

DATA ANALYTICS ASSESSMENT: ANALYSE A DATASET

By

Aaron Ward

Supervisor(s):

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
B.SC IN COMPUTING AND INFORMATION TECHNOLOGY
AT
INSTITUTE OF TECHNOLOGY BLANCHARDSTOWN
DUBLIN, IRELAND
2017

Declaration

I/We declare that this material, which I/We now submit for assessment, is entirely my/our own work and has not been taken from the work of others, except where otherwise stated. I/We have identified and included the source of all facts, ideas, opinions, and viewpoints of others in the assignment references. Direct quotations from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged and the source cited are identified in the assignment references.

I/We understand that plagiarism, collusion, and copying are grave and serious offences and accept the penalties that would be imposed should I/we engage in plagiarism, collusion or copying. I acknowledge that copying someone elses assignment, or part of it, is wrong, and that submitting identical work to others constitutes a form of plagiarism. I/We have read and understood the colleges plagiarism policy 3AS08.

This material, or any part of it, has not been previously submitted for assessment for an academic purpose at this or any other academic institution. I have not allowed anyone to copy my work with the intention of passing it off as their own work.

Dated: 2017

Author:

Aaron Ward

Business Understanding and Data Understanding

Business Understanding

This dataset provides data on subjects that with and without meningitis. It contains information such as age, gender, location, sum of health problems such as headaches, fevers and seizures. Additionally, it provides an attribute that indicates if the subject does or doesn't have meningitis, with a negative or positive value.

Business Objective

- Predict if someone is at risk of getting meningitis.

Data Mining objective

- The main objective is to create a model to predict the risk of a person getting meningitis
- This model will use the attributes provided in the dataset such as age, gender, seizure history etc to obtain an accurate prediction.
- The

Data Understanding

Describing Data

Data Exploration

Verifying Data Quality

Data Preparation

Modeling

Modeling technique

Test Design

Build and Assess the Model

Evaluation