**Data Scientist Associate Practical Exam Submission**

Use this template to complete your analysis and write up your summary for submission.

**Fitness Class**

* The dataset contains each record when a member registered for a fitness class.

[**Background**](https://app.datacamp.com/workspace/w/a7128a82-2f4c-4597-a985-082a7af372c1#background)

* GoalZone is a fitness club chain in Canada.

GoalZone offers a range of fitness classes in two capacities - 25 and 15. Some classes are always fully booked. Fully booked classes often have a low attendance rate. GoalZone wants to increase the number of spaces available for classes. They want to do this by predicting whether the member will attend the class or not. If they can predict a member will not attend the class, they can make another space available.

### **Features of the Dataset**

Column Name Criteria

* booking\_id Nominal. The unique identifier of the booking. Missing values are not possible due to the database structure.
* months\_as\_member Discrete. The number of months as this fitness club member, minimum 1 month. Replace missing values with the overall average month.
* weight Continuous. The member's weight in kg, rounded to 2 decimal places. The minimum possible value is 40.00 kg. Replace missing values with the overall average weight.
* days\_before Discrete. The number of days before the class the member registered, minimum 1 day. Replace missing values with 0.
* day\_of\_week Nominal. The day of the week of the class. One of “Mon”, “Tue”, “Wed”, “Thu”, “Fri”, “Sat” or “Sun”. Replace missing values with “unknown”.
* time Ordinal. The time of day of the class. Either “AM” or “PM”. Replace missing values with “unknown”.
* category Nominal. The category of the fitness class. One of “Yoga”, “Aqua”, “Strength”, “HIIT”, or “Cycling”. Replace missing values with “unknown”.
* attended Nominal. Whether the member attended the class (1) or not (0). Missing values should be removed.