



Team Final Project

Team 2

Fanshu Li, Jaymie Basilio, Vikram Singh,
Yuzhen He, Fang Zhang, & Nishant Sharma





Contents

1. Company Overview
2. Bus / Opportunity / Prioritization Matrices
3. Star Schema & Conformed Dimensions
4. Transformation Rules & Aggregate Tables
5. Cube
6. Visualizations / Dashboards

Company Introduction





Amazon

- Founded in 1994, it is the largest internet retailer in the world.
- Started as an online bookstore, they later diversified into electronics, apparels, toys, food, games and a lot more.
- It is the world's largest provider of cloud infrastructure services.
- With 566,000 employee strength, it is the second largest employer in the United States.
- As of 2017, it has 131.31 billion dollars in assets.



DW-BI System Implementation

How is a Data Warehouse - Business Intelligence System helpful to Amazon?

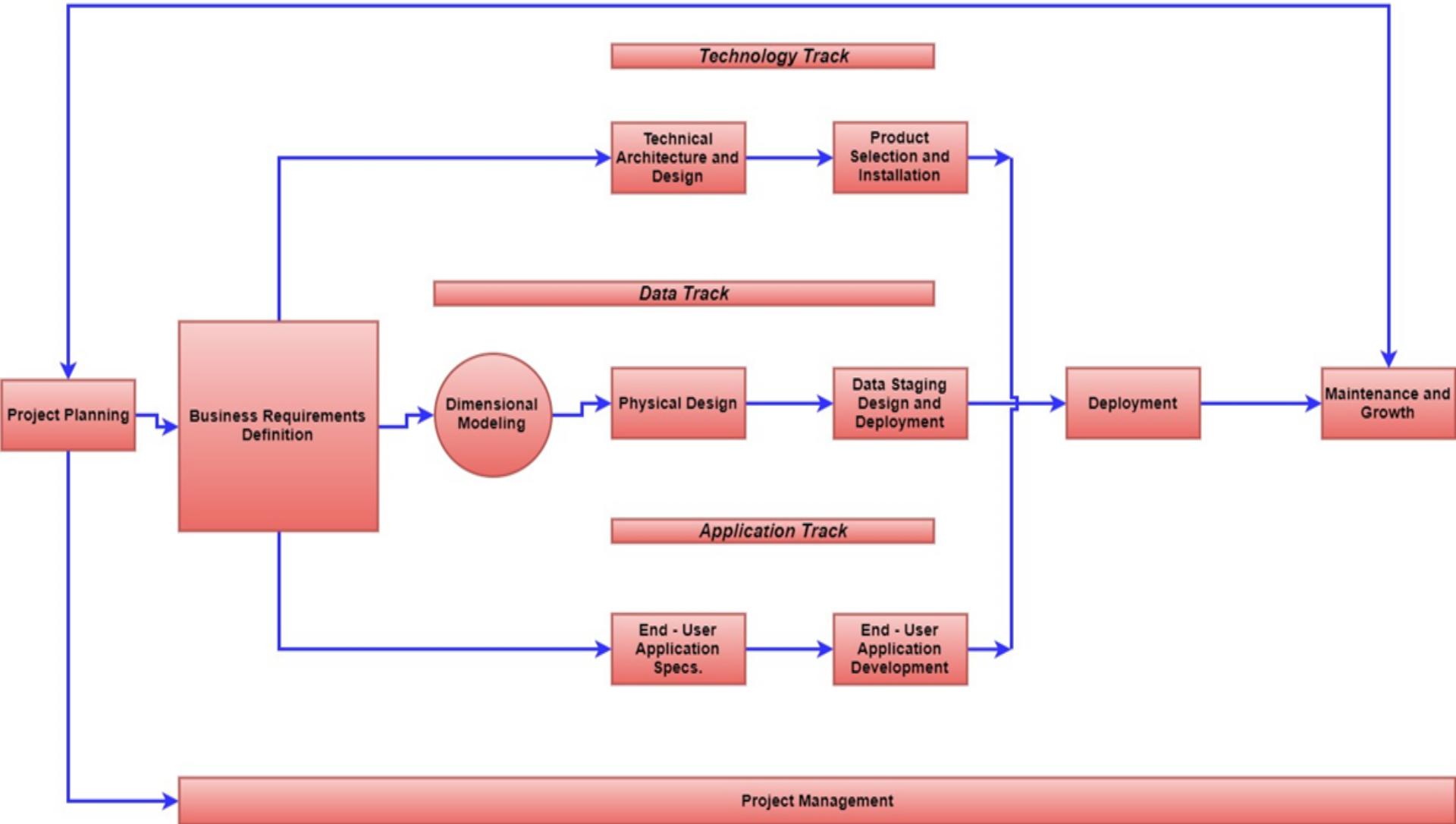
A Data Warehouse - Business Intelligence System can help Amazon in multiple ways discussed as follows:

- Help manage their inventory of vast variety of products
- Achieve proficiency in Strategic Decision making at all levels
- Promoting products and achieving sales goals
- Managing vendors
- Maximize business profits and analyzing their true potential
- Tracking progress using Dashboards and Visualizations



Bus / Opportunity / Prioritization Matrices

Business Dimension Lifecycle





Business Dimensional Lifecycle Bus Matrix

Business Process	Date & Time	Vendor	Product	Logistics	Transaction	Customer	Media	Service Policy	Status & Types	Courier
Procurement	X	X	X	X	X				X	X
Manufacturing	X	X	X	X						X
Sales	X	X	X		X	X		X	X	
Inventory	X		X					X	X	X
HR	X					X	X	X		
Financial	X		X		X				X	
Payroll	X							X		
Marketing	X						X			
Customer Services						X		X		
Infra / Operations	X	X	X		X	X	X		X	X

DETAILED BUS MATRIX

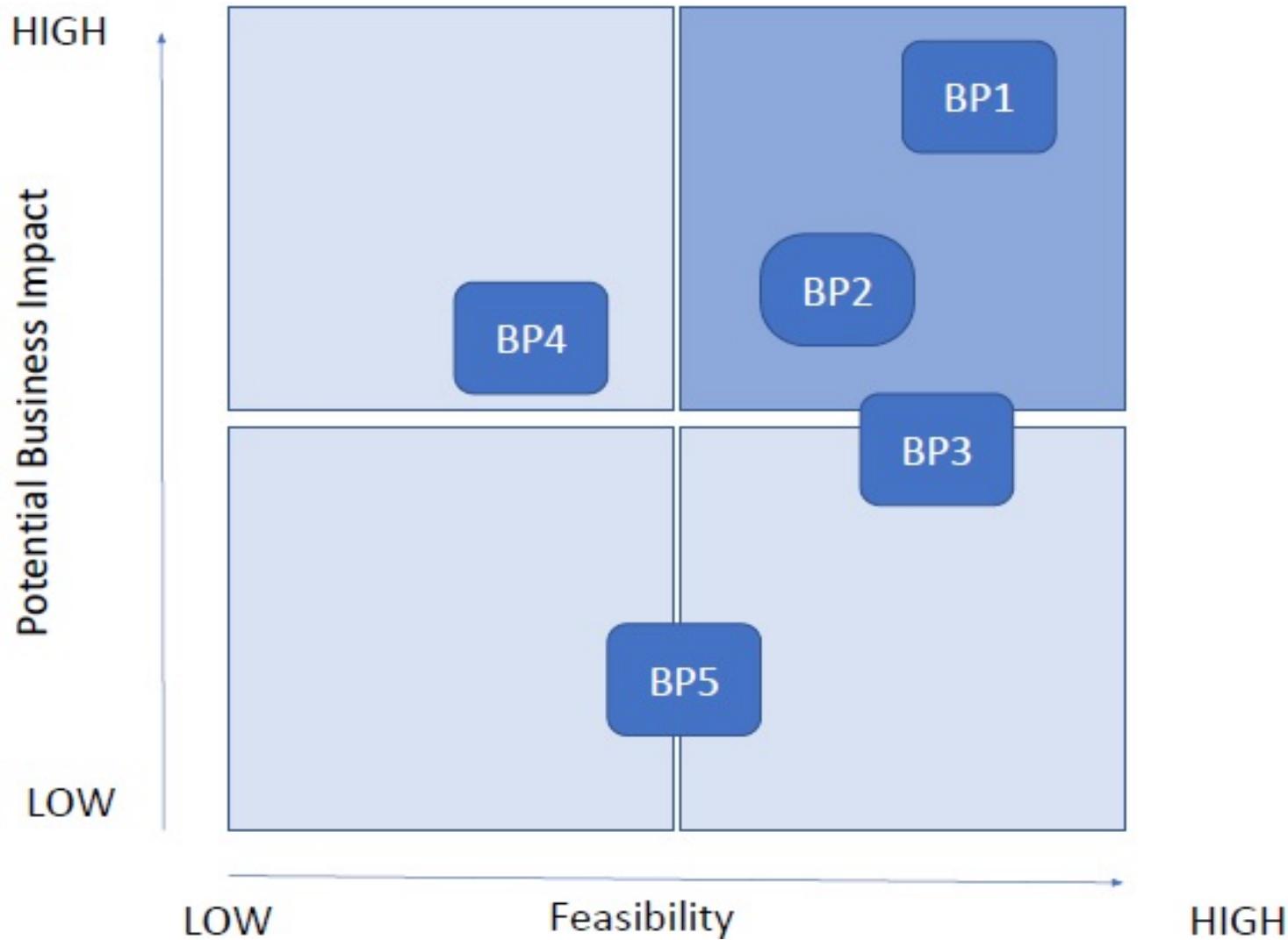
E-commerce business process	Fact tables	Granularity	Fact	Date	Product	Location	Owner	Service Policy	Account Status
Online sales transaction	Sales transaction	Per line item	purchase date key purchase amount key purchase unit price key transaction number	X	X		X	X	X
	Click stream	Per click	referral key session key time key event key page key	X	X			X	X
	Vendor information	Per vendor item	vendor number key vendor product amount key	X	X	X	X	X	
Order processing	Warehouse picking	Per warehouse receipt	ship date key requested date key product key vendor key	X	X		X	X	
	Billing and invoicing	Pre order	date key order number quantity key product key	X	X		X		X
Delivery Process	Shipping notice	Per line item	shipping date key shipping cost key tracking number delivery company key	X	X	X	X	X	X
	Delivery operation	Per line item	tracking number shipping date delivery cost key	X	X	X		X	X
Return Policy	Mail return process	Per line item	tracking number return date key	X	X		X	X	X



Opportunity Matrix - Sales

Business Process	Customer Services	Financial	Operations	Human Resources	Logistics	Strategy Management
Online sales transaction	X	X	X	X	X	X
Vendor Sales		X			X	
Order Processing			X			
Delivery Process			X			
Return Policy	X	X			X	

Prioritization Grid





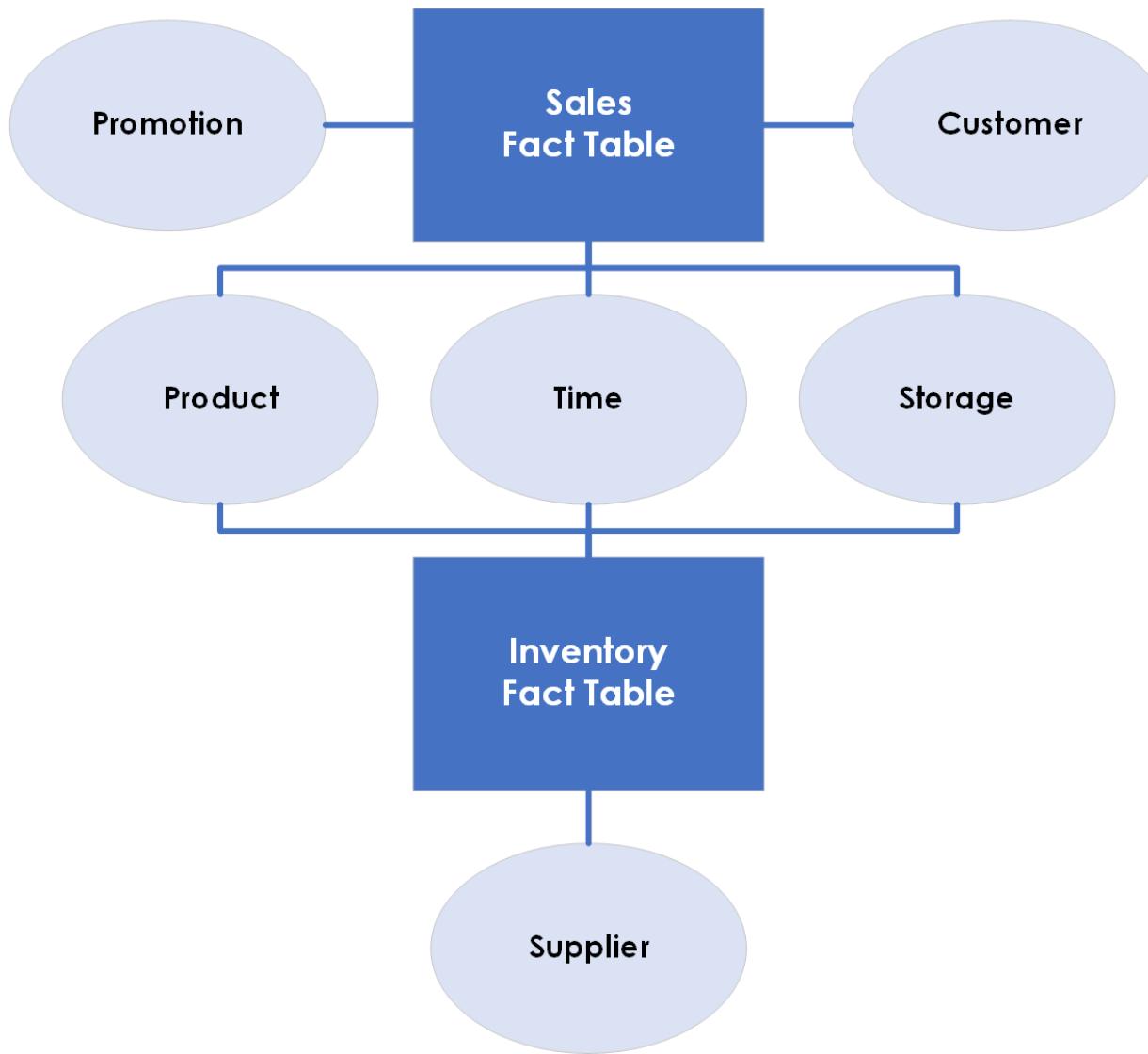
Dimensional Modelling: 4 Key Steps

- Identify the Business Process – There is a series of business processes that each and every organization follows (in this case, package tracking, debit/credit card payment transactions, product ratings, feedbacks). To collect the relevant data, there are various source systems which are relevant to each of the business processes.
- Identify the Grain – Grain is nothing but the lowest level of detail that a single row in the fact table is described in. (e.g., per product, per payment, per rating). For maximum flexibility, we should start with the lowest level of information that is related to the business process.
- Identify the Dimensions – How business people comprehend the data that arises out of a business process is represented by dimensions. A dimension table can be thought of as an entry point to the facts.
- Identify the Facts – Facts are nothing but continuously changing information (e.g., unit quantity, dollar sales, price, etc.). They may also be calculated and/or derived.

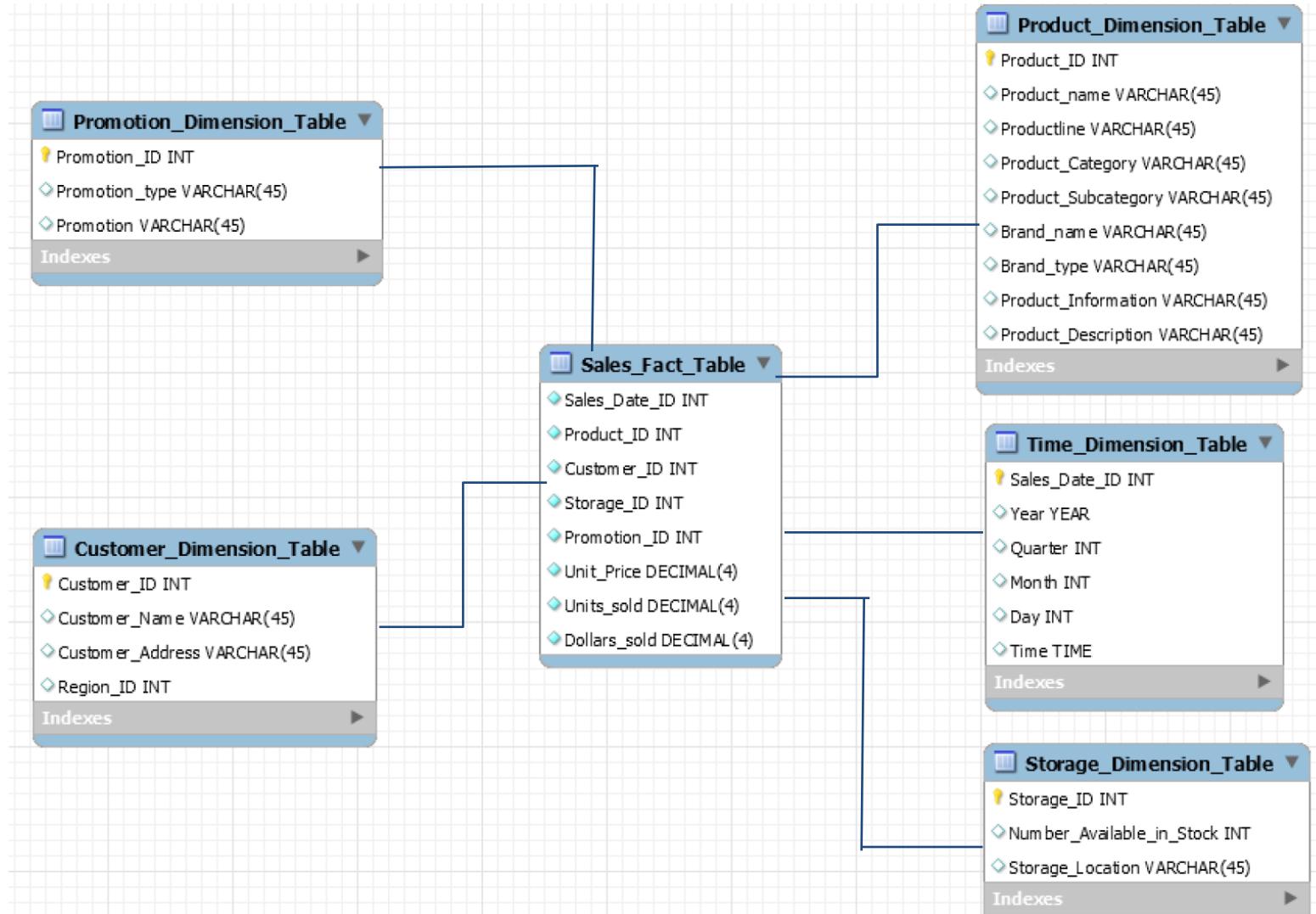


Star Schema & Conformed Dimensions

High Level Model Design

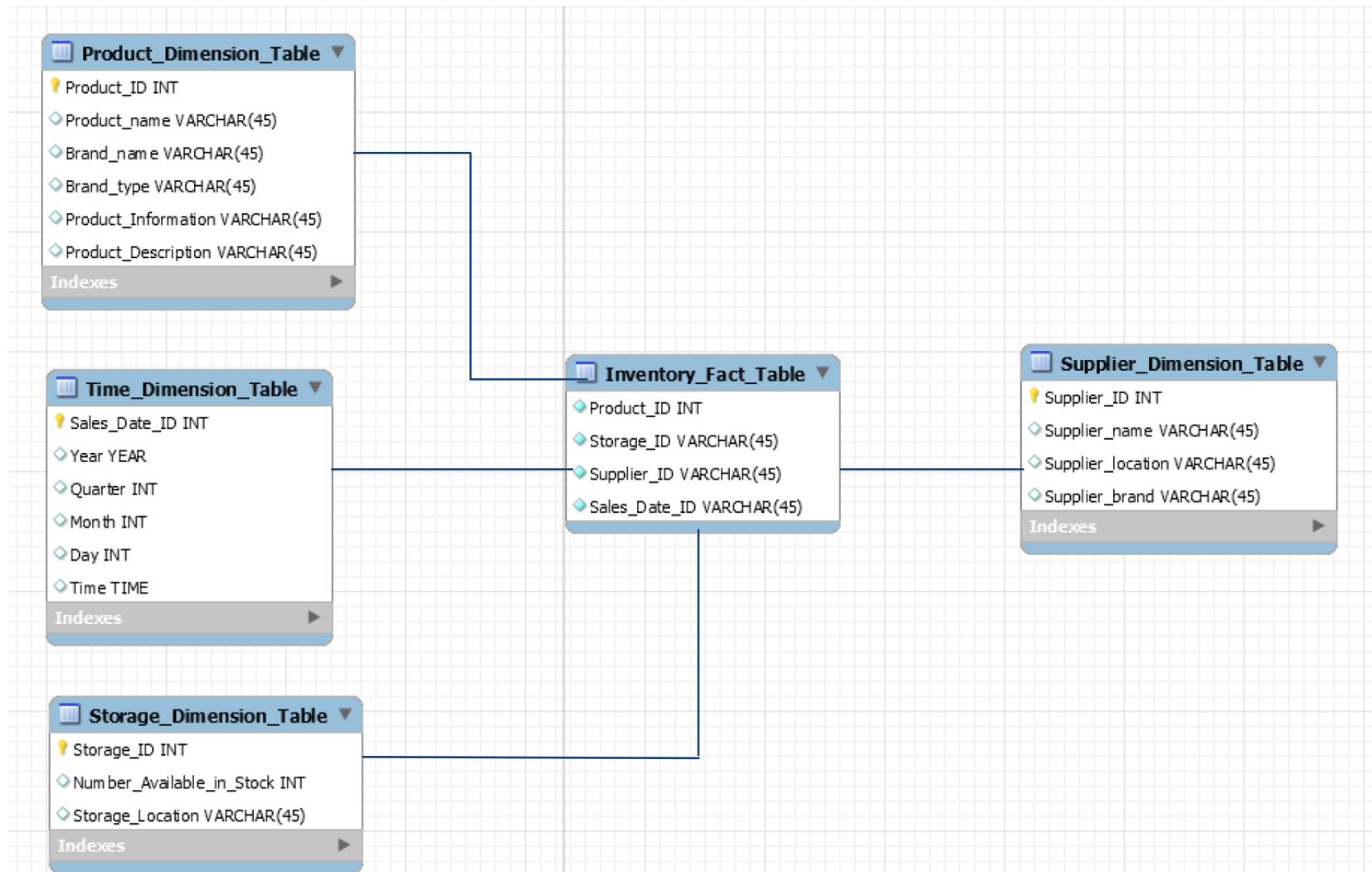


Sales Detailed Diagram

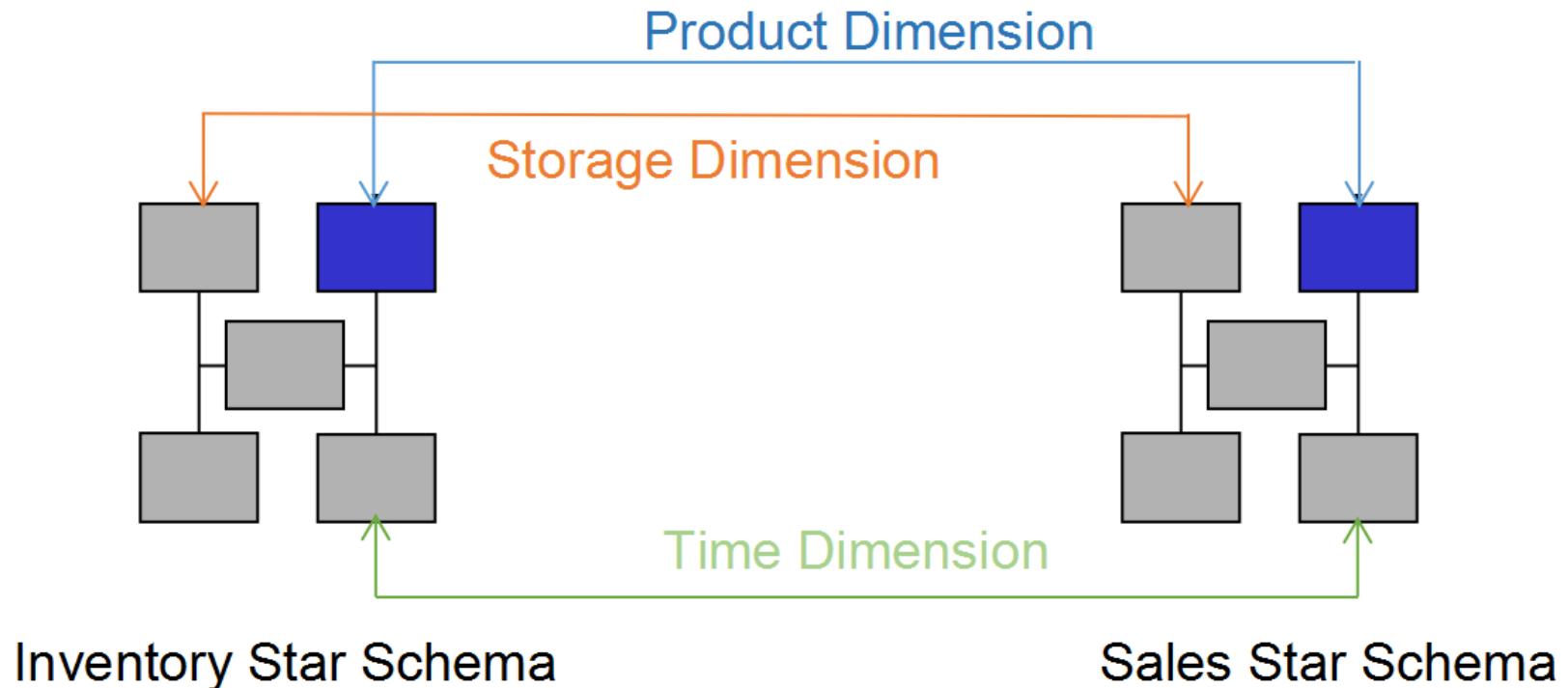




Inventory Detailed Diagram



Conformed Dimensions

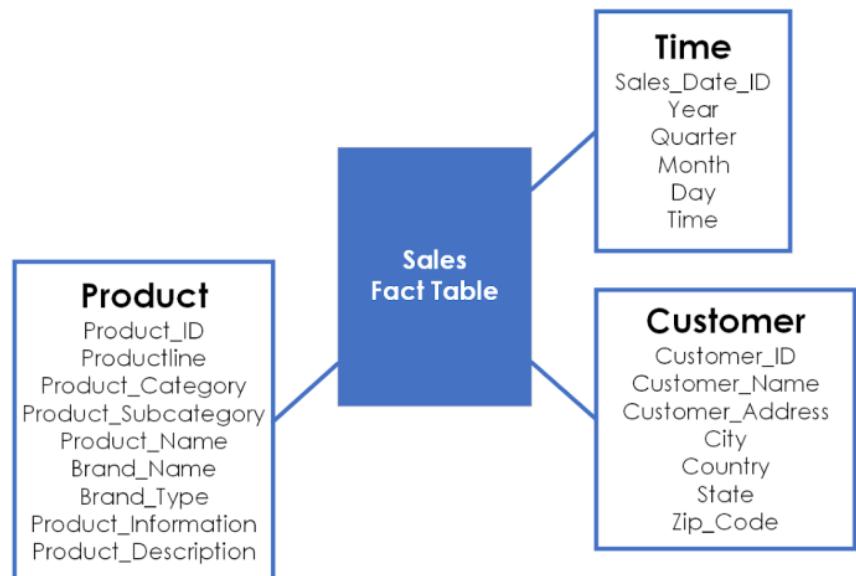




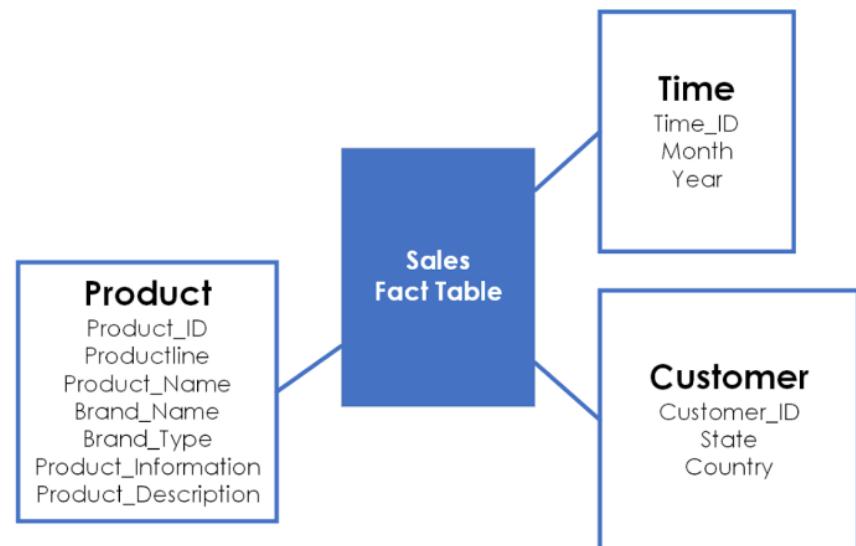
Transformation Rules & Aggregate Tables

Aggregate Tables

Base fact and dimension



Aggregation by region by month





Aggregate Table with Atomic Facts

Sales Transaction	Product				Customer				Time			
	Productline1		Productline2		Country				Year1		Year2	
	Product 1	Product 2	Product 1	Product 2	State 1		State 2		Month 1	Month 2	M1	M2
					City 1	City 2	City 1	City 2	D1	D2	D1	D2
Product SKU												
Warehouse ID												
Promotion ID												
Time ID												
Unit Price												
Sales Price												
Quantity												
...												
...												



Aggregate Table by Product Line by State by Month

Sales Transaction	Product		Customer		Time			
	Productline1	Productline2	Country		Year1		Year2	
			State 1	State 2	Month 1	Month 2	M1	M2
Product SKU								
Warehouse ID								
Promotion ID								
Time ID								
Unit Price								
Sales Price								
Quantity								
...								
...								



Aggregate Table by Product Line by Country by Year

Sales Transaction	Product		Customer	Time	
	Productline1	Productline2		Year1	Year2
Product SKU					
Warehouse ID					
Promotion ID					
Time ID					
Unit Price					
Sales Price					
Quantity					
...					
...					



Transformation Rules

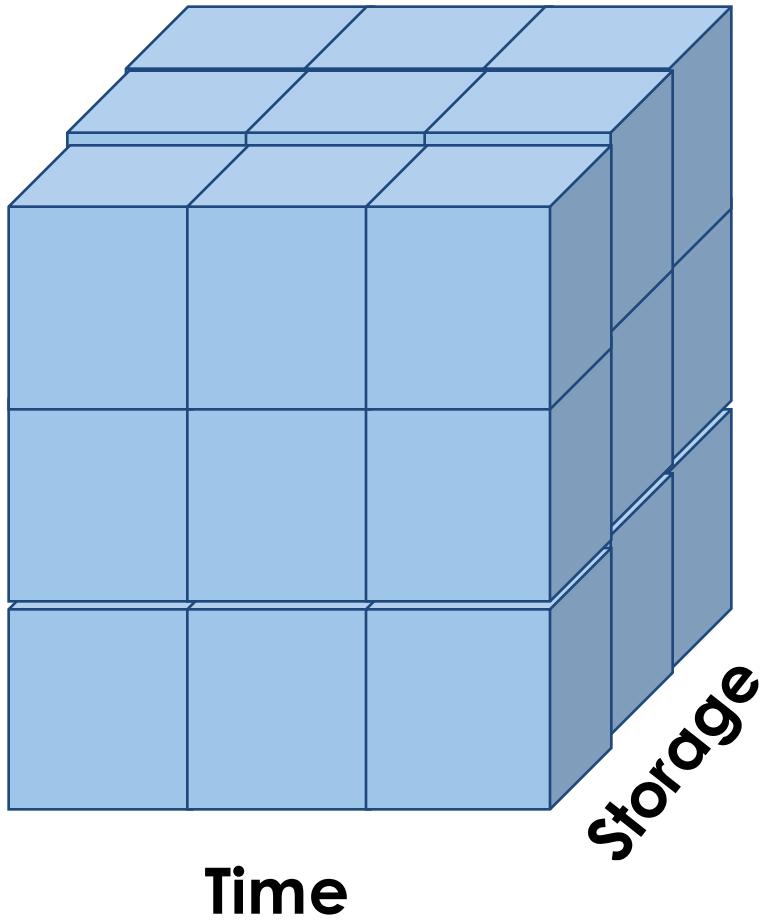
Rule Type	Description
Default	The initial field is empty for every record.
Transform	Source data value will be converted to match the destination format. Ex: Date: DDMMYYYY, Name: First Name & Last Name, Address, Contact, etc.
Consistency	Source data and destination data should have same value.
Range	Range is created into specific columns such as price (>0) or date (1-31) column to ensure value is valid.
Derive	Taxes (depending on local regulation) will be added to derive final amount. All purchased products amount will be summarized as total cost.
Drop	Source data field is not allowed to be inserted to the destination table.



Cube

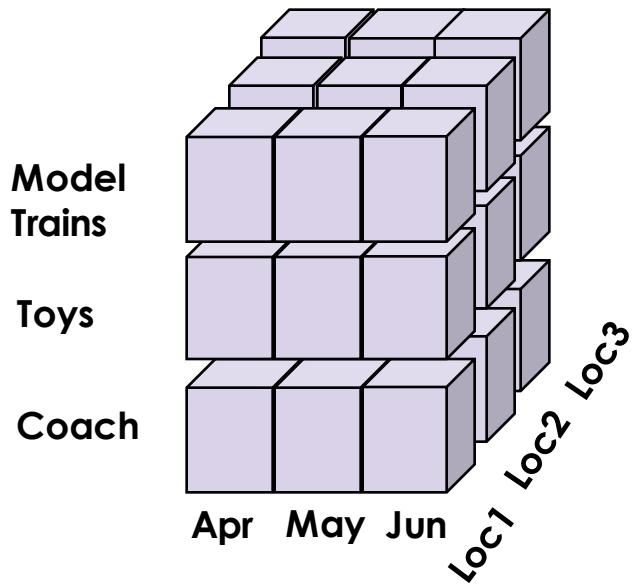
OLAP Cube

Product



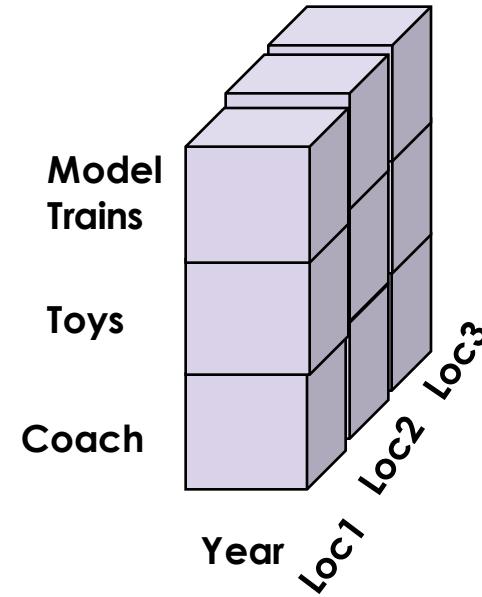
This OLAP Cube has 3 dimensional arrays of data:

- Product
- Time
- Storage



Drill-down

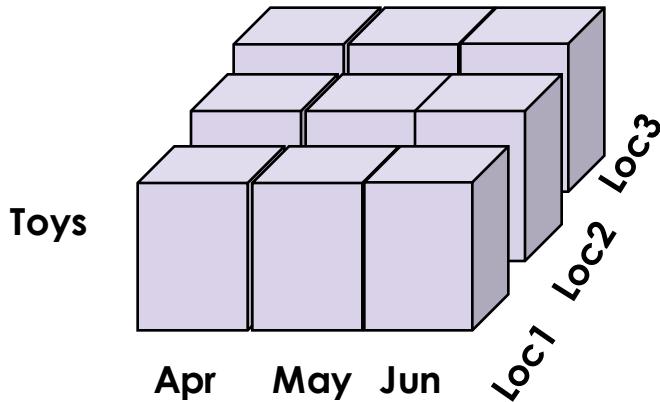
- Drill-down allow user to navigate data in the most detailed level.
- In this case, we could scan through particular products such as model trains, toys and coach, among with other dimensions.



Roll-Up

- Roll-up moves from detailed level to summary category.
- In this case, time dimension goes from month to year.

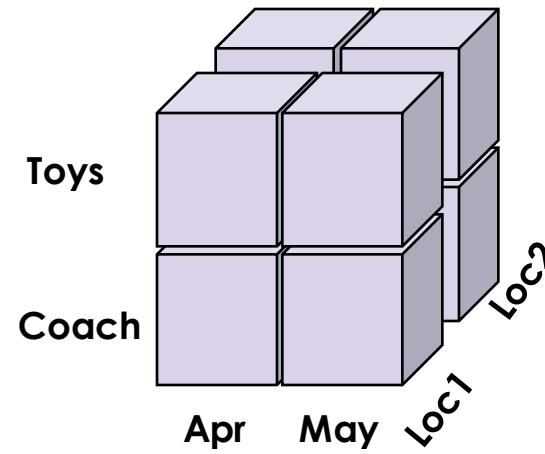
I II



Slice

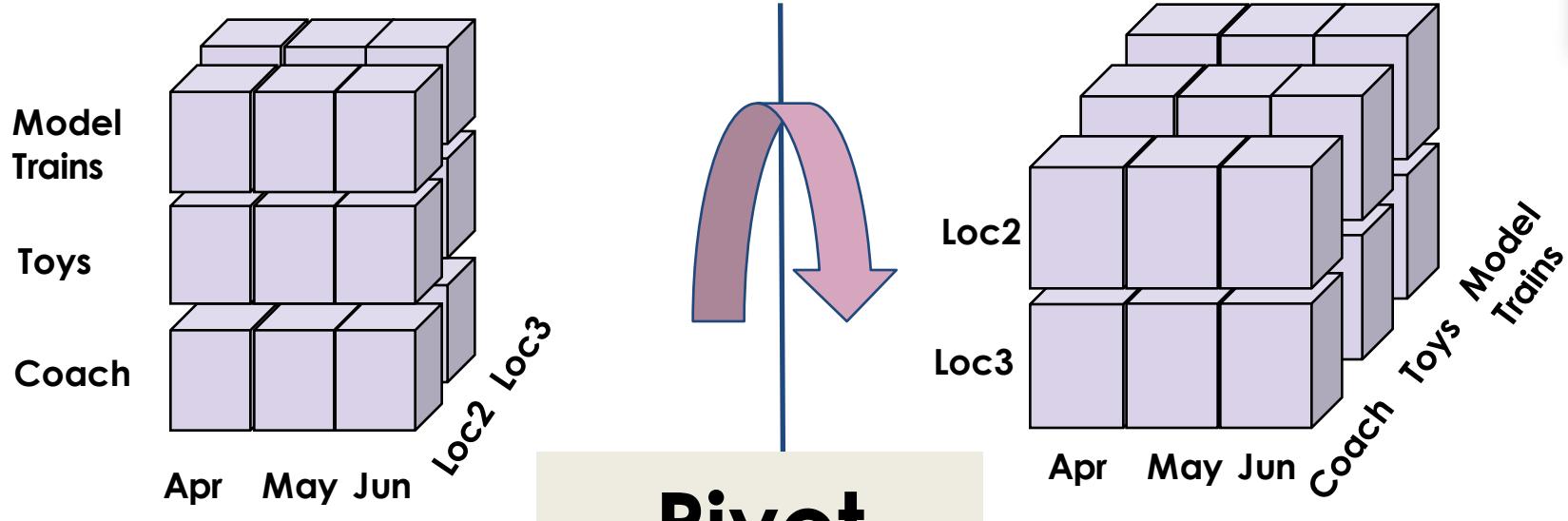
- Slice is the act of picking a rectangular subset of a cube by choosing a single value for one of its dimensions
- In this case, toys are chosen as unique dimension in process of slicing.

III IV



Dice

- The dice operation produces a subcube by allowing the analyst to pick specific values of multiple dimensions.
- In this case, user want to specifically look over products that are toys and coach, stored in location 1 and location 2, during April and May.



- Pivot allows user to rotate the cube in space to see its various faces. That's very useful in term of viewing data in different perspective.
- For example, you could check out three products in different period of time and in different location, or you could check out storage situate by viewing different products in different period of time.

Users & Tasks

- Task analysis is the process of learning about Amazon users by observing them in action to understand in detail how they perform their tasks and achieve their intended goals.

User Role	Delivery	Task Analysis
Executives/Manager/ Decision makers	Portal → report emailed Low effort/push button/ report pushed on regular basis	Company's overall situation/Total Business operation and earning performance
Business/ Data Analyst	Portal/ MOLAP Advanced ad-hoc query	Specific task analysis by extracting data from DB
Operational Staff	Portal / Operational Report Access to data details and create report	Intend to support day-to-day activities of Amazon company



Visualizations / Dashboards

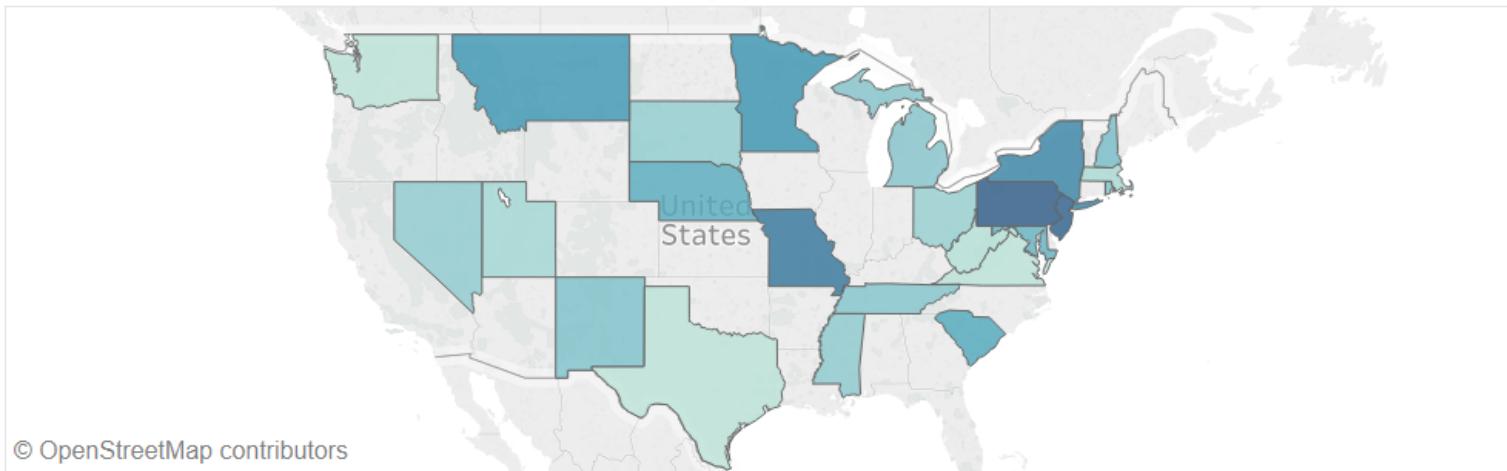


Tableau Public Dashboard URL

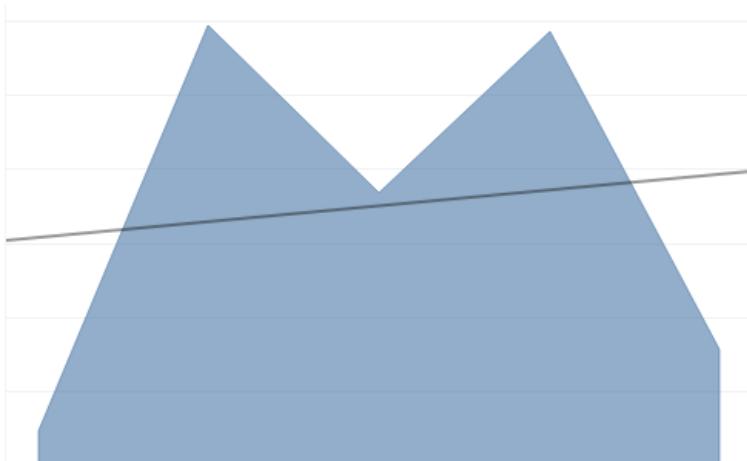
https://public.tableau.com/views/MIS636A-FinalProject/MgmtDashboard?:embed=y&:display_count=yes&publish=yes

U.S. Sales Management Dashboard

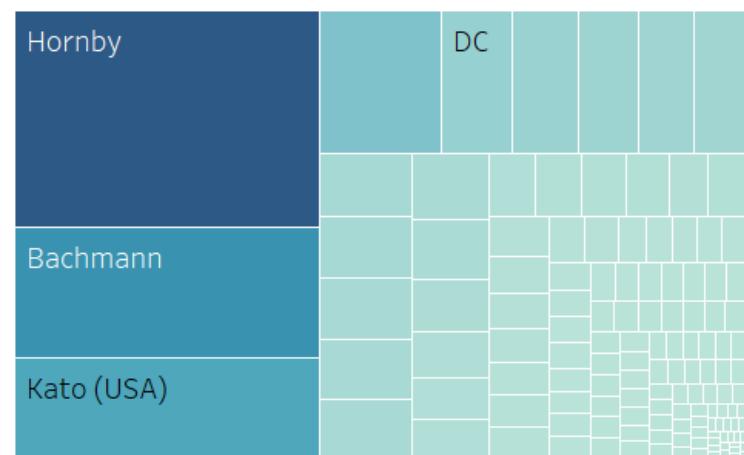
Total Sales by State



Total Sales by Year

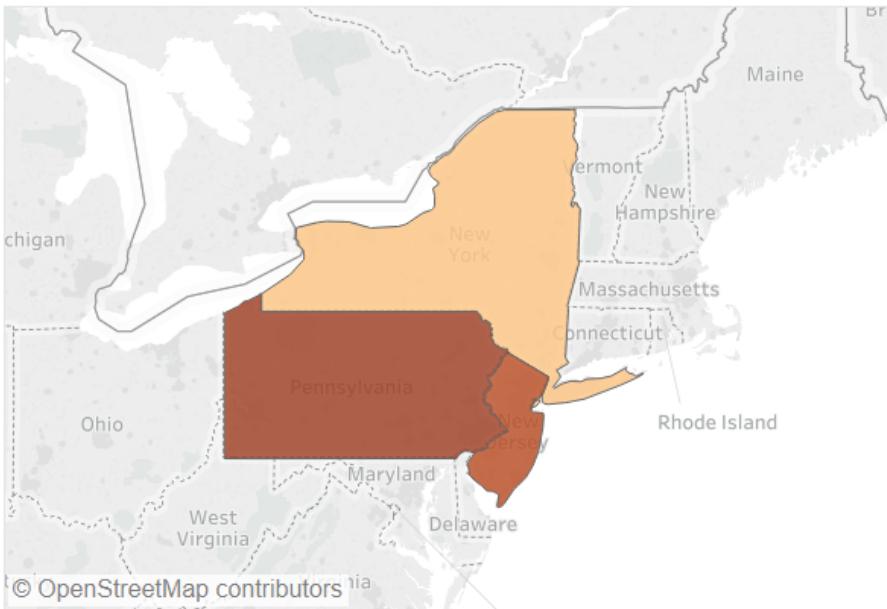


Top Suppliers

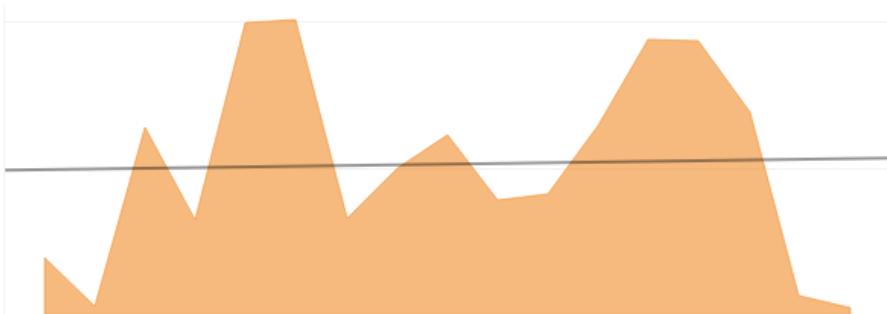


Regional Sales Manager Dashboard

Total Sales by State



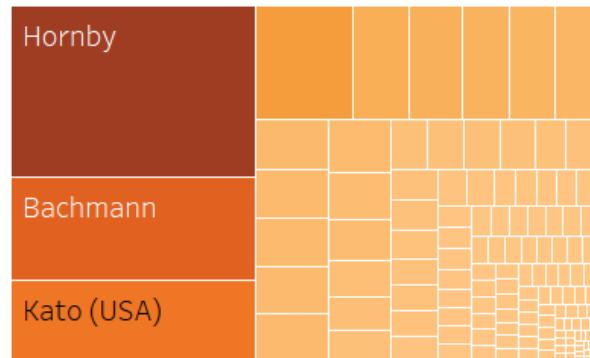
Total Sales by Quarter



Top Selling Product Categories



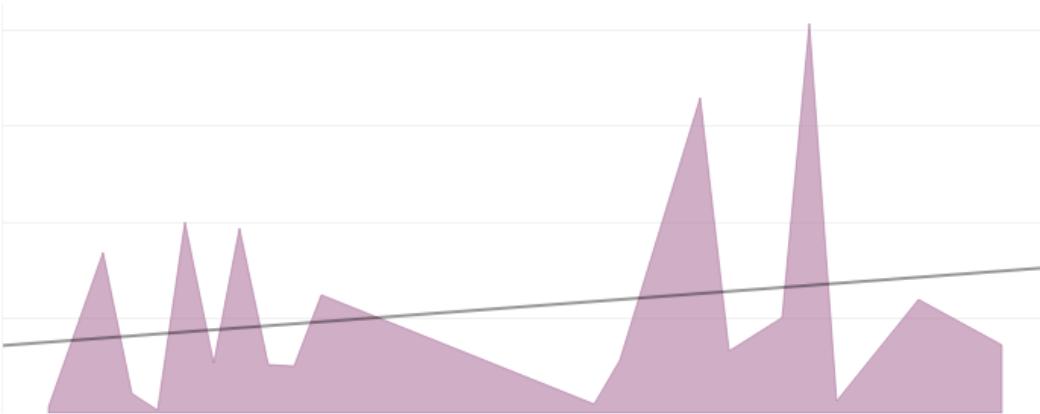
Top Suppliers



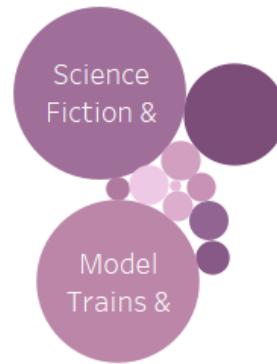


NJ Sales Analyst Dashboard

Total Monthly Sales in NJ



Top Product Sub-Categories

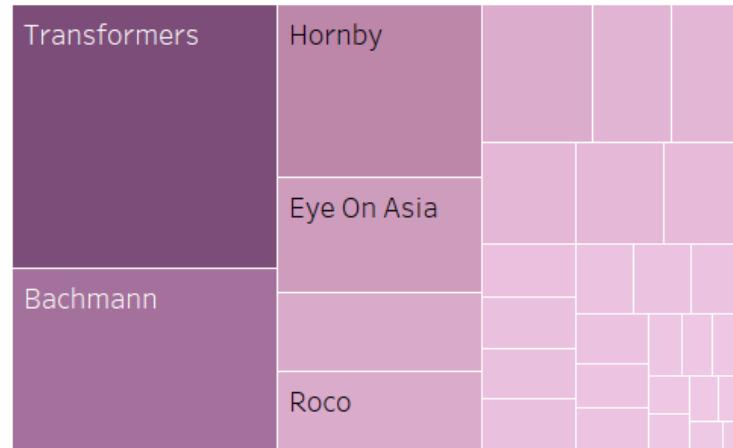


Top Selling Products in NJ

Product Name (S..)

Transformers M..	▲
Bachmann Class ..	
Docolla: Charact..	
Hornby R3086 R..	
Vintage Tractor ..	
Roco 64723 OBB ..	
Super Figure mo..	
HORNBY Coach R..	
Ford Capri 3.0S - ..	
Playmobil 5995 F..	
Preiser 30495 H..	
Forces of Valor 1..	
Transformers Ag..	
Pokemon - Seale..	
ixo silver nissan ..	
Plastoy Tubo My..	
Coloured Craft S..	

Top Suppliers in NJ





STEVENS
INSTITUTE *of* TECHNOLOGY
THE INNOVATION UNIVERSITY®

stevens.edu

Thank you.