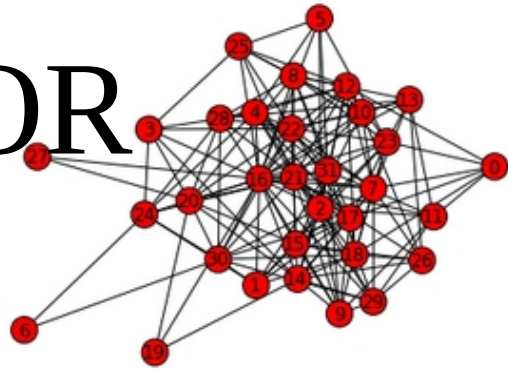


DATA EXPLORATION06



These are powerful new tools; used today to answer the unanswerable and start life-changing companies. One part math, one part programming, and a whole lot of creativity: data science is an art that cannot be learned from a textbook.

Dataset: GratefulDeadData.csv. This is a webscraping and categorization exercise. The csv file consists of setlists for every Grateful Dead concert from 1972 to 1993. The data was webscraped using the requests and BeautifulSoup libraries from <http://www.cs.cmu.edu/~./gdead/setlists.html>

- 1. Use BeautifulSoup to webscrape the data.** Data and webscraping code will be released on Sunday
- 2. Once the data is in python the depth of analysis will vary.** Test a model to predict shows from the 80's. From year 1987. Are there issues with using a standard KFold if testing a model from only one year? Develop a multi-class model for classifying shows into each decade. Is there another way besides date that the data can be grouped?
- 3. Data Visualization.**