KNOWLEDGE CHECKS — DAT208x

Module 1

Lecture: Python Basics
Review Question 1
Which of the following statements is correct?
Answer : The IPython Shell is typically used to work with Python interactively.

Review Question 2
Which file extension is used for Python script files?
Answer : .py

Review Question 3
You need to print the result of adding 3 and 4 inside a script. Which line of code should you write in the script?
Answer : print(3 + 4)

Lecture: Variables and Types
Review Question 1
Which line of code creates a variable x with the value 15?
Answer : x = 15

Review Question 2
What is the value of the variable z after executing these commands?
x = 5
y = 7
z = x + y + 1
Answer : 13

Review Question 3
You execute the following two lines of Python code:
x = "test"
y = False
Which of the following statements is correct?
Answer : x is a string, and y is a boolean.

Module 2

Lecture: Python Lists
Review Question 1
Which of the following is a characteristic of a Python list?
Answer : It is a way to name a collection of values, instead of having to create separate variables for each element.

Review Question 2
You use type(fam) and type(fam2) to reveal the type of a Python list. What is the result?
Answer : list

Review Question 3
Which command is invalid Python syntax to create a list x?
Answer : x = ["this", "is", "a" True "list"]

Review Question 4
Which three Python data types does this list contain?
x = ["you", 2, "are", "so", True]

Answer : int, str, bool

Lecture: Subsetting Lists
Review Question 1
Which pair of symbols do you need to do list subsetting in Python?
Answer : Brackets: []

Review Question 2
What Python command should you use to extract the element with index 1 from a Python list x?
More specifically, given a list x = [4, 5, 6, 7], which command do you need to extract the number 5?
Answer : x[1]

Review Question 3
You have a list y, containing 5 elements:
y = ["this", "is", "a", True, "test"]
Which two Python commands correctly extract the boolean value from this list?
Answer : y[3] and y[-2]

Review Question 4
You want to slice a list x. The general syntax is:
x[begin:end]
Answer : You need to replace begin and end with indexes according to the slice you want to make.

Which of the following statements is correct?
Answer : The begin index is included in the slice, the end index is not

Lecture: Manipulating Lists
Review Question 1
You have a list x that is defined as follows:
x = ["a", "b", "b"]
Answer : You need to change the second "b" (the third element) to "c".

Which command should you use?
Answer : x[2] = "c"

Review Question 2
You have a list x that is defined as follows:
x = ["a", "b", "c"]
Which line of Python code do you need to add "d" at the end of the list x?
Answer : x = x + ["d"]

Review Question 3
You have a list x that is defined as follows:
x = ["a", "b", "c", "d"]
You decide to remove an element from it by using del:
del(x[3])
How does the list x look after this operation?
Answer : ["a", "b", "c"]

Module 3

Lecture: Functions
Review Question 1
What is a Python function?
Answer : A piece of reusable Python code that solves a particular problem.

Review Question 2
You have a list named x. To calculate the minimum value in this list, you use the min() function.
Which Python command should you use?
Answer : min(x)

Review Question 3
What Python command opens up the documentation from inside the IPython Shell for the min function?
Answer : help(min)

Review Question 4
The function round has two arguments. Select the two correct statements about these arguments.
Answer : number is a required argument. and ndigits is an optional argument.

Lecture: Methods
Review Question 1
What is append() in Python?
Answer : append() is a method, and therefore also a function.

Review Question 2
You have a string x defined as follows:
x = "monty python says hi!"
Which Python command should you use to capitalize this string x?
Answer : x.capitalize()

Review Question 3
How does the list x look after you execute the following two commands?
x = [4, 9, 5, 7]
x.append(6)
Answer : [4, 9, 5, 7, 6]

Lecture: Packages
Review Question 1
Which of the following is a package for installation and maintenance system for Python?
Answer : pip

Review Question 2
Which statement is the most common way to invoke the import machinery?
Answer : import

Review Question 3
You import Numpy as foo as follows:
import numpy as foo
Which Python command that used the array() function from Numpy is valid if Numpy is imported as foo?
Answer : foo.array([1, 2, 3])

Review Question 4
You want to use Numpy's array() function.
You need to decide whether to import this function as follows:
from numpy import array
or by importing the entire numpy package:
import numpy
Select the two correct statements about these different import methods.
Answer : The from numpy import array version will make it less clear in the code that you're using Numpy's array() function.
Using import numpy will require you to use numpy.array(), making it clear that you're using a Numpy function.

Module 4

Lecture: Numpy
Review Question 1
Which Numpy function do you use to create an array?
Answer : array()

Review Question 2
Which two statements describe the advantage of Numpy Package over regular Python Lists?
Answer : The Numpy Package provides the array, a data type that can be used to do element-wise calculations.
Because Numpy arrays can only hold element of a single type,
calculations on Numpy arrays can be carried out way faster than regular Python lists.

Review Question 3
What is the resulting Numpy array z after executing the following lines of code?
```
import numpy as np
x = np.array([1, 2, 3])
y = np.array([3, 2, 1])
z = x + y
```
Answer : array([4, 4, 4])

Review Question 4
What happens when you put an integer, a Boolean, and a string in the same Numpy array using the array() function?
Answer : All array elements are converted to strings.

Lecture: 2D Numpy Arrays
Review Question 1
What charaterizes multi-dimensional Numpy arrays?
Answer : You can create a 2D Numpy array from a regular list of lists.

Review Question 2
You created the following 2D Numpy array, x
:
```
import numpy as np
x = np.array([["a", "b", "c", "d"],
              ["e", "f", "g", "h"]])
```
Which Python command do you use to select the string "g" from x?
Answer : x[1,2]

Review Question 3
What does the resulting array z contain after executing the following lines of Python code?

```
import numpy as np
x = np.array([[1, 2, 3],
              [1, 2, 3]])
y = np.array([[1, 1, 1],
              [1, 2, 3]])
z = x - y
Answer : array([[0, 1, 2],
                [0, 0, 0]])
```

Lecture: Basic Statistics with Numpy
Review Question 1
Which of the following statement about basic statistics with Numpy is correct?
Note: assume that the Numpy package is imported as np.
Answer : Numpy offers many functions to calculate basic statistics, such as np.mean(), np.median() and np.std().

Review Question 2
You are writing code to measure your travel time and weather conditions to work each day.
The data is recorded in a Numpy array where each row specifies the measurements for a single day.
The first column specifies the temperature in Fahrenheit. The second column specifies the amount of travel time in minutes.
The following is a sample of the code.

```
import numpy as np
x = np.array([[28, 18],
              [34, 14],
              [32, 16],
              ...
              [26, 23],
              [23, 17]])
```

Which Python command do you use to calculate the average travel time?
Answer : np.mean(x[:,1])

Review Question 3
As a wrap up, have a look at the statements below about Numpy in general. Select the three statements that hold.
Answer : Numpy is a great alternative to the regular Python list if you want to do Data Science in Python.
Numpy arrays can only hold elements of the same basic type.
Next to an efficient data structure, Numpy also offers tools to calculate summary statistics and to simulate statistical distributions.

Module 5

Lecture: Basic Plot with matplotlib
Review Question 1
What is the characteristic about data visualization?
Answer : Visualization is a very powerful tool for exploring your data and reporting results.

Review Question 2
What is the conventional way of importing the pyplot sub-package from the matplotlib package?
Answer : import matplotlib.pyplot as plt

Review Question 3

You are creating a line plot using the following code:
a = [1, 2, 3, 4]
b = [3, 9, 2, 6]
plt.plot(a, b)
plt.show()
Which two options describe the result of your code?
Answer : The values in a are mapped onto the horizontal axis.
The values in b are mapped onto the vertical axis.

Review Question 4
You are modifying the following code that calls the plot() function to create a line plot:
a = [1, 2, 3, 4]
b = [3, 9, 2, 6]
plt.plot(a, b)
plt.show()
What should you change in the code to create a scatter plot instead of a line plot?
Answer : Change plot() in plt.plot() to scatter().

Lecture: Histograms
Review Question 1
What is a characteristic of a histogram?
Answer : Histogram is a great tool for getting a first impression about the distribution of your data.

Review Question 2
You are working with a Python list with 10 different values. You divide the values into 5 equally-sized bins.
How wide will these bins be if the lowest value in your list is 0 and the highest is 20?
Answer : 4

Review Question 3
You write the following code:
import matplotlib.pyplot as plt
x = [1, 3, 6, 3, 2, 7, 3, 9, 7, 5, 2, 4]
plt.hist(x)
plt.show()
You need to extend the plt.hist() command to specifically set the number of bins to 4. What should you do?
Answer : Add a second argument to plt.hist(): plt.hist(x, bins = 4)

Lecture: Customization
Review Question 1
You are customizing a plot by labelling its axes. You need to do this by using matplotlib.
Which code should you use?
Answer : xlabel("x-axis title") and ylabel("y-axis title")

Review Question 2
Which matplotlib function do you use to build a line plot where the area under the graph is colored?
Answer : fill_between()

Review Question 3
Typically, you place all customization commands between the plot() call and the show() call, as follows:
import matplotlib.pyplot as plt
x = [1, 2, 3]
y = [4, 5, 6]
plt.plot(x, y)

```
# customization here
plt.show()
```
What will happen if you place the customization code after the show() function instead?
```
import matplotlib.pyplot as plt
x = [1, 2, 3]
y = [4, 5, 6]
plt.plot(x, y)
plt.show()
# customization here
```
Answer : Python doesn't throw an error, but you won't see your customizations.
The show() function displays the plot you've built up until then.
If the customizations come afterwards, there is no effect on the shown output.

Module 6

Lecture: Boolean Logic and Control Flow
Review Question 1
What is the result of the following comparison?
```
5 >= 5
```
Answer : True

Review Question 2
What is the result of the following comparison?
```
4 != 4
```
Answer : False

Review Question 3
What is the characteristic of the and operator?
Answer : The and operator returns True only if both operands are True.

Review Question 4
What is the characteristic of the or operator?
Answer : The or operator returns True if at least one of the operands is True.

Review Question 5
You write the following code:
```
x = 5
if x > 6 :
    print("high")
else :
    print("low")
```
What will be printed out if you execute the code?
Answer : low

Review Question 6
You write the following code:
```
x = 7
if x > 6 :
    print("high")
elif x > 3 :
    print("ok")
else :
    print("low")
```
What will be printed out if you execute the code?
Answer : high

Lecture: Pandas
Review Question 1
How is a Pandas DataFrame different from a 2D Numpy array?
Answer : In Pandas, different columns can contain different types.

Review Question 2
What are two characteristics that describe Pandas DataFrame?
Answer : The rows correspond to observations.
The columns correspond to variables.

Review Question 3
Which Pandas function do you use to import data from a comma-separated value (CSV) file into a Pandas DataFrame?
Answer : read_csv()

Review Question 4
Which technique should you use to select an entire row by its row label when accessing data in a Pandas DataFrame?
Answer : loc