



SNS COLLEGE OF TECHNOLOGY



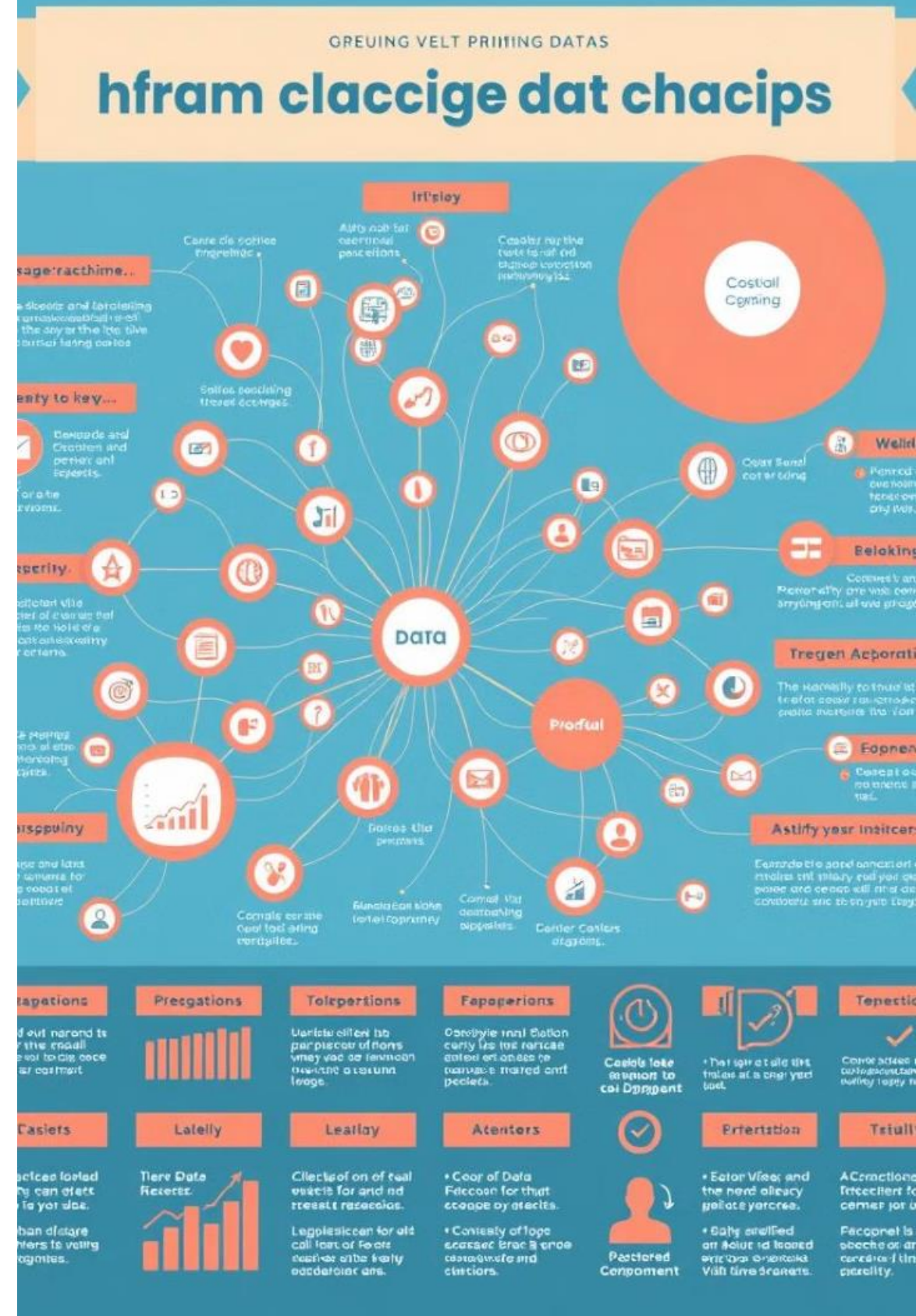
COIMBATORE-35

23BAE733-DATA- DRIVEN DECISION MAKING

Dr.P.Krishnaveni  
Design Thinker

# Data Transformations: Essential for Data Management

Data transformations are crucial for organizing and manipulating data, empowering effective analysis and reporting. Understanding these techniques unlocks powerful capabilities for data professionals.



# Union Operation: Merging Data Sets

## Combining Sets

Combines two data sets.

Includes all entries.

## Example: Customer Lists

Merging customer lists from two branches.

Creates a unified list of all customers.

Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147

UNION

Item	Date	Time	Price	Quantity
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285



Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285

set A

set B

Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324

INTERSECT

Item	Date	Time	Price	Quantity
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285



Item	Date	Time	Price	Quantity
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324



# Intersection Operation: Finding Common Entries

Shared Elements

Identifies entries in both sets.

Commonality Focus

Results in a set of shared entries.

# Difference Operation: Removing Overlap

1

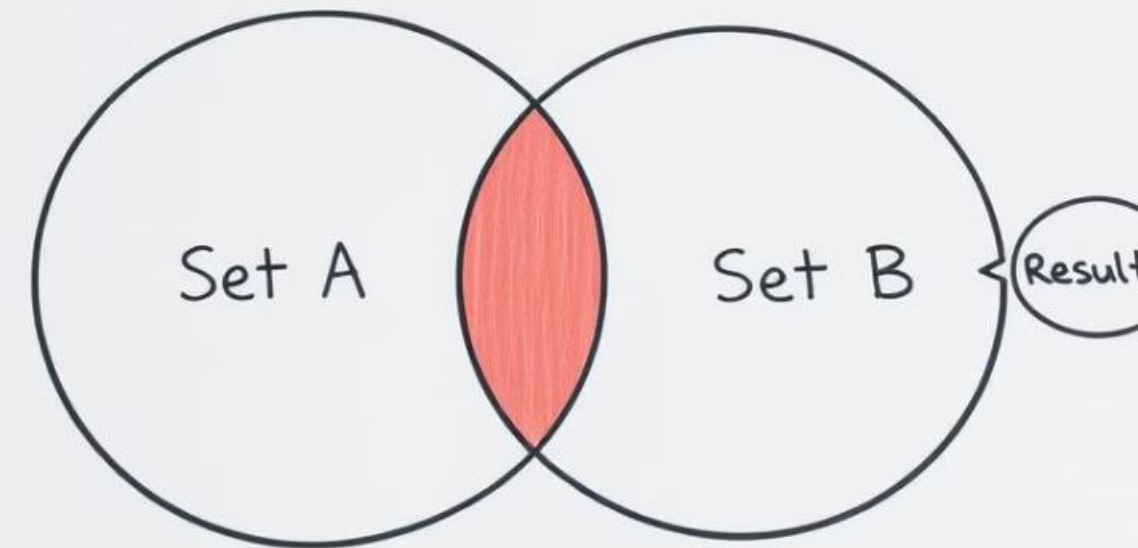
## Removal

Entries from the first set.  
Present in the second set.

2

## Unique Entries

Remaining entries.





Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285



**SELECT** Item, Date, Time, Price, Quantity **WHERE** Quantity < 200



Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147

# Selection: Filtering Data Based on Conditions

## Unmodified Data

Replicates the original data set.

## Specific Entries

Applies constraints for selection.

# Projection: Modifying and Limiting Attributes

1

## Same Entries

Maintain entry count.

Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285

SELECT Item, Price \* Quantity

Item	
233894	\$388.00
143938	\$1,029.00
134239	\$2,122.20
371864	\$1,567.5



Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285

SELECT Item, Price \* Quantity AS Revenue

Item	Revenue
233894	\$388.00
143938	\$1,029.00
134239	\$2,122.20
371864	\$1,567.5

2

## Attribute Changes

Exclude or calculate attributes.

3

## Example: Revenue

Price and quantity multiplied.

# Combining Selection and Projection: Precise Data Manipulation

Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285



`SELECT Item AS Product, Price * Quantity AS Revenue WHERE Quantity < 200`



<u>Product</u>	<u>Revenue</u>
233894	\$388.00
143938	\$1,029.00



1

Refinement

Refined data analysis.

2

Relevant Data

Only essential attributes.

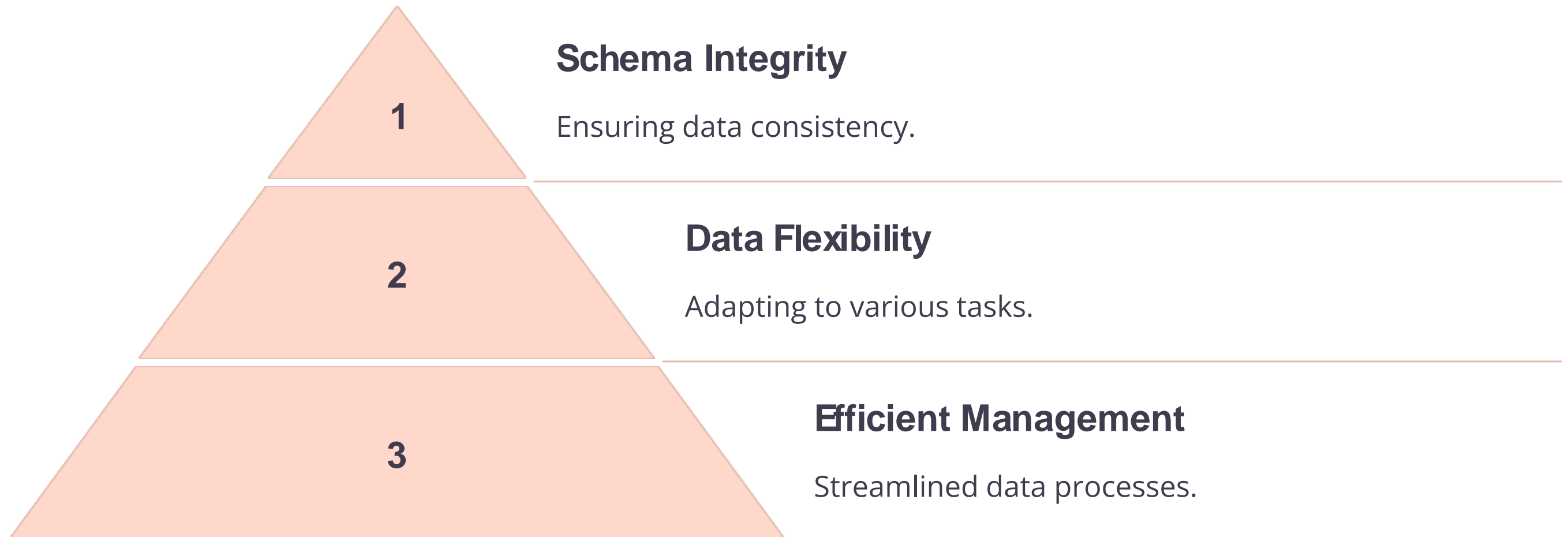
3

Calculated Values

Combined analysis results.



# Mastering Data Transformations: Benefits and Applications



# Knowledge Check:

The following examples shows Combining Selection and Projection and Precise Data Manipulation.

True/False

Item	Date	Time	Price	Quantity
233894	Sep. 7	12:00	\$4.00	97
143938	Sep. 12	14:00	\$7.00	147
134239	Sep. 19	09:00	\$6.55	324
371864	Oct. 1	17:00	\$5.50	285



```
SELECT Item AS Product, Price * Quantity AS Revenue WHERE Quantity < 200
```



<u>Product</u>	<u>Revenue</u>
233894	\$388.00
143938	\$1,029.00

