados:** Bases de Datos Tradicionales.
- **No Estructurados:** Correos Electrónicos, Videos, Audios, Imágenes, etc.



**Datos**



**Información/  
Evidencia**



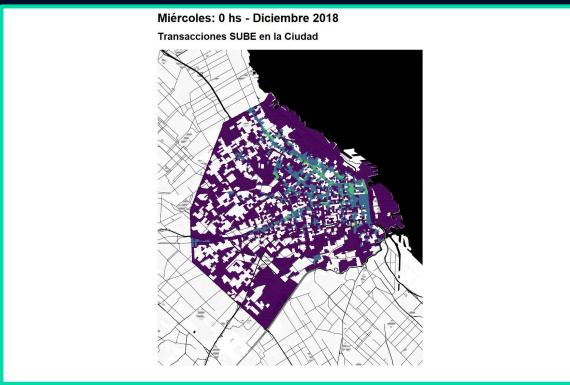
**Decisiones**



**Políticas  
públicas**



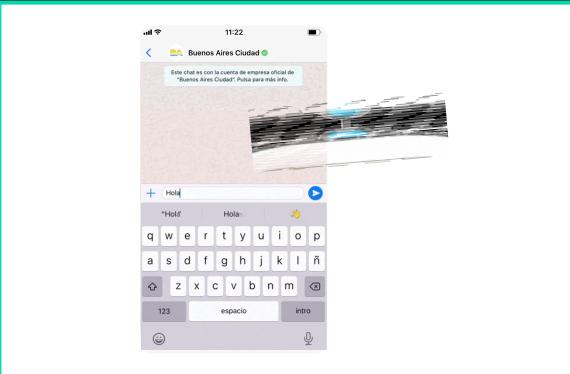
# **POLÍTICAS PÚBLICAS BASADAS EN EVIDENCIA**



## Movilidad en la Ciudad



## Ciudad On Demand



**Boti no te entiendo!**

BA Data

Datasets Organizaciones Historias APIs

Buenos Aires Data

Estos son datos públicos generados, guardados y publicados por el Gobierno de la Ciudad de Buenos Aires.

¿Qué dataset buscas?

Ordenar por Última modificación

Temas

- Administración pública (73)
- COVID-19 (28)
- Economía y finanzas (47)
- Desarrollo humano (37)
- Economía y finanzas (34)
- Desarrollo humano (30)
- Gobierno (42)
- Medio Ambiente (46)
- Seguridad (30)
- Salud (44)
- Seguridad (30)
- Urbanismo y territorio (87)

Organizaciones

- Secretaría Legal y Técnica
- Ministerio de Salud
- Agencia de Protección Ambiental (Apha)
- Secretaría de Ambiente, Jefatura de Gobierno

Calidad de Aire

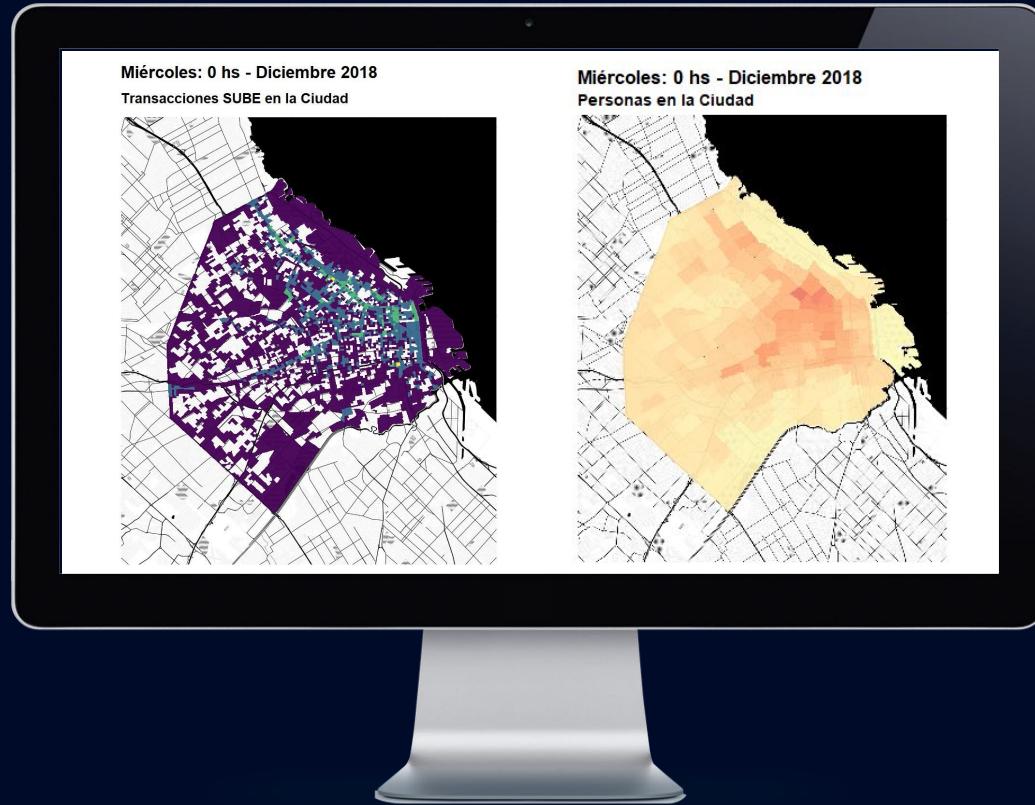
Información sobre la calidad de aire medida de niveles de carbono, dióxido de nitrógeno y material particulado registrada entre a 10 metros medidas en las estaciones de La Boca, Corrientes, Palermo,

**BA Data**

# CIRCULACIÓN DE PERSONAS

*Transporte público (SUBE)*

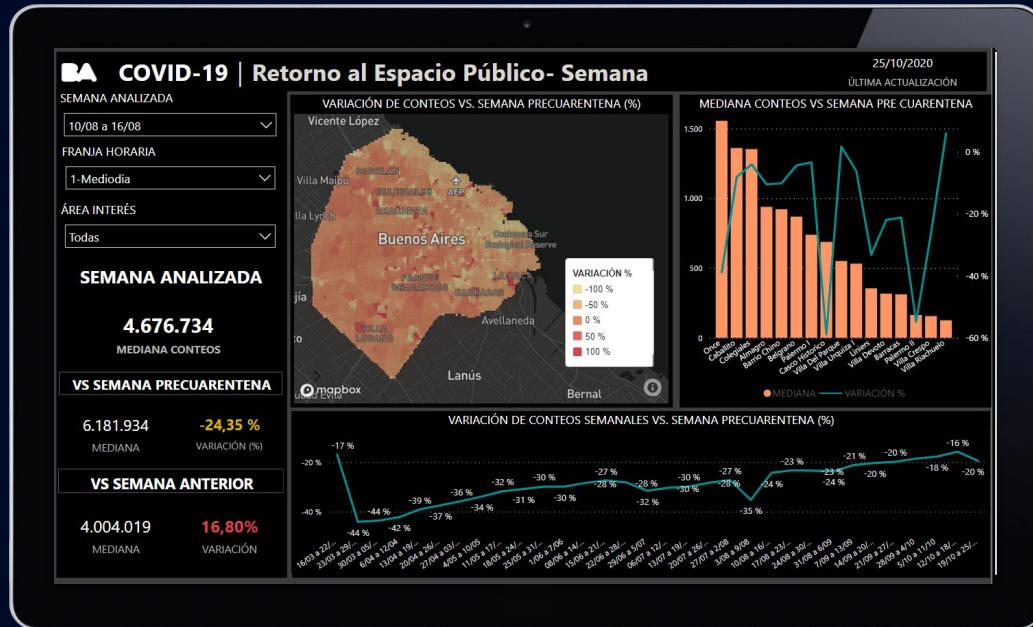
*Antenas de telefonía móvil*



# CIRCULACIÓN DE PERSONAS

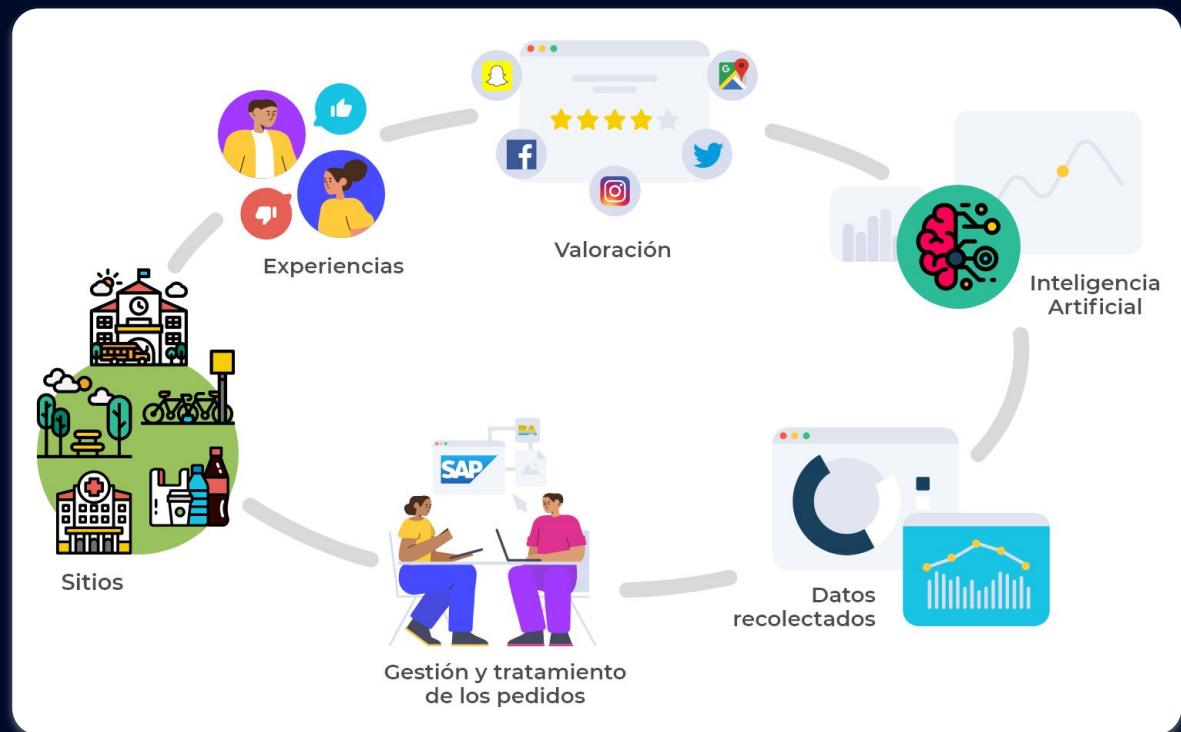
*Transporte público (SUBE)*

*Antenas de telefonía móvil*

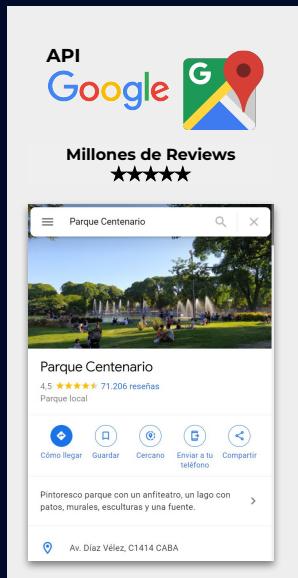


Los datos fueron reemplazados por datos ficticios con el fin de conservar la confidencialidad de la información.

# CIUDAD ON DEMAND



# CIUDAD ON DEMAND



# CIUDAD ON DEMAND

SENTIMIENTO

SITIO

RECLAMO

PROPUESTA



Julieta Vecina

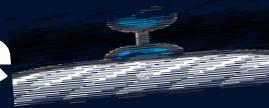
Local Guide · 36 reviews

★★★★★ una semana atrás

Excelente plaza para pasar la tarde, siempre muy limpia y agradable. Me preocupa un poco la oscuridad a la noche. Si hubiese luz nocturna, sería un lugar ideal para actividades físicas al aire libre o simplemente pasear el perro...



# Soy BOTI, el chatbot de la Ciudad de Buenos Aires



+549 11-5050-0147

¿En qué te puedo ayudar?

## CONTENIDO FALTANTE



12:01 🔍

Boti Whatsapp de la ciudad  
Buenos Aires Ciudad ✓

Hola! Quiero información de colonias de chicos

12:01 ✓

Creo que no te entendí bien. 😊

Intentemos de nuevo, ¿dale?

12:01

Escríbeme algunas palabras claves sobre tu consulta. Por ejemplo: 'inscripción escuelas', 'turismo ciudad'.

12:01

## FALSOS POSITIVOS



11:56 🔍

Boti Whatsapp de la ciudad  
Buenos Aires Ciudad ✓

Hola! Cómo puedo modificar los datos de mi Buenos Aires?

11:56 ✓

¡Sí, hablemos de turismo!

Es de mis temas favoritos. 😊 11:56

Seas turista o profesional del turismo, tengo mucho para contarte. 11:56

Don't understand Spanish? Don't worry!

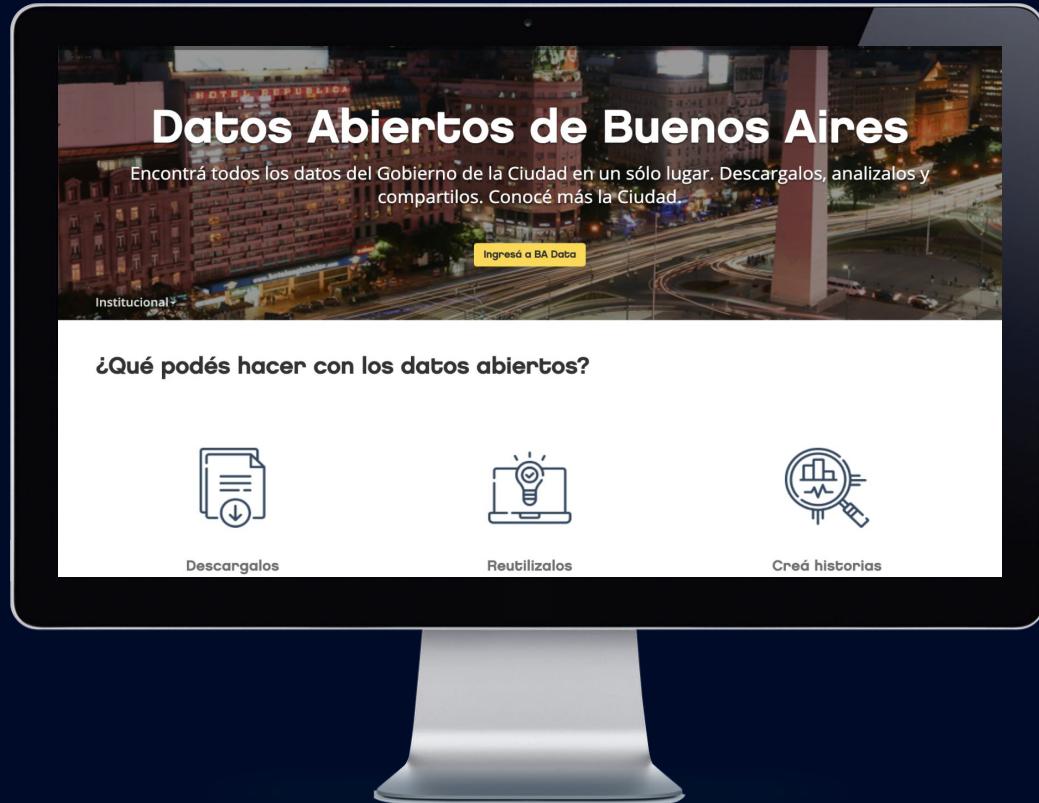
Choose 'I speak in English' and I'll help you. 😊 11:56

¿Cuál es tu caso? Elegí y sigamos hablando:

A. Soy turista  
B. I speak in English  
C. Trabajo en turismo 11:56

# BA DATA

[data.buenosaires.gob.ar/dataset](http://data.buenosaires.gob.ar/dataset)



# BA DATA

[data.buenosaires.gob.ar/dataset](http://data.buenosaires.gob.ar/dataset)



# ¿PREGUNTAS?

# CURSO **CIENCIA DE DATOS**

---

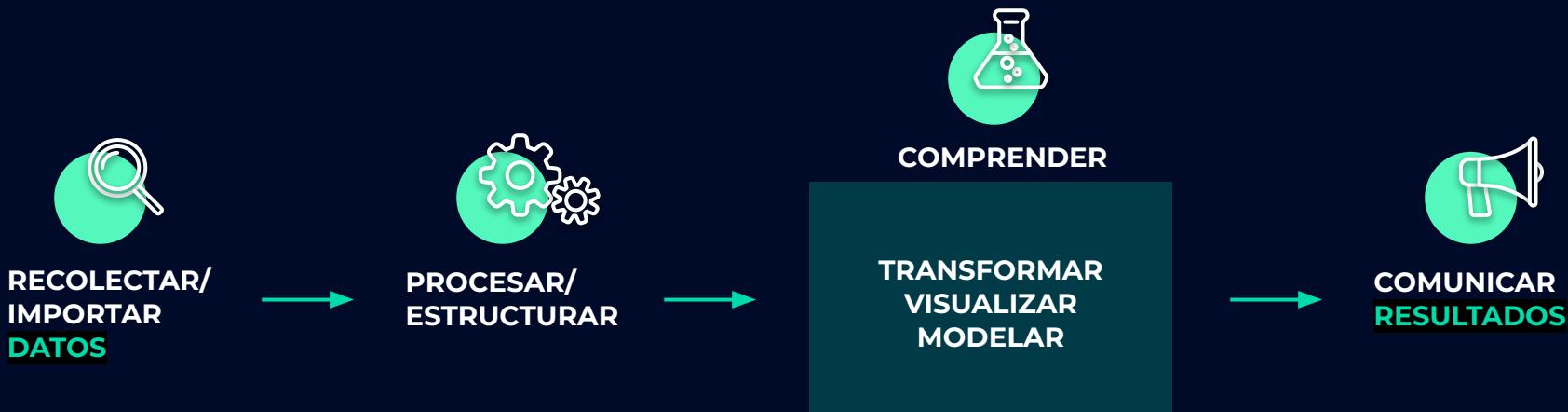
PROGRAMA 2022



UNIVERSIDAD  
TORCUATO DI TELLA

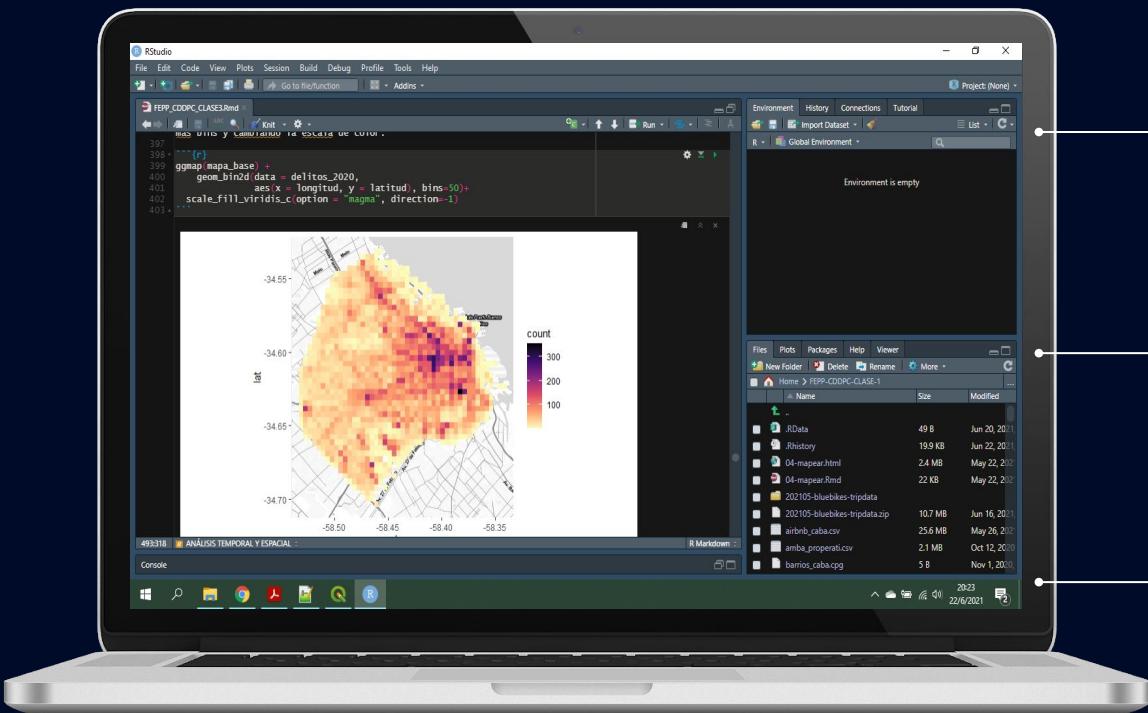


# ¿QUÉ VAMOS A APRENDER EN ESTE CURSO?



# ¿QUÉ HERRAMIENTA VAMOS A USAR?

## OPEN SOURCE



LENGUAJE DE PROGRAMACIÓN



INTERFAZ GRÁFICA



LIBRERÍA PARA GENERAR INFORMES



# ¿QUÉ DATOS VAMOS A ANALIZAR?

## PORTALES DE DATOS ABIERTOS

<https://data.buenosaires.gob.ar/>



<https://datos.gob.ar/>



<https://catalogo.datos.gba.gob.ar/>



<https://datos.bancomundial.org/>



<https://www.idera.gob.ar/>



<https://www.properati.com.ar/data>



<https://www.indec.gob.ar/>



y algunos más!

# PROGRAMA

DÍA Y HORARIO **Lunes 19h**

**CLASE 1** Introducción a la ciencia de datos y a la programación en R **PRESENCIAL**

**CLASE 2** Análisis exploratorio y transformación de datos

**CLASE 3** Pre procesamiento de datos

**ENTREGA PRIMER EJERCICIO GRUPAL** (*Grupos de 3 alumnos*)

**CLASE 4** Visualización de datos I **PRESENCIAL**

**CLASE 5** Visualización de datos II

**CLASE 6** Visualización de datos III

**CLASE 7** Detección de patrones temporales

**ENTREGA SEGUNDO EJERCICIO GRUPAL** (*Grupos de 3 alumnos*)

**CLASE 8** Información geográfica y mapas **PRESENCIAL**

**CLASE 9** Análisis y geoprocесamiento de datos

**CLASE 10** Descarga y análisis de datos espaciales

**ENTREGA TERCER EJERCICIO GRUPAL** (*Grupos de 3 alumnos*)

**¡BIENVENIDOS AL  
MUNDO DE LOS  
DATOS!**

CAMPUS  
VIRTUAL