Quiz 3 – Data Science 101

1. Which of the following are parts of the 5 P's of data science and what is the additional P introduced in the slides?

- **People**
- **Purpose**
- **Product**
- Perception
- **Process**
- **Programmability**
- **Platforms**

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

- **NoSQL Storage**
- **Remote Data**
- **Traditional Databases**
- Web Services
- **Text Files**

3. What are the steps required for data analysis?

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

4. Of the following, which is a technique mentioned in the videos for building a model?

- Investigation
- Validation
- Evaluation
- **Analysis**

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

- Assess the Situation
- Ask the Right Questions
- **Define the Problem**
- Define Goals

6. What is the first step in determining a big data strategy?

- **Business Objectives**
- Collect Data
- Build In-House Expertise
- Organizational Buy-In

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

Data exploration...

- **leads to data understanding which allows an informed analysis of the data.**
- enables a description of data which allows visualization.
- enables understanding of general trends, correlations, and outliers.
- enables histograms and others graphs as data visualization.

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.

9. What are the ways to address data quality issues?

- **Remove outliers.**
- **Generate best estimates for invalid values.**
- **Remove data with missing values.**
- Data Wrangling
- **Merge duplicate records.**

10. What is done to the data in the preparation stage?

- Retrieve Data
- Select Analytical Techniques
- Build Models
- Identify Data Sets and Query Data
- **Understanding Nature of Data and Preliminary Analysis**