

2024.04.16

크롤링 범위

조회기간: 2018-2022 각각

광역자치단체 서울

지자체 : 서울내 각각

차량종류: 전체

상세보기: 자동차 차량종류

위반유형별 - 파파야와수박

사고유형별 - 령우령우

기상상태별 - hazy

도로형태별 - 김춘식

운전경력별 - limit

연령대별 - variable

작업할때 컬럼명 꼭 명시하고, 꼭 데이터하나로 만들필요없어요.

column명

시도 SIDO_NM

지자체 JIJACE_NM

각자유형 TYPE_NAME

사고원인 BASE_NAME

차량종류 GUBUN_NM

발생건수 ACC_CNT

대형사고 BIG_CNT

중대가여객 PSGT_CNT

중대사화물 TRK_CNT

사망자수 DEATH_CNT

치사 CHISA

연도 YEAR

시도, 지자체,
년도,유형,발생건수,차량종류,대형사고,중대사고(여객건),중대사고(화물건),사망자수,치사
율

accident_class, year, cartype

-전체 체크하데 차량종류 체크하기

-컬럼명 년도 붙이기, 유형

-각자 맡은 업무 페이지 만들어서 하는과정 진행상화 보고

-전체 체크하지 마시고 차량종류 다양하게 해서 조사

-다들 잘 분석해옵시다(각자의 재량)

-값이 없을시 0값 넣기

variable 곳 -> 상세검색 항목

1차목표는 **bar** 차트

=>전체 **db** 합해서

2차목표 히트맵 차트-마커 **geotry**

3차목표 연도적으로 보았을때 사건이 감소하고있는 형태, 그 이유를 찾아보자.

연관관계를 찾자. 상관관계를 찾자. 찾자.

상관관계를 찾을 데이터가 필요하다.

전기차 비율도 알아봤음 좋겠다(전기차가 나온이후로 증가했냐)

1종 2종 자동차시험 시험난이도 증가로 교통사고관련이있는가

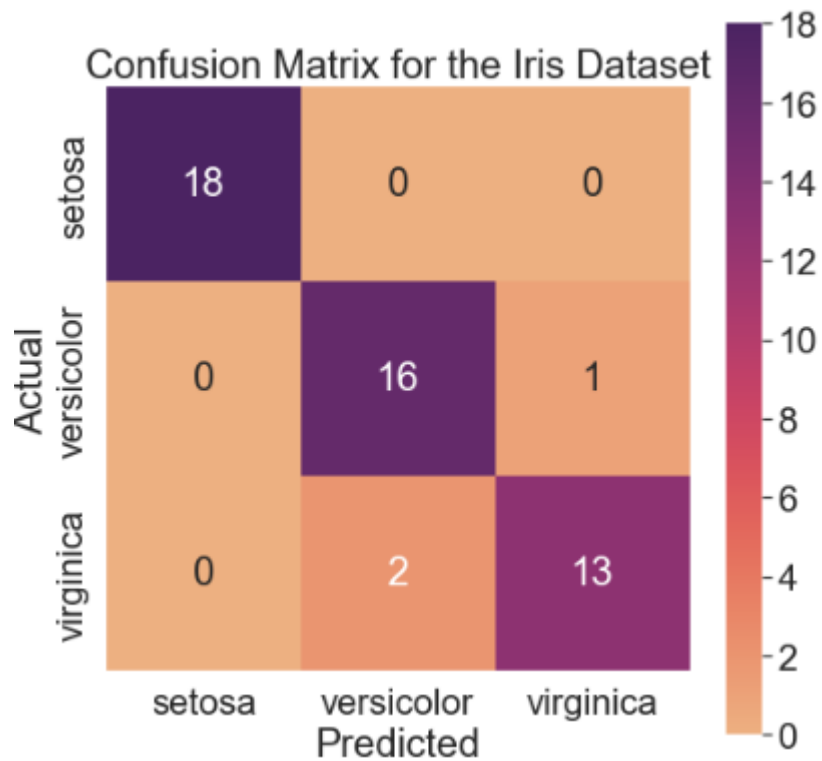
교통인프라

교통세금

교통량(혼잡도)

23개의 컬럼에다가 + **year** 컬럼 하나 추가해서 총 24개의 컬럼

지자체**name**으로 join



column명 유형 발생건수 대형사고 중대사고(여객건) 중대사고(화물건) 사망자수 시차율
유형종류 6개

4.17.오후2시

4.17 오후2시 db통합

운전면허시험 합격률

졸음운전휴게소 설치수

가로등,신호등,유도선 등 운전보조설비 설치현황

또는 교통관리부 세금할당에관하여

— 이 주제들은 현재데이터로봤을때 줄어드는 사고율에대하여

뒷받쳐줄수있는 것들입니다 데이터들을 가져올수있다면 사용가능하지않을까하네요

git 사용 예정

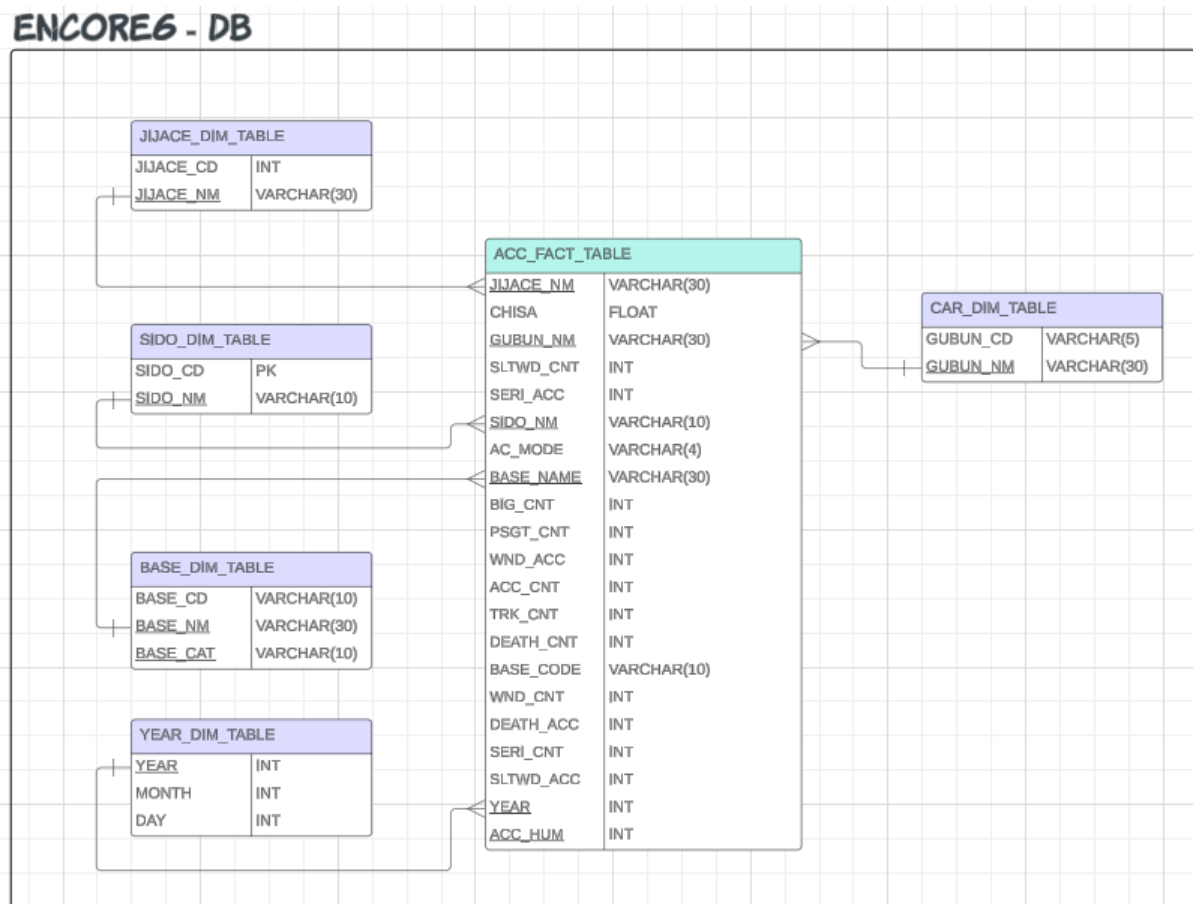
- 하나의 repository를 만들고,
- 해당 repo 공유 후, 각자 소스 코드 공유

2024.04.17

유형별 테이블 6개 -> 유형별 코드, 유형별 이름
메타데이터 테이블 1개 -> 유형,년도,지자체이름

pk

ERD



여기 적은

```
`JIJACE_NM` VARCHAR(30),  
`CHISA` FLOAT,  
`ACC_HUM` INT,  
`SIDO_CD` INT,  
`GUBUN_NM` VARCHAR(30),  
`SLTWD_CNT` INT,  
`SERI_ACC` INT,  
`SIDO_NM` VARCHAR(10),  
`AC_MODE` VARCHAR(4),  
`CRT_JIJACE_CD` INT ,  
`BASE_NAME` VARCHAR(30),  
`BIG_CNT` INT,  
`PSGT_CNT` INT,  
`WND_ACC` INT,  
`ACC_CNT` INT,  
`TRK_CNT` INT,  
`DEATH_CNT` INT,  
`BASE_CODE` VARCHAR(10),  
`WND_CNT` INT,  
`DEATH_ACC` INT,  
`GUBUN` VARCHAR(5),  
`SERI_CNT` INT,  
`SLTWD_ACC` INT,  
`YEAR` INT,  
`BASE_CATEGORY` VARCHAR(10)
```

JIJACE_DIM_TABLE

`JIJACE_NM` VARCHAR(30),

`CRT_JIJACE_CD` INT

SIDO_DIM_TABLE

`SIDO_CD` INT,

`SIDO_NM` VARCHAR(10),

BASE_DIM_TABLE

`BASE_NAME` VARCHAR(30),

`BASE_CODE` VARCHAR(10),

`BASE_CATEGORY` VARCHAR(10)

YEAR_DIM_TABLE

`YEAR` INT

`MONTH` INT

`DAY` INT

CAR_DIM_TABLE

`GUBUN_NM` VARCHAR(30),

`GUBUN` VARCHAR(5),

violation_accident – 설명좀 해주세요

진행상황

1. 따릉이 이용량에 따른 자전거교통사고의 연관관계 파악
2. 차량 속도에 따른 교통사고 연관관계 파악

—

월별 따릉이 **csv**파일 읽어서 하나의 데이터프레임으로 제작

각 1년치 데이터프레임에서 월별(**groupby()**) 이용자수 합계(**sum()**)