

Similarity Data Challenge

We have suggested using a dataset on travel insurance claims (see attachment; under [open database license](#)). It would allow us to assess candidates' abilities to work with data as well as to present key findings and results. Below is a short description of the data attributes of the data set.

You will be asked to analyze the data with the tool of your choice. Please present interesting facts, insights/patterns and key findings you have found in the data in a 15mins presentation. What additional data sources or features would help you to improve your analysis? Assume we are your clients, so focus on the business outcome.

The questions for the presentation:

- To present interesting patterns/insights found in the data set
- What practical suggestions would you give for our business based on your analysis?
- What additional data and features would help you to improve? What kind of additional features could you derive from the data?

Should you have any questions, don't hesitate to ask them in advance.

Data Description

- Target: Claim Status (Claim.Status)
- Name of agency (Agency)
- Type of travel insurance agencies (Agency.Type)
- Distribution channel of travel insurance agencies (Distribution.Channel)
- Name of the travel insurance products (Product.Name)
- Duration of travel (Duration)
- Destination of travel (Destination)
- Amount of sales of travel insurance policies (Net.Sales)
- Commission received for travel insurance agency (Commission)
- Gender of insured (Gender)
- Age of insured (Age)