

[데이터베이스 구축>데이터 전환>데이터 전환 기술]

[기출 예상 문제]

1. 다음은 무엇에 대한 설명인가?

새로운 정보 시스템을 개발하거나 기존의 시스템을 재 구축한 경우 기존의 시스템에서 사용하던 데이터를 새로운 시스템으로 이행하기 위해 이전 시스템(원천 시스템)에서 데이터를 추출, 변환하여 새로운 시스템(목표 시스템)으로 적재하는 일련의 작업을 말한다.

- ① 데이터 정제 ② 데이터 전환
- ③ 데이터 정규화 ④ 데이터 마이닝

[기출 예상 문제]

2. 데이터 전환 절차로 가장 적절한 것은?

- ① 전환 준비→전환 테스트→전환 설계/ 개발→최종 전환 및 검증→실데이터 전환
- ② 전환 준비→실데이터 전환→전환 테스트→전환 설계/개발→최종 전환 및 검증
- ③ 전환 준비→전환 설계/개발→전환 테스트→실데이터 전환→최종 전환 및 검증
- ④ 전환 준비→전환 테스트→전환 설계/개발→최종 전환 및 검증→실데이터 전환

[기출 예상 문제]

3. 초기 데이터 구축에 대한 설명으로 가장 거리가 먼 것은?

- ① 기존 운영 시스템에 대한 이해를 바탕으로 구축 범위를 명확하게 식별해야 한다.
- ② 발생할 수 있는 예상 문제점을 도출하여 사전에 대응 방안을 마련해야 한다.
- ③ 단시일에 대량의 데이터를 전환함에 있어 발생할 수 있는 위험을 최소화하기 위함이다.
- ④ 사전에 추출 및 변환해서 적재할 수 있도록 인터넷에 저장해 두는 것을 말한다.

[기출 예상 문제]

4. 다음은 무엇에 대한 설명인가?

– 조직 내외부의 복수의 데이터 소스(Source)들로부터 분석을 위한 저장소(Data Warehouse, Data Mart)로 데이터를 이동시키는 프로세스이다.
– 다양한 소스시스템(Source System)으로부터 필요한 데이터를 추출하여 변환 작업을 거쳐 타겟 시스템(Target System)으로 전송 및 로딩하는 모든과정을 말한다.

- ① ETL ② CRUD
- ③ SDLC ④ CDC

[데이터베이스 구축>데이터 전환>데이터 전환 기술]

[기출 예상 문제]

5. ETL(Extraction, Transformation, Loading)에 대한 설명으로 가장 거리가 먼 것은?

- ① ETL 변환 작업은 특수한 엔진에서 진행된다.
- ② 일반적으로 발생하는 데이터 변환에는 필터링, 정렬, 집계 등 등의 작업이 포함된다.
- ③ ETL의 3가지 단계 추출, 변환, 로드는 동시에 진행될 수 없다.
- ④ 변환 중인 데이터가 준비 테이블에서 임시로 보유되었다가 대상에 로드될 수 있다.

[기출 예상 문제]

6. 다음 중 ETL 프로세스별 방법으로 가장 거리가 먼 것은?

- ① 추출(Extraction) - flat file 생성
- ② 추출(Extraction) - CDC
- ③ 변형(Transformation) - Rule 적용
- ④ 적재(Loading) - 데이터필드 검사

[이전 기출 문제]

7. 파일 조직 기법 중 순차 파일에 대한 설명으로 옳지 않은 것은?

- ① 레코드 사이에 빈 공간이 존재하지 않으므로 기억장치 효율적 이용이 가능하다.
- ② 레코드들이 순차적으로 처리되므로 대화식 처리보다 일괄 처리에 적합한 구조이다.
- ③ 필요한 레코드를 삽입, 삭제하는 경우 파일을 재구성해야 하므로 파일 전체를 복사해야 한다.
- ④ 데이터 검색 시 검색 효율이 높다.

[이전 기출 문제]

8. 막대한 양의 자료를 각종 매체에 저장하는 기법을 파일 조직, 파일 편성 혹은 파일 구성 방법이라 한다. 일반적으로 많이 사용되는 파일 조직 방법 중에서 키 값에 따라 순차적으로 정렬된 데이터를 저장하는 데이터 지역(Data Area)과 이 지역에 대한 포인터를 가진 색인 지역(Index Area)으로 구성된 파일은?

- ① 링 파일(Ring File)
- ② 직접 파일(Direct File)
- ③ 순차 파일(Sequential File)
- ④ 색인 순차 파일(Indexed Sequential File)

[데이터베이스 구축>데이터 전환>데이터 전환 기술]

[이전 기출 문제]

9. VSAM 파일에 대한 설명으로 거리가 먼 것은?

- ① 기본 데이터 영역과 오버플로우 영역을 구분하지 않는다.
- ② 레코드를 삭제하면 그 공간을 재사용 할 수 있다.
- ③ 제어 구간에 가변 길이 레코드를 쉽게 수용할 수 있다.
- ④ 해시 함수를 사용하여 레코드를 저장할 위치를 결정한다.

[이전 기출 문제]

10. 파일에 대한 설명 중 옳지 않은 것은?

- ① 순차 파일(Sequential file)을 생성되는 순서에 따라 레코드를 순차적으로 저장하므로 저장 매체의 효율이 가장 높다.
- ② 직접파일(Direct file)은 특정 레코드에 접근하기 위해서 디스크의 물리적주소로 변환할 수 있는 함수를 사용한다.
- ③ 색인 순차 파일(Indexed sequential file)은 순차 및 직접 접근 형태를 모두 지원할 수 있으나 기억장소의 낭비를 초래한다.
- ④ VSAM 파일(Virtual Storage Access method file)은 검색 속도를 빠르게 하기 위하여, 기본 데이터구역과 오버플로우 구역을 구분하여 갖추어야 한다.

[기출 예상 문제]

11. 직접 접근 방식(DAM: Directed Access Method)에 설명으로 거리가 먼 것은?

- ① 어떤 레코드라도 평균 접근 시간(Access Time) 내에 검색이 가능하다.
- ② 기억공간의 효율이 좋다.
- ③ 접근 시간이 빠르고 레코드의 삽입, 삭제, 갱신이 용이하다.
- ④ 임의 접근이 가능한 자기 디스크나 자기 드럼을 사용한다.

[데이터베이스 구축>데이터 전환>데이터 전환 수행]

[기출 예상 문제]

1. 데이터 전환 절차에 따른 주요 목표로 가장 적절한 것은?

- ① 데이터전환 설계 - 전환 및 검증 프로그램 개발
- ② 데이터 전환 테스트 및 검증 - 반복 테스트를 통한 데이터 정합성 확보
- ③ 데이터 전환 - Target/Source 매핑 및 검증
- ④ 데이터 전환 계획 및 요건 정의 - 성공적인 데이터 전환 수행

[기출 예상 문제]

2. 다음 중 데이터 전환 설계 단계에서 수행하는 과정으로 가장 거리가 먼 것은?

- ① 전환 개발환경 구축
- ② 코드 매핑(Code Mapping)
- ③ 검증 규칙(Rule) 정의
- ④ 전환 설계서 작성

[기출 예상 문제]

3. 다음은 데이터 전환 절차 중 어느 단계에 대한 설명인가?

데이터 전환 수행을 위해서는 우선 데이터 전환 프로젝트의 계획을 수립하고 현행 정보시스템(As-Is) 및 원천 데이터의 구조를 분석해야 한다. 현행 정보시스템의 데이터 구조 분석은 IT 시스템을 활용하는 각 업무 영역들에 대한 분석으로 볼 수 있다. 현행 정보시스템의 분석이 끝나면 목표 시스템(To-Be)에 대한 분석을 진행한다.

- ① 데이터 전환 설계 단계
- ② 데이터 전환 개발 단계
- ③ 데이터 전환 계획 및 요건정의 단계
- ④ 데이터 전환 단계

[기출 예상 문제]

4. 다음 중 데이터 전환 계획서 항목으로 가장 거리가 먼 것은?

- ① 데이터 복구 대책
- ② 데이터 전환 환경 구성
- ③ 데이터 정비 방안
- ④ 데이터 보관 계정 정보

[데이터베이스 구축>데이터 전환>데이터 전환 수행]

[기출 예상 문제]

5. 데이터 전환 계획서 작성을 위한 데이터 전환 수행 절차에 대한 설명으로 가장 거리가 먼 것은?

- ① 데이터 전환 범위, 일정, 전환 절차를 명확하게 해야 한다.
- ② 데이터 전환 제외 대상은 기록하지 않아도 된다.
- ③ 팀별 업무를 분장하여 팀별 역할과 납기를 기록한다.
- ④ 데이터 전환이 완료된 후 원천 데이터베이스와 비교하여 데이터 오류를 확인한다.

[기출 예상 문제]

6. 다음 중 데이터 전환 계획서 작성 단계별 수행 방안에 대한 설명으로 가장 거리가 먼 것은?

- ① 1단계 전환 계획 - 단위 시스템별 데이터 이행 및 검증을 위해서 수행한다.
- ② 2단계 전환 계획 - 전체 시스템 데이터 전환 시나리오 초안 작성 및 검증을 위해 수행한다.
- ③ 최종 단계 전환 계획 - 최종 데이터 전환 모의 계획 작성 및 검증을 위해 수행한다.
- ④ 2단계 전환 계획 - 데이터 전환 시간 측정 및 순서를 확정하고, 이에 따른 최종 데이터 전환 모의 계획시나리오를 작성한다.

[기출 예상 문제]

7. 아래의 예시와 같이 데이터 전환 수행자가 수행할 작업의 상세 항목, 작업 내용, 예정 시작·종료 시각, 작업 담당자 항목을 포함한 문서는 무엇인가?

단계	작업내용	작업자	수행실적	작업 예정 시간	
				시작시간	종료시간
사전준비	1. 운영 환경에 대한 설정 및 점검				
	2. 전환 환경에 대한 사전 점검				
	3. DB 상태 점검 수행(가동·접속 여부, 스페이스 확인)				
데이터 전환	4. 수작업 테이블 이관 <총 000개>				
	5. 데이터 추출, 변환, 적재, 수행 <총 000개>				
	6. 인덱스 리빌드, 권한 재설정				
데이터 점검	7. 전환 검증 요건 항목 검증 <000개 항목 검사 수행>				
	8. 업무계 0개 테이블에 대한 후속 SQL 작업(응용 팀)				

- ① 검증 데이터 확인서
- ② 데이터 체크리스트
- ③ 데이터 전환 계획서
- ④ 데이터 모델 품질 검토서

[기출 예상 문제]

8. 다음 중 데이터 검증 방법이 아닌 것은?

- ① 추출 검증
- ② 전송 검증
- ③ 전환 검증
- ④ 동적 검증

[데이터베이스 구축>데이터 전환>데이터 전환 수행]

[기출 예상 문제]

9. 데이터 검증 방법 중 매핑 후 검증을 의미하는 방법으로, 오류 데이터를 분석하여 사전 정비 또는 전환 규칙 오류를 수정 반영하는 방법으로 가장 적절한 것은?

- ① 통합 검증 ② 적재 검증
- ③ 전송 검증 ④ 전환 검증

[데이터베이스 구축>데이터 전환>데이터 정제]

[기출 예상 문제]

1. 데이터 품질 분석에 대한 설명으로 가장 거리가 먼 것은?

- ① 데이터 품질 관리란 사용자의 기대를 만족시키기 위해 지속적으로 수행하는 데이터 관리 및 개선 활동이다.
- ② 원천 데이터의 품질이 보장되지 않으면 전환 후의 데이터도 정확하지 않다.
- ③ 원천 데이터의 품질을 검증함으로써 전환의 정확성을 보장할 수 있다.
- ④ 원천 데이터와 목적 데이터베이스의 속성(Column) 간 대응 관계는 N : M 대응보다 1 : 1 대응 관계가 더 많다.

[기출 예상 문제]

2. 전환 데이터의 대표적인 정합성 검증 항목으로 가장 거리가 먼 것은?

- ① 보고서 항목 또는 통계 수치
- ② 특정 기준으로 분류된 좌수 및 잔액
- ③ 검증팀의 역량 및 특이사항
- ④ 과목별 좌수 및 잔액

[기출 예상 문제]

3. 원천 데이터 품질 검증을 위한 대표적인 정합성 항목으로 가장 거리가 먼 것은?

- ① 데이터베이스 정규화가 이루어져 있는가?
- ② 코드값이 정해진 범위에 속하는가?
- ③ 보고서값과 실제 데이터값이 일치하는가?
- ④ 필수 항목의 데이터가 모두 존재하는가?

[기출 예상 문제]

4. 오류 데이터 측정에 대한 설명으로 가장 거리가 먼 것은?

- ① 데이터 품질 기준에 따라 정상 데이터와 오류 데이터를 분리한다.
- ② 정상 데이터의 건수를 업무 영역별, 테이블별로 세분화 하여 정상 데이터 수량을 정확히 측정하고 기록한다.
- ③ 해결 불가능한 오류는 오류 관리 목록에 기록하지 않는다.
- ④ 원천 데이터의 정합성 기준을 근거로 오류 데이터 유형과 위치를 파악한다.

[정답] [데이터베이스 구축>데이터 전환]

1. 데이터 전환 기술

1	2	3	4	5	6	7	8	9	10	11									
2	3	4	1	3	4	4	4	4	4	2									

2. 데이터 전환 수행

[illegible]

3. 데이터 정제

[illegible]

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 전환 기술]

데이터 전환 기술 1. ②

데이터 전환(데이터 이행)에 대한 설명이다.

데이터 전환 기술 2. ③

데이터 전환 절차

: 전환 준비 → 전환 설계/개발 → 전환 테스트 → 실행
데이터 전환 → 최종 전환 및 검증

데이터 전환 기술 3. ④

④ 원천시스템에서 특정 데이터들을 사전에 추출 및 변환해서 적재할 수 있도록 데이터베이스 혹은 데이터 파일로 구축해 두는 것을 말한다.

데이터 전환 기술 4. ①

ETL(Extraction, Transformation, Loading)에 대한 설명이다.

데이터 전환 기술 5. ③

③ 시간 절약을 위해 3가지 ETL 단계가 동시에 실행된다.

데이터 전환 기술 6. ④

ETL 프로세스별 방법

- 추출(E): JDBC, ODBC 기술의 이용, flat file 생성, CDC(Change Data Capture) 등
- 변형(T): 재구성, 정제(중복 제거, 일관성 확보), 변환(타겟 저장소에 적합한 형태), 데이터필드 검사, Rule 적용
- 적재(L): DBMS 고유 기능 이용, Utility(append, delete/insert, update) 이용

데이터 전환 기술 7. ④

④ 데이터 검색 시 처음부터 순차적으로 검색하기 때문에 검색 효율이 낮다.

데이터 전환 기술 8. ④

색인 순차 파일(Indexed Sequential File)에 대한 설명이다.

데이터 전환 기술 9. ④

④ 직접 파일(Direct File)에 대한 설명이다.

데이터 전환 기술 10. ④

④ ISAM(Indexed Sequential File)에 대한 설명이다.

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 전환 기술]

데이터 전환 기술 11. ②

② 기억공간의 효율이 저하될 수 있다.

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 전환 수행]

데이터 전환 수행 1. ②

데이터 전환 절차의 주요 목표

- 데이터 전환 계획 및 요건 정의: 원천데이터 구조에 대한 분석
- 데이터 전환 설계: Target/Source 매핑 및 검증
- 데이터 전환 개발: 전환 및 검증 프로그램 개발
- 데이터 전환 테스트 및 검증: 반복 테스트를 통한 데이터 정합성 확보
- 데이터 전환: 성공적인 데이터 전환 수행

데이터 전환 수행 2. ①

- ① 개발 표준과 개발방법론에 맞게 데이터 전환 개발 환경을 구축하는 과정은 데이터 전환 개발 단계이다.
- 데이터 전환 설계 단계: 로지컬 매핑(Logical Mapping), 코드 매핑(Code Mapping), 검증 규칙(Rule) 정의, 전환 설계서 작성

데이터 전환 수행 3. ③

- 데이터 전환 계획 및 요건정의 단계에 대한 설명이다.
- 데이터 전환 계획 및 요건정의 단계: 프로젝트 계획, 프로젝트 환경 구축, 현행 시스템(As-Is) 분석, 목표 시스템(To-Be) 분석, 데이터 전환 요건 정의

데이터 전환 수행 4. ④

데이터 전환 계획서 항목

- 데이터 전환 개요 (데이터 전환 목표 등)
- 데이터 전환 대상 및 범위 (테이블 수, 테이블 크기 등)
- 데이터 전환 환경 구성 (원천/목적 시스템 구성도, 디스크 용량 산정 등)
- 데이터 전환 조직 및 역할
- 데이터 전환 일정
- 데이터 전환 방안 (데이터 전환 계획/규칙/절차/방법 등)
- 데이터 정비 방안 (데이터 정비 대상 등)
- 비상 계획 (비상시 대처 방안 수립, 의사소통 체계 등)
- 데이터 복구 대책

데이터 전환 수행 5. ②

- ② 데이터 전환 범위 선정 시 데이터 전환 대상을 기록한다.

데이터 전환 수행 6. ④

- ④ 최종 단계 전환 계획에 대한 설명이다.

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 전환 수행]

데이터 전환 수행 7. ②

데이터 체크리스트에 대한 설명이다.

데이터 전환 수행 8. ④

데이터 검증 방법

- 추출 검증
- 전송 검증
- 전환 검증
- 적재 검증
- 통합 검증

데이터 전환 수행 9. ①

통합 검증에 대한 설명이다.

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 정제]

데이터 정제 1. ④

④ 원천 데이터와 목적 데이터베이스의 속성(Column) 간 대응 관계는 1:1 대응보다 N:M 대응 관계가 더 많다.

데이터 정제 2. ③

전환 데이터의 대표적인 정합성 검증 항목

- 과목별 좌수 및 잔액
- 특정 기준으로 분류된 좌수 및 잔액
- 보고서 항목 또는 통계 수치
- 계좌 및 고객을 Sampling하여 해당 사항을 모두 검증
- 특수한 관계가 있는 고객을 추출하여 관련된 데이터가 정확한지 검증

데이터 정제 3. ①

원천 데이터의 품질 검증을 위한 대표적인 정합성 항목

- 필수 항목의 데이터가 모두 존재하는가?
- 데이터 유형(문자형, 숫자형, 날짜형)이 정확하게 관리되고 있는가?
- 날짜의 경우 날짜로서 유효한 형태를 가지고 있는가?
- 금액의 경우 유효한 값의 범위인가?
- 계좌 번호 등이 정해진 법칙에 따른 번호인가?
- 영업점이 실제로 존재하는 영업점인가?
- 코드값이 정해진 범위에 속하는가?
- 개설 일자, 해지 일자 등 모든 일자의 시점이 업무 규칙에 위배되지 않고 정확하게 설정되어 있는가?
- 특히, 선후 관계가 정확하게 표현되어 있는가?
- 업무 규칙에 위배되는 잘못된 정보가 존재하는가?
- 잔액의 총합이 회계 정보와 동일한가?
- 보고서값과 실제 데이터값이 일치하는가?

데이터 정제 4. ③

③ 해결 불가능한 오류는 고객과 협의한 내용을 기술한다.

[정답 및 해설] [데이터베이스 구축>데이터 전환>데이터 정제]

데이터 정제 5. ②

- ② Classified: 보고된 오류가 프로젝트 팀에 의해 오류가 분류된 상태
- Fixed: 개발자가 오류를 수정한 상태

데이터 정제 6. ①

- ① 데이터 정제 항목을 정제 시점에 따라 전환 테스트 전, 전환 테스트 중, 최종 전환 중 3단계로 구분하여 데이터 정제 작업을 수행한다.

데이터 정제 7. ③

정제 유형

- 완전성: 고객 정보에 주민 등록 번호/사업자 등록 번호가 없는 경우처럼 업무적으로 반드시 있어야 하는 자료가 누락된 경우
- 유효성: 생년월일이 현재 일자보다 큰 경우처럼 항목의 값이 유효하지 않은 경우
- 일치성: 성별이 남자이면서 주민 등록 번호 뒤 7자리가 2 또는 4로 시작되는 경우와 같이 상호 관련이 있는 자료 항목이 서로 상이한 경우
- 유일성: 서로 다른 사람의 주민 등록 번호가 동일하다면 정제 대상