PROYECTO FINAL DATA ANALYTICS CODERHOUSE

IMPACTO DEL CAMBIO CLIMÁTICO EN LA AGRICULTURA

FRANCISCO CALAHORRA 16/11/2024

Contenido

Tabla de versiones	2
Descripción de la temática de los datos:	2
Objetivo del Análisis:	2
Alcance:	3
Usuario Final:	3
Tecnologías utilizadas:	4
Acerca del Dataset:	4
Diagrama entidad-relación	5
Listado de tablas:	6
Esquema del modelo de datos:	8
Transformaciones realizadas:	9
Tablas nuevas	10
Funciones DAX:	11
Información sobre solapas:	12
Manual de marca:	15
Conclusiones:	16

Tabla de versiones

Versión	Fecha
Versión 1.0	19/09/2024
Versión 2.0	07/10/2024
Versión 2.1	01/11/2024
Versión 3.0	17/11/2024

Descripción de la temática de los datos:

La temática de los datos se centra en explorar el impacto del cambio climático en la productividad de la agricultura a lo largo de varias regiones. Esto incluye datos sobre la temperatura, precipitaciones, y rendimientos entre otros, para entender cómo los cambios en los patrones climáticos afectan la agricultura.

Objetivo del Análisis:

El objetivo del análisis es identificar y cuantificar el impacto del cambio climático en la productividad agrícola de frutas, verduras, trigo y maíz de distintas regiones a nivel mundial. Esto permitirá entender cómo los patrones climáticos afectan los rendimientos de los diferentes cultivos, ayudando a la empresa a tomar decisiones estratégicas informadas sobre dónde y cómo invertir en prácticas agrícolas adaptativas. El análisis buscará optimizar la eficiencia de la producción en función de las condiciones climáticas esperadas y brindar recomendaciones específicas para la gestión de recursos y adaptación a eventos extremos en cada zona de operación.

Alcance:

Este tablero de control estratégico-operativo proporcionará una vista consolidada del rendimiento agrícola y del riesgo climático, facilitando así la planificación a corto y largo plazo para maximizar la rentabilidad y sostenibilidad.

El proyecto abarcará:

- 1. <u>Recopilación y limpieza de datos</u>: Obtención de datos climáticos (temperatura, precipitaciones) y de productividad agrícola desde 1990 hasta 2024, asegurando calidad y consistencia.
- 2. <u>Creación de un modelo de datos</u>: Desarrollo de un modelo analítico que correlacione factores climáticos con los rendimientos agrícolas en distintas regiones climáticas.
- 3. <u>Desarrollo de visualizaciones interactivas en Power BI:</u> Construcción de un dashboard en Power BI que permita explorar el impacto climático en la productividad
- 4. Documentación detallada del proceso.

El análisis abarca múltiples regiones climáticas y tipos de cultivos relevantes para la empresa, con el objetivo de brindar insights accionables para mejorar la resiliencia y sostenibilidad de la producción agrícola frente al cambio climático.

Usuario Final:

El usuario final de este análisis será una empresa global dedicada a la explotación agrícola de frutas, vegetales, trigo y maíz en diversas zonas climáticas. La empresa podrá utilizar el dashboard de Power BI para:

- Evaluar cómo los cambios en la temperatura afectan los rendimientos en distintas regiones.
- 2. Desarrollar estrategias de adaptación para mejorar la resiliencia de su producción ante el cambio climático.
- 3. Planificar y priorizar inversiones en nuevas tecnologías, infraestructuras o prácticas agrícolas adaptativas.
- 4. Tomar decisiones operativas sobre cultivos y prácticas de riego en función de las condiciones climáticas actuales.

Tecnologías utilizadas:

Excel: utilizado para la creación de las tablas de dimensiones.

SQL Server: utilizado para crear el modelo inicial.

Power BI: utilizado tanto para la limpieza de los datos con Power Query, así como para las modificaciones del modelo relacional y la creación del tablero de control.

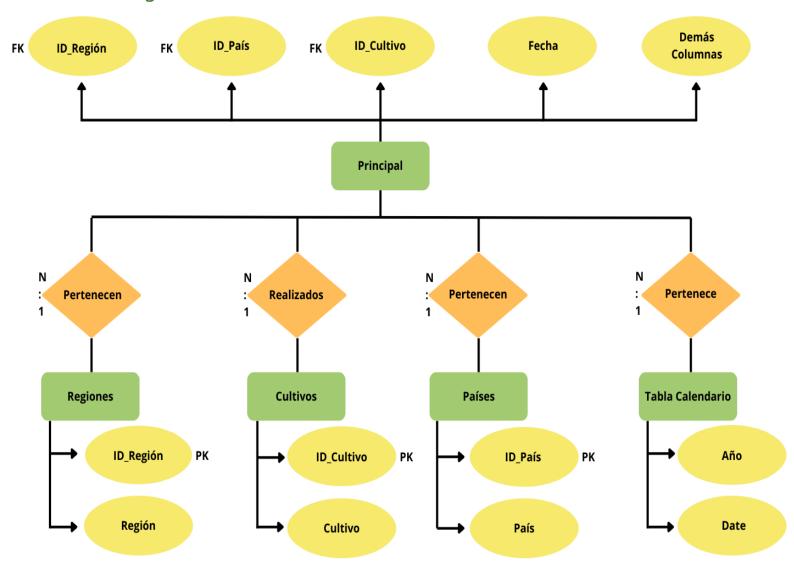
Canva: utilizado para la realización del diagrama entidad-relación.

Acerca del Dataset:

El dataset original es de dominio público y su información se recopiló de múltiples fuentes confiables, centrándose en registros recientes e históricos de temperatura, precipitación y rendimiento de los cultivos.

Link del conjunto de datos: https://www.kaggle.com/datasets/waqi786/climate-change-impact-on-agriculture/data

Diagrama entidad-relación:



Listado de tablas:

A continuación de presentan las tablas pertenecientes al modelo de datos. Fuera del modelo hay dos tablas más: una tabla de resumen de Regiones y una tabla de medidas.

Tabla "Principal"

Campo	Tipo	Descripción
Fecha	Date	Fecha del Cultivo.
Temperatura_promedio(C°)	Número decimal	Temperatura promedio del año pertinente al cultivo en grados centígrados.
Precipitaciones_totales(mm)	Número decimal	Precipitaciones totales en el año pertinente al cultivo medidas en milímetros.
Emisiones_CO2(MT)	Número decimal	Mide las emisiones de CO2 del cultivo en toneladas métricas.
Rendimiento_cultivo(TM_por_HA)	Número decimal	Mide el rendimiento del cultivo en toneladas métricas por hectárea.
Fenomenos_meteorologicos_extremos	Número entero	Mide en una escala del 1 al 10 el sufrimiento de fenómenos meteorológicos extremos.
Acceso_al_riego(%)	Número decimal	Representa el porcentaje de acceso al riego para el cultivo.
Uso_de_pesticidas(KG_por_HA)	Número decimal	Mide el uso de pesticidas en el cultivo en kilos por hectárea.
Uso_de_fertilizantes(KG_por_HA)	Número decimal	Mide el uso de fertilizantes en el

		cultivo en kilos por hectárea.
Indice_de_salud_del_suelo	Número decimal	Representa la salud del suelo en un índice del 1 al 100.
Estrategias_de_adaptacion	Texto	Representa las diferentes estrategias de adaptación utilizadas en los cultivos.
Impacto_economico(Milliones_USD)	Número decimal	Representa el impacto económico de los cultivos medido en millones de usd.
ID_region	FK Número entero	Foreign key que representa el ID de las regiones
ID_País	FK Número entero	Foreign key que representa el ID de los países
ID_Cultivo	FK Número entero	Foreign key que representa el ID de los cultivos

Tabla "Países"

Campo	Tipo	Descripción
ID_País	PK Número entero	Primary key que representa el ID de los países
País	Texto	Representa los países

Tabla "Regiones"

Campo	Tipo	Descripción
ID_Región	PK Número entero	Primary key que representa el ID de las regiones.
Región	Texto	Representa las regiones de los países.

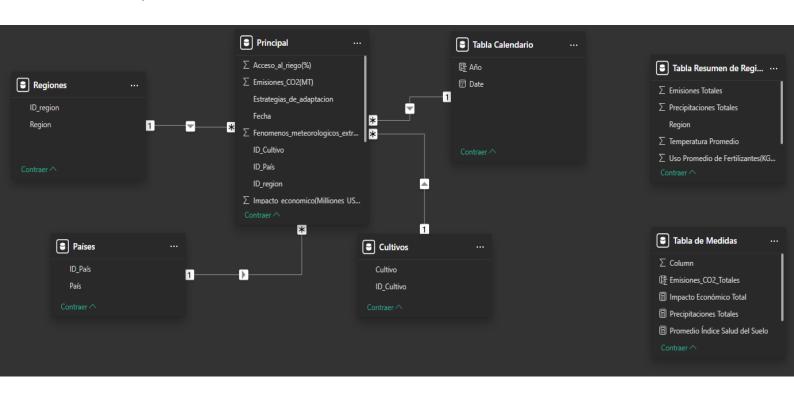
Tabla "Cultivos"

Campo	Tipo	Descripción
ID_Cultivo	PK Número entero	Primary key que representa el ID de los cultivos.
Cultivo	Texto	Representa los diferentes tipos de cultivos.

Tabla calendario

Campo	Tipo	Descripción
Date	Fecha	Fechas de los cultivos.
Año	Número entero	Año perteneciente a la fecha de los cultivos.

Esquema del modelo de datos:



Transformaciones realizadas:

Realizadas sobre la tabla "Países":

- Formato cambiado de texto a fecha #"Tipo cambiado" = Table.TransformColumnTypes(dbo_Paises,{{"Año", type date}})
- columna sin valores eliminada #"Columnas quitadas" = Table.RemoveColumns(#"Tipo cambiado",{"Regiones"})
- Cambio de nombre a columna "Año" por "Fecha"
- reemplazo de los nombres de los cultivos y estrategias de adaptación por sus respectivos nombres en español.

```
#"Valor reemplazado" = Table.ReplaceValue(#"Columnas con nombre
cambiado","Corn","Maíz",Replacer.ReplaceText,{"Cultivo"}),
#"Valor reemplazado1" = Table.ReplaceValue(#"Valor
reemplazado","Barley","Cebada",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado2" = Table.ReplaceValue(#"Valor
reemplazado1","Coffee","Café",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado3" = Table.ReplaceValue(#"Valor
reemplazado2","Cotton","Algodón",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado4" = Table.ReplaceValue(#"Valor
reemplazado3","Fruits","Frutas",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado5" = Table.ReplaceValue(#"Valor
reemplazado4","Rice","Arroz",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado6" = Table.ReplaceValue(#"Valor
reemplazado5","Soybeans","Soja",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado7" = Table.ReplaceValue(#"Valor
reemplazado6","Sugarcane","Caña de
Azúcar", Replacer. ReplaceText, {"Cultivo"}),
 #"Valor reemplazado8" = Table.ReplaceValue(#"Valor
reemplazado7","Vegetables","Vegetales",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado9" = Table.ReplaceValue(#"Valor
reemplazado8","Wheat","Trigo",Replacer.ReplaceText,{"Cultivo"}),
 #"Valor reemplazado10" = Table.ReplaceValue(#"Valor
reemplazado9","Crop Rotation","Rotación de
Cultivos", Replacer. ReplaceText, {"Estrategias_de_adaptacion"}),
 #"Filas filtradas" = Table.SelectRows(#"Valor reemplazado10", each true),
 #"Valor reemplazado11" = Table.ReplaceValue(#"Filas filtradas","Drought-
resistant Crops","Cultivos Resistentes a la
Sequía", Replacer. ReplaceText, {"Estrategias_de_adaptacion"}),
```

```
#"Valor reemplazado12" = Table.ReplaceValue(#"Valor reemplazado11","No Adaptation","Sin

Adaptación",Replacer.ReplaceText,{"Estrategias_de_adaptacion"}),

#"Valor reemplazado13" = Table.ReplaceValue(#"Valor reemplazado12","Organic Farming","Agricultura

Orgánica",Replacer.ReplaceText,{"Estrategias_de_adaptacion"}),

#"Valor reemplazado14" = Table.ReplaceValue(#"Valor reemplazado13","Water Management","Gestión del

Agua",Replacer.ReplaceText,{"Estrategias_de_adaptacion"})
```

Realizadas sobre la tabla "Regiones":

columna sin valores eliminada
 #"Columnas quitadas" = Table.RemoveColumns(dbo_Regiones,{"Paises"})

Tablas nuevas

- Creación de tabla calendario a partir de la función DAX "CALCULATE": CALENDAR(MIN(Paises[Fecha]), MAX(Paises[Fecha])).
- Nueva columna con el año de la fecha:

Año = 'Tabla Calendario'[Date].[Año]

- Creación de tabla de medidas calculadas.
- Creación de tabla resumen de regiones.
- Importé de Excel dos tablas de dimensiones: una para los países y la otra para los cultivos. Intenté crearlas en dax tomando los nombres de las dimensiones mediante la fórmula DISTINCT pero al intentar relacionar las tablas surgía el error "a circular dependency was detected" que ocurre cuando hay una dependencia cíclica entre tablas o columnas, lo cual impide que Power BI establezca la relación sin conflictos. Esto suele suceder cuando las columnas están definidas en términos de otras columnas que, directa o indirectamente, se refieren entre sí, o cuando el modelo de datos incluye relaciones o cálculos que dependen mutuamente.

Funciones DAX:

Utilizadas en tabla resumen de regiones:

```
SUMMARIZE(
Paises,

Regiones[Region],

"Temperatura Promedio",

ROUND(AVERAGE(Paises[Temperatura_promedio(C°)]),2),

"Precipitaciones Totales", ROUND(SUM(Paises[Precipitaciones_totales(mm)]),2),

"Emisiones Totales", ROUND(SUM(Paises[Emisiones_CO2(TM)]),2),

"Uso Promedio de Pesticidas(KG_por_HA)",

ROUND(AVERAGE(Paises[Uso_de_pesticidas(KG_por_HA)]),2),

"Uso Promedio de Fertilizantes(KG_por_HA)",

ROUND(AVERAGE(Paises[Uso_de_fertilizantes(KG_por_HA)]),2)

)
```

Utilizadas en tabla de medidas:

- Creada para calcular el impacto económico total Impacto Económico Total = SUM(Paises[Impacto_economico(Milliones_USD)])
- Creada para calcular las precipitaciones totales
 Precipitaciones Totales =
 - SUM(Paises[Precipitaciones_totales(mm)])
- Creada para calcular el promedio del índice de salud del suelo Promedio Índice Salud del Suelo = AVERAGE(Paises[Indice_de_salud_del_suelo])
- Creada para calcular los rendimientos totales de los cultivos Rendimiento_cultivos_total = SUM(Medidas[Rendimiento_cultivo(TM_por_HA)])
- 5. Creada para calcular la variación porcentual de la temperatura promedio del año actual (2024) con al anterior

```
Variación_temperatura =

VAR temp_promedio_2023 =

CALCULATE(AVERAGE(Principal[Temperatura_promedio(C°)]), 'Tabla

Calendario'[Año] = 2023)

VAR temp_promedio_2024 =

CALCULATE(AVERAGE(Principal[Temperatura_promedio(C°)]), 'Tabla

Calendario'[Año] = 2024)
```

RETURN
DIVIDE(temp_promedio_2023 - temp_promedio_2024, temp_promedio_2023, 0) * 100

Creada para calcular las emisiones totales de CO2
 Emisiones_CO2_Totales =
 SUM(Principal[Emisiones_CO2(MT)])

Información sobre solapas:

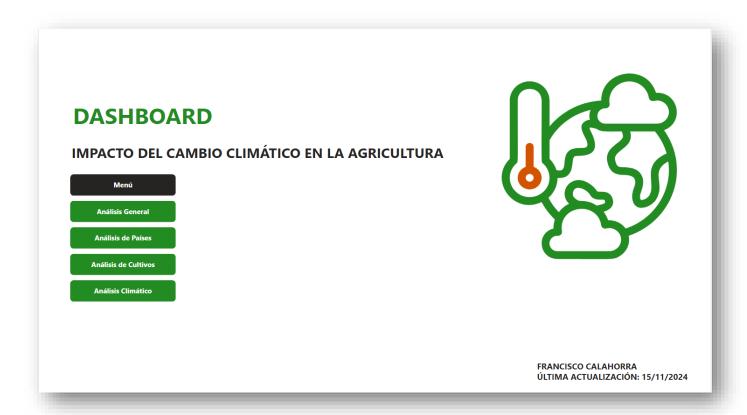
A continuación, se detalla un breve resumen de cada una de las páginas creadas:

Página "Menú": Es la página presentación del tablero de control, incluye información como el título, la última fecha de actualización y el menú para navegar a través de las diferentes solapas.

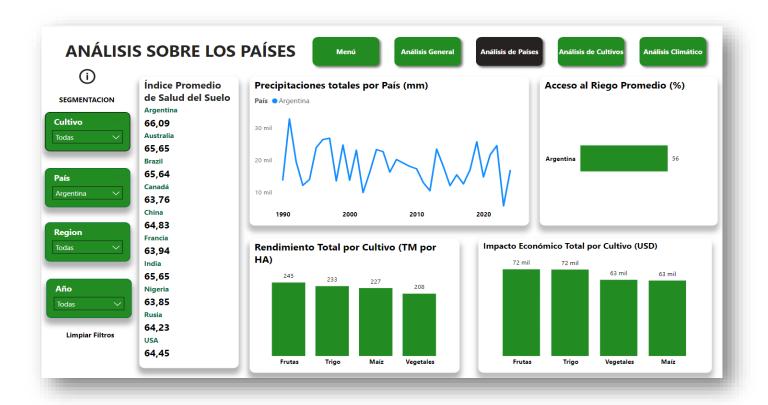
Página "Análisis General": esta página proporciona un análisis general abordando información útil para saber rápidamente cuáles son los países y los cultivos más rentables, así como las estrategias de adaptación más utilizadas, permitiendo segmentar por país, cultivo, región y año.

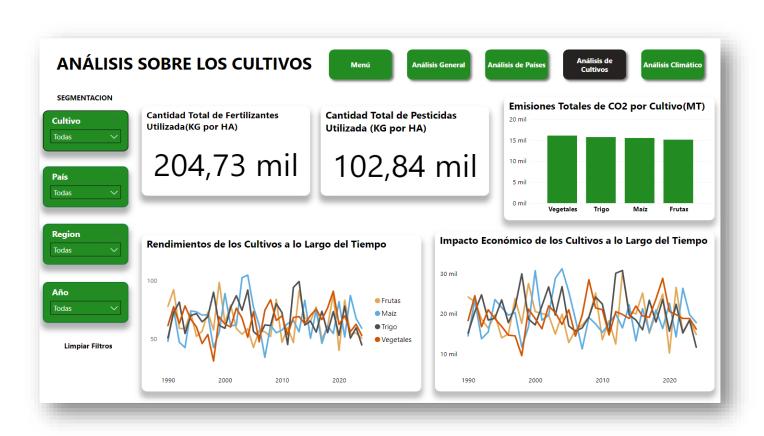
Página "Análisis de Países": esta solapa contiene información útil respecto al estado de situación de los países en cuestiones clave para el análisis como el índice de salud del suelo, las precipitaciones, el acceso al riego, y tanto el rendimiento que tienen los cultivos en los países como su impacto económico. La solapa cuenta con una segmentación por defecto a los efectos de mejorar la visualización del gráfico de líneas, se da aviso de esto mediante un botón de información.

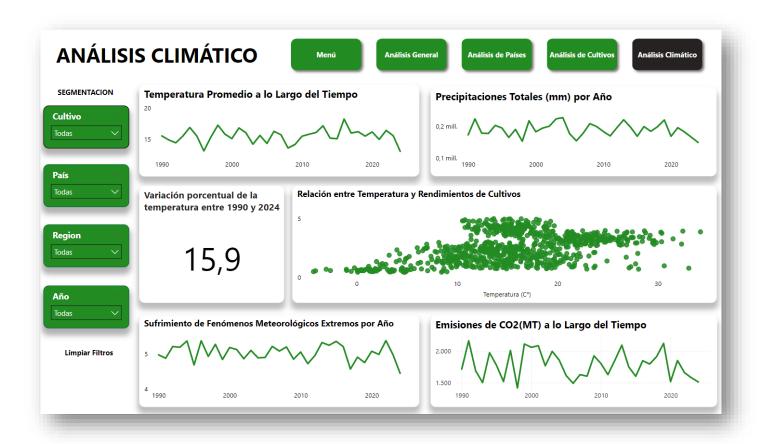
Página "Análisis de Cultivos": proporciona información sobre las cantidades utilizadas de pesticidas y fertilizantes, también sobre las emisiones de CO2 de cada cultivo, sumado a los rendimientos y el impacto económico que han tenido a lo largo del tiempo.











Manual de marca:

Tipografía: Segoe UI

Paleta de colores:

- (#228B22)
- (#E3A857)
- (#5DADE2)
- (#4F4F4F)
- (#D35400)

Conclusiones del análisis:

En términos generales los tres países donde más rinden los cultivos y consecuentemente más impacto económico tienen son Australia, la India y Canadá.

Australia

- Australia ocupa el segundo lugar en índices promedio de salud del suelo con un valor de 65,65, lo que indica buenas condiciones para la agricultura.
- Los cultivos de frutas son los más convenientes en términos de rendimiento e impacto económico.
- Las estrategias adaptativas más comunes incluyen el uso de frutales resistentes a la sequía como olivos, granados, higos, ciruelos y frutos secos como almendros, pistachos y nueces.
- La temperatura promedio anual fue de 12,56 °C, cercana al óptimo de 12,03 °C para el rendimiento de frutas.
- Los fenómenos meteorológicos extremos disminuyeron de 6,25 a 4,36 puntos respecto al año anterior, pero regiones como Western Australia (6,33) y Queensland (5) aún presentan índices altos.
- Se recomienda iniciar actividades en regiones como New South Wales (2,33) y Victoria (3,5), que son más seguras.

India

- India comparte el segundo lugar en índices promedio de salud del suelo con Australia (65,65).
- Los cultivos de vegetales tienen el mayor rendimiento, aunque los cultivos de trigo y frutas generan mayor impacto económico.
- La práctica adaptativa más común es la rotación de cultivos, combinando frutales y trigo para aprovechar sus ventajas.
- La temperatura promedio anual fue de 12,75 °C, favorable para los frutales, pero no para el trigo, que rinde mejor con 16,79 °C.
- Se sugiere comenzar con frutales y luego rotar al trigo cuando las temperaturas aumenten.

• West Bengal tiene el índice más alto de fenómenos meteorológicos extremos (7,25), por lo que sería mejor evitar esta región inicialmente y priorizar zonas como Tamil Nadu (3,75) y Maharashtra (3).

Canadá

- Canadá tiene el índice de salud del suelo más bajo entre los tres países (63,76), pero es el segundo en impacto económico total.
- El trigo es el cultivo de mayor rendimiento e impacto económico, favorecido por el uso de agricultura orgánica como método adaptativo.
- La temperatura promedio fue de 14,34 °C, cercana al óptimo de 16 °C para el trigo, y las temperaturas están en ascenso.
- Quebec es la región más segura ante fenómenos meteorológicos extremos (índice de 2), seguida de Ontario (4), Prairies (5) y British Columbia (5,75).
- Es recomendable enfocar las actividades en Quebec y Ontario por su menor riesgo climático.