Assignment 4

Q1

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.9544	0.6888	8.0502	0.0045
Age		1	-0.0303	0.0122	6.1345	0.0133
DistanceFromHome		1	0.0336	0.00949	12.5079	0.0004
Education	1	1	0.1647	0.5457	0.0911	0.7628
Education	2	1	0.1658	0.5261	0.0993	0.7527
Education	3	1	0.3925	0.5099	0.5926	0.4414
Education	4	1	0.1564	0.5170	0.0915	0.7622
EmployeeID		1	-0.00002	0.000131	0.0137	0.9068
EnvironmentSatisfact		1	-0.3614	0.0724	24.9185	<.0001
MonthlyIncome		1	-0.00006	0.000032	4.0950	0.0430
NumCompaniesWorked		1	0.1438	0.0319	20.2931	<.0001
TotalWorkingYears		1	-0.0444	0.0212	4.4141	0.0356
MaritalStatus	Divorced	1	-1.1959	0.2299	27.0649	<.0001
MaritalStatus	Married	1	-0.8974	0.1751	26.2560	<.0001
Gender	Female	1	-0.3105	0.1643	3.5697	0.0588
OverTime	No	1	-1.6396	0.1646	99.2184	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Age	0.970	0.947	0.994
DistanceFromHome	1.034	1.015	1.054
Education 1 vs 5	1.179	0.405	3.436
Education 2 vs 5	1.180	0.421	3.310
Education 3 vs 5	1.481	0.545	4.022
Education 4 vs 5	1.169	0.424	3.221
EmployeeID	1.000	1.000	1.000
EnvironmentSatisfact	0.697	0.605	0.803
MonthlyIncome	1.000	1.000	1.000
NumCompaniesWorked	1.155	1.085	1.229
TotalWorkingYears	0.957	0.918	0.997
MaritalStatus Divorced vs Single	0.302	0.193	0.475
MaritalStatus Married vs Single	0.408	0.289	0.575
Gender Female vs Male	0.733	0.531	1.012
OverTime No vs Yes	0.194	0.141	0.268

Code:

```
data ibm_data;
  set sasdata.ibm_data;
    if Attrition="Yes" then new_attrition=1;
     else new_attrition =0;

     if Gender="Female" then new_gender=0;
     else new_gender=1;

     if Overtime="Yes" then new_Overtime=0;
     else new_Overtime=1;


   run;


options nodate pageno=1 linesize=80 pagesize=60;

   proc means data=ibm_data;
      var age DistancefromHome  Education
          EmployeeID EnvironmentSatisfaction MonthlyIncome NumCompaniesWorked
             TotalWorkingYears
          new_attrition new_gender new_Overtime ;
   run;


   proc freq data=ibm_data;
       tables new_attrition new_gender;
```

```
   run;


   proc logistic data=ibm_data descending;
     class new_gender new_overtime education MaritalStatus Gender Overtime /
param=ref ;
     model new_ttrition = age DistancefromHome  Education
          EmployeeID EnvironmentSatisfaction MonthlyIncome NumCompaniesWorked
             TotalWorkingYears MaritalStatus
          Gender overtime;
   quit;
```

Q2)

The page rank is as follows:

rank_p50
0.000038
0.000098
0.3160516
0.0000527
0.3871196
0.2965495
0.0000906

Code:

```
data Node;
    infile datalines;
    input Node $ A B C D E F G  ;
    datalines;

A  0  1  0  0  0  0  0
B  1  0  0  1  0  0  1
C  1  0  0  1  0  1  0
D  1  1  0  0  0  0  0
E  0  0  1  0  0  0  0
F  0  0  0  0  1  0  0
G  0  1  0  0  0  0  1
;
run;

proc sql;
    create table matrix as
        select a/sum(a) as x1
              ,b/sum(b) as x2
              ,c/sum(c) as x3
              ,d/sum(d) as x4
              ,e/sum(e) as x5
              ,f/sum(f) as x6
              ,g/sum(g) as x7
        from Node
    ;
quit;


data rank;
    x1=1/7;
    x2=1/7;
    x3=1/7;
    x4=1/7;
      x5=1/7;
      x6=1/7;
      x7=1/7;
         output;
run;

proc iml;
 use matrix;
 read all var {x1 x2 x3 x4 x5 x6 x7} into M;
 PRINT M;

 use rank;
 read all var {x1 x2 x3 x4 x5 x6 x7} into rank_1;
 print rank_1;
 rank_2 = t(rank_1);
 print rank_2;
```

```
 rank_p2=M *rank_2;
 print rank_p2 ;
 rank_p50=(M**50)*rank_2;
 print rank_p50 ;

quit;
```