수원시 후생복지 분석 매뉴얼

분석 보고서 작성자: 배준호 | 매뉴얼 작성자: 배준호

개요

개발 기간

• 2021.05 ~ 2021.06

활용 기술

• Spark SQL, IRIS DSL, IRIS Studio

활용 데이터

- 연도 별 후생복지 참여자 (2016~)
- 후생복지 참여 희망자

설명

목적

• 수원시 후생복지 참여 희망자 중 이전에 복지 혜택을 적게 받은 희망자 우선 선별

상세 내용

- 한정된 복지 자원의 효율적 분배를 위해 희망자 선별 작업이 필요
- 과거 참여 데이터를 기반으로 신규 참여 희망자 중 우선순위 파악
- 사용자가 보고서 기능을 통해 직관적으로 조건에 맞는 희망자 추출 가능

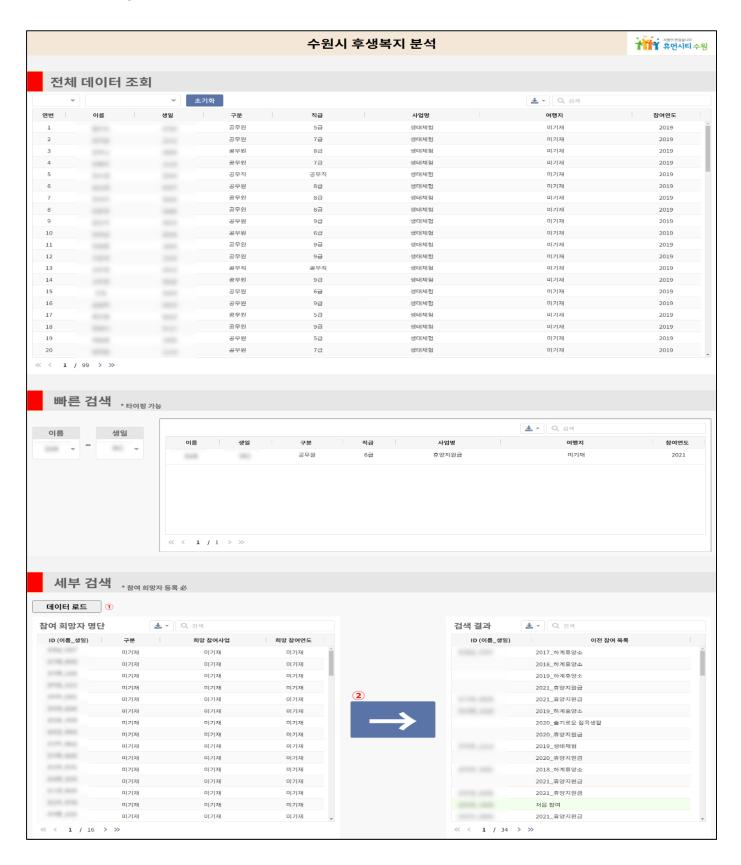
목차

- 1. 화면 구성
- Page.01
- 2. 사용자 매뉴얼
- Page.02~04
- 3. 유지관리 매뉴얼
- Page.05~07

1. 화면 구성

화면 1

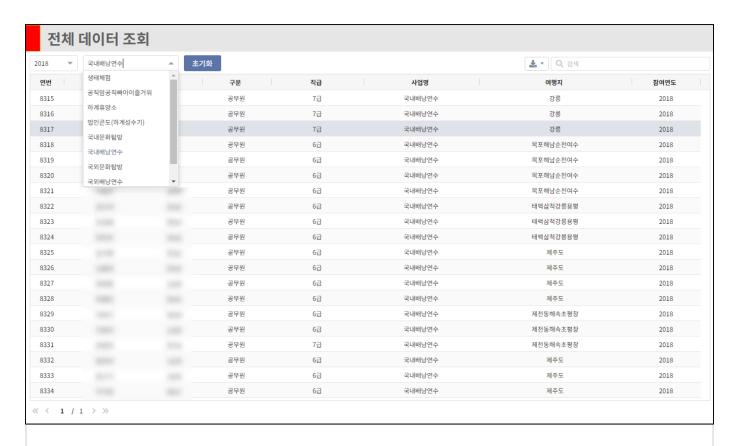
• 수원시 후생복지 분석



2. 사용자 매뉴얼

전체 데이터 조회

• 2016년 ~ 2021년 후생복지 참여 데이터



- 2016년부터 2021년까지 참여한 후생복지자원 명단을 테이블로 시각화
- 이름 오름차순 기준으로 전체 연도를 표출하는 기본 테이블 구성
- 효율적인 데이터 파악을 위해 콤보 박스를 이용해 연도, 사업 명 별로 필터링 가능하도록 구성
- 초기화 버튼 클릭을 통해 검색 후 기본 테이블로 전환 가능
- 검색 기능을 통해 이름, 직업 등 콤보 박스에 없더라도 각 칼럼에 존재하는 내용 검색 가능
- 우측 상단의 버튼을 통해 필터링된 데이터를 CSV 형식으로 해당 다운로드 가능

빠른 검색

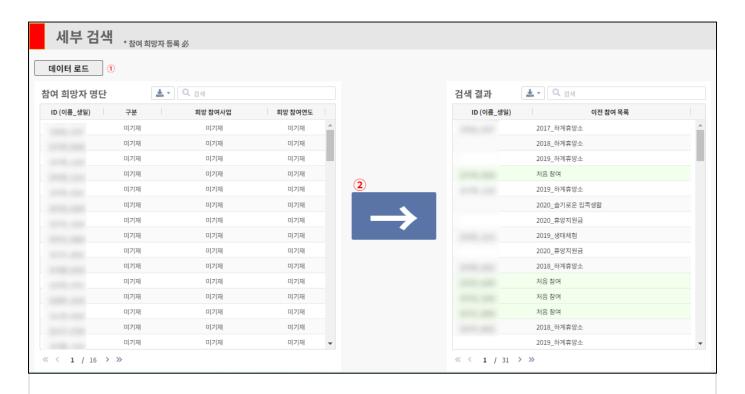
• 2016년 ~ 2021년 후생복지 참여 데이터



- 2016년부터 2021년 까지 참여한 후생복지자원 명단을 테이블로 시각화
- 콤보 박스에서 이름을 선택하거나 타이핑하면 우측 테이블에 필터링된 결과가 표출됨
- 검색 기능을 통해 참여 연도, 직업 등 콤보 박스에 없더라도 각 칼럼에 존재하는 내용 검색 가능
- 동명이인이 있을 경우 생년 선택을 통해 세부적으로 검색하여 필터링 가능
- 우측 상단의 버튼을 통해 CSV 형식으로 필터링된 데이터 다운로드 가능

세부 검색

• 후생복지 참여 희망자 데이터

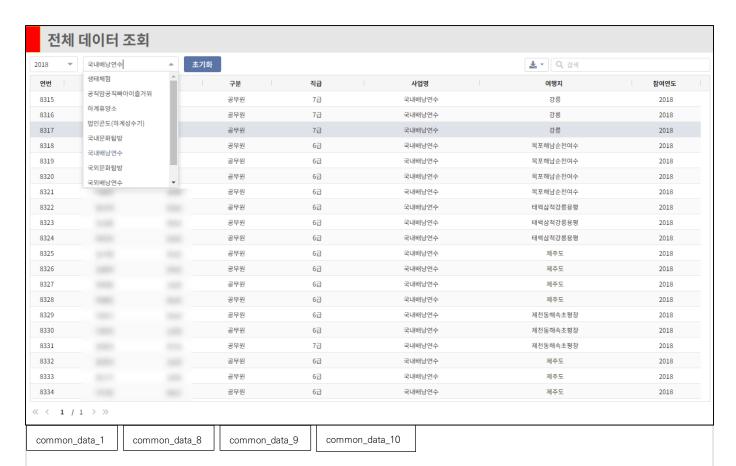


- 후생복지자원 신규 참여 희망자를 참여 희망자 명단 테이블로 시각화
- 반드시 참여 희망자 데이터를 등록해야 하며, 기존 참여 희망자 데이터가 존재하고 이를 갱신하는 경우 기존 참여자 명단과 칼럼 개수가 같아야 등록 가능
- 사용자가 참여 희망자 데이터 지표 등록 후 '데이터 로드' 버튼을 클릭하면 최신 데이터가 참여 희망자 명단 테이블에 표출
- 동명이인을 구분하기 위해 참여 희망자 이름과 생일을 합쳐 ID 칼럼이 자동으로 생성
- 데이터 로드 후 화살표 버튼을 클릭하면 기존 참여 데이터를 참조하여 검색 결과 테이블이 표출
- 각 ID가 이전에 참여했던 목록이 나타나며, ID 하나의 결과값이 여러 개일 경우 모든 이력이 행으로 이어서 표시됨
- 이전에 참여 경험이 없는 ID의 경우 '처음 참여'로 표시되며 초록색으로 강조됨
- 우측 상단의 버튼을 통해 CSV 형식으로 필터링된 데이터 다운로드 가능

3. 유지관리 매뉴얼

전체 데이터 조회

common_data_1, 8, 9, 10



- 트리거 기능을 이용할 시 이벤트가 발생하지 않는 첫 화면에는 아무 화면도 뜨지 않기 때문에 이를 막기 위해서 전체 데이터를 표출하는 모델 (common_data_1)을 생성하여 자동 실행을 지정해야 함
- 생년월일을 concat, substr로 정제하여 개인 식별을 위한 4자리의 '생년월' 파생변수 추출
- 여행지를 기재하지 않은 경우 모두 미기재로 표시하여 사용자의 검색을 도움 (fillna TOUR_PLACE 미기재)
- 참여 연도가 4자리 숫자로, IRIS 적재 시 자동으로 INT로 판단하기 때문에 TEXT로 변환 필요 (typecast PARTICIPATION_YEAR text)
- 사용자가 콤보 박스에서 참여 연도를 선택하면 트리거 기능을 통해 테이블 갱신 (where PARTICIPATION_YEAR = '\${combo_4}')
- 세부 검색을 위해 사용자가 참여 연도 \${combo_4} 선택 후에 추가적으로 콤보 박스에서 사업 명을 선택하면 트리거 기능을 통해 테이블 갱신 (where BUSINESS_NAME = '\${combo_5}')
- 클릭 시 전체 연도 테이블을 표출하도록 트리거가 적용된 초기화 버튼을 통해 사용자가 검색 후 원본 테이블로 다시 돌아갈 수 있도록 조치하며, 콤보 박스(\${combo_4}) 또한 트리거 연결을 통해 버튼 클릭 시 사용자가 연도를 다시 선택할 수 있도록 조치

빠른 검색

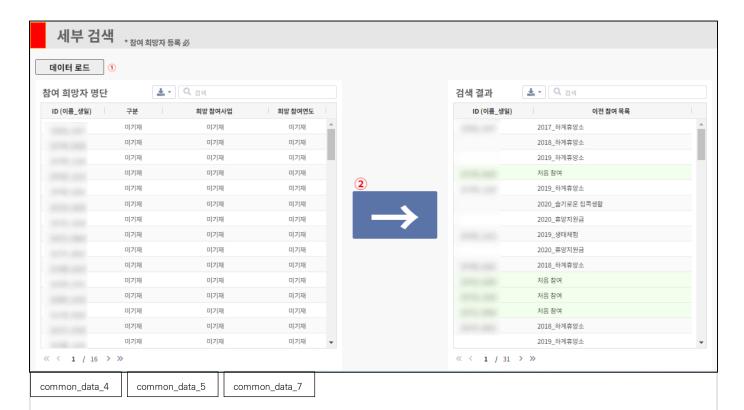
common_data_2, 3, 6



- 과거 5년 전체 참여자 이름을 오름차순으로 정렬하여 콤보 박스에 표시
- 사용자가 콤보 박스에서 참여자 이름을 선택하면 트리거 기능을 통해 테이블 갱신 (where NAME = '\${combo_2}')
- 동명이인에 대한 구별을 위해 사용자가 이름 \${combo_4} 선택 후에 추가적으로 콤보박스에서 생일을 선택하면 트리거 기능을 통해 테이블 갱신 (where BIRTHDAY ='\${combo_3}')

세부 검색

common_data_4, 5, 7



- 데이터 로드 버튼에 데이터 모델을 연동하여 사용자 클릭 시 캐시 초기화 구현 (* | command cache clear)
- 버튼을 누른다고 실제로 데이터가 업로드 되는 것은 아니며, 사용자가 참여 희망자 지표 등록을 하거나 기존 데이터를 갱신 할 시 프로세스가 바로 반영되지 않아 발생할 수 있는 데이터 로드 오류를 막기 위함
- 이름과 파생 변수 '생년'을 합쳐 개인을 식별한 수 있는 ID 칼럼 생성 (concat NAME, "_", BIRTH as ID)
- 사용자가 화살표 버튼을 클릭하면, 후생 참여 희망자 데이터와 기존 참여자 데이터에서 ID를 추출한 데이터 모델 뷰 '후생 참여자 ID'를 조인 (join outer '후생 참여자 ID' 후생 참여 희망자.ID = 후생 참여자 ID.ID)
- ID 오름차순으로 정렬한 뒤 SQL을 사용해 같은 ID를 하나의 Rank로 묶는 DIS 칼럼을 생성 (sql "select*, rank() over(order by ID) as DIS from angora")
- ID 오름차순으로 정렬된 기존 테이블에서 INDEX 순서로 숫자를 붙이는 NUM 칼럼을 생성 (numbering NUM)
- 같은 사람이 이전에 여러 번 참여 기록이 있을 경우 행에 연속해서 이름이 표출되기 때문에 이를 막기위해 NUM과 ID 별 랭크를 갖고 있는 DIS 칼럼으로 같은 이름을 판별 후 값을 공백으로 채우는 result 칼럼 생성 (case when NUM = DIS then ' ')
- 해당 result 칼럼은 NUM=DIS면 ''라는 공백 문자가 들어가 있으며, NUM과 DIS가 같지 않을 때는 Null 값을 가짐
- 위와 같은 조치는 IRIS DSL에서 concat이 칼럼에 Null을 포함하면 합쳐진 결과값이 Null로 나오기 때문에 NUM=DIS의 유효한 값 조차 사라지는 것을 막기 위함
- ID와 result를 합쳐 여러 값이 있을 경우 이름이 중복 표출되지 않고 공백으로 비워두는 RID 칼럼 생성 (concat ID, result as RID)
- 이전 참여 목록이 공백으로 존재하는 ID의 경우 이전에 참여 기록이 없는 ID이기 때문에 표출 후 테이블 시각화 기능으로 강조 표시 (fillna concated '처음 참여', RID'')

수원시 후생복지 분석 (추가 내용)

개요

개발 기간

• 2021.06 ~ 2021.07

활용 기술

• Spark SQL, IRIS DSL, IRIS Studio

활용 데이터

- 연도 별 후생복지 참여자 (2016~)
- 수원시 직원 인원 현황

설명_

목적

• 수원시 후생복지 참여자 현황을 사업별, 직류별로 확인할 수 있는 대시보드 필요

상세 내용

- 기존의 후생 참여 희망자 조회 페이지에서는 참여 현황에 대한 정보를 확인 할 수 없었음
- 관련 부서 요청에 의해 2016년부터 현재까지의 후생 복지 참여 정보를 시각화
- 기존 조회 페이지와 현황 페이지 연동으로 사용자 편의성 증대

목차

- 1. 화면 구성
- Page.09~12
- 2. 사용자 매뉴얼
- Page.13~16
- 3. 유지관리 매뉴얼
- Page.17~20

1. 화면 구성

화면 1

• 후생복지 사업현황

조회 페이지 수원시 후생복지 분석 후생복지 사업현황 (단위 : 명) 연도 총 인원 참가자 총계 생태체험 국내문화탐방 국외문화탐방 국내배낭연수 국외배낭연수 공직맘공… 휴양지원 법인콘도… 기타프로그램 0 33 40 521 38 (힐링프로그램) 2016 4,143 1053 (25.4%) 155 120 5 141 4,345 1550 (35.7%) 468 148 13 33 36 580 112 62 (힐링프로그램) 2017 98 59 (힐링프로그램) 2018 4,744 1956 (41.2%) 642 150 99 20 40 50 796 100 2019 5,273 1936 (36.7%) 110 20 (힐링프로그램) 2185 (39.8%) 709 (슬기로운 집콕생활) 2020 5,491 0 0 0 0 35 1,280 161 1182 (21.2%) 0 0 0 1,120 62 (힐링하우스) 5,587 0 0 0 2021 • 연도별 참여현황 ◆ 연간 참여자 ● 누적 참여자 10,000 2000 6,495 4,559 5000 1500 2500 1,053 1,05 1000 2017 2018 2019 2016 2020 2021 2016 2017 2018 2019 2020 2021

화면 2

• 분류별 후생복지 참여 현황



화면 3

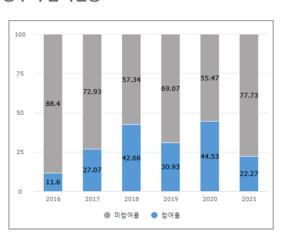
• 후생복지 참여 상세 현황

후생복지 참여 상세 현황

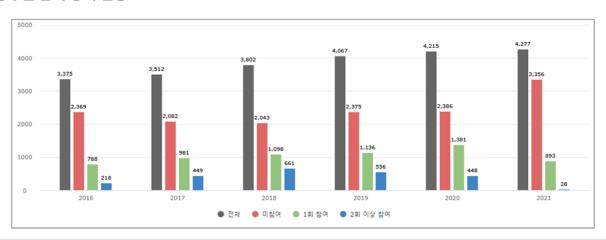
• 공무원 참여현황



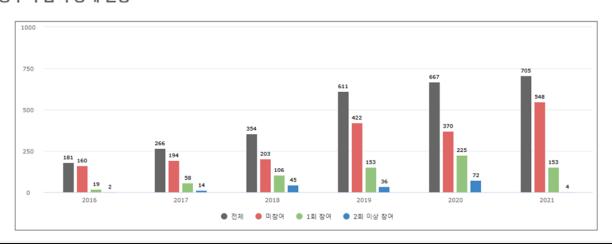
• 공무직 참여현황



• 공무원 참여 상세 현황



• 공무직 참여 상세 현황



화면 4

• 사업별 후생복지 참여 현황



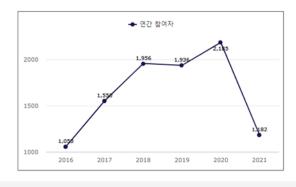
2. 사용자 매뉴얼

후생복지 사업 현황

- 2016년 ~ 2021년 후생복지 참여 데이터
- 수원시 직원 현황 데이터

| 조회 페0 | 기지 | | 수원시 후생복지 분석 | | | | | | | | |
|---------------------|-------|--------------|-------------|--------|--------|--------|--------|-------|-------|-------|-----------------|
| + | ᄱᆸᆉ | 미어취취 | | | | | | | | | |
| 후생복지 사업현황 (단위:명) | | | | | | | | | | | |
| 연도 | 총 인원 | 참가자 총계 | 생태체험 | 국내문화탐방 | 국외문화탐방 | 국내배낭연수 | 국외배낭연수 | 공직맘공… | 휴양지원 | 법인콘도… | 기타프로그램 |
| 2016 | 4,143 | 1053 (25.4%) | 0 | 155 | 120 | 5 | 33 | 40 | 521 | 141 | 38 (힐링프로그램) |
| 2017 | 4,345 | 1550 (35.7%) | 468 | 148 | 98 | 13 | 33 | 36 | 580 | 112 | 62 (힐링프로그램) |
| 2018 | 4,744 | 1956 (41.2%) | 642 | 150 | 99 | 20 | 40 | 50 | 796 | 100 | 59 (힐링프로그램) |
| 2019 | 5,273 | 1936 (36.7%) | 654 | 153 | 99 | 10 | 38 | 49 | 803 | 110 | 20 (힐링프로그램) |
| 2020 | 5,491 | 2185 (39.8%) | 0 | 0 | 0 | 0 | 0 | 35 | 1,280 | 161 | 709 (슬기로운 집콕생활) |
| 2021 | 5,587 | 1182 (21.2%) | 0 | 0 | 0 | 0 | 0 | 0 | 1,120 | 0 | 62 (힐링하우스) |
| | | | | | | | | | | | |

• 연도별 참여현황





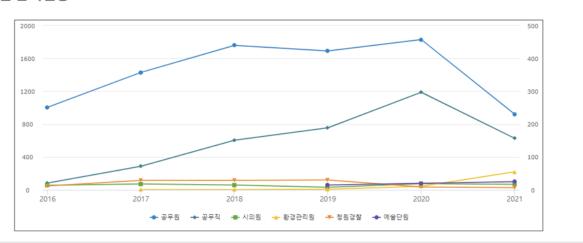
- 2016년부터 최근까지 참여한 후생복지자원 현황을 테이블로 시각화
- 연도 오름차순 기준으로 전체 프로그램을 표출하는 테이블 구성
- 수원시 직원 인원 현황 데이터를 통해 총 인원을 파악하여 참가자 통계 칼럼 도출
- 각 연도별 연간 참여자를 꺾은선 그래프로, 누적 참여자를 영역 그래프로 표출
- 좌측 상단의 버튼을 통해 후생복지 참여자 조회 페이지로 이동할 수 있음

분류별 후생복지 참여 현황

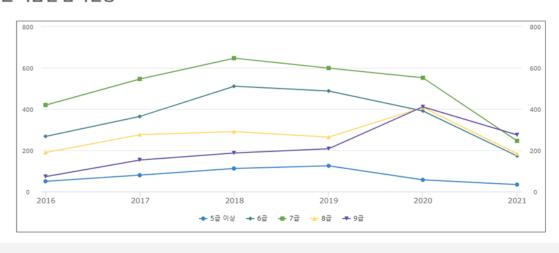
• 2016년 ~ 2021년 후생복지 참여 데이터

분류별 후생복지 참여 현황

• 직류별 참여현황



• 공무원 직급별 참여현황



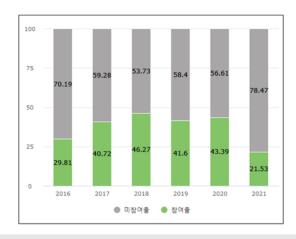
- 총 6개의 직류로 구분된 참여 현황을 꺾은선 그래프로 표출
- 타 직류에 비해 공무원 직류의 인원 수가 많기 때문에 좌우 Y축의 단위를 다르게 표현
- 공무원 직급별 참여 현황은 인원 수가 적은 5급 이상을 합쳐 하나의 꺾은선으로 표현
- 마우스를 그래프 위에 올려 각 연도의 정확한 참여 인원 수를 팝업으로 확인 가능

후생복지 참여 상세 현황

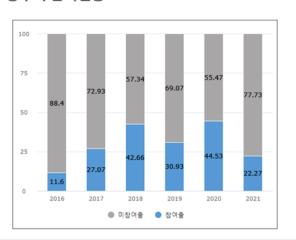
- 2016년 ~ 2021년 후생복지 참여 데이터
- 수원시 직원 현황 데이터

후생복지 참여 상세 현황

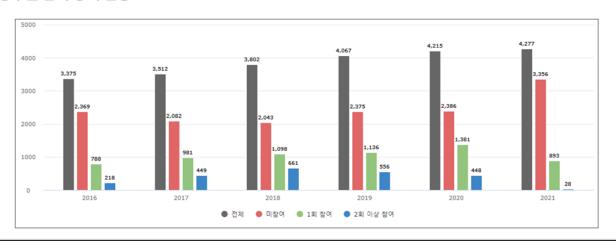
• 공무원 참여현황



• 공무직 참여현황



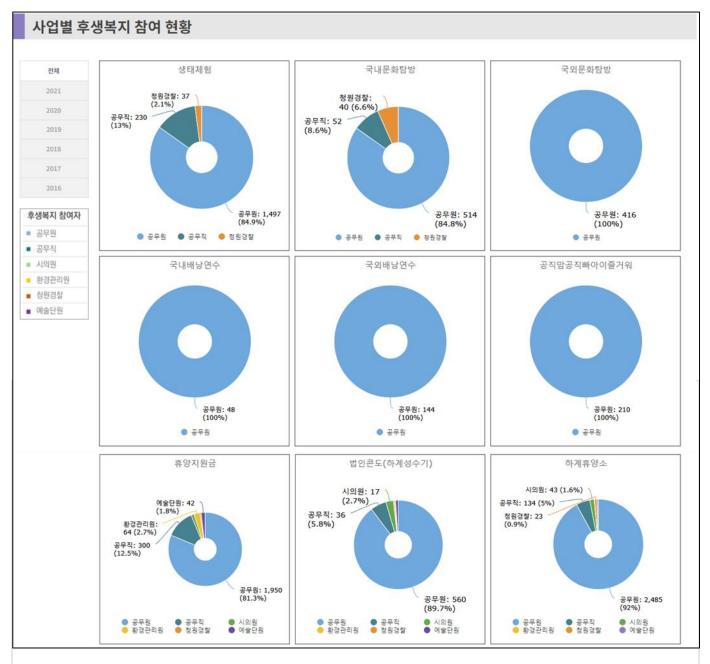
• 공무원 참여 상세 현황



- 공무원, 공무직 후생복지 참여 상세 현황을 연도별 세로 막대 그래프로 표출
- 막대그래프의 스택 기능을 사용해 미참여율과 참여율을 시각적으로 구성
- 이름, 생년, 직급을 합친 하나의 ID로 개인을 구별하고, 각 연도 참여 현황 데이터에서 조회하여 연도별 참여 횟수를 표출
- 참여 횟수가 2회를 초과하는 참여자는 거의 없기 때문에 현재는 2회 이상으로 통합하였고, 추후 대시보드 유지관리 시에 데이터 분포에 따라 세분화할 필요가 생길 수 있음

세부 검색

2016년 ~ 2021년 후생복지 참여 데이터



- 사업별 후생복지 참여 현황을 참여자 분류로 나누어 연도별 파이 그래프로 시각화
- 각 연도는 최신순으로 탭으로 구분하여 해당 연도에 시행했던 사업만 표출되고, 전체 탭은 각 연도 참여 현황을 합산하여 표출하였음
- 기존 사업에 대해서는 시간 흐름에 따라 연도가 넘어가거나, 추가적으로 데이터 등록할 시 변경된 데이터를 전체 탭에서 정상적으로 그래프 표출 가능
- 기타 프로그램과 같이 각 해마다 사업 자체가 추가되거나, 다음 연도를 탭으로 추가하여 표출할 시에는 탭 생성, 파이 그래프 생성과 같은 유지보수가 필요함

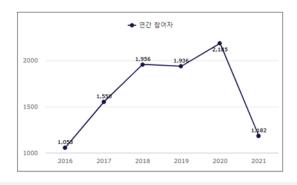
3. 유지관리 매뉴얼

후생복지 사업 현황

• common_data_1, 2, 3

조회 페이지 수원시 후생복지 분석 하면시티 수원 후생복지 사업현황 (단위 : 명) 휴양지원 법인콘도… 기타프로그램 연도 총 인원 참가자 총계 생태체험 국내문화탐방 국외문화탐방 국내배낭연수 국외배낭연수 공직맘공… 2016 4.143 1053 (25.4%) 0 155 120 5 33 40 521 141 38 (힐링프로그램) 2017 4,345 1550 (35.7%) 468 148 98 13 33 36 580 112 62 (힐링프로그램) 2018 4,744 1956 (41.2%) 150 50 100 59 (힐링프로그램) 2019 5,273 1936 (36.7%) 20 (힐링프로그램) 709 (슬기로운 집콕생활) 2020 5,491 2185 (39.8%) 0 0 0 35 1,280 161 2021 5.587 1182 (21.2%) 0 0 0 0 0 1.120 0 62 (힐링하우스)

• 연도별 참여현황





common data 1

common_data_2

common data 3

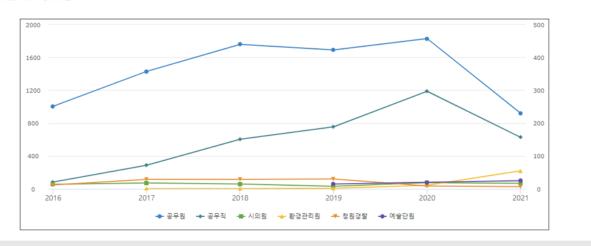
- 힐링 프로그램, 슬기로운 집콕 생활 등 연도마다 변경되는 프로그램들을 기타 프로그램으로 통합
- 사업명, 연도 기준으로 집계 후 피벗 함수를 통해 칼럼으로 변경 (pivot max(cnt) splitrow YEAR splitcol BUSINESS sortrow ASC)
- 모델의 연산을 줄이기 위해 데이터 모델 뷰로 각 연도 총계를 계산한 뒤 수원시 직원 현황 데이터와 조인하여 참가 비율 생성
- 생성한 데이터 모델 뷰를 기존 값과 연도 기준으로 결합 (join inner '후생복지 총계' as " 연도 = .YEAR)
- 참여 인원을 연도 기준으로 집계 하여 연도별 참여 현황 생성 (stats count(*) as '연간 참여자' by PARTICIPATION YEAR)
- 연도별 누적 참여자 계산 수행 (sql "select*, sum(CNT) over(order by PARTICIPATION_YEAR) as '누적 참여자' from angora")

분류별 후생복지 참여 현황

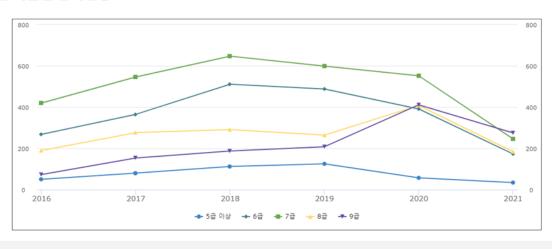
• common_data_4, 5

분류별 후생복지 참여 현황

• 직류별 참여현황



• 공무원 직급별 참여현황



common_data_4

common_data_5

- 직류 (Type)별, 연도 별 집계를 통해 꺾은선 그래프를 도출 (stats count(*) as CNT by PARTICIPATION_YEAR, type)
- 각 직류 꺾은선에 다른 색을 지정하기 위해, 다중축 그래프로 설정한 뒤 CASE 조건으로 직류 별로 칼럼 분할
- 직급 (Class)별 집계를 통해 꺾은선 그래프를 도출 (stats count(*) as CNT by PARTICIPATION_YEAR, CLASS_OF_POSITION
- 각 직급 꺾은선에 다른 색을 지정하기 위해, 다중축 그래프로 설정한 뒤 CASE 조건으로 직급 별로 칼럼 분할

후생복지 참여 상세 현황

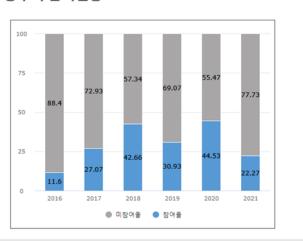
common_data_6, 7, 8, 9

후생복지 참여 상세 현황

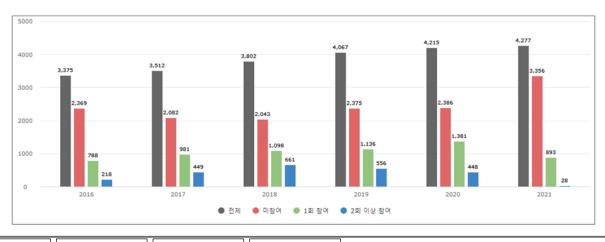
• 공무원 참여현황



• 공무직 참여현황



• 공무원 참여 상세 현황



common_data_6

common_data_7

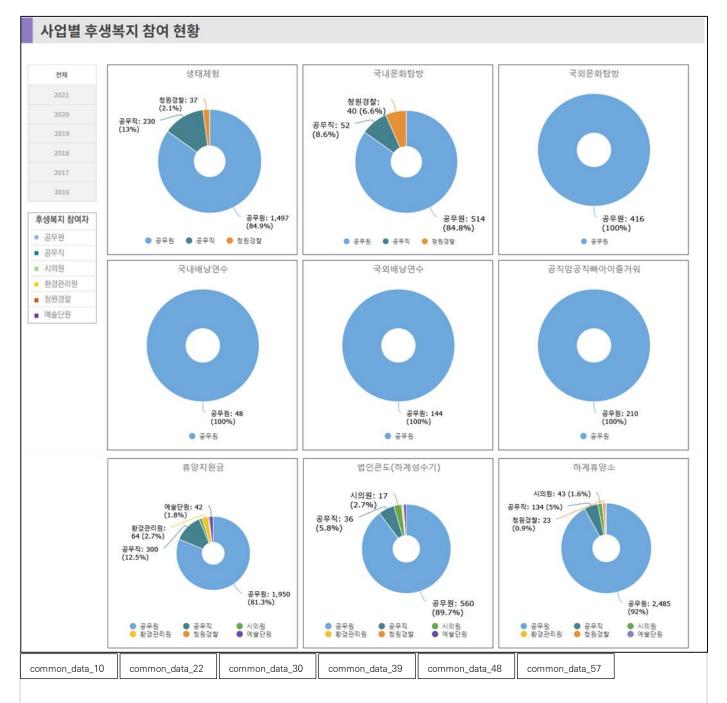
common_data_8

common_data_9

- 공무원, 공무직으로 필터링 후 연도 기준 집계 계산 (stats count(*) as 참여 by PARTICIPATION_YEAR)
- 수원시 직원 현황 데이터와 조인하여 공무원, 공무직의 참여율과 미참여율 계산
- 성명과 생년으로 ID를 추출하고 집계 (concat NAME, "_", BIRTH as ID | stats count(*) as CNT by ID, PARTICIPATION_YEAR)
- 연도, ID별 참여 횟수 추출(sql "select DISTINCT PARTICIPATION_YEAR, max('1회 참여') over(partition by PARTICIPATION_YEAR) as '1회 참여', max('2회 이상 참여') over(partition by PARTICIPATION_YEAR) as '2회 이상 참여' from angora)"
- 수원시 직원 현황 데이터와 조인하여 전체 인원을 추출하고, 전체 참여와 미참여 인원 수를 추출 (calculate '1회 참여' + '2회 이상 참여' as '전체 참여 | calculate 'NUMBER_OF_PERSON_COUNT' '전체 참여' as 미참여)

사업별 후생복지 참여 현황

common_data_10 ~ 64



- 전체 연도와 16년~21년을 각각 탭으로 구성하고, 사업별 파이 차트를 현황 테이블과 같은 순서로 배치
- 파이 차트 내 구분과 색상은 협의 기안의 범례와 같이 구성하고, 순서 설정을 위해 CASE문으로 직류에 따라 순서 생성 후 정렬 시행
- 모델의 성능 개선을 위해 집계를 수행한 '후생복지 사업별' 데이터 모델 뷰를 생성 (stats count(*) as CNT by PARTICIPATION_YEAR, BUSINESS_NAME, TYPE)
- 생성된 데이터 모델 뷰에 WHERE 조건을 이용한 필터링으로 각 연도, 사업별 데이터 추출하여 파이 그래프에 전달