CS 422: Data Mining (CRN: 201820)

Department of Computer Science
Illinois Institute of Technology
Vijay K. Gurbani, Ph.D.

Fall 2018: Homework 0 (0 points; not graded)

Due date: Wed, Aug 29, 2018 11:59:59 PM Chicago Time

1 Exercises

1.1 Describe in 2-3 sentences what you expect to achieve from this course.

2 Practicum problems

2.1 Problem 1

This exercise will ensure that you are comfortable with the mechanics of creating R notebooks and submitting the html from the notebook to Blackboard.

Create a new R notebook, name it hw0.Rmd. The first three lines of the notebook should be as follows (put your section number below as appropriate:

```
---
title: “CS 422 Section 04”
output: html_notebook
author: Me Student
---
```

1-A: Load the *cars* dataset and print it. The *cars* dataset is a built-in dataset in R. To access it, simply type “attach(cars)” from the R REPL in RStdio. You should subsequently see a “cars” object in your environment.

1-B: Plot the data in the *cars* dataset. Provide a title for the graph. The title should be “The Cars Dataset”.

1-C: Print a summary of the *cars* dataset using the R *summary()* command.

1-D: What is the maximum speed and minimum distance as shown in the *summary()* command?

2.2 Problem 2

Recall the “students.csv” file that you read in R-intro-1.r.

2-A: Read the file into a data frame and add a new row consisting of the following data to the existing rows:

name: Brad Pitt
id: 40051
gpa: 2.21
grade: C

Hint: search up *rbind()*.

Save the notebook. Generate the html file (it will be called hw0.nb.html). Add the hw0.Rmd and hw0.nb.html files into the archive and upload the archive to Blackboard before the due date.