C4

NumPy Basics

CHAPTER ASSIGNMENT

Name _____ Class _____ No. _____ Date _____

Answer every question in the space given. Show your working. You may use the chapter page to help you.

Questions from the Guru Jazzy chapter page (based on *Python for Data Analysis*, 3rd edition, Wes McKinney, O'Reilly).

4.1 The ndarray

1. Time [x * 2 for x in range(1_000_000)] against np.arange(1_000_000) * 2 using %timeit in a notebook. How many times faster is NumPy on your machine?

2. Make a 3×4 array of the numbers 1 to 12 with arange and reshape. Print its shape, dtype, and the array divided by 2.

4.2 Indexing and Slicing

1. Repeat the example with arr_slice = arr[5:8].copy(). Show that the original is now untouched.

```
arr = np.arange(10)
arr_slice = arr[5:8]
arr_slice[1] = 12345
print(arr)            # the original changed too
```

2. From a 4×4 array of 0–15, pull out: the last row, the middle 2×2 block, and the first column as a 1-D array.

4.3 Boolean and Fancy Indexing

1. Use the mask explorer below with each expression. Before pressing Select, predict how many rows or values come back.

```
names = np.array(["Bob", "Joe", "Will", "Bob", "Will", "Joe", "Joe"])
data = np.array([[4, 7], [0, 2], [-5, 6], [0, 0], [1, 2], [-12, -4], [3, 4]])

data[names == "Bob"]
data[~(names == "Bob")]
data[(names == "Bob") | (names == "Will")]
data[names == "Bob", 1]
data[data < 0]
data[data > 3]
```

2. Write the mask for "rows that are not Joe" two ways: with != and with ~. Check they select the same rows.

3. Set every value in data greater than 5 to 5, using one line with a mask.

4.4 Random Numbers and ufuncs

1. Simulate 20 dice rolls with rng.integers(1, 7, size=20). Run it twice with the same seed, then with no seed. What changes?

2. Make an array of 10 random normals and use np.abs, np.round and np.maximum(arr, 0) on it. Describe what each did.

4.5 Array-Oriented Programming

1. Given an array of test scores, use np.where to make a matching array of "pass" or "fail" (pass is 50 or more).

2. For a 5×3 array of random integers, find the mean of each column, the max of each row, and how many values are above 10.

4.6 Files, Linear Algebra and Random Walks

1. Simulate 5,000 walks at once with a 2-D array (size=(5000, 1000)) and cumsum(axis=1). How many ever reach 30 or -30?

2. Solve the equations 2x + y = 5 and x - y = 1 with np.linalg.solve.

End of assignment · Check your answers against the chapter at gurujazzy.com/subjects/books/python-for-data-analysis/chapter-4-numpy-basics/