HW 1: due Thurs, Sept 17

For this assignment, submit your answers as problems as a single pdf file named hwk1.pdf alongside the CSV files for the requested graphs (see below) and any code you may have written and required additional text files, using "provide" with the command:

```
homework% provide comp150NS hwk1 hwk1.pdf p1.csv p2.csv p3.csv p4.csv code-you-wrote
```

When asked to construct a graph, submit a comma-delimited adjacency matrix as a text file named p[problem number].csv. For the graphs in this assignment that are bipartite, include only one set of nodes on each axis of the adjacency matrix. Make sure to include node names in all matrices. For example, for problem 1 you will submit p1.csv, and the first few lines might look like this:

```
,Jumbo Studios,Jumbo Kingdom,World's Fair,Jumbo Golf Course,Trunk Water Park,Animal Planet Zoo
Mueller,1,1,0,0,0,1
Ayers,1,1,0,0,1,0
Rivas,1,0,1,0,0,0
```

For this problem, you will be using a dataset for families visiting the Jumbo World Resort, available as the Theme Park data set on the course website. Jumbo World contains six seperate parks and eight hotels, and you are given data for 1000 families visiting the park over the weekend. For each family,

you are given three entries, one for each day, indicating which hotel the family stayed at (which doesn't change during their visit) and which park they visited that day.

1. Construct the graph matching families and parks. Which park is the most popular? Which park is the least popular?

2. Construct the graph matching hotels and parks, with edges weighted by the number of unique families that visit a park while staying at a given hotel. Include a visualization of this graph using the graph visualization program of your choice.

3. Construct a graph of families, with edges between families if they visit at least two of the same parks.

4. Construct the same graph, but add edges only if two families visited at least two of the same parks on the same day. How does the resulting graph change?

5. Families' choices of which hotel to stay at and parks to visit may be influenced by park geography. What might you be able to infer about park geography using the graphs you have made?

Note: data available at http://www.cs.tufts.edu/comp/150NS/private/Datasets