



Análisis de retrasos y cancelaciones de Vuelos

EUA – 01/2023

Proyecto colaborativo: Jaqueline Mera y Frida Castillo

Objetivo:

"Evaluar y caracterizar los patrones de retraso y cancelaciones en los vuelos de Estados Unidos durante enero de 2023, con el fin de identificar rutas, aeropuertos, aerolíneas y factores específicos que contribuyen significativamente a estos retrasos, utilizando técnicas de análisis estadístico como las correlaciones y el riesgo relativo."





Metodología

Flujo de trabajo

Google
BigQuery
(SQL)

Google Colab
(Python)

Power BI

Procesamiento

- Identificación y manejo de nulos, duplicados, valores discrepantes, outliers, e imputación de datos.
- Creación de nuevas variables.

EDA

- Creación de tablas (medidas de tendencia central y dispersión)
- Gráficos de variables categóricas (g. barras, líneas, bivariados, histogramas y boxplots).
- Correlaciones.

Riesgo relativo

- Segmentación por variable categórica: Aerolíneas, aeropuertos, rutas, horas, motivos de retraso y cancelación.
- Segmentación por cuartiles, distancia.
- Cálculo del riesgo relativo.
- Validación de hipótesis.



Contexto: Retrasos y cancelaciones

Retraso \geq 15 min

- El Departamento de Transporte de los Estados Unidos (DOT) y la Administración Federal de Aviación (FAA), definen como retraso significativo una demora de 15 minutos o más.

Motivos:

- **Operador/Carrier:** Este tipo de retraso es causado por la aerolínea o el operador del vuelo. Puede incluir demoras por problemas de mantenimiento, tripulación, o gestión operativa de la aerolínea.
- **Retraso meteorológico*/Weather:** Este retraso es causado por condiciones meteorológicas adversas que afectan la salida o llegada del vuelo. Ejemplos incluyen tormentas, niebla, nieve, o cualquier otra condición climática que impacte la seguridad del vuelo. *Para los fines de este proyecto, será nombrado retraso por "Clima".
- **Sistema Aéreo Nacional/NAS:** Este retraso se debe a factores relacionados con el NAS. Incluye problemas como congestión del tráfico aéreo, limitaciones de control del tráfico aéreo, y restricciones del espacio aéreo.
- **Seguridad/Security:** Este tipo de retraso está relacionado con problemas de seguridad, como medidas adicionales de seguridad en el aeropuerto, inspecciones más rigurosas, o incidentes que requieren una mayor atención a la seguridad antes del despegue.
- **Aeronave tardía/Late aircraft:** Este retraso ocurre cuando un vuelo llega tarde y provoca un efecto dominó, retrasando otros vuelos posteriores que dependen de la misma aeronave. Es común en vuelos de conexión o cuando la misma aeronave realiza varios vuelos en un día.



EDA

Retrasos:
1 de cada 5

Cancelados:
2 de cada 100

11 de enero, día
con anomalías

La mayoría de
retrasos son
multifactoriales

Retraso
promedio

Vuelos EUA 01-2023

Vuelos
538.84 mil

Aeropuertos
339

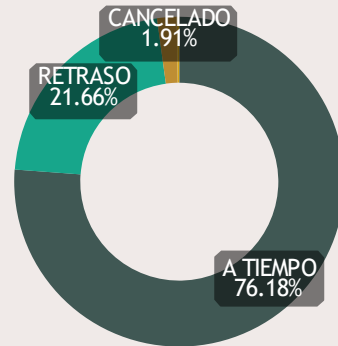
Retrasos
117 mil

Cancelados
10 mil

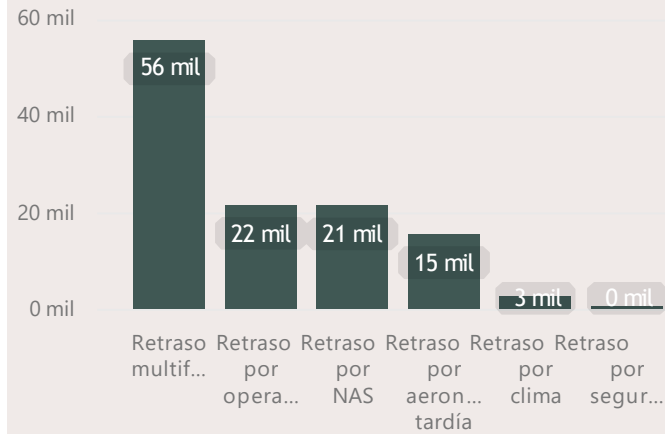
Rutas
5581

Aerolíneas
15

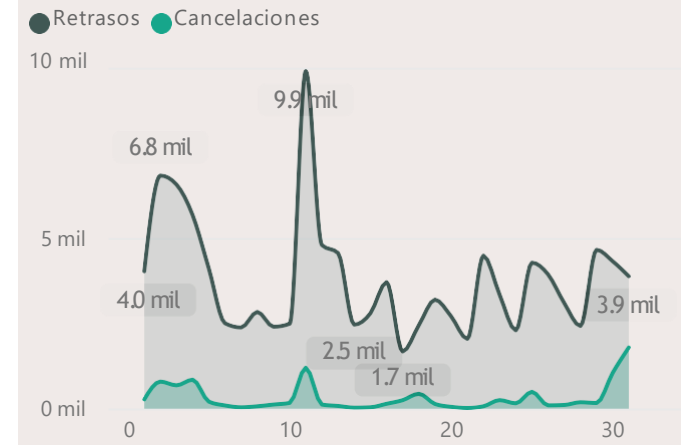
Porcentaje de vuelos por categoría



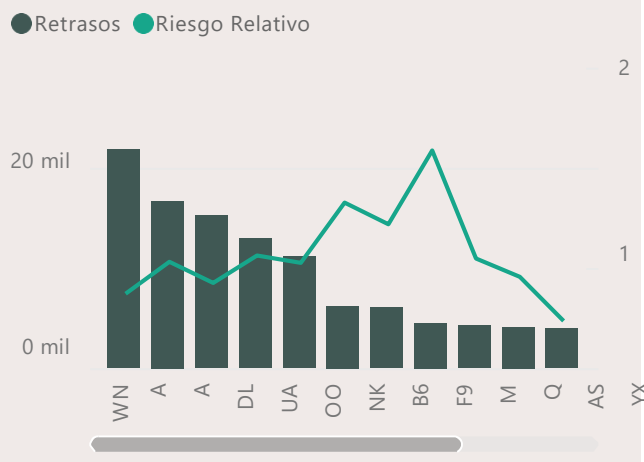
Vuelos por motivo de retraso



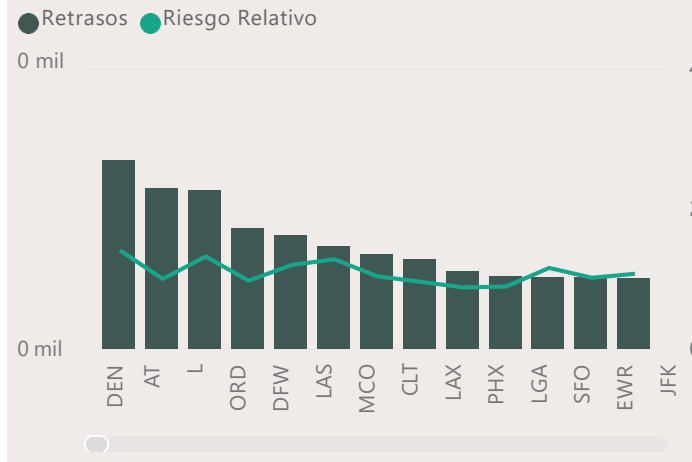
Retrasos y Cancelaciones por Día



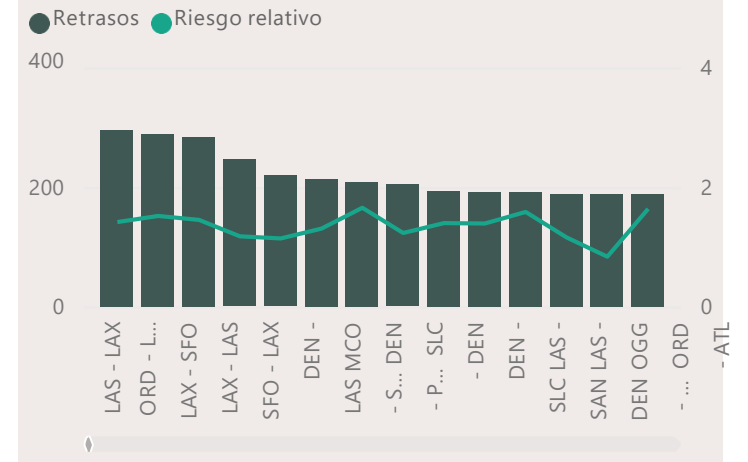
Retrasos por Aerolínea Vs Riesgo Relativo



Retrasos por Aeropuerto Vs Riesgo Relativo



Retrasos por Rutas Vs Riesgo Relativo





Riesgo relativo

- El riesgo relativo se calculó como:

$$\text{Riesgo relativo} = \frac{\text{Tasa de Incidencia en el Grupo Expuesto}}{\text{Tasa de Incidencia en el Grupo No Expuesto.}}$$

RR>1

Menor riesgo en
GE

RR=1

Sin asociación
significativa

RR<1

Mayor riesgo en
GE

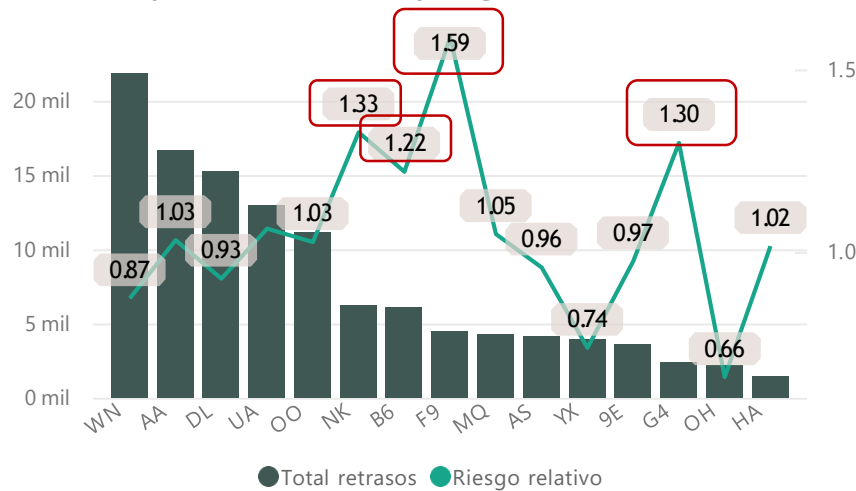
- Segmentos a evaluar:
 - Categóricos: Aerolíneas, Aeropuertos, Rutas, Motivos de retraso y cancelación, Horas.
 - Cuartiles: Distancia.



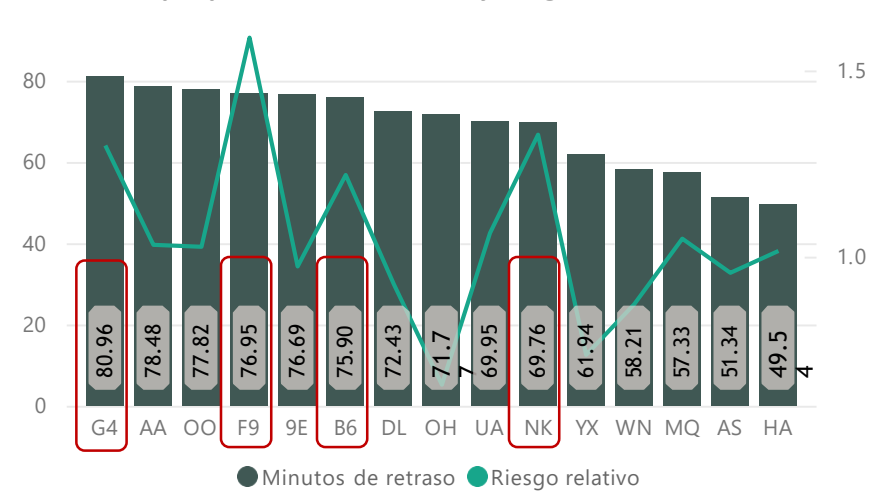
Aerolíneas y Aeropuertos

Hipótesis 1: Algunas aerolíneas tienen un historial de retrasos significativamente mayor que otras.

Aerolíneas por cantidad retrasos y riesgo relativo

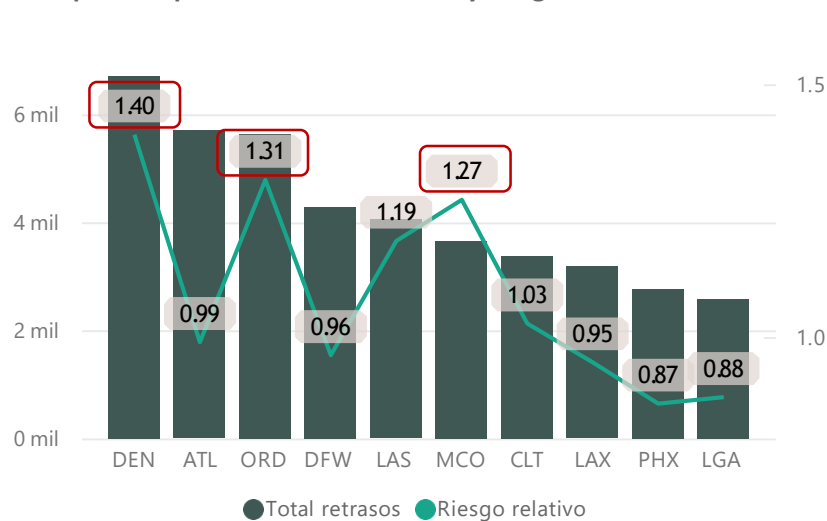


Aerolíneas por promedio de retrasos y riesgo relativo

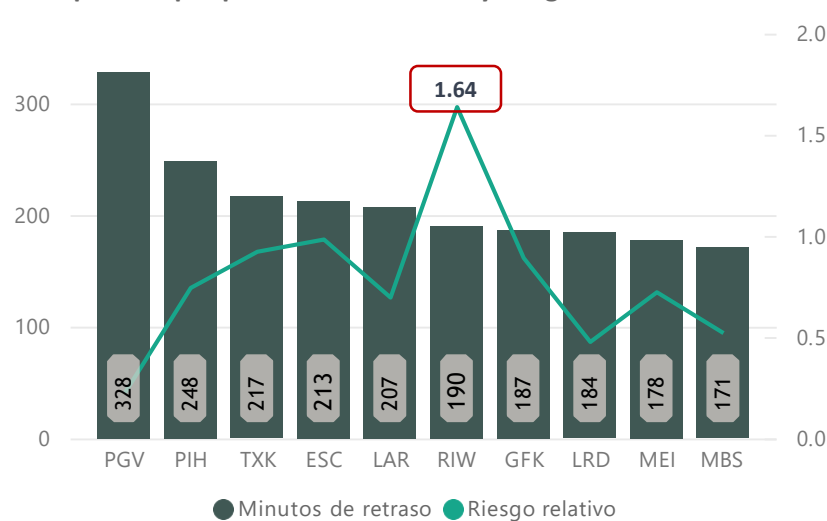


Hipótesis 2: Algunos aeropuertos tienden a tener retrasos más frecuentes o severos en comparación con otros.

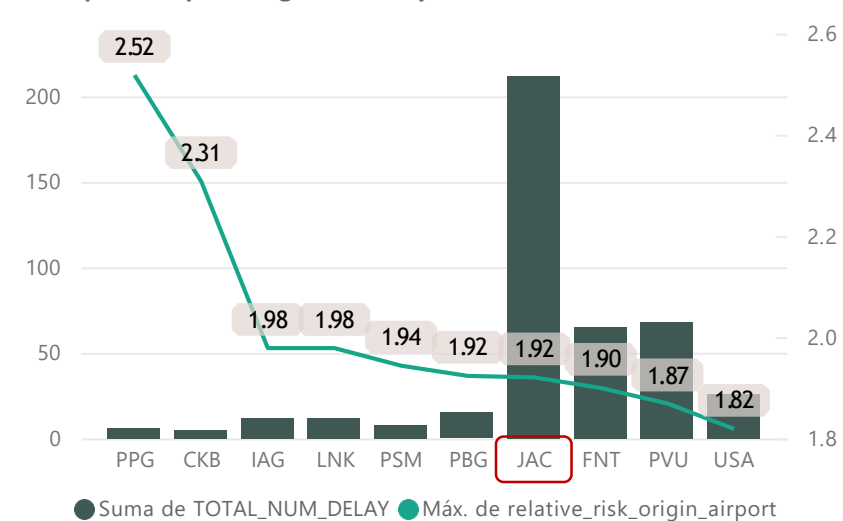
Aeropuertos por número de retrasos y riesgo relativo



Aeropuertos por promedio de retraso y riesgo relativo



Aeropuertos por riesgo relativo y número de retrasos

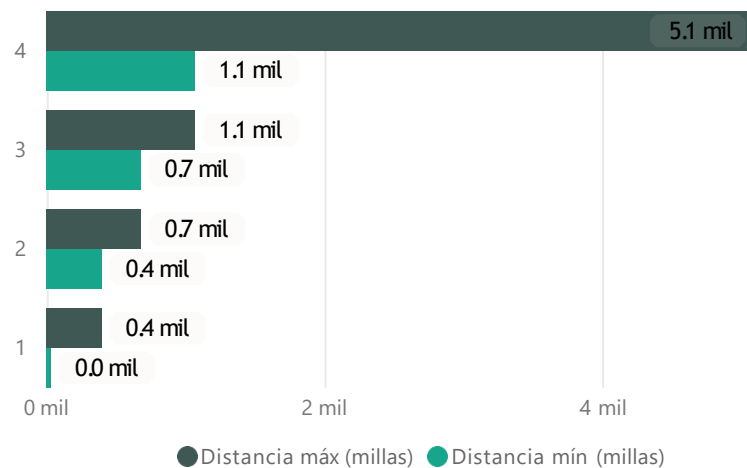




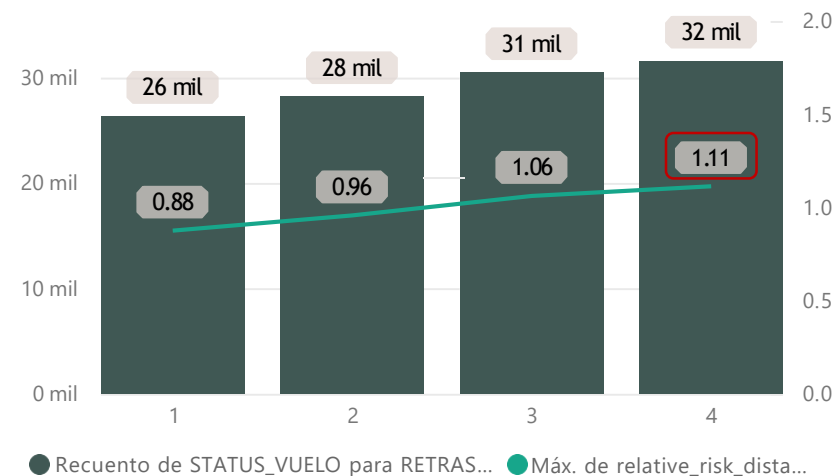
Distancia y Hora

Hipótesis 3: Los vuelos más largos tienen mayores tiempos de retraso en comparación con los vuelos más cortos.

Rango de cuartiles por Distancia (millas)

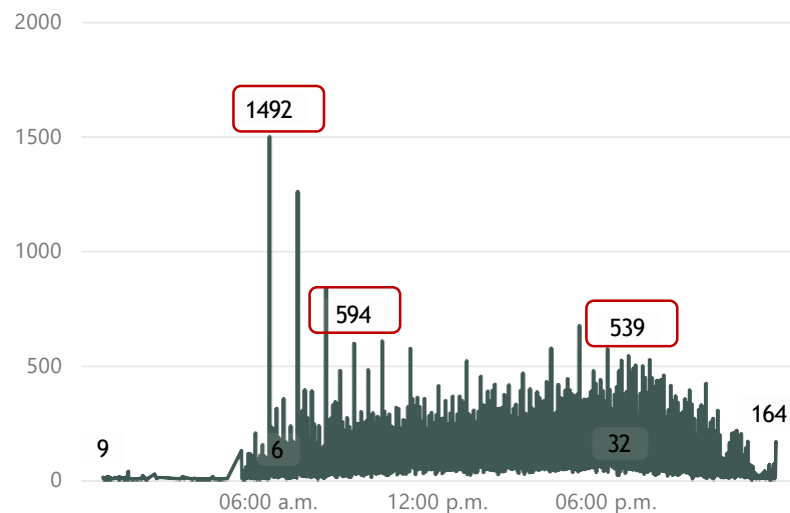


Total de retrasos por cuartil de Distancia y riesgo relativo

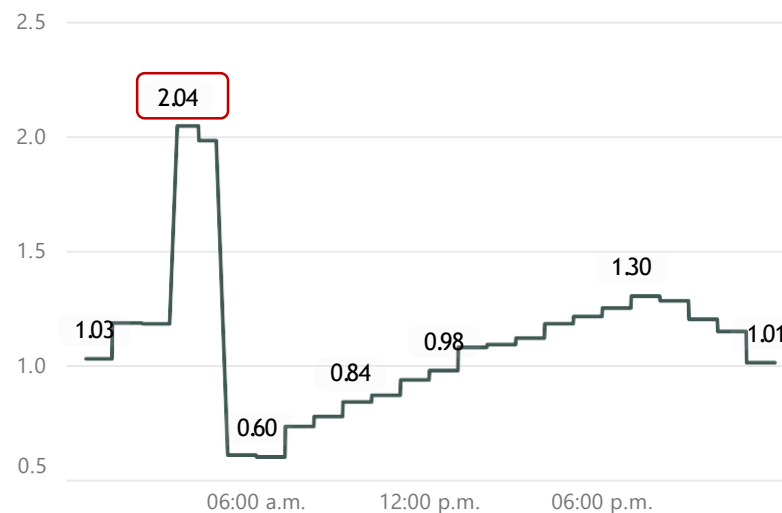


Hipótesis 4: Los retrasos en los vuelos son más comunes durante las horas punta del día.

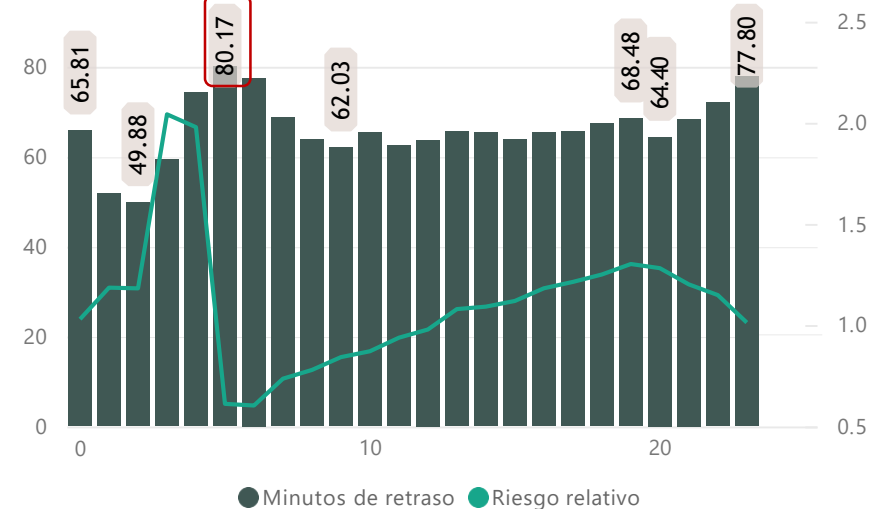
Retrasos por Hora



Riesgo relativo por Hora



Promedio de retraso por hora Vs Riesgo relativo

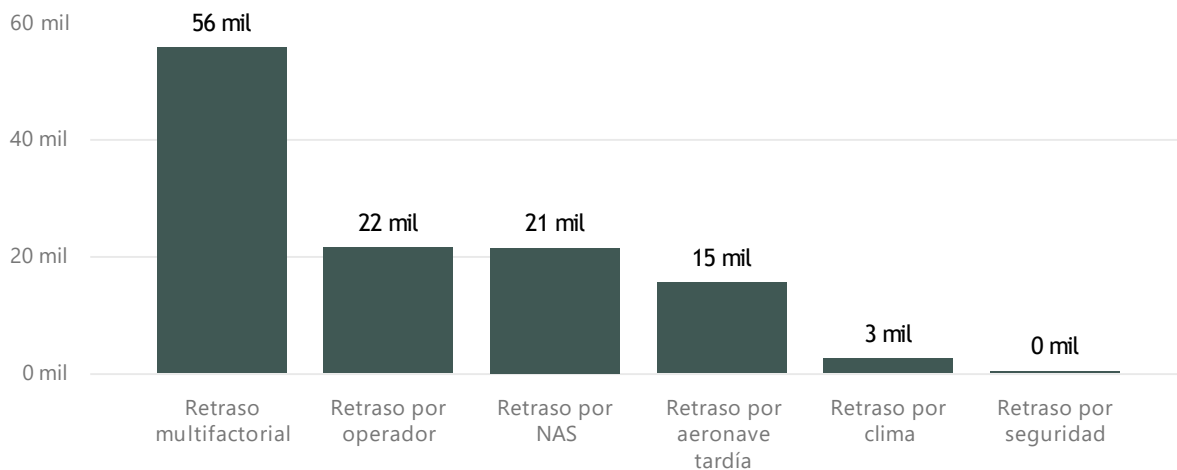




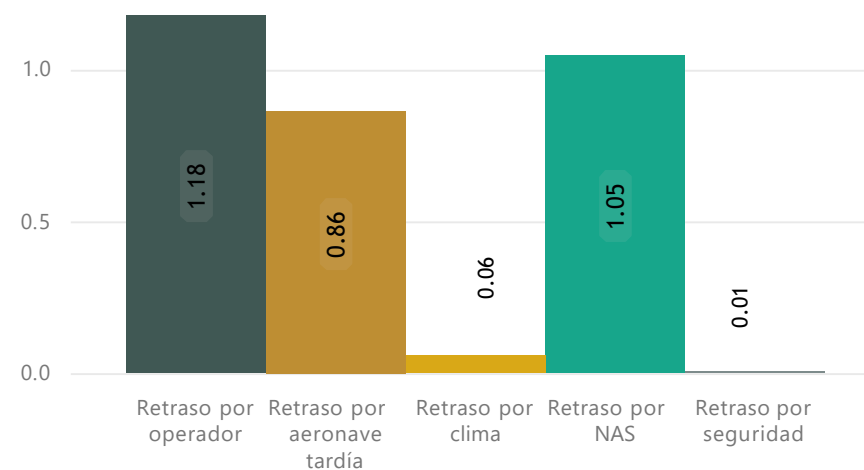
Motivos de Retrasos y cancelaciones

Hipótesis 5: Algunos motivos de retrasos son más prevalentes que otros, indicando causas específicas más comunes para el retraso de un vuelo.

Motivo de cancelación por Motivo

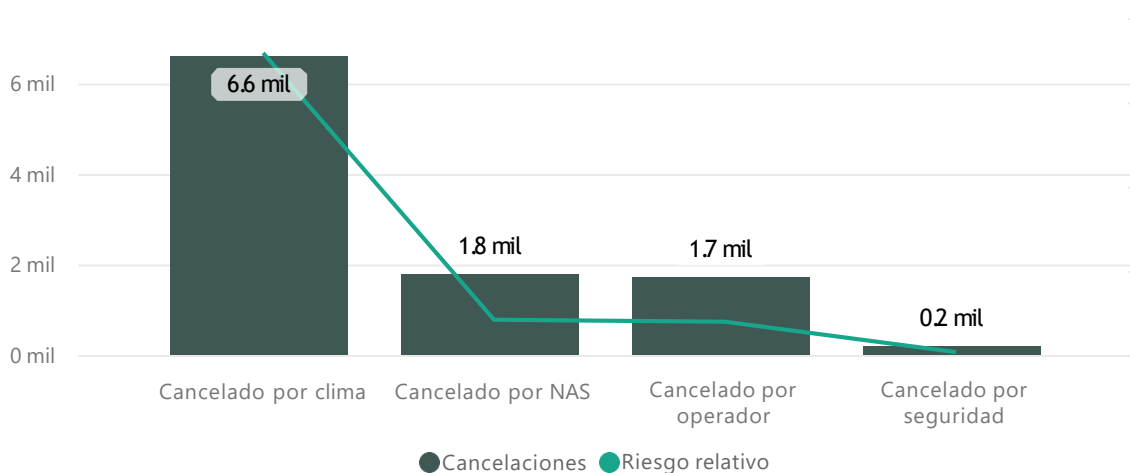


Riesgo relativo por motivo

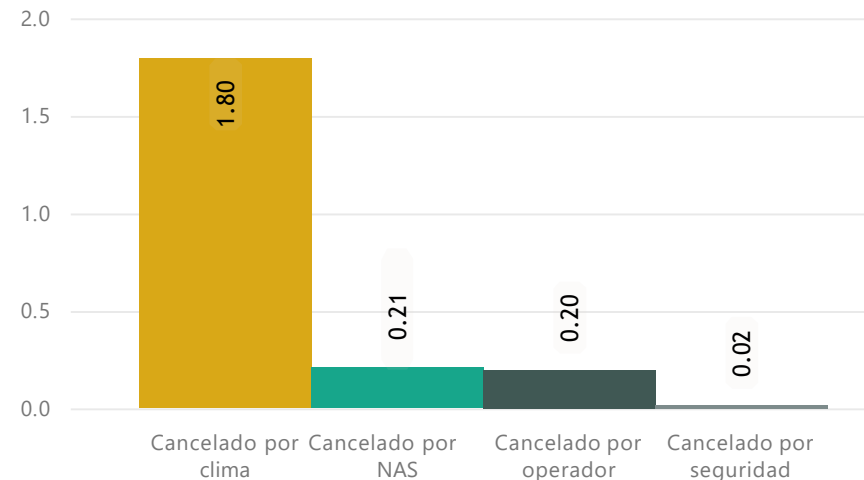


Hipótesis 6: Algunos códigos de cancelación son más prevalentes que otros, indicando causas específicas más comunes.

Cancelaciones y Riesgo relativo por Motivo



Riesgo relativo por Cancelaciones





Conclusiones

- **Variabilidad en el riesgo de retrasos entre aerolíneas:** No todas las aerolíneas presentan el mismo riesgo de retrasos, y las aerolíneas con más retrasos no siempre tienen los mayores riesgos relativos. Esto sugiere que la capacidad para gestionar las operaciones es crucial para mantener bajo control los retrasos.
- **Diferencias notables entre aeropuertos:** Aunque muchos aeropuertos tienen un riesgo relativo de retraso alto, estos no necesariamente coinciden con los que operan más vuelos. Aeropuertos pequeños, aunque menos concurridos, parecen enfrentarse a mayores desafíos en la gestión de operaciones.
- **Mayor probabilidad de retrasos en vuelos largos:** Los vuelos más largos tienen un riesgo más elevado de sufrir retrasos debido a la complejidad logística, las múltiples escalas y la exposición prolongada a factores externos.
- **Horarios pico son propensos a retrasos en cuanto a su frecuencia:** Las horas pico, especialmente las mañanas y tardes, tienen una mayor propensión a retrasos debido a la alta demanda y congestión operativa.
- **Primeras horas de la mañana tienen más riesgo:** La posibilidad de sufrir retrasos en la madrugada puede deberse a la acumulación de problemas operativos del día anterior.
- **Anomalías en el Operador y el sistema NAS son los principales motivos de retrasos:** Puede ser atribuible a el mantenimiento, la tripulación, la gestión operativa de la aerolínea, así como, complicaciones en el tráfico aéreo. Otros factores, como el clima o la seguridad, tienen un impacto relativamente menor, aún así estos deben ser considerados.
- **Cancelaciones mayoritariamente asociadas al clima:** Las cancelaciones por condiciones climáticas adversas son las más comunes entre los códigos de cancelación, superando ampliamente a otros motivos como problemas del operador o seguridad.



Recomendaciones

- **Optimización de las operaciones aéreas:** Las aerolíneas y aeropuertos con mayor riesgo relativo deberían optimizar sus operaciones, mejorando la planificación, la gestión del personal, y la asignación de recursos durante los períodos de mayor demanda.
- **Gestión de aeropuertos pequeños:** Dado que los aeropuertos más pequeños enfrentan un riesgo mayor de retrasos, es esencial que estos inviertan en mejorar su infraestructura operativa, como sistemas de gestión de tráfico aéreo y control en tierra.
- **Planificación en vuelos largos:** Optimizar las conexiones y prever potenciales problemas durante las rutas largas puede reducir las demoras.
- **Estrategias para mitigar retrasos en horas pico:** Las aerolíneas y los aeropuertos deben evaluar la programación de vuelos durante las horas pico. Recomendamos redistribuir los vuelos en horarios menos congestionados o aumentar la capacidad operativa durante estos períodos. Una mejor gestión del tráfico aéreo en las primeras horas del día y las tardes podría aliviar los retrasos.
- **Enfocar esfuerzos en causas operacionales:** La implementación de herramientas que permitan una respuesta más rápida y efectiva ante problemas operativos podría reducir significativamente los retrasos.
- **Preparación ante condiciones climáticas:** Mejorar la capacidad de reacción en situaciones climáticas severas, así como proporcionar flexibilidad a los pasajeros, podría reducir las cancelaciones y mejorar la experiencia del usuario.
- **Análisis continuo y adaptación:** Es crucial que las aerolíneas y los aeropuertos realicen un análisis continuo de sus operaciones, evaluando los cambios en el riesgo de retrasos y cancelaciones. Adaptar sus estrategias con base en datos actualizados permitirá a las empresas mantener una mejor competitividad y ofrecer un servicio más eficiente.

GRACIAS!

Conoce más en:

GitHub: https://github.com/JaquelineMera/Analisis_Retrasos_Vuelos

