Sample Questions on Business Intelligence

Note This list of questions is provided as a source of orientation only. You can use it for supervised learning ;) Other questions related to the course can be raised at the exam, as well.

BI Fundamentals

1. What is Business Intelligence? Why is it important for the business? How can the business benefit from implementing BI?
2. Which are the main stages of BI process? Which are the objectives and the programming tasks at each stage?
3. Which categories of data BI works with? What do they represent? Discrete and continuous data are represented differently. How? What is the difference between nominal and ordinal data?
4. What are the data ingestion and data wrangling? Which operations with the data are performed at that stage of BI process?
5. What are the data warehouse, data mart, and data lake? What is the difference between them? Which processes are referred by the abbreviations ETL and ELT? Name some of the difference between the two.
6. The quality of data is very important for successful data analytics. Remember GIGO? Which are the criteria of good quality? What can be done to improve the data quality?
7. What is known by data transformation? Give some examples of operations of transformation. In which cases is it important part of the data pre-processing?
8. What is known by features engineering? Give some examples of operations of this category. In which cases is it important part of the data pre-processing?

Exploration

9. In statistics, what are population and sample? How do they differ? What is a normal distribution? Which are the characteristics of a normal distribution?
10. What is Descriptive Statistics and where does it find place in BI? Which instruments of Descriptive Statistics are used for exploratory analysis of data in BI?
11. What are the central tendency and the variability? What are their measures? Explain the meaning of these measures, give examples.
12. What are the outliers. What is their role in the analysis? Name two methods for discovering outliers in a data set.
13. What is measured by the parameter correlation? How is the correlation coefficient used in the exploratory data analysis?
14. In traditional BI systems data cube is a fundamental data structure? How is it defined? What are the dimensions and the measures of a cube? Which other data structures are associated with it?
15. Which operations can be performed on the data cube? Give examples.
16. What is the role of visualisation in data exploration? What do diagrams, dashboards, and data stories show? What is the role of interaction with the data?
17. How to design a usable dashboard? Whish strategies to apply for building informative data story?
18. What are the advantages of 3D visualization versus 2D visualization in data science? Name some types of visuals you have used in your projects.

Prediction

19. What is called predictive analytics? How is it different from exploratory data analysis? Which methods of predictive analytics are you familiar with?
20. Which branch of AI is known as machine learning? Give your own explanation of the meaning of it. What are the goals of ML, which data related tasks are solved by ML?
21. Describe the process of machine learning. Which activities would you plan to solve a BI task by implementing machine learning methods? Draw a simple diagram for illustration.
22. What is the difference between a method and a model in machine learning? And algorithm? Which machine learning libraries and frameworks you are familiar with? Which range of functions each of them provides?
23. What is the difference between supervised and unsupervised machine learning? Compare classification, regression, and clustering. Give examples of appropriate real-life cases for each of the three categories of tasks.
24. If you program a robot that has to sort out big potatoes from small potatoes, which machine learning method would you use? Explain how it would work. Which alternative methods you could use and which you couldn't?
25. The probability theory and the rule-based decision making are powerful instruments in machine learning. Solving what kind of tasks they can be applied for? Which related algorithms you are familiar with? Explain the difference. Explain the use of measures entropy and information gain.
26. What is clustering in machine learning? How does it differ from classification? Name some methods for clustering. How to they differ? Give examples of appropriate implementations.
27. Explain overfitting and underfitting. Which techniques would you use to avoid them? Explain cross-validation. In which cases it is recommended for use?
28. How to validate machine learning models? What kind of errors can it produce? How are they measured? Give examples of error measurement.
29. If the accuracy of a model is not good enough, what would you try for improving it?

Multimodal Data Analysis

30. Which branch of AI is known as GenAI? Which are its implementations? What is known as Transformer, Attention, and LLM?
31. Explain how images are processed by AI methods. Name some programming tasks of processing images.
32. Explain how texts in natural language are processed by AI methods. Which language-dependant tasks can be solved by text processing?
33. What is known as text embedding? Explain the process of embedding. What are the vector databases? How are they used?
34. What is RAG? What is its role in GenAI? Explain how it works. Name some advantages and disadvantages of developing RAG in the AI-powered applications.
35. What, do you think, is the future of BI-AI in the modern business and society?