

# ***PRODUCCIÓN DE SOJA EN LA ZONA CENTRO DE ARGENTINA***

**CURSO DE DATA ANALYTICS – CODERHOUSE**

**27/11/2022**



## **Integrantes:**

- ✓ Azar María Laura
- ✓ Ferreyra María Gimena
- ✓ López Mariano
- ✓ Pardo Margot

## CONTENIDO

INTRODUCCIÓN .....	3
OBJETIVO DEL PROYECTO .....	3
ALCANCE.....	3
USUARIO FINAL Y NIVEL DE APLICACIÓN .....	3
TABLA DE VERSIONADO .....	4
HERRAMIENTAS UTILIZADAS.....	4
DATASETS .....	4
ESQUEMA ENTIDAD-RELACIÓN.....	5
LISTADO DE TABLAS .....	7
TIPOS DE DATOS .....	8
• TABLA PROVINCIA .....	8
• TABLA DEPARTAMENTO.....	8
• TABLA RINDE .....	9
• TABLA CAMPAÑA .....	10
• TABLA PRECIOS .....	10
• TABLA PRECIOS MAIZ.....	10
• TABLA PRECIOS TRIGO .....	11
• TABLA PI USD SOJA .....	11
• TABLA COORDENADAS.....	12
TRANSFORMACIONES REALIZADAS.....	12
MEDIDAS CALCULADAS .....	13
SEGMENTACIONES .....	14
VISUALIZACIÓN DE DATOS .....	15
• PORTADA.....	15
• ANÁLISIS GENERAL.....	15
• DIFERENCIAS REGIONALES.....	17
• PRECIOS.....	18
CONCLUSIÓN .....	21
FUTURAS LÍNEAS .....	22

## **INTRODUCCIÓN**

En el marco del proyecto final del curso de Data Analytics dictado por Coderhouse se analizarán diferentes cuestiones vinculadas al cultivo de Soja en la zona centro del país. En particular; se estudiará la superficie sembrada, superficie cosechada, producción y rinde de las diferentes campañas que fueron llevadas a cabo durante el período 2010 a 2020 en las provincias de Córdoba, Santa Fé, La Pampa y San Luis. Nuestra hipótesis general es que la producción sojera presenta marcadas diferencias entre regiones, y ha sufrido variaciones importantes por inundaciones o sequías. También se incorporan los precios al análisis para evaluar el impacto sobre los incentivos del sector.

## **OBJETIVO DEL PROYECTO**

### **1. OBJETIVO GENERAL**

El objetivo general del proyecto es analizar la evolución de la siembra, cosecha, producción y rinde del sector sojero en las distintas provincias del país; y como se ve afectada la decisión de los productores por la variación de precios en los mercados interno y externo.

### **2. OBJETIVOS ESPECIFICOS**

- ✓ Tener conocimiento de cuantas hectáreas son sembradas y cosechadas por campaña en las distintas provincias de la zona centro de Argentina
- ✓ Conocer la cantidad de toneladas producidas y rinde promedio de la hectárea en cada caso
- ✓ Analizar la evolución temporal para identificar períodos más o menos productivos
- ✓ Analizar si existen diferencias entre localidades de una misma provincia
- ✓ Analizar la evolución del precio de la soja en el mercado local y externo para poder entender el comportamiento de los productores y hacer proyecciones.

## **ALCANCE**

El alcance del estudio se reduce a la zona centro del país; provincias de Córdoba, Santa Fé, La Pampa y San Luis; durante el período 2010 a 2020.

## **USUARIO FINAL Y NIVEL DE APLICACIÓN**

El trabajo va dirigido principalmente a todos aquellos productores y/o futuros productores, industrias aceiteras y, en general, a todos los sectores económicos involucrados con el agro y la ganadería de nuestro país que necesiten tomar decisiones inteligentes basadas en el estudio de la coyuntura actual del sector sojero. También a los funcionarios gubernamentales encargados de tomar decisiones de política económica que repercutan sobre sector.

**TABLA DE VERSIONADO**

VERSIÓN	FECHA MODIFICACIÓN
V1	14/09/2022
V2	30/09/2022
V3	03/10/2022
V4	07/11/2022
V5	17/11/2022

**HERRAMIENTAS UTILIZADAS**

Para realizar el presente trabajo se utilizaron los siguientes programas:

- Microsoft Excel para la elaboración de la base de datos.
- Página erplus.com/standalone para crear el Diagrama Entidad Relación.
- Microsoft Power BI Desktop para crear y modificar el dashboard.

**DATASETS**

El datasets original incluía cinco tablas. Las primeras cuatro con estructura similar. Una tabla por provincia con datos de superficie sembrada (ha), superficie cosechada (ha), producción (tn) y rinde (kg/ha) en los distintos departamentos. Los datos se mostraban por campaña y correspondían a las efectuadas durante el período 2010 a 2020. La información se obtuvo de la Dirección de Estadísticas Agrícolas y está actualizada al 30 de agosto de 2022. La quinta tabla contenía datos mensuales de precios de la soja en el mercado interno (Cámara Arbitral de Rosario) y externo (FOB Golfo y puertos argentinos) durante el período enero 2009 - agosto 2022. La fuente de información era el MAGYP y puede consultarse en el siguiente link <https://www.argentina.gob.ar/economia/politicaeconomica/regionalysectorial/informesproductivos/datasets>

## TRABAJO FINAL-DATA ANALYTICS

provincia	id_departamento	departamento	cultivo	id_campaña	campaña	sup_sembrado	sup cosechado	produccion	rendimiento
Cordoba	7	Calamuchita	Soja total	42	2010/2011	94000	94000	227900	2424
Cordoba	7	Calamuchita	Soja total	43	2011/2012	93700	93700	158090	1687
Cordoba	7	Calamuchita	Soja total	44	2012/2013	93700	93700	185800	1983
Cordoba	7	Calamuchita	Soja total	45	2013/2014	69000	69000	178600	2588
Cordoba	7	Calamuchita	Soja total	46	2014/2015	73900	73900	205520	2786
Cordoba	7	Calamuchita	Soja total	47	2015/2016	58530	58530	157031	2683
Cordoba	7	Calamuchita	Soja total	48	2016/2017	42505	42505	127328	2996
Cordoba	7	Calamuchita	Soja total	49	2017/2018	39000	39000	94614	2421
Cordoba	7	Calamuchita	Soja total	50	2018/2019	34600	34600	102950	2975
Cordoba	14	Capital	Soja total	42	2010/2011	4500	4500	8820	1960
Cordoba	14	Capital	Soja total	43	2011/2012	4700	4700	10600	2255
Cordoba	14	Capital	Soja total	44	2012/2013	4700	4700	10290	2189
Cordoba	14	Capital	Soja total	45	2013/2014	4700	4700	16330	3474
Cordoba	14	Capital	Soja total	46	2014/2015	4800	4800	18120	3775
Cordoba	14	Capital	Soja total	47	2015/2016	4400	4400	13088	2975
Cordoba	14	Capital	Soja total	48	2016/2017	4300	4300	10085	2341
Cordoba	14	Capital	Soja total	49	2017/2018	4300	4300	10155	2362
Cordoba	14	Capital	Soja total	50	2018/2019	4100	4100	13788	3363
Cordoba	21	Colon	Soja total	42	2010/2011	113000	112500	231375	2057
Cordoba	21	Colon	Soja total	43	2011/2012	113000	113000	114800	2779
Cordoba	21	Colon	Soja total	44	2012/2013	120000	119000	145800	1225
Cordoba	21	Colon	Soja total	45	2013/2014	118000	118000	377900	3203
Cordoba	21	Colon	Soja total	46	2014/2015	145000	137000	518000	3785
Cordoba	21	Colon	Soja total	47	2015/2016	91293	91293	315026	3451
Cordoba	21	Colon	Soja total	48	2016/2017	85000	85000	268500	3159
Cordoba	21	Colon	Soja total	49	2017/2018	64790	63652	193602	3042
Cordoba	21	Colon	Soja total	50	2018/2019	67210	66161	260823	3942
Cordoba	35	General Roca	Soja total	42	2010/2011	383000	383000	795600	2082
Cordoba	35	General Roca	Soja total	43	2011/2012	398000	384000	794000	1990
Cordoba	35	General Roca	Soja total	44	2012/2013	508200	490200	1103310	2251
Cordoba	35	General Roca	Soja total	45	2013/2014	417000	413500	1132600	2739
Cordoba	35	General Roca	Soja total	46	2014/2015	395400	394700	1118228	3339
Cordoba	35	General Roca	Soja total	47	2015/2016	440200	425000	1555840	3648
Cordoba	35	General Roca	Soja total	48	2016/2017	373000	354680	1273916	3592
Cordoba	35	General Roca	Soja total	49	2017/2018	371000	340000	978300	2677
Cordoba	35	General Roca	Soja total	50	2018/2019	355000	350000	1278730	3654

bar_seja - Excel																			Inicio		Compartir									
Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda ¿Qué desea hacer?																			Inicio		Compartir									
Cortar Copiar Formato Copiar formato Fuente Alineación Número																			Formato condicional Dar formato como tabla		Normal Bueno Incorrecto Neutral		Insertar Eliminar Formato		Autosuma Relevar Borrar Ordenar y filtrar Buscar y seleccionar					
1/7/2009																			Cálculo		Celda de texto Celda vinculada Entidad		Celdas							
LB	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W							
1	sector	nomi	variable	id_actividad	pr	indicador	unidad	de	f_fuente	frecuencia	r_cobertura	n_alcance	tipo_alcance	id	alcance	non	indice	tiempo	valor											
2	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/1/2009		0	MUNDO	1/1/2009		395.81											
3	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/2/2009		0	MUNDO	1/2/2009		388.7											
4	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/3/2009		0	MUNDO	1/3/2009		380.07											
5	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/4/2009		0	MUNDO	1/4/2009		399.58											
6	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/5/2009		0	MUNDO	1/5/2009		446.00											
7	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/6/2009		0	MUNDO	1/6/2009		472.00											
8	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/7/2009		0	MUNDO	1/7/2009		429.00											
9	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/8/2009		0	MUNDO	1/8/2009		443.00											
10	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/9/2009		0	MUNDO	1/9/2009		387.00											
11	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/10/2009		0	MUNDO	1/10/2009		397.00											
12	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/11/2009		0	MUNDO	1/11/2009		406.00											
13	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/12/2009		0	MUNDO	1/12/2009		411.00											
14	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/1/2010		0	MUNDO	1/1/2010		417.00											
15	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/2/2010		0	MUNDO	1/2/2010		388.7											
16	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/3/2010		0	MUNDO	1/3/2010		373.00											
17	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/4/2010		0	MUNDO	1/4/2010		377.01											
18	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/5/2010		0	MUNDO	1/5/2010		373.00											
19	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/6/2010		0	MUNDO	1/6/2010		374.00											
20	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/7/2010		0	MUNDO	1/7/2010		480.00											
21	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/8/2010		0	MUNDO	1/8/2010		424.00											
22	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/9/2010		0	MUNDO	1/9/2010		415.00											
23	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/10/2010		0	MUNDO	1/10/2010		467.00											
24	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/11/2010		0	MUNDO	1/11/2010		492.65											
25	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/12/2010		0	MUNDO	1/12/2010		513.00											
26	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/1/2011		0	MUNDO	1/1/2011		518.00											
27	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/2/2011		0	MUNDO	1/2/2011		557.00											
28	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/3/2011		0	MUNDO	1/3/2011		540.00											
29	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/4/2011		0	MUNDO	1/4/2011		531.00											
30	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/5/2011		0	MUNDO	1/5/2011		520.00											
31	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/6/2011		0	MUNDO	1/6/2011		535.00											
32	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/7/2011		0	MUNDO	1/7/2011		527.00											
33	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/8/2011		0	MUNDO	1/8/2011		527.00											
34	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/9/2011		0	MUNDO	1/9/2011		528.00											
35	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/10/2011		0	MUNDO	1/10/2011		473.00											
36	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/11/2011		0	MUNDO	1/11/2011		455.6											
37	Oleaginosas	754	Soja	Precio	exteri	US\$/tn	MAGYP	Mensual	Mundial	MUNDO	0	MUNDO	1/12/2011		0	MUNDO	1/12/2011		443.4											
Integrantes																			Pcia de Córdoba		Pcia de Santa Fe		Pcia de La Pampa		Pcia de San Luis		Precios			
Escribe aquí para buscar																														

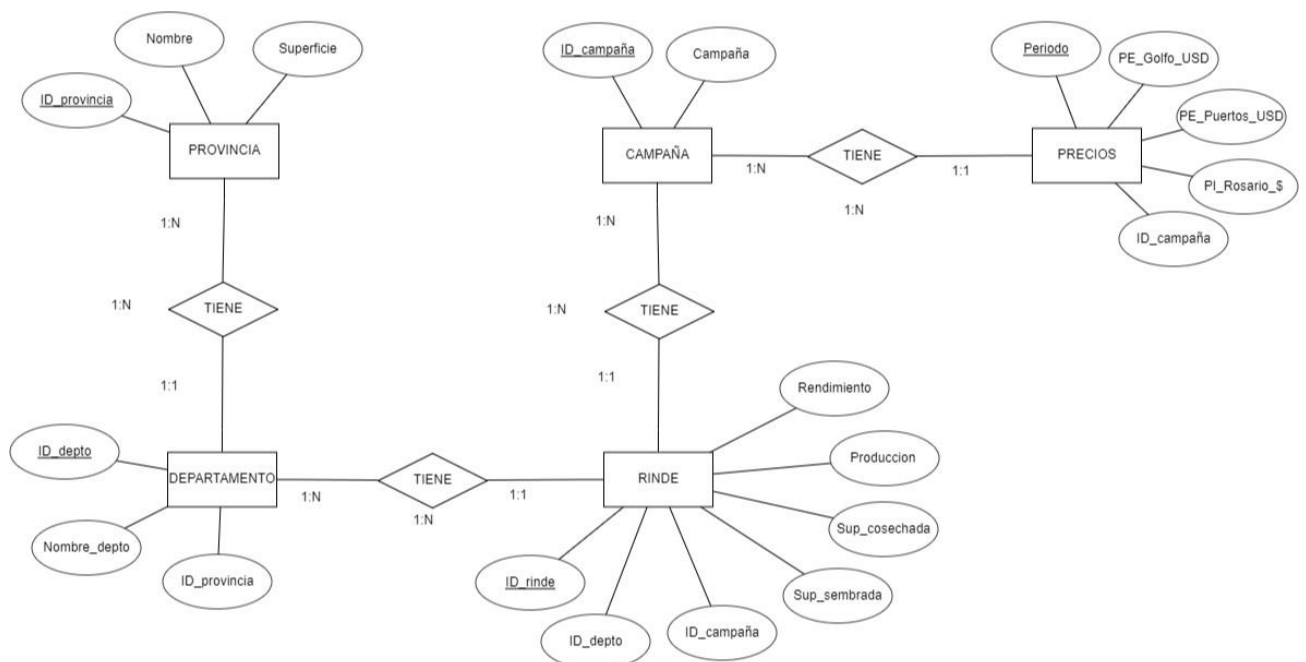
En la segunda preentrega fue necesario modificar un poco la estructura de la información para generar tablas que pudieran vincularse correctamente en el modelo, y realizar el diagrama Entidad-Relación. También se decidió acotar la información de precios al período mayo 2010-abril 2019 para que coincidiera con las campañas analizadas.

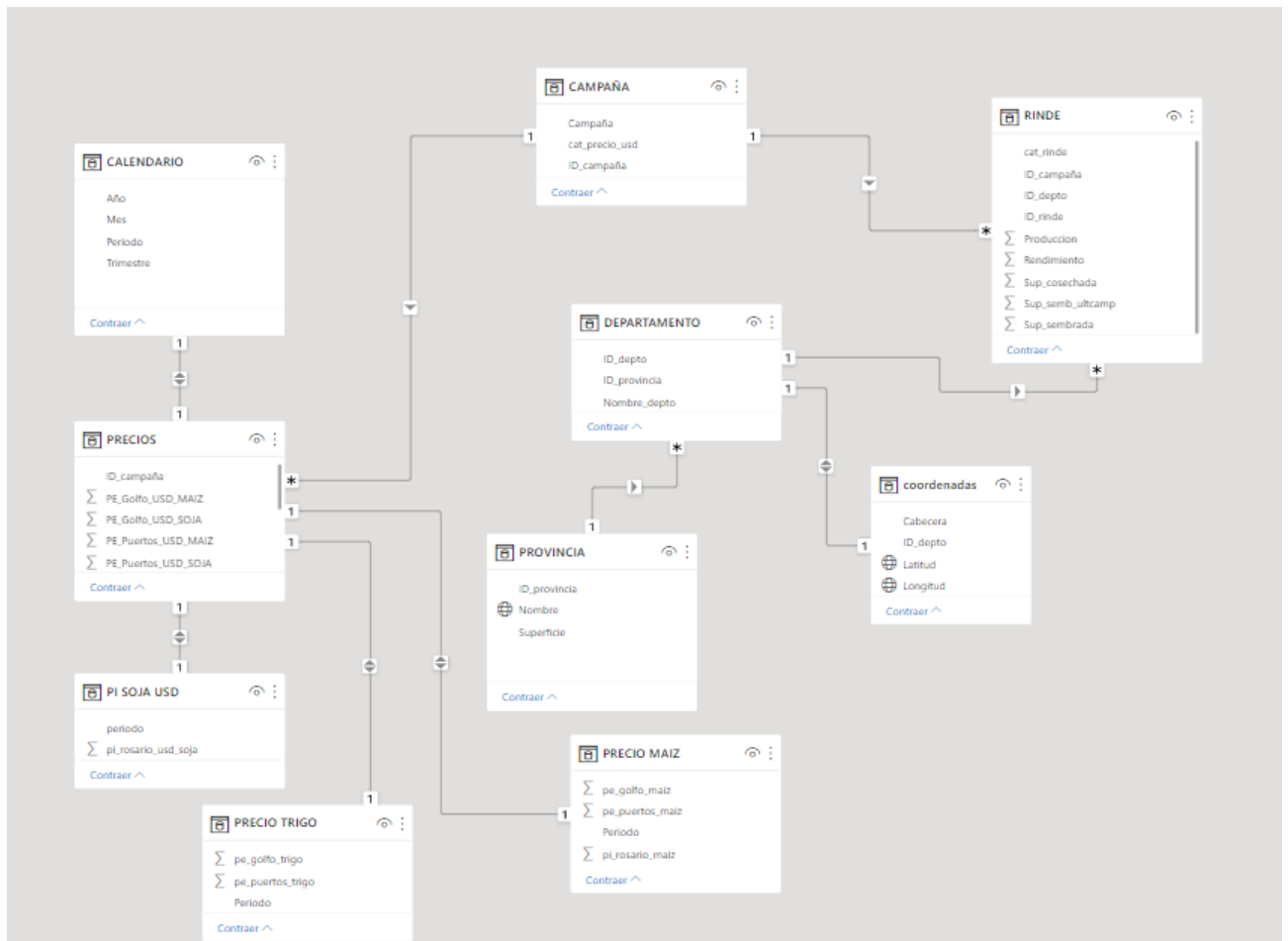
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID_depto	Nombre_depto	ID_provincia													
2		1 Calamuchita	1													
3		2 Capital	1													
4		3 Colon	1													
5		4 General Roca	1													
6		5 General San Martin	1													
7		6 Ischilin	1													
8		7 Juarez Celman	1													
9		8 Marcos Juarez	1													
10		9 Pocho	1													
11		10 Presidente Roque Saenz Pena	1													
12		11 Rio Cuarto	1													
13		12 Rio Primero	1													
14		13 Rio Seco	1													
15		14 Rio Segundo	1													
16		15 San Alberto	1													
17		16 San Javier	1													
18		17 San Justo	1													
19		18 Santa Maria	1													
20		19 Sobremonste	1													
21		20 Tercero Arriba	1													
22		21 Totoral	1													
23		22 Tulumba	1													
24		23 Union	1													
25		24 Atreuco	2													
26		25 Capital	2													
27		26 Catrilo	2													

En la última prentrega se incorporó al dataset información de precios de cultivos alternativos como el trigo y maíz, el precio interno en dólares de la soja y las coordenadas que permiten geolocalizar los departamentos. Ello posibilitó enriquecer el análisis y realizar mapas en POWER BI.

## ESQUEMA ENTIDAD-RELACIÓN

A continuación, se presenta el diagrama entidad-relación creado al principio del proyecto, y el modelo de datos utilizado en el tablero final luego de incorporar la información adicional de precios y coordenadas:





## LISTADO DE TABLAS

**TABLA PROVINCIA:** Esta tabla contiene las cuatro provincias de la zona centro del país que serán analizadas, con su superficie en hectáreas.

- ID\_provincia (PK)
- Nombre
- Superficie

**TABLA DEPARTAMENTO:** Esta tabla contiene los distintos departamentos de cada una de las provincias seleccionadas donde se va a realizar el análisis de la producción de soja.

- ID\_depto (PK)
- Nombre\_depto
- ID\_provincia (FK)

**TABLA RINDE:** Esta tabla contiene datos por departamento de superficie sembrada (ha), superficie cosechada (ha), producción (tn) y rendimiento (kg/ha) obtenido en las cosechas realizadas en cada campaña.

- ID\_rinde (PK)
- ID\_depto (FK)
- ID\_campaña (FK)
- Sup\_semrada
- Sup\_cosechada
- Produccion
- Rendimiento

**TABLA CAMPAÑA:** En esta tabla se detallan las campañas analizadas, que van desde el año 2010 al 2020.

- ID\_campaña (PK)
- Campaña

**TABLA PRECIOS:** Esta tabla contiene los precios externos e internos de los distintos mercados donde se comercializa la soja.

- Periodo (PK)
- PE\_Puertos\_USD
- PE\_Golfo\_USD
- PI\_Rosario\_\$
- ID\_campaña (FK)

**TABLA PRECIO MAIZ:** Esta tabla contiene los precios externos e internos del maíz

- Periodo (PK)
- pe\_puertos\_maiz
- pe\_golfo\_maiz
- pi\_rosario\_maiz

**TABLA PRECIO TRIGO:** Esta tabla contiene los precios externos del trigo

- Periodo (PK)
- pe\_puertos\_trigo
- pe\_golfo\_trigo

**TABLA PI SOJA USD:** Esta tabla contiene el precio interno de la soja en dólares

- Periodo (PK)
- pi\_rosario\_usd\_soja

**TABLA COORDENADAS:** Esta tabla contiene la geolocalización de cada departamento

- ID\_depto (PK)
- Cabecera
- Latitud
- Longitud

**TABLA CALENDARIO:** Se genera a partir de campo período



- Periodo (PK)
- Año
- Mes
- Trimestre

**TABLA AMEDIDAS:** Se genera para guardar las medidas calculadas

## TIPOS DE DATOS

**TABLA PROVINCIA**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
ID_provincia	Int	PK	Clave única de identificación de la provincia
Nombre	Varchar (50)	-	Nombre de la provincia
Superficie	Decimal (18,2)	-	Superficie de la provincia medida en hectáreas

**TABLA DEPARTAMENTO**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
ID_depto	Int	PK	Clave única de identificación del departamento
Nombre_depto	Varchar (100)	-	Nombre del departamento
ID_provincia	Int	FK	Clave de identificación de la provincia. Es la clave foránea que relaciona esta tabla con la tabla PROVINCIA

**TABLA RINDE**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
ID_rinde	Int	PK	Clave única de identificación de cada registro. Surge de concatenar los campos ID_Depto e ID_Campaña.

ID_depto	Int	FK	Clave de identificación del departamento. Es la clave foránea que relaciona esta tabla con la tabla DEPARTAMENTO
ID_campaña	Int	FK	Clave de identificación de la campaña. Es la clave foránea que relaciona esta tabla con la tabla CAMPAÑA
Sup_sembrada	Decimal (18,2)	-	Superficie sembrada medida en hectáreas
Sup_cosechada	Decimal (18,2)	-	Superficie cosechada medida en hectáreas
Produccion	Decimal (18,2)	-	Cantidad producida medida en toneladas
Rendimiento	Decimal (18,2)	-	Rendimiento promedio medido en kilogramos por hectárea

**TABLA CAMPAÑA**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
ID_campaña	Int	PK	Clave única de identificación de la campaña
Campaña	Varchar (50)	-	Nombre de la campaña

**TABLA PRECIOS**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
Periodo	Int	PK	Indica el año y el mes en que se observa cada precio. Es la Clave única de identificación de cada registro
PE_Golfo_USD	Decimal (18,2)	-	Precio externo FOB Golfo. Está medido en dólares por tonelada
PE_Puertos_USD	Decimal (18,2)	-	Precio externo FOB Puertos Argentinos. Está medido en dólares por tonelada
PI_Rosario_₺	Decimal (18,2)	-	Precio interno Cámara Arbitral de Rosario. Está medido en pesos por tonelada.
ID_campaña	Int	FK	Clave de identificación de la campaña. Es la clave foránea que relaciona esta tabla con la tabla CAMPAÑA

**TABLA PRECIO MAIZ**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
Periodo	Int	PK	Indica el año y el mes en que se observa cada precio. Es la Clave única de identificación de cada registro

pe_golfo_maiz	Decimal (18,2)	-	Precio externo FOB Golfo. Está medido en dólares por tonelada
pe_puertos_maiz	Decimal (18,2)	-	Precio externo FOB Puertos Argentinos. Está medido en dólares por tonelada
pi_rosario_maiz	Decimal (18,2)	-	Precio interno Cámara Arbitral de Rosario. Está medido en dólares por tonelada

**TABLA PRECIO TRIGO**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
Periodo	Int	PK	Indica el año y el mes en que se observa cada precio. Es la Clave única de identificación de cada registro
pe_golfo_trigo	Decimal (18,2)	-	Precio externo FOB Golfo. Está medido en dólares por tonelada
pe_puertos_trigo	Decimal (18,2)	-	Precio externo FOB Puertos Argentinos. Está medido en dólares por tonelada

**TABLA PI USD SOJA**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
Periodo	Int	PK	Indica el año y el mes en que se observa cada precio. Es la Clave única de identificación de cada registro
pi_rosario_usd_soja	Decimal (18,2)	-	Precio interno Cámara Arbitral de Rosario. Está medido en dólares por tonelada

**TABLA COORDENADAS**

CAMPO	TIPO DE DATO	TIPO DE CLAVE	DESCRIPCIÓN
ID_depto	Int	PK	Clave única de identificación del departamento
Cabecera	Varchar (100)	-	Nombre de la ciudad cabecera del departamento
Latitud	Decimal (18,5)	-	Latitud
Longitud	Decimal (18,5)	-	Longitud

**TRANSFORMACIONES REALIZADAS****TABLA CAMPAÑA**

- Se agrego una columna condicional que categoriza al precio promedio en una determinada campaña como Bajo, Medio o Alto. Para ello se calculó el precio externo FOB Puertos Argentinos medio por campaña y se las ordenó de menor a mayor según el indicador. Finalmente se clasificaron como “Bajo” las primeras tres campañas, “Medio” la tres del medio y “Alto” las últimas tres con mayores valores.

**TABLA RINDE**

- Se agrego una columna condicional que clasifica el rendimiento promedio por hectárea en Bajo, Medio o Alto
  - Si el rinde medio del departamento fue menor a 2.098 se clasifica como “Bajo”
  - Si el rinde medio del departamento estuvo entre 2.098 y 2.700 se clasifica como “Medio-Bajo”
  - Si el rinde medio del departamento estuvo entre 2.700 y 3.263 se clasifica como “Medio-Alto”
  - Si el rinde medio del departamento fue mayor a 3.263 se clasifica como “Alto”
- Se agregó una columna condicional “Sup\_sembrada\_ultcamp” que muestra la superficie sembrada sólo si se trata de la última campaña (id\_campaña=50)

**TABLA CALENDARIO**

- Se generaron columnas con el año, mes y trimestre a partir del campo período.

**TABLA PRECIOS**

- Se combinó con la TABLA PRECIO MAIZ por el campo período para traer los precios externos e interno del cultivo
- Se combinó con la TABLA PRECIO TRIGO por el campo período para traer el precio externo del cultivo

- Se combinó con la TABLA PI SOJA USD por el campo período para traer el precio interno en dólares de la soja
- Se calculó el precio relativo de la soja respecto del maíz en una columna separada "Personalizada agregada1" = Table.AddColumn("#Tipo cambiado2", "pe\_relativo", each [PE\_Puertos\_USD\_SOJA]/[PE\_Puertos\_USD\_MAIZ]),
- Se calculó el precio relativo de la soja respecto del trigo en una columna separada "Personalizada agregada2" = Table.AddColumn("#Columnas con nombre cambiado3", "pe\_rel\_TRIGO", each [PE\_Puertos\_USD\_SOJA]/[PE\_Puertos\_USD\_TRIGO]),

## MEDIDAS CALCULADAS

1. **Cant\_dptos:** Se creó esta medida para poder obtener el total de departamentos de las diferentes provincias céntricas del país que se analizaron en el presente trabajo. La fórmula de cálculo es la siguiente:

$$cant\_dptos = COUNT(DEPARTAMENTO[ID\_depto])$$

Dicha medida se utiliza en una de las tarjetas del tablero



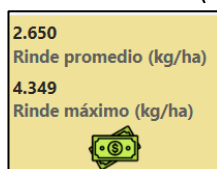
2. **Max\_rendimiento:** Esta medida se creó para conocer el rendimiento medio máximo teniendo en cuenta todas las campañas y departamentos. La fórmula de cálculo es la siguiente:

$$Max\_rendimiento = max(RINDE[Rendimiento])$$

3. **Prom\_rendimiento:** Esta medida se creó para obtener el rendimiento promedio.

$$Prom\_rendimiento = AVERAGE(RINDE[Rendimiento])$$

Dichas medidas (2 y 3) se utilizan en la siguiente tarjeta del tablero



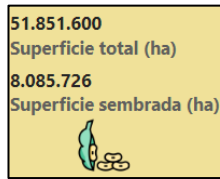
4. **Sup\_semb\_totuc:** Esta medida se creó para obtener la superficie sembrada total en la última campaña analizada (2018/2019)

$$Sup\_semb\_totuc = SUM(RINDE[Sup\_semb\_ultcamp])$$

5. **Sup\_total:** Debido a que en la tabla provincia se pueden visualizar las superficies de cada una de las provincias céntricas del país, con esta medida se buscó obtener la suma del total de las mismas y así poder comparar con el total de superficies sembradas (medida n°4)

$$Sup\_total = sum(PROVINCIA[Superficie])$$

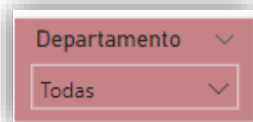
Dichas medidas (4 y 5) se utilizan en la siguiente tarjeta del tablero



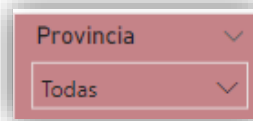
## SEGMENTACIONES

En el trabajo se utilizaron 7 tipos de segmentaciones

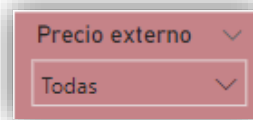
- Departamento: utilizado para filtrar los distintos departamentos de cada provincia.



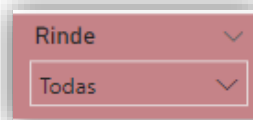
- Provincia: utilizado para filtrar las distintas provincias en análisis.



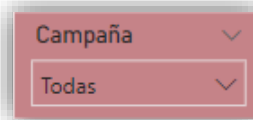
- Precio Externo: utilizado para filtrar por categoría de precio.



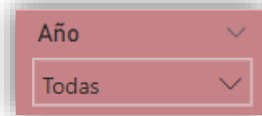
- Rinde: utilizado para filtrar por categoría de rinde.



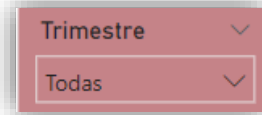
- Campaña: utilizado para filtrar las distintas campañas de soja.



- Año: utilizado para filtrar los años analizados.

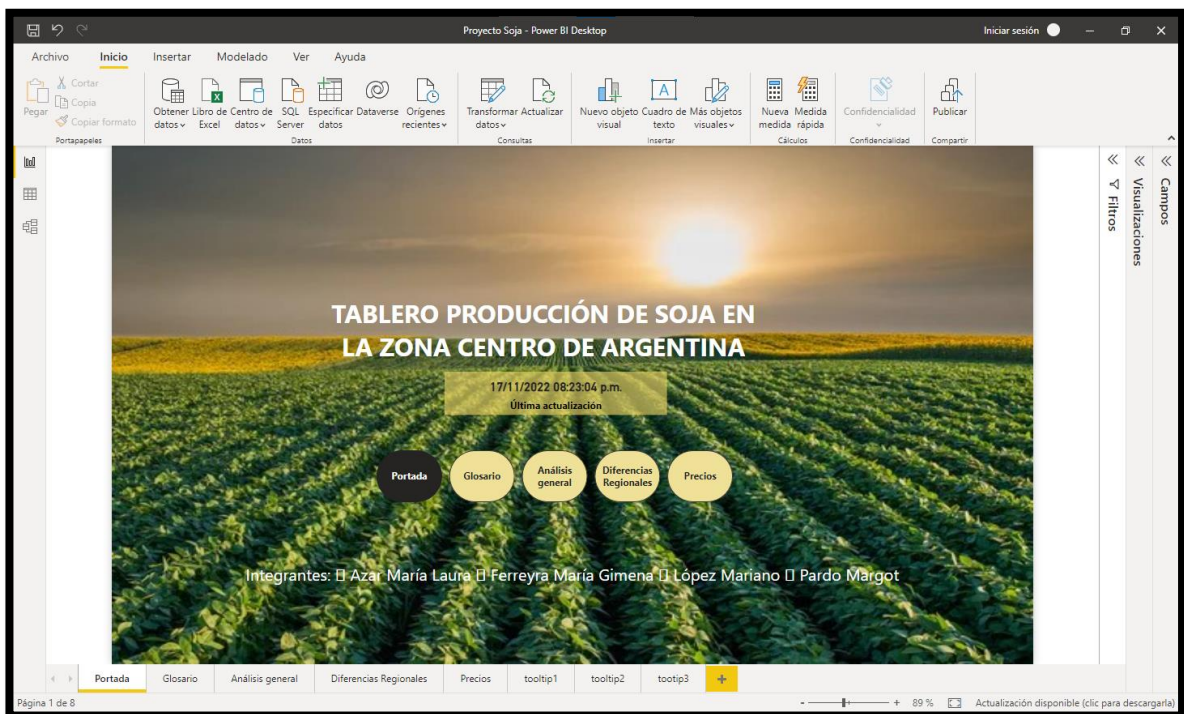


- Trimestre: utilizado para filtrar los trimestres.



## VISUALIZACIÓN DEL TABLERO

En la primera solapa se presenta la portada del dashboard. En la misma se crearon botones de navegabilidad, para que el usuario pueda ir a la página que desee, y se colocó la fecha de última actualización.



En la segunda solapa se incorporó un glosario explicando los conceptos más importantes.





En la tercera solapa se presenta un análisis general para que el usuario pueda tener un conocimiento de lo sembrado, cosechado y el rinde del cultivo de soja.

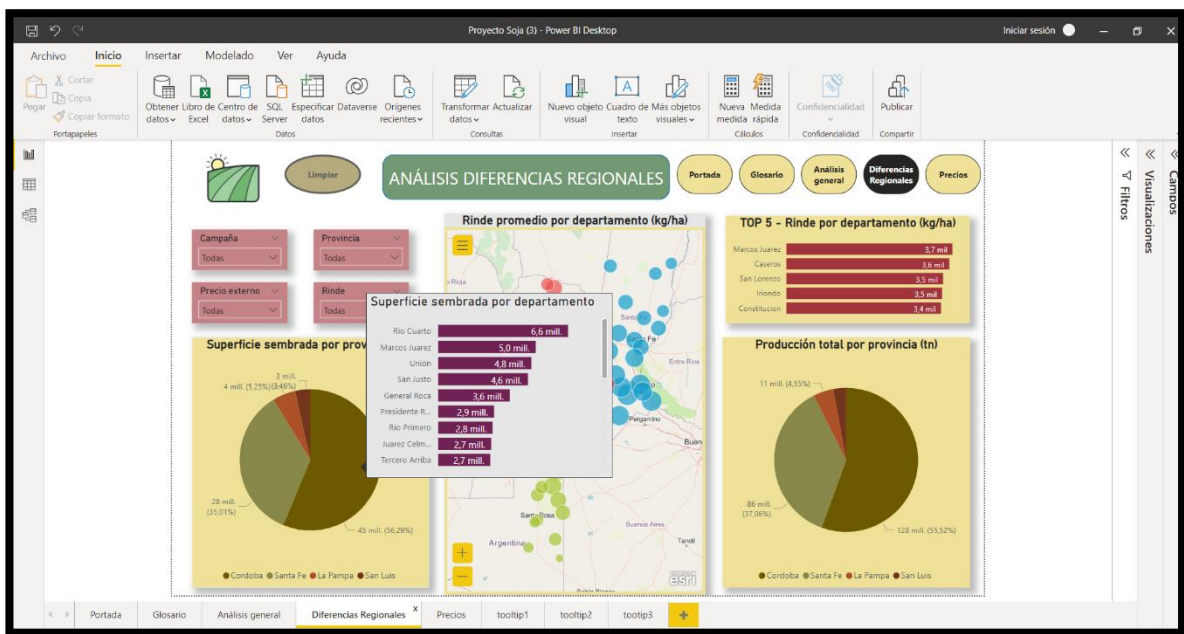
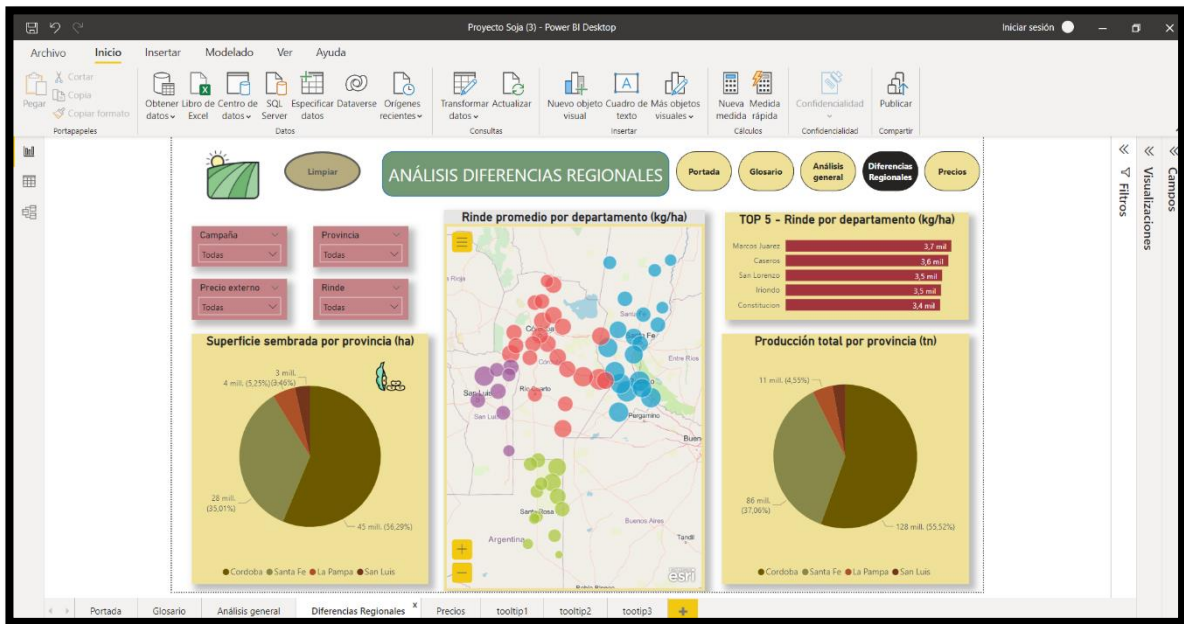
- En la primera tarjeta del margen superior izquierdo se muestra la superficie total y la superficie sembrada en la última campaña, ambas en hectáreas.
- En la tarjeta del medio se muestra el rinde máximo y promedio, para poder ver cuánto es el rinde en kilogramos por hectárea sembrada en los años analizados.
- En la tarjeta del margen superior derecho se puede ver el total de departamentos analizados en el trabajo.
- En el margen inferior izquierdo se presenta un gráfico de columnas agrupadas y de líneas, en donde se puede visualizar la evolución del área cosechada y sembrada en las distintas campañas en relación al rinde promedio del cultivo.
- En el margen inferior derecho se puede visualizar un gráfico de barras horizontales en donde se muestra un ranking de las campañas que mayor producción tuvieron. En este último se sumó un tooltip que permite ver lo producido en cada provincia.

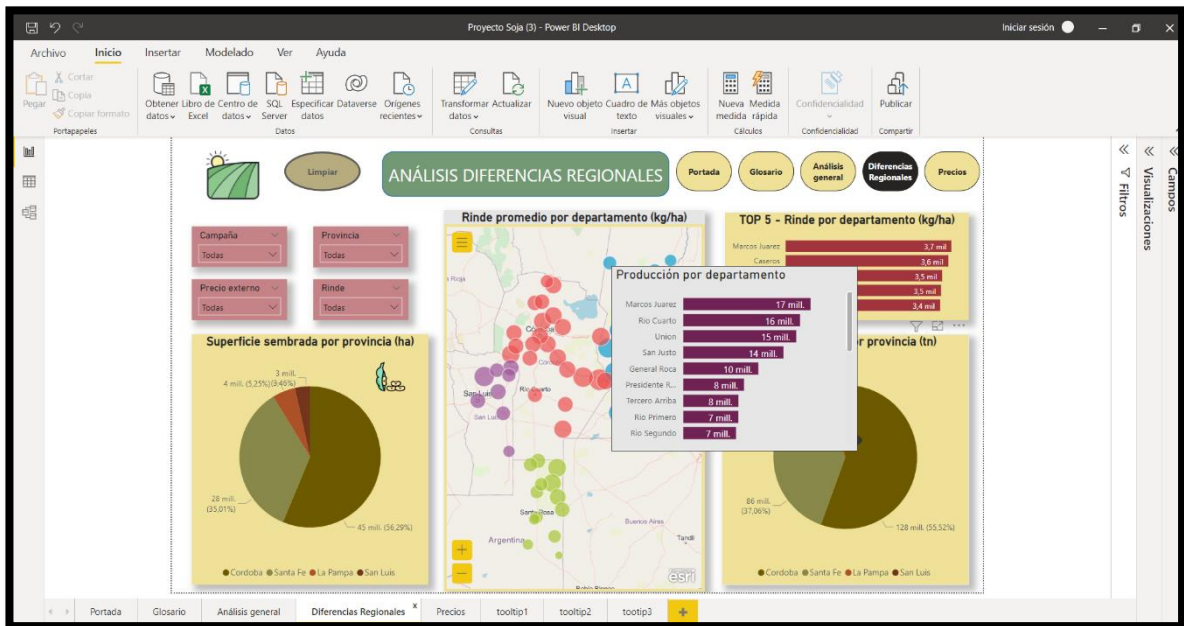


En la cuarta solapa se presenta un análisis por departamento que permite identificar diferencias y similitudes entre ellos

- En el margen superior derecho se muestra un gráfico de barras horizontales para detallar los cinco departamentos con mayor rinde.
- En el margen inferior izquierdo se visualiza un gráfico circular que muestra la superficie sembrada por provincia. Aquí se incorporó un tooltip que permite conocer el mismo dato por departamento.
- En el margen inferior derecho se visualiza un gráfico circular en donde se muestra la producción total por provincia en toneladas. Como en el caso anterior se aplicó un tooltip que detalla el mismo dato por departamento.
- En el centro de la solapa se generó un mapa donde pueden observarse claramente las áreas con mayor rendimiento.

# TRABAJO FINAL-DATA ANALYTICS





La última solapa presenta un análisis de precios. El objetivo es que el usuario conozca su evolución y el efecto que tienen en la decisión de los productores de soja.

- En las tarjetas se muestran los precios máximos y mínimos de comercialización en los mercados interno y externo. Se anuló la interacción con las segmentaciones para que estos valores permanezcan inalterados.
- En el gráfico de líneas del margen inferior izquierdo se analiza la evolución de los precios a lo largo de los años. Se anuló la interacción con las segmentaciones por provincia o departamento que carecían de sentido.
- En el gráfico de columnas y líneas del margen inferior derecho podemos ver la evolución de la superficie sembrada por campaña en relación al cambio de precios relativos. En este caso se anuló la interacción con las segmentaciones por año o trimestre.



En cada una de las páginas se incorporó el botón limpiar para que el usuario pueda volver a la visualización inicial del tablero, libre de segmentaciones, siempre que lo desee.

También se utilizaron iconos con imágenes para mejorar la experiencia al visualizar las páginas; tales como:



## CONCLUSIÓN

El tablero cumple con los objetivos general y específicos propuestos al inicio del proyecto. Permite al usuario analizar la evolución de la siembra, cosecha, producción y rinde del sector sojero en la zona centro del país; conocer la cantidad de hectáreas sembradas, cosechadas y el rinde medio de cada lugar; conocer las toneladas producidas y el grado de concentración de la actividad; identificar diferencias regionales y si existen períodos más o menos productivos; etc. También es posible efectuar un análisis de la evolución de los tres precios fundamentales en la determinación del valor del cultivo; y cómo pudo afectar las decisiones de siembra de los productores.

Entre las cuestiones a destacar que pueden observarse cabe mencionar:

- La evolución de la superficie sembrada presenta la forma de una “U” invertida. Si bien el número de hectáreas sembradas crece en forma sostenida durante las primeras campañas, desde el año 2016 viene mostrando un fuerte retroceso posiblemente en favor de otros cultivos alternativos que presentan mejor rentabilidad.
- El rendimiento medio de la hectárea muestra altibajos pronunciados, con una caída importante en la campaña 2017/2018 y un fuerte rebote posterior.
- El período de mayor producción fue 2014 a 2016, seguido de cerca por la última campaña analizada 2018/2019.
- En todas las campañas se observa una importante concentración de la superficie sembrada y la producción en las provincias de Córdoba y Santa Fé, con una participación conjunta cercana al 80% del total.
- Las diferencias regionales también se evidencian al interior de las provincias, especialmente en el rendimiento por hectárea de las distintas zonas.

Una última cuestión a señalar relativa a los precios es la caída observada en el precio internacional de la soja desde fines de 2012, y el achicamiento de la brecha entre los precios internos y externos desde 2015 debido posiblemente a la reducción en los derechos de exportación

Si bien tenemos varias ideas que podrían sumarse al tablero para mejorar la experiencia del usuario, creemos que la versión presentada permite extraer muchísimas conclusiones importantes y resulta de gran utilidad para el usuario interesado en el sector.

## **FUTURAS LÍNEAS**

A continuación, se enumeran un conjunto de mejoras que podrían sumarse al tablero en el futuro para incrementar su utilidad:

- Ampliar el alcance del análisis a todo el país incorporando información de las provincias que faltan
- Sumar las campañas recientes: 2019/2020, 2020/2021 y 2021/2022
- Incorporar datos de costos de producción: insumos; alquiler, compra y mantenimiento de maquinaria, arrendamiento y valor de la tierra, etc.
- Identificar shocks internos y externos como sequías, inundaciones, guerras y otros conflictos externos utilizando variables dummies.
- Identificar cambios en la política fiscal del gobierno asociada al sector (subsidios y/o impuestos) utilizando otra variable dummy específica.