

# 제7회 빅데이터분석기사 실기 응시 가이드

## 1. 시험개요

### 1-1. 일시 및 장소

구 분	주요 내용
시 험 명	제7회 빅데이터분석기사 실기
시험장소	dataq.or.kr에서 수험표를 출력하여 확인 ☞ 모든 고사장 주차 지원 불가
시험일시	2023. 12. 2(토) 입실 09:00~09:30 / 시험 10:00~13:00 · 10:30 이후 화장실 이용 가능 · 11:30 이전 퇴실 ☞ 0점 처리 · 11:30 이후 퇴실 ☞ 정상 채점

### 1-2. 유의사항

- 시험 당일, **09:30까지 고사실에 입실**하여 안내 사항 숙지 및 **시험환경 테스트**
- 시험 당일, 본인확인을 위해 **인정 범위 신분증 지참 필수**
- 시험 중, **전자·통신기기(계산기 등) 및 기타 프로그램(윈도우 메모장, 윈도우 계산기 등) 사용 불가**

**\* 윈도우 기본 메모장 사용 불가 ⇒ 시험 시작 후 시험환경 내에서 제공되는 별도 메모장 사용**

- 시험 중, 허가되지 않은 사이트(예:구글, 네이버 등 시험환경 외 모든 사이트)에 접속하거나, 이를 이용하여 답안을 작성하면 **부정행위 처리**(3년간 국가기술자격 시험 응시자격 정지)
- 시험 중, 휴대한 전자·통신기기에서 진동, 소음 등 발생하면 **시험 중지 및 무효**(0점)
- 그 밖의 유의사항 및 관련 상세 내용은 수험표 및 **dataq.or.kr 공지사항** 참조

## 2. 답안제출 및 채점기준

구 분	작업형 제1유형	작업형 제2유형	작업형 제3유형	합 계
문항수	3문항, 문항당 10점	1문항, 40점	2문항, 문항당 15점 (소문항 구성, 소문항 배점 합산)	6문항
점 수	30점	40점	30점	100점
답안 제출	코딩화면에서 문제풀이 후 답안제출 화면으로 이동하여 입력·제출	답안 CSV 파일을 생성하여 제출	코딩화면에서 문제풀이 후 답안제출 화면으로 이동하여 입력·제출	-

## 2-1. 작업형 제1유형

- 각 문항별로 코딩 화면에서 문제를 풀이한 후, 답안은 별도의 답안제출 화면으로 이동하여 제출하며, 지시된 제출 형식 준수

### ○ 채점 기준

- 각 문항별로 정답 여부에 따라 배점 기준 만점 또는 0점
- 작성 코드에 대한 부분 점수 없음

#### 【제1유형 0점 유형】

❶ 제출 형식 위반 : 0점

❷ 복수 답안 제출 : 0점

<점수 예시>

\* 제출형식 : 반올림하여 소숫점 둘째 자리까지 작성(산출된 값 : 0.117 / 정답 : 0.12)

제출답안	0.12	0.120	0.117	0.12, 0.117	0.11	0.1
기준	정답	❶ 제출 형식 위반		❷ 복수 답안 제출	오답	
획득점수	배점 기준 만점	0점		0점	0점	

## 2-2. 작업형 제2유형

- 평가용 데이터를 이용한 예측 결과를 지시된 제출 형식 준수하여 CSV 파일로 제출

- 예측 결과는 지시된 칼럼명을 사용하여 생성
- 자동 생성되는 index 칼럼 제거
- 답안 파일에는 예측 결과 칼럼 1개만 생성
- 답안 파일은 지시된 파일명을 사용하여 생성
- 답안 파일 별도 디렉토리 지정 금지

### ○ 채점 기준

- 평가지표에 따라 구간별 점수 부여
- 작성 코드에 대한 부분 점수 없음
- 평가지표에 따른 구간 점수를 획득하여도 제출 형식을 위반하면 득점 점수에서 감점하며, 감점 유형이 중복되면 누적하여 감점

#### 【제2유형 0점 유형】

❶ 예측값 정확도가 평가지표 최저 구간 미만 : 0점

❷ 평가용 데이터 개수와 예측 결과 데이터 개수 불일치로 평가지표 산출 불가 : 0점

**【제2유형 감점 유형】**

**① 지시 파일명 미사용 : 3점 감점**

- 지시된 파일명을 사용하지 않거나, 파일명에 '.csv'를 누락, 또는 파일명 내 디렉토리를 지정하면 감점

**② 지시 칼럼명 미사용 : 3점 감점**

- 예측 결과 칼럼명을 지시와 다르게 지정하면 감점

**③ 제출 칼럼 개수 1개 초과 : 3점 감점**

- 예측 결과 칼럼 외, 자동 생성되는 index 칼럼이나 다른 칼럼을 함께 제출하면 감점

<평가지표 40점 구간의 점수 예시>

\* 파일명 : result.csv / 칼럼명 : pred

파일명	000000.csv	result.csv	result.csv	result.csv	000000.csv
제출칼럼명	pred	predict	pred + 다른 칼럼	predict+ 다른 칼럼	predict+ 다른 칼럼
기준	① 지시 파일명 미사용	② 지시 칼럼명 미사용	③ 제출 칼럼 1개 초과	② 지시 칼럼명 미사용 ③ 제출 칼럼 1개 초과	① 지시 파일명 미사용 ② 지시 칼럼명 미사용 ③ 제출 칼럼 1개 초과
획득점수	37점	37점	37점	34점	31점

**2-3. 작업형 제3유형**

- 각 문항별로 코딩 화면에서 문제를 풀이한 후, 답안은 별도의 답안제출 화면으로 이동하여 각 문항별 소문항의 순서대로 제출하며, 지시된 제출 형식 준수

○ 채점 기준

- 각 문항의 소문항별로 정답 여부에 따라 배점 기준 만점 또는 0점
- 작성 코드에 대한 부분 점수 없음

**【제3유형 소문항별 0점 유형】**

**① 소문항별 제출 형식 위반 : 0점**

**② 소문항별 복수 답안 제출 : 0점**

<점수 예시>

\* 제출형식 : 반올림하여 소숫점 둘째 자리까지 작성(산출된 값 : 0.117 / 정답 : 0.12)

제출답안	0.12	0.120	0.117	0.12, 0.117	0.11	0.1
기준	정답	① 제출 형식 위반	② 복수 답안 제출	오답		
획득점수	배점 기준 만점	0점	0점	0점		

### 3. 시험환경

- 클라우드 기반 CBT(computer based test), 크롬(chrome) 브라우저 사용
- 문항별로 R 또는 Python 중 언어 선택 가능
- 제약사항
  - 코드 라인별 실행 불가
  - 그래프 기능, 단축키, 자동완성 기능 미제공
  - 코드 실행 시간은 1분으로 제한되며, 시간 초과 시 강제 실행 취소
  - 패키지는 시험환경에서 제공되는 것만 이용할 수 있으며, 시험 중 추가 설치 불가
  - 시험환경에 설치된 패키지는 코딩화면에 아래 코드를 실행하여 확인

#### 【패키지 확인 방법】

Python 패키지 확인 코드	<pre>import pkg_resources import pandas OutputDataSet = pandas.DataFrame(sorted(     [(i.key, i.version) for i in pkg_resources.working_set])) print(OutputDataSet)</pre>
R 패키지 확인 코드	<pre>as.data.frame(installed.packages()[,c(3:4)])</pre>

### 4. 시험환경 점검

#### 4-1. 사전 점검

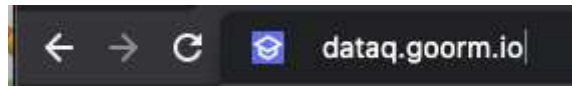
- 시험 당일 고사실 입실은 09:00부터 가능하며, 09:30까지 입실 완료하여 PC 및 시험사이트 등 시험환경 사전 점검

#### 【점검 항목】

- ① 크롬(chrome) 브라우저 설치 유무
- ② 검정용 PC 모니터, 키보드 및 마우스 작동
- ③ 시험 사이트 접속(<https://dataq.goorm.io>)
- ④ 작업형 제1유형 코드 실행 및 답안 제출 테스트
- ⑤ 작업형 제2유형 코드 실행 및 답안 제출 테스트
- ⑥ 작업형 제3유형 코드 실행 및 답안 제출 테스트

## 4-2. 점검 사이트 접속

- URL : 크롬(chrome) 브라우저를 이용하여 <https://dataq.goorm.io> 접속



- 로그인 : ID 수험번호 / PW 시험 당일 안내

## 4-3. 문의사항

- 시험사이트 우측 하단의 채팅 상담을 통해 문의
- PC(본체, 모니터, 키보드, 마우스 등) 장애는 고사실 감독관에게 문의

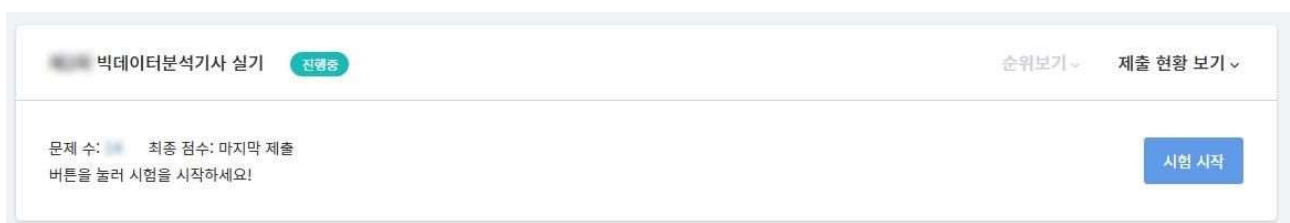


## 5. 시험응시

- 시험 페이지 입장 : 시험 목록에서 “빅데이터분석기사 실기시험” 클릭



- 시험 시작 : “시험 시작” 버튼 클릭
  - 10:00 시험이 시작되면 “시험 시작” 버튼 활성화
  - “시험 시작” 버튼을 클릭하면 수험자 유의사항이 표시되며, 문제 페이지로 이동



○ **작업형 시험화면 ① : 작업형 제1/2/3유형 공통**

- 문제영역 : 문제 풀이에 필요한 정보 확인
- 코딩영역 : 문제 풀이를 위한 코드 작성
- 실행결과영역 : 실행결과 및 제출결과 확인
- 시험끝내기 : 시험을 종료하며, 시험 종료 후에는 답안 제출 불가

**시험끝내기**

**문제영역**

(채점) 제2유형

\* 예측 결과는 ROC-AUC 평가지표에 따라 평가함

\* 성능이 우수한 예측 모델을 구축하기 위해서는 데이터 정제, Feature Engineering, 하이퍼 파라미터(hyper parameter) 최적화, 모델 비교 등이 필요할 수 있음. 다만, 과적합에 유의하여야 함.

**【제출 형식】**

① CSV 파일명 : result.csv (파일명에 디렉토리-폴더 지정불가)

② 예측 상별 칼럼명 : pred

③ 제출 칼럼 개수 : pred 칼럼 1개

④ 평가용 데이터 개수와 예측 결과 데이터 개수 일치 :  개

**제공 데이터**

■ 데이터 목록

①

②

☞ 평가용 데이터는 '성별' 칼럼 미제공

■ 데이터 설명

칼럼명	칼럼 설명
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>

**코딩영역**

```

1 # 출력을 원하실 경우 print() 함수 활용
2 # 예시) print(df.head())
3
4 # getcwd(), chdir() 등 작업 폴더 설정 불필요
5 # 파일 경로 상 내부 드라이브 경로(C: 등) 접근 불가
6
7 # 데이터 파일 읽기 예제
8 import pandas as pd
9 train = pd.read_csv("data/customer_train.csv")
10 test = pd.read_csv("data/customer_test.csv")
11
12 # 사용자 코딩
13 print(test)
14
15 # 답안 제출 참고
16 # 아래 코드 예측변수와 수험번호를 개인별로 변경하여 활용
17 # pd.DataFrame({'cust_id': X_test.cust_id, 'gender': pred}).to_csv('003000000.csv', index=False)
  
```

**실행결과영역**

실행결과 | 제출결과 | 테스트케이스

평가

○ **작업형 시험화면 ② : 작업형 제1/2/3유형 공통**

- **데이터셋 확인** : 제공 데이터 내용 확인
- **언어선택** : R 또는 Python 선택
- **초기화** : 작성 중인 코드를 모두 지우고 기본 코드로 돌아감
- **저장** : 작성 중인 코드를 저장. 소스코드는 "저장"을 해야 새로고침하거나 다시 접속하더라도 화면에 남아있으므로, 주기적으로 저장하면서 문제 풀기를 권장
- **실행** : 작성한 코드의 결과나 오류 메시지는 아래 "실행결과" 영역에 표시  
코드 실행시간은 1분으로 제한되며, 초과시 실행 강제 취소
- **제출** : 답안은 여러 번 제출 가능하며, 마지막 제출된 답안이 채점 대상

**\* 제1유형, 제3유형은 "제출" 버튼이 없고, 별도의 답안제출 화면으로 이동하여 제출**

**문제정보**

100%

(제형) 제2유형

\* 예측 결과는 ROC-AUC 평가지표에 따라 평가함

\* 성능이 우수한 예측 모델을 구축하기 위해서는 데이터 정제, Feature Engineering, 하이퍼 파라미터(hyper parameter) 최적화, 모델 비교 등이 필요할 수 있음. 다만, 과적합에 유의하여야 함.

**【제출 형식】**

① CSV 파일명 : result.csv (파일명에 디렉토리-폴더 지정불가)

② 예측 성별 칼럼명 : pred

③ 제출 칼럼 개수 : pred 칼럼 1개

④ 평가용 데이터 개수와 예측 결과 데이터 개수 일치 : [ ]개

**제공 데이터**

• 데이터 목록

① [ ]

② [ ]

평가용 데이터는 '성별' 칼럼 미제공

• 데이터 설명

칼럼명	칼럼 설명
[ ]	[ ]
[ ]	[ ]

**Python3** data/customer\_train.csv data/c Python3 초기화 저장 실행 테스트케이스 추가 테스트 제출

```

1 # 알렉을 전처리 করে print() 함수 출력
2 # 예시) print('alex')
3
4 # getcwd(), chdir() 등 작업 폴더 설정 불필요
5 # 파일 경로 상 내부 드라이브 경로(C: 등) 접근 불가
6
7 # 데이터 파일 읽기 예제
8 import pandas as pd
9 train = pd.read_csv("data/customer_train.csv")
10 test = pd.read_csv("data/customer_test.csv")
11
12 # 사용자 코딩
13 print(test)
14
15 # 답안 제출 참고
16 # 아래 코드 예측변수와 수평번호를 개인별로 변경하여 활용
17 # pd.DataFrame({'cust_id': X_test.cust_id, 'gender': pred}).to_csv('003000000.csv', index=False)
    
```

실행결과 제출결과 테스트케이스

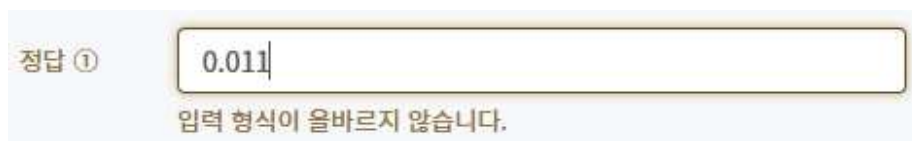
**채팅상담**

### ○ 작업형 시험화면 ③ : 작업형 제1/3유형 해당

- 작업형 제1유형/제3유형은 문제 코딩 화면과 답안 제출화면이 분리
- (제1유형) 각 문항의 코딩 화면에서 문제를 풀이한 후 별도의 답안제출 화면으로 이동하여, 지시된 제출 형식에 따라 답안 제출
- (제3유형) 각 문항의 코딩 화면에서 문제를 풀이한 후 별도의 답안제출 화면으로 이동하여, 각 문항별 소문항의 순서대로 지시된 제출 형식에 따라 답안 제출
- 제출 답안이 지시문을 위반하면, “입력 형식이 올바르지 않습니다” 안내 표시

<입력 형식 안내 예>

- \* 제출형식 : 반올림하여 소숫점 둘째 자리까지 작성
- \* 입력답안 : 0.011



정답 ① 0.011

입력 형식이 올바르지 않습니다.

- “제출”은 여러 번 가능하며, 마지막 제출된 답안이 채점 대상

(제1형) 제1유형 답안제출



**답안 입력 화면**

(제3형) 제3유형 답안제출

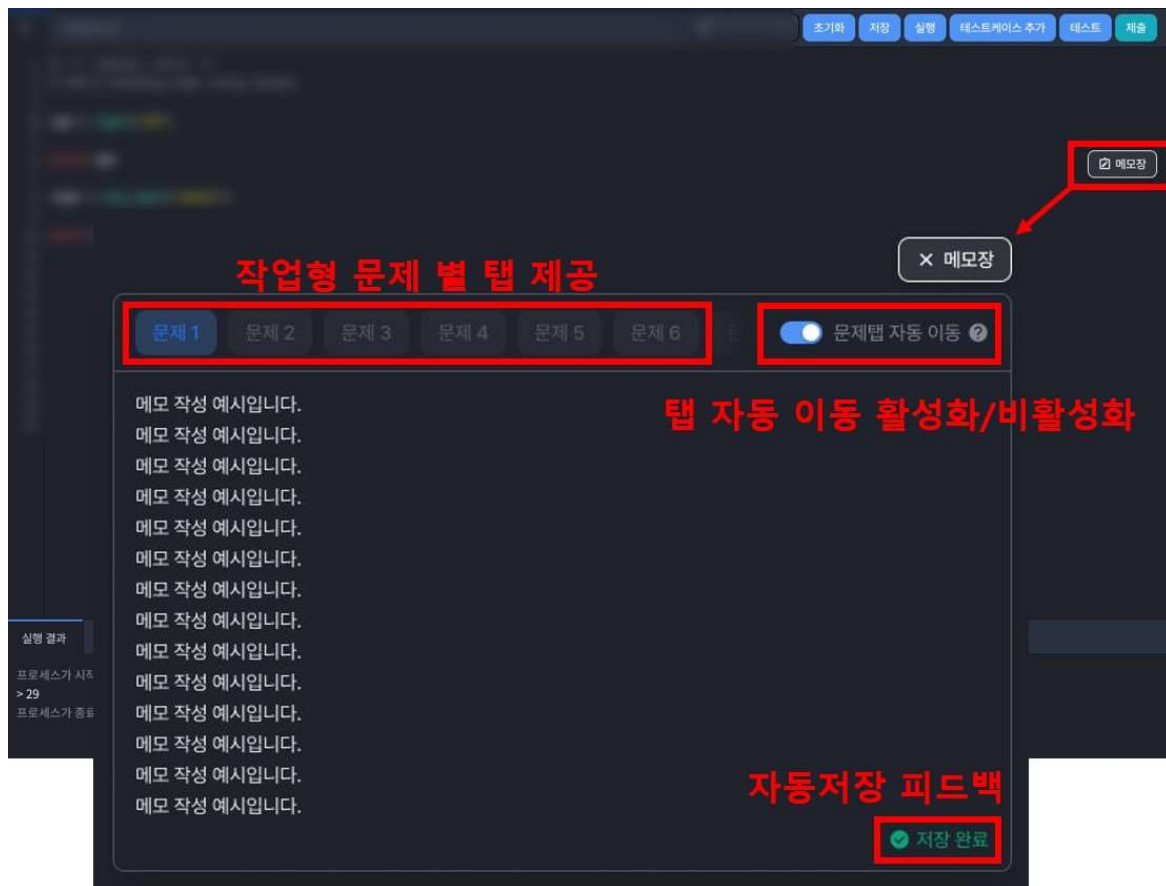


**제출 버튼**



○ **작업형 시험화면 ④ : 메모장 기능\***

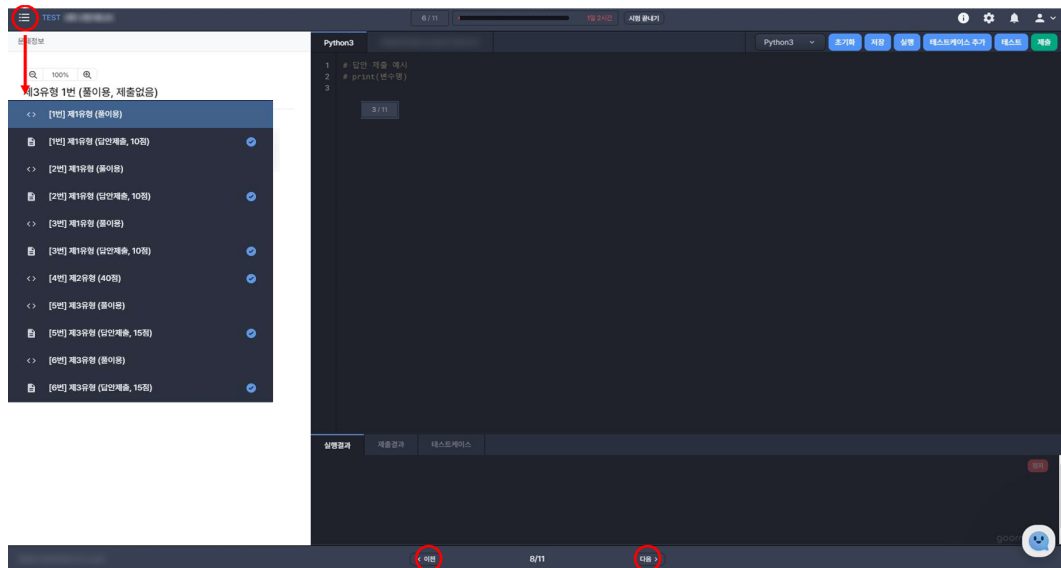
- 윈도우 기본 메모장은 사용할 수 없으며, 시험 시작 후 코딩화면에서 제공되는 메모장 기능 이용
- 코딩화면 내 "메모장" 버튼을 클릭하면 메모장이 활성화되며, 답안제출 페이지에서는 제공되지 않음
- 메모장 "문제탭 자동 이동" 옵션을 활성화하면 메모장 탭이 문제 번호를 따라 같이 이동하며, 비활성화하면 문제를 이동하여도 작성 중이던 탭을 그대로 사용할 수 있음
- 메모장 내용은 입력 후 1초 안에 자동 저장되며, 메모장 오른쪽 하단에 저장 완료 안내



\* 메모장 화면은 기능 및 예시가 일부 상이할 수 있음

## ○ 문제 이동

- 좌측 상단의 아이콘으로 문제를 선택하여 이동
- 또는 화면 하단의 “이전” 또는 “다음” 버튼으로 이동



## ○ 제출 여부 확인

- 좌측 상단의 아이콘을 클릭하여 문제 선택 리스트 내 “체크” 버튼으로 확인
- 또는 문제 내 제목 우측 “제출완료” 버튼으로 확인



- 화면 상단에서 제출한 총 문항수 확인(제출한 문항수/총 문항수)

\* 총 문항수에는 제1유형과 제3유형의 코딩화면(풀이용)이 포함되어,  
모두 제출시 6문항으로 표기

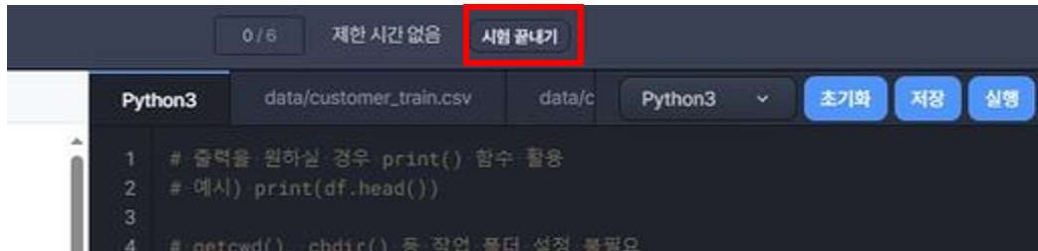
6 / 11

## ○ 시험 끝내기

- 화면 상단 중앙의 “시험 끝내기” 버튼을 클릭하여 시험을 종료하며, 종료 후에는 답안 제출 불가

**\* 시험시간 종료 전에 퇴실할 경우에도 “시험 끝내기”를 클릭해야 하며, 버튼 클릭 여부를 감독관에게 확인받은 후 퇴실**

### 시험끝내기



## ○ 시험 종료

- 종료 후 [제출현황보기]를 통해 문항별 답안 제출 여부 확인이 가능하며, 제출이 완료된 문항은 제출결과 “완료”로 표시
- 단, 제1유형과 제3유형의 풀이용 코딩화면은 “미제출”로 표시되고, 답안제출 화면이 “완료”로 표시되면 제출이 완료된 것임

문항이름 테스트 진행중 순위보기 제출 현황 보기

제출 현황 6 / 11

문제 이름	제출 결과	최종 제출 시각
[1번] 제1유형 (풀이용)	미제출	-
[1번] 제1유형 (답안제출, 10점)	완료	2024. 11. 20. 14:00:00
[2번] 제1유형 (풀이용)	미제출	-
[2번] 제1유형 (답안제출, 10점)	완료	2024. 11. 20. 14:00:00
[3번] 제1유형 (풀이용)	미제출	-
[3번] 제1유형 (답안제출, 10점)	완료	2024. 11. 20. 14:00:00
[4번] 제2유형 (40점)	완료	2024. 11. 20. 14:00:00
[5번] 제3유형 (풀이용)	미제출	-
[5번] 제3유형 (답안제출, 15점)	완료	2024. 11. 20. 14:00:00
[6번] 제3유형 (풀이용)	미제출	-
[6번] 제3유형 (답안제출, 15점)	완료	2024. 11. 20. 14:00:00

## 시험 중지·무효 및 부정행위

1. 아래 행위는 시험 중지 및 퇴실조치 사유이며 해당 시험이 중지 및 무효 처리 됩니다.

- 1) 휴대전화를 포함한 전자·통신기기로 인한 소음이 발생한 경우
- 2) 문제지의 전부 또는 일부를 잘라 소지한 채 퇴실하거나, 시험문제 및 작성답안을 옮겨적어 밖으로 가져나가는 경우
- 3) 시험장에서 소란을 피우거나 타인의 시험을 방해하는 경우
- 4) 시험장의 각종 시설, 장비 등을 파괴, 손괴, 오손하는 경우
- 5) 응시안내 등을 통해 공지한 시험 중지 및 무효사유에 해당하는 경우
- 6) 기타 시험장의 질서유지를 위해 퇴장 또는 응시를 중지시킬 필요가 있는 경우

2. 아래 행위는 부정행위로 간주하여 해당 시험을 중지 및 무효 처리되며, 국가기술자격법에 따라 시행되는 모든 국가기술자격시험에 3년간 응시자격이 정지됩니다.

- 1) 시험 중 다른 수험자와 시험과 관련된 대화를 하는 행위
- 2) 답안지를 교환하는 행위
- 3) 시험 중에 다른 수험자의 답안지 또는 문제지를 엿보고 자신의 답안지를 작성하는 행위
- 4) 다른 수험자를 위하여 답안을 알려주거나 엿보게 하는 행위
- 5) 시험 중 시험문제 내용과 관련된 물건을 휴대하여 사용하거나 이를 주고받는 행위
- 6) 시험장 내외의 자로부터 도움을 받고 답안지를 작성하는 행위
- 7) 미리 시험문제를 알고 시험을 치른 행위
- 8) 다른 수험자와 성명 또는 수험번호를 바꾸어 제출하는 행위
- 9) 대리 시험을 치르거나 치르게 하는 행위
- 10) 수험자가 시험시간에 통신기기 및 전자기기 등 을 사용하여 답안지를 작성하거나 다른 수험자를 위하여 답안을 송신하는 행위
- 11) 그 밖에 부정 또는 불공정한 방법으로 시험을 치르는 행위
  - 실기시험 중 허가되지 않은 웹사이트에 접속 또는 검색하거나, 이를 이용하여 답안을 작성하는 행위
  - 실기시험 중 USB 등 허가되지 않은 장치를 응시환경에 연결하거나, 이를 이용하여 답안을 작성하는 행위
  - 그 밖에 부정 또는 불공정한 방법으로 시험을 치르는 행위