9일차 과제

[Day9]

과제 소개

시각적 분석의 모범 사례(베스트 프랙티스)를 적용하는 연습을 해 보겠습니다.

지금까지 배운 여러 기능을 총정리해서 활용해 볼 수 있는 시간입니다.

여러 모범 사례가 오늘 학습 자료(▮9일차 학습 자료 안내)에 제시된 백서에 자세하게 설명되어 있으며 거기에서 설명된 원칙을 요약하면 다음과 같습니다.

- 질문으로 시작하기 (Start with questions)
- 적합한 차트 유형 선택 (Choose the right chart type)
- 효과적인 뷰 만들기 (Create effective views)
- 총체적인 대시보드 디자인 (Create holistic dashboards)
- 작업을 완벽하게 (Perfect your views)

또한 백서의 38페이지에는 작성된 시각화가 베스트 프랙티스에 맞도록 디자인되었는지 점검할 수 있는 질문 목록이 제공됩니다.

9일차 과제는 크게 '베스트 프랙티스 적용'과 '서버 게시'의 2개 파트로 구성되어 있습니다.

♣과제 1: 시각적 분석 베스트 프랙티스 적용해 보기

♣과제 1-1 : 데이터 추출하기

♣과제 1-2 : 베스트 프랙티스 적용하기

♣과제 2 : 서버(온라인)에 게시하기

♣과제 2-1 : 서버 혹은 온라인 환경 확인

♣과제 2-2 : 데이터 원본 게시하기

❖과제 2-3 : 통합 문서 게시하기

♣과제 2-4 게시된 데이터 원본에 연결하기

과제 1 제출 방법 : 베스트 프랙티스가 적용된 대시보드를 스크린샷 혹은 사진 촬영하여 제출해 주세요.

과제 2 제출 방법 : 서버(혹은 Tableau Online) 게시된 데이터 원본 및 대시보드를 웹 브라우저에서 열어서 스크린샷 혹은 사진 촬영하여 제출해 주세요.

!! 가용한 서버가 없으며 Tableau Online 활성화에 실패한 경우 2번은 생략하고 제출해도 좋습니다.

과제 1: 시각적 분석 베스트 프랙티스 적용해 보기

과제 1에서는 아래 첨부된 통합 문서를 어떻게 만들었는지 살펴보고 위 원칙과 맞지 않는 부분을 바꿔 보겠습니다.

⇒ 과제 시작 통합 문서 다운로드 : @ DAY9 COVID-19 Cases_Start.twbx

 → 과제 시작 통합 문서 다운로드 : ∅ DAY9 COVID-19 Cases_Start_v2018.1.twbx (최신 버전 다운로드를 권장합니다)

과제 1-1: 데이터 추출하기

데이터가 구글 스프레드시트에 저장되어 있으므로 매번 데이터를 읽어 오는 데 시간이 걸릴 수 있으니 데이터를 **추출**해서 태블로의 성능 향상 기술을 활용하도록 하겠습니다.

추출에 대한 자세한 정보가 필요하다면 도움말을 참고하시기 바랍니다.

<u>▲ 주의!</u> 아래 구글 시트 연결을 시도할 때 사내 인터넷 등의 환경에서 구글 스프레트시트에 연결이 막혀 있거나 Google ID가 없는 등의 이유로 접근이 안되는 경우라면

과제 1-1의 추출 작업을 하는 대신에 추출 작업이 완료된 아래 통합 문서를 다운로드해서 사용하며, 바로 ♣ 과제 1-2 : 베스트 프랙티스 적용하기로 이동해서 진행합니다.

⇒ 추출 완료본 통합 문서 : @ DAY9 COVID-19 Cases_Start_Extract.twbx

☞ 추출 완료본 통합 문서 : @ DAY9 COVID-19 Cases_Start_Extract_v2018.1.twbx (최신 버전 다운로드를 권장합니다)

1. 구글 스프레드시트 로그인

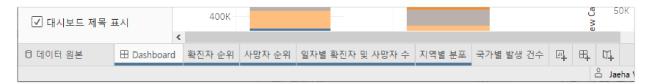
통합 문서를 열면 구글 스프레드시트에 연결하기 위해 웹 브라우저의 Google 로그인 창이 열립니다.



← 이와 같이 로그인된 계정을 선택하는 창이 뜨는 경우도 있지 만 Gmail 아이디와 패스워드를 처음부터 입력해서 로그인해야 할 수도 있습니다.

로그인 성공하여 통합 문서의 내용이 표시되면 하단에서 <u>1개의 대시보드</u>와 <u>4개의 파란색 표시된 시트</u> 및 <u>1개의 색상 표시</u> <u>없는 시트</u>를 확인할 수 있습니다.

파란색 표시된 시트는 대시보드에서 사용중인 시트이며, 색상 표시가 없는 시트는 대시보드에서 사용하지 않은 시트입니다.



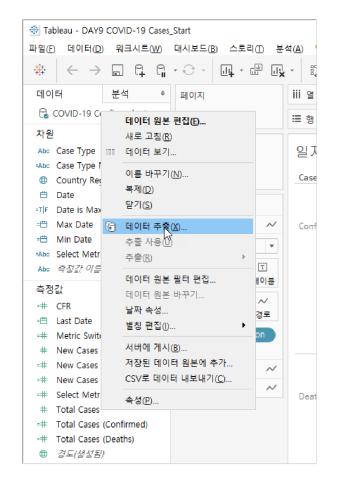
2. 데이터 추출하기

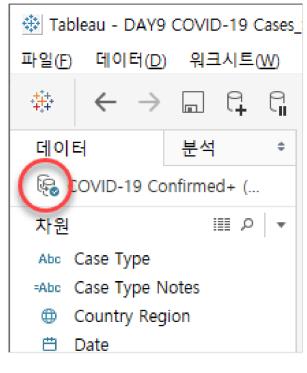
대시보드가 아닌 워크시트 중 아무 것이나 하나 선택하여 열어서 왼쪽 상단에 데이터 원본 이름이 보이게 합니다.

<u>단계1</u>. 데이터 원본에 우클릭하여 '데이터 추출'을 클릭합니다.

<u>단계2</u>. 데이터 추출 창이 뜨면 어떤 옵션도 손대지 말고 바로 추출 버튼을 눌러서 추출로 변환합니다.

추출이 완료되면 아래와 같이 원통 2개 모양으로 변경됩니다.





추출이 완료되면 통합 문서를 저장합니다.

이제는 어떤 작업을 해도 인터넷을 통해 구글 스프레드시트를 읽어오지 않고 추출된 hyper 데이터를 사용해서 빠르게 처리됩니다.

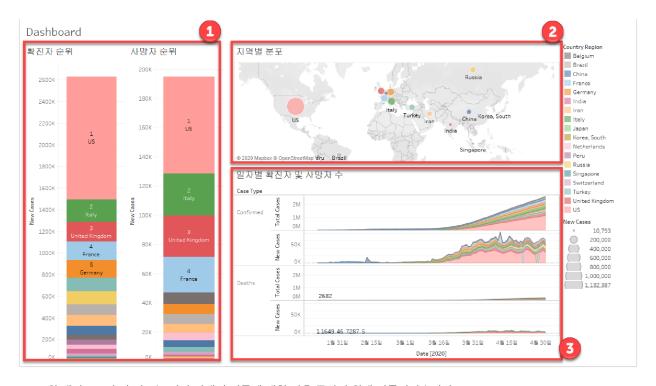
과제 1-2 : 베스트 프랙티스 적용하기

1. 대시보드 살펴보기

여기에 사용된 데이터는 Johns Hopkins 대학교에서 제공하는 코로나19 확진자 및 사망자 통계 중 일부 국가의 데이터를 발췌한 것입니다.

참고로 'New Cases' 측정값은 해당 일자의 신규 확진자 및 사망자 수, 'Total Cases' 측정값은 해당 일자까지의 누적 확진자 및 사망자 수를 의미합니다.

각 시트가 어떻게 작성되어 있는지 살펴 보고 각 시트를 클릭해 보면서 필터나 하이라이트 등의 작업(동작)이 적용되어 있는 지 확인해 봅니다.



위 대시보드의 각 시트는 각각 아래의 질문에 대한 답을 구하기 위해 만들어졌습니다.

- ① 각 국가별 확진자 및 사망자 수는 얼마나 되는가? 어느 나라가 더 많은가?
- ② 각 지역(대륙)별 확진자 분포는 어떠하며 어떤 특징을 보이는가
- ③ 각 국가별로 확진자는 증가세인가 감소세인가?

각 차트가 위 세 가지 질문에 대한 답을 구할 수 있도록 시각화되어 있는지 확인해 봅니다.

체크가 완료되었다면 시트별로 차례로 수정해 보도록 하겠습니다.

각 시트별로 적용할 기법은 한국어판 백서의 표제 및 페이지 번호와 함께 제시하였습니다.

베스트 프랙티스에는 완벽한 정답이 존재하지 않으므로 상상력을 발휘하여 본 과제 안내의 지시 사항에 없는 것까지 수정하거나 지시 사항과 다른 방향으로 수정하셔도 좋습니다.

2. '확진자 순위' 및 '사망자 순위' 시트 개선하기

<u>국가별 확진자 및 사망자 수를 순서대로 나열하고 비교</u>하기 위한 목적에 맞는 차트가 작성되어 있나요? 현재 표시된 차트는 오히려 국가별 비중을 보기에 좋은 차트인 것 같은데, 전세계 모든 국가가 포함되지 않았으므로 비중은 의미가 없습니다.

단계1. 적용할 기법 : 비교 및 순위 지정 (백서 8페이지)

순서대로 나열하고 값을 비교하기 좋은 차트는 막대 차트이니 거기에 맞게 수정해 봅니다.

Hint! '국가별 발생 건수' 시트가 이미 만들어져 있으니 이것으로 대체하면 되겠네요.

추가로 아래 기법도 적용해 봅시다.

단계2. 적용할 기법 : **알아보기 쉽도록 뷰 방향 조정** (16페이지)

국가의 수가 적지 않으며 표시 공간의 가로폭이 좁기 때문에 세로 막대보다는 가로 막대로 표현하는 것이 보기 쉽습니다. Hint! 툴바에 있는 '행과 열 바꾸기' 버튼을 사용해 보셨나요?

단계3. 적용할 기법 : **뷰 구성** (17 페이지)

가로 막대로 바꾼 상태에서 국가별 확진자와 사망자를 위 아래에 나란히 놓는 것은 의미가 없어 보입니다.

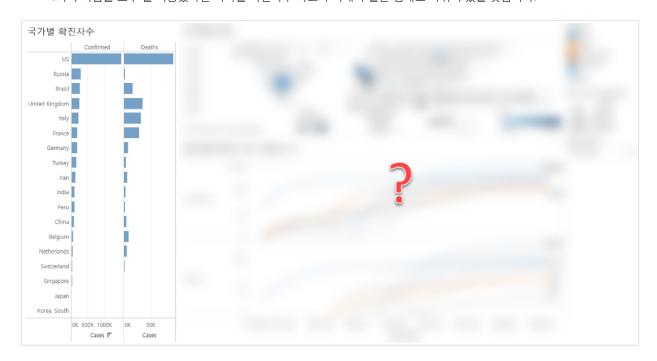
백서에서 적용한 방법을 거꾸로 사용하여 두 측정값이 좌우에 표시되게 바꾸는 것이 좋겠습니다.

단계4. 적용할 기법 : 지능형 축을 만들지 않으면 그냥 예쁜 직사각형만 남음 (36페이지)

Deaths 막대가 Confirmed 막대 대비하여 너무 작아서 잘 식별되지 않으므로 '**각 행 또는 열에 독립적인 축 범위**'를 지정하여 잘 구분되게 만듭니다.

Hint! 축 범위 옵션을 지정하려면 해당 축에 더블클릭하거나 우클릭 후 '축 편집'을 선택합니다.

4가지 기법을 모두 잘 적용했다면 국가별 확진자수 차트가 아래와 같은 형태로 바뀌어 있을 것입니다.



3. '지역별 분포' 시트 개선하기

<u>각 지역(대륙)별 확진자 분포와 특징을 확인</u>하기 위한 목적에 맞는 차트가 작성되어 있나요? 국가 이름을 색상으로 지정할 필요가 있을까요? 아니라면 색상에는 무엇을 사용할 수 있을까요?

Hint! 치명율(Case fatality rate)을 뜻하는 CFR 필드를 색상에 사용하면 어떨까요?

4. '일자별 확진자 및 사망자 수' 시트 개선하기

국가별 확진자 및 사망자 수가 증가세인지 감소세인지 확인하기 위한 목적에 맞는 차트가 작성되어 있나요? 아래 기법들을 적용해 볼 수 있습니다.

단계1. 적용할 기법 : 시간의 경과에 따른 트렌드 (7페이지)

영역 차트는 개별 국가의 숫자보다는 전체 합의 트렌드가 강조되는 차트 유형입니다.

마크 유형을 영역에서 라인으로 변경해 봅시다.

마크 유형을 바꾸면 차트가 어떻게 변화되는지 관찰해 봅시다.

단계2. 적용할 기법 : '단일 뷰의 색상 및 모양 개수 제한 (19페이지)

모든 국가별 색상이 지정되어 있어 식별이 되지 않다 보니 색상에 큰 의미가 부여되지 않습니다. 관심 있는 몇 개 국가에만 색상을 지정하는 방법을 사용하면 어떨까요? 이때 모든 국가별에 대한 선은 모두 표시하되, 기타 국가들은 회색으로 표시되게 합시다. Hint! 집합을 만들고 아래와 같이 계산된 필드를 만들면 관심 국가 이외의 국가들을 기타로 모을 수 있습니다. IF 조건에서 집합 이름만으로도 참/거짓을 구별해 주는 역할을 할 수 있음에 주목하세요.



단계3. 적용할 기법 : 공간이 부족하여 압축된 뷰 (31페이지)

차트 안에 너무 많은 정보가 들어 있어서 표시할 공간이 부족합니다.

대시보드 안에서 이리 저리 배치를 조정해서 제대로 표시되게 할 수 있으면 좋겠지만 어려워 보이네요. 3일차에 배웠던 매개 변수를 사용해서 Total Cases와 New Cases 측정값을 전환해 가며 볼 수 있게 만들어 봅시다. Hint! 'Select Metric' 매개 변수와 'Metric Switcher' 필드가 이미 작성되어 있으므로 바로 사용할 수 있습니다.

단계4. 적용할 기법 : 지능형 축을 만들지 않으면 그냥 예쁜 직사각형만 남음 (36페이지)

막대 차트에서 했던 것처럼 '각 행 또는 열에 독립적인 축 범위'를 지정하여 'Deaths' 라인이 잘 보이게 만듭니다.

단계5. 적용할 기법 : 모든 적합한 위치에 비주얼리제이션 레이블 지정 (37페이지)

모든 레이블이 다 표시되어 복잡해 보인다면 각 라인의 마지막 점에 대해서만 레이블을 표시해 보면 어떨까요?

□ 마크 i	베이블 표시				
레이블 모양					
텍스트:					
글꼴: 맑은 고딕, 9pt, 자동 색상 🗸					
맞춤: 자동	~				
레이블 마크					
전체	선택됨				
최소/최대	하이라이트됨				
라인 끝	가장 최근				
범위					
테이불	솉				
패널	라인/파이 차트				
음센					
☑ 레이블이 다른	마크와 겹치도록 허용				

 \rightarrow '라인 끝' 방식과 '가장 최근' 방식 두 가지 다 활용이 가능한데 어떤 차이가 있는지 직접 확인해 보세요.

이 때 '레이블이 다른 마크와 겹치도록 허용'을 하는 것이 좋을까요? 안 하는 것이 좋을까요? 직접 확인해 보세요.

5. 대시보드 개선하기

개별 시트에 대한 작업이 완료되었다면 대시보드에 대해서도 다음과 같은 기법을 적용해 보겠습니다.

단계1. 적용할 기법 : **일반 지침** (22페이지)

사용성을 고려한 시트 배치가 잘 되어 있는지 체크해 봅니다.

단계2. 적용할 기법 : **필터** (25페이지)

세 가지 측면에서 검토해 봅니다.

- 추가 하고 싶은 필터 작업(동작)이 있는가? → '국가별 발생 건수' 막대를 클릭하면 다른 시트에 국가 필터가 적용되게 합니다.
- 불필요한 필터 작업이 적용되어 있는가? → '**국가별 발생 건수**' 막대와 **지도** 간의 필터는 불필요해 보입니다. 필터 작업(동작)의 **대상 시트**를 조정하여 지도는 필터 대상에서 제외하면 됩니다.
- 필터 적용시 불필요한 필드가 전달되지 않는가? → '국가별 발생 건수' 막대를 클릭하여 필터 적용시에 라인 차트의 Confirmed 혹은 Deaths 둘 중 하나만 남는 것이 불만입니다. 막대를 선택할 때 선택한 국가로 라인 차트를 필터링하되 Confirmed와 Deaths 라인 둘 다 보이도록 합니다.

Hint! 필터 작업(동작)을 편집하여 아래와 같이 대상 필터 부분에서 국가 필드만 지정합니다.



단계3. 적용할 기법 : 강조 표시 (24페이지)

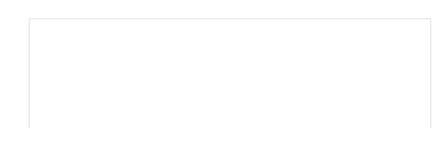
위에서 막대 차트와 지도 간에 필터가 걸리지 않게 했는데요. 그 대신에 선택한 국가를 다른 시트에서 하이라이트시켜 주면 어떨까요?

○ 툴바의 하이라이트 버튼에서 쉽게 설정할 수 있습니다.



단계4. 적용할 기법 : 스토리를 보여주는 툴팁 (35페이지)

기본 툴팁을 사용해도 좋지만 약간만 손을 보면 더 쓸모있고 압축된 정보를 사용자에게 제공할 수 있습니다.





이외에도 본인이 적용하고 싶은 기법들을 추가로 적용해도 좋습니다.

과제 2: 서버(온라인)에 게시하기

수정이 완료된 통합 문서를 Tableau Online 혹은 Tableau Server에 게시합니다.

과제 2-1 : 서버 혹은 온라인 환경 확인

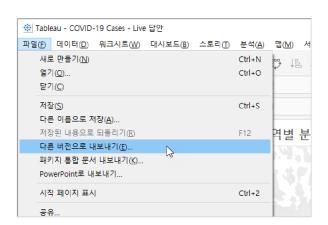
본 과제를 수행하기 위해서는 Tableau Server 혹은 Tableau Online에 게시자 권한을 부여 받아서 로그인이 가능해야 합니다.

1. 사내 서버의 로그인 아이디를 받은 경우

부여받은 아이디와 패스워드를 사용하시면 됩니다.

사내 Tableau Server의 버전이 Tableau Desktop보다 낮은 경우 데이터 원본 게시가 정상적으로 수행되지 않을 수 있습니다.

- 이 경우 두 가지 방법이 있습니다.
- 1) Tableau Desktop 2020.1 혹은 2020.2에서 '파일 > 다른 버전으로 내보내기' 메뉴 선택하여 서버와 동일 버전을 선택해서 내보내고 내보낸 통합 문서와 동일 버전의 데스크탑으로 통합 문서를 열고 과제 2-2 수행 2) 아래 안내에 따라 Tableau Online 평가판 사이트를 활성화해서 과제 2-2 수행



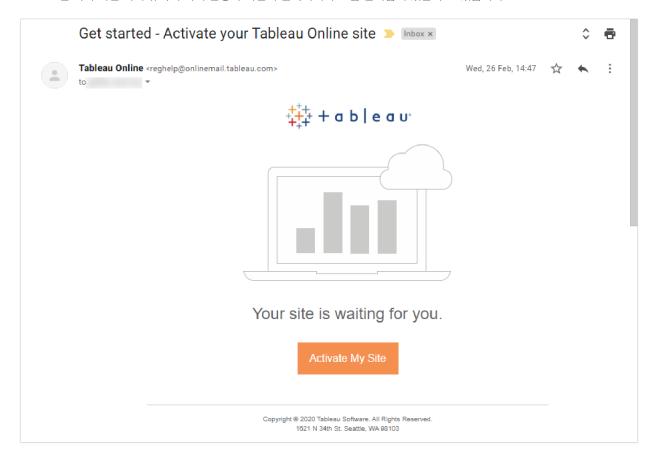
2. 사내 서버가 가용하지 않은 경우 - Tableau Online 평가판 사용

가용한 사내 서버가 없는 경우에는 Tableau Online의 평가판을 사용하여 게시해 봅니다. 아래 링크에서 Tableau Online 평가판 사이트를 활성화해서 사용하시기 바랍니다.

https://www.tableau.com/ko-kr/products/cloud-bi#form

위 링크에서 평가판 등록을 수행하면 등록한 메일 주소로 아래와 같은 활성화 메일이 수신됩니다. 메일 내용 중 Activate My Site 혹은 내 사이트 활성화 버튼을 눌러서 활성화를 수행합니다.

이 활성화 메일이 수신될 때까지 다소 시간이 걸리는 경우가 있습니다. 또한 회사 메일 서버 규칙에 의해 활성화 메일이 블록되거나 스팸 편지함에 있을 수도 있습니다.



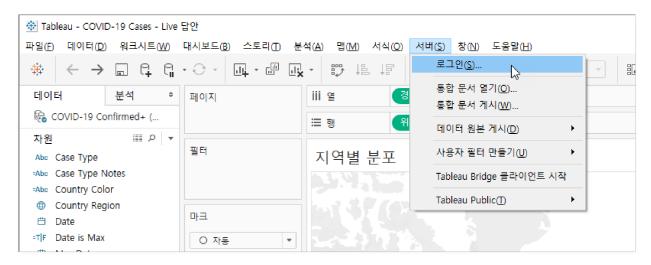
▲주의! 일부 회사 네트워크에서는 방화벽에 의해 Tableau Online에 접근 또는 업로드(게시)가 막혀 있을 수 있습니다. ▲주의! tableau.com에서 무료 교육 동영상 시청을 위해 로그인했던 아이디가 있다고 해서 바로 Tableau Online에 해당 아이디로 로그인 가능하지 않습니다. Tableau Online 평가판 활성화시에 지정한 패스워드로 로그인해야 합니다.

과제 2-2: 데이터 원본 게시하기

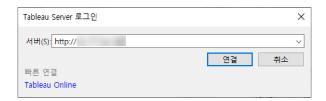
통합 문서와 데이터 원본을 한꺼번에 게시하는 방법도 있지만 여기서는 데이터 원본을 먼저 게시하도록 하겠습니다. 이렇게 하면 서버에 게시된 데이터 원본을 이용하여 다음에도 동일한 데이터에 접근하는 것이 간편해지며, 다른 사람들과도 데이터 원본을 쉽게 공유할 수 있으므로 누구나 동일한 데이터를 보고 있다는 확신을 가질 수 있습니다. 조금 거창하게 말하자면 **거버넌스**의 시작이라고 할 수 있겠습니다.

1. 서버(온라인) 로그인

서버 혹은 온라인에 로그인합니다.



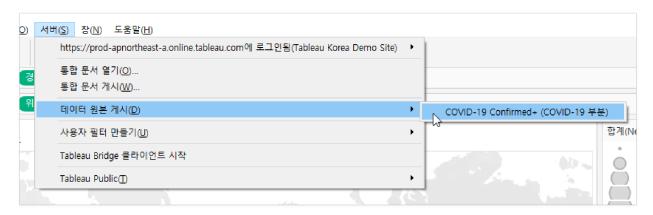
로그인 창에서 서버 주소를 입력하여 Tableau Server에 접속하거나 좌측 하단의 Tableau Online을 클릭하여 Tableau Online에 접속합니다.

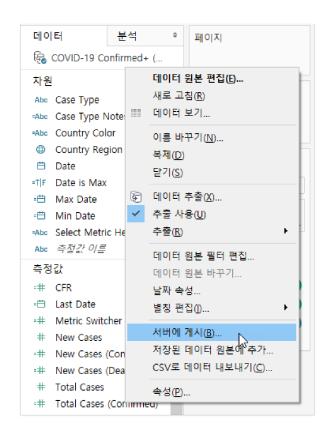


로그인 창이 뜨면 아이디와 패스워드를 입력하여 로그인합니다.

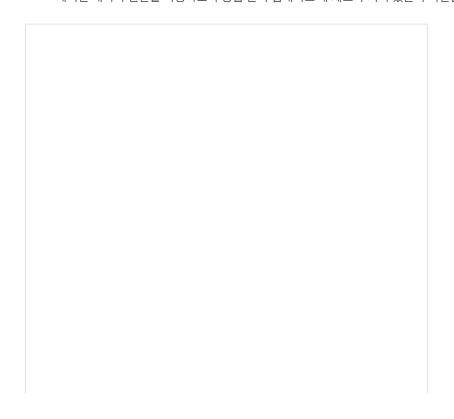
2. 데이터 원본 게시

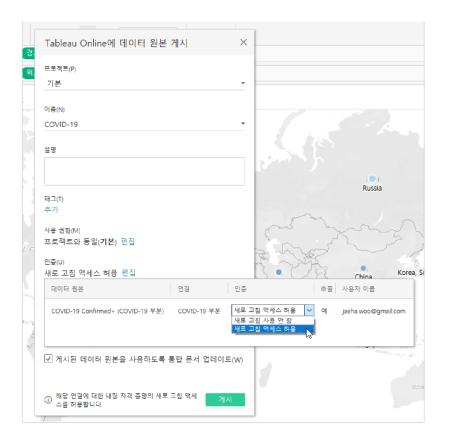
단계1. '서버 > 데이터 원본 게시' 메뉴를 선택하거나 데이터 원본 이름을 우클릭하여 '서버에 게시' 메뉴를 선택합니다.





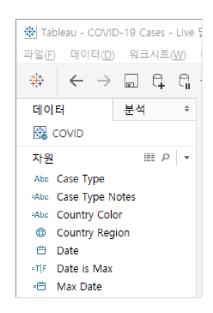
단계2. 프로젝트를 선택하고 서버에 저장될 이름을 지정한 뒤 필요하다면 설명과 태그를 붙입니다. 새로 고침 일정을 지정하기 위해서 '새로 고침 액세스 허용' 옵션을 선택합니다. '게시된 데이터 원본을 사용하도록 통합 문서 업데이트'에 체크가 되어 있는지 확인합니다.





단계3. 게시 버튼을 눌러서 게시가 완료되면 데이터 원본의 아이콘이 원통에서 아래와 같이 변하는 것을 확인할 수 있습니다.

또한 웹 브라우저 창이 열려서 서버 혹은 온라인에 게시된 데이터 원본이 뜹니다. 여기서 우측 하단의 '**일정**' 버튼을 눌러서 새로 고침 일정을 지정합니다.





단계4. 내일 아침에 새로운 데이터가 들어오는지 확인하려면 이른 아침 시간을 선택해서 새로 고침 일정을 지정합니다. 데이터 추출을 사용할 때는 대부분의 경우 새로 고침을 함께 활용하게 됩니다.

(Tableau Server를 사용하는 경우 원하는 일정이 없다면 서버 관리자가 만들어줘야 합니다)

"COVID" 데이터 원본에 대한 새로 고침 일정을 선택하십시오.						
O 검색						
매일 오전 5:00				A		
매일 오전 5:00 - 순차(느린 순)						
매일 오전 6:00						
매일 오전 7:00				- 1		
매일 오전 8:00						
매일 오전 9:00	Spu)					
매일 오전 10:00	O					
매일 오전 11:00						
매일 오후 12:00						

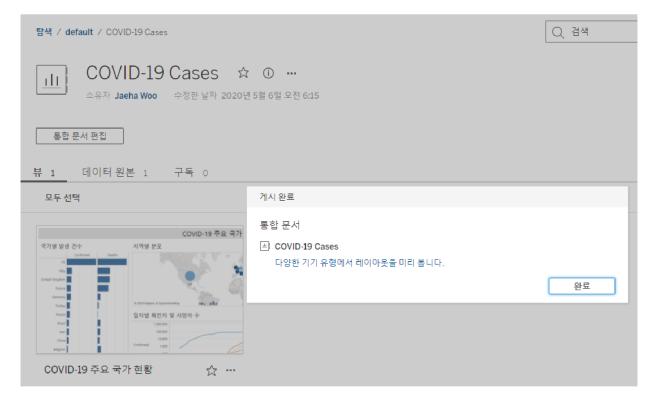
과제 2-3 : 통합 문서 게시하기

1. ▶9일차 학습 자료 안내에 안내된 'Tableau Server에 게시' 교육 동영상을 참고하여 통합 문서를 게시합니다.

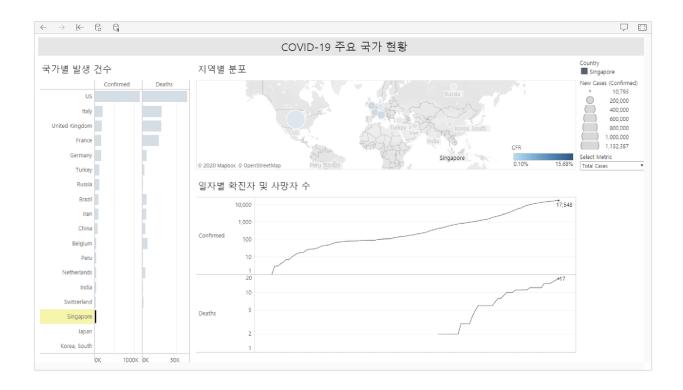
이때 모든 시트를 게시하지 않고 대시보드만 선택해서 게시하도록 합니다.



2. 게시가 완료되면 웹 브라우저 창이 열려서 서버 혹은 온라인에 게시된 통합 문서가 뜹니다.

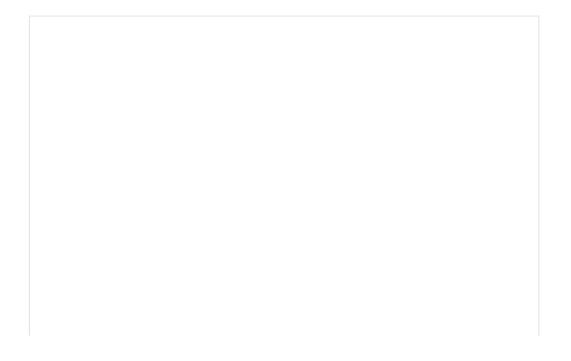


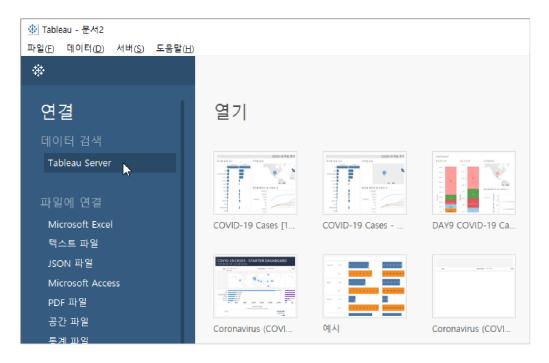
3. 대시보드를 클릭해서 데스크탑에서 본 것과 동일하게 작동하는지 확인해 봅니다.



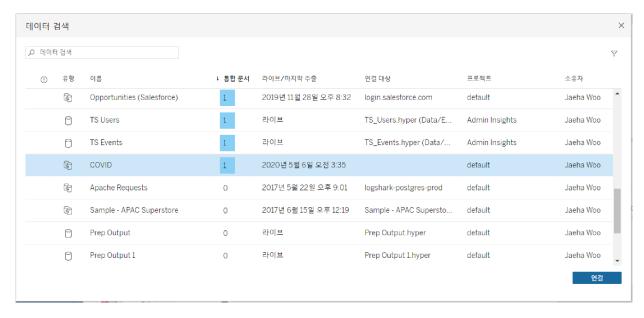
과제 2-4 게시된 데이터 원본에 연결하기

1. Tableau Desktop을 새로 실행하거나 열려 있는 통합 문서를 닫아서 시작 페이지에서 게시된 데이터 원본에 연결합니다. 연결할 데이터 종류는 데이터 검색 아래의 **Tableau Server**입니다.



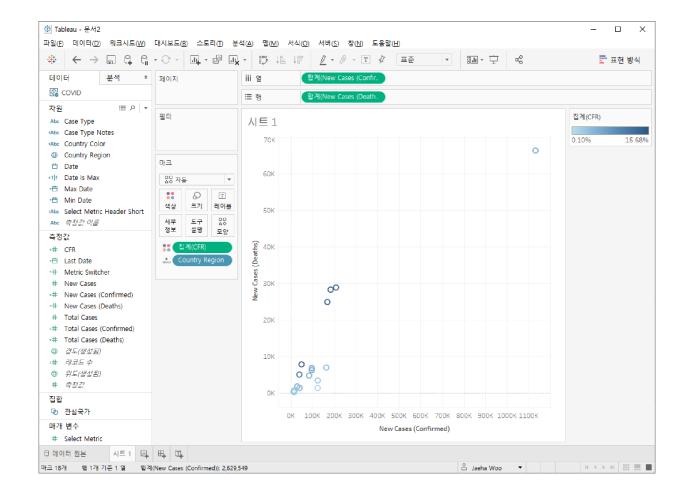


2. 서버 혹은 온라인에 로그인되어 있다면 게시되어 있는 데이터 원본 목록이 보입니다.



3. 방금 게시한 데이터 원본을 선택해서 연결한 뒤 아래와 같은 차트를 만들어 봅니다.

누구나 쉽게 서버에 공유된 데이터 원본을 사용해서 분석/시각화할 수 있게 되었음을 확인하셨나요?



맺음말

오늘 연습하셨던 베스트 프랙티스가 도움이 되셨나요? 익숙해지면 특히 여러 사람이 보는 대시보드를 한층 효과적으로 만들 수 있게 됩니다.

Tableau Server나 Online과 같이 사용한다면 공유와 협업을 위해 데이터 원본과 통합 문서의 게시 방법 혹은 게시된 데이터 원본의 사용 방법을 아는 것은 필수이겠죠.

추출을 사용하는 경우에는 어떻게 원하는 주기와 일정으로 새로운 데이터로 갱신하는지도 알아 보았습니다.

왠지 분량 조절 실패로 과제가 다소 많아진 것 같은데... ② 여기까지 오시느라 **수고 많으셨습니다!!**

다시 한번 과제 제출 방법을 안내해 드립니다.

과제 1 제출 방법 : 베스트 프랙티스를 적용한 대시보드를 스크린샷 혹은 사진 촬영하여 제출해 주세요.

과제 2 제출 방법 : 서버(온라인) 게시된 데이터 원본 및 대시보드를 웹 브라우저에서 열어서 스크린샷 혹은 사진 촬영하여 제출해주세요.

👭 가용한 서버가 없으며 Tableau Online 활성화에 실패한 경우 2번은 생략하고 제출해도 좋습니다.