Amro Dajani

Past Papers questions
key answers

Q1: Do you consider Amazon data as an example of Big data?

Yes, Amazon has large amounts of data gathered from many different sources which need to be stored and processed. Amazon's data is also characterised by Volume as it has large sizes of data, velocity as data is gathered on daily basis, and variety as there are different types of data in terms of Structuredness.

Q2: What is the relationship between data management sub-system and model management sub-system, how does this relation end up with a recommended decision?

Data management is the fuel that provides the model management with the data from various sources. Then the model management sub-system will use this data for analysis to be able to process it and provide the recommended decision.

Q3: What is the role of transactional information systems in the high-level architecture of BI?

The transaction processing systems are constantly involved in handling updates to what we might call operational databases.

Q4: List and briefly describe the three main types of business reports:

1) Metric Management Reports → Business Performance is managed through outcome oriented metrics such as Service-level agreements for external groups, and KPIs for internal management.

2) Dashboard-Type Reports → Present a range of different Performance indicators on one page, dashboard vendors would Provide a set of predefined reports with static elements and fixed structure.

3) Balanced score-card type reports → Present an integrated view of success in an organization Like financial Performance, customers, business Processes and Learning and growth Perspectives.

Q5: Give one example of using NLP:

using Alexa (Smart home), it will understand your Language and fullfill your orders.

Q6: Assume you are working in the Loan department at ABC bank.

Daily, you have many car Loan applications. Your Job is to study these applications and then decide whether to approve it or not:

A) Prepare a principle of choice for this decision including three factors that you can use to evaluate each applicant:

B) what are the alternatives to make the decision on each applicatio.

C) what would you do in the implementation phase?

D) in such a decision, what is the importance of monitoring phase?

②

A)

	5	3	1
Salary	>3000	1500 - 3000	≤ 1500
Age	≤ 40	25 - 40	≤ 25
Balance	≥10,000	5,000 - 10,000	≤ 5,000

B) The alternatives are (yes/No), we either give the Loan or don't give the Loan.

C) In the implementation Phase, I would check Past Loans of customers with the same characteristics and see whether they were committed or not.

D) The monitoring Phase is very important as we monitor the customer to see if he will be committed to this Loan to keep this data for future decision making.

Q7: Which Level of business analytics can be used in each of the following cases?

1) weather fore casting for the next 10 Days? Predictive.

2) measuring students' Performance during covid? Descriptive

3) Study the consequences of automating all the Processes in the company: Prescriptive

③

Q8: Decision making process:

A) why is it important to define the problem ownership?

B) If the decision is: "Should I do masters after my bachelor or not?" then what are the alternatives.

C) According to Simon, the monitoring phase is not a primary phase, however, other approaches consider it as a 5th phase, which approach is more suitable from your point of view?

A) Defining problem ownership is very important, because if it is not established, then someone is not doing his job and it has yet to be identified as belonging to anyone. It is important that someone takes the responsibility otherwise the problem will never be solved.

B) The alternatives are (yes/no) I either continue masters after my bachelor, or I don't.

C) From my point of view, the monitoring phase is very important because it provides the required feedback so we can know if our decision was right and optimal, it is done by monitoring as the intelligence phase applied to the implementation phase.

Q9: Match the datasets X and Y with the curves A and B then justify your answer:

$X = \{7, 14, 18, 19, 25, 25, 29, 31, 37, 45\}$
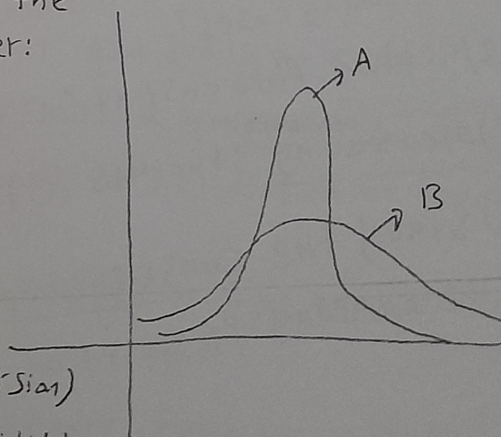
$Y = \{15, 16, 19, 25, 25, 25, 27, 29, 33, 36\}$

I will use range:

$X = 45 - 7 = 38$

$Y = 36 - 15 = 21$

$X \rightarrow B$ because the range is higher (dispersion)

$Y \rightarrow A$ because the range is lower (centrality)

④

Q10: Choose the most suitable data mining tool (classification, clustering, regression, time series analysis, association, outlier analysis):

1) To discover relationships between symptoms and illness in the medicine field: Association.

2) To define three groups of tourists in Jordan in order to serve them better according to their preferences: clustering.

3) To expect whether a corporate will have bankrupted during the next year or not: Regression.

4) To forecast the cost of moving an equipment using the cost data for the last six months: Time-Series analysis.

5) To detect fraudlent banking transactions: outlier analysis.

Q11: Choose the most suitable data mining tool (classification, clustering, regression, time series analysis, association, outlier analysis):

1) Given the number of students who applied to join the university for the last 15 academic years, you are asked to build a predictive model to estimate the number of students for the next academic year: Time series analysis.

2) There is a plan to start a new student club for the business technology school. You are asked to study the student's profiles and find the common interests to define the new club's Aim: clustering

3) PSUT is planning to study the <u>factors</u> affecting the students' performance in their exams. So that, they can work on these factors to improve the academic files of its students: Association.

4) The admission and registration dean ship is looking to build a <u>predictive model</u> using students' historical ~~the~~ data that can be used to tell whether a new applicant (student) fits into the <u>BIT</u> program or not: classification.

Good Luck.