# Understanding Trough Tweets

Introduction to the twitter meta-data
Mining Twitter API

By Dina Bavli


# What Answers Your Data Brings?

Image by Gerd Altmann from Pixabay

By Dina Bavli

# Drowning in Data

Image by Engin Akyurt from Pixabay

By Dina Bavli

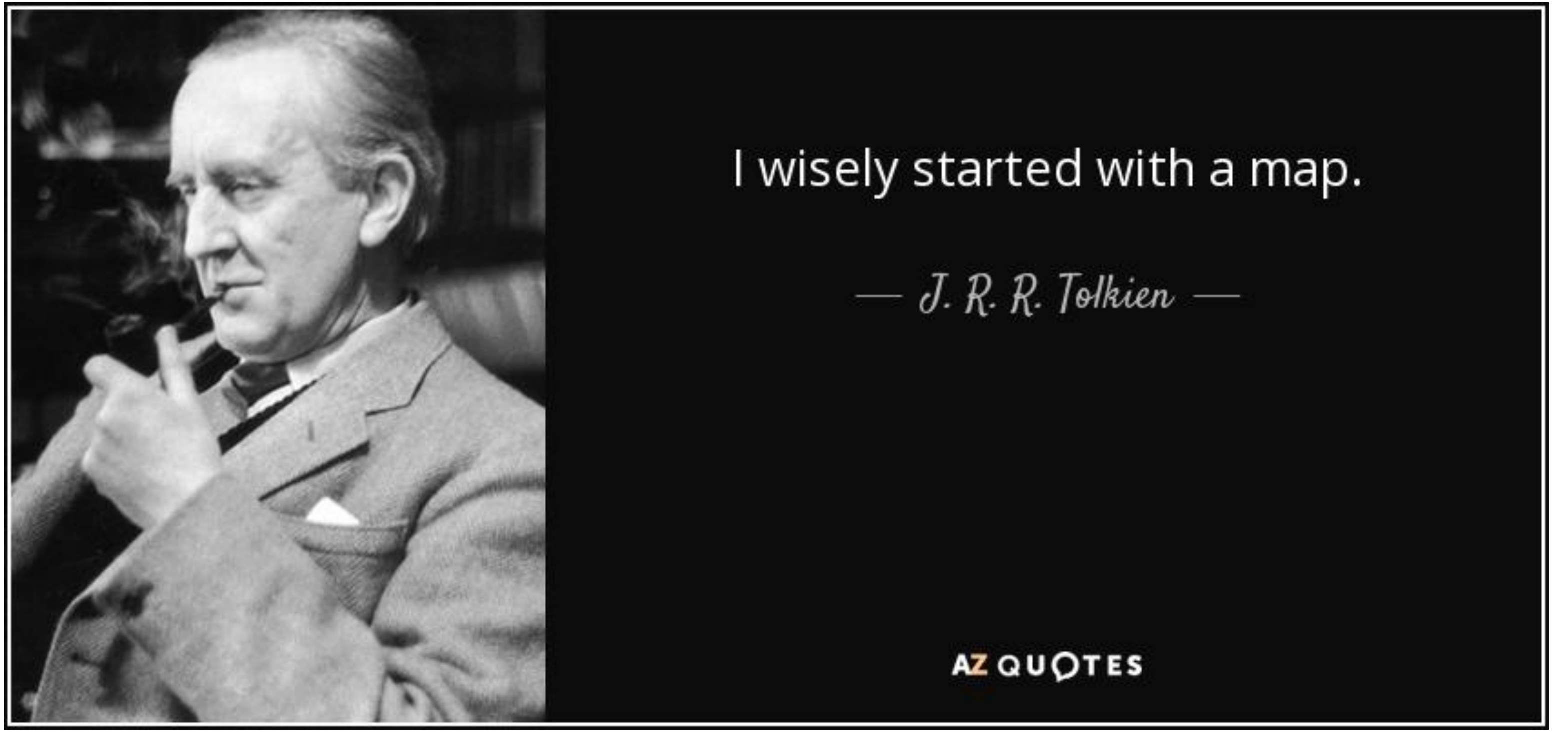

By Dina Bavli

Image by Gerd Altmann from Pixabay

By Dina Bavli


Image by photogrammer7 from Pixabay

By Dina Bavli




By Dina Bavli

# Viewing Trends By Location

## Putting Tweets on the Map Using Basemap

- Map Projections
- Latitude
- Longitude
- Weights
- Colors

By Dina Bavli

# Projections

- Pseudo-cylindrical projections
- Cylindrical projections
- Perspective projections

*Elaborate in your colab notebook*

By Dina Bavli

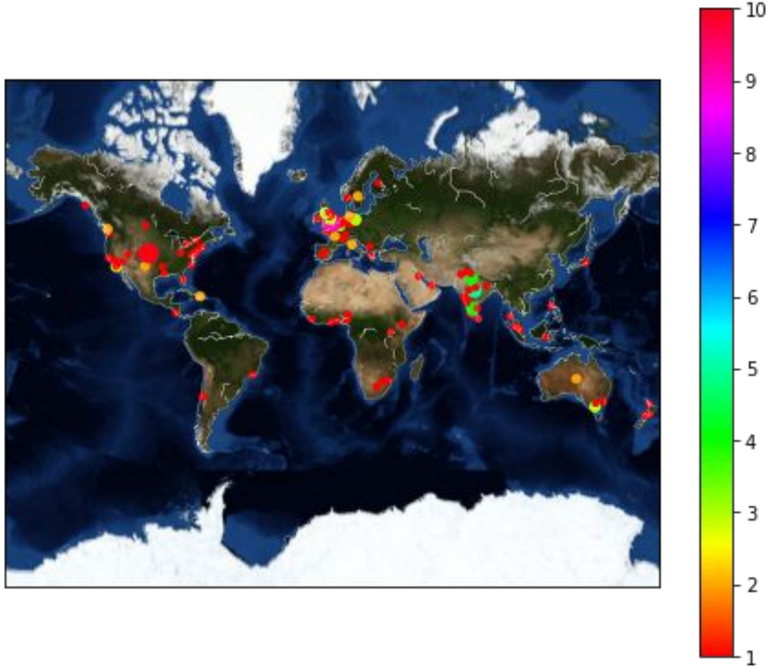
# Plotting Latitude and Longitude

By Dina Bavli

# Size by Weights

By Dina Bavli

# Adding Color




By Dina Bavli

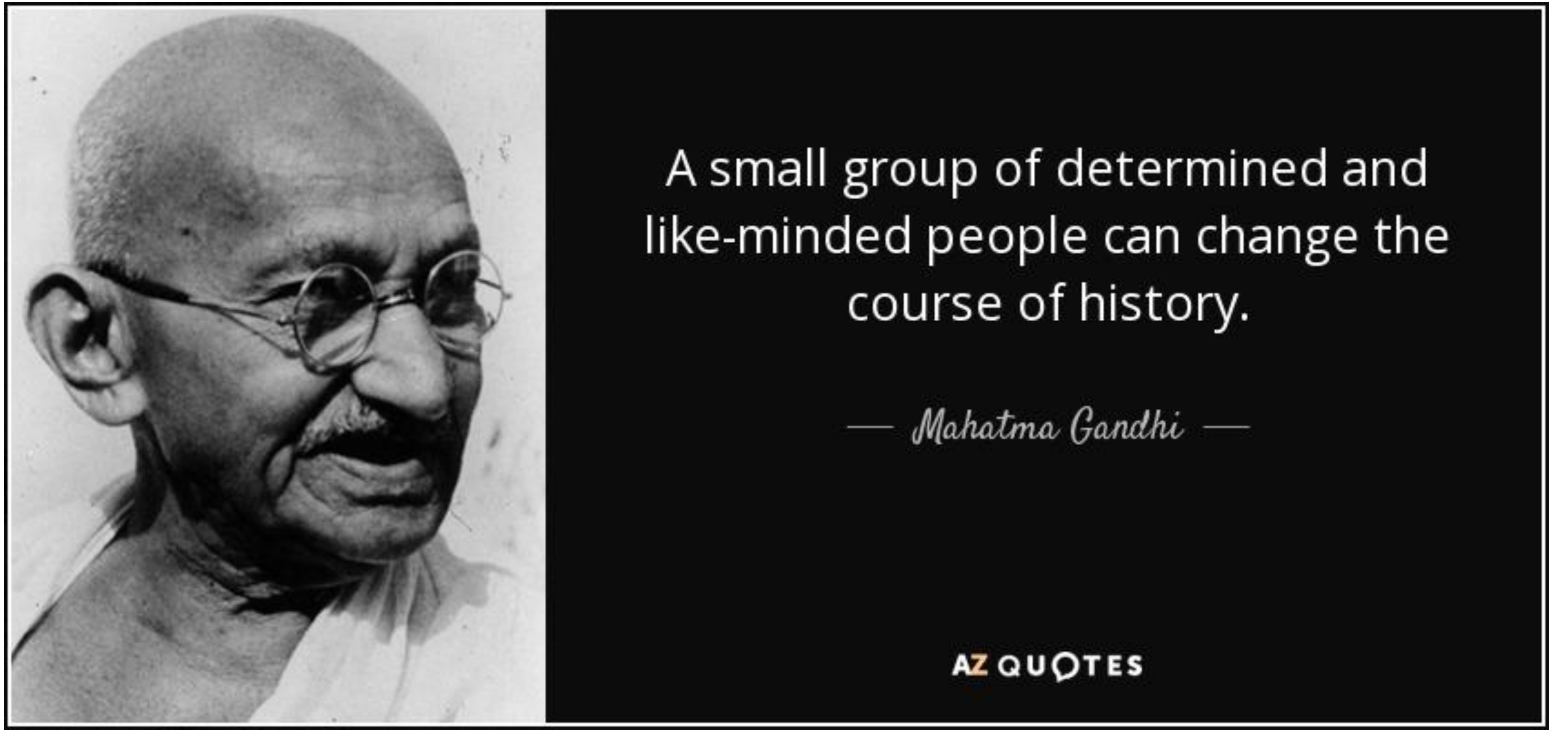
# Applying on Your Data

Image by Rudy and Peter Skitterians from Pixabay

By Dina Bavli


Image by Kevinsphotos from Pixabay

By Dina Bavli




By Dina Bavli


Image by Gerd Altmann from Pixabay

By Dina Bavli

# SNA
# Social Network Analysis

By Dina Bavli

# What is a social network?

| Points | Lines |
| --- | --- |
| People | Followers, friendships, kin |
| Companies | Acquire, trade, chain of supply |
| Disease | Chain of infection |
| Web pages | links |
| Countries | Flight routes, trade |
| Articles | Citation, writers |

By Dina Bavli

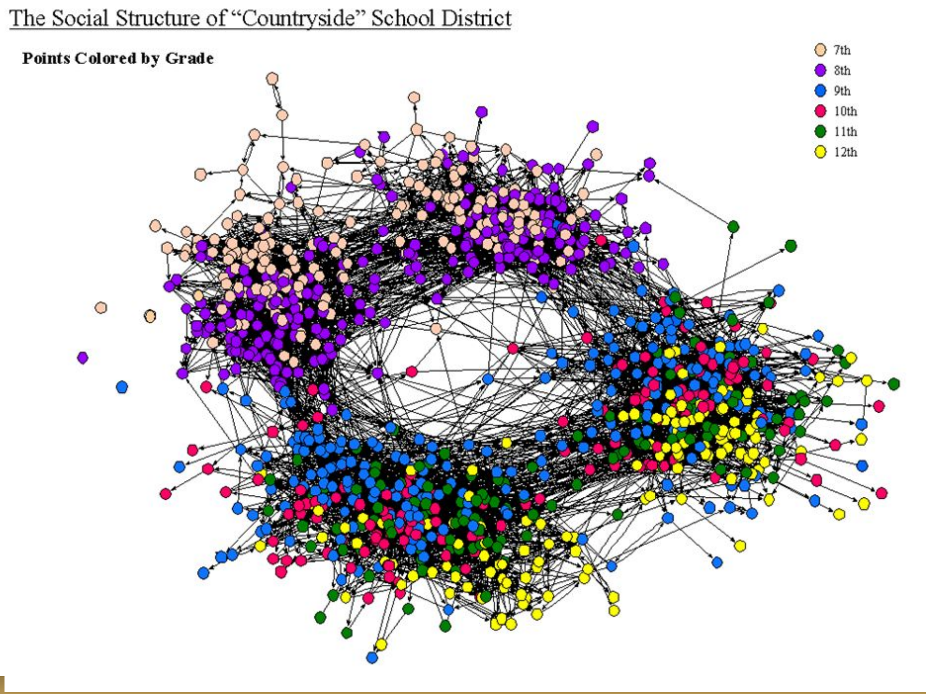
# Points and line have formal names that vary in different disciplines:

| Points | Lines | Discipline |
|---|---|---|
| Vertices | Edges, arcs | Math |
| Nodes | Linked, edges | Computer Science |
| Sites | Bonds | physics |
| Actors | Ties, relations | Sociology |

https://gawron.sdsu.edu/python_for_ss/course_core/book_draft/Social_Networks/Social_Networks.html#what-are-networks

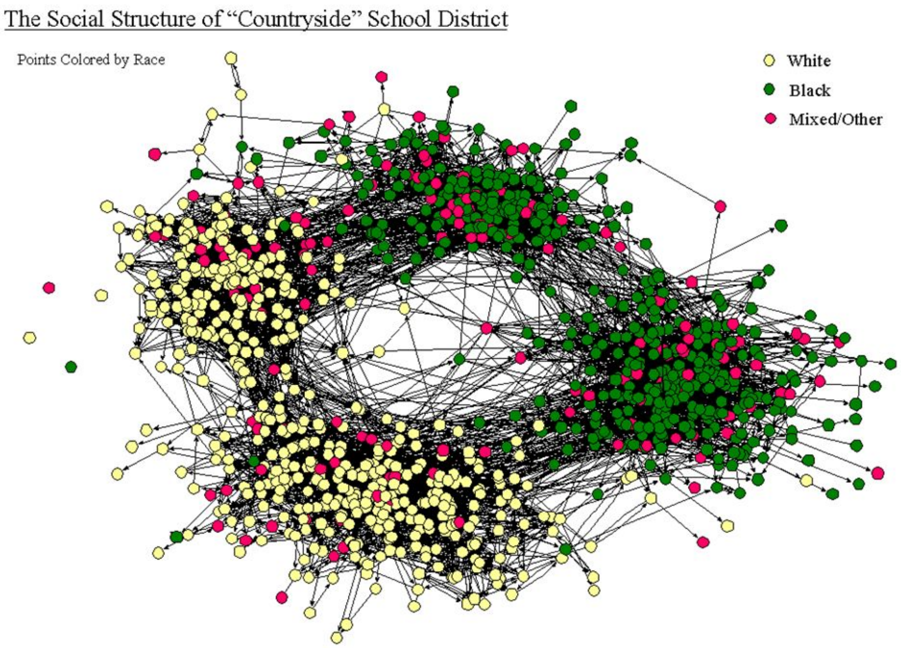
By Dina Bavli

# High Schools as Networks



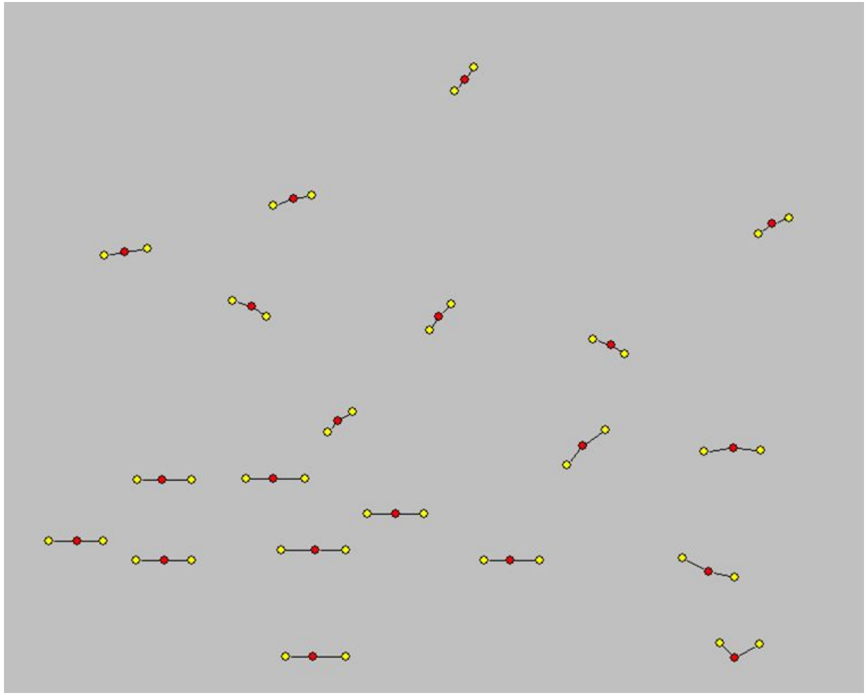

By Dina Bavli

# High Schools as Networks



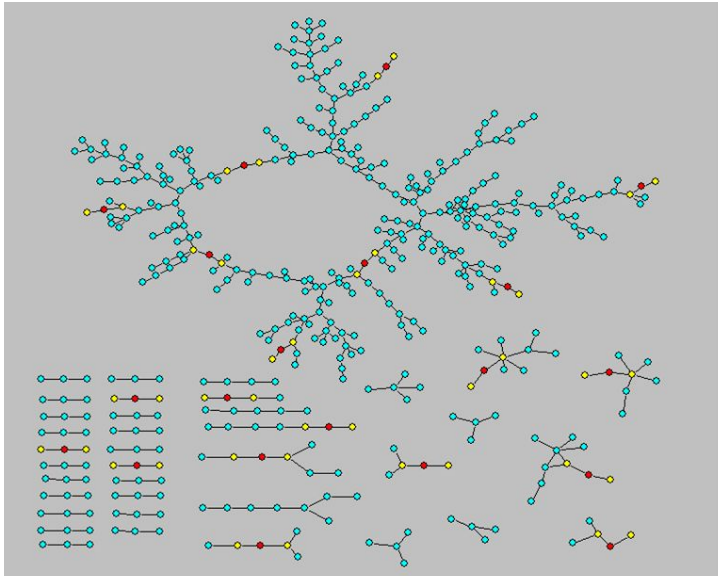

By Dina Bavli

# Why do Networks Matter?



By Dina Bavli

# Why do Networks Matter?

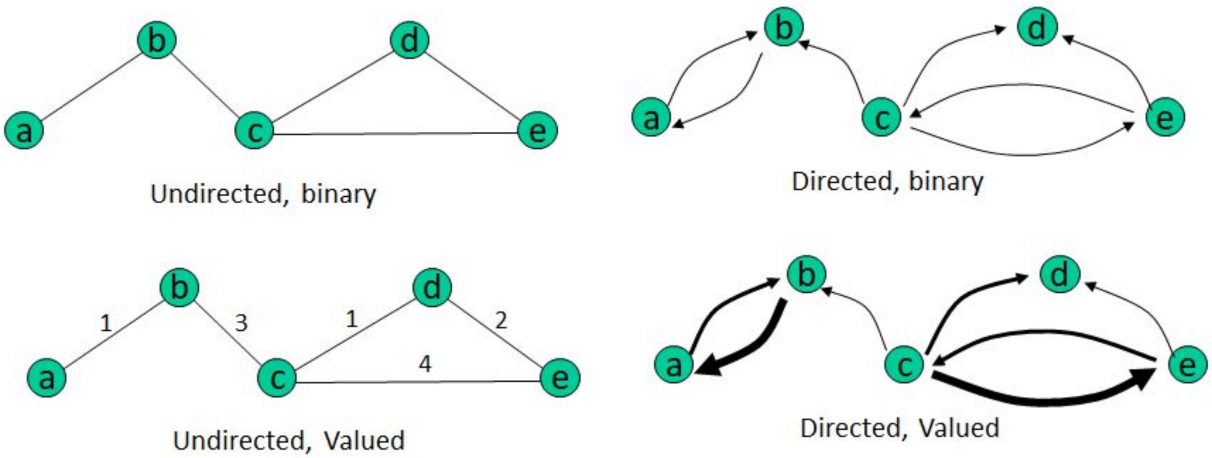


By Dina Bavli

# Why do Networks Matter?

In general, a relation can be:

- Binary or Valued
- Directed or Undirected



Two features of the network's topology are known to be important: connectivity and centrality

By Dina Bavli

# Connectivity

Connectivity refers to how actors in one part of the network are connected to actors in another part of the network.

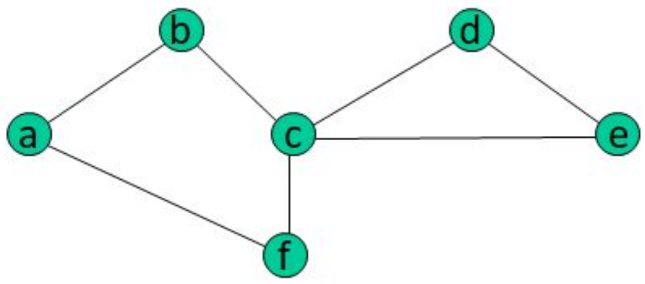
Reachability: Is it possible for actor i to reach actor j? This can only be true if there is a chain of contact from one actor to another.

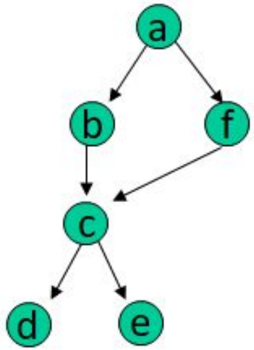
Distance: Given they can be reached, how many steps are they from each other?

Number of paths: How many different paths connect each pair?

# Reachability

Indirect connections are what make networks systems.  One actor can reach another if there is a path in the graph connecting them.



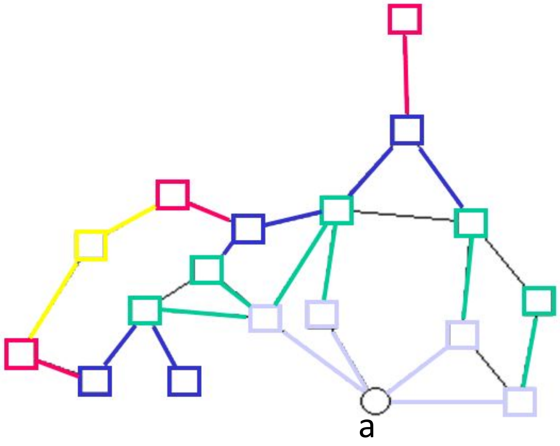




Indirect connections are what make networks systems.  One actor can reach another if there is a path in the graph connecting them.

By Dina Bavli

# Distance & number of paths

Distance is measured by the (weighted) number of relations separating a pair:




Actor “a” is:

1 step from 4- purple

2 steps from 5- green

3 steps from 4 - blue

4 steps from 3- red

5 steps from 1- yellow

By Dina Bavli

# Centrality

## Centrality refers to (one dimension of) location, identifying where an actor resides in a network.

For example, we can compare actors at the edge of the network to actors at the center.

In general, this is a way to formalize intuitive notions about the distinction between insiders and outsiders.

# Centrality

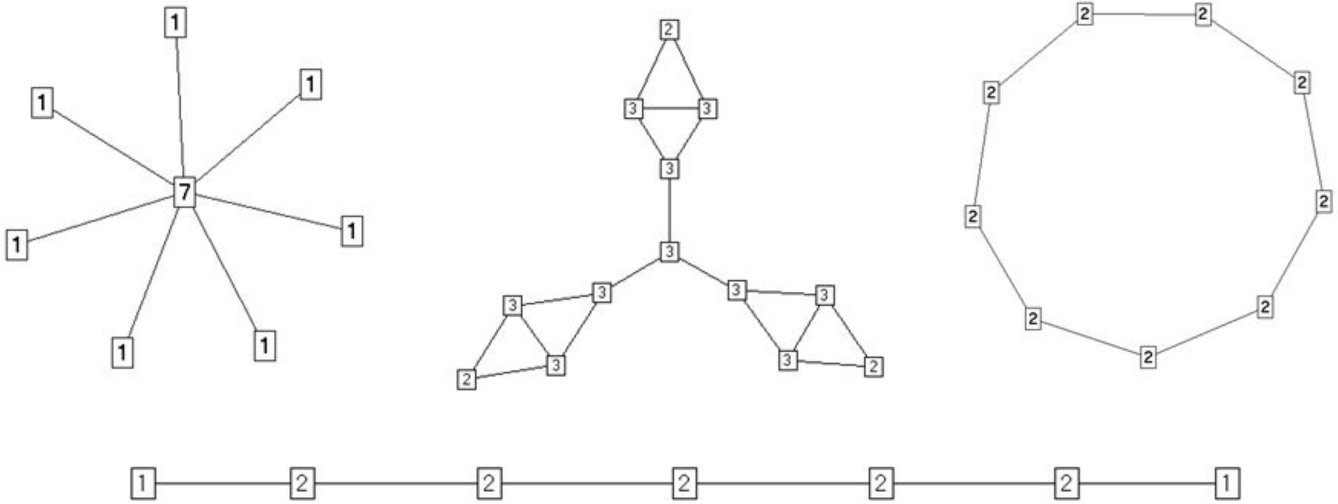
## Identifying which nodes are in the 'center' of the network.

Three standard centrality measures capture a wide range of "importance" in a network:

- Degree
- Closeness
- Betweenness

# Degree Centrality

Degree is the number of ties, and the actor with the most ties is the most important



$$C_D = d(n_i) = X_{i+} = \sum_j X_{ij}$$

# Closeness Centrality

An actor is considered important if he/she is relatively close to all other actors.

Closeness is based on the inverse of the distance of each actor to every other actor in the network.

Closeness Centrality:

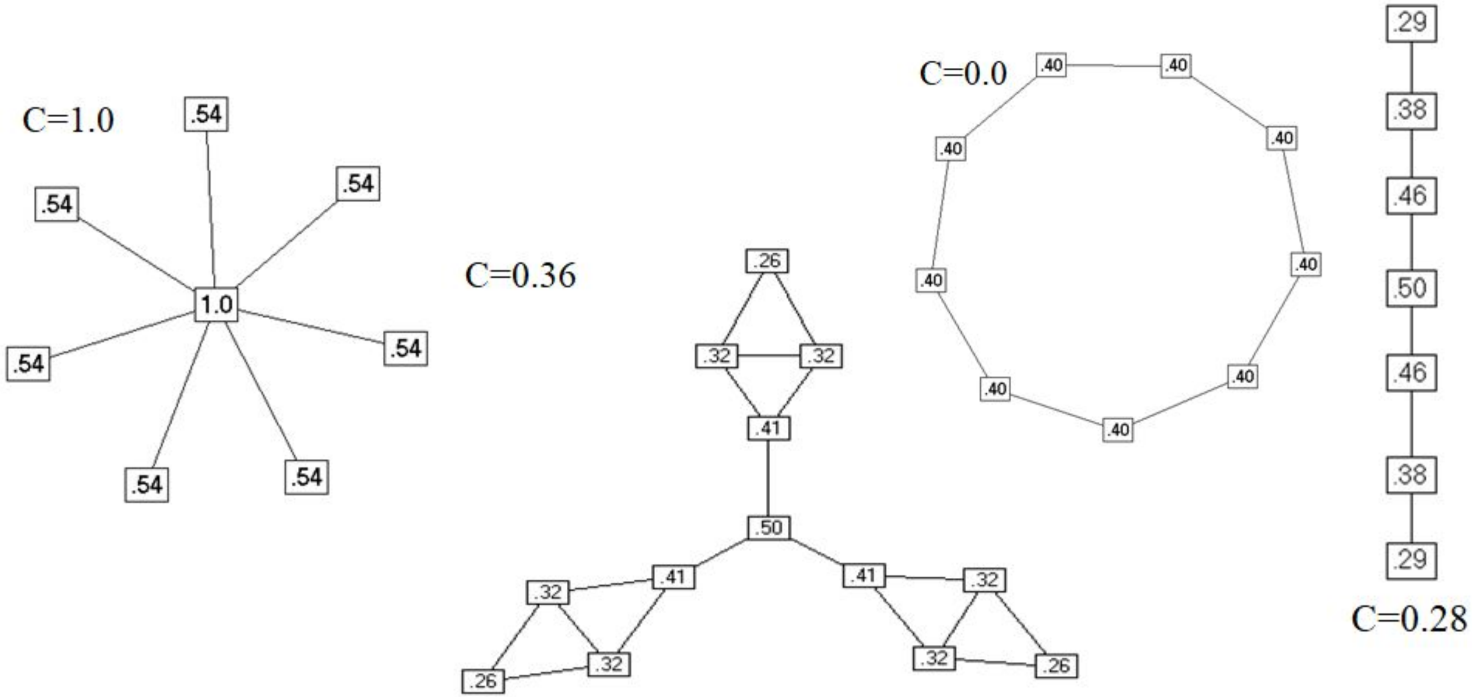
$$C_c(n_i) = \left[ \sum_{j=1}^{g} d(n_i, n_j) \right]^{-1}$$

Normalized Closeness Centrality

$$C'_C(n_i) = (C_C(n_i))(g-1)$$

By Dina Bavli

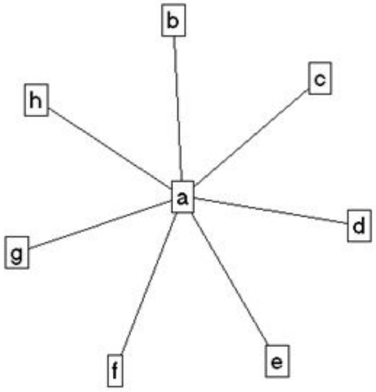
# Closeness Centrality examples




By Dina Bavli

# Betweenness Centrality

A person who lies on communication paths can control communication flow, and is thus important.



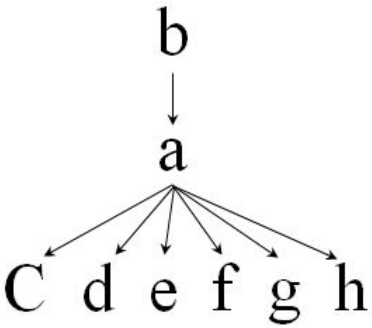




image: Freepik.com

By Dina Bavli

# Betweenness Centrality

A person who lies on communication paths can control communication flow, and is thus important.

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk}$$

Where $g_{jk}$ = the number of geodesics connecting jk, and

$g_{jk}(n_i)$ = the number of that actor i is on.

Usually normalized by:

$$C'_B(n_i) = C_B(n_i) / [(g-1)(g-2)/2]$$

image: Freepik.com

# What Kinds of Relations We Have?

By Dina Bavli

# Twitter Relations

Retweets

Mentions

Replies

By Dina Bavli


Image by 995645 from Pixabay

By Dina Bavli

# Understanding Groups

## There may be multiple relations of multiple types connecting your nodes.

Social network data are substantively divided by the number of modes in the data.

2-mode data represents nodes from two separate classes, where all ties are across classes.  Examples:

- People as members of groups
- People as authors on papers
- Words used often by people
- Events in the life history of people

The two modes of the data represent a duality: you can project the data as people connected to people through joint membership in a group, or groups to each other through common membership

By Dina Bavli


Image by Gino Crescoli from Pixabay

By Dina Bavli

# Sentiment Analysis

By Dina Bavli

Friedrich Nietzsche

We often refuse to accept an idea merely because the tone of voice in which it has been expressed is unsympathetic to us.

AZ QUOTES

# Sentiment analysis has many other names

- Opinion extraction
- Opinion mining
- Sentiment mining
- Subjectivity analysis

By Dina Bavli

# Why sentiment analysis?

- Movie:  is this review positive or negative?
- Products: what do people think about the new iPhone?
- Public sentiment: how is consumer confidence? Is despair increasing?
- Politics: what do people think about this candidate or issue?
- Prediction: predict election outcomes or market trends from sentiment

# Lexicon-based solutions

- Detect/extract the polarity of opinions, based on affective dictionaries [7,8]
- Word-lists where each token is annotated with an ‘emotional’ value
- e.g., positive/negative words or words that express anger, fear, happiness, etc.
- More to follow…
- Add syntactic and prose rules to estimate the overall polarity of text:
- Negation detection: “the movie wasn’t good”
- Exclamation detection: “great show!!”
- Emoticon detection: “went to the movies J”
- Emphasis detection: “You are goooooood”
- Intensifier, diminisher word detection: “Very good movie” vs. “good movie”
- Example of simplified process in next page…•

By Dina Bavli

# Lexicon-based solutions

- Detect emotion in two independent dimensions:
  - Positive: $D_{pos}$: {1, 2,… 5}
  - Negative: $D_{neg}$: {-5, -4,… -1}
- (optional) Predict overall polarity by comparing them :
  - If $D_{pos} > |D_{neg}|$ then positive
- Example: “He is brilliant but boring”
  - Emotion(‘brilliant’)=+3
  - Emotion(‘boring’)=-2

$D_{pos}$ =+3, $D_{neg}$=-2 => *positive*

- Negation detection: “H
  - Emotion(NOT ‘brilliant’) = -2
    - Decreased by 1 and sign reversed

$D_{pos}$ =+1 (default), $D_{neg}$=-3 => *negative*

- Exclamation detection: “H

# Lexicon-based solutions

- Detect/extract the polarity of opinions, based on affective dictionaries [7,8]
  - Word-lists where each token is annotated with an ‘emotional’ value
    - e.g., positive/negative words or words that express anger, fear, happiness, etc.
  - More to follow…
- Add syntactic and prose rules to estimate the overall polarity of text:
  - Negation detection: “the movie wasn’t good”
  - Exclamation detection: “great show!!”
  - Emoticon detection: “went to the movies J”
  - Emphasis detection: “You are goooooood”
  - Intensifier, diminisher word detection: “Very good movie” vs. “good movie”
- Example of simplified process in next page…•

By Dina Bavli


Image by Lubos Houska from Pixabay

By Dina Bavli

# Closer Look- Is It Mean the Same?

Image by Meli1670 from Pixabay

By Dina Bavli

# Challenges (I)

- Subtle ways of expressing private states
- "If you are reading this because it is your darling fragrance, please wear it at home exclusively and tape the windows shut" No negative words
- "Miss Austen is not a poetess" Fact or opinion?
- "Go read the book" Context
- "Yeah, sure!" Irony
- "I feel blue" vs "The sky is blue" Idioms
- "If you thought this was going to be a good movie, this isn't your day" Negation
- Informal language
- 90+% of language used in some social platforms deviates from standard English [3]
- "wuddup doe mah nigga juz droppin sum cuzz luv on u DeUcEz"
- As a result, even standard NLP processes need revisiting:
- Part-of-speech tagging in Twitter [4]

By Dina Bavli

# Challenges (II)

- "This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up" Opinion reversal
- "I bought an iPhone a few days ago. It was such a nice phone. The touch screen was really cool. The voice quality was clear too. Although the battery life was not long, that is ok for me.  However, my mother was mad with me… " Topic drift
- Lastly, in contrast to IR which is typically based on keywords, opinions are NOT easily conveyed by keywords.
- e.g. "unpredictable plot" vs. "unpredictable steering"

By Dina Bavli


Image by Amaya Eguizábal from Pixabay

By Dina Bavli


Image by Hans Benn from Pixabay
Questions?
By Dina Bavli

Image by Alexas_Fotos from Pixabay

https://www.linkedin.com/in/dina-bavli-502430158/

https://github.com/dinbav

By Dina Bavli


Image by Christine Sponchia from Pixabay

https://www.linkedin.com/in/dina-bavli-502430158/ https://github.com/dinbav

By Dina Bavli