

Course	<i>Professional Certificate Data Science & Analytics</i>	Trainer	<i>Dr Obay</i>
--------	--	---------	----------------

Trainee Details:

1	Name	Yong Zong Wei
2	Email Address	soyayong@gmail.com

In this Assignment, you will demonstrate your understanding of the videos and the readings by answering open-ended questions, defining data science and data scientist, and describing the different sections comprising a final deliverable of a data science project.

1. Give a descriptive title to this project.

Fantastic Data Scientist

2. Based on your understanding of the presentation and the readings, how can you define "Data Science" and "Data Scientist"?

- Data Science - a concept used to tackle big data that includes data cleansing, preparation, and analysis.
- Data Scientist - able to take data science projects from end to end and help to store large amounts of data, create predictively modeling processes, and present the findings.

3. What industry are you passionate about and would like to pursue a data science career in? and Why? What is the Data you think is useful?

Business industry that I would like to pursue a data science career.

This is because I can explore different types of data like Shopee, e-commerce, Lazada API for business data, and do some predict for business analysis.

The business data such as sales records, inventory data, invoice data, and traction sales data.

4. What are the ten main clear and logical components of a report that would be delivered at the end of a data science project?

- Cover page
- Table of content
- Introduction
- Methodology
- Results
- Discussion
- Conclusion
- Acknowledgement
- References
- Appendices

5. Define Big Data and its five Vs.

- Volume - Huge amount of data
- Variety - Different formats of data from various sources
- Value - Extract useful data
- Velocity - High speed of accumulation of data
- Veracity - Inconsistencies data uncertainty in data

6. There are many paths to a career in data science; most, but not all, involve a little math, a little science, and a lot of curiosity about data (True)

7. Cloud for Data Science has the limitations of not being able to deploy (False).

8. Hadoop is a collection of open-source software utilities using a network of many computers to solve problems involving Big Data (True).

9. Data Science is a part of Data Mining (False).