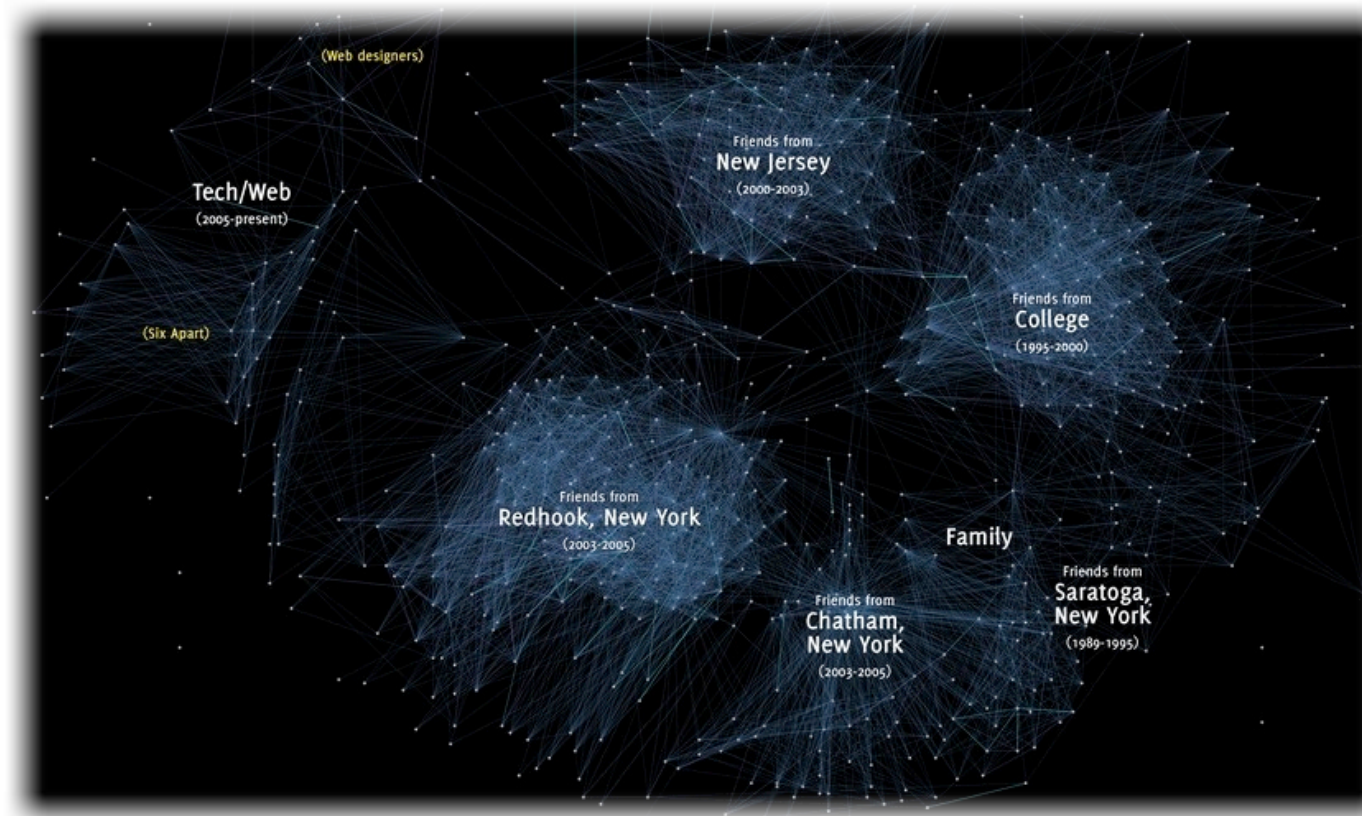


Using Python for Social Network Analysis



Matt Williams

CM1113 – Problem Solving with Python
Fri 8th Dec 2011

Overview

Concepts

Social networks

Graphs and their
properties

Dunbar's number

Small-world networks
(Milgram's small-world experiment)

Tools



NetworkX

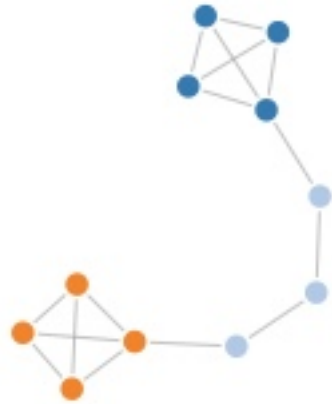


Data



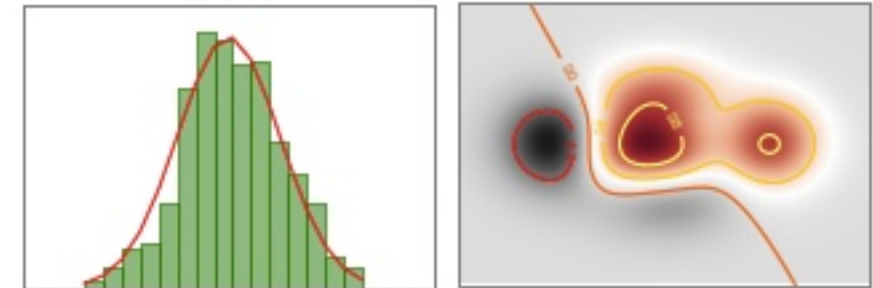
CS&I
undergraduates
friendships survey

Python packages



NetworkX

- Package for analysing the structure of networks (not just social!)
- Lots of complicated graph algorithms wrapped up in easy-to-use functions



matplotlib

- A comprehensive plotting package
- Usually easy-to-use
- Very flexible

Both included in the Enthought Python Distribution
And available (free) from their respective websites

What is a social network?

Examples...?

- A **collection of individuals** that are **connected** by one or more types of **‘social’ relationship**
 - Very broad definition!
- Can be **online** or **real-world** relationships



Linked 

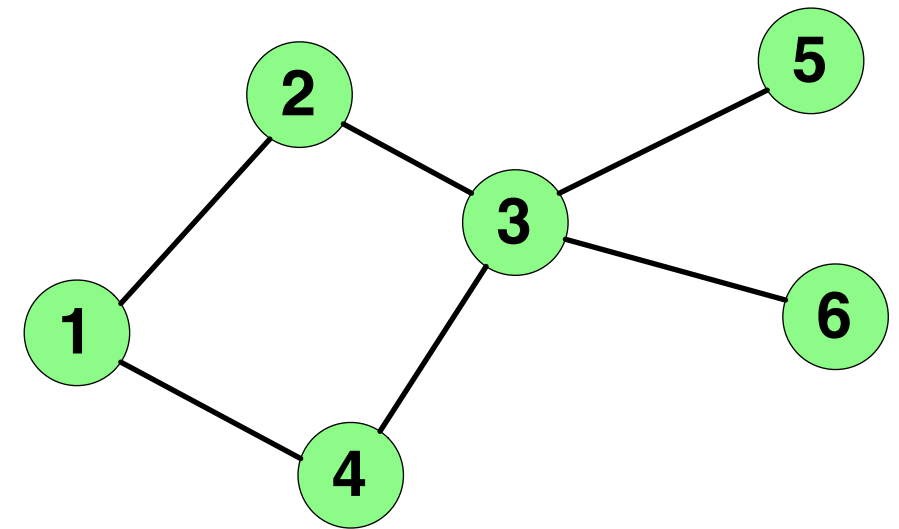


What is social network analysis (SNA)?

- Study of the **social relations** in a **group of people**
- Of interest to **academics** and **businesses**
- Brings together many fields:
 - sociology, anthropology, mathematics & statistics, computer science, and more
- Structure of relationships often reduced to a **graph representation**
- **Computers**, the **internet**, and the **web** have enabled large-scale analysis of social structures
 - Facebook: 800+ million active users
 - Massive computational problem

Graph theory recap

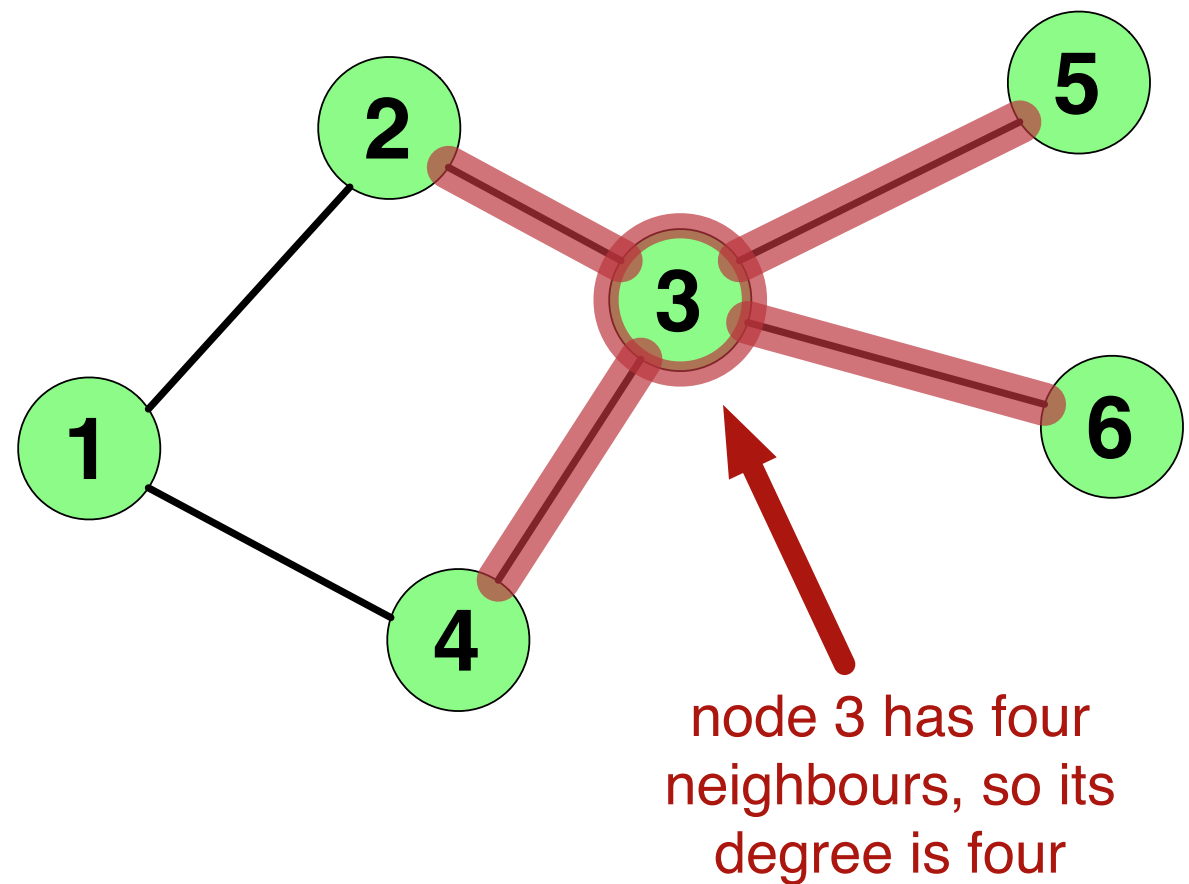
- **Node**: an entity (e.g., a person)
- **Edge**: some form of relationship between two nodes (e.g., Facebook friendship)
- Graphs may have **directed** or **undirected** edges. For example...
 - directed = Twitter followers
 - undirected = Facebook friendships
- We'll just focus on **undirected**
- There are other graph types: weighted graphs, multi graphs, ...



```
NetworkX:
nx.Graph()
nx.draw(g)
g.add_node(v)
g.add_edge(v, w)
g.nodes()
g.edges()
```

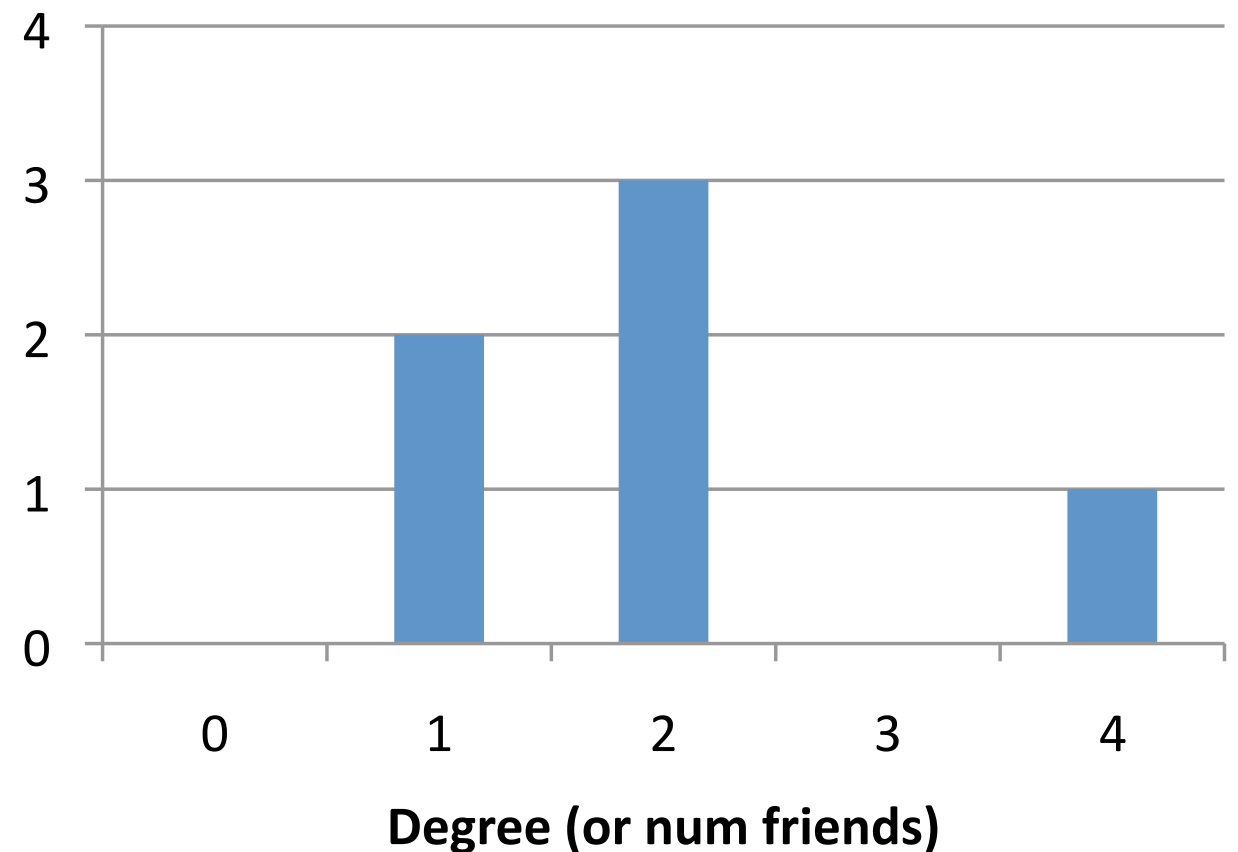
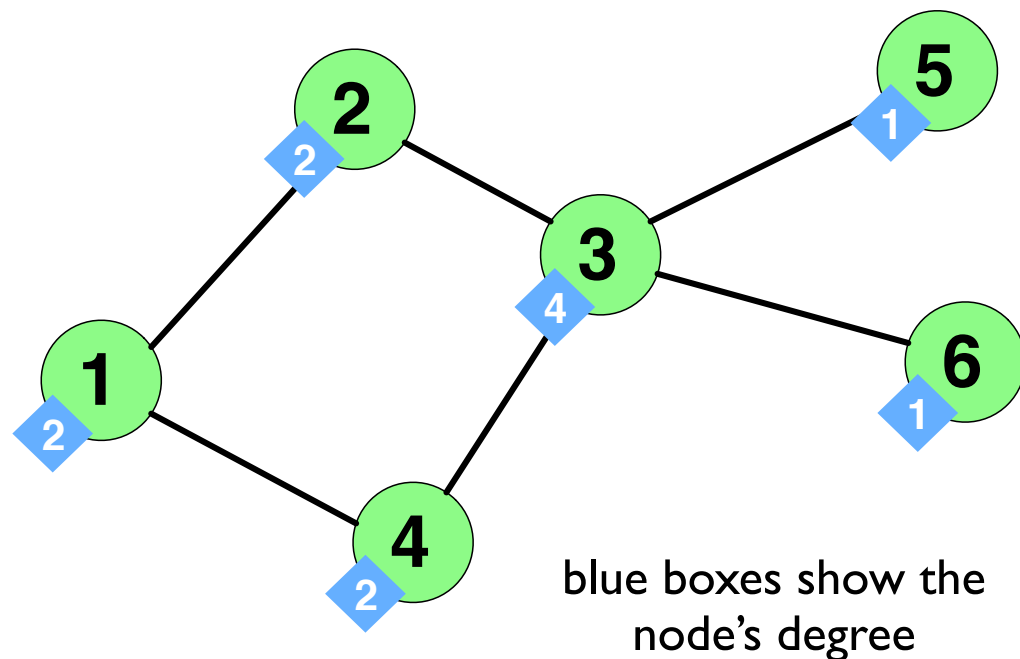
Some properties of graphs

- The **degree** of a node is the **number of neighbours** it has
- Think: number of friends a person has



```
NetworkX:  
g.degree()  
g.degree(v)
```

Some properties of graphs



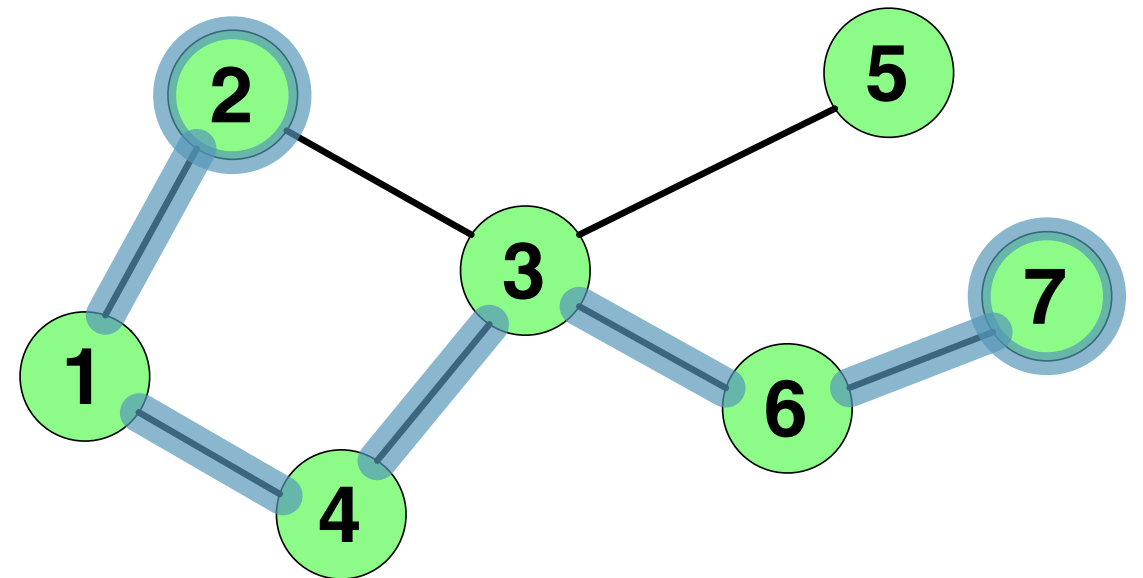
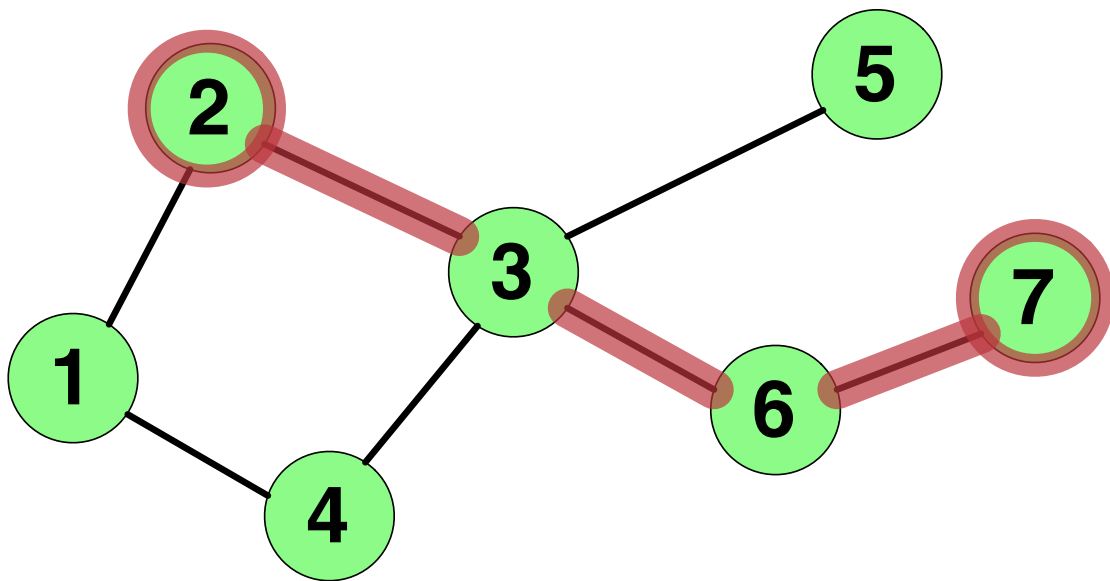
- **Degree distribution:**
 - How many people have...
 - One friend?
 - Two friends?
 - Three friends?
 - ...
 - One hundred friends?

```
matplotlib:  
plt.hist()
```

```
NetworkX:  
g.degree()  
g.degree(v)
```

Some properties of graphs

- The **shortest path** between two nodes is the path with the fewest number of ‘hops’ between them
 - There may be more than one shortest path



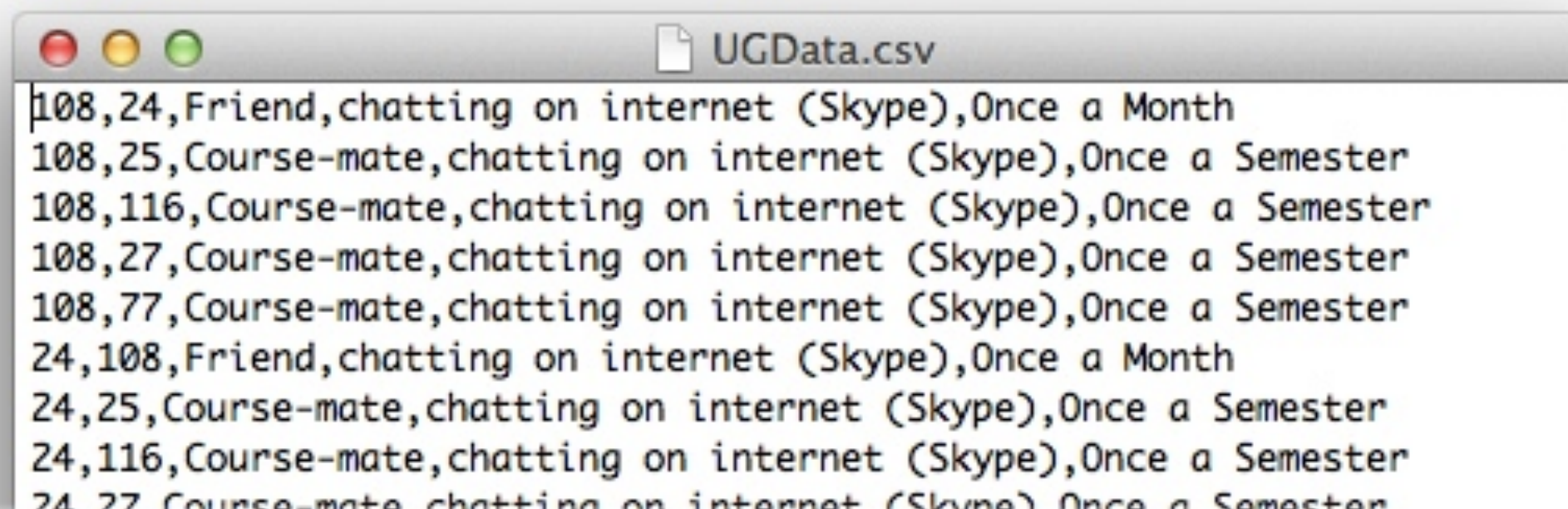
- What's the **shortest path** between **2** and **7**?
 - The **blue path** is **five hops** long
 - The **red path** is **three hops** long
 - What about other paths?

NetworkX:

```
nx.shortest_path(g, v, w)
nx.shortest_path_length(g, v, w)
nx.average_shortest_path_length(g)
```

A social network at CS&I

- Survey of the social ties between Computer Science & Informatics undergraduates
 - 130 participants
 - 514 social relationships mapped
 - Collected by Mona Ali, CS&I PhD student
- Stored as a comma-separated values (CSV) file



Milgram's small-world experiment

- The small-world experiment(s)
 - Began in 1967
 - 296 letters sent to individuals in the USA
 - Recipients told to **forward the letter** to someone they know **personally**, with the goal of it eventually reaching a designated person in Boston
 - Intermediaries repeat the procedure
 - Chain of hops recorded along the way
- 64 reached the destination
- Average path length was surprisingly short -- around **six hops**
- ...**six degrees of separation?**
- Recent experiments on online social networks have found figures of 6.7 and 4.0



Stanley Milgram

Milgram's small-world experiment

What is the average degree of separation in the CS&I Undergrad network?

Dunbar's number

- “**Cognitive limit** to the **number of individuals** with whom any one person can maintain **stable relationships**”
 - The maximum size of **social groups** in **primate species** is related to the species's neocortex size
 - Given the capacity of the human brain, we should only be able to maintain around **150 stable relationships** each!
- “But I have over 400 friends on Facebook!”
 - Actually...



Robin Dunbar

Dunbar's number

Neolithic Farming Villages	150-200
Maniples in the Roman Legion	120-130
Christmas Card Networks	154
Independent Units in Modern Armies	200
Average Number of Facebook Friends	130

Dunbar's number

How many friends do people tend to have in the CS&I
Undergrad network?

Summary

- Some **graph properties**:
 - degree, degree distribution, shortest path
- Some observations from **sociologists** and **anthropologists**:
 - Dunbar's number
 - Milgram's small-world experiment
- Don't forget: social network analysis is only one application of graph theory
 - Everything you've seen can be applied to other networks -- e.g., **routers** on the Internet, **neurons** in the brain, **road networks**, etc.
- Barely scratched the surface of **matplotlib** and **networkx**

Attribution

Social network visualised with Nexus Toolkit:

<http://nexus.ludios.net/>

<http://www.flickr.com/photos/plasticmind/3963424725/>

Facebook geographic network visualisation:

https://www.facebook.com/note.php?note_id=469716398919

Christmas card:

http://www.flickr.com/photos/the_justified_sinner/5225722803/

Old telephone:

<http://www.flickr.com/photos/ajc1/3367295141/>

Photograph of Stanley Milgram:

<http://www.stanleymilgram.com/>

Photograph of Robin Dunbar:

<http://re.allofus.com/post/12600355213/robin-dunbar-connecting-anthropology-and-user>

<http://www.lighttrick.co.uk/>