

Propuesta de Proyecto

Maximizar el impacto y la relevancia del turismo gastronómico, en el sector de la industria turística.

Análisis de mercado en Estados Unidos.



Roles a desempeñar

- Carlos Rivera: Data Analytics/ Data Scientist
- Darwin Moreno: Data Engineer/Data Scientist
- Soledad Rozzatti: Data Analytics
- Stefany Rojas: Data Analytics
- Wilson Waigant: Data Engineer/Data Scientist

Contexto y entendimiento del problema

El turismo gastronómico en Estados Unidos es una tendencia en auge que puede potenciar y enriquecer la experiencia turística en el país. Esta forma de turismo se centra en la exploración de la diversidad culinaria y ofrece a los viajeros la oportunidad de descubrir y disfrutar de la rica oferta gastronómica que Estados Unidos tiene para ofrecer.

Una de las características más destacadas de este tipo de turismo es la amplia variedad de platos típicos y especialidades regionales. Desde los sabores auténticos del sur hasta la fusión de influencias internacionales en las principales ciudades, el país ofrece una gama diversa de opciones culinarias para satisfacer todos los gustos. Los viajeros tienen la oportunidad de probar nuevos sabores, experimentar la cultura local a través de la comida y deleitarse con experiencias gastronómicas únicas.

Además, el turismo gastronómico no se limita solo a la comida, sino que también se trata de las experiencias y actividades que la rodean. Los turistas pueden participar en recorridos gastronómicos para explorar barrios históricos, probar platos tradicionales y aprender sobre la historia y la cultura local. También, pueden disfrutar de clases de cocina interactivas, visitas a viñedos y cervecerías artesanales, y participar en festivales y eventos culinarios que celebran la diversidad gastronómica del país. Por ello, el turismo gastronómico es un recurso fundamental para promover la industria turística en general, ya que se beneficia con el crecimiento de la demanda de hoteles, restaurantes y servicios relacionados.

En conclusión, el turismo gastronómico en Estados Unidos tiene el potencial de potenciar y enriquecer la experiencia turística al ofrecer una variedad de sabores, experiencias culinarias únicas y la posibilidad de sumergirse en la cultura local a través de la comida.



Contexto del cliente

TURFOOD es una empresa filial del conglomerado TURISK. LTD. La cual busca, ofrecer experiencias gastronómicas únicas y memorables en el sector del turismo gastronómico. Por ello, busca establecer alianzas sólidas con socios estratégicos, especializados en la analítica de datos, los cuales, brinden para su portafolio de servicios las mejores soluciones respecto a recomendaciones, análisis y evaluaciones de datos.

Con base en lo expuesto, TURFOOD decide emplear los servicios ofrecidos por Datum Tech, la cual es una consultora externa especializada en la analítica de datos, en busca de soluciones, que impacten de manera positiva su portafolio en el sector del turismo gastronómico.

Las principales razones para ello son:

1. Experiencia y conocimientos especializados: Datum Tech cuenta con un equipo de expertos altamente capacitados y con experiencia en el campo de la analítica de datos. Su amplio abanico de soluciones les permite añadir valor agregado a los servicios ofrecidos por TURFOOD en el sector del turismo gastronómico.
2. Mejora de la toma de decisiones: TURFOOD reconoce la importancia de tomar decisiones basadas en el análisis de datos. Al contratar a Datum Tech, esperan que la consultora pueda proporcionar información valiosa y de calidad, permitiendo a TURFOOD optimizar sus servicios y ofrecer experiencias gastronómicas únicas y memorables a sus clientes.
3. Recomendaciones estratégicas: Datum Tech tiene la capacidad de analizar y comprender los datos más complejos, y transformarlos en señales que identifican patrones, tendencias y oportunidades de mejora. Con esta información, se crean mejores sistemas de recomendaciones personalizadas para mejorar la experiencia del cliente, optimizar los procesos internos y maximizar la rentabilidad de TURFOOD.



Objetivos

Objetivo general:

Implementar una herramienta de análisis de datos escalable para comprender mejor a los consumidores de determinados productos o servicios en el ámbito de la hotelería y gastronomía, y en base a ello tomar las mejores decisiones de negocio.

Objetivo específicos:

- Disponibilizar información relevante y de calidad, extraída desde fuentes externas, necesaria para realizar el análisis de mercado solicitado.
- Evaluar las emociones, actitudes y opiniones de los usuarios, para obtener información que permita comprender la imagen que los mismos tienen respecto de un determinado producto o servicio ofrecido por nuestro cliente.
- Predecir qué rubros de negocios tendrán mayor y menor crecimiento asociado.
- Determinar potenciales oportunidades para futuros nuevos locales.
- Crear un sistema de recomendación de locales que pertenezcan al conglomerado de nuestro cliente.

Alcance

El análisis se enfocará en los datos obtenidos desde dos plataformas externas virtuales (Yelp y Google Maps). El mismo tendrá en cuenta el análisis cuantitativo del nivel de satisfacción de los usuarios, que permitirá un mejor conocimiento de los consumidores. De esta forma se logrará una mejor toma de decisiones y estrategias de mercado, basadas en información precisa y de calidad.

Adicionalmente, el proyecto se sustenta en la implementación de un sistema de recomendación de restaurantes y afines basado en el conocimiento del interés de los usuarios.



KPIs - Key performance indicator

Los KPIs son herramientas esenciales para la gestión y el control efectivos, ya que brindan una visión clara y medible del desempeño y rendimiento de una organización.

Sus características principales son : específico, medible, alcanzable, relevante, tiempo determinado.

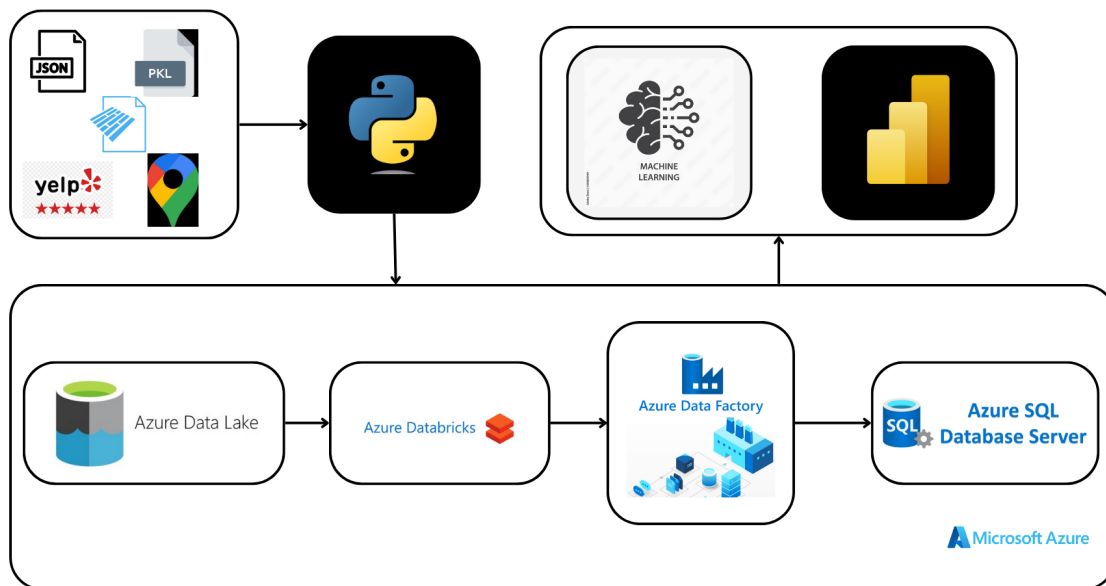
KPIs

KPIs	Característica	Descripción	Fórmula	Rango
Índice de satisfacción del cliente	Permite medir si un servicio cumplió con las expectativas del cliente	Crece un índice de satisfacción 5% cada mes	$\text{Reseñas Promedio} = \frac{\text{Reseñas sumadas}}{\text{Reseñas totales}}$	Mensual
Porcentaje de clientes que repiten el servicio	Permite mantener un flujo constante de clientes	se estima en el valor de 35%, siendo este el valor objetivo a superar	$\text{Porcentaje de clientes} = \left(\frac{\text{clientes que repiten el servicio}}{\text{total de clientes que califican}} \right) * 100$	Mensual
Tasa de crecimiento de nuevos clientes	Proporciona una visión general y una evaluación de la atracción de nuevos clientes en la industria de restaurantes	en este caso del 10% cada mes, se busca superar el valor estimado y lograr una mayor fidelización de los clientes	$\text{Tasa de crecimiento} = \left(\frac{\text{Nuevos clientes} - \text{Clientes anteriores}}{\text{Clientes anteriores}} \right) * 100$	Mensual
Tasa de crecimiento de reseñas	Permite evaluar el impacto de las experiencias de los clientes en la reputación y el éxito del negocio.	Aumentar la tasa de crecimiento de reseñas en un 15% cada mes	$\text{Tasa de crecimiento de reseñas} = \left(\frac{\text{Cantidad reseñas} - \text{Cantidad reseñas}}{\text{Cantidad reseñas}} \right) * 100$	Mensual



Sentimiento Promedio	Proporciona una medida cuantitativa de la satisfacción del cliente.	mantener un sentimiento promedio alto, lo más cercano posible al 100%.	Sentimiento Promedio = (suma de los valores de sentimiento de las opiniones / número total de opiniones) * 100	Mensual
----------------------	---	--	--	---------

Stack Tecnológico



Usaremos Python como lenguaje principal y Azure como proveedor de servicios en la nube.

1. Almacenamiento de datos:

Utilizaremos Azure Data Lake Storage (ADLS) como sistema de almacenamiento para crear un Data Lake. Es altamente escalable y compatible con múltiples formatos de datos, incluyendo Parquet y JSON. Que son, entre otros, formatos de archivos recibidos.



2. Procesamiento de datos:

Utilizaremos Azure Databricks, que es una plataforma de análisis y procesamiento de Big Data que admite Python y tiene integración directa con ADLS. Esto nos permitirá ejecutar el código necesario para transformar los datos.

3. Transformación y carga incremental:

Utilizaremos la biblioteca PySpark para leer los datos en los distintos formatos que recibimos. Aprovecharemos las capacidades de procesamiento distribuido de PySpark para acelerar el proceso.

Para la carga incremental, compararemos los nuevos datos con los datos existentes en el Data Lake utilizando alguna columna clave o timestamp.

4. Automatización:

Utilizaremos Azure Data Factory (ADF) para orquestar y automatizar el pipeline de la extracción, transformación y carga de datos. Data Factory nos permitirá definir y programar las actividades de ETL en un flujo de trabajo secuencial.

5. Disponibilización de datos:

Una vez que los datos se hayan procesado y cargado en el Data Lake, utilizaremos ADF para orquestar la copia de los datos limpios a una ubicación específica como Azure SQL Database accesible por Power BI en tiempo real y para ser consumidos por el modelo de ML planteado.



Metodología SCRUM

La metodología elegida para el desarrollo de este proyecto es la “ Metodología Scrum”. Scrum es un proceso en el cual se aplican de manera regular un conjunto de buenas prácticas para trabajar colaborativamente, en equipo, y obtener el mejor resultado posible de un proyecto. Estas prácticas se apoyan unas a otras y su selección tiene origen en un estudio de la manera de trabajar de equipos altamente productivos.

Consiste en realizar entregas parciales y regulares del producto final, priorizadas por el beneficio que aportan al receptor del proyecto. Donde los requisitos pueden ser cambiantes, donde la innovación, la competitividad, la flexibilidad y la productividad son fundamentales.

Planilla de metodología SCRUM:

Sprint Actual: Semana 1

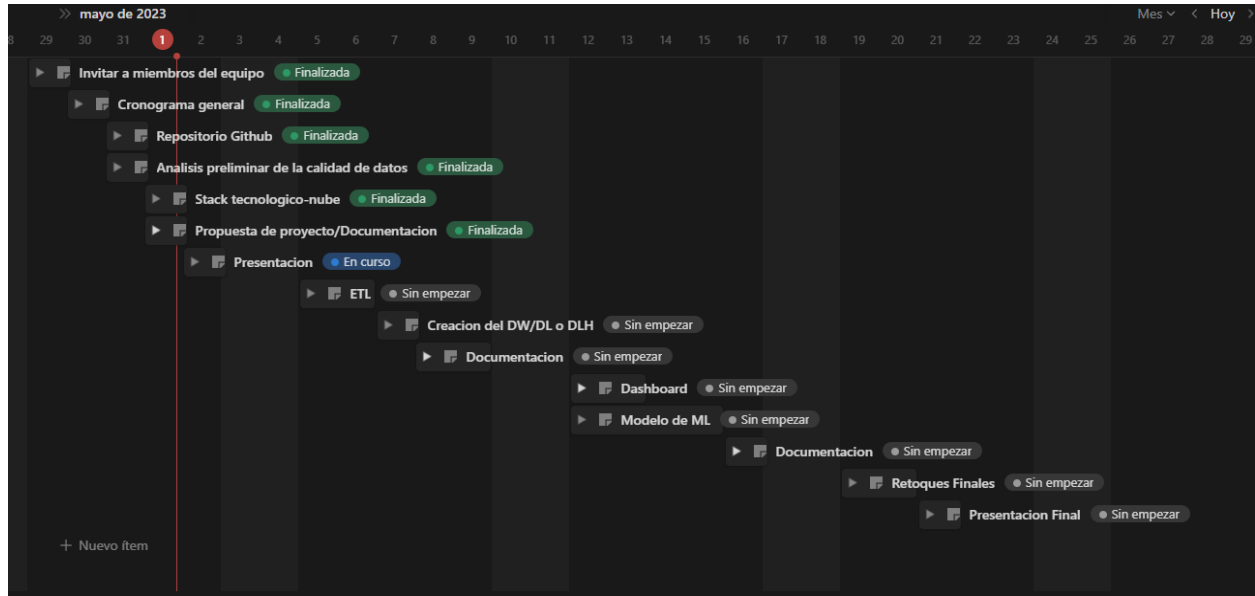
Fecha de Inicio: 30 de mayo

Fecha de Fin: 2 de junio

Objetivos del sprint:

- Contexto y entendimiento de la problemática
- Objetivos específicos
- Alcance
- KPIs asociados
- Repositorio de Github
- Stack tecnológico
- Asignación de roles



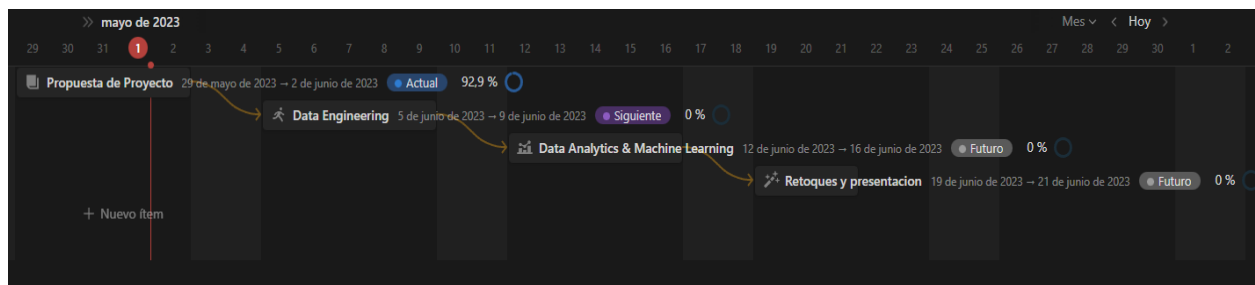


Tareas completadas:

1. Contexto y entendimiento de la problemática (100%)
2. Objetivos específicos (100%)
3. Alcance (100%)
4. KPIs asociados (100%)
5. Repositorio de Github (100%)
6. Stack tecnológico (100%)
7. Asignación de roles (100%)

Tareas pendientes:

- Sprint 2
- Sprint 3
- Sprint 4



Reuniones planificadas:

- Reunión de Planificación del Sprint: Diarias, 17-hrs
- Reunión Diaria de Scrum: Diarias, 11-hrs
- Revisión del Sprint: 2 de junio - 11 hrs
- Retrospectiva del Sprint: 2 de junio - 11 hrs

Productos a entregar:

Primera semana:

- Solución a la propuesta de proyecto.

Segunda semana:

- Modelo de almacenamiento de datos.

Tercera semana:

- Dashboard interactivo con KPIs para su análisis.
- Modelo de recomendación de lugares gastronómicos para ser consumido por el usuario.
Extra:
 - Modelo de recomendación de hospedaje con ruta gastronómica para ser usado por el cliente.



Conclusiones:

Creemos firmemente en la aplicación de data science en el turismo gastronómico para ofrecer experiencias excepcionales a nuestros clientes. Estamos emocionados por implementar este proyecto y confiamos en alcanzar nuestros objetivos, consolidándonos como líderes en el sector. La implementación de este proyecto beneficiará a nuestros clientes al ofrecer experiencias gastronómicas personalizadas, aumentar la satisfacción y generar mayor lealtad. Además, la adopción de data science nos brindará una ventaja competitiva al tomar decisiones informadas y mejorar continuamente nuestras ofertas y servicios. Estamos comprometidos a aprovechar al máximo el poder de los datos para transformar la forma en que se experimenta y disfruta la gastronomía, esperando con entusiasmo los resultados positivos y el éxito duradero que este proyecto puede brindarnos.

