

**Licenciatura en
Ciencia de Datos
ECYT_UNSAM**

Bienvididxs

A LCD

Campus Miguelete, 13 de marzo de 2023

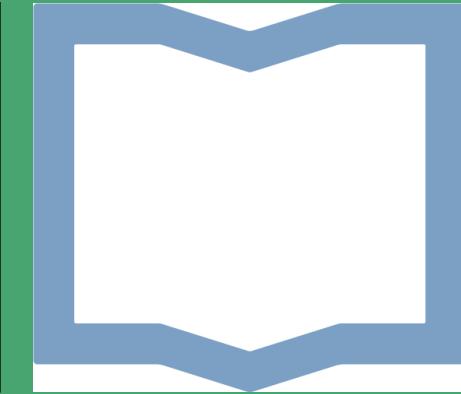
En conjunto



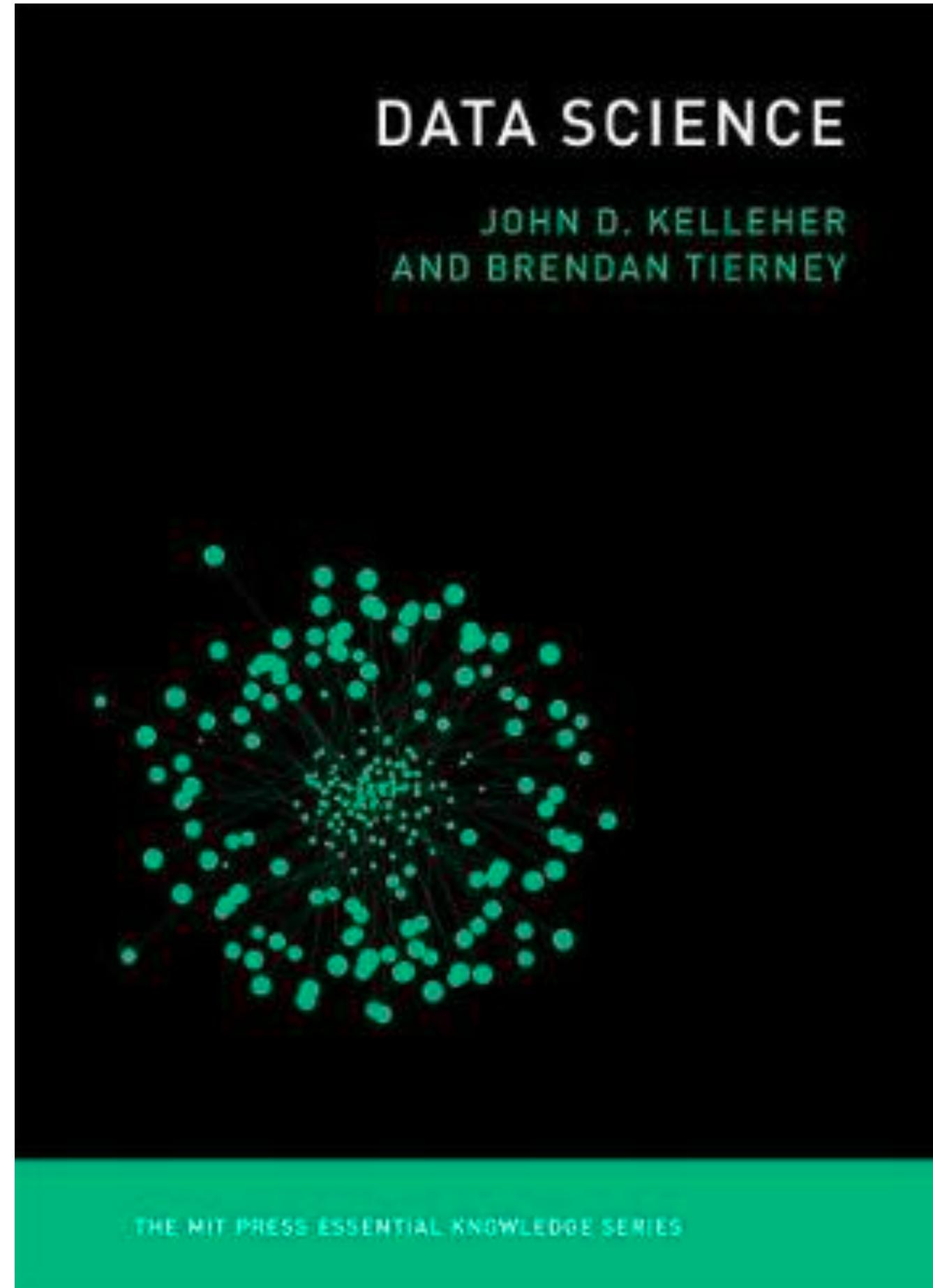
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

- Materia ICD
- Carrera LCD
- Proyecto de Datos/IA de la Universidad
- Sociedad

¿Para qué? ¿Qué?



Licenciatura en
Ciencia de Datos
ECYT_UNSAM



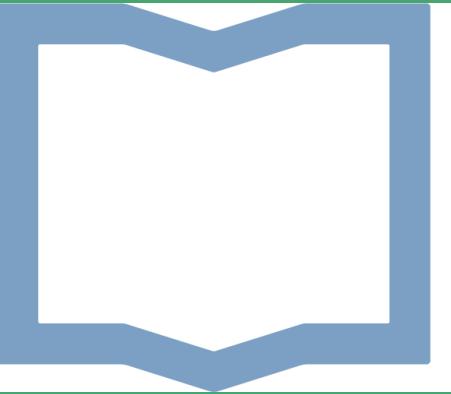
Para qué

"Mejorar la toma de decisiones basándose en los conocimientos extraídos de (grandes) conjuntos de datos."

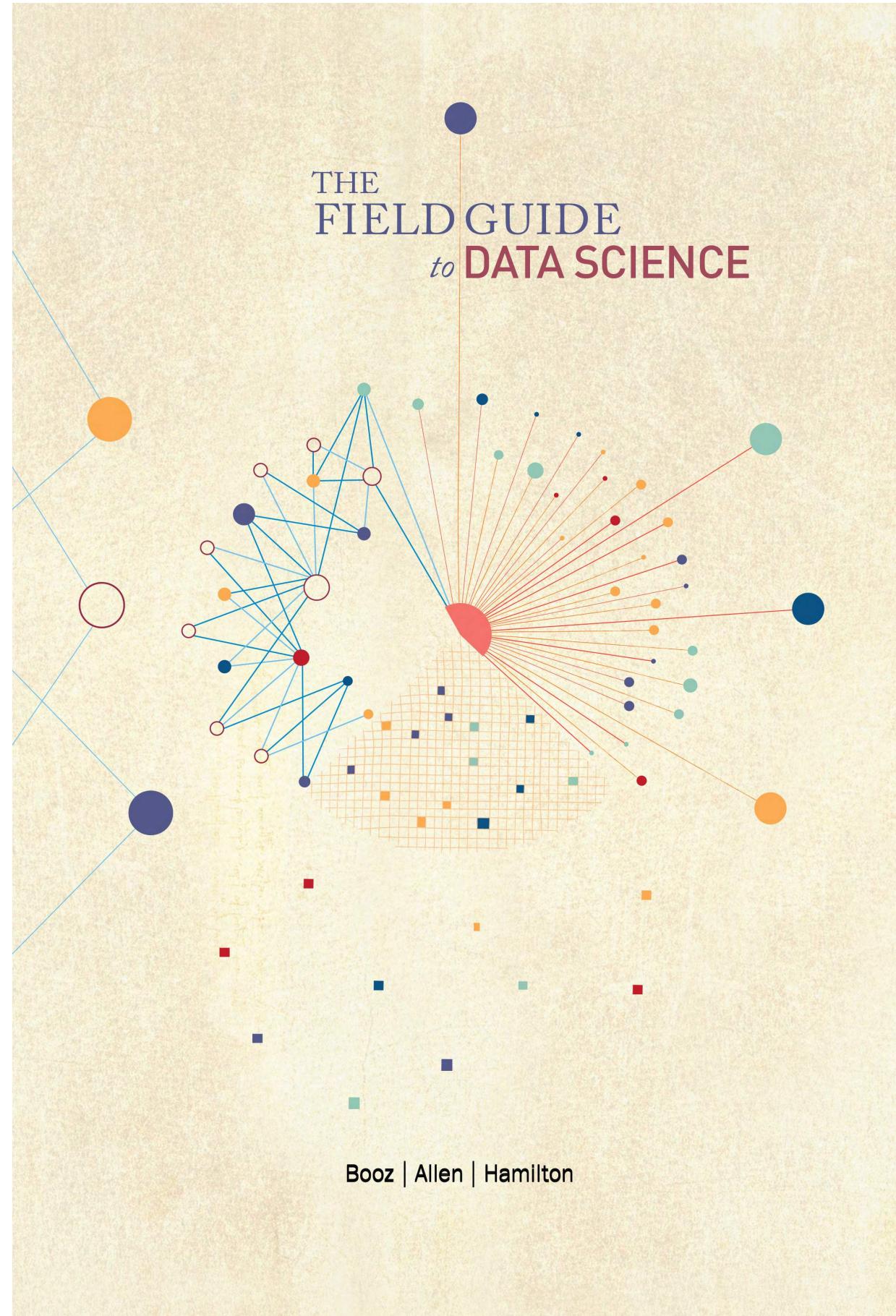
Qué

"Un conjunto de principios, definiciones de problemas, algoritmos, y procesos para extraer patrones no obvios y útiles a partir de grandes conjuntos de datos."

¿Para qué? ¿Qué?



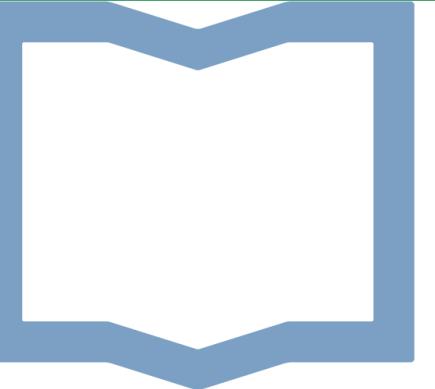
Licenciatura en
Ciencia de Datos
ECYT_UNSAM



"Data Science is the art of turning data into actions."

"La Ciencia de Datos es el arte de convertir a los datos en acciones"

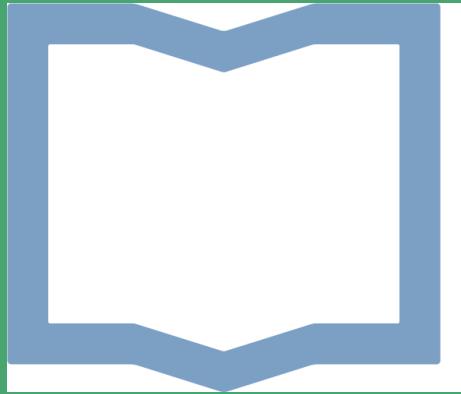
¿Qué es?



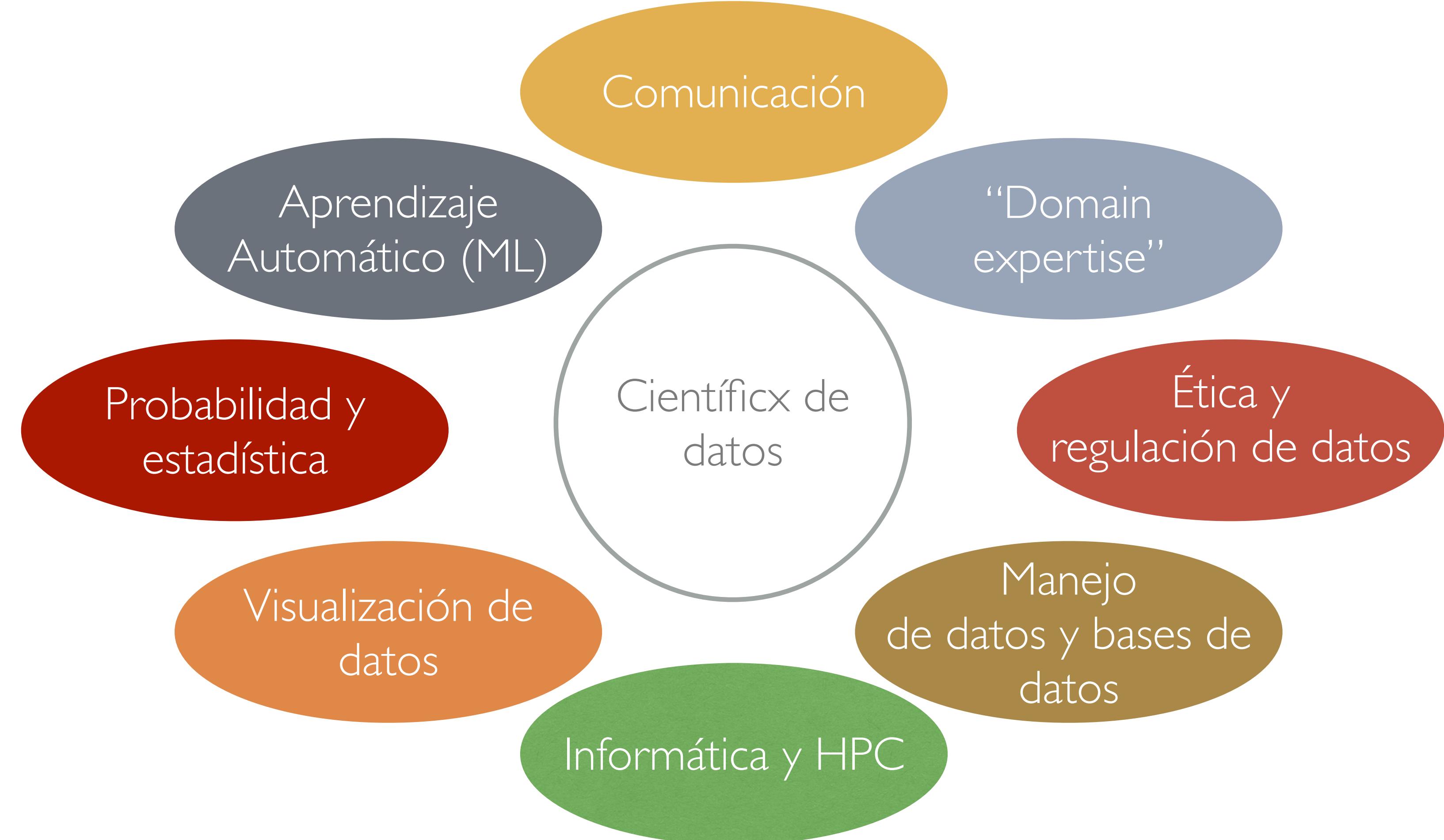
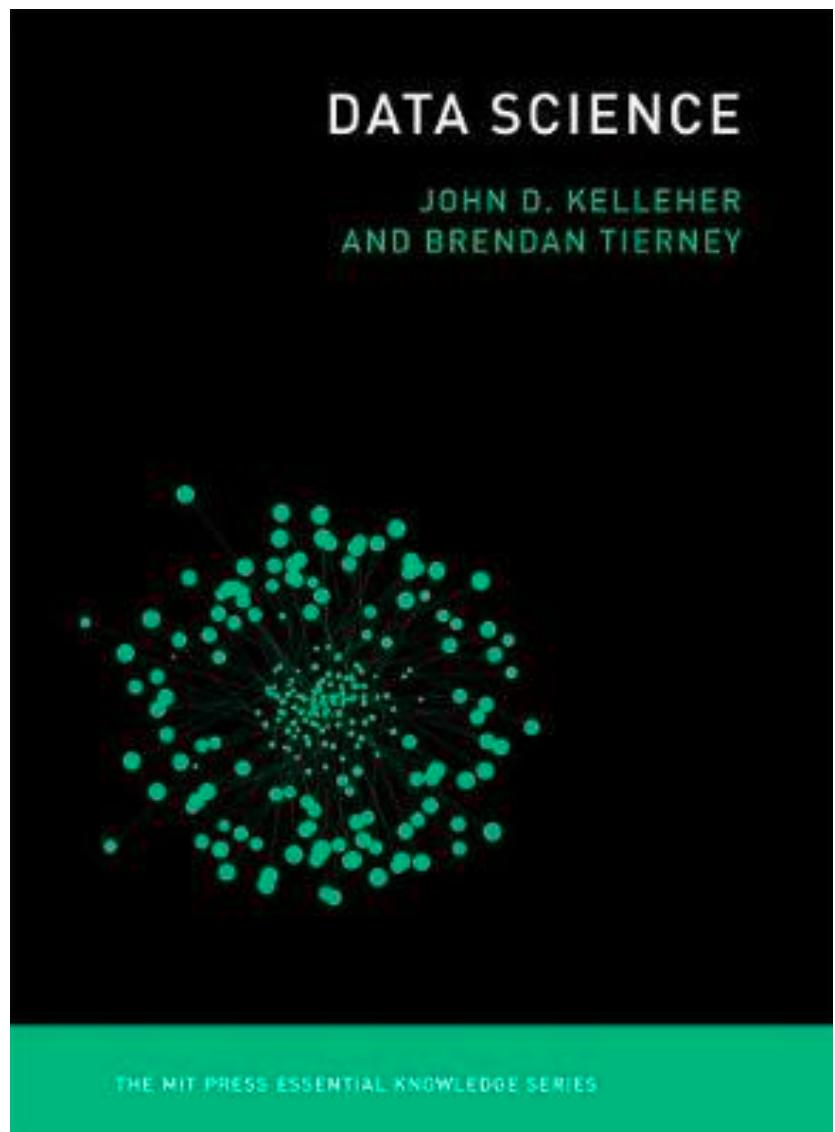
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

ES UNA
PRÁCTICA

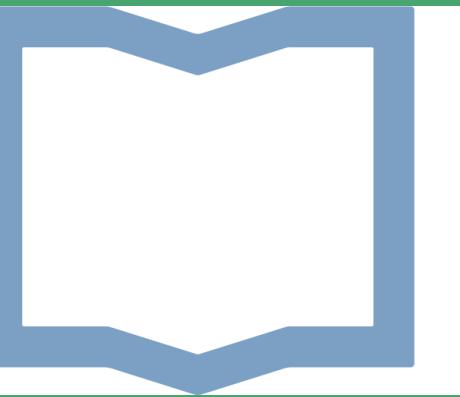
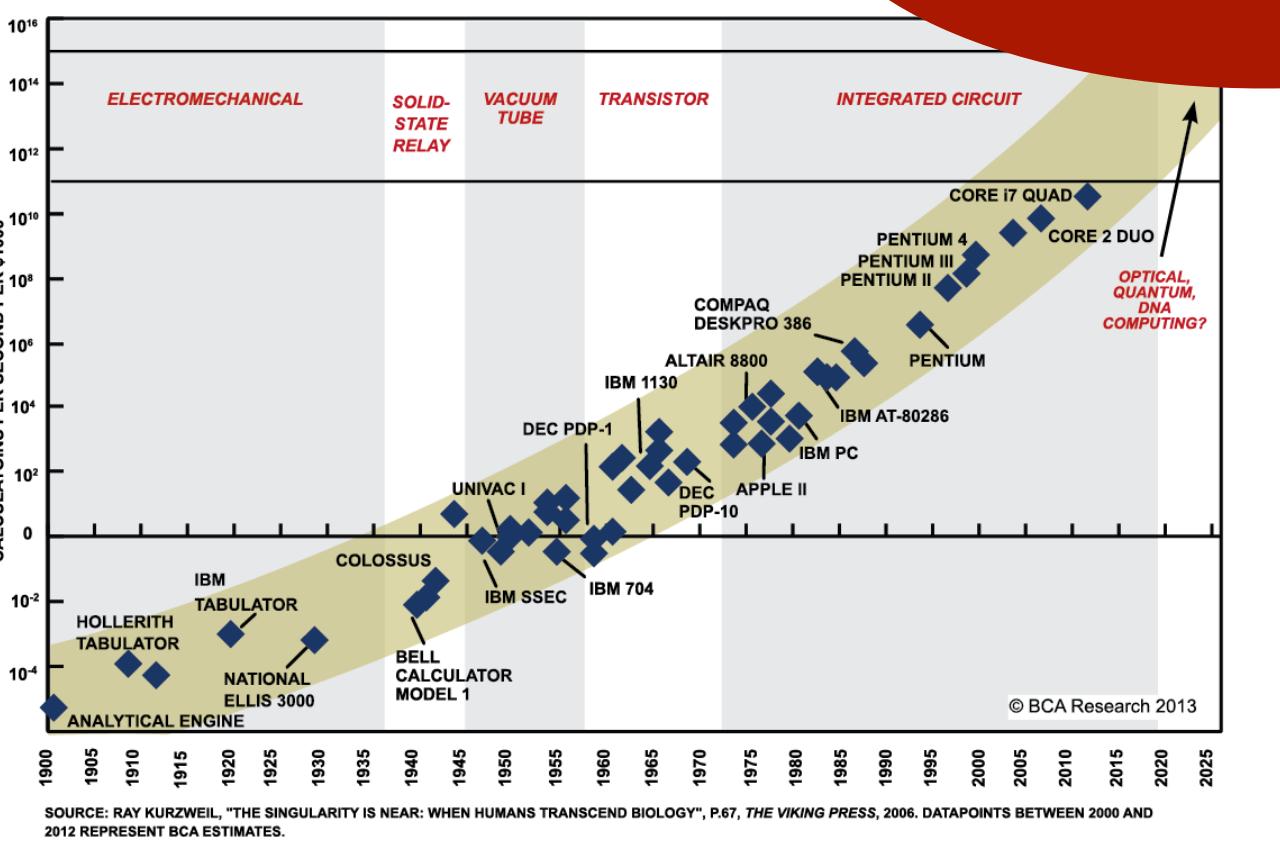
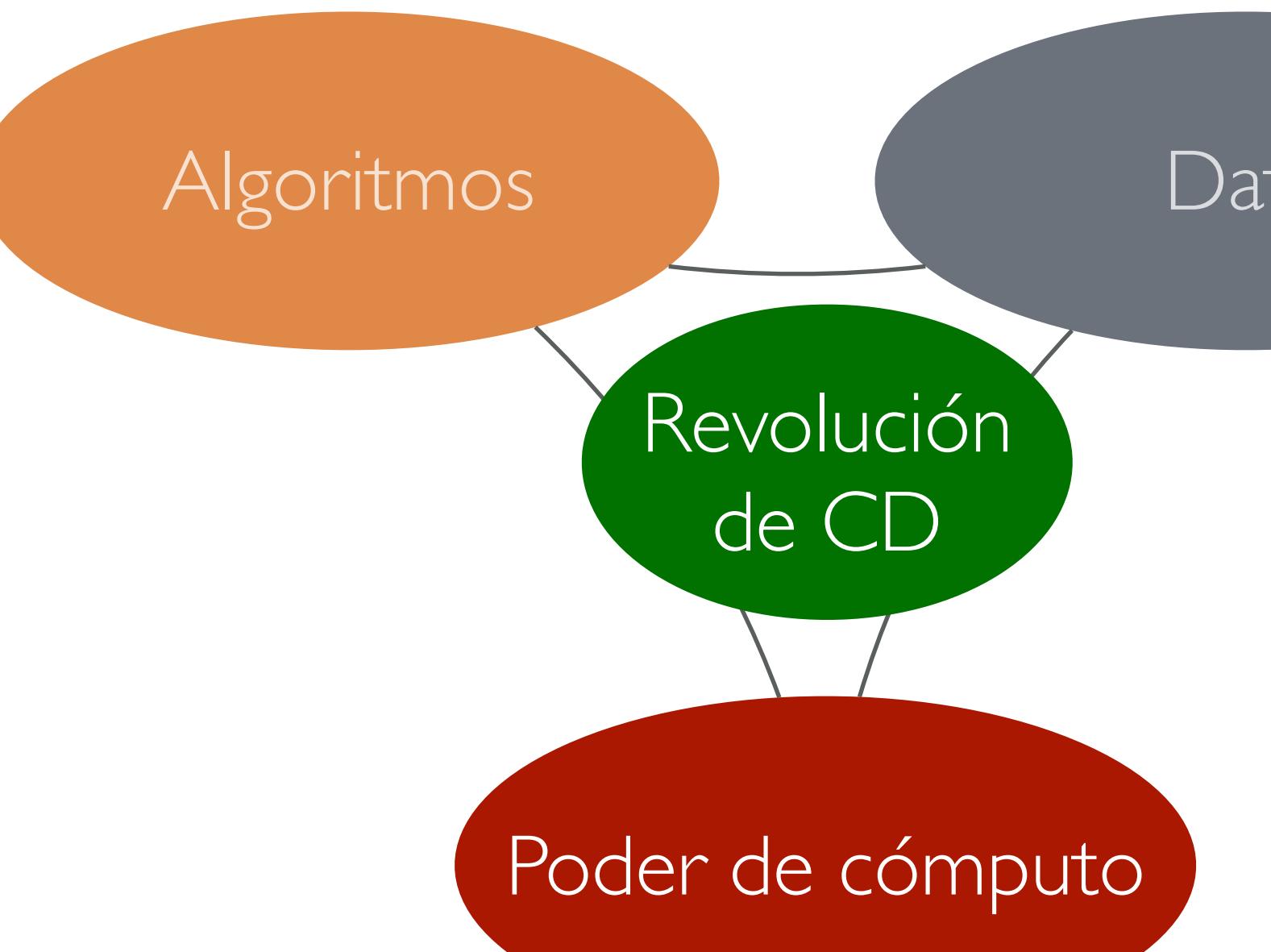
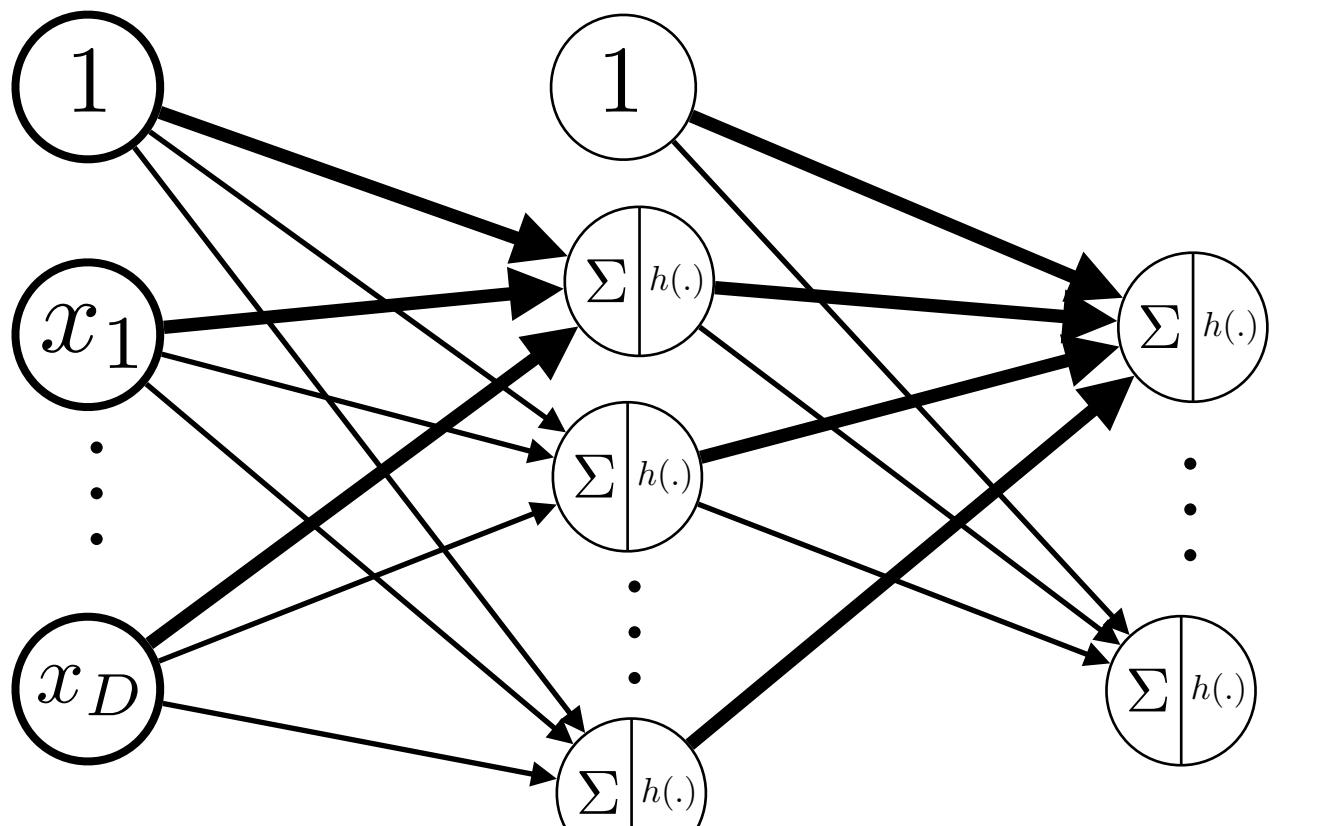
¿Cómo?



Licenciatura en
Ciencia de Datos
ECYT_UNSAM



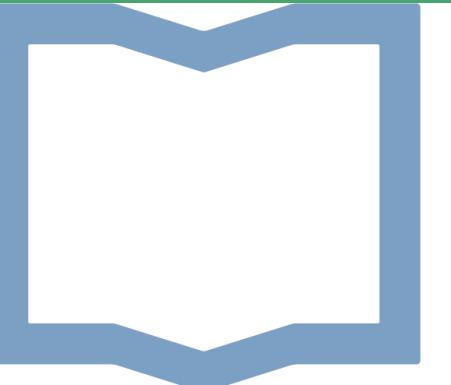
Por qué ahora



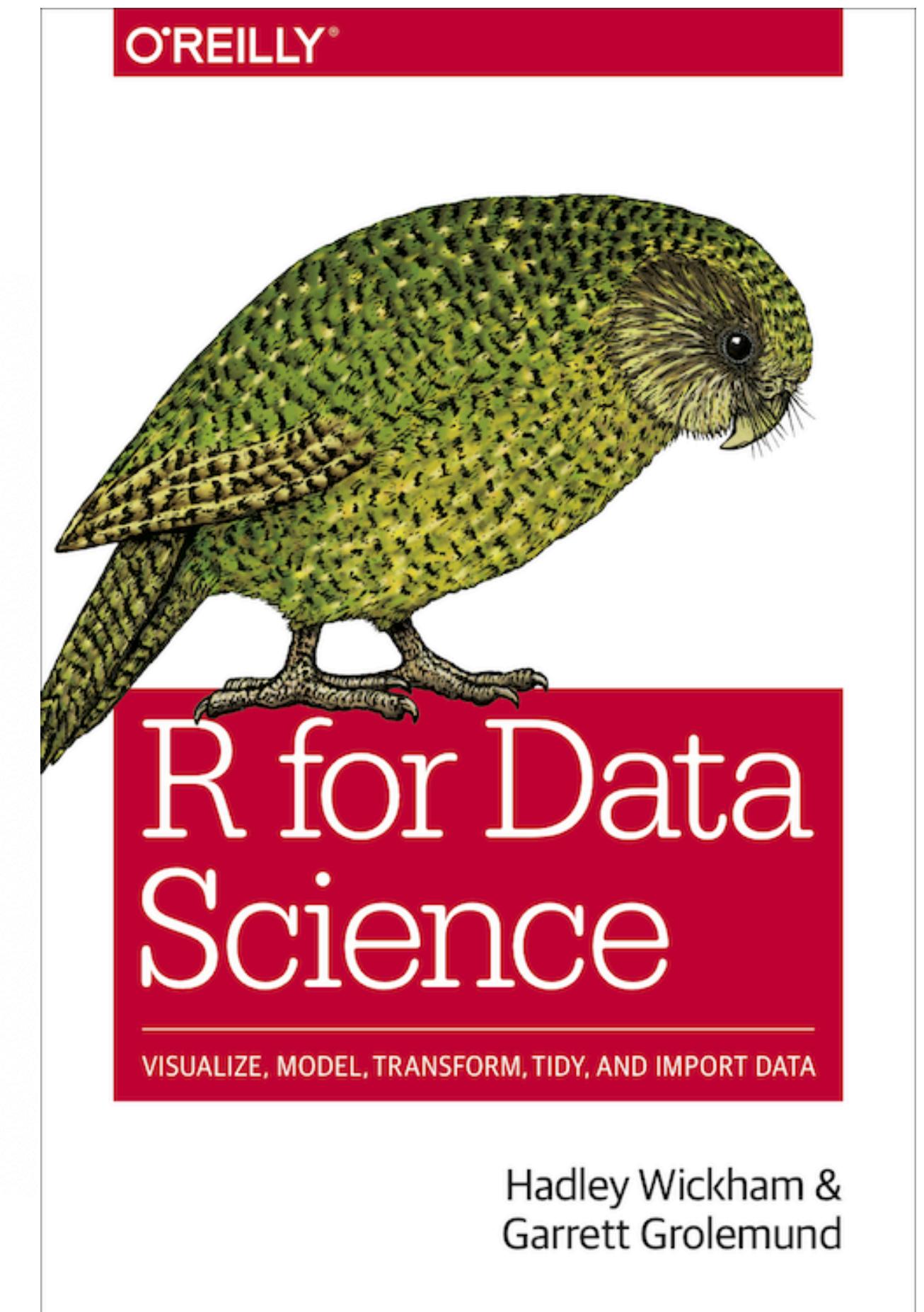
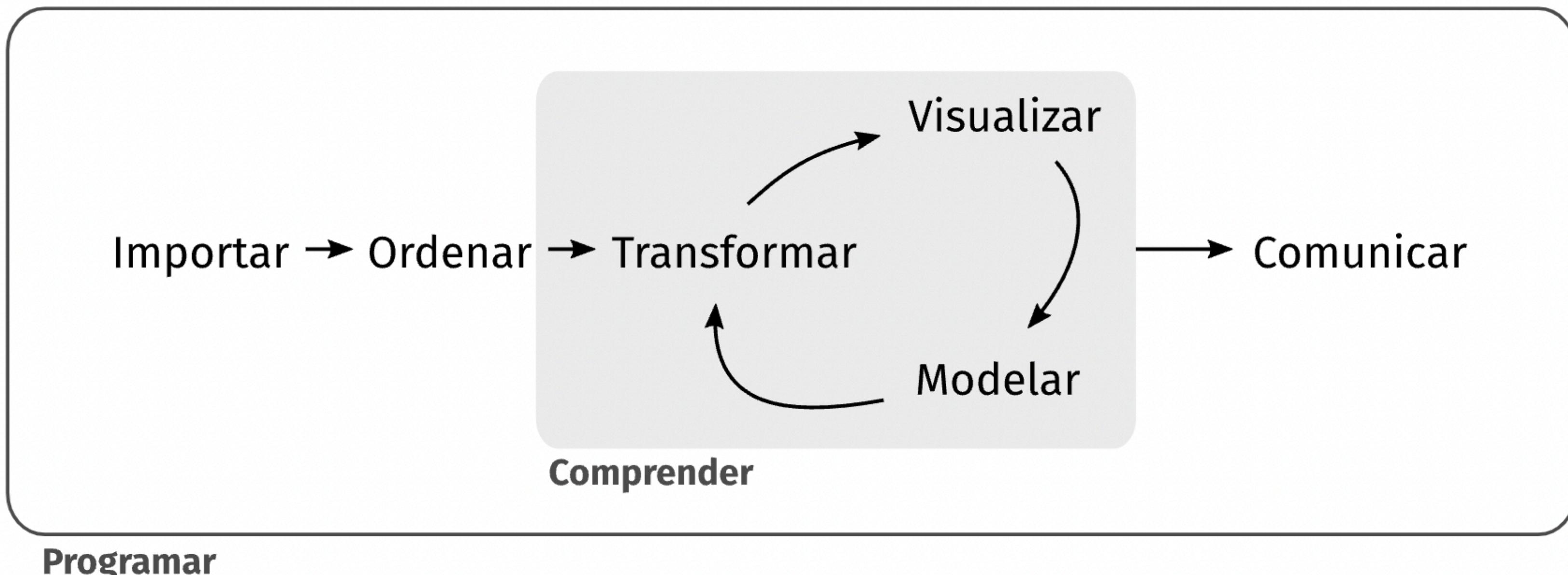
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

El proceso de la Ciencia de Datos

Un esquema básico



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

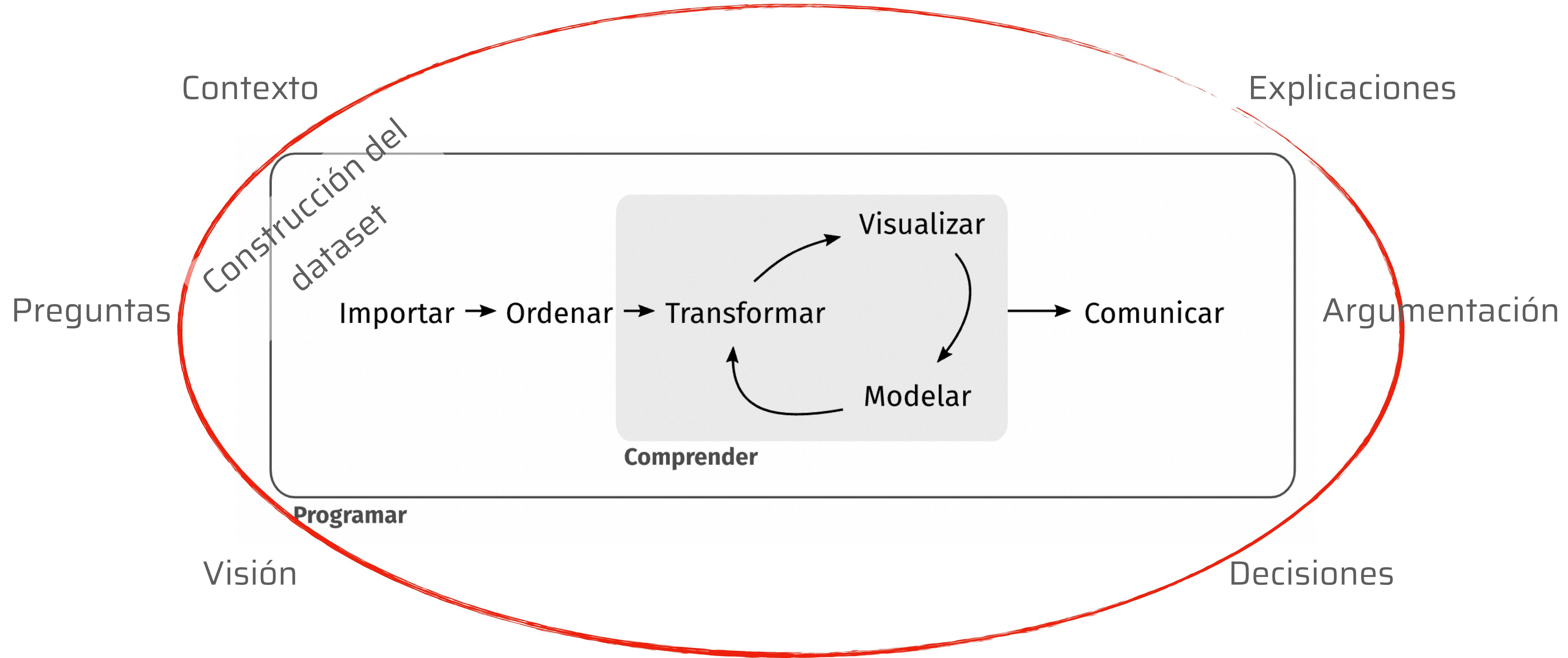


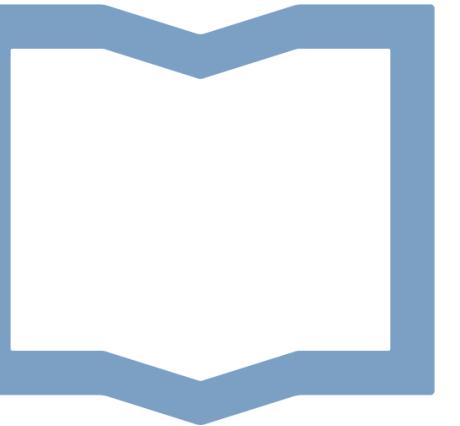
El proceso de la Ciencia de Datos

Un esquema básico



Licenciatura en
Ciencia de Datos
ECYT_UNSAM





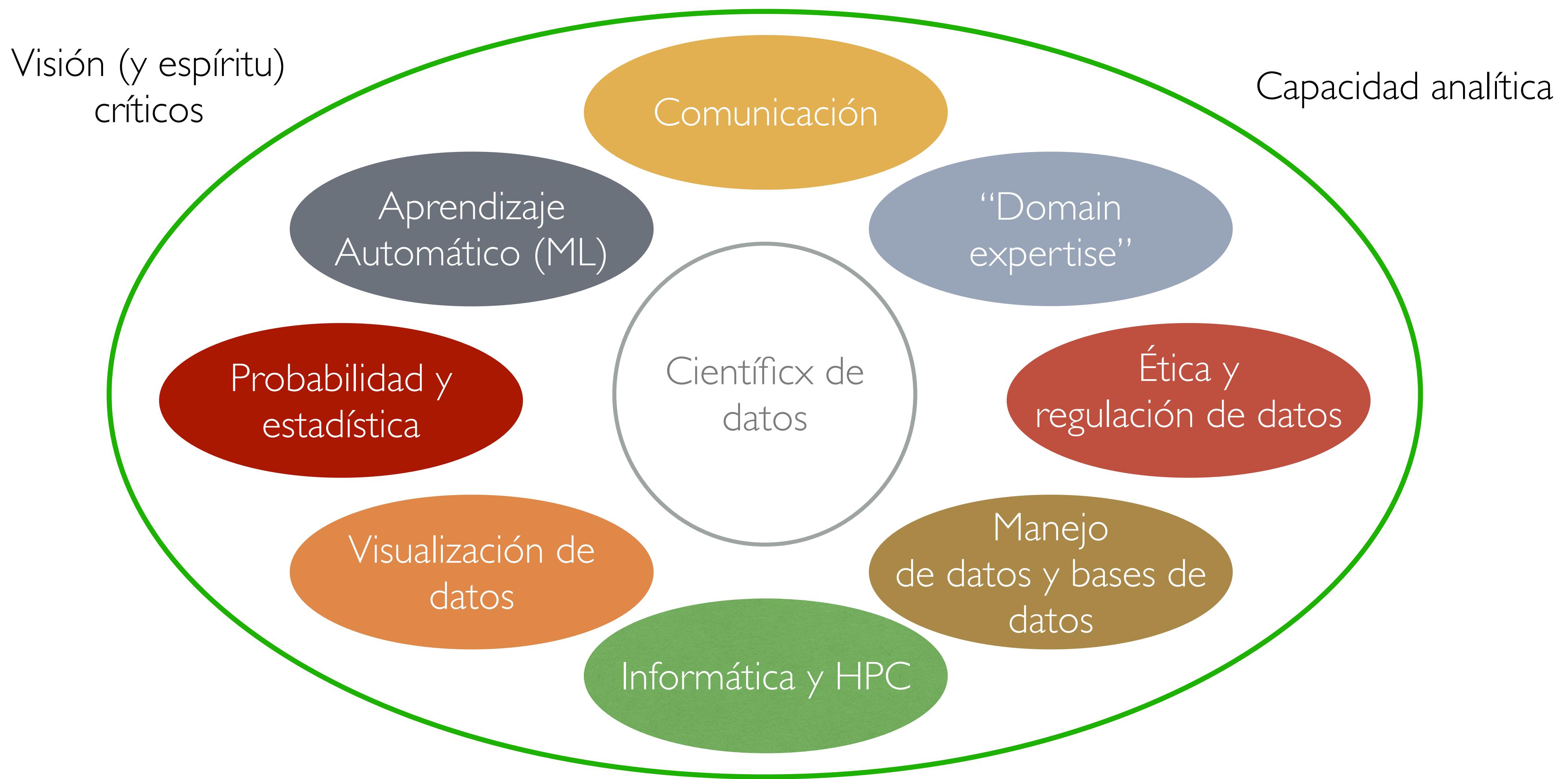
**Licenciatura en
Ciencia de Datos
ECYT_UNSAM**

Bienvenidxs

A ICD

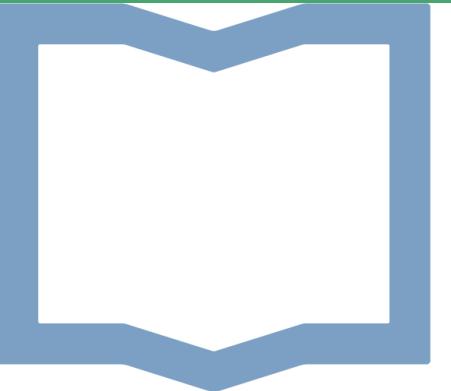
Campus Miguelete, 13 de marzo de 2023

Contenidos de ICD



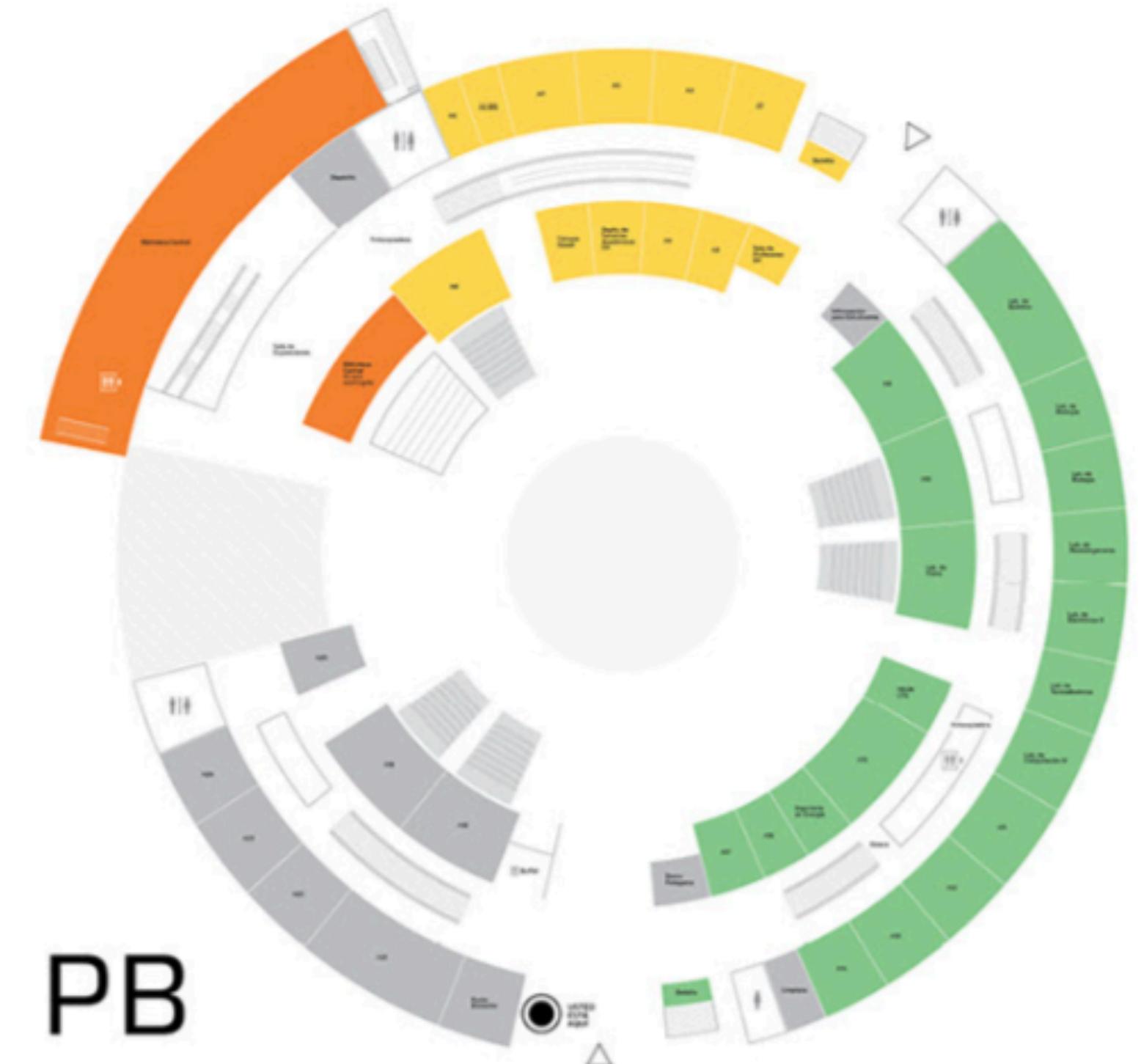
Asuntos prácticos

Clases presenciales



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

- **Horario:** lunes de 13h a 17h
- Estructura típica de una clase
 - Parte expositiva (~1 hora)
 - Trabajo práctico en laboratorios en grupo (**labos 1, 4**) (~2h30, con descansos)
 - Puesta en común al final (~30 minutos; **Aula ??**)



Asuntos prácticos

Clases virtuales



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

- **Asincrónico**
 - Cápsulas con contenido técnico en el Campus.
 - Util para tener a mano en caso de dudas.
- **Sincrónico:**
 - Horario: jueves a las 9h.
 - Conectarse con la cápsula estudiada.
 - Buen momento para consultas entre clases presenciales.

≡ [Campusvirtual ECyT - UNSAM](#) [Español - Internacional \(es\) ▾](#)

General

Semana 1

Semana 2 - Manipulación

Semana 3 - Descripción

Cápsulas

Bienven

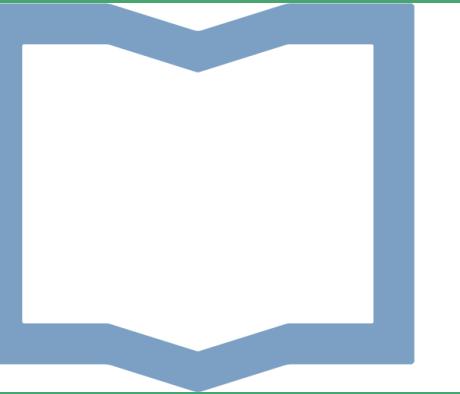
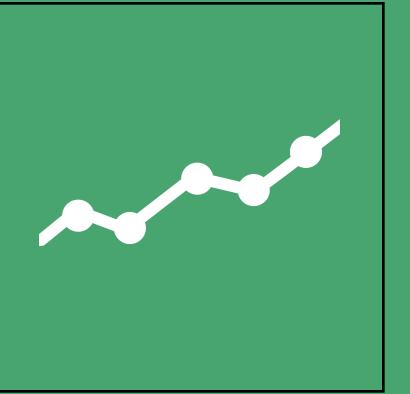
Introducción
(1er cuad)

En esta página encontrarán información general.
Las pestañas de la derecha iremos subiendo el

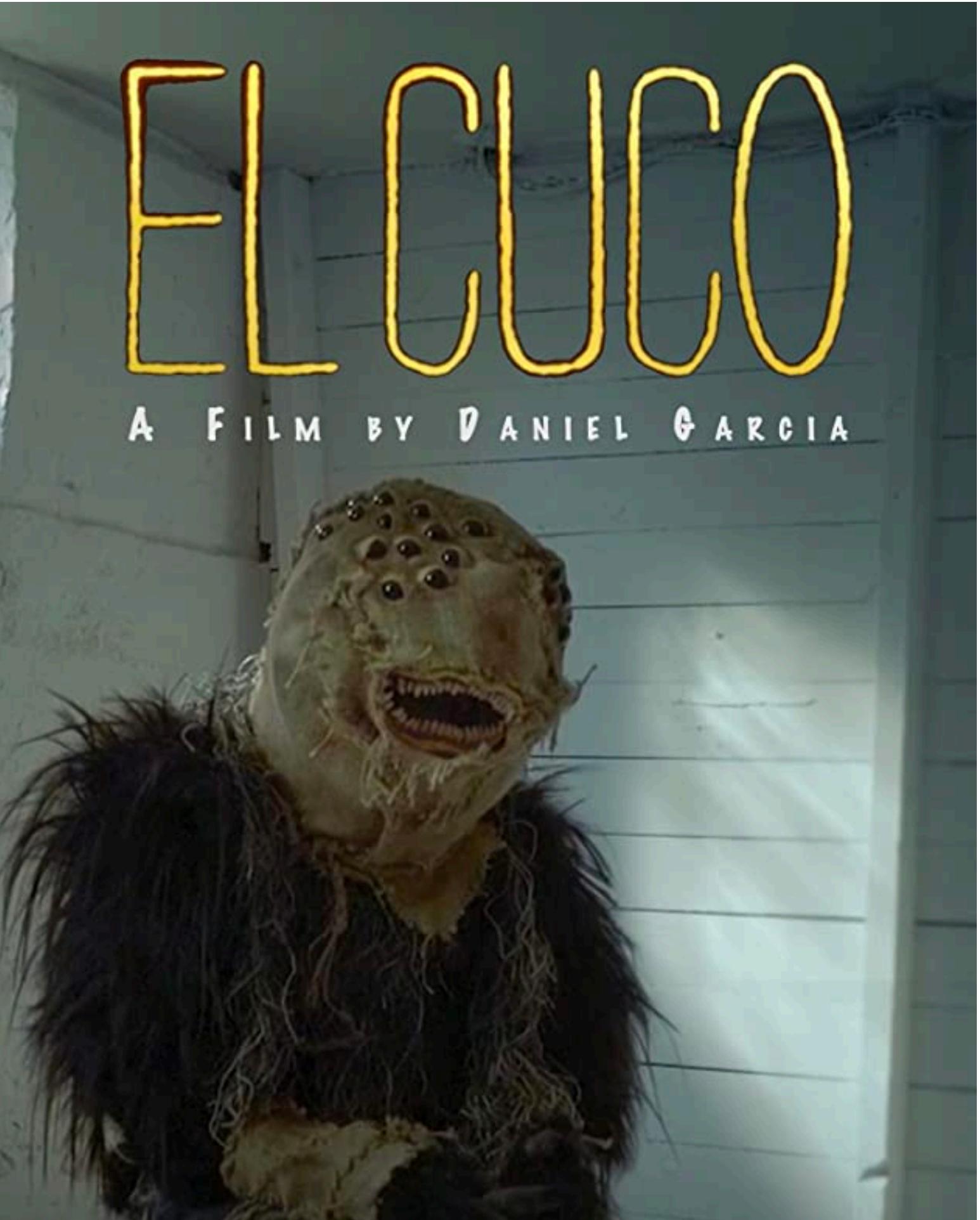
La comunicación se hará exclusivamente a tra

Asuntos prácticos

Evaluaciones

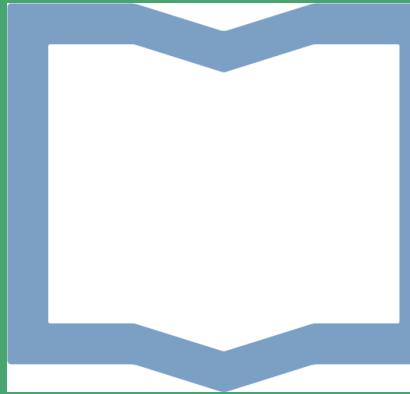


- **Entregas semanales.**
 - Entregas breves (típicamente un gráfico con una breve descripción)
 - Con nota, pero que no cuenta para el promedio.
 - No pueden adeudarse más de dos entregas.
- **Síntesis de mitad de camino.**
 - Una entrega más completa.
 - ¡Con nota!
 - **Fecha de entrega:** 30 de abril
- **Trabajo práctico final**
 - Se realiza durante las últimas seis semanas del cuatrimestre
 - Presentación final
- $\text{nota_cursada} = 0,4 * \text{nota_entrega} + 0,6 * \text{nota_tp}$
- **Criterio de aprobación**
 - 75% de presencialidad
 - Para regularizar: $\text{nota_cursada} \geq 4$
 - Para promocionar (ambas instancias con $\text{nota} \geq 6$ y promedio ≥ 7)
 - Instancia final (tipo entrevista), si $\text{nota_cursada} < 7$

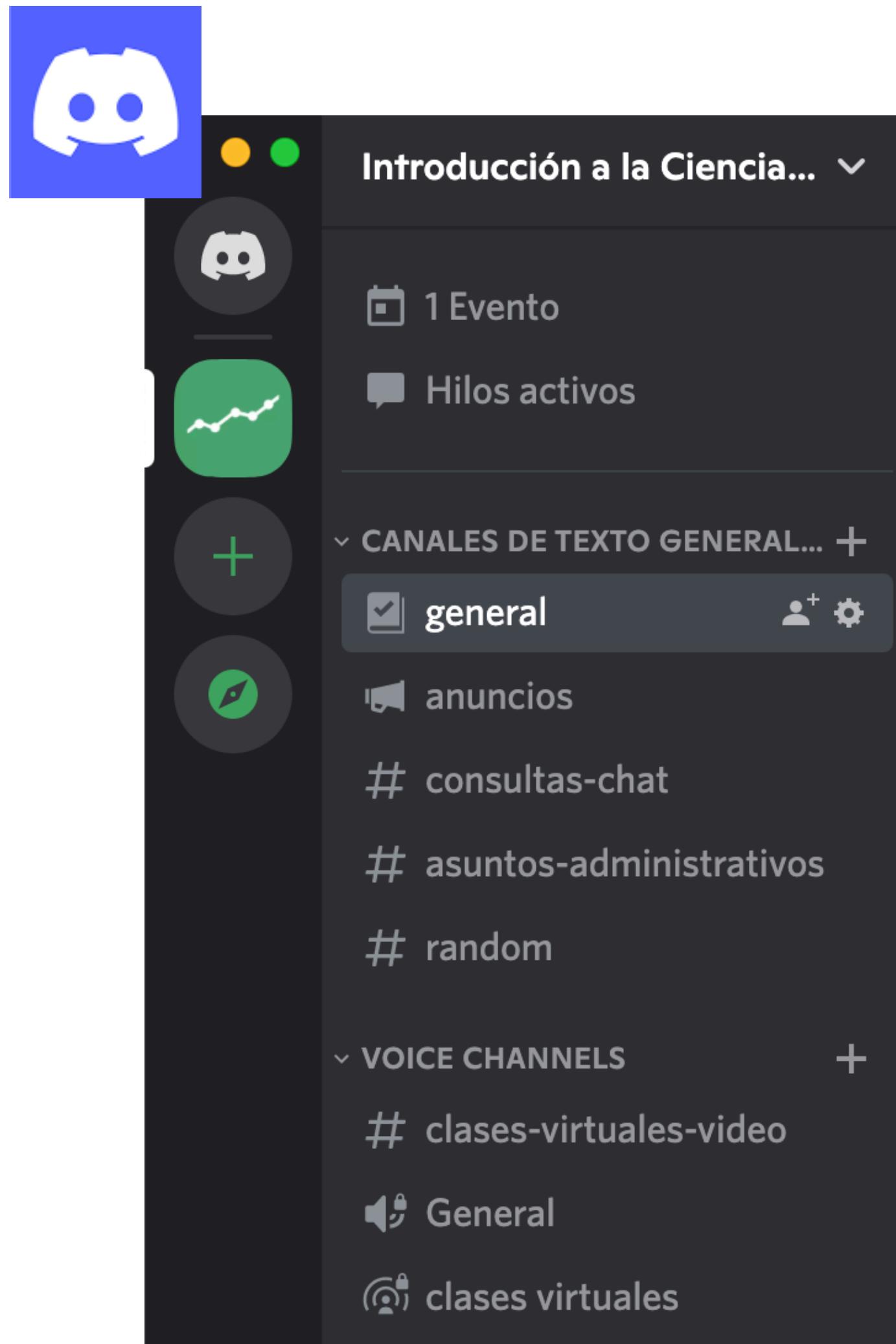


Asuntos prácticos

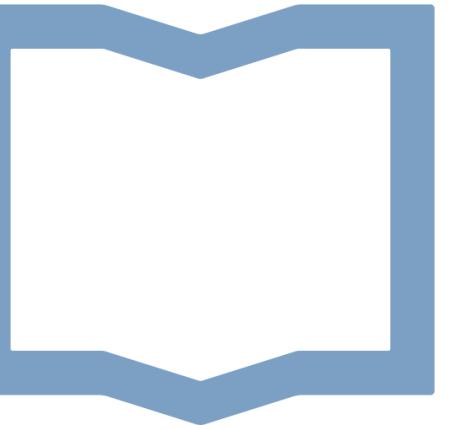
Comunicación



Licenciatura en
Ciencia de Datos
ECYT_UNSAM



- **Intercambio y discusiones**
 - Exclusivamente por Discord.
 - Usar el canal adecuado para cada cosa.
 - Siéntanse libres de armar canales.
- **Material:**
 - Mayormente disponible en el Campus.
 - El material se va actualizando semana a semana, así que es fundamental que se conecten frecuentemente.

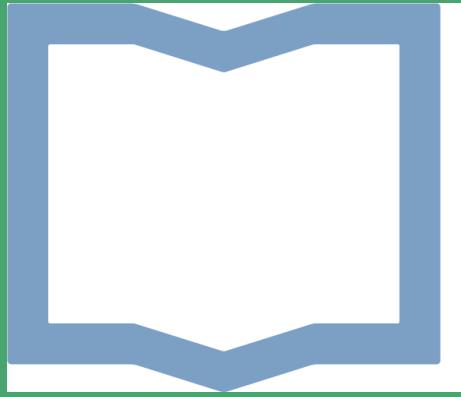


Licenciatura en
Ciencia de Datos
ECYT_UNSAM

Arrancaremos

2013,9,30,2009,2000,9,2238,2300,-22,VX,415,N844VA,JFK,LAX,303,2475,20,0,2013-10-01T00:00:00Z
2013,9,30,2009,2009,0,2126,2129,-3,UA,1049,N77258,EWR,BOS,38,200,20,9,2013-10-01T00:00:00Z
2013,9,30,2010,2012,-2,2143,2150,-7,EV,4106,N12921,EWR,GSO,64,445,20,12,2013-10-01T00:00:00Z
2013,9,30,2010,2016,-6,2119,2128,-9,EV,4312,N13969,EWR,DCA,35,199,20,16,2013-10-01T00:00:00Z
2013,9,30,2010,2015,-5,2100,2125,-25,B6,418,N249JB,JFK,BOS,35,187,20,15,2013-10-01T00:00:00Z
2013,9,30,2012,2015,-3,2117,2144,-27,B6,702,N339JB,JFK,BUF,51,301,20,15,2013-10-01T00:00:00Z
2013,9,30,2013,2000,13,2133,2120,13,US,2195,N747UW,LGA,DCA,41,214,20,0,2013-10-01T00:00:00Z
2013,9,30,2015,2030,-15,2221,2251,-30,EV,5209,N741EV,LGA,CHS,84,641,20,30,2013-10-01T00:00:00Z
2013,9,30,2015,2015,0,2244,2307,-23,UA,1545,N17730,EWR,IAH,174,1400,20,15,2013-10-01T00:00:00Z
2013,9,30,2016,2022,-6,2215,2246,-31,B6,135,N607JB,JFK,PHX,263,2153,20,22,2013-10-01T00:00:00Z
2013,9,30,2017,2022,-5,2132,2155,-23,B6,105,N298JB,JFK,ORD,112,740,20,22,2013-10-01T00:00:00Z
2013,9,30,2017,2017,0,2253,2320,-27,UA,771,N510UA,JFK,LAX,304,2475,20,17,2013-10-01T00:00:00Z
2013,9,30,2019,2023,-4,2225,2240,-15,EV,4085,N14188,EWR,OMA,149,1134,20,23,2013-10-01T00:00:00Z
2013,9,30,2021,2022,-1,2144,2200,-16,UA,1464,N76503,EWR,CLE,54,404,20,22,2013-10-01T00:00:00Z
2013,9,30,2022,2025,-3,2127,2140,-13,AA,2314,N3CUAA,JFK,BOS,37,187,20,25,2013-10-01T00:00:00Z
2013,9,30,2023,2025,-2,2315,2350,-35,B6,915,N712JB,JFK,SFO,327,2586,20,25,2013-10-01T00:00:00Z
2013,9,30,2026,2028,-2,2317,2340,-23,UA,1680,N31412,EWR,MIA,141,1085,20,28,2013-10-01T00:00:00Z
2013,9,30,2027,2034,-7,2129,2209,-40,9E,4127,N8790A,JFK,IAD,41,228,20,34,2013-10-01T00:00:00Z
2013,9,30,2028,1910,78,2255,2215,40,AA,21,N338AA,JFK,LAX,294,2475,19,10,2013-09-30T23:00:00Z
2013,9,30,2030,2045,-15,2141,2153,-12,B6,2680,N266JB,EWR,BOS,38,200,20,45,2013-10-01T00:00:00Z
2013,9,30,2030,2035,-5,2324,2340,-16,UA,374,N484UA,EWR,FLL,141,1065,20,35,2013-10-01T00:00:00Z
2013,9,30,2031,2040,-9,2228,2300,-32,9E,4033,N8924B,LGA,TYS,86,647,20,40,2013-10-01T00:00:00Z
2013,9,30,2032,1945,47,2115,2106,9,9E,3864,N602XJ,JFK,PHL,24,94,19,45,2013-09-30T23:00:00Z
2013,9,30,2032,2045,-13,2147,2225,-38,AA,371,N434AA,LGA,ORD,105,733,20,45,2013-10-01T00:00:00Z
2013,9,30,2033,2030,3,2143,2150,-7,WN,2520,N286WN,EWR,MDW,97,711,20,30,2013-10-01T00:00:00Z
2013,9,30,2034,2040,-6,2227,2248,-21,EV,4536,N25134,EWR,CVG,80,569,20,40,2013-10-01T00:00:00Z
2013,9,30,2034,2030,4,2204,2201,3,FL,354,N894AT,LGA,CAK,58,397,20,30,2013-10-01T00:00:00Z

Encabezados



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

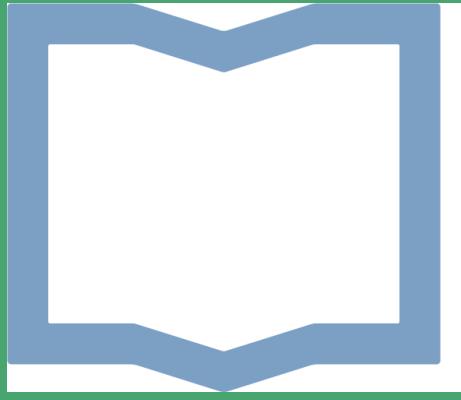
Encabezado
(Header)

Readme

Preguntar

	year	month	day	dep_time	sched_dep_time	dep_delay	arr_time	sched_arr_time	arr_delay	carrier	flight	tailnum	origin	dest	air_time	distance	hour	minute	time_hour
0	2013	1	1	517.0	515	2.0	830.0	819	11.0	UA	1545	N14228	EWR	IAH	227.0	1400	5	15	2013-01-01T10:00:00Z
1	2013	1	1	533.0	529	4.0	850.0	830	20.0	UA	1714	N24211	LGA	IAH	227.0	1416	5	29	2013-01-01T10:00:00Z
2	2013	1	1	542.0	540	2.0	923.0	850	33.0	AA	1141	N619AA	JFK	MIA	160.0	1089	5	40	2013-01-01T10:00:00Z
3	2013	1	1	544.0	545	-1.0	1004.0	1022	-18.0	B6	725	N804JB	JFK	BQN	183.0	1576	5	45	2013-01-01T10:00:00Z
4	2013	1	1	554.0	600	-6.0	812.0	837	-25.0	DL	461	N668DN	LGA	ATL	116.0	762	6	0	2013-01-01T11:00:00Z
5	2013	1	1	554.0	558	-4.0	740.0	728	12.0	UA	1696	N39463	EWR	ORD	150.0	719	5	58	2013-01-01T10:00:00Z
6	2013	1	1	555.0	600	-5.0	913.0	854	19.0	B6	507	N516JB	EWR	FLL	158.0	1065	6	0	2013-01-01T11:00:00Z
7	2013	1	1	557.0	600	-3.0	709.0	723	-14.0	EV	5708	N829AS	LGA	IAD	53.0	229	6	0	2013-01-01T11:00:00Z
8	2013	1	1	557.0	600	-3.0	838.0	846	-8.0	B6	79	N593JB	JFK	MCO	140.0	944	6	0	2013-01-01T11:00:00Z
9	2013	1	1	558.0	600	-2.0	753.0	745	8.0	AA	301	N3ALAA	LGA	ORD	138.0	733	6	0	2013-01-01T11:00:00Z
10	2013	1	1	558.0	600	-2.0	849.0	851	-2.0	B6	49	N793JB	JFK	PBI	149.0	1028	6	0	2013-01-01T11:00:00Z
11	2013	1	1	558.0	600	-2.0	853.0	856	-3.0	B6	71	N657JB	JFK	TPA	158.0	1005	6	0	2013-01-01T11:00:00Z
12	2013	1	1	558.0	600	-2.0	924.0	917	7.0	UA	194	N29129	JFK	LAX	345.0	2475	6	0	2013-01-01T11:00:00Z
13	2013	1	1	558.0	600	-2.0	923.0	937	-14.0	UA	1124	N53441	EWR	SFO	361.0	2565	6	0	2013-01-01T11:00:00Z
14	2013	1	1	559.0	600	-1.0	941.0	910	31.0	AA	707	N3DUAA	LGA	DFW	257.0	1389	6	0	2013-01-01T11:00:00Z
15	2013	1	1	559.0	559	0.0	702.0	706	-4.0	B6	1806	N708JB	JFK	BOS	44.0	187	5	59	2013-01-01T10:00:00Z
16	2013	1	1	559.0	600	-1.0	854.0	902	-8.0	UA	1187	N76515	EWR	LAS	337.0	2227	6	0	2013-01-01T11:00:00Z
17	2013	1	1	600.0	600	0.0	851.0	858	-7.0	B6	371	N595JB	LGA	FLL	152.0	1076	6	0	2013-01-01T11:00:00Z
18	2013	1	1	600.0	600	0.0	837.0	825	12.0	MQ	4650	N542MQ	LGA	ATL	134.0	762	6	0	2013-01-01T11:00:00Z
19	2013	1	1	601.0	600	1.0	844.0	850	-6.0	B6	343	N644JB	EWR	PBI	147.0	1023	6	0	2013-01-01T11:00:00Z
20	2013	1	1	602.0	610	-8.0	812.0	820	-8.0	DL	1919	N971DL	LGA	MSP	170.0	1020	6	10	2013-01-01T11:00:00Z
21	2013	1	1	602.0	605	-3.0	821.0	805	16.0	MQ	4401	N730MQ	LGA	DTW	105.0	502	6	5	2013-01-01T11:00:00Z
22	2013	1	1	606.0	610	-4.0	858.0	910	-12.0	AA	1895	N633AA	EWR	MIA	152.0	1085	6	10	2013-01-01T11:00:00Z
23	2013	1	1	606.0	610	-4.0	837.0	845	-8.0	DL	1743	N3739P	JFK	ATL	128.0	760	6	10	2013-01-01T11:00:00Z
24	2013	1	1	607.0	607	0.0	858.0	915	-17.0	UA	1077	N53442	EWR	MIA	157.0	1085	6	7	2013-01-01T11:00:00Z
25	2013	1	1	608.0	600	8.0	807.0	735	32.0	MQ	3768	N9EAMQ	EWR	ORD	139.0	719	6	0	2013-01-01T11:00:00Z
26	2013	1	1	611.0	600	11.0	945.0	931	14.0	UA	303	N532UA	JFK	SFO	366.0	2586	6	0	2013-01-01T11:00:00Z
27	2013	1	1	613.0	610	3.0	925.0	921	4.0	B6	135	N635JB	JFK	RSW	175.0	1074	6	10	2013-01-01T11:00:00Z
28	2013	1	1	615.0	615	0.0	1039.0	1100	-21.0	B6	709	N794JB	JFK	SJU	182.0	1598	6	15	2013-01-01T11:00:00Z
29	2013	1	1	615.0	615	0.0	833.0	842	-9.0	DL	575	N326NB	EWR	ATL	120.0	746	6	15	2013-01-01T11:00:00Z

Algunas definiciones



Unidades. Son los objetos principales de estudio. Por ejemplo, los vuelos en la tabla nycflights13. ¿Otras ideas? También llamadas *instancias, ejemplos, entidades, casos, sujetos, registros*, etc.

Variables. Son propiedades, cualidades o cantidades de las unidades que se pueden medir.

Valores. Son los estados de las variables una vez medida.

Observaciones. Es el conjunto de valores medidos en condiciones similares. En general, hechas en el mismo tiempo y para el mismo objeto. P.ej.: el alto y ancho del tronco de un árbol.

Un formato ordenado (tidy)



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

Observación /
Data point

Variables

	year	month	day	dep_time	sched_dep_time	dep_delay	arr_time	sched_arr_time	arr_delay	carrier	flight	tailnum	origin	dest	air_time	distance	hour	minute	time_hour
0	2013	1	1	517.0	515	2.0	830.0	819	11.0	UA	1545	N14228	EWR	IAH	227.0	1400	5	15	2013-01-01T10:00:00Z
1	2013	1	1	533.0	529	4.0	850.0	830	20.0	UA	1714	N24211	LGA	IAH	227.0	1416	5	29	2013-01-01T10:00:00Z
2	2013	1	1	542.0	540	2.0	923.0	850	33.0	AA	1141	N619AA	JFK	MIA	160.0	1089	5	40	2013-01-01T10:00:00Z
3	2013	1	1	544.0	545	-1.0	1004.0	1022	-18.0	B6	725	N804JB	JFK	BQN	183.0	1576	5	45	2013-01-01T10:00:00Z
4	2013	1	1	554.0	600	-6.0	812.0	837	-25.0	DL	461	N668DN	LGA	ATL	116.0	762	6	0	2013-01-01T11:00:00Z
5	2013	1	1	554.0	558	-4.0	740.0	728	12.0	UA	1696	N39463	EWR	ORD	150.0	719	5	58	2013-01-01T10:00:00Z
6	2013	1	1	555.0	600	-5.0	913.0	854	19.0	B6	507	N516JB	EWR	FLL	158.0	1065	6	0	2013-01-01T11:00:00Z
7	2013	1	1	557.0	600	-3.0	709.0	723	-14.0	EV	5708	N829AS	LGA	IAD	53.0	229	6	0	2013-01-01T11:00:00Z
8	2013	1	1	557.0	600	-3.0	838.0	846	-8.0	B6	79	N593JB	JFK	MCO	140.0	944	6	0	2013-01-01T11:00:00Z
9	2013	1	1	558.0	600	-2.0	753.0	745	8.0	AA	301	N3ALAA	LGA	ORD	138.0	733	6	0	2013-01-01T11:00:00Z
10	2013	1	1	558.0	600	-2.0	849.0	851	-2.0	B6	49	N793JB	JFK	PBI	149.0	1028	6	0	2013-01-01T11:00:00Z
11	2013	1	1	558.0	600	-2.0	853.0	856	-3.0	B6	71	N657JB	JFK	TPA	158.0	1005	6	0	2013-01-01T11:00:00Z
12	2013	1	1	558.0	600	-2.0	924.0	917	7.0	UA	194	N29129	JFK	LAX	345.0	2475	6	0	2013-01-01T11:00:00Z
13	2013	1	1	558.0	600	-2.0	923.0	937	-14.0	UA	1124	N53441	EWR	SFO	361.0	2565	6	0	2013-01-01T11:00:00Z
14	2013	1	1	559.0	600	-1.0	941.0	910	31.0	AA	707	LGA	DFW	257.0	1389	6	0	2013-01-01T11:00:00Z	
15	2013	1	1	559.0	559	0.0	702.0	706	-4.0	B6	1806	N708JB	JFK	BOS	44.0	187	5	59	2013-01-01T10:00:00Z
16	2013	1	1	559.0	600	-1.0	854.0	902	-8.0	UA	1187	N76515	EWR	LAS	337.0	2227	6	0	2013-01-01T11:00:00Z
17	2013	1	1	600.0	600	0.0	851.0	858	-7.0	B6	371	N595JB	LGA	FLL	152.0	1076	6	0	2013-01-01T11:00:00Z
18	2013	1	1	600.0	600	0.0	837.0	825	12.0	MQ	4650	N542MQ	LGA	ATL	134.0	762	6	0	2013-01-01T11:00:00Z
19	2013	1	1	601.0	600	1.0	844.0	850	-6.0	B6	343	N644JB	EWR	PBI	147.0	1023	6	0	2013-01-01T11:00:00Z
20	2013	1	1	602.0	610	-8.0	812.0	820	-8.0	DL	1919	N971DL	LGA	MSP	170.0	1020	6	10	2013-01-01T11:00:00Z
21	2013	1	1	602.0	605	-3.0	821.0	805	16.0	MQ	4401	N730MQ	LGA	DTW	105.0	502	6	5	2013-01-01T11:00:00Z
22	2013	1	1	606.0	610	-4.0	858.0	910	-12.0	AA	1895	N633AA	EWR	MIA	152.0	1085	6	10	2013-01-01T11:00:00Z
23	2013	1	1	606.0	610	-4.0	837.0	845	-8.0	DL	1743	N3739P	JFK	ATL	128.0	760	6	10	2013-01-01T11:00:00Z
24	2013	1	1	607.0	607	0.0	858.0	915	-17.0	UA	1077	N53442	EWR	MIA	157.0	1085	6	7	2013-01-01T11:00:00Z
25	2013	1	1	608.0	600	8.0	807.0	735	32.0	MQ	3768	N9EAMQ	EWR	ORD	139.0	719	6	0	2013-01-01T11:00:00Z
26	2013	1	1	611.0	600	11.0	945.0	931	14.0	UA	303	N532UA	JFK	SFO	366.0	2586	6	0	2013-01-01T11:00:00Z
27	2013	1	1	613.0	610	3.0	925.0	921	4.0	B6	135	N635JB	JFK	RSW	175.0	1074	6	10	2013-01-01T11:00:00Z
28	2013	1	1	615.0	615	0.0	1039.0	1100	-21.0	B6	709	N794JB	JFK	SJU	182.0	1598	6	15	2013-01-01T11:00:00Z
29	2013	1	1	615.0	615	0.0	833.0	842	-9.0	DL	575	N326NB	EWR	ATL	120.0	746	6	15	2013-01-01T11:00:00Z

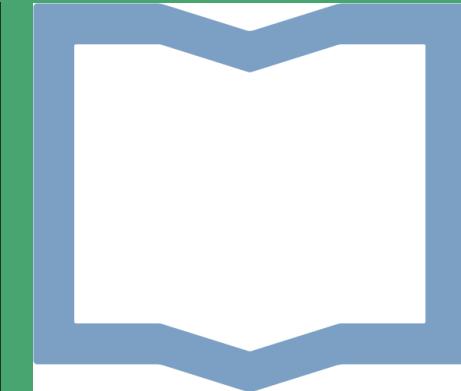
Valor

Archivos en la computadora

El formato CSV

CSV, su nuevo mejor amigo

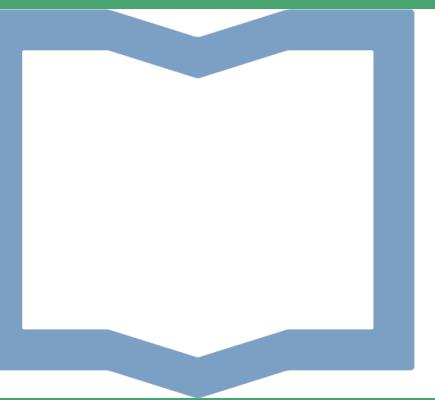
- comma-separated-values
- Son archivos de texto, donde las columnas están separadas por una coma.
- Formato estándar para transmitir datos ordenados.



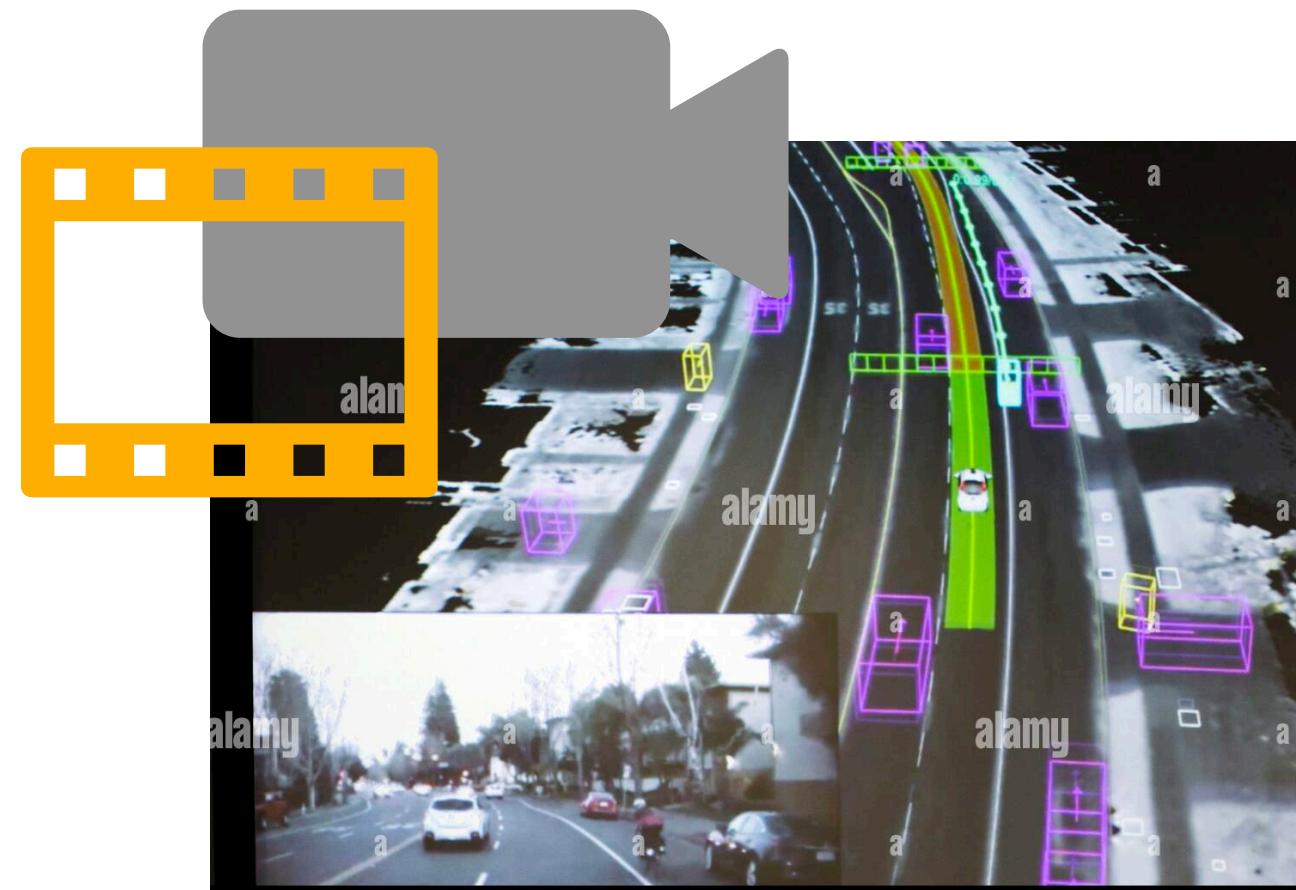
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

```
1 long,lat,nro_registro,comuna,calle_nombre,calle_altura,calle_chapa,  
    nombre_cientifico,ancho_acera,estado_plantera,ubicacion_plantera,  
    nivel_plantera,diametro_altura_pecho,altura_arbol  
2 -58.52246854385779,-34.635242540536105,248212,9,Alvarez Jonte Av.,6500,6509,  
    Lagerstroemia,4.4,Ocupada,Regular,Bajo nivel,4,2  
3 -58.522629550526105,-34.6354294300277,248235,9,Alvarez Jonte Av.,6500,6563,  
    Melia,4.4,Ocupada,Regular,A nivel,70,10  
4 -58.522684484956706,-34.635493195266,248260,9,Alvarez Jonte Av.,6500,6567,Melia  
    4.4,Ocupada,Regular,A nivel,73,10  
5 -58.5228272952872,-34.635658962044204,248278,9,Alvarez Jonte Av.,6500,6591,  
    Melia,4.4,Ocupada,Regular,A nivel,57,11  
6 -58.5221122386389,-34.634885150978604,248675,9,Alvarez Jonte Av.,6400,6481,  
    Melia,4,Ocupada,Regular,A nivel,58,9  
7 -58.522164769565705,-34.6349430221208,248701,9,Alvarez Jonte Av.,6400,6485,  
    Melia,4,Ocupada,Regular,A nivel,57,9  
8 -58.5218837652316,-34.634633450096196,248619,9,Alvarez Jonte Av.,6400,6457,  
    Lagerstroemia,4,Ocupada,Regular,Elevada,30,8  
9 -58.521941356652896,-34.634696896668004,248632,9,Alvarez Jonte Av.,6400,6465,  
    Melia,4,Ocupada,Regular,A nivel,77,8  
10 -58.522055938054294,-34.634823126844694,248655,9,Alvarez Jonte Av.,6400,6473,  
    Ficus,4,Ocupada,Regular,A nivel,35,6  
11 -58.51345241543199,-34.6258097317006,290862,10,Alvarez Jonte Av.,5400.0,5491,  
    Tilia,4.2,Ocupada,Regular,A nivel,25,6
```

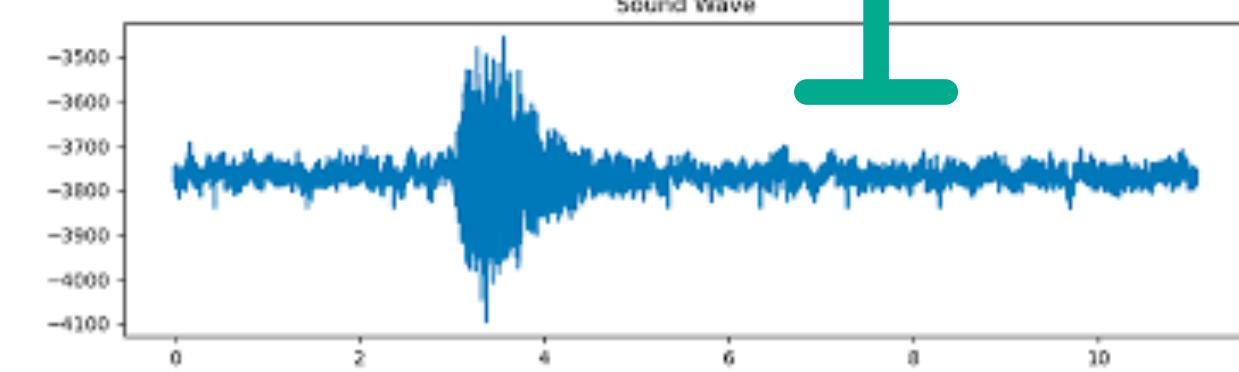
Datos no estructurados



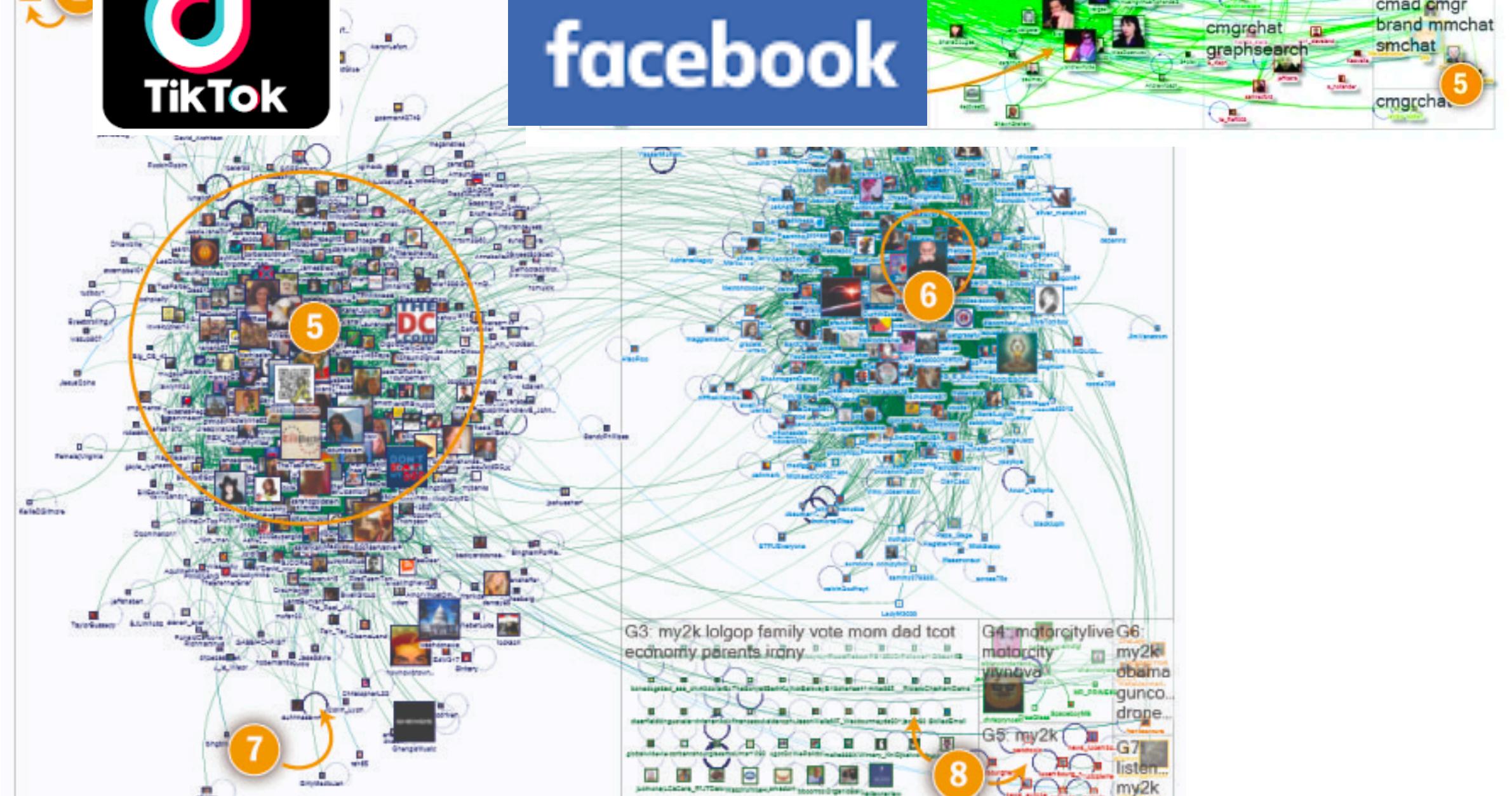
Licenciatura en
Ciencia de Datos
ECYT_UNSAM



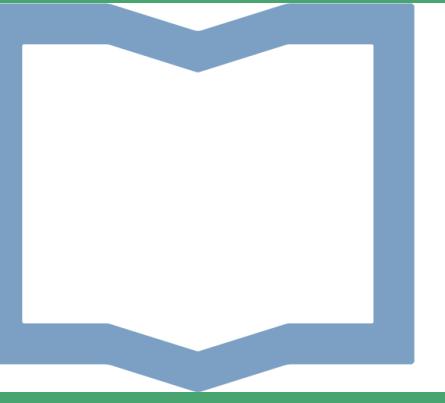
Exhaust
data



facebook



Tipos de variables



Las **variables** pueden ser de varios **tipos**

Números enteros
(int)

Números
de punto flotante
(float)

Tiempos
(date / datetime / time)

Cadenas de texto
(string)

Todas ellas
pueden ser

Factores /
Categóricas

sched_dep_time	dep_delay	arr_time	sched_arr_time	arr_delay	carrier	flight	tailnum	origin	dest	air_time	distance	hour	minute	time_hour
1030	8.0	1347.0	1340	7.0	AA	19	N328AA	JFK	LAX	344.0	2475	10	30	2013-01-02T15:00:00Z
700	-7.0	855.0	900	-5.0	9E	3769	N820AY	EWR	CVG	104.0	569	7	0	2013-09-16T11:00:00Z
1900	-10.0	2016.0	2014	2.0	US	2160	N953UW	LGA	BOS	34.0	184	19	0	2013-11-05T00:00:00Z
1810	-5.0	1922.0	1945	-23.0	WN	1368	N956WN	LGA	BNA	109.0	764	18	10	2013-09-23T22:00:00Z
600	-9.0	701.0	711	-10.0	EV	4533	N25134	EWR	BUF	48.0	282	6	0	2013-10-09T10:00:00Z
1745	-5.0	2047.0	2110	-23.0	AA	1611	N3FAAA	LGA	MIA	152.0	1096	17	45	2013-08-18T21:00:00Z
1100	171.0	1515.0	1210	185.0	MQ	3230	N532MQ	JFK	DCA	56.0	213	11	0	2013-07-22T15:00:00Z
1035	3.0	1148.0	1159	-11.0	EV	4322	N13123	EWR	MKE	111.0	725	10	35	2013-03-31T14:00:00Z
630	115.0	936.0	749	107.0	EV	4533	N12135	EWR	BUF	47.0	282	6	30	2013-03-28T10:00:00Z
1635	-7.0	1742.0	1810	-28.0	MQ	3695	N527MQ	EWR	ORD	107.0	719	16	35	2013-03-28T20:00:00Z

Armado del dataset

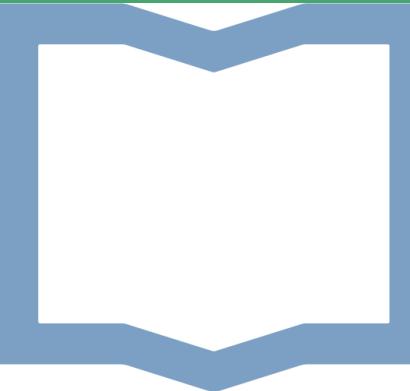
Algo un poco más filosófico



- A diferencia de otras ciencias (física, geología, química, ...), el objeto de estudio no viene dado por la Naturaleza. Es un producto humano.
- Como tal, los dataset son resultados de procesos y decisiones. Hay algo de abstracción, para arrancar, como mínimo.
- En esas decisiones, se eligen los atributos, etc. (p.ej. para los cumpleaños elegimos poner el día, mes y año, pero no la hora).
- En general, puede pasar que uno no tiene acceso a las personas u organizaciones que lo armaron, pero es importante tener conciencia de que se trata de un proceso *humano*.

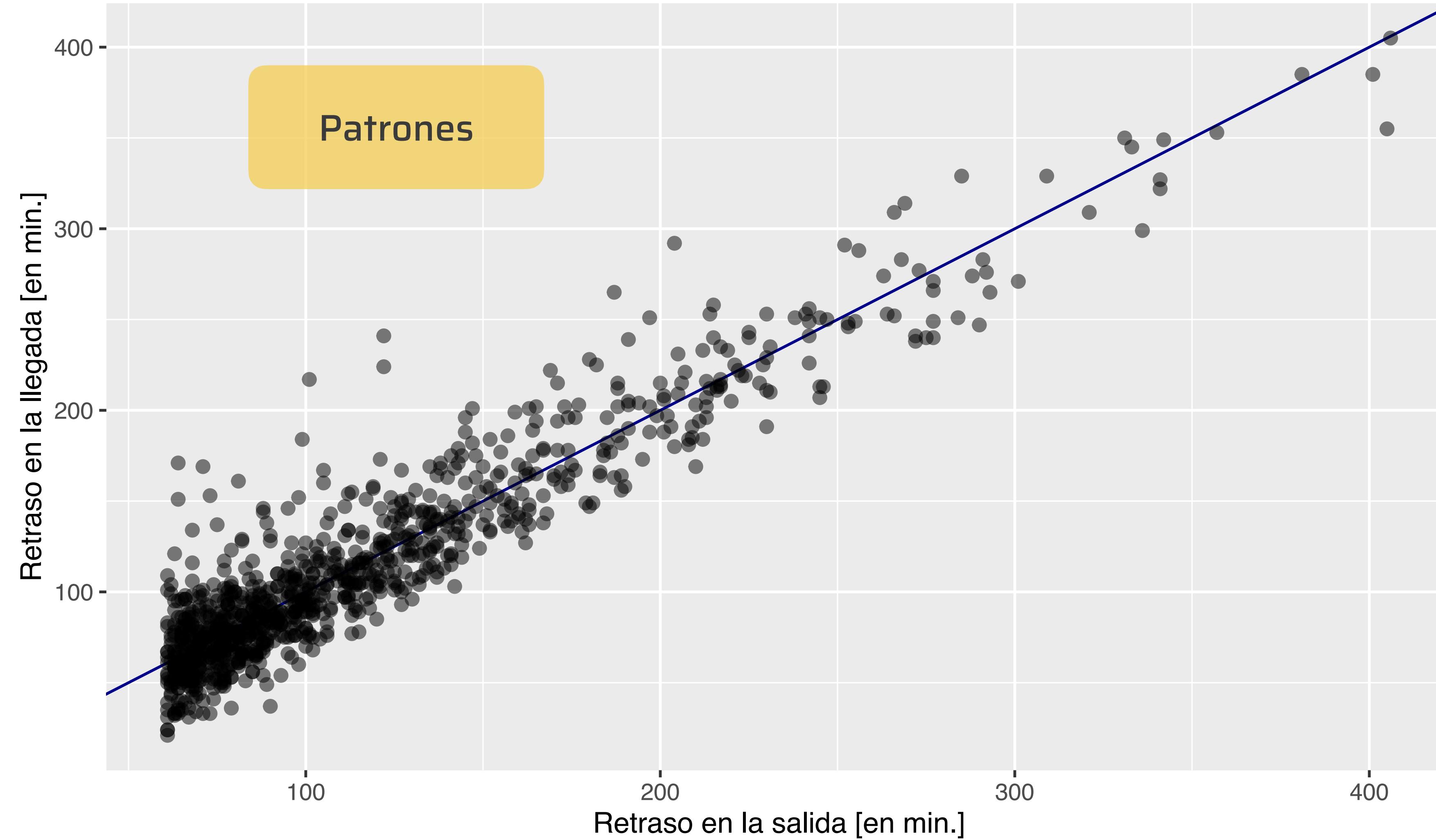
Visualización de Datos

Data Viz



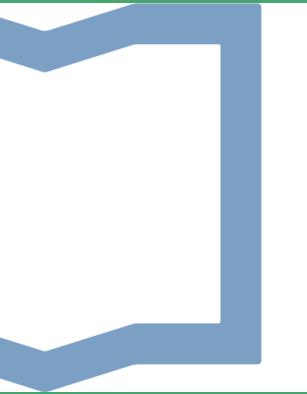
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

Vuelos desde NYC los días 19 del año 2013 retrasados más de 1 hora



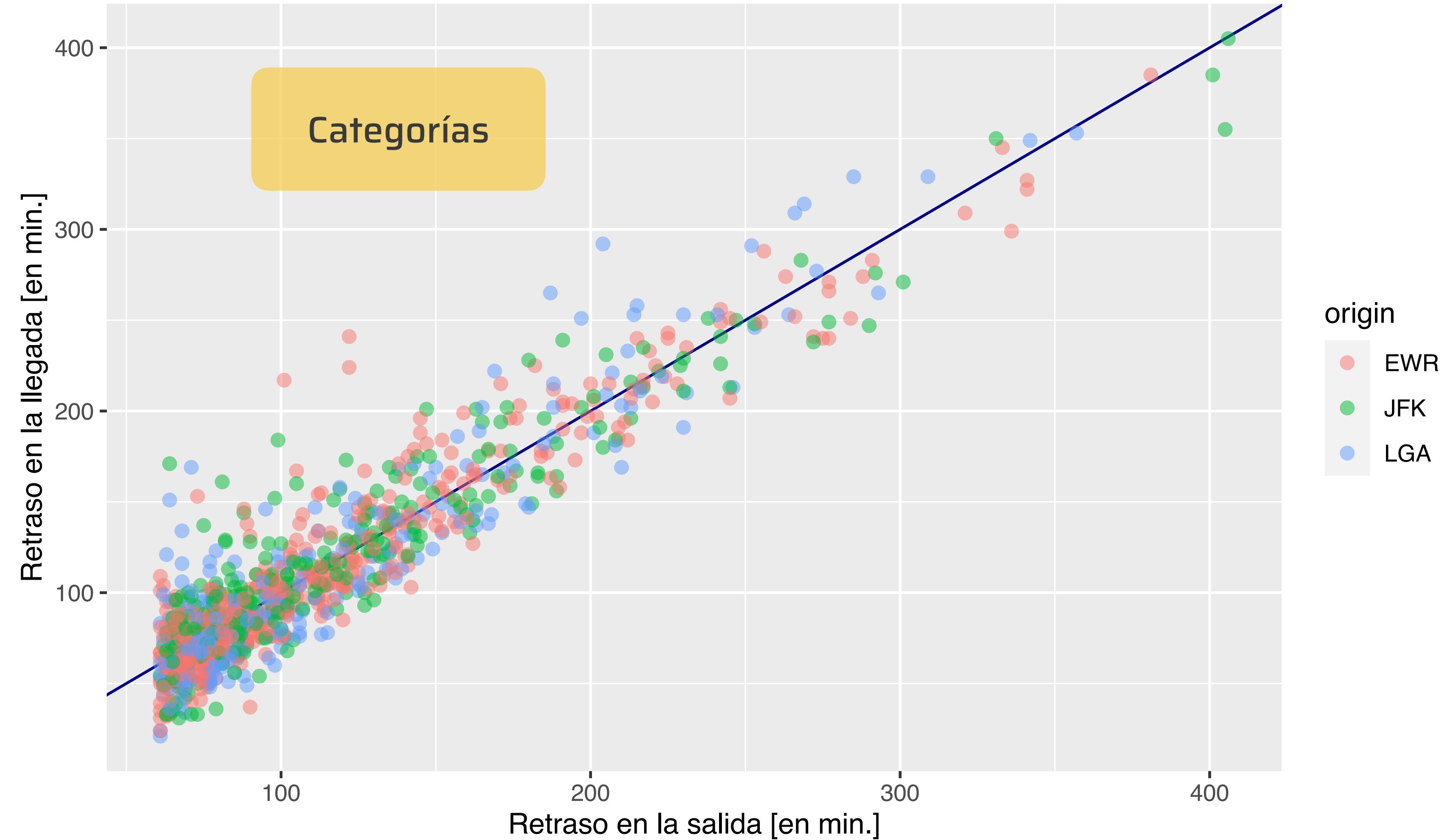
Visualización de Datos

Data Viz



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

Vuelos desde NYC los días 19 del año 2013 retrasados más de 1 hora



Bibliografía

Fundamental para el principio



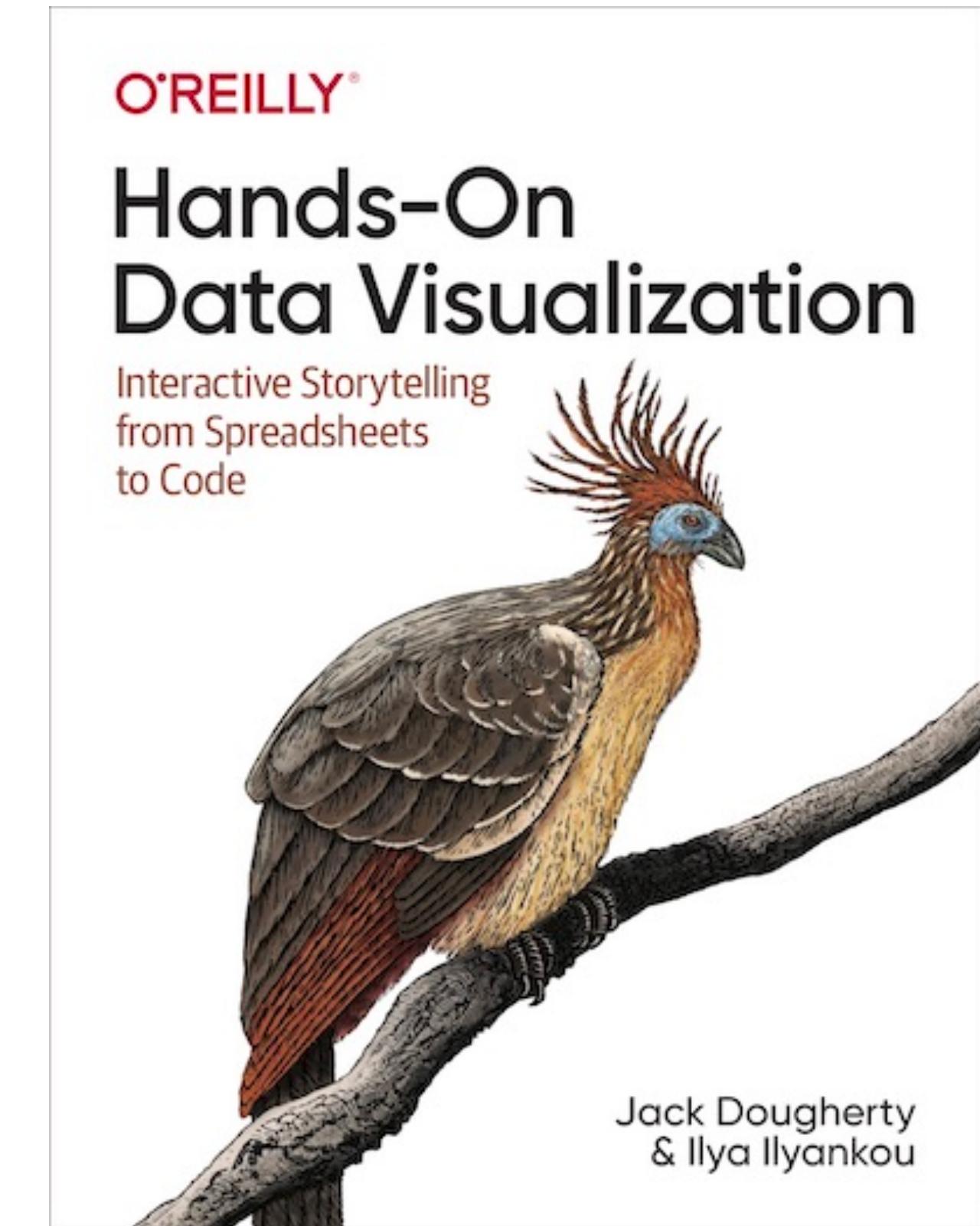
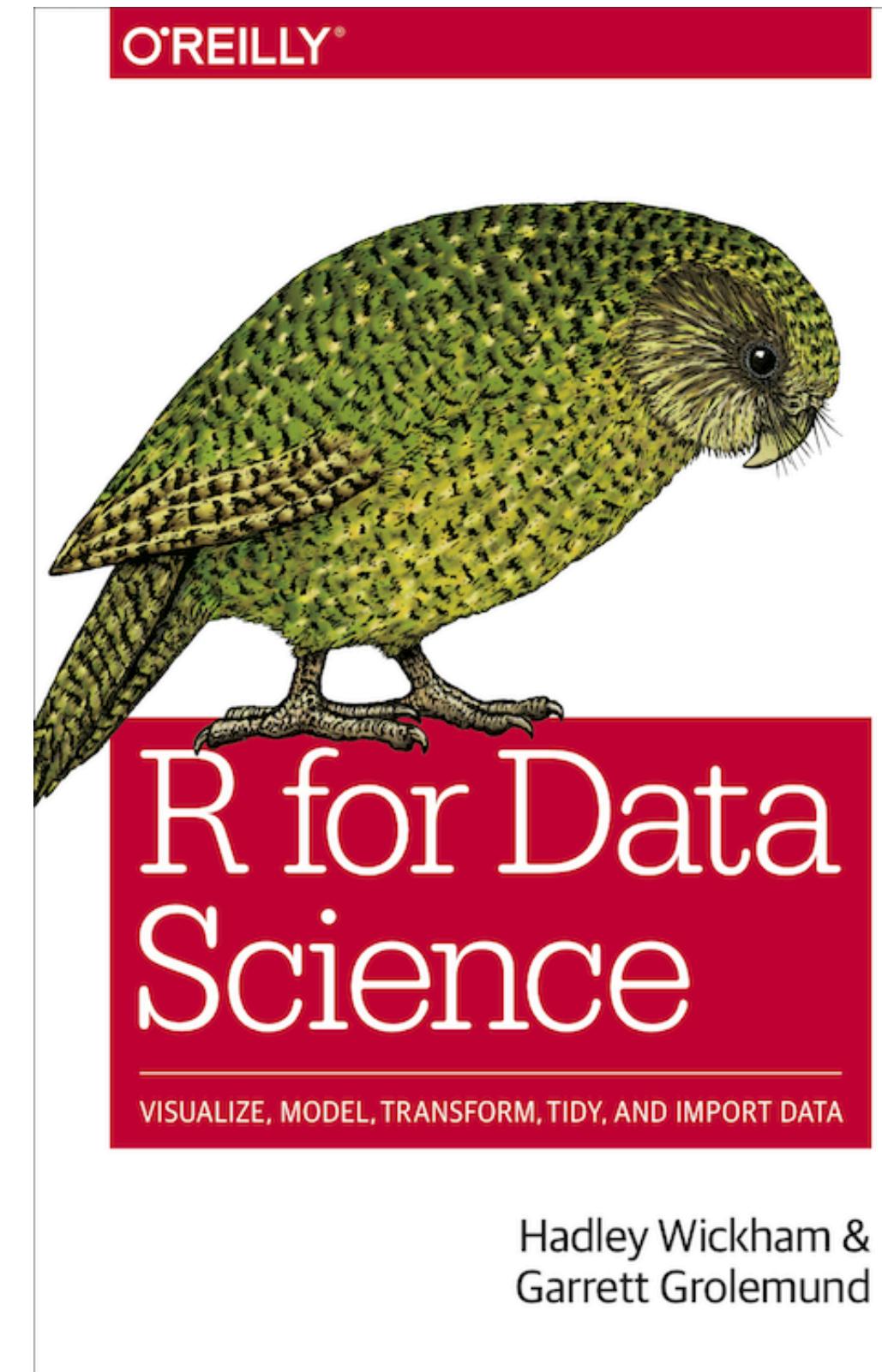
Licenciatura en
Ciencia de Datos
ECYT_UNSAM

En castellano

<https://es.r4ds.hadley.nz/>

En inglés

<https://es.r4ds.hadley.nz/>



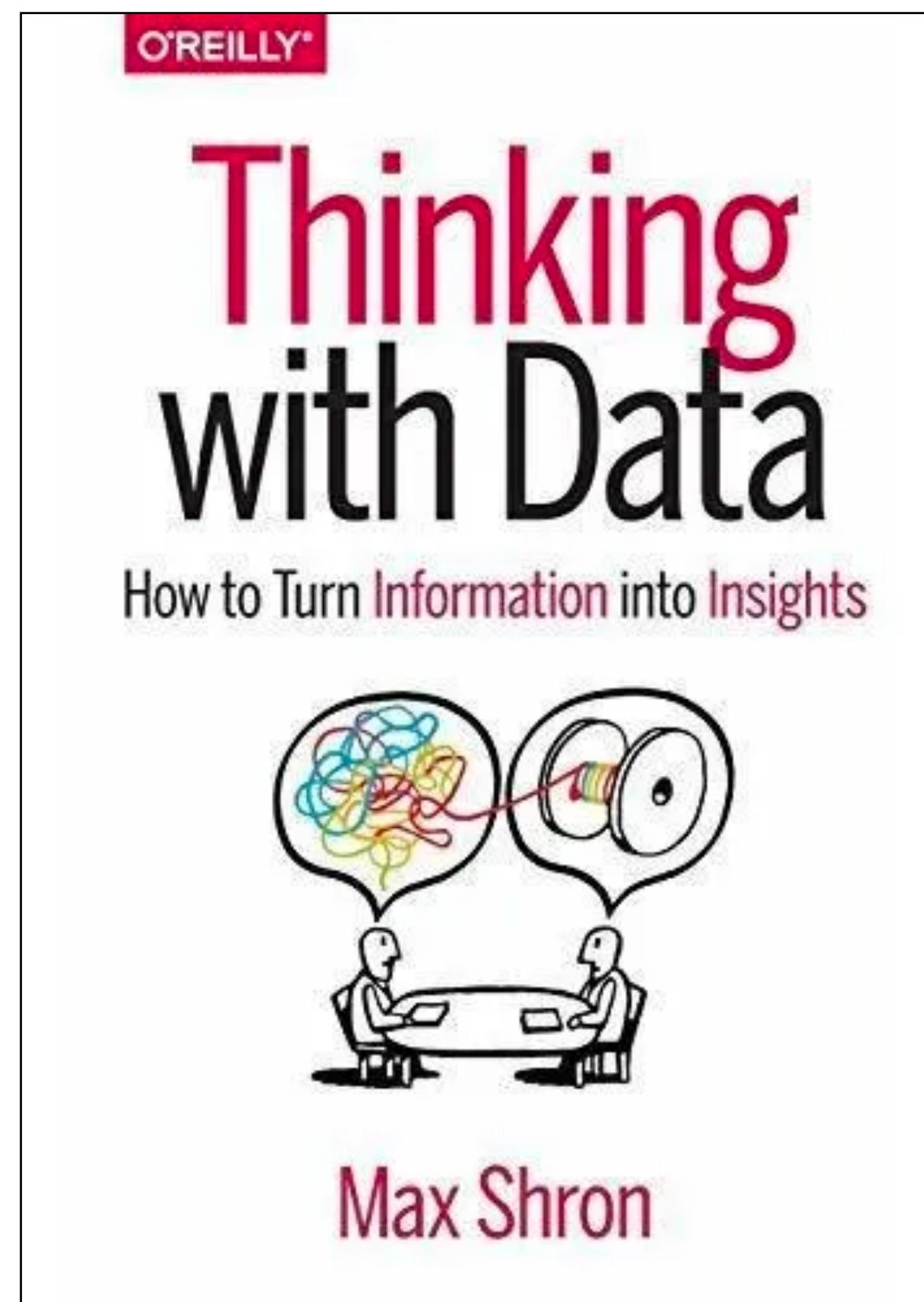
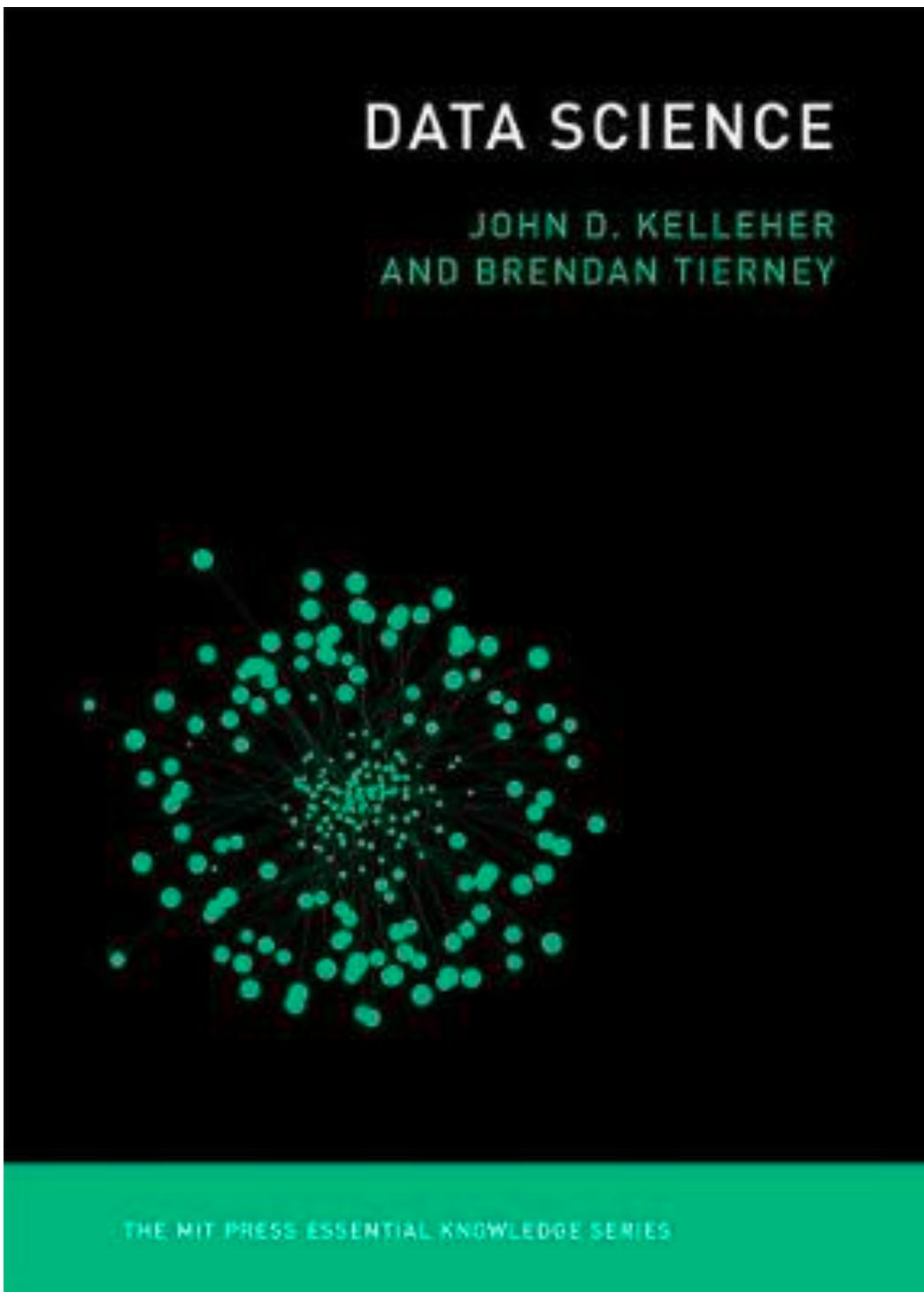
English only (sorry)

<https://handsondataviz.org/>

Bibliografía Complementaria



Licenciatura en
Ciencia de Datos
ECYT_UNSAM



En los laboratorios

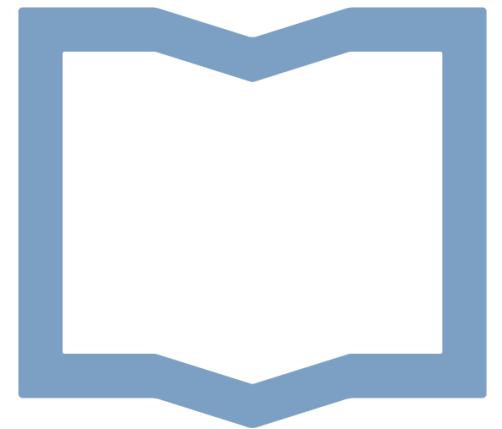
Algunas pautas a seguir



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

Laboratorios de computación 1 y 4

- Grupos de a dos (hoy a elección)
- Si trajiste tu laptop, mejor Labo 1 (wifi y espacio).
- Nosotros vamos a ir pasando para ver cómo vienen y para responder consultas.
- No duden en acercarse si se traban; **no esperen a que pasemos**.
- Libertad de ir y venir durante este tiempo.
- Puesta en común en Aula a definir a las 16h45



Licenciatura en
Ciencia de Datos
ECYT_UNSAM

¡A los laboratorios!

