

# Módulo 4: Visualización

## 1. Resumen

¡Bienvenidas al emocionante proyecto "Data Insights: ETL y Visualización Impactante en Tableau/Power BI" en nuestro bootcamp de análisis de datos!

En este desafío, nos sumergiremos en el apasionante mundo del *reporting* de datos, donde adquirireis las habilidades clave para transformar datos crudos en información significativa y visualmente impactante.

Comenzaremos explorando/limpiando un conjunto de datos proporcionado, identificando patrones y determinando las mejores prácticas para la limpieza y transformación de datos. A través de Pandas, abordaremos problemas comunes como datos faltantes o inconsistentes (es decir, afianzaremos todos los conocimientos adquiridos durante el Módulo 3).

Además, nos adentraremos en el poderoso Tableau o Power BI, donde crearemos visualizaciones interactivas y atractivas que contarán historias con datos. Gracias a estas herramientas, podréis comunicar de manera efectiva y persuasiva los *insights* obtenidos durante el análisis y limpieza previo.

Al finalizar, presentareis vuestros informes finales y visualizaciones ante la clase, demostrando vuestra habilidad para comunicar hallazgos clave y conclusiones a través de datos bien presentados.

Este proyecto os permitirá desarrollar habilidades esenciales en ETL y visualización, lo que supone una ventaja competitiva en el campo del análisis de datos y el *reporting*.

## 2. Datos

### Conjunto de datos de Finanzas

Este conjunto de datos contiene información sobre reservas de hoteles. A continuación, tenemos una descripción de lo que nos podemos encontrar en el conjunto de datos:

Columna	Descripción
hotel	Tipo de hotel
is_canceled	Indica si la reserva fue cancelada (True) o no (False)
lead_time	Número de días entre la fecha de reserva y la fecha de llegada al hotel
arrival_date_year	Año de llegada al hotel
arrival_date_month	Mes de llegada al hotel
arrival_date_week_number	Número de la semana de llegada al hotel
arrival_date_day_of_month	Día del mes de llegada al hotel
stays_in_weekend_nights	Número de noches que el cliente se quedó durante el fin de semana
stays_in_week_nights	Número de noches que el cliente se quedó durante la semana
adults	Número de adultos que acompañaban al cliente en la reserva
children	Número de niños que acompañaban al cliente en la reserva
babies	Número de bebés que acompañaban al cliente en la reserva
meal	Tipo de comida incluida en la reserva (BB: Desayuno, HB: Media Pensión, FB: Pensión Completa)
country	País de origen del cliente
market_segment	Segmento de mercado al que pertenece la reserva
distribution_channel	Canal de distribución utilizado para realizar la reserva
is_repeated_guest	Indica si el cliente es un huésped repetido (1) o no (0)
previous_cancellations	Número de reservas canceladas por el cliente antes de esta reserva
previous_bookings_not_canceled	Número de reservas no canceladas por el cliente antes de esta reserva
reserved_room_type	Tipo de habitación reservada
assigned_room_type	Tipo de habitación asignada en la reserva
booking_changes	Número de cambios realizados en la reserva
agent	Identificador del agente involucrado en la reserva
company	Identificador de la compañía involucrada en la reserva
days_in_waiting_list	Número de días que la reserva estuvo en lista de espera
customer_type	Tipo de cliente que realizó la reserva (Transient, Contract, Group, Transient-Party)
adr	Tarifa promedio diaria pagada por la reserva
required_car_parking_spaces	Número de espacios de estacionamiento requeridos por el cliente
total_of_special_requests	Número total de solicitudes especiales realizadas por el cliente
reservation_status	Estado de la reserva (Check-Out: Salida, Canceled: Cancelada)
reservation_status_date	Fecha del estado de la reserva

#### Problema de este conjunto de datos: Análisis y Optimización de Cancelaciones de Reservas

El hotel ha notado un aumento en el número de cancelaciones de reservas en los últimos meses y necesita comprender las causas y patrones detrás de estas cancelaciones para tomar medidas correctivas. El objetivo es llevar a cabo un análisis exhaustivo de las cancelaciones de reservas y crear visualizaciones impactantes para que la gerencia pueda identificar áreas de mejora y tomar decisiones informadas para reducir la tasa de cancelación. Algunas de las tareas que podríamos hacer son:

- Limpieza de Datos:** Realizar una limpieza completa del conjunto de datos, abordando posibles valores atípicos, datos faltantes y errores tipográficos que puedan afectar el análisis.
- Preguntas que nos podemos hacer:**
  - ¿Cuál es la tasa de cancelación de este hotel?
  - ¿Ha cambiado la tasa de cancelación con el tiempo?
  - ¿Se cancelan más las reservas de verano que de invierno?. ¿Se cancelan más las reservas de entre semana que las de los fines de semana?
  - Los clientes que cancelan, ¿presentan algún tipo de característica demográfica?
  - Las reservas que se hacen con mayor anticipación, ¿tienen mucho riesgo de cancelarse?
  - Las reservas que incluyen hijos, ¿tienen menor riesgo de cancelación?
  - Los usuarios que realizaron algún cambio en su reserva, ¿tienen menor riesgo de cancelación?
  - ¿Cuándo el usuario ha realizado una solicitud especial, ¿el riesgo de cancelación es menor?
  - Las reservas que tienen un "adr" bajo, ¿tienen un riesgo menor de cancelación?
  - Relación con otras variables: Analizar cómo otras variables, como el tipo de comida incluida en la reserva o el número de solicitudes especiales, están relacionadas con las cancelaciones.

  
21MB

finanzas-hotel-bookings.csv

## 3. Objetivos

- Consolidar los conocimientos de visualización básicos.
- Implementar Scrum como marco de referencia para el desarrollo del producto, basándonos siempre en los valores de *Agile* como puntos clave del trabajo en equipo y la mejora continua.
- Mejorar la comunicación entre los miembros del equipo.
- Mejorar vuestras habilidades de comunicación en público al exponer el proyecto en la sesión final.

## 4. Planificación del proyecto

### 4.1. Sprints

Para la realización de este proyecto trabajaremos en 2 *sprints* (2 iteraciones). Siguiendo los principios ágiles, estableceremos pequeños ciclos iterativos de forma que al final de cada uno generemos valor perceptible por nuestros usuarios. Dedicaremos el primer día a la planificación del *sprint* (*sprint planning*) y el resto a trabajar en el desarrollo del proyecto. Al final de cada *sprint* haremos un *Sprint Review* del proyecto para presentar los resultados conseguidos y recoger *feedback*.

También haremos una retro corta revisando los *working agreements* que hemos acordado al inicio del proyecto y añadiendo cualquier otro *feedback* que nos permita mejorar el proyecto.

Al final del segundo **sprint** (que coincidirá con el final del proyecto y del módulo), haremos una sesión de presentación más completa, más allá de lo que sería un *Sprint Review*.

### 4.2. Criterios de aceptación

- Crear la infraestructura necesaria: repositorio en GitHub y con acceso para todos los miembros del equipo.
- Crear un *dashboard* donde presentéis los resultados de vuestro análisis .
- Tener en el repositorio de GitHub todo el código del desarrollo del proyecto.

### 4.3. Historias de usuario

Como hasta ahora, usaremos historias de usuario para la gestión del proyecto. Recordemos que, las historias de usuario son descripciones breves y concretas de las funcionalidades o características que un usuario espera encontrar en un producto o sistema. Recordemos que, las historias de usuario son una herramienta importante para asegurarse de que el equipo de desarrollo (es decir, nosotras) entienda las necesidades de los usuarios y construya el producto de manera efectiva.

Algunas de las historias de usuario de este proyecto pueden ser:

- Seleccionar el conjunto de datos sobre el que queremos hacer el proyecto.
- Limpieza del conjunto de datos y prepararlo para el reporte.
- Realización de uno o más *dashboard* en Tableau o Power BI para la presentación de los resultados.

## 5. Entrega

El formato de entrega de este proyecto será mediante la subida de éste a la plataforma de GitHub. Para subirlo, se creará un repositorio en la organización de Adalab. El nombre del repositorio deberá estar compuesto de las siguientes partes, todo ello separado por guiones:

- La palabra: project-da.
- Letra de la promoción: promo-X.
- Número del módulo: modulo-4.
- Número del equipo: team-Y.

Por ejemplo:

- Adalab/project-da-promo-X-modulo-4-team-Y

En las *sprint review* se revisará que se hayan solucionado todas las tareas técnicas asociadas a la entrega de esos *sprints*, si algo quedara pendiente se arrastraría al siguiente *sprint*.

## 6. Presentación

El último día del módulo presentaréis la versión final de este proyecto a vuestras compañeras y al equipo de Adalab. Cada equipo realizará una presentación de 5 minutos y posteriormente habrá 5 minutos de *feedback* por parte del público. En este caso, la audiencia podría ser más variada pues no sólo estarán los profesores.

El objetivo es que practiquéis la realización de las demos de los proyectos que habéis desarrollado, explicándolo desde un punto de vista técnico y también desde la perspectiva de producto, mejorando además vuestras habilidades de exposición, objetivo de desarrollo profesional del curso.

Para que la presentación salga bien es imprescindible una buena preparación. Por ello, durante el segundo *sprint* del módulo tendréis que asignar responsabilidades dentro del equipo relacionadas con la preparación de ésta. Algunas tips para preparaos este demo son:

- Todas las participantes del equipo deben hablar en la presentación (sin práctica no hay mejora).
  - Identificad los objetivos de la presentación: Debemos tener claro qué es lo que queremos lograr con la presentación. ¿Queremos demostrar la funcionalidad de un producto? ¿Queremos mostrar los resultados de un experimento? ¿Queremos atraer inversores? Dependiendo del objetivo, deberemos enfocar la presentación de manera distinta.
  - Conoce a tu audiencia: La presentación debe estar adaptada al tipo de audiencia que se espera. Si se presentara ante posibles inversores, la presentación debe estar enfocada en los beneficios y la rentabilidad del producto. Si es una presentación para usuarios, deberá enfocarse en la usabilidad y la facilidad de uso.
  - Debéis ser claras y concisas: La presentación debe ser fácil de entender y no debe ser demasiado larga. Es importante presentar la información de manera breve y clara. Debemos tener en cuenta que los detalles técnicos pueden ser interesantes, pero no deberían opacar la idea principal.
  - Demostrad la funcionalidad del producto: Si la presentación es para demostrar un producto, es importante que mostremos su funcionalidad. Podemos hacerlo mediante un video o demostrando el producto en vivo. Es importante asegurarse de que el producto funciona correctamente antes de la presentación.
  - Resaltad los aspectos más importantes: En cualquier presentación, hay aspectos que son más importantes que otros. Debéis resaltar los aspectos que sean más relevantes para su objetivo. Por ejemplo, si queréis atraer inversores, deberéis resaltar los beneficios y la rentabilidad del producto.
  - Practicad la presentación: Es importante que practiquéis la presentación varias veces antes del evento. Debéis aseguraos de que la presentación esté bien estructurada y que os sentís cómodas hablando frente a la audiencia.
  - Preparad respuestas a preguntas frecuentes: Es probable que la audiencia tenga preguntas después de la presentación. Debéis preparaos para responder preguntas frecuentes y tener la información necesaria a mano.
- Además de esto, para mejorar vuestras habilidades de exposición en público y hacer la presentación más rica, podréis incorporar otros elementos adicionales (son solo ideas, sentíos libres de innovar y ser creativas):
- Dejar muy claro quién ha sido vuestro cliente y qué fue lo que os pidió.
  - Explicar qué necesidades cubre o qué problemas soluciona el producto, cuál es el beneficio principal que aporta y qué lo hace único comparado con otros productos parecidos del mercado.
  - Aportaciones "únicas y diferenciadoras" de cada equipo al proyecto.
  - Cómo ha sido la organización del equipo, el reparto de tareas y la coordinación a la hora de trabajar todas en el mismo código.
  - Cuál de las tareas o los puntos ha sido el que más esfuerzo ha requerido.
  - Cuál de las tareas o partes del producto es la que hace que el equipo esté más orgulloso.
  - Las tecnologías que habéis utilizado y para qué sirven, y algunas partes del código que habéis desarrollado que merezca la pena resaltar.
  - La presentación debe tener un "buen inicio y un buen cierre" que nos haga a todos estar atentos y aplaudir... ah! os dejamos que echéis a volar vuestra imaginación.
  - No habléis en primera persona de lo que habéis hecho, hablad del equipo.
  - No mencionéis problemas, sino "retos" que os han hecho aprender y crecer.