Let Us Start With The Concepts of Dimensional Modeling

SkyEss Techno Solutions Pvt. Ltd.

- Dimensional Modeling is A Set of Techniques And Concepts Used in Data Warehouse Design.
- Dimensional Modeling is Considered To Be Different From Entity-Relationship Modeling As Dimensional Modeling Does Not Necessarily Involve A Relational Database.
- Dimensional Modeling is A Design Technique For Databases Intended To Support End-User Queries in A Data Warehouse.
- Dimensional Modeling is Oriented Around Understandability of The Patterns of Data And Query Performance.
- Dimensional Modeling Always Uses The Concepts of Facts Called As Measures, And Dimensions Called As Context.
- Facts Are Typically, (But Not Always) Numeric Values That Can Be Aggregated.
- Dimensions Are Groups of Hierarchies And Descriptors That Define The Facts.
- Dimensional Models Are Built By Business Process Area Called As Subject Areas of The Business Analysis.
- Dimensional Modeling Should Conceptualize And Make Data Models Visualized As A Set of Measures That Are Described By Common Aspects of The Business.
- Dimensional Modeling is Useful For Summarizing And Rearranging The Data And Presenting Views of The Data To Support Data Analysis.
- Dimensional Modeling Focuses Mainly on Numerical Data.

SkyEss Techno Solutions Pvt. Ltd.

Dimensional Modeling Process

- The Dimensional Model is Built on A Star-Like Schema, With Dimensions Surrounding The Fact Table.

Stages To Build A Dimensional Modeling Schema

- Choose The Business Process
- Declare The Grain
- Identify The Dimensions
- Identify The Fact

Choose The Business Process

- Identify The Actual Business Process Upon Which The Data Warehouse Should Be Covered.
- Describe The Business Processes That Are All Important To The Pilot Process of The Data Warehouse.
- To Describe The Business Process, We Can Choose Plain Text Format OR Use Basic Business Process Modeling Notation (BPMN) OR Any Other Design Guides Like The Unified Modeling Language (UML).

Declare The Grain

- Declare The Grain of The Model Only After The Actual Business Process Has been Identified.
- The Grain of The Model is The Exact Description of What The Dimensional Model Should Be Focusing on For its Implementation in The Data Warehouse.

- To Clarify What The Grain Means, We Should Pick The Central Process And Describe it With One Sentence.
- The Grain Which is The Actual Sentence That is Built is What We Are Going To Build Into Dimensions And Facts in A Table From.
- We Might Find it Necessary To Go Back To This Step To Alter The Grain Due To New Information Gained on What The Model is Supposed To Be Able To Deliver.

Identify The Dimensions

- The Dimensions Must Be Defined Within The Grain From The Second Step of The Four Step Process.
- Dimensions Are The Foundation of The Fact Table, And is Where The Data For The Fact Table is Collected.
- Typically Dimensions Are Nouns, And These Dimensions Are The Actual Store Houses Where All The Data To Be Described is Stored.

Identify The Facts

- After All The Dimensions Are Defined Then Only We Can Make The Process of Keys For The Fact Table To Be Initiated.
- We Identify The Numeric Facts That Will Populate Each Fact Table Row.
- This Step is Closely Related To The Business Users of The System, As This is Where They Get Access To Data Stored in The Data Warehouse.

Benefits of Dimensional Modeling

Understandability

- The Dimensional Model is Easier To Understand And More Intuitive.
- Dimensional Models Group Information into Coherent Business Categories OR Dimensions, Making it Easier To Read And Interpret.
- Dimensional Models Simplicity Allows Software To Navigate Databases Efficiently.

Query Performance

- Dimensional Models Are More De-Normalized And Optimized For Data Querying.
- The Predictable Framework of A Dimensional Model Allows The Database To Make Strong Assumptions About The Data That Aid in Performance Allowing Effective Handling of Complex Queries.

Extensibility

- Dimensional Models Are Extensible And Easily Accommodate Unexpected New Data.
- Existing Tables Can Be Changed in Place Either By Simply Adding New Data Rows into The Table OR Executing SQL Alter Table Commands.
- No Queries OR Other Applications That Sit on Top of The Warehouse Need To Be Reprogrammed To Accommodate Changes, Old Queries Continue To Run As They Exist.

Basic Concepts of Dimensional Modeling

SkyEss Techno Solutions Pvt. Ltd.

- The Basic Concepts of The Dimensional Modeling
  - Facts
  - Dimensions
  - Measures (Variables)

Facts

- A Fact is A Collection of Related Data Items, Consisting of Measures And Context Data.
- Each Fact Typically Represents
  - A Business Item
  - A Business Transaction
  - An Event That Can Be Used in Analyzing The Business OR Business Processes
- In A Data Warehouse, Facts Are Implemented in The Core Tables in Which All of The Numeric Data is Stored.

Dimension

- A Dimension is A Collection of Members OR Units of The Same Type of Views.
- In A Dimensional Modeling Diagram, A Dimension is Usually Represented By An Axis on The Basic Unit Called Cube.
- In A Dimensional Model, Every Data Point in The Fact Table is Associated With One And Only One Member From Each of The Multiple Dimensions.
- Dimensions Determine The Contextual Background For The Facts.

- Many Analytical Processes Are Used To Quantify The Impact of Dimensions on The Facts.
- Dimensions Are The Parameters Over Which The OLAP Users Perform Online Analytical Processing (OLAP).
- Dimensions Can Usually Be Mapped To Non-Numeric, Informative Entities Called As
  - Dimension Members
  - Dimension Hierarchies

Dimension Members

- A Dimension Contains Many Dimension Members.
- A Dimension Member is A Distinct Name OR Identifier Used To Determine A Data Item OR Data Items Position.

Illustrative Example

- Time Dimension → Months, Quarters, And Years
- Geography Dimension → All Cities, Regions, And Countries

Dimension Hierarchies

- We Can Arrange The Members of A Dimension into One OR More Hierarchies.
- Each Hierarchy Can Also Have Multiple Hierarchy Levels.
- Every Member of A Dimension Does Not Locate on One Hierarchy Structure.

Illustrative Example

- Year → Half, Quarter, Month And Day
- Week → Day, Hours, Minutes And Seconds

Measure

- A Measure is A Numeric Attribute of A Fact, Representing The Performance OR Behavior of The Business Relative To The Dimensions.
- The Actual Numbers in The Measures Are Called As Variables.
- A Measure is Determined By Combinations of The Members of The Dimensions And is Located on Facts.

Basic Operations for OLAP

Drill Down And Roll Up

- Drill Down And Roll Up Are The Operations For Moving The View Down And Up Along The Dimensional Hierarchy Levels.
- With Drill-Down Capability, Users Can Navigate To Higher Levels of Detail.
- With Roll-Up Capability, Users Can Zoom Out To See A Summarized Level of Data.
- The Navigation Path is Determined By The Hierarchies Within Dimensions.

Slice And Dice

- Slice And Dice Are The Operations For Browsing The Data Through The Visualized Cube.
- Slicing Cuts Through The Cube So That Users Can Focus on Some Specific Perspectives.
- Dicing Rotates The Cube To Another Perspective So That Users Can Be More Specific With The Data Analysis.

Star And Snowflake Models

- Dimensional Modeling Presents The System of The Data Warehouse To Be Presented in Two Ways
  - Star Model
  - Snowflake Model

Note

- The Constellation Model OR Multistar Model is As An Extension of Star And Snowflake Models.

Star Model

- He Star Schema Structure Looks Like A Star And The Logical Diagram Looks Like The Physical Schema.
- The Star Model is The Basic Structure For A Dimensional Model That is Designed For Data Warehouse.
- Star Model Typically Has One Large Central Table Called As Fact Table) And A Set of Smaller Tables Called Dimension Tables Arranged in A Radial Pattern Around The Fact Table.
- The Fact Table in The Dimensional Model is Joined With All The Other Dimension Tables, And There Will Be Only A Single Join Line Connecting The Fact Table To The Dimension Tables.

Snowflake Model

- Dimensional Modeling Typically Begins By Identifying Facts And Dimensions, After The Business Requirements Have Been Gathered.

- The Initial Dimensional Model is Usually Star Like in Appearance, With One Fact in The Center And One Level of Several Dimensions Around it.
- The Snowflake Model is The Result of Decomposing One OR More of The Dimensions, Which Sometimes Have Hierarchies Themselves.
- We Can Define The Many-To-One Relationships Among Members Within A Dimension Table As A Separate Dimension Table, Forming A Hierarchy.
- The Decomposed Snowflake Structure Visualizes The Hierarchical Structure of Dimensions Very Well.
- The Snowflake Model is Easy For Data Modelers To Understand And For Database Designers To Use For The Analysis of Dimensions.
- The Snowflake Structure Seems More Complex And Could Tend To Make The Business Users Feel More Uncomfortable Working With it Than With The Simpler Star Model.
- Developers Can Elect The Snowflake in The Data Warehouse As it Typically Saves Data Storage.
- Snowflake Model Saves Space, it is Generally Not Significant When Compared To The Fact Table.
- Most Database Designers Do Not Consider The Savings in Space To Be A Major Decision Criterion in The Selection of A Modeling Technique While Developing Data Warehouse.