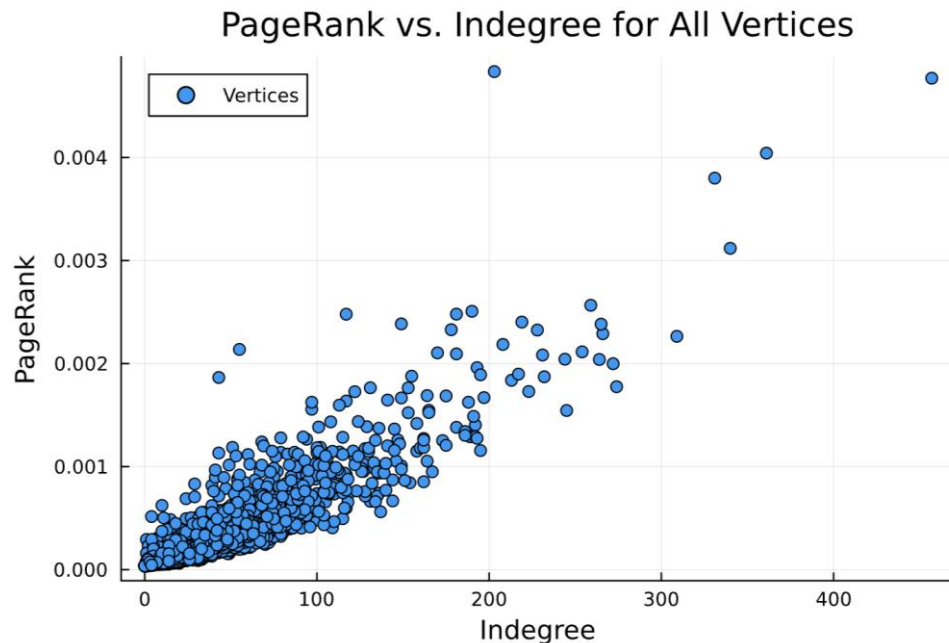


Problem 1: Done in assignment4_handout.jl.

Problem 2: Done in assignment4_handout.jl.

Problem 3:



a) Provide the top 5 and bottom 5 vertices in the graph according to PageRank

```
edges = load_directed_graph()
tol = 1e-6
v = page_rank(edges, tol)
top_indices = sortperm(v, rev=true)[1:5]
println("Top 5 Vertices According to PageRank")
for i in top_indices
    println("Vertex $(i) has PageRank $(v[i])")
end

bot_indices = sortperm(v)[1:5]
println("Bottom 5 Vertices According to PageRank")
for i in bot_indices
    println("Vertex $(i) has PageRank $(v[i])")
end
```

```
Top 5 Vertices According to PageRank
Vertex 5807 has PageRank 0.004833702077926967
Vertex 3650 has PageRank 0.004769747272235577
Vertex 13   has PageRank 0.0040423793705081056
Vertex 2411 has PageRank 0.003799802906569877
Vertex 2205 has PageRank 0.0031187003180348266
```

```
Bottom 5 Vertices According to PageRank
Vertex 2   has PageRank 3.6572739853054447e-5
Vertex 3   has PageRank 3.6572739853054447e-5
Vertex 5   has PageRank 3.6572739853054447e-5
Vertex 7   has PageRank 3.6572739853054447e-5
Vertex 9   has PageRank 3.6572739853054447e-5
```

b) Provide an explanation for why the incoming degree and the PageRank of each user are not perfectly correlated.

In the context of network analysis, the incoming degree (indegree) of a vertex is defined by the number of direct connections it receives from other vertices, showing how often it is referenced by others. Conversely, the outgoing degree (outdegree) of a vertex shows the count of its connections directed towards other nodes, representing how often it refers to other vertices within the network. PageRank is a more complex algorithm that considers not only the quantity of incoming links, but also their quality. PageRank operates on the idea that a link from a highly ranked page is more influential than one from a lower-ranked page. This process is iterative and the initial PageRank values are recalculated over several passes until the values converge to a stable state. As a result, a page with a high number of incoming links (high indegree) may not necessarily have a high PageRank if those links come from less reputable or lower-ranked vertices. Conversely, a page with fewer incoming links (low indegree) can have a high PageRank if those links are from highly ranked and quality vertices. Discrepancies arise because PageRank reflects the probability distribution of a random surfer's location after many steps, which is influenced by both the link structure and the rank of the linking pages. Therefore, although there is a connection between indegree and PageRank, they don't always align perfectly. This is because PageRank delves deeper, considering not just how many links point to a page, but also the ranking and quality of those linking pages, all within the broader context of the network's link structure.