

Facultad de Ingeniería



TALLER DE SOFTWARE

Trabajo Final - Foodset Dataset

Bloque

FC-PREISI02A01T

Profesora:

DEISY LIZBETH, ACOSTA TICSE

Integrantes

- Acosta Pérez, Rodrigo Alessandro
- Conde Soto, Dalo Josué
- Criado Nalvarte, Nicolás Patricio
- Curi Jaramillo, José Leonardo
- Espinosa Castellanos, Sebastián Daniel
- Guarniz Morales, José Jesús

Lima - Perú

2024-2

TABLA DE CONTENIDOS

Introducción	3
<i>Profundización sobre Food Tech</i>	3
<i>Idea del proyecto</i>	3
<i>Importancia y Justificación</i>	
Resumen	4
<i>Propósito</i>	4
<i>Metodología</i>	4
<i>Resultados</i>	4
<i>Conclusiones</i>	
Marco Teórico:	5
<i>Fundamentos de Programación</i>	5
<i>Tecnologías y Herramientas Utilizadas</i>	5
<i>Metodología de desarrollo</i>	5
Objetivos	6
<i>Objetivo General</i>	6
<i>Objetivos específicos</i>	6
Planificación y Diseño	7
<i>Planificación Business Model Canvas</i>	7
<i>Identificación de Requerimientos</i>	7
<i>Fuentes Literarias</i>	8
<i>Similitudes</i>	9
Desarrollo del Programa en Streamlit	11
<i>Requerimientos del programa</i>	11
<i>Desarrollo del programa</i>	12
<i>Ejemplo de Uso</i>	14
Comparación con otros programas	15
<i>Seaborn y matplotlib comparado con Streamlit</i>	15
<i>Power BI comparado con Streamlit</i>	16
<i>Tableau comparado con Streamlit</i>	16
Conclusiones	7
<i>Resumen de logros</i>	19
<i>Recomendaciones</i>	19
<i>Impacto en el Aprendizaje</i>	19
Referencias	20

INTRODUCCIÓN:

En el sector de alimentos, el uso de la tecnología ha transformado radicalmente la manera en que los alimentos son producidos, distribuidos y consumidos. Una de estas herramientas claves es FoodTech, esta consiste en una combinación entre la tecnología y la ciencia de los alimentos para resolver problemas críticos como la sostenibilidad, la seguridad y el desperdicio de alimentos. A través del uso de Python podemos dar soluciones a los problemas, ya que este lenguaje de programación es muy flexible y accesible. Podemos utilizar sus librerías como: Pandas que es fundamental para la manipulación y análisis de datos, NumPy proporciona el soporte para el cálculo de números, Matplotlib y Seaborn que nos ayudarían en la visualización de gráficos de los datos.

- **Profundización sobre Food Tech:**

El FoodTech está en el centro de la innovación en toda la cadena de valor de los alimentos, desde la agricultura digital hasta los alimentos cultivados en laboratorio y las aplicaciones inteligentes para reducir el desperdicio de alimentos. La agricultura sostenible, por ejemplo, se basa en el uso de datos y sensores para optimizar el uso de recursos como el agua y los fertilizantes. En la producción de alimentos, tecnologías como la fermentación pueden producir una variedad de leche y productos alimenticios sin el uso de animales. A través de modelos predictivos desarrollados con Python y otras herramientas de programación, las empresas pueden predecir la demanda de productos, optimizar la logística y mejorar la calidad. Esto incluye analizar big data de múltiples fuentes, incluidos sensores, bases de datos agrícolas y aplicaciones de clientes, para comprender y tomar mejores decisiones.

- **Idea del proyecto:**

El proyecto tiene como objetivo desarrollar una aplicación de entrega a domicilio FoodTech que mejore el estilo dietético de las personas a través de recomendaciones personalizadas, utilizando conjuntos de datos para ofrecer planes de alimentación, recetas y algoritmos de optimización.

- **Importancia y Justificación:**

La razón principal por la cual estamos investigando sobre Foodtech es porque creemos que esta tecnología de la industria alimenticia podría revolucionar de manera fundamental la calidad de alimentación de las personas, especialmente en este mundo donde los problemas alimenticios como la obesidad, la diabetes y varios problemas estomacales van aumentando drásticamente conforme pasan los años. y la mayoría de aplicaciones similares no logran poder ajustarse a las necesidades alimenticias de cada usuario.

Creemos que el desarrollo de la plataforma digital basada en Foodtech. que puede utilizar múltiples datasets y algoritmos para poder conseguir las mejores recomendaciones para cada usuario puede ser clave para solucionar este problema. Con este programa buscamos permitir a nuestros usuarios poder llevar a cabo una alimentación más ideal en base a sus objetivos y su salud

RESUMEN:

- **Propósito del proyecto:**

El propósito de este proyecto es optimizar la cadena de suministro de Impossible Foods, una organización del sector *FoodTech*, a través del análisis y la mejora de sus procesos de gestión de inventarios y distribución. El objetivo es identificar inefficiencias en la logística y el manejo de productos perecederos, implementando soluciones tecnológicas que permitan mejorar la visibilidad en tiempo real, reducir el desperdicio y minimizar los costos operativos.

El desafío abordado se centra en la gestión de su cadena de suministro global y en la logística de distribución de productos alimentarios a base de plantas. Esto incluye garantizar que los productos lleguen a tiempo a su destino y se mantengan frescos y en óptimas condiciones, al mismo tiempo que se minimizan los costos asociados al almacenamiento y transporte.

La implementación de esta solución generará un impacto directo en la eficiencia operativa, permitiendo a Impossible Foods mejorar la gestión de inventarios, reducir el desperdicio alimentario y aumentar la sostenibilidad en sus operaciones. Además, esto fortalecerá su posición en el mercado y contribuirá a cumplir su misión de reducir el impacto ambiental del sistema alimentario.

- **Metodología:**

Para desarrollar la solución tecnológica que optimizará la cadena de suministro de Impossible Foods, se utilizará la metodología ágil. Este enfoque se caracteriza por trabajar en ciclos cortos, comúnmente conocidos como *sprints*, lo que permite una mayor flexibilidad y adaptación a los cambios, así como la entrega continua de funcionalidades.

Entre los recursos a utilizar se encuentran los lenguajes de programación, encabezados por Python, cuya función será el análisis de datos y la implementación de algoritmos de optimización en la gestión de inventarios y distribución. También se emplea JavaScript para el desarrollo de interfaces de usuario web que permitan una interacción óptima con el sistema, y finalmente SQL para la gestión y consulta de bases de datos relacionadas con la logística y el inventario.

Además, se utilizarán herramientas como GitHub para el control de versiones y la colaboración entre desarrolladores, Power BI para la visualización de datos y análisis de KPIs relacionados con la cadena de suministro, y Jira para la gestión de proyectos, coordinando los *sprints* y el seguimiento de tareas.

- **Resultado:**

Los resultados clave que se obtendrán tras la implementación de la solución tecnológica serán: la mejora en la visibilidad en tiempo real, la reducción del desperdicio alimentario, la optimización de la logística de distribución mediante tecnologías avanzadas de gestión de rutas y transporte (TMS), la reducción de costos operativos, y, finalmente, una mayor sostenibilidad al minimizar el desperdicio y optimizar los recursos.

- **Conclusiones:**

Mediante el desarrollo del software se obtendrá la optimización integral de la cadena de suministro y además, mediante el enfoque ágil se permitirá una solución tecnológica flexible y adaptable que se

ajustara continuamente a las necesidades cambiantes del negocio, lo que garantizará su relevancia y efectividad a largo plazo.

MARCO TEÓRICO:

- **Fundamentos de Programación:**

Uno de los conceptos claves de este proyecto es el uso de Dashboards y Streamlight, los cuales serán esenciales para la creación de nuestro programa de Foodtech. Los Dashboards serán esenciales para visualizar los datos de forma comprensible y atractiva, mientras que Streamlit ofrece una interfaz simple y poderosa para desarrollar aplicaciones web de ciencia de datos sin la necesidad de conocimientos avanzados en desarrollo web. Con estas herramientas podemos hacer que la presentación de datos sea más sencilla y más fácil de comprender para los usuarios

- **Tecnologías y Herramientas Utilizadas:**

Este proyecto será desarrollado con el lenguaje Python, debido a que es un lenguaje excelente por sus capacidades de manejo de datos, además junto a sus bibliotecas externas como Panda, Numpy, Matplotlib, etc. Python se convierte en uno de los idiomas de programación más usados para la ciencia y gestión de datos e inteligencia artificial. También utilizaremos Github como la plataforma donde trabajaremos grupalmente el código del programa y tener acceso a versiones anteriores, lo que facilita la colaboración y la revisión de código.

- **Metodologías de Desarrollo:**

Para el desarrollo de este proyecto se utilizará la metodología ágil, una metodología caracterizada por trabajar en sprints, lo que permite una mayor flexibilidad y adaptación a los cambios, así como la entrega continua de funcionalidades. Con este tipo de metodología, el proyecto puede adaptarse fácilmente a los cambios en los requisitos que pueden presentarse a medida que vamos continuando con el proyecto, además de mejorar continuamente en base a la retroalimentación. Se trabajará en múltiples sprints que incluyen la planificación, diseño, desarrollo y pruebas, las cuales aseguran de que logremos crear un producto estable en cada iteración

OBJETIVOS:

- **Objetivo General:**

El objetivo general de este proyecto será diseñar una aplicación de delivery basada en Foodtech que pueda buscar mejorar el estilo alimenticio de las personas en base a sus preferencias y necesidades. La plataforma utilizará múltiples datasets para poder ofrecer las mejores recomendaciones adecuadas como diferentes planes de alimentación, recetas y varias sugerencias alimenticias, tomando en cuenta varios factores del consumidor como sus hábitos de consumo, consumo, las restricciones dietéticas, los objetivos nutricionales y la accesibilidad de los alimentos. Aparte los algoritmos también se irán cambiando en base a la experiencia del consumidor para ofrecerle lo más indicado para este.

- **Objetivos Específicos:**

- **Desarrollar un sistema de recomendación personalizado:**

El sistema debe tener un reconocimiento estará basado en algoritmos de aprendizaje automático que analiza información alimentaria del usuario como, sus hábitos de consumo, preferencias alimentarias, restricciones dietéticas y objetivos nutricionales.

- **Analizar múltiples tipos de datasets con diferentes tipos de datos:**

El sistema debe poder ser capaz de integrar múltiples tipos de datasets para poder mejorar y enriquecer las recomendaciones y garantizar una buena calidad de servicio

- **Implementar una interfaz amigable e intuitiva:**

Con un interfaz más simple y amigable de entender, los usuarios no tendrán problemas a la hora de registrar sus diferentes metas alimenticias (preferencias, objetivos y restricciones), etc

- **Desarrollar una funcionalidad de monitoreo de progreso:**

El sistema debe tener un reconocimiento estará basado en algoritmos

PLANIFICACIÓN Y DISEÑO:

- **Planificación Business Model Canvas:**

La planificación inicial del proyecto de software fue planificado mediante el siguiente enfoque estructurado y metodológico, basado en el Business Model Canvas y complementado con técnicas de análisis y diseño de sistemas. Con el modelo BMC pudimos alinear las metas del proyecto con las necesidades de los usuarios y las oportunidades del mercado

Componentes del BMC:

1. **Clientes:** Reconocimos que el tipo de público que buscamos atraer consistirá en diferentes tipos de personas como vegetarianos, fisicoculturistas, gente con alergia alimentaria
2. **Propuesta de Valor:**
La principal propuesta de valor de nuestro sistema consiste en la creación de una plataforma digital que pueda crear una dieta adecuada en base a las condiciones y preferencias alimenticias
3. **Canales de Distribución:**
Nuestro producto se conectara con el público de varias maneras como un sitio web oficial, cuentas de redes sociales y anuncios en Radio o TV
4. **Relaciones con Clientes:**
Buscamos poder establecer una relación con nuestros consumidores al darles una Experiencia personalidad y Comunicación efectiva
5. **Fuentes de Ingresos:**
Exploramos cómo nuestras fuentes de ingresos serán definidas por la venta directa de productos, ofertas y paquetes especiales y la publicidad de parte de marcas de alimentos
6. **Recursos Clave:**
Identificamos los datos alimenticios, algoritmos de aprendizaje automático, y un equipo interdisciplinario (desarrolladores, expertos en nutrición, y especialistas en experiencia del usuario) como los recursos esenciales para el éxito del proyecto.
7. **Actividades Clave:**
Definimos que las actividades clave durante el desarrollo del negocio consistirán en el mantenimiento del sistema de delivery, gestionar las relaciones con proveedores y la asociación con otros servicios de delivery
8. **Socios Clave:**
Establecimos que nuestros socios claves de nuestro proyecto consistirá en proveedores de alimentos orgánicos, múltiples empresas de delivery y varios nutricionistas especializados en las alimentaciones saludables
9. **Presupuesto:**
Reconocimos que los principales costos de nuestro proyecto consistirán en el desarrollo y mantenimiento de la compra, los costos de los ingredientes del menú, gastos de delivery, marketing y publicidad y entre otros costos administrativos

Socios Clave	Actividades Clave	Propuesta de Valor	Relacion con el cliente	Cilientes
<p>Los socios claves de nuestro proyecto consistirían en</p> <ul style="list-style-type: none"> • Proveedores de alimentos frescos y organicos • Diferentes empresas de delivery • Nutricionistas y otros expertos especializados en la alimentaciones sanudables <p>Agregue texto</p>	<p>Las actividades clave para el desarrollo del negocio consistiran en lo siguiente:</p> <ul style="list-style-type: none"> • Mantenimiento del sistema de delivery • Gestionar las relaciones con proveedores • Asociación con otros servicios de delivery <p>Recursos Clave</p> <p>Nuestra Servicio de Food Tech consistira en los siguientes recursos clave:</p> <ul style="list-style-type: none"> • Red de proveedores de alimentos • Algoritmo personalizado para conseguirte la mejor recomendación mas sana • Base de datos con la informacion 	<p>La principal propuesta de valor de nuestro SISTEMA:</p> <ul style="list-style-type: none"> • Creación de una plataforma digital que pueda crear una dieta adecuada en base a sus preferencias • Ofrecimiento de comidas saludables • Experiencia personalidad en base a las costumbres de los clientes <p>Agregue texto</p>	<p>Buscamos establecer una relacion con los clientes mediante la siguiente manera</p> <ul style="list-style-type: none"> • Experiencia personalidad • Comunicación efectiva • Plataforma web • Comunicación con el clientes mediante redes sociolas y notificaciones <p>Canales</p> <p>Nuestro producto se conectara con nuestros clientes de la siguiente manera:</p> <ul style="list-style-type: none"> • Sitio Web Oficial • Aplicación Móvil • Varias Redes Sociales (Facebook, Instagram, Twitter y Tiktok) • Anuncios comerciales en Radio o TV 	<p>Al publico que buscamos atraer consistirá en estas siguientes características</p> <ul style="list-style-type: none"> • Familias que buscan mantener un estilo de vida Saludable • Personas con problemas de salud o alimentación que buscan comer sano • Personas que buscan mejorar su fisicoculturismo • Personas que buscan opciones veganas adecuadas
Presupuesto		Fuente de Ingreso		
<p>Los diferentes costos que nuestro proyecto buscara consistiran en lo siguiente:</p> <ul style="list-style-type: none"> • Desarrollo y mantenimiento de la plataforma digital. • Costos de los ingredientes y preparacion de alimentos • Gastos de delivery • Inversión en marketing y publicidad. • Gastos generales y administrativos, incluyendo el personal de atención y soporte. 		<ul style="list-style-type: none"> • Venta Directa de Productos • Ofertas y Paquetes Especiales • Publicidad Dirigida y Colaboraciones • Publicidad con marcas de alimentacion saludable 		

● Identificación de Requerimientos:

- Creación de una plataforma digital:

Los usuarios deben poder crear un perfil donde especifique sus preferencias dietéticas, restricciones alimentarias (como alergias o intolerancias) y objetivos de salud (por ejemplo, pérdida de peso, ganancia muscular, dieta vegana).

- Ofrecimientos de comida saludables:

Crear un algoritmo que, basado en las preferencias y restricciones del usuario, genere planes de dieta diarios o semanales con comidas saludables.

- Experiencia personalidad:

Con la información de cada usuario proporcionar un catálogo de comidas saludables que cumplan con diversos tipos de dietas respectivas.

➔ Fuentes Literarias:

◆ (+) Ventajas del Foodtech, la revolución tecnológica de la alimentación - Ekos:

- Este sector está siendo promovido como el futuro de la industria alimentaria, gracias a su enfoque en la sostenibilidad, la eficiencia y la innovación tecnológica. Entre sus argumentos principales a favor serian poder aumentar la producción de alimentos y protección del medio ambiente,

◆ (+) FoodTech avanza hacia alimentos más saludables y proteínas alternativas innovadoras - The FoodTech:

- Otro dato a tomar en cuenta es de qué Food Tech busca mejorar la salud y nutrición de los alimentos que consumimos. Food Tech busca llegar a esta meta mediante la creación de alimentos más saludables y la reducción de ingredientes no deseados como azúcares y grasas sin comprometer el sabor

◆ **(=) Los Pros Y Los Contras De Foodtech Ventures - FasterCapital:**

- Al igual que otros negocios Foodtech tienes sus propias ventajas y desventajas, de las cuales destacan las siguientes
- **VENTAJAS:**
 - Existe un gran mercado en crecimiento para la tecnología alimentaria
 - La tecnología alimentaria puede ayudar a abordar los desafíos globales
 - La industria de la tecnología alimentaria tiene el potencial de tener un impacto positivo en algunos de los problemas más acuciantes del mundo, como el cambio climático y la inseguridad alimentaria.
- **DESVENTAJAS:**
 - La industria de la tecnología alimentaria es altamente competitiva
 - Muchas empresas de esta tecnología suelen fracasar
 - Muchas empresas emergentes de tecnología alimentaria fracasan, lo que dificulta que nuevas empresas alimentarias puedan navegar por el panorama regulatorio

◆ **(-) Aranca - Why Foodtech Investments Are Declining:**

- Foodtech originalmente atraía una gran inversión significativa, llegando a su máximo valor de en 2021, Sin embargo en 2023 la financiación del capital de riesgo ha decaído drásticamente debido a varios factores como la inflación, costos elevados de insumos y altas tasas de interés, lo que ha perjudicado el crecimiento y desarrollo de varios negocios start up the Food Tech

◆ **(-) Food Tech Startups - The one big thing that everyone seems to be missing.: A**

- Una de las principales críticas está relacionada con la sustentabilidad y viabilidad económica de muchos negocios startups de foodtech, puesto a que este modelo de negocio suele presentar malas unidades económicas. Por ejemplo, negocios como Foodpanda, Swiggy y Tiny Owl se encuentran con grandes dificultades que han dificultado su rentabilidad.

Similitudes

Las similitudes entre las fuentes muestran que las foodtech son buenas para el medio ambiente y mejoran la salud al proveer mejores alimentos, pero también expresan que es una industria competitiva y es afectada por la inflación, específicamente la subida de precios, mientras continúan siendo una industria rentable. También estresan que hay un pesimismo hacia estas ya que se asocia el fallo de una iniciativa al fallo de la industria, sin embargo, explican la industria es sana y con oportunidades tanto para mejorar, cómo hacer dinero.

Diferencia entre las librerías:

Para librerías, como ya antes mencionado, se utilizarán Numpy, Pandas, Matplotlib y Seaborn. Y a continuación se mencionan las principales diferencias y características de estas:

- **Pandas:** Está centrado en la manipulación de datos y en cuanto a cálculos matemáticos está muy por detrás de Numpy. Se especializa en bases de datos, hojas de cálculo. etc.

- **Numpy:** Se especializa en grandes conjuntos de datos numéricos y cálculos tanto científicos como matemáticos. A pesar de que destaca en datos numéricos, este es la base de Pandas para muchas operaciones ya que Pandas se construye sobre Numpy.
- **Matplotlib:** Permite crear muchísima variedad de gráficos, tales como líneas, dispersión, histogramas, gráficos de barras, etc. Destaca sobre Seaborn en situaciones donde se necesita un control preciso sobre cada aspecto de la visualización.
- **Seaborn:** Está construida sobre Matplotlib. Aunque no es tan avanzada para los detalles como Matplotlib, ofrece una mayor facilidad para la creación de gráficos complejos y con menos código. Se especializa más en visualizaciones estadísticas como los gráficos de distribución, gráficos de relaciones y gráficos categóricos. Puede integrarse fácilmente con Pandas, lo cual permite trabajar fácilmente con Dataframes.

A menudo Pandas y Numpy se usan juntas para proyectos de análisis de datos, así como Matplotlib y Seaborn se complementan en la visualización de datos.

Datasets

❖ Datasets encontrados:

- [Online Food Dataset](#)
- [ESTADO NUTRICIONAL DE LA POBLACIÓN PERUANA - CENAN](#)
- [Producción de Cultivos](#)
- [Food dataset](#)
- [Food Nutrition Dataset](#)
- [Food Delivery Dataset](#)
- [Zomato Dataset](#)
- [Restaurant Food Consumption](#)
- [USDA ERS - U.S. Food Imports](#)

❖ Cómo se identificaron los Datasets:

- Estos datasets fueron identificados por que contenían información vital para nuestro trabajo de Food Tech, Con la 1ra identificamos varias características de los consumidores para adaptarnos a sus preferencias, con la 2da identificamos el estado nutricional de las personas para ver qué podemos mejorar para su salud y nutrición y con la 3ra tenemos un record de cultivos a nivel global con el que podemos identificar con qué cultivos podemos trabajar en cada región del mundo.

DESARROLLO DEL PROGRAMA EN STREAMLIT

Requerimientos del Programa:

Carga de Datos:

- Los usuarios cargan datasets en formatos como CSV o Excel en la interfaz.
- El sistema procesa los datos automáticamente para análisis.

Preprocesamiento de Datos:

- **Limpieza:** Eliminación de duplicados, valores vacíos o inconsistentes.
- **Normalización:** Formatos homogéneos (texto en minúsculas, fechas uniformes).
- **Conversión:** Adaptación de datos a los tipos necesarios (números, fechas, texto).

Análisis de Datos:

- **Descriptivo:** Tablas y resúmenes estadísticos para entender las características generales.
- **Exploratorio:** Visualizaciones interactivas que identifican patrones, tendencias y anomalías.

Visualización de Gráficos:

- Uso de Streamlit para generar gráficos dinámicos.
- Los usuarios ajustan las visualizaciones mediante filtros para analizar variables específicas.

Generación de Reportes:

- Resúmenes de análisis presentados en un formato claro y exportable.

Dashboards Interactivos:

- Plataforma integrada en Streamlit para interactuar con los datasets, mostrando gráficos y resultados en tiempo real.

Desarrollo del Programa

Utilizamos **Streamlit** porque permite crear aplicaciones interactivas de manera rápida y sencilla, ideal para transformar análisis de datos en herramientas accesibles para usuarios no técnicos. Su interfaz intuitiva facilita cargar archivos, procesar datos y generar visualizaciones dinámicas en tiempo real, lo que es especialmente útil para explorar tendencias, presentar resultados y colaborar fácilmente al compartir la aplicación a través de un enlace. Es una solución eficiente para convertir scripts en Python en aplicaciones prácticas y funcionales.

Streamlit

El programa comienza importando las librerías necesarias: **Streamlit** para la interfaz interactiva, **Pandas** para manipulación de datos y **Matplotlib** para gráficos. Se establece un título principal para la aplicación: "Visualización de datos con Streamlit".

```
import streamlit as st
import pandas as pd
import matplotlib.pyplot as plt

# Título
st.title("Visualización de datos con Streamlit")
```

La aplicación permite a los usuarios cargar un archivo en formato **CSV** o **Excel** mediante un componente interactivo. Si el archivo es un Excel con varias hojas, se proporciona la opción de seleccionar una hoja específica. Una vez cargado, se muestra una vista previa de las primeras cinco filas de los datos.

```
# Subir un archivo
uploaded_file = st.file_uploader("Carga tu archivo Excel o CSV", type=["csv", "xlsx"])

if uploaded_file is not None:
    # Verificar tipo de archivo
    if uploaded_file.name.endswith(".csv"):
        # Leer archivo CSV
        data = pd.read_csv(uploaded_file)
        st.write("Vista previa de los datos:", data.head())
    elif uploaded_file.name.endswith(".xlsx"):
        # Leer archivo Excel
        excel_file = pd.ExcelFile(uploaded_file)
        sheet_names = excel_file.sheet_names # Obtener nombres de las hojas

        if len(sheet_names) == 1:
            # Si solo hay una hoja, cargarla directamente
            data = excel_file.parse(sheet_names[0])
            st.write(f"Vista previa de la única hoja '{sheet_names[0]}':", data.head())
        else:
            # Si hay varias hojas, permitir seleccionar
            selected_sheet = st.selectbox("Selecciona una hoja:", sheet_names) # Seleccionar hoja
            data = excel_file.parse(selected_sheet) # Cargar hoja seleccionada
            st.write(f"Vista previa de la hoja '{selected_sheet}':", data.head())
```



Después de cargar el archivo, el usuario puede seleccionar una columna del dataset usando un menú desplegable. Esto permite enfocar el análisis en una variable específica.

```
# Seleccionar columna para graficar
column = st.selectbox("Selecciona una columna para graficar", data.columns)
```

Selecciona una hoja:

Contenido

Vista previa de la hoja 'Contenidos':

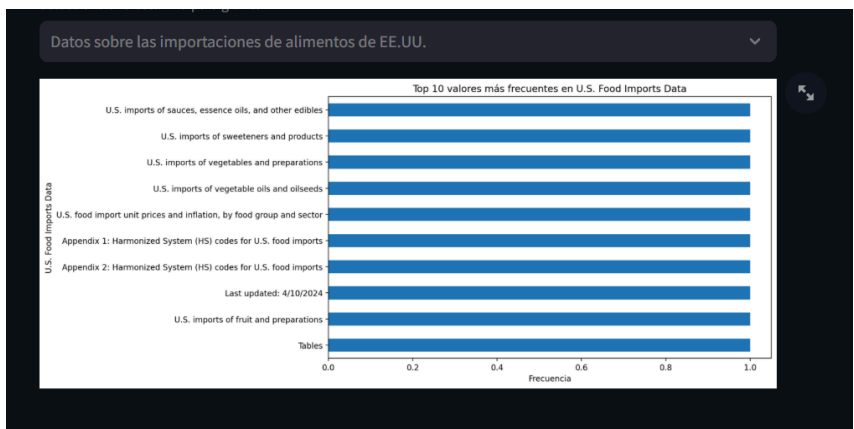
	U.S. Food Imports Data
0	None
1	Tables
2	U.S. Food Import Value
3	U.S. food import value, by food group
4	U.S. food import value, by food group, percent change from last year

Selecciona una columna para graficar

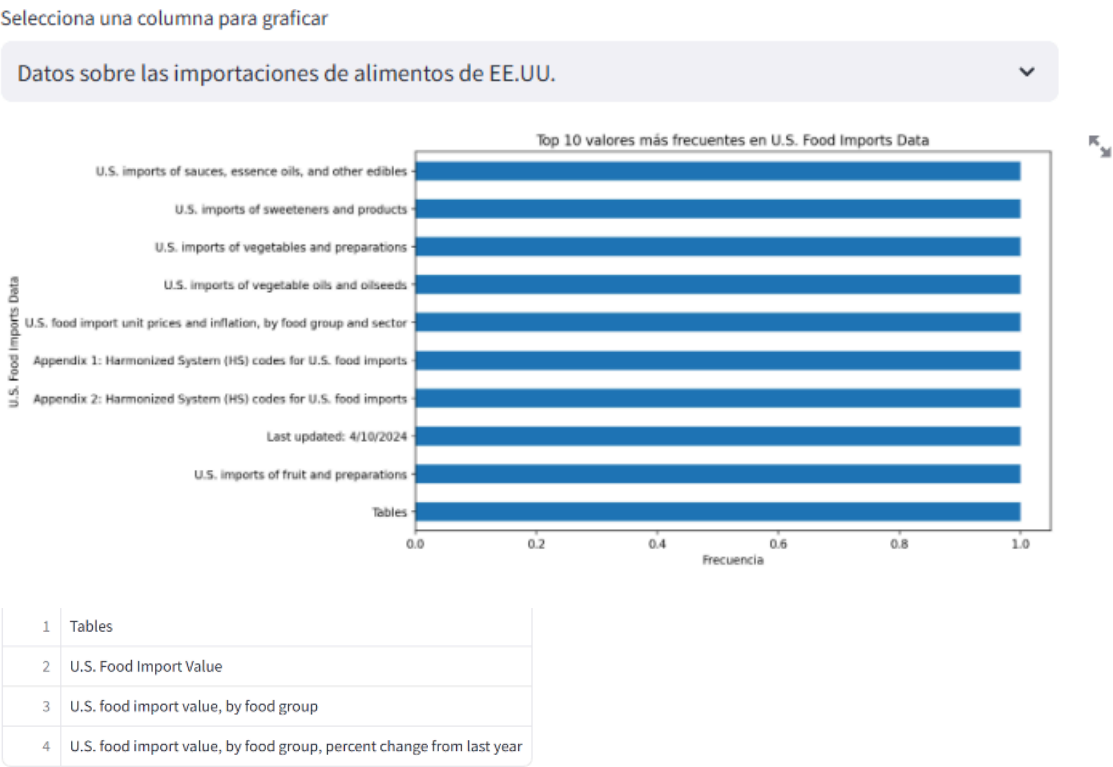
Datos sobre las importaciones de alimentos de EE.UU.

El programa calcula los 10 valores más frecuentes de la columna seleccionada y genera un gráfico de barras horizontal. Este gráfico muestra la frecuencia de aparición de cada valor, ordenado para facilitar su interpretación.

```
# Graficar
if column:
    # Ordenar y seleccionar los valores más importantes
    value_counts = data[column].value_counts().head(10) # Los 10 más frecuentes
    fig, ax = plt.subplots(figsize=(10, 6)) # Tamaño ajustado del gráfico
    value_counts.sort_values(ascending=True).plot(kind="barh", ax=ax) # Gráfico de barras horizontal
    ax.set_title(f"Top 10 valores más frecuentes en {column}")
    ax.set_xlabel("Frecuencia")
    ax.set_ylabel(column)
    st.pyplot(fig)
```



Ejemplo de uso de Streamlit



Usar Streamlit en este proyecto fue una decisión estratégica debido a su capacidad para manipular y visualizar datos de manera intuitiva y eficiente. La plataforma permite a los usuarios cargar archivos en formatos como CSV o Excel, eliminando la necesidad de herramientas externas o conocimientos avanzados de programación. Además, Streamlit ofrece visualizaciones dinámicas, como gráficos de barras horizontales, que facilitan la identificación de patrones y tendencias en tiempo real. Su simplicidad, flexibilidad y capacidad para manejar grandes volúmenes de datos hacen de Streamlit una herramienta ideal para analizar y presentar resultados de manera clara y profesional, especialmente en proyectos relacionados con datos alimentarios

COMPARACIÓN CON OTROS PROGRAMAS

Compararemos nuestro programa con Streamlit con otros programas de análisis de datos .

Streamlit es una librería para construir aplicaciones web interactivas en Python. Está diseñada para mostrar visualizaciones en un entorno interactivo.

Las ventajas son:

- **Interactividad:** Los usuarios pueden ajustar parámetros, como filtros, sliders, botones y ver actualizaciones en tiempo real.
- **Fácil creación de aplicaciones:** No se necesitan habilidades avanzadas en desarrollo web.
- **Integración con otras herramientas:** Puedes incorporar gráficos de Matplotlib, Seaborn, y bibliotecas como Plotly o Altair.
- **Despliegue web:** Fácil de compartir con otras personas mediante un navegador.

Sus desventajas son:

- **Más complejo que otros programas:** Si solo necesitas gráficos estáticos, Streamlit puede ser innecesariamente avanzado.
- **Dependencia de un servidor:** Necesitas ejecutar la aplicación en un servidor o en tu computadora para verla.

Seaborn y Matplotlib comparado con Streamlit:

- **Descripción de Matplotlib:** Es la librería más básica y versátil para visualización en Python. Muchas otras bibliotecas (como Seaborn) están construidas sobre Matplotlib. Las ventajas que tiene Matplotlib.

Ventajas:

- Permite un control total sobre todos los aspectos del gráfico (colores, tamaños, posiciones, etc.).
- Soporta una variedad de formatos y tipos de gráficos.
- Sirve para gráficos simples y avanzados.

Desventajas:

- La creación de gráficos avanzados pueden requerir más código y tiempo.
- Los gráficos por defecto no son tan atractivos como los de Seaborn.
- Los gráficos que se producen son estáticos (no interactivos).

- **Descripción de Seaborn:** Es una extensión de Matplotlib diseñada para visualización estadística. Ofrece gráficos estéticamente atractivos con menos esfuerzo.

Ventajas:

- Para gráficos avanzados no se requiere tanto código.
- Es ideal para analizar distribuciones, relaciones, y patrones en datos.
- Los gráficos por defecto tienen un estilo moderno y profesional.
- Trabaja muy bien con DataFrames de Pandas, lo que facilita la visualización de grandes volúmenes de datos estructurados.

Desventajas:

- Es más fácil de usar que Matplotlib, pero tiene menos flexibilidad para personalizaciones específicas.
- Los gráficos que se producen son estáticos (no interactivos).

Power BI comparado con Streamlit:

- **Descripción de Power BI:** Es una herramienta de visualización y análisis de datos que permite a los usuarios conectar, transformar y visualizar datos de diversas fuentes para crear dashboards e informes interactivos.

VENTAJAS DE POWER BI:

- **Facilidad de uso:** Está diseñado principalmente para usuarios de negocios sin requerir un nivel avanzado de programación. La interfaz en buena parte consiste en arrastrar y soltar íconos/objetos para crear visualizaciones.
- **Soporte empresarial y seguridad:** Está diseñado para cumplir con requisitos importantes para una empresa, como el hecho de ofrecer capacidades de seguridad, tales como el control de acceso, la autenticación mediante Active Directory, y la posibilidad de gestionar usuarios y permisos a nivel granular.
- **Visualizaciones preconfiguradas:** Tiene un amplio conjunto de visualizaciones predefinidas, como gráficos de barras, líneas, mapas, tablas y más, y permite personalizarlas mediante una interfaz gráfica. Además, las visualizaciones pueden ser altamente interactivas, lo que es muy útil para los informes ejecutivos y de negocios.

DESVENTAJAS DE POWER BI:

- **Flexibilidad y personalización:** Tiene limitaciones en la personalización y el control de diseño y la interactividad.
 - **Capacidad de desarrollo:** Aunque tiene algunas capacidades de scripting con DAX y Power Query, no permite el mismo nivel de personalización avanzada como Streamlit, especialmente cuando se trata de integrar modelos predictivos o lógicos complejos directamente en las visualizaciones.
 - **Costo Alto:** La versión gratuita de Power BI ofrece muchas funcionalidades, pero muchas características avanzadas requieren una licencia 'Power BI Pro' o 'Power BI Premium', lo que puede resultar costoso, especialmente para pequeños equipos o empresas.
- **Comparación gráfica:**

[Restaurant Food Consumption en Power BI](#)

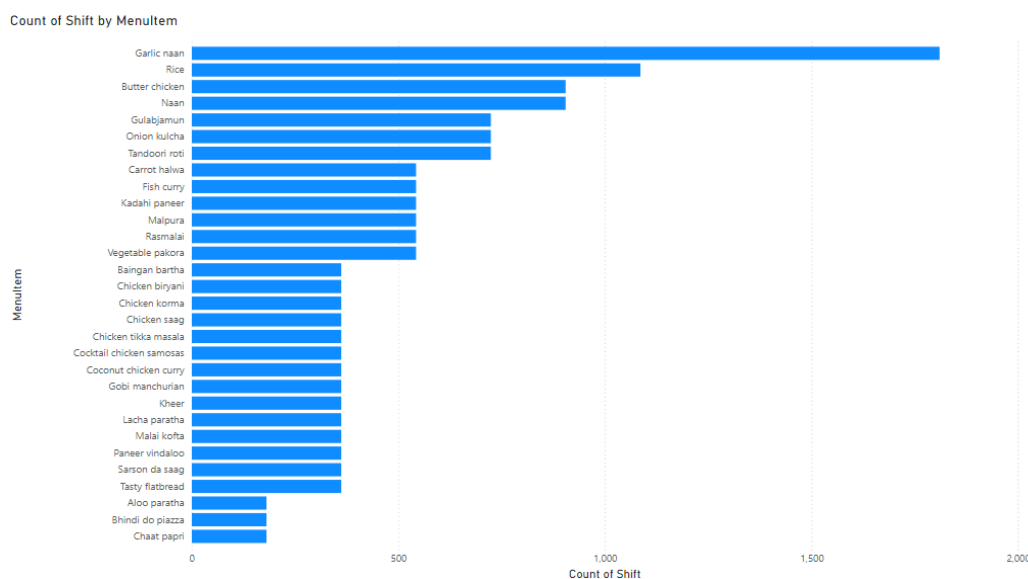


Tableau comparado con Streamlit:

- **Descripción de Tableau:** Es una plataforma de análisis visual que fue diseñada para ayudar a las organizaciones o personas a comprender y tomar decisiones basadas en información visual. Analizar y visualizar datasets de manera más sencilla. Esta funciona en base a tomar largas cantidades de datos y poder transformarlos en tablas, estadísticas y datos complejos. Además es bastante sencilla de usar

VENTAJAS DE TABLEAU:

- **INTERFAZ SENCILLA:** Con su interfaz simplista y fácil de utilizar, hace que Tableau no requiere demasiada experiencia para la creación de dashboards y visualizaciones de datos
- **HERRAMIENTAS COLABORATIVAS:** Los dashboards creados pueden compartirse para poder trabajar en grupo de manera similar como un Google Drive
- **COMPATIBILIDAD CON VARIOS LENGUAJES:** Tableau integra algunos lenguajes como R o Python para poder hacer análisis más complejos

DESVENTAJAS DE TABLEAU:

- **ALTO COSTO DE USO:** Las licencias de Tableau son caras, lo que no puede hacerse viable a empresas pequeñas o recomendable para uso personal.
- **LIMITACIÓN DE TRANSFORMACIÓN DE DATOS:** Tableau no es la herramienta ideal para realizar procesos ETL (Extract, Transform, Load), lo que significa que los datos ya deben de estar limpios y estructurados
- **PROBLEMAS DE RENDIMIENTOS:** Tableau puede experimentar problemas de rendimiento al manejar datasets demasiado grandes

- **Comparación Gráfica**

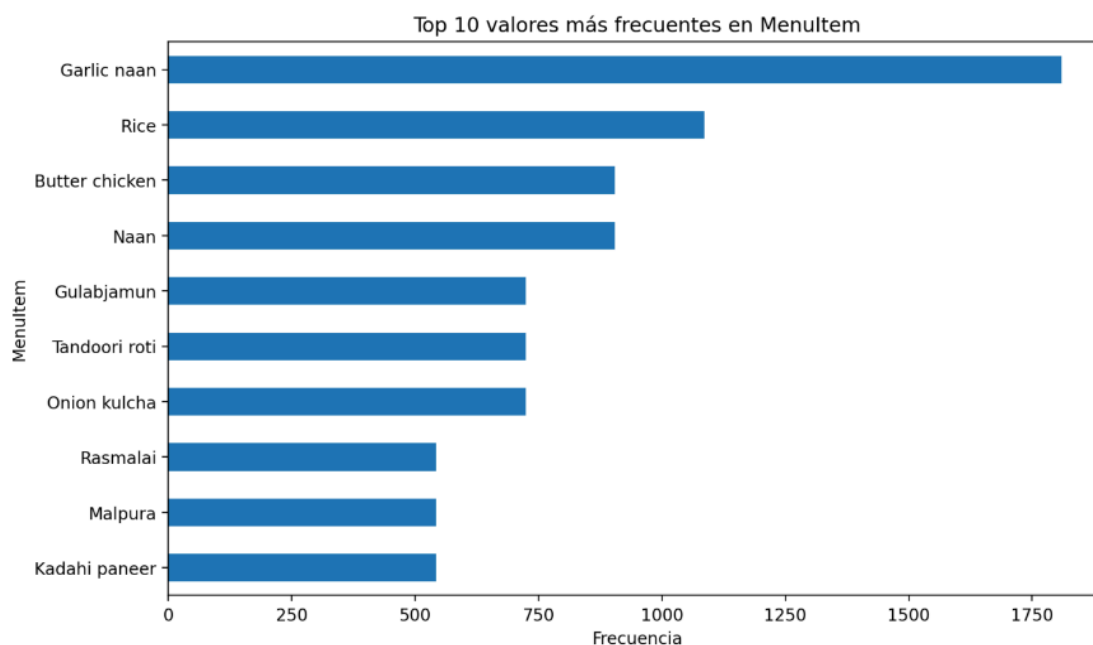
Ejemplo de Uso de Tableau con el dataset [Restaurant food consumption.xlsx](#)



Ejemplo de Uso de Streamlit con el mismo dataset

Selecciona una columna para graficar

MenuItem



Para la comparación de ambas estadísticas dadas por ambos problemas hemos hecho la siguiente gráfica:

ASPECTO	ESTADÍSTICA DE TABLEAU	ESTADÍSTICA DE STREAMLIT
Nivel de Detalle	Alto detalle (Clasificado por día y turno)	Resumen General del dataset
Facilidad de Comprensión	Un poco complejo, se requiere mayor tiempo de comprensión y análisis	Sencillo, se enfoca en el top 10 elementos
Aplicación	Análisis por tiempo y segmentación	Identificación de Elementos principales
Información Clave	Patrones Diarios y por turno	Popularidad general de los productos

Comparación con r

descripción del lenguaje: R es un lenguaje de programación especializado en uso estadístico creado para trabajar con base de datos y operaciones estadísticas. creado en 1992. Su misión es mejorar y facilitar el uso de estadística avanzada en las ciencias de la computación.

Ventajas de r

- Diseñado específicamente para uso estadístico.
- Librerías especializadas

Fácil de aprender al inicio

Desventajas de r:

- menos flexible que python
- más complicado de aprender a niveles avanzados
- Rendimiento más pesado

```
library(shiny)
library(readr) # Para leer CSV
library(readxl) # Para leer Excel
library(ggplot2) # Para gráficos

# Interfaz de usuario
ui <- fluidPage(
  titlePanel("Visualización de datos con Shiny"),

  sidebarLayout(
    sidebarPanel(
      fileInput("file", "Carga tu archivo CSV o Excel",
        accept = c(".csv", ".xlsx")),
      uiOutput("sheet_selector"), # Selector de hojas dinámico
      uiOutput("column_selector") # Selector de columnas dinámico
    ),

    mainPanel(
      tableOutput("data_preview"), # Vista previa de datos
      plotOutput("bar_plot")      # Gráfico de barras
    )
  )
)
```

```

# Servidor
server <- function(input, output, session) {
  # Leer archivo cargado
  data <- reactive({
    req(input$file) # Asegurarse de que el archivo esté cargado

    ext <- tools::file_ext(input$file$name)

    if (ext == "csv") {
      read_csv(input$file$datapath)
    } else if (ext == "xlsx") {
      readxl::excel_sheets(input$file$datapath) # Devuelve nombres de las hojas
    }
  })

  # Mostrar selector de hojas si el archivo es Excel
  output$sheet_selector <- renderUI({
    req(data())

    if (inherits(data(), "character")) { # Si es un Excel (lista de hojas)
      selectInput("sheet", "Selecciona una hoja:", choices = data())
    }
  })
}

```

```

# Obtener datos finales según la hoja seleccionada
final_data <- reactive({
  req(data())

  if (inherits(data(), "data.frame")) {
    data() # Retorna el DataFrame directamente (para CSV)
  } else if (inherits(data(), "character") && !is.null(input$sheet)) {
    read_excel(input$file$datapath, sheet = input$sheet)
  }
})

# Mostrar selector de columnas
output$column_selector <- renderUI({
  req(final_data())
  selectInput("column", "Selecciona una columna para graficar:",
    choices = names(final_data()))
})

# Vista previa de los datos
output$data_preview <- renderTable({
  req(final_data())
  head(final_data(), 10) # Muestra las primeras 10 filas
})

```

```

# Generar gráfico de barras
output$bar_plot <- renderPlot({
  req(final_data(), input$column)

  data_to_plot <- final_data()[[input$column]]
  data_to_plot <- as.data.frame(table(data_to_plot))
  data_to_plot <- data_to_plot[order(data_to_plot$Freq, decreasing = TRUE), ]

  ggplot(data_to_plot[1:10, ], aes(x = reorder(data_to_plot, Freq), y = Freq)) +
    geom_bar(stat = "identity") +
    coord_flip() + # Gráfico horizontal
    labs(x = input$column, y = "Frecuencia",
         title = paste("Top 10 valores más frecuentes en", input$column))
})
}

# Ejecutar aplicación
shinyApp(ui, server)

```

CONCLUSIONES

Resumen de los logros:

- **Optimización de la cadena de suministro:** Se logró desarrollar una solución tecnológica que mejora significativamente la visibilidad en tiempo real de inventarios y procesos logísticos.
- **Impacto en la sostenibilidad:** La adopción de tecnologías avanzadas no solo disminuyó los gastos operativos, sino que también disminuyó el impacto ecológico, alineándose con las metas de sostenibilidad de la industria FoodTech.
- **Desarrollo de capacidades tecnológicas:** Se utilizaron recursos como Python, Streamlit y Power BI para el análisis de datos y la representación interactiva de información.

Recomendaciones:

- **Escalabilidad:** Incrementar la capacidad de la plataforma para incorporar más datasets y manejar cantidades más grandes de datos facilitará la atención a un mayor número de usuarios y situaciones.
- **Pruebas de usuario:** Implementar etapas de prueba con usuarios reales para verificar la eficacia del algoritmo y hacer modificaciones basándose en sus observaciones.
- **Educación y divulgación:** Fomentar la formación de los usuarios acerca de la relevancia de adoptar costumbres alimenticias sustentables y saludables mediante campañas incorporadas en la plataforma.

Impacto en el aprendizaje:

- **Fortalecimiento de habilidades técnicas:** La implementación de bibliotecas sofisticadas de Python y recursos como Streamlit promovió un aprendizaje práctico en ciencia de datos, creación de software y visualización interactiva.
- **Comprensión de metodologías ágiles:** La experiencia con sprints y herramientas como Jira fomenta habilidades fundamentales en la administración de proyectos y trabajo en equipo.
- **Visión integral del FoodTech:** El proyecto brindó un entendimiento más detallado de cómo la tecnología puede revolucionar la industria de la alimentación, desde la sostenibilidad hasta el mejoramiento de la salud de los consumidores.

- **Desarrollo de pensamiento crítico:** La combinación de datos variados y la creación de algoritmos a medida demandaron un análisis exhaustivo y la solución de problemas complicados.

REFERENCIAS:

- TraceX Tech. (2024). 2024 food industry trends: Tech and sustainability insights. <https://tracextech.com/food-industry-trends-2024/>
- Negocios, E., & Negocios, E. (2022, 13 junio). 5 ventajas del Foodtech, la revolución tecnológica de la alimentación. Ekos Negocios. <https://ekosnegocios.com/articulo/5-ventajas-del-foodtech-la-revolucion-tecnologica-de-la-alimentacion>
- Santiago, Judith. (2024, julio 11). Siete tendencias que marcan el futuro del foodtech. THE FOOD TECH - Medio de noticias líder en la Industria de Alimentos y Bebidas; THE FOOD TECH. <https://thefoodtech.com/tendencias-de-consumo/siete-tendencias-que-marcan-el-futuro-del-foodtech/>
- World Economic Forum. (2024). *Innovators reimagine the future of food systems*. <https://www.weforum.org/agenda/2024/05/innovators-reimagine-the-future-of-food-systems/>
- Tech, A. (2022, 6 Agosto). FoodTech: Qué es, usos y ventajas de esta nueva revolución de la alimentación. <https://www.giztab.com/foodtech-que-es-usos-ventajas/>
- AVP, Publications. (2024, Julio 22). Why foodtech investments are declining. Aranca. <https://www.aranca.com/knowledge-library/articles/business-research/why-foodtech-investments-are-declining>
- Author, G. (2021, 8 Abril). Food Tech startups - the one big thing that everyone seems to be missing. YourStory. <https://yourstory.com/2015/11/foodtech-startups-missing>