

Business Understanding

Objetivos analíticos

1. Entender la tasa de puntualidad de cada aerolínea comparando tiempos programados y reales
2. Comprender la relevancia de las diferentes causas de demora y su distribución
3. Identificar las aerolíneas y aeropuertos con mayor concurrencia de uso.
4. Identificar las rutas más afectadas por desvíos y cancelaciones

KPI's Iniciales

1. Porcentaje de vuelos puntuales
Fórmula: $(\text{Vuelos sin retraso} / \text{Total de vuelos}) \times 100$
2. Porcentaje de desvíos
Fórmula: $(\text{Vuelos desviados} / \text{Total de vuelos}) \times 100$
3. Porcentaje de cancelación
Fórmula: $(\text{Vuelos cancelados} / \text{Total de vuelos}) \times 100$
4. Porcentaje de desvíos
Fórmula: $(\text{Vuelos desviados} / \text{Total de vuelos}) \times 100$
5. Porcentaje de vuelos con retraso por aerolínea
Fórmula: $(\text{Vuelos con retraso} / \text{Total de vuelos por aerolínea}) \times 100$

Lista de requerimientos analíticos

- Software: Python, SQL Server, Power BI, Docker, Office (Excel)
- Infraestructura: Computadora con al menos 10GB de espacio en disco disponible y 8GB de RAM, conexión a internet.

Preguntas de negocio

- ¿Qué aerolíneas y aeropuertos presentan mejor o peor puntualidad?
- ¿Cuáles son las causas de demora más relevantes y cómo se distribuyen por aerolínea/aeropuerto?
- ¿Qué rutas son más vulnerables a retrasos, cancelaciones o desvíos?
- ¿Qué tan diferentes son los tiempos programados vs reales y qué factores influyen en esa brecha?
- ¿Qué aeropuertos con mayor concurrencia muestran congestión operacional?
- ¿Las rutas largas tienen mejor o peor puntualidad que las rutas cortas?

Canvas Data Product

Problema <ul style="list-style-type: none">Se desconoce qué aerolíneas presentan mejor o peor desempeño en puntualidad.No se sabe qué causas de demora son más frecuentes o relevantes.No está claro cuáles aeropuertos y rutas concentran mayor tráfico y riesgo operativo.Las cancelaciones y desvíos impactan la eficiencia del sistema y requieren identificación.	Datos <p>Se tiene un dataset de Bureau of Transportation Statistics (BTS) que contiene horarios, retrasos, causas de demora, aerolíneas, aeropuertos, tiempos operacionales y cancelaciones.</p> Hipótesis <ul style="list-style-type: none">Algunas aerolíneas son más puntuales que otras.NAS y LateAircraftDelay serán las causas dominantes.Aeropuertos de alta concurrencia tendrán más retrasos.Algunas rutas presentan más cancelaciones y desvíos	Solución <p>La solución será un dashboard analítico que permita visualizar patrones de puntualidad, causas de retraso, concurrencia y rutas con problemas, entre otros, siendo un consolidado de los problemas encontrados en los vuelos.</p>	KPI's <ul style="list-style-type: none">Porcentaje de vuelos puntualesPorcentaje de desvíosPorcentaje de cancelaciónPorcentaje de desvíosPorcentaje de vuelos con retraso Actores <ul style="list-style-type: none">Cliente general: Entidades reguladoras de vuelos/tráfico aéreoStakeholders: Aerolíneas, aeropuertos, reguladores.Usuarios: Analistas e investigadoresImpacto en: planificación, eficiencia operativa y experiencia del pasajero.	Acciones <p>Las acciones derivadas del dashboard permitirán identificar patrones, segmentar aerolíneas y aeropuertos y visualizar horarios o rutas vulnerables para realizar mejores planificaciones en el tráfico aéreo</p>
Valores / Riesgos <p>El resultado aportará un gran valor al generar visibilidad sobre la puntualidad y causas de demora Los riesgos se encuentran en los límites que pueda dar el dataset debido a la falta de datos externos.</p>			Rendimiento / Impacto <p>Capacidad de incrementar la tasa de vuelos puntuales, reducir el número de vuelos cancelados, entender cómo reducir la variabilidad entre tiempos programados y reales.</p>	