

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

## ✓ Numpy

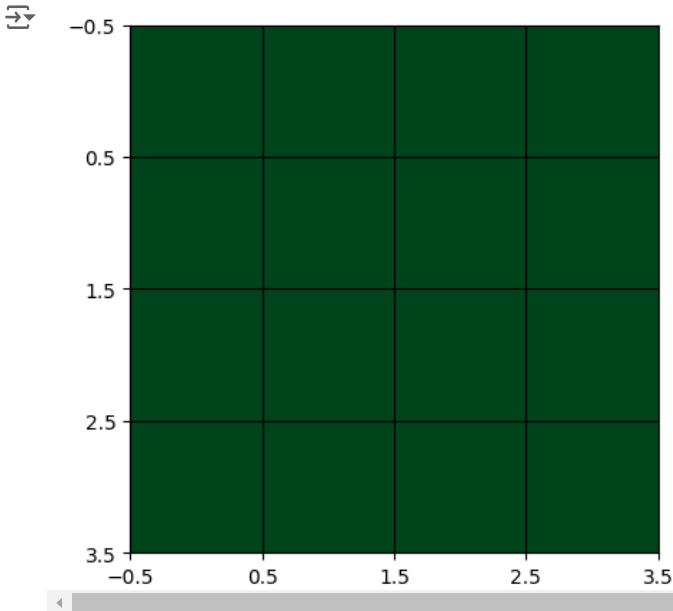
```
def visualize_array(data):

    fig, ax = plt.subplots()
    ax.imshow(data, cmap='Greens', vmin=0, vmax=1)

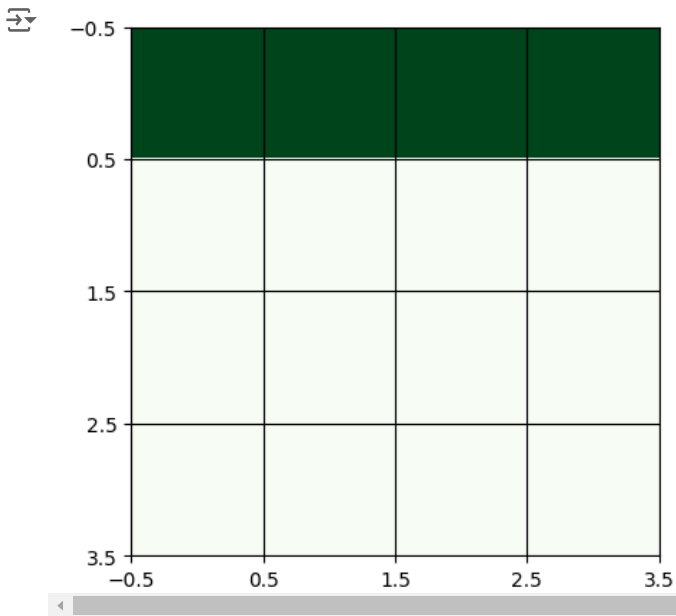
    ax.set_xticks(np.arange(-0.5, 4, 1))
    ax.set_yticks(np.arange(-0.5, 4, 1))
    ax.grid(color='black')

    plt.show()
```

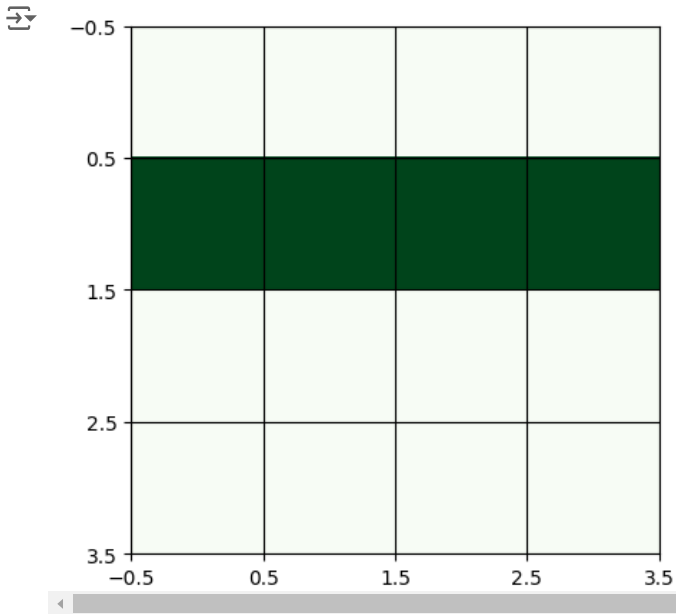
```
data = np.zeros((4, 4))
data[:, :] = 1
visualize_array(data)
```



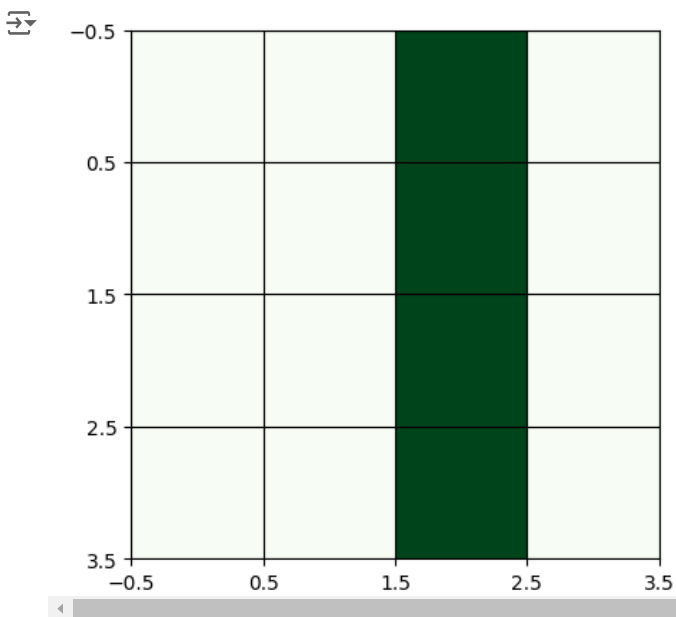
```
data = np.zeros((4, 4))
data[0] = 1
visualize_array(data)
```



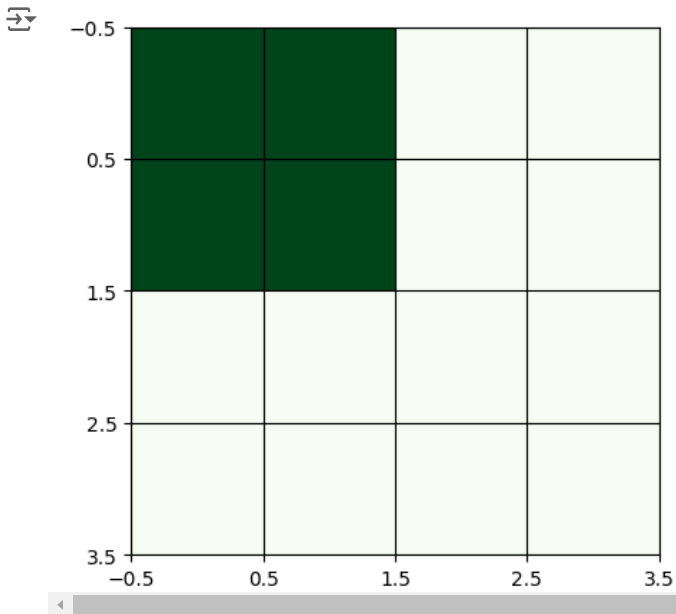
```
data = np.zeros((4, 4))  
data[1] = 1  
visualize_array(data)
```



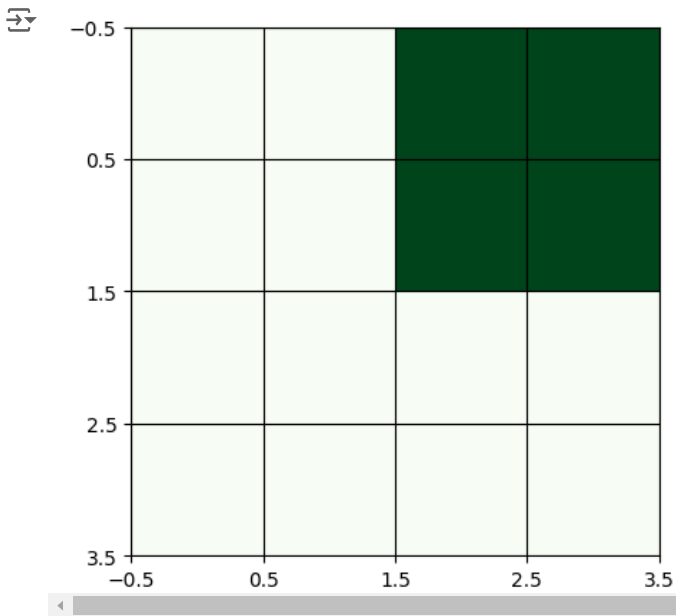
```
data = np.zeros((4, 4))  
data[:, 2] = 1  
visualize_array(data)
```



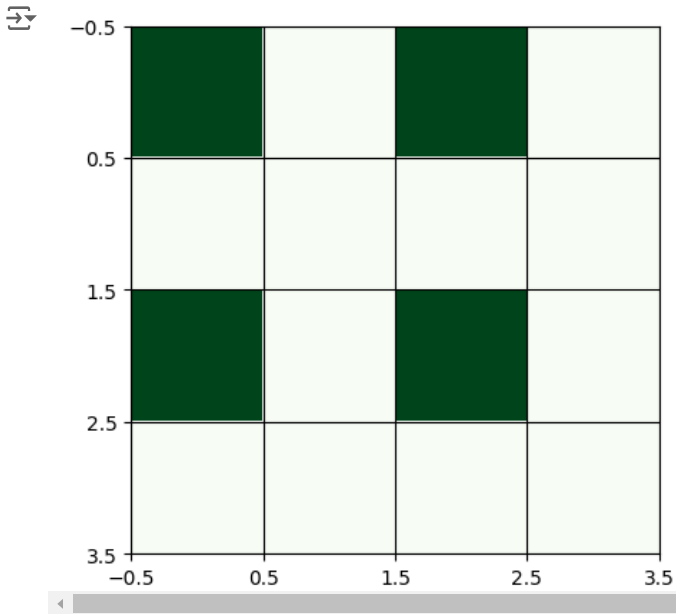
```
data = np.zeros((4, 4))
data[:, :2] = 1
visualize_array(data)
```



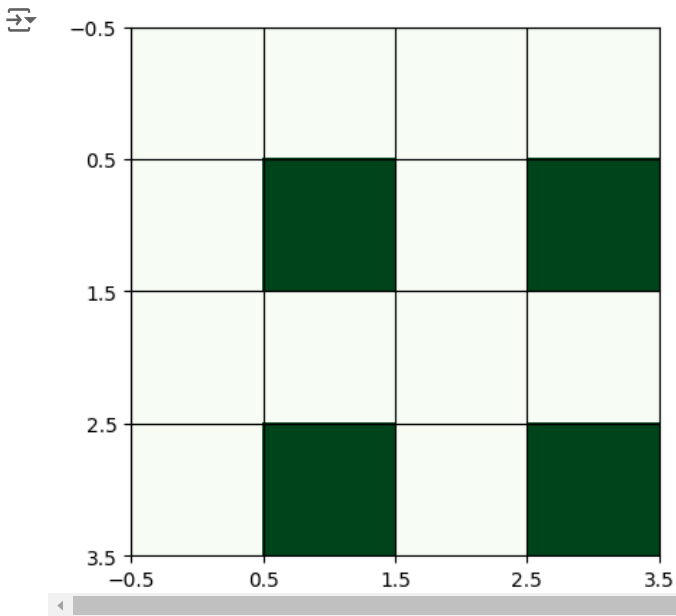
```
data = np.zeros((4, 4))
data[2, 2:] = 1
visualize_array(data)
```



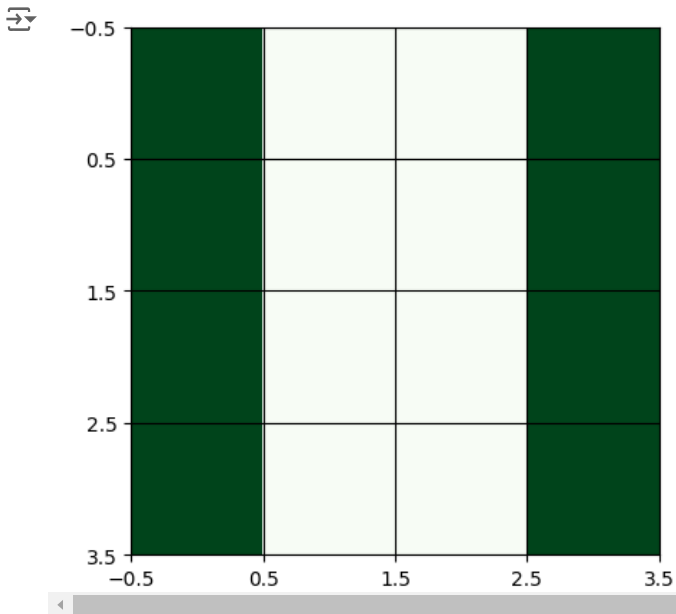
```
data = np.zeros((4, 4))  
data[0:3:2, 0:3:2] = 1  
visualize_array(data)
```



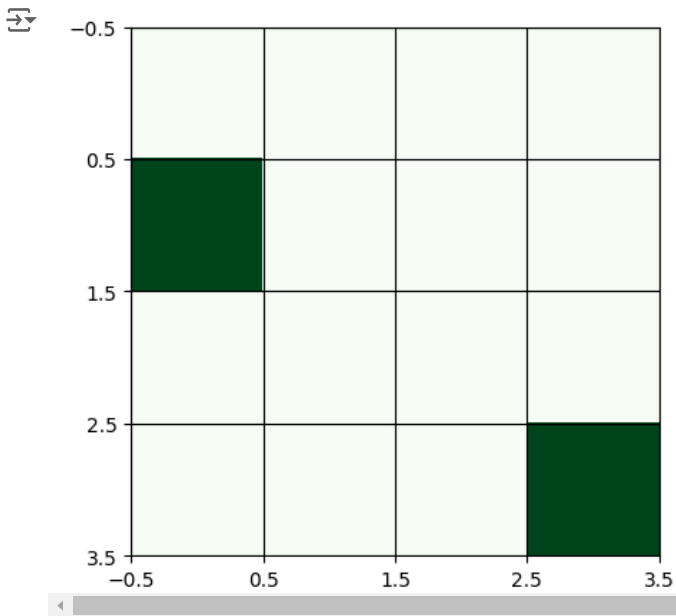
```
data = np.zeros((4, 4))  
data[1::2, 1::2] = 1  
visualize_array(data)
```



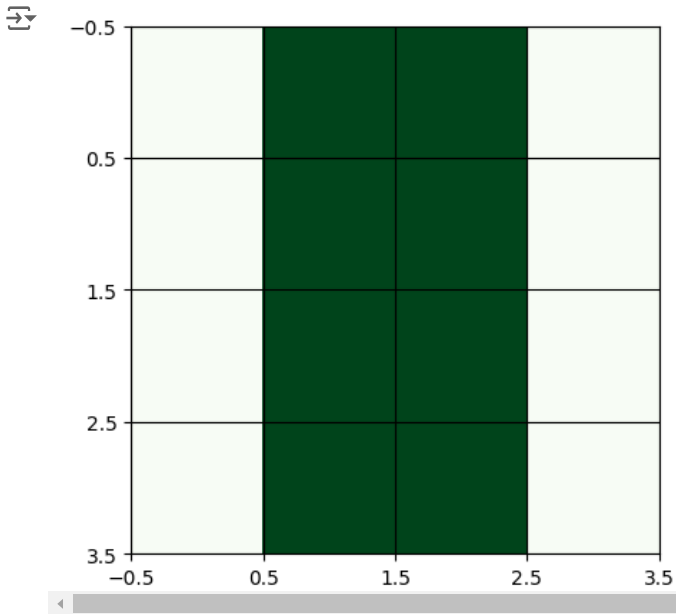
```
data = np.zeros((4, 4))
data[:, [0, 3]] = 1
visualize_array(data)
```



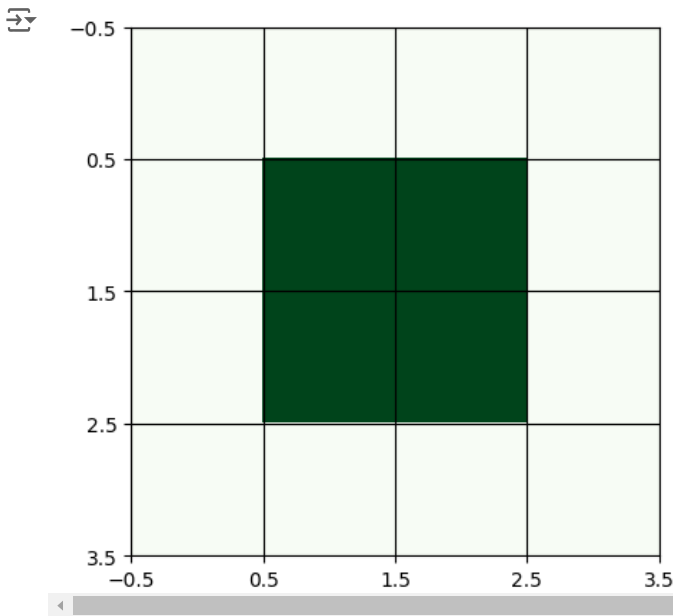
```
data = np.zeros((4, 4))
data[[1,3] , [0, 3]] = 1
visualize_array(data)
```



```
data = np.zeros((4, 4))  
data[:, 1:3] = 1  
visualize_array(data)
```




```
data = np.zeros((4, 4))  
data[1:3 , 1:3] = 1  
visualize_array(data)
```



## ▼ Pandas

```
df = pd.read_csv('https://raw.githubusercontent.com/holtzy/data_to_viz/master/Example_dataset/10_0')
print(df.head(10))
```



	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
5	25.29	4.71	Male	No	Sun	Dinner	4
6	8.77	2.00	Male	No	Sun	Dinner	2
7	26.88	3.12	Male	No	Sun	Dinner	4
8	15.04	1.96	Male	No	Sun	Dinner	2
9	14.78	3.23	Male	No	Sun	Dinner	2

```
# Calculate total bill, total tips, statistic value of them
```

```
total_bill = df['total_bill'].sum()
total_tips = df['tip'].sum()
```

```
print(f"Total bill: {total_bill}")
print(f"Total tips: {total_tips}")
```

```
bill_stats = df['total_bill'].describe()
tips_stats = df['tip'].describe()
```

```
print("\nBill statistics:")
print(bill_stats)
```

```
print("\nTips statistics:")
```

```
print(tips_stats)
```

```
➞ Total bill: 4827.77  
Total tips: 731.5799999999999
```

Bill statistics:

```
count    244.000000  
mean      19.785943  
std        8.902412  
min        3.070000  
25%       13.347500  
50%       17.795000  
75%       24.127500  
max       50.810000  
Name: total_bill, dtype: float64
```

Tips statistics:

```
count    244.000000  
mean      2.998279  
std        1.383638  
min        1.000000  
25%        2.000000  
50%        2.900000  
75%        3.562500  
max       10.000000  
Name: tip, dtype: float64
```

```
# Number of man/women    book the table?
```

```
man_count = df[df['sex'] == 'Male']['sex'].count()  
woman_count = df[df['sex'] == 'Female']['sex'].count()  
  
print(f"Số lượng người đàn ông đã đặt bàn: {man_count}")  
print(f"Số lượng phụ nữ đã đặt bàn: {woman_count}")
```

```
➞ Số lượng người đàn ông đã đặt bàn: 157  
Số lượng phụ nữ đã đặt bàn: 87
```

```
# Probability that a women will book a table?
```

```
total_count = df['sex'].count()  
probability_women_booking = woman_count / total_count  
print(f"Xác suất một người phụ nữ sẽ đặt bàn: {probability_women_booking}")
```

```
➞ Xác suất một người phụ nữ sẽ đặt bàn: 0.35655737704918034
```

```
# Number of smoker/ non-smoker
```

```
smoker_count = df[df['smoker'] == 'Yes']['smoker'].count()  
non_smoker_count = df[df['smoker'] == 'No']['smoker'].count()  
  
print(f"Số lượng người hút thuốc: {smoker_count}")  
print(f"Số lượng người không hút thuốc: {non_smoker_count}")
```

```
➞ Số lượng người hút thuốc: 93  
Số lượng người không hút thuốc: 151
```



```
# Probability that a smoker will book a table?
```

```
total_count = df['smoker'].count()
probability_smoker_booking = smoker_count / total_count
print(f"Xác suất một người hút thuốc sẽ đặt bàn: {probability_smoker_booking}")
```

➡ Xác suất một người hút thuốc sẽ đặt bàn: 0.38114754098360654

```
# Probability that a smoker is a women will book a table
```

```
women_smoker_count = df[(df['sex'] == 'Female') & (df['smoker'] == 'Yes')]['sex'].count()
probability_women_smoker_booking = women_smoker_count / total_count
print(f"Xác suất một người phụ nữ hút thuốc sẽ đặt bàn: {probability_women_smoker_booking}")
```

➡ Xác suất một người phụ nữ hút thuốc sẽ đặt bàn: 0.13524590163934427

```
# Classify by time
```

```
lunch_count = df[df['time'] == 'Lunch']['time'].count()
dinner_count = df[df['time'] == 'Dinner']['time'].count()
```

```
print(f"Số lượng người ăn trưa: {lunch_count}")
print(f"Số lượng người ăn tối: {dinner_count}")
```

```
# Probability that a people will have lunch
```

```
probability_lunch = lunch_count / total_count
print(f"Xác suất một người sẽ ăn trưa: {probability_lunch}")
```

```
# Probability that a people will have dinner
```

```
probability_dinner = dinner_count / total_count
print(f"Xác suất một người sẽ ăn tối: {probability_dinner}")
```

➡ Số lượng người ăn trưa: 68  
Số lượng người ăn tối: 176  
Xác suất một người sẽ ăn trưa: 0.2786885245901639  
Xác suất một người sẽ ăn tối: 0.7213114754098361