

# Data Warehouse Design

# Knowledge Objectives

---

1. Distinguish demand and data driven approaches
2. Enumerate project phases
3. Enumerate steps in dimensional modeling
4. Enumerate factual requirements
5. Explain compatibility of facts

# DESIGN METHOD

# Approaches

---

- ❑ Data-driven/Supply-driven
- ❑ Requirement-driven/Demand-driven

# Project phases

---

## I. Technology

## II. Data

### 1. Dimensional modeling

- Logical schema
- Relationships with sources

### 2. Physical design

- Deployment
- Optimization
  - Indexing
  - Partitioning

### 3. Data staging design and implementation

- Extraction
- Transformation
  - Quality improvement
  - Data preparation
- Load

## III. Applications

# Dimensional modeling

---

1. Analyse sources
  - a. Understand available source schema
  - b. Reengineer them to discover unexpressed relationships
  - c. Identify data useful for decisión making
  - d. Assess data quality
  - e. Align different sources
2. Requirement analysis
  - a. Identify facts
    - ▣ Determine granularities
  - b. Determine volumes and workloads
3. Create a star schema for each fact
  - Includes measures, dimensions and hierarchies
4. Validate the queries
5. Translate the schema into relational tables

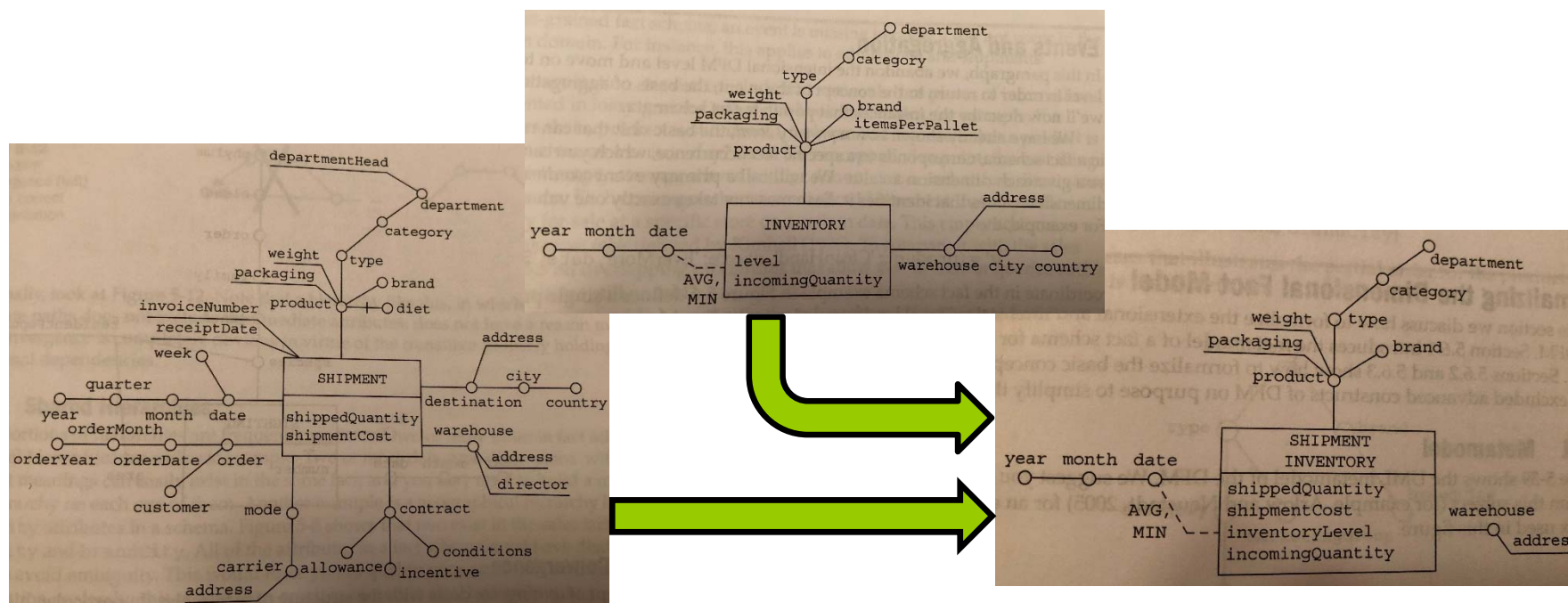
# Factual requirements

---

- Measures
  - Additivity
- Dimensions
  - Hierarchies
- Queries
  - Granularities
- History length

# Overlapping Compatible Fact Schemata

- Conditions:
  - a) Shared hierarchies
  - b) Comparable fact schemata
    - Share a piece of multidimensional space
- Result:
  - a) Unite measures
  - b) Merge hierarchies
  - c) Intersect dimensional instances





# CLOSING

# Summary

---

- ❑ Demand-driven vs Data-driven
- ❑ Project phases
- ❑ Dimensional modeling

# Bibliography

---

- ▣ M. Golfarelli and S. Rizzi. *Data Warehouse Design*. McGraw-Hill, 2009