

훈련과정	파이썬과 R을 활용한 빅데이터 분석 실무자 양성_혼합훈련				훈련기간	2021-02-15 ~ 2021-06-30(1회차)		
평가일시	2021.04.01	훈련생	왕지영(인)			훈련교사	손 형 진 (인)	
교과목명	능력단위명		평가방법		배점	점수	총점	성취수준
탐색적 데이터 분석	탐색적 데이터 분석		직무 수행 평가	사례연구				
				포트폴리오				
성취수준	매우높음(5수준)	높음(4수준)	보통(3수준)		낮음(2수준)		매우낮음(1수준)	
성취기준	90-100점	80-89점	70-79점		60-69점		59점 이하	
난이도수준	5(아주어려움)	4(어려움)	3(보통)		2(쉬움)		1(아주쉬움)	

〈 정기 평가의 목적 〉

〈 평가자 유의사항 〉

능력단위 요 소	기본통계 확인하기	평가 방법	사례연구	난이도	3	배점	40
	데이터 분포 분석하기						
	변수간 관계 확인하기						
평가문항 ①	제공된 데이터셋은 세계은행(World Bank)에서 집계한 세계 246개국 인구와 관련된 다양한 통계량이다. 제공된 데이터셋을 이용하여 R code로 다음을 수행하고, 결과를 팀별로 리포트하시오.						
	팀명	A1팀		B1팀		C1팀	
	팀원	김나현, 이상언, 유하영		최진우, 김호중, 이혜수		김정훈, 전해영, 정승호	
	1) 2019년 기준 나라별 총 인구수, 그나라에서 가장 큰 도시에 사는 인구수, 나라별 총 인구수 대비 그 나라에서 가장 큰 도시에 사는 인구의 비율을 구하시오.						
	2) 총 인구수 대비 그 나라에서 가장 큰 도시에 사는 인구의 비율이 가장 높은 나라 10개국 리스트를 작성하시오.(비율 항목 포함)						
정 답	※ 조사한 내용을 카페 내 '과제제출' 게시판에 업로드 할 것. (사례연구 보고서, R 코드를 제출할 것)						
	채점기준						
	평가문항 ②						
능력단위 요 소	기본통계 확인하기	평가 방법	포트폴리오	난이도	3	배점	30
	데이터 분포 분석하기						
	변수간 관계 확인하기						
평가문항 ②	아래 문제를 R 코드로 작성하여 제출하시오.						
	dply 패키지와 iris 데이터 넷을 대상으로 아래의 문제를 실행하는 R 코드를 작성하여 제출하시오.						
	1) iris의 꽃받침의 폭(sepal,Width)이 3.7 이상의 값만 필터링하여 화면 출력하시오.						
	2) 1)의 결과에서 2,4,5번째 컬럼을 선택하시오.						
	3) 2)의 결과에서 2번 컬럼의 값에서 4번 컬럼의 값을 뺀 diff 파생변수를 만들고, 앞부분 10개만 출력하시오.						
정 답	※ R 코드 자료를 카페 내 '평가답안제출' 게시판에 업로드 할 것. (R 코드 안에 본인의 성명을 기입하고 R 파일명에 본인의 영문이니셜을 추가할 것)						
	채점기준						
	평가문항 ③						

이력단위 교 소	기본통계 확인하기	평가 방법	포트폴리오	난이도	4	배점	30
	데이터 분포 분석하기						
	변수간 관계 확인하기						
평가문항 ③	아래 문제를 R 코드로 작성하여 제출하시오.						
	<div>“user_data.csv”와 “return_data.csv” 파일을 이용하여 고객별 반품사유코드(return_code)를 대상으로 다음과 같이 단계별로 실행하여 파생 변수를 추가하시오.</div> <div>반품사유코드에 대한 파생변수 컬럼명 설명: 제품이상(1) → return_code1 원인불명(2) → return_code2 변심(3) → return_code3 기타(4) → return_code4</div> <div>1) 고객 정보 파일 가져오기</div> <div>2) 반품 정도 파일 가져오기</div> <div>3) 고객별 반품사유코드에 따른 파생변수(customer_return) 생성하기</div> <div>4) 고객정보(customer_data)에 파생변수 (customer_return, 반품사유 컬럼)를 추가하여 고객반품정보(customer_return_data)를 만들고 맨 앞 6개 데이터 화면 출력하기</div> <div>5) 고객반품정보(customer_return_data) 테이블에서 맨 밑에서 10개를 화면 출력하기</div>						
정 답	※ R 코드 자료를 카페 내 ‘평가답안제출’ 게시판에 업로드 할 것. (R 코드 안에 본인의 성명을 기입하고 R 파일명에 본인의 영문이니셜을 추가할 것)						
채점기준	<div>〈채점기준표〉 - 평가항목을 수행하지 못하였을 경우 해당 항목 0점 처리</div> <div>각 항목별 미수행 시 항목별 감점</div>						