# Instructions for Review Questions

1. Time allowed: **Unlimited**

- We encourage you to go back and review the materials to find the right answer

- Please remember that the Review Questions are worth 50% of your final mark.

3. Attempts per question:

- One attempt - For True/False questions

- Two attempts - For any question other than True/False

3. Clicking the "**Final Check**" button when it appears, means your submission is **FINAL**. You will **NOT** be able to resubmit your answer for that question ever again

4. Check your grades in the course at any time by clicking on the "Progress" tab

---

# REVIEW QUESTION 1 (1/1 point)

What is Data Science?

○ It uses less complex statistics and generally tries to identify patterns that can improve an organization.

○ It is a set of instructions we give a computer so it can take values and manipulate them into a usable form.

⦿ Data science is focused on creating understanding among messy and disparate data. ✔

○ Mathematically, it is the average difference between individual values and the mean for the set of values.

○ This machine learning method uses a line of branching questions or observations about a given data set to predict a target value.

*You have used 2 of 2 submissions*

---

# REVIEW QUESTION 2 (1/1 point)

The Hadoop Distributed File System (HDFS) allows for storage, retrieval, and analysis of very large data sets using distributed hardware.: True or False. Enter question

○ False

Cookie Preferences

*You have used 1 of 1 submissions*

## REVIEW QUESTION 3 (1/1 point)

Which definition(s) of Open Data are true? Select all that apply.

☑ Open Data is defined as structured data that is machine-readable, freely shared, used and built on without restrictions

☑ Open Data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet.

☑ When using Open Data, there should be no discrimination against fields of endeavor or against persons or groups.

☑ Open data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.

☐ Open Data is free for academics only

✔

*You have used 2 of 2 submissions*

Cookie Preferences