Feedback — Optional programming assignment 2

For this optional assignment you'll need a GML file that is a non-random subset of English-language wikipedia pages from 2007 and their hyperlinks. 6MB file here. It is also recommended that you work in R with igraph, as I have provided a mostly filled-in template you can work with.

Question 1

What is the power-law exponent of the degree distribution of the fragment of English-language wikipedia pages from 2007? Your answer must be accurate to within 0.2

You entered:

Your Answer		Score	Explanation
	✖	0.00	
Total		0.00 / 2.00	

Question Explanation

See http://tuvalu.santafe.edu/~aaronc/powerlaws/ for an explanation of how to fit power-laws and to download R code.

Question 2

Looking at the degree distribution of the above Wikipedia page network, what can you say?

Your Answer		Score	Explanation
○ the distribution has xmin > 1			
○ the distribution has xmin = 1			
○ the distribution has a pronounced exponential cutoff			

○ the distribution has no exponential cutoff

Total		0.00 / 2.00

Question Explanation

Look at the xmin returned by the plfit function. Also check the plot for deviation from the pure power-law distribution.

Question 3

Calculate the indegrees and outdegrees of the nodes in the Wikipedia graph. Which of the following is true.

Your Answer	Score	Explanation
○ the node with the highest outdegree is a Wikipedia hub page pointing to many articles on the same topic		
○ the node with the highest outdegree pertains to computer science		
○ the node with the highest indegree pertains to a broad range of materials		
○ the node with the highest outdegree pertains to sociology		
Total	0.00 / 2.00	

Question Explanation

The empirical results should be intuitive: broad concepts should have high indegree, hub pages should have high outdegree, but you need to compute them in order to see which hub/concept pages these are.

Question 4

Find the highest (undirected) betweenness node in the Wikipedia graph. This node likely has high betweenness because

Your Answer	Score	Explanation
○ It has high outdegree		

○ It is in the periphery of the graph

○ It has high indegree

○ It links different mathematical concepts

Total	0.00 / 2.00

Question 5

Use igraph's page.rank() function with the default damping factor (damping = 0.85). The node with the highest PageRank in the graph is:

Your Answer	Score	Explanation
○ Set		
○ Government		
○ Education		
○ Metabolism		
Total	0.00 / 2.00	