Q1.
An important and large-scale application of stable marriage is in assigning users to servers in a large distributed Internet service. Billions of users access web pages, videos, and other services on the Internet, requiring each user to be matched to one of (potentially) hundreds of thousands of servers around the world that offer that service. A user prefers servers that are proximal enough to provide a faster response time for the requested service, resulting in a (partial) preferential ordering of the servers for each user. Each server prefers to serve users that it can with a lower cost, resulting in a (partial) preferential ordering of users for each server.

Q2.
Given nodes in the network, you can identify the nodes which form a critical part of the network. These can be identified by centrality measures such as betweenness centrality, degree centrality and closeness centrality. Also identify the nodes on critical lines and hence place more sensors on the transmission lines which will help when a sensor goes down other sensors take care.

Q3.

Q3. Local clustering coefficient

This measures how well connected a node is with its neighbors. It also quantifies the proportion of connections among a vertex's neighbors out of all possible connections that could exist between them.

The local clustering coefficient for a vertex v is calculated as follows.

$$C_v = \frac{2E_v}{K_v(K_v-1)} \rightarrow eq(1)$$

where C_v = local clustering coefficient of vertex v

E_v = number of actual edges between the neighbors of vertex v

K_v = degree of vertex v

Redundancy (R_v)

Redundancy focuses on the average number of connections that a neighbor of vertex v has to other neighbors of v. In the other words, it quantifies how well the neighbors of v are interconnected.

The redundancy of vertex v can be calculated as follows,

$$R_v = \frac{\sum e_v}{k_v} \qquad R_v = \frac{2e_v}{K_v}$$

where R_v - redundancy of vertex v

N_v - set of neighbors of vertex v

E_u - no. of actual edges b/w neighbors of vertex u

K_u - degree of vertex u.

Relation between C_v and R_v

$$R_i = C_i (K_i - 1)$$

$$R_i \propto C_i$$

This means that the redundancy of a node is the product of the clustering coefficient of the node with the number of neighbouring nodes and ~~is the~~ directly proportional to clustering coefficient. coefficient increases, the redundancy would also increase and vice-versa.

Q4.

a)

We can provide a solution based on network science ideas to improve traffic flow on the D.C. metro map while keeping the same stations and network topology. The goal would be to reduce traffic, increase connectivity, and improve the metro system's overall effectiveness. We need to have the following data before we can proceed with solving the problem.

Find out how many people use a particular route between two stations at various times of the day, month, and year. Count the number of transfer stations and identify the locations with the most transfers.Capture seasonal patterns to analyze trends in metro journeys during various times of the year. Collect data on the capacity of each station, taking into account elements like station space, platform size, and the number of platforms.

Investigate ideas such as clustering coefficients, network diameter, node centrality, and degree of connection. To determine which transfer locations have the highest centrality, create various node centralities, such as edge betweenness. To remove stations that are too near to one another and improve network performance, use geodesic pathways. Maintain network connectivity to avoid stations and nodes that are separated. Reassign stations to lines using graph theory techniques, aiming to minimize travel time, transfers, and uniform passenger load. Use algorithms like Breadth-First Search, Dijkstra's, Bellman-Ford, or Floyd-Warshall to identify the fastest routes between stations while taking traffic and time into account.Before putting the suggested network into practice on a bigger scale, simulate it and assess its performance on a smaller scale. Keep an eye on the network for any problems and have a backup plan ready to handle any difficulties. Continually evaluate the network's performance and make adjustments in response to shifting circumstances. On a bigger scale, gradually introduce the improved network to ensure a seamless transition. Maintain ongoing observation to spot and quickly resolve any problems. Implement a robust backup and recovery system to handle unexpected disruptions.

B.
A data-driven strategy and the application of network science ideas are required to determine the best start and stop locations for three new express trains on the D.C. metro system in order to maximize passenger flow and meet particular demands. The following actions need be conducted in order to ascertain the start and stop locations for these express trains:
Data-Driven Express Train Stops: Based on ridership data for each line, use the WMATA dataset to detect busy stations.
Identify stations where passengers on a particular line tend to congregate and have a look at these as potential Express train stops.
Identify the primary crossings where commuters change from one line to another. See also Node Centrality. To identify stations with the most people moving between them, evaluate node centrality using edge betweenness.
Peak Hour and Seasonal Considerations:

To reduce congestion, optimize Express train timetables by allocating extra trains during busy times and various seasons.

When demand is at its peak, strategically deploy Express trains based on rider data.
Continuous Monitoring and Backup Plans:

Establish a reliable monitoring system to address any potential concerns as soon as they appear.
Routes of Express Trains Between Lines: The Orange Line Express stops at several important stations, including East Falls Church, Rosslyn, Metro Center, L'Enfant Plaza, and Stadium Armory.
The Orange and Red Line Express begins at the Fort Totten transfer station on the Red Line, with stops at Gallery Place and Metro Center before merging with the Orange Line at Rosslyn and coming to an end at East Falls Church.
Blue and Yellow Line Express:Departs from the King Street-Old Town transfer station on the Blue Line. Major stops include the Pentagon, where it joins the Yellow Line, L'Enfant Plaza, Gallery Place, and Fort Totten transfer station.
Design Express routes that cover significant transfer stations and locations with high traffic density. Ensure thorough coverage on all lines to give commuters effective transportation options. Work together with WMATA to alter Express train services depending on shifting commuter trends through ongoing data sharing and feedback. WHY Encourage teamwork in order to improve the metro system in Washington, DC, generally.

Q5.

The paper presents us with densification laws where we observe that as the density of the graph increases the average degree and hence the number of edges increase super linearly with nodes and hence the average degree of the node increases and it follows a power law. And also the diameter of the dense graph is smaller which suggests that the graphs are shrinking over time. So, the paper presents 2 families of probabilistic generative models. The first model which is called Community guided attachment which says that nodes are divided into nested sets of communities such that the difficulty increases for forming links between communities. Another model which explains the phenomenon is the Forest fire model which says that a new node gets added to the network by burning the existing set of edges in epidemic fashion. These models help us in several ways. In graph generation, these models help us in generating synthetic graphs with those cases which we want to simulate on. These also help in sampling of the graph as the network size of real world data increases, it's difficult to collect a part of the graph which suits our needs. This helps in identifying the abnormal behavior in a graph which helps us in preventing fraud, spam or Ddos attack. Here, this paper considered that the outdegree of the node increases polynomial with time, hence the results seem to be more accurate than the previous work in the area rather than bounding the degrees by constants. With almost any network dataset,one does not have data reaching all the way back to the network’s birth. Due to this, there will be some effect of increasing outdegree simply because some nodes point back to nodes prior to the beginning of observation period. Even in the presence of noise, changing external conditions, and disruptions to the Internet we observe strong super-linear growth. The problems that this problem in the paper faces presents with new observations that need to be noted down while carrying out our research. Computing the shortest path between various nodes is expensive and hence for sampling, a subset of nodes is chosen and an exhaustive BFS to all other nodes is carried out. Also the connectivity needs to be taken care of. Using the parameters computed from the single giant component does not change the results presented in the paper. Observing the diameter with time shows that there is no effect of the “missing past”. The density power law has a condition where the nodes should form communities within communities and if the cost of cross community edges is scale free. Few of the observations from forest fire model are: heavy tailed in degree where the highly linked nodes follow rich get richer. The model also has copying flavor where a newcomer copies several of the neighbors of his or her starting point from where the current node is reached from. When the link formation is recursive, then the out degrees are also heavily tailed. The intuition behind the community guided attachment is that each node has a center of gravity and the probability of linking to other nodes decreases with their distance from the center of gravity. There’s also that, when a new node is added it adds a lot of outlinks and hence these nodes cause skewness in the out degree distribution and they also serve as bridges that connect formerly disparate parts of the network bringing the diameter down. There are also cases where the “orphan” nodes and multiple ambassadors are considered. In the basic model of Forest fire, we chose an ambassador w and then a new node chooses this node at random and forms a link and then with a probability choses x number of inlinks and outlinks with probability of outlinks being r times less that the probability of inlinks. Then recursively start from each outlink node of w preventing visiting of previously visited nodes and also preventing cycling. There are several theorems that are presented in the paper which are useful in dealing with similar problems.

In the community guided attachment model (when generalized by making it dynamic) , the in degree distribution follows a Zipf distribution. This depends on the difficulty constant and branching factor of the graph.

Q6.

The study article investigates how a city's street network, or the arrangement of its streets, might provide insight into the spatial organization and order of the city. Using information from OpenStreetMap, the author, Geoff Boeing, examines the orientation, configuration, and entropy (a measure of disorder) of street networks in 100 cities across the globe.
In his study, Geoff Boeing analyzes the spatial arrangement and organization of roadway networks in 100 cities around the world using data from OpenStreetMap. According to the study, cities in the US and Canada have less turmoil and a more grid-like street network than cities in other parts of the world. The creation of new techniques for calculating and representing the global spatial hierarchy of cities is an example of how the topic is being applied to urban planning and transportation. The analysis takes into account a number of variables, such as the direction of the streets, the width of the street segments, the number of intersections, and the general layout of the street network. To measure how well a city's street network adheres to the geometric ordering logic of a grid, Boeing develops a new indicator termed "orientation-order". The redundancy of the indicators, network order in the streets, high entropy in the cities, and circuity are significant findings. The cities with the lowest direction entropies were Chicago, Miami, and Minneapolis, all of which are known for having efficient road networks. Street networks are less organized in cities with significant entropy, such as Charlotte. The average length of a roadway section was found to vary significantly between locations, with Helsinki, Mogadishu, and Venice reporting the shortest median street segment lengths. The networks in Circuity were found to be the least complicated, while the most complex networks were found in Caracas, Hong Kong, and Sarajevo. Street orientation entropy and circuity were found to be lowest in American and Canadian cities and highest in European cities. The study revealed patterns of order, entropy, and other network features in many cities, revealing trends in urban design and development across regions. The findings provided insight into the variety of urban roadway networks and their distinguishing characteristics.This study provides a unique metric, as an indication of street orientation order, to enable significant conclusions about the topologies of various metropolitan networks. Charlotte's low of 0.002 exemplifies disorder and indicates a lack of clear orientation, while Chicago's high of 0.90 illustrates the city's almost perfect grid. Cities in the United States and Canada have highly organized networks and low entropy, whereas their European equivalents have greater entropy and less of a grid-like organization, according to a cluster study of areas. The top tertile is primarily made up of the United States and Canada.
The lowest direction entropies were found in three American cities, Chicago, Miami, and Minneapolis, indicating well structured street networks. The 16 cities with the lowest entropies, it turns out, were all located in the US or Canada. Cities outside of this region with low orientation entropies were Mogadishu, Kyoto, and Melbourne. The US city of Charlotte scored the highest entropy, which denotes a more disorganized roadway network. The next two cities with the highest entropy were Rome and So Paulo. Chicago has the closest resemblance to a single

ideal grid, according to the indicator, which represents orientation order, with a of 0.90. Charlotte, on the other hand, had almost perfect disorder in street orientations, with a of 0.002. The shortest median street segment lengths were found in cities like Venice, Mogadishu, and Helsinki, showing fine-grained networks. The longest were found in locations like Kiev, Moscow, and Pyongyang, which suggests coarse-grained networks. The least convoluted networks were found in Buenos Aires, Detroit, and Chicago, whereas the most circuitous networks were found in Caracas, Hong Kong, and Sarajevo.

US/Canadian cities have the largest median street segment lengths and average node degrees, as well as the lowest average street orientation entropy and circuity. While the proportion of dead ends and street orientation entropy were highest in European towns.

The study greatly advances knowledge of urban spatial organization and its influence on transportation dynamics. With the advent of the measure, it is now possible to quantify the order of orientation, and the thorough research of 100 cities reveals global trends. The study has practical implications for urban planning and design in addition to improving our understanding of urban morphology. The study lays the framework for further investigation into regional variations and the improvement of indicators for a more complex comprehension of urban spatial organization.