# CS 696 Intro to Big Data <br> Fall Semester, 2017 <br> Assignment 1 <br> © 2017, All Rights Reserved, SDSU \& Roger Whitney San Diego State University -- This page last updated 8/28/17 

Due Sept 8 23:59
In each of the problems below you are asked to implement a function. Make sure that your function uses the exact name given in the problem. Test cases will be run in grading your assignment. If your functions use different names the test cases will fail and you will lose points.

1. If we list all the natural numbers below 20 that are multiples of 3 or 5 but not multiples both of 3 and 5 , we get $3,5,6,9,10,12,18$. The sum of these multiples is 63 . Write a function, sum_multiples_3_5, that returns the sum of the multiples of 3 and 5 less than N.
2. Write a function, patternCount with two arguments. The first arguments is a string, lets call it text, and the second argument is also a string, call it pattern. The function patternCount returns the number of times the pattern occurs in the text. For example

$$
\begin{aligned}
& \text { pattern_count("abababa", "aba") == } 3 \\
& \text { pattern_count(""aaaaa", "aa") == } 4 \\
& \text { pattern_count("Abcde", "abc") }==0
\end{aligned}
$$

3. Write a function product(s : String) that returns the product of the Unicode codes of all letters in string s without using a loop.
4. Write a function noDuplicates(ints : Array[Int]) that returns an array all values from the array ints with duplicates removed.
5. Create an array that contains the value 0.00001 f repeated 1000000 times. Sum up all the elements of the array. What is the result? How big is the error?
6. Using Scala compute the sums $0.0001 f+8000.0 f$ and $0.000000001+90000000$. What is the result? How big is the error?
7. Write a function randomInts( $n$ :Int, range:Int) that returns an array of $n$ integers all between 0 and range.
8. Create three files with random integers in the range 0 to Int.MaxValue/2. The first file contains 1,000 integers, the second 100,000 integers, the third 10,000,000 integers. Write scala code to find the average and median value for each file. What are the average and median value for each file? How long does it take compute those values for each file?
9. Write a function minmax(values: Array[Int]) that returns a pair containing the smallest and the largest values in the array.
10. Make a class Car with read-only properties for manufacturer, model name, and model year, and a read-write property for the license plate. Supply four constructors. All require the manufacturer and model name. Optionally, model year and license plate can also be speci-
fied in the constructor. If not, the model year is set to -1 and the license plate to the empty string. Which constructor are you choosing as the primary constructor? Why?

## Grading

Each problem is worth 5 points. Four of the points are for correct solutions. One point is for style. That is did you use proper indentation, naming convention, good names and reasonable Julia constructs. Providing good tests for the problems is worth another 5 points total.

## What to turn in

You are to turn in a Jupyter Jupyter-scala notebook containing the answers to the questions above. Since Jupyter notebooks can contain text and code, before each problem indicate which problem it is in text, not in code comment.

To turn in your assignment download your Jupyter notebook as an Notebook (.ipynb). See image below. This will allow me to run your assignment in Jupyter. Do not download it as a Scala file (.ji) as this will not run in Jupyter and removes all the text (markdown).


Once you have downloaded the assignment zip it up and then upload the zip file to the course portal.

## Late Penalty

An assignment turned in 1-7 days late, will lose $5 \%$ of the total value of the assignment per day late. The eight day late the penalty will be $40 \%$ of the assignment, the ninth day late the penalty will be $60 \%$, after the ninth day late the penalty will be $90 \%$. Once a solution to an assignment has been posted or discussed in class, the assignment will no longer be accepted. Late penalties are always rounded up to the next integer value.

