



Ministerio de Agricultura,
Ganadería y Pesca
Argentina

ESTIMACIONES AGRÍCOLAS - INFORME 1º 2022

Reporte de Data Analytics

Proyecto de Data Analytics:
Producción Nacional de girasol, avena y mijo de la
República Argentina

Ariel Fernando Marquez Agüero
15/06/2022

Tabla de contenido

Tabla de contenidos.....	1
Introducción.....	2
Objetivo	2
Alcance.....	2
Destinatarios	3
Diagrama Entidad - Relación.....	3
Definición de tablas (clave primaria [PK] y clave foránea [FK])	3
Producción.....	3
Calendario	4
Medidas.....	4
Columnas y definición de tipos de datos	4
Modificaciones realizadas en los dataset	5
Trasformaciones realizadas en Power BI	5
Columnas y tablas generadas – Detalle	6
Medidas calculadas.....	6
Análisis funcional.....	6
Conclusiones.....	7
Puntos a futuro.....	8
Referencias	9

Introducción

El conocimiento es un factor esencial tanto para alcanzar objetivos y metas estratégicas como elevar en el nivel de productividad y rentabilidad, sin embargo, a pesar de que la información se genera de manera automática, existen muchas organizaciones que no están aprovechándola para la correcta toma de decisiones **informada**. Frecuentemente la información disponible es un limitante para la toma de decisiones acertada, es por eso que hoy en día las organizaciones están apostando por la implementación de distintas herramientas y/o soluciones que le ayuden a alcanzar sus metas y simplificar la obtención de **datos confiables** para poder llevar a cabo un análisis válido y tomar decisiones informadas.

Antes, los directores recurrían a la intuición para decidir el rumbo de su organización. Ahora, con el **análisis de datos** para la toma de decisiones ahora pueden experimentar con los productos, servicios o conocer los beneficios que puede traer invertir en nuevos negocios y responder a las necesidades, tanto internas como externas. Ese mayor acceso y transparencia informativa influye en la **democratización de la toma de decisiones**. Con un crecimiento potencial en el sector público, por el reclamo ciudadano de más transparencia en las cuentas públicas.

Objetivo

El objetivo del siguiente trabajo es analizar los datos disponibles del gobierno nacional, mediante MAGyP (Ministerio de Agricultura, Ganadería y Pesca), en la producción nacional de cultivos de girasol, avena y mijo desde los primeros registros recuperados hasta el año más actual de actualización con sus correspondientes campañas, con el fin de comprender cuál es la realidad nacional y provincial de dicha producción, haciendo énfasis en la cantidad de superficie sembrada y cosechada.

A su vez, buscamos reconocer la provincia con mayor rendimiento por superficie tanto por cultivo como para tales periodos anuales.

Como propósito final, nuestra función es asistir al gobierno nacional según MAGyP en el delineado de una estrategia para aumentar la tasa de rendimiento de las cosechas en dichos cultivos.

Alcance

Se determinó como banda temporal de análisis a los registros de los años 1920 hasta el año 2020, tomando en cuenta todas las provincias más significativas para este estudio que componen la República Argentina.

Destinatarios

Este análisis está orientado a asistir a la mesa de dirección del gobierno nacional, MAGyP y a los distintos funcionarios de las distintas provincias del país, con el fin de aportar información valiosa para la gestión y el accionar económico, tecnológico y político como plan productivo. A trabajadores agropecuarios y dueños de terrenos cultivables rentables para generar índices productivos más ajustados a los parámetros reales. Y, por último, al público en general por tratarse de información de libre acceso acerca de los cultivos analizados a modo de información como interés múltiple y medio didáctico analítico.

Diagrama Entidad - Relación



Definición de tablas (clave primaria [PK] y clave foránea [FK])

Producción

Contiene la información principal de los cultivos, las provincias, fechas, campañas, superficies cultivadas y cosechadas, con su respectivo rendimiento. Su clave principal es el “año” y su clave foránea es “año” en la tabla *Calendario* respectivamente, para vincularla con sus respectivos atributos correspondientes.

Calendario

Aquí datamos la información de forma cronológica por año según los registros. Su clave foránea es “año” respectivamente a la tabla *Producción*, donde se hace su correspondiente medida siguiendo la línea de código DAX según cálculo de la función:
Calendario = CALENDAR(MIN('Producción'[año]), MAX('Producción'[año])).

Medidas

Contiene las medidas correspondientes, que para este caso posee el cálculo en sumatoria de los cultivos de girasol, avena y mijo proporcionados, siguiendo la línea de código DAX según cálculo de función: Cultivos = COUNT('Producción'[cultivo]).

Columnas y definición de tipos de datos

TABLA: Producción	
COLUMNAS	TIPO DE DATOS
cultivo	Texto
año	Fecha
campana	Texto
provincia	Texto
provincia id	Número entero
departamento	Texto
departamento id	Número entero
superficie cultivada en ha	Número entero
superficie cosechada en ha	Número entero
producción en tm	Número entero
rendimiento en kgxha	Número entero

TABLA: Calendario	
COLUMNAS	TIPO DE DATOS
año	Fecha

TABLA: Medidas	
COLUMNAS	TIPO DE DATOS
Cultivo	Número entero

Modificaciones realizadas en los dataset

Se descargo los archivos originales con sus correspondientes tablas de fuentes gubernamentales, tales como el MAGyP: <https://www.argentina.gob.ar/agricultura> , proporcionados por la fundación desde su plataforma: <http://fesat.mendoza.edu.ar/>

Con el fin de la protección de datos personales, las tablas descargadas no contenían *Primary Keys [PK]* por lo que utilizamos las herramientas correspondientes de PowerQuery y estructuras relacionales en función de Power BI.

Trasformaciones realizadas en Power BI

Todas las tablas fueron renombradas y se les quito el guion bajo y remplazados por espacio

- Se agrupo y unifico las tablas en una única llamada Producción:
 - Se tomo las tablas de girasol, avena y mijo anexándose en dicha nueva tabla
- Se cambio los valores erróneos de fechas de campaña con coincidentes a sus años:
 - Se modifiko las filas de dichos periodos en la tabla de avena para rectificar dichas fechas en los gráficos.
- Se filtro los valores nulos correspondientes, antes de los cálculos a razón de sus promedios:
 - Datos como superficie sembrada, superficies cosechadas, producción en tm, entre otras, poseían datos nulos (null) que fueron filtrados paralelamente al hacer los cálculos.
- Se creo tabla Calendario:
 - Se tomo como fecha de inicio la primera del año que se registro y como fecha de fin la más actualizada en actualidad.
- Se creo tabla Medidas:
 - Se tomo como medida la contabilización como suma del total de cultivos

Columnas y tablas generadas – Detalle

Las columnas calculadas introducidas en este estudio fueron:

- Calendario:

En esta columna se obtiene los años de forma cronológica según los registros, donde se hace su correspondiente medida siguiendo la línea de código DAX según cálculo de la función: $\text{Calendario} = \text{CALENDAR}(\text{MIN}(\text{'Producción'[\text{año}]}, \text{MAX}(\text{'Producción'[\text{año}]})$)).

Medidas calculadas

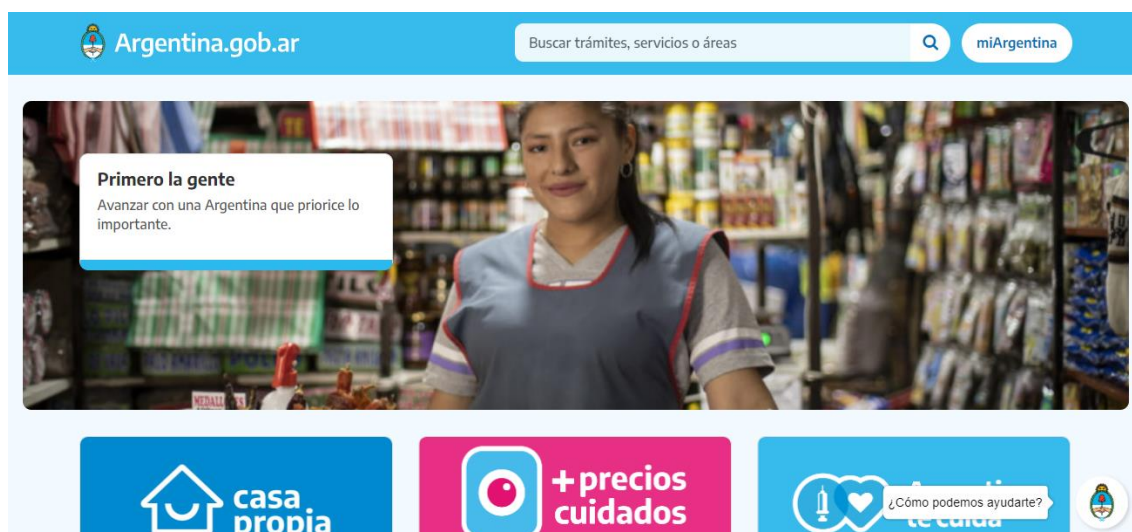
- Cultivo:

Esta columna posee el cálculo en sumatoria de los cultivos de girasol, avena y mijo proporcionados respectivamente, siguiendo la línea de código DAX según cálculo de función: $\text{Cultivos} = \text{COUNT}(\text{'Producción'[\text{cultivo}]})$.

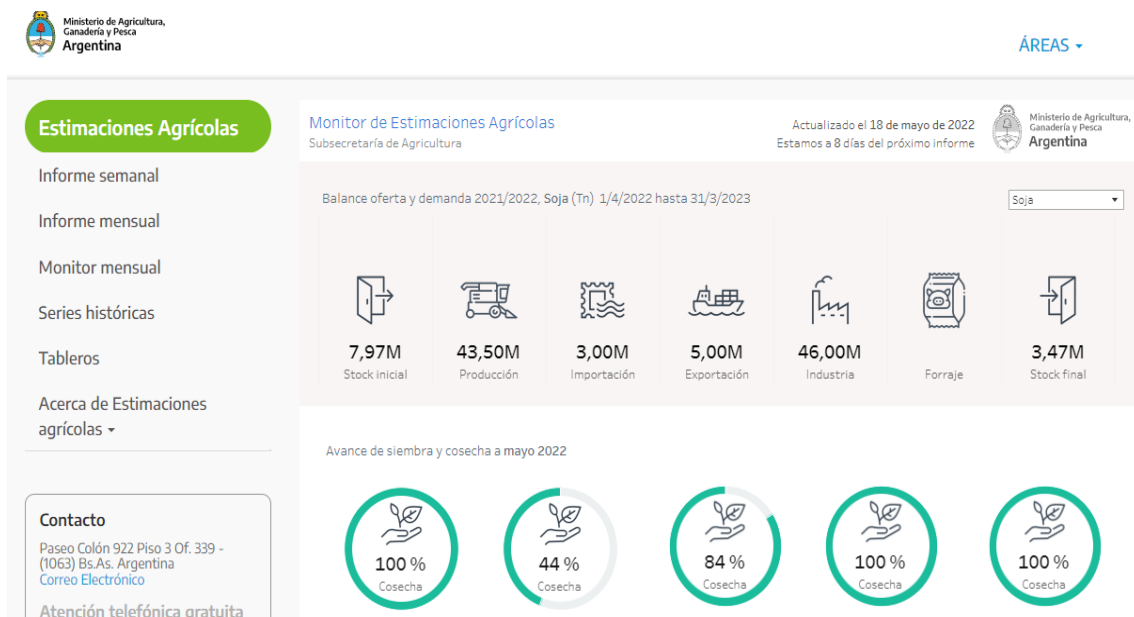
Análisis funcional

Para realizar la visualización de los datos se tomó como plantilla los colores oficiales del gobierno nacional, tal cual se formaliza en los informes del MAGyP, el detalle de los mismos es el siguiente:

- Banner superior: #049DD9
- Botones: #049DD9
- Fondo: #E9EFF2
- Texto: #261E17



- Banner superior: #000000
- Botones: #000000
- Fondo: #F2F2F2
- Texto: #261E17



Conclusiones

Este estudio no se continua en las estimaciones actualizadas, ya que no llega al presente año, por lo que se debería incrementar el alcance y actualizar dicho estudio a un periodo de tiempo más actual.

A razón de lo expuesto en este proyecto, se ve y concluye que se tendría que observar y ajustar más en detalle los datos erróneos y/o faltantes a razón de dicha información requerida, como parte de las competencias censadas a evaluar en el trabajo de campo para la toma de muestra a nivel estadístico, por el organismo censal correspondiente al presente estudio y sus atributos correspondientes.

Puntos a futuro

- Incrementar el alcance y actualizar dicho estudio a un periodo de tiempo más actual
- Ajustar datos nulos y/o en blanco
- Obtener medidas más acordes a lo pertinente y rectificadas según los atributos, observaciones y evaluaciones propuestas
- Mejorar y actualizar el tablero (Dashboard) a razón de las especificaciones mandatorias de los interesados.

Referencias

- Dataset: Producción
- Apache2 Debian Default Page: It works. (s. f.). FESAT. Recuperado 15 de junio de 2022, de <http://fesat.mendoza.edu.ar/>

- Tema fuente:

Argentina.gob.ar. (s. f.). Gobierno de la República Argentina. Recuperado 15 de junio de 2022, de <https://www.argentina.gob.ar/>

- Tema fuente:

Prensa, C. D. D. Y. (s. f.). MAGyP. MAGyP. Recuperado 15 de junio de 2022, de <https://www.magyp.gob.ar/sitio/areas/estimaciones/monitor/index.php>
<https://www.magyp.gob.ar/sitio/areas/estimaciones/monitor/index.php>