## Week 1 – From Problem to Approach

1. The first stage of the data science methodology is Business Understanding.
   1. True.
2. The main purpose of the analytic approach is identifying what type of patterns will be needed to address the posed questions most effectively.
   1. True.

## Week 1 – From Requirements to Collection

1. The Data Requirements stage of the data science methodology involves identifying the necessary data content, formats and sources for initial data collection.
   1. True.
2. Data scientists:
   1. Determine how to collect the data.
   2. Identify the data that is required for data modelling.
   3. Determine how to prepare the data.
3. In the Data Collection stage, the data requirements are revised and decisions are made as to whether or not more data is needed.
   1. True.

## Week 2 – From Understanding to Preparation

1. During the Data Preparation stage, clients and stakeholders aggregate the data and merge them from different sources, enabling data scientists to use clean data in the analysis.
   1. False. (Building a data set is an iterative process. The methods for defining and collecting the data can be refined until all the required information is accurately captured, even if that means looping back to a previous stage in the model.)
2. Data Preparation involves dealing with missing improperly coded data and can include using text analysis to structure unstructured or semi-structured text data.
   1. True.

## Week 2 – From Modelling to Evaluation

1. Modelling may require testing multiple algorithms and parameters.
   1. True.
2. Model evaluation includes ensuring that the data are properly handled and interpreted.
   1. True.
3. The ROC curve is a useful diagnostic tool for determining the optimal classification method.
   1. True.

## Week 2 – From Deployment to Feedback

1. Feedback is not required once the model is deployed because the Model Evaluation stage would have assessed the model and made sure that it performed well.
   1. False.
2. A data scientist determines that building a recommender system is the solution for a particular business problem at hand. What stage of the data science methodology does this represent?
   1. Analytic Approach.
3. A data scientist, John, was asked to help reduce readmission rates at a local hospital. After some time, John provided a model that predicted which patients were more likely to be readmitted to the hospital and declared that his work was done. Which of the following best describes this scenario?
   1. Even though John only submitted one solution, it might be a good one. However, John needed feedback on his model from the hospital to confirm that his model was able to address the problem appropriately and sufficiently.
4. What represents the two important characteristics of the data science methodology?
   1. It is a highly iterative process and it never ends.
5. For predictive models, a test set, which is similar to – but independent of – the training set, is used to determine how well the model predicts outcomes. This is an example of what step in the methodology?
   1. Model evaluation.
6. What are three important reasons that data scientists should maintain continuous communication with business sponsors throughout a project?
   1. They can provide domain expertise.
   2. They can review intermediate findings.
   3. They can ensure the work remains on track to generate the intended solution.