ASSIGNMENT 4

1. The dataset "assignment4.sav" includes scores of student academic portfolios provided by 4 different raters. Take the data and conduct a G-study using a single facet (rater) design. Report the proportion of variance in the scores are due to each of the model terms.

ANSWER:

```
Summary of G-Study

                  Systems      Queries Interaction
              ----------- ------------ -----------
Variance          1.3934     0.010793     0.69453
Variance(%)       66.393      0.51424      33.093
---
Mean Sq.          6.2682       1.7414     0.69453
Sample size           97            4         388
```

1. Raters (Systems) accounted for 66.393% of item response variance
2. Items (Queries) accounted for 0.51424%
3. Error (Interaction) accounted for 33.093%

2. What interpretations can you make regarding the measurement process, given the percentages associated with each of the terms in the model? .

The majority of item variance is associated with raters (Systems) (66.393%) with the second highest proportion associated with error (the interaction) (33.093%). Very little variance was

associated with the items themselves (0.51424%), indicating that the item response patterns among the raters are very consistent.

3. **Using the results from Q1, conduct a D-study using between 1 and 10 raters. Report your results for both norm and criterion reference-based assessment. Do they differ? .**

```
> dstudy(assignment4.g, queries = c(10))

Summary of D-Study

Call:
    queries = 10
  stability = 0.95
      alpha = 0.025

Stability:
                              Erho2                                    Phi
              ------------------------------     ------------------------------
    Queries   Expected    Lower     Upper        Expected    Lower     Upper
              ------------------------------     ------------------------------
         10    0.95252   0.93321   0.96695         0.95182   0.9212   0.96563

Required number of queries:
                              Erho2                                    Phi
              ------------------------------     ------------------------------
  Stability   Expected    Lower     Upper        Expected    Lower     Upper
              ------------------------------     ------------------------------
       0.95         10        7        14              10        7        17
```

ANSWER:

G coefficient = 0.95252

Phi coefficient = 0.95182

The two results look like the same

4. **Given the results that you obtained in answering Q1 and Q2, what would be the number of raters needed to have the coefficient in exceed of .80**

```
> dstudy(assignment4.g, queries = c(1,2,3,4,5,6,7,8,9,10,11,12,13))

Summary of D-Study

Call:
     queries = 1 2 3 4 5 6 7 8 9 10 11 12 13
   stability = 0.95
       alpha = 0.025

Stability:
```

			Erho2				Phi
Queries	Expected	Lower	Upper	Expected	Lower	Upper	
1	0.66736	0.58286	0.74527	0.66393	0.53895	0.7375	
2	0.8005	0.73647	0.85404	0.79803	0.70041	0.84892	
3	0.85753	0.80739	0.89772	0.85563	0.77812	0.89394	
4	0.8892	0.84824	0.92128	0.88767	0.82382	0.91829	
5	0.90935	0.87479	0.93601	0.90807	0.8539	0.93354	
6	0.9233	0.89343	0.9461	0.9222	0.87522	0.944	
7	0.93353	0.90724	0.95344	0.93256	0.8911	0.95161	
8	0.94135	0.91789	0.95903	0.94049	0.9034	0.9574	
9	0.94752	0.92634	0.96341	0.94675	0.9132	0.96196	
10	0.95252	0.93321	0.96695	0.95182	0.9212	0.96563	
11	0.95665	0.93891	0.96986	0.95601	0.92784	0.96866	
12	0.96012	0.94372	0.97231	0.95953	0.93346	0.97119	
13	0.96307	0.94782	0.97438	0.96252	0.93826	0.97335	

```
Required number of queries:
```

			Erho2				Phi
Stability	Expected	Lower	Upper	Expected	Lower	Upper	
0.95	10	7	14	10	7	17	

a. 1) if you were using the scores to rank examinees => **We need 2 raters (G coefficient = 0.8005) in exceed of .80**
b. 2) if you were using scores to determine a passing grade, respectively? => **We need 3 raters (Phi coefficient = 0.85563) in exceed of .80**