Data as Sparse Matrices

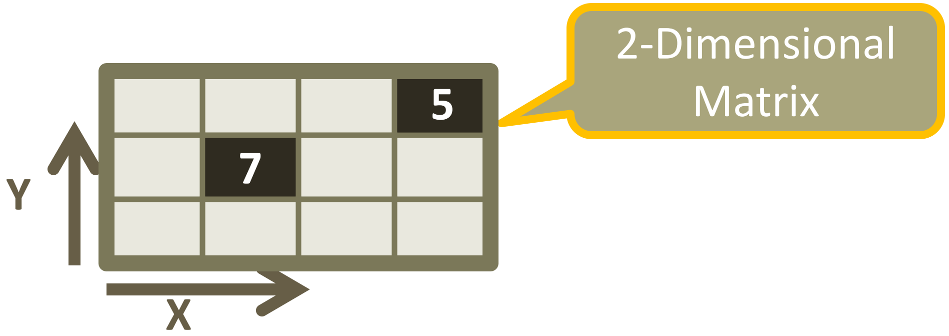
Cartesian Product

- Cartesian product
- http://en.wikipedia.org/wiki/Cartesian_product
- The Cartesian product of two sets A and B is the set of all ordered pairs ab, where a is element of A and b is element of B.

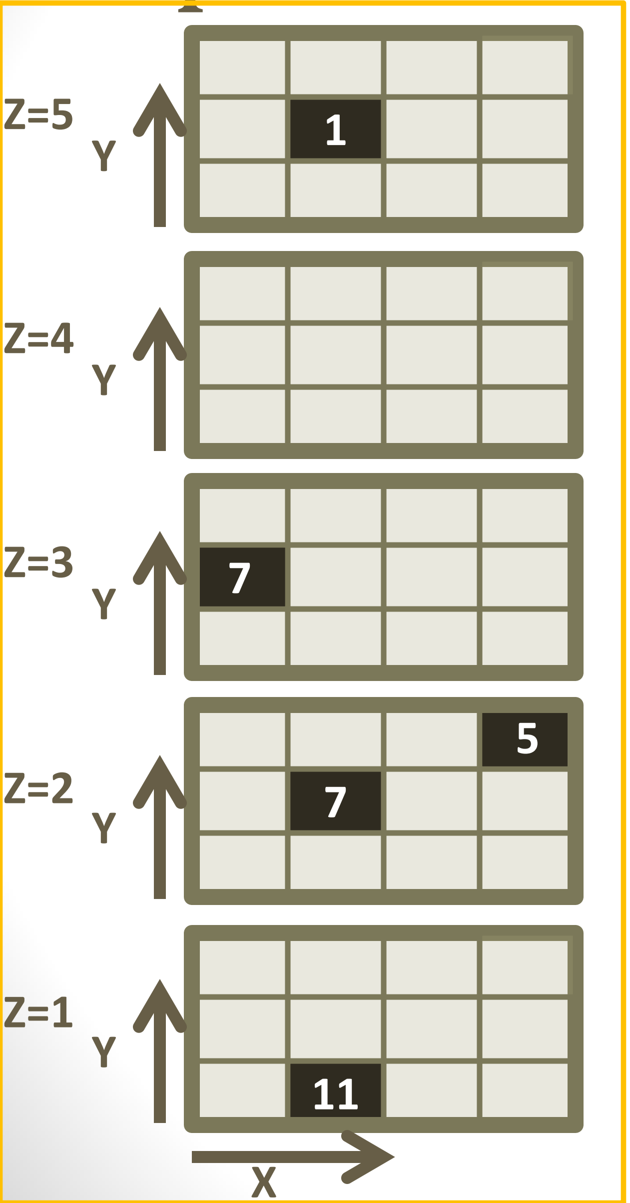
- Relational Algebra
- http://en.wikipedia.org/wiki/Relational_algebra
- In Relational Algebra we need the Cartesian product to combine tuples into a single tuple. The Cartesian product creates a new schema (relation) from other relations.

- Hyperrectangle (Sparse Multi-Dimensional Matrix)
- http://en.wikipedia.org/wiki/Hyperrectangle
- Hyperrectangle is the generalization of a rectangle for higher dimensions and is defined as the Cartesian product of intervals

Sparse Matrices

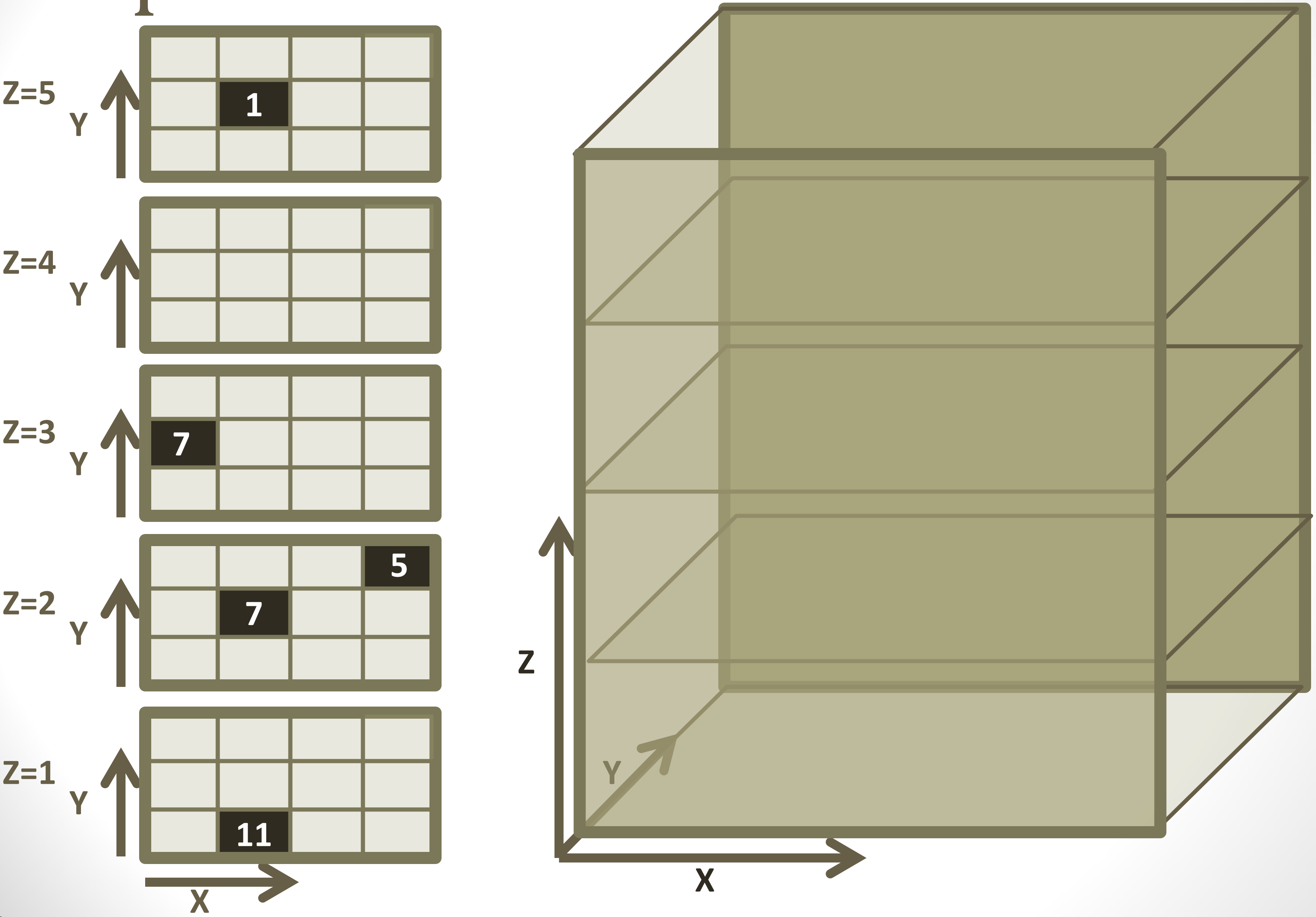


Sparse Matrices

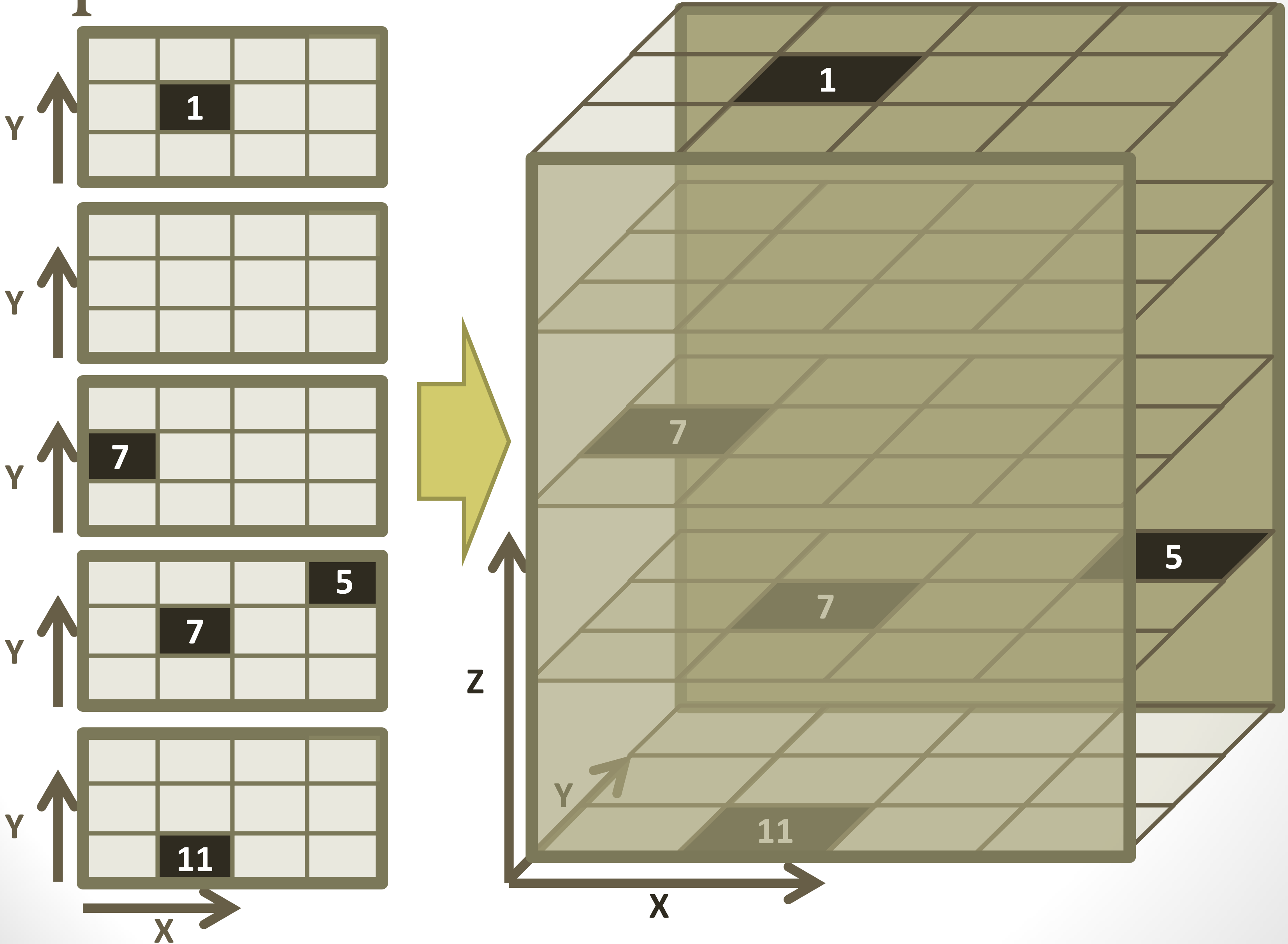


A series of equal-sized 2-dimensional matrices is a 3-dimensional matrix

Sparse Matrices



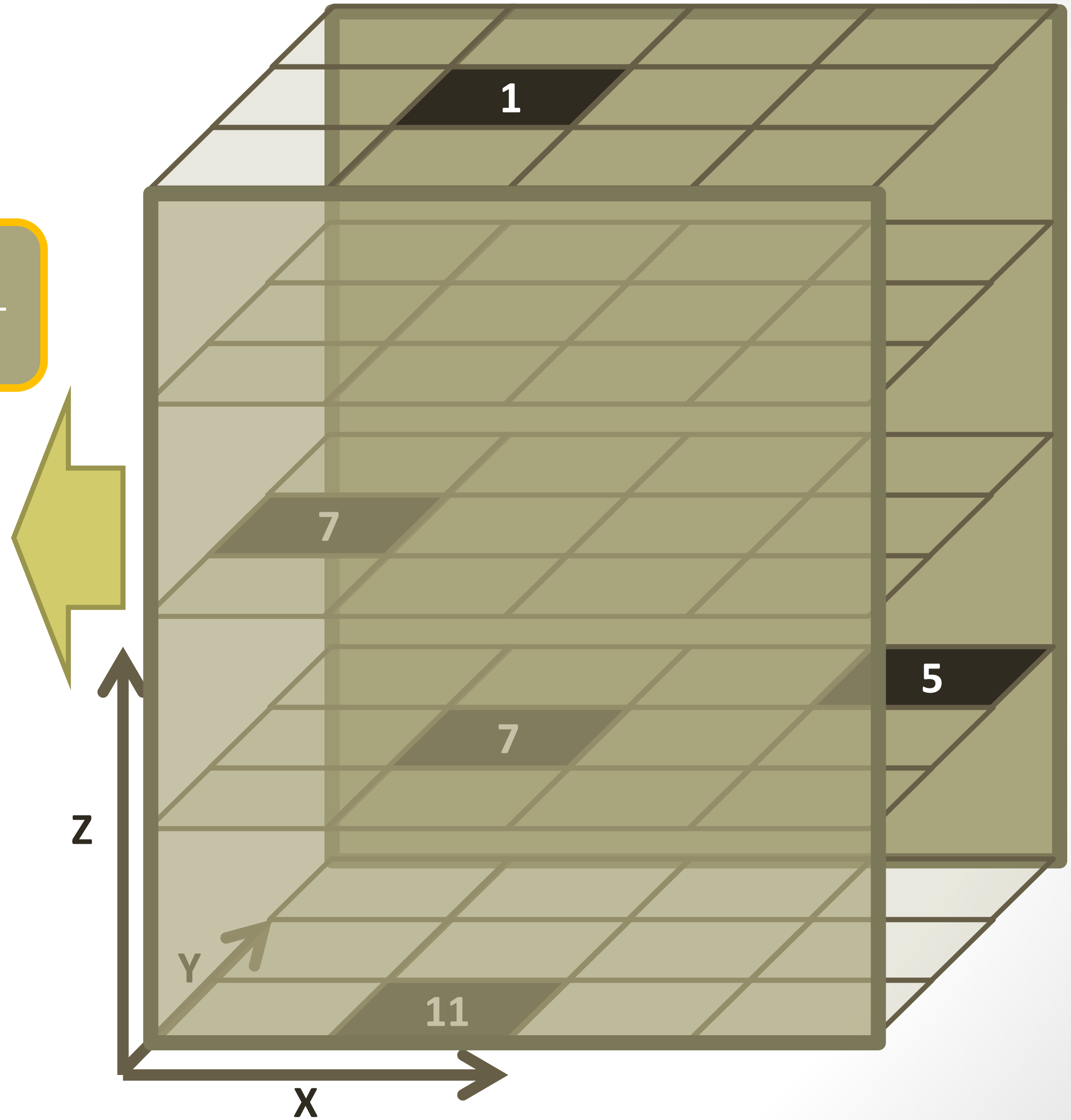
Sparse Matrices



Sparse Matrices

A table with n columns represents values in an n-1 dimensional matrix

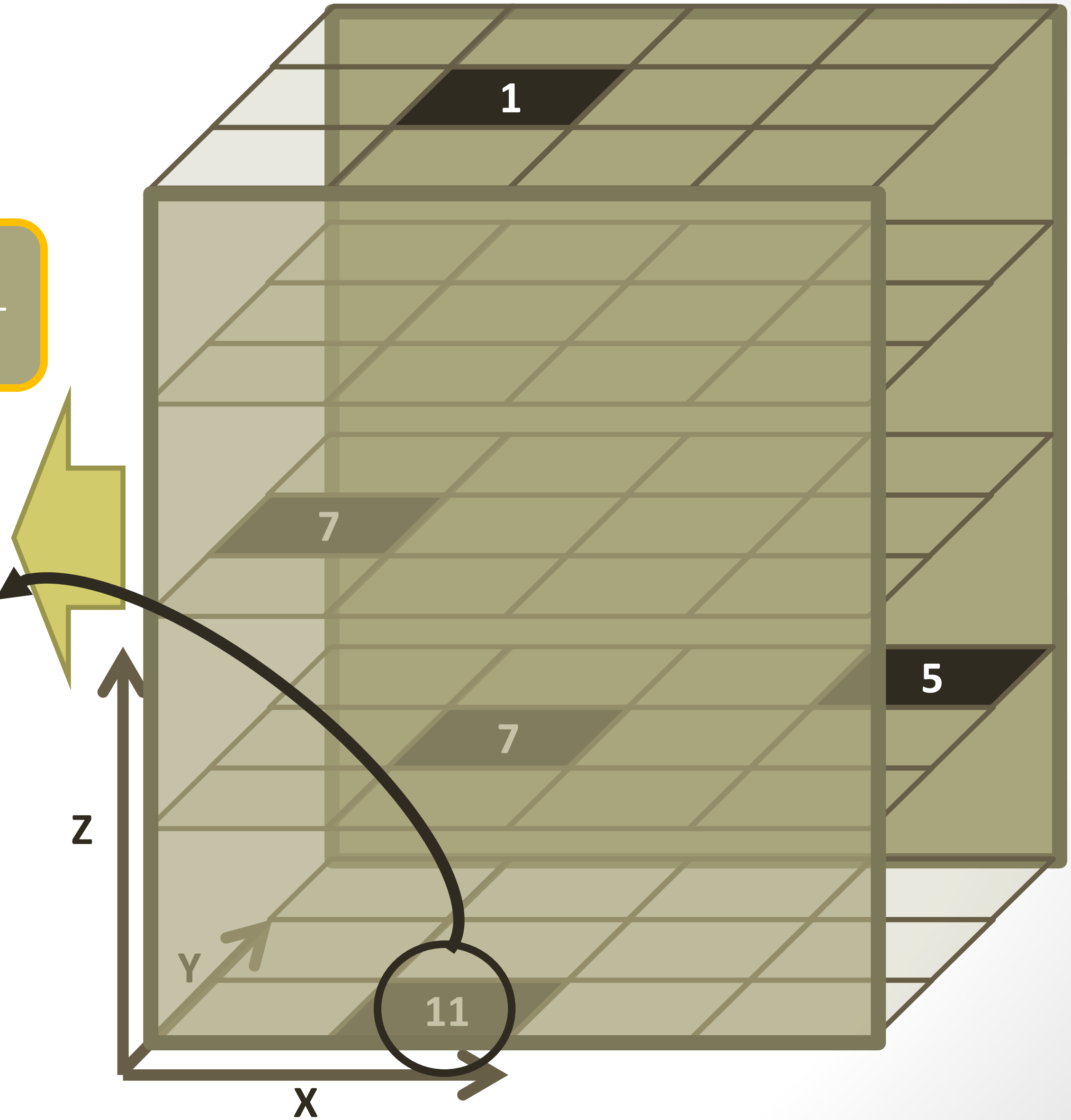
X	Y	Z	V



Sparse Matrices

A table with n columns represents values in an n-1 dimensional matrix

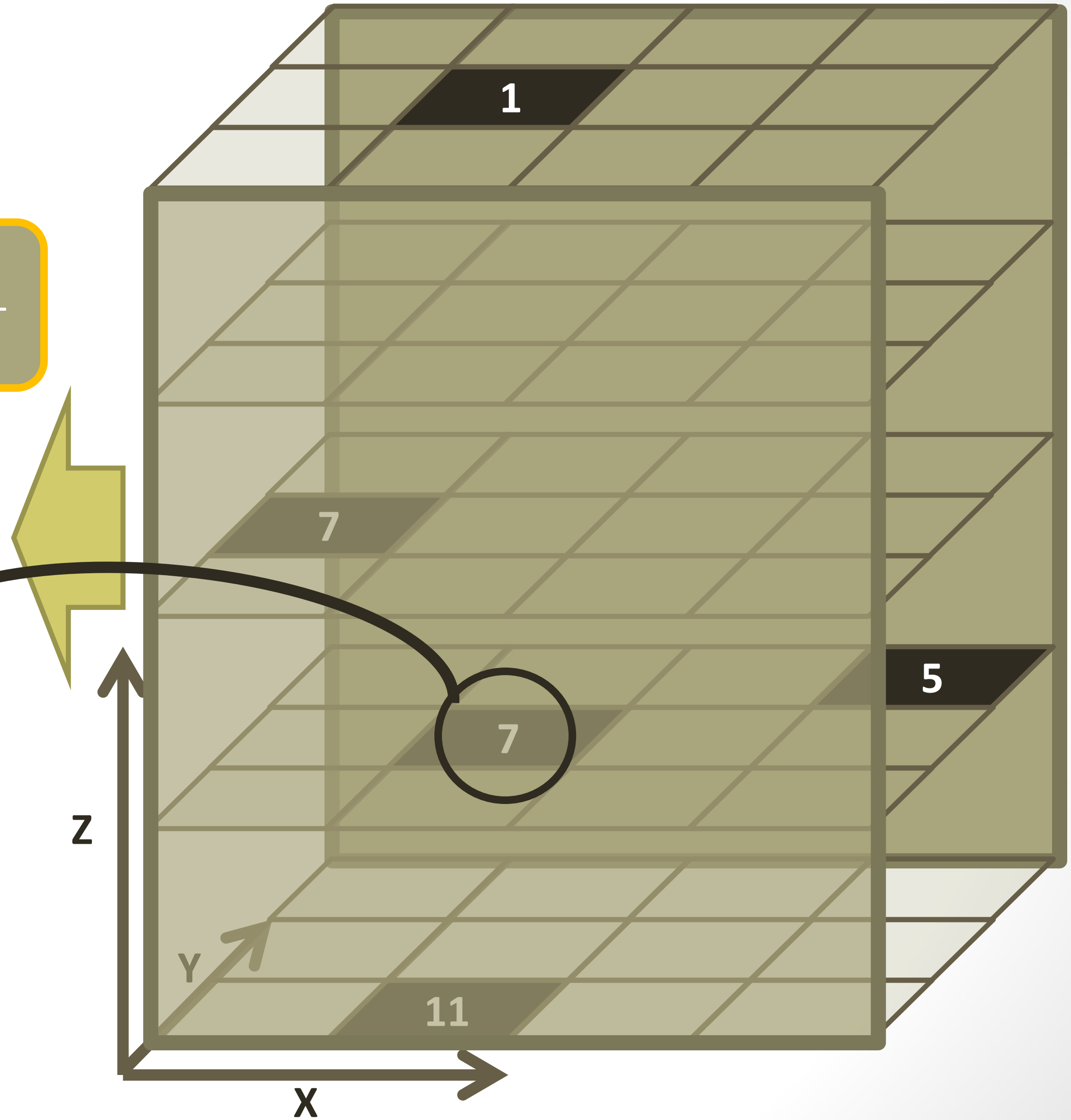
X	Y	Z	V
2	1	1	11



Sparse Matrices

A table with n columns represents values in an n-1 dimensional matrix

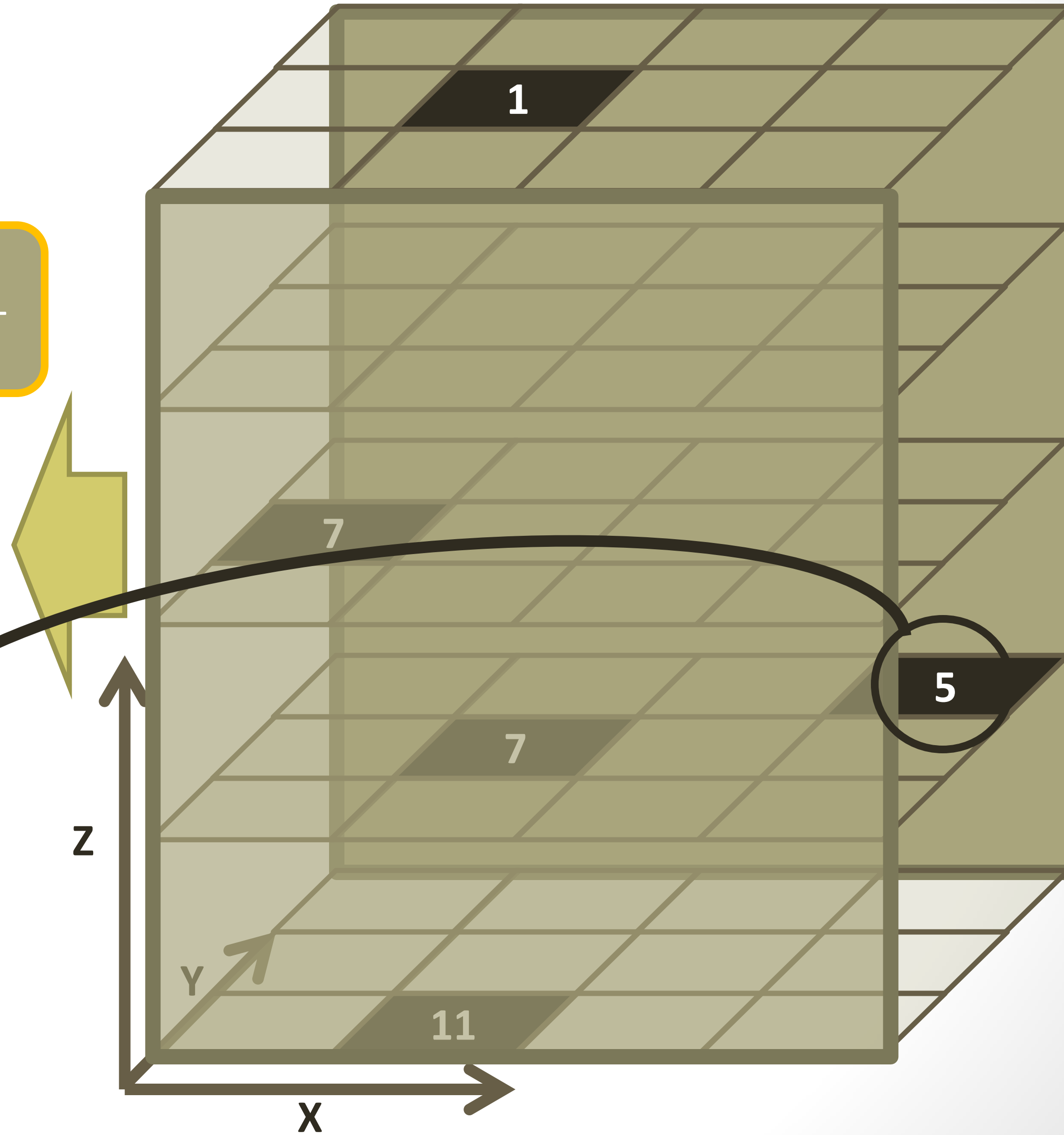
X	Y	Z	V
2	1	1	11
2	2	2	7



Sparse Matrices

A table with n columns represents values in an n-1 dimensional matrix

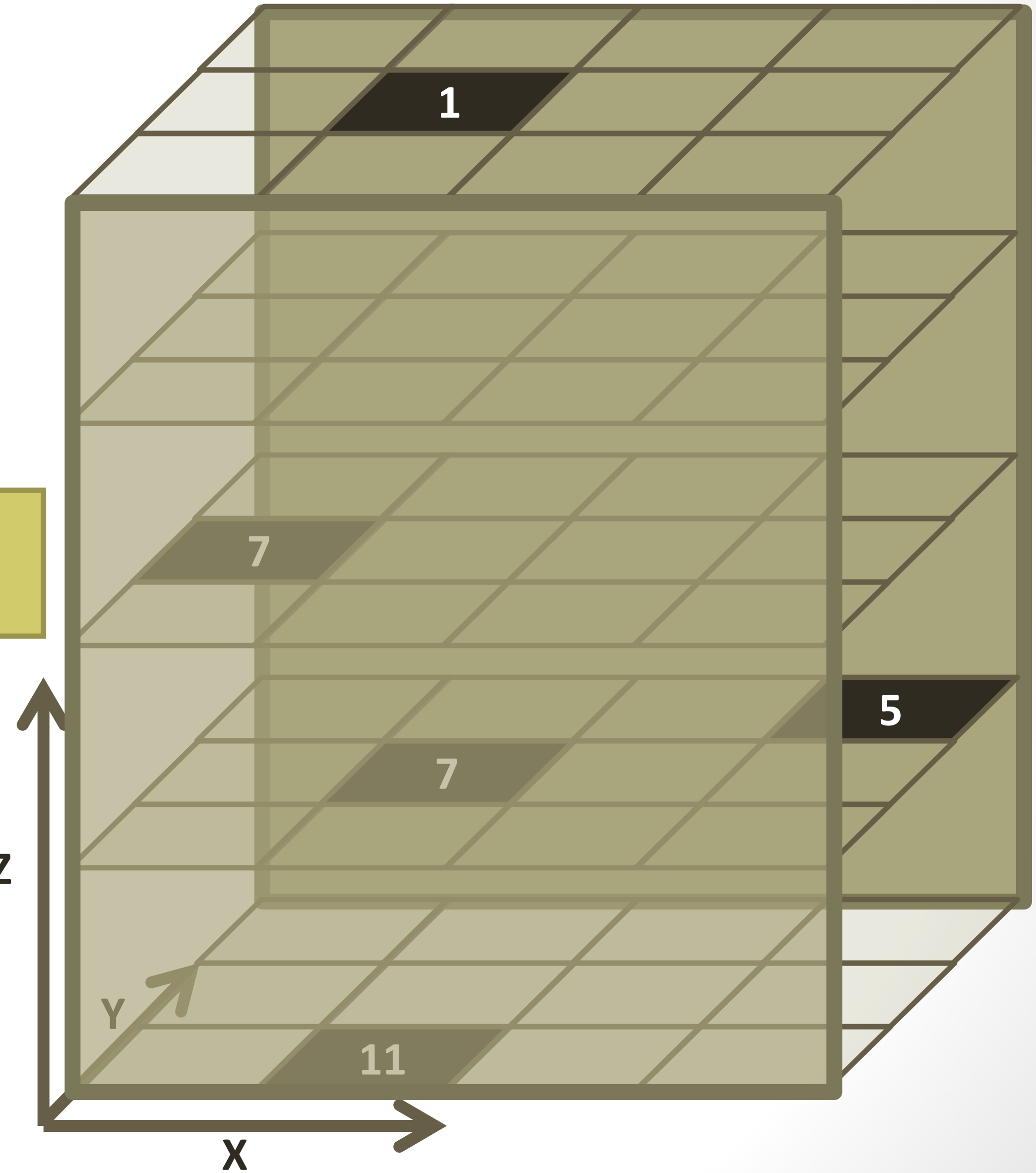
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5



Sparse Matrices

A table with n columns represents values in an n-1 dimensional matrix

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

Sparse Matrices

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

I can think of V as just another dimension

Sparse Matrices

A table with n columns represents points in an n-dimensional matrix

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

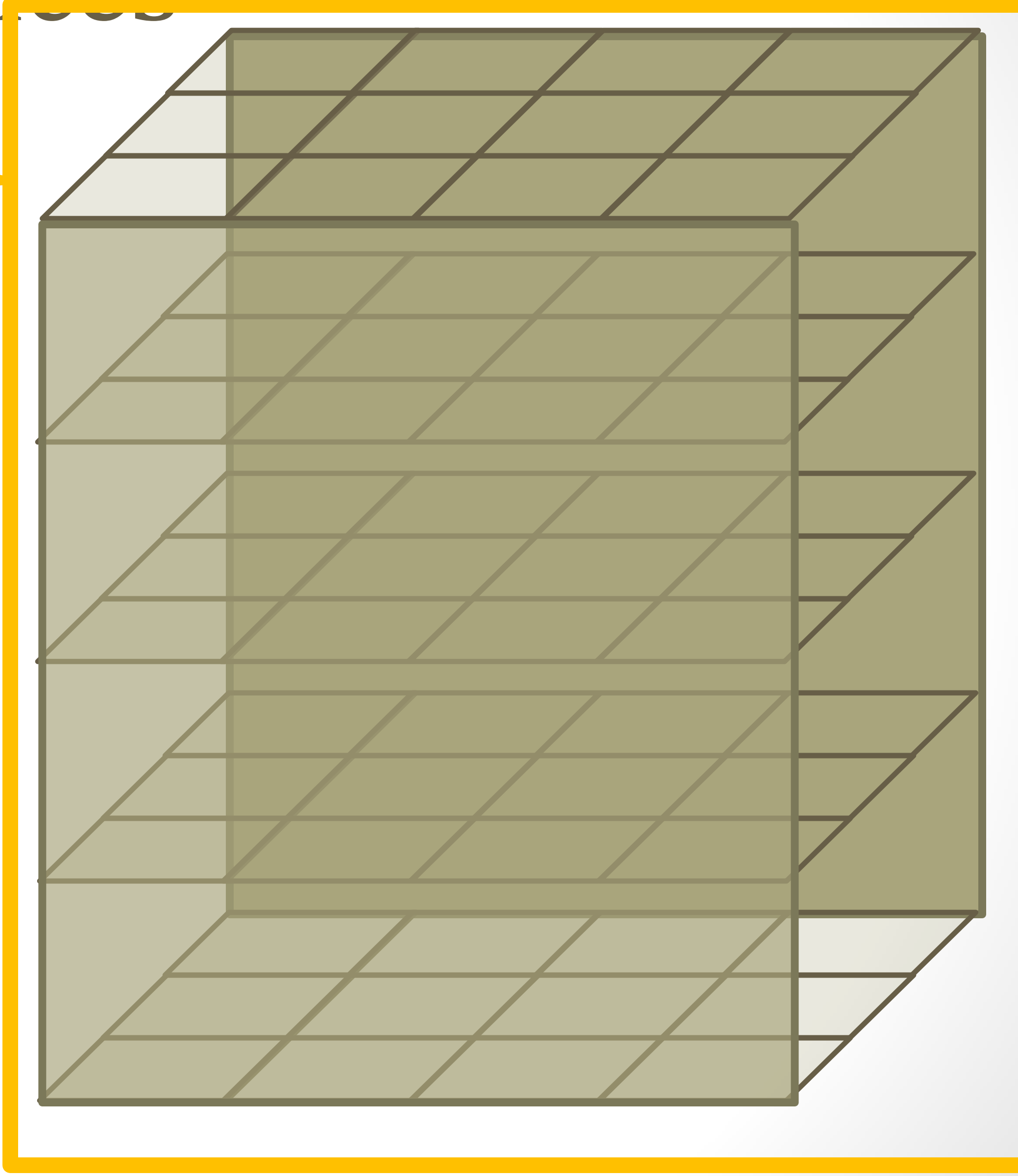
I can think of V as just another dimension

Sparse Matrices

3-Dimensional Space.

This table represents points in 4-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

4-Dimensional Space.

This table represents points in 4-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

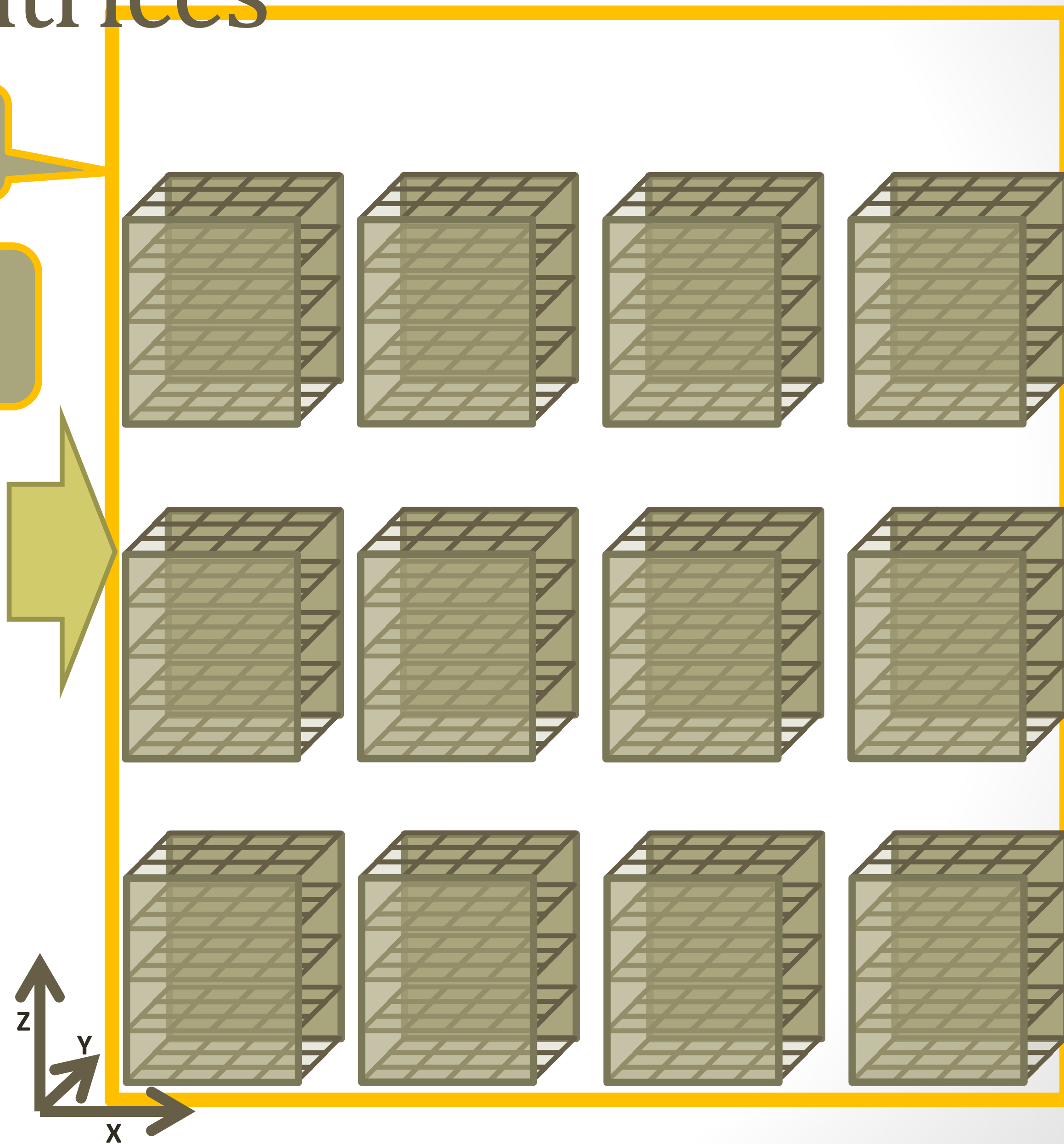
?

Sparse Matrices

4-Dimensional Space with 12 discrete states.

This table represents points in 4-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

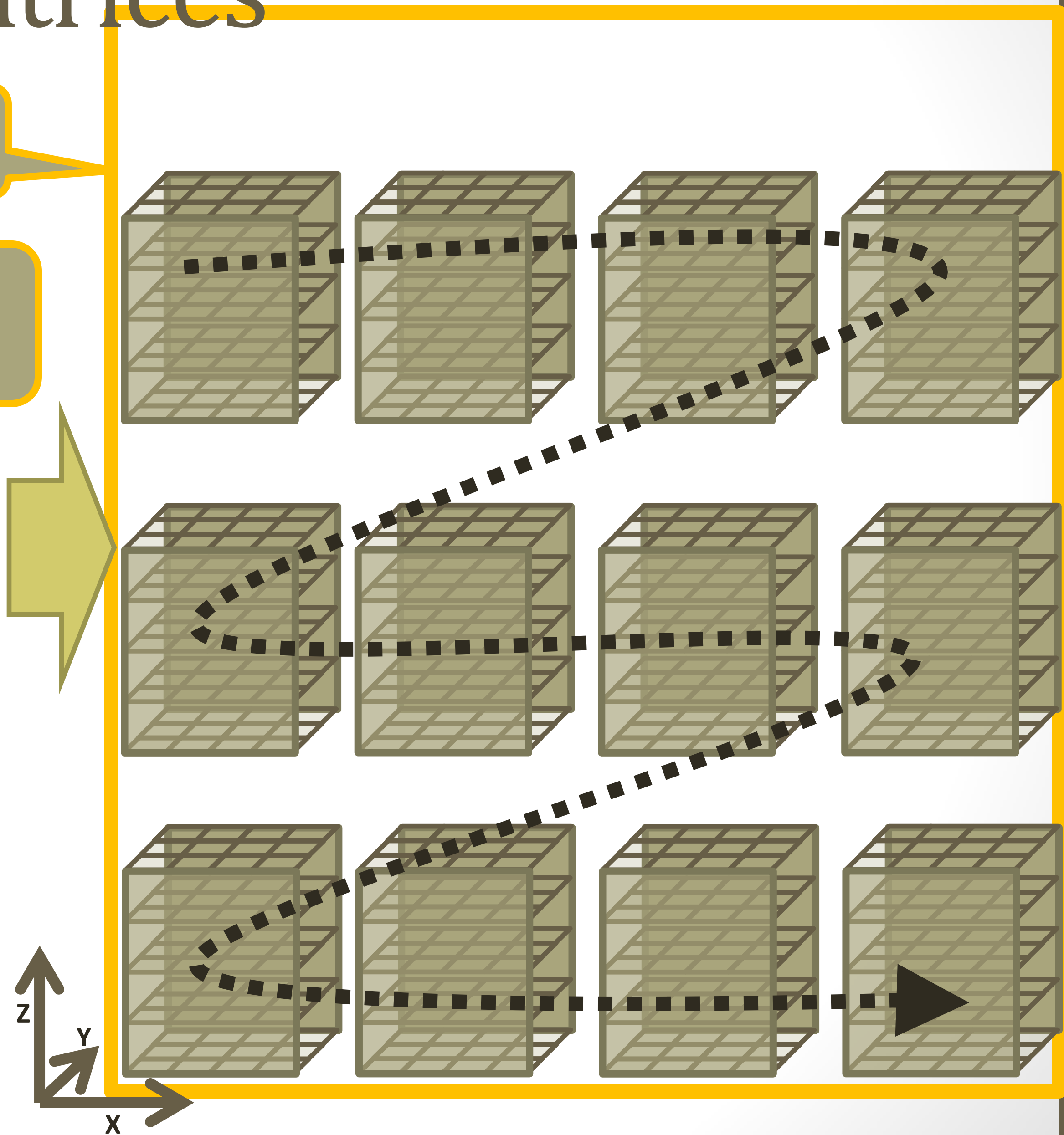


Sparse Matrices

4-Dimensional Space with 12 discrete states.

This table represents points in 4-Dimensional Space.

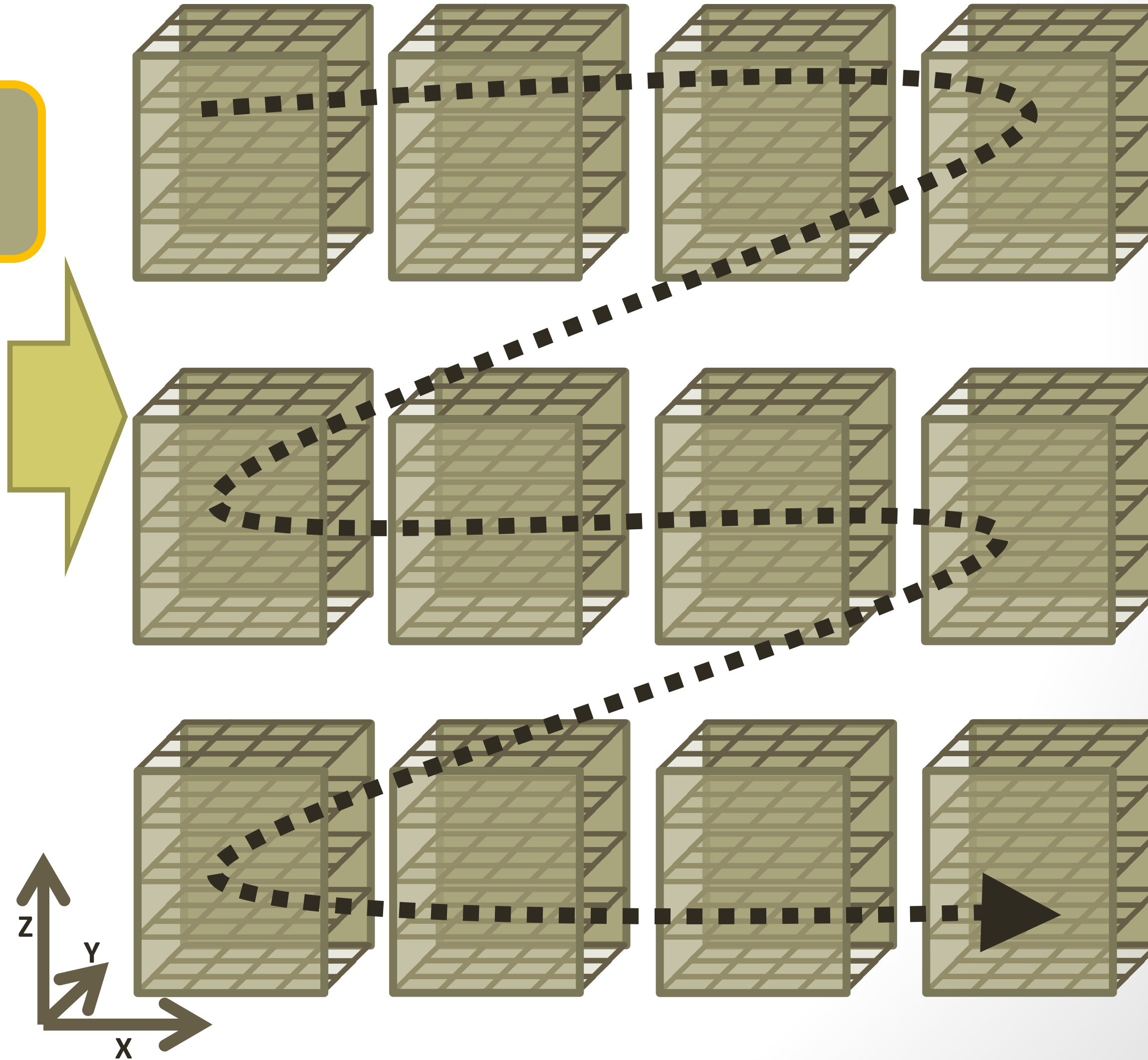
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

This table represents points in 4-Dimensional Space.

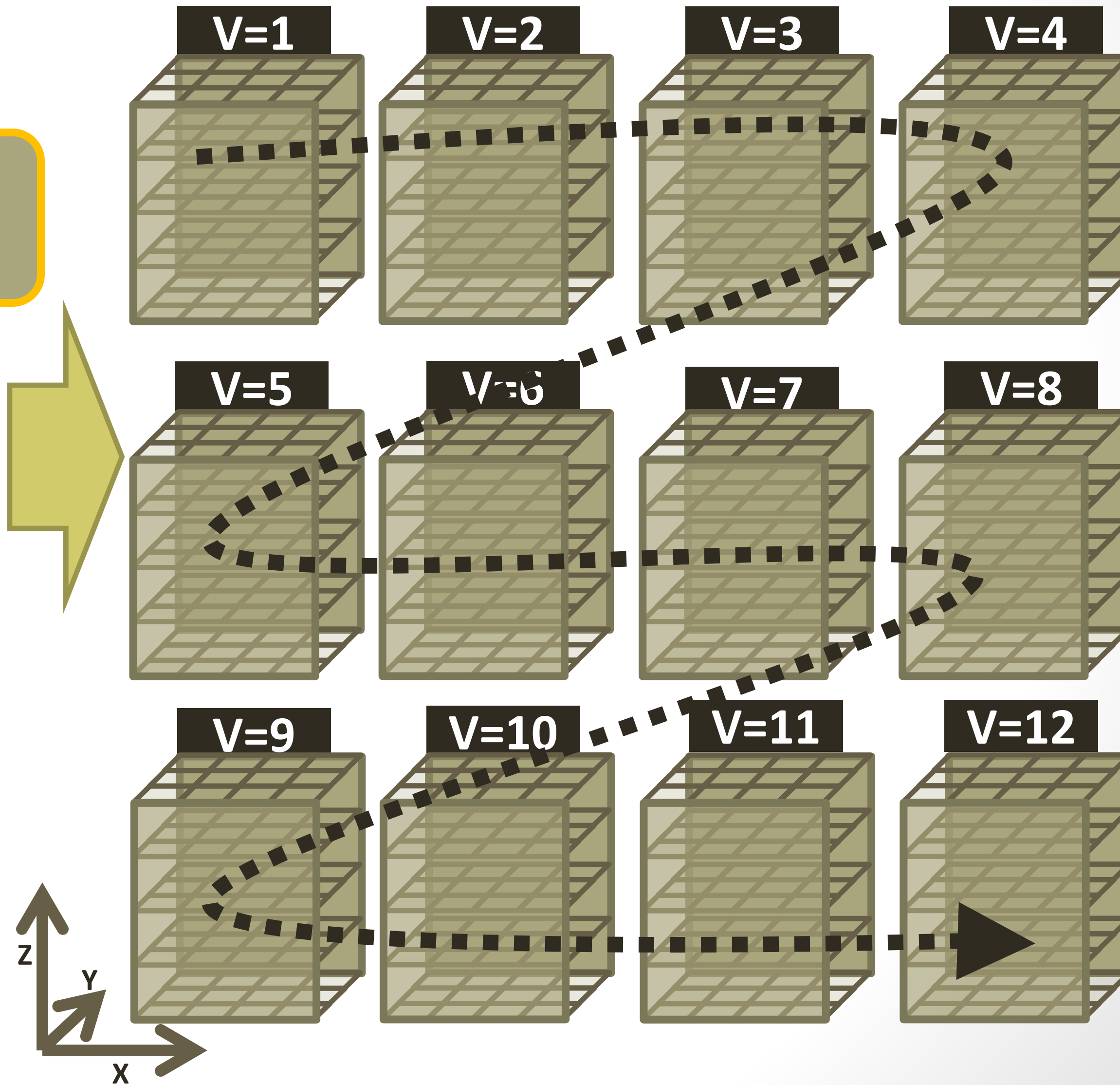
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

This table represents points in 4-Dimensional Space.

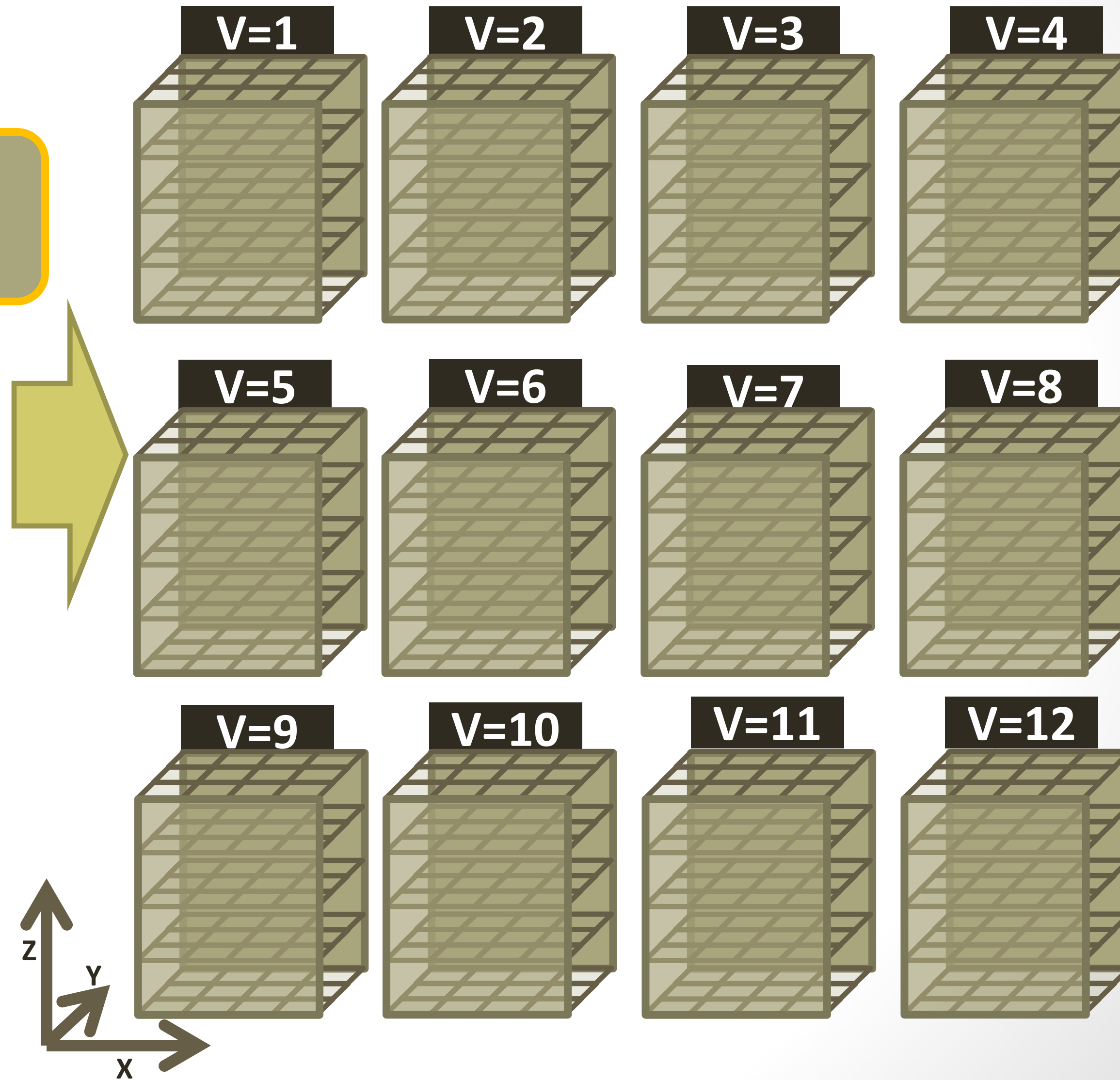
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

This table represents points in 4-Dimensional Space.

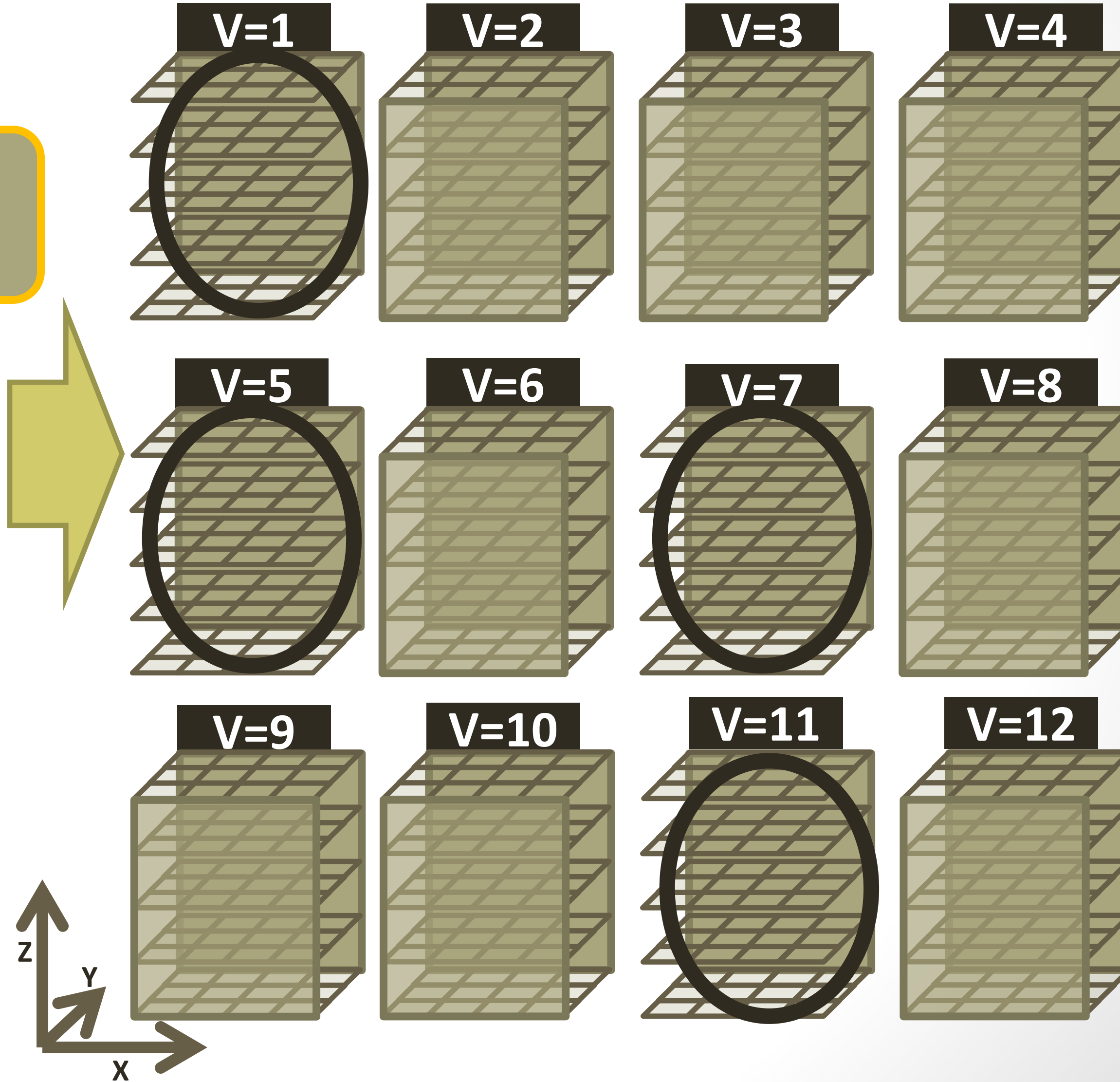
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

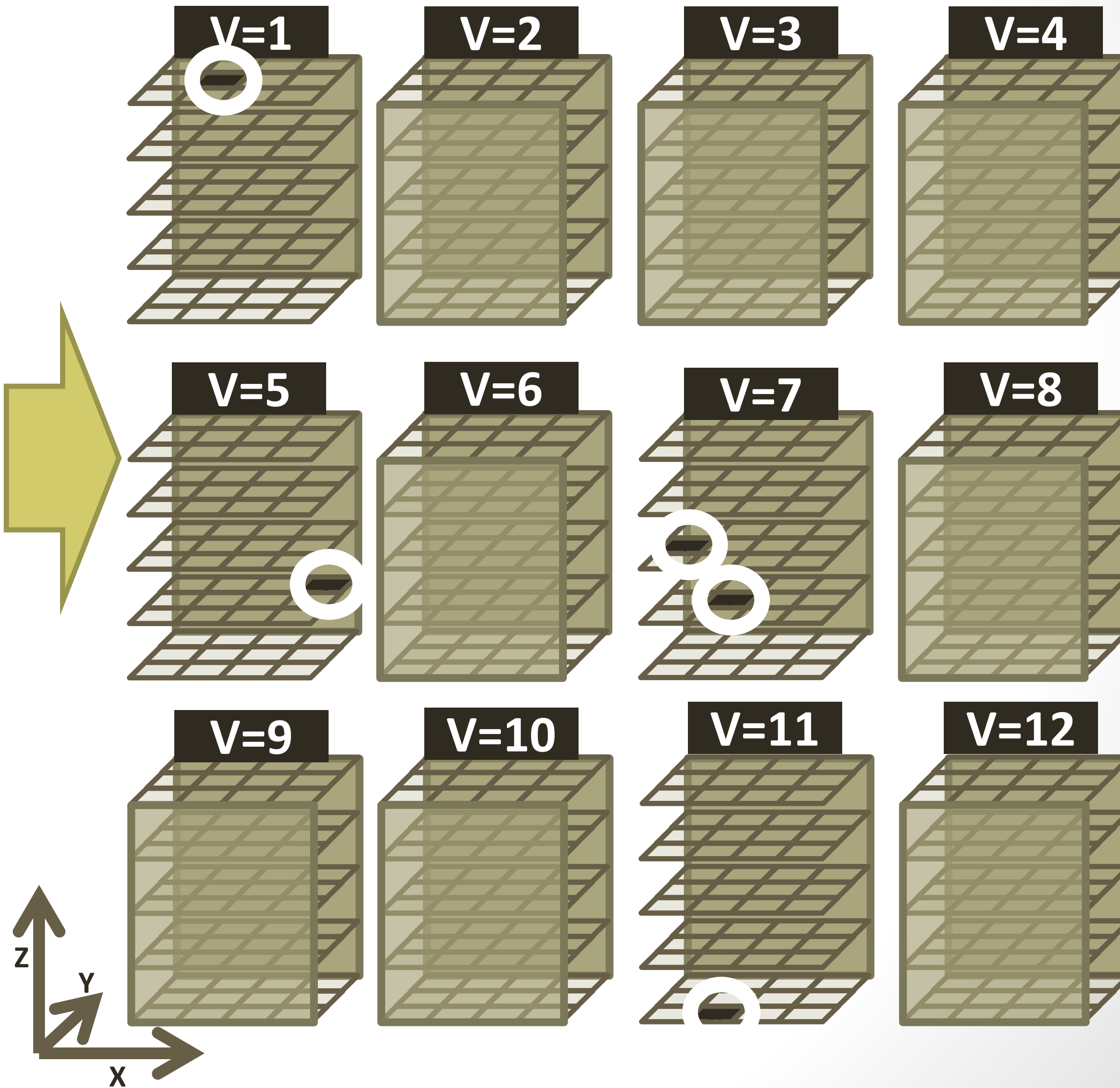
This table represents points in 4-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices: EAV

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

Sparse Matrices: EAV

A table represents points in n-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

Sparse Matrices: EAV

A table represents points in n-Dimensional Space.

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

- Can we represent all tables in a single schema?
- Any table or matrix cell can be described by row, column and value.
- Represent each cell of a table in its own row.
- Entity–attribute–value model

Sparse Matrices: EAV

Column Name

Row ID. Needs to be unique for a given row in the original table. Does not need to be a number or sequential (ordinal) number

Cell Values

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

R	C	M

- Can we represent all tables in a single schema?
- Any table or matrix cell can be described by row, column and value.
- Represent each cell of a table in its own row.
- **[Entity–attribute–value model]**

Sparse Matrices: EAV

Row ID. Needs to be unique for a given row in the original table. Does not need to be a number or sequential (ordinal) number

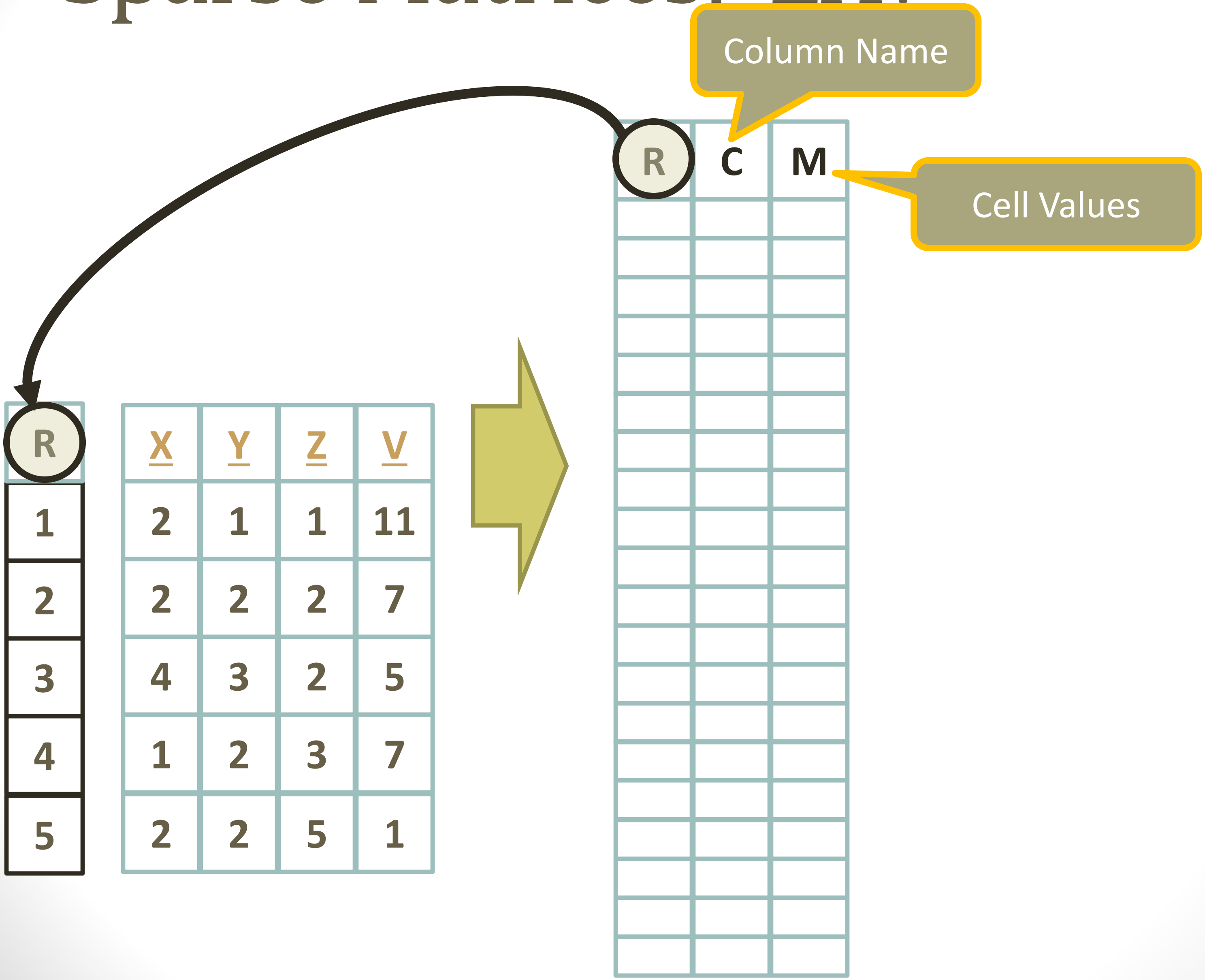
Column Name

Cell Values

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

R	C	M

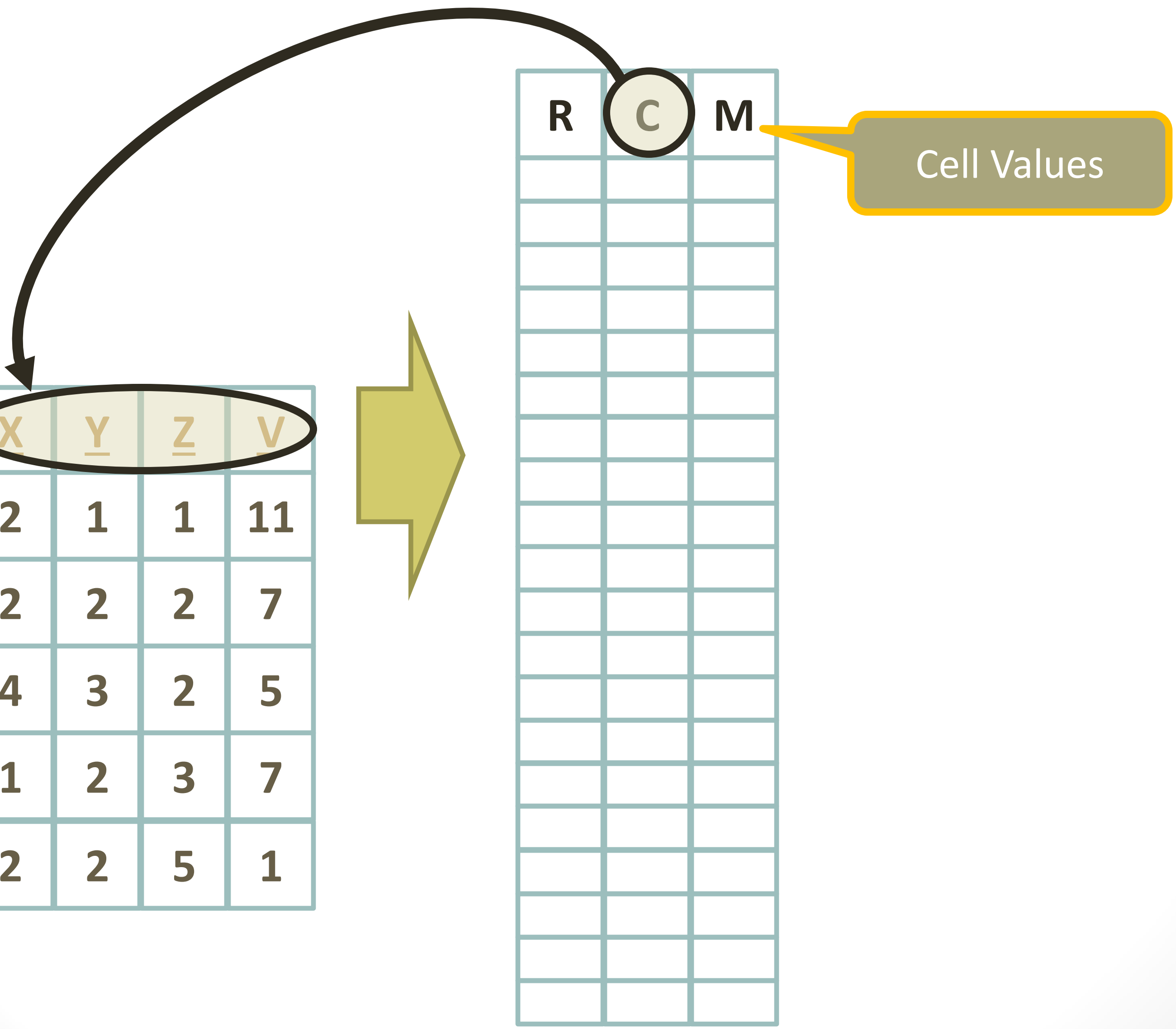
Sparse Matrices: EAV



R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

Sparse Matrices: EAV

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1



Sparse Matrices

R
1
2
3
4
5

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

R	C	M

Sparse Matrices

R
1
2
3
4
5

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

R	C	M
1		

Sparse Matrices

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

R	C	M
1	X	

Sparse Matrices

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

R	C	M
1	X	2

Sparse Matrices

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

R	C	M
1	X	2
1		

Sparse Matrices

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

R	C	M
1	X	2
1	Y	

Sparse Matrices

R	X	Y	Z	V
1	2	1	1	11
2	2	2	2	7
3	4	3	2	5
4	1	2	3	7
5	2	2	5	1

R	C	M
1	X	2
1	Y	1

Sparse Matrices

R
1
2
3
4
5

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

R	C	M
1	X	2
1	Y	1
1	Z	1

Sparse Matrices

X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

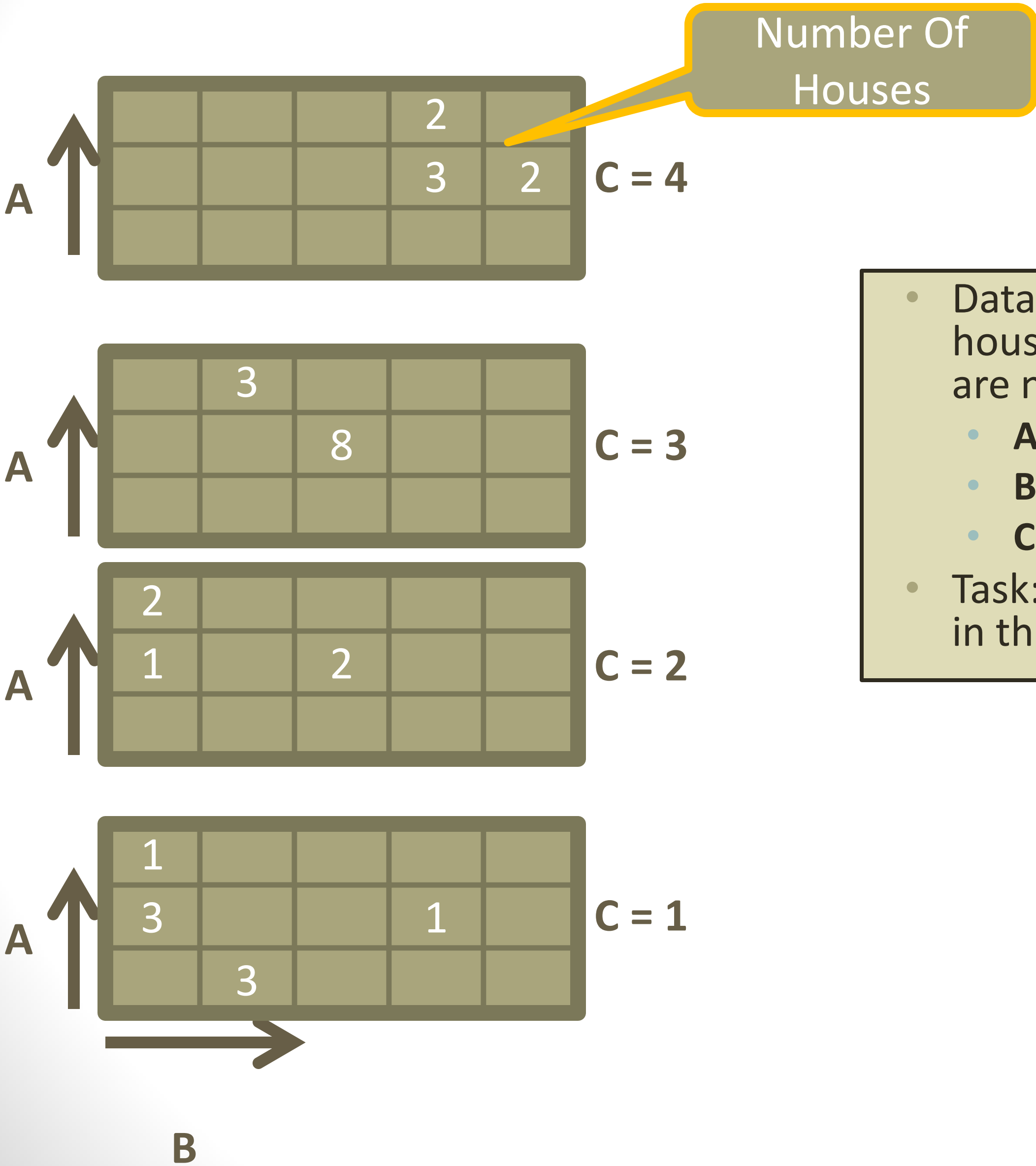
Column Name

Row ID.

Cell Values

R	C	M
1	X	2
1	Y	1
1	Z	1
1	V	11
2	X	2
2	Y	2
2	Z	2
2	V	7
3	X	4
3	Y	3
3	Z	2
3	V	5
4	X	1
4	Y	2
4	Z	3
4	V	7
5	X	2
5	Y	2
5	Z	5
5	V	1

Sparse Matrices: Exercise (1)



- Data: Real estate survey of single-family houses in downtown Seattle. Cell values are number (**N**) of houses found for sale.
  - **A**: Area in 1000's of square feet
  - **B**: Number of Bathrooms
  - **C**: Cost in $100,000.-
- Task: Create sparse matrices of the type in the previous slide.

Sparse Matrices: Exercise (2)

C = 4

			2	
			3	2

A

C = 3

	3			
		8		

A

C = 2

2				
1		2		

A

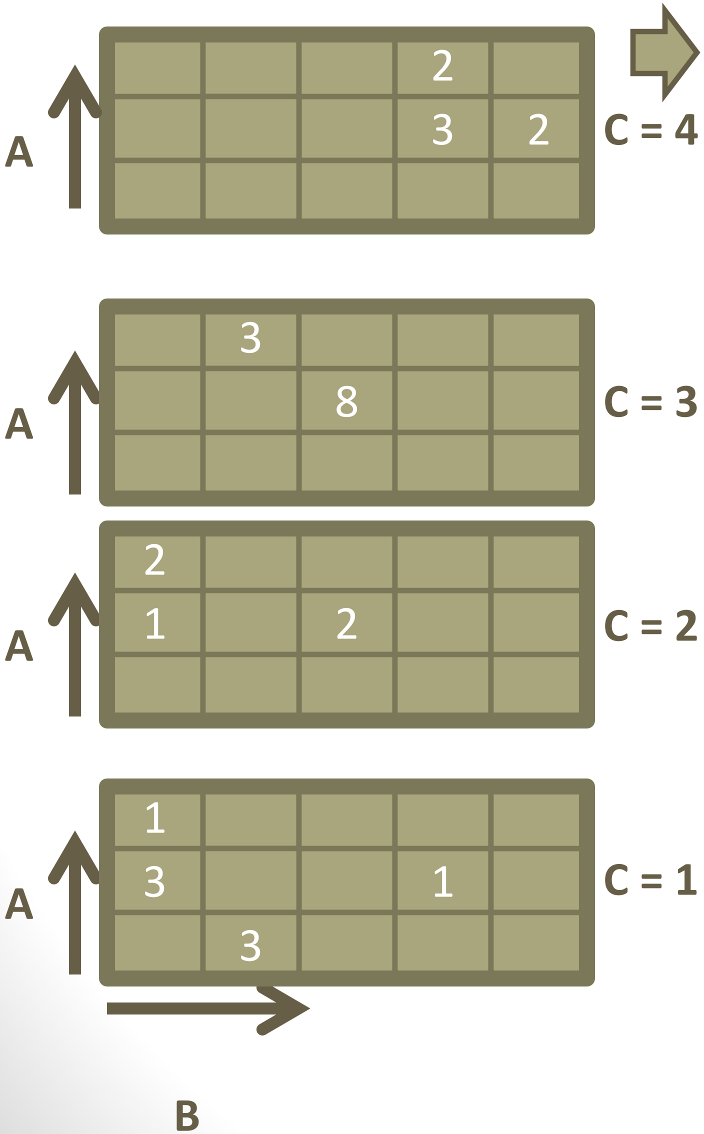
C = 1

1				
3			1	
	3			

A

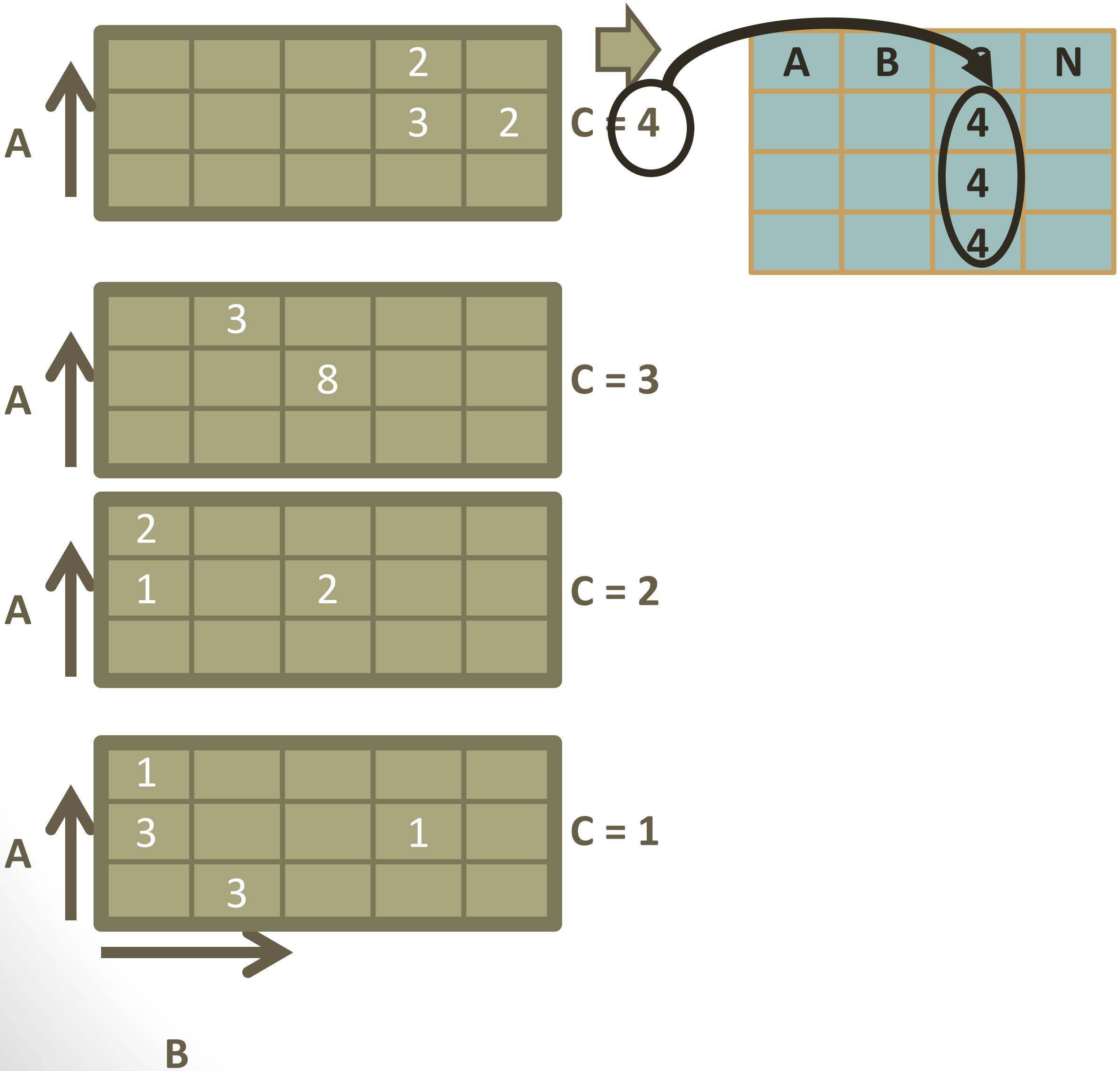
B

Sparse Matrices: Exercise (3)

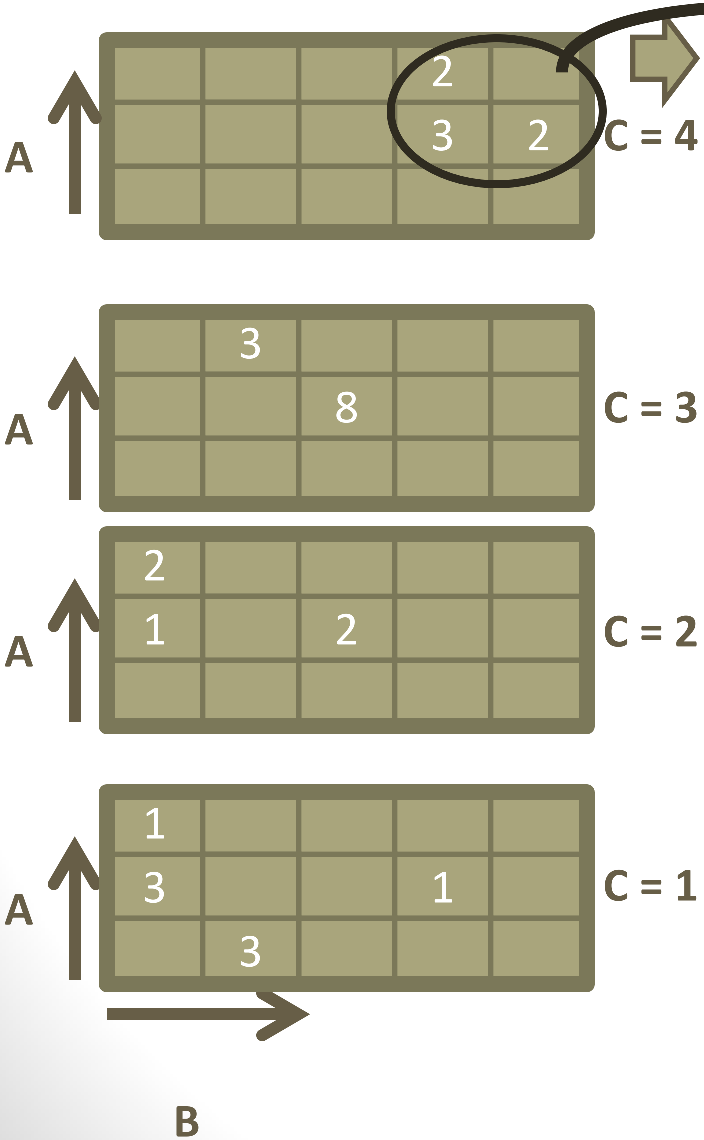


A	B	C	N

Sparse Matrices: Exercise (4)

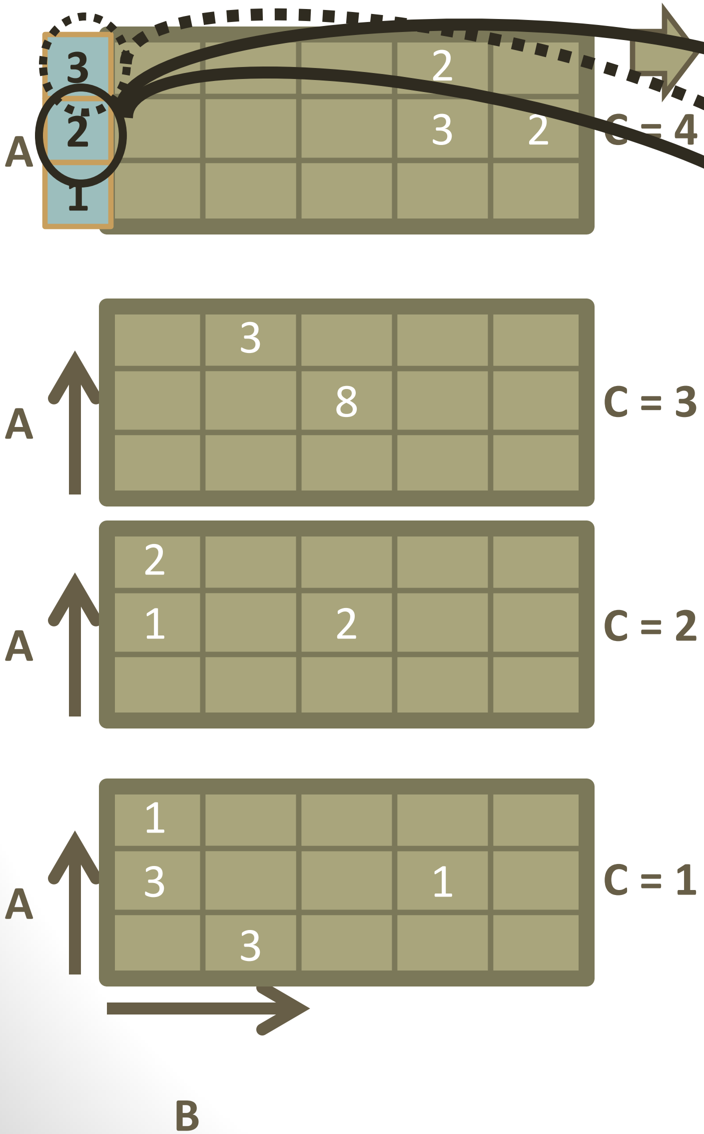


Sparse Matrices: Exercise (5)



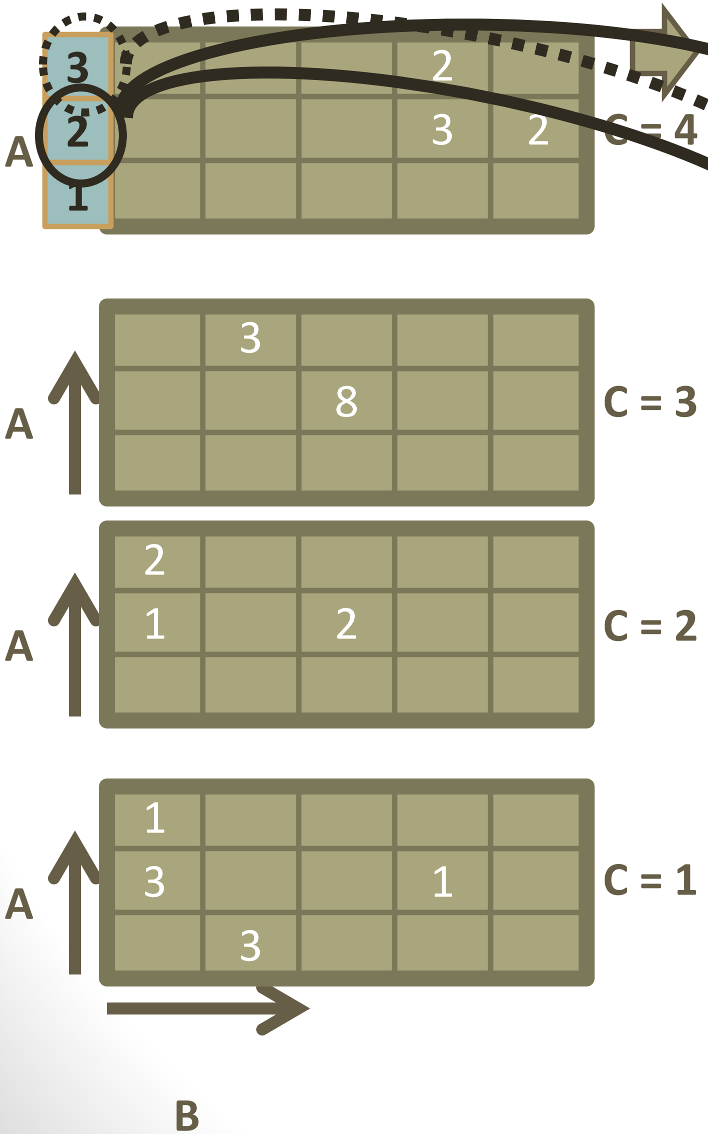
A	B	C	N
		4	2
		4	2
		4	3

Sparse Matrices: Exercise (6)



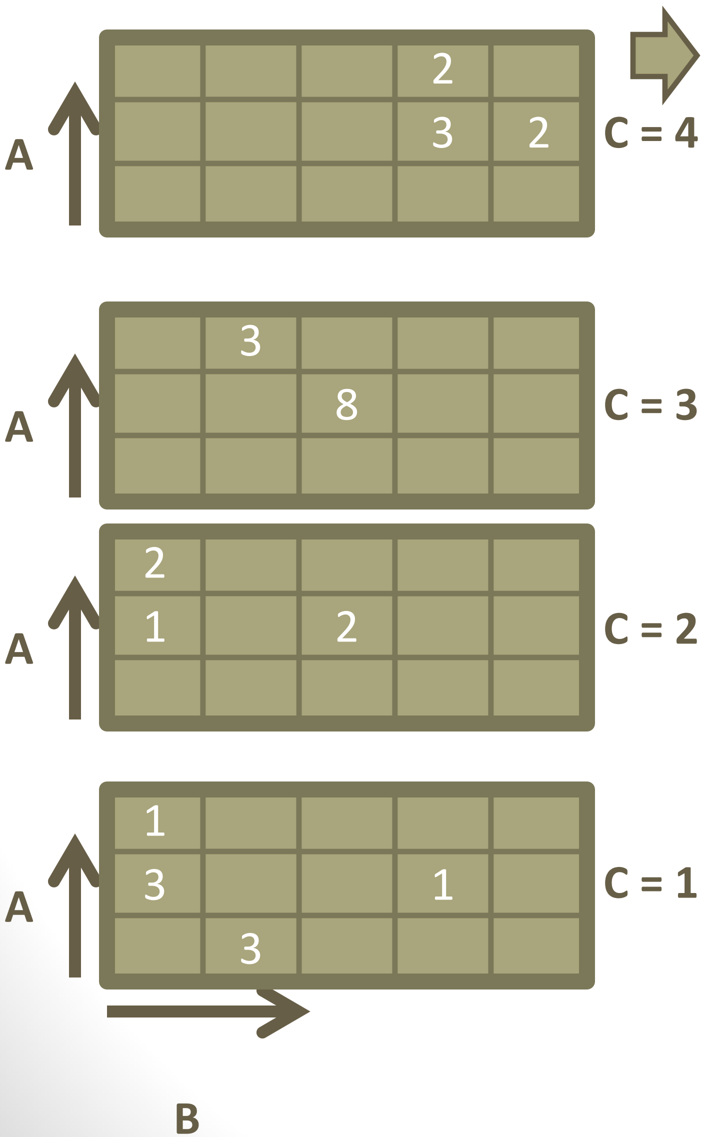
A	B	C	N
2		4	2
3		4	2
2		4	3

Sparse Matrices: Exercise (7)



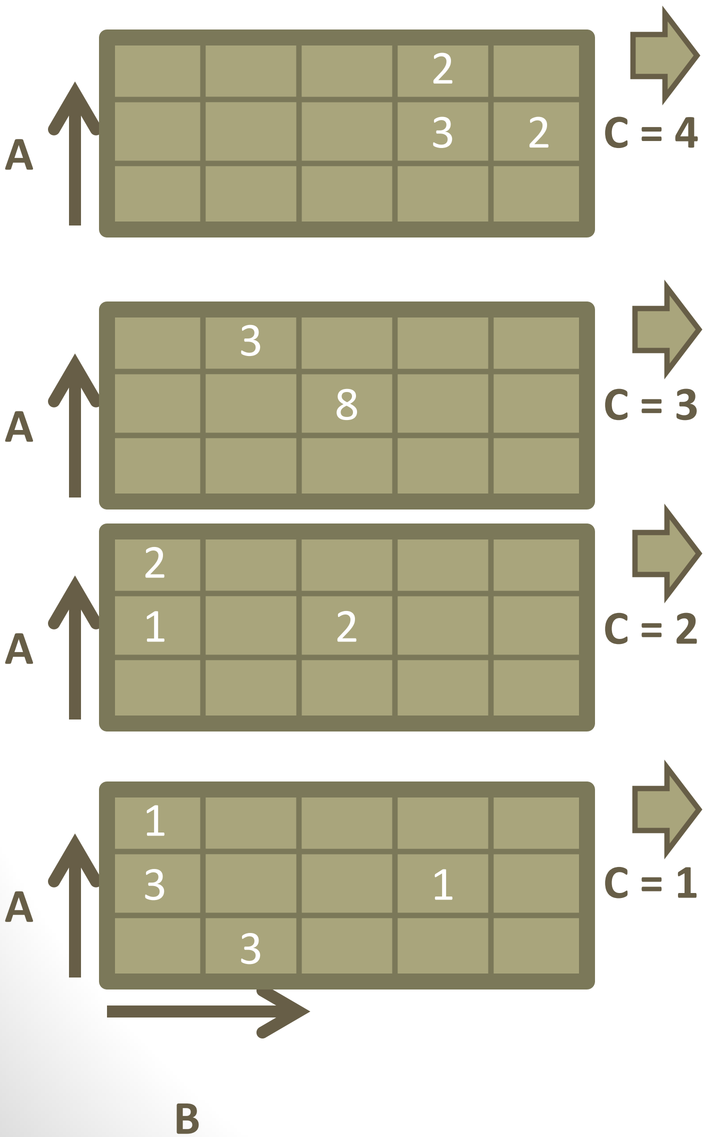
A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3

Sparse Matrices: Exercise (8)



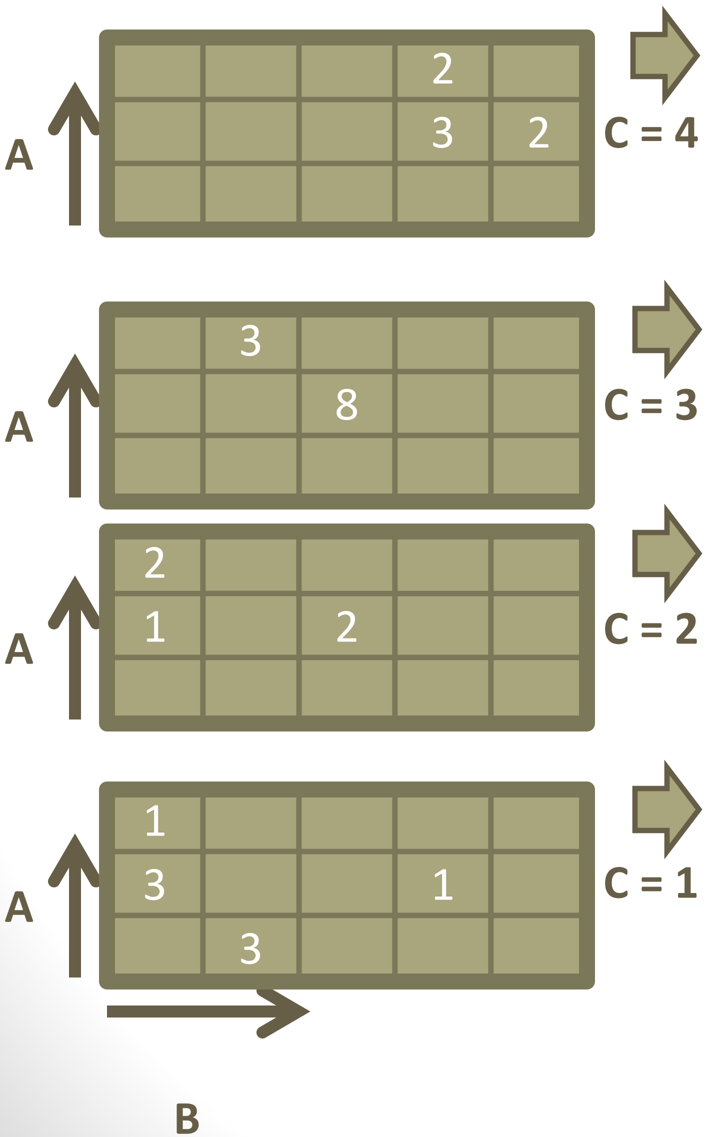
A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3

Sparse Matrices: Exercise (9)



A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3

Sparse Matrices: Exercise (10)



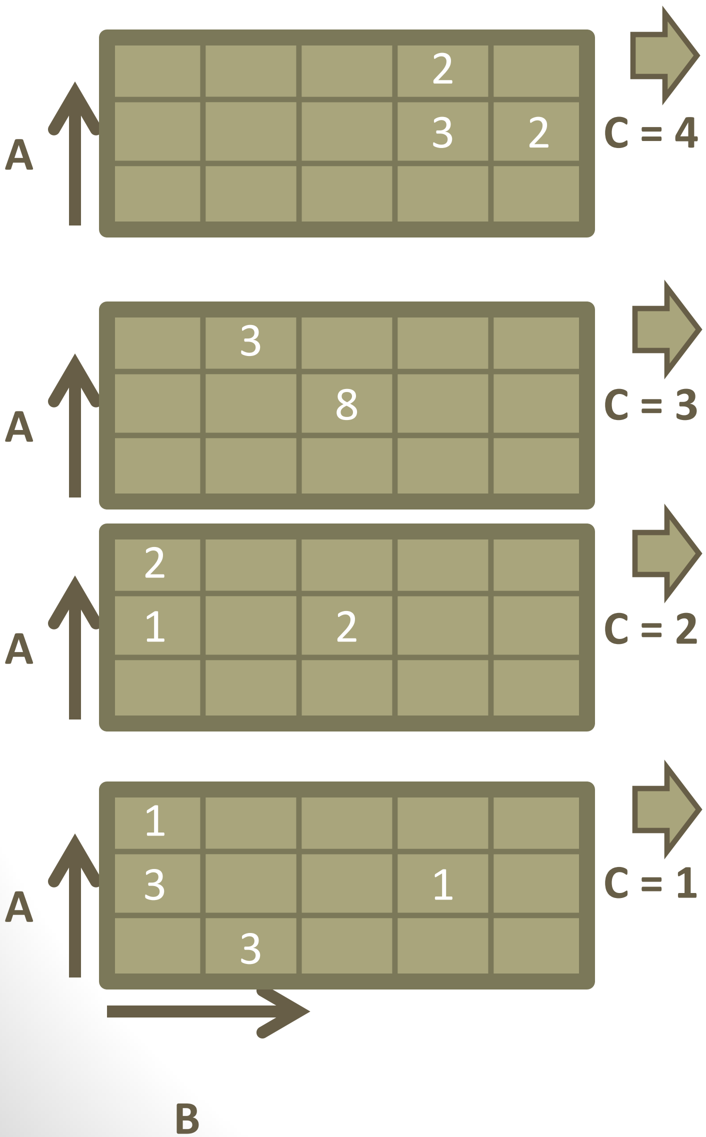
A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3

A	B	C	N
2	3	3	8
3	2	3	3

A	B	C	N
2	3	2	2
3	1	2	2
2	1	2	1

A	B	C	N
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

Sparse Matrices: Exercise (11)



A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3

A	B	C	N
2	3	3	8
3	2	3	3

A	B	C	N
2	3	2	2
3	1	2	2
2	1	2	1

A	B	C	N
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2
3	1	2	2
2	1	2	1
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

Sparse Matrices: Exercise (12)

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2
3	1	2	2
2	1	2	1
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

Sparse Matrices: Exercise (13)

	A	B	C	N
1	2	5	4	2
2	3	4	4	2
3	2	4	4	3
4	2	3	3	8
5	3	2	3	3
6	2	3	2	2
7	3	1	2	2
8	2	1	2	1
9	2	4	1	1
10	1	2	1	3
11	3	1	1	1
12	2	1	1	3

R	C	M
1		
1		
1		
1		
2		
2		
2		
2		
3		
3		
3		
3		
4		
4		
4		
4		
5		
5		
5		
5		

Sparse Matrices: Exercise (14)

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2
3	1	2	2
2	1	2	1
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

R	C	M
1	A	
1	B	
1	C	
1	N	
2	A	
2	B	
2	C	
2	N	
3	A	
3	B	
3	C	
3	N	
4		
4		
4		
4		
5		
5		
5		
5		

Sparse Matrices: Exercise (15)

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2
3	1	2	2
2	1	2	1
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4		
4		
4		
4		
5		
5		
5		
5		

Sparse Matrices: Exercise (16)

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2
3	1	2	2
2	1	2	1
2	4	1	1
1	2	1	3
3	1	1	1
2	1	1	3

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4		
4		
4		
4		
5		
5		
5		
5		

Homework: Matrices (17)

- Main Point:
  - Condensing information from multi-dimensional entity is good but not the main point.
  - The main point is to convince you that the last two tables represent multi-dimensional matrices (Hyper-rectangles, or Cartesian products of their intervals)
- Further Lessons:
  - These tables abide by the rules of relational algebra
    - Rows are unique
    - Columns have headers
    - Row order is irrelevant
  - Relaxed Layout / Schema
  - Extensible: New tables can be added without disrupting the schema

Schema Change: add a column

- Schema change can happen by adding rows (tuples) to a table that indexes another table

Schema Change: add a column

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

Schema Change: add a column

This Relation represents a sparse 3-D Matrix

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

This Relation represents a sparse 4-D Matrix

Schema Change: add a column

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

This Relation represents a sparse 3-D Matrix

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

This Relation represents a sparse 4-D Matrix

Schema Change: add a column

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

Represent Relation by indexing Row, Column, and Value

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

R	C	M
1	A	2
1	B	5
1	C	4

Represent Relation by indexing Row, Column, and Value

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2

Adding new rows to second table with a new index

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2
1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

Adding new rows to second table with a new index

A	B	C
2	5	4
3	4	4
2	4	4
2	3	3
3	2	3
2	3	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2
1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

Adding new rows to second table with a new index

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2
1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

Adding new rows to second table with a new index

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2
1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

Rows may be resorted without changing the relation

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
2	A	3
2	B	4
2	C	4
3	A	2
3	B	4
3	C	4
4	A	2
4	B	3
4	C	3
5	A	3
5	B	2
5	C	3
6	A	2
6	B	3
6	C	2

1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

Rows may be resorted without changing the relation

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4

2	A	3
2	B	4
2	C	4

3	A	2
3	B	4
3	C	4

4	A	2
4	B	3
4	C	3

5	A	3
5	B	2
5	C	3

1	N	2
2	N	2
3	N	3
4	N	8
5	N	3
6	N	2

6	A	2
6	B	3
6	C	2

Rows may be resorted without changing the relation

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4	A	2
4	B	3
4	C	3
4	N	8
5	A	3
5	B	2
5	C	3
5	N	3
6	A	2
6	B	3
6	C	2
6	N	2

Rows may be resorted without changing the relation

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4	A	2
4	B	3
4	C	3
4	N	8
5	A	3
5	B	2
5	C	3
5	N	3
6	A	2
6	B	3
6	C	2
6	N	2

Schema Change Proved

	A	B	C	N
1	2	5	4	2
2	3	4	4	2
3	2	4	4	3
4	2	3	3	8
5	3	2	3	3
6	2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4	A	2
4	B	3
4	C	3
4	N	8
5	A	3
5	B	2
5	C	3
5	N	3
6	A	2
6	B	3
6	C	2
6	N	2

Schema Change Proved

A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

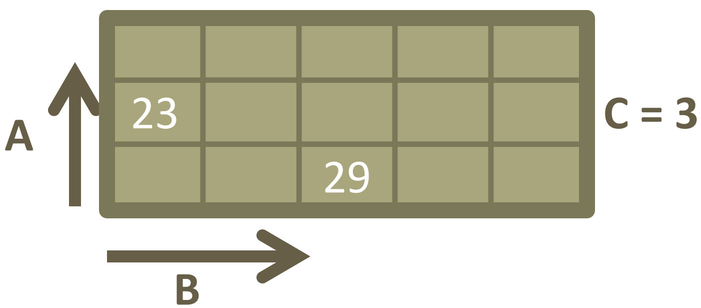
R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4	A	2
4	B	3
4	C	3
4	N	8
5	A	3
5	B	2
5	C	3
5	N	3
6	A	2
6	B	3
6	C	2
6	N	2

Schema Change Proved

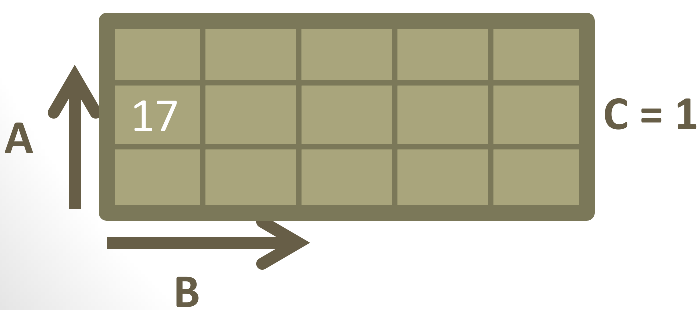
A	B	C	N
2	5	4	2
3	4	4	2
2	4	4	3
2	3	3	8
3	2	3	3
2	3	2	2

R	C	M
1	A	2
1	B	5
1	C	4
1	N	2
2	A	3
2	B	4
2	C	4
2	N	2
3	A	2
3	B	4
3	C	4
3	N	3
4	A	2
4	B	3
4	C	3
4	N	8
5	A	3
5	B	2
5	C	3
5	N	3
6	A	2
6	B	3
6	C	2
6	N	2

Sparse Matrices: Assignment



- Data: Real estate survey of single-family houses in downtown Seattle.   Cell values are number (**N**) of houses found for sale.
  - **A**: Area in 1000's of square feet
  - **B**: Number of Bathrooms
  - **C**: Cost in $100,000.-
- Task: Create sparse matrices of the type in the previous slide. See the following Assignment slide for elaboration.



Sparse Matrices Manipulation

Examples of Sparse Matrix Manipulation in a database

(see MatrixAlgebra.sql)

- Matrix Addition
- Scalar Multiplication
- Matrix Multiplication
  - Inner Product (Dot Product, Scalar Product)
  - Outer Product (Cartesian Product)
- Matrix Transposition

Sparse Matrices Assignment

1. Create the two tables that result from the “Sparse Matrices: Assignment” slide.
   a) Table 1 will have as headers: A, B, C, & N.
   b) Table 2 will have as its headers: R, C, & M.
2. Change the schema by changing Table 2. The new values will represent Cost per Square Foot.
4. SQL on Sparse Matrices. Given that sparse matrices are encoded with the EAV schema do the following:
   a) Write SQL for scalar multiplication of a Sparse Matrix (See Exercise 5 in MatrixAlgebra.sql
   b) Write SQL for transposition of a Sparse Matrix (See Exercise 6 in MatrixAlgebra.sql
   c) Optional: Write SQL for addition of two matrices
5. Submit Completed Assignment

Data as Sparse Matrices