**IS 665 Data Analysis For Information Systems**

**Technical Assignment 1**

**Part I. Statistics (40 pts.)**

1. The file “assignment1\_actors.xls” contains information on 66 movie stars. This data set contains five variables:
   1. Gender
   2. DomesticGross: average domestic gross of star’s last few movies (in $ million)
   3. ForeignGross: average foreign gross of star’s last few movies (in $ million)
   4. Salary: current amount the star asks for a movie (in $ million)

**Question 1.** Using bin width of 2 (0-2, 2-4, 4-6, etc.), draw a histogram using the “salary” variable using Excel

**Question 2.** Is there an association between a star’s domestic and foreign gross? Choose an appropriate tool to analyze this problem. Show your work in Excel, and put some brief analysis under the chart / table / picture you have in Excel

1. The file “Assignment1\_Commute.xls” contains average time it takes for a citizen of each metropolitan area to commute to work and back home each day.

**Question 3.** Generate a histogram for distribution of daily commute times. Are shorter or longer commute times generally more likely for these commuters? (Hint: analyze the **shape** of the distribution).

1. The file “Assignment1\_Mortgage.xls” contains annual interest rates on 30-year fixed mortgages over the years.

**Question 4.** What conclusions can be drawn through time series analysis generated from this data?

PLEASE USE ONE EXCEL FILE THAT CONTAINS FOUR WORKSHEETS TO ANSWER THE ABOVE QUESTIONS. EACH QUESTION SHOULD HAVE ITS OWN WORKSHEET.

**Part II. Data Structure and Algorithms (60 pts.)**

**In class, we discussed array and linked list’s operations such as insertion and deletion. However, I left out an important operation: SEARCH**

**Question 1.**  Assume that you have an SORTED array of records. Assume that the length of the array (n) is **known**. Give TWO different methods to SEARCH for a specific value in this array. You can use English or pseudo-code for your algorithm. What is the time complexity for each algorithm and why?

**Question 2.**  Assume that you have a linked list of records. Assume that you have a **head**, a **current**, and a **tail** pointer. Write an algorithm that **swaps the data in the current node and the node after it**. You can use pseudo-code, English or drawing to describe your solution.

**Question 3.**  Assume that you have a linked list of records. Assume that you have a **head**, a **current**, and a **tail** pointer. Write an algorithm that **DELETES the node BEFORE the current node**. You can use pseudo-code, English or drawing to describe your solution.( this was, and remains to be, a popular technical interview question)