

CSE574 Introduction to Machine Learning
Programming Assignment 3 (in R-Studio)

Unsupervised Learning

Due Date: **May 16, 2018 before midnight**

For programming assignment 3, we will read *An Introduction to Statistical Learning with Applications in R* (ISLR) chapter 10 *Unsupervised Learning*; and, replicate the Lab sections. Chapter 10 of ISLR is available on Piazza.

TASK 1 Lab 1: Principal Components Analysis, 30 points

Replicate section 10.4, turn in output (R-markdown output is probably easiest; pdf, docx, or html is fine)

TASK 2 Lab 2: Clustering, 30 points

Replicate section 10.5, turn in output (pdf, docx, or html is fine)

TASK 3 Lab 3: NCI60 Data example, 30 points

Replicate section 10.6, turn in output (pdf, docx, or html is fine)

TASK 4 Summary 10 points

Report time required for homework, difficulties encountered, and collaborators.

Submission requirements

Each student should submit a zip file name named <ubperson#>-project3.zip

The zip file should contain a single file (pdf, docx, html) containing your output for Task 1-3, plus your comments for Task 4.

Remember: submit your own work; collaboration is ok; accessing libraries / other resources on web is ok. The report must be your own, meaning: if we ask you about your work, you must be able to explain it in detail to us.