Data Science 101
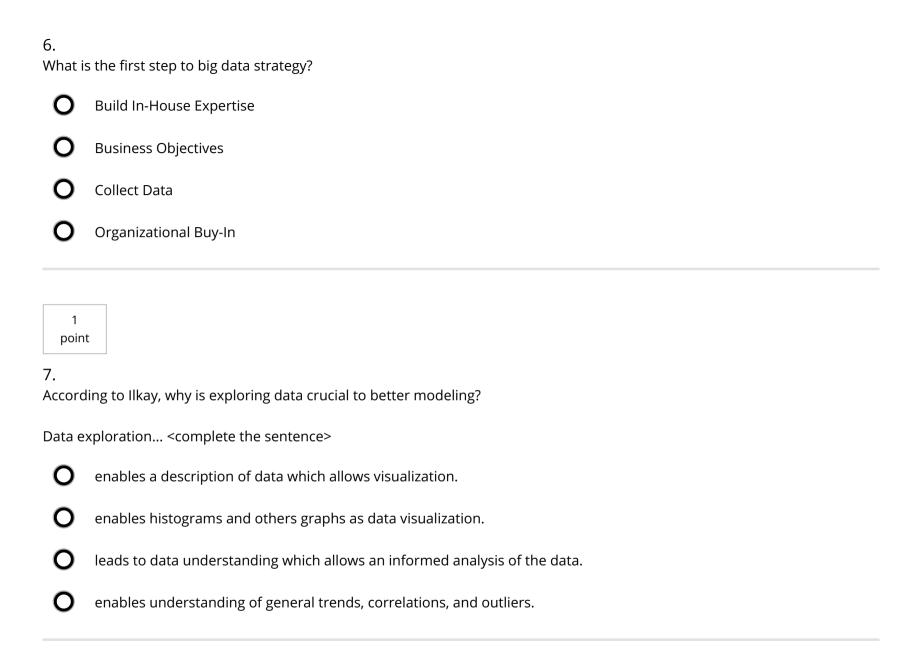
10 questions

1.
Which of the follow are parts of the 5 P's of data science and part of an additional P introduced in the slides?

- [ ] Process

- [ ] Programmability

- [ ] Product

- [ ] Platforms

- [ ] Purpose

- [ ] Perception

- [ ] People

2.

Which of the following are part of the four main categories to acquire, access, and retrieve data?

- [ ] Remote Data
- [ ] NoSQL Storage
- [ ] Text Files
- [ ] Web Services
- [ ] Traditional Databases

3.

What are the steps required for data analyzation?

- ◯ Select Technique, Build Model, Evaluate
- ◯ Regression, Evaluate, Classification
- ◯ Investigate, Build Model, Evaluate
- ◯ Classification, Regression, Analysis

4.

Of the following, what is a technique mentioned in the videos for building a model?

○   Analysis

○   Evaluation

○   Validation

○   Investigation

5.

What is the first step in finding a right problem to tackle in data science?

○   Assess the Situation

○   Ask the Right Questions

○   Define Goals

○   Define the Problem

6.

What is the first step to big data strategy?

○ Build In-House Expertise

○ Business Objectives

○ Collect Data

○ Organizational Buy-In

1
point

7.

According to Ilkay, why is exploring data crucial to better modeling?

Data exploration... <complete the sentence>

○ enables a description of data which allows visualization.

○ enables histograms and others graphs as data visualization.

○ leads to data understanding which allows an informed analysis of the data.

○ enables understanding of general trends, correlations, and outliers.

1
point

8.

Why is data science mainly about teamwork?

○ Analytic solutions are required.

○ Engineering solutions are preferred.

○ Data science requires a variety of expertise in different fields.

○ Exhibition of curiosity is required.

1
point

9.

What are the ways to address data quality issues?

☐ Remove outliers.

☐ Data Wrangling

☐ Remove data with missing values.

☐ Generate best estimates for invalid values.

☐ Merge duplicate records.

1
point