
과목 : 데이터 이해

1 다음 중 빅데이터 분석에 경제성을 제공해 준 결정적인 기술로 가장 적절한 것은?

- ① 텍스트 마이닝
- ② 클라우드 컴퓨팅
- ③ 저장장치 비용의 지속적인 하락
- ④ 스마트폰의 급속한 확산

2 다음 중 빅데이터의 가치 산정이 어려운 이유를 나타내는 사례로 가장 부적절한 것은?

- ① 구글 검색에서 나타나는 것과 같은 데이터의 반복적 재사용
- ② 전기차 배터리 정보를 충전소 최적지 선정과 같은 2차적 목적에 활용
- ③ 독자의 전자책 독서 순서 정보가 저자의 글쓰기 방식에 영향을 주는 것
- ④ 은행 대출심사 알고리즘 작동 원리 이해의 어려움

3 다음 중 빅데이터 분석의 특성에 대한 설명 중 가장 부적절한 것은?

- ① 더 많은 정보가 더 많은 가치를 창출할 수 있는 것은 아니다.
- ② 비즈니스의 핵심에 대해 보다 객관적이고 종합적인 통찰력을 줄 수 있는 데이터를 찾는 것이 중요하다.
- ③ 빅데이터 과제와 관련된 주된 걸림돌은 비용이 아니다.
- ④ 데이터의 크기가 커질수록 분석을 많이 사용하는 것이 중요해진다.

4 다음 중 데이터 분석에 기초한 가치 창출과 관련된 설명으로 가장 부적절한 것은?

- ① 핵심적인 비즈니스 이슈에 답을 주는 분석은 기업의 경쟁전략과 밀접하게 연관된다.
- ② 복잡하고 다양한 데이터를 최적화 하는 능력이 최고의 가치를 창출하는 가장 중요한 기준이다.
- ③ 전략적 분석과 통찰력의 창출은 빅데이터 프로젝트에서 핵심적인 역할을 한다.
- ④ 기존 성과를 유지하고 업계를 따라잡는 것이 전략적 가치 기반 분석의 가장 중요한 목표는 아니다.

5 다음 중 비즈니스에 분석을 적용할 때 효과적인 분석 적용 대상에 대한 검토와 관련된 설명으로 가장 부적절한 것은?

- ① 사업적 잠재력만이 아니라 데이터 및 필요한 자원이 이용 가능한지를 고려해 대상을 택한다.
- ② 조직이 분석을 배우는 동안에는 분석의 노력을 가능한 다양한 대상에 기울이도록 한다.
- ③ 업계 상황에 한정해서 바라보지 말고 더 넓은 시야에서 차별화를 고려한다.
- ④ 무엇이 가능한지 발견하기 위한 실험을 망설이지 말고 분석기회에 대한 직관을 무시하지 않는다.

6 데이터 사이언티스트가 효과적인 분석 모델 개발을 위해 고려해야 하는 사항으로 가장 부적절한 것은?

- ① 분석모델이 예측할 수 없는 위험을 살피기 위해 현실세계를 돌아보고 분석을 경험과 세상에 대한 통찰력과 함께 활용한다.
- ② 가정들과 현실의 불일치에 대해 끊임없이 고찰하고 모델의 능력에 대해 항상 의구심을 가진다.
- ③ 분석의 객관성에 의문을 제기하고 분석 모델에 포함된 가정과 해석의 개입 등의 한계를 고려한다.
- ④ 넓은 시각에서 모델 범위 바깥의 요인들을 판단할 수 있도록 가능한 한 많은 과거 상황 데이터를 모델에 포함한다.

7 분석의 전형적인 의사결정 오류를 로직 오류와 프로세스 오류로 나뉘 보았을 때, 각 오류에 대한 설명으로 가장 부적절한 것은?

- ① 부정확한 가정을 하고 테스트를 하지 않는 것은 로직 오류이다.
- ② 결정에서 분석과 통찰력을 고려하지 않는 것은 프로세스 오류이다.
- ③ 데이터 수집이나 분석이 너무 늦어 사용할 수 없게 되는 것은 로직 오류이다.
- ④ 대안을 진지하게 고려하지 않는 것은 프로세스 오류이다.

- 8 아래 내용은 데이터, 정보, 지식의 차이점을 예시로 설명한 것이다. 각각의 용어와 예시가 적절하게 연결된 것은?

아 래

(가) A마트는 100원, B마트는 200원에 연필을 판매한다.
 (나) A마트의 연필가격이 더 싸다.
 (다) 상대적으로 저렴한 A마트에서 연필을 사야겠다.
 (라) A마트의 다른 상품도 B마트보다 저렴할 것이다.

- ① 데이터-(가), 정보-(나), 지식-(다)
- ② 데이터-(가), 정보-(나), 지식-(라)
- ③ 데이터-(가), 정보-(다), 지식-(라)
- ④ 데이터-(가), 정보-(라), 지식-(다)

- 9 다음 중 기업내부 데이터베이스의 활용과 가장 관련이 없는 것은?

- ① CRM(Customer Relationship Management)
- ② ERP(Enterprise Resource Planning)
- ③ ITS(Intelligent Transport Systems)
- ④ KMS(Knowledge Management Systems)

- 10 다음 중 빅데이터가 만들어 내는 변화와 가장 거리가 먼 것은?

- ① 가치가 있을 것이라고 예상되는 특정한 정보만 모아서 처리하는 것이 아니라 최대한 많은 데이터를 모으고 그 데이터를 다양한 방식으로 조합해 숨은 정보를 찾아내는 방식이 중요해진다.
- ② 데이터의 규모가 증가함에 따라 사소한 몇 개의 오류 데이터는 분석결과에 영향을 미치지 않기 때문에 데이터세트에 포함하여 분석해도 상관없는 경우가 많아진다.
- ③ 인과관계의 규명 없이 상관관계 분석결과만으로도 인사이트를 얻고 이를 바탕으로 수익을 창출할 수 있는 기회가 점차 늘어나고 있다.
- ④ 데이터의 양이 증가하고 유형이 복잡해짐에 따라 수많은 데이터 중에서 분석에 필요한 데이터를 선정하기 위한 정교한 표본조사 기법의 중요성이 대두되고 있다.

※ 숫자, 영문 등의 맞춤법을 준수하여 작성하여 주십시오. (부분점수 없음)

답1 아래 ㉠에 들어갈 내용을 작성하십시오.

아 래

IoT(Internet of Things) 시대가 되면 수많은 물건에 센서가 부착되어 데이터가 끊임없이 생산된다. 이 때 빅데이터 사업자들은 IoT를 통해 생산된 데이터를 저장해 두고 이 데이터를 3rd party에 API(Application Programing Interface)를 통해 공개하여 활용하도록 할 수 있다. 이러한 기능을 빅데이터의 (㉠) 기능이라고 한다.

공유

답2 아래 제시된 데이터양의 표시 단위를 작은 것부터 큰 것의 순서로 나열하십시오.

아 래

- 가. 엑사바이트(Exabyte, EB)
- 나. 페타바이트(Pettabyte, PB)
- 다. 요타바이트(Yottabyte, YB)
- 라. 제타바이트(Zettabyte, ZB)

11 아래는 빅데이터로 달성할 수 있는 경영혁신의 단계를 나열한 것이다. 빅데이터 경영혁신의 시작 단계에서 궁극적인 단계까지 바르게 나열된 것은?

아 래

- (가) 생산성 향상
- (나) 발견에 의한 문제해결
- (다) 의사결정 향상
- (라) 새로운 고객가치와 비즈니스 창출

- ① (가) → (나) → (다) → (라)
- ② (나) → (다) → (라) → (가)
- ③ (다) → (라) → (가) → (나)
- ④ (라) → (가) → (나) → (다)

12 다음 중 미국 연방거래위원회(FTC)가 제시한 '소비자 프라이버시 보호 3대 권고사항'에 해당되지 않는 것은?

- ① 상품 개발 단계에서부터 소비자 프라이버시 보호 방안을 적용
- ② 소비자 행동 예측 알고리즘에 대한 이해당사자의 접근권 보장
- ③ 소비자에게 수집된 정보 내용 공개 및 접근권 부여
- ④ 소비자에게 공유 정보 선택 옵션 제공

13 데이터 사이언티스트가 갖춰야 할 역량은 빅데이터의 처리 및 분석에 필요한 이론적 지식과 기술적 숙련과 관련된 능력인 하드 스킬(Hard Skill)과 데이터 속에 숨겨진 가치를 발견하고 새로운 발전 기회를 만들어 내기 위한 능력인 소프트 스킬(Soft Skill)로 나누어진다. 다음 중 소프트 스킬에 가장 가까운 것은?

- ① Machine Learning
- ② Statistics Modeling
- ③ Data Visualization
- ④ Distributed Computing

14 다음 중 데이터에 대한 설명으로 가장 부적절한 것은?

- ① 1바이트는 0에서 255까지의 정수 값을 표현할 수 있는 데이터의 크기를 의미한다.
- ② 영어 한 글자는 1바이트로 저장할 수 있지만 한글 한 글자는 2바이트로 저장하는 것이 일반적이다.
- ③ 수치 데이터는 용량이 증가하더라도 텍스트 데이터에 비해 DBMS에서 관리하기 용이하다.
- ④ 설문조사 데이터는 그 형태와 형식이 정해져 있지 않아 비정형 데이터라고 한다.

15 데이터웨어하우스는 기업 내의 의사결정지원 어플리케이션에 정보 기반을 제공하는 하나의 통합된 데이터 저장 공간을 말한다. 다음 중 데이터웨어하우스의 고유한 특성이 아닌 것은?

- ① 데이터웨어하우스에서는 데이터의 지속적 갱신에 따른 데이터의 무결성 유지가 무엇보다 중요하다.
- ② 데이터웨어하우스의 데이터들은 전사적 차원에서 일관된 형식으로 정의된다.
- ③ 데이터웨어하우스에서 관리되는 데이터들은 시간의 흐름에 따라 변화하는 값을 저장한다.
- ④ 데이터웨어하우스에서는 특정 주제에 따라 데이터들이 분류, 저장, 관리된다.

16 다음 중 빅데이터로 인한 변화방향을 기술한 것으로 가장 부적절한 것은?

- ① 필요한 정보만을 수집하던 시대에서 가능한 많은 정보를 수집한 후 용도를 찾아내는 시대로의 변화
- ② 단순한 상관관계보다 이론적 인과관계를 중시하는 방향으로 변화
- ③ 표본조사 보다 전수조사가 보편화되는 시대로의 변화
- ④ 데이터의 질보다 양을 중시하는 시대로의 변화

17 다음 중 빅데이터의 출현 배경과 가장 거리가 먼 것은?

- ① 하둡과 같은 분산처리 기술의 발전
- ② M2M, IoT과 같은 통신 기술의 발전
- ③ 의료정보 등 공공데이터의 개방 가속화
- ④ 트위터, 페이스북 등 SNS의 급격한 확산

18 구글 번역 서비스에서 활용된 빅데이터의 특성으로 가장 적절한 것은?

- ① Volume
- ② Variety
- ③ Velocity
- ④ Veracity

19 빅데이터 시대가 도래하면서 발생될 수 있는 부정적인 측면의 하나로 "책임원칙 훼손"의 이슈가 있다. 다음 중 이에 대한 사례로 가장 적절한 것은?

- ① 범죄 예측 프로그램에 의해 범행 전에 체포
- ② 빅브라더가 개인의 일상생활 전반을 감시
- ③ 여행 사실을 트윗한 사람의 집에 강도가 침입
- ④ 검색엔진의 검색결과의 차별적 누락에 따른 피해 발생

20 다음 중 빅데이터에 관한 설명으로 가장 적절한 것은?

- ① 빅데이터 분석을 통한 가치창출 여부는 데이터의 규모에 의해 크게 좌우된다.
- ② 비즈니스 핵심에 대해 보다 객관적이고 종합적인 통찰을 줄 수 있는 데이터를 확보해야 한다.
- ③ 빅데이터 프로젝트를 추진하는데 있어 가장 큰 걸림돌은 막대한 소요비용이다.
- ④ 성과가 높은 기업의 대부분은 폭넓은 가치 분석적 통찰력을 갖추고 있는 것으로 밝혀졌다.

※ 숫자, 영문 등의 맞춤법을 준수하여 작성하여 주십시오. (부분점수 없음)

답3 아래는 데이터의 이용과 분석에 대한 시대별 용어와 그것의 의미를 서로 연결한 것이다. ㉠에 들어갈 적절한 용어는?

아 래

O L A P - 다차원의 데이터를 대화식으로 분석하기 위한 소프트웨어
(㉠) - 데이터 기반 의사결정을 지원하기 위한 리포트 중심의 도구
Analytics - 의사결정을 위한 통계적이고 수학적인 분석에 초점을 둔 기법

답4 아래에서 언급한 이것은 무엇인가?

아 래

이것은 컴퓨터공학, 통계학, 수학 등의 학문적 지식은 물론 시각화 및 해커로서의 소양에 이르는 관련 분야의 전문지식을 종합한 학문을 일컫는다. 기존의 통계학과 이것이 다른 점은 총체적 접근법을 사용한 다는 점이다.

21 다음 중 정량적 데이터(Quantitative data)의 예로서 부적절한 것은?

- ① 직업
- ② 강우량
- ③ 영업이익
- ④ GDP 증가율

22 일반적으로 지적 및 지식 기반 자산은 형식지(Explicit knowledge)와 암묵지(Tacit knowledge)로 나누어진다. 다음 중 암묵지의 예로서 부적절한 것은?

- ① 고려청자를 만드는 비법
- ② 간장의 맛을 결정한다는 시어머니의 손맛
- ③ 연봉이 수억에 달한다는 보험설계사의 영업비밀
- ④ 회계규칙에 입각하여 대차대조표를 작성하는데 요구되는 지식

23 다음 중 최대의 시청률을 얻으려면 어떤 프로그램을 어떤 시간대에 방송해야 하는가?라는 문제의 해결을 위해 사용 될 분석방법으로 가장 적절한 것은?

- ① 소셜 네트워크 분석(Social network analysis)
- ② 유전 알고리즘(Genetic algorithms)
- ③ 기계 학습(Machine learning)
- ④ 유형 분석(Classification tree analysis)

24 다음 중 비정형데이터로 분류하기에 가장 부적절한 것은?

- ① IE 웹브라우저로 Gmarket에 접속한 로그기록
- ② Naver의 검색창에 입력한 키워드
- ③ Facebook 친구에게 게시한 댓글
- ④ Galaxy S5로 찍은 셀카 사진

25 다음 중 빅데이터 분석의 특성에 대한 설명으로 가장 부적절한 것은?

- ① 더 많은 정보가 더 많은 가치를 창출할 수 있는 것은 아니다
- ② 비즈니스의 핵심에 대해 보다 객관적이고 종합적인 통찰력을 줄 수 있는 데이터를 찾는 것이 중요하다.
- ③ 빅데이터 과제와 관련된 주된 걸림돌은 비용이 아니다.
- ④ 데이터의 크기가 커질수록 분석을 많이 사용하는 것이 경쟁우위를 가져다주는 원천이 된다.

26 빅데이터 시대에 가치 패러다임이 변화하는 단계를 올바르게 나열한 것은?

- ① 연결(Connection) → 디지털화(Digitalization) → 에이전시(Agency)
- ② 디지털화(Digitalization) → 연결(Connection) → 에이전시(Agency)
- ③ 에이전시(Agency) → 연결(Connection) → 디지털화(Digitalization)
- ④ 연결(Connection) → 에이전시(Agency) → 디지털화(Digitalization)

27 다음 중 데이터베이스의 특징에 대한 설명으로 가장 부적절한 것은?

- ① 데이터베이스는 통합된 데이터(integrated data)이다.
- ② 데이터베이스는 다양한 방법으로 필요한 정보를 검색할 수 있는 검색가능성을 가진다.
- ③ 데이터베이스는 여러 사용자가 서로 단일한 목적으로 데이터를 공동으로 이용할 수 있도록 구성되어야 한다.
- ④ 데이터베이스는 변화되는 데이터로 데이터의 삽입, 삭제, 갱신에도 항상 현재의 정확한 데이터를 유지해야 한다.

28 다음 중 기업내부 데이터베이스에 대한 설명으로 가장 부적절한 것은?

- ① 2000년대 들어서면서 기업 DB구축의 화두는 CRM과 SCM에서 경영정보시스템(MIS)과 ERP로 바뀌었다.
- ② '실시간 기업'은 기업의 비즈니스 프로세스를 투명하고 민첩하게 유지하여 환경 변화에 따른 적응 속도를 최대화하여 지연시간을 없애는 정보화 전략이다.
- ③ 2000년대 들어 금융부문에도 DW를 적극적으로 도입하여 관련 DB마케팅을 증대시키기 위한 노력이 가시화되었다.
- ④ 제조부문은 데이터베이스 기술의 가장 중요한 적용분야로 부품 테이블이나 재고관리 등의 영역에서 데이터베이스 활용이 중점을 이루다가, 부품의 설계, 제조, 유통 전 공정을 포함하는 범위로 확대되었다.

29 아래와 같은 SQL 문장을 사용할 때, 다음 중 출력되는 결과로 옳은 것은?

아 래

```
SELECT CUSTOMER_NAME 고객명, E_CUSTOMER_NAME 고객영문명
FROM CUSTOMER
WHERE E_CUSTOMER_NAME LIKE '_A%';
```

- ① 영문명이 A로 시작하는 고객들의 이름
- ② 영문명의 두 번째 문자가 A인 고객들의 이름
- ③ 영문명이 A나 a로 시작하는 고객들의 이름
- ④ 위치에 상관없이 영문명에 A를 포함하는 고객들의 이름

30 아래 SQL 명령 중 DML(Data Manipulation Language)에 해당하는 항목을 모두 나열한 것은?

아 래

- (A) DELETE
- (B) INSERT
- (C) SELECT
- (D) UPDATE
- (E) CREATE

- ① (A), (B)
- ② (A), (B), (C)
- ③ (A), (B), (C), (D)
- ④ (A), (B), (C), (D), (E)

※ 숫자, 영문 등의 맞춤법을 준수하여 작성하여 주십시오. (부분점수 없음)

단답5 아래 ㉠에 들어갈 말로 적절한 것은?

아 래

빅데이터가 기업에 주는 영향으로는 혁신 촉진, 경쟁력 제고, 그리고 (㉠) 향상을 들 수 있다.

단답6 아래 (㉡) 안에 들어갈 용어를 작성하십시오.

아 래

(㉡)은(는) 데이터베이스의 구조와 제약조건에 관한 전반적인 명세(Specification)를 의미하는 것으로서, 데이터베이스를 구성하는 데이터 개체(Entity), 속성(Attribute), 관계(Relationship) 및 데이터 조작 시 데이터 값들이 갖는 제약 조건 등에 관해 전반적으로 정의한다.

31 아래는 용어와 의미를 서로 연결한 것이다. 다음 중 용어-의미가 잘못 연결된 것을 모두 나열한 것은?

아 래

- | | |
|---------------------------|---------------------------------------|
| (A) OLTP | - 다차원의 데이터를 대화식으로 분석하기 위한 소프트웨어 |
| (B) Business Intelligence | - 경영 의사결정을 위한 통계적이고 수학적인 분석에 초점을 둔 기법 |
| (C) Business Analytics | - 데이터 기반 의사결정을 지원하기 위한 리포트 중심의 도구 |
| (D) Data Mining | - 대용량 데이터로부터 의미 있는 관계, 규칙, 패턴을 찾는 과정 |

- ① (A)
- ② (A), (B)
- ③ (A), (B), (C)
- ④ (A), (B), (C), (D)

32 다음 중 일반적으로 통용되고 있는 빅데이터의 정의와 가장 거리가 먼 것은 무엇인가?

- ① 빅데이터는 일반적인 데이터베이스 소프트웨어로 저장, 관리, 분석할 수 있는 범위를 초과하는 규모의 데이터다.
- ② 빅데이터는 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고 데이터의 초고속 수집·발굴·분석을 지원하도록 고안된 차세대 기술 및 아키텍처다.
- ③ 빅데이터는 데이터의 양(Volume), 데이터 유형과 소스 측면의 다양성(Variety), 데이터 수집과 처리 측면에서 속도(Velocity)가 급격히 증가하면서 나타난 현상이다.
- ④ 빅데이터는 기존의 작은 데이터 처리 분석으로는 얻을 수 없었던 통찰과 가치를 하둡(Hadoop)을 기반으로 하는 대용량 분산처리 기술을 통해 창출하는 새로운 방식이다.

33 다음 중 구글의 Ngram Viewer는 빅데이터의 어떤 기능을 보여주는 사례라고 할 수 있는가?

- ① 현미경
- ② 21세기의 오일(oil)
- ③ 차세대 산업혁명의 에너지원
- ④ 플랫폼

34 다음 중 빅데이터의 위기요인 통제를 위해 논의되고 있는 방안에 대한 설명으로 가장 부적절한 것은?

- ① 개인정보 활용에 대한 동의제를 책임제로 전환하는 것은 책임원칙 훼손 위기에 대한 통제 방안으로 효과적일 수 있으며 개인정보 사용자가 적극적인 보호 장치를 강구하게 하는 효과가 기대된다.
- ② 개인정보의 가치가 증가하고 기술발전으로 사생활침해 가능성도 함께 증가하고 있기 때문에 개인정보 활용에 대한 가이드라인 제정에 대한 요구가 급증하고 있다.
- ③ 데이터 오용의 위기 요소에 대한 대응책으로 알고리즘에 대한 접근권 보장이 필요하며 알고리즘리스트(algorithmist)의 필요성도 증대되고 있다.
- ④ 특정인이 채용이나 대출, 신용카드 발급 여부 결정 등에서 예측 자료에 의해 불이익을 당할 가능성을 최소화하는 장치를 마련하는 것이 필요하다.

35 빅데이터의 시대가 도래함에 따라 CRM(Customer Relationship Management) 환경이 바뀌고 있다. 다음 중 이와 관련된 설명으로 옳지 않은 것은 무엇인가?

- ① 다양한 채널로 CRM 수행이 가능하게 되었다.
- ② CRM의 목적과 방법이 변화하고 있다.
- ③ 기업과 고객 간 양방향 소통이 가능한 서비스 플랫폼이 활성화되고 있다.
- ④ 실시간으로 고객 트렌드를 파악할 수 있게 되었다.

36 다음 중 전략적 통찰력을 얻기 위해 분석을 사용하는 방법으로 가장 부적절한 것을 2개 고르시오.

- ① 경영진의 직관적 결정을 뒷받침하기 위해 분석을 사용한다.
- ② 비즈니스의 핵심가치와 관련된 분석 프레임워크와 평가지표를 개발하여 사용한다.
- ③ 사업 상황을 확인하기 위해 업계 내부의 문제들에 집중하여 분석을 사용한다.
- ④ 분석이 경쟁의 본질을 제대로 바라볼 수 있도록 큰 그림을 그린다.

37 데이터(data)와 정보(information)의 차이를 구분하는 것은 중요하다. 다음 중 정보에 대한 예로 가장 부적절한 것은?

- ① 평균 구매액
- ② 주문수량
- ③ 베스트셀러
- ④ 우량고객

38 다음 중 데이터 분석 테크닉에 대한 설명 중 가장 부적절한 것은?

- ① 개인의 신용도 평가에 가장 많이 활용되는 것은 연관 규칙 학습이다.
- ② 기계 학습은 대규모 데이터를 처리 분석할 때 상당한 분석 인프라와 많은 시간이 소요될 수 있다.
- ③ 한국어의 경우 그 언어적 특성으로 인해 감정 분석에 상대적으로 어려운 측면이 있다.
- ④ 소셜 네트워크 분석은 최근 핀테크 기업에서 대출을 제공할 때 활용되고 있다.

39 다음 중 뛰어난 데이터 사이언티스트에 대한 설명으로 가장 적절한 것은?

- ① 통계학에 대한 이론적 지식과 최적의 분석 설계 노하우를 축적하는 것이 가장 중요하다.
- ② 정량 분석이라는 과학과 인문학적 통찰에 근거한 합리적 추론을 조합한다.
- ③ 분석 결과를 인간의 해석이 개입하지 않도록 객관적으로 사용한다.
- ④ 분석에 현실이 개입하지 않도록 최대한 긍정적인 가정과 모델들에 기초한다.

40 사용자와 데이터베이스 사이에서 사용자의 요구에 따라 정보를 처리해주고 데이터베이스를 관리해주는 소프트웨어는?

- ① DBMS
- ② Data Dictionary
- ③ SQL
- ④ ERD

※ 숫자, 영문 등의 맞춤법을 준수하여 작성하여 주십시오. (부분점수 없음)

답답7 아래에서 언급한 이것은 무엇인가?

아 래

이것은 데이터에 포함된 개인 식별 정보를 삭제하거나 알아볼 수 없는 형태로 변환하는 것을 말한다. 가명(pseudonym), 일반화(generalization), 치환(permutation), 섭동(perturbation) 등을 포함한 다양한 방법으로 이것을 구현한다.

데이터 이해

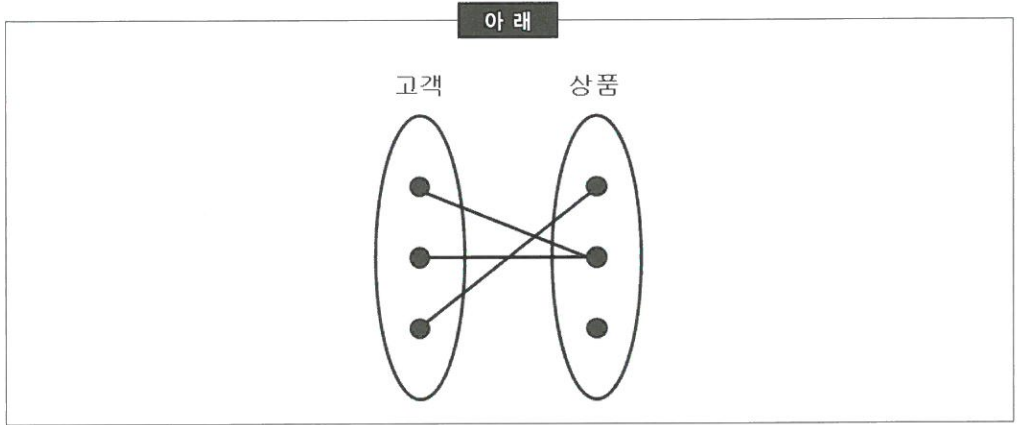
답8

데이터 사이언스는 데이터 공학, 수학, 통계학, 컴퓨터공학, 시각화, 해커의 사고방식, 해당분야의 전문 지식을 종합한 학문으로 정의하기도 한다. 이런 정의의 연장선에서 데이터 사이언스가 기존의 통계학과 다른 접근법을 사용하는데 이는 무엇인가?

통계학 장려함...

41

아래는 고객과 상품의 대응관계를 도식한 것이다. 대응비(Cardinality Ratio) 관점에서 둘 간의 관계를 가장 잘 표현한 것은?



- ① 1 : 1
- ② 1 : N
- ③ N : 1
- ④ N : N

42

다음 중 데이터웨어하우스에 대한 설명으로 가장 부적절한 것은?

- ① 데이터웨어하우스는 기업 내의 의사결정지원 어플리케이션을 위한 정보 기반을 제공하는 하나의 통합된 데이터 저장 공간을 말한다.
- ② ETL은 주기적으로 내부 및 외부 데이터베이스로부터 정보를 추출하고 정해진 규약에 따라 정보를 변환한 후에 데이터웨어하우스에 정보를 적재한다.
- ③ 데이터웨어하우스에서 관리하는 데이터들은 시간의 흐름에 따라 변화하는 값을 유지한다.
- ④ 일반적으로 데이터웨어하우스는 전사적 차원에서 접근하기 보다는 재무, 생산, 운영과 같이 특정 조직의 특정 업무 분야에 초점을 맞추어 구축된다.

43 다음 중 내부 프로세스와 관련된 전형적인 분석 활용기법에 대한 설명으로 가장 부적합한 것은?

- ① 미래가치 분석 - 증거를 관찰하기 전과 후의 가설에 대한 믿음의 정도를 나타내는 추정치로 수익예측에 사용
- ② 최적화 - 일부 변수의 값이 반드시 정수여야 할 때, 한정된 자원으로 특정 목적을 달성하는 솔루션을 내놓은 효과적인 할당을 의미
- ③ 몬테카를로 시뮬레이션 - 특정 결과 혹은 위험이 발생할 확률을 평가하는 계산 기법으로 수학적 모델을 통해 가설 사건을 여러 차례 실험하고 미리 정해진 확률 분포와 비교
- ④ 신경망 분석 - 투입 요소들을 구분할 때까지 반복해서 학습이 이루어지며 주로 대규모 데이터베이스에 사용

44 다음은 비즈니스 문제와 이를 해결하기 위해 주로 사용되는 기법을 연결한 것이다. 잘못 연결된 것은?

- ① 맥주를 사는 사람은 콜라도 같이 구매하는 경우가 많은가? - 연관규칙학습
- ② 고객의 만족도가 충성도에 어떤 영향을 미치는가? - 회귀분석
- ③ 친분관계가 승진에 어떤 영향을 미치는가? - 소셜네트워크분석
- ④ 택배차량을 어떻게 배치하는 것이 가장 비용 효율적인가? - 유형분석

45 아래는 데이터베이스를 기반으로 기업 내에 구축되는 주요 정보시스템 중 하나를 설명한 것이다. 다음 중 아래의 내용에 적합한 것은?

아 래

기업 전체를 경영자원의 효과적 이용이라는 관점에서 통합적으로 관리하고 경영의 효율화를 기하기 위한 시스템

- ① ERP
- ② CRM
- ③ SCM
- ④ KMS

46 최근에 인문학적 사고의 필요성이 대두하게 된 사회경제적 환경의 변화를 기술한 것 중 가장 관련이 없는 것은?

- ① 단순한 세계화에서 복잡한 세계화로 변화되고 있다.
- ② 비즈니스의 중심이 제품생산에서 서비스로 이동하고 있다.
- ③ 경제와 산업의 논리가 생산에서 시장창조로 바뀌고 있다.
- ④ 직관 보다는 데이터에 기반한 의사결정이 더욱 중요해지고 있다.

※ 숫자, 영문 등의 맞춤법을 준수하여 작성하여 주십시오. (부분점수 없음)

단답9 DIKW 피라미드에서 상호 연결된 정보 패턴을 이해하여 이를 토대로 예측한 결과물을 무엇이라고 하는가?

단답10 아래(㉠)에 적합한 데이터베이스 용어는?

아 래

데이터 (㉠)은(는) 데이터베이스 내의 데이터에 대한 정확성, 일관성, 유효성, 신뢰성을 보장하기 위해 데이터 변경 혹은 수정 시 여러 가지 제한을 두어 데이터의 정확성을 보증하는 것을 말한다.

The background of the header is a teal color with various white icons representing data and technology. These include a globe, a bar chart, a line graph, a target, a network diagram with nodes and lines, and a circular arrow. The icons are arranged in a way that suggests a professional or semi-professional data analysis environment.

데이터 분석

Professional · Semi-Professional