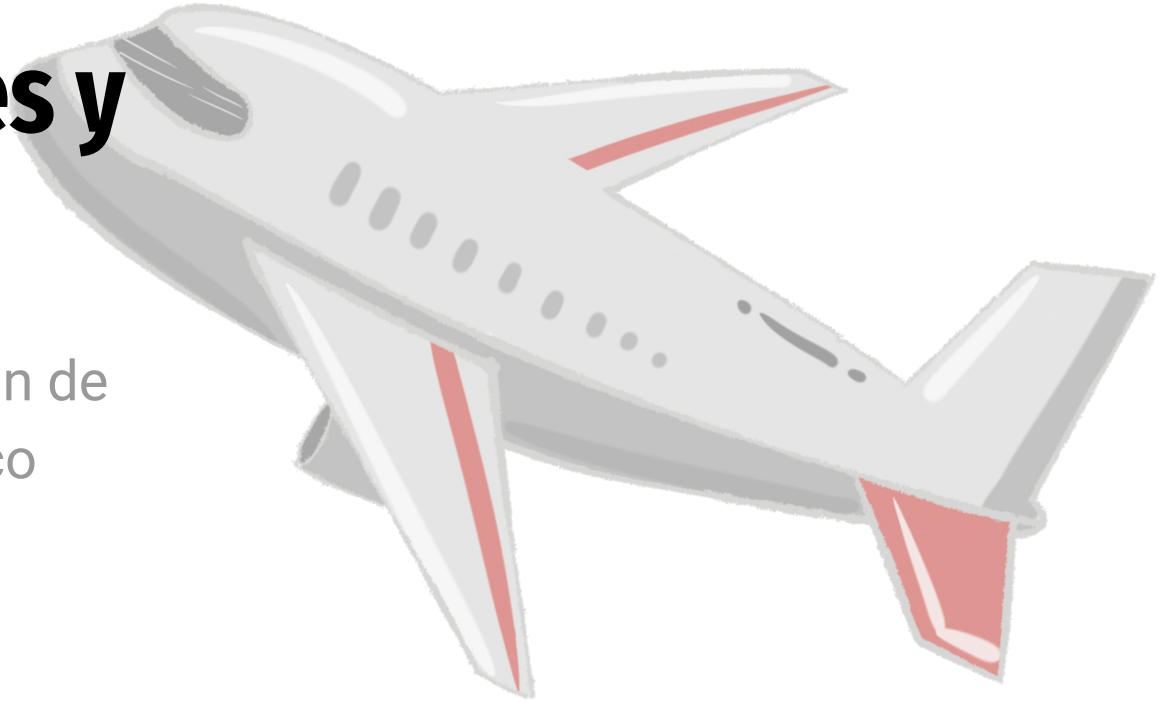


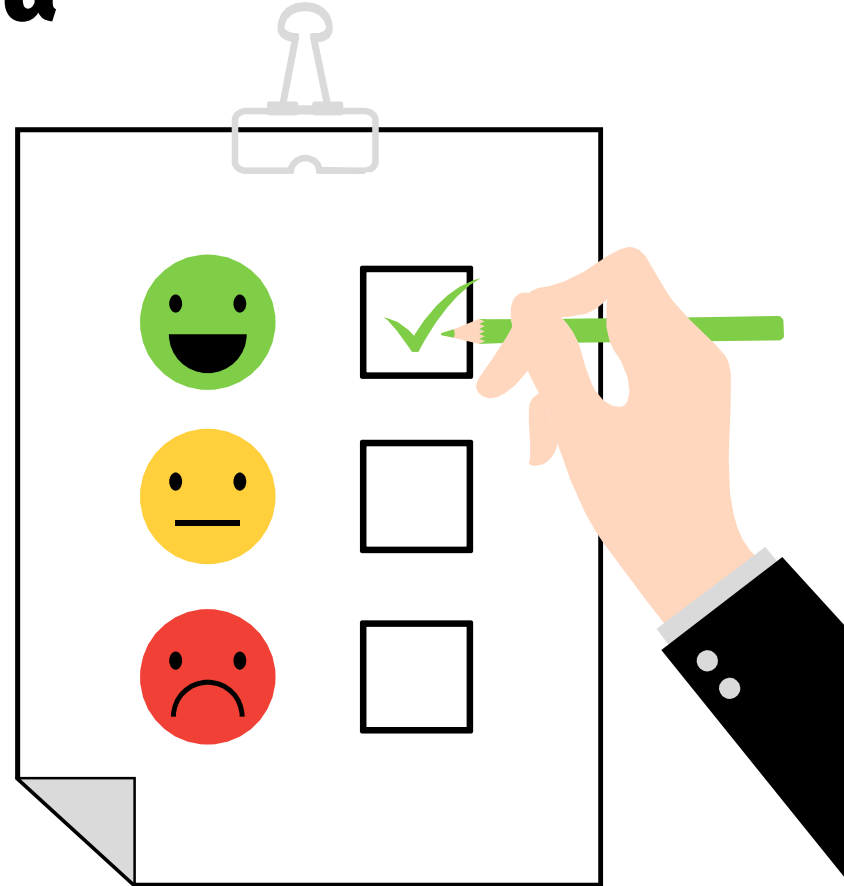
# Entre Opiniones y Vuelos

Explorando la Satisfacción de los Pasajeros con Invistico Airlines



# Agenda

- 1 Contexto
- 2 Resumen Metadata
- 3 Industria Aérea
- 4 Invistico Airlines - EDA
- 5 Encoding y Feature Engineering
- 6 Selección del modelo
- 7 Insights y Recomendaciones
- 8 Bibliografía





# Contexto



## Motivo

La premisa inicial consistía en la creación de un modelo destinado a anticipar las decisiones de compra de pasajeros de buses de larga distancia, basándose en su historial de comportamiento o del comportamiento de usuarios con atributos similares.

Dada la ausencia de un conjunto de datos específico para este propósito, se utilizó información proveniente de una aerolínea como sustituto temporal, con la posibilidad de replicar el proceso en el eventual acceso a la información del servicio mencionado.

## Audiencia

Puede ser de utilidad para cualquier persona iniciándose en programación/ciencia de datos.



## Invistico Airlines

- La información fue suministrada por una aerolínea que, por diversos motivos, el nombre real de la empresa no se revela, adoptando en su lugar el nombre ficticio de Invistico Airlines.
- La base de datos comprende información detallada de clientes que ya han experimentado sus servicios y se han recopilado opiniones de estos, junto con sus datos de vuelo.

*"Airlines Customer satisfaction - Customer satisfaction with various other factors"* creado por Sayantan Jana y actualizado en 2019. Se puede encontrar el conjunto de datos en la siguiente URL:  
<https://www.kaggle.com/datasets/sjleshrac/airlines-customer-satisfaction>



## Preguntas de interés

¿Cómo **mejorar el índice de satisfacción actual** de los usuarios de la aerolínea para cumplir con el objetivo establecido del 70%?

- ¿Cuál es el nivel general de satisfacción de los clientes y cuáles son los factores clave que influyen en ello?
- ¿Cómo se ven afectados los niveles de satisfacción por lealtad o tipo de viaje, así como por la duración y tipo de vuelo?
- ¿Cuál es el impacto de los diversos servicios, tanto a bordo como en línea, en la satisfacción de los pasajeros, y cómo influyen la puntualidad y la gestión de retrasos en la experiencia global del cliente?



# Resumen Metadata

## Datos analizados

---

La base de datos comprende información detallada de clientes que ya han experimentado sus servicios y se han recopilado opiniones de estos, junto con sus datos de vuelo.

- 129.880 filas
- 23 columnas

## Tipos de Variables

---

### STRING

- Satisfaction
- Gender
- Customer Type
- Type of Travel
- Class

### INTEGER

- Age
- Flight Distance
- Seat comfort
- Departure/Arrival time convenient
- Food and drink
- Gate location
- Inflight wifi service
- Inflight entertainment
- Online support
- Ease of Online booking
- On-board service
- Leg room service
- Baggage handling
- Checkin service
- Cleanliness
- Online boarding
- Departure Delay in Minutes
- Arrival Delay in Minutes

## Columnas

---

Satisfaction: Satisfied, Dissatisfied  
Gender: Female, Male  
Customer Type: Loyal Customer, Disloyal Customer  
Type of Travel: Business travel, Personal Travel  
Class: Business, Eco, Eco Plus

Age: [7; 85]  
Flight Distance: [50; 6.951]  
Departure Delay in Minutes: [0;1.592]  
Arrival Delay in Minutes: [0;1.584]  
Services: [0; 5]



# Industria Aérea



## Importancia

---

La satisfacción del cliente es **crucial** para el éxito y **reputación** de las aerolíneas.



## Desafíos

---

- **Costos** y tarifas elevados.
- **Retrasos** de vuelos, **pérdida** de equipaje y restricciones en la **disponibilidad** de vuelos.
- **Eficiencia** operativa, **incomodidad** de asientos y **calidad** del servicio.



## Estrategia

---

- Buenas prácticas en la **experiencia** del cliente.
- **Diseño** de servicios impactante.
- **Atención a las preferencias** emergentes de los pasajeros

## Caso de éxito

Aerolíneas Argentinas alcanzó un destacado índice de satisfacción del cliente en mayo, con un promedio de 8.25 puntos y destacadas calificaciones en aspectos clave como atención en el check-in (8.88) y proceso de compra (8.85). Además, la aerolínea ha sido reconocida a nivel global, obteniendo premios por puntualidad, siendo distinguida como la mejor línea aérea de Latinoamérica según Trip-Advisor y destacada entre las tres más puntuales del mundo por Flightstats.



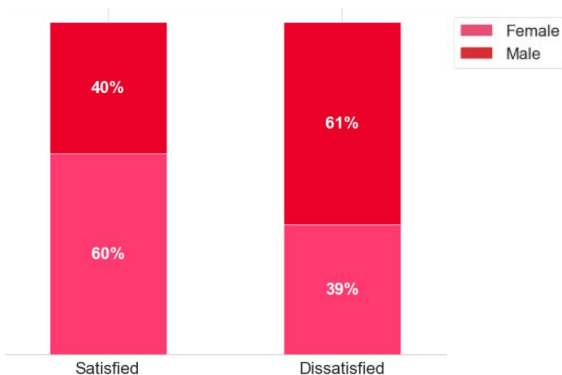
# Invistico Airlines

Índice de satisfacción de 55%, 15 p.p. por debajo del objetivo de 70%\*

4

## Actualidad

En Invistico **servicio overall** es **bueno** dado que **hay más usuarios satisfechos (55%) que insatisfechos (45%)**, estando **15 p.p. por debajo del objetivo**. Las mujeres suelen estar más satisfechas que los hombres.



## Edad y Satisfacción

La **edad media** es de **40 años** en **ambos casos**, dato que coincide con la edad media y la moda de toda la base. Los valores se encuentran dispersos dado que el coeficiente de variación es superior al 20% en ambos estados de satisfacción, sin embargo, el coeficiente de asimetría de los datos para ambos estados se acerca al 0, por lo que hay una concentración de usuarios cercanos a los 40 años.

## Sexo y satisfacción

La mayor cantidad de **mujeres** satisfechas se encuentran entre los 40 y 60 años, mientras que la mayoría de las **insatisfechas están entre los 20 y los 40 años**. Para el caso de los **hombres** se ve un comportamiento similar, sin embargo, los usuarios **insatisfechos** se encuentran más dispersos, encontrándose una gran parte **por encima de los 40 años**.

\* Assumption: la aerolínea quiere conseguir un mínimo de 70% en el índice de satisfacción



# Invistico Airlines

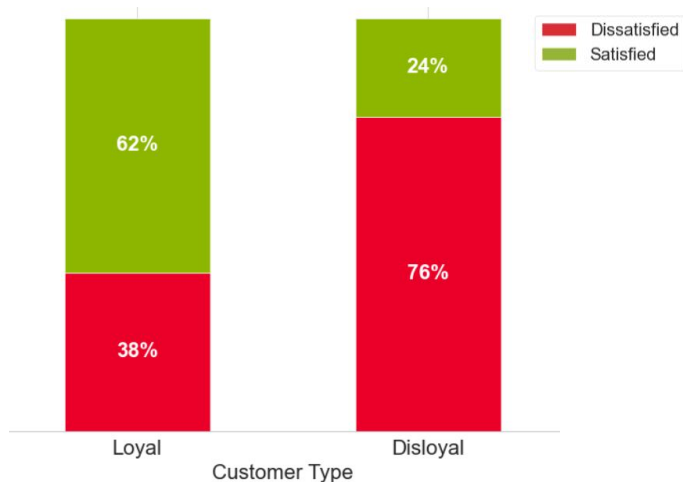
4

El índice de satisfacción para usuarios leales está 8 p.p. por debajo del objetivo.

## Loyalty and Satisfaction

Los usuarios **Loyal** son el **80%** de la encuesta.

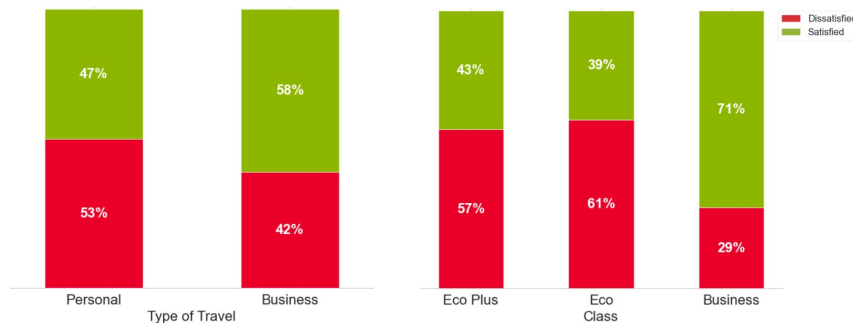
Los usuarios Loyal son más propensos a encontrarse satisfechos, sin embargo, los niveles de satisfacción **se encuentran 8 p.p. por debajo del objetivo** establecido de 70%, habiendo espacios de mejora.



## “Business or Pleasure?”

Aunque ambos servicios (Negocios y Personal) están por debajo del nivel de satisfacción deseado del 70%, **el servicio de viaje Personal es el más afectado**. La composición de los tipos de viaje por clase revela que **la mayoría de los viajes de negocios se realizan en Business Class**, explicando así por qué este grupo presenta una satisfacción superior: tienen servicio Premium.

Los usuarios de **clase económica (Eco)**, que constituyen **casi el 50% de la clientela total**, experimentan la satisfacción más baja (menos del 40%).





# Invistico Airlines

4

Los "servicios online" y "servicios del aeropuerto" son cruciales para impulsar la satisfacción del cliente.

## Factores en la Experiencia Aérea

- La segmentación "servicios a bordo," "servicios online," y "servicios del aeropuerto" revela patrones clave.
- La satisfacción general se ve fuertemente afectada por la **eficacia y calidad de "servicios online" y "servicios del aeropuerto"**.
- Los "servicios a bordo" no tienen una relación directa con la satisfacción general.

Al alinearse con las tendencias de la industria y priorizar estas áreas, se puede abordar eficazmente las expectativas y necesidades del cliente, logrando una mejora significativa en la satisfacción general.







# Encoding y Feature Engineering

## Encoding

---

- Se creó una columna 'satisf\_binaria' que indica 1 si la satisfacción es 'Satisfied' y 0 en caso contrario.
- Se creó una columna 'gender\_binaria' que indica 1 si el género es 'Female' y 0 en caso contrario.
- Se creó una columna 'loyal\_binaria' que indica 1 si el tipo de cliente es 'Loyal Customer' y 0 en caso contrario.
- Se creó una columna 'travel\_type' que indica 1 si el tipo de viaje es 'Personal Travel' y 2 si es 'Business Travel'.
- Se creó una columna 'class\_type' que indica 1 si la clase es 'Business', 2 si es 'Eco Plus', y 3 si es 'Eco'.
- Se agregaron columnas adicionales 'Es\_Satisfied', 'Es\_Female', 'Es\_Male', 'Es\_LoyalC', 'Es\_PersonalT', 'Es\_BusinessT', 'Es\_Class1', 'Es\_Class2', y 'Es\_Class3' con valores binarios basados en las columnas originales.

## Feature Engineering

---

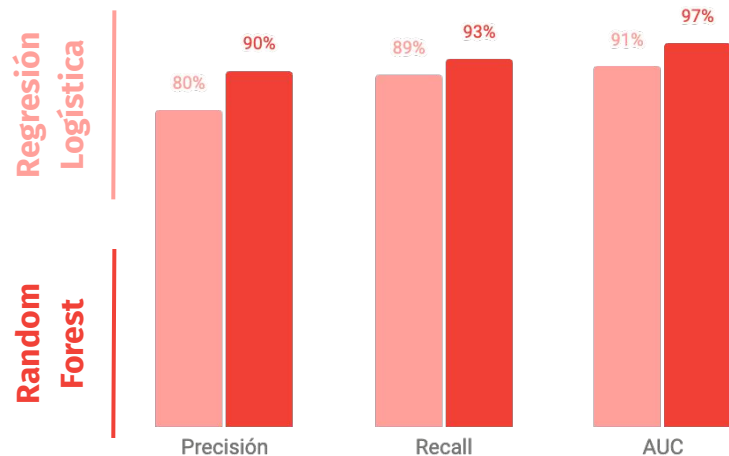
- Se realizó un análisis de correlación entre las variables de la encuesta.
- Se identificaron variables con correlación absoluta superior a 0.7, como 'Seat comfort', 'Food and drink', etc.
- Se seleccionaron solo las variables con mayor importancia para evitar multicolinealidad.
- Se encontró que 'Inflight entertainment' representaba el 25% de la importancia, seguido de 'Seat comfort'.
- Se relaciona 'Inflight entertainment' con la cantidad de usuarios en la clase económica y la facilidad de reserva de vuelos en línea.
- Se excluyeron del modelo variables con importancia menor al 1%, como 'Gate Location', 'Clase 2', y 'Arrival delay'.



# Selección del modelo

## Cross Validation: K-fold

Para ambos modelos, los resultados de la validación cruzada y las métricas del modelo en los conjuntos de entrenamiento y prueba sugieren que el modelo de Random Forest tiene un buen rendimiento y es consistente en su capacidad para clasificar las muestras en diferentes conjuntos de datos.



El modelo de *Random Forest* posee:

- un AUC medio más alto → **mejor discriminación**;
- baja desviación estándar → **estabilidad**;
- una precisión y un recall superiores → **mejor capacidad para identificar instancias positivas**.

En resumen posee un **rendimiento general superior en la clasificación de datos**.



## Algoritmo elegido

**Random Forest** para **mejor capacidad predictiva**, capacidad de discriminación y manejo de datos complejos.

## Análisis de Componentes Principales (PCA)

Se conservan todas las variables individuales sin sintetizarlas dado que es crucial para comprender completamente todos los factores que afectan a nuestro modelo. Esta decisión se basa en la granularidad de los datos, la capacidad de ajuste del modelo, la capacidad de descubrir relaciones no lineales y la transparencia del modelo.



# Insights y Recomendaciones

## Insights

### 1. Nivel General de Satisfacción y Factores Clave

- El **nivel general de satisfacción de los clientes es positivo**, con un 55% satisfechos. Sin embargo, está 15 p.p. por debajo de la meta del 70%.
- Los factores clave que influyen en la satisfacción son **la lealtad del cliente, la edad y género, tipo de viaje, clase de vuelo, y servicios específicos** como entretenimiento a bordo, comodidad del asiento, y facilidad de reserva en línea.

### 2. Impacto de Lealtad, Tipo de Viaje y Clase

- La **lealtad** del cliente está **estrechamente relacionada con la satisfacción**, pero está 8 p.p. por debajo de la meta.
- **Ambos tipos de viaje (negocios y personal) tienen niveles de satisfacción por debajo del objetivo.** La clase de vuelo impacta, siendo la satisfacción más baja en económica y más alta en negocios.

### 3. Impacto de Servicios

- Los **servicios en línea y del aeropuerto tienen una fuerte correlación con la satisfacción general.**

## Recomendaciones

### 1. Tipo de Usuarios: Loyal (80% de la encuesta)

- Desarrollar **programas de fidelización.**
- Obtener **retroalimentación directa** de estos clientes para comprender sus expectativas y áreas específicas de mejora.

### 2. Tipo de Viaje: Foco en Business (70% de la encuesta)

- Mejorar la **experiencia para viajeros de negocios**, considerando servicios como salas VIP, check-in rápido y comodidades específicas para el trabajo durante el vuelo.
- Colaborar con empresas para ofrecer **programas corporativos** que mejoren la eficiencia y comodidad.

### 3. Clase de Vuelo: Eco está en decadencia (50% de la encuesta)

- Mejorar la **experiencia** brindando mejores opciones de entretenimiento y alimentación.
- Ofrecer **promociones y descuentos** para la clase económica.

### 4. Servicios: Online y del Aeropuerto

- **Optimizar la plataforma en línea** para una navegación fácil y rápida, mejorando la experiencia de reserva.
- **Implementar tecnologías innovadoras en el aeropuerto**, como check-in automático y servicios móviles, para agilizar las interacciones en el aeropuerto.



# Bibliografía

- **Dataset:** "Airlines Customer satisfaction - Customer satisfaction with various other factors" creado por Sayantan Jana y actualizado en 2019. Se puede encontrar el conjunto de datos en la siguiente URL: [Airlines Customer satisfaction | Kaggle](#)
- **Notebooks:** [https://github.com/florderudder/proyectofinal\\_coderhouse](https://github.com/florderudder/proyectofinal_coderhouse)
- **Caso de éxito:** "Avion Revue. 'Aerolíneas Argentinas mejora el índice de satisfacción de sus pasajeros.' Disponible en: [Aerolíneas Argentinas mejora el índice de satisfacción de sus pasajeros – Avion Revue Internacional](#)