# Data handling workshop:

## Get the basic programming skills for handling and analyzing data for scientific research and general purpose

Time: 9:30 am – 12:30 pm, July 31, 2021 (Saturday) Mode: via ZOOM

By Dr Guang Ouyang, Assistant Professor, Academic Unit of Human Communication, Development, and Information Sciences, Faculty of Education, HKU

### Perception

### Information

Bird
Blue
Animal
Fly
Beauty
Bipedalism
…

### Data

00111010101011010100
11101010010101010010
10010101011101110001
01001110101010110101
00111010100101010100
10100101010111011100
0101001

# Outline

1. Data types, organization and structures
2. Data operation (in Matlab)
3. Examples:
   - ❑Basic data/file operation
   - ❑Basic matrix operation
   - ❑Simple image processing
   - ❑Simple EEG & eyetracking data processing
   - ❑Simple AI application

# Data types

## 1. Numerical

8

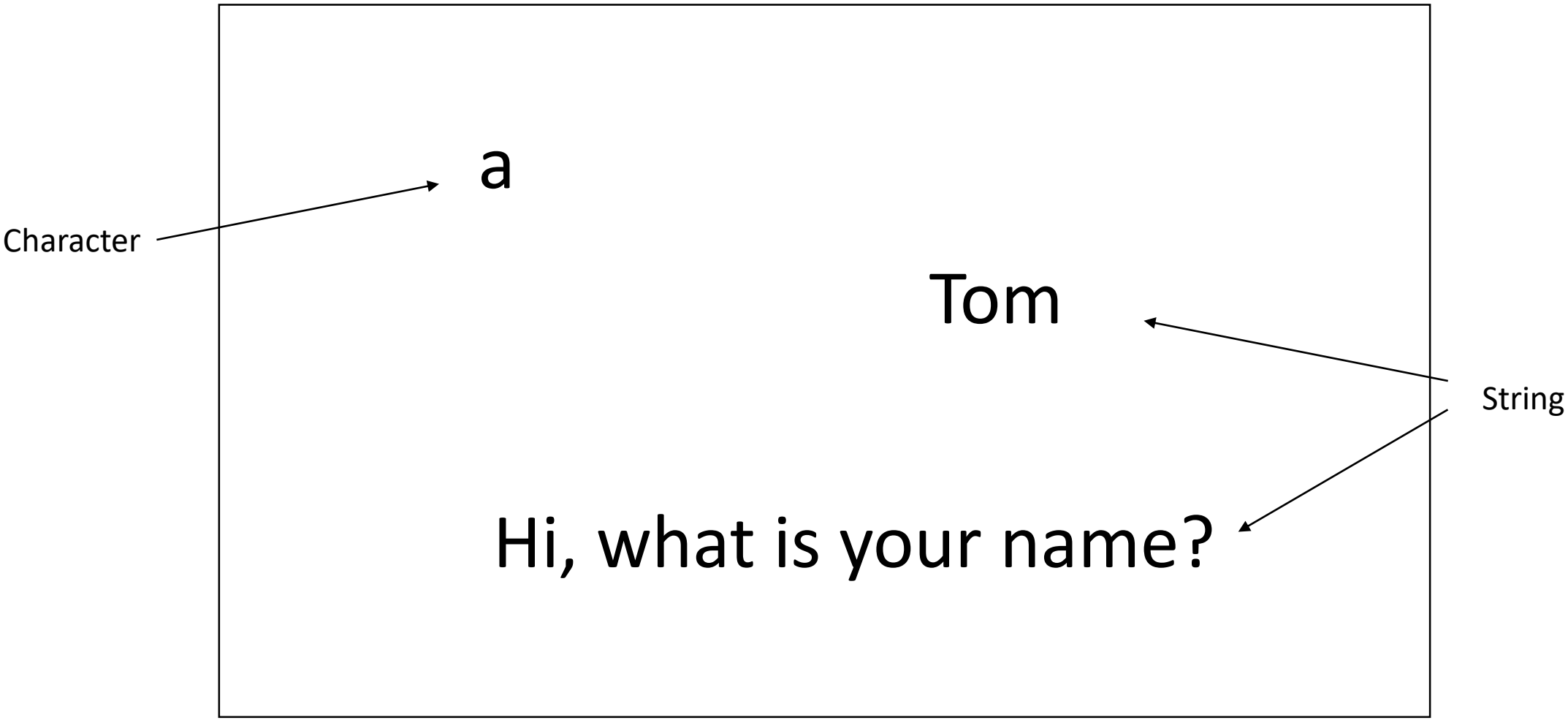
2.5

3.1415926

999,999,999

# Data types

## 2. Character/String



# 3. Boolean

| | |
|---|---|
| Yes | No |
| True | False |
| 1 | 0 |

# Data types

1. Numerical
2. Character/String
3. Boolean
4. …

Can be operated arithmetically

Carries semantic/categorical information

Carries logical information

In a computer, they are all represented as 01001011010101…

# Data organization

8

a

2.5

Tom

No

3.1415926

Variable = Value

Yes

False

Hi, What is your name?

True

999,999,999

# Data organization

Single element:

| 1 |
|---|

Vector:

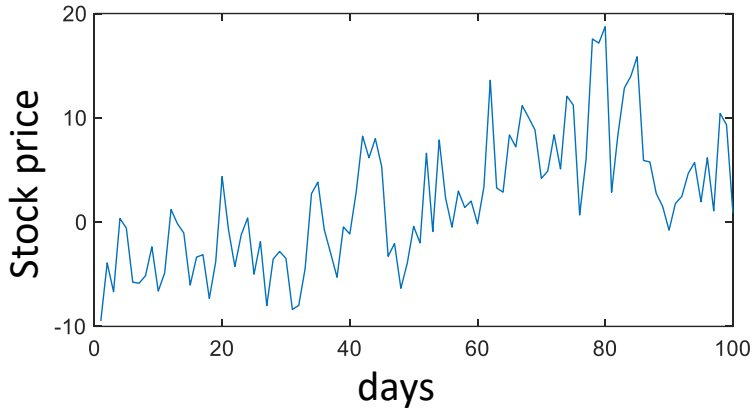
| 1 | 3 | 6 | 3 | 1 | 0 | 32 | 9 |
|---|---|---|---|---|---|---|---|

Matrix:

| 1 | 3 | 6 | 3 | 1 | 0 | 32 | 9 |
|---|---|---|---|---|---|---|---|
| 45 | 2 | 8 | 4 | 665 | 77 | 7 | 0 |
| 43 | 4 | 9 | 7 | 4 | 9 | 90 | 2 |
| 3 | 1 | 22 | 5 | 8 | 55 | 6 | 3 |

1-d array, 2-d array, 3-d array, n-d array, …

# Data organization




1d

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| -9.5 | -3.9 | -6.7 | 0.3 | -0.6 | -5.8 | -5.9 | -5.1 | -2.4 | -6.6 | -4.9 | 1.2 | -0.2 | -1.0 | -6.0 | -3.4 | -3.1 | -7.3 | -3.8 | 4.4 | -0.7 | -4.3 | -1.2 | 0.4 | -5.0 | -1.9 | -8.0 | -3.5 | ..... |

2d

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 95 | 95 | 96 | 97 | 97 | 98 | 99 | 99 | 101 | 101 | 102 | 103 | 103 |
| 95 | 95 | 96 | 96 | 97 | 98 | 99 | 99 | 101 | 101 | 102 | 102 | 103 |
| 95 | 95 | 95 | 96 | 97 | 98 | 98 | 99 | 101 | 101 | 101 | 102 | 103 |
| 94 | 94 | 95 | 96 | 97 | 97 | 98 | 98 | 100 | 100 | 101 | 102 | 103 |
| 94 | 94 | 95 | 95 | 96 | 97 | 98 | 98 | 100 | 100 | 101 | 101 | 102 |
| 93 | 94 | 94 | 95 | 96 | 97 | 97 | 97 | 99 | 100 | 100 | 101 | 102 |
| 93 | 93 | 94 | 95 | 96 | 96 | 97 | 97 | 99 | 99 | 100 | 101 | 102 |
| 93 | 93 | 94 | 95 | 95 | 96 | 97 | 97 | 99 | 99 | 100 | 101 | 101 |
| 94 | 94 | 94 | 95 | 95 | 96 | 96 | 96 | 97 | 97 | 98 | 99 | 99 |
| 94 | 94 | 94 | 95 | 95 | 95 | 96 | 96 | 97 | 97 | 98 | 98 | 99 |
| 94 | 94 | 94 | 94 | 95 | 95 | 95 | 96 | 97 | 97 | 97 | 98 | 99 |
| 93 | 93 | 94 | 94 | 94 | 95 | 95 | 95 | 96 | 96 | 97 | 98 | 99 |
| 93 | 93 | 93 | 94 | 94 | 94 | 95 | 95 | 96 | 96 | 97 | 97 | 98 |
| 92 | 93 | 93 | 93 | 94 | 94 | 94 | 94 | 95 | 96 | 96 | 97 | 98 |
| 92 | 92 | 93 | 93 | 93 | 94 | 94 | 94 | 95 | 95 | 96 | 97 | 98 |
| 92 | 92 | 92 | 93 | 93 | 94 | 94 | 94 | 95 | 95 | 96 | 97 | 97 |
| 90 | 90 | 91 | 92 | 92 | 93 | 94 | 94 | 96 | 96 | 95 | 96 | 97 |
| 90 | 90 | 91 | 92 | 92 | 93 | 94 | 94 | 96 | 96 | 95 | 96 | 97 |

.....

⋮

# Data organization

- Variable with heterogeneous elements
- In Matlab, it is called ‘cell’ variable type.

```
a = {'a',1.2,'Jason',299,'How are you'};
```

a

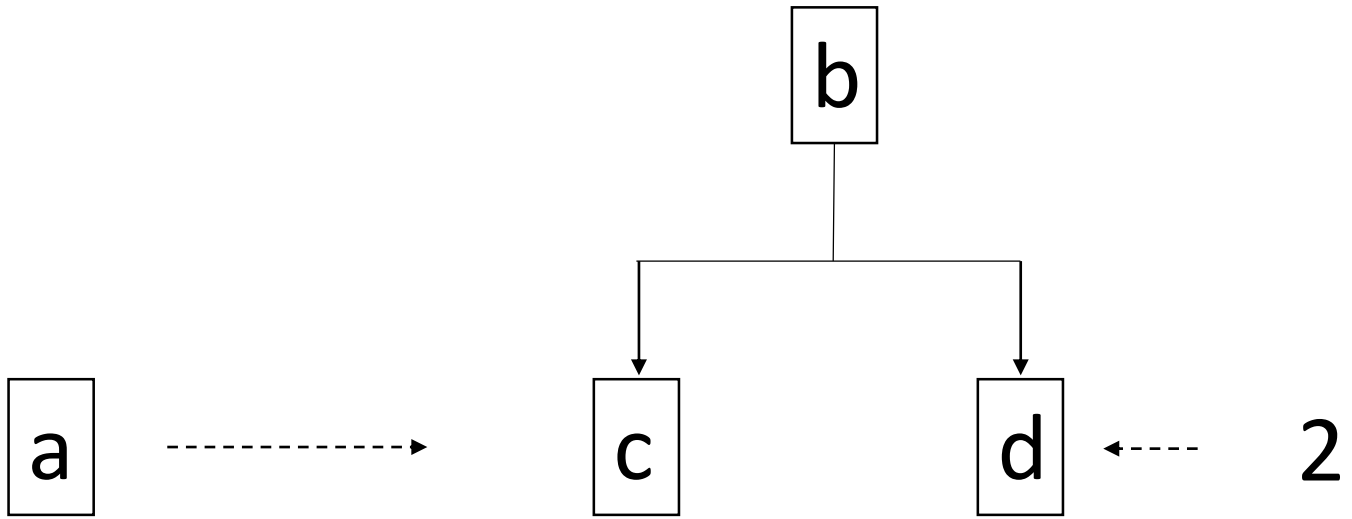
1x5 cell

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 'a' | 1.2000 | 'Jason' | 299 | 'How are you' | | |
| 2 | | | | | | | |
| 3 | | | | | | | |

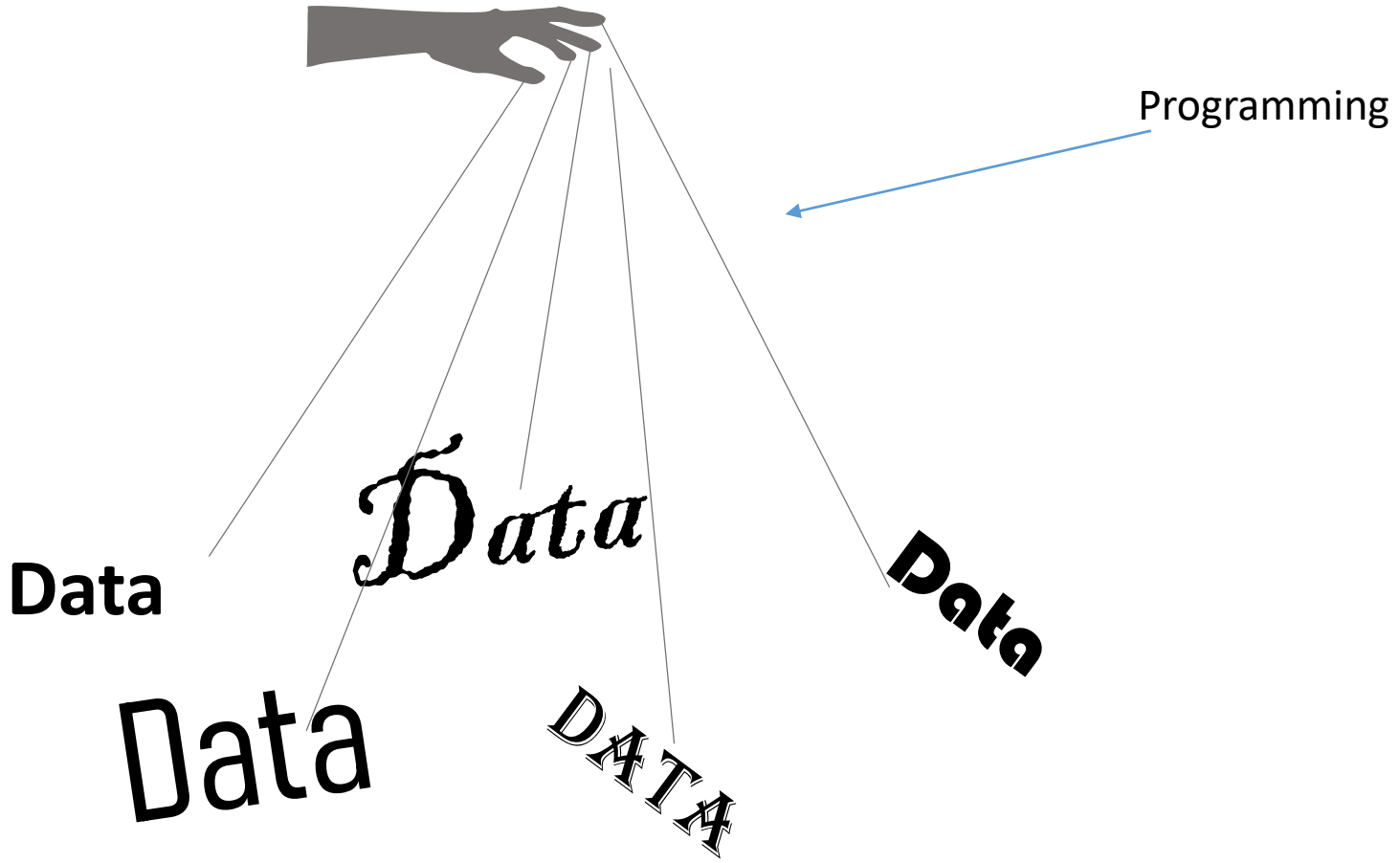
# Data organization

- Variable with hierarchical structures
- In Matlab, it is called ‘structure’ variable type.

```
a = {'a',1.2,'Jason',299,'How are you'};
b.c = a;
b.d = 2;
```



# Data operation



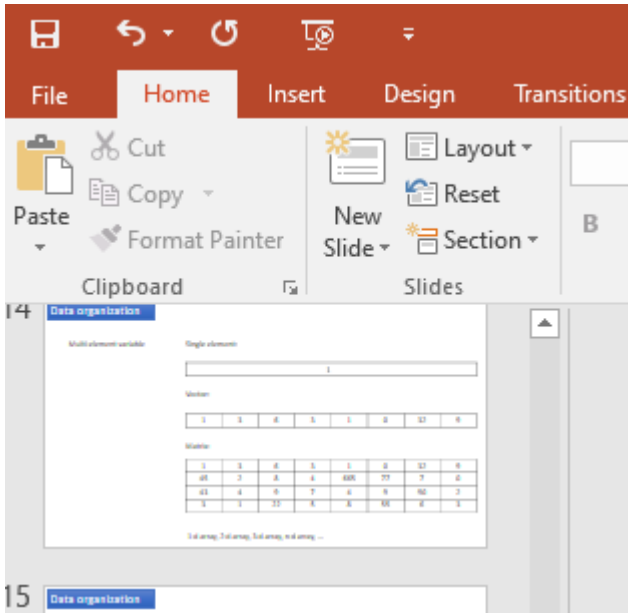
# Data operation

- Value assignment
- Value extraction
- Operations
  - Plotting; data importing/exporting; operating hardware, etc
  - Function
- Automatic operations
  - If statement;
  - For statement;

# Data operation

The essence of programming:

Using codes to tell the computer to do something




Step 1

```
clear;
eeglab nogui
```

Step 2

```
%% data loading
data_path = 'D:\EEG_workshop\data\';
%need to change to your actual data path

parti = 'sample_visual_oddball';
EEG = pop_loadset('filename',[parti,'.set'],'filepath',data_path);
```

```
%% pre-processing
```

Step 3

```
%bandpass filtering
EEG = pop_eegfiltnew(EEG, 'locutoff',1,'hicutoff',45,'plotfr
%detect bad channel and interpolate
std_temp= std(EEG.data,1,2);
ol = find(isoutlier(std_temp,'ThresholdFactor',4
if ~isempty(ol) EEG = pop_interp(EEG, ol, 'spher
%average referencing
EEG = pop_reref( EEG, []);
%run ica (the stop criterion usually needs to be lower)
EEG = pop_runica(EEG, 'icatype', 'runica', 'extended',1,'stop',0.0
%automatically detec artifacts and clean the data
[comps,info] = MARA(EEG);
EEG = pop_subcomp( EEG, comps, 0);
```

.
.
.

# Data operation

- **Value assignment**
- Value extraction
- Operations
  - Plotting; data importing/exporting; operating hardware, etc
  - Function
- Automatic operations
  - If statement;
  - For statement;

# Data operation

8

a

Tom

3.1415926

No

2.5

Variable = Value

Yes

False

Hi, What is your name?

True

999,999,999

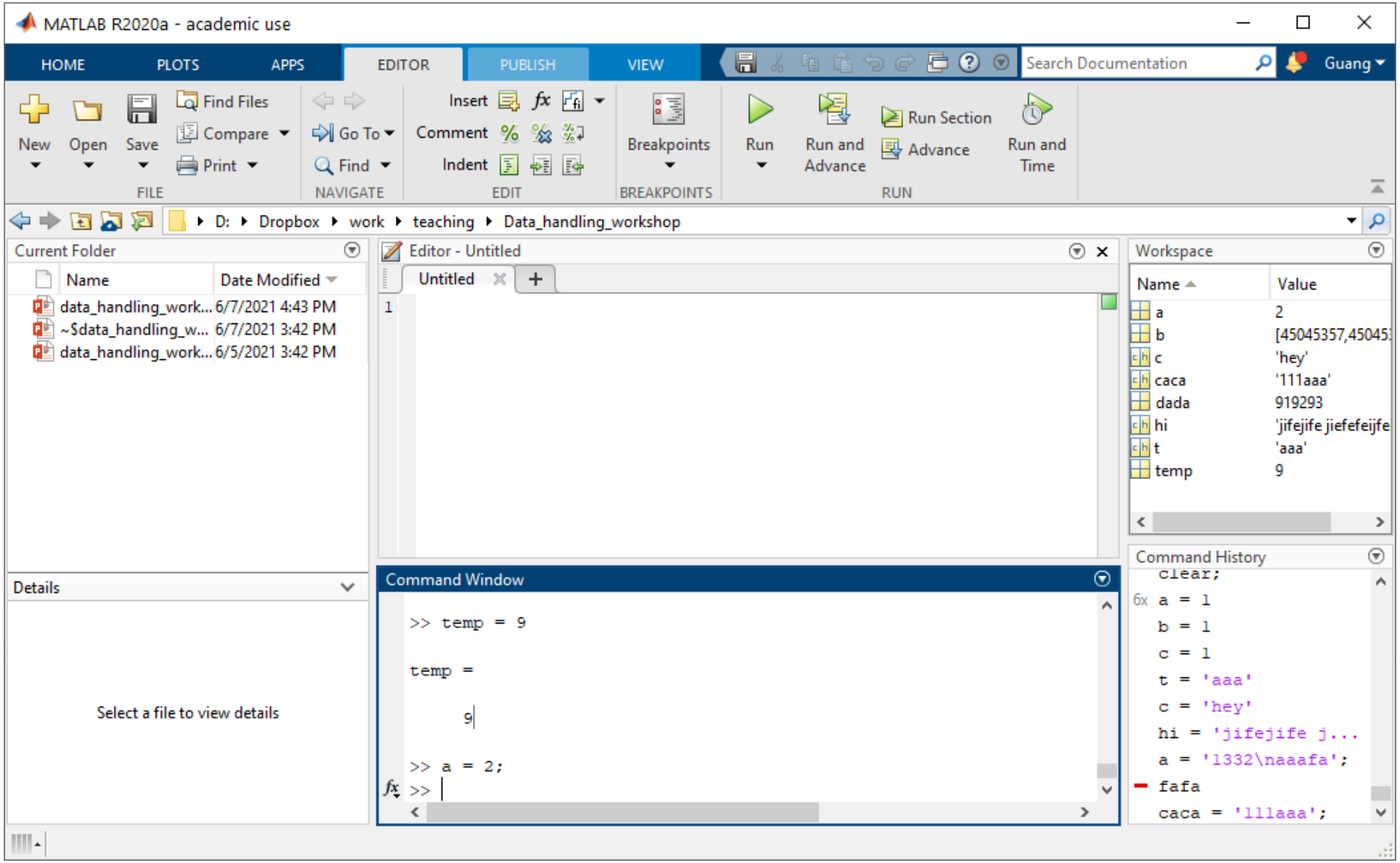
# Data operation

Variable = Value

A host/an object that we can assign a value to
We can also retrieve its value, or change its value.

# Data operation

## Try it



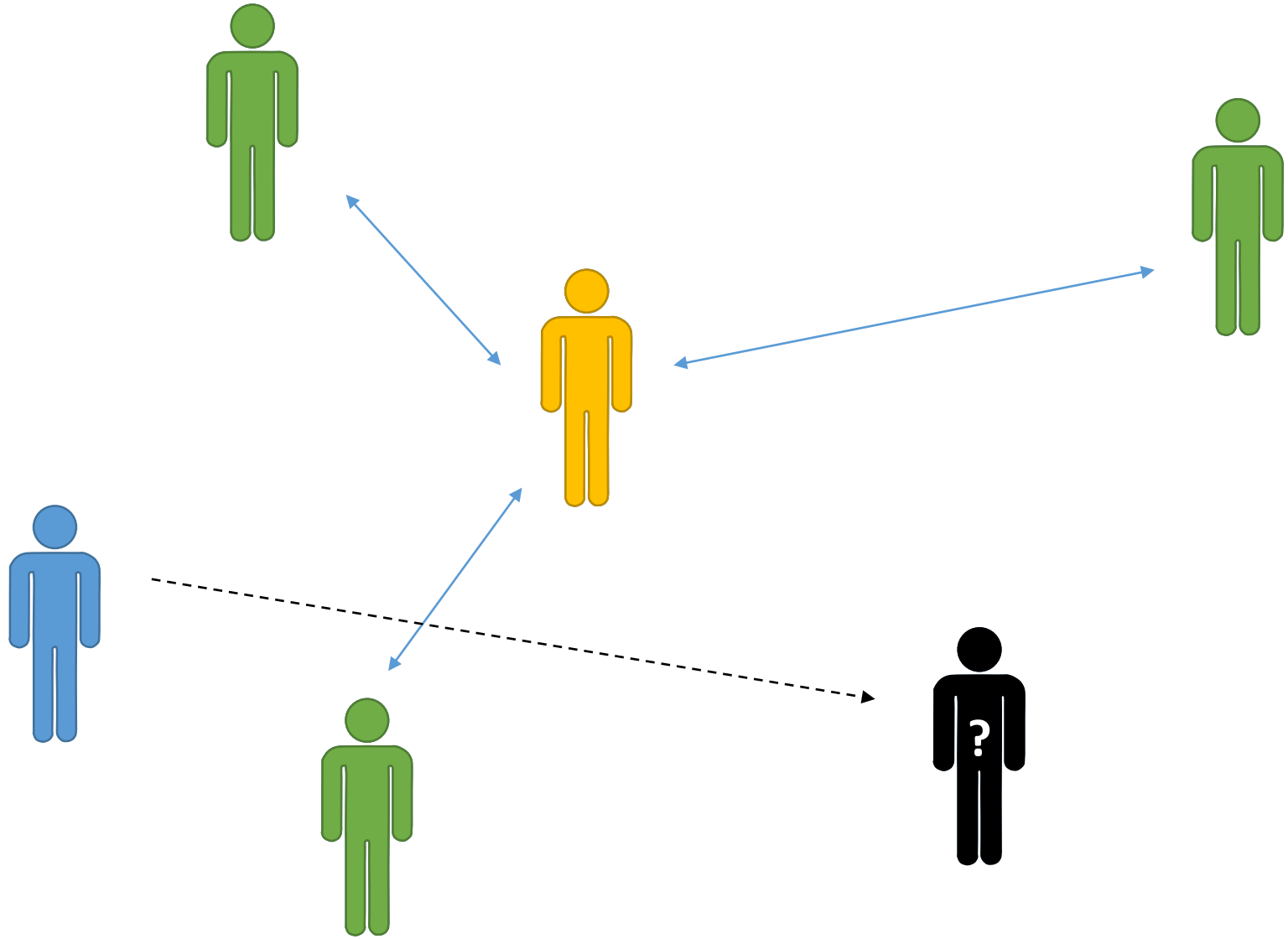
# Data operation

Variable = Value

A host/an object that we can assign a value to
We can also retrieve its value, or change its value.

# Data operation

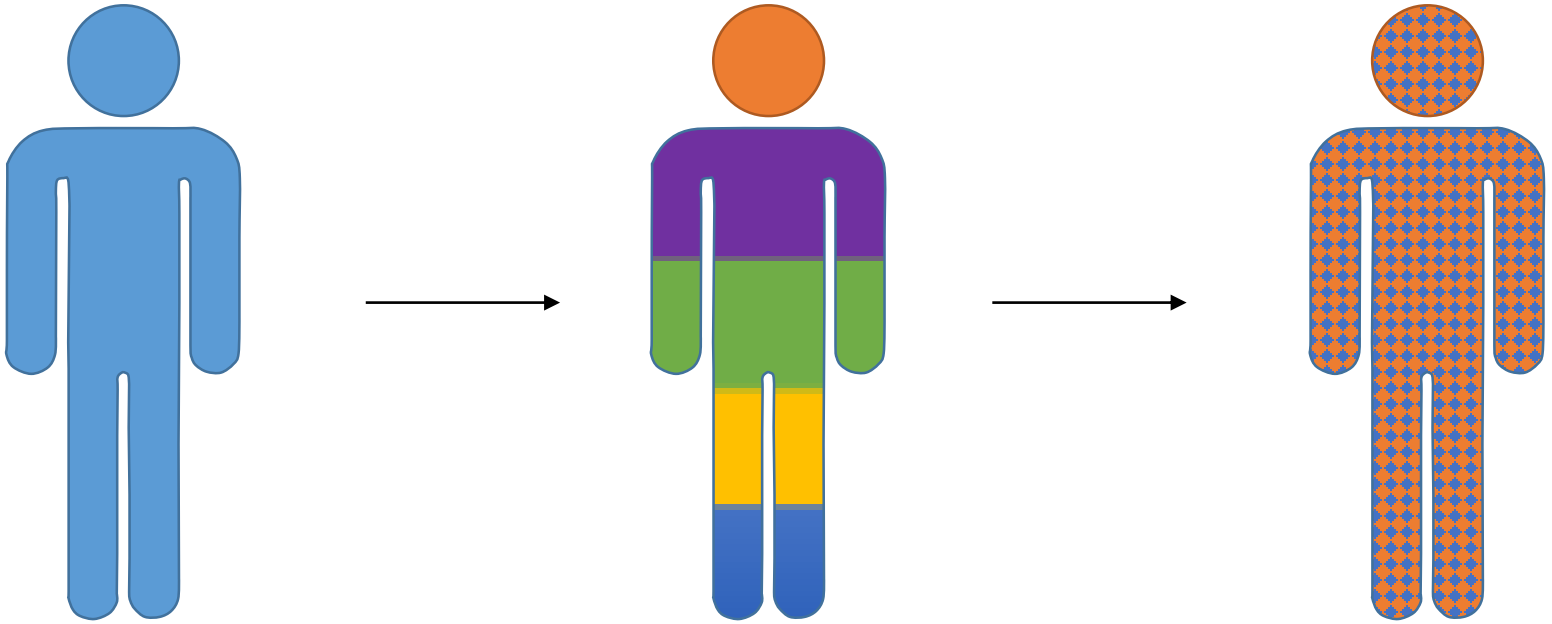
- Understanding variable and value
- The essence of programming




**Variable:** the host
**Value:** the content

# Data operation

Multi-element variable



a = 1;

a = [1,2,4,100];

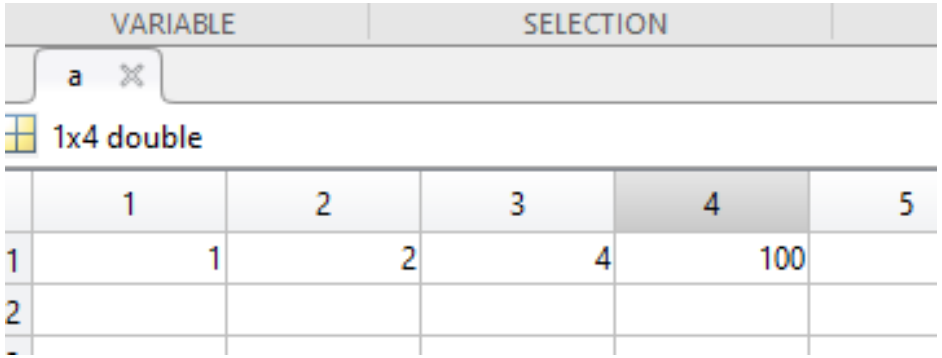
a = {1,'hi', 0.39, 'c'};

# Data operation

- Value assignment
- **Value extraction**
- Operations
  - Plotting; data importing/exporting; operating hardware, etc
  - Function
- Automatic operations
  - If statement;
  - For statement;

# Data operation

```
a = [1,2,4,100];
```




```
b = a(3);
```

```
b = a(3:4);
```

# Data operation

- Value assignment
- Value extraction
- **Operations**
  - Plotting; data importing/exporting; operating hardware, etc
  - Function
- Automatic operations
  - If statement;
  - For statement;

# Data operation

Play with plotting; data importing/exporting; operating hardware, etc

misc.m

# Data operation

## Function

- An input-output system
- Packed operation
- Allows your to achieve certain data processing goal without knowing how it works

Applying a function:

# Data operation

## Function

- An input-output system
- Packed operation
- Allows your to achieve certain data processing goal without knowing how it works

```
y = fun(x);
```

```
a = [1,2,4,100];

b = a(3);
```

```
b = max(a);
```

# Data operation

- Value assignment
- Value extraction
- Operations
  - Plotting; data importing/exporting; operating hardware, etc
  - Function
- **Automatic operations**
  - if statement;
  - for statement;

# if statement

Let computer make decision for you

```
if condition
    do something;
end
```

(try it out)

# for statement

Automatize repetitive operations

```
for j = 1:n
    do something;
end
```

(try it out)

# Data operation

**while statement**
**switch statement**
**others …**

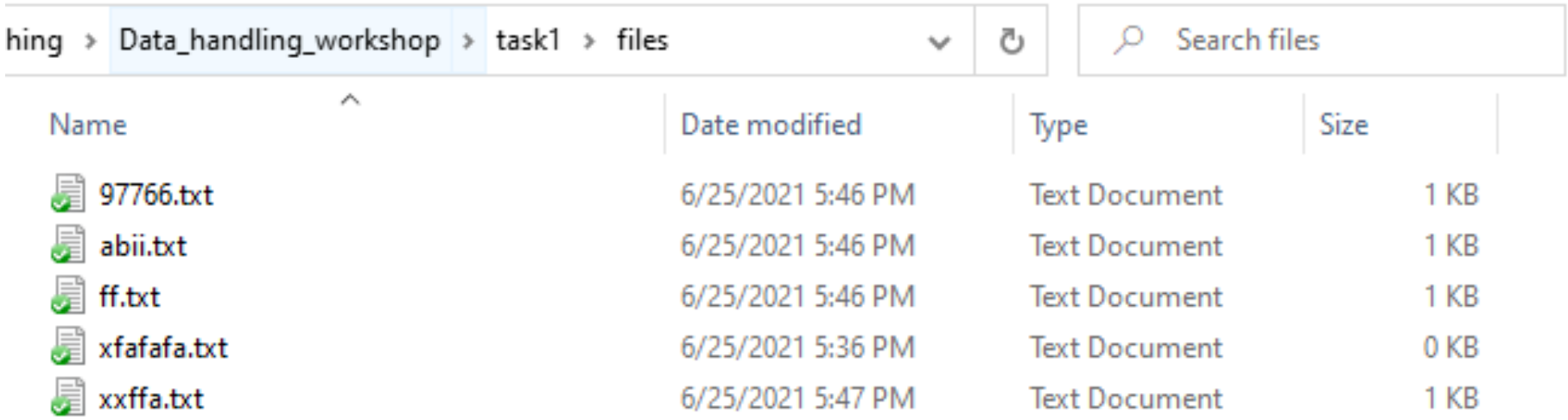
Check them out by yourself from google

# Outline

1. Data types, organization and structures
2. Data operation (in Matlab)
3. Examples:
   - ❑Basic data/file operation
   - ❑Basic matrix operation
   - ❑Simple image processing
   - ❑Simple EEG & eyetracking data processing
   - ❑Simple AI application

# Examples

Basic data/file operation



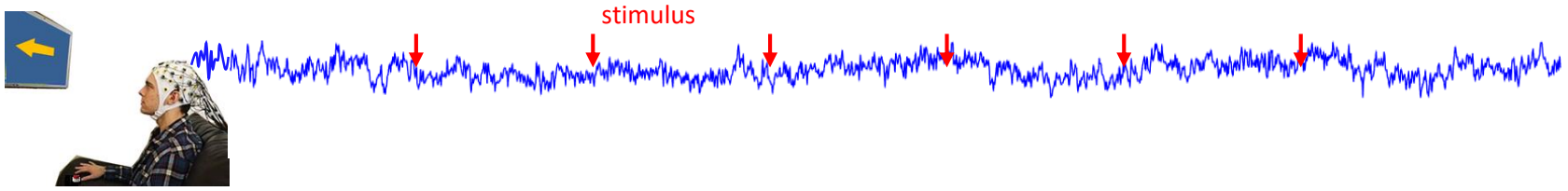
# Examples

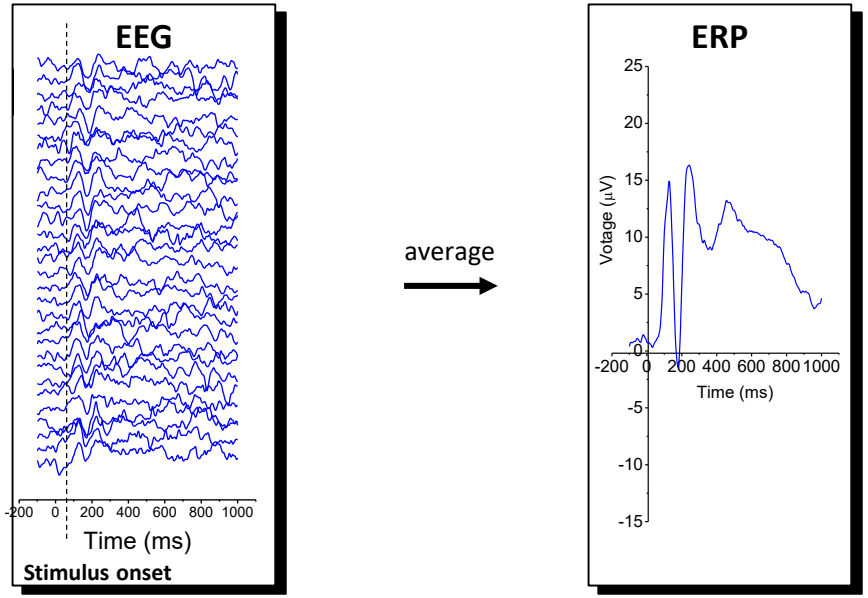
Basic matrix operation

| | | | |
|---|---|---|---|
| 4 | 7 | 9 | -5 |
| -9 | -3.3 | 5 | 0 |
| 22 | 3 | 4 | -1 |
| 10 | 2 | 9 | 3 |

# Examples

## Analyzing brain EEG signal

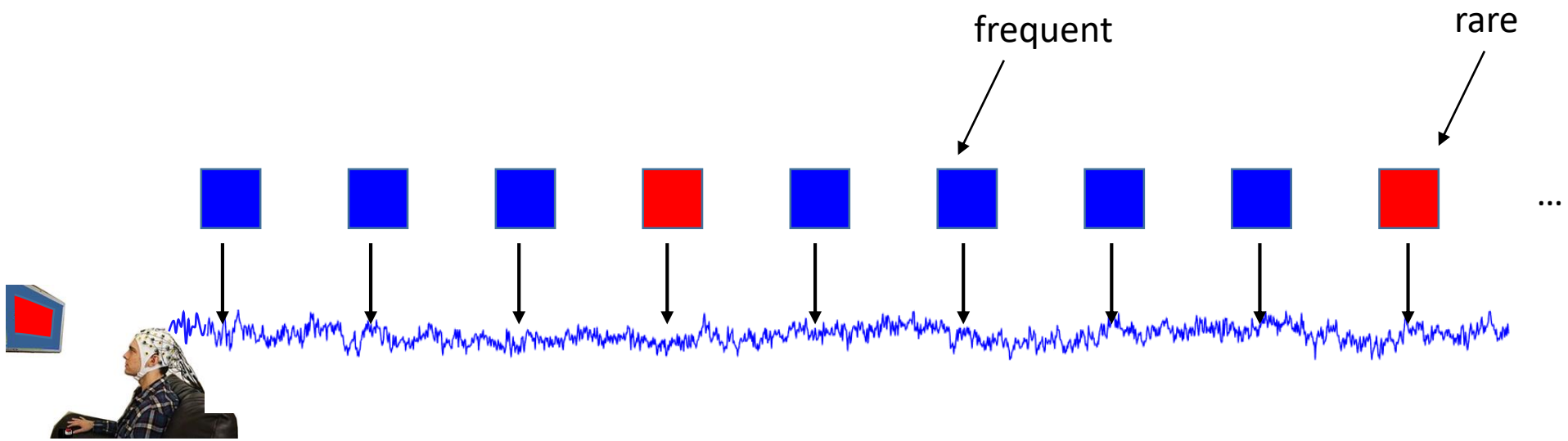



ERP (Event-related Potential)

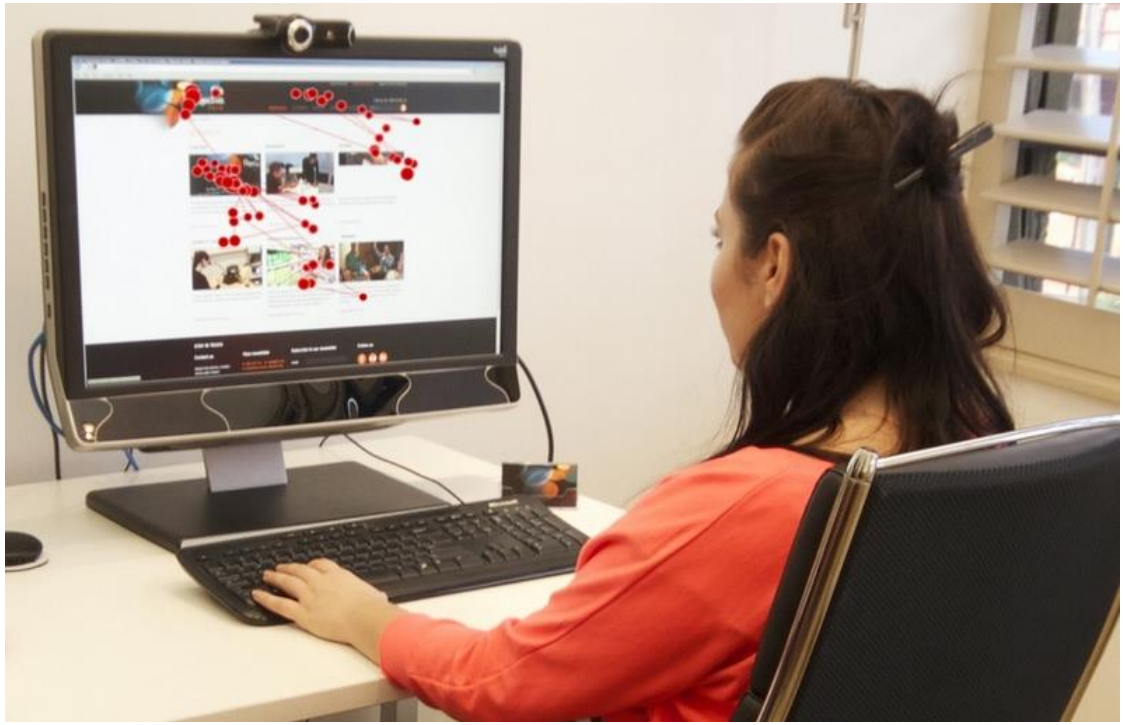
# Examples

Analyzing brain EEG signal



# Examples

Analyzing eye-tracking data



https://www.pantechsolutions.net/blog/wp-content/uploads/2015/02/eye-tracking.png

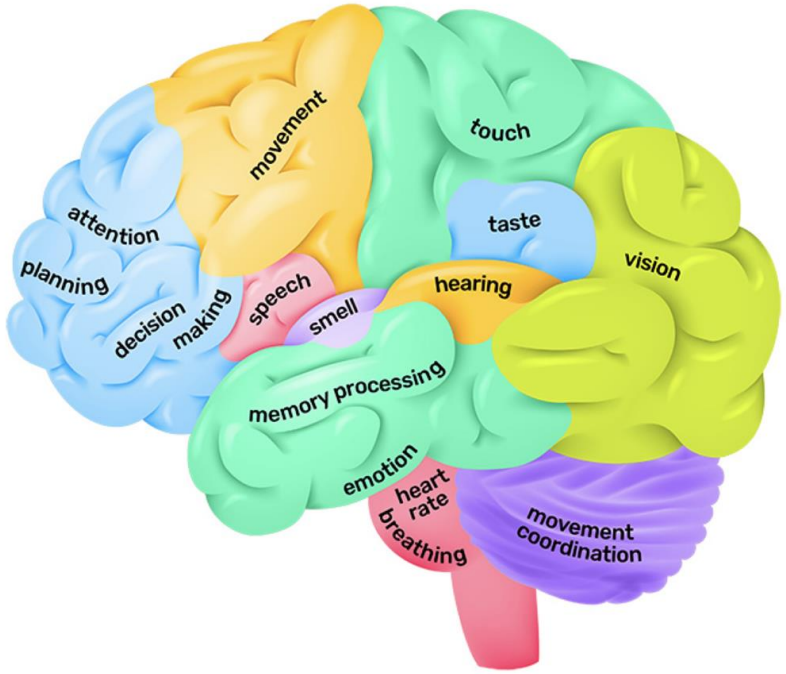

https://i.marketingprofs.com/assets/images/daily-data-point/heat-map-of-FB-results-page-mediative.jpg

# Examples

## Simple AI application

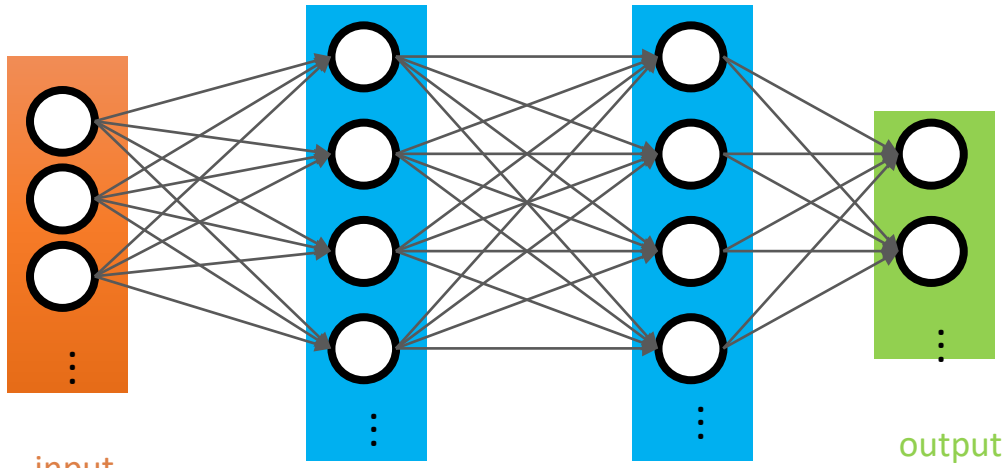
output

**brain**




input

**AI**



input layer

hidden layer

hidden layer

output layer

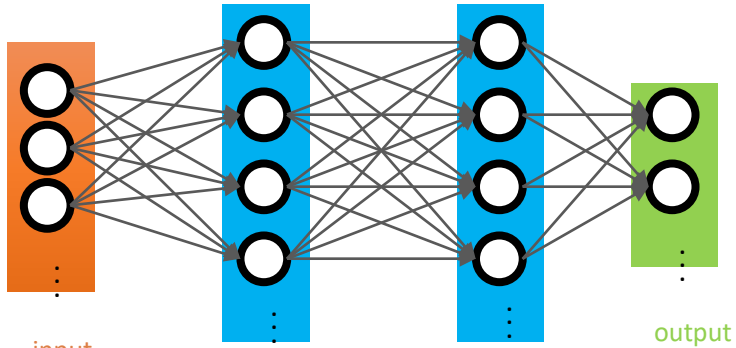
Can have many hidden layers

# Examples

cat dog dog cat dog dog cat dog dog dog cat dog ....

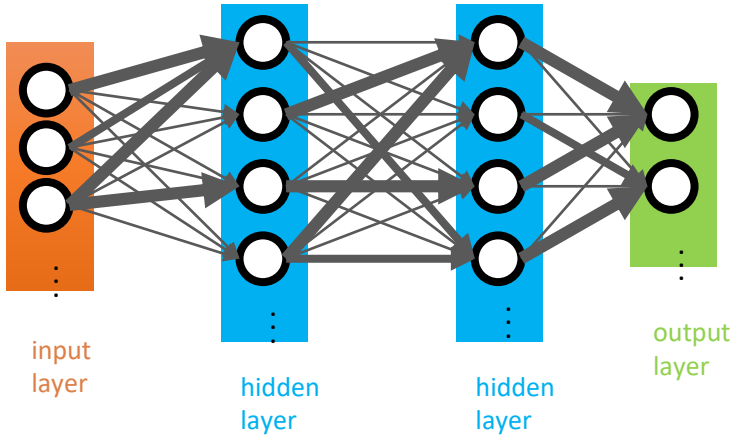
.....

feed for training




Can have many hidden layers

after training

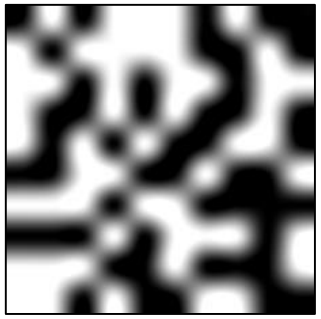


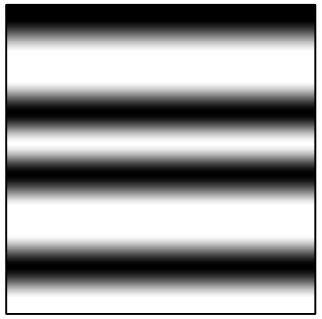

P(cat) = 0.954
P(dog) = 0.006

P(dog) = 0.921
P(cat) = 0.079

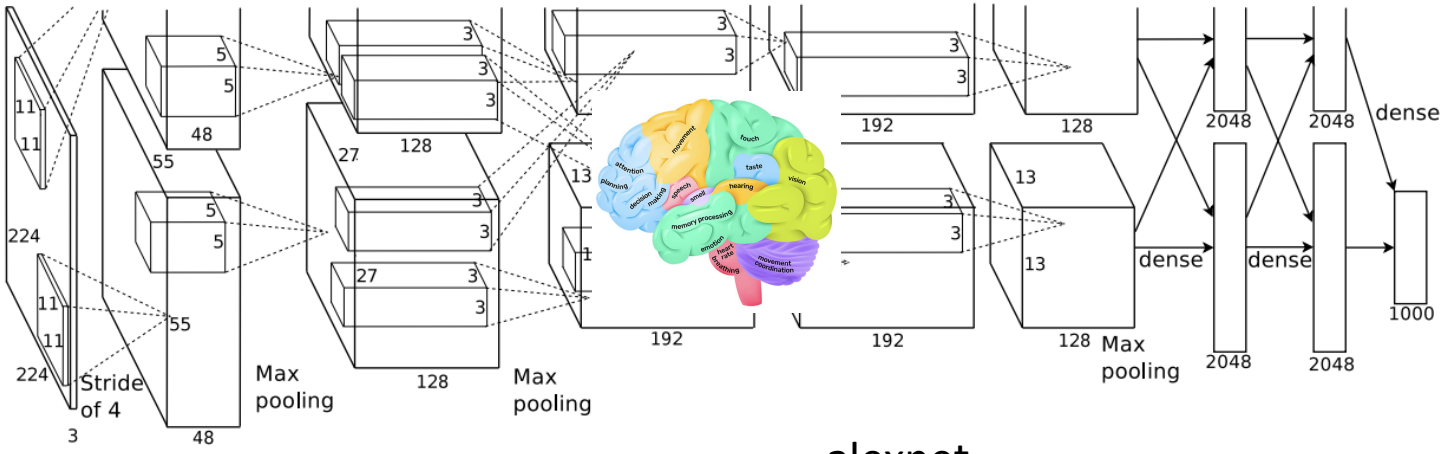
# Examples

→ Take hours to train





→ Take minutes to train




alexnet

# Wrap-up

- Embark on this journey as early as possible
- Data handling skill is accumulative (over years)
- Task-driven
- Google
- Every problem can be turned into a programming problem
- Everybody can do it

Give your feedback

https://forms.gle/YC9Pv8AbeN8PpTPd8