

# Data Science

Code in Place 2021



# Final Project



# Why is this so fast?



mantis shrimp colors



All

Videos

Shopping

Images

News

More

Settings

Tools

About 1,870,000 results (0.54 seconds)

Humans and many other primates have three; some birds and reptiles have four photoreceptors. Certain butterflies can even have six. But the mantis shrimp has **12** different types of photoreceptors in their eyes – and scientists haven't understood why until now. Jan 27, 2014



Study Offers Insights into Unique Color Vision of Mantis Shrimp ...

[www.sci-news.com/biology/science-color-vision-mantis-shrimp-01719.html](http://www.sci-news.com/biology/science-color-vision-mantis-shrimp-01719.html)





# Today

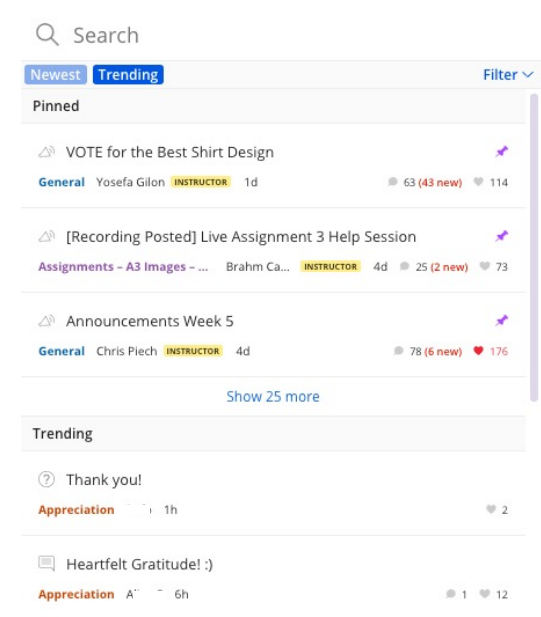
## Review



## Ultimate Problem



## Data Explore on Code in Place Ed





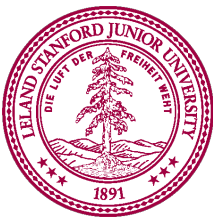
Review

# Dict Review

```
# 1. Make a new Dict  
animal_sounds = {}
```

```
# 2. Put things into the Dict  
animal_sounds["dog"] = "woof"  
animal_sounds["cat"] = "meow"  
animal_sounds["seal"] = "ow ow ow"
```

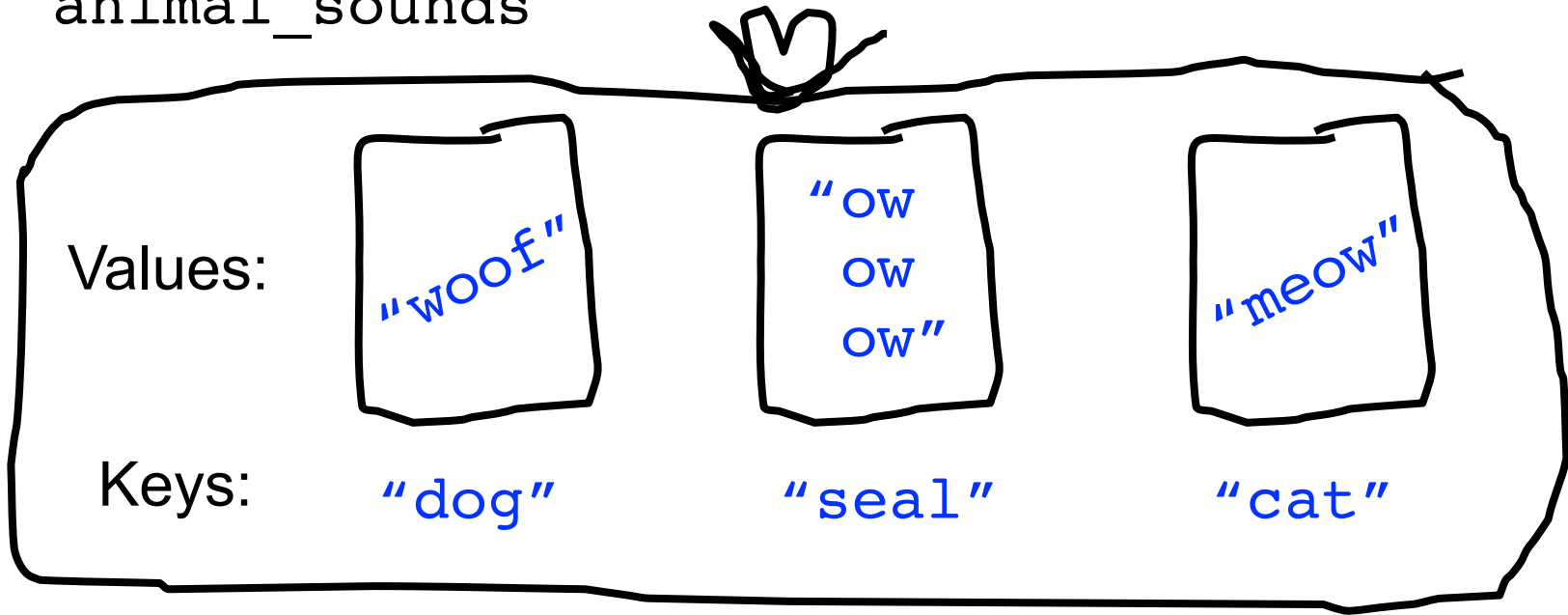
```
# 3. Get things out of the Dict  
dog_sound = animal_sounds["dog"] # "woof"
```





# Dict Review

animal\_sounds



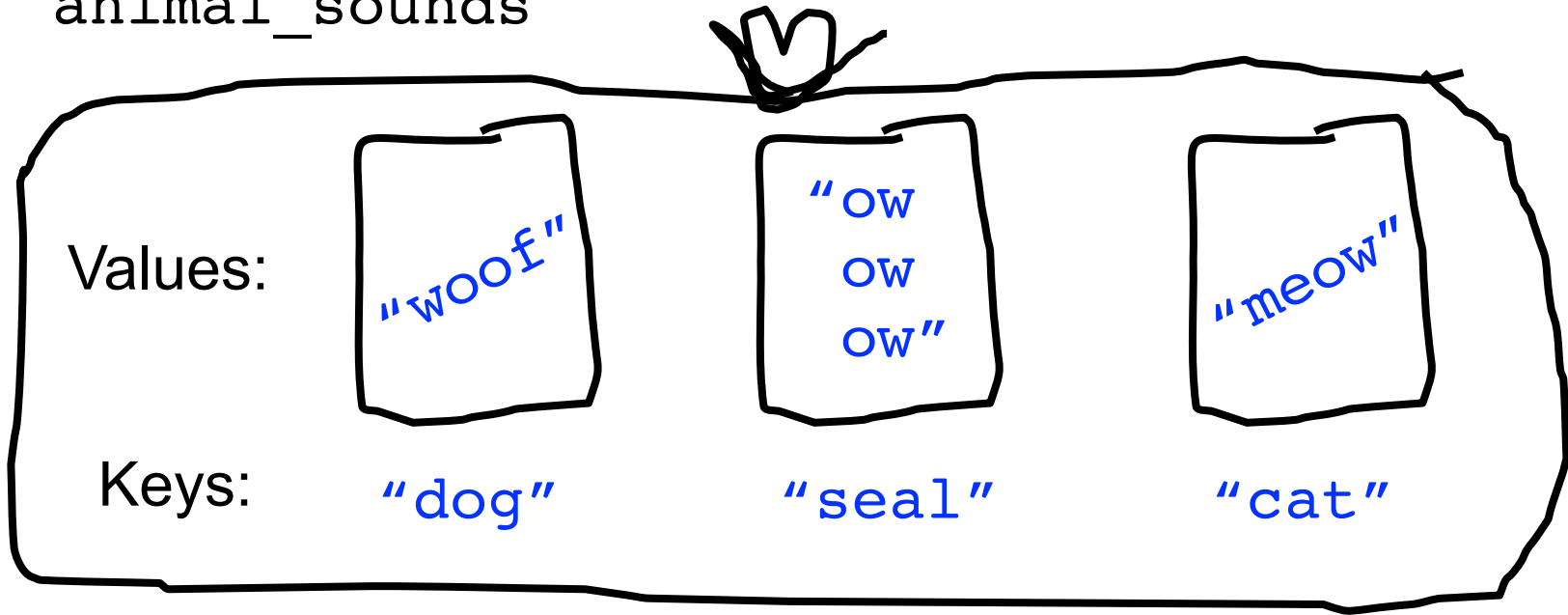
```
# 1. Make a new Dict
animal_sounds = {}
```

```
# 2. Put things into the Dict
animal_sounds["dog"] = "woof"
animal_sounds["cat"] = "meow"
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict
dog_sound = animal_sounds["dog"] # "woof"
```

# Dict Review

animal\_sounds



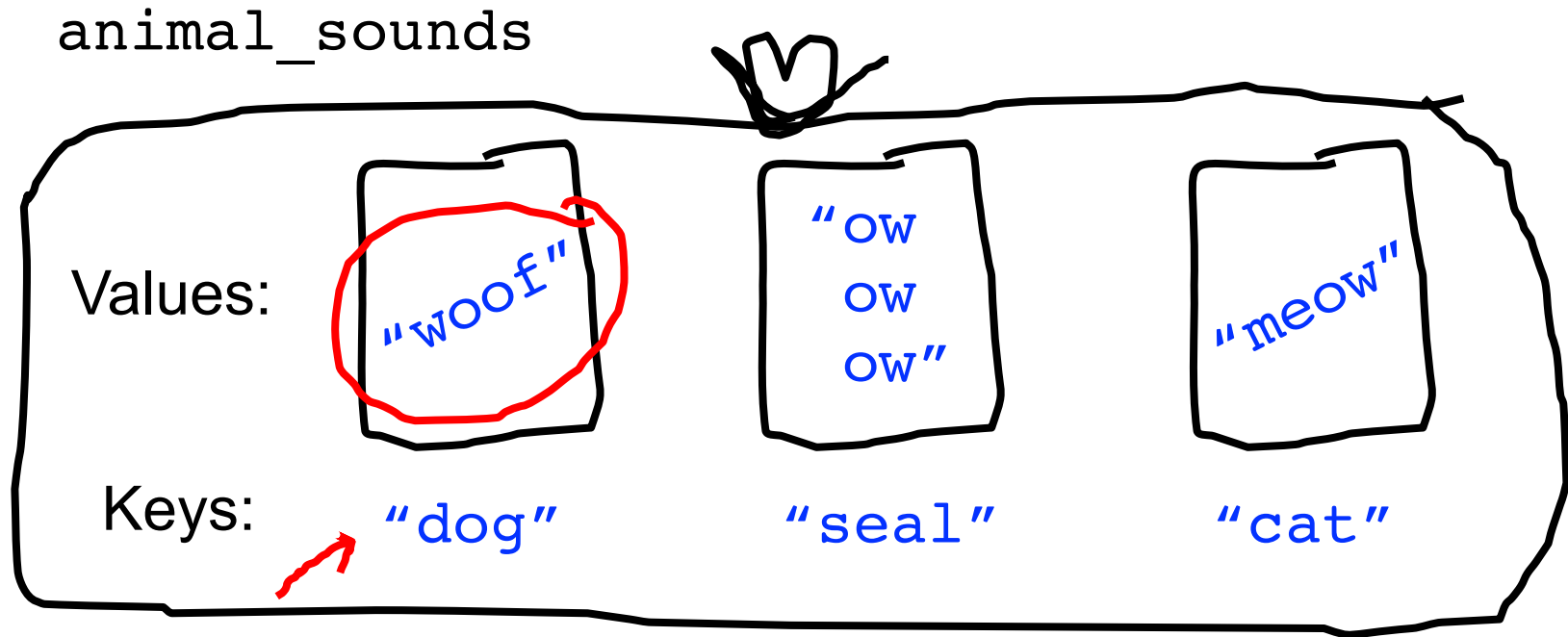
```
# 1. Make a new Dict  
animal_sounds = {}
```

```
# 2. Put things into the Dict  
animal_sounds["dog"] = "woof"  
animal_sounds["cat"] = "meow"  
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict  
dog_sound = animal_sounds["dog"] # "woof"
```



# Dict Review



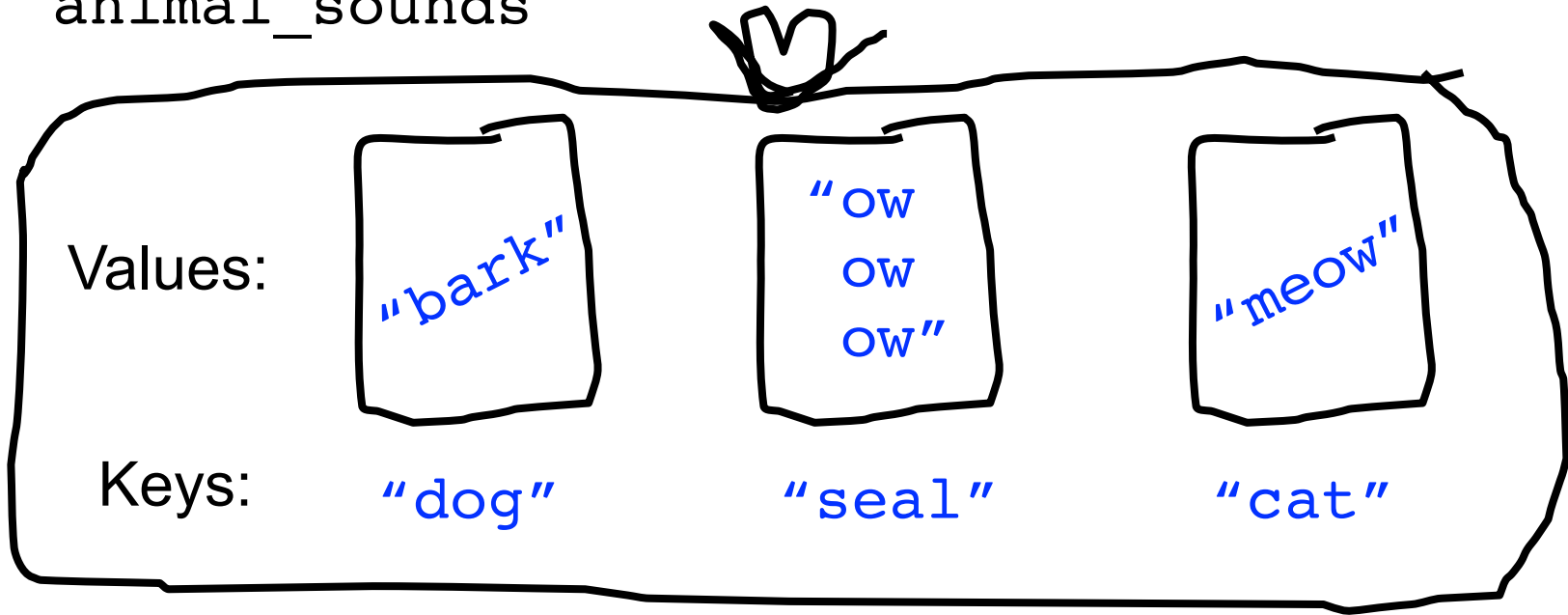
```
# 1. Make a new Dict  
animal_sounds = {}
```

```
# 2. Put things into the Dict  
animal_sounds["dog"] = "woof"  
animal_sounds["cat"] = "meow"  
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict  
dog_sound = animal_sounds["dog"] # "woof"
```

# Dict Review

animal\_sounds



```
# 1. Make a new Dict
animal_sounds = {}
```

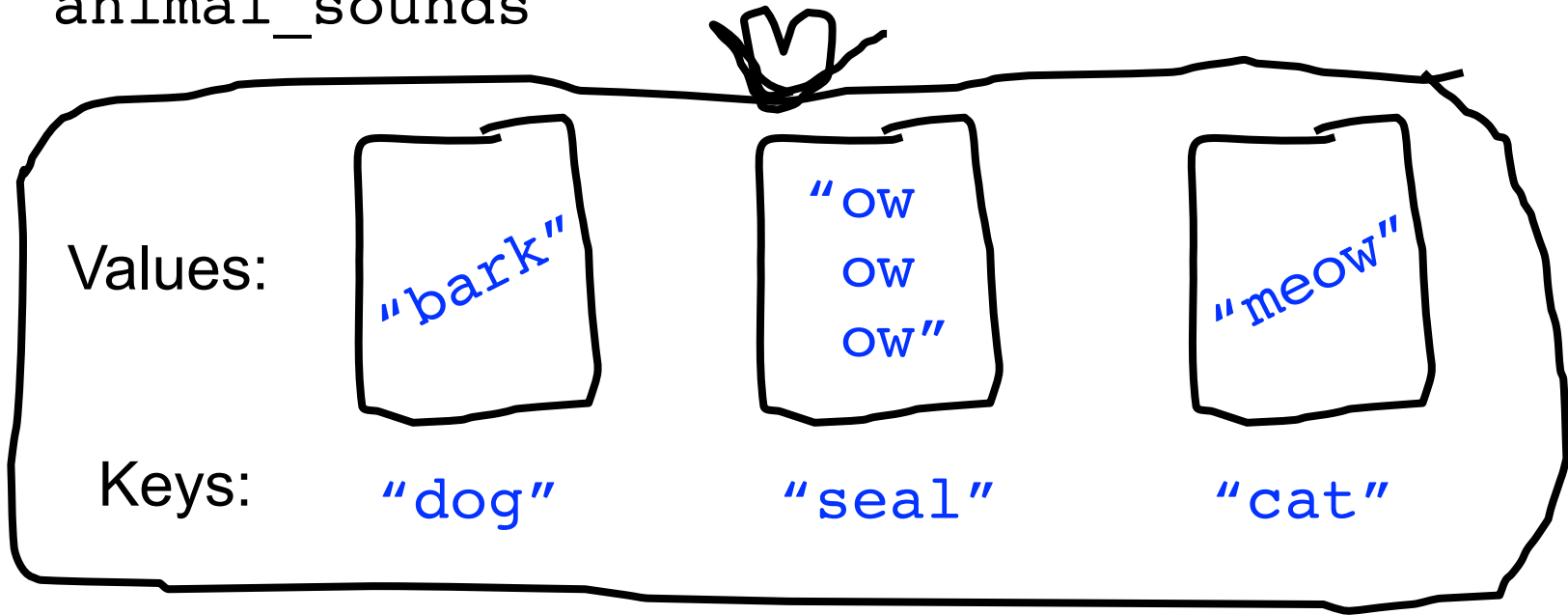
```
# 2. Put things into the Dict
animal_sounds["dog"] = "woof"
animal_sounds["cat"] = "meow"
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict
dog_sound = animal_sounds["dog"] # "woof"
fox_sound = animal_sounds["fox"]
```



# Dict Review

animal\_sounds



```
# 1. Make a new Dict
animal_sounds = {}
```

```
# 2. Put things into the Dict
animal_sounds["dog"] = "woof"
animal_sounds["cat"] = "meow"
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict
dog_sound = animal_sounds["dog"] # "woof"
fox_sound = animal_sounds["fox"]  # KeyError: 'fox'
```



brothers Vegard  
and Bård Ylvisåker

Circa 2013



*But there's one sound*

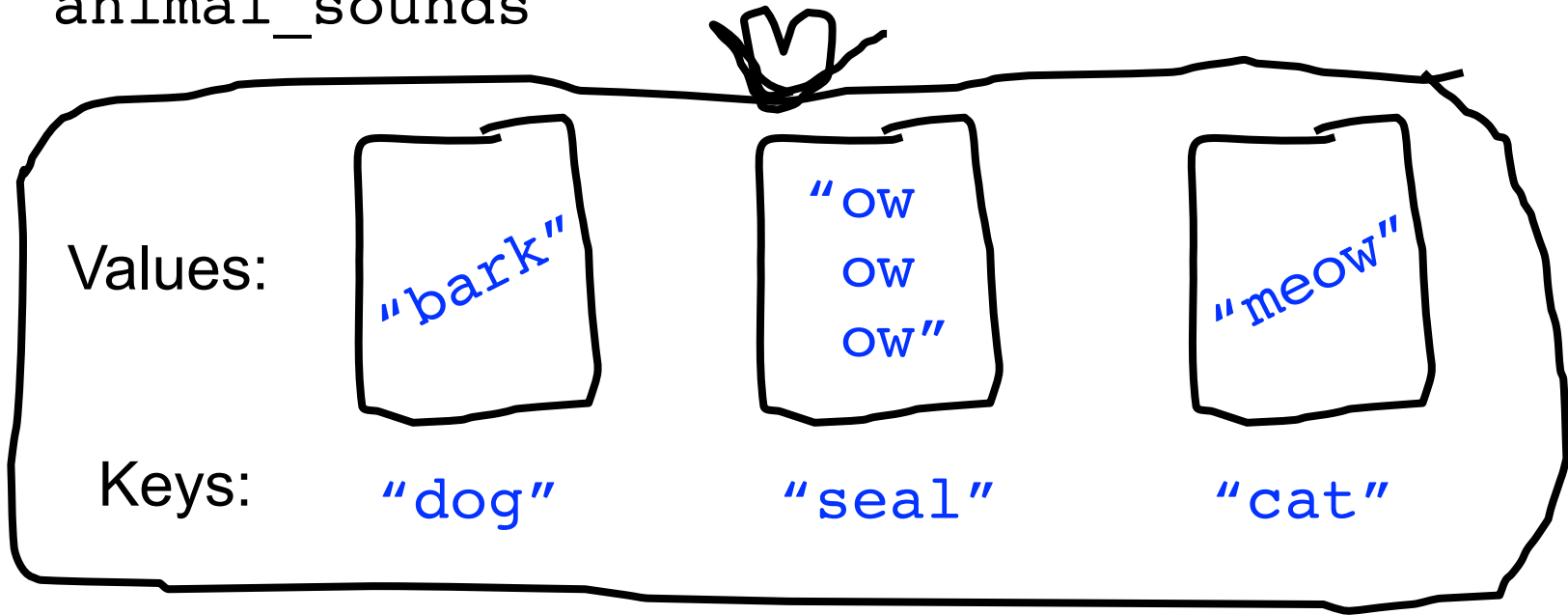
C Ylvis – “The Fox”. Permission asked. Pending.





# Dict Review

animal\_sounds



```
# 1. Make a new Dict
animal_sounds = {}
```

```
# 2. Put things into the Dict
animal_sounds["dog"] = "woof"
animal_sounds["cat"] = "meow"
animal_sounds["seal"] = "ow ow ow"
```

```
# 3. Get things out of the Dict
dog_sound = animal_sounds["dog"] # "woof"
fox_sound = animal_sounds["fox"] # KeyError: 'fox'
```



Dictionary  
key -> value

List

index -> value



# List

```
my_list = ['a', 'b', 'c']
```

```
print(my_list[1])
```

```
for i in range(len(my_list)):
    value = my_list[i]
    print(i, value)
```

my\_list

a	b	c
---	---	---

0

1

2

*indices*

# Dictionary

```
my_dict = {
    'x': 'a',
    'y': 'b',
    'z': 'c'
}
```

```
print(my_dict['y'])
```

```
for key in my_dict:
    value = my_dict[key]
    print(key, value)
```

my\_dict

a	b	c
---	---	---

'x'

'y'

'z'

*keys*



# List

```
my_list = [  
    'a',  
    'b',  
    'c'  
]
```

```
print(my_list[1])
```

```
for i in range(len(my_list)):  
    value = my_list[i]  
    print(i, value)
```

my\_list

a	b	c
---	---	---

0

1

2

*indices*

# Dictionary

```
my_dict = {  
    'x': 'a',  
    'y': 'b',  
    'z': 'c'  
}
```

```
print(my_dict['y'])
```

```
for key in my_dict:  
    value = my_dict[key]  
    print(key, value)
```

my\_dict

a	b	c
---	---	---

'x'

'y'

'z'

*keys*



# List

```
my_list = ['a', 'b', 'c']
```

```
print(my_list[1])
```

```
for i in range(len(my_list)):
    value = my_list[i]
    print(i, value)
```

my\_list

a	b	c
---	---	---

0

1

2

*indices*

# Dictionary

```
my_dict = {
    'x': 'a',
    'y': 'b',
    'z': 'c'
}
```

```
print(my_dict['y'])
```

```
for key in my_dict:
    value = my_dict[key]
    print(key, value)
```

my\_dict

a	b	c
---	---	---

'x'

'y'

'z'

*keys*



# List

```
my_list = ['a', 'b', 'c']
```

```
print(my_list[1])
```

```
for i in range(len(my_list)):
    value = my_list[i]
    print(i, value)
```

my\_list

a	b	c
---	---	---

0

1

2

*indices*

# Dictionary

```
my_dict = {
    'x': 'a',
    'y': 'b',
    'z': 'c'
}
```

```
print(my_dict['y'])
```

```
for key in my_dict:
    value = my_dict[key]
    print(key, value)
```

my\_dict

a	b	c
---	---	---

'x'

'y'

'z'

*keys*





# List

```
my_list = ['a', 'b', 'c']
```

```
print(my_list[1])
```

```
for i in range(len(my_list)):  
    value = my_list[i]  
    print(i, value)
```

my\_list

a	b	c
---	---	---

0

1

2

*indices*

# Dictionary

```
my_dict = {  
    'x': 'a',  
    'y': 'b',  
    'z': 'c'  
}
```

```
print(my_dict['y'])
```

```
for key in my_dict:  
    value = my_dict[key]  
    print(key, value)
```

my\_dict

a	b	c
---	---	---

'x'

'y'

'z'

*keys*



# Dictionaries are one way!



In dictionaries you can only look up values by keys. You can't look up keys by value.

```
animal_sounds = {  
    'dog': 'woof',  
    'cat': 'meow'  
}
```

`animal_sounds['meow']`  `KeyError: 'meow'`



# Each key gets has only one value!



If you put a key in the dictionary twice, it will overwrite

```
animal_sounds['dog'] = 'bark'  
animal_sounds['dog'] = 'woof'
```



```
animal_sounds = {'dog': 'woof'}
```



# Common Bug



**You can use variables as keys! Be careful of quotes**

```
animal_sounds = {  
    'dog': 'woof',  
    'cat': 'meow'  
}
```

A

```
animal = input('? ')  
print(animal_sounds["animal"])
```

B

```
animal = input('? ')  
print(animal_sounds[animal])
```





# Common Bug



**Key lookups can be literals or variables. Don't confuse the two**

```
animal_sounds = {  
    'dog': 'woof',  
    'cat': 'meow'  
}
```

A

```
animal = input('? ')  
print(animal_sounds["animal"])
```

B

```
animal = input('? ')  
print(animal_sounds[animal])
```



# Core Datastructures

All datasets can  
be represented by:



**Dictionaries,**  
**Lists,**  
strings,  
floats,  
integers  
booleans



# Example Google Maps Query Result

```
{  
  "markers": [  
    {  
      "name": "Rixos The Palm Dubai",  
      "position": [25.1212, 55.1535],  
    },  
    {  
      "name": "Shangri-La Hotel",  
      "location": [25.2084, 55.2719]  
    },  
    {  
      "name": "Grand Hyatt",  
      "location": [25.2285, 55.3373]  
    }  
  ]  
}
```

Welcome to the wild west of data



# JSON

results.json

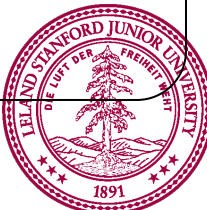
```
{
  "markers": [
    {
      "name": "The Palm Dubai",
      "position": [25.1212, 55.1535],
    },
    {
      "name": "Shangri-La Hotel",
      "location": [25.2084, 55.2719]
    },
    {
      "name": "Grand Hyatt",
      "location": [25.2285, 55.3373]
    }
  ]
}
```

program.py

```
import json

def main():
    # load data from file
    my_file = open('results.json', 'w')
    data = json.load(my_file)

    print(data)
```



End Review

Are you ready?



For...

The ULTIMATE cs106a question?

# Ultimate CS106A: Reverse a Dict



**Normal Dict:**

Key -> Value

**Reversed Dict:**

Value -> Keys

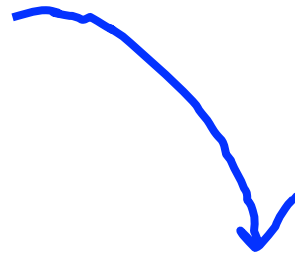


Claim: understanding this  
single example is most  
indicative of mastery in  
CS106A



# Ultimate CS106A: Reverse a Dict

```
ages = {  
    'Mehran':51,  
    'Gary':70,  
    'Chris':33,  
    'Freya':1,  
    'Adele':33,  
    'Lionel':33,  
    'Rihanna':33,  
    'Stephen':33  
}
```



```
reversed = {  
    51: ['Mehran'],  
    70: ['Gary'],  
    33: ['Chris', 'Adele', 'Lionel', 'Rihanna', 'Stephen'],  
    1 : ['Freya']  
}
```



What is the strategy?

# Ultimate CS106A: Reverse a Dict

Gary -> 70

reversed = {

}





# Ultimate CS106A: Reverse a Dict

Gary -> 70

```
reversed = {  
    70 : ['Gary'],
```

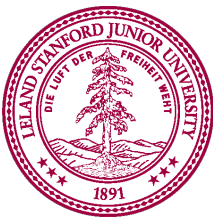
```
}
```



# Ultimate CS106A: Reverse a Dict

Chris -> 33

```
reversed = {  
    70 : ['Gary'],  
  
    33 : ['Chris'],  
  
}
```



# Ultimate CS106A: Reverse a Dict

Mehran -> 51

```
reversed = {  
    70 : ['Gary'],  
  
    33 : ['Chris'],  
  
}
```



# Ultimate CS106A: Reverse a Dict

Mehran -> 51

```
reversed = {  
    70 : ['Gary'],  
    51 : ['Mehran'],  
    33 : ['Chris'],  
}
```



# Ultimate CS106A: Reverse a Dict

Rihanna -> 33

```
reversed = {  
    70 : ['Gary'],  
    51 : ['Mehran'],  
    33 : ['Chris'],  
}
```



# Ultimate CS106A: Reverse a Dict

Rihanna -> 33

```
reversed = {  
    70 : ['Gary'],  
    51 : ['Mehran'],  
    33 : ['Chris', 'Rihanna'],  
}
```



# Ultimate CS106A: Reverse a Dict

```
reversed = {  
    70 : ['Gary'],  
    51 : ['Mehran'],  
    33 : ['Chris', 'Rihanna', 'Stephen', ... ]  
    1 : ['Freya']  
}
```



something awesome



# Data Analysis on our Ed?

Search

Newest Trending Filter

Pinned

VOTE for the Best Shirt Design  
General Yosefa Gilon INSTRUCTOR 1d 63 (43 new) 114

[Recording Posted] Live Assignment 3 Help Session  
Assignments - A3 Images - ... Brahm Ca... INSTRUCTOR 4d 25 (2 new) 73

Announcements Week 5  
General Chris Piech INSTRUCTOR 4d 78 (6 new) 176

Show 25 more

Trending

Thank you!  
Appreciation 1h 2

Heartfelt Gratitude! :)  
Appreciation 6h 1 12



Analytics CSV

Threads JSON

```
original_ed.json
1 [
2   {
3     "url": "https://us.edstem.org/courses/10000/discussion/4505177c",
4     "votes": 9,
5     "created_at": "2021-05-20T04:43:22.348474+10:00",
6     "user": {
7       "name": "Chris Piech",
8       "email": "piech@cs.stanford.edu",
9       "role": "admin"
10    },
11    "text": "Happy birthday Mehran! You are a fantastic teacher, me",
12    "document": "<document version='2.0'><paragraph>Happy birthda",
13    "comments": []
14  },
15  {
16    "url": "https://us.edstem.org/courses/10000/discussion/451906",
17    "type": "post",
18    "title": "Took the diagnostic? Can you help us test our diagnos",
19    "category": "Diagnostic",
20    "subcategory": "",
21    "votes": 54,
22    "private": false,
23    "anonymous": false,
24    "created_at": "2021-05-20T04:37:49.800578+10:00",
25    "user": {
26      "name": "Chris Piech",
27      "email": "piech@cs.stanford.edu",
28      "role": "admin"
29    },
30    "text": "Hi all, \n\nIf someone is online right now could you h",
31    "document": "<document version='2.0'><paragraph>Hi all, </par",
32    "comments": ["many comments"]
33  },
34  {
35    "url": "https://us.edstem.org/courses/10000/discussion/4333227c",
36    "votes": 5,
37    "created_at": "2021-05-12T06:42:01.885227+10:00",
38    "user": {
39      "name": "Mehran Sahami",
40      "email": "sahami@cs.stanford.edu",
41      "role": "admin"
42    },
43    "text": "Great work! It looks like a flower made from burritos",
44    "document": "<document version='2.0'><paragraph>Great work!",
45    "comments": []
46  }
47 ]
```



Mystery:

What time of day is  
Code in Place most active?

# The Ed Data

A list of "posts". Each post is a dictionary.

```
[
  {
    "created_at": "2021-05-21T01:20:39.296044+10:00",
    "votes": 0,
    "user": {
      "name": "Anonymous",
      "role": ""
    }
  },
  {
    "created_at": "2021-05-21T01:21:25.225994+10:00",
    "votes": 0,
    "user": {
      "name": "Anonymous",
      "role": "admin"
    }
  },
  ...
]
```



# How to get the hour of a post

I will give you this function

```
def get_hour(time_string):  
    """  
    Given a time string, returns the day of the week (in pacific time).  
    >>> get_hour('2021-05-21T01:20:39.296044+10:00')  
    5  
    """  
    date_time = parser.parse(time_string)  
    # change to my timezone  
    date_time = date_time.astimezone(timezone('US/Pacific'))  
    # get the hour out of the time object  
    return date_time.hour
```

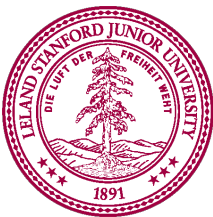


# Analyze Ed... From Ed!

The screenshot displays the edstem.org Code in Place 2021 interface. The browser address bar shows the URL `edstem.org/us/courses/10000/lessons/14518/slides/73858`. The page title is "ed Code in Place 2021 - Lessons". The navigation bar includes links for Lessons, Slides, Prev, Next, Explore Ed, Challenge, Submissions, Solution (hidden), Edit Slide, and a 20+ notification. The left sidebar shows "Lecture 14: Data Science" with a list of activities: Explore Ed (selected), Reverse Dictionary, and a plus icon. The main editor area shows a file explorer with `ed.json`, `ed_small.json`, `post_times.py`, and `student_to_staff.py` (selected). The code editor displays the following Python code:

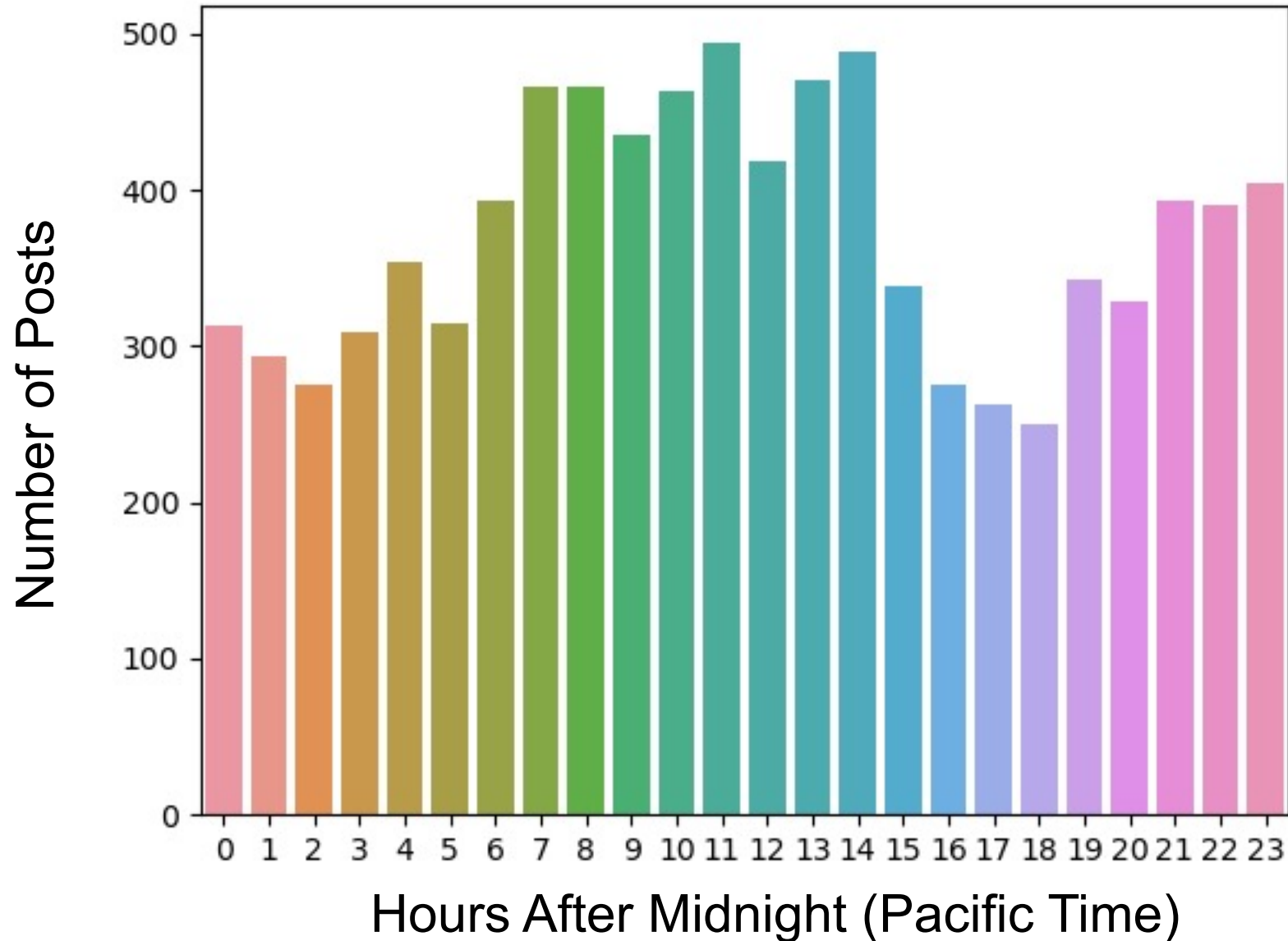
```
1 # for loading files with nested data
2 import json
3
4 def main():
5     ed_data = json.load(open('ed.json'))
6     n_posts = len(ed_data)
7     print('n_posts', n_posts)
8
9     n_teacher_posts = 0
10    for post in ed_data:
11        user_role = post['user']['role']
12        if user_role == 'tutor':
13            n_teacher_posts += 1
14        elif user_role == 'admin':
15            n_teacher_posts += 1
16
17    print(n_teacher_posts / n_posts)
18
19 if __name__ == '__main__':
20    main()
```

The status bar at the bottom indicates the file path `/home/student_to_staff.py`, the time `11:10`, and the number of spaces `Spaces: 4 (Auto)`. It also shows "All changes saved" with a green dot. The bottom right corner features a "Run" button (green play icon) and a "Submit" button (blue checkmark icon).

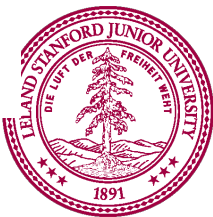


20% of posts on the the  
Code in Place forum are from  
Teachers!

# Code in Place Activity



Code in Place Teaching Team, 2021, Global



# Why is this so fast?



mantis shrimp colors



All

Videos

Shopping

Images

News

More

Settings

Tools

About 1,870,000 results (0.54 seconds)

Humans and many other primates have three; some birds and reptiles have four photoreceptors. Certain butterflies can even have six. But the mantis shrimp has **12** different types of photoreceptors in their eyes – and scientists haven't understood why until now. Jan 27, 2014



Study Offers Insights into Unique Color Vision of Mantis Shrimp ...

[www.sci-news.com/biology/science-color-vision-mantis-shrimp-01719.html](http://www.sci-news.com/biology/science-color-vision-mantis-shrimp-01719.html)



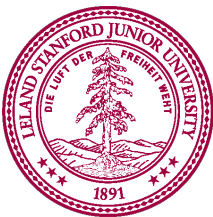


# Why is this so fast?

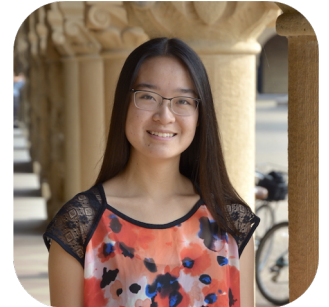


```
hash_int = hash(key);
```

\* There is always more to learn



# Thank You So Much



Code in Place Teaching Team, 2021, Global



Coming up:  
Final week of  
Code in Place  
2021

