

# APACHE HIVE

Part 5

by Sumit Mittal



# IMPORTANT

## **Copyright Infringement and Illegal Content Sharing Notice**

All course content designs, video, audio, text, graphics, logos, images are Copyright© and are protected by India and international copyright laws. All rights reserved.

Permission to download the contents (wherever applicable) for the sole purpose of individual reading and preparing yourself to crack the interview only. Any other use of study materials – including reproduction, modification, distribution, republishing, transmission, display – without the prior written permission of Author is strictly prohibited.

**Trendytech Insights** legal team, along with thousands of our students, actively searches the Internet for copyright infringements. Violators subject to prosecution.



# Trainer Introduction



**Mr. Sumit Mittal**, CEO & founder of **TrendyTech**. He has a Master's degree in Computer Applications from NIT Trichy & have a total of 7+ years of industry experience. He has worked for top MNC's like **Cisco** & **VMware**.



Consistent 5 star **Google** rated  
Bigdata course



---

# Normalized Storage in Traditional Databases

---

# Traditional Database Design



## Normalized data

Data is stored in a granular form to minimize redundancy



# Employee Information

**name**

**address**

**id**

**subordinates**

**department**

**grade**





id grade name department

id subordinates

id address





## Employee Details

Id	Name	Department	Grade
1	Emily	Finance	6

## Employee Details

Id	Subordinate Id
1	2
1	3

## Employee Address

Id	City	Zip Code
1	Palo Alto	94305
2	Seattle	98101





Id	Name	Department	Grade
1	Emily	Finance	6

Id	Subordinate Id
1	2
1	3

Id	City	Zip Code
1	Palo Alto	94305
2	Seattle	98101

## Normalization





Id	Name	Department	Grade
1	Emily	Finance	6

join

Id	Subordinate Id
1	2
1	3

Query for Emily's department  
and her subordinates



# Normalization



Minimizes redundancy, optimizes storage

Foreign keys to ensure valid joins

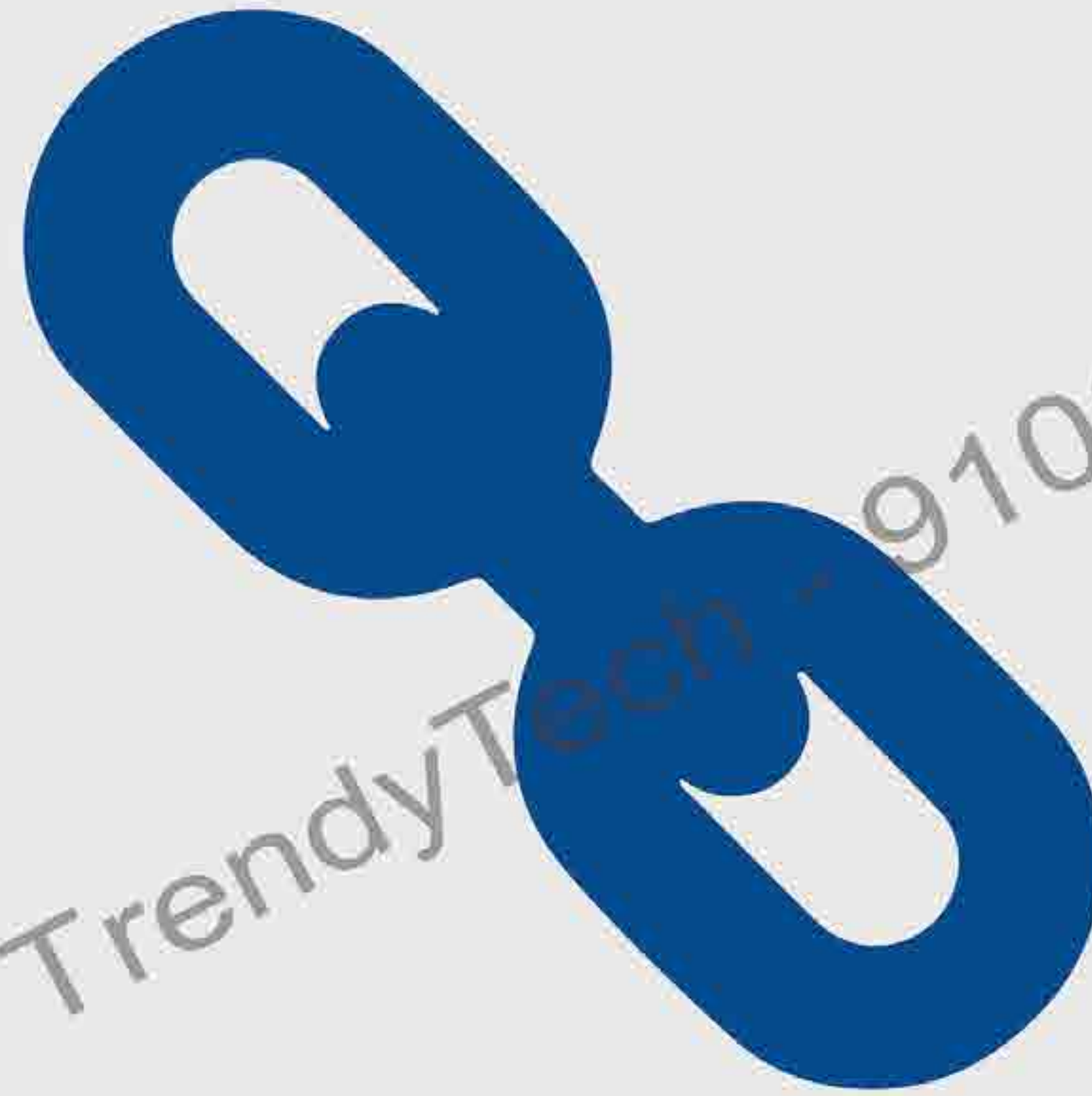
Updates in one location, no duplication of data



# Denormalized Storage in Hive



# Denormalized Storage



## Denormalized data

Data is compressed into one table to be read in a single operation



# Denormalized Storage



Disk space is very cheap

No foreign key constraints

Read operations, no data updates



# Denormalized Storage

Optimize the number of disk seeks

Store data for an entity in one location

Ignore redundancy, minimize joins





# Denormalized Storage

Id	Name	Function	Grade
1	Emily	Finance	6
2	John	Finance	3
3	Ben	Finance	4

Id	Subordinate Id
1	2
1	3



Id	Name	Function	Grade	Subordinates
1	Emily	Finance	6	2, 3
2	John	Finance	3	
3	Ben	Finance	4	



# Denormalized Storage

Id	Name	Function	Grade
1	Emily	Finance	6
2	John	Finance	3
3	Ben	Finance	4

Id	City	Zip Code
1	Palo Alto	94305
2	Seattle	94305



Id	Name	Function	Grade	Subordinates	Address
1	Emily	Finance	6	2, 3	Palo Alto, 94305
2	John	Finance	3		
3	Ben	Finance	4		



# Denormalized Storage

Id	Name	Function	Grade	Subordinates	Address
1	Emily	Finance	6	2, 3	Palo Alto, 94305
2	John	Finance	3		
3	Ben	Finance	4		

Store everything related to an employee in the same table



# Denormalized Storage

Id	Name	Function	Grade	Subordinates	Address
1	Emily	Finance	6	2, 3	Palo Alto, 94305
2	John	Finance	3		
3	Ben	Finance	4		

Get all details about an  
employee in one read operation





**5** Star Google Rated  
Big Data Course  
**LEARN FROM THE EXPERT**



**9 1 0 8 1 7 9 5 7 8**

Call for more details





**Follow US**

<b>Trainer</b>	Mr. Sumit Mittal
<b>Phone</b>	9108179578
<b>Email</b>	trendytech.sumit@gmail.com
<b>Website</b>	<a href="https://trendytech.in/courses/big-data-online-training/">https://trendytech.in/courses/big-data-online-training/</a>
<b>Linked In</b>	<a href="https://www.linkedin.com/in/bigdatabysumit/">https://www.linkedin.com/in/bigdatabysumit/</a>
<b>Twitter</b>	@BigdataBySumit
<b>Instagram</b>	bigdatabysumit
<b>Facebook</b>	<a href="https://www.facebook.com/trendytech.in/">https://www.facebook.com/trendytech.in/</a>
<b>Youtube</b>	TrendyTech