데이터 분석 과제 및 프로젝트

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 직업 만족도 설문 조사 ANOVA 분석 | 프로젝트기간 | 1개월 |
| 주요역할 및 담당 | 다중 회귀 모델 테스트 적용, R-Square, F-score, 유의확률, 잔차 분석 및 상관 관계 분석 실시 | | |
| 업무 성과 | - 다중 회귀 모델 테스트 수행 방법 및 분석 보고서 작성 능력 향상  - 가설 검증 결과 해석 능력 향상 | | |
| 참고 자료 | <https://github.com/marvin-jung/Employee-Job-Satisfaction-Survey>  Diagram  Description automatically generated  A screenshot of a cell phone  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 온라인 잡화 쇼핑몰 마케팅 KPI 대시보드 구축 | 프로젝트 기간 | 14일 |
| 주요역할 및 담당 | - Google Merchandise Store 실시간 Log 데이터를 이용하여 KPI 계산  - Google Analytics 도구로 핵심 데이터 KPI 지표 시각화 | | |
| 업무 성과 | - 이커머스 쇼핑 비즈니스 이해도 증가  - 대시보드 구성 방법 및 비즈니스 성과 지표 해석 능력 향상 | | |
| 참고 자료 | <https://github.com/marvin-jung/Google-Merchandise-Store-Analysis>  Graphical user interface, application  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 월드뱅크 오픈 데이터 EDA 및 Data 시각화 | 프로젝트 기간 | 20일 |
| 주요역할 및 담당 | - Python을 이용하여 미국과 중남미 국가의 사회, 경제, 교육, 보건, 기반시설 지표의 통계 분포 확인  - Histogram, Boxplot 검토 작업으로 수집데이터의 결측치, 이상치 보정  - 국가 별 변수 격차 정도를 Heatmap으로 시각화 | | |
| 업무 성과 | - 파이썬 데이터 시각화 라이브러리 사용 능력 향상  - 적절한 파이썬 조건문, 반복문, 사용자 정의 함수 사용으로 읽기 쉽고 효율적인 코드 작성 | | |
| 참고 자료 | <https://github.com/marvin-jung/Exploratory-Data-Analysis>  Chart, histogram  Description automatically generated  Chart, box and whisker chart  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | Air France SEM 캠페인 분석 과제 | 프로젝트 기간 | 14일 |
| 주요역할 및 담당 | - R, RStudio Connect를 이용하여 SEM 입찰 업체별 매출, ROA, 클릭 당 순수익 시각화  - Logistic 모델, GINI 트리 모델 사용하여 주요 ROA 영향 변수 파악 | | |
| 업무 성과 | - R markdown, flexdashboard 라이브러리 활용 스킬 향상  - Logistic Regression 및 GINI Tree 모델 등 지도학습의 필수 개념 이해 | | |
| 참고 자료 | <https://github.com/marvin-jung/Air-France-Flexdashboard>  Table  Description automatically generated with medium confidence | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | Meal Kit 배달 브랜드 고객 데이터 머신러닝 분석 과제  (Python, Scikit-learn 라이브러리) | 프로젝트 기간 | 21일 |
| 주요역할 및 담당 | - 전년도 플랫폼 이용객 누적 데이터를 4개의 회귀 ML 모델에 Train 시켜 매출에 가장 영향을 주는  소비자 행동 변수 파악  - 5개의 분류 ML 모델에 이용객 데이터를 Train 시켜 와인 구독 서비스 홍보 효과를 극대화 시키는  소비자 행동 변수 파악 | | |
| 업무 성과 | - 회귀, 분류에 사용되는 주요 모델들을 현실 데이터에 적용 가능  - 데이터 전처리, 정규화, 하이퍼파라미터 튜닝을 반복하여 0.7이상으로 예측 정확도 상승 | | |
| 참고 자료 | <https://github.com/marvin-jung/Apprentice-Chef-Analysis>  Diagram  Description automatically generated  Chart, bar chart  Description automatically generatedChart, line chart  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 비지도학습을 이용한 대학생 컴퓨터 구매 패턴 분석  (Python, Scikit-learn 라이브러리) | 프로젝트 기간 | 10일 |
| 주요역할 및 담당 | - 대학생 대상, 개인 성격 특성 및 노트북 브랜드 선호에 대한 설문조사 결과를 PCA분석 및  K-means 클러스터링 하여 MZ 세대의 컴퓨터 구매 패턴 및 잠재적인 고객 성향 확인 | | |
| 업무 성과 | - 차원 축소 모델링을 이용한 데이터 압축 기술 이해  - 클러스터링을 통해 변수들간의 새로운 관계 파악 및 직관적인 데이터 해석으로 유의미한 결론 도출 | | |
| 참고 자료 | <https://github.com/marvin-jung/Windows-or-Mac>  Table  Description automatically generated with medium confidence  Chart, bar chart  Description automatically generated  • Cluster 1, 2, and 3 are most likely to purchase a Mac.  • Cluster 4 is most likely to switch to a PC. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | Twitter 트윗 키워드를 활용한 대형 음원 스트리밍 사이트 비교 | 프로젝트 기간 | 12일 |
| 주요역할 및 담당 | - #Spotify, #Apple, #YouTube, #Music 태그가 있는 트윗을 크롤링  - 대표 음원 스트리밍 기업 3사 간 키워드 상관관계를 시각화  - 빈출 단어들을 Network Diagram으로 연결하여 3사 서비스 비교 분석 | | |
| 업무 성과 | - Tidyverse 라이브러리 활용 스킬 향상  - 텍스트 데이터 시각화 방법 및 분석 능력 향상 | | |
| 참고 자료 | <https://github.com/marvin-jung/Tweets-Text-Analytics>  Chart, scatter chart  Description automatically generated  Map  Description automatically generated with low confidence | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | Speech to Text 자연어 처리 활용 설문 응답자 감정 분석 및 성격 예측 과제 | 프로젝트 기간 | 10일 |
| 주요역할 및 담당 | 37명의 설문 인터뷰 음성을 Text로 변환하고 응답자의 빈출 단어를 이용하여 내향적인 사람들과 외향적인 사람들의 여러가지 감정 분류 결과를 시각화 하였으며 데이터를 8:2 혹은 임의 비율로 Train / Test Set로 나눈 다음 Naive Bayes 성격 예측 모델을 만들어 모델링 정확도 측정 | | |
| 업무 성과 | - 설문조사를 적용한 실무 분석 경험을 통해 NLP 모델의 무한한 확장 발전 가능성을 확인  - 인터액티브한 그래프를 R Shiny를 통해 작성하고 외부 인터넷에 배포하는 방법 이해 | | |
| 참고 자료 | <https://github.com/marvin-jung/Speech-to-Text-Survey-Analysis>  Chart, bar chart  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | Intel APAC 국가 대상 온라인 마케팅 광고 예산 최적화 | 프로젝트 기간 | 14일 |
| 주요역할 및 담당 | 아시아 국가 별 Facebook 광고 마케팅 계획안을 선형계획법(LP), Excel Solver Add-in function을 이용하여 모든 타겟 국가의 Click-throughs를 최대화 할 수 있게끔 Impression 수량 변경 | | |
| 업무 성과 | - Linear Programming, Slack, Sensitivity Analysis 개념 및 모델링 절차 이해  - 소셜 마케팅 비즈니스 용어 이해 및 마케팅 예산 최적화 가능 | | |
| 참고 자료 | <https://github.com/marvin-jung/Intel-Asia-Pacific-The-Catch-Win-Campaign>  Table  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 컨설팅 기업 2019년 재무제표, 손익계산서 페이지 구축 | 프로젝트 기간 | 10일 |
| 주요역할 및 담당 | 5개년 기업 회계 장부 데이터베이스를 MySQL Query 및 Store Procedure 커맨드로 호출 후 `19 회계년도의 재무제표, 손익계산서 완성 | | |
| 업무 성과 | - 회계 장부 데이터 원본에서 돈의 흐름에 따라 어떻게 재무제표를 만들 수 있는지 파악 가능  - ER Diagram 분석 능력 향상 및 이해관계자들의 이해를 돕는 사용자 맞춤 테이블 제작 가능 | | |
| 참고 자료 | <https://github.com/marvin-jung/Accounting-Database>  Table  Description automatically generated  Graphical user interface  Description automatically generated with medium confidence | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 싱가포르 국립대 (NUS) Tableau 스토리보드 홍보물 작성 | 프로젝트 기간 | 14일 |
| 주요역할 및 담당 | 싱가포르 전체 대학 졸업생 취업 현황 설문조사 데이터를 기반으로 싱가포르 국립대 (NUS)가 높은 취업률, 고임금 졸업생들을 많이 배출했음을 알리는 Tableau 스토리보드 홍보물 작성 | | |
| 업무 성과 | - 공공 데이터 자료를 마케팅 캠페인에 적용하는 요령 이해  - Tableau Chart 및 Storyboard 주요 기능 사용법 이해 | | |
| 참고 자료 | <https://github.com/marvin-jung/NUS-Visual-Data-Storytelling>  <https://public.tableau.com/app/profile/ching.chih.chang/viz/Team7-VisualDataStorytelling2_16192431166070/Story>  Chart  Description automatically generated  Chart  Description automatically generated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 프로젝트 명 | 대형 유통 기업 미국 시장 매출 및 수익 분포 현황 대시보드 구축 | 프로젝트 기간 | 14일 |
| 주요역할 및 담당 | Tableau 인터액티브 대시보드를 활용, 대형 유통 기업의 미국 시장의 2개년도 Revenue 및 Profit의 시계열 그래프, 제품군별 그래프, 지도 그래프를 시각화 | | |
| 업무 성과 | - 지리적 데이터 시각화를 통해 보다 높은 수준의 대시보드 완성  - 2년간의 기업 시계열 데이터를 분석할 수 있고 수익률이 낮은 제품군에 대한 문제점 및 발생 원인을 추론 할 수 있는 능력 향상 | | |
| 참고 자료 | <https://github.com/marvin-jung/Superstore-Interactive-Dashboard>  <https://public.tableau.com/app/profile/marvinjung/viz/Data-DrivenStrategicRecommendationsandBusinessReport_16194722823160/StoryPoint>  Graphical user interface, chart, application  Description automatically generated  Graphical user interface, map  Description automatically generated | | |