

## **DATA SCIENCE QUESTIONS**

**By Elsayed Ahmed**

### **1) What is Data science?**

- a) Data science is a multi-disciplinary field that uses (scientific methods, processes, algorithms and systems) to extract knowledge and insights from many structural and unstructured data which are typically large.
- b) A concept to unify statistics, data analysis, machine learning and their related methods in order to understand and analyze actual phenomena with data.

### **2) What is the Impact of Data Science?**

- a) Big data is very quickly becoming a vital tool for businesses and companies of all sizes.
- b) Data-driven businesses are worth \$1.2 trillion collectively in 2020, an increase from \$333 billion in the year 2015.

### **3) What are the skills data science should know?**

- a) Mathematics
- b) Software Engineering
- c) Machine Learning

### **4) What are Data Science Steps?**

- a) Data Gathering
- b) Data Preparation
- c) Data Exploration
- d) Building Model

### **5) What is the difference between data science and business intelligence?**

- a) Data science in term of future
- b) BI interim of past and present

### **6) What are 3V's?**

- a) Volume
- b) Velocity
- c) Variety

### **7) What is the role of machine learning in data science?**

- a) is about how systems learn from data.
- b) Use to train on data to make decisions.

**8) What are the two ways systems learn from?**

- a) Supervised
- b) Unsupervised

**9) List three data science goals?**

- a) Prediction (predict a value based on inputs)
- b) Classification (e.g., spam or not spam)
- c) Recommendations (e.g., Amazon, Wuzzuf and Netflix recommendations)

**10) What is the most famous container-orchestration system in computer application deployment?**

- a) Kubernetes

---

### MCQ Questions

**1) Point out the correct statement**

- a) Raw data is original source of data
- b) Preprocessed data is original source of data
- c) Raw data is the data obtained after processing steps
- d) None of the mentioned

**2) Which of the following is performed by a data scientist?**

- a) Define the question
- b) Create reproducible code
- c) Challenge results

d) All of the mentioned

**3) Which of the following is the most important language for Data Science?**

a) Java

b) Ruby

c) R

d) None of the mentioned

**4) Raw data should be processed only one time.**

a) True

b) False

**5) Which of the following is one of the key data science skills?**

a) Statistics

b) Machine Learning

c) Data Visualization

d) All of the mentioned

**6) Data science is the process of a diverse set of data through?**

a) Organizing data

b) Processing data

c) Analysing data

d) All of the above

**7) Which of the following is correct skills for a Data Scientist?**

a) Probability & Statistics

b) Machine Learning / Deep Learning

c) Data Wrangling

d) All of the above

**8) Which of the following are the Data Sources in data science?**

- a) Structured
- b) UnStructured
- c) Both A and B
- d) None Of the above

**9) Which of the following is not an application for data science?**

- a) Recommendation Systems
- b) Image & Speech Recognition
- c) Online Price Comparison
- d) Privacy Checker

**10) Supervised learning is a process of reorganizing and enhancing the inputs in order to place structure on unlabeled data?**

- a) True
- b) False