BIG DATA MODELING
AGENDA

1. Big Data
2. Data Modeling
3. Big Data Modeling
Session Objectives

• Big Data Fundamentals
  – Components of Big Data
  – Structure & Schemas
  – Tools & Architecture

• Data Modeling
  – Integration & History
  – Data Warehousing & BI
  – Conceptual to Physical

• Big Data Modeling
  – Focus on Meaning

• Ensemble Modeling
  – The Blended Architecture
Big Data

Typical Data

- Structured
- Very Complex

Big Data

- "Huge" Data Volumes
- n-Structured & Very Complex
- Streaming & Shape-Shifting
Big Data

• Volume
  Huge Volumes of Data

• Velocity
  Drinking from a Fire Hose

• Variety
  \( n \)-Structured Data

• Veracity
  Quality, Accuracy, Reliability, Trustworthiness

• Value
  Business Value and Value Potential
Modeling and Understanding

- Schema on Write
- Schema on Read
- Dismantled Schema on Write
- Schema on Focus
- Schema on Leverage

MODEL  EXPLORE  APPLY
LOAD
Modeling and Understanding

- Big Data Possibilities

LOAD EXPLORE MODEL APPLY
"If we don't understand the data, it can only be less useful"
- Silverston (paraphrased)

DATA MODELING
Data Modeling

Mans Search for Meaning...

- Conceptual Modeling
- Logical Modeling
- Information Modeling
- Physical Data Modeling
# Three Paths for Modeling

<table>
<thead>
<tr>
<th>Structured / Known</th>
<th>N-Structured / NVP</th>
<th>N-Structured / KVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBC</td>
<td>CBC</td>
<td>CBC</td>
</tr>
<tr>
<td>NBR</td>
<td>NBR</td>
<td>NBR</td>
</tr>
<tr>
<td>Attribution</td>
<td>Attribution</td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CBC: Core Business Concepts  
NBR: Natural Business Relationships
APPLYING THE ENSEMBLE

Integration across Platforms

Entity \rightarrow Ensemble \rightarrow Dim
Logical business model

• Leveraged for all **logical model** needs including the data warehouse, big data lake, master data management (MDM) and operational integration initiatives

• Closely aligned to DV physical model
Summary

Ensemble in the Big Data World

• Conceptual Modeling
• Logical Modeling
• Information Modeling
• Physical Data Modeling
• Integration Platform
Estimated 1400 + Data Vault based Data Warehouses around the world
Links and Information

CDVDM Training & Certification
www.GeneseeAcademy.com
Hans@GeneseeAcademy.com
go hans go
Book  DataVaultBook.blogspot.com
HansHultgren.WordPress.com
LinkedIn  HansHultgren
YouTube  DataVaultAcademy

Online video-lesson training
DataVaultAcademy.com