< 미니 프로젝트 중간발표 스크립트 >

안녕하십니까? 노인 교통사고 감소를 위한 데이터 분석이란 주제로 프로젝트를 진행한 OOOO 조

조장 홍해림입니다, 발표 시작하겠습니다.

(다음장)

프로젝트의 목차를 설명하자면, 첫번째에는 팀원 소개 및 프로젝트 수행기간, 두번째는 프로젝트를 진행하게된 배경과 목표, 세번째 프로젝트 데이터 수집에서부터 시각화까지 진행한 수행과정, 네번째에는 수행결과에 대한 설명과 마지막엔 피드백 대한 설명 순으로 발표는 진행됩니다.

(다음장)

팀원 소개부터 진행하겠습니다. 발표를 맡게 된 저는 팀장 홍해림입니다. 그리고 팀원으로는

김소영, 김세정, 박범철, 이유라, 이혁진으로 총 6명입니다.

저희 팀은 프로젝트 순서에 따라 6명이 각각 아이디어 대로 자료를 처리한 후 합치는 과정으로

프로젝트를 진행하였습니다.

(다음장)

프로젝트 기간에 대해서 설명 드리겠습니다

주제 선정을 5월 3일에 완료하였고, 7~8일엔 연구 배경 및 목표 설정을 완료했으며, 8~10일까진 데이터 수집을, 13~14일엔 데이터 전처리를 진행했습니다.

그리고 16~17일에는 탐색적 데이터 분석과 스트림릿을 이용한 시각화 작업을 동시에 진행, 20~24일까진 중간 발표를 위한 자료 제작 및 다듬기 작업을 진행하였고, 24일인 오늘에 중간 발표를 진행하게 되었습니다.

(다음장)

연구 배경부터 설명하겠습니다.

이 신문 기사들을 함께 보실까요? 주요 핵심 내용은 초고령 사회로 진입함으로서 노인 운전자가 증가되면서, 그들의 교통사고 위험성이 증가되고 있다는 점입니다.

초고령사회는 이제 곧 우리가 직면하게 되는 사회적 이슈입니다. 우리는 앞으로 수년 내에 초고령화 사회에 진입하게 됩니다. 이에따라 노인운전자의 교통사고 감소 관련 대처 방안은 시급해졌습니다.

(다음장)

다음 그래프들은 65세 이상의 고령 인구 예상 추이를 나타내고 있습니다

노인 인구가 지속적으로 증가하고 있음을 확인 할 수 있습니다, 핵심은 2000년도 부터 고령화 사회로 진입을 했는데,

시간이 지날수록 노인 인구가 늘어나더니 2026년 그러니까 앞으로 2년 후에 초고령화 사회에 진입한다는 사실을 확인 할 수 있습니다.

노인운전자는 앞으로 더 증가할 것이고 노인운전자의 교통사고는 더 늘어난다는 지표가 되는것입니다.

그럼 저희가 주제로 선정한 노인운전자란 무엇인지 확인해보겠습니다.

(다음장)

노인운전자의 정의에 대해서 설명하겠습니다.

노인 운전자란 고령의 나이에 자동차를 운전하는 운전자를 뜻하는 말로, 일반적으론 65세 이상의 운전자를 의미합니다.

이들은 신체적, 인지적 능력이 젊은 운전자에 비해 저하 될 수 있고, 이는 운전 능력에 영향을 미칠 수 있습니다.

노인 운전자의 특징은 첫번째는 반응 시간의 지연, 두번째 시력 및 청각 저하, 세번째 인지 능력의 감소, 네번쨰 약물복용, 다섯번쨰 오랜 운전 경험으로 정리 됩니다.

그러나 고령 운전자에 대한 사회의 견해는 특징과 일치하는 것도 있고, 불일치 하는것도 있습니다.

불일치하는 것만 짚고 넘어가겠습니다. 안전 수칙과 교통법규의 이해도가 높다라는 물음에 61 %가 동의하지 않는다라는 답변을

젊은 운전자에 비해 안전 지향적이다라는 질문엔 57%가 동의하지 않는다라고 답변 했습니다

위의 조사로 알 수 있는건 사회적으로 노인 운전자가 안전 수칙과 교통 법규의 이해도가 무조건적으로 높고, 젊은 운전자에 비해 안전 지향적일거라는 기존의 선입견을 깨는 내용이며, 노인 운전자에 대한 대처 방안이 필요한 이유를 설명해주는 설문 조사 같습니다.

(다음장)

그렇다면 이런 노인운전자에 대처하는 정부의 정책은 무엇이 있을까요?

과거부터 현재까지 진행 된 정부의 노인 운전자 지원 정책을 한번 살펴보겠습니다.

첫번째는 2019년 부터 시행된 면허 자진 반납 정책입니다.

노인운전자에게 면허증 자진반납을 유도하며, 자진 반납한 노인운전자에게는 지역 화폐 및 교통비 지원을 해주는 정책입니다.

그리고 2021년 6월부터는 차량에 스마일 실버마크 스티커를 부착해, 다른 운전자에게 노인 운전자를 인지 시키는 정책입니다.

또한 2022년에는 고령 운전자에 대한 교육을 의무화 시켜, 안전에 대한 교육을 강화한 정책들이 진행되고 있습니다.

그렇다면 정부의 이러한 정책들로 인해 교통사고는 감소하였을까요?

(다음장)

다음 이미지와 그래프는 정부 지원 정책의 효과에 대해 나타내고 있습니다.

정부 정책이 진행되어도, 매년 노인 운전자의 교통사고는 줄어들지 않고, 인명 피해와 연결 되었습니다. 다시 한번 노인 운전자에 대한 대처 방안의 필요성이 강조 되는 부분 같습니다.

(다음장)

그렇다면 노인운전자는 매년 증가하고 있을까요?

다음 그래프들을 보면 노인운전자는 매년 증가하고 있습니다. 수원시로만 봐도 증가되는 추세입니다.

정부에서 노인 운전자 면허 자진 반납 정책을 시행하고 있음에도 불구하고 증가합니다.

노인운전자의 면허 자진 반납율은 어떨까요?

(다음장)

이 페이지에서의 그래프들은 노인 운전자 증가에 비해 자진 반납은 현저히 적다는게 보여집니다.

비율로 봐도 자진 반납율이 3% 이하입니다.

정부에서 여러가지 정책들을 시행하게 되었는데도 왜 이런 결과가 나올까요?

(다음장)

고령 운전자가 면허 반납을 고려하지 않는 이유를 확인해보았더니

1. 직접 운전시 시간 단축 및 이동 편의성 때문에

2. 충분히 안전 운전할 수 있다고 판단하기 때문에 ,

3. 긴급 상황을 대비해서,

4. 대중 교통의 불편함으로 대답하였습니다

반납을 하지 않는 이유들은 복합적이겠지만, 일반운전자에 비해 노인운전자들의 사고 위험성이 증가하고 있기 때문에 다양한 방안을 모색하여 대처 준비해야 할 것입니다.

(다음장)

다음 내용의 핵심은 65세 이상 운전자의 교통 사고로 인한 사망자 증가가 되고 있다는 점과

정부가 고령 운잔자를 대상으로 운전 평가를 실시하고, 결과에 따라 야간, 고속도로 운전을 금지하는 조건부 면허제 도입까지 검토 중이라는 점입니다.

모든 운동과 신체적 활동에는 에이징 커브라는 것이 있습니다. 노화로 인한 신체 능력의 감소는 노력으로는 극복하기 어렵습니다. 다만

노인 운전자이기 때문에 무조건 제한을 할 수 만도 없고, 분명히 안전 운전이 가능한 노인 운전자도 있을 것입니다.

이러한 배경 속에서 다음 장에서 설명할 목표를 설정하게 되었습니다.

(다음장)

앞으로 초고령 사회로 진입 될 예정이며, 노인 운전자수는 폭발적으로 늘어날 예정입니다.

예방 차원의 목적으로 "노인 운전자 교통사고 예보시스템" 구축을 시도 했습니다.

노인 운전자의 인지 및 신체 능력을 제외 한 나머지 외부 요소인 기상, 노면상태, 주야간, 지역 등을 고려하여 “요인별 위험지수"를 정립하고 데이터 정제를 통해 계산 된 ‘조건에 따른 위험지수’를 기준으로 지역별 실시간 사고 발생 위험도 시각화 및 예보 서비스를 구현하여 수원시 한정으로 제공하는 것을 목표로 삼았습니다..

(다음장)

데이터 수집 및 탐색 부분부터 설명 드리겠습니다.

저희 프로젝트의 메인 데이터는 도로 교통 공단의 교통사고분석시스템인 TAAS와 GIS에서

21~23년 / 지역은 수원시 한정 / 사고 분류엔 노인 운전자 사고에 한해서만 데이터를 분류하여 수집하였고, 기상청에서 수원시의 기상 정보와 데이터 포털에선 고령 운잔자의 자진반납 현황,

경기도 행정 구역별 주민등록 인구 자료를 수집하였고, 경찰청에서는 경기도 운전면허 소지현황을 별도로 정보공개 요청하여 수집하였습니다.

(다음장)

* 데이터 수집 자료들 예시는 슬라이드를 보여주면서 애드립으로 진행

(다음장)

데이터 전처리에 대해 설명 드리겠습니다.

사고 일시 데이터를 datetime을 통해 연도와 월, 일 시간으로 분리하였습니다.

시군구 열을 문자열 분리를 통해 구와 동으로 분리 하였습니다

사고 문자열을 분리하여 두 유형으로 나누었습니다.

도로형태 또한 문자열을 분리하여 두 유형으로 나누었습니다.

1,880개의 데이터 중 47개의 누락 데이터를 확인하였습니다. 누락 된 정보를 확인해 보니 차량 단독 사고의 경우였고

피해 운전자의 차종, 성별 연령, 상해정도의 데이터가 누락되었고, 그것을 0 으로 치환하였습니다.

가해 운전자 연령이 65세로 문자열인 것을 정수로 변환하였고, 피해 운전자 연령 중 '미분류'는 0으로 치환하였습니다.

(다음장)

분석 데이터 추출 및 '사고건수' 추가를 진행 하였습니다. 1880개의 데이터 행은 각각 1개의 개별 사고를 뜻함으로

1로 잡아 사고 건수를 계산 하였습니다.

그리고 ECLO라는 개념을 설명하겠습니다. ECLO는 인명피해 심각도 지수입니다.,

간단하게 말씀 드리면 부상자의 유형에 따라 가산점을 붙여 계산한다 생각하시면 됩니다.

공식으로는 (사망자X10) + (중상자X5)+(경상자X3)+부상신고자수 입니다만 저희는 공식을 약간 추가하여

(사망자X10) + (중상자X5)+(경상자X3)+부상신고자수 / 사고건수 로 계산하여 사고건당 인명피해 심각도로 이름지었습니다.

(다음장)

조건별 사고 건수 및 ECLO 지수 그래프들을 보여드리겠습니다.

(다음장)

연도, 및 기간에 관한 그래프들입니다.

(다음장)

지역(구) 와 지역(동) 별 그래프, 기상 상태, 노면 상태 별 그래프들입니다.

(다음장)

사고유형(대분류), 사고유형(소분류) , 도로형태(대분류), 도로형태(소분류)별 그래프들입니다.

(다음장)

가해운전자 연령, 성별, 차종 / 법규 위반별 그래프입니다.

(다음장)

위의 그래프들의 결과 값에 대한 요약 페이지이고, 조건별 사고 건수와 ECLO 지수의 최대 최소 값에 대한 요약입니다.

ECLO와 사고 건수가 제일 높았던 연도는 2023년이고

ECLO와 사고 건수가 제일 높은 구는 권선구이며

의외였던 점은 사고 건수가 제일 높은 기상 상태는 맑음이였던 점이고

사고유형(대분류) 부분은 차대차 사고가 건수와 ECLO 지수가 제일 높았고

사고유형(소분류) 부분에서 정면충돌이 ECLO 지수가 제일 높았습니다.

가해 운전자 차종 별로 봤을 때, 사고 건수가 제일 많은 것은 승용차였으나

ECLO가 높은 것은 농기계였습니다.

이렇게 핵심적인 내용을 짚어드리고 다음으로 넘어가겠습니다.

(다음장)

사고 발생 위험 예보를 위한 조회 조건 선정 표입니다.

저희가 분류한 다양한 조건 중 실시간 조회 가능성에 초점을 두고 조건을 분류한 결과

시간대와 지역, 기상상태, 노면상태 4가지로 선별하였습니다.

시간대는 주/야간으로 변환하여 사용하였고, 지역은 구 단위로 사전 분석 진행

기상 상태와 노면상태는 웹 스크래핑 + 조건문을 통해 자료를 추출하였습니다.

선별되지 못한 조건들은 대부분 지역별 구분이 불가한 조건이였습니다.

(다음장)

앞에서 선별 한 조건들로 예보 시스템 구현한 과정에 대한 설명 드리겠습니다.

발생 가능한 모든 조건 조합 200개를 생성 하였고

사고 발생 위험 지표를 생성 하였는데, 발생 빈도와 ECLO 지수를 이용하였습니다.

지역별 조건에 해당하는 사고 발생 위험 지표를 추출하여 수원시 지도 영역에 단계 구분도로 표시 할 수 있게 만들었습니다.

단계 구분도란 행정 구역과 같이 지도 상의 어떤 경계에 둘러 싸인 영역에 색을 칠하거나 음영 등으로 정보를 나타내는 시각화 방법입니다.

(다음장)

코딩 관련 된 설명을 드리려고 합니다.

첫번째로는 발생 가능한 모든 조건 조합의 데이터 프레임 생성

Timelist(시간대) : 주간, 야간

gu\_list(지역\_구): 권선구, 장안구, 팔달구, 영통구

weather\_list(기상상태): 맑음, 흐림, 비, 눈, 기타

road\_list(노면상태): 건조, 젖음/습기, 적설, 서리/결빙, 기타

조합 된 200가지를 생성했습니다.

(다음장)

2. 사고 발생 위험 지표 생성(발생빈도)

수집한 2021~2023년 노인 운전자 교통사고 데이터 전처리 및 ‘사고 건수’ 값(1) 추가

사고 발생 빈도 지표에 필요한 열 (조회 조건) 추출

조건 항목별 데이터가 많을수록 큰 값을 가지도록 지표 계산

각 조건의 항목별 총 사고 건수(데이터 개수) 계산을 value\_counts()함수를 통해 진행하였고

총 사고 건수가 적은 순으로 지표 계산하였습니다. Rank(ascending=True,오름차 순)

(다음장)

2. 사고 발생 위험 지표 생성(ECLO 지수)

‘사고 건수’ 값(1) 추가

ECLO 지수에 필요한 열 (조회 조건 및 사고 건수, 사망자수, 중상자수, 경상자수, 부상신고자수) 추출

조건 항목별로 ECLO 지수 계산

각 조건 별 그룹화를 통해 사고 건수 및 사상자 수 총합계 계산 groupby().sum()

ECLO 지수를 계산한 후 사고 건수로 나누는 함수 생성: cal\_eclo()

(다음장)

2. 사고 발생 위험지표 생성(데이터 병합)

앞의 과정으로 만든 “발생 가능한 모든 조건 조합 데이터 프레임”에 조건 항목별 ‘발생 빈도’ 와 ‘ECLO’ 지수 값 취합을 진행한 과정입니다.

(다음장)

2. 사고 발생 위험 지표 생성(지표 계산)

조건 항목 별 ‘발생 빈도’ 와 ‘ECLO’ 지표 값을 활용하여 ‘사고 발생 위험 지표’ 계산

1. 발생 빈도 지표의 합계
2. 발생 빈도 지표와 발생 지표 합계의 상관계수를 곱한 결과의 합계
3. 발생 빈도 지표와 ECLO 지수의 합계
4. 발생 빈도 지표와 ECLO 지수를 곱한 결과의 합계 -> 시각화 활용 지표로 사용

(다음장)

3. 조회 조건 자동 반영

1. 주간/야간 구분 함수 timesplit() 이용

datetime.datetime.now()로 현재 시간 추출

현재 시간 기준으로 주간. 야간 구분

주간: 오전 7시부터 오후 8시까지 (13시간)

야간: 오후 8시부터 다음 날 오전 7시까지(11시간)

2. 수원시 날씨 데이터 수집 함수 weather\_info()

Requests.get(), BeautifulSoup()을 통해 수원시 지역별 네이버 날씨 정보

스크래핑

스크래핑한 데이터 중 현재 날씨, 현재 온도, 현재 체감온도/습도/풍속

현재 강수량 텍스트 추출

3. 기상상태 구분 함수 weathe\_result()

wather\_info() 함수를 통해 추출한 데이터 기준

if–elif-else 문을 통해 기상상태 (맑음, 흐림, 비,눈,기타) 구분

1. 노면상태 구분 함수 road\_result()

Weather\_info() 함수를 통해 추출한 현재 온도, 현재 강수량 데이터 및

Weather result() 함수를 통해 구분한 기상상태 기준

If-elif-else 문을 통해 노면상태(건조, 젖음/습기, 적설, 서리/결빙, 기타) 구분

(다음장)

4. 수원시 지도 영역에 단계 구분도 표시

1. 사용자가 교통사고 위험도를 조회하는 시점의 시간, 날씨 기준 지역(구)별 사고 발생

위험 지표 추출

1. 지역별 조회 조건 및 사고 발생 위험 지표 데이터프레임 생성

(다음장)

4. 수원시 지도 영역에 단계 구분도 표시

1. 수원시 법정경계 파일(geojson)을 기준으로 수원시 지도 영역 생성

2. 지역(구)별 사고 발생 위험 지표로 단계 구분도 표시

(다음장)

지역(구) 별 노인 교통사고 현황 웹 서비스 구현

1. 수원시 지역(구) 선택 후  
   ‘데이터 가져오기’ 클릭 시  
   “지역, 사고일자, 도로형태, 사고유형 등” 교통사고 정보 추출
2. 2021~2023년까지의 해당 지역 교통사고 정보 출력
3. 2번에서 출력 된 데이터 프레임을 CSV 혹은 EXCEL 파일로 다운로드
4. 2021~2023년의 해당 지역 교통사고 정보 시각화  
   - 지역별/ 일별 교통사고 건수  
   - 연도별, 월별 교통사고 건수

(다음장)

지역(구) 별 노인 교통사고 위험도 웹 서비스 구현

1. 조회 시점의 일자 및 시간 출력  
   -datetime.datetime.now()
2. 지역(구)별 조회 조건 및 사고 발생 위험 지표 데이터프레임 출력  
   - 지역(구)  
   - 위험지수 : ‘사고 발생 위험 지표’  
   - 날씨, 온도, 습도, 강수량: 네이버 날씨  
   - 주야간, 노면상태, 기상상태 : 조건문
3. 지역(구)별 ‘사고 발생 위험 지표’로 수원시 단계 구분도 표시  
   - 위험지수가 높을수록 붉은 색으로 표시  
   - ‘구’ 클릭 시 위험지수 표시 팝업 출력

이로서 스트림릿을 통해 웹 서비스 구현한 것까지 보여드리면서 수행 결과 발표를 마무리하고

다음 장에서 발표할 피드백 부분으로 넘어가겠습니다.

(다음장)

프로젝트를 진행하면서 아쉬웠던 부분과 청취자 분들의 의견을 듣는 시간입니다.

아쉬운 점부터 말씀을 드리자면

1. 사고 발생 위험 지표 생성

- 머신러닝을 통해 사고 발생율을 예측한 것이 아니라 임의로 지표를 생성하여 시각화하여

예보의 효율성이 있을지 불확실함

2. 지표 생성을 위해 활용한 조건

- 지역별로 추출하지 못하는 조건을 반영하지 못함

- 수원시 내 기상관측 지점이 2군데 뿐이어서 전 지역의 날씨 데이터가 동일함

- 스크래핑한 날씨를 기준으로 조건에 따라 '기상상태', '노면상태'를 구분함

- '기상상태'는 비교적 정확하게 구분한 것같으나 '노면상태'의 조건이 미흡함

- 수집한 과거 교통사고 데이터 내의 '기상상태', '노면상태'의 '기타'가 불분명함

3. 조회 조건에 따른 결과 출력

- 예보 서비스를 구현하려 했으나 실시간으로만 적용되어 일정 시간 뒤의 위험도를 확인할 수

없음

※ 열심히 배웠지만, 실제로 적용하기엔 너무 여러가지 복잡 요소로 인해 한계를 느끼게 되었습니다. 그리고 많은 것을 배울 수 있던 경험이었습니다.

이상으로 프로젝트를 수행하면서 느꼈던 한계점과 소감에 대한 얘기를 마지막 발표로 마무리 짓고 들어주신 분들의 고명한 의견을 듣고자 합니다, 들어주셔서 감사합니다.