Six degrees of wikipedia
Assignment 1
Sociology 204 (Social Networks)
Matthew J. Salganik
Due: Wednesday, February 11, 2015 at 10am
This assignment has 22 questions for a total of 100 points.

Name: _____

Preceptror: _____

Precept Time: _____

1 Short questions

1. (8 points) Draw a simple example that highlights the difference between number of intermediaries and number of degrees of separation.

2. (4 points) List two limitations of the Travers and Milgram experiment.

3. (10 points) What is a question or issue that you would like to discuss in precept?

2 Six degrees of wikipedia

So far we have read about experiments that show the existence of short path lengths between randomly chosen pairs of people. And, we have read models that suggest that the short paths might exist in many kinds of networks, not just social networks. For this assignment you will look for those short paths in the links between entries in Wikipedia. For example, imagine trying to get from the page of Cornell University to the page Princeton University moving only through links on Wikipedia. That one is pretty easy. One path is:

- Cornell University → Ivy League → Princeton University

This path has two links. Another maybe less obvious path is:

- Cornell University → Supreme Court → Sonia Sotomayor → Princeton University

This path has three links. In general, there may be many paths, but always try to find the shortest. Finding a path between Cornell University and Princeton University is pretty easy. Sometimes things are harder and you make wrong turns. For example, trying to go from Lehman Brothers to Great Depression I went:

- Lehman Brothers → Investment Bank → Lehman Brothers → New York City → Wall Street → Great Depression

Note that you can go backwards in your search, as above. This path had 5 links.
Please describe the paths that you find for the following pairs (and keep track of wrong turns).

1. (2 points) Princeton University → Kevin Bacon

2. (2 points) Kevin Bacon → Princeton University

3. (2 points) Princeton University → Henry VIII of England

4. (2 points) Henry VIII of England → Princeton University

5. (2 points) Princeton University → Carnegie Hall

6. (2 points) Carnegie Hall → Princeton University

7. (2 points) Princeton University → White House

8. (2 points) White House → Princeton University

9. (2 points) Princeton University → Austrailia

10. (2 points) Australia → Princeton University

11. (2 points) Princeton University → Gretna, Nebraska

12. (2 points) Gretna, Nebraska → Princeton University

13. (1 point) Go to the website `http://beta.degreesofwikipedia.com/` and find the actual shortest paths.[1] Put your data into a format like Table 1 so that another scientist (e.g., your preceptor) can easily review it.

14. (6 points) How did you do at finding the shortest paths? Be specific and cite data.

---

[1] If you are having trouble getting the website to recognize the entities, try using capital letters.

15. (6 points) Were you closer to the true shortest paths when you going from Princeton University or when you were going to Princeton University? Be specific and cite data.

16. Travers and Milgram (1969) noted that of the 64 chains that reached the target 16 were sent by Mr. Jacobs, a clothing merchant in the town.

    (a) (6 points) Did you observe funneling in your searches to Princeton University? Why do you think this happened?

    (b) (6 points) Did you funnel in your searches from Princeton University? Why do you think this happened?

17. (7 points) Imagine that you were going to advise a friend about how to get to the Princeton University page from a random Wikipedia page. What algorithm/strategy/approach would you advise them to use?

18. (7 points) Imagine that you were going to advise a friend about how to get from the Princeton University page to a random Wikipedia page. What algorithm/strategy/approach would you advise them to use?

19. The diameter of a graph is the longest possible shortest path in a graph. For example, if you calculated the shortest path between all points on Wikipedia, the diameter would be the longest of these shortest paths. Try to find two pages that have a very long shortest path connecting them (remember you can find the degrees of separation here: `http://beta.degreesofwikipedia.com/`. However, there are a few rules: you can?t use the pages of asteroids; you can't use any of the paths that are already listed on the web; and you can't use brute force trial and error (or you might crash the website).

   (a) (5 points) What is the longest path you found? Extra credit for the person in each precept group who finds the longest shortest path.

   (b) (10 points) What was the process you went through to find this pair? What strategy did you use to find this pair?

Pair	# of links in your path	# of links in shortest path
Princeton University → Kevin Bacon		
Kevin Bacon → Princeton University		
Princeton University → Carnegie Hall		
Carnegie Hall → Princeton University		
Princeton University → White House		
White House → Princeton University		
Princeton University → Australia		
Australia → Princeton University		

Table 1: Table to store your data.