**문제해결 빅데이터 활용 프로젝트 기획안**

기획안 작성일자 : 2023.05.09

|  |  |
| --- | --- |
| 조 명 | 3조 : 술이문제조 |
| 조 원 | 조장 : 최전석 조원: 김진영, 성나영, 양정아 |
| 프로젝트 주제 및 개요 | 음주운전 데이터 분석을 통해 제작된 시각화 자료를 활용한 보고서 작성 |
| 프로젝트 수행 방향 | ● 프로젝트 목적 설정  1. 분석 목적 및 목표 설정  - 음주운전은 교통사고 발생의 큰 원인 중 하나이며, 사고 발생 시 많은 사회적 비용이 발생  - 문제를 해결하기 위해 음주운전과 관련된 데이터를 수집하고 분석하여 대책을 마련  - 음주운전 데이터를 분석하여 경각심을 높이고 처벌 및 단속을 강화하기 위한 보고서를 작성  2. 보고서 분량 및 내용 설정  - 음주운전의 사고 및 단속 현황, 음주운전의 위험성과 그로 인한 사회적 비용을 시각화 자료를 통해 강조  ● 데이터 수집 및 분석  1. 데이터 수집  - 경찰청, 도로교통공단의 통계와 교통사고분석시스템, 기상청의 시간단위 기상 데이터, 질병관리청을 통한 국민 음주실태조사 등의 시스템을 활용하여 다음과 같은 데이터를 수집  - 음주운전 사고 발생건수, 부상자수, 사망자수, 음주운전 단속과 사고 현황 등의 음주운전과 관련된 데이터  - 음주 소비 현황 및 음주운전 경험률 등 음주문화와 관련된 데이터  - 음주운전 사고 당시의 발생지역, 기상, 사고 다발 구역 간의 관계를 조사하기 위한 데이터  2. 데이터 전처리  - 수집된 데이터 중 결측치가 있는 경우, 평균값, 중앙값, 최빈값 등을 활용하여 결측치를 채우되, 그 과정을 기록  - 수집된 데이터 중 이상치가 있는 경우, 삭제, 대체, 보정 등의 방법을 활용하여 대처하되, 그 과정을 기록  - 데이터 파일들을 파악한 뒤 필요시 데이터의 단위(scale)을 조정  - 데이터 전처리 과정은 jupyter notebook을 사용하며 전처리 코드에 주석을 달아 기록  3. 데이터 분석 및 시각화  - 데이터 분석에는 상관 분석, 회귀 분석 등 통계 기법을 활용  - 데이터 간 관계에 주목하고 연관성을 찾아 새로운 데이터를 생성  - 분석된 자료를 가시성 있게 이해할 수 있게 시각화를 진행  - 지도를 통해 음주운전 적발 지역, 사고발생 지역을 시각화하고, 음주운전 단속/적발 현황, 국민 음주율 증감 추이 등을 시각화하여 음주운전에 대한 대한민국 사회의 인식과 위험 지역 등을 조명  ● 보고서  1. 보고서 기획  - 보고서를 통해 독자가 음주운전에 대한 경각심을 갖고 음주운전에 대한 사회적 관심을 가질 수 있게 보고서의 내용과 구성을 기획  - 음주운전 사고와 지역의 현황, 음주운전에 대한 사회적 관심 등.  2. 보고서 작성  - 음주운전 데이터 분석 결과와 결론을 작성  - 차트, 그래프, 표 등을 활용하여 데이터를 시각화  3. 보고서 편집 및 검토  - 보고서의 가독성과 일관성 높임  - 오탈자와 문법을 검토  - 내용의 출처와 정확성을 검토.  - 수정 사항을 반영하고 최종 검토 과정 진행 |
| 프로젝트 조직  (구성원 및 역할) | ● 역할분담  조장 최전석: 보고서 작성, 데이터 분석 및 시각화  조원 김진영: 데이터 분석 및 시각화  조원 성나영: 데이터 분석 및 시각화  조원 양정아: 데이터 분석 및 시각화 |
| 프로젝트 추진 일정 | ● 일정  5/03 ~ 5/04 : 프로젝트 주제 선정  5/08 ~ 5/09 : 기획안 및 WBS 작성  5/09 ~ 5/09 : 데이터 수집 및 데이터 분석 분량 분배  5/10 ~ 5/12 : 데이터 전처리  5/12 ~ 5/15 : 데이터 전처리 보완  5/15 ~ 5/18 : 데이터 분석 및 시각화  5/19 ~ 5/19 : 보고서 기획  5/19 ~ 5/22 : 보고서 작성, 편집 및 검토  5/23 ~ 5/23 : 발표 준비 및 보고서 최종 검토  5/24 ~ 5/24 : 발표 |