Sparse 2D Matrix Format

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

2D Matrix

- A 2D Matrix is a rectangular data structure with rows and columns
- Each matrix element is uniquely identified by a row and column
- All matrix elements have the same type
  - Not so in a table (dataframe)!

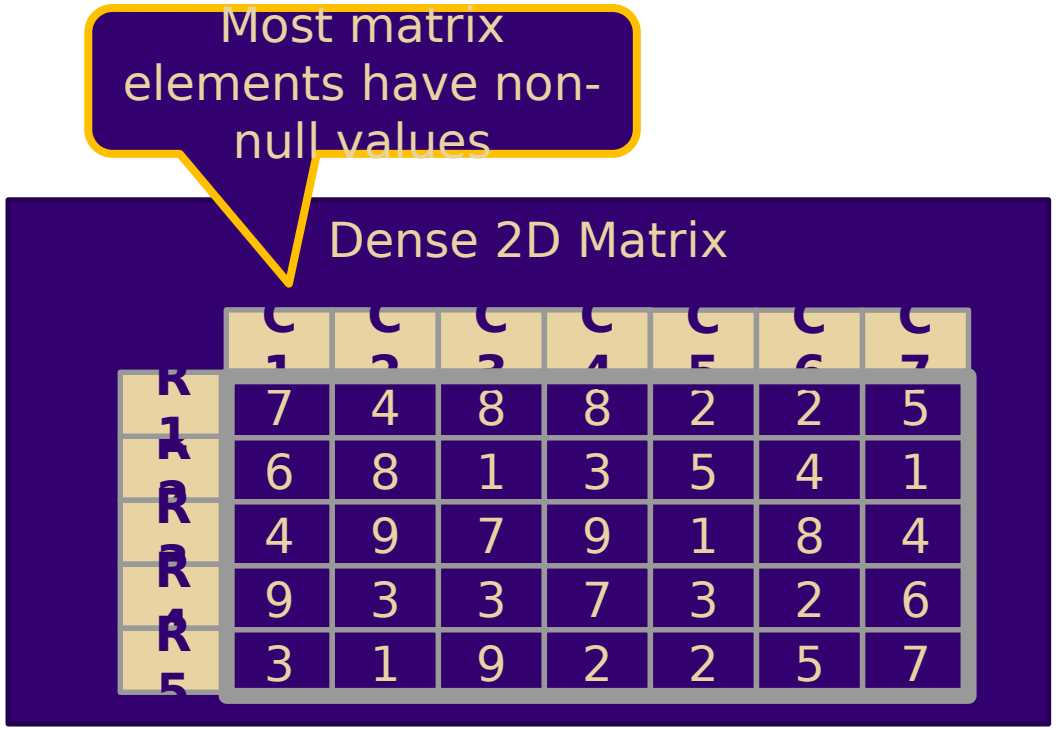
```
>>> np.random.random_integers(low=1, high=9, size=[5, 7])
```

Here is Python code to generate a 2D matrix

Result of the Python code. The row and column indices are not shown

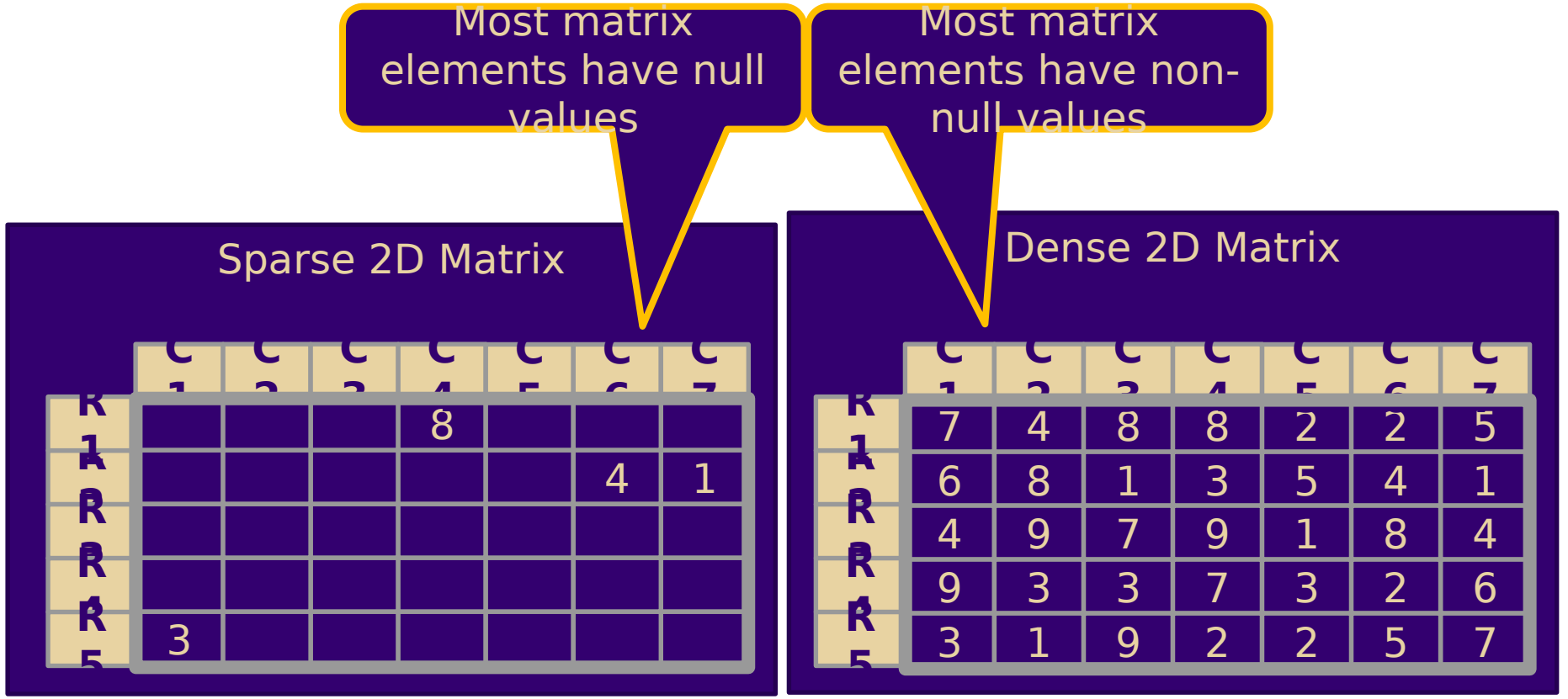
```
array([
  [7,  4,  8,  8,  2,  2, 5],
  [6,  8,  1,  3,  5,  4, 1],
  [4,  9,  7,  9,  1,  8, 4],
  [9,  3,  3,  7,  3,  2, 6],
  [3,  1,  9,  2,  2,  5,
```

2D Matrix - dense



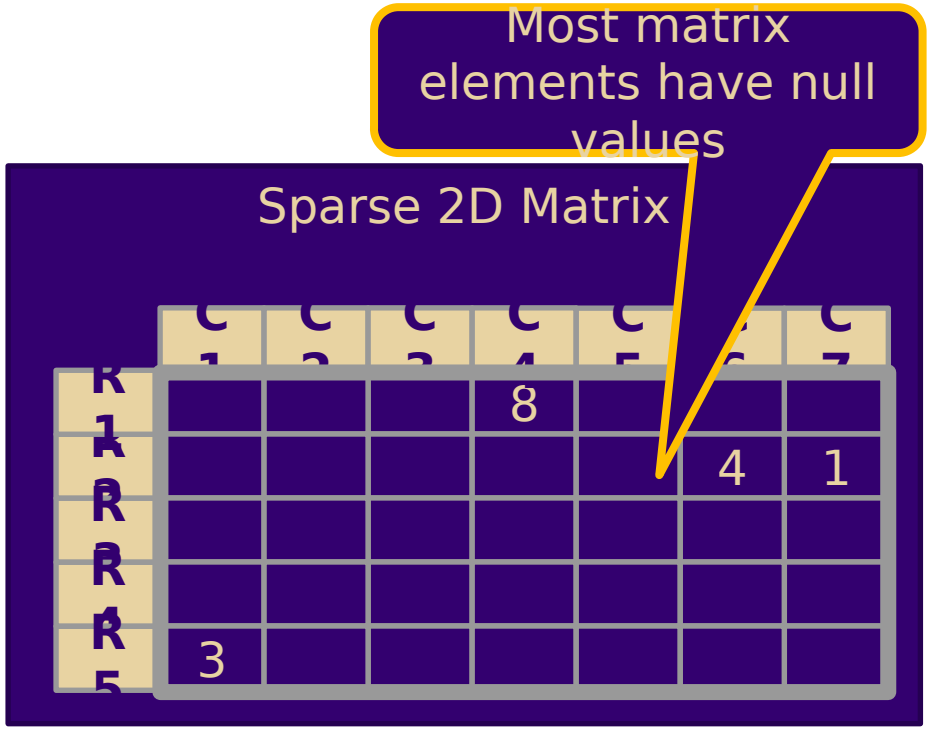
Sparse 2D Matrix Format

In a Sparse 2D Matrix most of the matrix elements are null



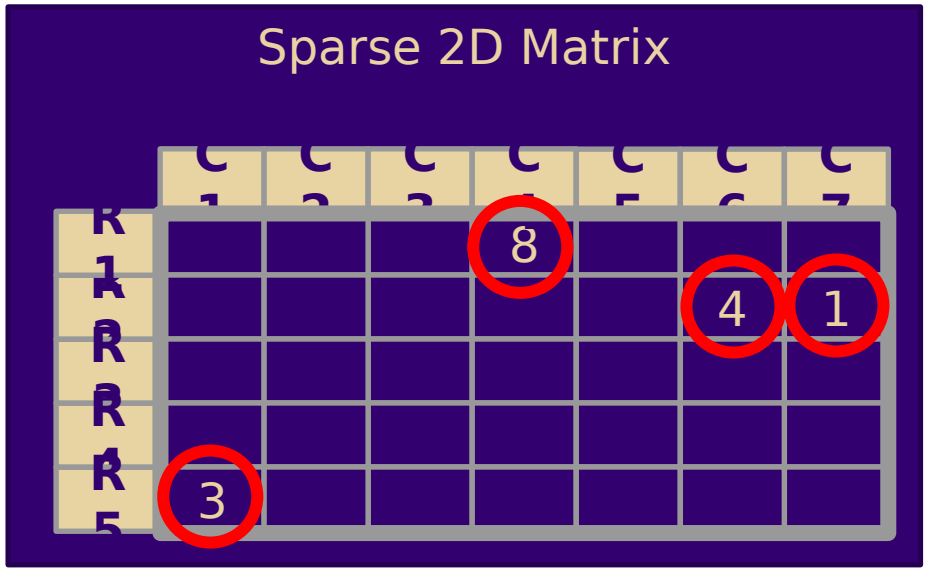
Sparse 2D Matrix Format

In a Sparse 2D Matrix most of the matrix elements are null
The traditional matrix layout is wasteful for large sparse matrices



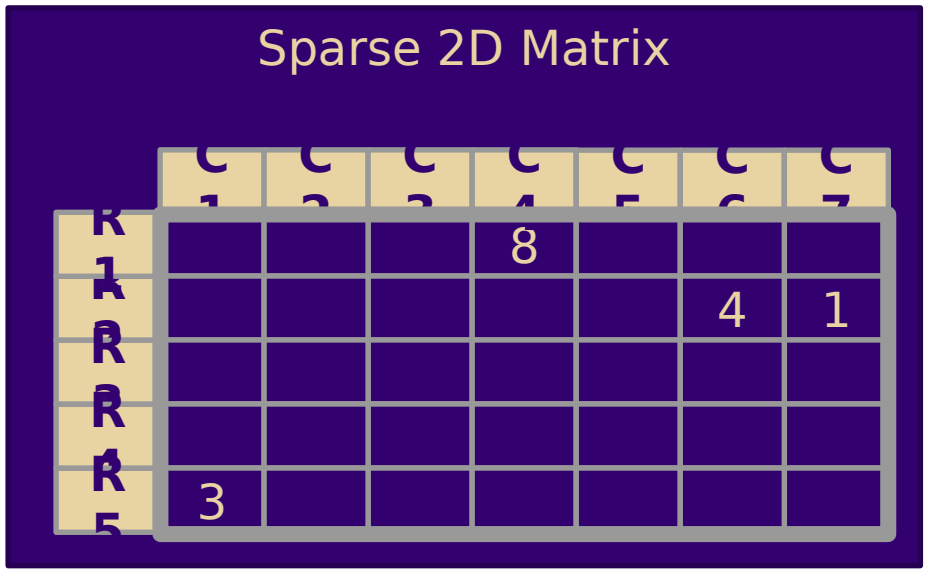
Sparse 2D Matrix Format

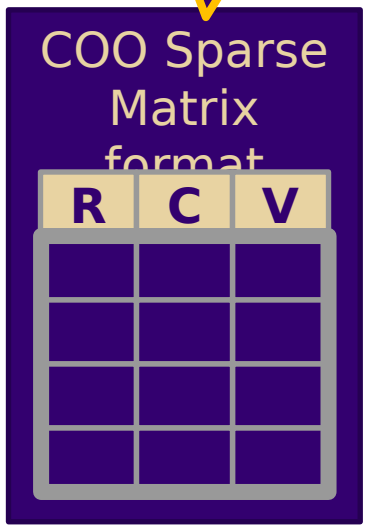
In a Sparse 2D Matrix most of the matrix elements are null.
The traditional matrix layout is wasteful for large sparse matrices
An efficient representation would only reference non-null values.



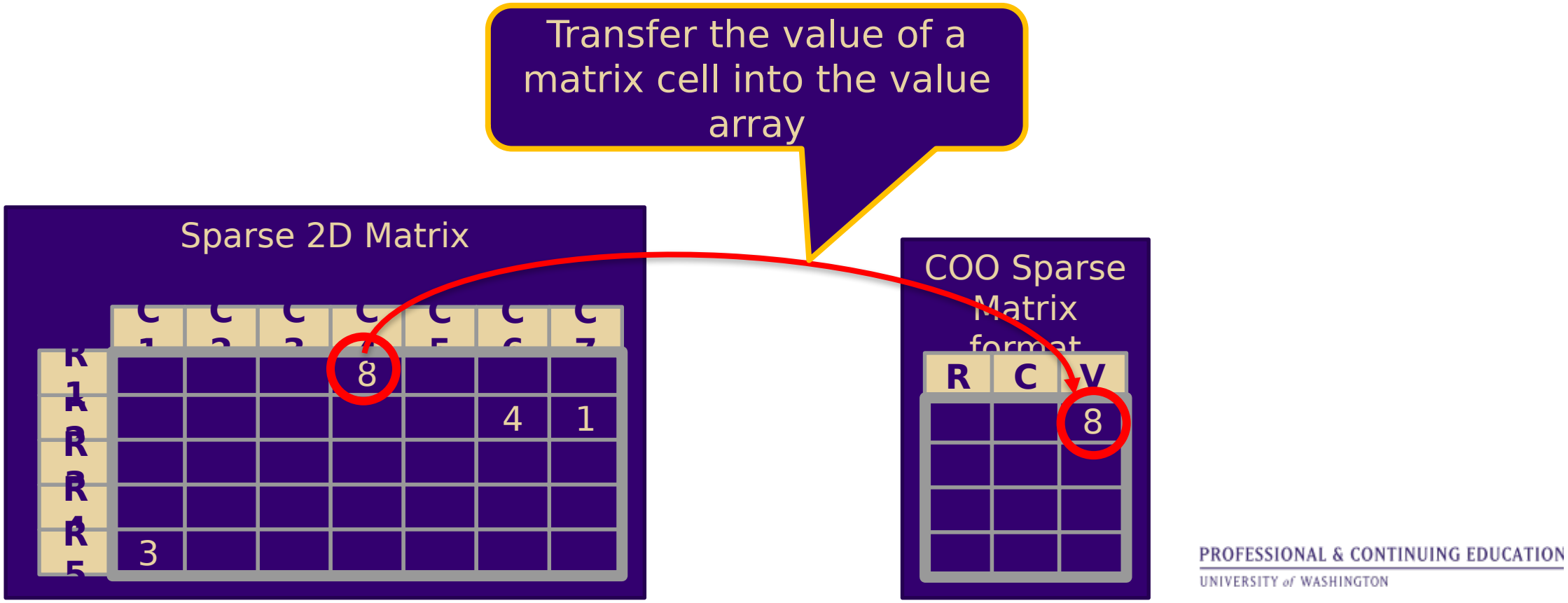
Sparse 2D Matrix Format

This Sparse 2D Matrix format has three arrays: Row (R), Column (C), and Value (V)

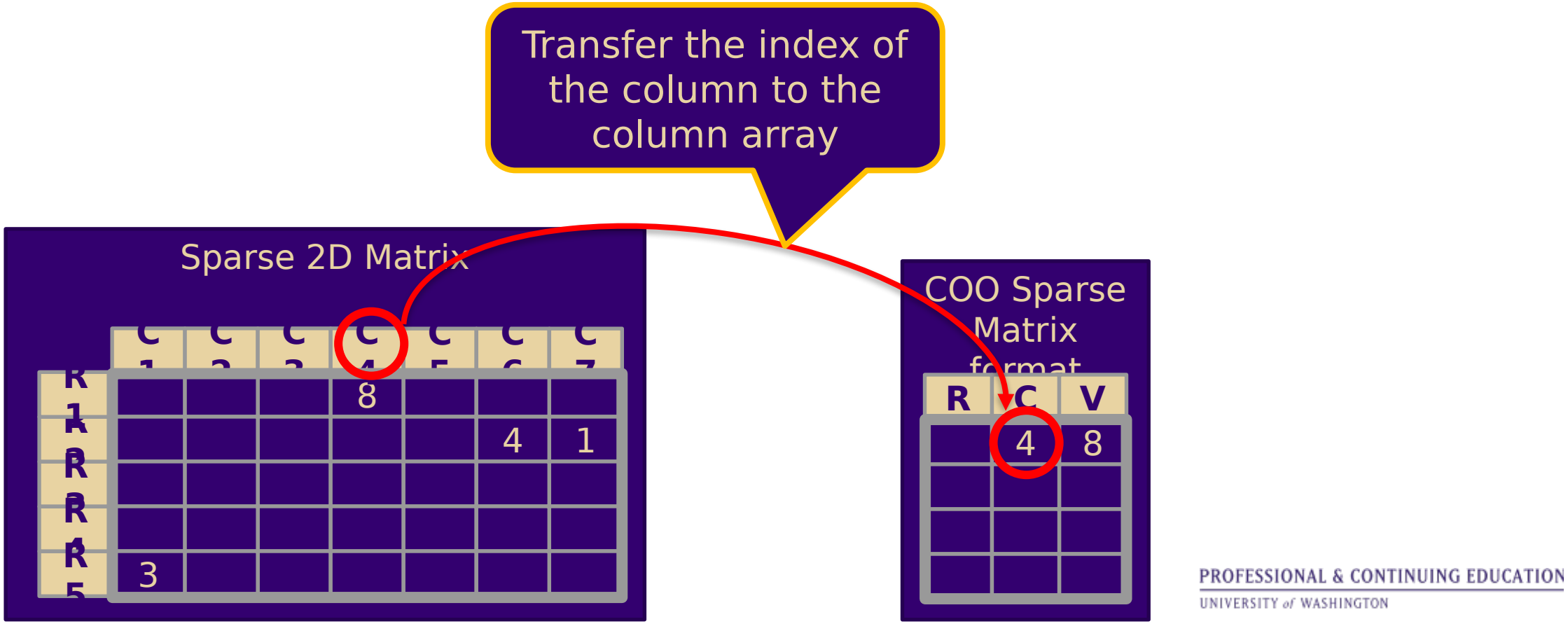




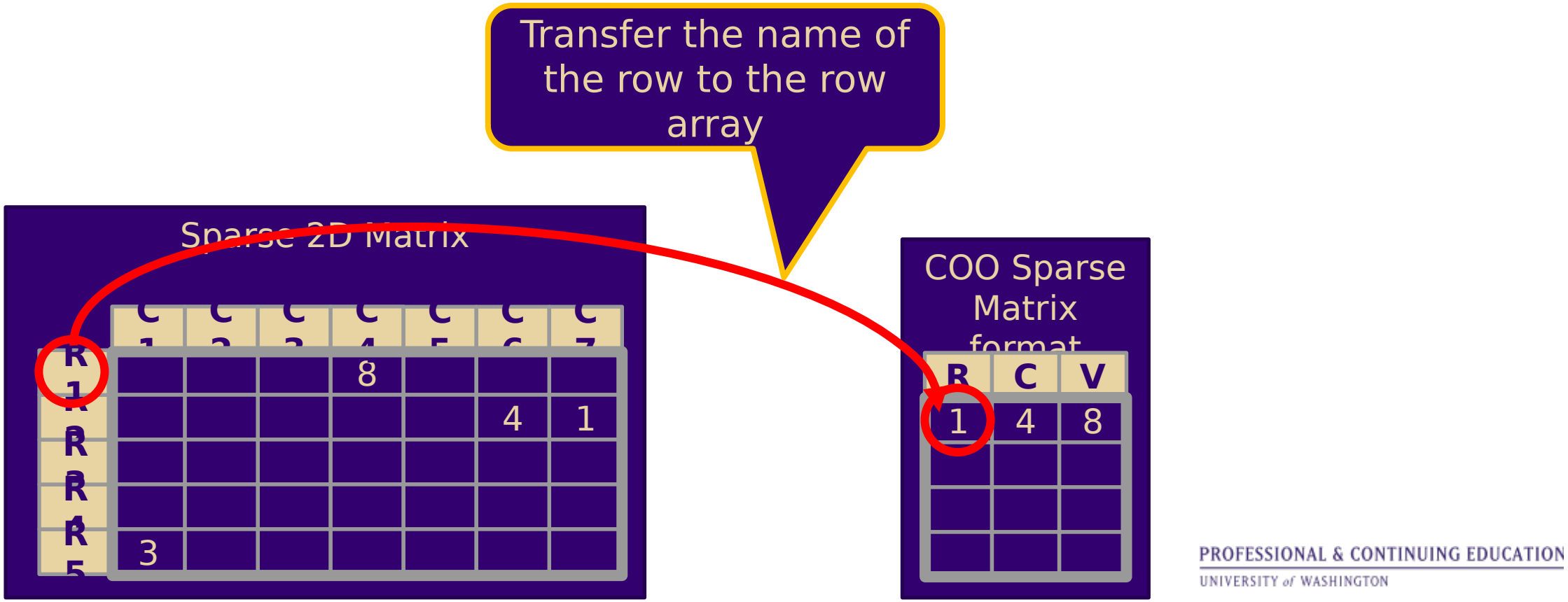
Sparse 2D Matrix Format



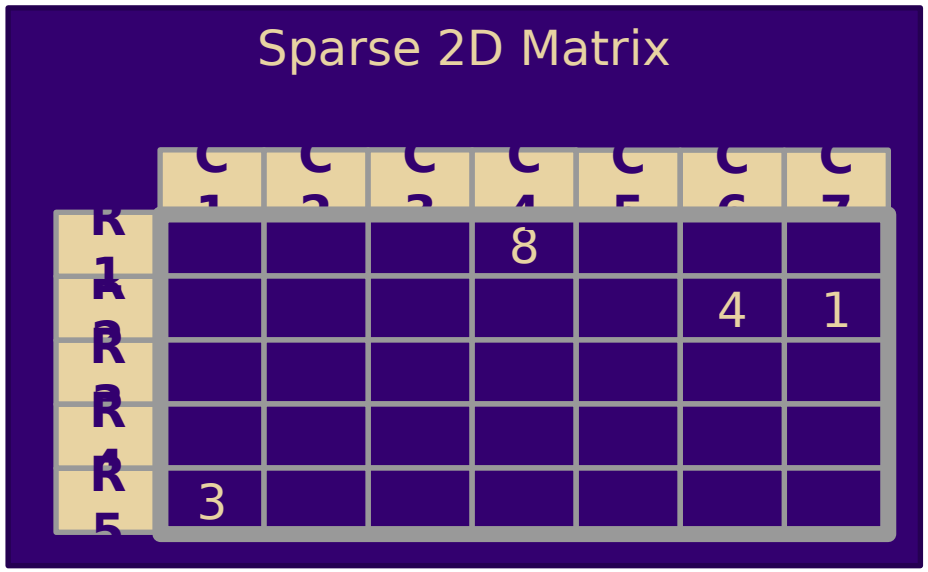
Sparse 2D Matrix Format



Sparse 2D Matrix Format



Sparse 2D Matrix Format

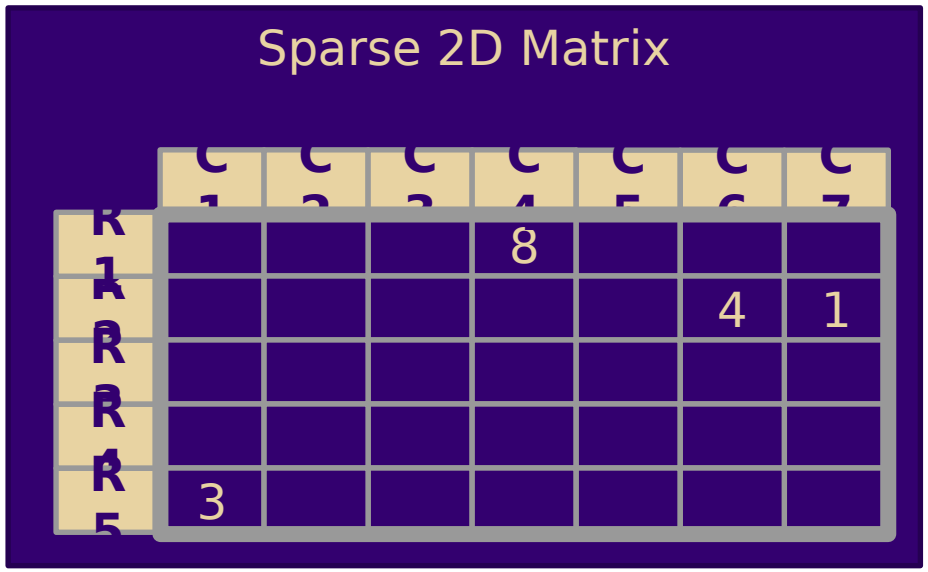


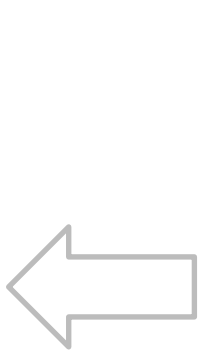
Sparse 2D Matrix Format. Does not contain null values

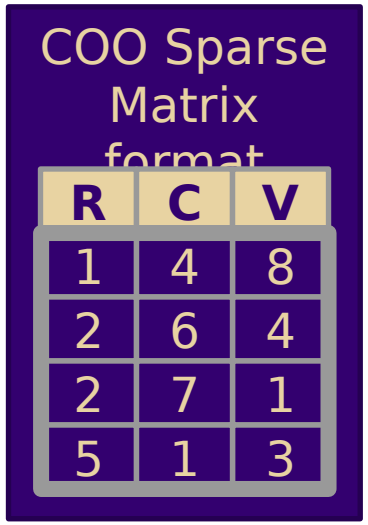
COO Sparse Matrix format

R	C	V
1	4	8
2	6	4
2	7	1
5	1	3

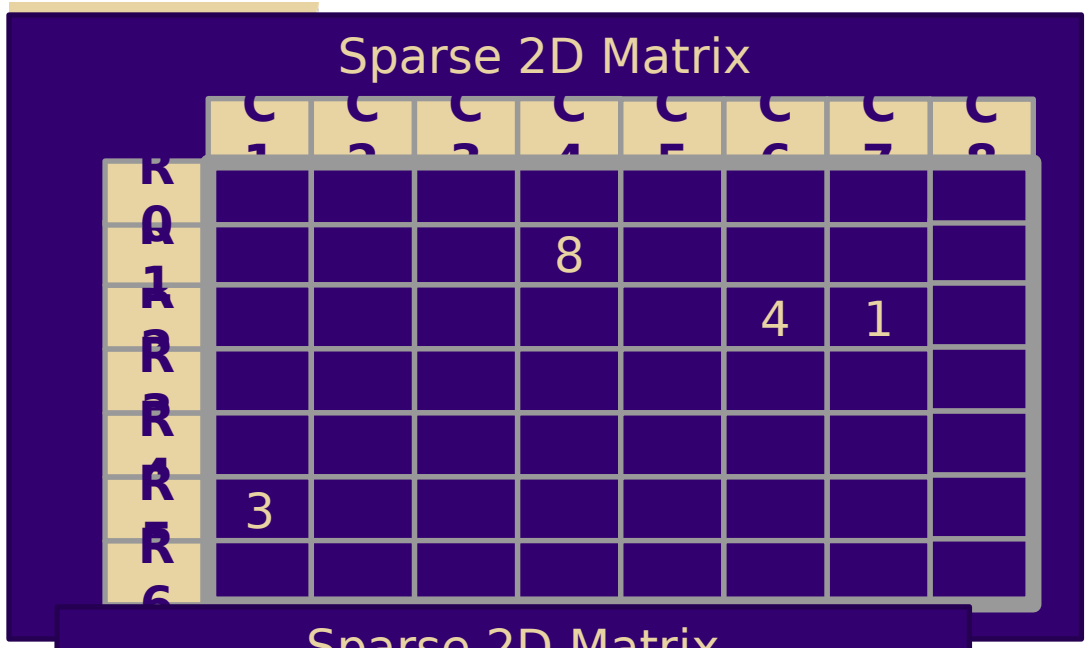
Sparse 2D Matrix Format



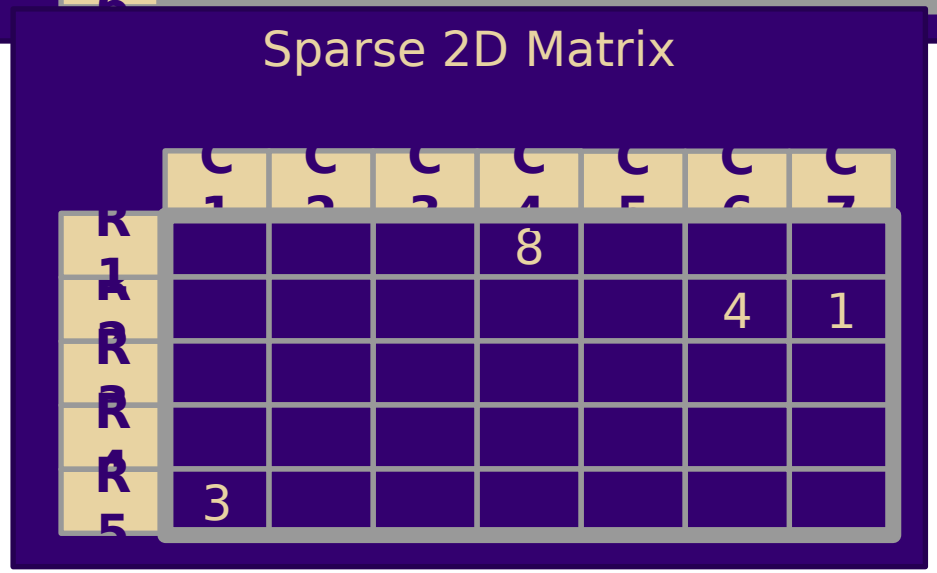




Sparse 2D Matrix Format



Note, the ambiguity!
Need dimension limits or vocabulary to be unambiguous.



COO Sparse Matrix format

R	C	V
1	4	8
2	6	4
2	7	1
5	1	3

Sparse 2D Matrices

Examples of Sparse 2D-Matrix Manipulation in a relational database

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

http://www.scipy-lectures.org/advanced/scipy_sparse/coo_matrix.html

Sparse 2D Matrix Format

Data as Multi-Dimensional Sparse Matrices

How algorithms view tables

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

Sparse Matrices

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.

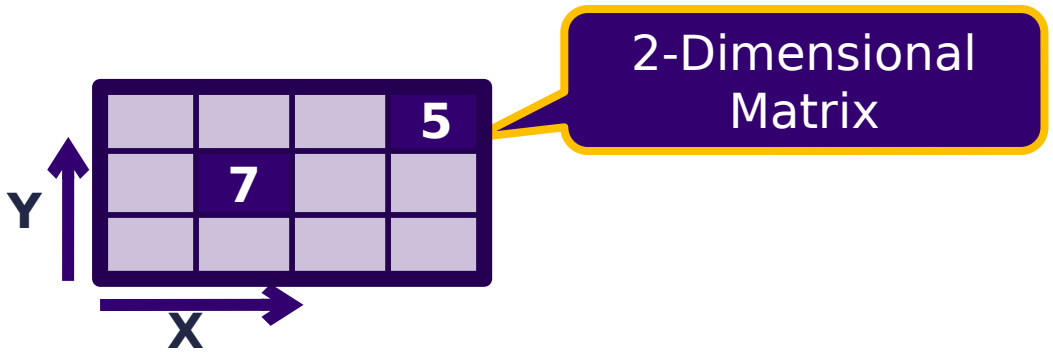
Relational Algebra

- 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.

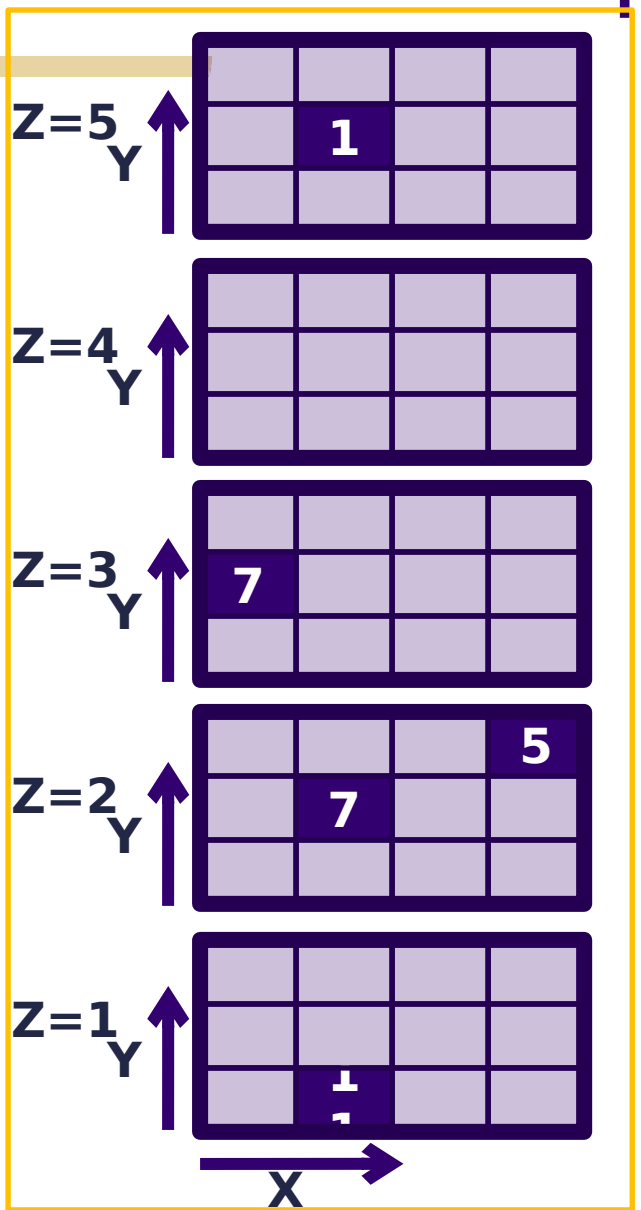
Hyperrectangle (Sparse Multi-Dimensional Matrix)

- 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

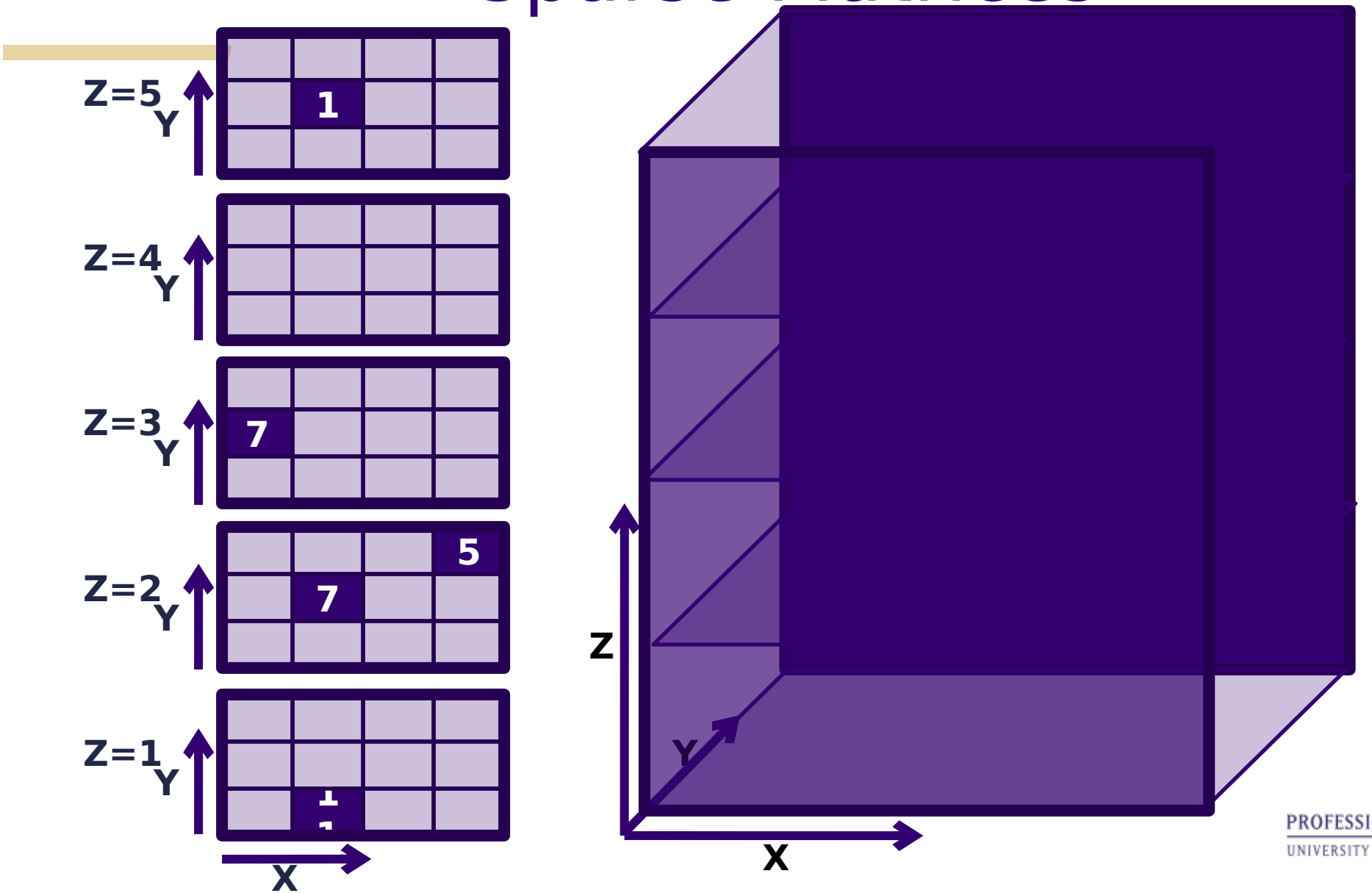


Sparse Matrices

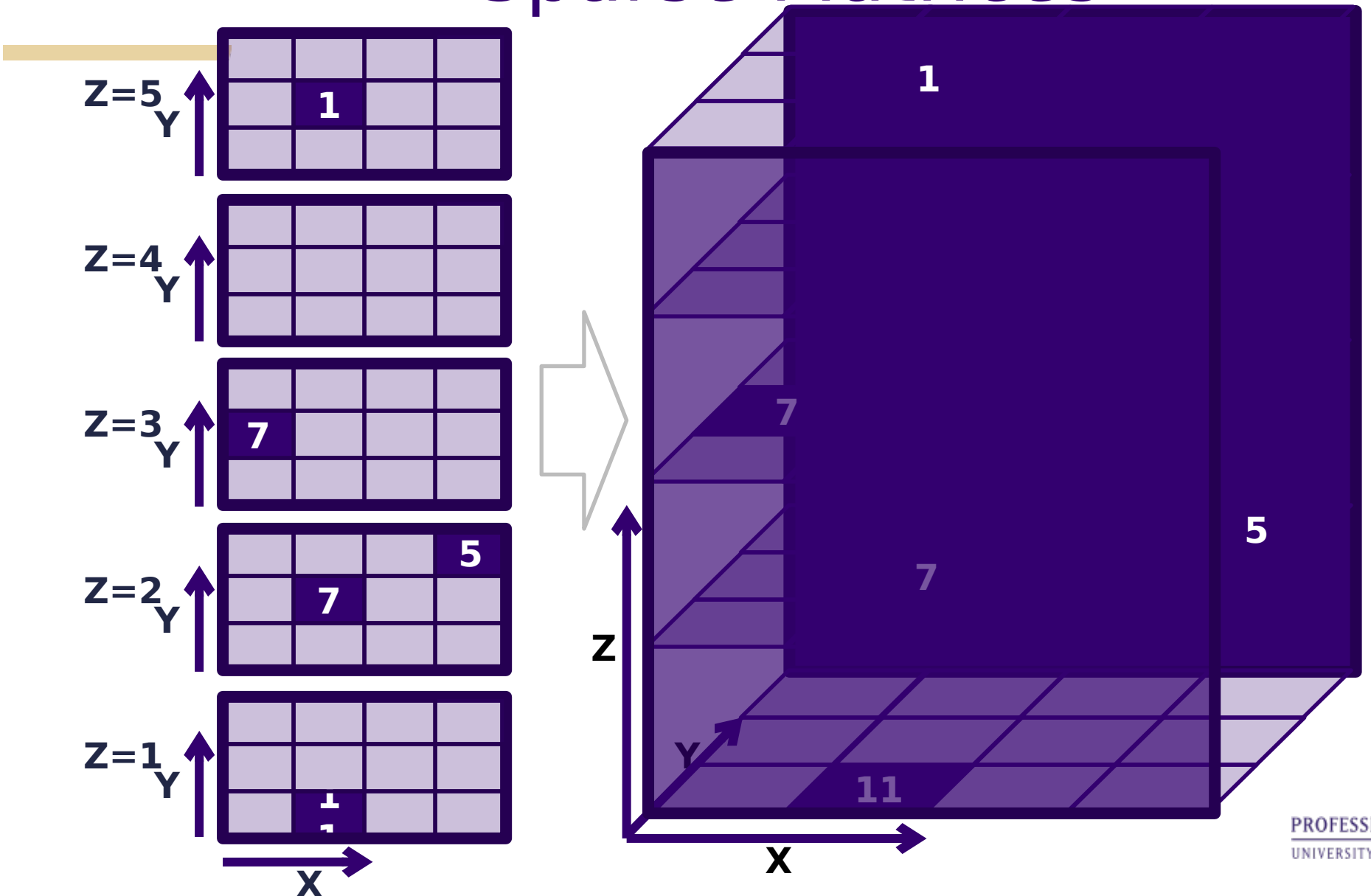


A series of equal-sized 2-dimensional matrices is a 3-dimensional matrix

Sparse Matrices



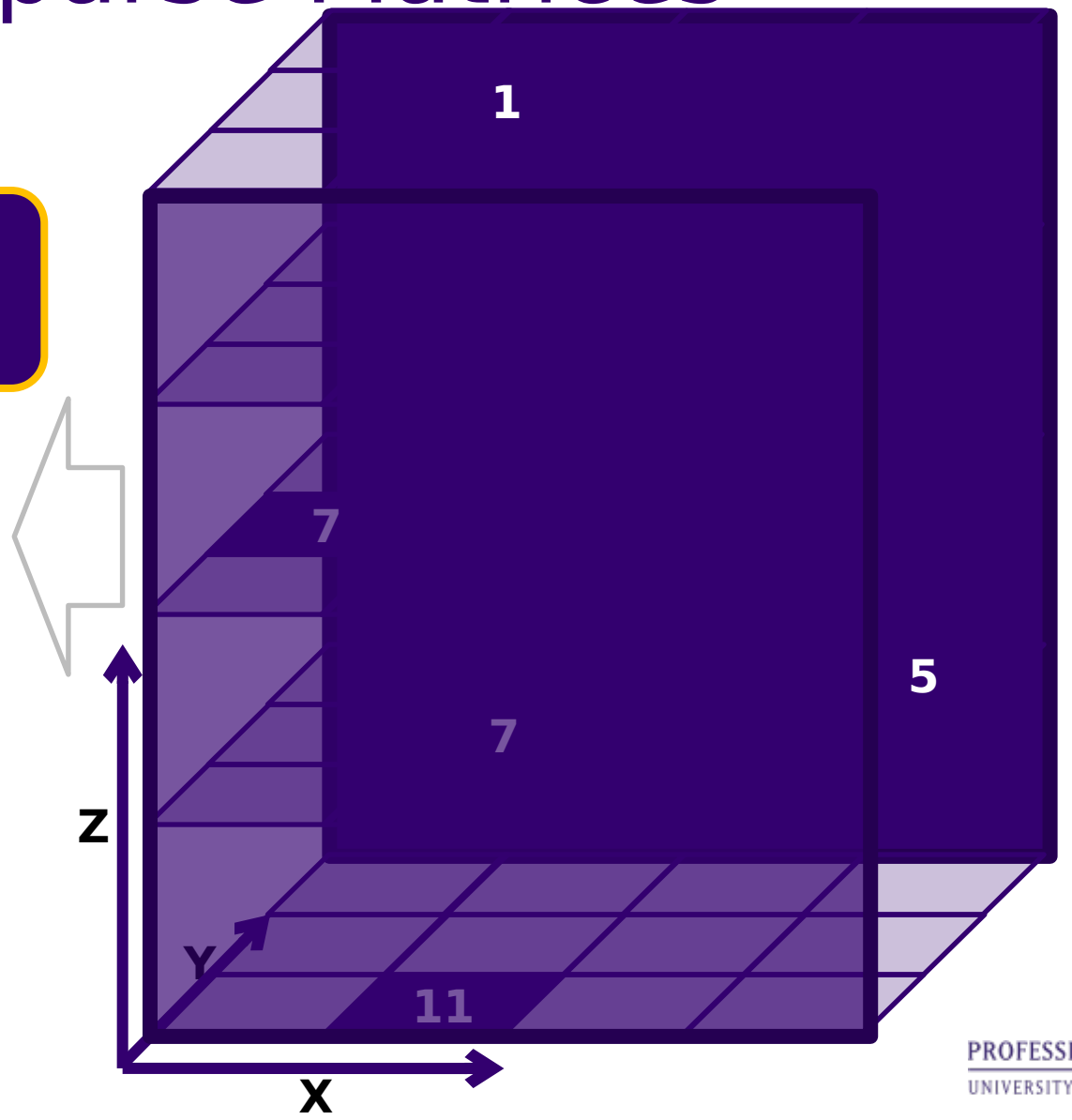
Sparse Matrices



Sparse Matrices

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

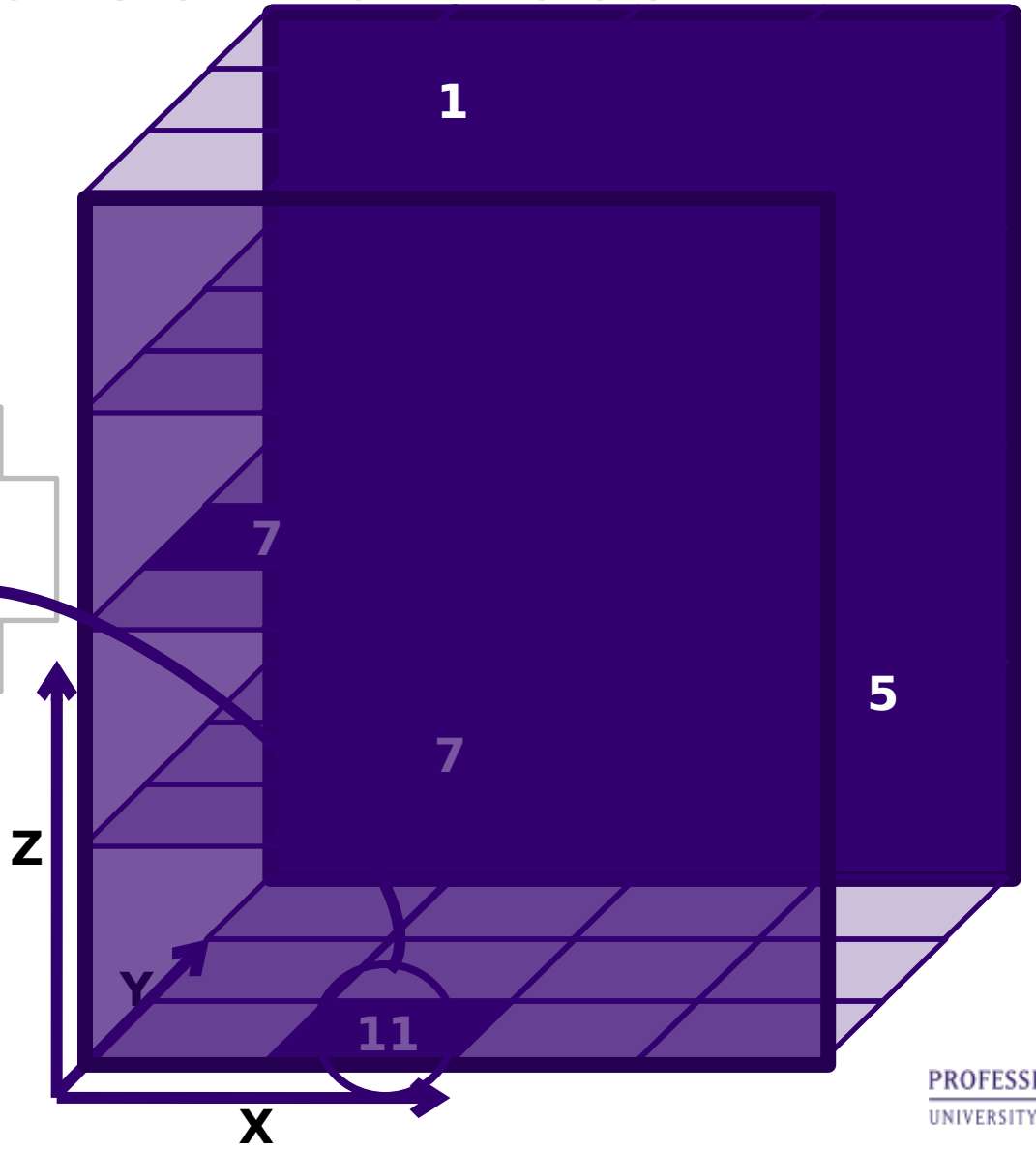
X	Y	Z	V



Sparse Matrices

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

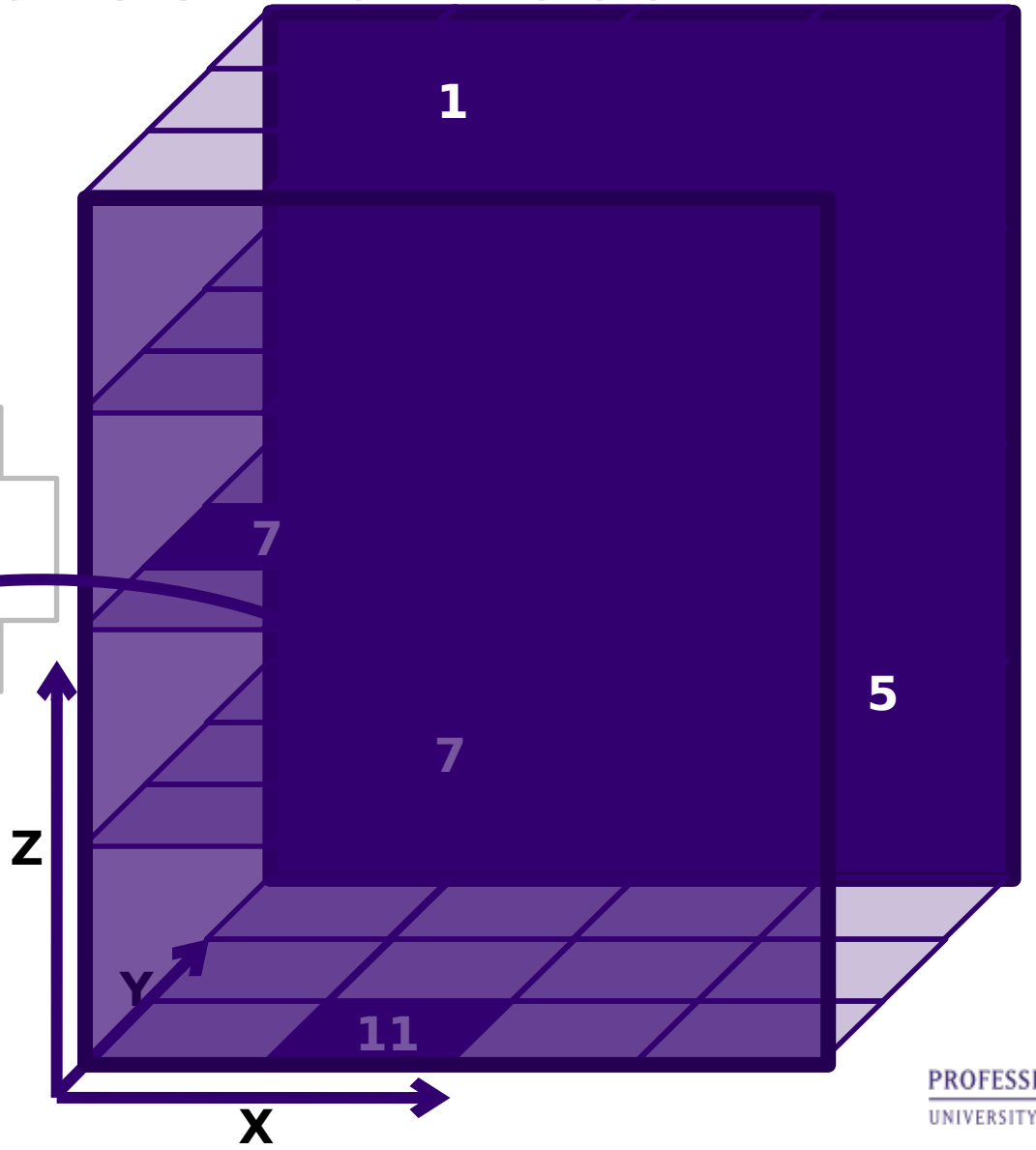
X	Y	Z	V
2	1	1	1
			1



Sparse Matrices

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

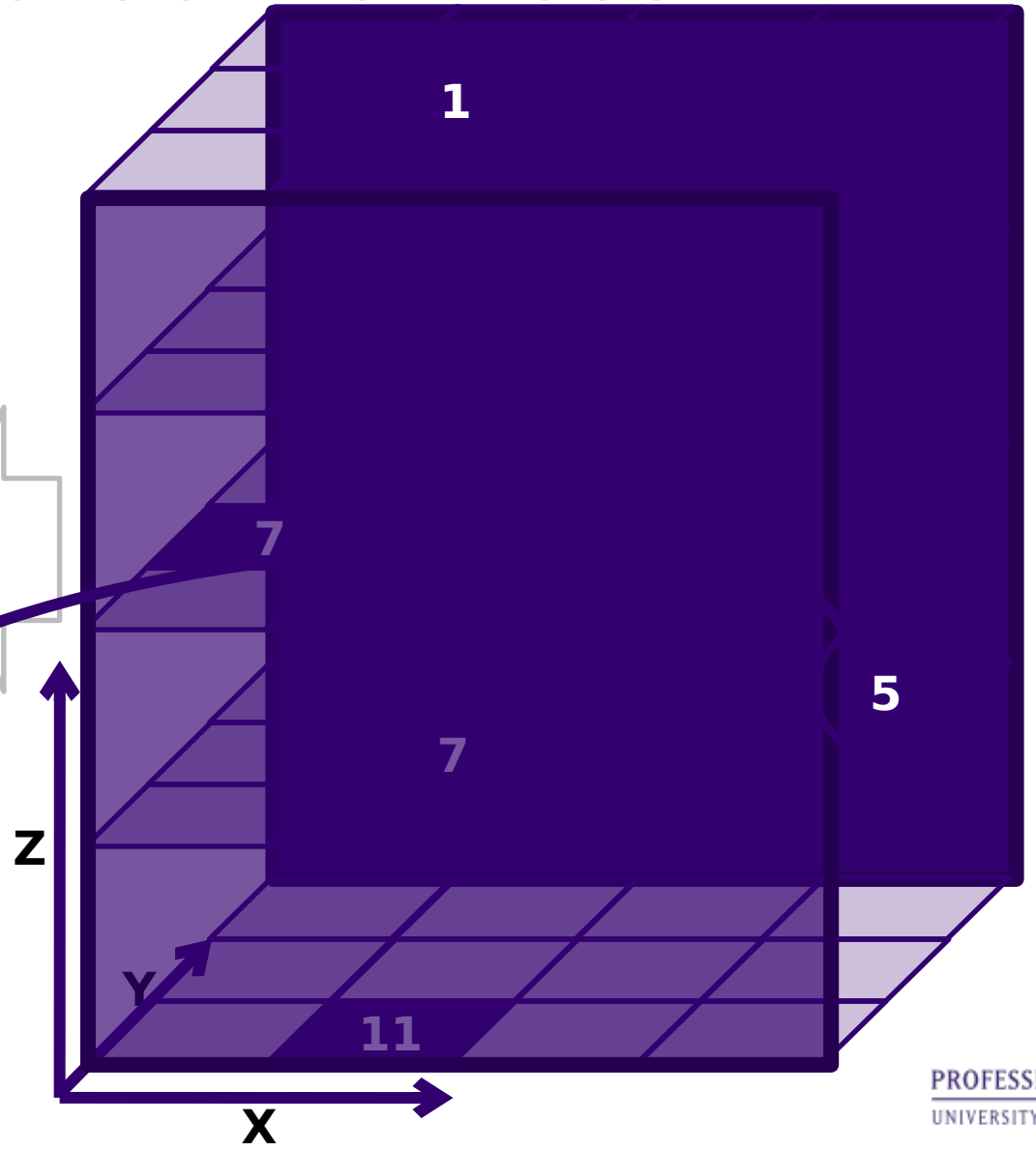
X	Y	Z	V
2	1	1	1
2	2	2	7



Sparse Matrices

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

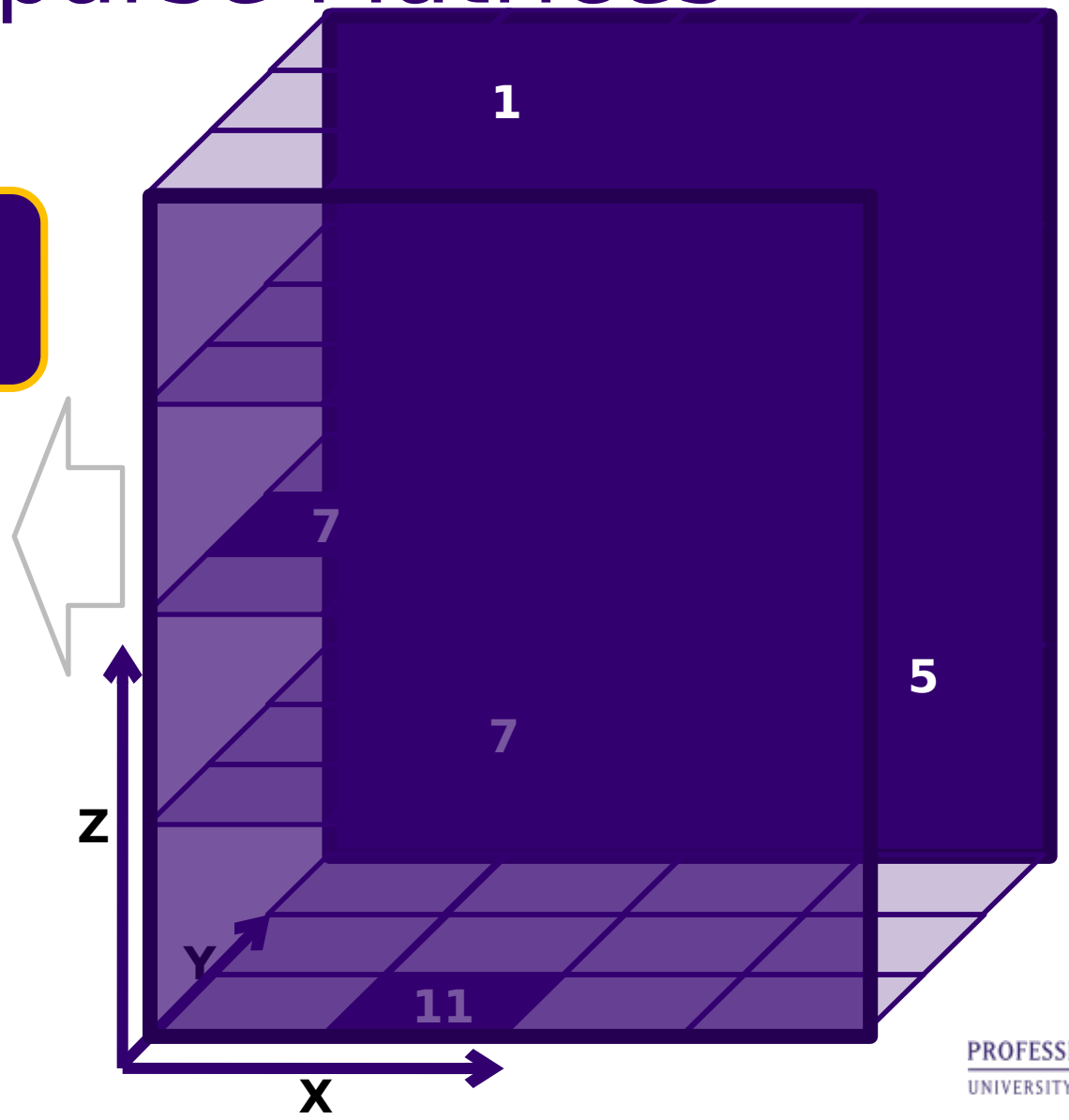
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5



Sparse Matrices

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

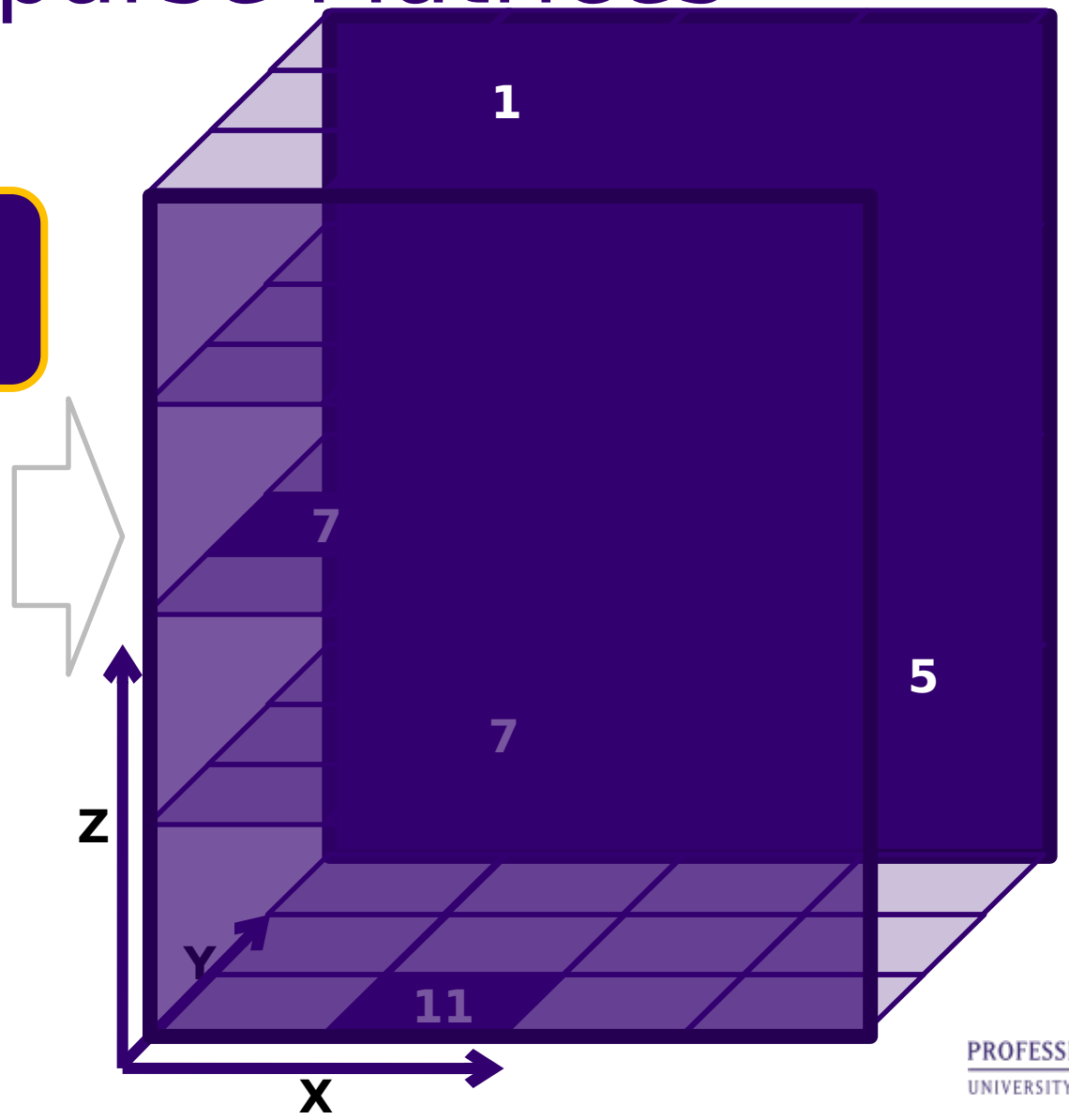
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

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

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

Think of V as just another dimension

Sparse Matrices

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

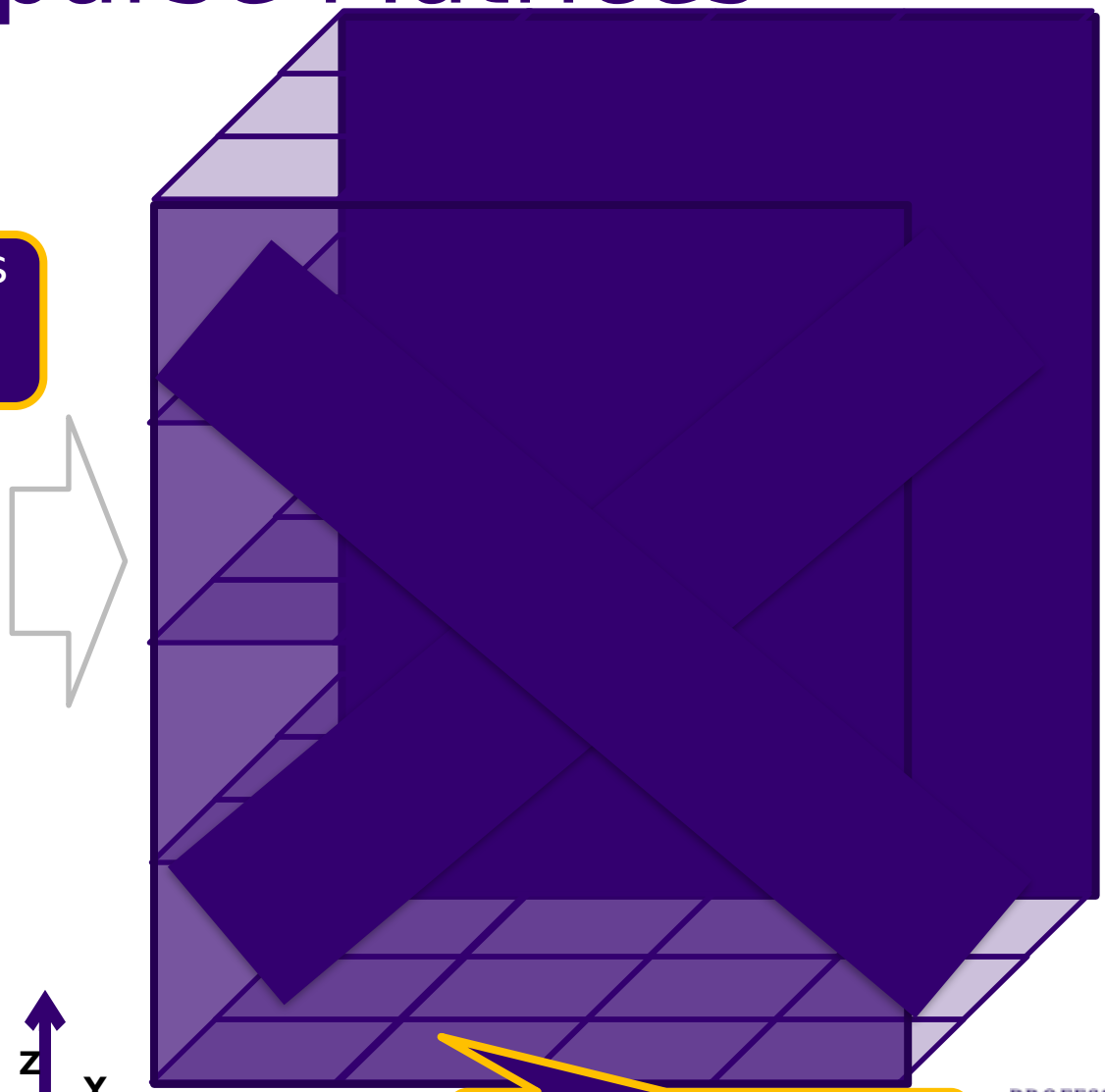
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

Think of V as just another dimension

Sparse Matrices

This table represents points in 4-Dimensional Space.

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

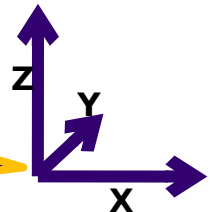


Sparse Matrices

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

3 Dimension

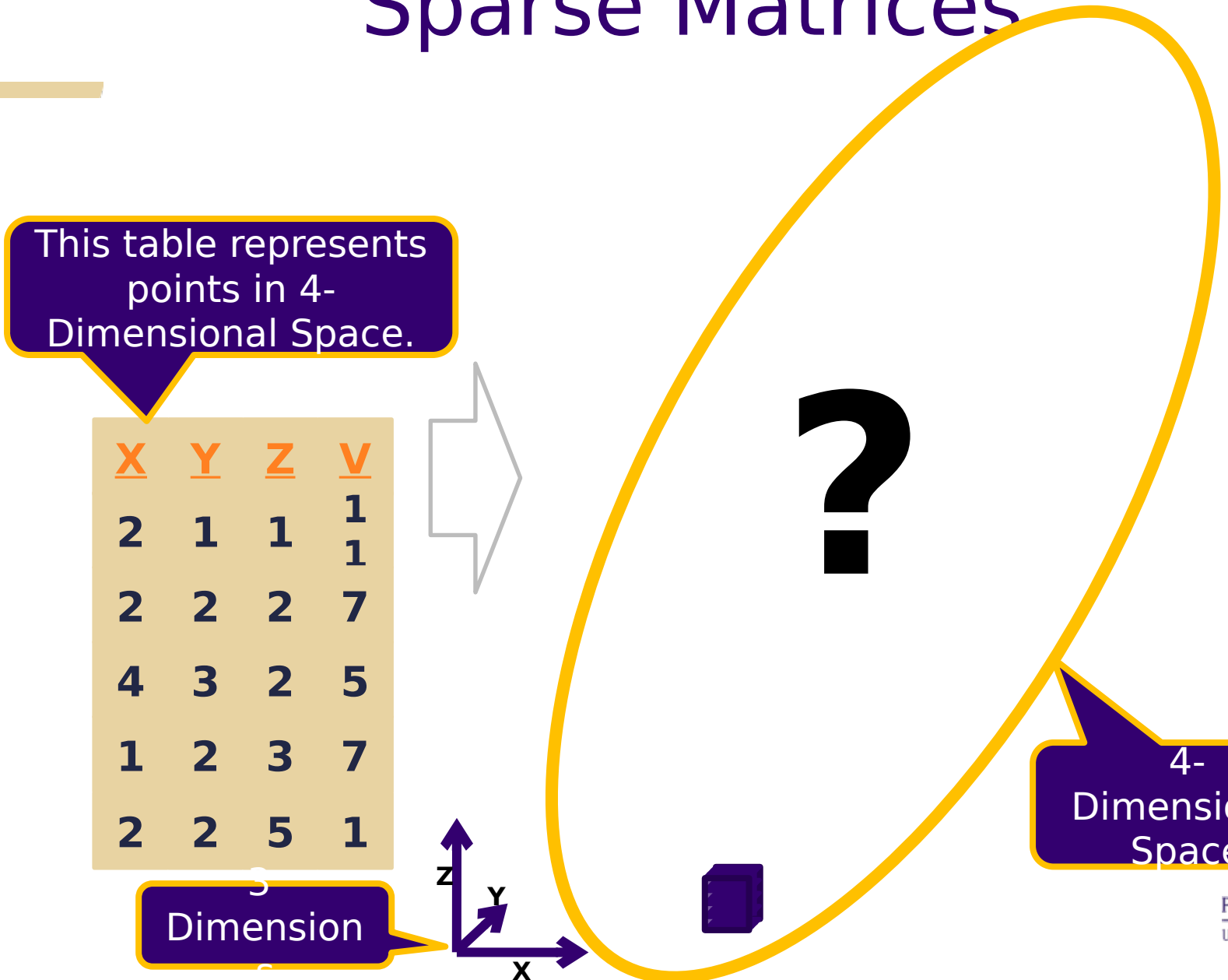


3-Dimensional Space

Sparse Matrices

This table represents points in 4-Dimensional Space.

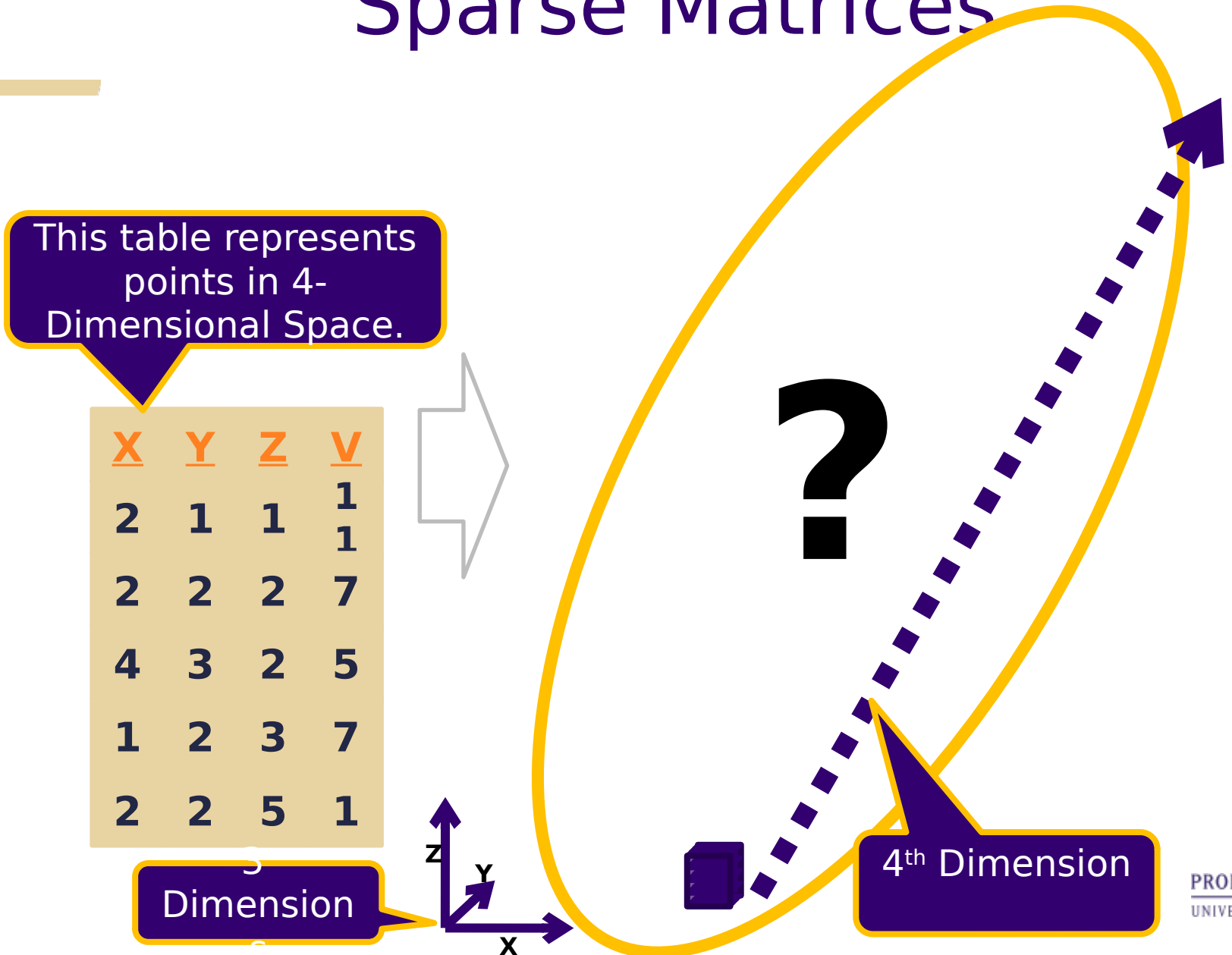
X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



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

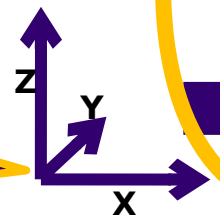


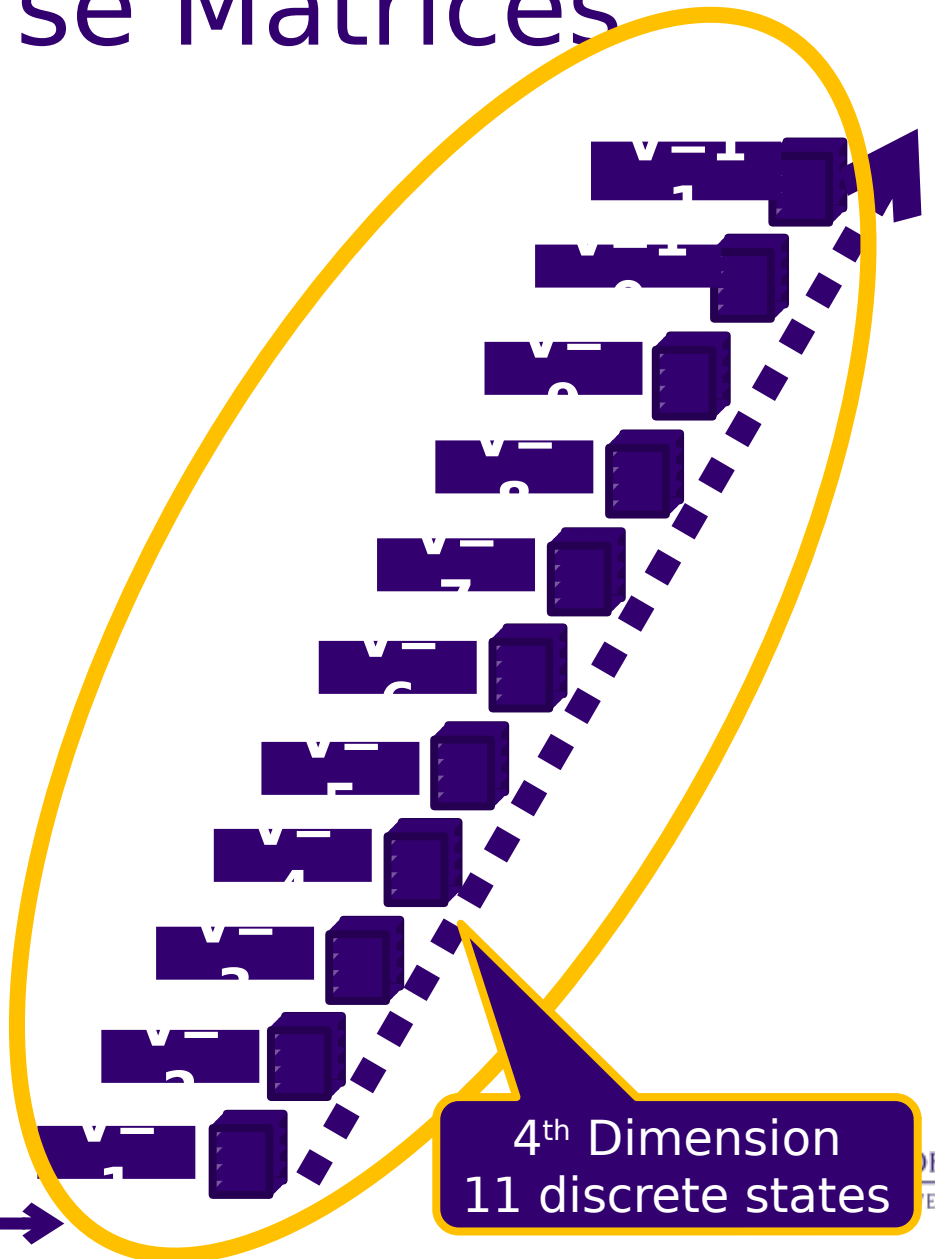
Sparse Matrices

This table represents points in 4-Dimensional Space.

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

3 Dimension



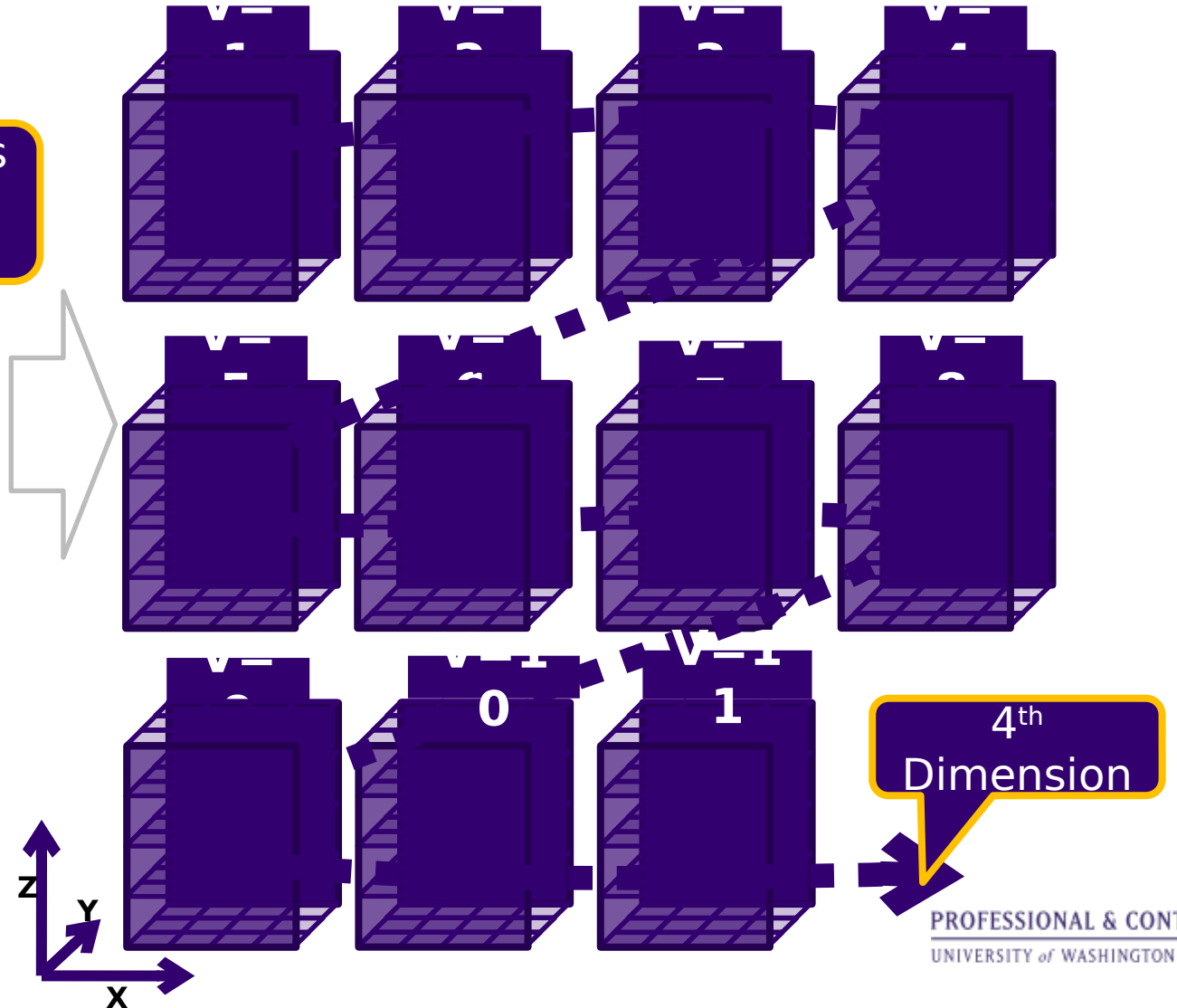


4th Dimension
11 discrete states

Sparse Matrices

This table represents points in 4-Dimensional Space.

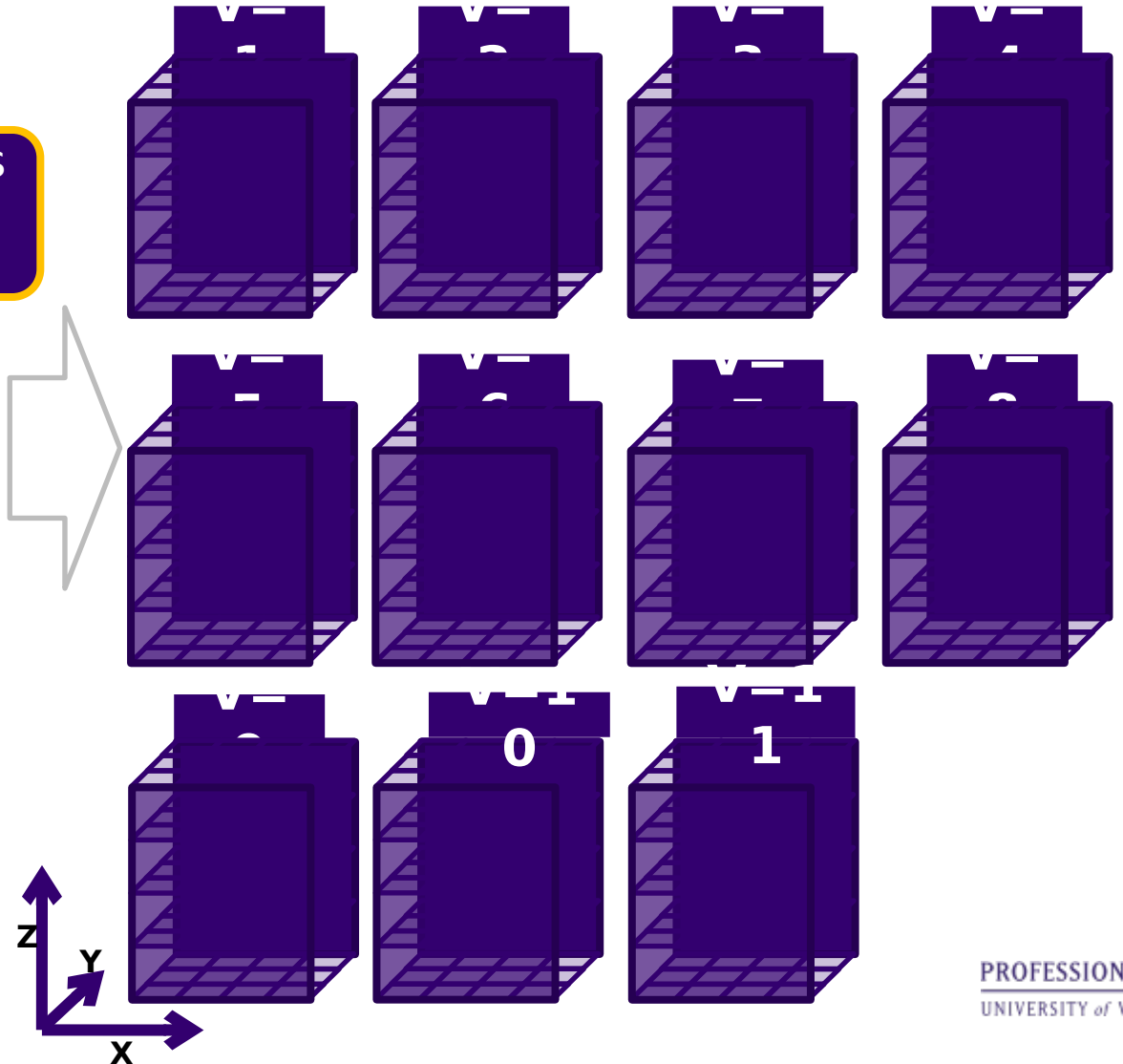
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices

This table represents points in 4-Dimensional Space.

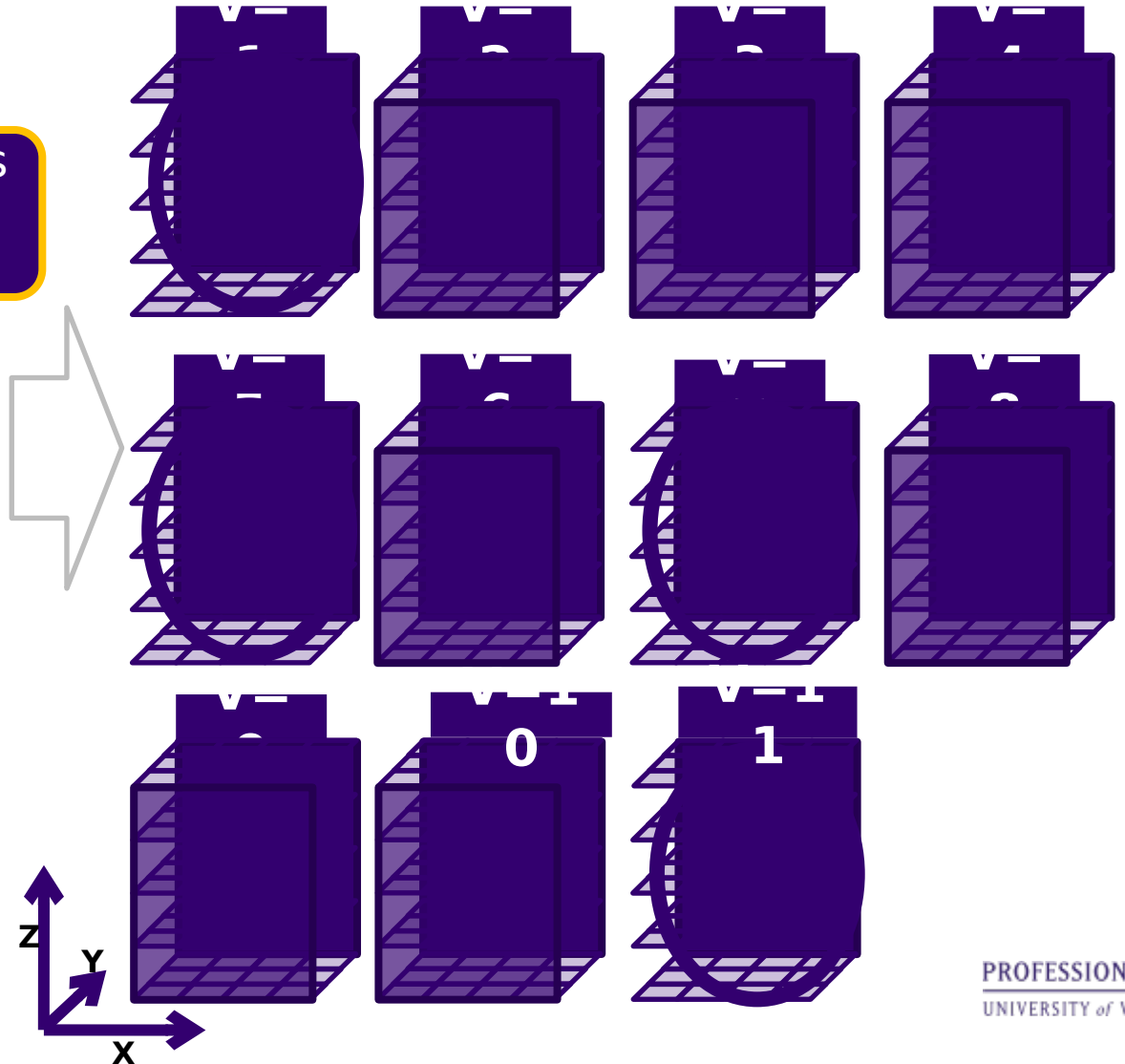
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



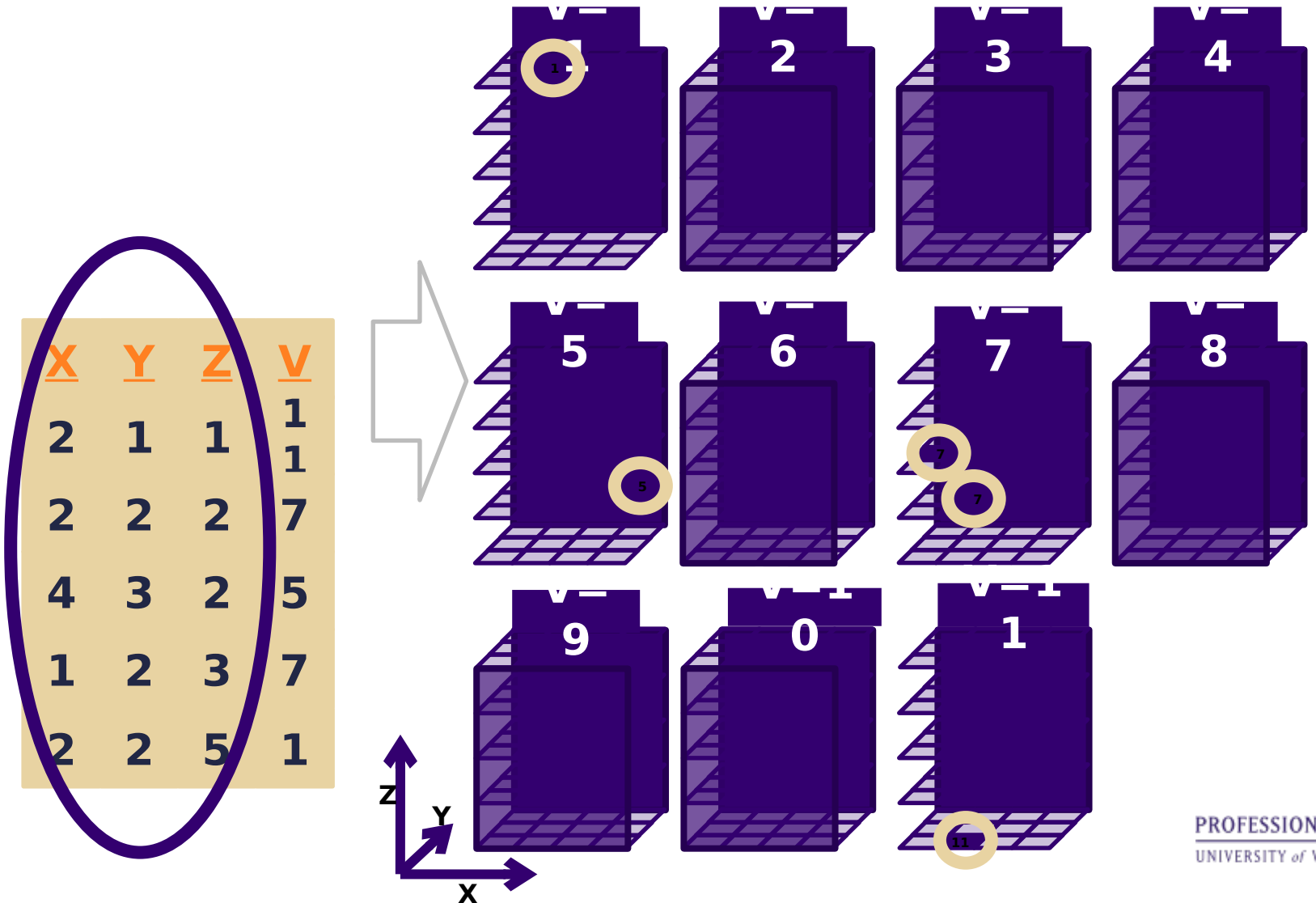
Sparse Matrices

This table represents points in 4-Dimensional Space.

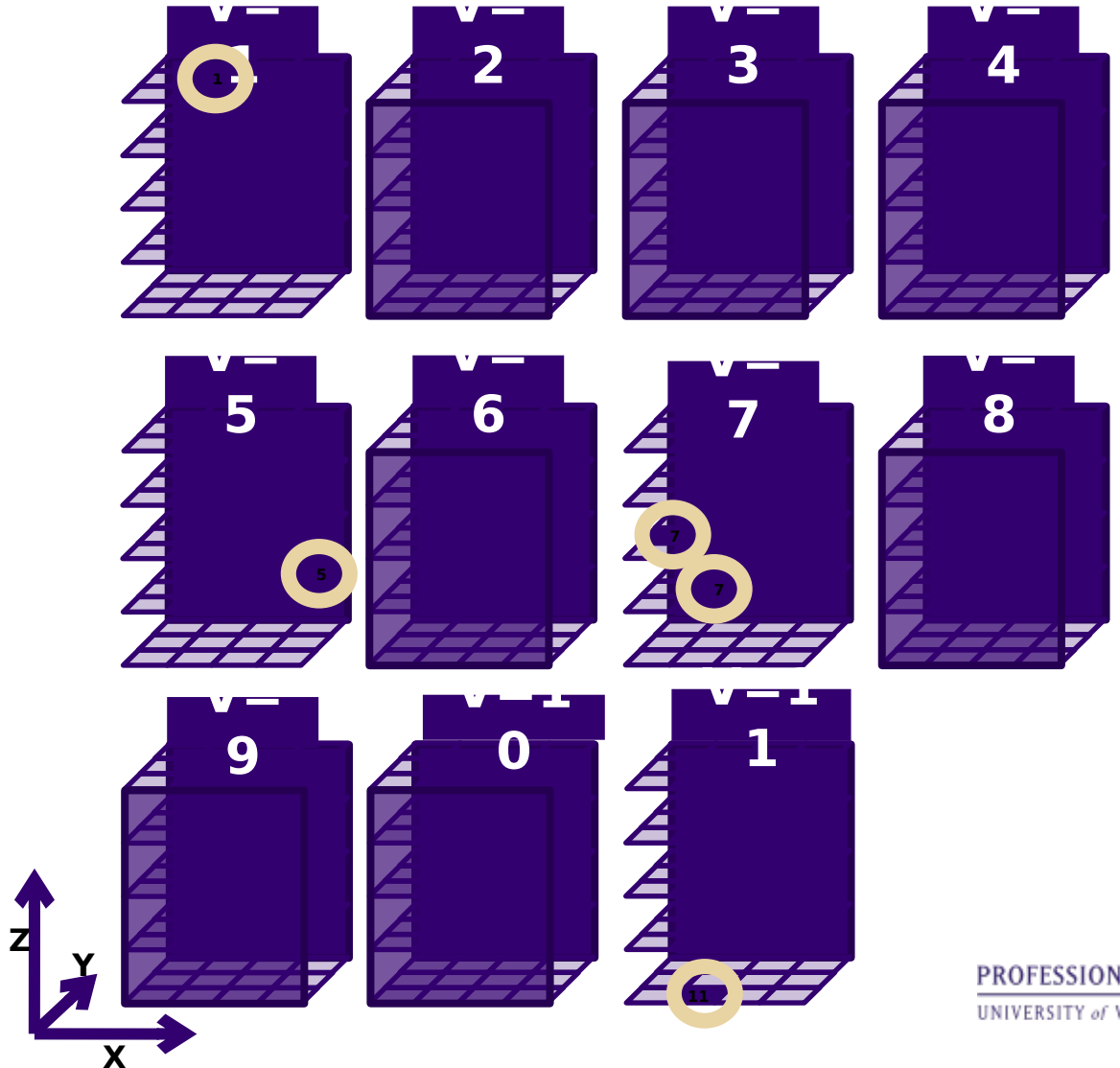
X	Y	Z	V
2	1	1	1
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1



Sparse Matrices



Sparse Matrices



Multi-dimensional Sparse Matrix

- Machine learning algorithms use tables as their primary structure
- A row in a table = a point in multi-D space
- A table = a multi-D sparse matrix
- Use these matrices to manipulate (transform) the data

Data as Multi-Dimensional Sparse Matrices