**Diccionario correspondiente a raw\_data\_sample5perc.zip. Buscarlo en el basecampo y no publicarlo, por favor.**

**Tendré que normalizar de alguna manera el precio con las características del vuelo**

**Precio. Centrarnos en sólo euros para simplificar. Esto es del booking reservadas. Como hay varias, hacer la media**

**Sum of all distances**

**Variable => Precio / Km**

**Busqueda:**

* **Filtramos**

**Table 1.** Data dictionary for matching row.

| **Field name** | **Type** | **Description** | **Units** | **Examples** |
| --- | --- | --- | --- | --- |
| **search\_date** | Date | Date of the search including time |  | %Y%m%d\_%H%M |
| **origin** | String | Origin of the search trip - city and airport IATA codes mixed |  | ORD, CHI |
| **destination** | String | Destination of the search trip - city and airport IATA codes mixed. OJO, que aquí puede el usuario poner nombre de ciudad o de aeropuerto. |  | NCE |
| **advance\_purchase** | Int | Days between search and first flight. Puedes segmentar mercado para tiempo de antelación | days | 34 |
| stay\_duration | Int | Days between round trip (ida y vuelta) | days | 3 |
| **departure\_day\_of\_week** | String | Week day of departure |  | Mon, Tue, ... |
| **departure\_date** | Date | Departure date (HH:MM missing) |  | %Y%m%d\_%H%M |
| return\_date | Date | Return date |  | %Y%m%d\_%H%M |
| **itinerary\_type** | String | Computation of itinerary type. Oneway or Roundtrip |  | R, O |
| **geography** | String | Trip range: Domestic/Continental/Intercontinental. 99% Domestic. |  | D, C, I |
| **persona** | String | Classification of pax Business/Holidays/Weekend/Oneway |  | O, H, W, B |
| **distance** | Float | Sum of all distances | km | 1200 |
| **currency** | String | Price currency. Vamos a tenerlas mezcladas. |  | EUR/USD |
| number\_of\_recos | Int | Number of recommendations |  | 50 |
| min\_price | Float | Minimum price given in the recommendations (in currency) |  | 234 |
| max\_price | Float | Maximum price given in the recommendations |  | 1203 |
| min\_eft | Float | Minimum elapsed flying time | hours | 5.2 |
| max\_eft | Float | Maximum elapsed flying time | hours | 9 |
| recos\_booked | List[String] | Te dice que REcomendaciones que ha sido reservadas. |  |  |
| **recos** | List[[reco](https://rndwww.nce.amadeus.net/confluence/display/TIU/CA+Data+Processing" \l "CADataProcessing-reco)] | Recommendations given (see Table 2) |  | reco |
| office | String | Office ID |  | WASPC3103 |
| office\_country | String | Office country |  | DE |
| ab\_group | String | A/B test group assignation: *A, B, DEFAULT and ABT\_IN* |  | A |
| requested\_cabins | List[String] | requested cabins from TFI |  | ['C'] |
| message\_id | String | Message ID |  |  |

**Table 2.** Recommendation (RECO) dictionary.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field name** | **Type** | **Description** | **Units** |
| **rank** | String | Rank by price, from 0 to the number of recos. Orden por precio (0 más barata, 1 más cara). A mismo precio, pone rankins diferentes. |  |
| **number\_of\_bookings** | Int | Number of bookings for the reco |  |
| bookings | List[booking] | As many bookings as number of bookings (see Table 3) | booking |
| **price** | Float | Total price of this reco (base amount + taxes + fees (including surcharges)). Assume currency from search. | currency |
| **eft** | Float | Elapsed Flight time | hours |
| **ground\_time** | Float | Ground time (some bad data - negative numbers) | hours |
| **number\_of\_connections** | Int | Number of connections of the reco |  |
| **number\_of\_passengers** | Int | Number of passengers |  |
| **number\_of\_children** | Int | Number of children |  |
| **flights** | List[flight] | As many flights as needed for each of the proposed bookings (need exploration) | flight |
| **taxes** | Float | Price of taxes |  |
| **surcharges** | Float | Price of surcharges |  |
| **fees** | Float | Amount of fees |  |

**Table 3.** Booking dictionary.

| **Field name** | **Type** | **Description** | **Units** | **Examples** |
| --- | --- | --- | --- | --- |
| booking\_date | String | Date of booking |  |  |
| delta | Float | Booking time difference wrt search time | hours | 0.139 |
| number\_of\_passengers | Int | Number of passengers |  | 2 |
| source | String | 1A\_midt, or null. We can drop this field. |  |  |
| id | String | Booking ID |  |  |
| office | String | Office name |  |  |
| office\_country | String | Office country |  |  |
| price | Float | Price of the booking |  | 320.2 |
| currency | String | Currency of the given price |  | EUR |
| taxes | Float | Taxes amount |  |  |
| taxes\_currency | String |  |  |  |
| cabins | String | Requested cabins, nasty field, should be a list. Could we drop it? |  |  |
| rbds | String | Bookings classes, nasty field, should be a list. Could we drop it? |  |  |
| origin | String | Origin of the trip |  |  |
| destination | String | Destination of the trip |  |  |

**Table 4.** Flight dictionary.

| **Field name** | **Type** | **Description** | **Units** | **Examples** |
| --- | --- | --- | --- | --- |
| departure | String | Fight Departure airport IATA code |  | DUS |
| arrival | String | Flight Arrival airport IATA code |  | BOS |
| departure\_date | String | Date of departure DDMMYY |  | '250315' |
| arrival\_date | String | Date of arrival DDMMYY |  | '250315' |
| **marketing\_carrier** | String | Marketing carrier |  | IB |
| flight\_number | String | Flight number |  |  |
| **cabin** | String | Cabin: F, M, C, W |  | C |
| **rbd** | String | Booking class |  | L |
| **segment** | Int | Segment number: only from 0-4 in the dataset. |  | 0 |
| **eft** | Float | Elapsed flight time | hour | 3.4 |
| **operating\_carrier** | String | Operating carrier (significant number missing) |  |  |
| **aircraft** | String | Aircraft model.  CHG - means Changing aircraft type in 2+ stops with same flight number  TRN - Train |  | All available: 788, ER3, FRJ, 789, 737, BUS, 7M8, E70, 320, 73H, [ERJ], [J31], CNA, 32A, 359, [772], [73J], [77W], [736], 753, [NON], [CR7], [CRJ], [DH3], [319], [AT4], [738], [ER4], E75, [744], [M80], [E7W], [73W], [32S], [BET], [ATR], [739], 32B], [SF3], [CNC], [7S8], [CN1], [777], [ERD], [757], [J32], S20], [TRN], [M90], [763], [DH4], [CHG], A@@, [752], [E90], 73G], [CR9], [77L], [M88], [75W], [764], [333], [73C], [PA2], 767], [321], [76W], [CRK], [PL2], 717, [330], [332]) |
| fare\_break\_point | Float | Fare break point (all nulls) |  | Y |
| fare\_basis | Float | Fare basis. All nulls |  |  |

Cómo incluye , a precios parecidos el orden. para comprar.

Sistema de recomendación de precios a aerolíneas.

Basándonos en antelacíon, tipo de viajes., si lleva niños.