**Column description**

In the data\_analysis\_case\_study\_part2.csv file provides you with all the data needed

to complete the second part of the Big Data Analyst case study.

Dataset description: each row of the dataset corresponds to a click-out from a user.

Columns:

• User\_id: anonymized cookie id of the visitor

• Session\_id: anonymized id of the session

• Clicked\_item\_id: id of the clicked hotel.

• Displayed\_position: position, relative to the page, of the item into the list (0=first

position)

• Page\_num: cardinal number of the page where the clicked hotel was displayed

(0=first page)

• Sort\_order: categorization of the sorting order of the list (e.g. “sort by popularity”,

“focus on rating”, etc)

• Search\_type: categorization of the type of search (e.g. “city search”, “item search”,

etc)

• Path\_id: id of the location where the listed items are situated.

• Arrival\_days: number of days from the date of the search and the date of the checkin.

• Departure\_days: number of days from the date of the search and the date of the

check-out.

• Traffic\_type: categorization of the type of the traffic (e.g. “SEM”, “Branded”, etc)

• Impressed\_items\_ids: ids of all the items which were displayed on the same page

where the click-out happened. They are in the order of display.