







Data Specifications

expected_sales_20xx.csv

서울시 상권분석서비스(추정매출-상권)







 expected_sales_2019	2025-05-11 오후 10:32	Microsoft Excel ...	36,409KB
 expected_sales_2020	2025-05-11 오후 10:32	Microsoft Excel ...	37,366KB
 expected_sales_2021	2025-05-11 오후 10:33	Microsoft Excel ...	41,481KB
 expected_sales_2022	2025-05-11 오후 10:32	Microsoft Excel ...	41,402KB
 expected_sales_2023	2025-05-11 오후 10:32	Microsoft Excel ...	41,176KB
 expected_sales_2024	2025-05-11 오후 10:32	Microsoft Excel ...	40,641KB

: 서울시 상권 영역 내의 점포들의 추정 매출 정보를 제공합니다(단위 : 원).

- Observations: 2024년 데이터 기준 87,179개. 근데 분기별로 측정했기 때문에 보통 동일한 범주에 대해 4개의 관측치가 있음. 물론 그렇지 않은 범주도 있는 것 같음.
 - e.g. '골목상권_이북5도청사_한식음식점'에 대한 추정 매출 데이터가 분기별로 네 개 존재.
- Columns: 55개. 분기, 상권구분(e.g. 골목상권), 구체적인 상권(e.g. 문래역3번), 업종명('한식음식점'), 당월 매출액/건수, 주중매출액, 주말매출액, 요일별매출액, 시간대별매출액, 성별별매출액, 연령대별매출액, 주중매출건수, 주말매출건수, 요일별매출건수, 시간대별매출건수, 성별별매출건수, 연령대별매출건수.
- Missing Values: 2024년 데이터만 봤을 때는 없어보임.

the_number_of_shops_20xx.csv

서울시 상권분석서비스(점포-상권)

 the_number_of_shops_2019	2025-05-11 오후 10:31	Microsoft Excel ...	33,570KB
 the_number_of_shops_2020	2025-05-11 오후 10:31	Microsoft Excel ...	35,048KB
 the_number_of_shops_2021	2025-05-11 오후 10:32	Microsoft Excel ...	34,311KB
 the_number_of_shops_2022	2025-05-11 오후 10:32	Microsoft Excel ...	34,504KB
 the_number_of_shops_2023	2025-05-11 오후 10:31	Microsoft Excel ...	34,761KB
 the_number_of_shops_2024	2025-05-11 오후 10:31	Microsoft Excel ...	34,656KB

: 서울시 상권 영역 내의 점포 정보 및 개/폐업, 프랜차이즈 정보를 제공합니다.

- Observations: 2024년 데이터 기준 306,889개. 애도 마찬가지로 분기별로 측정하긴 했음.
- Columns: 14개. 분기, 상권구분, 구체적인 상권, 업종명, 점포 수, 유사업종 점포 수, 개업율, 개업점포수, 폐업률, 폐업점포수, 프차점포수
 - 개업율, 폐업률이 정수로 표기되어있는데, 이거 그냥 비율 맞는데 앞에 0만 빼 듯. 파생변수 하나 새로 파는게 나올 수도 있음. **폐업점포수 / 전체점포수** 이런 식으로.

- Missing Values: 2024년 데이터만 봤을 때는 없어보임.