

Reto | Análisis de Tráfico en Aeropuertos

Nombre: Carlos Alejandro Hernández Gómez

1. Introducción donde debes de mencionar el objetivo del análisis considerando el rol que hayas seleccionado.

Rol seleccionado: **Público en general.**

El objetivo de este análisis es el de ayudar a la gente a tomar las mejores decisiones al viajar por avión. Planeo informar, por ejemplo, sobre qué aerolíneas brindan la mayor satisfacción, cuáles se retrasan más, qué aerolíneas ofrecen mejores precios, en qué horarios hay menos atrasos, qué horarios son más baratos, cuál es la relación precio-satisfacción para saber qué rango ofrece más satisfacción por menos dinero y demás. Mi intención no es decirle a la gente qué hacer, sino ofrecer la información para que las personas puedan decidir qué opción se adapta mejor a sus necesidades y preferencias personales.

Análisis exploratorio de datos y ligeros cambios a la base de datos en Python:

https://github.com/CalHdzg/Reto-Analisis-de-Trafico-en-Aeropuertos/blob/main/Aerolineas_EDA.ipynb

2. Tabla descriptiva de la Fuente de datos (desarrollada en el Paso 1).

Nombre	Descripción	Tipo de Dato	Valores
Nivel Satisfacción	Rango de satisfacción del viajero	Número entero Categórico ordinal	Enteros del 1 al 5
Airline Status	Membresía o categoría del pasajero basado en lealtad o programas de viajero frecuente, basados en la cantidad de millas o segmentos de vuelo acumulados.	Texto Categórico nominal	Blue, Silver, Gold, Platinum
Edad	Edad del pasajero	Número entero Cuantitativo discreto	Enteros del 15 al 85
Rango_edad	Rangos de edades para facilitar la clasificación	Texto Categórico ordinal	0-19, 20-29, 30-39 ..., 70-79, 80+
Género	Para fines de esta base de datos, se toma como sinónimo de sexo del pasajero o pasajera.	Texto Categórico nominal	Female, Male
Tipo Viaje	Motivo por el que viaja el pasajero	Texto Categórico nominal	Business travel, Personal Travel, Mileage tickets



Consumo en Aereopuerto	Cantidad de dinero gastado en el aeropuerto. (El nombre de la clase está mal escrito, esto lo cambiaré luego, pero por lo pronto lo dejo tal cual como viene).	Número Cuantitativo continuo	Enteros entre 0 y 879
Consumo en Alimentos	Cantidad de dinero gastado en alimentos (presumiblemente ya en el avión).	Número Cuantitativo continuo	Enteros entre 0 y 895
Clase	Tipo de asiento que compró el pasajero	Texto Categórico ordinal	Eco, Eco Plus y Business (en orden de precios)
Flight date	Fecha en la que se realizó el vuelo	Fecha Cuantitativo discreto	2021/01/01 a 2021/03/31
Airline Code	Código de la aerolínea	Texto Categórico nominal	'MQ', 'EV', 'FL', 'AS', 'F9', 'HA', 'B6', 'OO', 'OU', 'AA', 'DL', 'WN', 'US', 'VX'
Airline Name	Nombre de la aerolínea	Texto Categórico nominal	'EnjoyFlying Air Services', 'FlyFast Airways Inc. ', 'FlyHere Airways', 'FlyToSun Airlines Inc.', 'GoingNorth Airlines Inc. ', 'West Airways Inc. ', 'OnlyJets Airlines Inc. ', 'Northwest Business Airlines Inc. ', 'Oursin Airlines Inc. ', 'Paul Smith Airlines Inc. ', 'Sigma Airlines Inc. ', 'Cheapseats Airlines Inc. ', 'Southeast Airlines Co. ',



			'Cool&Young Airlines Inc. '
CD. Origen	Ciudad de origen del vuelo	Texto Categórico nominal	Son 295 ciudades diferentes de Estados Unidos
Edo. Origen	Estado de origen del vuelo	Texto Categórico nominal	Los 50 estados de EE UU y además 'Puerto Rico' y 'U.S. Pacific Trust Territories and Possessions'
CD. Destino	Ciudad de destino del vuelo	Texto Categórico nominal	Son 296 ciudades diferentes de Estados Unidos
Edo. Destino	Estado de destino del vuelo	Texto Categórico nominal	Los 50 estados de EE UU y además 'Puerto Rico' y 'U.S. Pacific Trust Territories and Possessions'
Horario Salida	Hora a la que parte el avión	Número entero Cuantitativo continuo	De 1 a 23
Retraso Salida (Min)	Cantidad de minutos que tuvo de retraso la salida avión	Número entero Cuantitativo continuo	De 0 a 1592
Retraso Llegado (Min)	Cantidad de minutos que tuvo de retraso la llegada del avión	Número entero Cuantitativo continuo	De 0 a 1584
Cancelación	Si el vuelo se canceló	Booleano Categórico	No, Yes
Tiempo vuelo (min)	Cuanto duró el vuelo	Número entero Cuantitativo continuo	De 8 a 669 (aunque no me creo que haya vuelos

			de menos de 20 minutos, de los cuales hay varios enlistados)
Distancia Vuelo	Distancia de viaje en millas	Número entero Cuantitativo continuo	31 a 4983
Retraso Mayor 5 Min	Verdadero si hubo un retraso de más de 5 minutos.	Booleano Categórico nominal	No, yes

3. Requerimientos de Análisis (desarrollado en el Paso 2).

Rol seleccionado: **Público en general.**

- ¿Qué aerolíneas tienen la mayor satisfacción promedio?
- ¿Qué aerolíneas tienen mayor índice de retrasos (de más de 5 minutos) y qué tan largos son en promedio?
- ¿Cuál es la relación entre la clase y el nivel de satisfacción?
- ¿Qué días de la semana tienen mayor tráfico de viajeros?
- ¿Cuáles son los meses donde hay más tráfico de viajeros?
- ¿En qué horario hay más tráfico de viajeros?
- ¿Qué aerolíneas se caracterizan por tener más pasajeros?

4. Acondicionamiento y Configuración de Datos (documentar lo desarrollado en el Paso 3).

Hice los siguientes cambios utilizando Python: https://github.com/CalHdzg/Reto-Analisis-de-Trafico-en-Aeropuertos/blob/main/Aerolineas_EDA.ipynb

En orden de aparición de las clases:

- Cambié el nombre de la clase "Consumo de Aereopuerto" a "Consumo de Aeropuerto" porque estaba mal escrita.
- Cambié los valores de la clase "Cancelación" de "No" y "Yes" a "False" y "True" respectivamente como valores Booleanos.
- Cambié el nombre de la clase "Distancia Vuelo" a "Distancia Vuelo (mi)" para enfatizar que la distancia está en millas.
- Cambié los valores de la clase "Retraso Mayor 5 Min" de "no" y "yes" a "False" y "True" respectivamente como valores Booleanos.

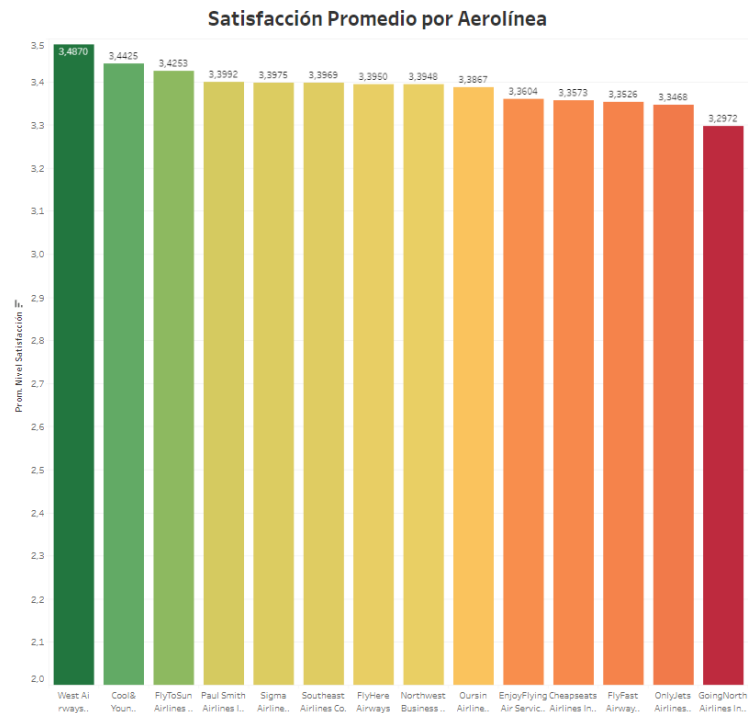
5. Herramienta de Análisis desarrollada (liga del sitio público de Tableau donde quedó guardado tu trabajo):

https://public.tableau.com/views/AnlisisdeTrficodeAeropuertos/ApoYoalosviajerosenaerolneas?:language=es-ES&publish=yes&:display_count=n&:origin=viz_share_link



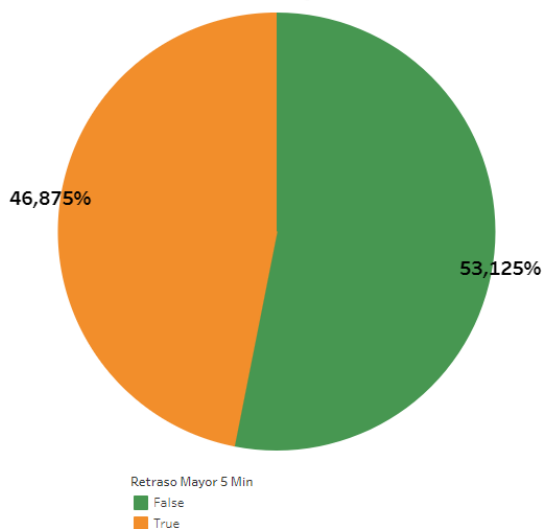
6. Resultados del Análisis y Conclusiones (desarrollado en el Paso 7).

¿Qué aerolíneas tienen la mayor satisfacción promedio? **West Airways, con un promedio de 3.487**



¿Qué aerolíneas tienen mayor índice de retrasos (de más de 5 minutos) y qué tan largos son en promedio? **GoingNorth Airlines Inc. Con un 46.875% de retrasos (de 5 minutos o más) y con 16.27 minutos promedio.**

Índice de retraso por aerolínea



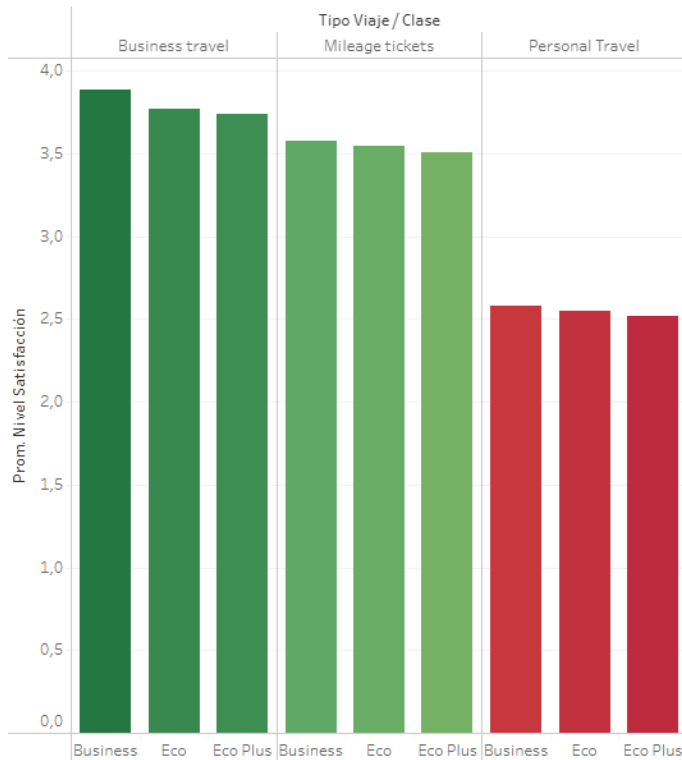
Promedio de tiempo de retraso

Airline Name	Min
FlyFast Airways Inc.	21.25
OnlyJets Airlines Inc.	19.33
Cheapseats Airlines Inc.	18.08
GoingNorth Airlines Inc.	16.27
Oursin Airlines Inc.	16.06
EnjoyFlying Air Services	15.28
Sigma Airlines Inc.	12.93
Paul Smith Airlines Inc.	12.79
Northwest Business Airlines Inc.	12.76
Cool&Young Airlines Inc.	12.48
FlyHere Airways	11.84
Southeast Airlines Co.	8.72
FlyToSun Airlines Inc.	6.30
West Airways Inc.	2.84

Tip: Presiona en las diferentes aerolíneas para ver a la izquierda la proporción de vuelos con retraso que tiene.



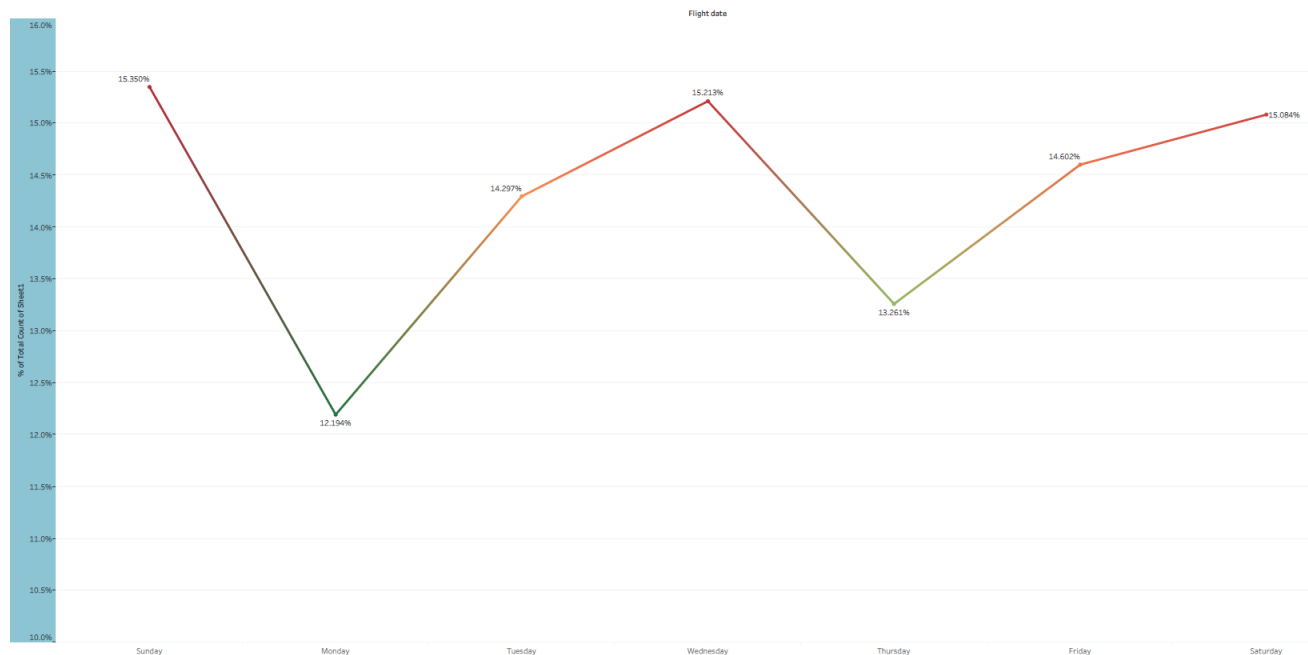
Tipo de Viaje vs Satisfacción



¿Cuál es la relación entre la clase y el nivel de satisfacción? **En general, Business da más satisfacción que Económica y ésta da más que la económica Plus. Además, la gente que viaja por negocio tiene una satisfacción mayor y la gente que viaja por motivos personales tienen la peor satisfacción por mucho.**

¿Qué días de la semana tienen mayor tráfico de viajeros? **Los domingos, los miércoles y los sábados son los peores, en ese orden. Los lunes y jueves son los mejores, en ese orden.**

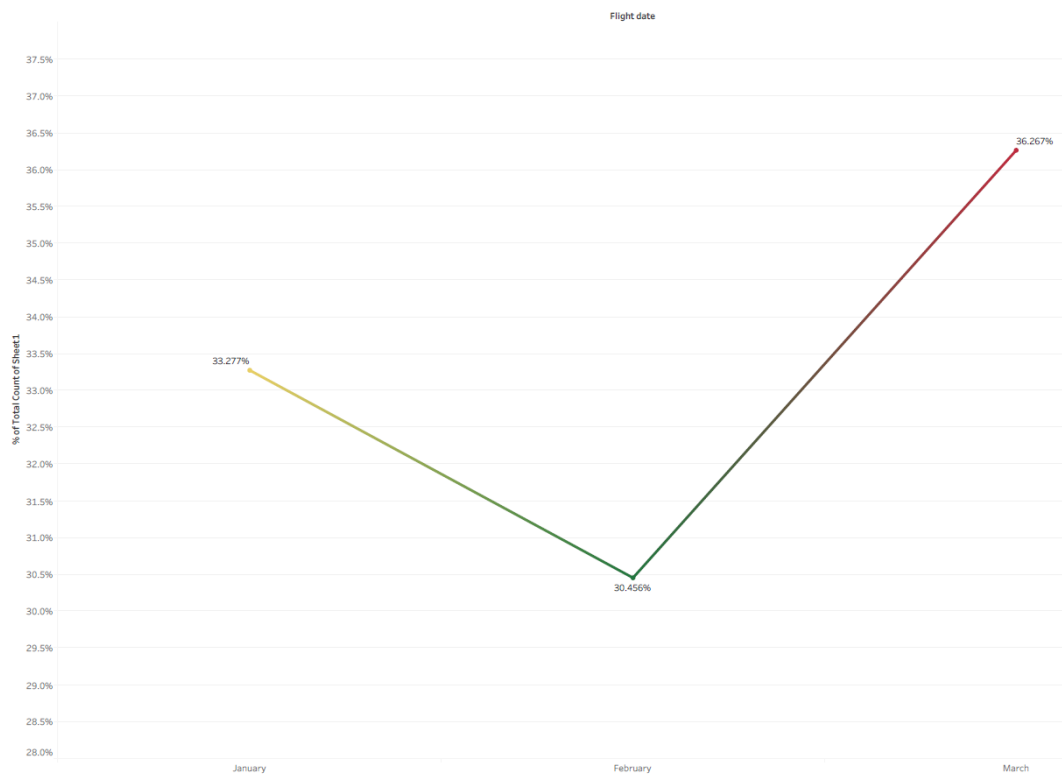
Táfico de gente por día de la semana





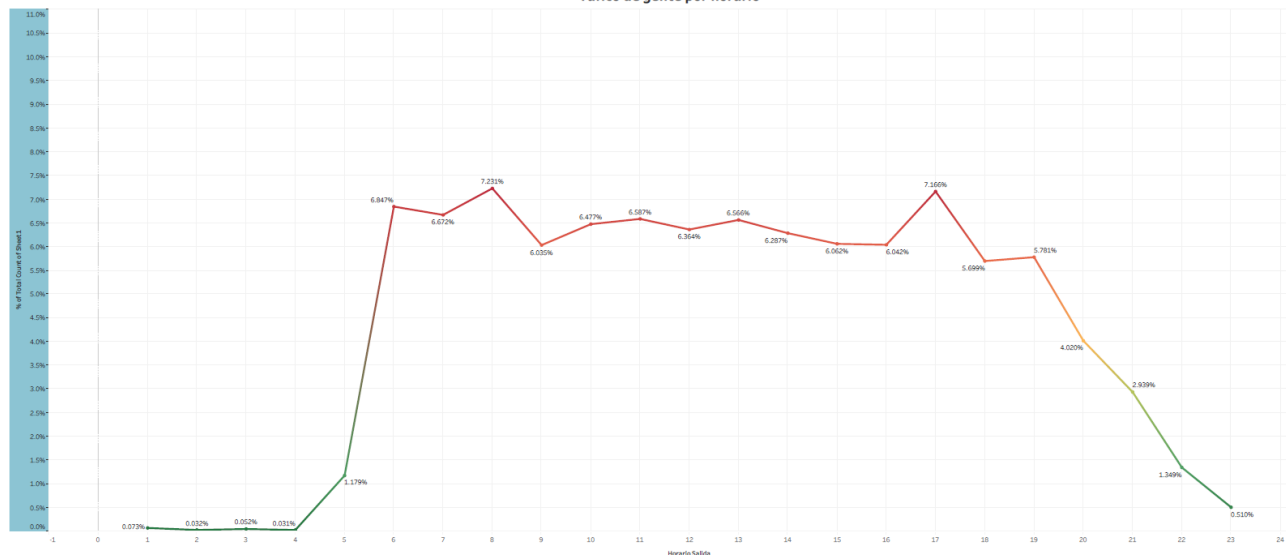
¿Cuáles son los meses donde hay más tráfico de viajeros? **Sólo tenemos datos de los primeros tres meses del año 2022, pero hasta ahora mes con más tráfico es marzo, luego enero y febrero ha sido el que ha tenido menos tráfico de viajeros.**

Táfico de gente por mes



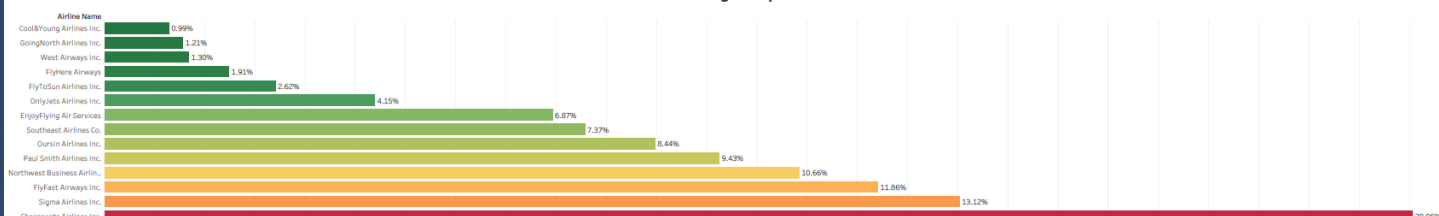
¿En qué horario hay más tráfico de viajeros? **En general, entre las 6 y las 19 horas hay mucha más gente y hay mucha menos entre las 23 y las 4 horas del día siguiente.**

Táfico de gente por horario



¿Qué aerolíneas se caracterizan por tener más pasajeros? **Cheapseats Airlines Inc. tiene por mucho la mayoría con 20.06% de los pasajeros, Young, North y West son las aerolíneas con menos pasajeros.**

Táfico de gente por aerolínea



Uno de los descubrimientos más impactantes para mí fue ver que la clase Económica Plus brinda peor satisfacción que la Económica, obviamente no vale la pena pagar el extra por la clase Plus.

Claramente a no tanta gente le gusta comprar sus boletos para horarios en la madrugada, pero si no te molesta ese horario y prefieres que no te toque tanta gente, puede valer la pena. También viajar en lunes o jueves es mejor y, extrañamente, el miércoles es uno de los días con más gente. En algunas aerolíneas hay un índice de vuelos retrasados bastante alto, definitivamente vale la pena ver las estadísticas por aerolínea antes de tomar una decisión. En general parece que comprar un boleto a conciencia puede ayudar a tener una experiencia significativamente mejor que si sólo se compra al azar, ya que no hay bastante homogeneidad en los resultados obtenidos.