Quiz-Course 3. IBM Data Science Methodology (Coursera)

Week 1. Quiz- From Problem to Approach

1.Select the correct statement:

- **The first stage of data science methodology is Business Understanding**
- The first stage of data science methodology is Modeling
- The first stage of data science methodology is Data Collection
- The first stage of data science methodology is Data Understanding

2.The main purpose of the analytic approach is identifying what type of patterns will be needed to address the posed question most effectively

- **True**
- False

3. For the case study, a decision tree classification model was used to identify the combination of conditions leading to each patient's outcome

- **True**
- False

Week 1- Quiz- From Requirements to Collection

1.The Data Requirements stage of the data science methodology involves the identifying the necessary data content, formats, and sources for initial data collection

- **True**
- False

2.Which of the following statements are correct?

- **Data scientists determine how to collect data**
- **Data scientists identify the data is required for data modelling**
- **Data scientists determine how to prepare the data**
- None of the above

3. In the data collection stage, the data requirements are revised and decisions are made as to whether or not more data is needed

- **True**
- False

Week 2- Quiz- From Understanding to Preparation

1.In the case study, during the Data Understanding stage, data scientists discovered that not all the congestive heart failure admissions that were expected were being captured. What action did they take to resolve the issue?

- The data scientists looped back to the Business Understanding stage to redefine the requirements
- **The data scientists looped back to the Data collection stage, adding secondary and tertiary diagnoses and building a more comprehensive definition of congestive heart failure admission**
- The data scientists added the missing data manually
- The data scientists did not need to do anything. In this case, expectations for data were incorrect.

→ *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 mode*

2.During the Data Preparation stage, clients and stakeholders aggregate the data and merge them for different sources, enabling data scientists to use clean data in the analysis

- True
- **False**

3.The Data Preparation stage is very iterative and complicated stage that cannot be accelerated through automation

- True
- **False**

Week 2. Quiz- From Modeling to Evaluation

1.Which statement best describes the Modeling stage of the data science methodology?

- Modeling always uses training and test sets
- The Modeling stage is followed by the Analytic Approach stage
- **Medeling may require testing multiple algorithms and parameters**
- Modeling is always based on predictive models

2.Model evaluation includes ensuring that the data are properly handled and interpreted

- **True**
- False

3. The ROC curve is a useful diagnostic tool for determining the optimal classification model

- **True**
- False

Quiz-Course 3. IBM Data Science Methodology (Coursera)

Week 3 . Quiz- 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.

- True
- 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?

- Analytic Approach.
- Deployment
- Model Evaluation.
- Modeling.

→ The selection of a model to use should happen in the Analytic Approach stage

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?

- John only provided one model as a solution and he should have provided multiple models.
- **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.**
- John's mistake is that he lied in the Analytic Approach step of the data science methodology.
- John still needed to collect more data.

4.Data scientists may frequently return to a previous stage to make adjustments, as they learn more about the data and the modeling.

- **True.**
- False.

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?

- Analytic Approach.
- Data Requirements.
- Deployment.
- **Model Evaluation.**

6. What are three important reasons that data scientists should maintain continuous communication with business sponsors throughout a project?

- **So that business sponsors can provide domain expertise.**
- **So that business sponsors can ensure the work remains on track to generate the intended solution.**
- **So that business sponsors can review intermediate findings.**
- Actually, data scientists do not need to maintain a continuous communication with business sponsors and stakeholders.

Final Exam (Timed Quiz)

1. Select the correct statement.

- The first stage of the data science methodology is Data Understanding.
- **The first stage of the data science methodology is Business Understanding.**
- The first stage of the data science methodology is Data Collection.
- The first stage of the data science methodology is Modeling.

2. _____ is an important stage in the data science methodology because it clearly defines the problem and the needs from a business perspective.

- Modeling
- Data Collection
- **Business Understanding**
- Data Understanding

3.Which of the following analogies is used in the videos to explain the Data Requirements and Data Collection stages of the data science methodology?

- You can think of the Data Requirements and Data Collection stages as building an outpatient clinic for patients with congestive heart failure, where the medical condition is the data and the patients are the ingredients.
- **You can think of the Data Requirements and Data Collection stages as a cooking task, where the problem at hand is a recipe, and the data to answer the question is the ingredients.**

4. In what stage can techniques such as descriptive statistics and visualization applied to the data set, to assess the content, quality, and initial insights about the data?
- The Data Requirements stage
- **The Data Collection stage**
- The Data Analysis stage
- The Business Analysis stage

5. A _____ is used for predictive modeling.
- **Training set**
- Technique set
- Modeling set
- Analysis set

6. A type I error is a _____.
- False-alarm error
- False-negative error
- Hypothesis error
- **False-positive error**

7. The Data Understanding stage encompasses _____.
- Transforming data.
- Sorting the data.
- Removing redundant data.
- **All activities related to constructing the dataset.**

8. In what stage would you correct invalid values and address outliers?
- **The Data Preparation stage**
- The Data Understanding stage
- The Modeling stage
- The Data Requirements stage

9. Which of the following is NOT one of the final stages of the data science methodology?
- Evaluation
- Deployment
- Feedback
- **Data Preparation**

10. Deploying a model into production represents the beginning of an iterative process from _____, then Model Refinement, and to Redeployment.
- Scalability
- Data Storage
- **Feedback**
- None of the above

11. The data science methodology provides the data scientist with a framework on how to proceed to do what?
- Obtain data storage
- None of the above
- **Obtain answers**
- Obtain data

12. As a data scientist what is typically NOT used for exploratory analysis of data?
- Vector machines, Descriptive statistics
- Deep Learning, Data visualization
- **Clustering, Deep learning**
- Clustering, Data visualization