

# Business Intelligence and Data Warehousing (ANL408)

- By Sabarish Nair

# Recap from last week....

- Relational Database
- Primary Key – Foreign Key Relationship
- Dimensional Modeling
- Fact table
- Dimensions
- Star Schema
- Snowflake Schema
- Star Schema vs Snowflake Schema
- Install PostgreSQL

# Addivity

---

- Property of measures in fact table.
- Determines how measures aggregate or sum up when querying the fact table at different levels of granularity.
- Three Types:
  - Additive
  - Semi-additive
  - Non-additive



# Additive Measure

- Can be added/Summed up across all dimensions
- Flexible and Useful
- Most Facts are Fully Additive

Sales_id	Product_id	Units	Price
1	3	2	5
2	2	3	5
3	4	3	2
4	3	2	1

Product_id	Name	Category
2	Bat	Toy
3	Ball	Toy
4	Chips	Food

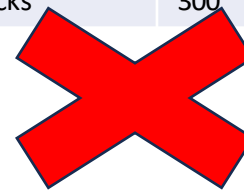
# Semi-additive measure

- Can be added/Summed up across a few dimensions
- Less Flexible
- Alternative: Average

Balance_id	Portfolio_id	Date	Balance
1	1	29-Feb	\$50
2	1	29-Feb	\$100
3	2	1-March	\$100
4	2	2-March	\$200

Portfolio	Type
1	Cash
2	Stock

Type	Balance
Cash	150
Stocks	300



# Non-additive measure

- Cannot be added up across any dimension
- E.g. Ratios, Percentages, etc.

Sales_id	Product_id	Units	Price
1	3	2	5
2	2	3	5
3	4	3	2
4	3	2	1

Product_id	Name	Category
2	Bat	Toy
3	Ball	Toy
4	Chips	Food

Category	Price
Toy	11
Food	2



Revenue  
= Units \* Price



# Types of Fact Tables

- Transactional
- Periodic Snapshot
- Accumulating Snapshot



# Transactional Fact table



Also known as an event fact table or atomic fact table.



Stores detailed transactional data at the lowest level of granularity.



Each row represents a single transaction or event.




Contains foreign keys to link to various dimension tables.

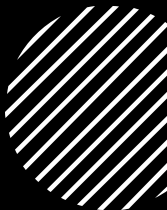



Examples include sales transactions, purchase orders, website clicks, etc.





# Periodic Snapshot Fact Table



Stores data captured at specific intervals or periods (1 week, year, etc.).



Typically contains aggregated data rather than individual transactions (No events are null or 0).



Represents a snapshot of the business at predefined time intervals.



Contains foreign keys to link to time dimension and other relevant dimension tables.



Examples include daily sales totals, monthly inventory levels, quarterly revenue, etc.

# Accumulating Snapshot



Tracks the progression or status of an ongoing process over time.



Contains data that is updated periodically to reflect the state of the process.



Contains a mix of aggregated and detailed data related to the process.



Typically contains multiple timestamp columns to capture different stages of the process.



Examples include order fulfillment process (order received, order shipped, order delivered), etc.

# Factless Fact table

---

- Events without any metrics
- Different combinations of dimension keys are analyzed to identify patterns, trends
- Types of Relationships:
  - Events or occurrences: e.g., student enrollment, customer visits, order placement.
  - Conditions or constraints: e.g., product availability, eligibility criteria.
  - Processes or workflows: e.g., order processing stages, customer journey stages.

Emp ID	Emp Name	Joined Date	Manager ID
1	John	2 March	2
2	Jill	3 March	4

Employee Registration

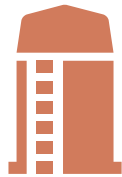
**No Metrics, Just Events!!**

# Steps to create a fact table

---



Identify business  
process for analysis



Define the grain  
(level of detail)



Identify relevant  
dimensions



Identify the facts  
for measurement

# Natural vs Surrogate Key

- Artificial Key
- Generated by System
- Purpose:
  - Uniqueness
  - Stability
  - Performance
  - Simplicity

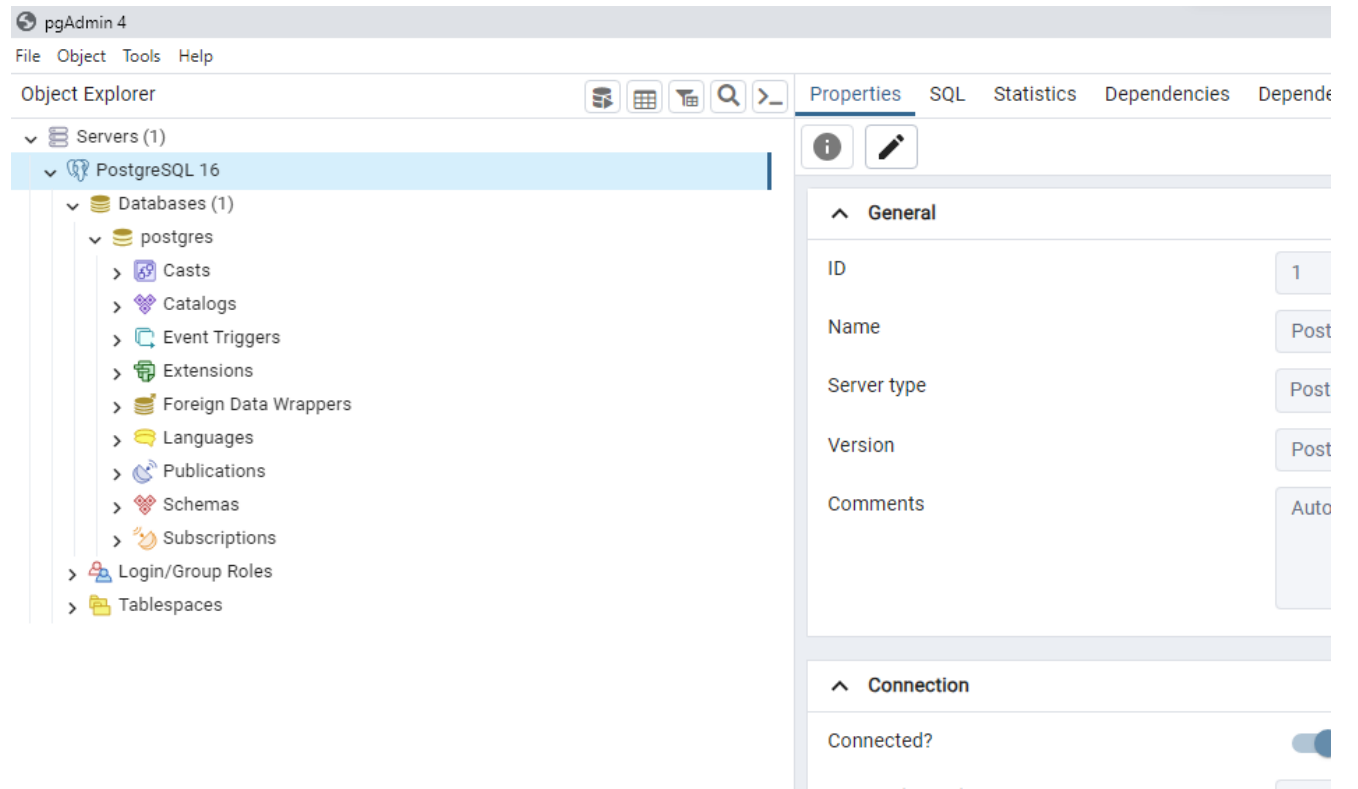
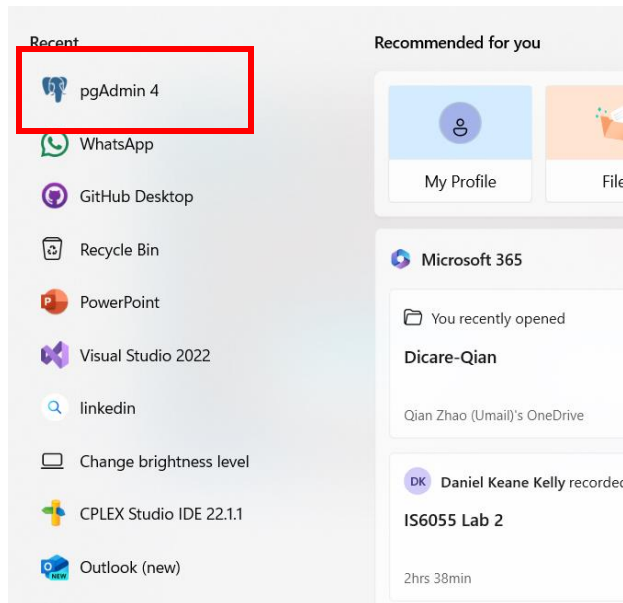
Product_ID	Product_Name	Category
PRD_101	Bat	Toy
PRD_102	Ball	Toy

Natural key

Product_PK	Product_ID	Product_Name	Category
1	PRD_101	Bat	Toy
2	PRD_102	Ball	Toy

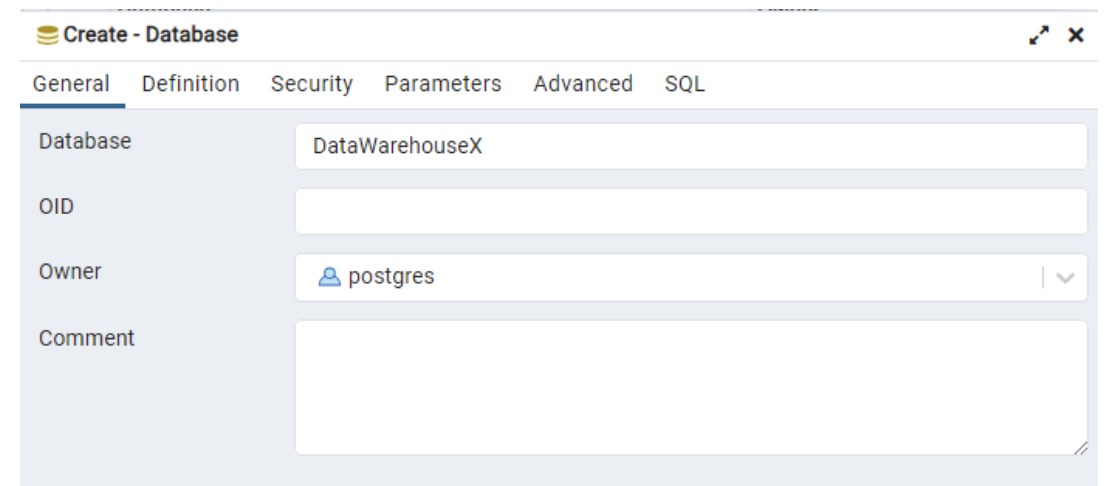
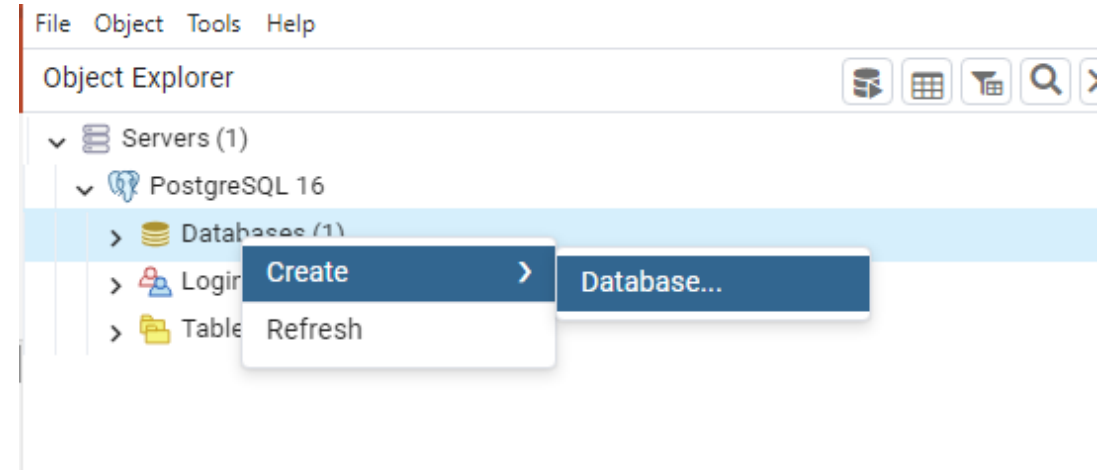
Surrogate key

# Launch Postgres SQL

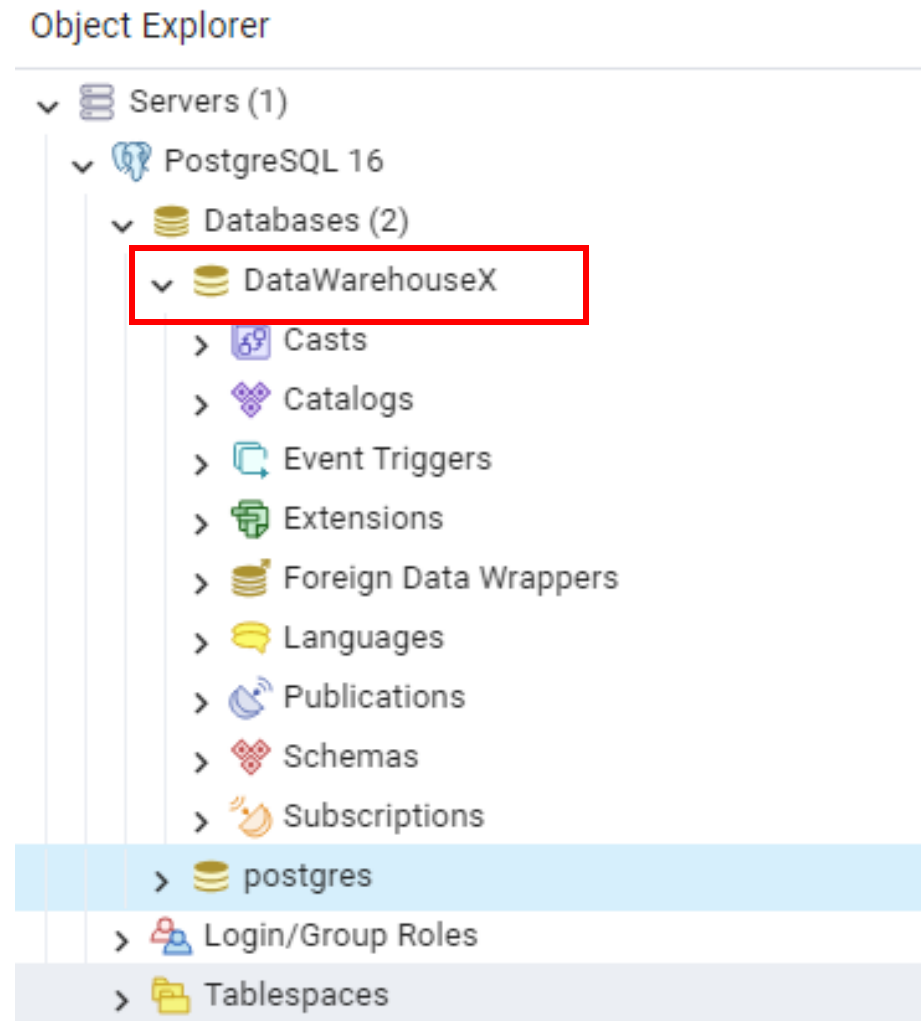


# Create a new database

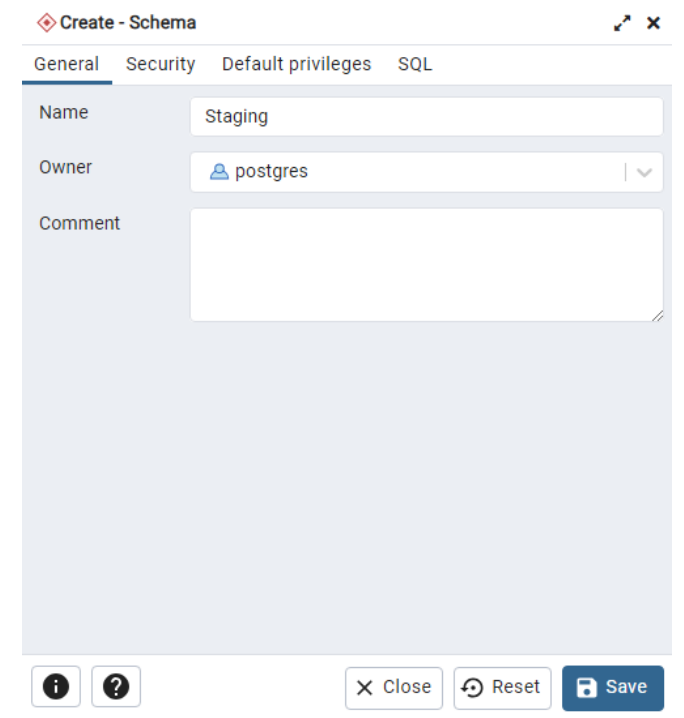
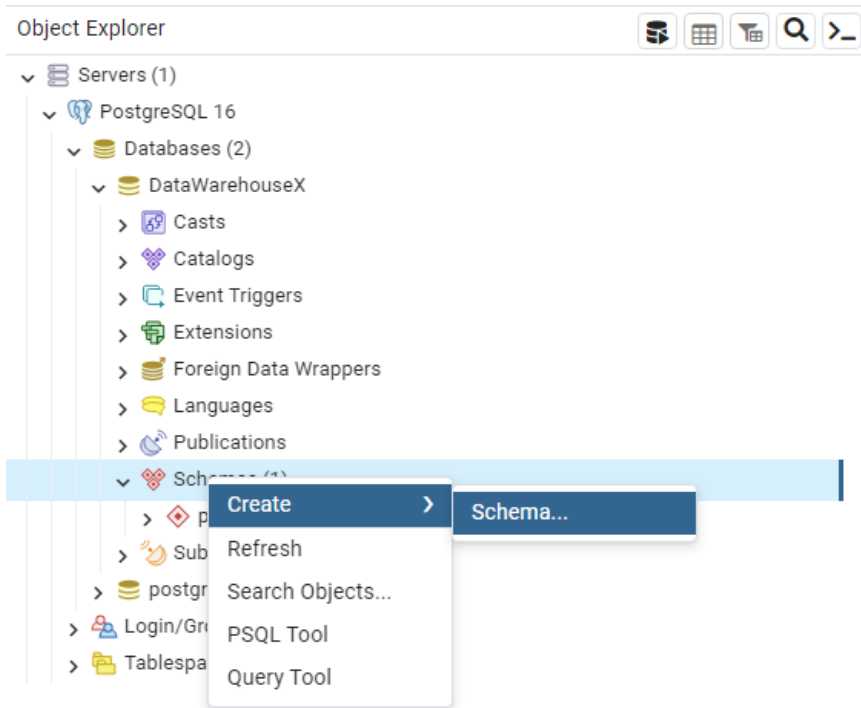
- Right-click on the databases and create a new database
- Give a name to the database



# Verify the new database







Add a staging Schema

Right-click on "Schemas" under the newly created database and create a new schema



*Thank you*