

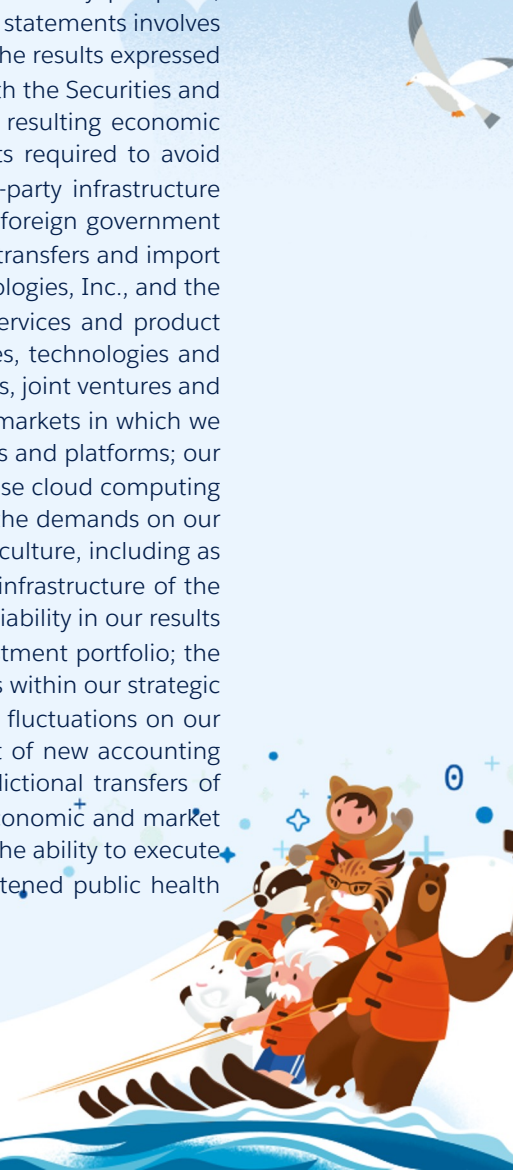
# 세상에서 가장 쉽게 배우는 태블로 데이터 모델링

명완식 솔루션 엔지니어, 태블로  
[wmyung@salesforce.com](mailto:wmyung@salesforce.com)

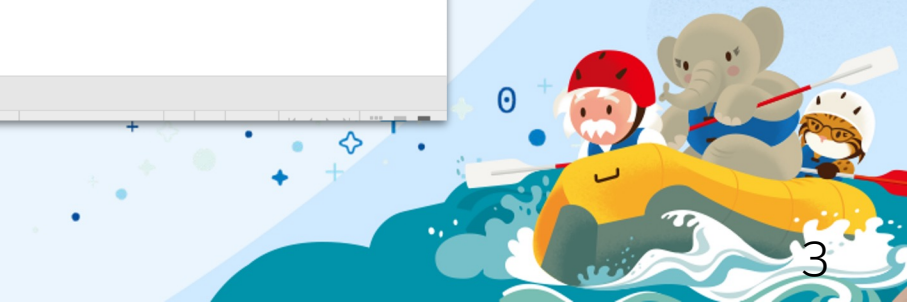
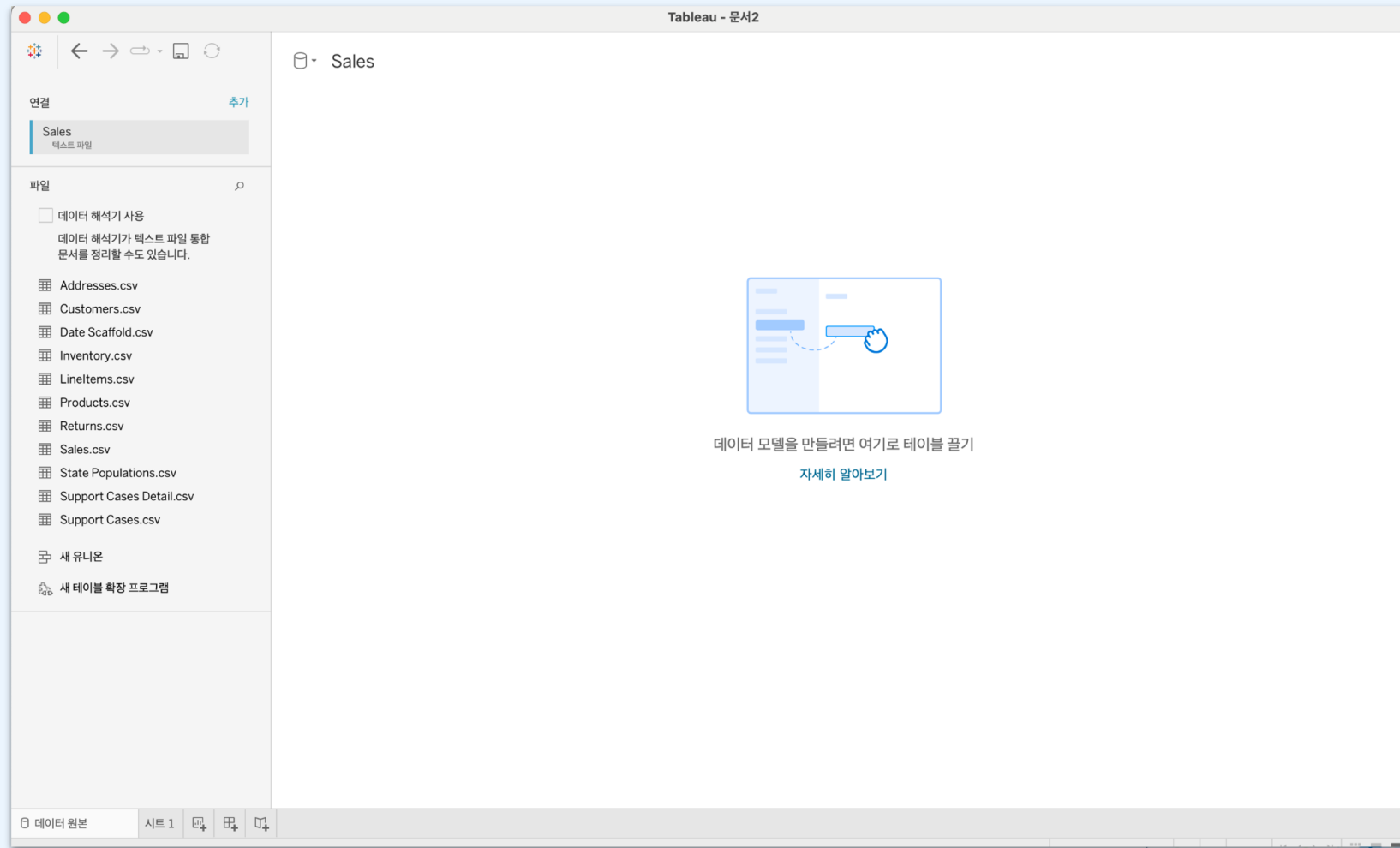


# Forward Looking Statements

This presentation contains forward-looking statements about, among other things, trend analyses and future events, future financial performance, anticipated growth, industry prospects, environmental, social and governance goals, and the anticipated benefits of acquired companies. The achievement or success of the matters covered by such forward-looking statements involves risks, uncertainties and assumptions. If any such risks or uncertainties materialize or if any of the assumptions prove incorrect, Salesforce's results could differ materially from the results expressed or implied by these forward-looking statements. The risks and uncertainties referred to above include those factors discussed in Salesforce's reports filed from time to time with the Securities and Exchange Commission, including, but not limited to: impact of, and actions we may take in response to, the COVID-19 pandemic, related public health measures and resulting economic downturn and market volatility; our ability to maintain security levels and service performance meeting the expectations of our customers, and the resources and costs required to avoid unanticipated downtime and prevent, detect and remediate performance degradation and security breaches; the expenses associated with our data centers and third-party infrastructure providers; our ability to secure additional data center capacity; our reliance on third-party hardware, software and platform providers; the effect of evolving domestic and foreign government regulations, including those related to the provision of services on the Internet, those related to accessing the Internet, and those addressing data privacy, cross-border data transfers and import and export controls; current and potential litigation involving us or our industry, including litigation involving acquired entities such as Tableau Software, Inc. and Slack Technologies, Inc., and the resolution or settlement thereof; regulatory developments and regulatory investigations involving us or affecting our industry; our ability to successfully introduce new services and product features, including any efforts to expand our services; the success of our strategy of acquiring or making investments in complementary businesses, joint ventures, services, technologies and intellectual property rights; our ability to complete, on a timely basis or at all, announced transactions; our ability to realize the benefits from acquisitions, strategic partnerships, joint ventures and investments, including our July 2021 acquisition of Slack Technologies, Inc., and successfully integrate acquired businesses and technologies; our ability to compete in the markets in which we participate; the success of our business strategy and our plan to build our business, including our strategy to be a leading provider of enterprise cloud computing applications and platforms; our ability to execute our business plans; our ability to continue to grow unearned revenue and remaining performance obligation; the pace of change and innovation in enterprise cloud computing services; the seasonal nature of our sales cycles; our ability to limit customer attrition and costs related to those efforts; the success of our international expansion strategy; the demands on our personnel and infrastructure resulting from significant growth in our customer base and operations, including as a result of acquisitions; our ability to preserve our workplace culture, including as a result of our decisions regarding our current and future office environments or work-from-home policies; our dependency on the development and maintenance of the infrastructure of the Internet; our real estate and office facilities strategy and related costs and uncertainties; fluctuations in, and our ability to predict, our operating results and cash flows; the variability in our results arising from the accounting for term license revenue products; the performance and fair value of our investments in complementary businesses through our strategic investment portfolio; the impact of future gains or losses from our strategic investment portfolio, including gains or losses from overall market conditions that may affect the publicly traded companies within our strategic investment portfolio; our ability to protect our intellectual property rights; our ability to develop our brands; the impact of foreign currency exchange rate and interest rate fluctuations on our results; the valuation of our deferred tax assets and the release of related valuation allowances; the potential availability of additional tax assets in the future; the impact of new accounting pronouncements and tax laws; uncertainties affecting our ability to estimate our tax rate; uncertainties regarding our tax obligations in connection with potential jurisdictional transfers of intellectual property, including the tax rate, the timing of the transfer and the value of such transferred intellectual property; uncertainties regarding the effect of general economic and market conditions; the impact of geopolitical events; uncertainties regarding the impact of expensing stock options and other equity awards; the sufficiency of our capital resources; the ability to execute our Share Repurchase Program; our ability to comply with our debt covenants and lease obligations; the impact of climate change, natural disasters and actual or threatened public health emergencies; and our ability to achieve our aspirations, goals and projections related to our environmental, social and governance initiatives.



# 태블로 데스크탑이 첫화면에서 지원하는 연결



# Agenda

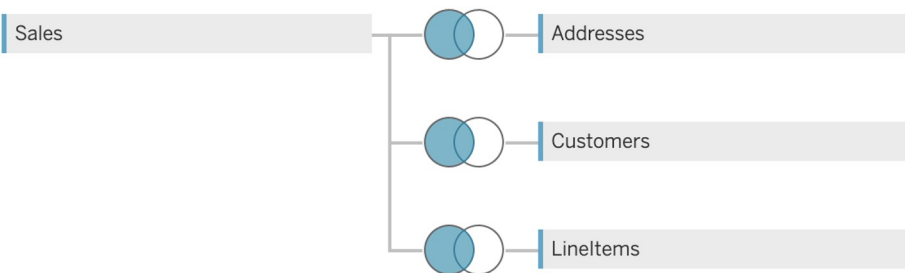
- ❖ 조인 (Join)
- ❖ 관계 (Relationship)
- ❖ 다중 관계 (Shared Dimension)



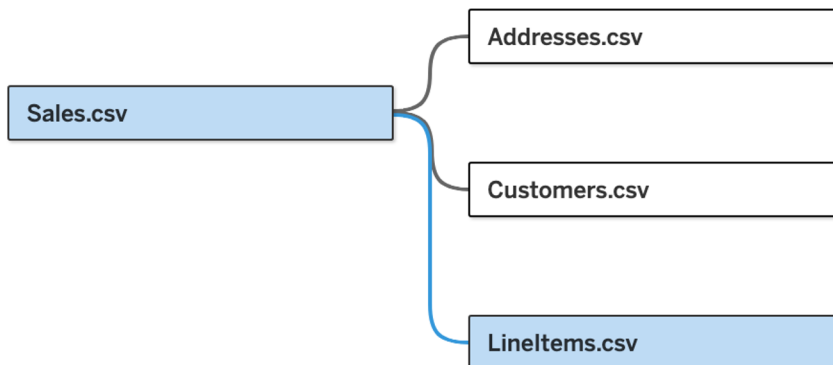
# 테블로 데스크탑이 첫화면에서 지원하는 연결

## 조인(Join)

Sales 4개 테이블로 구성되어 있습니다. ①



## 관계(Relationship)



# 어떤 기능을 써야 시원하게 데이터 준비를 잘할 수 있을까?



조인(Join)

그냥 드래그앤 드랍만 하면 되는거 아니까요?  
뭘 알아야해요?



관계 (Relationship)

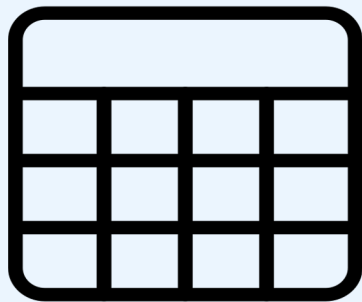


# 조인 (Join)

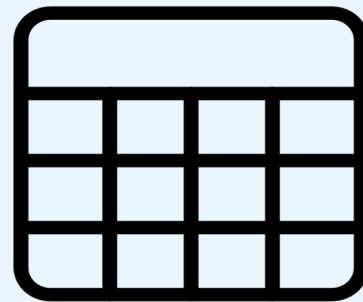


# 조인 (JOIN)

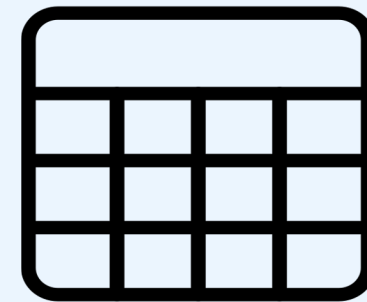
“데이터 분석을 할 경우 다른 데이터끼리 데이터를 결합해야 하는 경우가 많습니다.  
테이블 간의 키를 중심으로 분석 목적에 따라 여러 방법으로 테이블을 결합하여  
새로운 테이블을 생성하는 것을 기본적으로 조인(JOIN)이라고 합니다.”



A Table



B Table



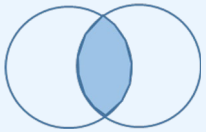
New Table



# INNER JOIN

“INNER JOIN은 두 개의 테이블에서 공통된 데이터를 가진 행들만 결합하는 방법입니다.

두 테이블에서 **같은 값**을 가지고 있는 부분만 가져옵니다.”



↓ key

지역	매출
서울	4200000
울산	3000000
대구	5000000
부산	4600000



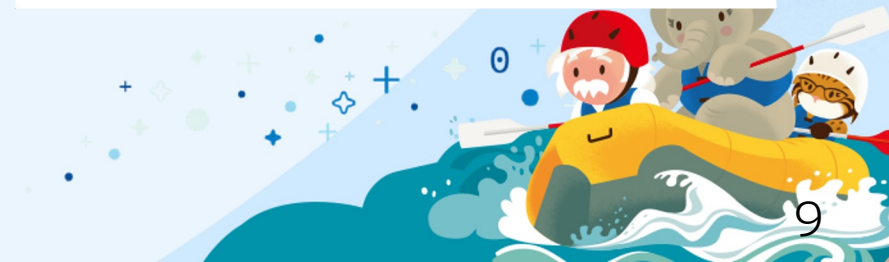
↓ key

지역	매니저
서울	박지현
울산	김성국
대구	이상욱
대전	김영건



Inner Join

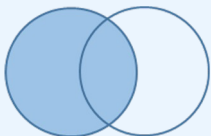
지역	매출	매니저
서울	4200000	박지현
울산	3000000	김성국
대구	5000000	이상욱



# LEFT JOIN

“LEFT JOIN은 두 개의 테이블을 결합할 때 **왼쪽 테이블의 모든행을 포함**하고  
오른쪽 테이블에서 일치하는 데이터가 있는 경우만 데이터를 결합합니다.

오른쪽 테이블에 데이터가 없는 경우 NULL로 표현합니다.”



↓ key

지역	매출
서울	4200000
울산	3000000
대구	5000000
부산	4600000



↓ key

지역	매니저
서울	박지현
울산	김성국
대구	이상욱
대전	김영건



Left Join

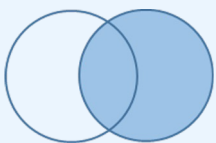
지역	매출	매니저
서울	4200000	박지현
울산	3000000	김성국
대구	5000000	이상욱
부산	4600000	<i>null</i>



# RIGHT JOIN

“RIGHT JOIN은 두 개의 테이블을 결합할 때 **오른쪽 테이블의 모든행**을 포함하고  
왼쪽 테이블에서 일치하는 데이터가 있는 경우만 데이터를 결합합니다.

왼쪽 테이블에 데이터가 없는 경우 NULL로 표현합니다.”



↓ key

지역	매출
서울	4200000
울산	3000000
대구	5000000
부산	4600000

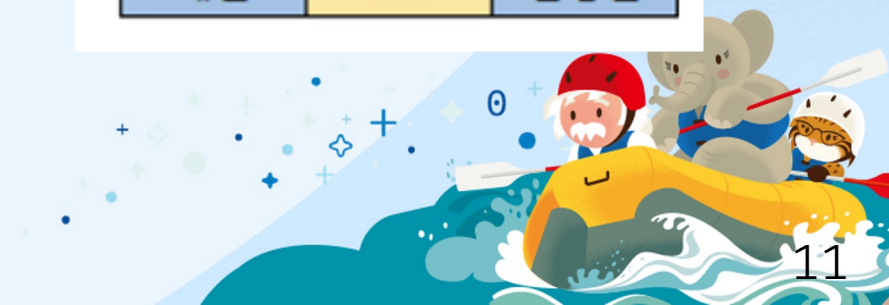
+

↓ key

지역	매니저
서울	박지현
울산	김성국
대구	이상욱
대전	김영건

=

Right Join		
지역	매출	매니저
서울	4200000	박지현
울산	3000000	김성국
대구	5000000	이상욱
대전	null	김영건



# FULL OUTER JOIN

“FULL OUTER JOIN은 두 개의 테이블을 결합할 때,  
두 테이블의 모든 데이터를 포함하는 방법입니다.

왼쪽과 오른쪽 테이블의 모든 행을 가져오며,  
어느 쪽에 일치하는 데이터가 없으면 NULL값으로 표현합니다.”



↓ key	매출	+	↓ key	매니저
지역			지역	
서울	4200000		서울	박지현
울산	3000000		울산	김성국
대구	5000000		대구	이상욱
부산	4600000		대전	김영건

=

Full Outer Join		
지역	매출	매니저
서울	4200000	박지현
울산	3000000	김성국
대구	5000000	이상욱
부산	4600000	null
대전	null	김영건

# 태블로는 자동으로 SQL을 만듭니다.

“태블로에서 JOIN을 활용하는 경우 먼저 키 기준으로 JOIN을 수행하여 완전한 테이블을 만든 후 필요 데이터를 뽑아서 분석 결과를 보여줍니다.

즉 Customer 테이블의 정보가 필요 없다하더라도 일단 조인 수식에 따라 테이블을 먼저 만들고 결과를 뽑아냅니다.”

빅쿼리에서 뽑아낸 실제 SQL

```
SELECT `LineItems`.`Product_ID` AS `Product_ID`,  
       SUM(`LineItems`.`Sales`) AS `sum_Sales_ok`  
FROM `secret-descent-288907.Tableau_Data_Modeling`.`Sales` `Sales`  
  INNER JOIN `secret-descent-288907.Tableau_Data_Modeling`.`Addresses` `Addresses` ON  
    (`Sales`.`Address_ID` = `Addresses`.`Address_ID`)  
  INNER JOIN `secret-descent-288907.Tableau_Data_Modeling`.`Customers` `Customers` ON  
    (`Sales`.`Customer_ID` = `Customers`.`Customer_ID`)  
  INNER JOIN `secret-descent-288907.Tableau_Data_Modeling`.`LineItems` `LineItems` ON  
    (`Sales`.`Order_ID` = `LineItems`.`Order_ID`)  
  INNER JOIN `secret-descent-288907.Tableau_Data_Modeling`.`Products` `Products` ON  
    (`LineItems`.`Product_ID` = `Products`.`Product_ID`)  
GROUP BY 1
```

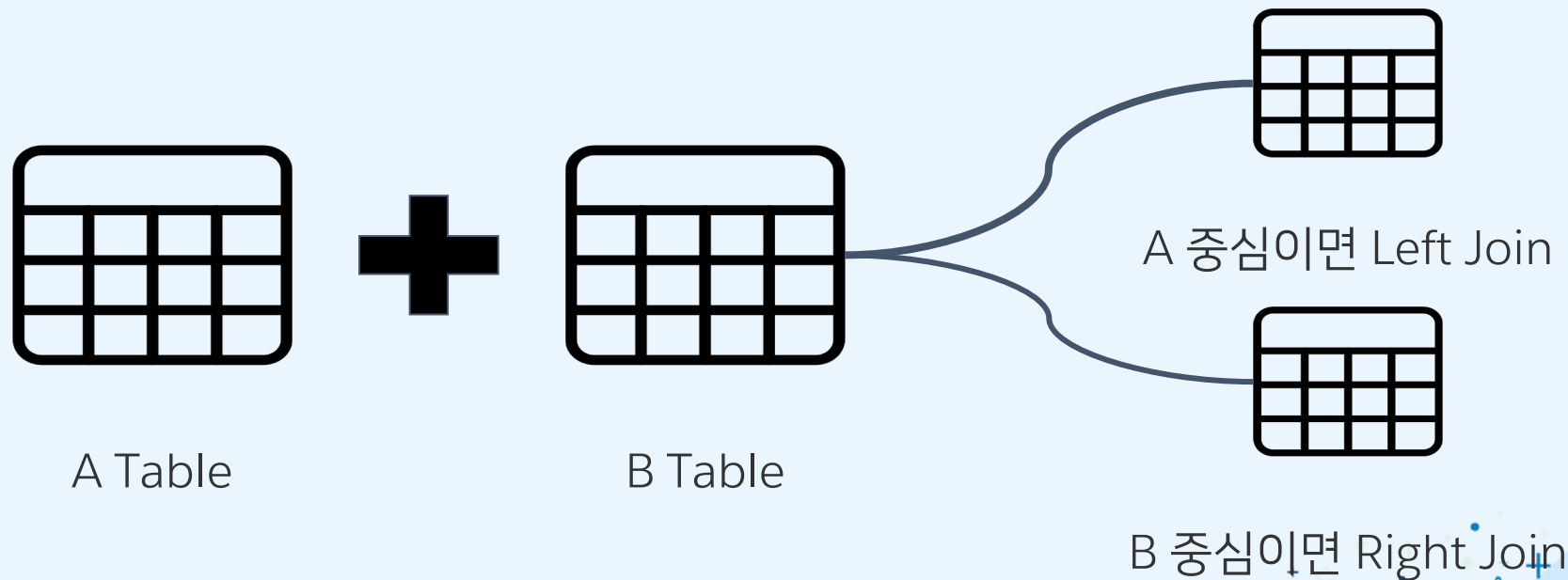


# 관계 (Relationship)



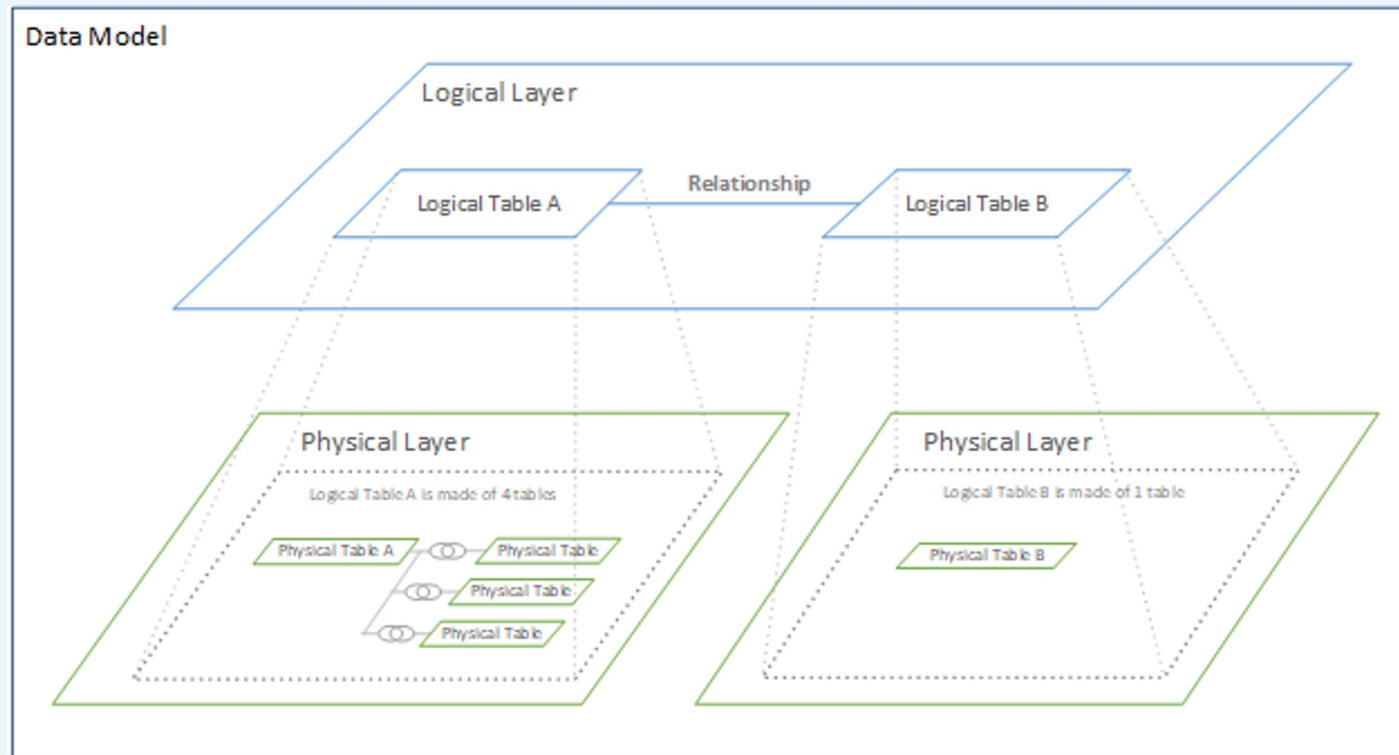
# 관계 (Relationship)

“관계는 분석을 위해 여러 테이블에서 데이터를 결합하는 동적이고 유연한 방법이며, 관계는 조인 유형을 정의하지 않고 분석하고자 하는 중심 테이블에 따라 조인 유형이 변경됩니다.”



# 관계 (Relationship)

“Join은 데이터 결합시, 미리 정의한 결합 조건에 따라 완전한 테이블을 먼저 생성하고 데이터 분석을 하지만, **관계는 분석 중심 테이블을 먼저 판단한 후 데이터를 결합하고 분석합니다.**”

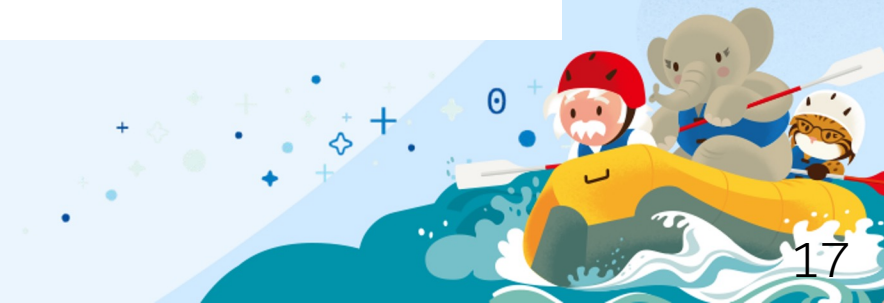


# 관계는 필요한 데이터만 가지고 SQL을 만듭니다.

“관계는 필요한 데이터 중심으로만 데이터를 준비시키기 때문에, 여러 테이블과 관계를 맺었다고 해서 모든 데이터를 조인한 후 데이터를 조회하지 않습니다.”

## 빅쿼리에서 뽑아낸 실제 SQL

```
SELECT `t0`.`Product_ID` AS `Product_ID__Products_`,  
       SUM(`LineItems`.`Sales`) AS `sum_Sales_ok`  
FROM `secret-descent-288907.Tableau_Data_Modeling`.`LineItems` `LineItems`  
LEFT JOIN (  
  SELECT `Products`.`Product_ID` AS `Product_ID`  
  FROM `secret-descent-288907.Tableau_Data_Modeling`.`Products` `Products`  
  GROUP BY 1  
) `t0` ON (`LineItems`.`Product_ID` = `t0`.`Product_ID`)  
GROUP BY 1
```



Relationship은 어떤 테이블의 데이터를 쓰는지 먼저 판단하고 분석수준에 필요한 데이터만 쓰기 때문에 **데이터 복제를 예방**할 수 있습니다.

## JOIN

주문아이디	주문자명	상품명	옵션	금액	상담 만족도	상담 피드백
A0901901	명완식	제네시스 G90	풀옵션	130,000,000	5	차 설명 너무 좋아요. 바로 계약 했어요
				130,000,000	2	설명이 어려웠어요
				130,000,000	3	차보고 싶은데 전화가 잘안되네요

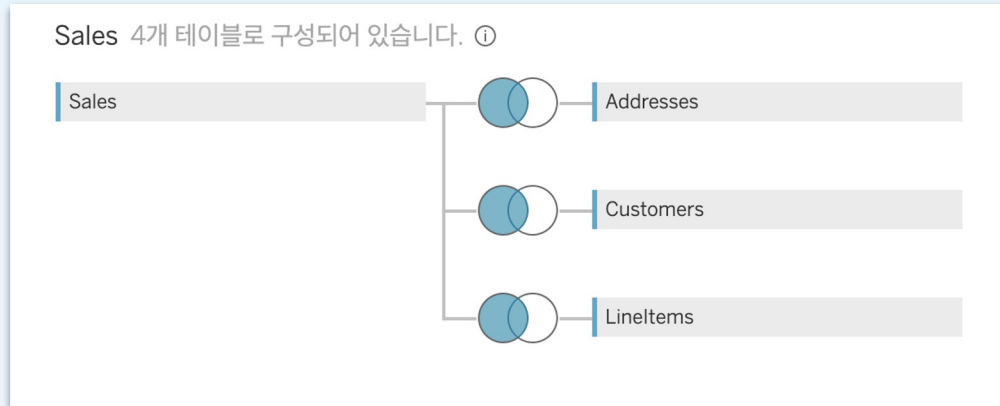
## Relationship

주문아이디	주문자명	상품명	옵션	금액
A0901901	명완식	제네시스 G90	풀옵션	130,000,000

상담 만족도	상담 피드백
5	상담 잘했어요
1	상담 불만족
3	너무 전화를 높게 받으시네요

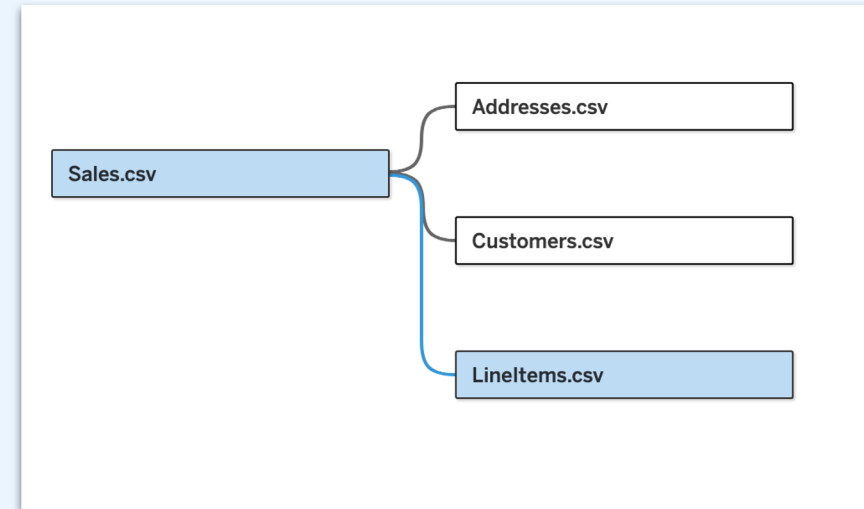
# JOIN과 RELATIONSHIP 차이점

## Traditional Physical Models



- Joins을 이용한 모델
- 모델의 결과는 단일 물리적 테이블 형태로 됨
- 서로 다른 세부수준의 데이터를 연결하면 복잡도가 증가함

## Logical – Relationship Models



- 조인 유형을 지정할 필요가 없음
- 개별 테이블을 별도로 저장
- 필요한 경우에만 테이블간 쿼리를 수행함
- 다양한 관점으로 테이블을 살펴볼 수 있음
- 서로 다른 세부수준의 데이터를 손쉽게 분석

# 다중 관계 (Shared Dimension)



아래 관계에서 고객의 피드백을 제품 기준과  
고객 기준 동시에 분석하고 싶다면?



! Support Cases.csv

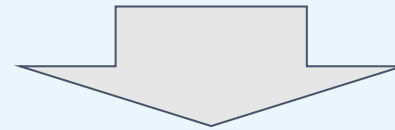
# 고객의 피드백을 제품 기준과 고객 기준 동시에 분석하고 싶다면?

오늘 야근각?



담당자 A씨

- 1 임시 테이블을 여러개 만들어서 SQL로 붙이고 시각화 한다.
- 2 그냥 처음부터 새로운 관계 하나 더 설정을 한다.
- 3 태블로 프렘으로 데이터를 어떻게 잘 말아서 가지고 온다.
- 4 피드백 테이블을 고객 테이블과 제품 테이블 오른쪽에 그냥 둘 다 붙인다.

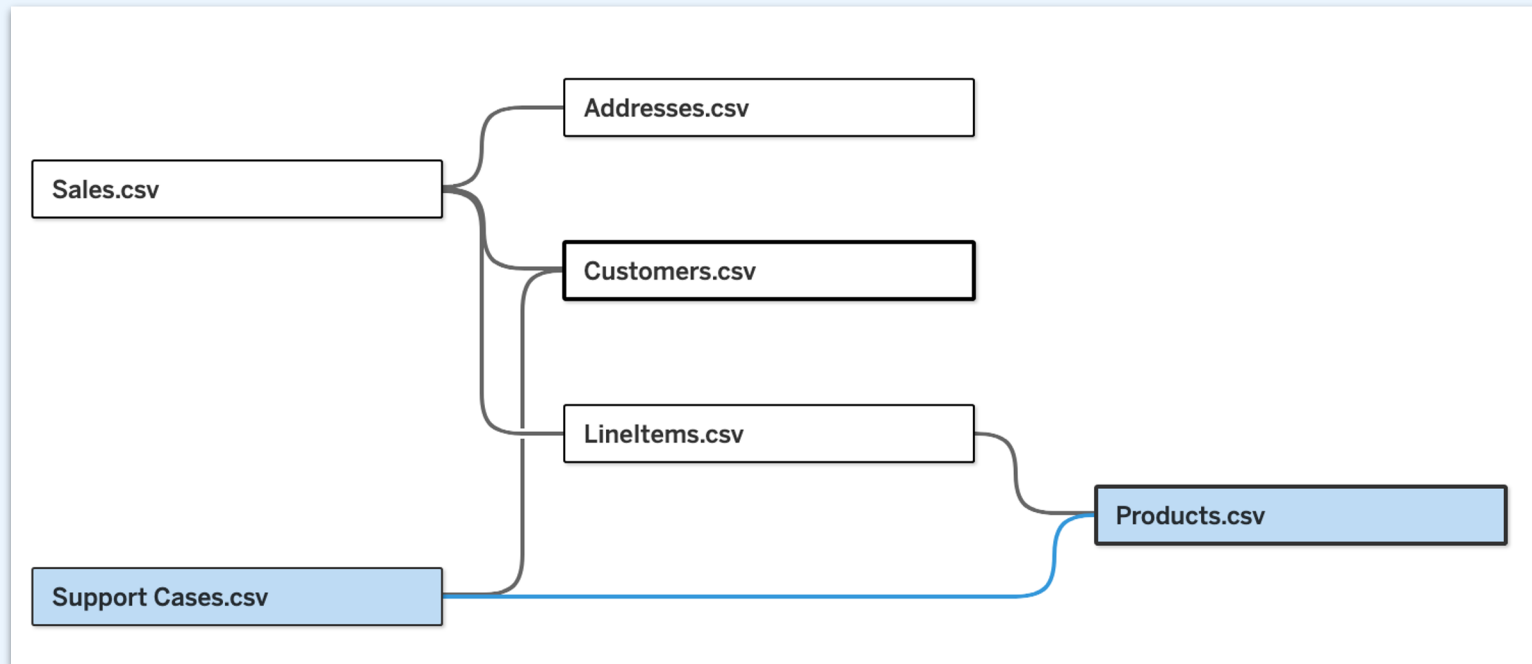


무엇을 해도 단점이 명확한데...



“다중 관계는 하나의 테이블이 여러 테이블과 결합을 맺을 수 있다면,  
공통으로 연결을 맺고 관계 기능으로 유연하게 데이터를 분석할 수 있습니다.”

다중관계 관계 다이어그램 예시



“다중 관계는 하나의 연결을 유지한채로 테이블간의 중심을 변경하면서 더 다양한 분석을 수행할 수 있도록 합니다.”



- 1 “재고가 많이 남는 제품이 피드백이 안좋은가?”
- 2 “지금 판매하고 있는 \*\*\* 제품 개발 개선시점은 언제지?”
- 3 “지금 \*\*\* 품목 너무 잘팔리는데 생산은 몇대가 들어갔지?”
- 4 “올해랑 지난년도 \*\*\* 제품 매출이 좋은데 AS는 얼마나 들어오고 제품 후기 피드백은 어떻게?”

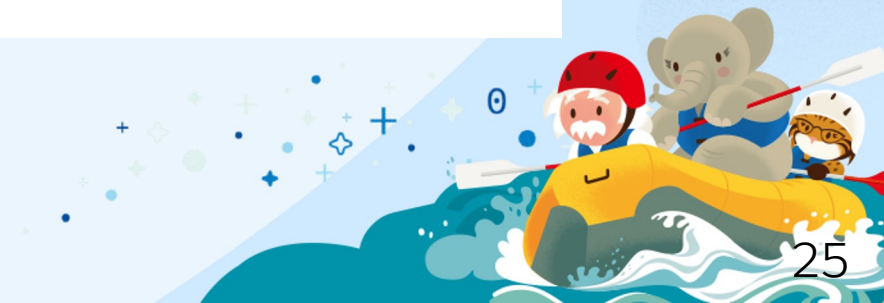


# 다중 관계도 필요한 데이터 기반으로 SQL을 만듭니다.

“다중 관계는 여러 팩트 테이블을 분석할때 먼저 선택한 중심 테이블 기반으로 데이터를 결합합니다. 따라서 중심 테이블에서 없는 데이터는 NULL 처리를 진행합니다.”

빅쿼리에서 뽑아낸 실제 SQL

```
SELECT `t0`.`Product_ID` AS `Product_ID__Products_`,  
       SUM(`LineItems`.`Sales`) AS `sum_Sales_ok`  
FROM `secret-descent-288907.Tableau_Data_Modeling`.`LineItems` `LineItems`  
LEFT JOIN (  
  SELECT `Products`.`Product_ID` AS `Product_ID`  
  FROM `secret-descent-288907.Tableau_Data_Modeling`.`Products` `Products`  
  GROUP BY 1  
) `t0` ON (`LineItems`.`Product_ID` = `t0`.`Product_ID`)  
GROUP BY 1
```



# 어떤 테이블을 중심으로 분석하는지 주의가 필요합니다.

빅쿼리에서 뽑아낸 실제 SQL

```
SELECT `t0`.`Product_Name` AS `Product_Name`,  
       `t0`.`Sentiment` AS `Sentiment`,  
       SUM(CAST(1 AS INT64)) AS `cnt_Support_Cases_22CC6A68699D4076BF0468323ABD2479_ok`  
FROM `secret-descent-288907.Tableau_Data_Modeling`.`Support_Cases` `Support_Cases`  
INNER JOIN (  
  SELECT `Support_Cases`.`Product_ID` AS `Product_ID__Support_Cases_`,  
         `Support_Cases`.`Case_ID` AS `Case_ID`,  
         `Products`.`Product_Name` AS `Product_Name`,  
         `Support_Cases_Detail`.`Sentiment` AS `Sentiment`  
  FROM `secret-descent-288907.Tableau_Data_Modeling`.`Support_Cases` `Support_Cases`  
  LEFT JOIN `secret-descent-288907.Tableau_Data_Modeling`.`Products` `Products` ON  
    (`Support_Cases`.`Product_ID` = `Products`.`Product_ID`)  
  LEFT JOIN `secret-descent-288907.Tableau_Data_Modeling`.`Support_Cases_Detail`  
    `Support_Cases_Detail` ON (`Support_Cases`.`Case_ID` = `Support_Cases_Detail`.`Case_ID`)  
  GROUP BY 2, 1, 3, 4  
) `t0` ON ((IFNULL(`Support_Cases`.`Product_ID`, "") = IFNULL(`t0`.`Product_ID__Support_Cases_`, "")) AND  
(IF(`Support_Cases`.`Product_ID` IS NULL, 0, 1) = IF(`t0`.`Product_ID__Support_Cases_` IS NULL, 0, 1)) AND  
(IFNULL(`Support_Cases`.`Case_ID`, "") = IFNULL(`t0`.`Case_ID`, "")) AND (IF(`Support_Cases`.`Case_ID` IS  
NULL, 0, 1) = IF(`t0`.`Case_ID` IS NULL, 0, 1)))  
GROUP BY 1,
```



# Summary

태블로는 간단한 UI/UX로 쉽게 관계를 설정할 수 있다.

JOIN은 관계를 정의하고 완전한 테이블을 먼저 만들고 시작할 수 있다.

Relationship은 유연하게 내가 분석하는 중심에 따라 관계를 설정한다.

Shared Dimension은 두개 이상의 팩트 테이블과 다수의 마스터 테이블이 존재한다면 다중 연결을 통해 분석에 바로 활용할 수 있다.



# 담당자 A씨, 태블로 고수가 된 후..

태블로를 잘 썼더니!?



- 1 야근을 하지 않고 칼퇴 했는데 새로운 인연을 만났습니다
- 2 다중 분석으로 인사이트를 찾고 팀장님을 만족시켰습니다.
- 3 새로운 데이터 모델링 스킬을 전파하여 회사내 주역이 되었습니다.
- 4 데이터 활용을 통한 이익창출과 비용감소의 효과로 승진하였습니다.

담당자 A씨



# Thank you

