EDA Exercise 2

P. McDonald

August 25, 2016

There is a great deal of data available in the form of Extensible Markup Language (xml) files. There is a good reference for xml files:

https://www.sitepoint.com/really-good-introduction-xml/

Navigate to http://www.cs.washington.edu/research/xmldatasets/www/repository.html#auctions

Prepare an R Markdown document which documents the use of R tools/code to address the following

1. Download the 277KB xml file associated to Reed courses.

2. How many distinct subjects are listed in the document?

3. How many courses have a NULL instructor listing?

4. How many distinct instructors are listed?