The Data:
Downloaded from https://www.movebank.org/, a repository for geolocation data on species. The specific files that were downloaded are from the Life Track project and consist of geopoints of tracked individuals of three bird species, Great Blue Heron, Great Egret, and Bald Eagle. Trackers were affixed to all birds and recorded geolocation during times of movement (based on accelerometer readings from the tracker).

Great Blue Heron: 334,513 entries for 5 individuals containing timestamp, latitude, longitude, accelerometer readings, and bird Id
Great Egret: 730,015 entries for 9 individuals containing timestamp, latitude, longitude, accelerometer readings, and bird Id
Bald Eagle: 1315 entries for 4 individuals containing timestamp, latitude, longitude, accelerometer readings, and bird Id

The Plan
Step 1: Read all data into dataframes
Step 2: Remove entries where there is no coordinate data
Step 3: Group by individual
Step 4: Group days and calculate number of recordings per day (proxy for number of movements because of accelerometers)
Step 5: Calculate distance traveled between all recordings
Step 6: Calculate distance traveled per day and statistics for that
Step 7: Break down distance traveled into time categories within the day to see if there are patterns for when the birds are most active
Step 8: Use clustering analysis and regression to see if individuals can be categorized to the proper bin based on their movements
Steps 1-3, and 5 have been completed.

Hypothesis:
Different species should have different movement patterns as determined by the number of movements per day, the timing of those movements, and the distance traveled during those movements. With regression and clustering we should be able to make predictions about which species and individual belongs to based on its movement patterns.

Preliminary analysis, mean distance traveled between recordings for all individuals (km):

Great Egret

ID	Distance (km)
2689	0.040669
2690	0.016758
2691	0.067915
2692	0.074887
2693	0.104122
2694	0.028772
2695	0.015980
2696	0.038352
2697	0.053197

Blue Heron

ID	Distance (km)
3258	0.034831
3259	0.017409
3260	0.010644
3262	0.014434
3263	0.029287

Eagle

ID	Distance (km)
2689	3.930932
2696	1.657986
2698	0.034350
66582	0.495760

There were entries that when the bird ID changed to the next one the distance calculated was an outlier due to different geographic locations of the tagged animals. These were filtered out by removing points with a distance traveled greater than a certain threshold (different for each species, as determined by examining normal maximum distances traveled).