

## **75.06/95.58 Organización de Datos**

### **Primer Cuatrimestre de 2024**

### **Trabajo Práctico 1: Enunciado**

El trabajo práctico 1 de la materia se basa en el análisis de datos del registro histórico de consumos de la línea aérea Airways de <https://www.kaggle.com/datasets/ghassenkhaled/airways-customer-data>. Los datos incluyen reservas, venta de pasajes y reviews. Datos

#### **Customer\_airways\_data**

- Num\_passengers: Cantidad de pasajeros que viajaron
- Sales\_channel: Canal de venta por el cual fue realizada la reserva.
- Trip\_type: Puede ser Round Trip (ida y vuelta), One way (solo ida) o Circle trip (secuencia de tramos)
- Purchase\_lead: Número de días entre el momento de la reserva y el momento del viaje.
- Length\_of\_stay: Número de días en el destino antes del regreso.
- Flight\_hour: Horario de partida del vuelo
- Flight\_day: Día de la semana de partida del vuelo
- Route: Destino
- Booking\_origin: País desde donde se realizó la reserva.
- Wants\_extra\_bags: Si el cliente solicitó equipaje extra al hacer la reserva.
- Wants\_preferred\_seat
- wants\_in\_flight\_meals
- flight\_duration
- booking\_complete

#### **Cleaned-reviews**

- Reviews: Texto conteniendo el review del usuario
- Rate: Puntaje (de 1 a 10) otorgado
- Date: Fecha
- Country: País de origen.
- Verified: Si el registro fue validado y se puede considerar un review legítimo
- Comments: Texto ingresado por el usuario.

## **Objetivo**

El objetivo del primer TP es realizar un análisis exploratorio del set de datos. Queremos ver qué cosas podemos descubrir sobre los datos que puedan resultar interesantes.

## Requisitos

Los requisitos de la primera entrega son los siguientes:

- El análisis debe estar hecho en Python Pandas.
- El análisis debe entregarse en formato pdf, conteniendo el análisis realizado, explicaciones, hipótesis, resultados, y visualizaciones que puedan resultar de interés.
- Debe informarse también el link a los notebooks / un repositorio Github en donde pueda bajarse el código completo para generar el análisis.
- El trabajo no debe limitarse a analizar cada archivo por separado, se espera que puedan buscar relaciones entre los datos de distintos archivos tratados en forma conjunta.
- El trabajo debe incluir al menos 10 visualizaciones, orientadas a entender mejor los datos y sus relaciones, y al menos una enfocada a mostrar alguna de las conclusiones a la que llegaron. Ésta no puede ser un plot simple, deben incluirse elementos de diseño y trabajarse la estética para que la misma sirva para ser presentada a un público no necesariamente experimentado en el mundo de la aviación.

## Evaluación

La evaluación del TP se realizará en base al siguiente criterio:

- Originalidad del análisis exploratorio.
- Calidad del reporte. ¿Está bien escrito? ¿Es claro y preciso?
- Calidad del análisis exploratorio: qué tipo de preguntas se hacen y de qué forma se responden, ¿es la respuesta clara y concisa con respecto a la pregunta formulada?
- Calidad de las visualizaciones presentadas.
  - ¿Tienen todos los ejes su rótulo?
  - ¿Tiene cada visualización un título?
  - ¿Es entendible la visualización sin tener que leer la explicación?
  - ¿El tipo de plot elegido es adecuado para lo que se quiere visualizar?
  - ¿Es una visualización interesante?
  - ¿El uso del color es adecuado?
  - ¿Hay un exceso o falta de elementos visuales en la visualización elegida?
  - ¿La visualización es consistente con los datos?
  - Conclusiones presentadas.
  - ¿Presenta el grupo un listado de "insights" aprendidos sobre los datos en base al análisis realizado? ¿Es interesante?

- ¿Descubrieron cosas que pueden ser de interés a las empresas de aviación, o a futuros pasajeros en busca de pasajes?

## Criterio de Reentrega

En caso de que la entrega no esté en condiciones de aprobar, se considerará la posibilidad de realizar una reentrega siempre y cuando el trabajo presentado cumpla con todas las condiciones antes descriptas para la entrega, y solamente se trate de mejorar algún punto que pueda requerir más desarrollo. No se considerarán para reentrega trabajos que hayan sido presentados en forma incompleta.

## Asignación Ayudantes

| <b>Grupo</b> | <b>Apellido y nombre Integrante 1</b> | <b>Padrón Integrante 1</b> | <b>Apellido y nombre Integrante 2</b> | <b>Padrón Integrante 2</b> | <b>Ayudante</b>       |
|--------------|---------------------------------------|----------------------------|---------------------------------------|----------------------------|-----------------------|
| <b>1</b>     | Flores Sosa<br>Zoraida Yurico         | 87039                      | Choque Carmona Luis<br>Alberto        | 93738                      | <b>Nacho B</b>        |
| <b>2</b>     | de San Vicente,<br>Hernán Martín      | 108800                     | Converso, Lara Daniela                | 107632                     | <b>Julieta</b>        |
| <b>3</b>     | Jonathan Leonel<br>Pepe               | 94692                      | Henseler Santiago                     | 110732                     | <b>Nacho B</b>        |
| <b>4</b>     | Marianela<br>Fernanda Gareca<br>Janko | 109606                     | Beltran Malbran                       | 110036                     | <b>Julieta</b>        |
| <b>5</b>     | Tomás Ezequiel<br>Apud Mora           | 107968                     | Juan Angel Gomez                      | 87943                      | <b>JuanMa/Alejo</b>   |
| <b>6</b>     | Hsieh Cindy<br>Teresa                 | 108051                     | Dominguez Gonzalo<br>Alejo            | 109759                     | <b>Nacho A</b>        |
| <b>7</b>     | Nemirovsky<br>Federico                | 108560                     | Czerwiak Juan Joaquin                 | 109640                     | <b>Nico</b>           |
| <b>8</b>     | Batallan David<br>Leonardo            | 97529                      | Fontana Maria Agustina                | 108090                     | <b>Joaco</b>          |
| <b>9</b>     | Cristian Eduardo<br>Roldan Montes     | 96713                      | Nicolas Rizzo<br>Ehrenbock            | 109756                     | <b>Martín</b>         |
| <b>10</b>    | Pelliciari Agustin                    | 108172                     | Dellisola Florencia                   | 109897                     | <b>Matías Rotondo</b> |
| <b>11</b>    | Simone Abril                          | 110939                     | Fiorotto Camila                       | 108239                     | <b>Lucas Correa</b>   |
| <b>12</b>    | Brandan, Esteban                      | 109186                     | Alonso Schneider,                     | 109551                     | <b>Matías Rotondo</b> |

|           |                               |        |                                 |        |                     |
|-----------|-------------------------------|--------|---------------------------------|--------|---------------------|
|           | Ariel                         |        | Micaela Jazmin                  |        |                     |
| <b>13</b> | Martina Gualdi                | 110513 | Cruz Marcela Jazmin             | 110066 | <b>Nacho A</b>      |
| <b>14</b> | Sol Guadalupe Urbano          | 109525 | Federico Martin Forte           | 92389  | <b>Nati</b>         |
| <b>15</b> | Leith Cristian                | 103494 | Guttlein Gareis Alexis Daniel   | 104431 | <b>Juan Pablo</b>   |
| <b>16</b> | Franco Bricchi                | 104438 | Matías Morales                  | 106793 | <b>Nico</b>         |
| <b>17</b> | Pablo Ezequiel Ruiz           | 109407 | Lucas Couttulenc                | 109726 | <b>Lucas Correa</b> |
| <b>18</b> | Galarza Adrian                | 106194 | Vargas Rodrigo                  | 97076  | <b>Lucas Fiegl</b>  |
| <b>19</b> | Santellán Felipe Martin       | 109022 | Brandan Ricardo Ezequiel        | 96970  | <b>JuanMa/Alejo</b> |
| <b>20</b> | Lucas Incerti                 | 106972 | Martin Quilis                   | 105005 | <b>Juan Pablo</b>   |
| <b>21</b> | Brizuela Lopez, Mariano Jesus | 105079 | Fernandez Moreno, Tomas Agustin | 109951 | <b>Nati</b>         |
| <b>22</b> | Axel Aparicio Rotermund       | 96283  | Ignacio Orona                   | 108013 | <b>Lucas Fiegl</b>  |
| <b>23</b> | Aizenscher Tomas              | 105350 | Facundo Lucio Brondo            | 97640  | <b>Damu</b>         |
| <b>24</b> | Graziosi German               | 104623 | Condori Luis                    | 98237  | <b>Martín</b>       |
| <b>25</b> | Idoia Larrayoz                | 183652 | Cortiñas Sabrina                | 97052  | <b>Joaco</b>        |
| <b>26</b> | Emanuel Sagman                | 103439 | Ariel Folgueira                 | 109473 | <b>JuanMa/Alejo</b> |