

International Institute of Health Management Research (IIHMR)
NEW DELHI
HIT 712 Datamining and Data warehouse in Healthcare
Batch- 2018- 2020
Term Exam

TOTAL TIME: 3.0 HOURS

MAXIMUM MARKS: 70

Part A (Marks 35)

- All answers must be brief, about 3-4 lines at most.
 - Answer any Ten questions. Each question in this section is 2 marks.
1. What is the purpose of an Operating system (OS)? Can you name some parts of an OS.
 2. What do web browsers do? What happens internally when we open a web site in a browser, i.e. what communication takes place?
 3. What service do Amazon Web Services (AWS) and Google Cloud provide?
 4. What is a virtual machine? Write briefly how can virtual machines help you as a Hospital administrator or as a Developer.
 5. a) What is the need for a Datawarehouse?
b) How can a datawarehouse help a hospital?
 6. What is a Data mart?
 7. Name any two architectures a Datawarehouse can be organized. Briefly write the features of these architectures.
 8. Would Date & Time be stored with data in datawarehouse. Why?
 9. What is a Relational Database (RDBMS)?
 10. What are the different ways data can be stored in a database. Name a few architectures or products you know.
 11. What is an Operational database in an organization. What does it store?
 12. What is an OLTP system?
 13. What is a Cube?

Long Answer (mandatory question). 3 marks

14. If you are IT Head of a Hospital which has good internet facility – would you use local servers to install your Hospital management software or would you look for cloud based solutions. Justify your choice. (You can write more than 3-4 lines for this question)

Objective Type & Short Answer (Tick or write at most one line). One mark each.

15. What is a web server?
16. What is the job of an ETL tool
17. In what format would you export the data from a MySQL database (or other kinds of databases) to import into a Datawarehouse.
18. What is Meta data?
19. What is a Fact table?
20. What is a dimension?
21. Do Transactions happen in a Datawarehouse? (Y/N)
22. Typically, how frequently is a Datawarehouse data updated? What does it depend on?
23. A Datawarehouse has Sales data of a company. It has 1000 products, 50 stores and we need to store 2 year monthly data. How many rows would this small Datawarehouse have?
24. A hospital management wants year on year change of daily flow of patients to each department. What would be your grain for the time dimension?
a. Quarterly b. Monthly c. Weekly d. Daily e. Hourly
25. If you hire web server space from a service provider (any company) on Linux, would the provider give you a physical computer connected on the internet or a virtual machine?
26. What factors would you consider while hiring a web server for a hospital where your Hospital management application will run? (Hint: what do you consider when you buy a mobile phone)

Part B (Marks 35)

1. You have started as the Chief Digital Officer in a leading Hospital group. The nationwide group has 15 multi-specialty hospitals in tertiary care with over 8,000 beds. The group serves nearly 15 million patients.

Your mandate is to use state-of-the-art technologies to bring improvements in -

- Top line i.e. revenue, patient volume, procedures
- Bottom line i.e. operational and cost efficiency
- Clinical outcomes

1. What data science and/or machine learning projects do you propose to undertake? In other words, what key problems are you solving? **(10 marks)**
2. How do you see the life cycle of any of these projects? What preparatory steps and resources are needed before you start any data modeling? **(10 marks)**
3. Read the data and answer the following 15 marks

Age	Gender	IOP	Telomere
31	F	16	8,800
41	F	17	6,500
25	M	14	9,000
51	F	20	5,700
35	M	17	7,900
45	M	14	5,600
51	M	20	4,100
37	M	19	7,700
43	M	18	6,000
55	F	18	5,000
61	F	25	4,000
30	F	14	9,800
38	M	18	7,000
46	M	17	5,000
56	F	21	5,000
34	F	17	8,300
44	F	20	6,000
29	F	16	9,000
58	M	24	3,800
60	M	24	3,100
31	F	15	8,700
41	F	18	6,600
25	F	14	9,400
51	F	19	5,800
35	F	15	7,900
45	F	19	6,200
49	M	16	4,900
37	F	16	8,000
43	F	17	6,400
55	M	21	3,900
61	F	23	4,200
30	M	12	8,500
38	M	17	7,300
46	M	18	5,300
56	F	17	5,100
34	M	14	7,600
44	F	18	6,500
29	F	12	8,900
58	F	18	4,900

60	F	21	4,700
28	F	15	9,200
61	F	22	4,600
39	M	18	6,500
55	F	18	5,500
31	M	17	8,300
41	F	18	7,200
41	F	16	8,100
48	M	15	4,600
62	M	23	2,900
32	M	15	8,000

Perform Exploratory data analysis and identify outliers