

경제총조사 추진 상황 매뉴얼

분석 보고서 작성자: 배준호 | 매뉴얼 작성자: 배준호

개요

개발 기간

- 2021.06 ~ 2021.07

활용 기술

- Python, Spark SQL, IRIS DSL, IRIS Studio

활용 데이터

- 경제총조사 통합 시스템 조사 현황 (시도, 시군구, 읍면동)

설명

목적

- 수원시 경제 총조사 조사 현황을 시각적으로 상시 확인할 수 있는 대시보드

상세 내용

- 경제 총조사 기간 내 매일 변경되는 조사 현황을 문서로 작성하여 배포해야 함
- 상시 모델링을 적용할 경우 매일 시행해야 하는 반복 작업을 줄일 수 있음
- 시각적인 대시보드 형태로 실무자, 의사결정자의 사업 이해를 도움

목차

1. 화면 구성

- Page.01

2. 사용자 매뉴얼

- Page.02~04

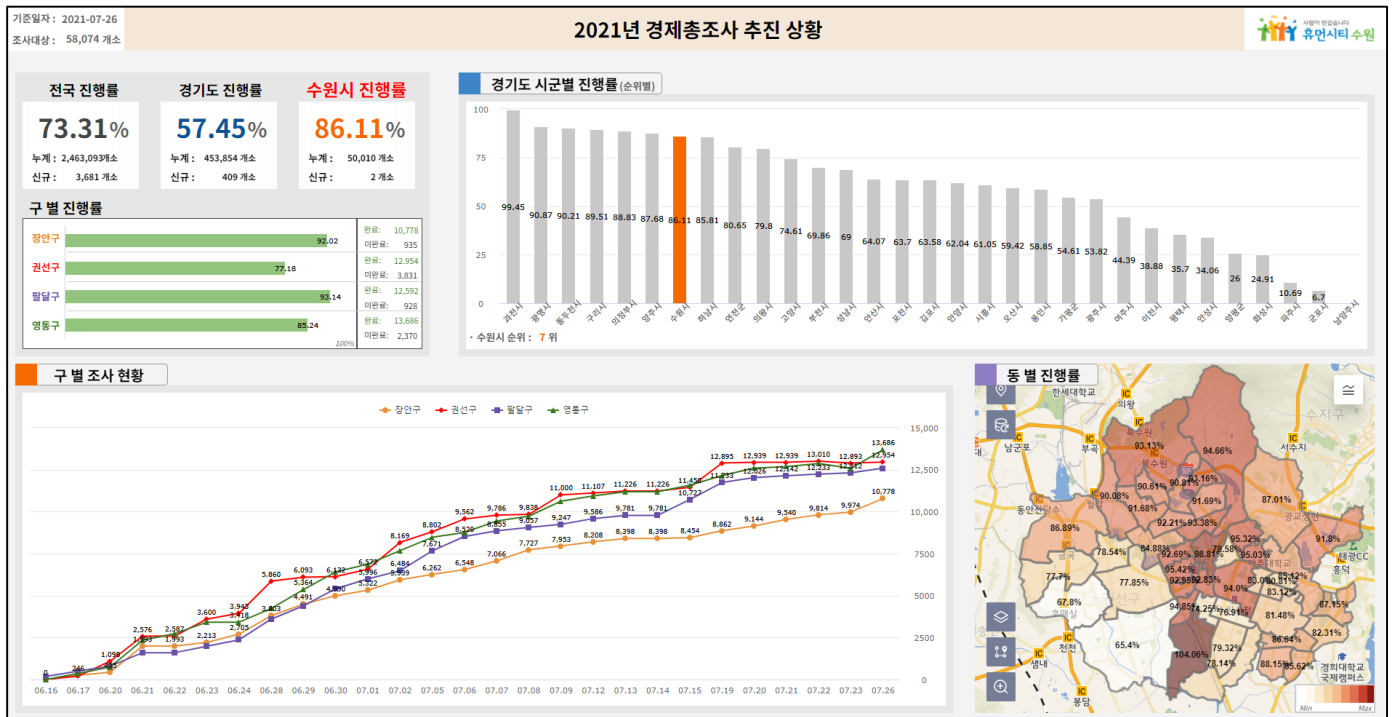
3. 유지관리 매뉴얼

- Page.05~07

1. 화면 구성

화면 1

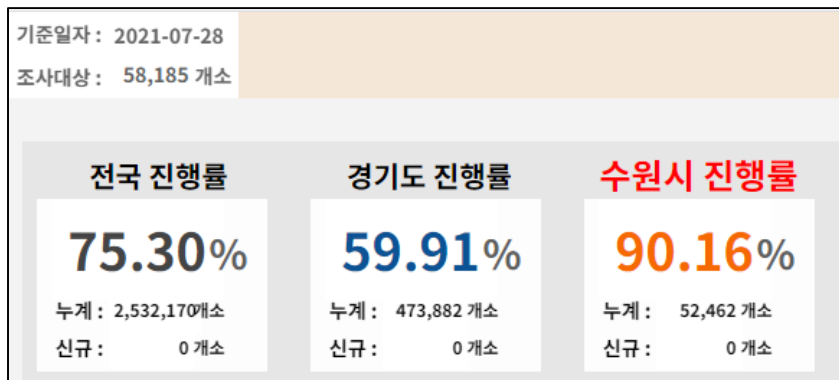
- 2021년 경제총조사 추진 상황



2. 사용자 매뉴얼

전국, 경기도, 수원시 진행률

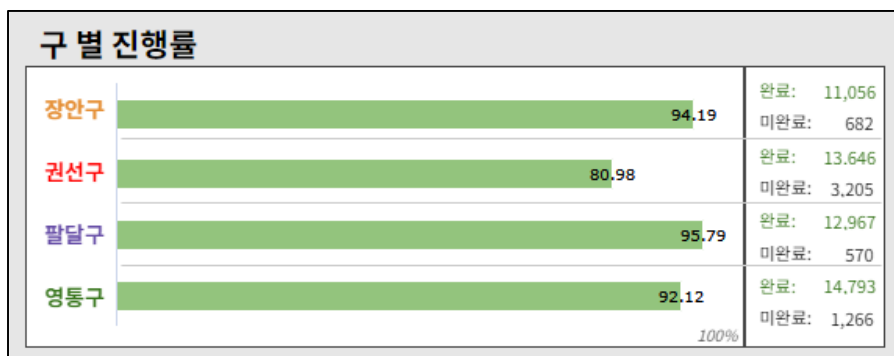
- 경제총조사 데이터



- 기준일자를 통해 데이터 갱신여부를 상시 확인할 수 있도록 단일값 그래프 구성
- 전국, 경기도, 수원시 진행률은 전처리된 조사 현황 데이터를 기반으로 단일값 그래프 표출
- 누계는 기준일자까지 조사된 합계 전체이며, 신규는 해당 일자에 새로 등록된 개소를 의미

수원시 구 별 진행률

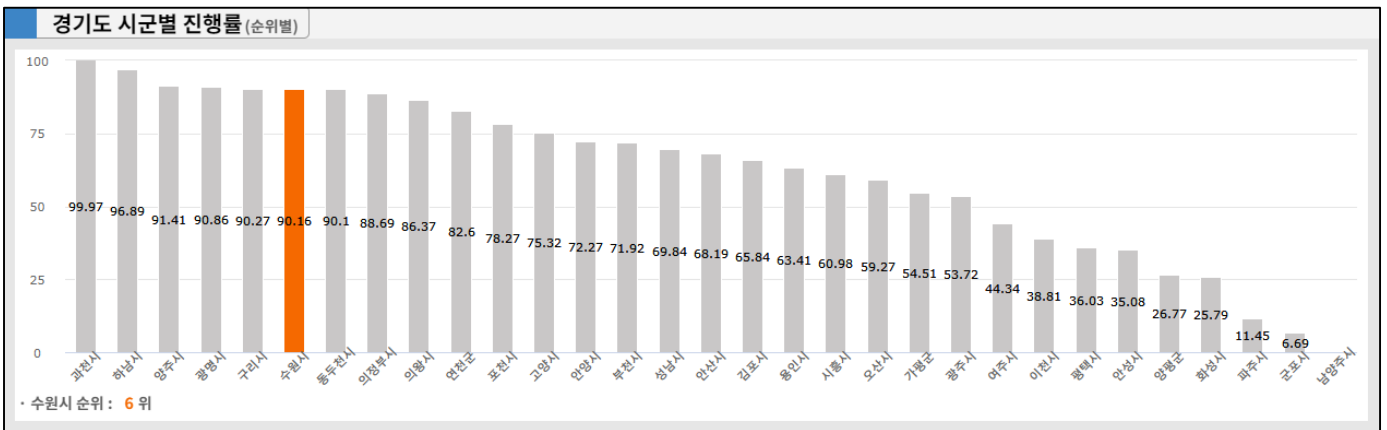
- 경제총조사 데이터



- 4개 구의 조사 진행률을 시각적으로 표현하기 위해 백분율 기반의 가로 막대 그래프로 표출
- 각 구의 완료와 미완료 개소는 단일값 그래프로 구성되어 데이터에 맞게 매일 변경됨
- 그래프에 표현된 선을 텍스트로 구현하였기 때문에 익스플로러 환경에서는 다르게 보일 수 있음

경기도 시군별 진행률

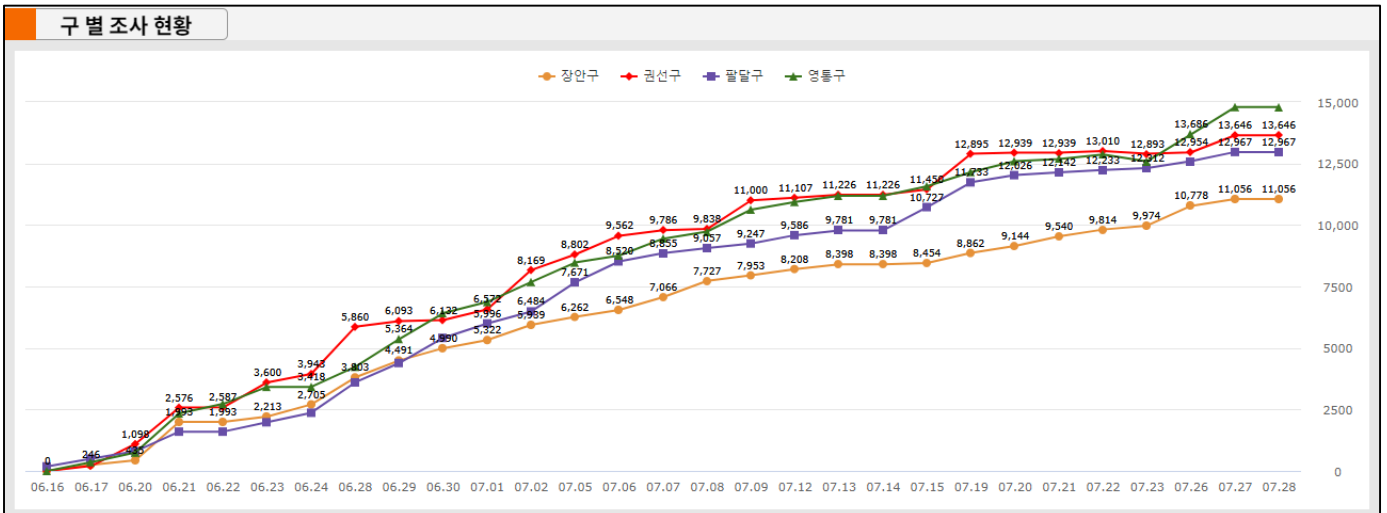
- 경제 총조사 데이터



- 경기도 전체 31개 시군 별 진행률을 다중축 세로 막대 그래프로 표출
- 그래프의 순서는 백분율 순위로 정렬되어 있고, 수원시 순위는 막대 그래프 밑에 단일값으로 표출
- 수원시 진행률의 경우 색상을 별도로 표시하여 강조 처리

구 별 조사 현황

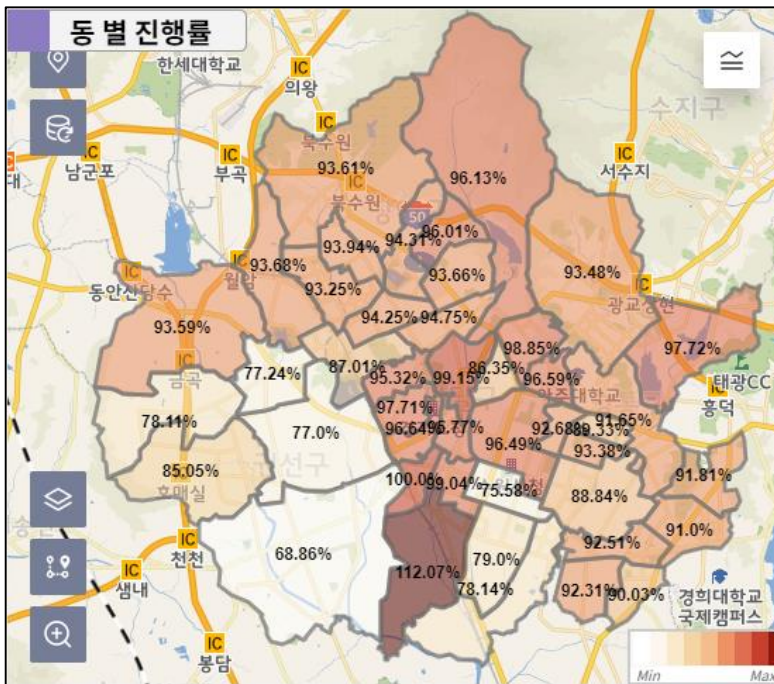
- 경제 총조사 데이터
- 장안구, 권선구, 팔달구, 영통구 경제총조사 데이터 모델 뷰



- 수원시 4개 구 별 조사 현황을 기준 일자 별 누적 값을 표현하는 꺾은선 그래프로 표출
- 원본 데이터의 특징 상 통합 시스템에서 조회하지 못하는 날짜는 데이터가 표출되지 않음
- 각 구의 꺾은선 색상을 다르게 지정하기 위해 데이터 모델 뷰 기반의 다중축 그래프를 이용

동 별 진행률 지도

- 경제 총조사 데이터

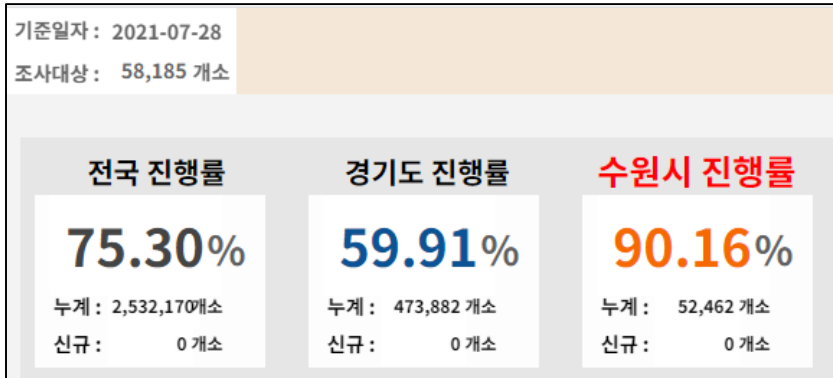


- 수원시 행정동 별 조사 진행률을 지도 시각화 표출
- 각 동의 색상은 우측 하단 범례와 같이 해당 동의 진행률이 높을 수록 진한 빨간색으로 표현됨
- 대시보드 공간 상 제약으로 지도의 크기가 한정되어있기 때문에 툴팁을 이용하여 정보 표출
- 각 행정동 위에 마우스를 올리면 해당 동의 조사 진행률, 전체 조사 대상, 조사 완료 수 확인 가능

3. 유지관리 매뉴얼

전국, 경기도, 수원시 진행률

- common_data_1 ~ 11

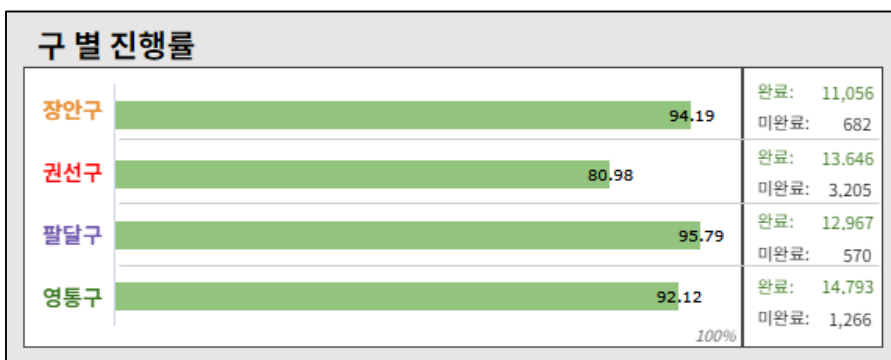


common_data_1 common_data_3 common_data_6 common_data_9

- 가장 최신 데이터 기준일자 추출 (top 1 -REFERENCEDATE) 및 최신 일자 조사대상 추출
(sql "select *, max(REFERENCEDATE) over () as DATE from angora" | where REFERENCEDATE = DATE)
- 행정구역 코드를 기준으로 전국, 경기도, 수원시 조사 진행률을 WHERE LIKE 정규 표현식 함수로 추출
- 누계는 각 일자 별 조사 완료 전체 합으로 구하고, 신규는 당일 조사 수를 표출

수원시 구 별 진행률

- common_data_12 ~ 21

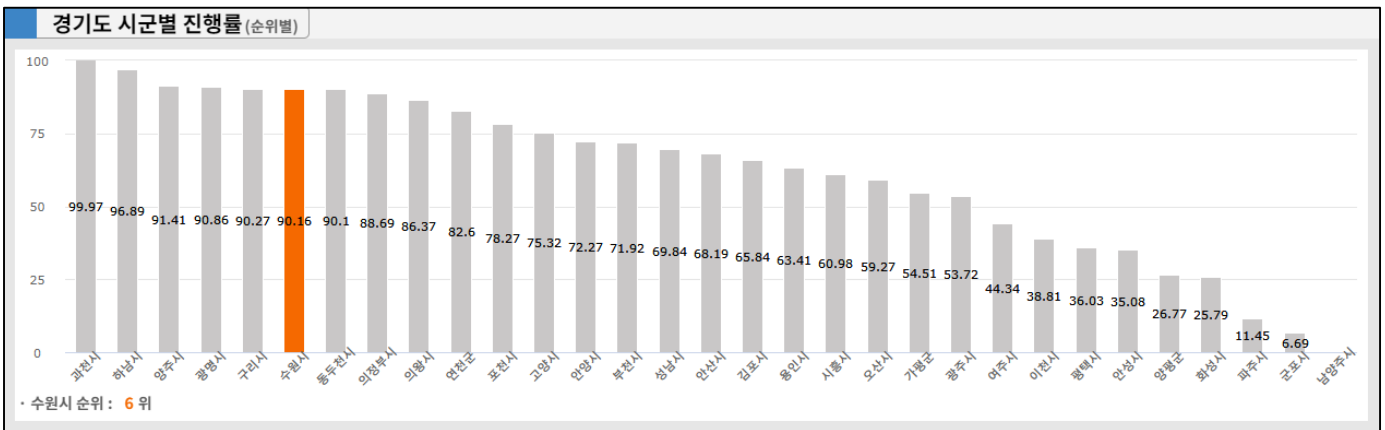


common_data_12 common_data_13 common_data_14

- WHERE LIKE 정규표현식을 이용해 수원시 4개 구를 필터링 후 막대그래프 표출
- 각 구의 조사 대상 업체 수에서 조사 완료 업체 수의 차이로 미완료 개소 수를 파악
(calculate TARGET_BUSINESS_COUNT - COMPLETION_BUSINESS_ACCUMULATE_TOTAL as CAL)
- 현재 Studio에서 선 도형을 지원하지 않기 때문에 빈 텍스트에 밑줄 옵션을 적용하여 선을 표현하였음, 추후 개선이 필요

경기도 시군별 진행률

- common_data_22, 23



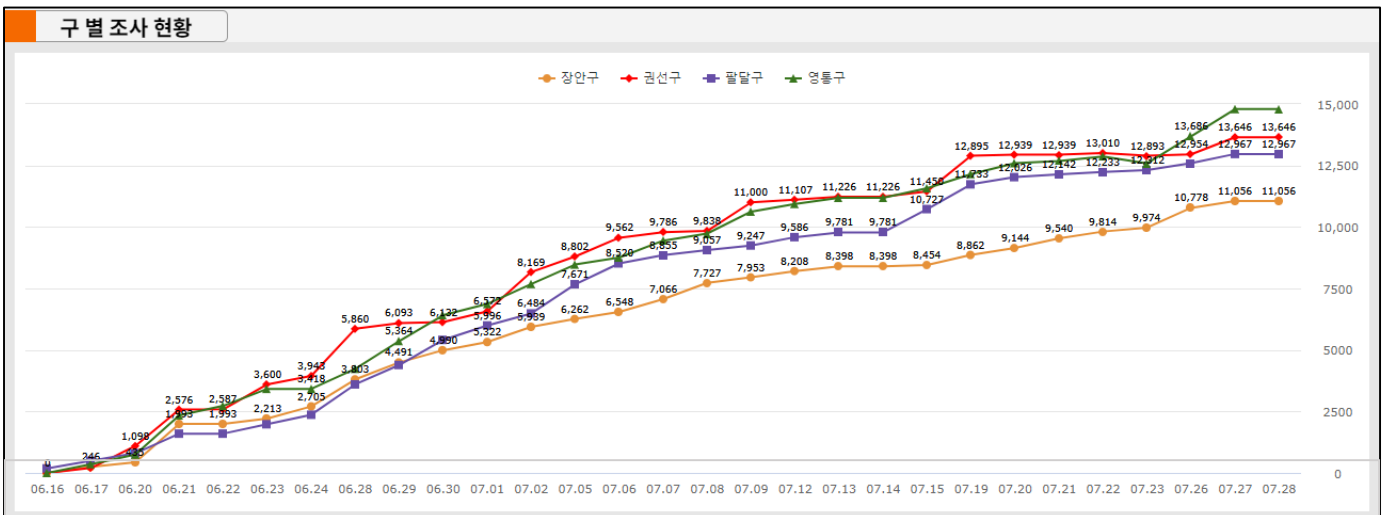
common_data_22

common_data_23

- WHERE LIKE 정규표현식을 이용해 경기도 내 시군 별 조사 현황 데이터 필터링
- 수원시 진행률 막대의 색상을 다르게 하기 위해 칼럼 분리를 시행해야 함
(case when ADMINISTRATION_ZONE_NAME = '수원시' then EXAMINATION_PROGRESS_RATE as 수원시
| case when ADMINISTRATION_ZONE_NAME = '수원시' then null otherwise EXAMINATION_PROGRESS_RATE as 진척률)
- 시군 별 조사율을 내림차순으로 정렬한 뒤, 순위를 부여 (numbering | where ADMINISTRATION_ZONE_NAME = '수원시')

구 별 조사 현황

- common_data_24

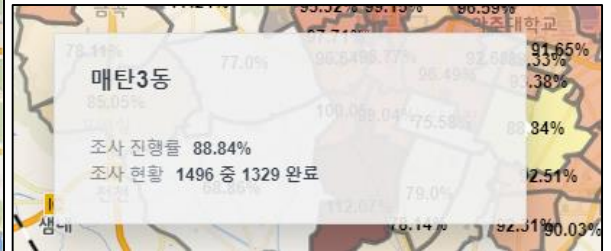
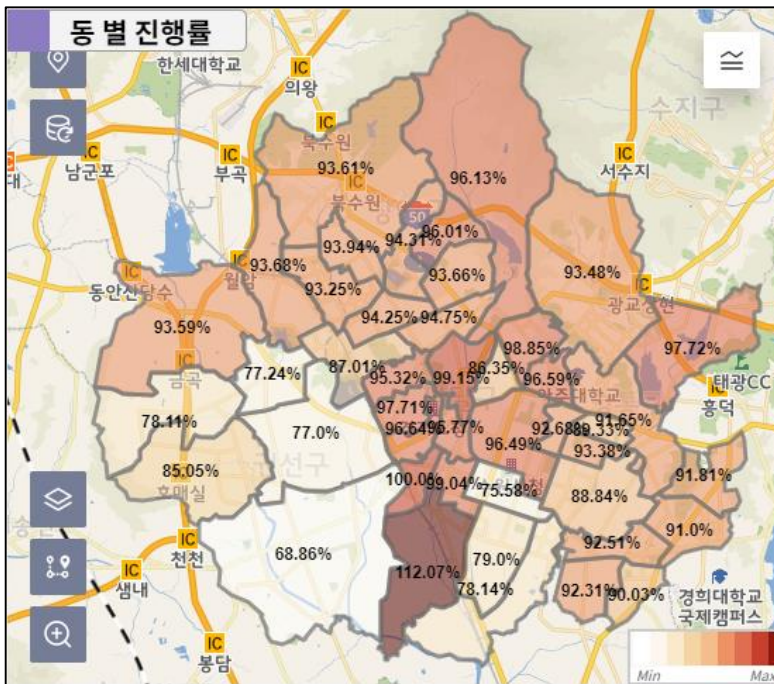


common_data_24

- 현재 꺾은선 그래프 기능에서는 칼럼 기준으로만 색을 지정할 수 있기 때문에 데이터 구조를 변경해야 함
- 모델의 연산을 줄이기 위해 4개 구를 각각 데이터 모델 뷰로 구성 (sql "select ADMINISTRATION_ZONE_NAME, REFERENCEDATE, EXAMINATION_COMPLETION_BUSINESS_ACCUMULATE_TOTAL as TODAY from angora")
- 생성된 각 구 별 데이터 모델 뷰를 조인하여 하나의 누적 조사율 집계 테이블을 생성하고 이를 다중축 꺾은선 그래프로 구성

동 별 진행률 지도

- 경제 총조사 IMA 지도



- SQL을 이용해 최신 기준일자 데이터 필터링 (sql "select *, max(REFERENCEDATE) over () as DATE from angora" | where REFERENCEDATE = DATE)
- CASE문을 이용해 IMA에서 지원하는 공간 데이터 셋 도로명 주소 형태에 맞게끔 데이터를 변환 (시도, 시군구, 읍면동)
- WHERE LIKE 정규 표현식을 이용해 수원시 행정동만 필터링
- 데이터에 따라 변화하는 툴팁 내용을 출력하기 위해 칼럼 생성 (concat EXAMINATION_PROGRESS_RATE, "%" as 조사 진행률 | concat EXAMINATION_TARGET_BUSINESS_COUNT, " 중 ", EXAMINATION_COMPLETION_BUSINESS_ACCUMULATE_TOTAL, " 완료" as 조사 현황)