|  |
| --- |
| **Week#03 | 태블로 필수 기능, 하나씩 익혀보기** |

**★ 3주차 학습이 본격적으로 시작되기 전 다음 사항을 체크해주세요!**

**[2주차 주요 내용 리뷰]**

* 태블로에서 모든 필드들은 차원과 측정값으로 나뉘게 된다. 측정값은 우리가 관심을 가지는 그 자체이며, 차원은 측정값을 나누어서 살펴보는 관점 또는 기준을 의미한다
* 태블로의 모든 필드들은 연속형과 불연속형으로 나뉘게 된다. 연속형은 측정값, 불연속형은 차원, 이런 공식은 성립하지 않는다
* 태블로는 항상 데이터를 집계하기 때문에, 데이터셋 한 줄 한 줄의 의미를 파악하는 것이 중요하다. 화면이 차원에 의해 쪼개져 있는 수준을 ‘View Level of Detail(VLOD)라고 부르며, 각 화면의 VLOD를 파악하는 것이 중요하다. 태블로는 항상 VLOD에서 집계하기 때문이다

|  |
| --- |
|  |

**들어가는 말. 학습 목표 및 3주차 학습 요약**

3주차 **태블로 필수 기능, 하나씩 익혀보기** 학습이 시작되었습니다.

**태블로 베이직 클래스 3주차의 학습 목표는 업무 현장에서 태블로 화면을 만드는데 가장 빈번하게 활용되는 주요 기능들의 사용법을 정확하게 익히는 것입니다.** ‘태블로를 배웠다’ 라고 하면 기본적으로 알아야 하는 핵심 기능 5개를 선별해서 살펴볼 것입니다. 1-2 주차 학습이 태블로의 기초를 다지는 시간이었다면, 3주차에는 본격적으로 실전적인 화면을 만드는 작업에 돌입하게 됩니다

먼저 태블로가 제공하고 있는 조인과 유니온 기능을 학습할 것이고, 2-4일차에는 이번주 학습의 핵심인 필터, 이중축, 분석 패널, 이렇게 세 가지를 살펴볼 예정입니다. 5일차에서는 태블로에서 지도는 어떻게 활용할 수 있는지 공부할 것입니다. 지난 2주차에 익혔던 개념들을 충분히 숙지하고 있다는 전제 하에 학습이 진행되기 때문에, 차원/측정값, 연속형/불연속형, VLOD 개념이 모호하신 분들은 반드시 2주차 주요 개념을 충분히 이해한 후 그 바탕 위에 3주차를 진행하시길 권해드립니다

3주차 주간 과제는 실전적 이슈가 포함되어 있어 이전에 비해 다소 어렵게 느껴지실 수 있습니다. 막히는 부분은 언제든 질문 남겨주십시오

기분 좋은 출발이 되길 기원합니다!

|  |
| --- |
|  |

**Day 1. 조인과 유니온**

* **(필수) 조인 (1) ([VizLab](https://vizlab.tistory.com/56))**
* **(필수) 데이터 조인 (2020.2 버전 이상) (**[**VizLab**](https://youtu.be/WkIZQtVk-H8)**)**
* **(필수) 유니온 ([VizLab](https://vizlab.tistory.com/58))**
* (추천) 블렌딩 (1) ([VizLab](https://vizlab.tistory.com/60))
* (추천) 블렌딩 (2) ([VizLab](https://vizlab.tistory.com/61))
* (추천) 블렌딩 Hands-On ([VizLab](https://vizlab.tistory.com/62))
* (참고) 조인 (2) ([VizLab](https://vizlab.tistory.com/57))
* (참고) 각각의 지점에서 다른 지점까지 거리 구하기 ([VizLab](https://youtu.be/HF858ThZa9w))
* (참고) Tableau’s New Relationships and What They Mean ([Spencer Bauke](https://www.tessellationtech.io/understanding-tableau-relationships/))

*다음의 Day 1 Quiz를 통해 오늘의 학습 내용을 정리해봅시다. 정답은 함께 제공되는 태블로 워크북에 나와있으니, 결과를 스스로 한 번 체크해보세요!*

* Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 태블로 위로 불러와봅시다. 01\_Orders 테이블과 02\_Returns 테이블을 Left 조인으로 연결해봅시다. 두 테이블을 조인하기 위해 사용하는 Key Field는 무엇인가요?
* 전체 주문 건수는 몇 건인가요? Order ID 필드를 고유 카운트로 집계하여 확인해봅시다. 반품 여부를 확인하기 위해 활용할 수 있는 필드는 무엇인가요? 해당 필드를 활용하여 반품된 주문과 정상 판매된 주문을 구분해봅시다. 주문 건수를 기준으로 반품 비중은 몇 %인가요? (\* 힌트: 퀵테이블 계산 활용)
* ★고난이도★ 이번에는 매출(Sales)을 기준으로 반품률을 살펴봅시다. 각 시도(Address SD)의 매출을 확인하는 바 차트를 만든 후 반품 여부를 확인할 수 있는 필드를 마크 카드 색상으로 올려봅시다. 반품 매출이 가장 큰 시도는 어디인가요? 각 시도에서 발생한 매출을 100%로 봤을 때 반품률이 가장 큰 시도는 어디인가요?
* Superstore KR - VizLab 2021 파일의 03\_Managers 테이블을 새롭게 연결해봅시다. 데이터 준비창의 아래쪽 프리뷰 화면을 보면 총 몇 개의 필드가 확인되고 있나요? 그 필드들의 이름은 무엇인가요?
* ★중요★ Managers 테이블을 화면에서 내린 후 이번에는 Orders 테이블을 올려봅시다. 그리고 Orders 테이블과 Managers 테이블을 Left 조인으로 연결해봅시다. 붉은색 느낌표가 나타나면서 조인이 이루어지지 않는 이유가 무엇일까요? 두 테이블을 조인하기 위해 사용하는 Key Field는 무엇인가요?
* 가장 낮은 수익(Profit)을 기록하고 있는 지역(Region)은 어디이며, 그 지역의 담당자(Sales Rep)는 누구인가요?
* Stocks.xlsx 파일을 태블로 위로 불러와봅시다. ‘Stocks 2010-2013’ 테이블과 ‘Stocks 2014’ 테이블을 유니온 기능을 활용하여 하나의 데이터 테이블로 결합해봅시다. 2014년 마지막 거래일 아마존의 종가(close)는 얼마인가요?

|  |
| --- |
|  |

**Day 2. 필터**

* **(필수) 태블로 코어 드릴다운 - ① 필터 (Filters) ([VizLab](https://youtu.be/i_60_1RNYk8))**
* (참고) 태블로 베이직 시리즈 - 첫걸음 ② 데이터 준비하기 ([VizLab](https://youtu.be/aRVccLYfyL4))
* (참고) 데이터 원본 필터 ([VizLab](https://vizlab.tistory.com/47))
* (참고) 태블로 마스터 클래스 - Order of Operations ④ 컨텍스트 필터 ([VizLab](https://youtu.be/DsGnLd0oZOY))
* (참고) 태블로 마스터 클래스 - Order of Operations ① 테이블 계산 필터: LAST ([VizLab](https://youtu.be/k41o1m9xsR8))
* (참고) 태블로 마스터 클래스 - Order of Operations ② 테이블 계산 필터: LOOKUP ([VizLab](https://youtu.be/IRZAbkrkj60))

*다음의 Day 2 Quiz를 통해 오늘의 학습 내용을 정리해봅시다*

* Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 태블로 위로 불러와봅시다. 데이터 원본 필터를 활용하여 Customer Segment 필드에서 기업 고객만 화면으로 불러와봅시다. 전체 데이터는 (Row 기준으로) 총 몇 건인가요?
* 위의 기업 고객 데이터를 계속해서 탐색해봅시다. 수도권 지역(서울특별시, 인천광역시, 경기도)의 수익(Profit)을 모두 합한 값은 얼마인가요?
* 역시 기업 고객 데이터를 계속해서 사용합니다. 매출과 수익을 활용하여 스캐터 플랏을 만든 후 시도와 시군구로 화면을 분기해봅시다. 측정값 필터를 활용하여 매출이 1,500만원보다 적고, 수익이 -50만원에서 50만원 사이에 있는 시군구만 화면에 남겨봅시다. 총 몇 개의 시군구가 화면에 남아있나요?
* Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 새롭게 연결해봅시다. 01\_Orders 테이블과 02\_Returns 테이블을 Left 조인으로 연결합니다. 연도별 반품 매출(Sales)을 바 차트로 표현해봅시다. 반품 매출이 가장 낮은 연도부터 높은 연도 순서대로 나열해보세요. 2017년의 반품 매출은 얼마인가요?
* 세계 지표(World Indicators) 데이터셋을 태블로 위로 불러와 봅시다. 연도, 국가/지역, 유아 사망률 필드를 활용하여 하이라이트 테이블을 만들어봅시다 (\* 힌트: 표현방식을 활용하면 쉽게 만들 수 있습니다). 중동 지역 국가 중에 2000년대 중반까지 압도적으로 높은 유아 사망률을 기록한 나라는 어디인가요? 해당 국가의 2000년 유아 사망률은 얼마인가요?
* 2010년 각 국가가 배출하는 이산화탄소 배출량을 (가로 형태의) 바 차트로 표현해봅시다. 배출량이 가장 많은 국가를 제일 위쪽에 배치하여 내림차순으로 정렬해보세요. 우리나라의 이산화탄소 배출량은 전 세계 이산화탄소 배출량의 몇 %를 차지하고 있나요? 이는 전 세계 국가 중 몇 위에 해당되나요?
* ★고난이도★ 위의 문제에서 만들었던 바 차트에서 일체의 수치 변화 없이 아시아 지역 국가들만 보여주고 싶습니다. 지역 필드를 필터에 올려 아시아 지역만 선택하면 어떤 결과가 나오나요? 이산화탄소 배출량 비중(%)은 전 세계 대비 수치를 보여주면서 아시아 지역만 나타나게 하고 싶다면 어떤 방법이 있을까요? (\* 주의: 태블로 마스터 클래스에서 다루는 내용으로 문제의 초점만 확인하면 되겠습니다)

|  |
| --- |
|  |

**Day 3. 이중축 + Combined Axis Chart**

* **(필수) 태블로 코어 드릴다운 - ② 이중축 (Dual Axis) ([VizLab](https://youtu.be/Zf4MABLbqhw))**
* **(필수) 도넛 차트 그리는 방법 ([VizLab](https://youtu.be/AtyvWeNOgjo))**
* **(필수) 도넛 차트에 레이블 넣기 ([VizLab](https://youtu.be/bvbcs4tGfeg))**
* (참고) 대시보드 설계와 데이터 시각화: 차트와 그래프를 활용한 비즈니스 데이터 분석 시나리오와 디자인 사례, 제34장 “파이와 도넛의 달콤한 유혹” (Steve Wexler, Jeffrey Shaffer, and Andy Cotgreave)

*다음의 Day 3 Quiz를 통해 오늘의 학습 내용을 정리해봅시다.*

* 이중축 차트를 만든 후 왼쪽과 오른쪽의 축의 값을 일치시키기 위해 하는 작업은 무엇인가요?
* ★고난이도★ Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 태블로 위로 불러와봅시다. Orders 테이블과 Returns 테이블을 Left 조인으로 연결한 후 반품된 매출(Sales)의 비중을 보여주는 도넛 차트를 만들어봅시다
* 도넛 차트의 강점은 무엇인가요? 또 약점은 무엇인가요?
* Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 새롭게 연결해봅시다. 이중축을 활용하여 아래 이미지의 bar-in-bar 형태의 화면을 만들어봅시다

Chart, bar chart

Description automatically generated

* ★중요★ 세계 지표(World Indicators) 데이터셋을 태블로 위로 불러와 봅시다. 우리나라의 연도별 유소년 인구 비율(모집단 0~14명)과 노인 인구 비율(모집단 65명 이상)을 이중축 차트와 결합축 차트로 각각 표현해봅시다. 두 화면에서 공통점과 차이점에 대해 설명해봅시다

|  |
| --- |
|  |

**Day 4. 분석 패널**

* **(필수) 태블로 코어 드릴다운 - ③ 분석 패널 (Analytics Pane) ([VizLab](https://youtu.be/lvCfNBpL5h8))**
* (추천) Tableau Essentials: Chart Types - Bullet Graph ([Carly Capitula](https://interworks.com/blog/ccapitula/2014/12/29/tableau-essentials-chart-types-bullet-graph/))
* (참고) 복수의 시트에서 바 차트 스케일링 ([VizLab](https://youtu.be/KaNKzYQNrNg))
* (참고) WINDOW\_MAX 함수와 참조선을 활용하여 여유 공간 만들기 ([VizLab](https://youtu.be/u38axt6Fvp4))
* (참고) 뷰에 고급분석 적용 ([태블로 온라인 헬프](https://help.tableau.com/current/pro/desktop/ko-kr/environ_workspace_analytics_pane.htm))

*다음의 Day 4 Quiz를 통해 오늘의 학습 내용을 정리해봅시다.*

* ★중요★ Superstore KR - VizLab 2021 파일의 01\_Orders 테이블을 태블로 위로 불러와봅시다. 각 시도별 매출(Sales)을 나타내는 바 차트를 만들어봅시다. 분석 패널에서 평균 라인을 테이블 형태로 집어 넣어 봅시다. 그 값은 얼마인가요? 여기에서 말하는 “평균”은 구체적으로 무엇을 의미하나요? 이 평균이 계산된 원리를 생각해봅시다
* 주문 일자(Order Timestamp)를 연속형 월로 화면에 올린 후 수량(Quantity)의 트렌드를 살펴봅시다. 태블로의 예측 기능을 사용하여 2020년 12월에 예상되는 수량을 확인해봅시다. 예측값은 얼마인가요? 예측이 시작되는 시기를 2021년 1월로 조정해봅시다. 이때 2020년 12월의 값은 실제값인가요, 예측값인가요? 태블로는 왜 마지막 값이 실제로 존재함에도 불구하고 예측값을 주는 것일까요?
* 평균 할인율(Discount)과 수익(Profit) 필드를 각각 열 선반과 행 선반에 올린 후 제품(Product Name) 레벨로 VLOD를 내려 스캐터 플랏을 만들어봅시다. 총 몇 개의 제품이 확인되고 있나요? 여기에 선형 트렌드 라인을 넣은 후 라인 위에 마우스를 올려봅시다. 할인율 앞에 있는 계수가 얼마인가요? 그 값의 의미는 무엇인가요? 모델의 R제곱(R-Squared)은 얼마인가요?
* ★고난이도★ Order Timestamp 필드를 연속형 월로 화면에 올린 후 평균 할인율의 트렌드를 살펴봅시다. 평균 할인율의 최대값과 최소값에 대해 각각의 참조선을 넣고 레이블을 통해 화면에 표시해봅시다. 각각의 값은 얼마인가요?
* ★고난이도★ 세계 지표(World Indicators) 데이터셋을 태블로 위로 불러와 봅시다. 유럽 국가들 중에서 평균 유소년 인구 비율는 중앙값(Median) 이상에 있지만 평균 노인 인구 비율은 중앙값 이하에 있어서 비교적 젊은 국가로 판단할 수 있는 국가는 몇 개인가요?
* 중동 국가들 중 평균 GDP가 1000억불 이상을 기록한 국가는 몇 개인가요? 바 차트에 상수 라인을 넣어서 확인해봅시다

|  |
| --- |
|  |

**Day 5. 지도**

* **(필수) 태블로 코어 드릴다운 - ④ 지도 (Map) ([VizLab](https://youtu.be/dadOqMBVUn0))**
* (추천) 3 Ways to Make Magnificent Maps in Tableau ([Ryan Sleeper](https://playfairdata.com/3-ways-to-make-magnificent-maps-in-tableau/))
* (추천) Vworld 지도 배경 Tableau에서 사용하기 ([우재하](https://douluvviz.tumblr.com/post/150711864227/using-vworld-background-map-in-tableau))
* (참고) [셔니의 5분 태블로] 나만의 멋진 맵 만들기! (맵박스 테마 5가지 파일 공유) ([WeViz](https://youtu.be/quZfx68_erE))

*다음의 Day 5 Quiz를 통해 오늘의 학습 내용을 정리해봅시다.*

* 태블로에서 지도를 그리기 위해서는 시도 필드와 시군구 필드에 \_\_\_\_\_\_\_\_\_\_을 부여해야 한다
* Superstore KR - VizLab 2021 파일의 Orders 테이블을 태블로 위로 불러와봅시다. 각 시군구별 수익(Profit)을 지도에 표현해봅시다. 수익이 가장 낮은 시군구는 어디인가요? 해당 시군구의 수익은 얼마인가요?
* ★고난이도★ 세계 지표(World Indicators) 데이터셋을 태블로 위로 불러와 봅시다. 아프리카 대륙에서 우간다(Uganda)를 찾아보세요. 우간다의 평균 유아 사망률은 몇 %인가요? 우간다보다 남쪽에 위치한 국가 중 평균 유아 사망률이 가장 높은 아프리카 국가는 어디인가요? 해당 국가의 유아 사망률은 몇 %인가요?

|  |
| --- |
|  |

**Day 6. Weekly Challenge 3**

**1. 다음의 대시보드를 만든 후 태블로 퍼블릭에 업로드 해주세요!** ([https://tabsoft.co/2Yafiob](https://www.google.com/url?q=https://tabsoft.co/2Yafiob&sa=D&ust=1584534665790000&usg=AFQjCNH6rlkRX_VBBPIjaICz-IVPx5byww))

A screenshot of a cell phone

Description automatically generated

* **[0] 대시보드 기본 서식**
  + 대시보드 사이즈 1200 \* 800, 시트 6장
  + 대시보드 제목: 폰트 태블로 볼드 / 사이즈 20
  + 대시보드 제목 우측의 회사 로고는 온라인에서 자유롭게 다운로드 받아서 넣어주세요
  + 시트 제목: 폰트 태블로 북 / 사이즈 12 / 볼드 및 언더라인
  + 화면 전체에 쓰인 보라색은 컬러 코드는 #3C0061입니다
* **[1] 데이터 조인**
  + Sample - Superstore 데이터의 Orders 테이블과 Returns 테이블을 Left Join으로 연결하여 사용합니다. 어떤 필드를 키 필드로 활용하여 두 테이블은 연결할 수 있을까요?(Day1 Quiz 참고)
  + 반품된 주문과 정상 판매된 주문을 구분해줄 수 있는 필드는 무엇인가요?
* **[2] BAN (Big Ass Number)**
  + BAN은 비즈니스 대시보드에서 많이 활용되는 데이터 디스플레이 방식으로 현재의 비즈니스 현황을 한 눈에 파악할 수 있도록 도와줍니다
  + 만드는 방식 또한 매우 간단합니다. 보여주고자 하는 필드를 마크 카드 레이블 위로 올립니다. 중앙 정렬, 크기 및 색상 조정 등을 적절히 진행하면 됩니다
  + BAN을 만들 때에는 도구 설명을 off 하는게 좋습니다. 독자에게 똑같은 정보를 반복적으로 보여줄 필요는 없습니다 (마크 카드 레이블 🡪 왼쪽 하단 도구 설명 표시 체크 마크 해제)
  + 동일한 포맷의 BAN을 두 개 생성할 것이기 때문에 첫번째 BAN을 복제하여 사용할 수 있습니다. 두번째 BAN에서는 반품 주문과 정상 판매 주문을 구분해주는 필드를 필터로 활용합니다
* **[3] 반품 비율 도넛 차트**
  + 3일차 학습 자료를 기반으로 반품 비율을 보여주는 도넛 차트를 구성합니다 (참고 영상: <https://youtu.be/AtyvWeNOgjo> / <https://youtu.be/bvbcs4tGfeg>)
  + 상식적인 선에서 생각해볼 때 반품 총액이 매출 총액보다 큰 상황이 발생할 수 있나요? 그런 면에서 봤을 때 반품 비율을 도넛 차트로 표현하는 것이 적절하다고도 생각되나요?

* **[4] 주(State)별 매출 현황**
  + 태블로에서 지도를 그리기 위해서는 지리 정보를 가지고 있는 필드에 ‘지리적 역할’을 부여해주는 것이 핵심입니다. 다행히 sample-superstore.xls 파일의 orders 테이블을 태블로에 불러오면 필요한 필드에 지리적 역할이 이미 부여되어 있습니다
  + State 필드를 더블 클릭 했을 때 (미국) 지도가 제대로 표현되지 않는 경우가 있다면 위치 정보가 제대로 인식되어 있지 않기 때문입니다. 우측 하단의 ‘OO 알 수 없음’을 클릭하여 위치 편집을 통해 국가/지역 부분을 미국으로 설정하면 해결됩니다
  + 지도의 배경은 상단 메뉴 맵 🡪 맵 계층 🡪 백그라운드 투명도 🡪 100% 설정을 통해 없앨 수 있습니다
  + 마우스 휠을 통한 지도의 줌 인/아웃 기능을 제한하기 위해서는 상단 메뉴 맵 🡪 맵 옵션 🡪 이동 및 확대 축소 허용 🡪 체크 마크 해제할 수 있습니다
  + 마우스 휠을 활용한 줌 인/아웃 기능을 제한하는 대신 지역(Region) 필터를 활용하여 줌 인/아웃 기능을 구현할 수 있습니다
* **[5 | 고난이도] 제품 서브카테고리별 반송율 바 차트**
  + 바 차트에 표시된 비중은 각 서브카테고리 내에서 반품된 금액의 비중을 의미합니다. 각 카테고리 내에서의 비중 합산이 100%가 되는 것이 아님을 유의하세요
  + 매출을 서브카테고리 레벨에서 보여준 후 반품된 금액과 정상 판매된 금액을 색상으로 구분합니다
  + 서브카테고리 간 업무 성과를 비교하기 위해서는 각각의 서브카테고리를 100%로 잡은 후(전체 1,700%) 각각의 서브카테고리 내에서 반품 비중을 살펴봐야 합니다. 퀵테이블 계산 ‘구성 비율’을 활용할 것인데, 퀵테이블 계산을 어떤 필드에 대해 어떻게 설정할 것인지가 관건입니다 (참고 영상: <https://youtu.be/7lDSFCMABUc>)
  + 퀵테이블 계산을 성공적으로 만들었다면 마지막으로 정상 판매된 부분을 화면에서 날리고, 반품된 부분만 화면에 남기면 됩니다. 여기에서는 우리가 배웠던 필터 기능 대신 숨기기 기능을 활용할 것입니다. 그 이유는 태블로의 Order of Operation([이미지 다운로드](https://1drv.ms/u/s!At3ucxamQMPYgcUKa8eGZ1XGzasDFQ?e=Jb7M6k))와 관련 있는데, 이 부분은 주간 미팅에서 상세히 설명하겠습니다

* **[6 | 고난이도] 월별 반송율 바 차트**
  + 바 차트에 표시된 비중은 각 월 내에서 반품된 금액의 비중을 의미합니다. 각 분기 또는 연도 내에서의 비중 합산이 100%가 되는 것이 아닙니다
  + 위의 제품 서브카테고리별 반송율 바 차트와 만드는 방식이 동일합니다
  + 차트가 완성되었으면, 마지막으로 분석 패널의 평균 라인을 각 분기별로 넣어주세요

**2. 해당 대시보드의 링크를 복사한 후 붙여주세요**

**3. 대시보드 화면을 만드는 것보다 더 중요한 것은 완성된 대시보드로부터 스토리텔링 포인트를 짚어내는 것입니다. 만들어진 화면을 탐색해보면서 눈에 띄는 포인트를 세 가지만 찾아보세요. 이 부분을 가다듬어 인사이트로 정리해주세요**

**★ 위의 1 – 3번 문제에 대한 답을 다음 링크를 통해 제출해주세요!**

<https://forms.gle/4MxfMj1s6P5KJfUA8>

|  |
| --- |
|  |

**Day 7. Weekly Challenge 리뷰 및 Q&A 세션**

작성해주신 과제를 바탕으로 주간 과제 화면을 리뷰하며, 한 주간의 학습 관련된 내용에 대한 질문 및 응답 시간을 가집니다