**[ 7차 훈련생 실습형 평가지 ]**

|  |  |  |  |
| --- | --- | --- | --- |
| **훈련과정명** | 텐서플로우(Tensorflow)를 활용한  빅데이터 분석 양성과정 | **훈련기간** | 2017-09-18 ~ 2018-02-26 |
| **강 사 명** | 설 진 욱 | **평가일자** | 2017. 12. 20. |
| **교과목 명** | 빅데이터 결과 분석 및 활용 | **총 평가시간** | 75 분 |
| **총 문항수** | 15 문항 | **총 배점** | 75 점 |
| **훈련생명** | 하 승 원 | **총 점** | 점 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 결과 스토리텔링하기 | **평가방법** | 서술형 | **난이도** | ★★★★★ | **배점** | 10 | **평가**  **시간** | 10분 |
| **문 항** | **6. 진행한 팀 프로젝트의 전반적인 개요를 작성하시오.**  **1) 주제**  **2) 목적**  **3) 내용**  **4) 기획 의도**  **5) 기타** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  진행 중인 프로젝트에 대한 기획 및 설계 능력 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   1) 주제  -> 우리나라의 항우울제 사용량 빅데이터 분석  2) 목적  -> 우울증 처방약 빅데이터를 활용하여 우울증 발생 원인이 되는 요인을 파악하여 우울증의 효율적인 예방 방법을 탐색해본다.  3) 내용  -> 국민건강보험자격자의 약 2%에 해당하는 100만명의  2012년부터 2015년까지 4개년의 약품처방내역정보를 중  항우울제의 처방기록 분석  4) 기획 의도  -> 제약회사의 영업활동 전략을 수립과 우울증 발생빈도 및  발생조건들을 추측할 수 있는 근거자료로 활용  5) 기타  -> 수업시간에 학습한 R 프로그래밍 능력 향상 | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 결과 스토리텔링하기 | **평가방법** | 서술형 | **난이도** | ★★★★☆ | **배점** | 5 | **평가**  **시간** | 5분 |
| **문 항** | **7.** **프로젝트를 위해서 수행한 요구 사항 정의/분석 내용을 서술하시오.** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  하고자 하는 프로젝트의 업무 요구 사항을 정확히 인지하고 있는 가에 대한 체크 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   - 효율적인 데이터 조회 및 정제를 위해 데이터셋을 DB화 -> Oracle DB 마이그레이션 작업  - 우울증과 기상과의 상호관계 및 항우울제의 소비량과 의약품 관련 배경지식 조사  - 통계적 지식을 기반으로 분석방향 설정  - 항우울제 데이터들을 가공한 후, 여러가지 종류의 그래프와 그림으로 시각화  - 지역별(수도권, 비수도권) 집단간의 항우울제 판매총액 평균 독립표본 t-검정  - 항우울제 판매량(처방량)과 기상 조건과의 관계 분석을 위한 다중 회귀 분석 | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 결과 스토리텔링하기 | **평가방법** | 서술형 | **난이도** | ★★★☆☆ | **배점** | 5 | **평가**  **시간** | 5분 |
| **문 항** | **8. 프로젝트에서 담당한 역할 및 주요 개발 내역을 서술하시오.** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  프로젝트 내에서의 역할 및 주요 개발 내역 체크 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   - Oracle DB 마이그레이션 작업  - 항우울제 N06AB, N06A 성분명코드로 빈도순 워드클라우드 작성  - 각 시도별 항우울제 N06AB 처방량에 대한 막대그래프 작성  - 지역별(수도권, 비수도권) 집단간의 항우울제 판매총액 평균 독립표본 t-검정 | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 결과 스토리텔링하기 | **평가방법** | 서술형 | **난이도** | ★★★★☆ | **배점** | 10 | **평가**  **시간** | 10분 |
| **문 항** | **9. 수집된 데이터들에 대하여 간략히 설명하시오. 또한 이를 가공하기 위하여 수행된 전처리 내역을 서술하시오.** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  결측치와 이상치 데이터들의 전처리 능력에 대한 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   -> 데이터는 국민건강보험공단 홈페이지에서 국민 의약품 처방내역 csv 파일을 다운로드하여 활용하였으며, Oracle sql developer 에서 '데이터 임포트' 기능을 활용하여 데이터베이스에 데이터를 넣어서 R에서 sql 쿼리문을 작성함으로써 데이터를 불러옴으로써 분석에 활용하였습니다. | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각화  하기 | **평가방법** | 서술형 | **난이도** | ★★★☆☆ | **배점** | 5 | **평가**  **시간** | 5분 |
| **문 항** | **첨부 파일 : member.csv, board.csv**  **10. 회원의 급여와 적립 포인트에 대한 상자 수염 그래프를 그리시오.**  **(단, 1개의 차트에 그리도록 하고, 적립 포인트는 곱하기 1000을 이용하여 값의 균형을 맞추시오.)**  회원의 급여와 적립 포인트.png | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  상자 수염 그래프에 대한 이해도 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   member <- read.csv('member.csv', header = TRUE)  mymember <- data.frame(급여 = member$급여, 포인트 = member$적립포인트\*1000)  windows()  boxplot(mymember, xlab = '급여', ylab = '포인트', col = c("red","cyan")) | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각화  하기 | **평가방법** | 서술형 | **난이도** | ★★★☆☆ | **배점** | 5 | **평가**  **시간** | 5분 |
| **문 항** | **11. 회원별 적립 포인트와 급여를 이용하여 다음과 같이 산점도를 그리시오.**  회원의 급여와 적립 포인트에 대한 산점도.png | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  산점도 그래프에 대한 이해도 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   member <- read.csv('member.csv', header = TRUE)  windows()  plot(member$급여, member$적립포인트, xlab = '급여', ylab = '포인트', type = 'l')  windows()  plot(member$급여, member$적립포인트, xlab = '급여', ylab = '포인트', type = 'o')  windows()  plot(member$급여, member$적립포인트, xlab = '급여', ylab = '포인트', type = 'h')  windows()  plot(member$급여, member$적립포인트, xlab = '급여', ylab = '포인트', type = 's') | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각화  하기 | **평가방법** | 서술형 | **난이도** | ★★★☆☆ | **배점** | 10 | **평가**  **시간** | 10분 |
| **문 항** | **12. 성별로 게시물 작성 비율을 파이 차트로 그리시오.**  성별 게시물 작성 비율.png | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  데이터의 집계 능력과 Pie 그래프에 대한 이해도 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   board <- read.csv('board.csv', header = TRUE)  member <- read.csv('member.csv', header = TRUE)  head(board)  member  mydata <- merge(board, member, by='아이디')  head(mydata)  mytable <- table(mydata$성별)  windows()  pie(mytable, label = paste(paste(names(mytable), mytable), "건"), col = c("red","cyan"), main = "성별 게시물 작성 비율") | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각화  하기 | **평가방법** | 서술형 | **난이도** | ★★★★☆ | **배점** | 10 | **평가**  **시간** | 10분 |
| **문 항** | **13 .** **회원별 게시물 작성 건수가 많은 것부터 역순으로 정렬하여 막대 차트를 그리시오.**  회원별 게시물 작성 건수.png | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  데이터의 정렬과 막대 그래프에 대한 이해도 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   board <- read.csv('board.csv', header = TRUE)  member <- read.csv('member.csv', header = TRUE)  mydata <- merge(board, member, by='아이디')  head(mydata)  mytable2 <- table(mydata$이름)  mytable2  myframe <- sort(mytable2, decreasing = TRUE)  windows()  barplot(myframe, xlab = '회원 이름', ylab = '게시물 건수', main = '회원별 게시물 작성 건수', col = rainbow(4)) | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각표현하기 | **평가방법** | 서술형 | **난이도** | ★★★★☆ | **배점** | 10 | **평가**  **시간** | 10분 |
| **문 항** | **14. 프로젝트를 하면서 도출된 주요 기능들의 화면을 출력하시오.**  **(화면이라 함은 차트, 표, 워드 클라우드 등의 이미지 데이터를 의미한다.)** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  개발 내역에 대한 주요 기능을 표현할 수 있는 능력 점검 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   C:\work\2nd_project\2015 N06A 성분명 빈도수 워드클라우드.png  C:\work\2nd_project\2015년 시도별 항우울제(N06AB) 판매량(수정).png | | | | | | | | | | |
| **관련**  **능력단위** | 빅데이터 분석 결과 시각화 | **관련**  **능력단위**  **요소** | 분석 정보 시각표현하기 | **평가방법** | 서술형 | **난이도** | ★★★★☆ | **배점** | 5 | **평가**  **시간** | 5분 |
| **문 항** | **15. 프로젝트를 통한 기대 효과 및 앞으로 나아가야 할 방향에 대하여 제시하시오.** | | | | | | | | | | |
| **평가 시**  **유의 사항** | [출제 의도]  프로젝트 결과물 관리 능력 및 향후 활용 방안에 대한 능력 평가 | | | | | | | | | | |
| **정 답** | |  |  | | --- | --- | | **득 점** |  |   질환별 대한민국 국민이 약에 지출하는 비용 절감을 위한 방안을 수립하고,  질환의 발생빈도를 추정하여 역학연구의 근거자료로 활용할 수 있다. | | | | | | | | | | |