Question 7:

With the additional data added, the queries reluctantly ran slower. Graph databases are useful when dealing with unbounded queries, meaning the path to the desired data is unknown. The queries generally run slower anyways because the queries are all unbounded. When the larger amounts of data were add, this just made the paths to search for the desired data longer, which in turn made the queries execute more slowly. When rearranging the triples and or the nodes, we did not notice much of a speed difference. The reason for this is that even when the desired data is arranged to be the first or one of the first things examined, the query is still unbounded. This means that even if the desired data is found within the first portions of the search, the query engine does not know that no more relevant data exists after the point at which it found the data, so it must continue its search until it essentially reaches a bound or the end of the graph. In other words until the search examines all possible nodes and edges of the graph to determine that the data it found was all the relevant data pertaining to a particular query. The way that this could be sped up, is by specifying a known path to the desired data. If the path to the data is known and specified, then the search is essentially now bounded, it knows what path to follow to get to the desired data, rather than trying all possible paths until the data is found.