빅기정 기말발표 대본 20173250 정경현

안녕하세요 저는 빅데이터 기초분석과 정보활용 기말프로젝트 발표를 맡은 삼더우먼조 금융정보통계학과 17학번 정경현입니다.

저희는 도로 형태별 계절별 교통사고에 대해서 분석해 보았습니다.

저희는 주제 선정동기와 데이터 탐색 데이터 분석 결론 및 해결방안 순서로 발표를 진행하도록 하겠습니다.

저희의 주제 선정 동기입니다.

먼저 앞에 두자료를 보면 교통사고 건수와 사망자 수가 감소하는 추세지만 사망자수는 여전이 4천명대라는 적지 않은 수치를 띄고 있습니다. 교통사고는 많은 위험요소를 가지고 있습니다.

외형상으로 보이는 상처가 치료되어도 지속되는 교통사고 후유증이 제일 대표적이고 어린이의 사망자수 구성비 추이를 살펴보면 운수사고인 교통사고가 사망원인으로 20년째 제일 많은 요인으로 지목되고 있습니다. 또한 어린이 교통사고의 경우 사망에 이르지 않더라도 정신적인 충격으로 트라우마로 남게 되어 상처로 남을 수 있습니다.

이에 따라서 저희는 교통사고를 최대한 예방할 수 있게 대책을 마련할 필요가 있다고 생각했습니다.

데이터 탐색 전 기본적인 용어의 정의에 대해 설명해 드리겠습니다.

교통사고란 도로교통법 제 2조의 규정에 의한 도로에서 차량의 운행 중 인적인 피해가 발생한 사고로 정의되어있습니다. 교통사고인것 같아 보이지만 제외되는 경우는 도로 이외의 장소에서 발생한 사고나 자살이라고 인정되는 사고 또는 확정적 고의에 의해서 발생한 사고나 건물 육교 등에서 추락해 차량등에 충돌하거나 접촉하여 또는 깔려서 사상한 사고, 높은 곳으로부터 떨어지는 물건에 맞은 사고나 벼락붕괴,도로함몰,유실등에 말려든 사고 ,주차된 차량이 스스로 굴러간 사고,그 밖에 적당치 않다고 판단되는 사고는 제외됩니다.

또한 사망사고는 사망자가 1명 이상이며 교통사고의 경우 교통사고 발생일로부터 30일 이내에 사망한 경우 사망사고라고 말합니다. 저희 데이터에서 쓰인 사망자 데이터는 인수 10만명당 사망자수를 나타내고 있습니다.

저희가 사용한 데이터는 가장 최근 5년동안의 자료를 얻었습니다.

이 데이터에는 연도 월별로 나눠진 데이터가 있고 단일로 교차로 주차장 철길건널목 기타불명의 자료와 발생건수 사망자수 부상자수 중상자수 경상자수 부상신고자 수로 분류 되어 있었습니다.

저희는 이 데이터에서 다음과 같이 전처리를 진행하였습니다.

일단 2017년 자료에는 도로형태별 자료에서 누락된 자료들이 많아서 데이터를 분석할 수 없을 것같아서 제외하고 16년까지의 자료를 자연지리학사전에 나온 자료에 따라 봄여름 가을 겨울을 월별로 나누어 정리한뒤 저희는 주제선정 배경단계에서 주목했던 발생건수와 사망자수의 데이터만 불러와서 데이터를 정리하였습니다.

저희가 사용한 데이터의 도로형태를 설명해 드리겠습니다.

첫번째로는 단일로 중 터널 안이 있습니다. 터널은 운전자의 시야가 좁아지고 속도감이 줄어들면서 발생하는 과속운전으로 인해 일반 교통사고 대비 치사율이 2.3배입니다.

교량은 다리위를 지나는 도로로 교량 밑에 주로 강이 흐르고 있어 빙판이 얼거나 안개가 끼는 일이 잦습니다.

횡단보도는 횡단보도 상과 부근 두가지로 나누어져 있습니다.

횡단보도 상은 흔히말하는 횡단보도의 위에서 일어난 사고를 말하고 신호위반으로 인한 사고가 많이 일어납니다.

횡단보도 부근은 횡단보도가 아니라 횡단보도 근처의 일반 도로를 뜻합니다. 이 곳에서는 보통 사람대 차 사고인 무단횡단사고가 자주일어납니다.

고가도로는 공중에 구조물을 설치하여 그 위에 입체적으로 조성한 도로입니다. 일반도로 위에 설치되있는 경우가 많아 사고시에 추락위험이 있습니다.

지하차도는 일반도로아래에 개통된 도로로 일반도로에서 하차도 입구로 추락하는 사고가 발생하는 경우가 있습니다.

교차로도 교차로 내와 부근으로 나뉘게 되는데

교차로는 두개 이상의 길이 교차된 도로를 말합니다.   
전체 도로형태를 통틀어 가장 많이 교통사고가 발생한 곳은 교차로 내입니다. 교차로 부근은 횡단보도 부근과 마찬가지로 무단횡단으로 인한 사고가 많이 발생합니다.

데이터를 분석하여 그래프로 나타내어 보았는데요 도로형태별로 발생건수와 사망자수 그래프 입니다.

이 그래프는 터널 안 사고 발생건수와 사망자수의 그래프 입니다. 앞서 말했듯이 사망자 수는 인구 10만명당 사망자 수의 그래프 입니다. 사고 발생건수는