Network Properties in Spark GraphFrames

Instruction to run the degree_powerlaw.py:

Please run the degree_powerlaw.py keeping all the .csv files in the same folder. This python script checks if a degree distribution (graph) follows power law and outputs the degree distributions (graphs) that follow power law.

Solutions to all the Questions:

- **Degree Distribution:**
    1) gnm1.csv having gamma = 2.89 follows power law and hence is scale free
    2) amazon.graph.small.csv having gamma = 2.39 follows power law and hence is scale free
- **Centrality:**
    1) Nodes from highest to lowest closeness centrality:
       F = C > H = D > B = E > G = A > I > J

    2) Two machines that would be the best candidates to hold this data based on other machines having few hops to access this data are the machines F and C

- **Articulation Points:**
    1) Below are the nodes that are articulation points and hence need to be disconnected:
        - Mohamed Atta
        - Usman Bandukra
        - Mamoun Darkazanli
        - Essid Sami Ben Khemais
        - Djamal Beghal
        - Nawaf Alhazmi
        - Raed Hijazi