

# Diploma SWD Group Project 1 – Part A Semester 1 - 2020

In groups of 3 – 4 students.

Dod&Gy Health (a newly formed health insurance company) has hired you to analyse and present findings relating to heart disease. For economic and timing purposes, they are interested in the three most important conditions that they should be screening for in applicants to their insurance.

They have provided the attached dataset on hospital patients and data points on 14 different conditions/tests, including a test to ascertain presence and type of heart disease (provided via <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>).

1. Prepare the datasets for analysis.

Describe the decision making used to prepare the data.

2. Use the dataset/s to ascertain a combination of the three most important conditions that indicate the highest probability of heart disease.

You can use any method you choose to develop the algorithm, but it may be an idea to check out a machine learning platform as part of your investigation,

3. Find ways and other data they may support your findings.

4. Visualise and present your findings to Dod&Gy Health.

## Submission Requirements:

1. Week 3 - Presentation and Report to the student group that includes:
  1. Decision making process in preparing data
  2. Description of process used to determine the three key conditions.
  3. Any other research/findings that support/challenge your findings
2. Week 5 - Make a 8 - 10 minute presentation to Dod&Gy Health. This presentation must include elements of the above and some visualisations to convince Dod&Gy that your findings are reliable. Every member must present roughly equal time fraction.