

Dashboard para el análisis sobre la efectividad de los cultivos y su relación con el clima a partir de datos abiertos en Colombia

Autores

Javier Eduardo Solano Mesa
John Fredy Orjuela Herrera

Director

Elio H. Cables Pérez, PhD.

Universidad Antonio Nariño
Facultad de Ingeniería de Sistemas
Especialización en Gobierno de Datos
Bogotá D.C
2022

Índice de contenidos

1. Introducción.....	1
2. Descripción y formulación del problema	2
2.1 Objetivo General	3
2.2 Objetivos Específicos	3
3. Marco referencial.....	4
3.1 Marco teórico	4
3.2 Ministerio de Agricultura	4
3.3 IDEAM.....	4
3.4 EVA.....	5
3.5 SQL Server	5
3.6 ETL – SQL Server Integration Service.....	5
3.7 Postman.....	5
3.8 Python.....	6
3.9 Docker.....	6
3.10 Power BI.....	6
4. Estado del arte	7
4.1 Impacto.....	8
4.2 Componente de Innovación	8
5. Metodología.....	9
6. Desarrollo de la propuesta.....	11
6.1 Identificación de Datos.....	12
6.2 Consulta De Los Datos	13
6.3 Alistamiento de datos	15
6.4 Almacenamiento	15
6.5 Procesamiento	16
6.6 Resultado	40
7. Conclusiones	48
8. Referencias	49

Índice de Figuras

Ilustración 1. Metodología para el procesamiento de los datos.....	9
Ilustración 2. Diagrama de Flujo de datos	11
Ilustración 3. Petición Servicio Api Web	13
Ilustración 4. Petición Servicio Api por Postman.....	14
Ilustración 5. Imagen Docker SQL SERVER	15
Ilustración 6. Detalle campos Api	16
Ilustración 7. Detalle campos Api 2.....	17
Ilustración 8. Resumen Datos de Origen	17
Ilustración 9. Definición datos de Origen	17
Ilustración 10. Normalización de variables	18
Ilustración 11. Normalización Variables 2	19
Ilustración 12. Resumen Apis IDEAM.....	20
Ilustración 13. Validación URL en Python	22
Ilustración 14. Normalización de datos con Pandas.....	22
Ilustración 15. Normalización de datos con Pandas 2	23
Ilustración 16. Normalización de datos con Pandas 3	24
Ilustración 17. Normalización de datos con Pandas 4	24
Ilustración 18. Normalización de datos con Pandas 5	25
Ilustración 19. Normalización de datos con Pandas 6	25
Ilustración 20. Normalización de datos con Pandas 7	26
Ilustración 21. Normalización de datos con Pandas 8	26
Ilustración 22. Normalización de datos con Pandas 9	27
Ilustración 23. Creación archivo csv.....	27
Ilustración 24. Configuración Archivos MDF	28
Ilustración 25. Resumen de Archivos MDF.....	28
Ilustración 26. Configuración Data Tools	29
Ilustración 27. Configuración Data Tools 2	30
Ilustración 28. Configuración Data Tools 3	31
Ilustración 29. Configuración Data Tools 4	32
Ilustración 30. Configuración Data Tools 5	32
Ilustración 31. Configuración Data Tools 6	33
Ilustración 32. Configuración Data Tools 7	33
Ilustración 33. Configuración Data Tools 8	34
Ilustración 34. Configuración Data Tools 9	34
Ilustración 35. Configuración Data Tools 10.....	35
Ilustración 36. Configuración Data Tools 11.....	35

Ilustración 37. Normalización de datos con SQL SERVER	36
Ilustración 38. Normalización de datos con SQL SERVER 2	37
Ilustración 39. Validación de datos con SQL de cultivos	37
Ilustración 40. Validación de datos con SQL de clima	38
Ilustración 41. Modelo lógico de bases de datos	38
Ilustración 42. Modelo físico de bases de datos	39
Ilustración 43. Creación usuario para Power BI.....	40
Ilustración 44. Resumen Dashboard	41
Ilustración 45. Producción x Cultivo	42
Ilustración 46. Cultivo por departamento	43
Ilustración 47. Producción por departamento	44
Ilustración 48. Resumen Dashboard	45
Ilustración 49. Resumen Dashboard 1	45
Ilustración 50. Resumen Dashboard 2	46

Resumen

El dashboard es una tecnología muy utilizada en el mundo ya que permite ver cómo está, cómo fue y cómo sería un dato bien procesado, almacenado y gestionado en el proceso de análisis de información.

Es por esto por lo que se implementa un análisis de información expuesta por (MINAGRICULTURA e IDEAM) para verificar como Colombia un país productor y único en el mundo con diferentes climas por regiones o departamentos, puede ser una despensa alimentaria tanto local como internacional.

Es por esto por lo que se quiere analizar los datos de cultivos que se dan en Colombia e identificar el clima que se tuvo en el momento del cultivo, con esto poder evidenciar que la producción agrícola puede ser muy productiva si se hacen los análisis correctos.

Con la experiencia y conocimiento adquirido durante la especialización de gobierno de datos y experiencia técnica y funcional de los integrantes del equipo de investigación, se manejaron diferentes herramientas tecnológicas, metodologías ágiles, se dió de forma satisfactoria el resultado de un dashboard de la información capturada y gestionada.

Palabras claves:

- Dashboard
- Power Bi

Abstract

The dashboard is a technological methodology widely used in the world since it allows you to see how it is, how it was and how it would be if data were well processed, stored and managed in the information analysis process.

Therefore, we want to implement an analysis of information exposed by (MINAGRICULTURA and IDEAM) to verify how Colombia as a producing country and unique in the world with different climates by region or department, can be a food pantry both local and international.

Therefore, we want to analyze the data of crops that occur in Colombia and identify the climate that existed at the time of cultivation, with this to be able to show that agricultural production can be very productive if the correct analyzes are made.

With the experience and knowledge acquired during the data governance specialization and the technical and functional experience of the members of the research team, different technological tools, agile methodologies were used, resulting in a dashboard of the information captured and managed.

Keywords:

- Dashboard
- Power Bi

1. Introducción

Los diferentes departamentos de Colombia cuentan con una gran variedad de pisos térmicos, climas y estaciones climáticas que genera tierras productivas, ciclos de cultivo y productos agrícolas durante todo el año, repartidos en diferentes meses, lo que genera que el país se convierta en una potencia agrícola a nivel mundial en la producción de alimentos. Actualmente el país se encuentra trabajando para ser un referente en cultivos como maíz, papa, yuca, plátano, hortalizas, diferentes frutas y verduras de gran variedad.

Aun así, Colombia cuenta con un nivel de crecimiento y desarrollo de tierras muy bajo, esto hace que la producción de alimentos y el aporte del sector agrícola no genere un impacto alto en (PIB) interno. Un método para apalancar este crecimiento es a través de las Tics, en especial una solución que nos permita realizar seguimientos de los cultivos, su porcentaje de éxito por departamento y municipio y su relación con el clima de la región.

Todo esto apoyándonos en los datos abiertos que se tienen de las entidades gubernamentales del país como lo son el Ministerio de Agricultura y el (IDEAM). Sin embargo, estos datos, aunque disponibles para el público en general, no cuentan con un procesamiento, calidad de información o producto final que sea usado por estudios, entidades u organizaciones.

Es por esto, que como estudiantes de la especialización en gobierno de datos decidimos dirigir nuestros esfuerzos sobre esta línea, con el objetivo de realizar un análisis de los datos expuestos por sistemas públicos, procesando la información y generando conocimiento enriquecido al cruzar los datos de cultivos con los valores meteorológicos. Esto se generaría a través de un tablero o dashboard para el análisis de los datos expuestos por estas dos instituciones, el resultado de este análisis en investigación se puede evidenciar en el entregable final que es el resultado de la siguiente investigación.

2. Descripción y formulación del problema

Colombia es un país único en Sudamérica por su rica geografía, climas e hidrografía, cuya diversidad genera una serie de condiciones climáticas que permite tener cultivo de diferentes alimentos durante todo el año. En consecuencia, Colombia puede ser un país potencia en agricultura global y potencia alimentaria del mundo.

Actualmente el ministerio de agricultura cuenta con información histórica del uso de suelo para agricultura que se ha realizado desde el año 2009 hasta el año 2021, conocido como Evaluaciones Agropecuarias (EVA), que permite identificar la fecha de cultivo, departamento y municipio, el tipo de alimento cultivado, las hectáreas cultivadas y el porcentaje de éxito de esas hectáreas. Así mismo el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) cuenta con información histórica de las condiciones meteorológicas de todas las regiones del país, como temperaturas, presión atmosférica, cantidad de agua lluvia, etc.

Aunque estas bases de datos se encuentran abiertas para el acceso público y existen una serie de reportes realizados por el ministerio de agricultura y el IDEAM, estos no son dinámicos y no permiten aplicar filtros o comparaciones de manera fácil. Así mismo varios reportes se encuentran distribuidos en diferentes páginas gubernamentales, lo que dificulta su interpretación y aumenta los tiempos de consultas.

De igual manera, se ha identificado una oportunidad de mejora, al analizar la data disponible en las fuentes de datos gubernamentales o en la base de datos abiertos, no se encontró una base relacionada entre las variables meteorológicas históricas por municipio y la base (EVA – Evaluaciones Agropecuarias Municipales) en especial mostrando el porcentaje de rendimiento por cultivo. Esto generaría información valiosa para identificar patrones, correlación de eventos y permitiría una mejor toma de decisiones.

El no tener la información centralizada, procesada y lista para análisis en cuadros de mandos o dashboard que permitan identificar data de manera rápida y eficaz dificulta la toma de decisiones, la identificación del mal uso de las tierras fértiles e identificación de zonas históricamente agrícolas que han sido migradas a

la priorización de la ganadería, minería ilícita, cultivos ilícitos o tierras improductivas (sin uso específico).

De acuerdo con lo expresado anteriormente, se identifica el siguiente planteamiento problemático:

¿Cómo mejorar el análisis sobre la efectividad de los cultivos y su relación con el clima a partir de los datos abiertos durante el periodo 2009-2021 en Colombia?

2.1 Objetivo General

Elaborar un dashboard para el análisis sobre la efectividad de los cultivos y su relación con el clima a partir de los datos abiertos del ministerio de agricultura y el IDEAM durante el periodo 2009-2021 en Colombia.

2.2 Objetivos Específicos

- Identificar las fuentes de datos, sus metadatos y las características de la data meteorológica del IDEAM y la data EVA del ministerio de agricultura.
- Generar los modelos lógicos, físicos y diseño de la solución a implementar.
- Implementar los servicios, sistemas y modelos que permitan realizar los procesos de extracción, transformación, cruce de datos y muestra de resultados.
- Implementar un Dashboard con la herramienta Power BI, que visualice el comportamiento de la información histórica y la relación entre la data meteorológica del IDEAM y la data EVA del ministerio de agricultura.

3. Marco referencial

Para lograr el objetivo planteado, se hizo un análisis de las diferentes tecnologías, lineamientos que están disponibles para este fin. Se analizaron sus ventajas y desventajas de tener un proceso de cultivos exitosos y el tipo de clima que se dio en ese momento.

3.1 Marco teórico

3.2 Ministerio de Agricultura

Es una entidad de Colombia la cual se encarga de garantizar políticas agrícolas que garanticen procesos alimentarios estables para el país, por ser una entidad pública. Esta, se encarga de la formulación, coordinación y adopción de las políticas, planes, programas y proyectos del Sector Agropecuario, Pesquero y de Desarrollo Rural.

Esta entidad también se encarga de monitorear el estado de los campos y cultivos a nivel nacional; más los que hace parte de una producción agrícola agroindustrial, ya que por ser un ministerio público garantiza que los insumos para el campo lleguen de manera efectiva y lo menos costosa posible a los agricultores a nivel nacional (República De Colombia Ministerio De Agricultura Y Desarrollo Rural, 1999).

3.3 IDEAM

Es un instituto de carácter público, el cual se encarga de hacer un monitoreo constante de las condiciones climáticas de Colombia, su labor principal es analizar y hacer estudios geográficos que evidencian unos cambios climáticos que puedan afectar al país. Colombia como es un país que no tiene estaciones, es un plus para garantizar cultivos todo el año, pero también es importante tener definido las épocas del año que llueve o hace demasiado calor, ya que con estas definiciones se puede garantizar procesos agrícolas satisfactorios. (Suárez P., 1999).

3.4 EVA

Las evaluaciones agropecuarias municipales (EVA) son la base de información y conocimiento de la oferta productiva agropecuaria de los municipios del país, en este caso en Colombia. EVA es el encargado a nivel Colombia de crear módulos temáticos para la recolecta de productos agrícolas. (*Evaluación Agropecuarias Municipales (EVA)*, 2022).

3.5 SQL Server

Es una base de datos de la casa de tecnología Microsoft, se trabajó con la versión Express Edition, que es la versión gratuita de la misma y por la gran cantidad de procesos que se manejan en muchas empresas es una base de datos muy conocida de Microsoft creado en 1980 de la casa matriz Sybase Inc. ahora SAP, es tan conocido y funcional que muchas de las empresas a nivel mundial utilizan una versión de esta base de datos. (*Wanumen S.*, 2018).

3.6 ETL – SQL Server Integration Service

Es la plataforma integrada con el producto SQL Server de Microsoft, para crear soluciones de integración y transformación de datos a nivel empresarial. Comprende la suite ETL de Microsoft On-premise. Este tipo de procesos de análisis de datos permite migrar, analizar y transformar grandes volúmenes de datos. (*Ladjel B.*, 2017).

3.7 Postman

Es una aplicación creada en el 2012 con la función de hacer eficaz las pruebas de servicios Apis y tener colecciones de manera ordenada, fue y es un programa de tanto éxito que más de 5 millones de personas entre desarrolladores y perfiles encargados de pruebas lo utilizan. Es un gestor de archivos JSON el cual permite analizar el modelo del código, los parámetros con los que están expuestos y cómo almacenarlo en una base de datos. (*Yang W.*, 2022).

3.8 Python

Es un lenguaje de programación creado en 1989 con el fin de solucionar problemas y tener opciones de código abierto para diferentes procesos como soluciones de capas de datos y soluciones web, este lenguaje ha tenido tanta evolución hasta nuestros días que es uno de los principales para el manejo de datos, transformación y procesamiento de datos que otros lenguajes no soportan. Es un lenguaje de programación libre, el cual es muy utilizado para la captura y transformación de datos, este permite un proceso ágil en el consumo de las APIS por los framework que ya están en la red para este proceso. (Michael A., 2011).

3.9 Docker

Es una aplicación creada en el año 2000 con la idea de tener servicios en la nube llamados servicios de contenedores, el cual es tener un servidor configurado listo para la gestión, con el fin de ayudar a los perfiles de infraestructura y desarrolladores de software de tener una imagen y ponerla productiva en tiempos muchos más cortos. Este es una aplicación de despliegue de aplicaciones muy ágil y de fácil despliegue, además por su gran servicio de tener bases de datos en contenedores nos garantiza que la base de datos de SQL server este en la nube y con esto no tenemos que instalar software sobre los equipos, ya que son software muy pesados con solo tener el cliente de SQL server podemos conectar al contenedor maestro y verificar los datos guardados que están sobre la base de datos. (Anderson C., 2015).

3.10 Power BI

Es un sistema de análisis de datos que permite la creación de informes de manera profesional creado en el año 2006, con el fin de hacer analítica de datos demasiados grandes que programas como Excel no pueden procesar, el éxito de este software ha sido tan grande que muchas organizaciones en el mundo lo utilizan para tener sus informes comerciales, financieros, entre otros, a la mano, de manera mucho más rápida, su gran calidad de análisis es debido a su función de conectarse a varios sistemas de servidores de almacenamiento de datos. Es un gestor de datos el cual permitió hacer informes muy detallados y de gran visibilidad para las personas que desean ver el resultado del dashboard, este gestor se conecta a SQL

server y genera el informe sobre las tablas que se definan, asegurando un proceso rápido y eficaz para la creación del dashboard. (Clark D., 2017).

4. Estado del arte

Se hizo el análisis de las diferentes fuentes, base de datos de internet y sitios gubernamentales para evidenciar proyectos, investigaciones o resultados similares al planteado, que a continuación se expondrán. El objetivo es identificar trabajos similares con datos históricos de áreas cosechadas (no solo del territorio colombiano) por los diferentes cultivos .

De las fuentes consultadas, no se evidencio un estudio, producto final o investigación que tenga un resultado similar al planteado en el proyecto. Se han evidenciado estudios sobre el impacto del cambio climático en los cultivos que se están produciendo, pero no está enfocada a la producción histórica por región específica.

Uno de estos estudios realizado por la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) indica cómo el cambio climático está afectando la producción de alimentos.

Se evidencio un análisis de cultivos en relación con las variables meteorológicas, es el sistema integrado de monitoreo (SIMA). Este es un software de pago, el cual lleva un control de la producción agrícola, pero este software está enfocado a ser utilizado por grandes terratenientes o empresas con extensiones de gran área de cosecha. Adicional por el costo del licenciamiento, no puede ser accedido fácilmente por pequeños y medianos agricultores. La data que genera la aplicación permite llevar un control del área cosechada, analizando diferentes variables meteorológicas y generando reportes históricos.

Esta aplicación requiere un conocimiento técnico alto, tanto en conocimiento agrícola, meteorológico, electrónico (sensores meteorológicos) y de manejo de sistemas de información, lo que limita el grupo de personas que puedan acceder al software. (SIMA, 2022)

Para el territorio colombiano, se identificaron reportes propios de las entidades gubernamentales, como el ministerio de agricultura a través de la unidad de planeación de red agropecuaria (UPRA), donde exponen los resultados de las evaluaciones agropecuarias municipales (EVA), comparando las cosechas vs años anteriores.

Esto se genera en un informe detallado, que se sube a las páginas gubernamentales (EVA,2022). Los informes van acompañados de gráficas de

rendimiento y distribución de cultivos, pero en un reporte plano, que no permite aplicar filtros o explorar la data de manera dinámica. De igual manera, la data está agrupada a nivel nacional y no muestra resultados por zonas, departamentos o municipios.

Por último, esta data no está enriquecida con otras variables como información meteorológica de las zonas, eventos sociales o políticos del periodo de tiempo analizado.

Un estudio similar es realizado por el Departamento Administrativo Nacional de Estadística (DANE), donde se expone el censo realizado de los cultivos que tiene el país, pero, igual al anterior, no está enriquecido con variables climáticas, políticas o sociales.

4.1 Impacto

Como resultado del dashboard se espera tener una idea más clara sobre la soberanía alimentaria del país, las mejores regiones de agricultura y el mejor clima donde este producto se dio, ya que con esto se puede garantizar un análisis de cómo el país es productivamente efectivo.

Otro resultado que se quiere tener es que se pueda trabajar con los datos que se tienen desde los diferentes orígenes (MINAGRICULTURA e IDEAM) es que el dato sea una forma de mejorar muchos de los procesos que se tienen estipulados para el tema de agricultura.

Por último, se espera tener un impacto social, al permitir que esta información sea de acceso público, para que pueda ser accedido por cualquier persona.

4.2 Componente de Innovación

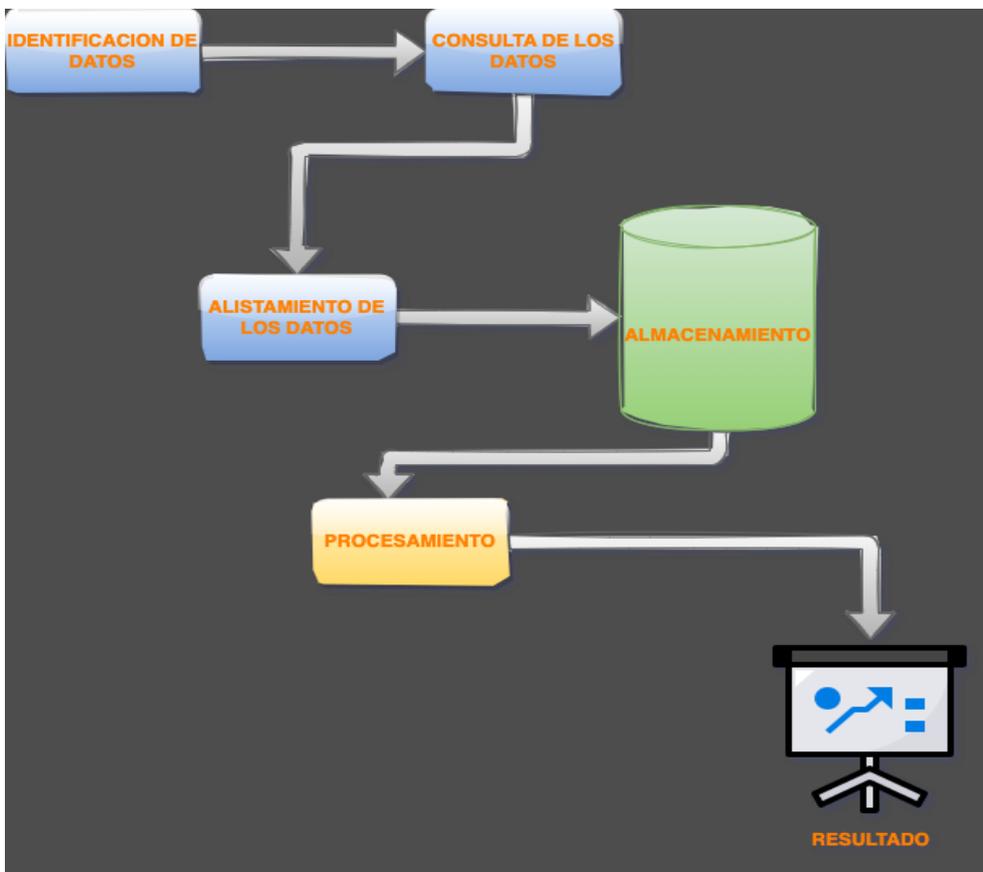
El dashboard que se propone es el resultado de la investigación y análisis de las fuentes de datos origen (MINAGRICULTURA e IDEAM), donde se podrá evidenciar un resultado como una vista la cual se pueda ver la región o departamento donde se cosechó y el tipo de clima que se tuvo en su momento.

5. Metodología

La siguiente metodología fue la utilizada dentro del proceso de investigación, análisis y desarrolló para dar como resultado el dashboard que permite analizar los datos de resultados agrícolas históricos del país y los datos meteorológicos, históricos, del IDEAM. Durante el desarrollo del proyecto, se definió un flujo de datos, en el cual se identifica la captura, procesamiento, almacenamiento y visualización de los datos de manera efectiva en un dashboard.

Basados en el DAMA-DMBOK (*Data Management (DAMA)*, 2022) en la metodología de proceso de un buen gobierno de datos y normalización de los datos, se identificó un modelo para el proceso de creación del dashboard, el cual se muestra en el siguiente gráfico de proceso:

Ilustración 1. Metodología para el procesamiento de los datos



Fuente: Elaboración Propia

A continuación, se describen las etapas de la metodología.

Identificación de datos: en esta etapa se define el proceso de cómo buscar los datos de una manera efectiva en los sistemas origen.

Consulta De Los Datos: ya identificados los datos y las URL de las entidades estatales que garantizan los datos que se necesita para el desarrollo del dashboard, se definió la tecnología de consultar los datos mediante peticiones de consumo de servicios este proceso fue resuelto bajo tecnologías de petición y consumo de servicios.

Alistamiento de datos: en esta etapa teniendo los datos consultados y garantizando las peticiones a los servicios de las entidades del gobierno, se definió un modelo de alistamiento de datos para guardarlos en una base de datos relacional.

Almacenamiento: ya el proceso de alistamiento como es efectivo los datos quedarán almacenados en un software de Microsoft que es muy efectivo para modelos relacionales, con este se garantiza que los datos queden guardados y seguros y tener la seguridad que no se vayan a perder en caso de que los sistemas que los tiene expuestos sufran una caída.

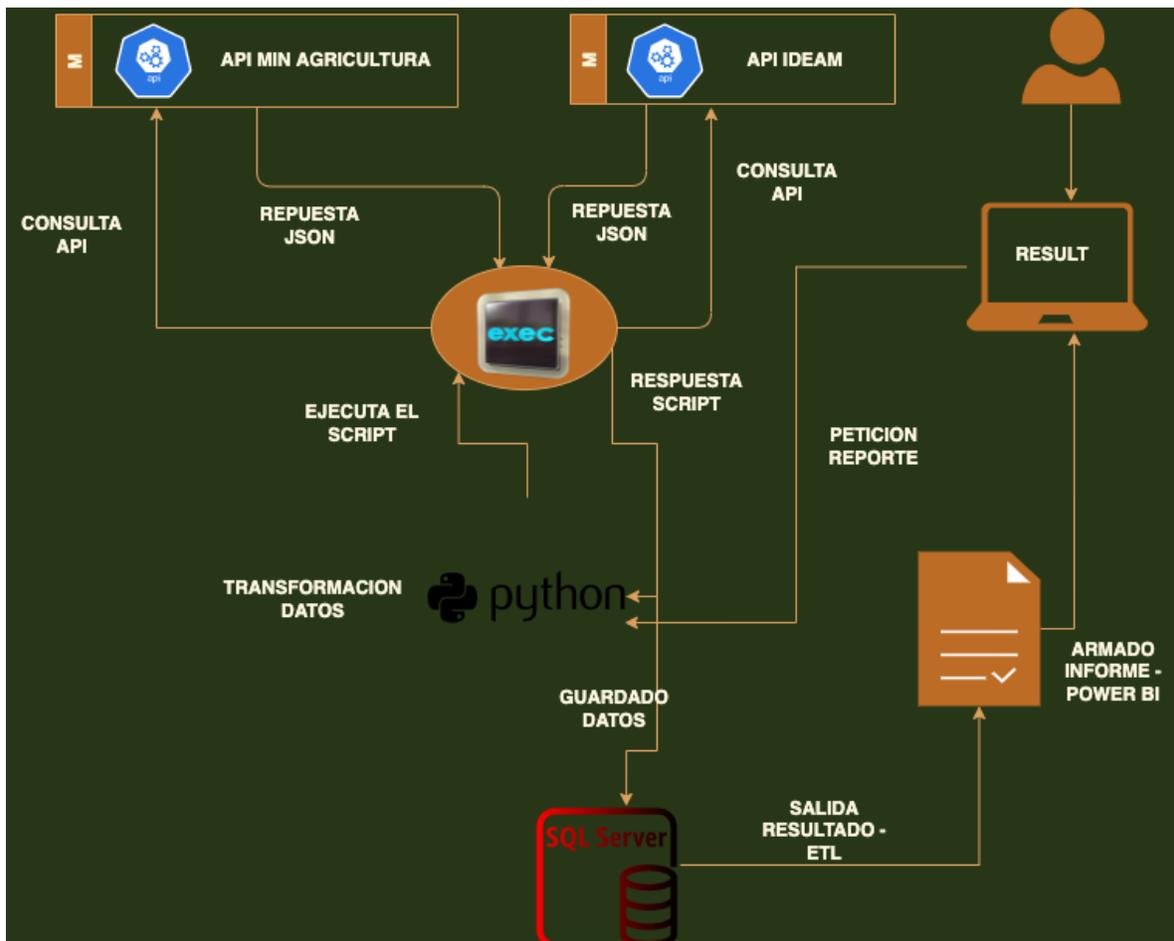
Procesamiento: esta etapa es una de las principales del proceso, en esta como ya se tienen los datos almacenados en una base de datos, se hace un análisis y transformación de los datos para garantizar que el dashboard muestre los datos limpios y las gráficas sean las más profesionales posibles.

Resultado: en la etapa de resultado ya es el proceso final donde se puede evidenciar el dashboard terminado, con esto se puede ver el proceso final del dato.

6. Desarrollo de la propuesta

Al aplicar la metodología planteada, se identificó un modelo de datos que ejemplifica el flujo de información desde su origen al proceso final. En este modelo se identifican las fuentes, el método de consumo de los datos, el proceso de depuración y almacenamiento y tratamiento para ser llevado a la herramienta de explotación de datos para el entregable final. Este proceso se define en el siguiente diagrama (Ver imagen 2). Las generalidades del proceso se expondrán a continuación:

Ilustración 2. Diagrama de Flujo de datos



Fuente: Desarrollo Drawio

6.1 Identificación de Datos

Dentro de la investigación, lectura de trabajos anteriores y verificación de fuentes, se identificó que los diferentes ministerios y organizaciones gubernamentales por la Ley 1712 de 2014, ley de datos abiertos, deben permitir el acceso a la información pública dispuesta en formatos que permiten su uso y reutilización bajo licencia abierta y sin restricciones legales para su aprovechamiento. Esto nos permite acceder a una gran cantidad de conjuntos de datos, que son del alcance y objetivo del proyecto.

A continuación, se describe el origen de los datos y sus características.

Datos EVA (Evaluaciones Agropecuarias Municipales)

De la página de datos abiertos, se encontraron el histórico de las evaluaciones realizadas por el ministerio, desde el año 2019 al año 2022. Estas se encuentran divididas en dos fuentes de años distintos, de 2019 a 2020 y de 2007 a 2018.

Esta información se encuentra en (Datos Abiertos, 2022). Allí se identifican un pequeño resumen del conjunto de datos, un diccionario de datos básico. Este data set se compone de 18 columnas y con un tamaño de más de 230K filas cuando ambos sets de datos están unificados.

Dentro de la información que trae el set de datos, se identifica los departamentos y municipios donde se realizó la medición. Se encuentra una clasificación y dos subclasificaciones de la cosecha y por último se identifica las hectáreas plantadas, cultivadas y el porcentaje de efectividad de dicha cosecha en el periodo de tiempo.

En este último se encuentran alimentos con periodo de tiempo de semestre o año.

De los datos de IDEAM se identificaron datos de Precipitación diaria, temperatura diaria, presión atmosférica y calidad del aire. Todos estos por municipio y departamento y con mediciones diarias y en algunos casos por hora. Como se evidencia en (Datos Abiertos, 2022). Las cuatro fuentes comparten una misma estructura de 12 columnas. Las fuentes cuentan con un tamaño de más de 800 millones de filas, debido a que son mediciones históricas desde el año 2005 a la fecha, en algunos casos de mediciones por hora. Para finalizar el análisis, se identificó que los campos más relevantes para el proyecto son Departamento y municipio de medición, valor observado y la unidad de medida del valor.

Por último, en cada una de las fuentes nos especifican el método de consumo para extraer la data por medio de un JSON o un CSV, a través del consumo de un API.

[Evaluaciones Agropecuarias Municipales EVA – Histórico 2007 a 2018](#)

- <https://www.datos.gov.co/resource/2pnw-mmge.json>

[Evaluaciones-Agropecuarias-Municipales-EVA 2019 a 2021](#)

- <https://www.datos.gov.co/resource/p5fp-pay3.json>

[Datos Hidrometeorológicos Crudos - Red de Estaciones IDEAM: Temperatura](#)

- <https://www.datos.gov.co/resource/sbwg-7ju4.json>

[Presión Atmosférica](#)

- <https://www.datos.gov.co/resource/62tk-nxj5.json>

[Precipitación](#)

- <https://www.datos.gov.co/resource/62tk-nxj5.json>

6.2 Consulta De Los Datos

Como resultado de la consulta de las APIS se puede evidenciar que arroja un resultado satisfactorio (Ver imagen 3 y 4).

Ilustración 3. Petición Servicio Api Web

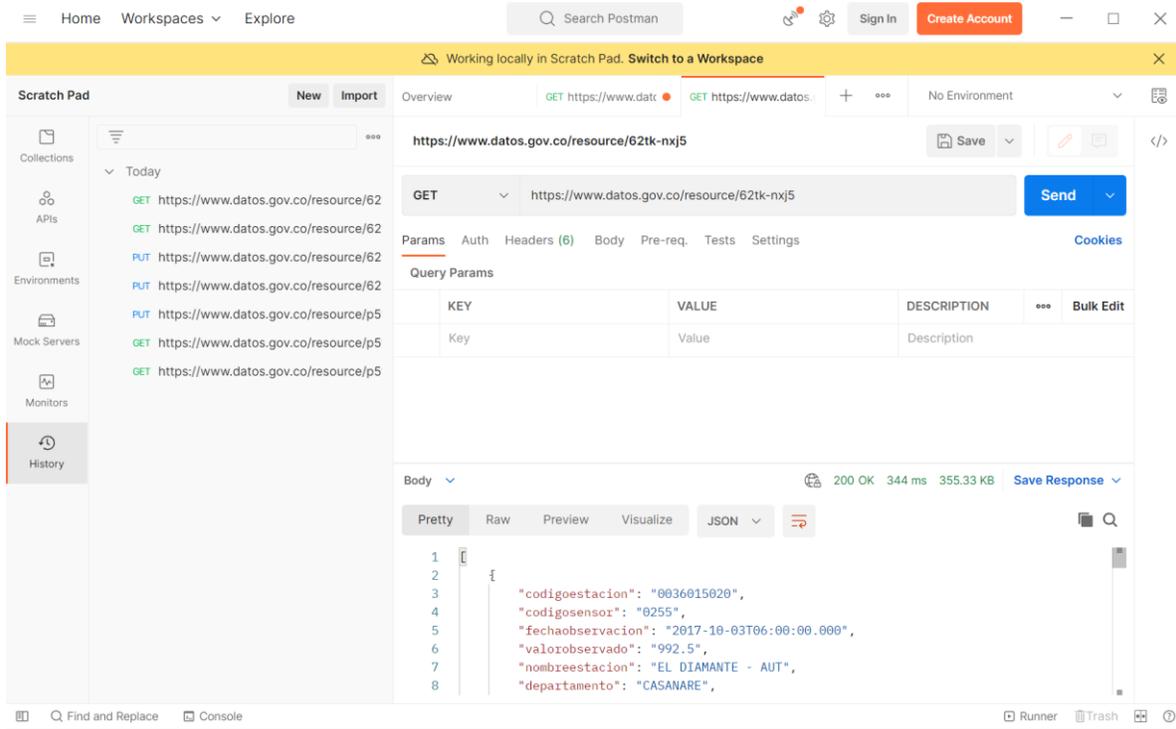
```

[{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Leguminosas Y Oleaginosas", "subgrupo_cultivo_seg_n_especie": "Leguminosas", "cultivo": "Frijol", "a_o": "2019", "periodo": "2019B", "area_semrada_ha": "110", "area_cosechada_ha": "105", "producci_n_t": "157.5", "rendimiento_t_ha": "1.5", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Grano Seco", "nombre_cient_fico": "Phaseolus Vulgaris"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Frutales", "subgrupo_cultivo_seg_n_especie": "Mora", "cultivo": "Mora", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "6", "area_cosechada_ha": "6", "producci_n_t": "48", "rendimiento_t_ha": "8", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Fruto Fresco", "nombre_cient_fico": "Rubus Glaucus"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Frutales", "subgrupo_cultivo_seg_n_especie": "Solanáceas", "cultivo": "Lulo", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "5", "area_cosechada_ha": "5", "producci_n_t": "25", "rendimiento_t_ha": "5", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Fruto Fresco", "nombre_cient_fico": "Solanum Quitoense"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Frutales", "subgrupo_cultivo_seg_n_especie": "Solanáceas", "cultivo": "Tomate de árbol", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "5", "area_cosechada_ha": "5", "producci_n_t": "125", "rendimiento_t_ha": "25", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Fruto Fresco", "nombre_cient_fico": "Solanum Betaceum"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Frutales", "subgrupo_cultivo_seg_n_especie": "Mango", "cultivo": "Mango", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "20", "area_cosechada_ha": "20", "producci_n_t": "120", "rendimiento_t_ha": "6", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Fruto Fresco", "nombre_cient_fico": "Mangifera Indica"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Leguminosas Y Oleaginosas", "subgrupo_cultivo_seg_n_especie": "Leguminosas", "cultivo": "Frijol", "a_o": "2019", "periodo": "2019A", "area_semrada_ha": "115", "area_cosechada_ha": "110", "producci_n_t": "165", "rendimiento_t_ha": "1.5", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Grano Seco", "nombre_cient_fico": "Phaseolus Vulgaris"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Plátanos", "subgrupo_cultivo_seg_n_especie": "Plátanos", "cultivo": "Plátano", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "2036.5", "area_cosechada_ha": "2036", "producci_n_t": "9180", "rendimiento_t_ha": "4.51", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Fruto Fresco", "nombre_cient_fico": "Musa Spp."},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Tubérculos Y Plátanos", "subgrupo_cultivo_seg_n_especie": "Tubérculos", "cultivo": "Papa", "a_o": "2019", "periodo": "2019A", "area_semrada_ha": "95", "area_cosechada_ha": "90", "producci_n_t": "900", "rendimiento_t_ha": "10", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Tubérculo Fresco", "nombre_cient_fico": "Solanum Tuberosum"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Tubérculos Y Plátanos", "subgrupo_cultivo_seg_n_especie": "Tubérculos", "cultivo": "Papa", "a_o": "2019", "periodo": "2019B", "area_semrada_ha": "80", "area_cosechada_ha": "75", "producci_n_t": "750", "rendimiento_t_ha": "10", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Tubérculo Fresco", "nombre_cient_fico": "Solanum Tuberosum"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Cereales", "subgrupo_cultivo_seg_n_especie": "Maíz", "cultivo": "Maíz", "a_o": "2019", "periodo": "2019A", "area_semrada_ha": "105", "area_cosechada_ha": "100", "producci_n_t": "200", "rendimiento_t_ha": "2", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Grano Seco", "nombre_cient_fico": "Zea Mays"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Cereales", "subgrupo_cultivo_seg_n_especie": "Maíz", "cultivo": "Maíz", "a_o": "2019", "periodo": "2019B", "area_semrada_ha": "80", "area_cosechada_ha": "75", "producci_n_t": "150", "rendimiento_t_ha": "2", "ciclo_del_cultivo": "Transitorio", "estado_f_sico_de_la_producci": "Grano Seco", "nombre_cient_fico": "Zea Mays"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Cultivos Tropicales Tradicionales", "subgrupo_cultivo_seg_n_especie": "Cacao", "cultivo": "Cacao", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "15", "area_cosechada_ha": "5", "producci_n_t": "2", "rendimiento_t_ha": "0.4", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Grano Seco", "nombre_cient_fico": "Theobroma Cacao"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Cultivos Tropicales Tradicionales", "subgrupo_cultivo_seg_n_especie": "Café", "cultivo": "Café", "a_o": "2019", "periodo": "2019", "area_semrada_ha": "2978", "area_cosechada_ha": "2326", "producci_n_t": "3093.58", "rendimiento_t_ha": "1.33", "ciclo_del_cultivo": "Permanente", "estado_f_sico_de_la_producci": "Café Verde Equivalente", "nombre_cient_fico": "Coffea Arabica"},
{"codigo_del_departamento": "5", "departamento": "Antioquia", "c_digo_del_municipio": "5002", "municipio": "Abejorral", "grupo_cultivo_seg_n_especie": "Cultivos Tropicales

```

Fuente: Datos.gov.co

Ilustración 4. Petición Servicio Api por Postman



Fuente: Postman

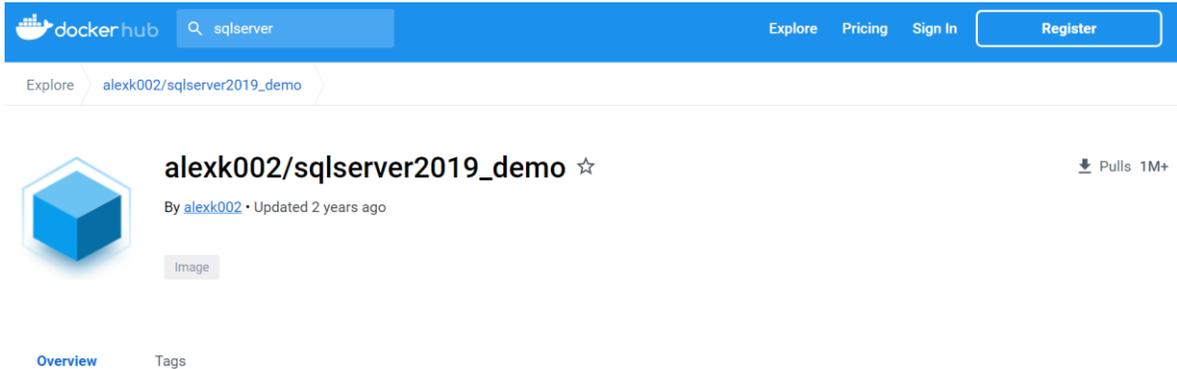
6.3 Alistamiento de datos

Se realiza un preprocesamiento de la data con un script en Python en el archivo `PreparacionDatos.ipynb` se encuentra el análisis realizado a los datos de las Apis y sus tipos de datos.

6.4 Almacenamiento

El almacenamiento de los datos se realizará en una base de datos de SQL Server, la cual está en un contenedor de Docker (ver imagen 5).

Ilustración 5. Imagen Docker SQL SERVER



Fuente: Hub Docker

6.5 Procesamiento

Al revisar las fuentes orígenes se evidencia que vienen en un formato de texto sencillo como lo es JSON, el cual permite evidenciar que significa cada uno de los parámetros y el significado de este (Ver imagen 6 y 7).

Ilustración 6. Detalle campos Api

Columnas en este Conjunto de Datos

Nombre de Columna	Descripción	Tipo
CÓD. DEP.	Código del departamento, según lo establecido por el DANE	Número #
DEPARTAMENTO	Departamento Colombiano	Texto simple T
CÓD. MUN.	Código del municipio, según lo establecido por el DANE	Número #
MUNICIPIO	Municipio Colombiano	Texto simple T
GRUPO DE CULTIVO	Categoría del cultivo	Texto simple T
SUBGRUPO DE CULTIVO	Tipo de cultivo según categoría	Texto simple T
CULTIVO	Nombre del cultivo	Texto simple T
DESAGREGACIÓN REGIONAL Y/O SISTEMA PRODUCTIVO	Nombre generico del cultivo	Texto simple T
AÑO	Año de producción	Número #

Fuente: Framework Python Pandas

Ilustración 7. Detalle campos Api 2

PERIODO	Periodo medico	Texto simple T	▼
Área Sembrada (ha)	Area sembrada en hectáreas	Número #	▼
Área Cosechada (ha)	Area cosechada en hectáreas	Número #	▼
Producción (t)	Tiempo de producción	Número #	▼
Rendimiento (t/ha)	Rendimiento de la cosecha	Número #	▼
ESTADO FISICO PRODUCCION	Estado del producto	Texto simple T	▼
NOMBRE CIENTIFICO	Nombre científico del cultivo	Texto simple T	▼
CICLO DE CULTIVO	Ciclo del cultivo en el país	Texto simple T	▼

Fuente: Desarrollo Python Pandas

También se logra identificar el contenido de los datos expuestos, si se tiene datos duplicados, nulos entre otros (Ver imagen 8 y 9).

Ilustración 8. Resumen Datos de Origen

Pandas Profiling Report Overview Variables Interactions Correlations Missing values

Overview

Dataset statistics		Variable types	
Number of variables	17	CAT	9
Number of observations	246013	NUM	8
Missing cells	6291		
Missing cells (%)	0.2%		
Duplicate rows	0		
Duplicate rows (%)	0.0%		
Total size in memory	155.5 MIB		
Average record size in memory	662.9 B		

Fuente: Framework Python Pandas

Ilustración 9. Definición datos de Origen

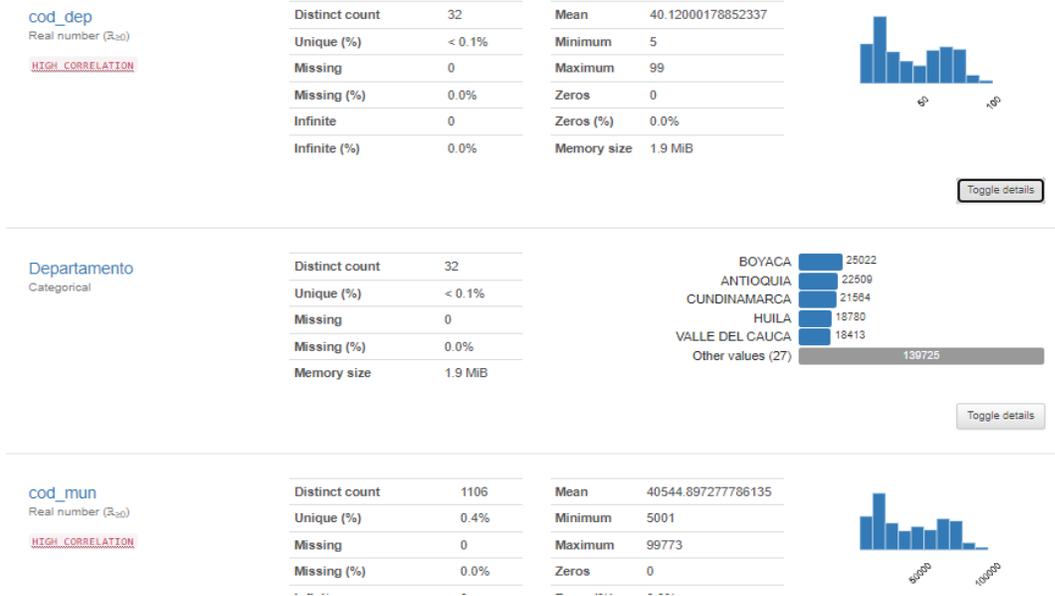
Overview

Overview	Reproduction	Warnings 13
<code>municipio</code> has a high cardinality: 1072 distinct values		High cardinality
<code>subgrup_cultivo</code> has a high cardinality: 132 distinct values		High cardinality
<code>Cultivo</code> has a high cardinality: 225 distinct values		High cardinality
<code>nombre_cientifico</code> has a high cardinality: 215 distinct values		High cardinality
<code>cod_mun</code> is highly correlated with <code>cod_dep</code>		High correlation
<code>cod_dep</code> is highly correlated with <code>cod_mun</code>		High correlation
<code>Area_Cosechada</code> is highly correlated with <code>Area_Sembrada</code>		High correlation
<code>Area_Sembrada</code> is highly correlated with <code>Area_Cosechada</code>		High correlation
<code>Rendimiento</code> has 3433 (1.4%) missing values		Missing
<code>nombre_cientifico</code> has 2857 (1.2%) missing values		Missing
<code>Produccion</code> is highly skewed ($\gamma_1 = 61.27679096$)		Skewed
<code>Area_Cosechada</code> has 5065 (2.1%) zeros		Zeros
<code>Produccion</code> has 4544 (1.8%) zeros		Zeros

Fuente: Framework Python Pandas

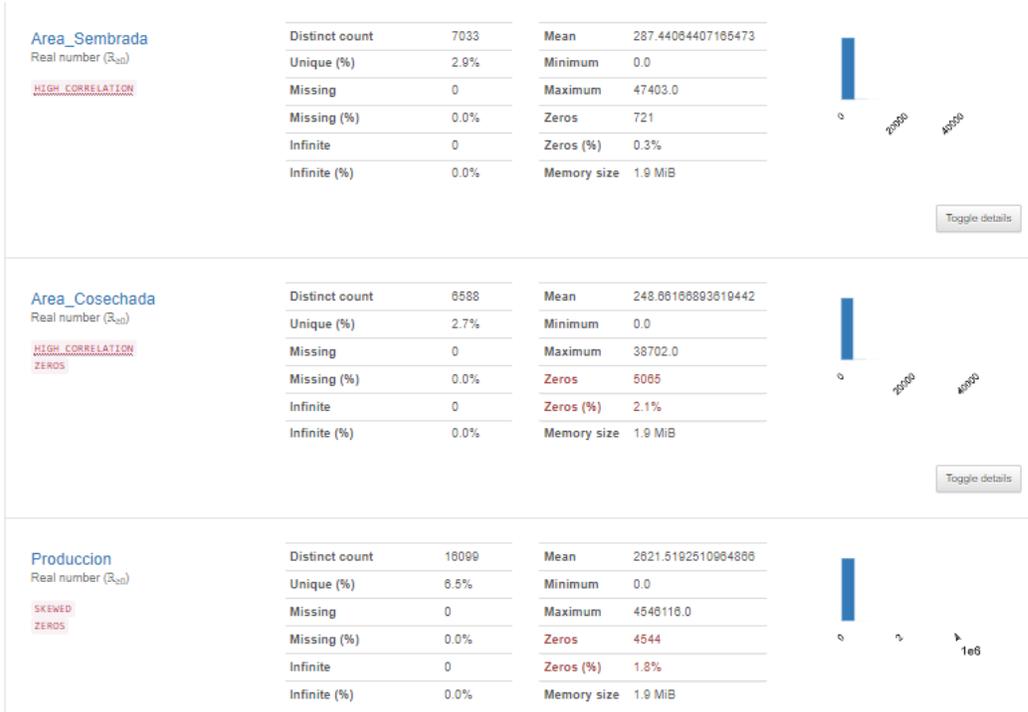
Al tener una idea más completa de cómo están los datos, mediante el framework de Python pandas, se realiza un análisis de los datos con el fin de validar la calidad de los datos expuestos por el ministerio de agricultura e IDEAM, esto da un apoyo para entender mejor el proceso que se debe tener encuentra en él un dashboard completo (Ver imagen 10 y 11).

Ilustración 10. Normalización de variables



Fuente: Framework Python Pandas

Ilustración 11. Normalización Variables 2



Fuente: Framework Python Pandas

Se identificó que la data del IDEAM tiene un tamaño superior a las 60 Millones de registros en su información histórica. Dado que trae datos por minutos y segundos (Ver imagen 12).

Ilustración 12. Resumen Apis IDEAM

Vista previa de la tabla

Explorar los datos

Crear visualización

Codig...	Codl...	FechaObservacion ↑	Valor...	Nom...	Depa...	Muni...	Zona...	Latit...	Longl...	Desc...	Ur
00261353...	0068	2001 Jan 01 12:19:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	CodigoEstacion	2001 Jan 01 11:59:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 01 11:39:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 02 11:19:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 02 10:59:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 03 10:39:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 03 10:19:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 04 09:59:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 04 09:39:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 05 09:19:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 05 08:59:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 06 08:39:05 AM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C
00261353...	0068	2001 Jan 06 08:19:05 PM	0	PNN QUI...	RISARALDA	PEREIRA	CAUCA	4,761916...	-75,74491...	Temp Air...	°C

Fuente: Framework Python Pandas

Se realizará el cargue de la información a través de una ETL en SQL Server Integration Server, a través de las herramientas SSDT - SQL Server Data Tools.

El proceso por realizar es:

- 1) Descargar los datos desde un script de Python que consulte el API.
- 2) En el mismo script, se realizará depuración de información y preprocesamiento + transformación de datos.
- 3) Se descarga a CSV.
- 4) Este archivo servirá de repositorio ante cualquier solicitud.
- 5) Se cargará el CSV a través de las herramientas Data Flow a la base de datos del ambiente Stage.
- 6) De allí se procesan los registros para dejarlos listos para ser consultados por el Power BI.

Como ya se tienen definidos y relacionados las Apis donde están expuesto los datos que se necesita para el dashboard, desde Python se hace el relacionamiento de estos para que quede como un proceso automático y relacionado de donde tiene que hacer la petición para traer la data y procesarla (Ver imagen 13).

Ilustración 13. Validación URL en Python

```
#Librerias a utilizar
import pandas as pd
from sodapy import Socrata
import datetime
from datetime import datetime

#Variables de API
url = 'https://www.datos.gov.co/resource/uejq-wxrr.csv'
client = Socrata("www.datos.gov.co", None)

#Sabemos la cantidad de registros
record_count = client.get("uejq-wxrr", select='COUNT(*)')

#variables para bajar pedazos pequeños y no saturar el server
start = 0
chunk_size = 1500
results = []
while True:
    #descargamos por pedazos o chunks
    results.extend(client.get("uejq-wxrr", limit=chunk_size))
    start = start+chunk_size
    if (start > int(record_count[0]['COUNT'])):
        #if (start > record_count):
            break
#Convertimos a un dataframe
df=pd.DataFrame(results)
#Variables para repositorio archivo
fecha = datetime.today().strftime('%Y%m%d')
ruta = "C:\\UAN\\Repositorio\\EVA\\EVAHist_"
nombre_archivo = ruta + fecha + ".csv"

#Bajamos a csv
df.to_csv(nombre_archivo)
```

Fuente: IDE Python

Después de haber consumido las Apis de los datos de los sistemas origen se hace una normalización y procesamiento de los datos con el framework de Python llamado pandas, este permite que los datos sean exportados en un formato plano listos para cargarlos a una base de datos en este caso SQL SERVER (Ver imagen 14, 15, 16 y 17).

Ilustración 14. Normalización de datos con Pandas

Revisamos nombre y orden de columnas

[1] dfEVA.columns

```
Index(['Código del departamento', 'Departamento', 'Código del Municipio',
      'Municipio', 'Grupo cultivo según especie',
      'Subgrupo cultivo según especie', 'Cultivo', 'Año', 'Periodo',
      'Área Sembrada (ha)', 'Área Cosechada (ha)', 'Producción (t)',
      'Rendimiento (t/ha)', 'Ciclo del cultivo',
      'Estado físico de la producción', 'Nombre científico'],
      dtype='object')
```

```
[ ] # change column name
dfEVA = dfEVA.rename(columns={"Código del departamento": "cod_dep"})
dfEVA = dfEVA.rename(columns={"Código del Municipio": "cod_mun"})
dfEVA = dfEVA.rename(columns={"Grupo cultivo según especie": "grup_cultivo"})
dfEVA = dfEVA.rename(columns={"Subgrupo cultivo según especie": "subgrup_cultivo"})
dfEVA = dfEVA.rename(columns={"a_o": "año"})
dfEVA = dfEVA.rename(columns={"Estado físico de la producción": "estado_produccion"})
dfEVA = dfEVA.rename(columns={"Nombre científico": "nombre_cientifico"})

dfEVA = dfEVA.rename(columns={"Área Sembrada (ha)": "Área_Sembrada"})
dfEVA = dfEVA.rename(columns={"Área Cosechada (ha)": "Área_Cosechada"})
dfEVA = dfEVA.rename(columns={"Producción (t)": "Produccion"})
dfEVA = dfEVA.rename(columns={"Rendimiento (t/ha)": "Rendimiento"})
dfEVA = dfEVA.rename(columns={"Ciclo del cultivo": "Ciclo_cultivo"})
```

Fuente: Framework Python Pandas

Ilustración 15. Normalización de datos con Pandas 2

[] dfEVA.columns

```
Index(['cod_dep', 'Departamento', 'cod_mun', 'Municipio', 'grup_cultivo',
      'subgrup_cultivo', 'Cultivo', 'Año', 'Periodo', 'Área_Sembrada',
      'Área_Cosechada', 'Produccion', 'Rendimiento', 'Ciclo_cultivo',
      'estado_produccion', 'nombre_cientifico'],
      dtype='object')
```

[] dfEVAHist.columns

```
Index(['CÓD. \nDEP.', 'DEPARTAMENTO', 'CÓD. MUN.', 'MUNICIPIO',
      'GRUPO \nDE CULTIVO', 'SUBGRUPO \nDE CULTIVO', 'CULTIVO',
      'DESAGREGACIÓN REGIONAL Y/O SISTEMA PRODUCTIVO', 'AÑO', 'PERIODO',
      'Área Sembrada\n(ha)', 'Área Cosechada\n(ha)', 'Producción\n(t)',
      'Rendimiento\n(t/ha)', 'ESTADO FISICO PRODUCCION',
      'NOMBRE \nCIENTIFICO', 'CICLO DE CULTIVO'],
      dtype='object')
```

Fuente: Framework Python Pandas

Ilustración 16. Normalización de datos con Pandas 3

```

dFEVAHist.columns
Index(['CÓD. \nDEP.', 'DEPARTAMENTO', 'CÓD. MUN.', 'MUNICIPIO',
      'GRUPO \nDE CULTIVO', 'SUBGRUPO \nDE CULTIVO', 'CULTIVO',
      'DESAGREGACIÓN REGIONAL Y/O SISTEMA PRODUCTIVO', 'AÑO', 'PERIODO',
      'Área Sembrada\n(ha)', 'Área Cosechada\n(ha)', 'Producción\n(t)',
      'Rendimiento\n(t/ha)', 'ESTADO FÍSICO PRODUCCION',
      'NOMBRE \nCIENTIFICO', 'CICLO DE CULTIVO'],
      dtype='object')

[ ] dFEVAHist = dFEVAHist.rename(columns={"CÓD. \nDEP.": "cod_dep"})
dFEVAHist = dFEVAHist.rename(columns={"DEPARTAMENTO": "Departamento"})
dFEVAHist = dFEVAHist.rename(columns={"CÓD. MUN.": "cod_mun"})
dFEVAHist = dFEVAHist.rename(columns={"MUNICIPIO": "Municipio"})
dFEVAHist = dFEVAHist.rename(columns={"GRUPO \nDE CULTIVO": "grup_cultivo"})
dFEVAHist = dFEVAHist.rename(columns={"SUBGRUPO \nDE CULTIVO": "subgrup_cultivo"})
dFEVAHist = dFEVAHist.rename(columns={"CULTIVO": "Cultivo"})
dFEVAHist = dFEVAHist.rename(columns={"DESAGREGACIÓN REGIONAL Y/O SISTEMA PRODUCTIVO": "desagregacion_regional"})
dFEVAHist = dFEVAHist.rename(columns={"AÑO": "Año"})
dFEVAHist = dFEVAHist.rename(columns={"PERIODO": "Periodo"})
dFEVAHist = dFEVAHist.rename(columns={"Área Sembrada\n(ha)": "Area_Sembrada"})
dFEVAHist = dFEVAHist.rename(columns={"Área Cosechada\n(ha)": "Area_Cosechada"})
dFEVAHist = dFEVAHist.rename(columns={"Producción\n(t)": "Produccion"})
dFEVAHist = dFEVAHist.rename(columns={"Rendimiento\n(t/ha)": "Rendimiento"})
dFEVAHist = dFEVAHist.rename(columns={"ESTADO FÍSICO PRODUCCION": "estado_produccion"})
dFEVAHist = dFEVAHist.rename(columns={"NOMBRE \nCIENTIFICO": "nombre_cientifico"})
dFEVAHist = dFEVAHist.rename(columns={"CICLO DE CULTIVO": "Ciclo_cultivo"})

```

Fuente: Framework Python Pandas

Ilustración 17. Normalización de datos con Pandas 4

```

Final EVA
cod_dep  Departamento  cod_mun  Municipio  grup_cultivo  subgrup_cultivo  Cultivo  Año  Periodo  Area_Sembrada  Area_Cosechada  Produccion  Rendimiento  estado_produccion  nombre_cientifico  Ciclo_cultivo
0      5      Antioquia    5002  Abejorral  Leguminosas Y Oleaginosas  Leguminosas  Frijol  2019  2019B    110.0      105.0      157.5      1.5      Grano Seco  Phaseolus Vulgaris  Transitorio
1      5      Antioquia    5002  Abejorral  Frutales  Mora  Mora  2019  2019    6.0      6.0      48.0      8.0      Fruto Fresco  Rubus Glaucus  Permanente
2      5      Antioquia    5002  Abejorral  Frutales  Solanáceas  Lulo  2019  2019    5.0      5.0      25.0      5.0      Fruto Fresco  Solanum Quitoense  Permanente
3      5      Antioquia    5002  Abejorral  Frutales  Solanáceas  Tomate de árbol  2019  2019    5.0      5.0      125.0      25.0      Fruto Fresco  Solanum Betaceum  Permanente
4      5      Antioquia    5002  Abejorral  Frutales  Mango  Mango  2019  2019    20.0      20.0      120.0      6.0      Fruto Fresco  Mangifera Indica  Permanente

[ ] print("Final EVAHist")
dFEVAHist.head()

Final EVAHist
cod_dep  Departamento  cod_mun  Municipio  grup_cultivo  subgrup_cultivo  Cultivo  desagregacion_regional  Año  Periodo  Area_Sembrada  Area_Cosechada  Produccion  Rendimiento  estado_produccion  nombre_cientifico  Ciclo_cultivo
0      15      BOYACA    15114  BUSBANZA  HORTALIZAS  ACELGA  ACELGA  ACELGA  2006  2006B    2      1      1      1.00  FRUTO FRESCO  BETA VULGARIS  TRANSITORIC
1      25      CUNDINAMARCA  25754  SOACHA  HORTALIZAS  ACELGA  ACELGA  ACELGA  2006  2006B    82      80      1440  18.00  FRUTO FRESCO  BETA VULGARIS  TRANSITORIC
2      25      CUNDINAMARCA  25214  COTA  HORTALIZAS  ACELGA  ACELGA  ACELGA  2006  2006B    2      2      26  17.33  FRUTO FRESCO  BETA VULGARIS  TRANSITORIC
3      54      NORTE DE SANTANDER  54405  LOS PATIOS  HORTALIZAS  ACELGA  ACELGA  ACELGA  2006  2006B    3      3      48  16.00  FRUTO FRESCO  BETA VULGARIS  TRANSITORIC
4      54      NORTE DE SANTANDER  54518  PAMPLONA  HORTALIZAS  ACELGA  ACELGA  ACELGA  2006  2006B    1      1      5  10.00  FRUTO FRESCO  BETA VULGARIS  TRANSITORIC

```

Fuente: Framework Python Pandas

Se verifican los datos origen y se evidencia la cantidad de registros que se recuperaron y se realiza el proceso de normalización de los datos (ver imagen 18, 19, 20, 21 y 22).

Ilustración 18. Normalización de datos con Pandas 5

Revisamos columna desagregacion_regional que solo esta en historico para validar si es igual a cultivo

```
[ ] # Revisamos columna desagregacion_regional que solo esta en historico para validar si es igual a cultivo
dfEVAHist['igual'] = np.where(dfEVAHist['Cultivo']==dfEVAHist['desagregacion_regional'],1,0)

[ ] dfEVAHist[['Cultivo','desagregacion_regional']].loc[dfEVAHist['igual'] == 0]
```

	Cultivo	desagregacion_regional
221	MALANGA	ACHIN
222	MALANGA	ACHIN
223	MALANGA	ACHIN
224	MALANGA	ACHIN
225	MALANGA	ACHIN
...
206063	CALABACIN	ZUCCHINI
206064	CALABACIN	ZUCCHINI
206065	CALABACIN	ZUCCHINI
206066	CALABACIN	ZUCCHINI
206067	CALABACIN	ZUCCHINI

43760 rows x 2 columns

```
[ ] dfEVAHist['igual'].value_counts()
1    162308
0     43760
Name: igual, dtype: int64
```

Fuente: Framework Python Pandas

Ilustración 19. Normalización de datos con Pandas 6

la columna desagregacion_regional tiene identificadores unicos que no estan en los datos de 2019 y 2021. para dejar todo en un solo dataframe se toma la desicion de eliminarla -->

```
[ ] desagregacion_regional = dfEVAHist.pop('desagregacion_regional')
```

Fuente: Framework Python Pandas

Ilustración 20. Normalización de datos con Pandas 7

▼ Cambiamos nombres de Minusculas a Mayusculas

```
[ ] dfEVAHist = dfEVAHist.apply(lambda x: x.str.upper() if(x.dtype == "object") else x)
[ ] dfEVA = dfEVA.apply(lambda x: x.str.upper() if(x.dtype == "object") else x)
```

▼ Reemplazamos Acentos

```
[ ] dfEVA = dfEVA.apply(lambda x: x.str.normalize('NFKD').str.encode('ascii', errors='ignore').str.decode('utf-8') if(x.dtype == "object") else x)
```

Fuente: Framework Python Pandas

Ilustración 21. Normalización de datos con Pandas 8

dfEVA.head()

cod_dep	Departamento	cod_mun	Municipio	grup_cultivo	subgrup_cultivo	Cultivo	Año	Periodo	Area_Sembrada	Area_Cosechada	Produccion	Rendimiento	estado_produccion	nombre_cientifico	Ci	
0	5	ANTIOQUIA	5002	ABEJORRAL	LEGUMINOSAS Y OLEAGINOSAS	LEGUMINOSAS	FRIJOL	2019	2019B	110.0	105.0	157.5	1.5	GRANO SECO	PHASEOLUS VULGARIS	TI
1	5	ANTIOQUIA	5002	ABEJORRAL	FRUTALES	MORA	MORA	2019	2019	6.0	6.0	48.0	8.0	FRUTO FRESCO	RUBUS GLAUCUS	PE
2	5	ANTIOQUIA	5002	ABEJORRAL	FRUTALES	SOLANACEAS	LULO	2019	2019	5.0	5.0	25.0	5.0	FRUTO FRESCO	SOLANUM QUITOENSE	PE
3	5	ANTIOQUIA	5002	ABEJORRAL	FRUTALES	SOLANACEAS	TOMATE DE ARBOL	2019	2019	5.0	5.0	125.0	25.0	FRUTO FRESCO	SOLANUM BETACEUM	PE
4	5	ANTIOQUIA	5002	ABEJORRAL	FRUTALES	MANGO	MANGO	2019	2019	20.0	20.0	120.0	6.0	FRUTO FRESCO	MANGIFERA INDICA	PE

```
[ ] dfEVAHist = dfEVAHist.apply(lambda x: x.str.normalize('NFKD').str.encode('ascii', errors='ignore').str.decode('utf-8') if(x.dtype == "object") else x)
```

dfEVAHist.head()

cod_dep	Departamento	cod_mun	Municipio	grup_cultivo	subgrup_cultivo	Cultivo	Año	Periodo	Area_Sembrada	Area_Cosechada	Produccion	Rendimiento	estado_produccion	nombre_cientifico	Ciclo_cultivo
0	15	BOYACA	15114	BUSBANZA	HORTALIZAS	ACELGA	2006	2006B	2	1	1	1.00	FRUTO FRESCO	BETA VULGARIS	TRANSITORIO
1	25	CUNDINAMARCA	25754	SOACHA	HORTALIZAS	ACELGA	2006	2006B	82	80	1440	18.00	FRUTO FRESCO	BETA VULGARIS	TRANSITORIO
2	25	CUNDINAMARCA	25214	COTA	HORTALIZAS	ACELGA	2006	2006B	2	2	26	17.33	FRUTO FRESCO	BETA VULGARIS	TRANSITORIO
3	54	NORTE DE SANTANDER	54405	LOS PATIOS	HORTALIZAS	ACELGA	2006	2006B	3	3	48	16.00	FRUTO FRESCO	BETA VULGARIS	TRANSITORIO
4	54	NORTE DE SANTANDER	54518	PAMPLONA	HORTALIZAS	ACELGA	2006	2006B	1	1	5	10.00	FRUTO FRESCO	BETA VULGARIS	TRANSITORIO

Fuente: Framework Python Pandas

Ilustración 22. Normalización de datos con Pandas 9

```

Unificamos en un solo dataframe

[ ] #Concatenamos en un solo dataframe uno debajo del otro
    dfEVAUnic = pd.concat([dfEVAHist, dfEVA], axis=0)

[ ] dfEVA.shape

    (39945, 16)

[ ] dfEVAHist.shape

    (206068, 16)

[ ] # Evidenciamos que queda un solo dataframe con las mismas columnas
    dfEVAUnic.shape

    (246013, 16)

[ ] dfEVAUnic.to_csv('EVAFinal.csv',index=False,sep='|')

```

Fuente: Framework Python Pandas

La captura de los datos se hizo por medio de Python bajo un framework de trabajo llamado pandas, después de la captura de los datos de los sistemas origen se almacena en un archivo de trabajo de texto plano en formato csv, ya que este formato permite hacer cargues masivos muchos más rápidos a los motores de bases de datos (Ver imagen 23).

Ilustración 23. Creación archivo csv

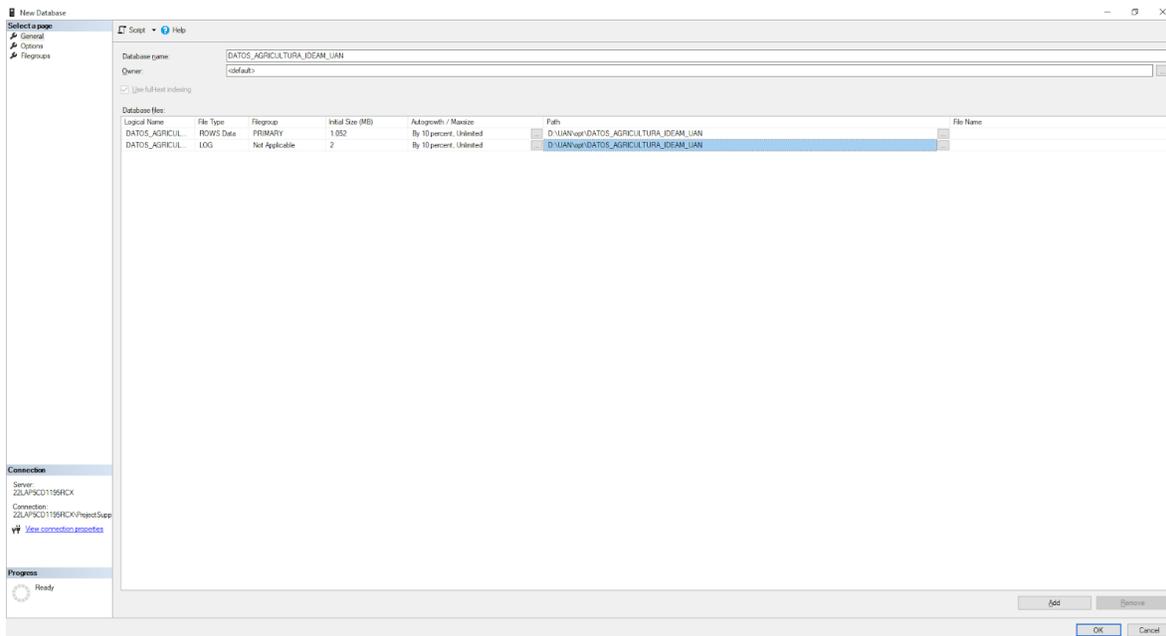
Este equipo > Windows (C:) > UAN > Repositorio > EVA

Nombre	Fecha de modificación	Tipo	Tamaño
EVA_20221024.csv	24/10/2022 10:27 p. m.	Archivo de valores...	28,799 KB
EVAHist_20221024.csv	24/10/2022 10:47 p. m.	Archivo de valores...	4,052 KB

Fuente: Archivo Plano csv

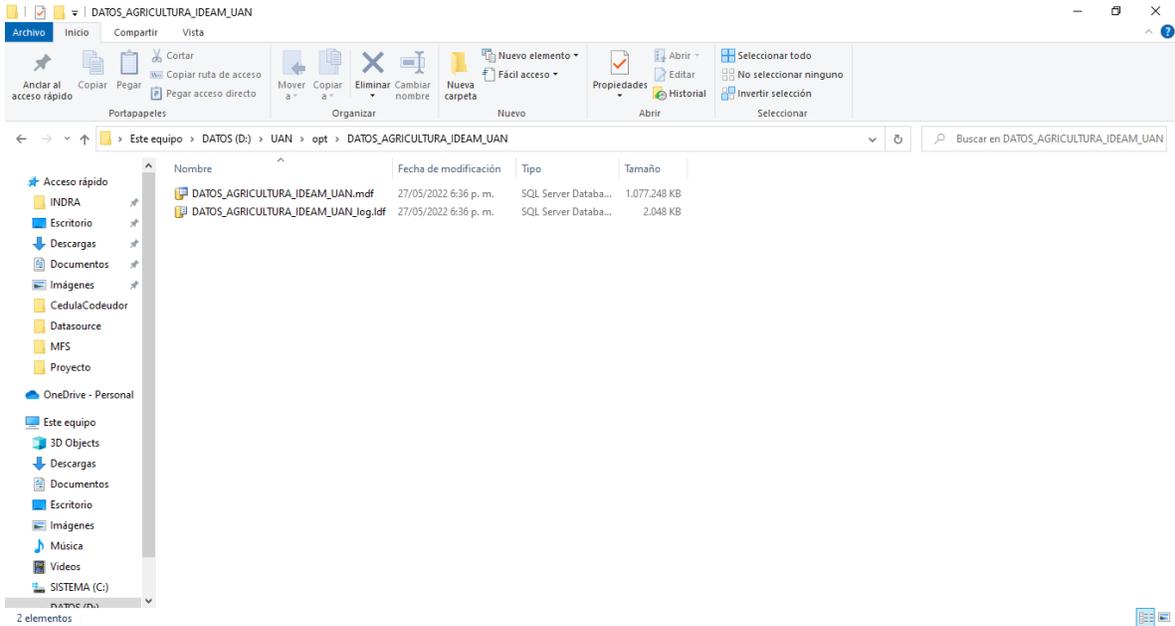
Para garantizar el correcto almacenamiento de los datos capturados de la fuente origen se trabajó mediante archivos mdf los cuales son los archivos donde SQL SERVER almacena los todos y con este se garantiza un respaldo de seguridad en caso de una contingencia (Ver imagen 24 y 25).

Ilustración 24. Configuración Archivos MDF



Fuente: Sql Server – Microsoft

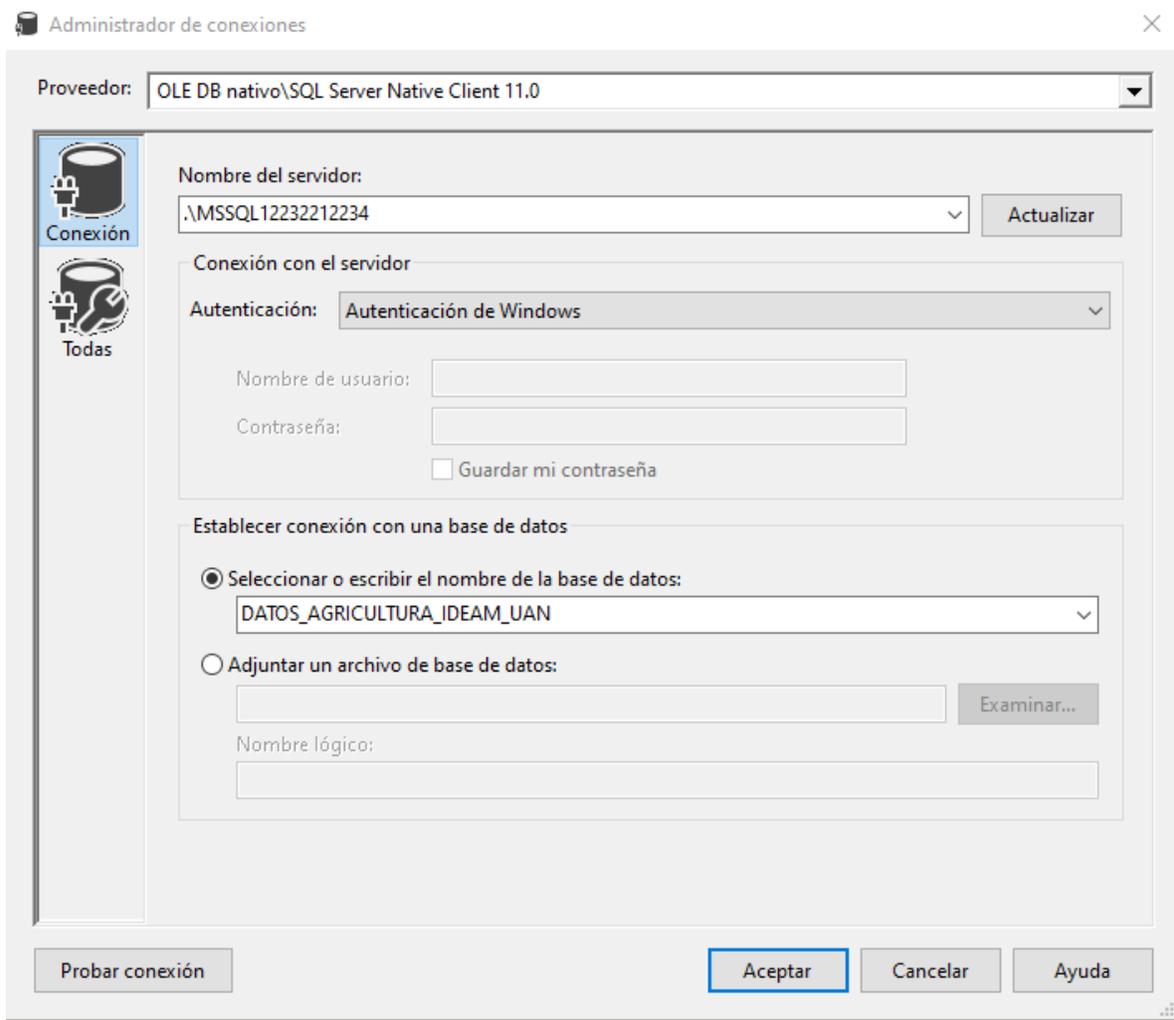
Ilustración 25. Resumen de Archivos MDF



Fuente: *Sql Server - Microsoft*

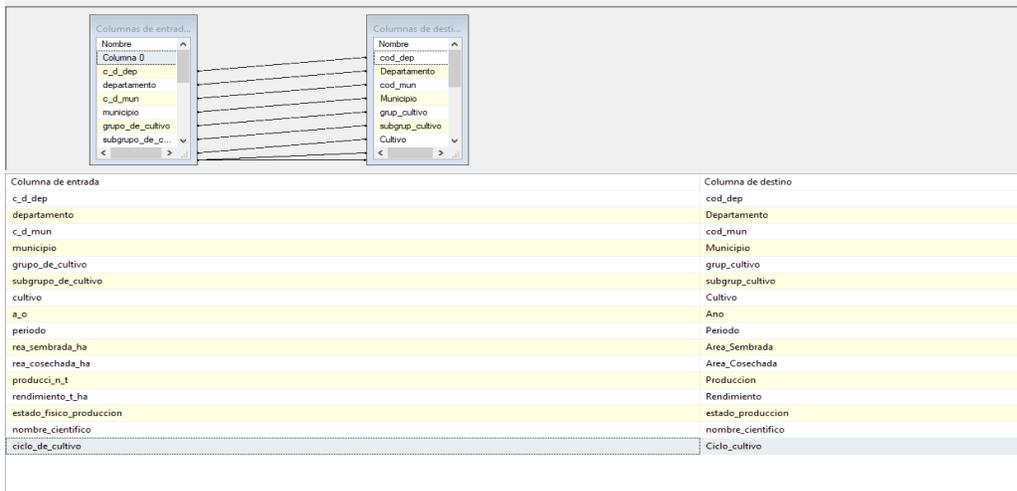
Para la ETL, se utilizó la herramienta SSDT - SQL Server Data Tools de Microsoft, donde se puede evidenciar el proceso de configuración de esta (Ver imagen 26 y 27).

Ilustración 26. Configuración Data Tools



Fuente: *Sql Data Tools - Microsoft*

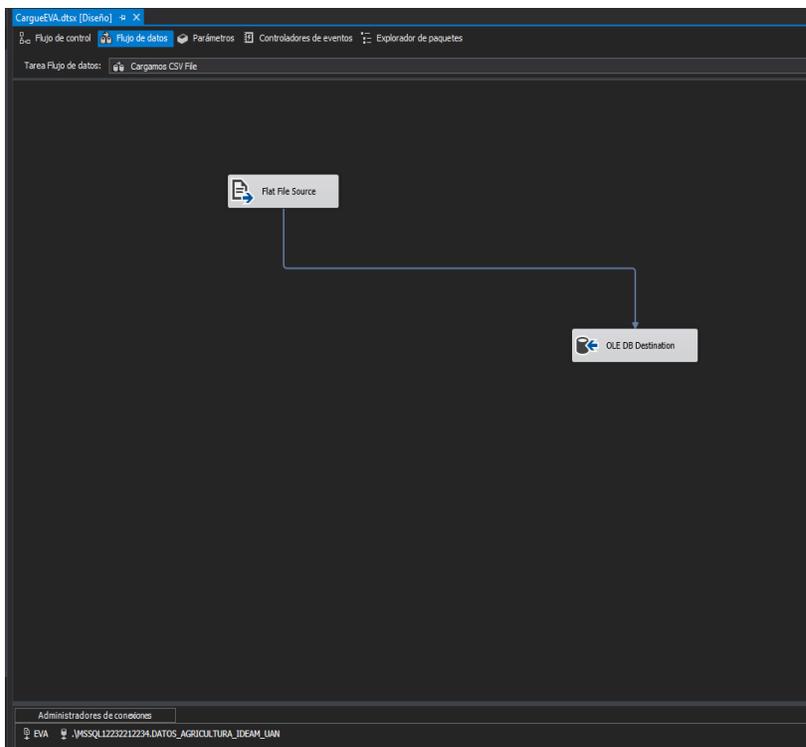
Ilustración 27. Configuración Data Tools 2



Fuente: *Sql Data Tools – Microsoft*

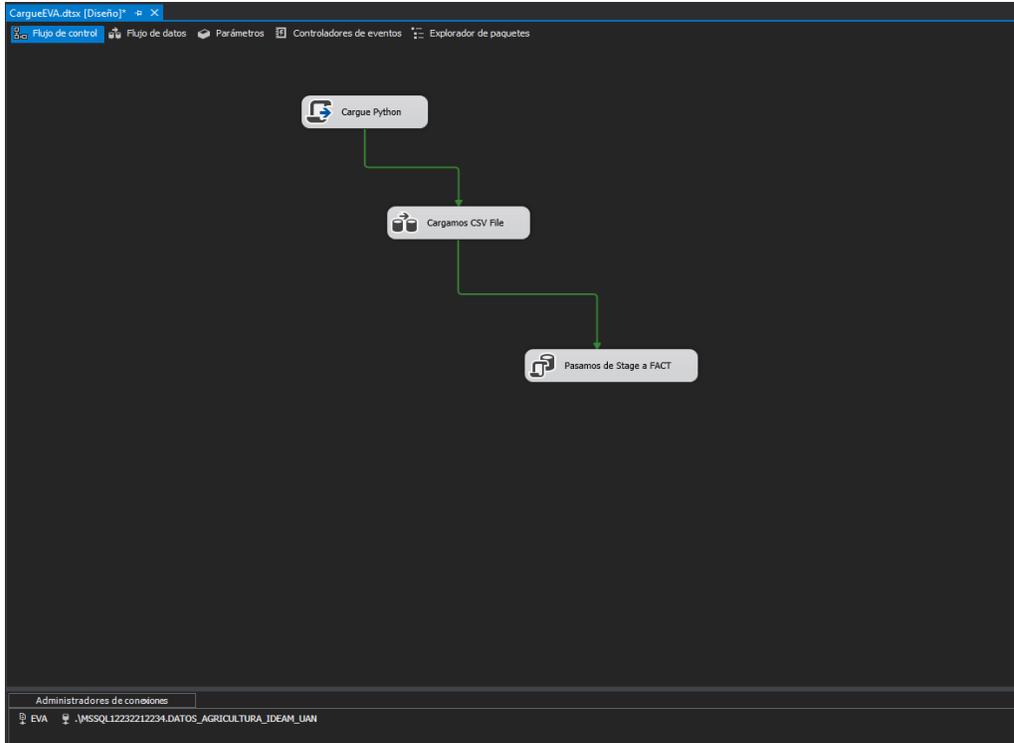
Se hace la configuración del diagrama del proceso el garantizar el dashboard (ver imagen 28, 29, 30, 31 y 32).

Ilustración 28. Configuración Data Tools 3



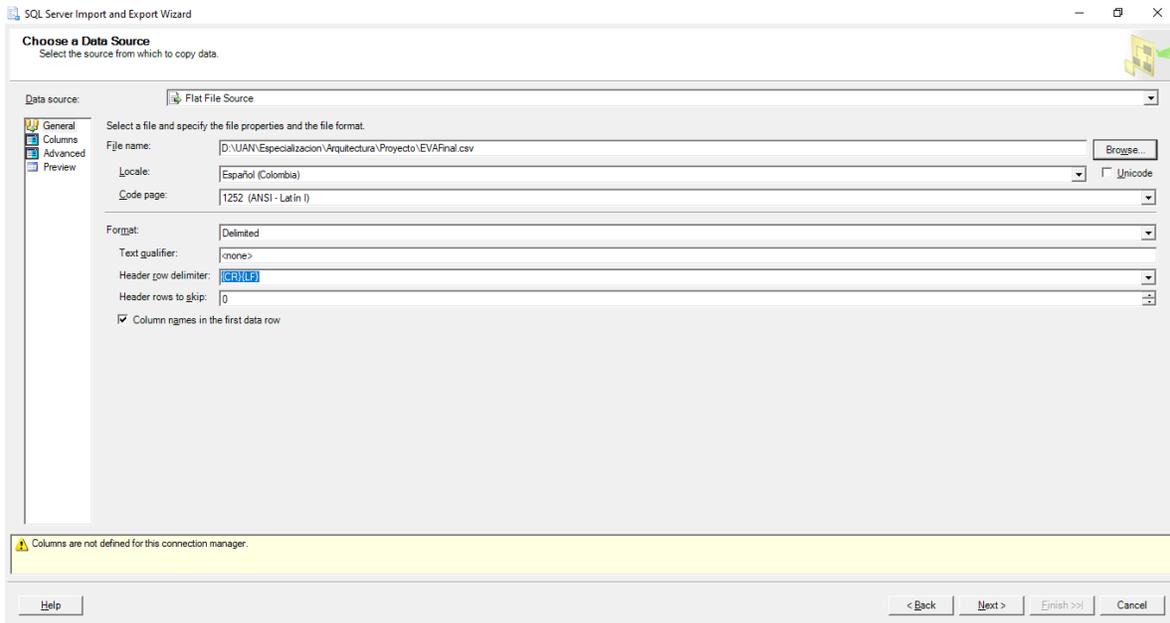
Fuente: *Sql Data Tools - Microsoft*

Ilustración 29. Configuración Data Tools 4



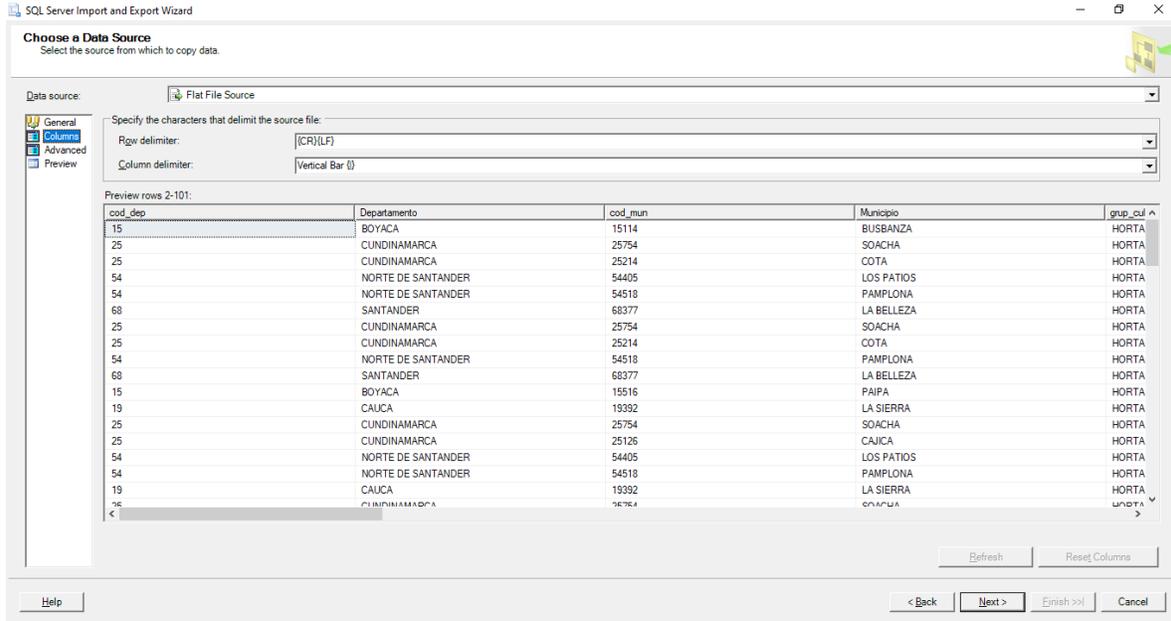
Fuente: *Sql Data Tools - Microsoft*

Ilustración 30. Configuración Data Tools 5



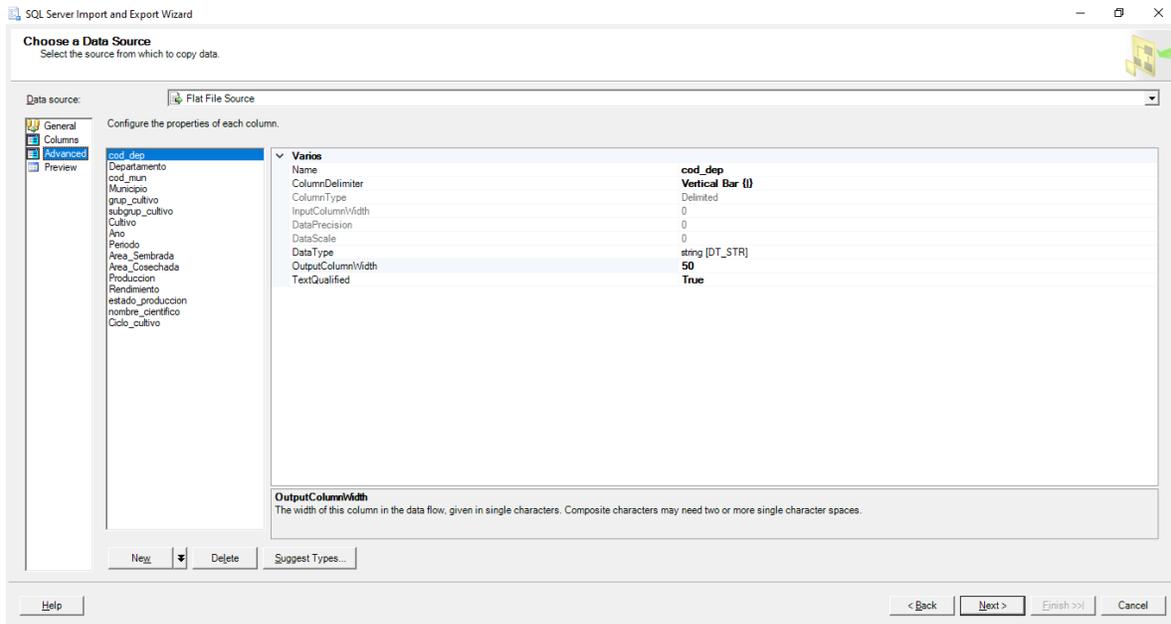
Fuente: Sql Data Tools - Microsoft

Ilustración 31. Configuración Data Tools 6



Fuente: Sql Data Tools - Microsoft

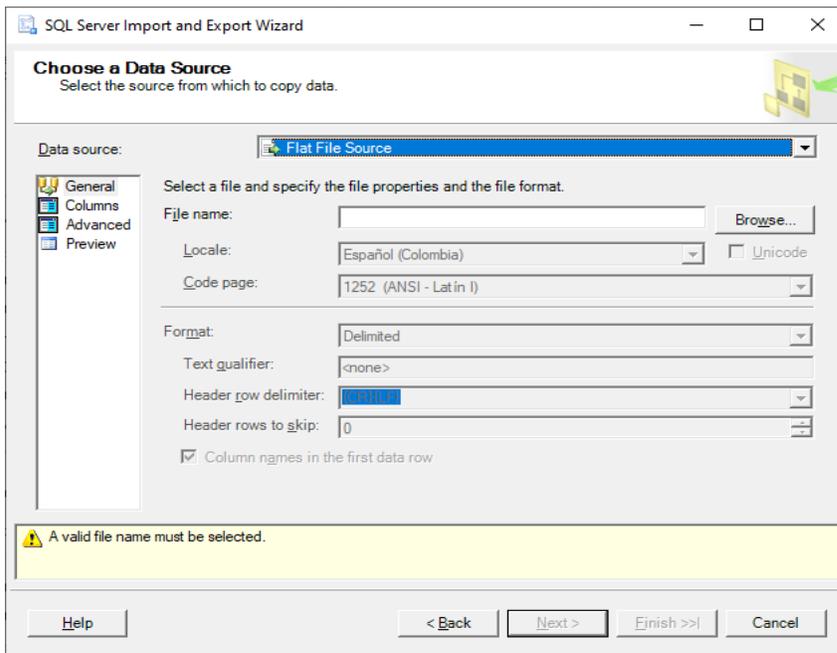
Ilustración 32. Configuración Data Tools 7



Fuente: *Sql Data Tools - Microsoft*

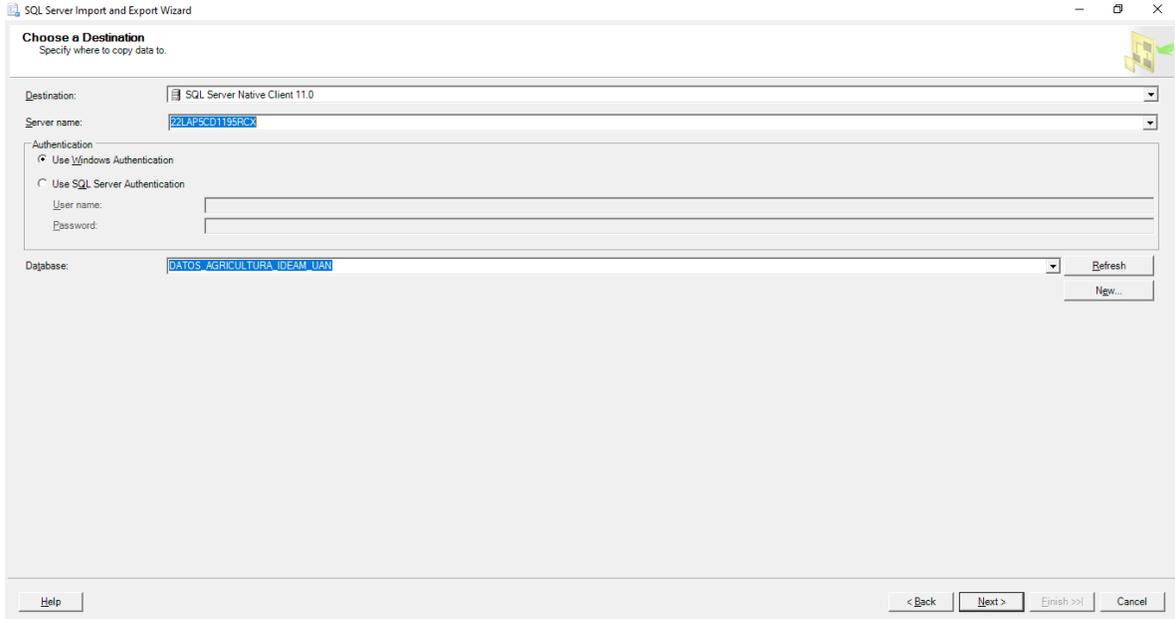
Se hace la configuración de la data tool para garantizar el proceso de lectura de los datos (ver imagen 33, 34, 35 y 36).

Ilustración 33. Configuración Data Tools 8



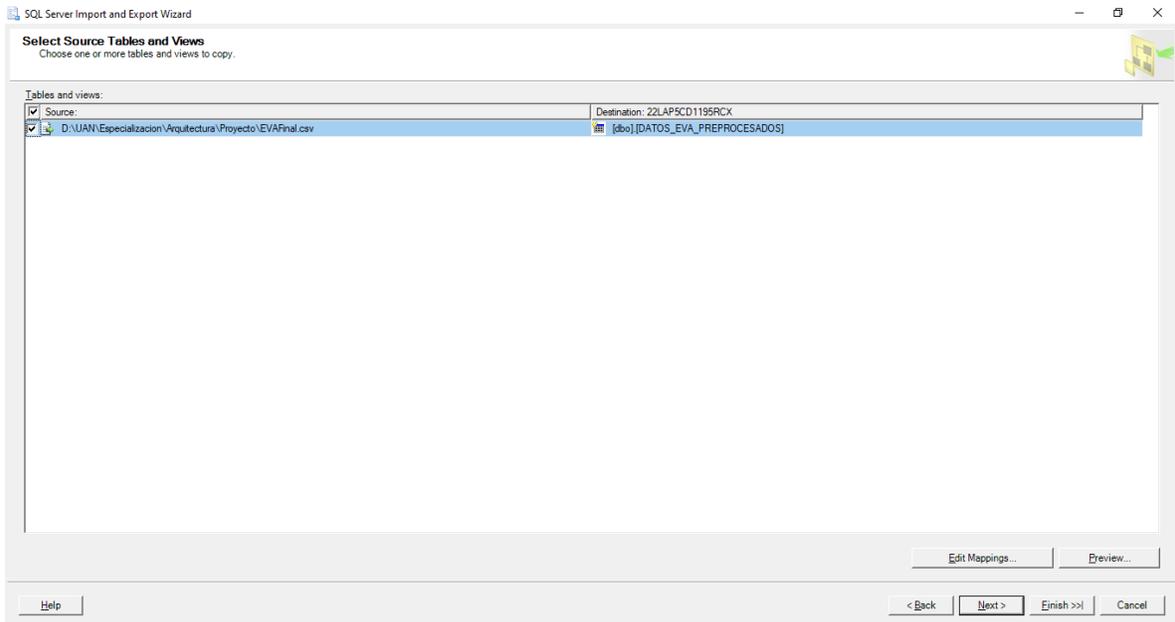
Fuente: *Sql Data Tools - Microsoft*

Ilustración 34. Configuración Data Tools 9



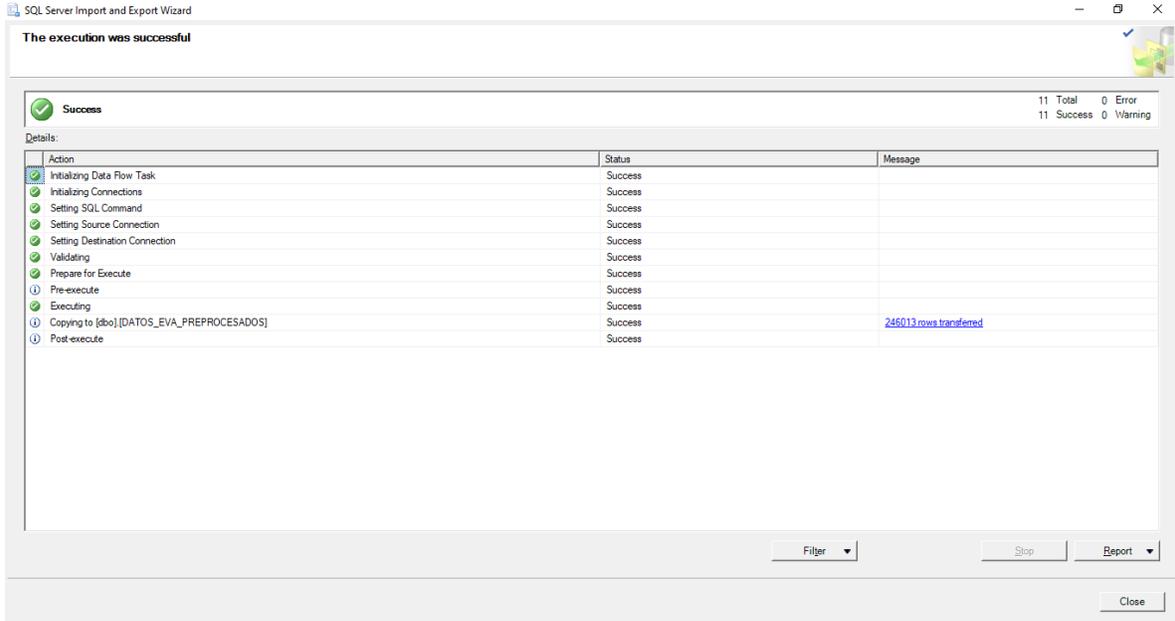
Fuente: *Sql Data Tools - Microsoft*

Ilustración 35. Configuración Data Tools 10



Fuente: *Sql Data Tools - Microsoft*

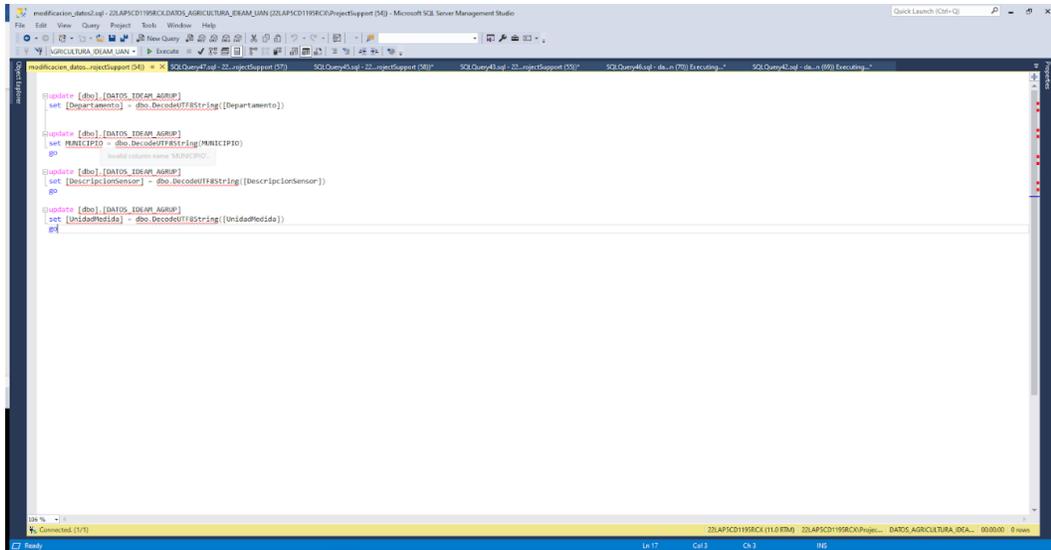
Ilustración 36. Configuración Data Tools 11



Fuente: *Sql Data Tools - Microsoft*

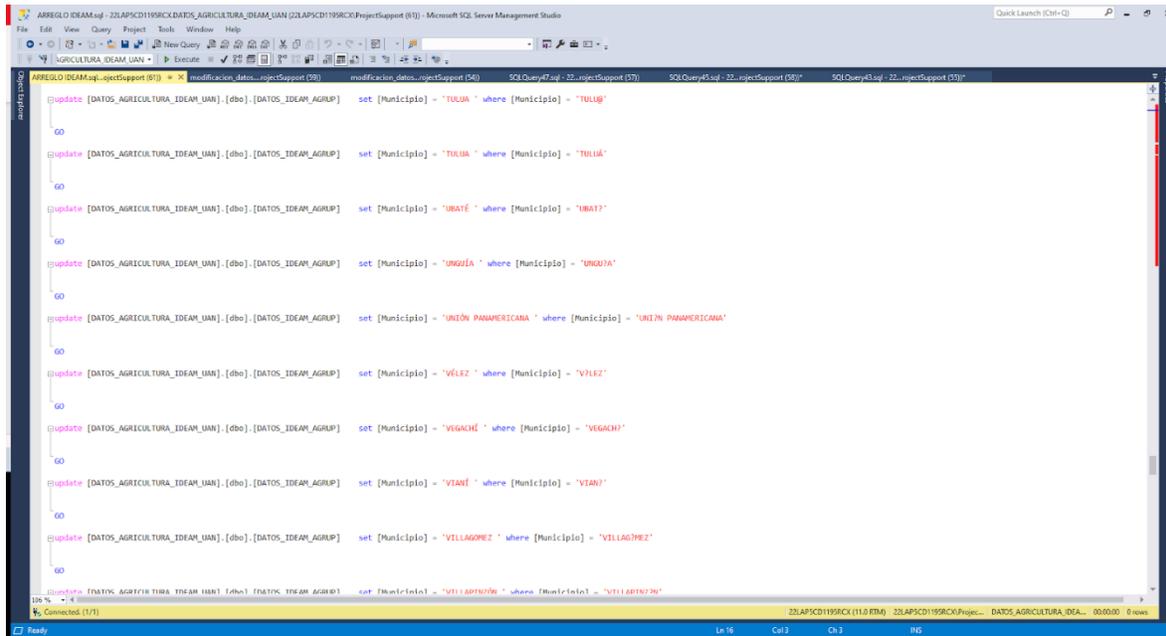
Para los preprocesamientos de los datos se realizaron con la herramienta de SQL SERVER (Ver imagen 37 y 38).

Ilustración 37. Normalización de datos con SQL SERVER



Fuente: *Sql Server - Microsoft*

Ilustración 38. Normalización de datos con SQL SERVER 2



Fuente: Sql Server - Microsoft

Después de hacer la corrección de los datos se verifica que los mismos hayan quedado bien definidos y transaccionales (Ver imagen 39).

Ilustración 39. Validación de datos con SQL de cultivos

	cod_dep	Departamento	cod_mun	Municipio	grup_cultivo	subgrup_cultivo	Cultivo	Año	Periodo	Area_Sembrada	Area_Cosechada	Produccion	Rendimiento	estado_produccion	nombre_cliente
1	15	BOYACA	15778	SUTATENZA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	5	5	27	6	HORTALIZA FRESCA	CUCURBITA MOSHATA
2	17	CALDAS	17867	VICTORIA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	3	3	50	20	HORTALIZA FRESCA	CUCURBITA MOSHATA
3	19	CAUCA	19845	VILLA RICA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	53	47	94	2	HORTALIZA FRESCA	CUCURBITA MOSHATA
4	19	CAUCA	19532	PATIA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	30	21	192	9.14	HORTALIZA FRESCA	CUCURBITA MOSHATA
5	19	CAUCA	19300	GUACHENE	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	27	21	105	5	HORTALIZA FRESCA	CUCURBITA MOSHATA
6	19	CAUCA	19573	PUERTO TEJADA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	22	22	281	13	HORTALIZA FRESCA	CUCURBITA MOSHATA
7	19	CAUCA	19075	BALBOA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	4	4	32	8	HORTALIZA FRESCA	CUCURBITA MOSHATA
8	20	CESAR	20001	VALLEDUPAR	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	100	100	1500	15	HORTALIZA FRESCA	CUCURBITA MOSHATA
9	20	CESAR	20787	TAMALAMEQUE	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	80	50	750	15	HORTALIZA FRESCA	CUCURBITA MOSHATA
10	20	CESAR	20060	BOSCONIA	HORTALIZAS	AHUYAMA	AHUYAMA	2009	2009A	10	9	135	15	HORTALIZA FRESCA	CUCURBITA MOSHATA

Fuente: Sql Server - Microsoft

Debido a la cantidad de registros del IDEAM, se aplica un agrupamiento de datos para poder manejar la data de manera correcta, se agrupó por día, departamento, municipio, medida y el promedio de medición (Ver imagen 40).

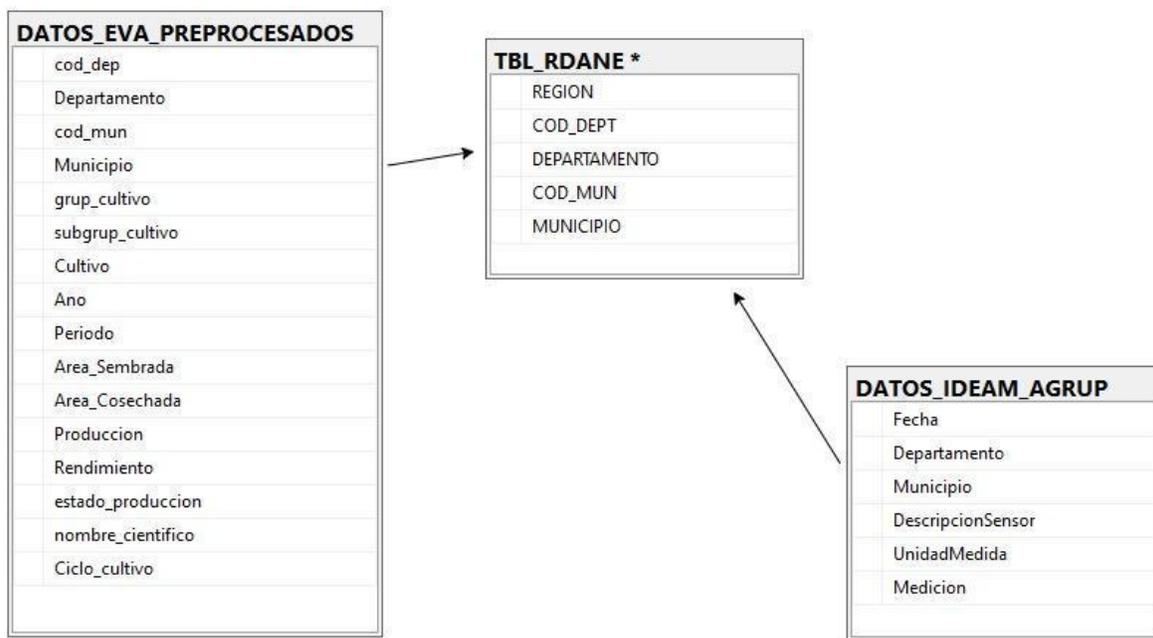
Ilustración 40. Validación de datos con SQL de clima

	Fecha	Departamento	Municipio	DescripcionSensor	UnidadMedida	Medicion
1	2005-01-08	CUNDINAMARCA	ZIPAQUIRA	Humedad del aire 2 mt	%	85.6
2	2005-01-09	ANTIOQUIA	JERICO	Humedad del aire 2 mt	%	89.08333333333333
3	2005-01-09	LA GUAJIRA	URIBIA	Humedad del aire 2 mt	%	70
4	2005-01-10	ANTIOQUIA	SANTA FE DE ANTIOQUIA	Humedad del aire 2 mt	%	80.54166666666667
5	2005-01-11	ANTIOQUIA	TURBO	Humedad del aire 2 mt	%	83.46875
6	2005-01-11	CORDOBA	LORICA	Humedad del aire 2 mt	%	74.94791666666667
7	2005-01-11	CAUCA	PAEZ	Humedad del aire 2 mt	%	58.77777777777778
8	2005-01-13	ANTIOQUIA	JERICO	Humedad del aire 2 mt	%	70.54166666666667
9	2005-01-15	ANTIOQUIA	VEGACHI	Humedad del aire 2 mt	%	88
10	2005-01-15	CUNDINAMARCA	ZIPAQUIRA	Humedad del aire 2 mt	%	89
11	2005-01-16	ANTIOQUIA	CHIGORODO	Humedad del aire 2 mt	%	81.54166666666667
12	2005-01-16	LA GUAJIRA	URIBIA	Humedad del aire 2 mt	%	69.6
13	2005-01-17	ANTIOQUIA	SANTA FE DE ANTIOQUIA	Humedad del aire 2 mt	%	65.33333333333333
14	2005-01-18	ANTIOQUIA	TURBO	Humedad del aire 2 mt	%	83.79166666666667
15	2005-01-18	CORDOBA	LORICA	Humedad del aire 2 mt	%	73.89583333333333
16	2005-01-19	ANTIOQUIA	RIONEGRO	Humedad del aire 2 mt	%	78.75

Fuente: Sql Server – Microsoft

Se define un diagrama lógico para la base de datos, donde se puede garantizar una correcta trazabilidad y transaccionalidad al momento de hacer consultas sobre la base de datos que va a estar almacenados (Ver imagen 41).

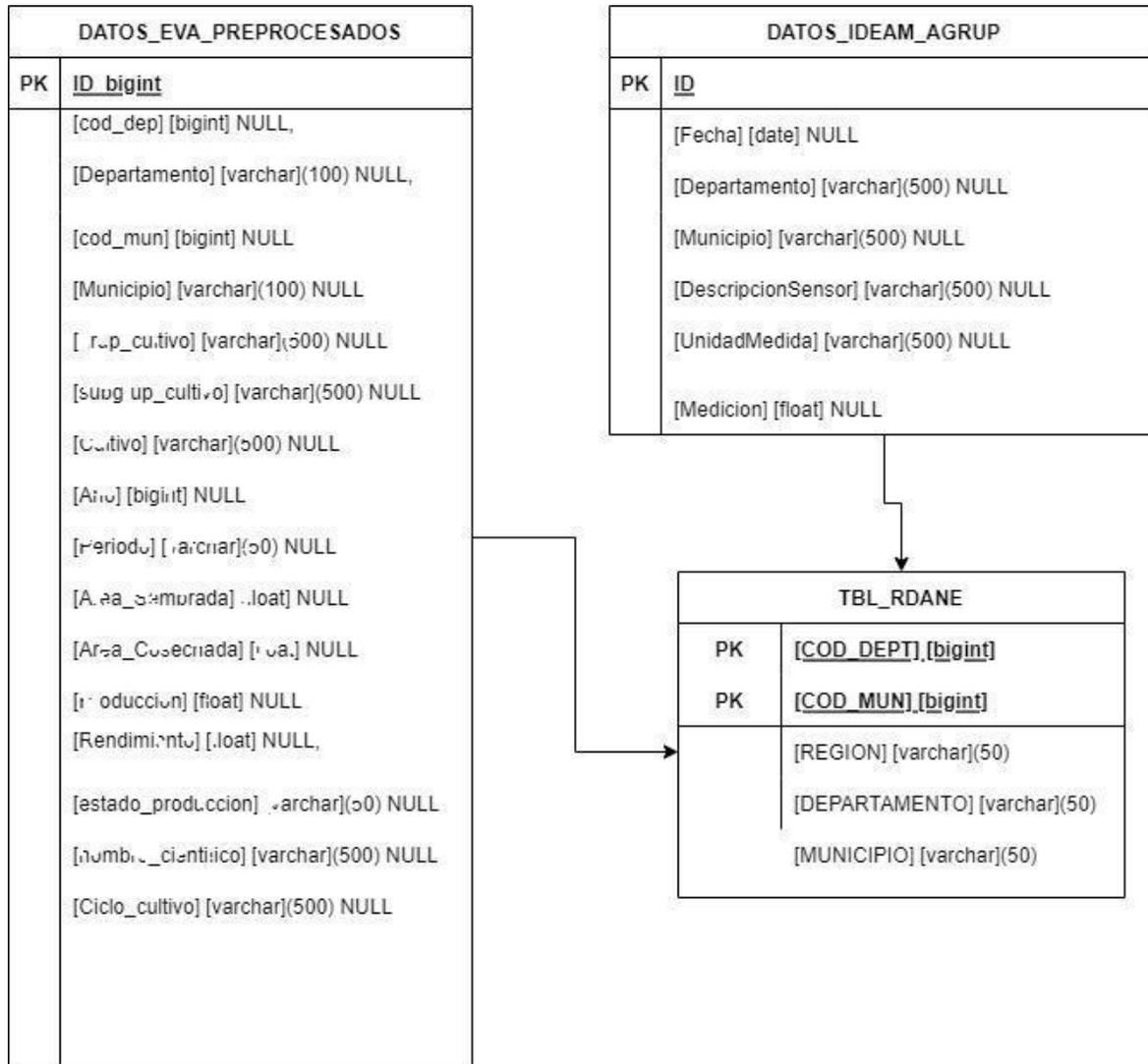
Ilustración 41. Modelo lógico de bases de datos



Fuente: *Sql Server – Microsoft*

Después de tener un modelo lógico definido se desarrolla el modelo físico, este asegura la capacidad del almacenamiento, el tipo de dato que se debe asegurar en la transformación y las llaves principales de cada objeto para garantizar una correcta ejecución de los scripts, ya que los datos almacenados van a hacer alimentados por días y puede que se tengan muchos datos (Ver imagen 42).

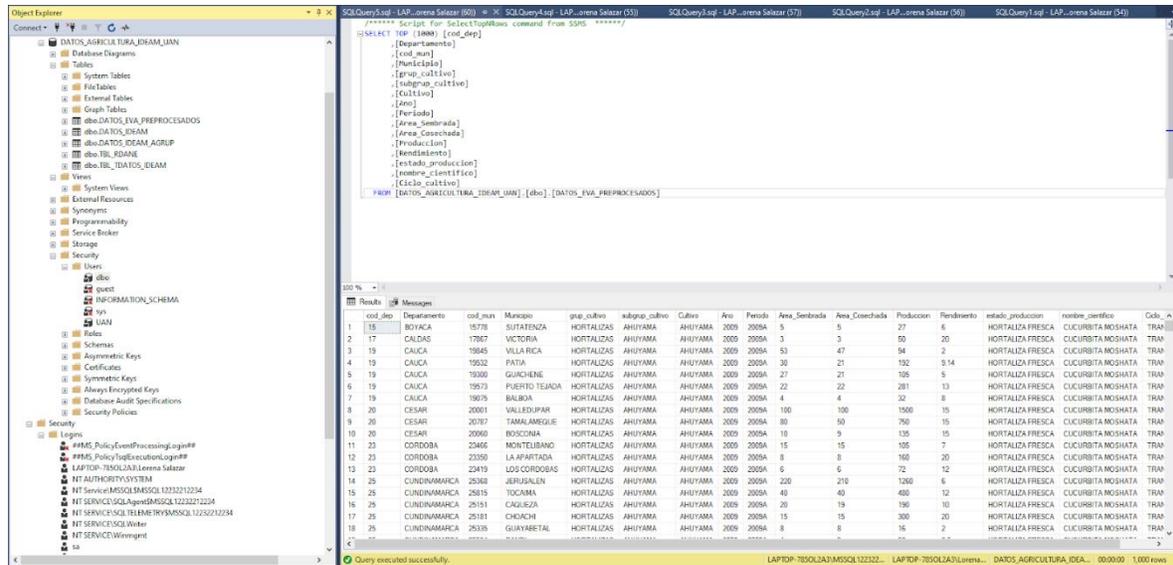
Ilustración 42. Modelo físico de bases de datos



Fuente: *Sql Server – Microsoft*

Como ya se tiene definido el modelo de datos y los objetos creados sobre la base de datos, se crea un usuario donde el dashboard se va a conectar a esta para empezar a leer los objetos ya definidos para los mismos y empezar a crear los cuadros de medición y generar los reportes (Ver imagen 43).

Ilustración 43. Creación usuario para Power BI



Fuente: Sql Server - Microsoft

6.6 Resultado

El dashboard se compone de lo siguiente:

- **Encabezado**
- Logo de UAN
- **Filtros de Búsqueda**

El tablero tendrá los siguientes filtros:

- Grupo Cultivo
- Cultivo
- Año

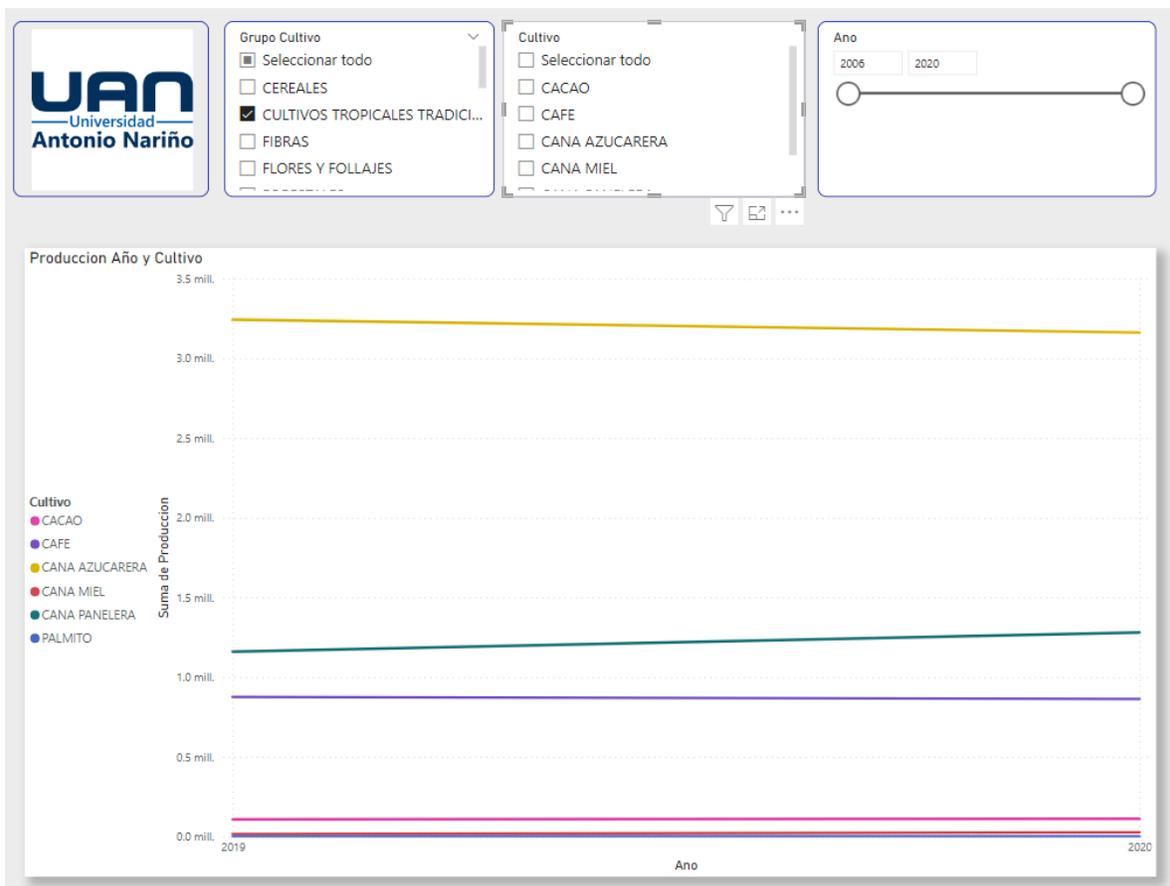
- Estos filtros serán transversales a todos los tableros.
- Todos los tableros muestran mostrar la eficacia del set de datos.

A continuación, se identifica cada tablero, su funcionalidad y filtros adicionales:

- Comportamiento Por Cultivo Por Año

Este reporte indica el comportamiento histórico por año de grupo cultivo y el cultivo en todo el territorio nacional. Los filtros permiten identificar el tiempo de grupo de cultivo y cultivo (ver imagen 44).

Ilustración 44. Resumen Dashboard

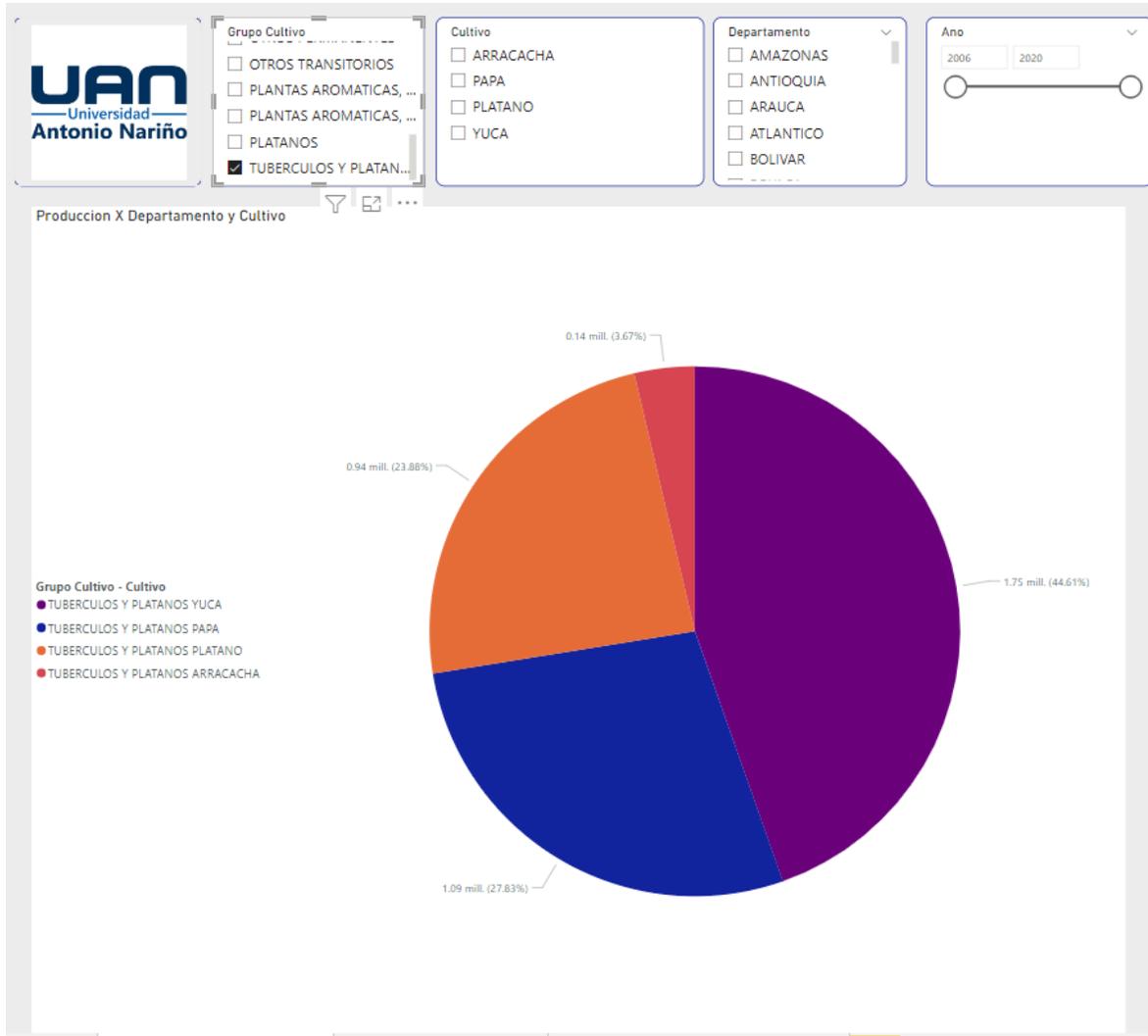


Fuente: Power BI - Microsoft

- **Total, Producción X Cultivo**

Este reporte contiene un filtro adicional por departamento que permite identificar el % de participación de cada uno de los cultivos (ver imagen 45).

Ilustración 45. Producción x Cultivo

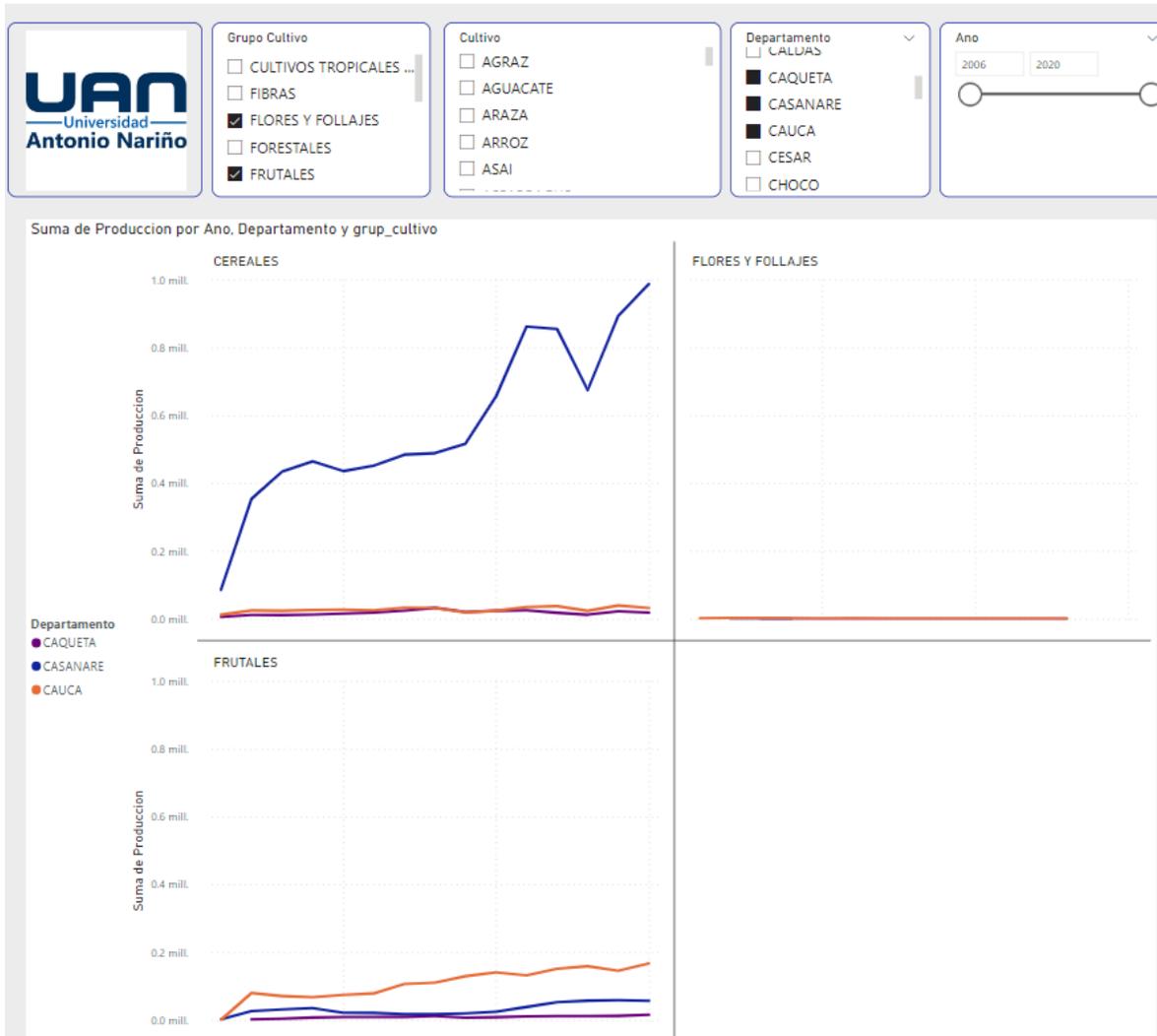


Fuente: Power BI - Microsoft

- **Cultivo x Departamento**

El siguiente reporte muestra el comportamiento de cada tipo de grupo de cultivos, sus filtros permiten identificar varios departamentos, cultivos y grupos de cultivo, así como el año (ver imagen 46).

Ilustración 46. Cultivo por departamento

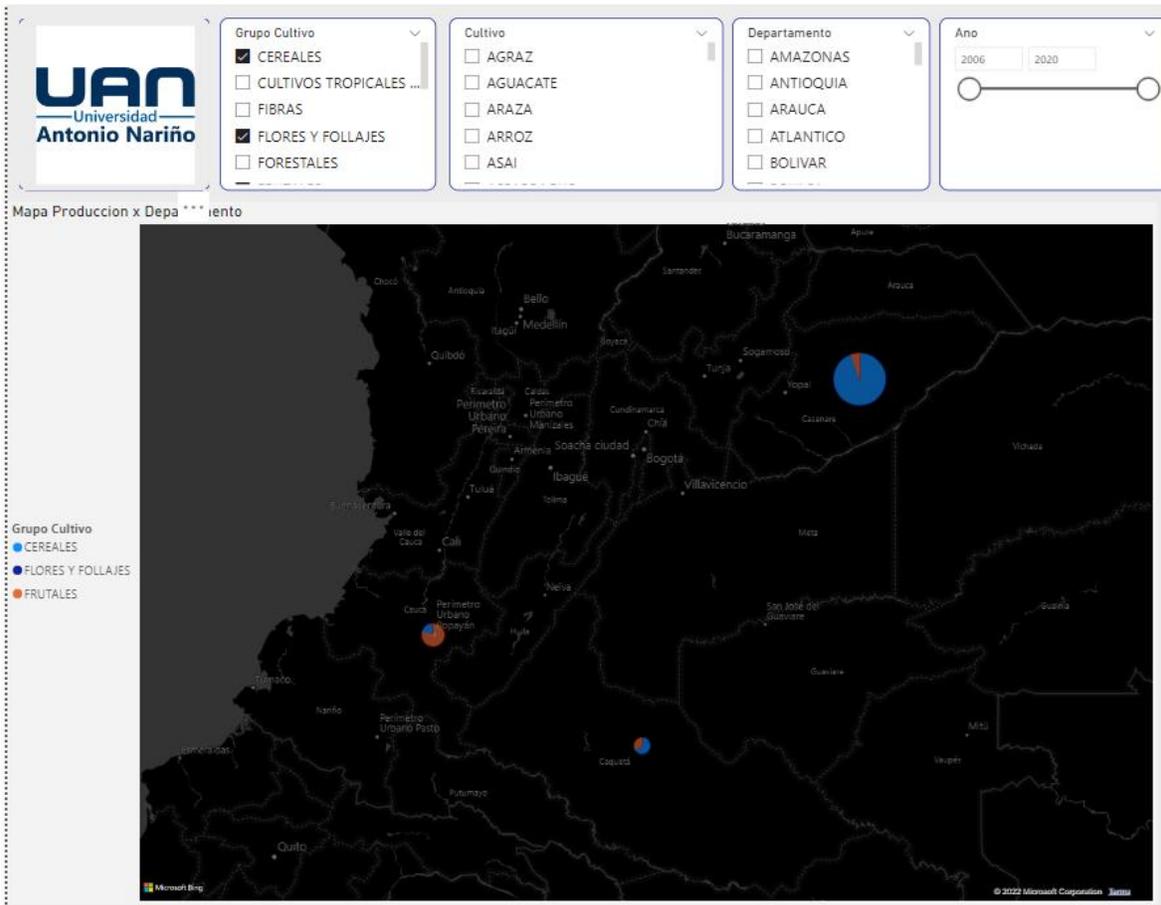


Fuente: Power BI - Microsoft

- **Mapa Producción Por Departamento**

Este nos permite identificar por departamento el cultivo y el % de efectividad en un mapa interactivo, al igual que los demás reportes cuenta con los filtros de grupo de cultivo, cultivo, departamento y año, que nos permite filtrar y visualizar los datos como deseemos (ver imagen 47).

Ilustración 47. Producción por departamento



Fuente: Power BI - Microsoft

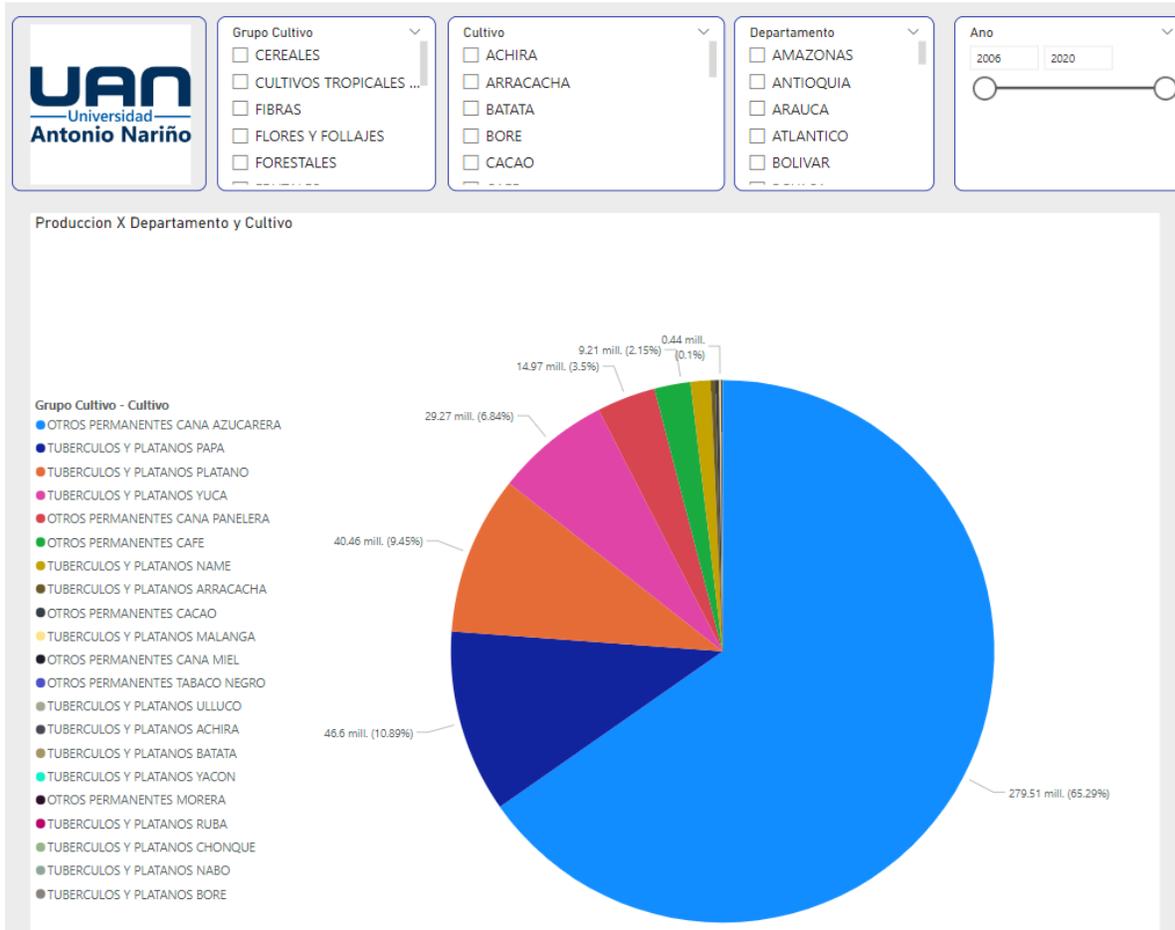
Como ya se tiene definido el usuario de conexión que se va a conectar a la base de datos, se configura sobre la herramienta de Power BI para poder crear los tableros del dashboard y así poder evidenciar que los procesos y procedimientos anteriormente definidos hayan sido efectivos para tener un resultado satisfactorio (ver imagen 48, 49 y 50).

Ilustración 48. Resumen Dashboard



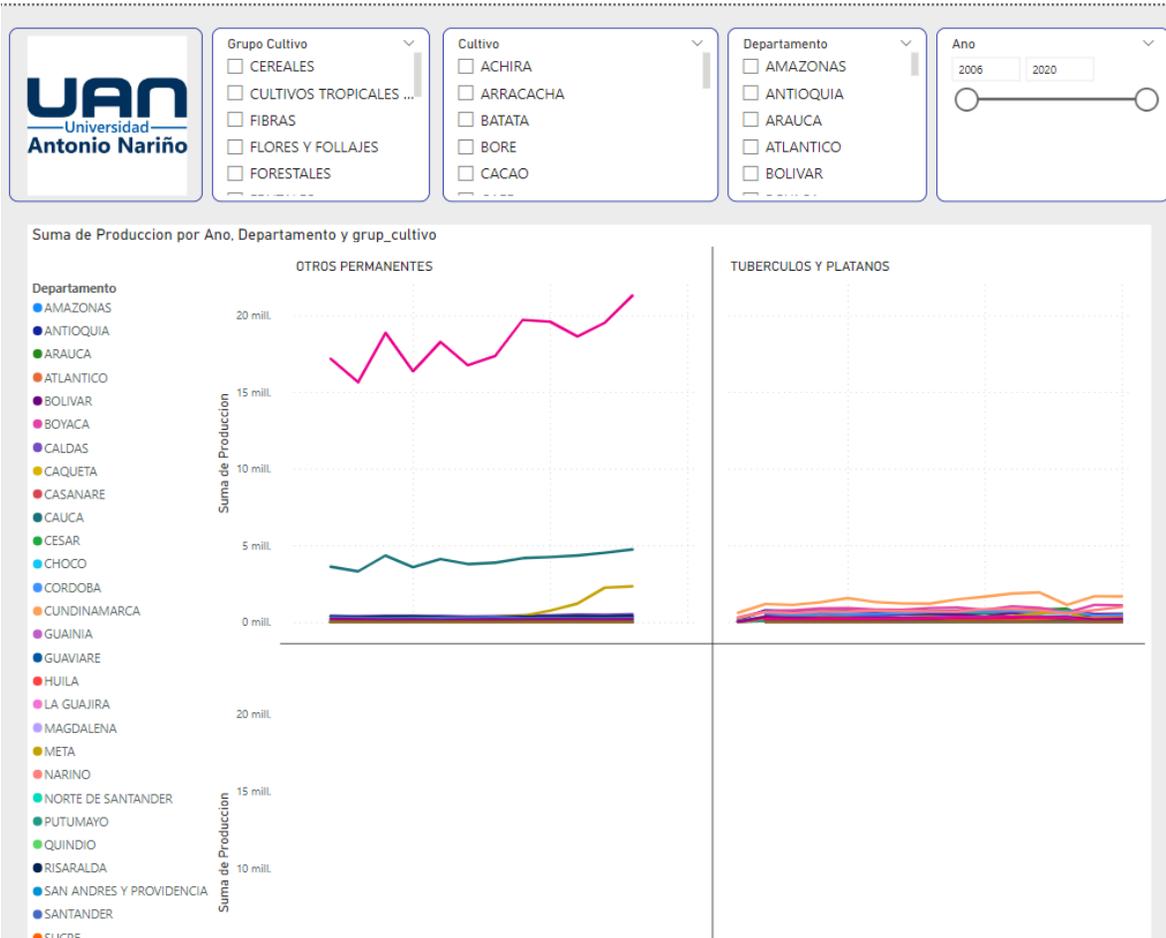
Fuente: Power BI - Microsoft

Ilustración 49. Resumen Dashboard 1



Fuente: Power BI - Microsoft

Ilustración 50. Resumen Dashboard 2



Fuente: Power BI - Microsoft

7. Conclusiones

- El resultado de la investigación y análisis de los datos que se desarrolló con el objetivo de hacer entrega de dashboard, fue satisfactorio.
- Se logro el objetivo del proyecto dado que el entregable responde la pregunta planteada al inicio de este. La generación del dashboard, en conjunto con toda la arquitectura y flujo de datos planteados, permite realizar un mejor análisis de toda la data, permitiendo realizar análisis cruzados entre municipio, tipo de cultivo, % de efectividad y clima de cada región.
- Se garantizó la correcta definición sobre la trazabilidad que se le debía hacer a los datos origen para garantizar el dashboard.
- Se logró identificar los datos expuestos por las entidades del gobierno de Colombia.
- Se crearon modelos lógicos y físicos para garantizar la correcta trazabilidad de los datos capturados en las Apis de Min. Agricultura e IDEAM.
- Se configuraron sistemas y gestores de tecnología que garantizo el correcto proceso de análisis, almacenamiento y transformación de estos.
- Se puede evidenciar mediante software de reportería en este caso Power BI que las configuraciones de los sistemas quedaron correctamente.
- Por último, se identificó que actualmente se cuenta con una cantidad gigante de datos de origen nacional, que pueden ser explotados para generar planes de mejora, acción o nuevos desarrollos en diferentes ámbitos de la sociedad colombiana.
- Las diferentes organizaciones gubernamentales deben promover y dar a conocer que existen los conjuntos de datos abiertos, así como promover el uso de herramientas de explotación para analizar esta información de manera efectiva, en cada uno de los procesos internos de cada organización.

8. Referencias

Wanumen Silva, Luis Felipe, Edwin Rivas Trujillo, and Darín Jairo Mosquera Palacios. *Bases De Datos En SQL Server*. Bogotá: Ecoe Ediciones, 2018. *Ingeniería Y Salud En El Trabajo*. Informática. Web.

República De Colombia Ministerio De Agricultura Y Desarrollo Rural. (1999). *Descentralizacion en el agro*. S.l.: Impreandespresencia.

Suárez Pérez, Sandra, Óscar Darío Amaya Navas, and Carlos Costa Posada. *Protocolo Para La Emisión De Los Pronósticos Hidrológicos*. Bogotá: IDEAM, 1999. Print.

Berkani, Nabila, and Ladjel Bellatreche. *Database and Expert Systems Applications*. Cham: Springer International, 2017. *Lecture Notes in Computer Science*. Web.

Niu, Yipei, Panpan Jin, Jian Guo, Yikai Xiao, Rong Shi, Fangming Liu, Chen Qian, and Yang Wang. "PostMan: Rapidly Mitigating Bursty Traffic via On-Demand Offloading of Packet Processing." *IEEE Transactions on Parallel and Distributed Systems* 33.2 (2022): 374-87. Web.

Millman, K. Jarrod, and Michael Aivazis. "Python for Scientists and Engineers." *Computing in Science & Engineering* 13.2 (2011): 9-12. Web.

Anderson, Charles. "Docker [Software Engineering]." *IEEE Software* 32.3 (2015): 102-C3. Web.

Clark, Dan. *Beginning Power BI: A Practical Guide to Self-Service Data Analytics with Excel 2016 and Power BI Desktop, Second Edition*. Second ed. Place of Publication Not Identified: Apress, 2017. Web.

Evaluación Agropecuarias Municipales (EVA). (s. f.). Recuperado 29 de octubre de 2022, de <https://www.upra.gov.co/web/guest/evaluaciones-agropecuarias-municipales-eva>

Data Management (DAMA). (s. f.). Recuperado 28 de octubre de 2022, de <https://www.dama.org/cpages/dmbok-images>

SIMA - Sistema Integrado de Monitoreo Agrícola. (s. f.). Recuperado 8 de noviembre de 2022, de <https://www.sima.ag/>

Datos Abiertos. (Julio de 2022). Datos Abiertos EVA Historicos. Obtenido de <https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Evaluaciones-Agropecuarias-Municipales-EVA-2019-20/p5fp-pay3>

Datos Abiertos. (Julio de 2022). Datos Abiertos EVA Históricos 2019 a 2020. Obtenido de <https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Evaluaciones-Agropecuarias-Municipales-EVA-2019-20/p5fp-pay3>

Datos Abiertos. (Julio de 2022). Datos Abiertos EVA Históricos 2007 a 2018. Obtenido de <https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Evaluaciones-Agropecuarias-Municipales-EVA/2pnw-mmge>

Datos Abiertos. (Julio de 2022). Calidad Aire IDEAM. Obtenido de <https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/DATOS-DE-CALIDAD-DEL-AIRE-EN-COLOMBIA-2011-2018/ysq6-ri4e>

Datos Abiertos. (Julio de 2022). Temperatura IDEAM. Obtenido de <https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/Datos-Hidrometeorol%C3%B3gicos-Crudos-Red-de-Estaciones/sbwg-7ju4>

Datos Abiertos. (Julio de 2022). Presión Atmosférica IDEAM. Obtenido de <https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/Presi-n-Atmosf-rica/62tk-nxj5>

Datos Abiertos. (Julio de 2022). Precipitaciones IDEAM. Obtenido de <https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/Precipitaci-n/s54a-sqyq>