Assignment: Centrality measures can be used to predict (positive or negative) outcomes for a node.

Your task in this week’s assignment is to identify an interesting set of network data that is available on the web (either through web scraping or web APIs) that could be used for analyzing and comparing centrality measures across nodes.  As an additional constraint, there should be at least one categorical variable available for each node (such as “Male” or “Female”; “Republican”, “Democrat,” or “Undecided”, etc.)

In addition to identifying your data source, you should create a high level plan that describes how you would load the data for analysis, and describe a hypothetical outcome that could be predicted from comparing degree centrality across categorical groups.

For this week’s assignment, you are not required to actually load or analyze the data.  Please see also Project 1 below.

One interesting set of network data available on the web that could be used for analyzing and comparing centrality measures across nodes is data from Eventbrite, an event management and ticketing website that allows users to browse, create, and promote local events.

Data from Eventbrite is available through connection to their API: <https://www.eventbrite.com/platform/api>

I would connect to the Eventbrite API through authentication with a private token in order to access and load the data into python. The key object to build the analysis on would be the Attendee. The Attendee object represents the details of Attendee (ticket holder to an Event). The model is one Attendee per each sold ticket. The Attendee Profile Fields contains elements such as gender, which can be used as nodes.

From these nodes, we can predict the type of events to be attended based on the gender of the attendee. Event details are available in the Event object, and contains information such as event category (“music”, “sports”), venue, team, address. One can possibly also predict what type of events individuals in certain regions would attend, as location of the attendees is also available.