



2023/07/25

📌 유형	팀 주간 회의
👥 참석자	 공수빈  나영 이

< 할 일 >

전처리 짬막!!!!!!!

1. 추정매출 데이터

a. 점포수 처리

→ 업종별로 점포수와 상관관계가 가장 높은 변수를 뽑아 해당 업종의 ex)분기당매출금액의 중위값보다 결측치가 있는 행의 분기당매출금액이 높을 경우 2로, 낮을 경우 1로 치환하기

b. 필요없는 열 삭제(서비스업종코드, 기준분기코드, 기준년코드)

c. 데이터타입 object 타입인거 → int 로 변환

2. ~~생활인구~~ 데이터

a. 여성/남성, 주중/주말, 시간대, 연령대 만 남기고 나머지 삭제

b. 데이터타입 object 타입인거 → numeric으로 변환

3. ~~유동인구~~ 데이터

a. 여성/남성, 주중/주말, 시간대, 연령대 만 남기고 나머지 삭제

b. 데이터타입 object 타입인거 → numeric으로 변환

4. 집객시설 데이터

a. 여성/남성, 주중/주말, 시간대, 연령대 만 남기고 나머지 삭제

b. 데이터타입 object 타입인거 → numeric으로 변환

5. 데이터프레임 하나로 만들기

1. 업종별 분기별 나눠서 하나로 만들기

2. 단위 맞추기 - 모든 값을 맞춰줘야 하는건지?? 우선 보류

a. 매출 금액

< 머신 러닝 >

1. 파이썬 머신러닝 완벽 가이드 책 보면서 공부
 2. 위 책으로 만들어진 인프런 강의 들으면서 공부
 - a. <https://www.inflearn.com/course/파이썬-머신러닝-완벽가이드#>
- 각자 맞는 방법으로 결정?