Keep Learning

Data Acquisition and Understanding Review

TOTAL POINTS 2

1. Which of the following statements is true?

    `1 / 1 point`

    ○ You don't need to worry about how the data is transferred from the business to you--that responsibility lies solely with the business.

    ○ Exploratory Data Analysis (EDA) is a quick process.

    ⦿ Part of data cleaning involves aligning multiple tables from the business.

    ○ There is no need for data cleaning, the data should be in perfect condition by the time you've received it from the business.

    > ✓ **Correct**
    > Correct! To start making sense of your data you will very likely have to align different datasets or multiple tables. When a business builds a database system, they normalize the databases so that they have unique small subtables that are easier to program and to store. But for data analysis and data science you want to denormalize these tables to aggregate and create a big table where all the data is in one place for your queries.

2. **Which of the following statements is true?**

    `1 / 1 point`

    ○ Exploratory Data Analysis (EDA) is meant to give you insights about the data, but it will not provide any added value to the business itself.

    ⦿ It is important to set up a data pipeline so that you can regularly refine any data coming in.

    ○ Because we want to use as much information as we can, your cleaned and refined dataset will necessarily be the same size as your raw dataset.

    ○ Because you're now in the Data Acquisition and Understanding phase of the MLPL, you will not need to worry about every returning to the Business Understanding and Problem Discovery phase.

    > ✓ **Correct**
    > Correct! This is an important step in the data acquisition and understanding phase because you're going to iterate through this process of raw data to clean data many times--so it is much better to have it as a streamlined process.