

# 习题2.1 ndarray 快速入门

这里是2.1节的习题。我们先在这里 import 所有需要用到的库。

```
import numpy as np
import pandas as pd

arr1 = np.load("data/1-task-arr1.npy")
arr2 = np.load("data/1-task-arr2.npy")

iris_df = pd.read_csv("data/iris.csv")
```

## 基础练习

1. 请创建形状为 (3, 4, 5) 的全0, 全1, 全2 数组.

# 请在这里输入你的代码

2. 请创建长度为16的一维数组, 其中数据为 24,22,20,18...,-6.

# 请在这里输入你的代码

3. 请获取数组 arr1 的维度、形状与类型.

# 请在这里输入你的代码

4. 请获取 arr2 的简略字符串与详细字符串.

# 请在这里输入你的代码

5. 请计算将 arr2 中的每个元素加100的平方根后的数组, 并求出它的最大值、均值与标准差.

# 请在这里输入你的代码

6. 请遍历 arr2 中的所有元素并存储到一个集合 set 中.

# 请在这里输入你的代码

7. 请获取一个数值为1 ~ 9, 形状为 (3, 3)的数组, 将其存储至 arr3 中.

# 请在这里输入你的代码

8. 请计算 arr2 与 arr3 在第一个轴上的堆叠, arr2 在 arr3 之前.

# 请在这里输入你的代码

## 实战演练

这里我们使用了经典的鸢尾花数据集 (Iris Dataset)。已知数组 `iris_arr` 存储的是 `iris_df` 的特征, 即其中第2-5列的数据(按照顺序), 数组 `iris_index` 存储的是这四列的名称, `iris_target` 存储的是目标特征, 即 `species` 列中的数据。请完成下述问题。

```
iris_arr = iris_df.to_numpy()[:,1:5]
iris_index = iris_df.columns[1:5].to_list().copy()
iris_target = iris_df.to_numpy()[:, 5].copy()
iris_df.head()
```

```
.dataframe tbody tr th {
    vertical-align: top;
}

.dataframe thead th {
    text-align: right;
}
```

	Unnamed: 0	sepal_length	sepal_width	petal_length	petal_width	species
0	0	5.1	3.5	1.4	0.2	Iris-setosa
1	1	4.9	3.0	1.4	0.2	Iris-setosa
2	2	4.7	3.2	1.3	0.2	Iris-setosa
3	3	4.6	3.1	1.5	0.2	Iris-setosa
4	4	5.0	3.6	1.4	0.2	Iris-setosa

1. 请计算每个特征的均值和方差

```
# 请在这里输入你的代码
```

2. 请将 `iris` 数据标准化 (均值为0, 方差为1)

```
# 请在这里输入你的代码
```

3. 请计算每个类别的平均特征值

# 请在这里输入你的代码

# 提示: 可以通过`np.unique(arr)`获取所有不同的值