Sparse 2D Matrix Format

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

W

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!)

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 index.
- All matrix elements have the same type. (Not so in a table!)

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

Here is Python code to generate a 2D matrix

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 index.
- All matrix elements have the same type. (Not so in a table!)

```
>>> 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,  7]
])
```

2D Matrix Format

- 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!)

Most matrix elements have non-null values

Dense 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1	7	4	8	8	2	2	5
R2	6	8	1	3	5	4	1
R3	4	9	7	9	1	8	4
R4	9	3	3	7	3	2	6
R5	3	1	9	2	2	5	7

& CONTINUING EDUCATION
HINGTON

Sparse 2D Matrix Format

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

Most matrix elements have null values

Most matrix elements have non-null values

Sparse 2D Matrix

|	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

Dense 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1	7	4	8	8	2	2	5
R2	6	8	1	3	5	4	1
R3	4	9	7	9	1	8	4
R4	9	3	3	7	3	2	6
R5	3	1	9	2	2	5	7

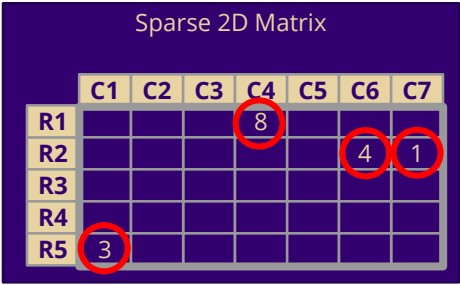NAL & CONTINUING EDUCATION
WASHINGTON

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

Most matrix elements have null values

Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						
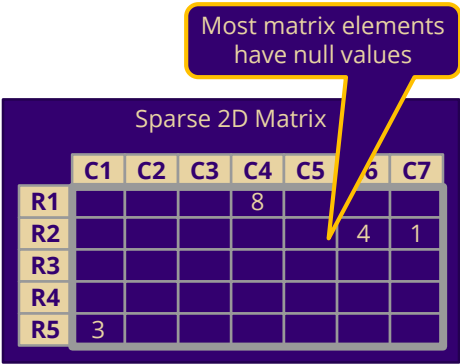
PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

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.
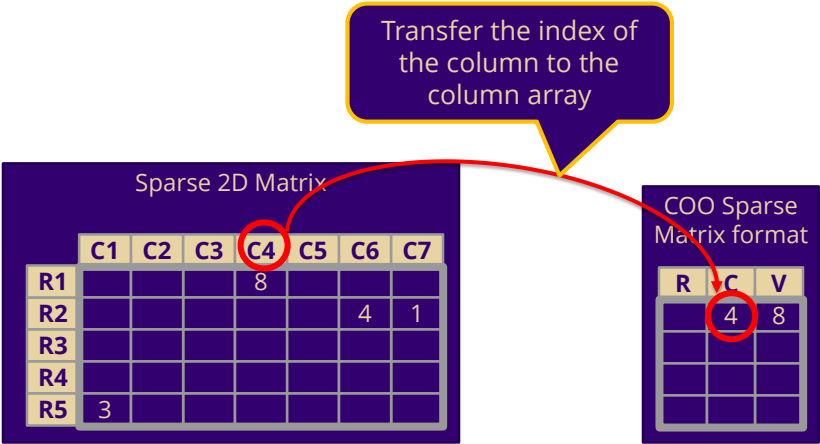
Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

Sparse 2D Matrix Format

Transfer the index of the column to the column array

Sparse 2D Matrix

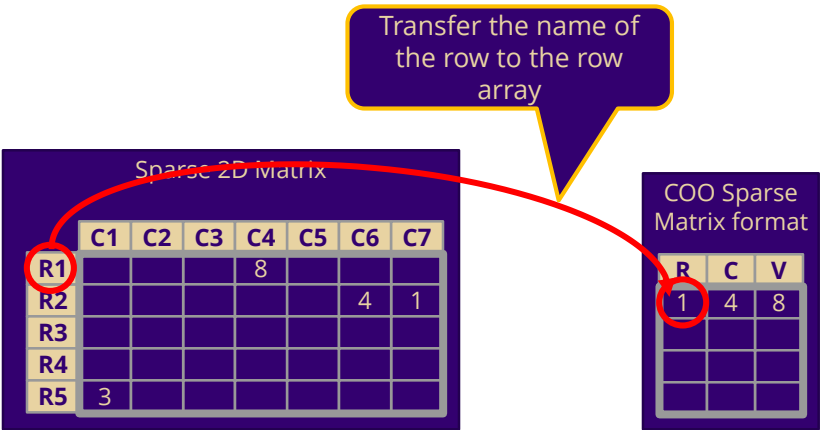|	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

COO Sparse Matrix format

R	C	V
	4	8

Sparse 2D Matrix Format

Transfer the name of the row to the row array

Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

COO Sparse Matrix format

R	C	V
1	4	8

Sparse 2D Matrix Format

Sparse 2D Matrix Format. Does not contain null values

Sparse 2D Matrix
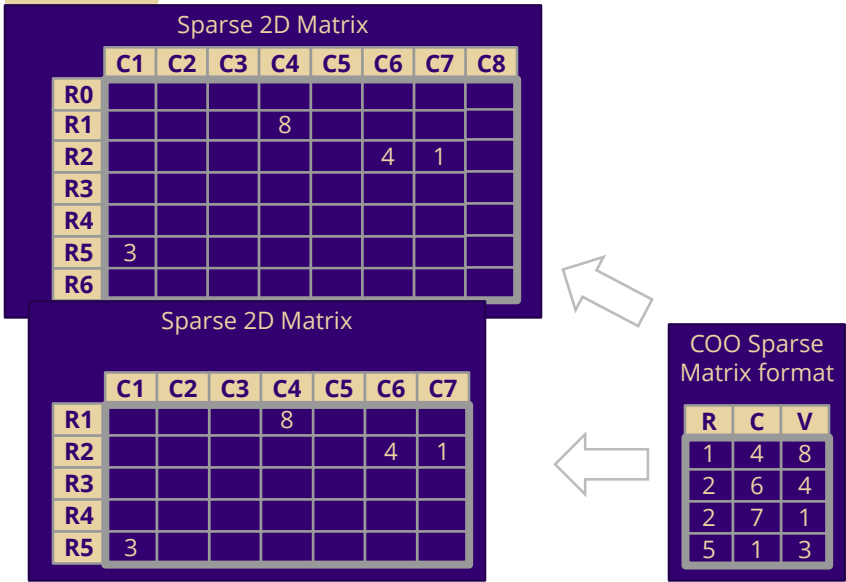
	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

COO Sparse Matrix format

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse 2D Matrix Format

Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

COO Sparse Matrix format

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse 2D Matrix Format

Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7	C8
R0								
R1				8				
R2						4	1	
R3								
R4								
R5	3							
R6								

Sparse 2D Matrix

	C1	C2	C3	C4	C5	C6	C7
R1				8			
R2						4	1
R3							
R4							
R5	3						

COO Sparse Matrix format

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

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

W

Sparse 2D Matrix Format

Use COO Sparse Matrix Format

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

W

Data as Multi–Dimensional Sparse Matrices

How algorithms view tables

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

W

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
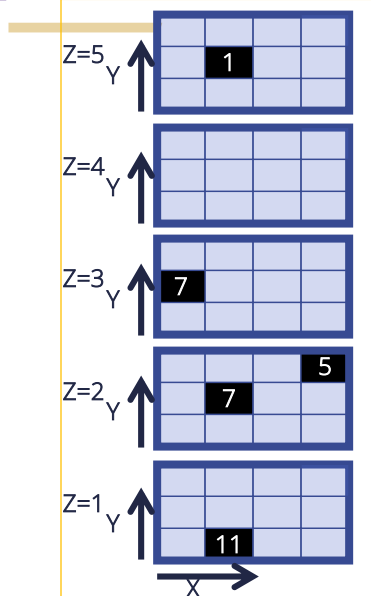
PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse Matrices

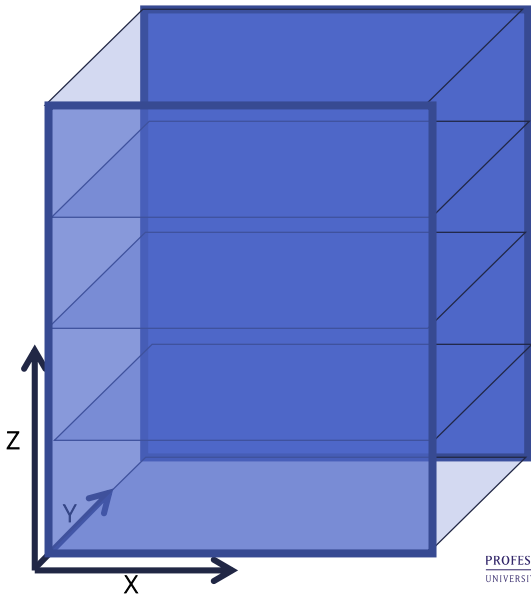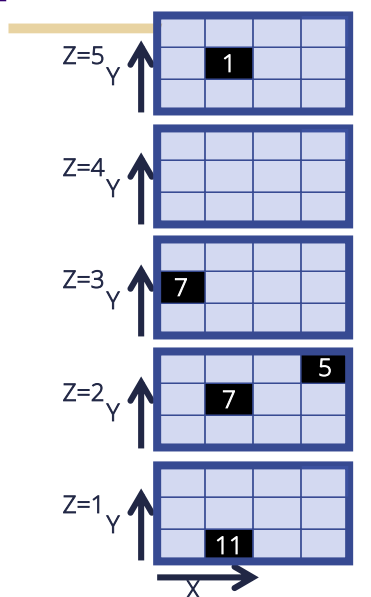2-Dimensional Matrix

5

7

Y

X
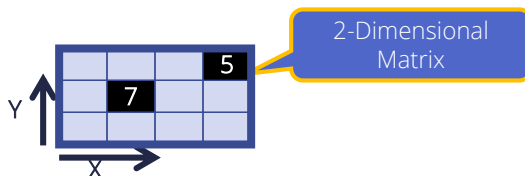
PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse Matrices

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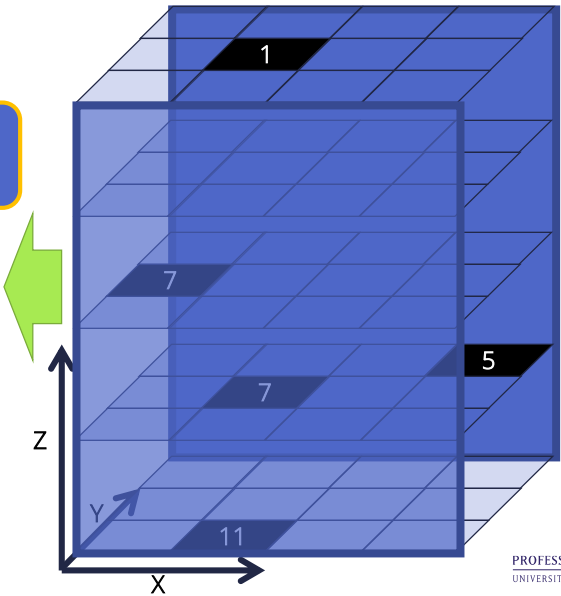
Sparse Matrices

Sparse Matrices

Z=5
Z=4
Z=3
Z=2
Z=1

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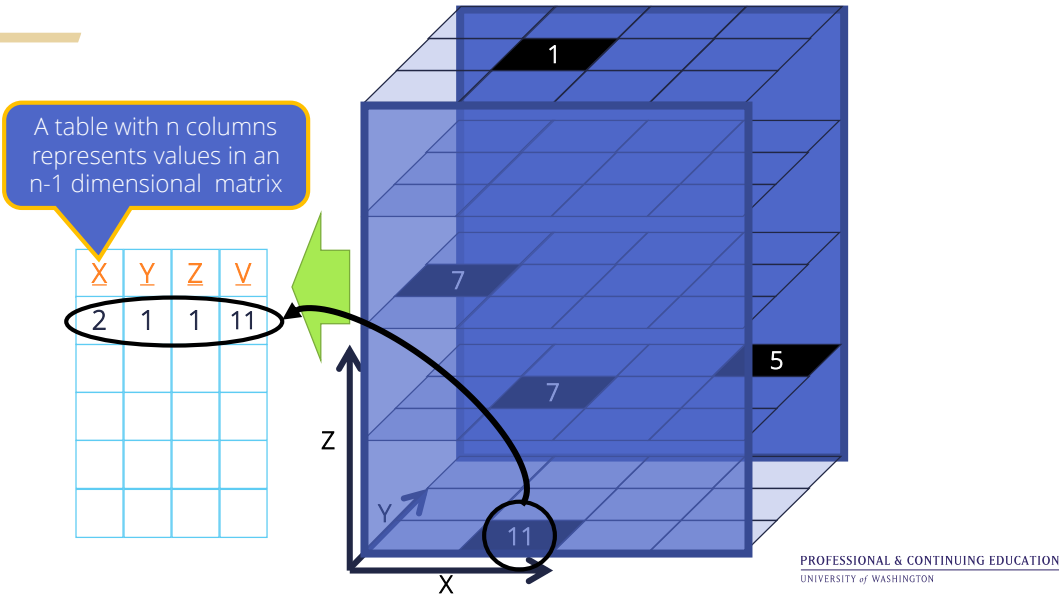
A table with n columns represents values in an n-1 dimensional matrix

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

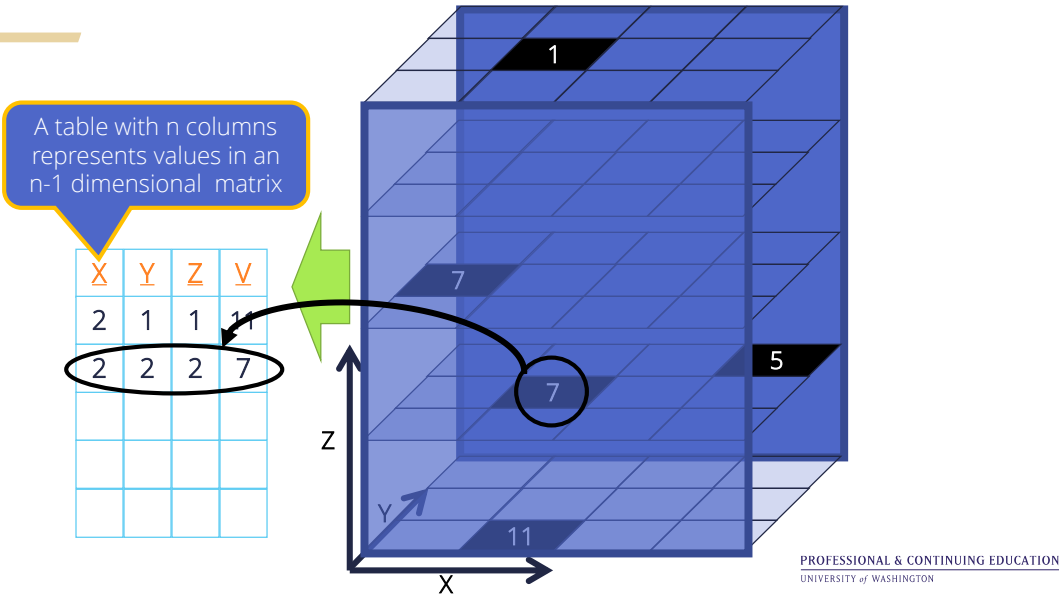| X	Y	Z	V
2	1	1	11
2	2	2	7

Sparse Matrices

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

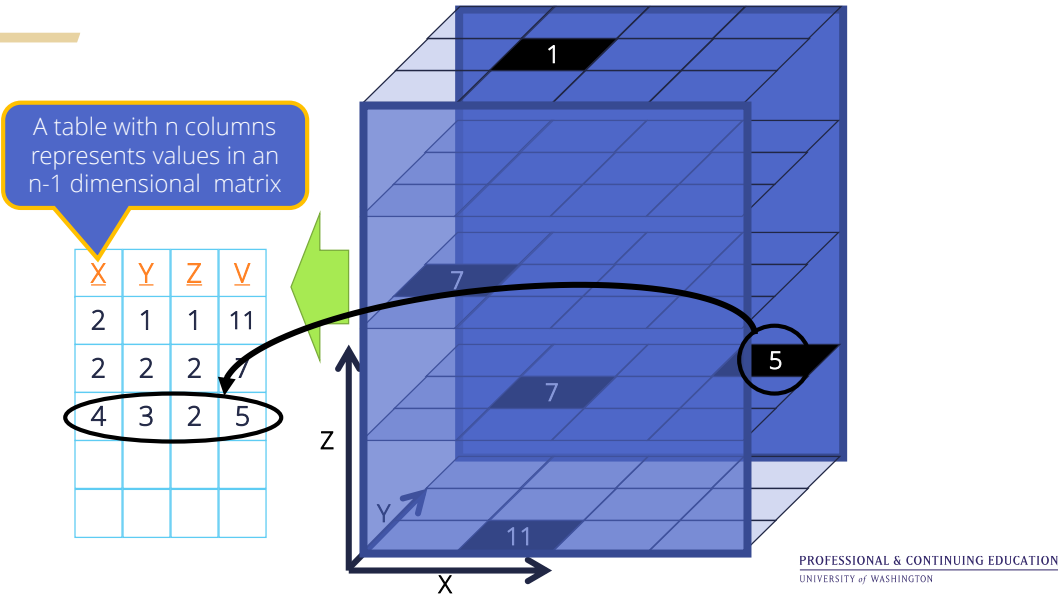| X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse Matrices

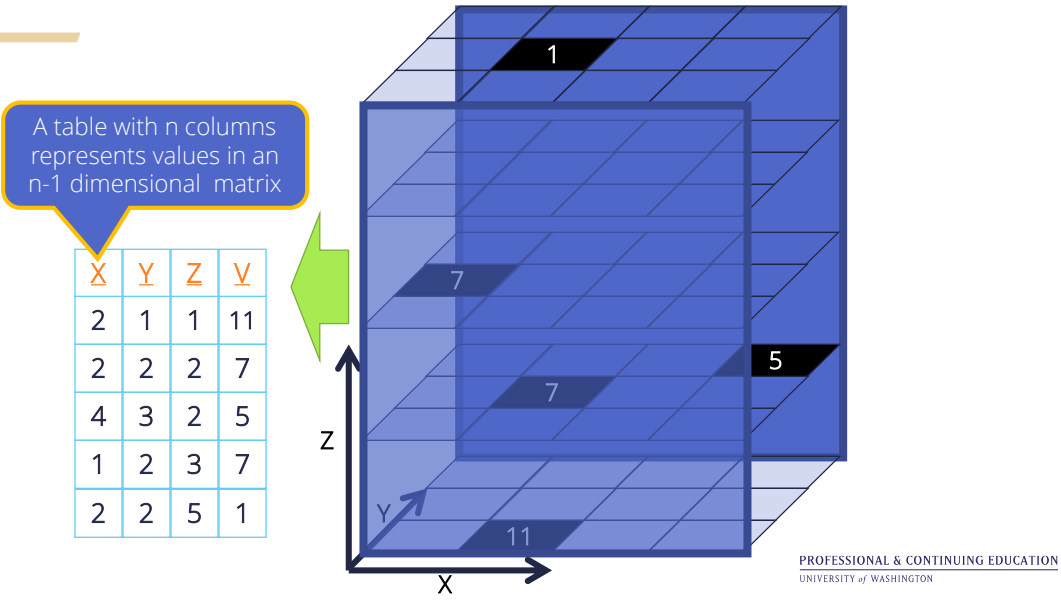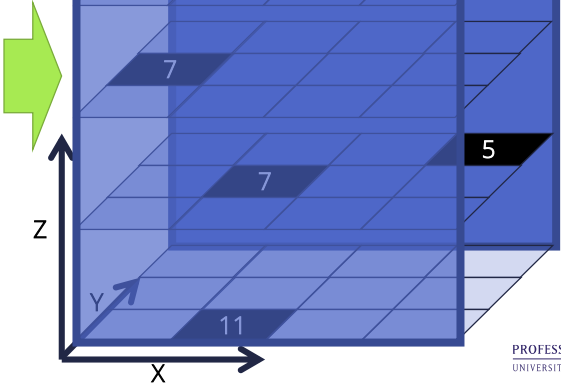A table with n columns represents values in an n-1 dimensional matrix

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

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

Think of V as just another dimension

Sparse Matrices

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

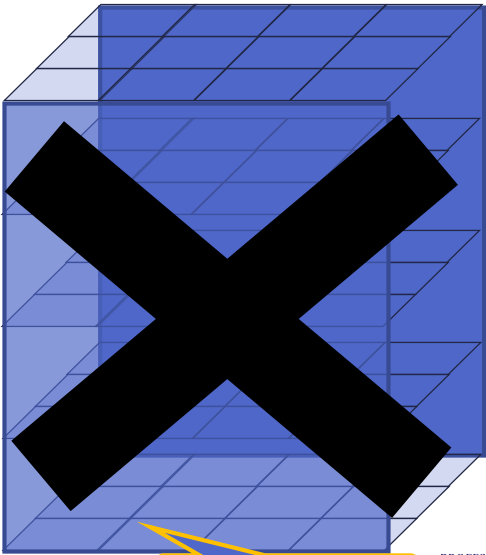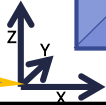| X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

Think of V as just another dimension

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 Dimensions

3-Dimensional Space

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

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 Dimensions

3-Dimensional Space

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

Sparse Matrices

This table represents points in 4-Dimensional Space.

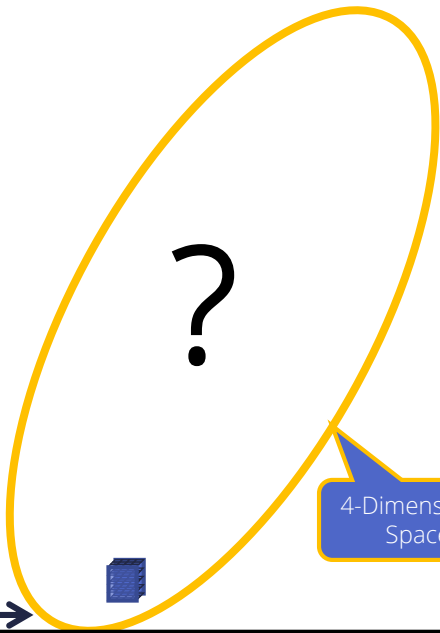| X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

3 Dimensions

?

4-Dimensional Space

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

Sparse Matrices

This table represents points in 4-Dimensional Space.

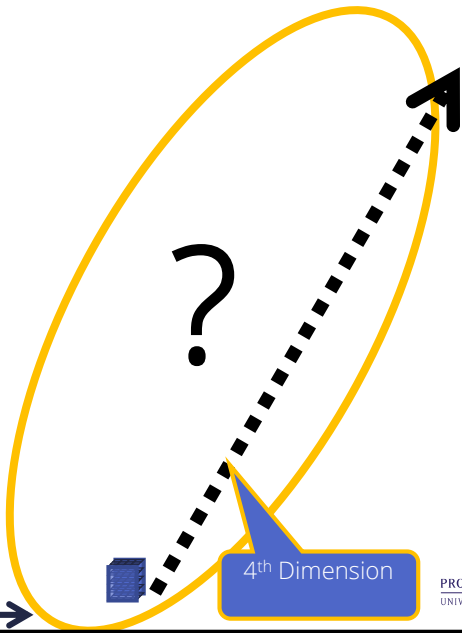| X	Y	Z	V
2	1	1	11
2	2	2	7
4	3	2	5
1	2	3	7
2	2	5	1

3 Dimensions

?

4th Dimension

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

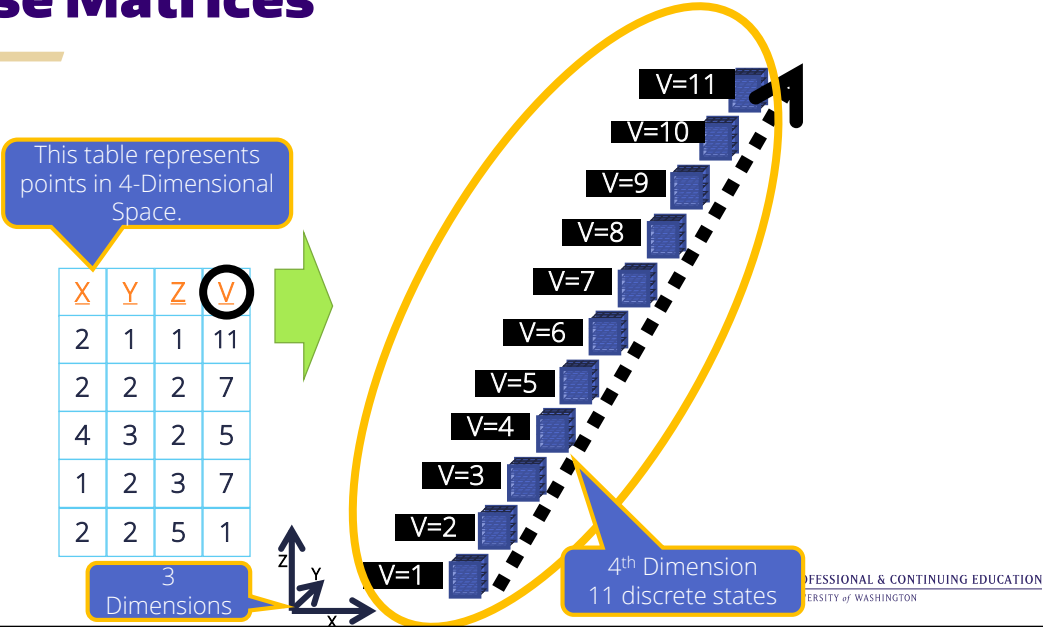# Sparse Matrices

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 Dimensions

V=11
V=10
V=9
V=8
V=7
V=6
V=5
V=4
V=3
V=2
V=1

4th Dimension
11 discrete states

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

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

V=1 V=2 V=3 V=4
V=5 V=6 V=7 V=8
V=9 V=10 V=11

4th Dimension

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

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

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

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
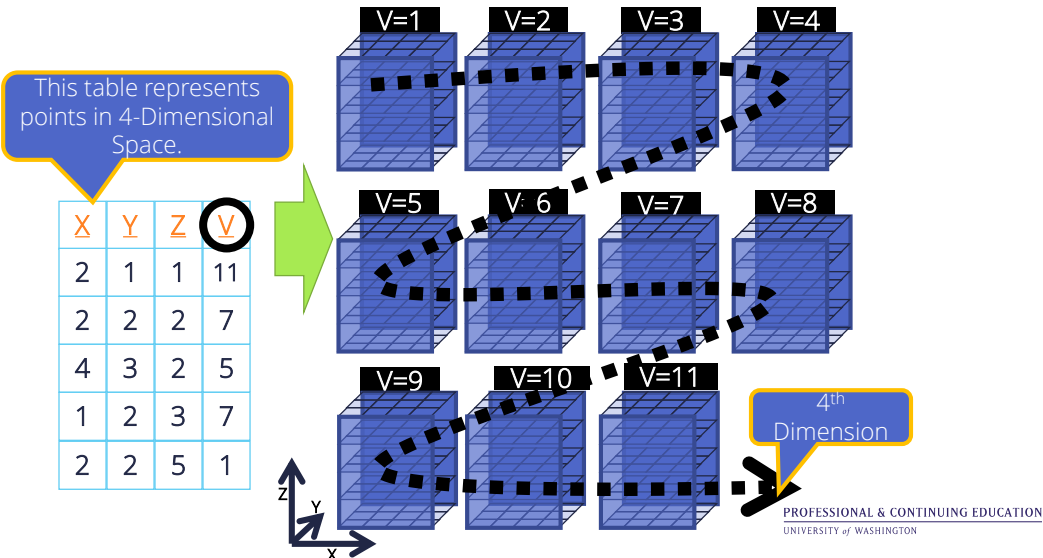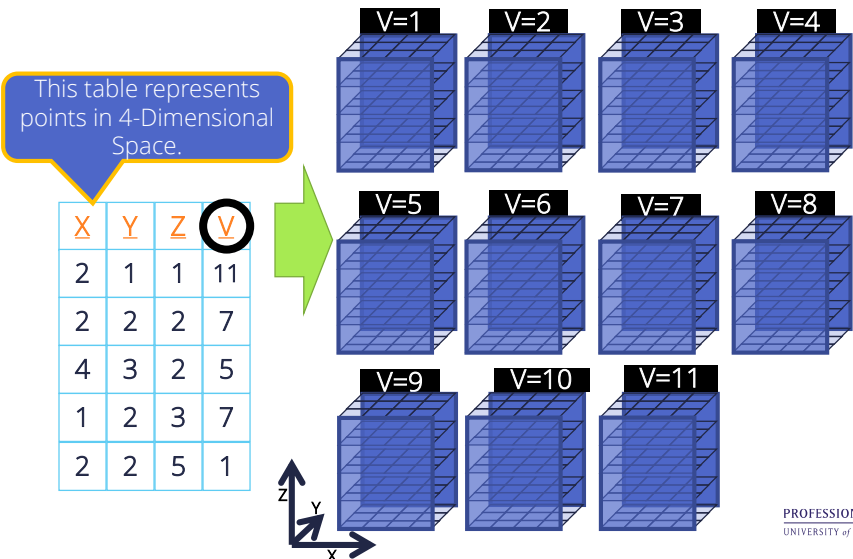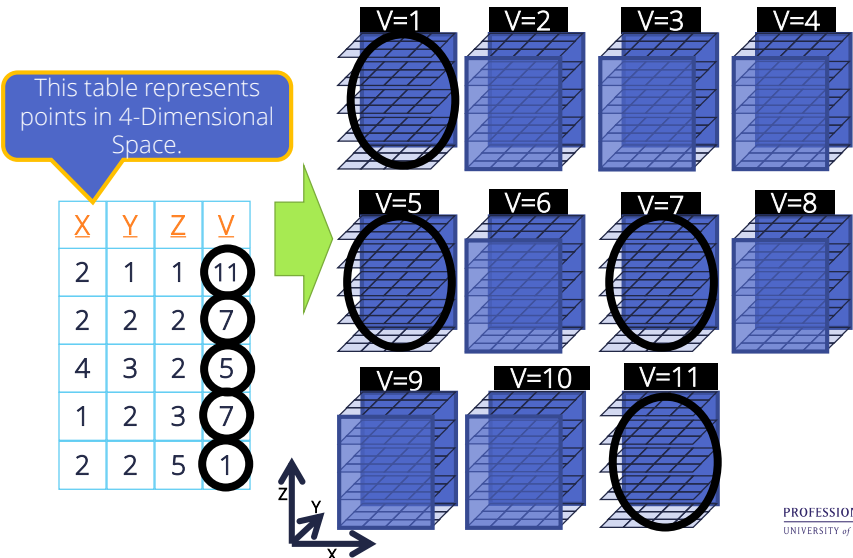
Sparse Matrices

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

How Algorithms view Tables

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON