

Bar plot

Bar plot mô tả dạng dữ liệu phân loại như:

- Lượng mưa theo tháng
- Số dân ở các quốc gia
- Lương trung bình theo ngành nghề

Thư viện matplotlib hỗ trợ vẽ bar plot sử dụng hàm `bar()` với các thuộc tính:

- `x`: tọa độ các bar hoặc danh sách các category
- `height`: giá trị tương ứng của mỗi bar
- `width`: độ rộng của bar, mặc định 0.8
- `bottom`: tọa độ Oy của các bar (mặc định bằng 0)
- `color`: chọn màu cho các bar
- `orientation`: hướng các bar (ngang/dọc)

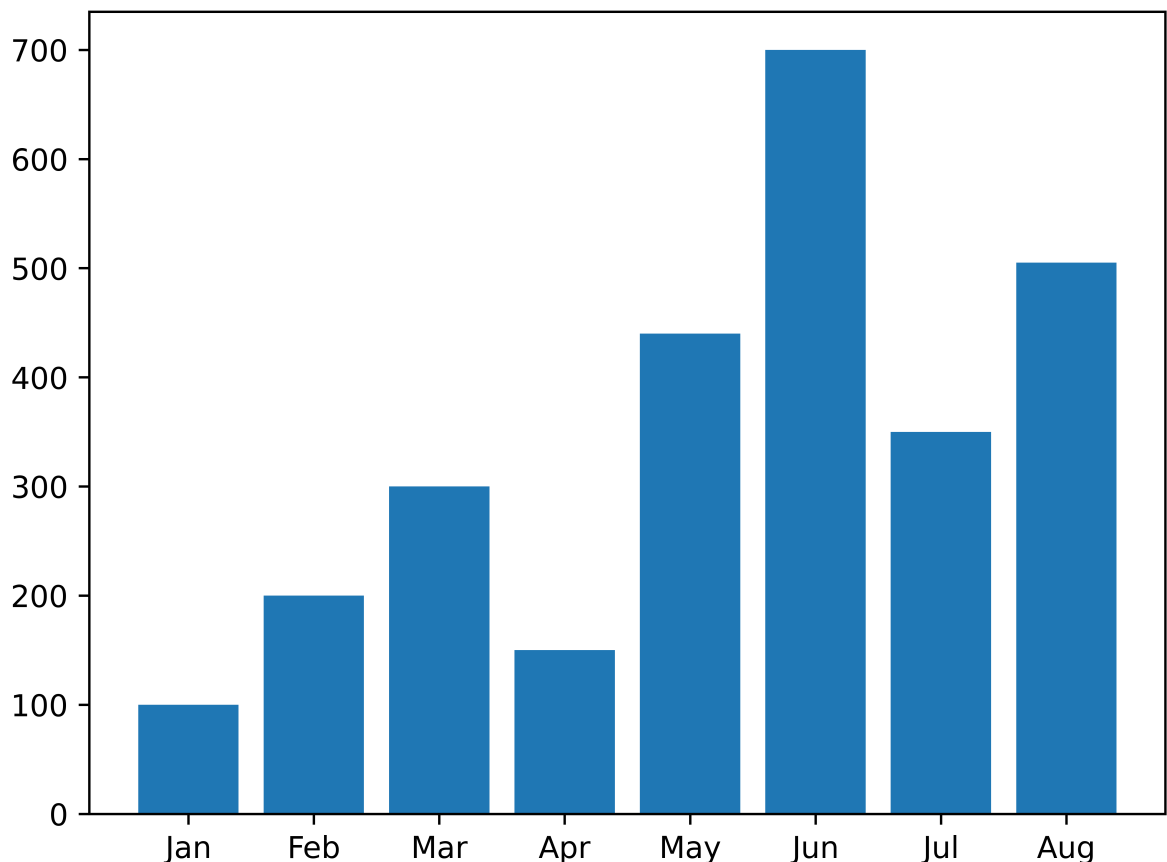
```
In [1]: import matplotlib.pyplot as plt

fig, axes = plt.subplots(dpi=1000)

label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]
values = [100, 200, 300, 150, 440, 700, 350, 505]

axes.bar(label, values)
```

Out[1]: <BarContainer object of 8 artists>



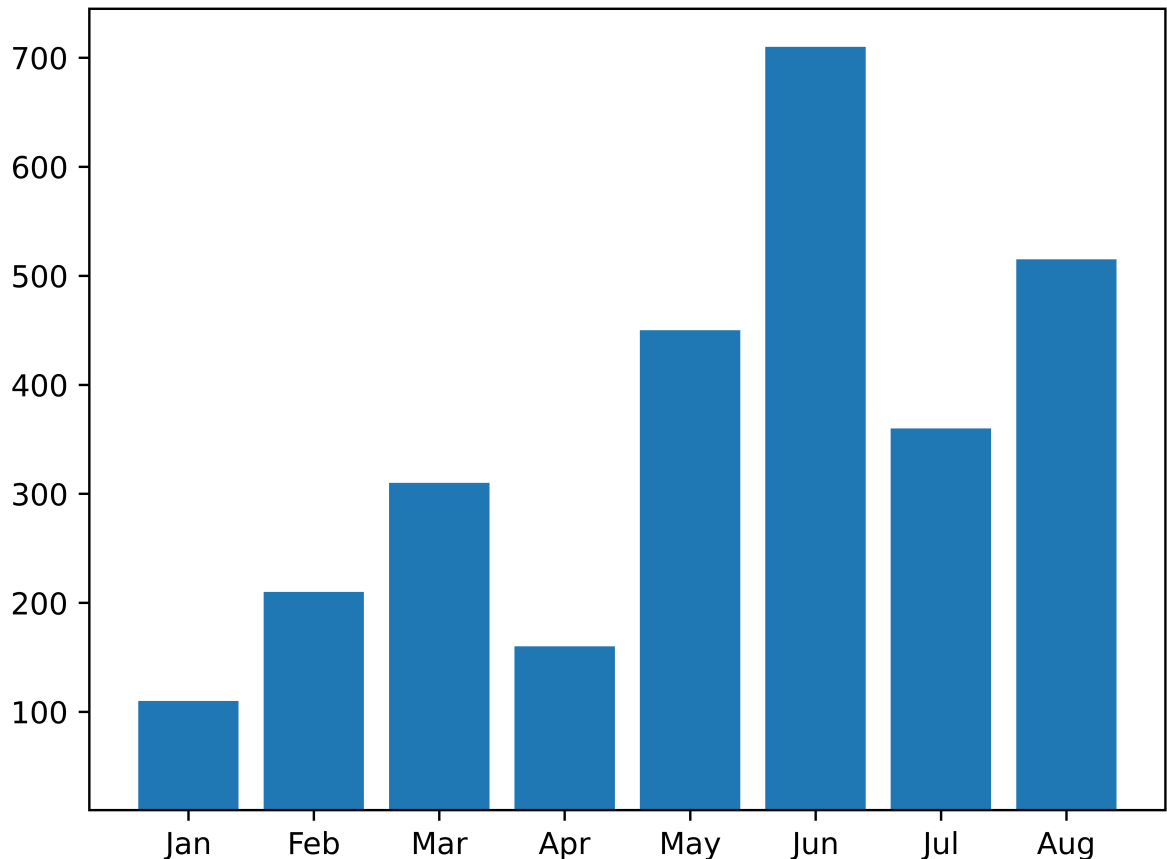
```
In [2]: import matplotlib.pyplot as plt

fig, axes = plt.subplots(dpi=1000)

label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]
values = [100, 200, 300, 150, 440, 700, 350, 505]

axes.bar(label, values, bottom=10)
```

Out[2]: <BarContainer object of 8 artists>



Vẽ nhiều cột trên cùng một label. Ví dụ muốn biểu diễn cả số tiền điện và tiền nước theo tháng trên cùng một biểu đồ

```
In [5]: import matplotlib.pyplot as plt
import numpy as np

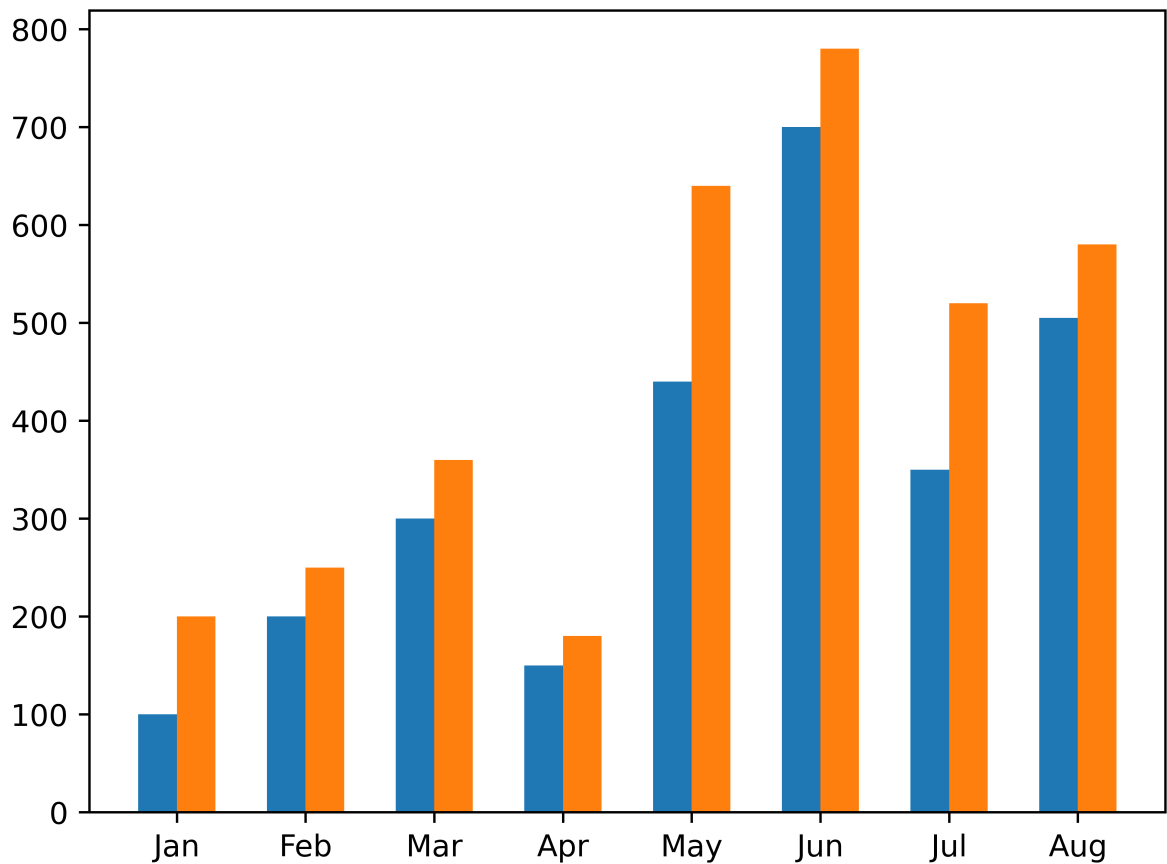
fig, axes = plt.subplots(dpi=800)

label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]
index = np.arange(len(label))
electricity = [100, 200, 300, 150, 440, 700, 350, 505]
water = [200, 250, 360, 180, 640, 780, 520, 580]

axes.bar(index, electricity, width=0.3)
axes.bar(index + 0.3, water, width=0.3)

axes.set_xticks(index + 0.15)
axes.set_xticklabels(label)

plt.show()
```



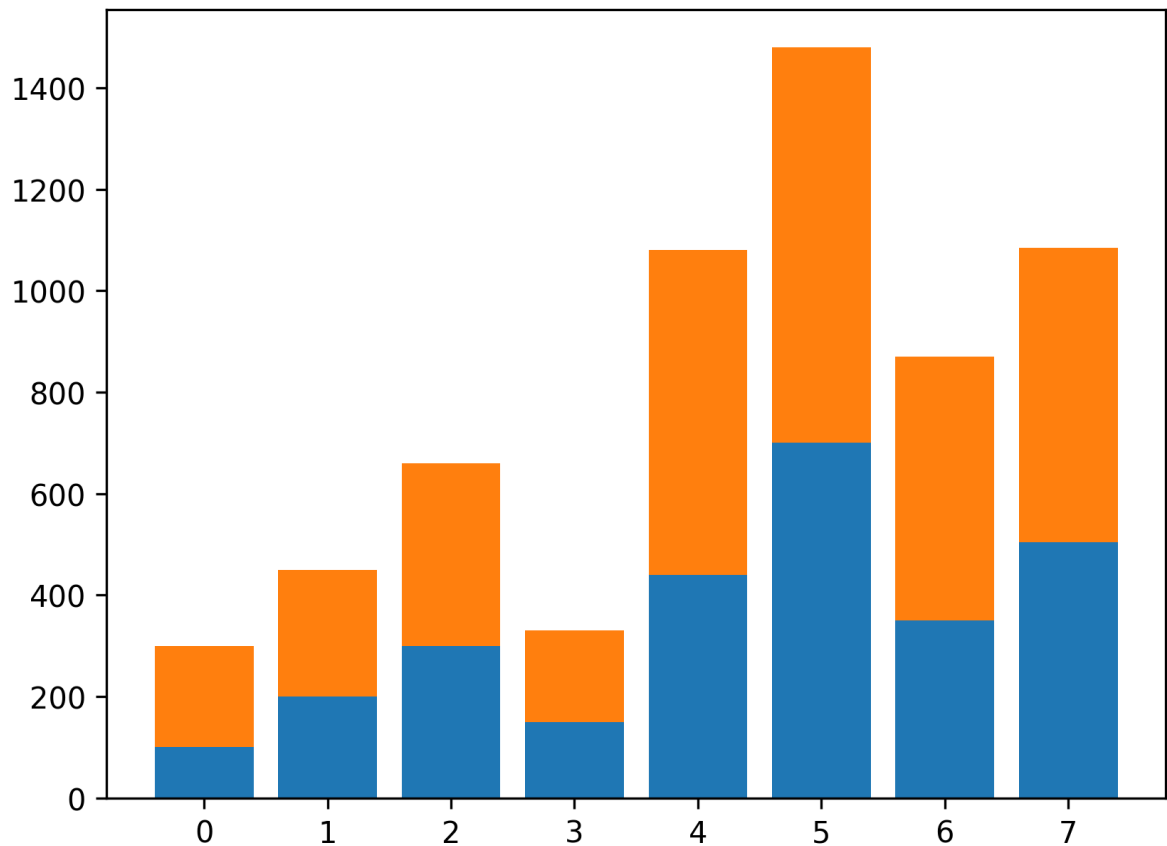
Vẽ stack bar mục đích xếp chồng các cột dữ liệu lên nhau

```
In [6]: import matplotlib.pyplot as plt
import numpy as np

fig, axes = plt.subplots(dpi=300)
label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]

index = np.arange(len(label))
electricity = [100, 200, 300, 150, 440, 700, 350, 505]
water = [200, 250, 360, 180, 640, 780, 520, 580]
axes.bar(index, electricity)
axes.bar(index, water, bottom=electricity) # ở đây bottom được sử dụng để c

plt.show()
```



Vẽ horizontal bar nghĩa là quay biểu đồ nằm ngang. Matplotlib có thể sử dụng hàm `barh` để vẽ horizontal bar, trong đó thuộc tính `bottom` đổi tên thành `left`

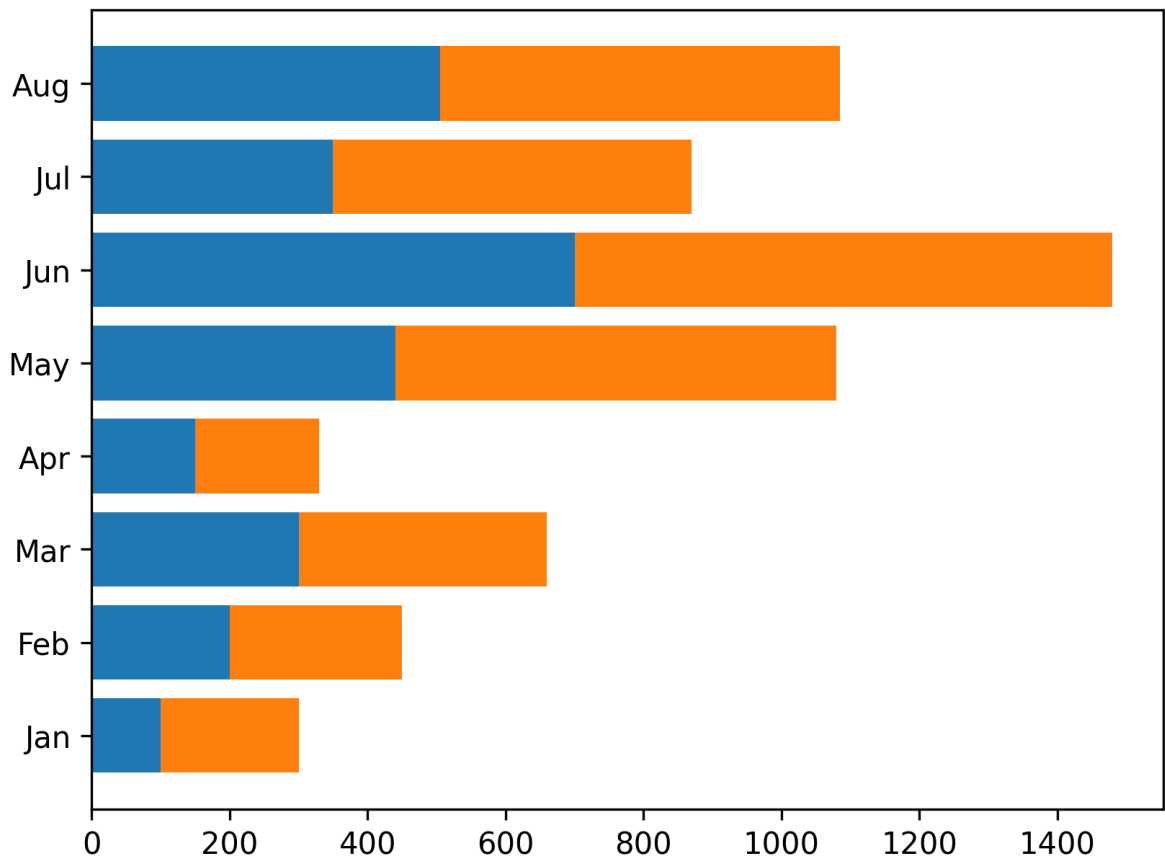
```
In [8]: import matplotlib.pyplot as plt
import numpy as np

fig, axe = plt.subplots(dpi=300)

label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]

index = np.arange(len(label))
electricity = [100, 200, 300, 150, 440, 700, 350, 505]
water = [200, 250, 360, 180, 640, 780, 520, 580]
axe.barh(index, electricity)
axe.barh(index, water, left=electricity)

axe.set_yticks(index)
axe.set_yticklabels(label)
plt.show()
```



Để biểu diễn sai số hoặc sự thay đổi data để đánh giá số liệu, ta sử dụng hàm `errorbar()` trong matplotlib như sau:

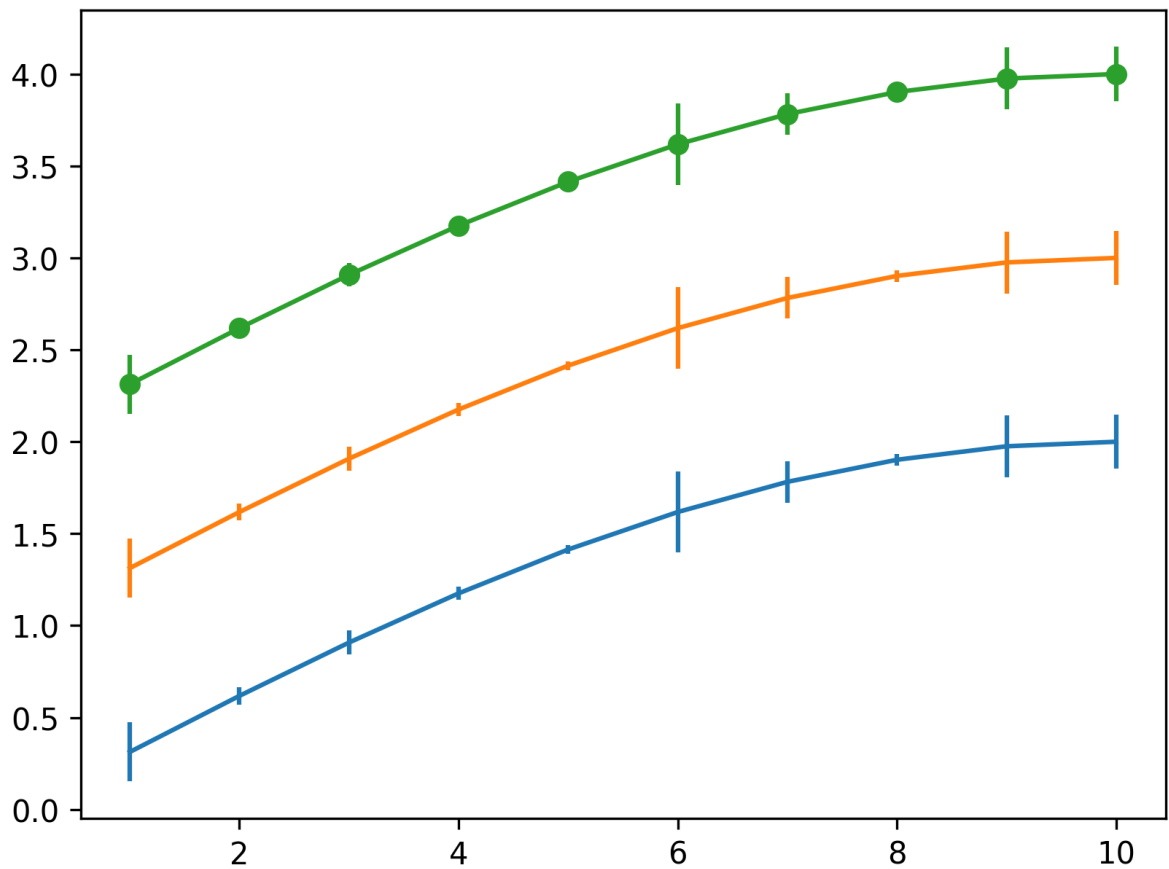
```
In [13]: import matplotlib.pyplot as plt
import numpy as np

x = np.linspace(1, 10, num=10)
y = 2 * np.sin(x/20 * np.pi)
yerr = np.random.normal(0.1, 0.1, 10)

fig, axes = plt.subplots(dpi=300)

axes.errorbar(x, y, yerr=yerr)
axes.errorbar(x, y+1, yerr=yerr)
axes.errorbar(x, y+2, yerr=yerr, fmt="-o")
```

Out[13]: <ErrorbarContainer object of 3 artists>



Đối với bar plot, muốn biểu diễn thay đổi của dữ liệu, ta sử dụng thuộc tính `yerr` trong hàm `bar()` như sau:

```
In [12]: import matplotlib.pyplot as plt
import numpy as np

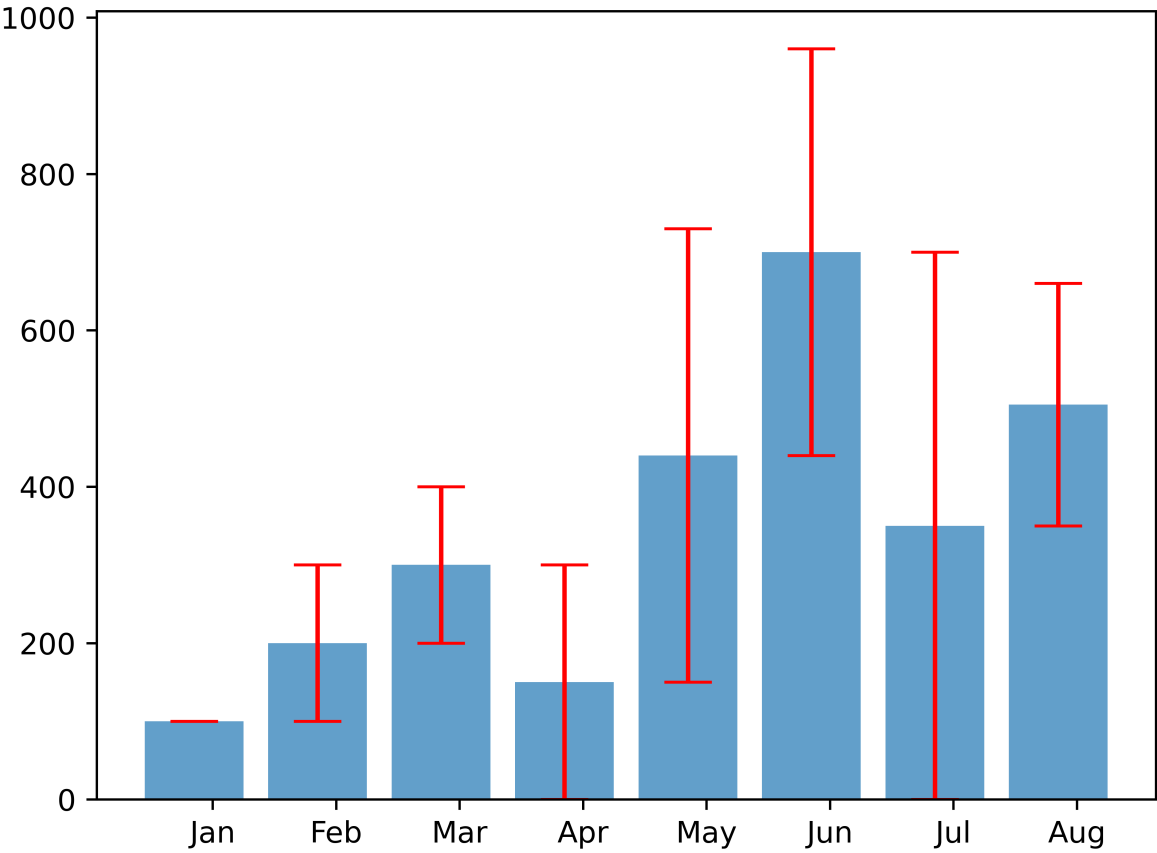
fig, axes = plt.subplots(dpi=800)

label = ["Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug"]
index = np.arange(len(label))
electricity = [100, 200, 300, 150, 440, 700, 350, 505]
electricity_err = [0, 100, 100, -150, 290, 260, -350, 155]

axes.bar(np.arange(0, len(label)), electricity, label=label,
         yerr=electricity_err, alpha=0.7, ecolor='r', capsize=8)

axes.set_xticks(index + 0.15)
axes.set_xticklabels(label)

plt.show()
```



In []: