Shannon McNish
MSAN 502 HW 4
8/6/17

For this project, I used a dataset of the complete US airport network in 2010. The data was found on https://toreopsahl.com/datasets but originally downloaded from the Bureau of Transportation Statistics (BTS) Transtats site with the following filters: Geography=all; Year=2010; Months=all; and columns: Passengers, Origin, Dest. Airport codes were converted to ID numbers.

The data is a text file with column one equal to the origin node and column 2 equal to the destination node. The third column is a weighted value which is ignored.

- The data in a text file can be found here: http://opsahl.co.uk/tnet/datasets/USairport_2010.txt
- The metadata in a text file can be found here: http://opsahl.co.uk/tnet/datasets/USairport_2010_codes.txt

To run my program, use the above two data files as script parameters. For example, with the files in the same directory as my script, I ran this from my terminal:

```
python McNish_HW4.py USairport_2010.txt USairport_2010_codes.txt
```

My is code well-commented, but the outline is as follows:

1. Read in data and metadata text files
2. Create matrix A based on the highest node id
3. Calculate the google matrix M
4. Calculate the pagerank vector, M^k*v, using a user-defined k
5. Check to see if that k was sufficiently large enough by testing for convergence
6. Merge metadata onto pagerank vector
7. Print top 10 and bottom 10 nodes

I chose this dataset because I travel a lot and I thought it would be interesting to see the most well-connected airports. I was not surprised to see the results. The top 10 airports are all well-known and well-traveled with JFK taking the top spot. I had to google what the bottom 10 airport codes were. Ugnu–Kuparuk Airport (UUK), the bottom ranking airport is is a private-use airport located in Kuparuk, Alaska.

Top 10 Ranking Airports

Node	Score	Airport Code
766	0.000321	JFK

114	0.000307	ATL
1016	0.000304	MIA
877	0.00029	LAX
709	0.000278	IAD
1200	0.00027	ORD
685	0.000265	HPN
711	0.000261	IAH
88	0.000258	ANC
389	0.000256	DEN

Bottom 10 Ranking Airports

Node	**Score**	**Airport Code**
1407	0.000003	SAW
958	0.000003	LWV
453	0.000003	ECU
1125	0.000003	NKG
1124	0.000003	NKC
276	0.000003	CEW
275	0.000003	CEV
1606	0.000003	TN7
1605	0.000003	TN2
1668	0.000003	UUK