

-환경데이터 분석,AI 전문가-

환경정보 융합 빅데이터 플랫폼 데이터셋 분석하기







목 차

시스템	개요		

나라애모기	1
1. 시작하기 및 데이터 연결	1
2. 작업 창 설명	
3. 뷰 작성	-3
4. 뷰 정렬	6
5. 마크 기능사용	7
6. 새 시트 및 복제	9
7. 표현 방식 사용 1	0
8. 사용자 정의 계산식 작성 1	1
9. 집계 방식 변경 1	3
10. 이중축 사용 1	4
11. 시트에서 데이터 다운로드1	6
12. 대시보드 구성1	8
12. 대시보드 구성1 13. 대시보드에서 데이터 다운로드 2	2





시스템 개요

- Tableau는 프로그래밍 과정을 최소화하여 자신의 데이터를 다양한 방식으로 시각화 할 수 있는 도구로, 시각화를 통하여 사용자가 데이터를 쉽고 직관적으로 확인하고 이해할 수 있도록 지원하는 시각화 분석 도구입니다.
- 빅데이터 분석 탭에서 분석도구 태블로를 클릭합니다.



• 하단으로 스크롤 하여 태블로 이용하기를 클릭해 분석 화면으로 들어갑니다.

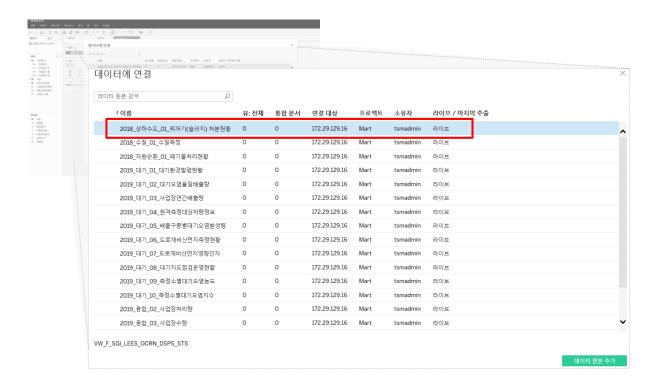


1 시작하기 및 데이터 연결

- Tableau 분석 페이지에 접속하면 처음 표시되는 페이지가 아래 이미지와 유사합니다.
- 만들기 버튼을 클릭합니다
- 통합문서 버튼을 클릭합니다.

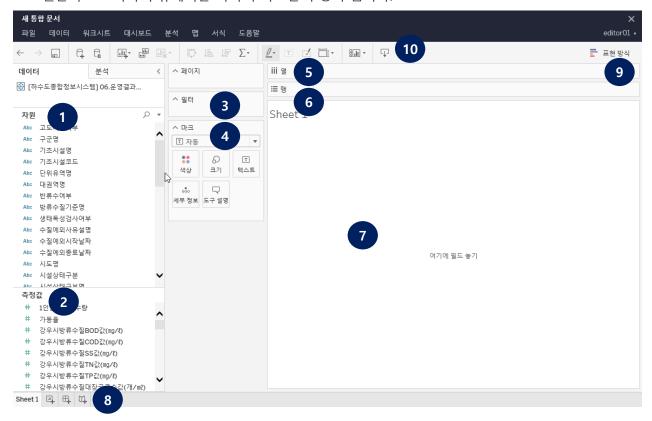


- 데이터에 연결 창이 나타납니다
- 분석하고자하는 데이터셋을 클릭하여 선택(2018 상하수도 01 찌꺼기(슬러지)처분현황)
- 하단의 데이터원본 추가를 클릭.



2 작업 창 설명

- 데이터에 연결하면 분석 할 수 있습니다. Tableau에서는 시트 각각을 뷰라고 합니다.
- 일반적으로 하나의 뷰에서는 하나의 차트만 구성이 됩니다.



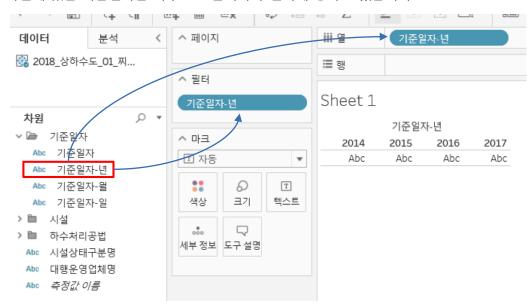
- ① 차원: 일반적으로 날짜, 구분, 업체번호 등의 범주형 데이터를 말합니다.
- ② 측정값: 처리량,발생량, 건수 등 수치 값으로 표현되는 숫자형 데이터를 말합니다.
- ③ 필터: 차원이나 측정값을 끌어다 놓아 뷰의 데이터를 필터링 할 수 있습니다.
- ④ 마크: 뷰의 마크 유형을 선택하거나 색상, 크기, 텍스트를 정의합니다.
- ⑤ 행: 차원이나 측정값을 끌어다 놓아 뷰의 행으로 지정합니다.
- ⑥ 열: 차원이나 측정값을 끌어다 놓아 뷰의 열로 지정합니다.
- ⑦ 뷰 : 시각화 분석의 결과 입니다. 행 과 열 마크에 의해 형태가 결정됩니다.
- ⑧ 시트 : 하나의 뷰를 시트라고도 부릅니다. 시트 하나에는 하나의 뷰가 포함됩니다.
- ⑨ 표현방식: 태블로에서 가장 자주 사용되는 차트 형태를 퀵메뉴로 구성되어있습니다
- ⑩ 다운로드: 현재 보이는 뷰를 이미지나 PDF 또는 데이터를 csv로 내려받을 수 있습니다.

Tip!

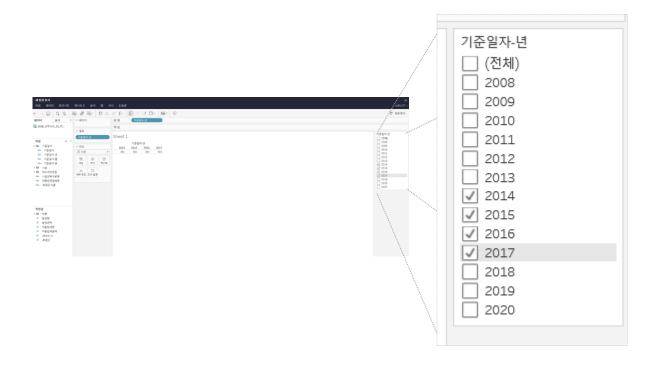
- ❖ 작업중 Undo 기능이 지원됩니다. CtI+Z 키로 실행 취소를 할 수 있습니다.
- ❖ 효율적인 작업을 위해서 보고자하는 데이터를 필터링하여 데이터 양을 줄이는것이 유효합니다.

3 뷰 작성 – 필터 추가

• 차원에 있는 기준일자를 마우스로 끌어다가 필터에 놓아 보겠습니다.



- 우측에 필터 선택기가 생성 됩니다.
- 선택기를 통해 기준년도를 선택할 수 있습니다.

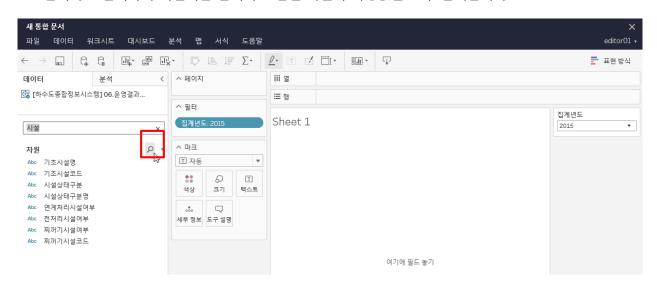


3 뷰 작성

- Tableau는 거의 대부분의 개체에 마우스를 가져다 대면 ▼ 형태의 확장메뉴 버튼이 나타납니다.
- 방금 추가한 우측의 집계년도 선택기에 마우스를 가져가 ▼ 메뉴를 눌러 확장 메뉴를 열고 단일값(드롭다운)을 클릭하여 선택기를 드롭다운 형태로 변경합니다.

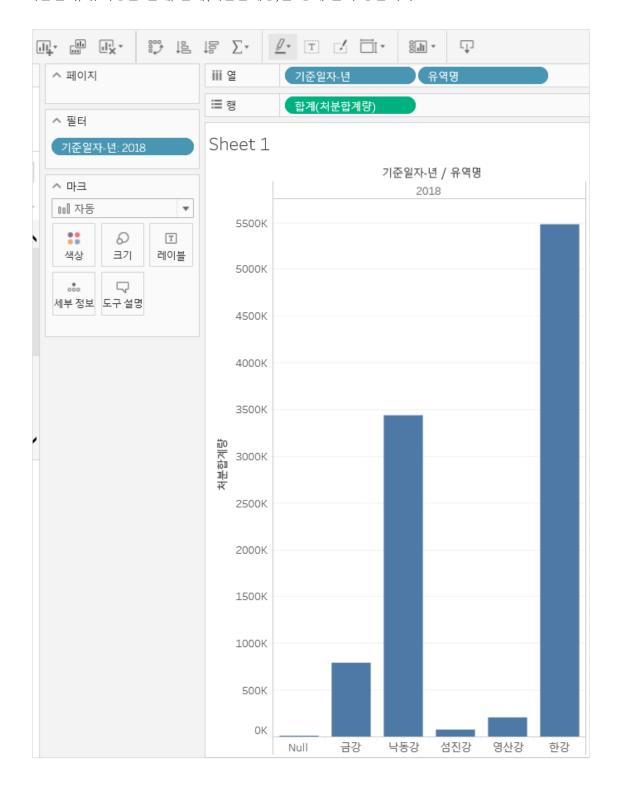


• 차원 옆 쪽에 있는 돋보기 버튼을 클릭하면 검색창이 나타납니다. 이 검색창에 '시설' 이란 단어를 입력하고 검색하여 시설이란 단어가 포함된 차원과 측정값을 모두 검색합니다.



3 뷰 작성

• 기준일자, 유역명을 열에, 합계(처분합계량)을 행에 끌어 놓습니다.

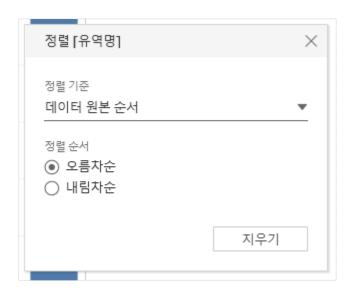


4 뷰 정렬

● ▼ 표시에 클릭하면 나오는 확장메뉴에서

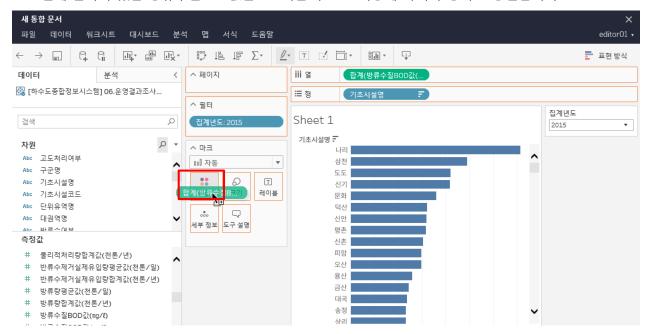


• 정렬 방식 변경 가능

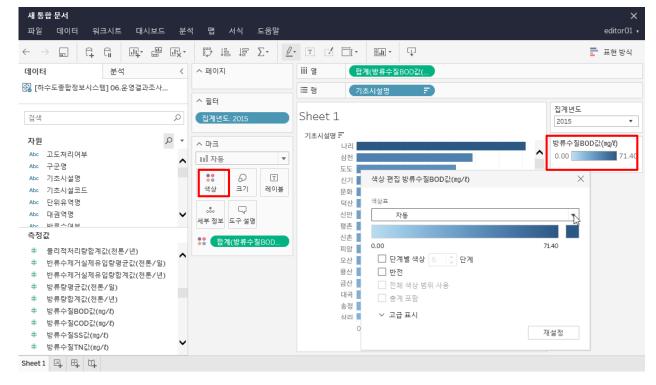


5 마크 기능 사용

- 방류수질BOD값을 측정값에서 끌어다가 마크에 있는 색상에 가져다 놓습니다.
- 방류수질BOD값을 기준으로 막대의차트의 색이 정해집니다.
- 열에 올라가 있는 방류수질BOD값을 Ctrl키를 누르고 색상에 가져다 놓아도 동일합니다.

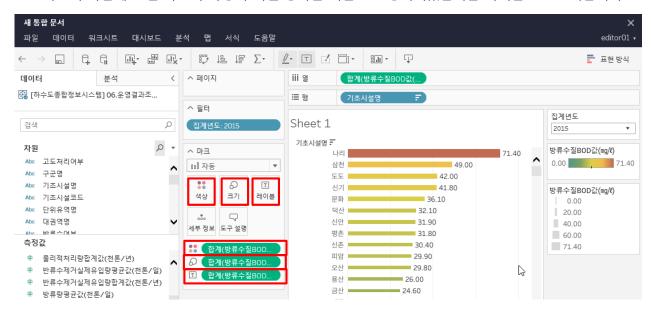


 범례에 마우스를 가져가 확장 버튼을 클릭하거나, 마크의 색상을 클릭하여 색상 편집을 눌러 색상을 변경 할 수 있습니다.

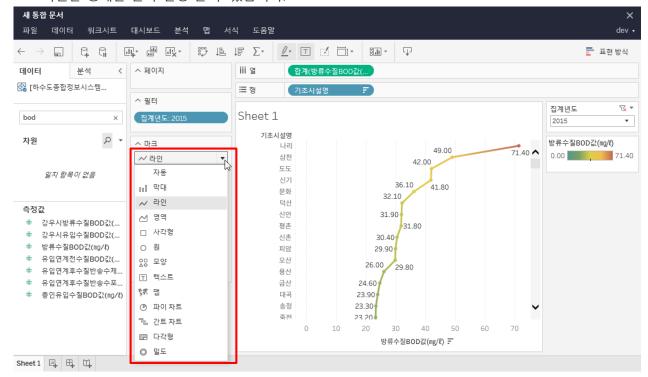


5 마크 기능 사용

- 마크에서 크기의 경우 차트의 크기를 결정합니다. 마찬가지로 현재 차트에 수치를 표시하고자 한다면, 레이블에 원하는 항목을 끌어다 놓으면 됩니다.
- 색상과 마찬가지로 크기와 레이블에 방류수질BOD값을 끌어다 놓았습니다.
- 마크의 하단에 보면 마크의 속성이 어떤 항목을 기준으로 정의되었는지를 아이콘으로 보여줍니다.

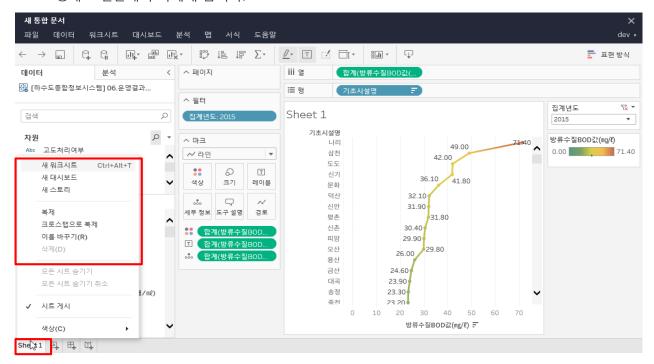


 태블로에서 기본적으로 차트의 형태는 자동으로 골라주지만 원하는 형태가 있다면, 마크에서 적절한 형태를 골라 변경 할 수 있습니다.

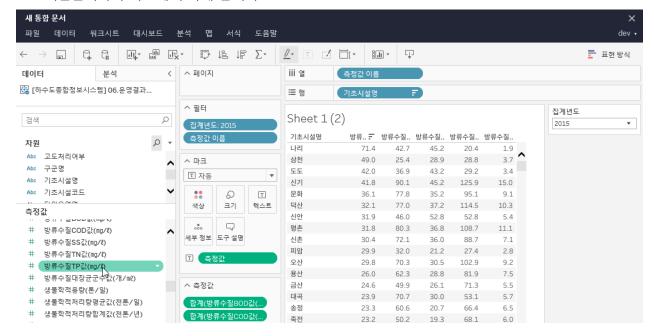


6 새 시트 및 복제

 하단에 현재 작업중이 sheet 1을 우클릭을 해보면 새워크시트나 새 대시보드를 열수 있습니다.
 복제를 누르면 현재 시트는 보존한 채로 새로 복제 됩니다. 크로스탭으로 복제는 현재 시트를 표형태로 변환해서 복제해 줍니다.



 크로스탭으로 복제를 눌러 표형태로 다시 돌아왔습니다. 이번엔 방류수질 COD, SS, TN, TP 를 더블클릭하여 차트에 추가해 줍니다.

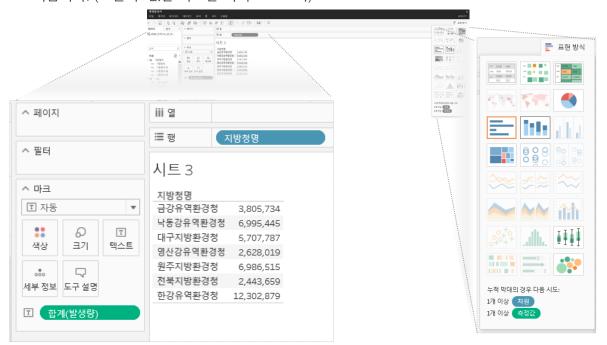


7 표현 방식 사용

 하단에 현재 작업중인 각각의 시트명칭을 더블클릭하여 변경합니다. 시트를 우클릭하여 이름바꾸기를 선택해도 동일합니다.

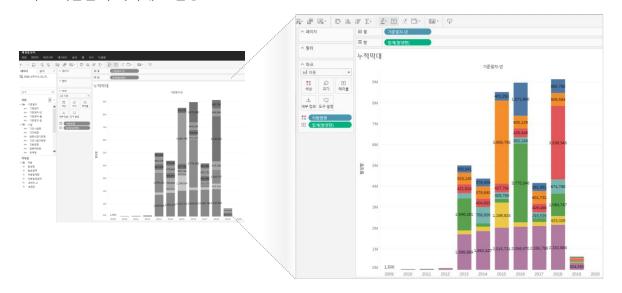


- 새로운 시트를 열고 지방청명과 합계를 더블 클릭으로 추가합니다.
- 우측 상단에 있는 표현방식을 클릭하여 누적마대 생성
- 표현방식에서 원하는 형태에 마우스를 가져다 대면, 해당 차트를 그리기 위한 데이터가 무엇인지 보여줍니다. (그릴 수 없는 차트는 회색으로 표시)

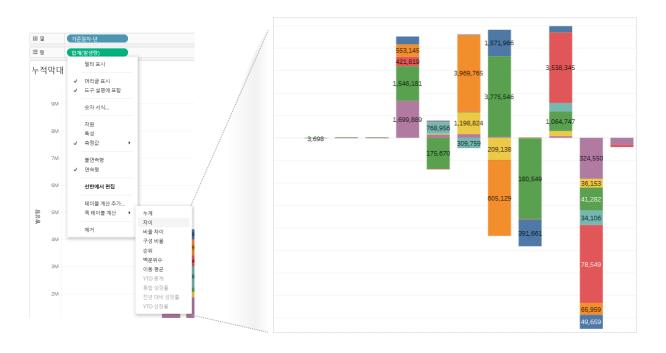


8 사용자 정의 계산식 작성

- 생성된 차트에 열에 집계 년도를 끌어다 놓아 년도별로 차트를 생성
- 측정값의 마크 레이블 설정
- 시트 이름을 누적막대로 변경



• 상단의 분석 탭에서 계산된필드 만들기를 선택하여 계산식 필드를 추가합니다.

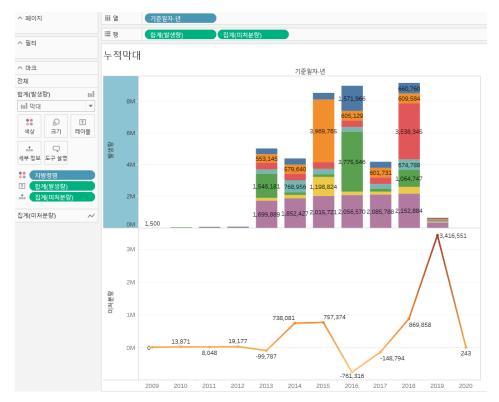


8 사용자 정의 계산식 작성

- 계산이 유효한지 확인 후 이름을 알아보기 쉽게 적어줍니다.
- 계산식 창에서 ▶ 표시를 눌러 태블로에서 사용할 수 있는 함수 목록을 볼 수 있습니다.

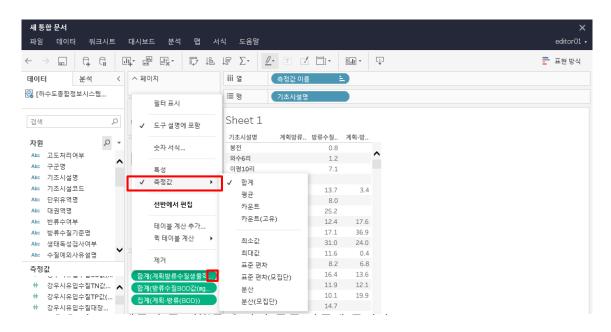


- ※ 각 항목을 SUM(합계)으로 계산 후 빼는 이유는 SUM을 하지 않으면 Low레벨에서 계산이 이루어진 후 사용자한테 보여줄 때 집계 되어 보여집니다. 이때 시스템상 null이 있을 경우 계산이 정상적으로 되지 않기 때문에 계산식을 사용할 땐 SUM등의 집계 함수를 자주 활용합니다.
- 정상적으로 계산이 잘 되었는지 다시 한번 확인해 봅니다.

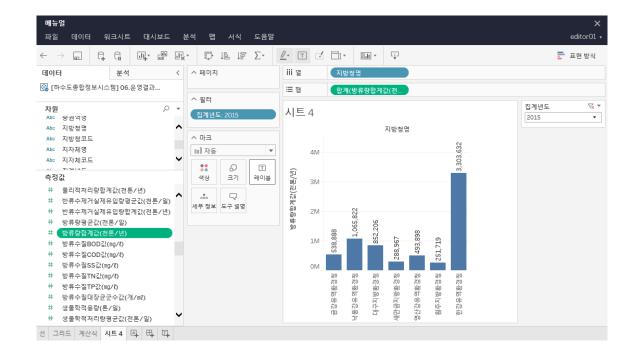


9 집계 방식 변경

- 태블로에서 측정값은 화면에 보여질때 집계방식이 기본은 SUM(합계)입니다.
- 만약 계획방류 수질의 집계방식을 SUM이 아닌 평균이나 최대값 등으로 변경하려면 해당 측정값에
 마우스를 가져대 ▼를 눌러 확장메뉴에서 측정값을 눌러 집계방식을 변경 할 수 있습니다.

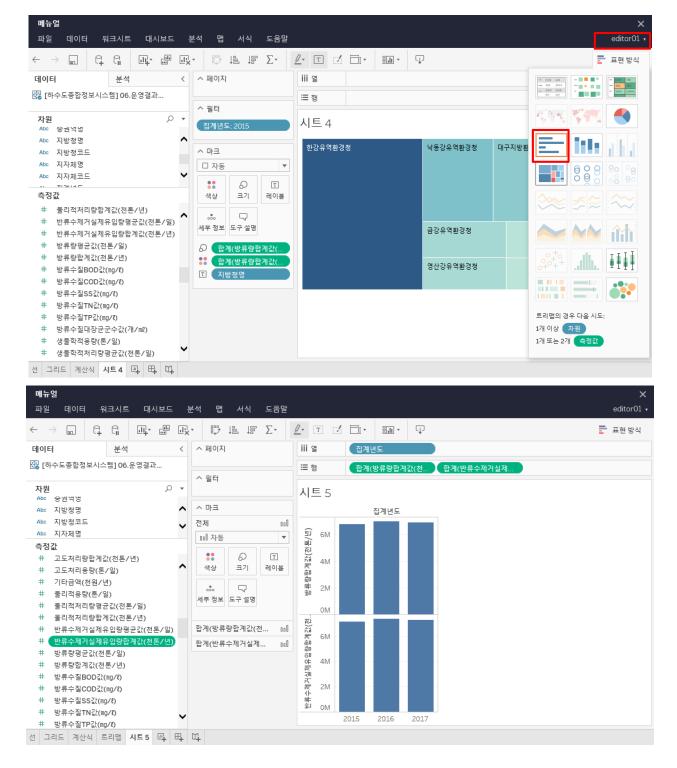


• 새로운 시트를 열어서 차원은 지방청명 ,측정값, 필터(집계년도)를 설정하고 레이블도 표시합니다.



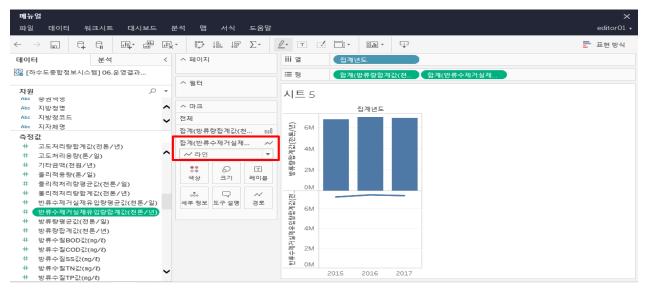
10 이중축 사용

- 우측 상단에 있는 표현방식을 클릭하여 트리맵을 선택해 줍니다.
- 차트 형태가 변경되면서 색과 영역의 넓이로 차트가 구성됩니다.

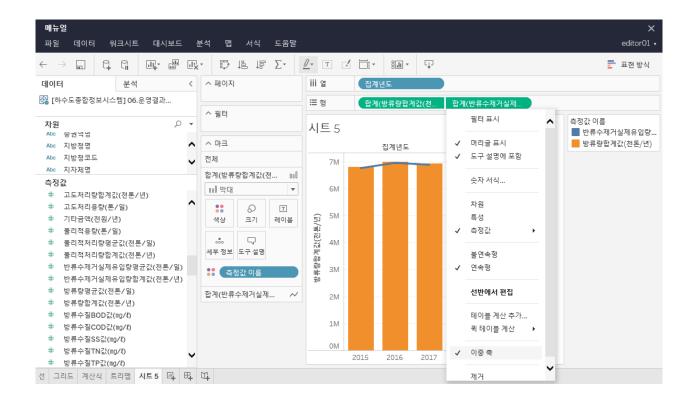


10 이중축 사용

- 마크를 보면 기존과는 다르게 전체 탭이 있고, 각각 에 해당하는 차트의 속성을 결정합니다.
- 탭을 클릭하여 자동 막대 대신에 라인으로 변경됩니다.

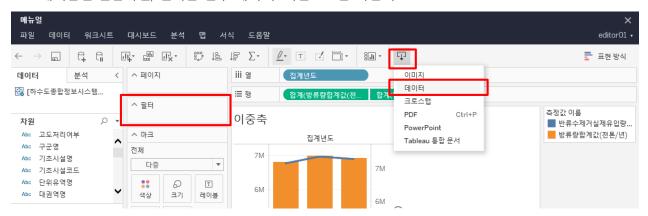


- 행에 마우스를 가져가 ▼ 버튼을 클릭 후 이중축 을 선택합니다.
- 막대에서 다른 형태로 변경이 되면 마크 자동대신에 막대로 변경합니다.

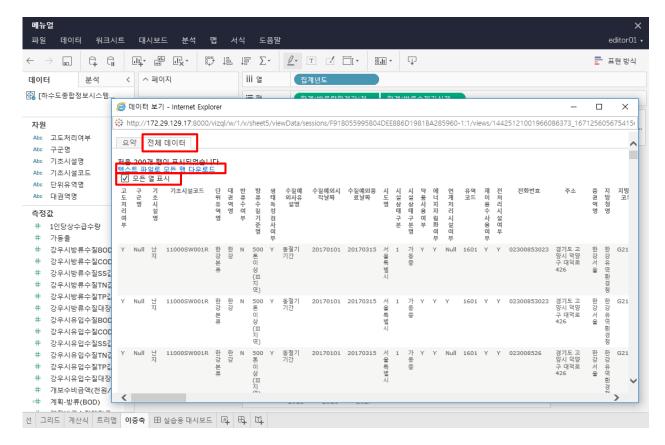


11) 시트에서 데이터 다운로드

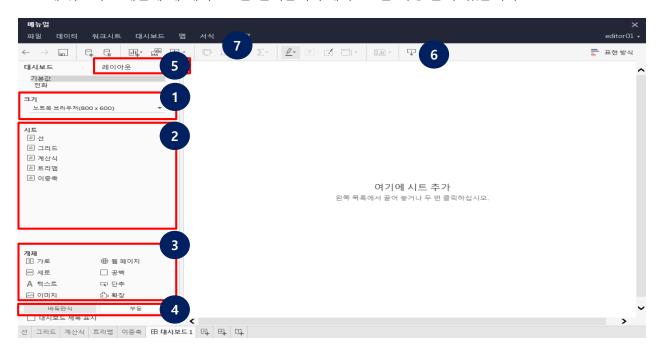
- 시트에서 상단에 다운로드 버튼을 클릭하여 데이터를 다운로드 받을 수 있습니다.
- 이때 필터에 특정 항목이 들어가 있다면, 필터링 된 항목만 다운로드가 됩니다. 만일 모든 데이터를 내려받길 원한다면, 필터를 전부 제거 후 다운로드를 하십시오.



- 처음에 데이터를 누르면 요약데이터를 보여줍니다. 전체 데이터 탭을 클릭 하면 현대 뷰에 구성되어있는 항목들 기준으로 데이터를 받을 수 있습니다.
- 모든열 표시를 체크 하여 데이터셋에 있는 모든 항목을 내려받을 수 있습니다. ※데이터 사이즈에 따라 시간이 많이 소요될 수 있습니다.



• 새 워크시트 대신에 새 대시보드를 선택합니다.대시보드를 작성 할 수 있습니다.

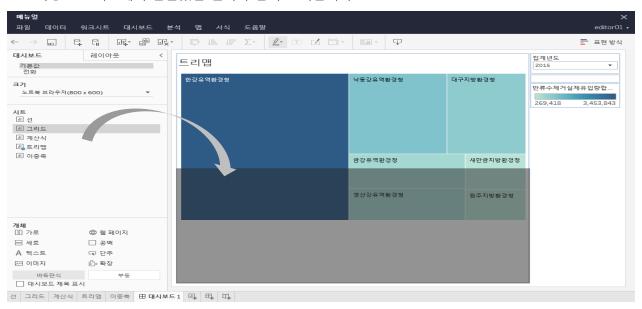


번호	구분	내용	
1	크기	대시보드의 크기결정(픽셀단위로 조정가능)	
2	시트	앞 단계에서 생성한 시트들을 표시	
3	개체	텍스트나 이미지를 넣거나 차트 구성을 가로 또는 세로로 할 수 있 는 컨테이너 박스가 있습니다.	
4	대시보드구성 방식	바둑판식 (가로,세로 개체들을 이용하여 마우스로 간단하게 끌어다 놓으면 사이즈를 태블로에서 조정하는 방식) 부동 방식 (각 사이즈를 픽셀 단위로 사용자가 지정하는 방식) 선택	
5	레이아웃	각 개체들의 테두리나 백그라운드 색, 위치나 너비, 높이 등을 지정	
6	다운로드	대시보드는 이미지나 PDF 등으로 저장	
7	서식	통합 문서 전체에 적용되는 서식을 정의	

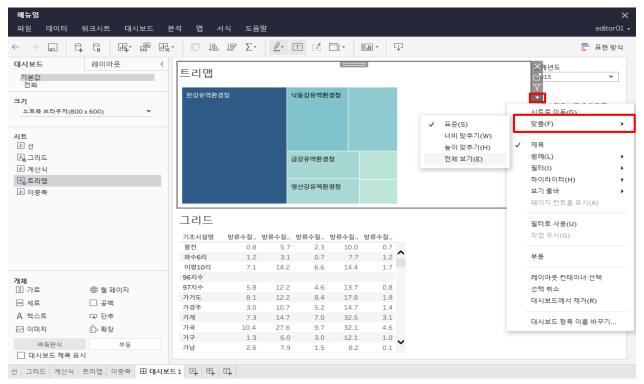
Tip!

- ◆ 크기를 자동이나 범위로 사용하면 화면에 따라 대시보드 크기가 변경이 되지만 안에 있는 차트들의 모양이 흐트러질 수도 있습니다
- ❖ 처음에는 부동으로 사용하여 위치와 사이즈를 전부 수작업으로 변경하는게 편할 수도 있습니다

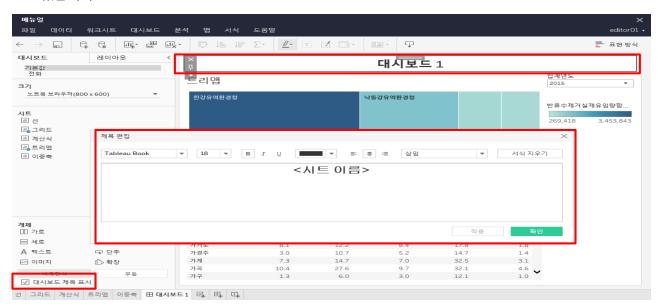
- 바둑판식으로 사용 할 경우(기본값) 원하는 시트를 좌측에서 끌어다가 원하는 위치에 가져다 대면 아래와 같이 어디에 들어가는지 표시가 됩니다.
- 자동으로 시트에서 만들었던 필터가 같이 표시됩니다



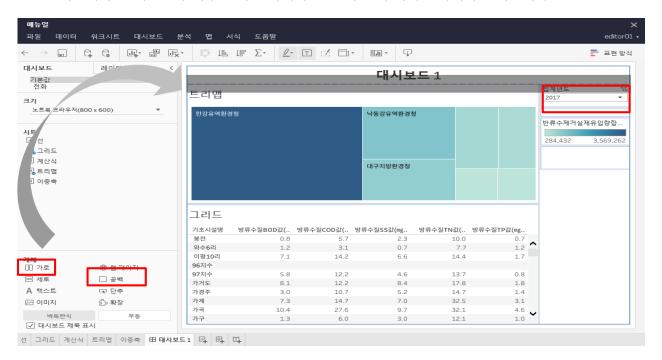
트리맵 차트를 화면에 가득 차게 사이즈를 변경하려면, 트리맵 차트를 클릭 후 ▼ 버튼을 눌러 맞춤
 옵션에서 너비 맞춤이나 전체 보기 등을 선택합니다.



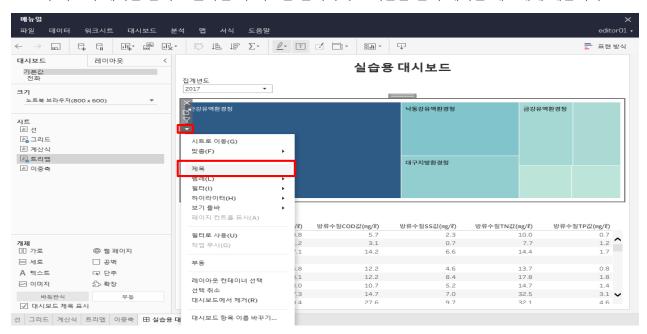
- 하단의 그리드는 아래로 스크롤이 길기 때문에 전체 보기 대신 너비 맞춤으로 해줍니다.
- 대시보드 제목 표시를 눌러 대시보드 제목을 표시해 주고 더블클릭 하여 편집을 합니다.
- 글자 크기나 색상 정렬 등을 변경 할 수 있고 시트 이름 대신 사용자가 새로 텍스트를 입력 할 수 있습니다.



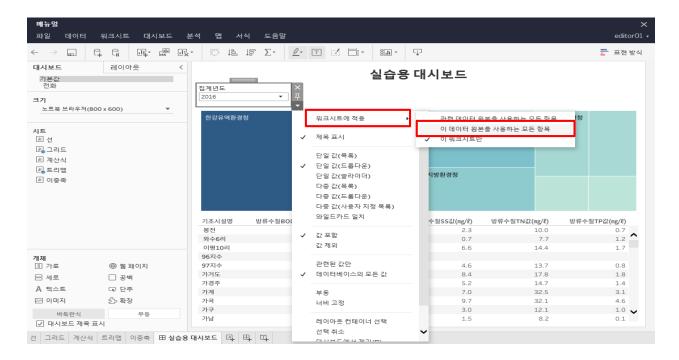
 먼저 개체에서 가로를 끌어다가 제목아래에 넣어주고 해당 가로 개체에 집계년도 필터를 끌어다가 넣어줍니다. 집계년도가 너무 길게 표시되니 그 옆에 공백 개체를 가져다 넣어줍니다.



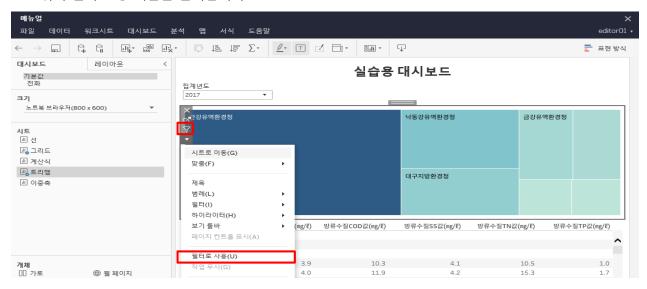
- 범례의 경우 필요 없으면 해당 항목을 선택하여 X 표시를 눌러 삭제합니다.
- 각 시트의 제목을 숨기고 싶다면 각 시트를 클릭하여 ▼버튼을 눌러 제목을 체크해제 해줍니다.



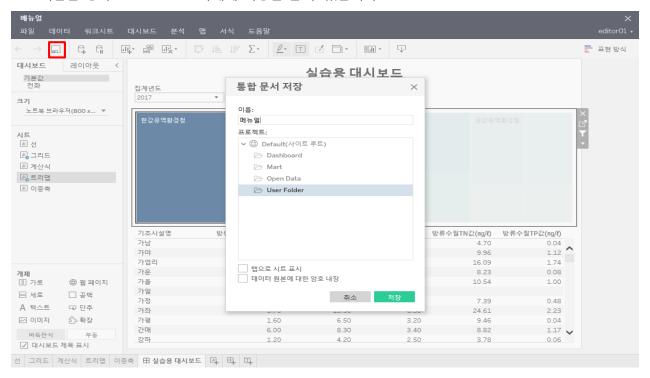
- 집계년도 필터를 눌러서 년도를 변경해보면 원래 필터가 있던 하나의 시트만 변경이 됩니다.
 ▼버튼을 눌러 이 데이터원본을 사용하는 모든 항목을 누릅니다.
- 이 데이터셋을 사용하는 모든시트에 집계년도 필터가 적용됩니다.



 트리맵 차트에서 환경청명을 클릭하면 하단의 그리드 차트가 해당 환경청에 소속된 시설들만 보이게 하려면, 트리맵 차트를 클릭 후 ▼ 버튼을 눌러 필터로 사용 버튼을 클릭하거나, ▼ 버튼 위의 필터 모양 버튼을 클릭합니다.



- 한강유역환경청을 눌러 하단의 그리드가 정상적으로 변경이 되는지 확인합니다.
- 저장버튼을 눌러 작업한 통합문서를 저장할 수 있습니다.
- 이름을 정하고 User Folder아래에 저장을 할 수 있습니다.

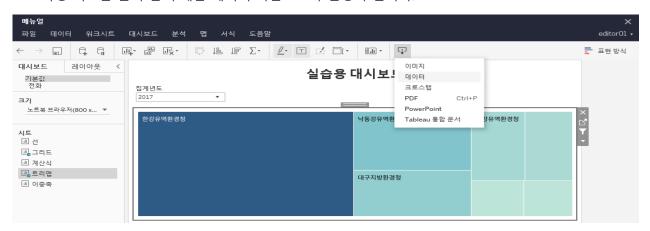


13 대시보드에서 데이터 다운로드

• 대시보드에서 아무것도 클릭하지 않은 상태에서는 데이터 다운로드 창이 활성화 되지 않습니다.



• 특정 시트를 클릭 한 후에는 데이터 다운로드가 활성화 됩니다.



● 특정항목을 선택한 후 데이터 다운로드를 누르면 선택된 항목의 데이터만 내려받을 수 있습니다.

