David Kinney DSC 550

Professor Deitenbeck Winter 2019

5.4 Reflection: Week 5 Reflection

Week 5 was pretty enlightening, as I had never encountered this type of graph analysis before. Well, I should clarify that—apparently, I *had* come across this before, I just didn’t know that is what it is called. In Joel Grus’s book, “Data Science from Scratch” (Grus 2015 O’Reilly Media), he opens up with the following case study: you are a new Data Scientist at a hypothetical company, given the task of identifying “key connectors”—such as who is friends with whom—between the employees at the company. I made the mental connection, post facto, that this is representative of graph analysis.

users = [

{ "id": 0, "name": "Hero" },

{ "id": 1, "name": "Dunn" },

{ "id": 2, "name": "Sue" },

{ "id": 3, "name": "Chi" },

{ "id": 4, "name": "Thor" },

{ "id": 5, "name": "Clive" },

{ "id": 6, "name": "Hicks" },

{ "id": 7, "name": "Devin" },

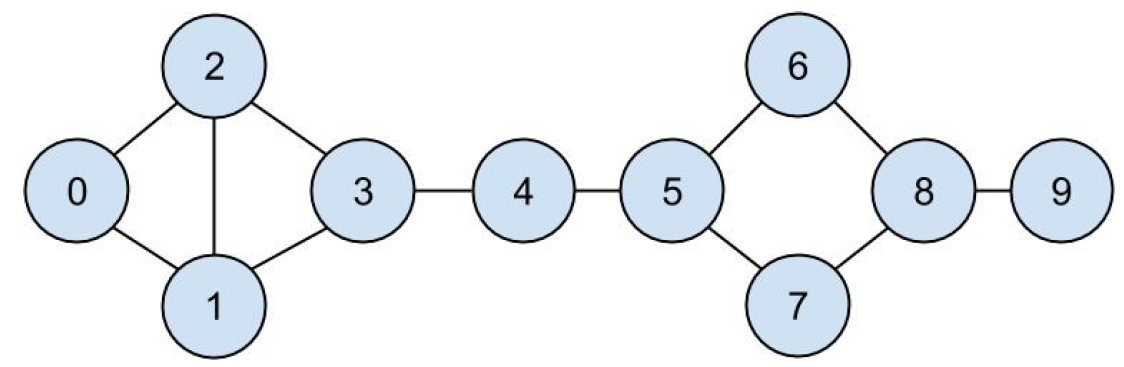
{ "id": 8, "name": "Kate" },

{ "id": 9, "name": "Klein" }

]

friendships = [(0, 1), (0, 2), (1, 2), (1, 3), (2, 3), (3, 4),

(4, 5), (5, 6), (5, 7), (6, 8), (7, 8), (8, 9)]



I get the concept now, but unlike descriptive/representation graphing, I don’t have any experience with the popular libraries (NetworkX and Graph-tool). I recognize that I need to set aside some time to embrace at least one of these and gain proficiency. As our textbook, “Applied Text Analysis” (Bengfort et. al., 2018, O’Reilly Media) recommends NetworkX as being somewhat simpler to grasp. I will start there.