

PART 3. 데이터 분석 - 5장. 정형 데이터 마이닝

# 데이터 준전문가

**ADSP, Advanced Data Analytics semi-Professional**

류영표 강사

ryp1662@gmail.com

## 25회 기출문제\_1과목

# 연습문제

1. 아래에서 빅데이터 거버넌스에 대한 설명으로 올바른 것끼리 묶은 것은?

- A. 빅데이터 분석은 다양한 데이터를 활용하기 위하여 회사 내 모든 데이터를 활용해야 한다.
- B. 빅데이터 분석은 고품질의 데이터 확보가 필요하므로 수명주기 관리보다는 품질관리가 중요하다.
- C. ERD는 운영중인 데이터베이스와 일치하기 위하여 철저한 변경관리가 필요하다.
- D. 빅데이터 거버넌스 산업분야별, 데이터 유형별, 정보 거버넌스 요소별로 구분하여 작성한다.

① A, B

② C, D

③ A, B, C

④ B, C, D



# 연습문제

2. 빅데이터 활용 시 필요한 기본 테크닉과 예시를 나열하였다. 이중 테크닉과 예시가 맞지 않은 것은?

- ① 소셜 네트워크 분석 - A 대학교 B학과의 학생들 간 관계망은 어떻게 구성되어 있는가?
- ② 연관분석 - 커피를 사는 사람들은 브라우니를 사 먹는다.
- ③ 유전자 알고리즘 - 최고의 시청률을 얻기 위해 프로그램을 어느 시간대에 방송해야 하는가?
- ④ 회귀 분석 - 월 100만원 이상 사용고객은 어떤 특성을 가진 집단에 속하는가?



3. 다음 중 데이터베이스 설계 절차가 적절하게 배치된 것은?

- ① 요구사항 분석 → 개념적 설계 → 논리적 설계 → 물리적 설계
- ② 요구사항 분석 → 객체적 설계 → 논리적 설계 → 물리적 설계
- ③ 요구사항 분석 → 물리적 설계 → 개념적 설계 → 논리적 설계
- ④ 요구사항 분석 → 개념적 설계 → 객체적 설계 → 논리적 설계



4. 다음 중 딥러닝과 가장 관련 없는 분석 기법은?

- ① CNN
- ② LSTM
- ③ SVM
- ④ Autoencoder



5. 다음 중 빅데이터 시대에 발생할 수 있는 위기 요인 중 사생활 침해 문제를 해결하기 위한 방법으로 가장 적절한 것은?

- ① 알고리즘 접근 허용
- ② 결과기반 책임 원칙 고수
- ③ 데이터 오용 방지
- ④ 정보 사용자 책임제로 변환



# 연습문제

6. 다음은 데이터베이스의 구성요소들을 설명한 것이다. 각 설명에 해당하는 구성요소를 가장 적절하게 나열한 것은?

(A) - 데이터에 관한 구조화된 데이터로, 다른 데이터를 설명해주는 데이터

(B) 데이터베이스 내의 데이터를 신속하게 정렬하고 탐색하게 해주는 구조

① (A) - 메타데이터, (B) - 인덱스

② (A) - 데이터 모델, (B) - 트리거

③ (A) - 백업데이터, (B) - 저장된

④ (A) - 스키마구조, (B) - 데이터 마트





7. 다음 DIKW 단계를 설명하는 것 중 다른 하나는 무엇인가?

- ① 7월 A상품을 구매하는 고객의 60%가 30대 남성 고객이다.
- ② 작년 매출은 2월에서 7월 까지 증가하였고, 10월에 다시 증가했다.
- ③ 날씨가 추워지고, 지점이 늘어나 11월 매출액은 5,000만원으로 예상한다.
- ④ 작년 매출액의 70%는 2월에 집중되어 있다.



# 연습문제

8. 아래의 보기 중 빅데이터의 본질적인 변화가 아닌 것을 고른 것은?

(ㄱ) 빅데이터 시대에는 데이터 획득 비용이 기하급수적으로 감소하고 모든 곳에서 데이터가 넘쳐나 사용자 전수조사가 가능해졌다.

(ㄴ) 다양한 샘플링 기법이 확산되었다.

(ㄷ) 데이터의 수가 증가함에 따라 사소한 몇 개의 오류 데이터가 대세에 영향을 주지 못하는 경향이 증가하고 있다.

(ㄹ) 상관관계보다 인과관계를 중시하게 되었다.

① (ㄱ), (ㄴ)

② (ㄷ), (ㄹ)

③ (ㄱ), (ㄷ)

④ (ㄴ), (ㄹ)



## 31회 기출문제\_1과목

# 연습문제

1. 사물끼리 정보를 주고받는 사물인터넷 시대를 빅데이터의 관점에서 바라볼 때, 다음 중 사물인터넷과 관련이 가장 큰 것은?

- ① 인공지능(AI)
- ② 스마트 데이터(Smart Data)
- ③ 데이터화(Datafication)
- ④ 지능적 서비스(Intelligent Service)



# 연습문제

2. 다음 데이터 분석 조직의 유형 중 별도의 분석 조직이 없고 해당 업무부서에서 분석을 수행하는 방식에 해당하는 것은?

- ① 기능형
- ② 분산형
- ③ 복합형
- ④ 집중형



## 3. 데이터베이스의 일반적인 특징으로 가장 부적절한 것은?

- ① 데이터베이스는 여러 사용자가 서로 다른 목적으로 데이터를 공동으로 이용할 수 있도록 구성되어 있다.
- ② 데이터베이스는 자기 디스크나 자기 테이프 등과 같이 컴퓨터가 접근할 수 있는 저장 매체에 저장된 데이터이다.
- ③ 데이터베이스는 변화하는 데이터로 데이터의 삽입, 삭제, 갱신을 한다고 하더라도 항상 현재의 정확한 데이터를 유지해야 한다.
- ④ 데이터베이스는 한곳에 통합된 데이터(integrated data)이므로 동일한 내용이라도 데이터의 중복을 허용한다.



# 연습문제

4. 데이터에서 가치를 찾아내는 과정을 피라미드의 계층구조로 나타낸다. 다음 예시를 알맞게 설명한 것은 고르시오.

- (a) A마트는 100원에, B마트는 200원에 연필을 판매한다.
- (b) A마트의 연필이 더 싸다.
- (c) 상대적으로 저렴한 A마트에서 연필을 사야겠다.

① (a) : 데이터, (b) : 정보, (c) : 지식

② (a) : 데이터, (b) : 지식, (c) : 지혜

③ (a) : 데이터, (b) : 정보, (c) : 지혜

④ (a) : 정보, (b) : 지식, (c) : 지혜



# 연습문제

5. 일차원적 분석을 통해서도 해당 부서나 업무 영역에서는 상당한 효과를 얻을 수 있다. 다음 중 업무 영역과 분석 사례의 연결이 가장 부적절한 것은?

- ① 마케팅 관리 - 상점과 가게 위치 선정
- ② 재무관리 - 거래처 선정
- ③ 공급체인 관리 - 적정 재고량 결정
- ④ 인력관리 - 이직 인력 예측





# 연습문제

6. 아래에서 빅데이터 시대의 위기와 통제에 대한 설명으로 가장 타당한 것끼리 묶은 것은?

(가) 데이터 익명화(anonymization)는 사생활 침해에 대한 근본요인을 차단할 수 있어 빠른 기술발전이 필요하다.

(나) 빅데이터 분석은 일어난 일에 대한 데이터에 의존하므로 예측 정확도는 높지만 항상 맞을 수 없어 데이터 오용의 피해가 발생할 수 있다.

(다) 개인정보 사용자 정보사용에 대한 무한책임의 한계로 개인정보사용 책임제 보다 동의제를 더욱 강화해야 한다.

(라) 민주주의에서 '행동결과'에 따른 처벌의 모순을 교훈삼아 빅데이터 사전 '성향' 분석을 통한 통제가 강화될 필요가 있다.

(마) 빅데이터가 발생시키는 문제를 중간자 입장에서 중재하며 해결해 주는 알고리즘미스트(algorithmist)도 새로운 직업으로 부상하게 될 것이다.

① (가), (다)

② (나), (다)

③ (가), (라)

④ (나), (마)



7. 다음 중 데이터베이스의 특징과 가장 거리가 먼 것은?

- ① 응용프로그램 종속성
- ② 데이터의 무결성 유지
- ③ 프로그래밍 생산성 향상
- ④ 데이터 중복성 최소화



8. 다음 중 데이터 관리 체계에 대한 설명으로 가장 거리가 먼 것은?

- ① ERD는 운영 중인 데이터베이스와 일치하기 위하여 철저한 변경관리가 필요하다.
- ② 빅데이터 거버넌스는 산업 분야별, 데이터 유형별, 정보 거버넌스 요소별로 구분하여 작성한다.
- ③ 빅데이터는 고품질의 데이터 확보가 필요하므로 데이터 수명주기 관리보다는 품질 관리가 중요하다.
- ④ 데이터 정합성 및 활용의 효율성을 위하여 표준 데이터를 포함한 메타데이터와 데이터 사전의 관리 원칙을 수립해야 한다.



## 32회 기출문제\_1과목

# 연습문제

1. 다음 SQL의 명령어 중 DML이 아닌 것은 무엇인가?

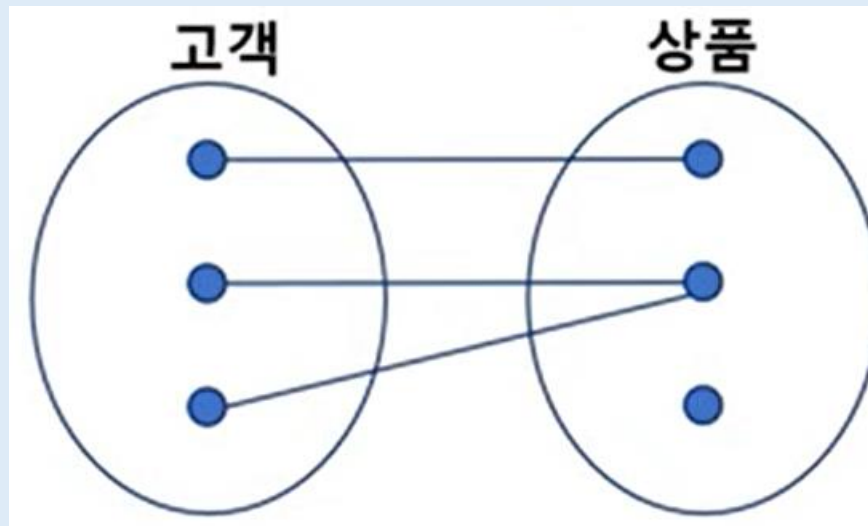
가 - SELECT , 나 - UPDATE, 다 - INSERT, 라-DELETE, 마 - CREATE

- ① 가, 나
- ② 다
- ③ 라
- ④ 마



# 연습문제

2. 고객과 상품의 대응관계를 도식화 한 것이다. 대응비 관점에서 고객과 상품 간의 관계가 옳은 것은?



- ① 1:1
- ② N:1
- ③ N:M
- ④ 1:N



# 연습문제

3. 다음 어떤 기업내부 데이터베이스 솔루션에 대한 설명인가?

제조업을 포함한 다양한 비즈니스 분야에서 생산, 구매, 재고, 주문, 공급자와의 거래, 고객 서비스 제공 등 주요 프로세스 관리를 돕는 여러 모듈로 구성된 통합 애플리케이션

- ① ERP
- ② CRM
- ③ SCM
- ④ KMS



4. 다음 중 빅데이터 활용 요소 3가지로 올바른 것으로만 구성된 것은?

- ① 데이터, 인프라, 기술
- ② 데이터, 인력, 기술
- ③ 인프라, 프로세스, 인력
- ④ 프로세스, 인력, 기술





5. 다음 중 딥러닝의 종류가 아닌 것은?

- ① SVM
- ② DNN
- ③ CNN
- ④ RNN



# 연습문제

6. 지도 학습은 입력과 출력 모두 주어진 상태에서의 학습이며, 비지도학습은 입력만 주어진 상태에서의 학습이다. 다음의 보기에서 그 종류가 다른 것은?

가. 회귀 분석    나. 군집 분석    다. 감정 분석    라. 분류분석

① 가

② 나

③ 다

④ 라



7. 빅데이터와 인문학 열풍과 관련된 내용으로 상관 없는 것은?

- ① 단순 세계에서 복잡한 세계로의 변화
- ② 비즈니스의 중심이 제품생산에서 서비스로 이동
- ③ 빅데이터 분석 기법 및 방법론의 확대
- ④ 경제와 산업의 논리가 생산에서 시장창조로 바뀜.



8. 데이터 사이언티스트의 역량에 대한 설명이 아닌 것은?

- ① 통찰력 있는 분석, Data Technical Skill
- ② 설득력 있는 전달
- ③ 다분여간 협력
- ④ 데이터 분석 알고리즘으로 인해 피해를 입은 사람을 구제함



## 9. 다음은 무엇에 대한 설명인가?

구글의 인터넷에 연결된 네스트는 날씨와 기온 정보 그리고 집주인의 평소 온도 DATA를 기반으로 사용자의 Context를 인식해 자동으로 온도를 설정해 주며, 아마존은 Dash라는 작은 장치는 Wifi가 내장된 바코드 인식기로 상품에 인쇄된 바코드를 Dash를 비추게 되면 그 상품을 아마존 장바구니에 저장할 수 있도록 해준다. 나이키의 경우 애플과 제휴하여 스마트한 운동 관리를 할 수 있도록 해주는 서비스로 자리 매김했다. 굳이 우리가 기계를 조작하지 않아도 모든 것이 사람을 위해 자동으로 돌아가는 세상이 이것이 보여줄 미래이다.



10. 다음은 어떤 기업내부데이터 솔루션에 대한 설명인가?

제조, 물류, 유통업체 등 유통공급망에 참여하는 모든 업체들이 협력을 바탕으로 정보 기술 (Information Technology)를 활용, 재고를 최적화하기 위한 솔루션이다.





# Thank you.

ADSP / 류영표 강사  
ryp1662@gmail.com