**CSCI 1360 Final Exam Study Questions**

**I: Python Language Basics**

* What kind of language is Python? What are its strengths and weaknesses?
* What kind of variable is “something”? What kind of variable is 5? What kind of variable is True?
* What is it called to convert the same value between different variable types?
* What is an operator? What is the “=” operator called, and what does it do?
* What is the type of the output of 6 / 2 ? What is the type of the output of 5 + 5 ? What is the type of the output of x < 5 ?
* What kind of typing mechanism does Python use? What other kinds of typing mechanisms are there in other languages?
* What are the different types of variable scope?
* What happens when you have two variables in different scopes that have the same names?

**II: Data Structures**

* What are the similarities and differences between lists, sets, tuples, and dictionaries?
* What is a comprehension? Show how you would create a list comprehension of all even numbers between 0 and 10, inclusive.
* Show how you would test if a certain value x is in a set s without using any loops.
* What is a generator? When would you use one?
* How can you access the 0th element of a list? What about the last element? What about every other element, starting with the 0th? What about all the elements except the first and last?

**III: Loops, Conditionals, Exceptions, Files**

* What does the zip() function do? When would you want to use it?
* What does the enumerate() function do? When would you want to use it?
* How would you compute the largest element of a list? The smallest element? The median element?
* When would you use a “break” statement? What about a “continue” statement?
* What kind of loop would you use to repeat a task a specific number of times? What kind of loop would you use to repeat a task an unknown number of times? Show how you can convert between these two kinds of loops.
* What is the different between file.read() and file.readlines() ?
* When opening up a file descriptor, you have to specify a *mode.* What is this “mode” mean? What are the possible modes?
* Show how you could write an exception handler to catch *only* SyntaxErrors. Show how you could write an exception handler to catch *any* error.
* What are some possible text-based structured file formats for storing data? What are some possible binary-based structured file formats for storing data?

**IV: Object-Oriented Programming**

* Name the critical differences between procedural and object-oriented programming.
* What is the difference between a “class” and an “object”?
* What is a constructor? What is a method? What is an attribute?
* What is it called when a class’ constructor is invoked?
* Write a class that defines a Car. Name some possible child classes and how they would differ from the parent class.
* Why is encapsulation important in object-oriented programming?

**V: Functions**

* What is a function? What are the core constituents of a function?
* What differentiates a Python function from a method?
* What are the inputs to a function called? What are the outputs called?
* How can I set some of the inputs to take on default values unless overridden by the programmer?
* I pass an integer into a function. Inside the function, I multiply the integer by 2. What value does the integer have once the function ends?
* I pass a list into a function. Inside the function, I add an element to the list. What does the list look like once the function ends?
* Describe the difference in behavior observed in the two previous bullet points.

**VI: NumPy Arrays**

* Show how you would add a constant value to every element of a Python list. Next, convert the list to a NumPy array. Show how you can add that same constant value again to every element of the NumPy array. What’s the difference?
* The number of dimensions of a NumPy array are its *axes*. How do you access this information from a NumPy array? How do you get the number of rows (1st axis)? Columns (2nd axis)? Frames (3rd axis)? Time points (4th axis)?
* You have a 2D NumPy array. Show how you can slice out the 10th *row*.
* You have a 2D NumPy array. Show how you can slice out the 10th *column*.
* You have a 3D NumPy array. Show how you can slice out the 10th *row* from the 10th *frame.*
* You have a 3D NumPy array and a 2D NumPy array. Show how you can add the 2D array to the 10th frame of the 3D array.
* Show how you can use fancy indexing to reverse the ordering of a NumPy array.
* Show how you can use masking to set any value less than 0 in a NumPy array to exactly 0.

**VII: Linear Algebra, Probability, Statistics**

* Define a scalar, a vector, and a matrix. What do vectors have that scalars do not?
* What is a “space” in linear algebra?
* What Python data structure do we use when performing vector computations? What does this data structure look like if the vector is three-dimensional?
* Differentiate between how we’ve used the term “dimensions” when referring to linear algebra, and how we’ve used “dimensions” when referring to NumPy arrays **(this is important)**.
* What determines the “order” of a summary statistic? List some first-order and second-order statistics.
* What is a “random variable”? What is its “distribution”? Name some of the popular distributions.
* Do we usually know what kind of variable a random variable is? How do we refer to the underlying process of a random variable?
* Name the three axioms of probability.
* What is independence? What is conditional probability?
* What is Bayes’ Theorem? What are the individual terms of Bayes’ Theorem?

**VIII: Natural Language Processing, Computer Vision, Machine Learning**

* Define the “bag of words” model of representing text. What implicit assumptions does it make about text?
* What are some common preprocessing techniques one would undertake before creating a bag of words model? What possible issues do these preprocessing steps mitigate?
* What is a feature vector? What is term-frequency? What is inverse-document-frequency? What is TF-IDF?
* What is the typical data structure used for representing an RGB image?
* What are some preprocessing techniques for analyzing images? What issues do they help mitigate?
* What is the goal of machine learning?
* What is supervised learning? What is unsupervised learning? Give example algorithms of both.
* What is overfitting? What is underfitting? How can you avoid overfitting?
* Define bias. Define variance. Why does decreasing one usually increase the other? Give an example of an algorithm where this tradeoff can be observed directly.

**IX: Data Science, Open Science**

* Define data science. What makes or doesn’t make Python a good language to study the field?
* Name the six subgroups of Greater Data Science and give examples of each.
* Define open science. What makes or doesn’t make Python a good language to use for open science?
* What are the advantages and disadvantages of open science?
* Name the six general principles of adhering to open science and give examples of each.