



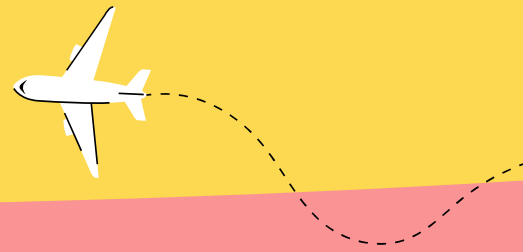
PROYECTO B.I.

VUELOS COMERCIALES

DS - COH3 - GRUPO 17



INTEGRANTES



Alan Sánchez	Data Analyst
Alejandro Aguilera W	Data
Christian Fajardo	Data Engineer
Eduardo Carhuaricra	Data Analyst







SITUACION ACTUAL



Retraso y suspensión de vuelos por varios factores:

- Incertidumbre sobre el mercado de vuelos comerciales(post-covid).
 - Estado del clima y otros.
 - Falla en los itinerarios.
 - Problemas técnicos
- 
- 

Objetivos

- **Describir la variabilidad en la cancelación de vuelos** y factores que intervienen en la misma.
- **Describir a través de mapas la realidad**, la concentración de vuelos, principales destinos y las rutas mas frecuentes teniendo como punto de origen EE.UU



- **KPI's Propuestos:**

Reducir cancelación de vuelos.

Determinar mejoras de logística de despegue de vuelos.

Reducir destinos poco frecuentes.

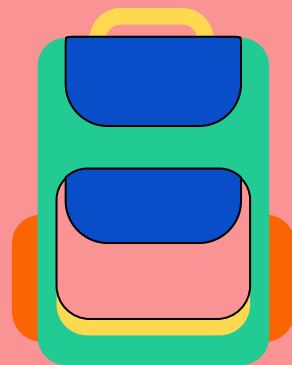
Prever cancelaciones de vuelos derivadas de condiciones climáticas.



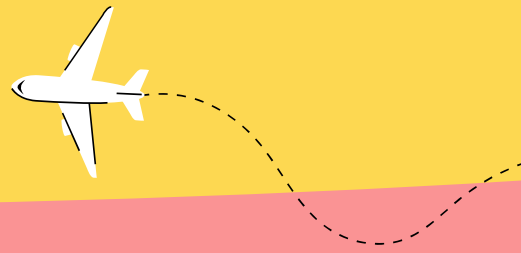
ALCANCES

El presente estudio enfocará su análisis en vuelos nacionales tomando como punto de origen **únicamente EEUU**, por concentrar el mayor tránsito aéreo.

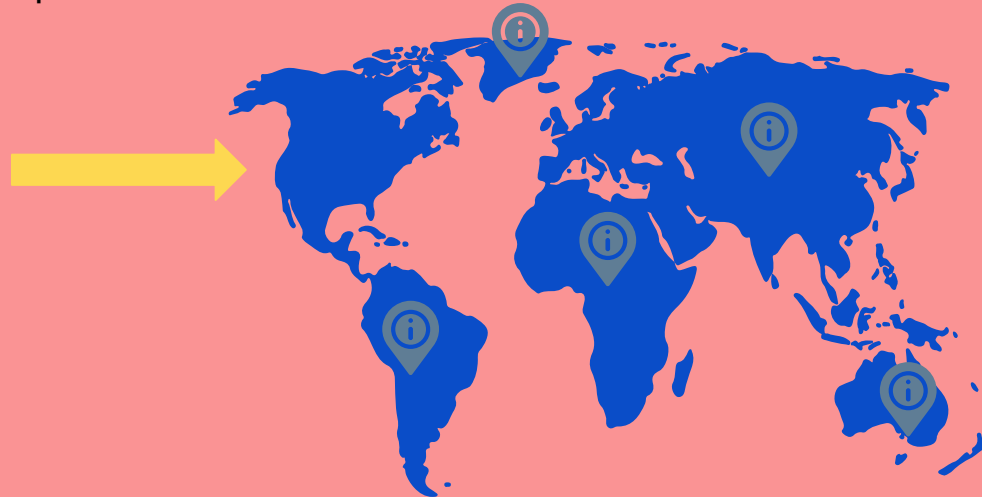
El estudio además incluirá un **análisis en base a datos respecto a los factores externos** como los climatológicos, tráfico aéreo, cierre de aeropuertos, etc. y factores internos como fallas en los itinerarios, problemas técnicos.



FUERA DE ALCANCE



El proyecto no abordará información sobre vuelos pertenecientes a países fuera de EEUU, ni en los itinerarios por problemas técnicos.



PROPUESTA DE SOLUCIÓN



La propuesta sería implementar una arquitectura de datos compuesta por un Cloud Data Lake(GCS) como repositorio único de los datos iniciales(raw data) con pipelines ETL automatizados y orquestados para depositar los datos que estén listos (processed data) en un Data Warehouse (GBQ) para aplicarle BI y ejecutar los dashboards(PowerBI) para así tener con claridad los datos analizados.



PROPUESTA DE SOLUCIÓN



Herramientas

A

Apache Airflow

Vênus é o segundo planeta mais próximo do Sol

P

Python

Júpiter é o maior planeta do Sistema Solar

M

Google Cloud Platform

Apesar de ser vermelho, Marte é verdadeiramente frio

P

PowerBi

Saturno é composto por hidrogênio e hélio



Metodología de trabajo

Durante el desarrollo de este proyecto usaremos la metodología ágil Scrum, estableciendo dailys que durarán 15 min, donde cada integrante informará de avances, oportunidades, y la tarea diaria.



Etapas del proyecto

Semana 1

Kickoff del proyecto
Indagación del tema
Objetivos

1



Semana 3

Reportes
Storytelling

3

Creación del Data Lake
Creación del Data
Warehouse
EDA

Semana 2

2

4

Demo final
Documentación

Semana 4





GRACIAS

