



7 결측치(missing value) 처리를 위한 방법으로 옳지 않은 것은?

- ① 결측치를 무시하고 분석을 진행한다.
- ② 결측치가 있는 데이터 포인트를 제거한다.
- ③ 유사성이 있는 데이터의 대푯값으로 대체한다.
- ④ 남아 있는 모든 데이터의 평균값으로 대체한다.
- ⑤ 남아 있는 모든 데이터의 최빈값으로 대체한다.

정답 | ①

해설 | 결측치를 무시하고 분석을 진행하면 분석 결과에 오류가 생기거나 왜곡될 수 있다. 결측치는 제거하거나 대체하는 방법으로 처리해야 한다.

8 이상치(outlier) 탐지를 위해 사용하는 방법으로 볼 수 없는 것은?

- ① 논리적 기준
- ② 상자그림(box plot)
- ③ 산점도(scatter plot)
- ④ 파이 차트(pie chart)
- ⑤ 히스토그램(histogram)

정답 | ④

해설 | 파이 차트는 데이터의 구성 비율을 시각화하는 데 사용되며, 이상치 탐지에는 적합하지 않다. 상자그림, 히스토그램, 산점도 등이 이상치를 탐지하는 데 유용하다.

9 정규화(normalization)의 목적은?

- ① 결측치를 대체하기 위해
- ② 데이터의 양을 줄이기 위해
- ③ 데이터의 타입을 변경하기 위해
- ④ 데이터의 구조를 변경하기 위해
- ⑤ 데이터의 범위를 일정하게 조정하기 위해

정답 | ⑤

해설 | 정규화는 서로 다른 속성값의 범위를 일치하게 조정하여 분석이나 모델링에서 특정 속성이 과도하게 영향을 미치는 것을 방지한다.

10 데이터 시각화의 주된 목적으로 가장 적절한 것은?

- ① 결측치를 대체하기 위해
- ② 데이터의 양을 줄이기 위해
- ③ 데이터의 정확성을 높이기 위해
- ④ 데이터의 타입을 변경하기 위해
- ⑤ 데이터의 패턴과 추세를 명확하게 파악하기 위해

정답 | ⑤

해설 | 데이터 시각화는 복잡한 데이터셋을 그래프나 차트 등으로 시각적으로 표현하여 데이터를 쉽게 이해하고, 숨겨진 패턴이나 추세를 명확하게 파악하는 데 도움을 준다.

11 파이 차트(pie chart)의 한계에 대한 설명으로 가장 적절한 것은?

- ① 이상치를 쉽게 파악할 수 있다.
- ② 시간에 따른 변화를 나타내기 어렵다.
- ③ 상대적인 비율을 쉽게 파악할 수 있다.
- ④ 두 속성 간의 관계를 나타내기 어렵다.
- ⑤ 비교해야 할 항목이 많을 때 비율을 파악하기 어렵다.

정답 | ⑤

해설 | 파이 차트는 데이터의 상대적인 비율을 시각적으로 쉽게 나타내지만, 비교해야 할 항목이 많아지면 각 항목의 비율을 파악하기 어려워진다.

12 상자그림(box plot)으로 확인할 수 없는 것은?

- ① 데이터의 중앙값
- ② 데이터의 평균값
- ③ 데이터의 이상치
- ④ 데이터의 최솟값과 최댓값
- ⑤ 데이터의 제1 사분위수와 제3 사분위수

정답 | ②

해설 | 상자그림은 데이터의 중앙값, 최솟값, 최댓값, 제1 사분위수, 제3 사분위수와 이상치를 시각적으로 나타내지만, 데이터의 평균값은 제공하지 않는다.