# Data Scientist Interview Test, EarlySense.

**Instructions:** The test has 2 independent sections. Read the description below and use the attached files to prepare your solution.

Please send back functioning code, outcomes and any other code you used in the process (only if you believe it’s relevant). Use any programming language you would like. If it’s not Python/MATLAB, please include a running/compiling quick guide.

**Section 1.** Predict a patient outcome from a given set of features.

The CSV attached includes 7 columns, the first 6 are features based on patient’s vitals, and the last column is an outcome to predict (1 = deteriorated/ 0 = got better).

The CSV includes 150 documented cases.

Please use any method you’ll like to predict the outcome, and provide the performance you’ve achieved.

**Section 2. Data Manipulation:** We are provided with patients’ info from 2 sources:

1. JSON files that holds patient’s ID & diagnosis code.
2. CSV files with 4 rows each: patient’s ID, weight[Kg], height[cm] and age.

Not every patient appears in both.

Write the code to extract the following basic statistics:

1. Mean of weight, height, age.
2. Count and patients with Code ‘X’ over age ‘Y’ (X & Y are parameters).