






# Proyecto Final – Google & Yelp

## Web recomendación de hoteles



### Team

-  Delfina Longo
-  Miguel Dallanegra
-  Carlos Hidalgo
-  Jonathan
-  Angel Prieto

### Sprint – 1

29 enero – 09 febrero – 2024

# Entendiendo la situación propuesta

En un mercado saturado de plataformas de reservas de hoteles, HotelWise se destaca como un oasis de calidad y confiabilidad. No nos conformamos con ofrecer una simple comparación de opciones de alojamiento; nuestra misión es proporcionar una experiencia de usuario excepcional y de primer nivel que va más allá de las expectativas tradicionales.

¿Qué nos diferencia de las demás plataformas?

**Selección Cuidadosa de Hoteles:** En HotelWise, nos enorgullece ofrecer una cuidadosa selección de hoteles de alta calidad. No se trata solo de proporcionar opciones; se trata de garantizar que cada hotel en nuestra plataforma cumplirá a cabalidad las expectativas del usuario.

**Reseñas Auténticas y Detalladas:** Entendemos la importancia de la transparencia y la confianza en la industria hotelera. Por eso, cada reseña en HotelWise es auténtica y detallada, proporcionada por viajeros reales que han experimentado de primera mano lo que cada hotel tiene para ofrecer. Creemos en la honestidad y la integridad en cada paso del camino.

**Información Confiable sobre Zonas Seguras:** La seguridad de nuestros usuarios es una prioridad absoluta. Por eso, ayudaremos a encontrar el hotel perfecto, sino que también te proporcionamos información confiable sobre las zonas seguras de cada destino. Queremos que nuestros usuarios viajen con tranquilidad y confianza, sabiendo que están en buenas manos.

En HotelWise, no solo estamos en el negocio de ofrecer habitaciones de hotel; estamos en el negocio de crear recuerdos inolvidables y experiencias de viaje extraordinarias. Si estás buscando invertir en una plataforma que se enorgullece de su compromiso con la calidad, la transparencia y la satisfacción del cliente, HotelWise es la elección perfecta.

# Alcance del proyecto

El proyecto parte con la base de datos ubicada en el país de Estados Unidos, abarcará de 51 Estados, donde se centrará la atención principalmente en el rendimiento del tiempo de retorno de información y el valor otorgado por la analítica para el convencimiento del Usuario al momento de la elección del Hotel recomendado.

La ejecución de este constará de 6 semanas donde se utilizarán Sprint's para la división de tareas a realizar y metas a cumplir, donde en el primer sprint semana 1 y 2 se definen los objetivos, metodologías, tecnologías, tareas y roles. La Semana 3 y 4 se realizará la extracción, disposición y análisis de datos; y por último la semana 5 y 6 el desarrollo del sistema de recomendación mediante machine learning. En paralelo a las semanas 3, 4, 5 y 6 se llevará a cabo el diseño y despliegue de la web.

El mismo será escalable gracias a que se implementará un flujo de extracción y disposición de datos automatizado, por lo que permitirá en un futuro ir agregando países junto con su división política.

# Objetivo General

Presentar una plataforma web que realice recomendaciones de hoteles, mediante machine learning utilizando las comodidades y zona geográfica preferidas por el usuario, donde el tiempo de muestra de la recomendación sea menor a 90 segundos, utilizando herramientas en la nube con un tiempo de implementación no mayor a 6 semanas.

## Objetivos Específicos

- Crear dos bases de datos una para filtrar las zonas utilizando Cloud SQL y para el modelo de machine learning con Firestore (Firebase).
- Crear Análisis de datos con Looker para medir rendimiento de la plataforma y mostrar indicadores clave para convencer al usuario de la recomendación del hotel.
- Crear un modelo de recomendación utilizando AutoML para la generación de valor obtenido de los datos al usuario.
- Crear una plataforma web intuitiva y eficiente al momento de realizar recomendaciones al usuario.

## KPI's plataforma

- Tiempo de obtención de datos de Cloud SQL.
- Tiempo de obtención de datos de Firestore.
- Tiempo de proceso de ML

## KPI's Usuario

- Porcentaje de similitud de amenidades del hotel con relación a las amenidades preferidas del usuario.
- Distancia del punto o zona de interés del usuario.
- Porcentaje de reseñas positivas sobre las amenidades del hotel.
- Índice de seguridad de la zona del hotel.
- demanda por meses o época.

# Solución propuesta

Implementación del Stack Tecnológico.



OBS: de no funcionar Composer se estaría reemplazando por Workflow.

# Descripción Stack



Cloud Storage: se almacenarán en los Buckets los datos en brutos a ser procesados por los flujos de datos y los DAG's junto con los operadores de Python.



Composer: Airflow – Orquestador donde estará deployado los DAG's, encargado de transmitir los datos limpios y extraídos de otras fuentes.



Cloud API's: conexión a otras bases de datos para la obtención de datos significativos para enriquecer la base de datos.



Big Query: encargado de realizar la conexión entre las Bases de datos y otros flujos de trabajo.



Looker: Herramienta para la elaboración y visualización de Dashboards.



AutoML: Herramienta para la elaboración del producto de Machine Learning, recomendación de hoteles.



Cloud SQL y Firestore: Base de datos.

# Metodología de trabajo

El proyecto se divide en 3 etapas denominadas Sprints, cada sprint constará de 2 semanas, donde en cada una tendrá 3 reuniones mediante google meet, en esta los miembros del equipo harán reporte sobre el trabajo realizado, los inconvenientes y la labor que realizarán de cara al próximo meet; mientras que el Mentor dará feedback a las tareas realizadas.

## **Sprint 1** del 29 de enero al 09 febrero 2024

- Entendimiento de la problemática.
- Documentación de Alcance, Objetivos.
- KPI's.
- Implementación del Stack.
- Github y EDA preliminar.
- Planificación – Gantt.

## **Sprint 2** del 12 al 23 de febrero 2024

- Creación de entornos GCP.
- Creación de flujo de datos automatizados en Composer.
- Creación del Data Warehouse en Cloud SQL y Firestore.
- Análisis de datos preliminar.
- MVP Dashboards.
- MVP producto ML.

## **Sprint 3** 26 de febrero al 09 de marzo 2024

- Dashboards final.
- Reportes geográficos.
- Storytelling (relacionar KPI's e Insights).
- Producto ML Ajustado.
- Demo general del proyecto.
- Correcciones y retoces en base a feedback de Mentor y Product Owner.
- Entregable final.
- Documentación.

# Roles y responsabilidades

## Data Engineer



Jonathan Gutiérrez (Reviews) | Angel Prieto (Meta-Data)

Las responsabilidades incluyen la creación y gestión de entornos en Google Cloud Platform (GCP), configuración y administración de flujos de datos automatizados en Apache Airflow, y la creación y mantenimiento de un Data Warehouse en Cloud SQL y Firestore, abarcando el diseño de esquemas, desarrollo de procesos ELT y gestión continua para optimización y calidad de los datos.

## Data Analyst

Delfina Longo | Carlos Hidalgo



Las responsabilidades abarcan la implementación y administración de bases de datos en BigQuery, el desarrollo de paneles interactivos en Looker Studio para análisis de datos, y la organización de tareas y seguimiento del progreso mediante la plataforma ClickUp. Además, se incluye la creación y mantenimiento de KPIs específicos.

## Data Cientist | Web Master

Miguel Dallanegra



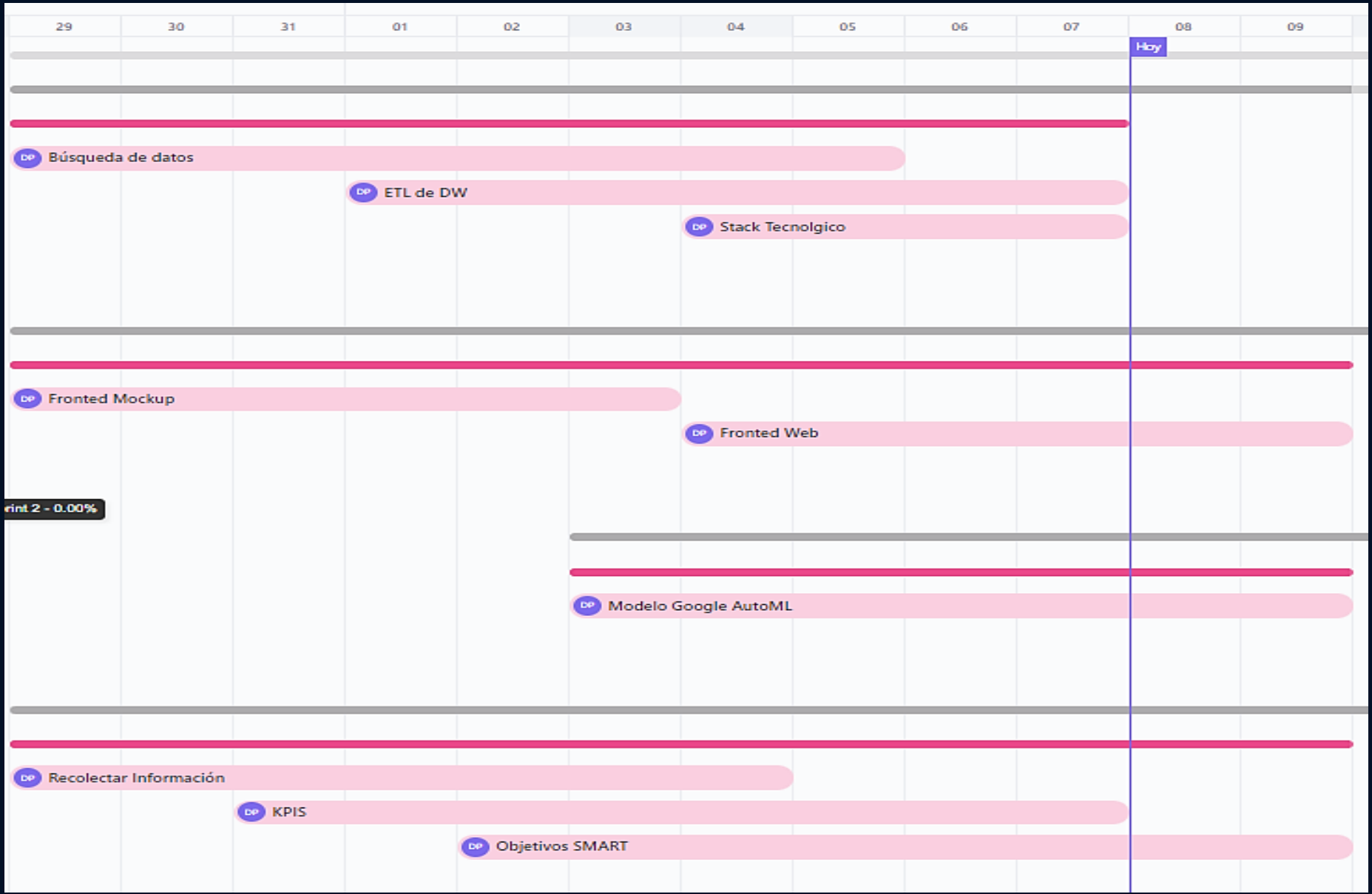
Las responsabilidades abarcan la selección del entorno de trabajo, biblioteca y modelo de machine learning, para luego implementarlo, entrenarlo, evaluarlo y validarlo, de manera que finalmente pueda ser desplegado y puesto en producción, devolviendo recomendaciones de hoteles.

Diseño y desarrollo del sitio web del proyecto, implementando las características y funcionalidades necesarias para poder hacer las mejores recomendaciones de hoteles. Dando a los clientes una experiencia superadora y otorgando información clara sobre sus consultas, haciendo hincapié en la facilidad para navegar el sitio.

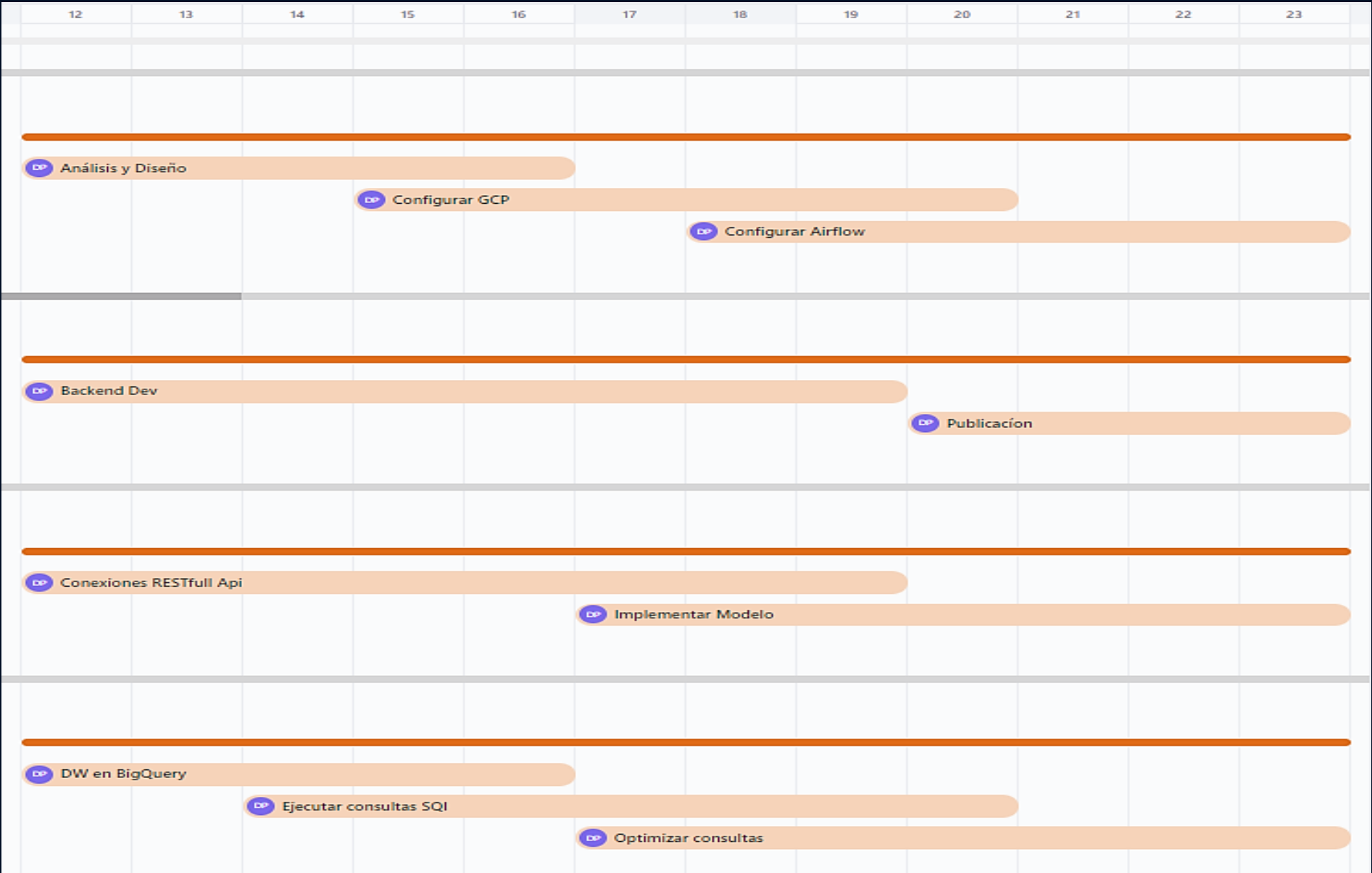


# Cronograma General Gantt

## Sprint 1



# Sprint 2



# Sprint 3

