**Hypothesis Testing Lab**

**ANSWER KEY**

For this lab we will use four datasets:

CAMPNET:

This is a dichotomous adjacency matrix of 18 participants in a qualitative methods class. Ties are directed and represent that the ego indicated that the nominated alter was one of the three people with which s/he spent the most time during the seminar.

ZACKAR & ZACHATTR:

ZACKAR is another stacked dataset, containing a dichotomous adjacency matrix, ZACHE, which represents the simple presence or absence of ties between members of a Karate Club, and ZACHC, which contains valued data counting the number of interactions between actors. ZACHATTR is a rectangular matrix with three columns of attributes for each of the actors from the ZACKAR datasets.

KRACK-HIGH-TEC & HIGH-TEC-ATTRIBUTES

KRACK-HIGH-TEC is another stacked dataset, containing three dichotomous relations (REPORTS\_TO, ADVICE, FRIENDSHIP). HIGH-TEC-ATTRIBUTES contains several attributes about the nodes in KRACK-HIGH-TEC, including Age, Level (CEO, Manager, Staff), Tenure, and Department.

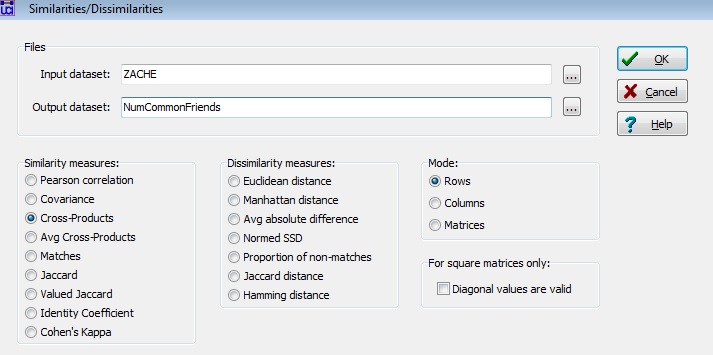
WIRING

This is a stacked dataset that includes many different files. This is a dichotomous adjacency matrix of 14 employees of the bank wiring room of Western Electric used in the famous Hawthorne Studies. Ties are symmetric and represent participation in games during work breaks. RDGAM records people playing games together, RDCON records conflict between people, RDPOS is positive interactions, RDCON is negative interactions.

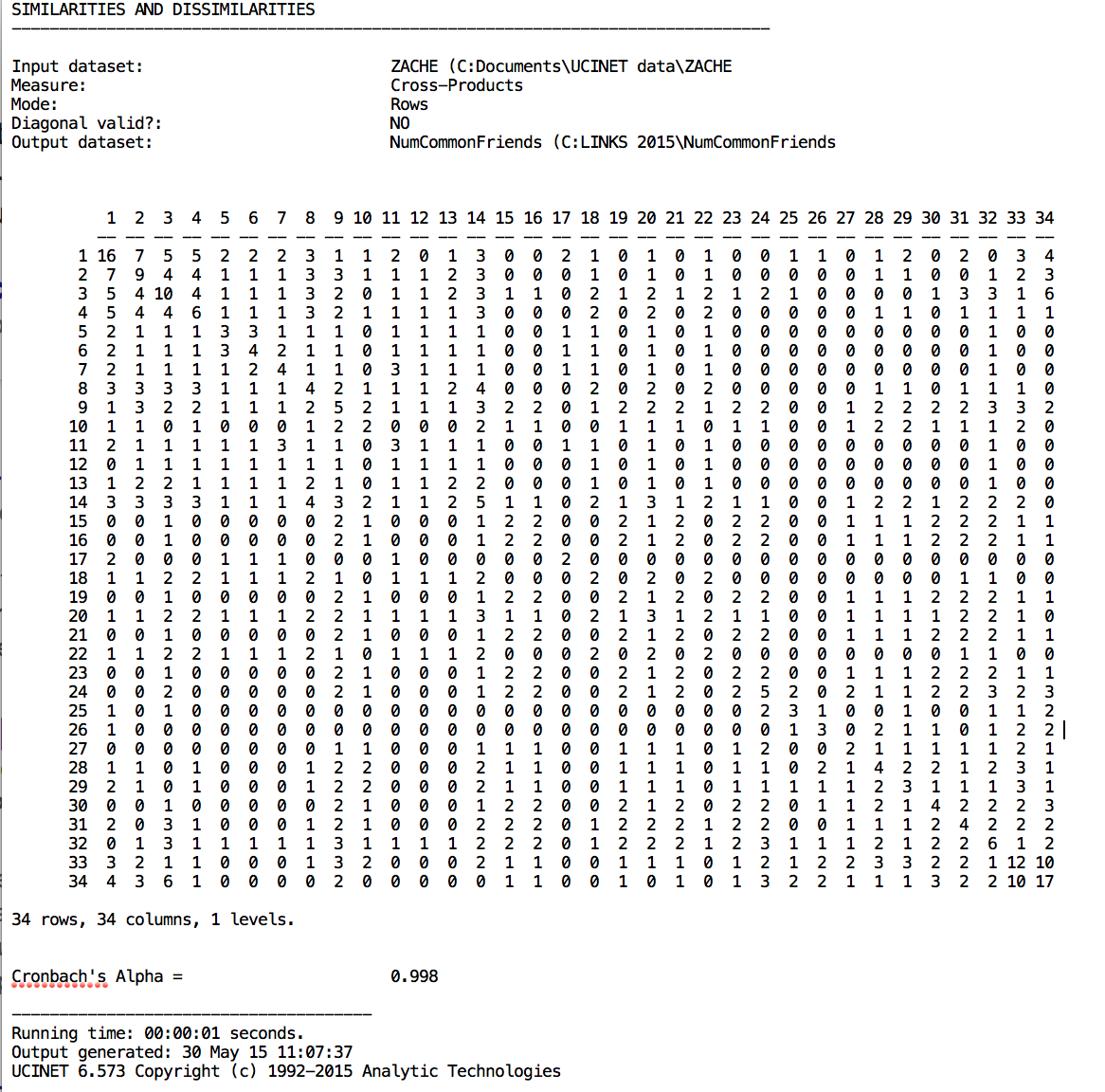
**EXERCISES**

1. Testing dyadic hypothesis
   1. Run Data | Unpack on ZACKAR (if you have not yet), which will create ZACHE and ZACHC. ZACHE has dichotomous data about the ties and ZACHC has valued data (the strength of ties).
   2. Run Tools | Similarities and use the cross-product measure to compute similarities among the rows of ZACHE. (The cross product is a very powerful and common matrix operation that, in this case, will count how many friends each pair of actors have in common.) Call the output **NumCommonFriends**.

The cross-product operation counts the common third parties between actors, specifically how many common friends they have. For Similarities & Distances the input data is ZACHE, the output dataset is called “NumCommonFriends”. Under Similarities measures, choose “Cross-Products” (we are using similarities, so no need to select Dissimilarity measures). Finally, for Mode, select Rows (because we are interested in how each person in the row has common ties with those in the columns). Diagonal can be left on or off, it won’t change the results here (the diagonal, in this case, would indicate that i is friends with j and j is friend with i—or the number of i’s ties that are reciprocated. Note: the ZACHE network is undirected, so all ties are by definition symmetric. The diagonal is then simply i’s degree centrality).



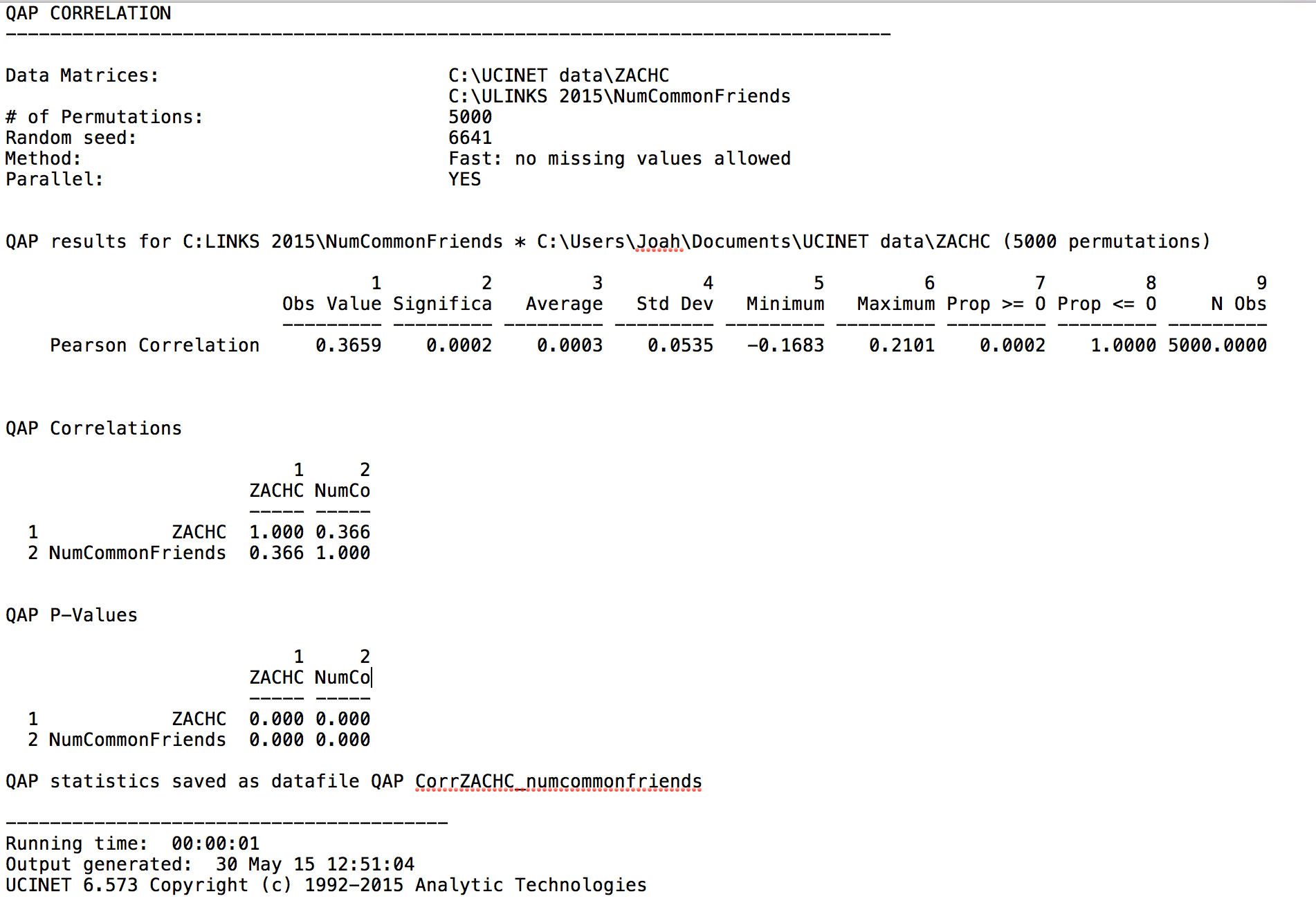
The operation counts the common third parties (shared ties) by multiplying the matrix (ZACHE) by its transpose (ZACHE transposed), thus each ij cell in NumCommonFriends give the number of common friends that i and j share with each other. For example, a cell with a value of 3 would indicate that i and j have 3 common friends. See the resulting matrix below:



* 1. Go to Tools | Testing Hypotheses | Dyadic (QAP) | QAP Correlation and browse to include both ZACHC and **NumCommonFriends** to be correlated and click okay. What do the results mean?

The correlation for ZACHC (valued friendship ties) and NumCommonFriends is .3659 (p = .0002). This means that there is a positive correlation between the strength of ties and the number of shared friends. Thus, as friendship between two people becomes stronger, there are likely to be more instances of shared friends. See QAP correlation output below:

­­­­­

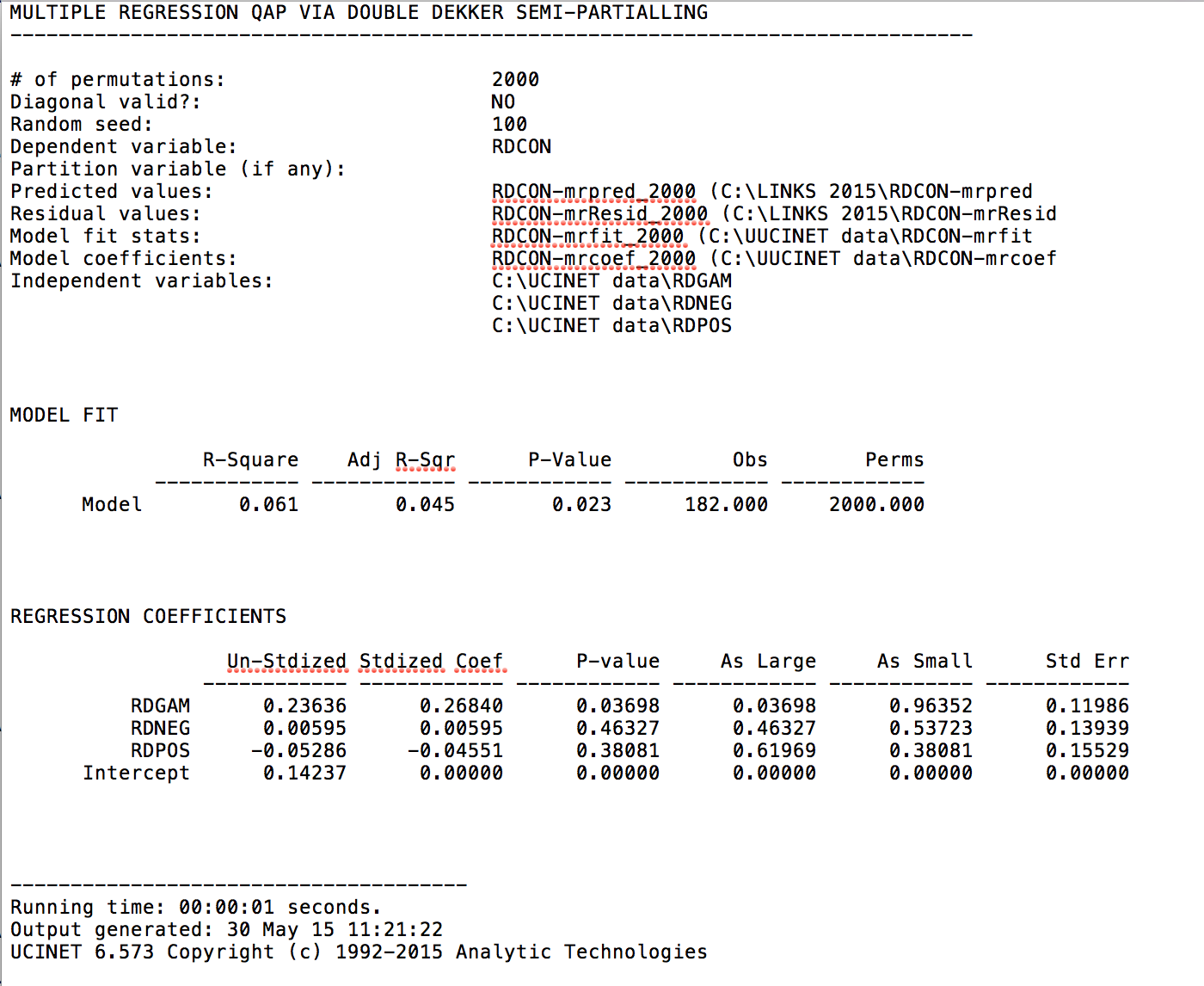


* 1. Congratulations, you have just statistically demonstrated the first part of Granovetter’s famous “strength of weak ties” theory, which states that I have stronger ties (ZACHC) with those people with whom I share more friends in common (**CommonFriends**).

1. Testing multivariate dyadic hypotheses
   1. Unpack the WIRING dataset if you have not done so yet.
   2. Go to Tools | Testing Hypotheses | Dyadic (QAP) | MR-QAP Linear Regression | Double-Dekker Semi-Partialling MRQAP. Put RDCON (conflict between members about whether the windows should be open or shut) in as the dependent variable. Put in RDPOS (positive relationships), RDNEG (negative relationships), and RDGAM (playing games together) in as independent variables. Before running it, what do you think would most significantly predict conflict? After running it, are your results what you expected? How would you explain the results?

A key question in making such a prediction is what is the nature of the conflict? Answering that question will help guide your predications. Having conflict about whether windows are open or closed seems petty at first blush. Is such a conflict based on negative emotions or feelings, or based on simple disagreement about room temperature preferences? Before running this analysis, it might be reasonable to think that negative relationships would be related to having such conflict. As individuals increase in the animosity or negative feelings toward others, they are likely to have additional disagreements that can manifest as conflict, which could be over simple things such as the state of the windows. While it is true that positive relationships can also relate to certain kinds of conflict (such as constructive conflict due to interdependencies within positive relationships), it isn’t clear from what we know about the data whether this would be the case. Playing games together might also relate to conflict, particularly in the case of competitive games. If playing games together is by nature competitive, its reasonable to guess that such rivalry might spill over into tangential issues, such as windows being open or closed. An argument could be made for each relationship being related to conflict, but which makes the strongest case? A good guess would be negative relationships as the strongest, then game-playing, then finally positive relationships. See results of the MRQAP below:

MRQAP 2000 permutations run #1

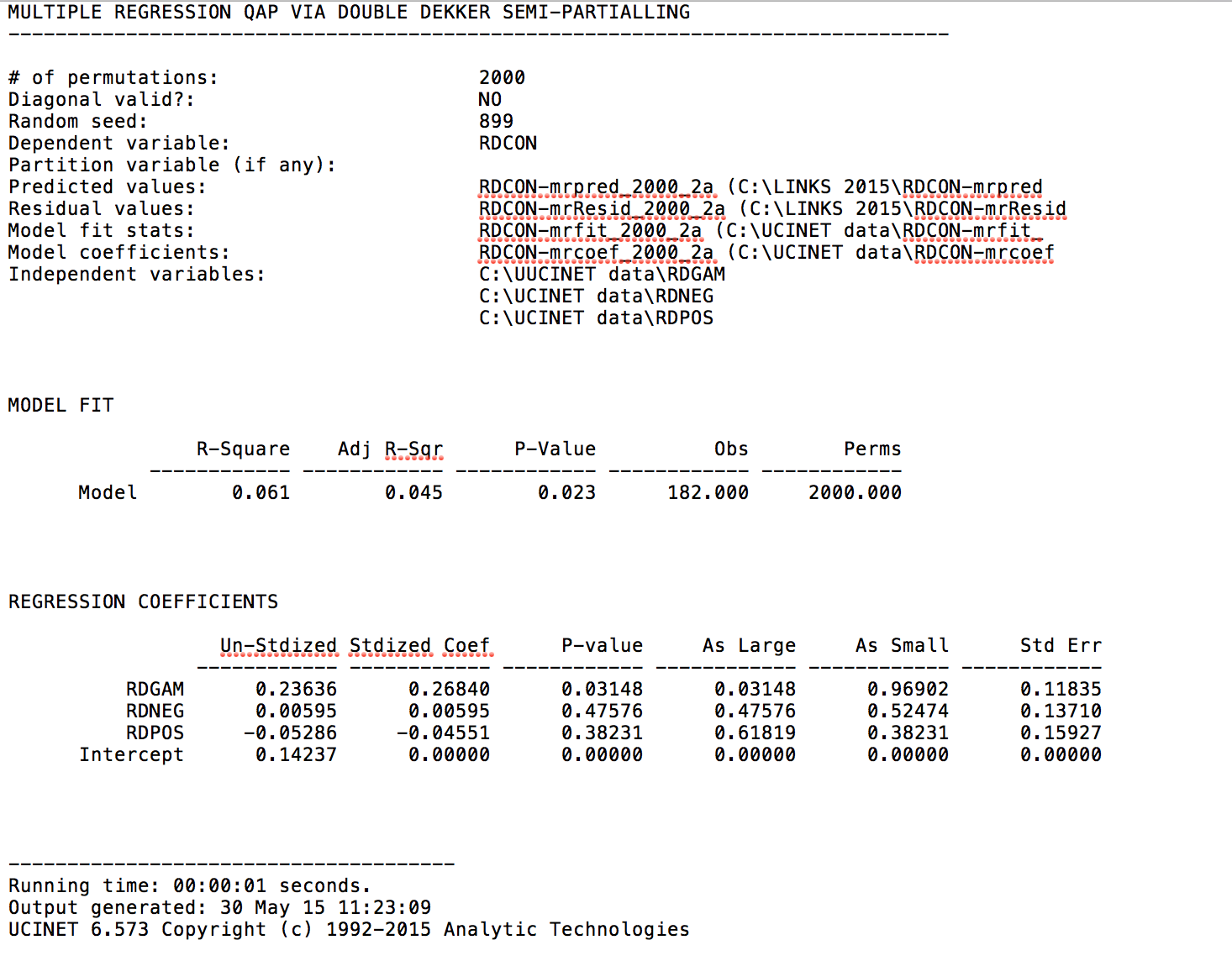


The predictor RDGAM is significant at the p < .05 level (b = 0.236, p = .0369) meaning that playing games with another person was related to having conflict about the windows with that person. Depending on your original prediction, you might have found support. However, it is interesting that neither negative relationships nor positive relationships were significantly related to conflict. This illustrates an important point that knowledge of the data, specifically the relationships and context of the sample, is extremely important in making predictions. The “good natured” conflict associated with playing games sparked conflict concerning whether the windows were opened or closed.

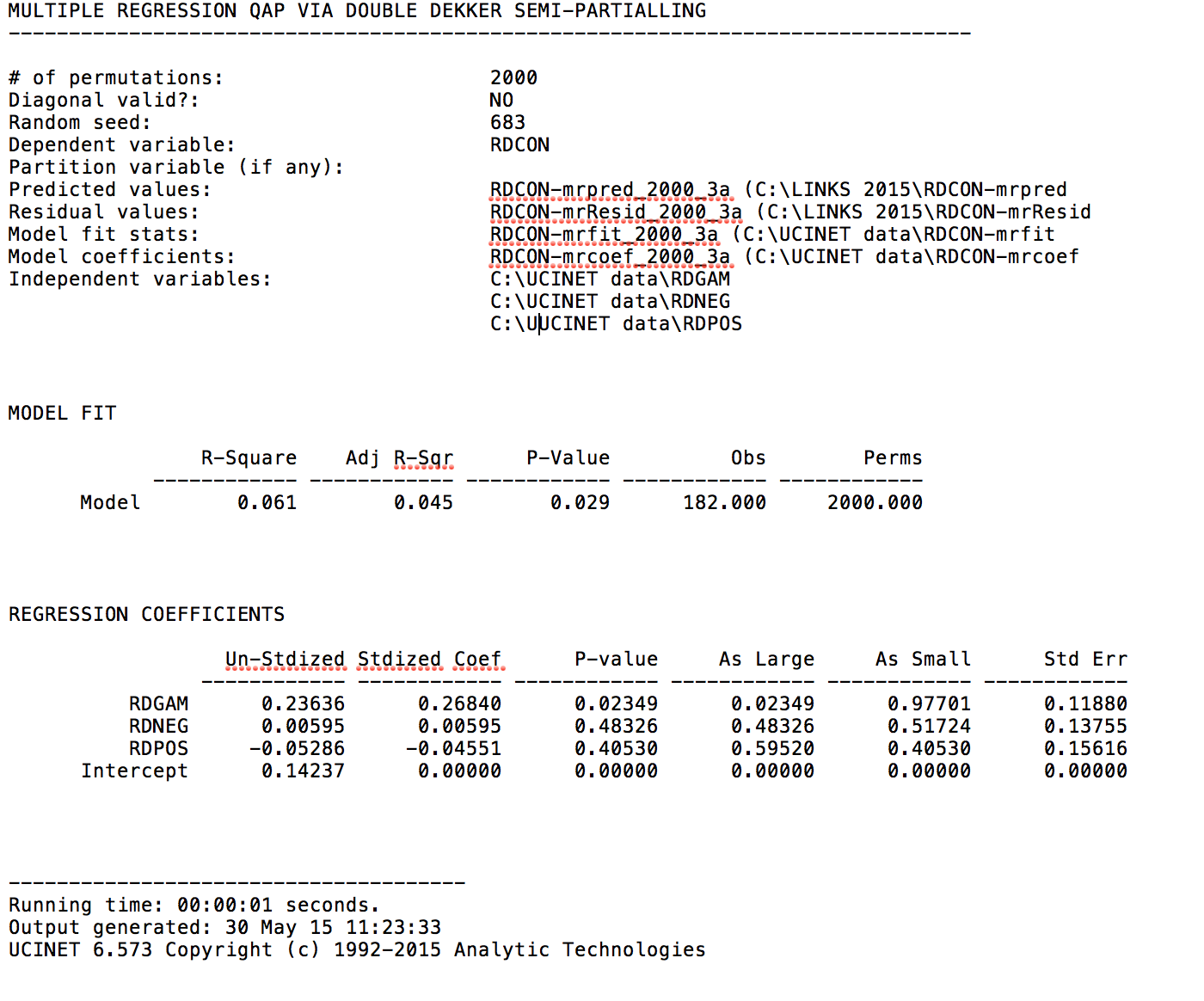
1. Record the standardized coefficient and significance for any significant predictor, and run the same procedure two more times (still using the default value of 2000 for the number of permutations) and record the same results. Now, run the same procedure three more times setting the number of random permutations set to 100000. Record the same results. How did the parameter affect the results? Why?

See results of additional MRQAP runs below:

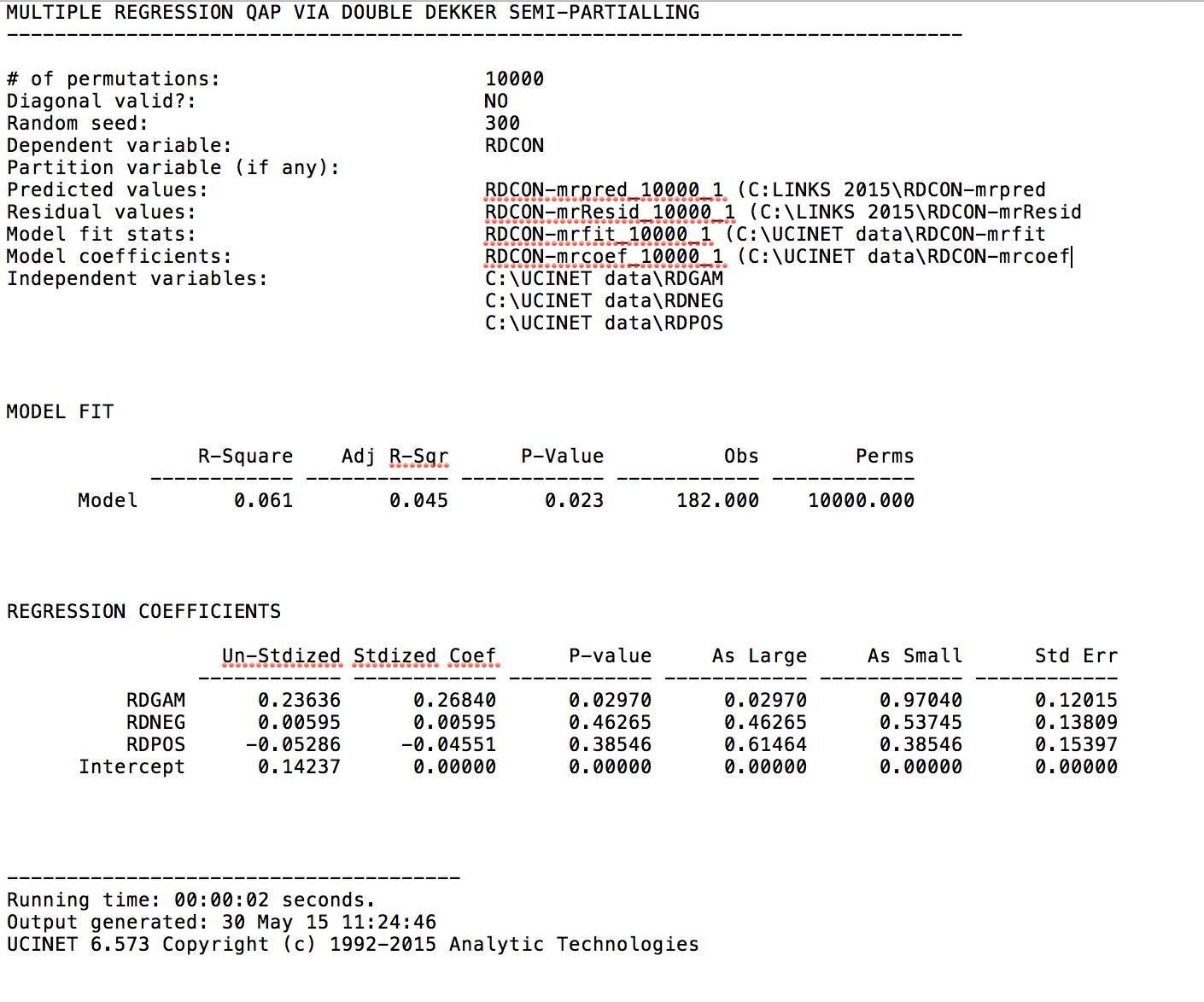
MRQAP 2000 permutations run #2



MRQAP 2000 permutations run #3



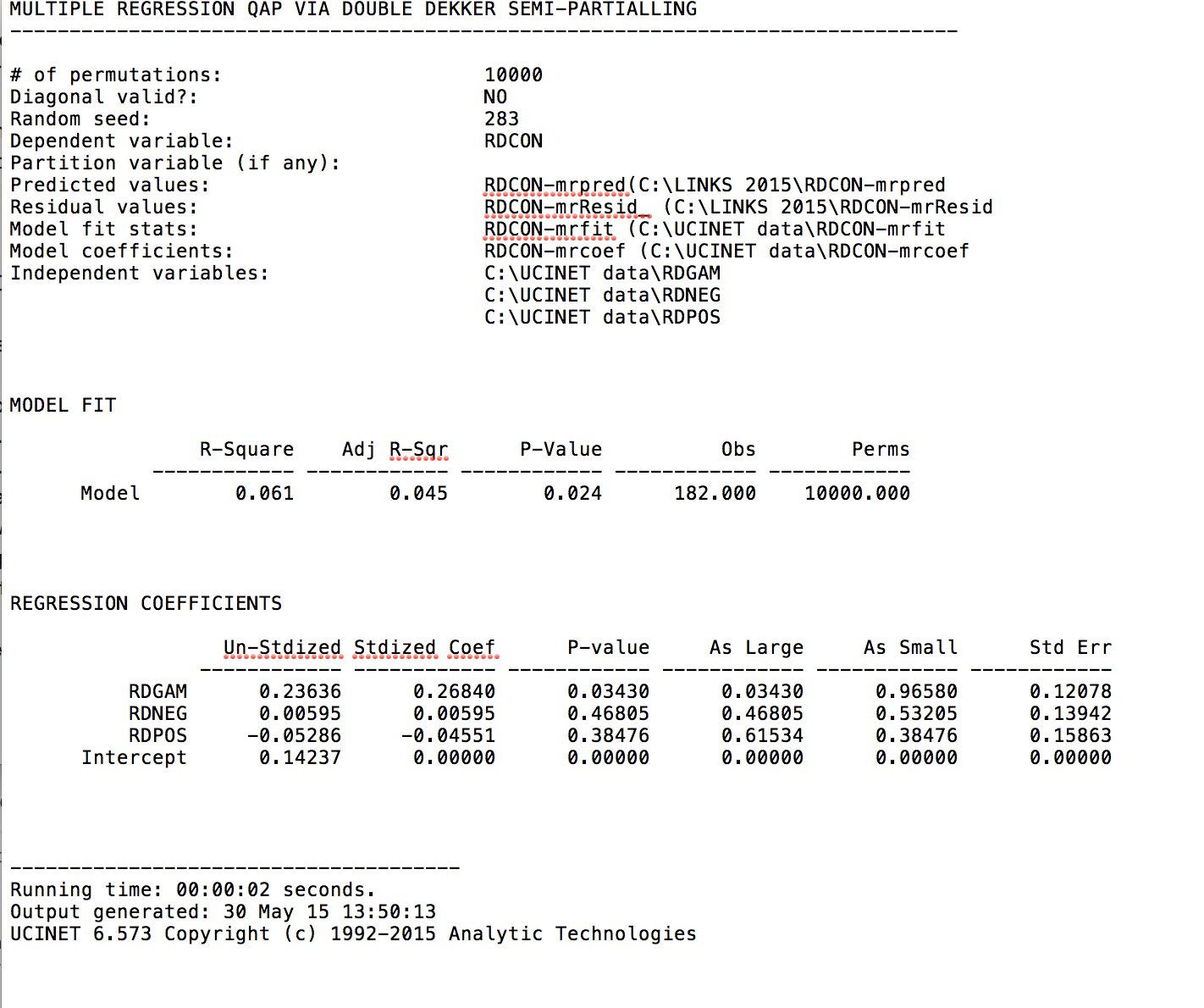
MRQAP 10000 permutations run #1



MRQAP 10000 permutations run #2



MRQAP 10000 permutations run #3



**Results Table (Note: p-values will differ slightly from your results)**

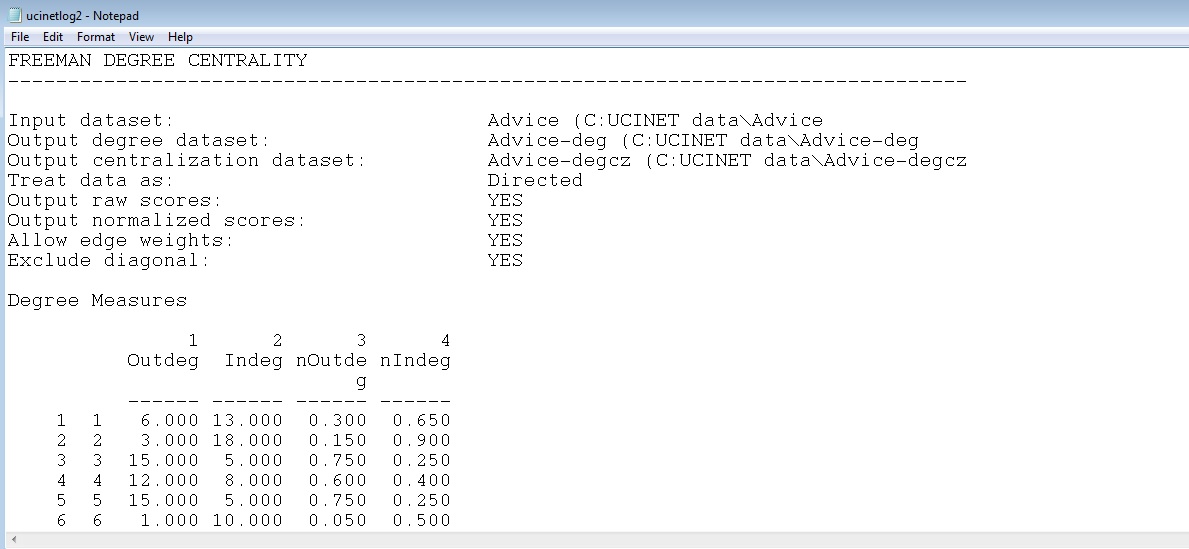
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MRQAP Model | Predictor | Coeff | P-value |  |
| 2000 permutations 1 | **RDGAM** | **0.236** | **0.0369** | **\*** |
|  | RDNEG | 0.005 | 0.4632 |  |
|  | RDPOS | -0.052 | 0.3808 |  |
| 2000 permutations 2 | **RDGAM** | **0.236** | **0.0314** | **\*** |
|  | RDNEG | 0.005 | 0.4757 |  |
|  | RDPOS | -0.052 | 0.3823 |  |
| 2000 permutations 3 | **RDGAM** | **0.236** | **0.0234** | **\*** |
|  | RDNEG | 0.005 | 0.4832 |  |
|  | RDPOS | -0.052 | 0.4053 |  |
| 10000 permutations 1 | **RDGAM** | **0.236** | **0.0297** | **\*** |
|  | RDNEG | 0.005 | 0.4626 |  |
|  | RDPOS | -0.052 | 0.3854 |  |
| 10000 permutations 2 | **RDGAM** | **0.236** | **0.0310** | **\*** |
|  | RDNEG | 0.005 | 0.4925 |  |
|  | RDPOS | -0.052 | 0.3823 |  |
| 10000 permutations 3 | **RDGAM** | **0.236** | **0.0343** | **\*** |
|  | RDNEG | 0.005 | 0.4680 |  |
|  | RDPOS | -0.052 | 0.3847 |  |

The significance levels changed slightly with some with the 2000 permutation tests (ranging from .0234 to .0369, or .0135). This is because it randomly creates different permutations and therefore it might randomly create more random assignments as extreme as the observed value. But, as we increase the number of permutations, the effect of any such outliers should be reduced, so the variability in the significance should be reduced (for example, the 10,000 permutation tests, the p-values only ranged from .029 to .034, or about .0046; considerably less than for the 2000 permutation tests). However, the RDGAM variable was significant at the p < .05 level for all models.

Also, the variability in the significance values should be reduced the most for truly significant coefficients. But, for ones that are not significant, the effects really are random and the variability in the significance measure will remain higher at higher numbers of permutations.

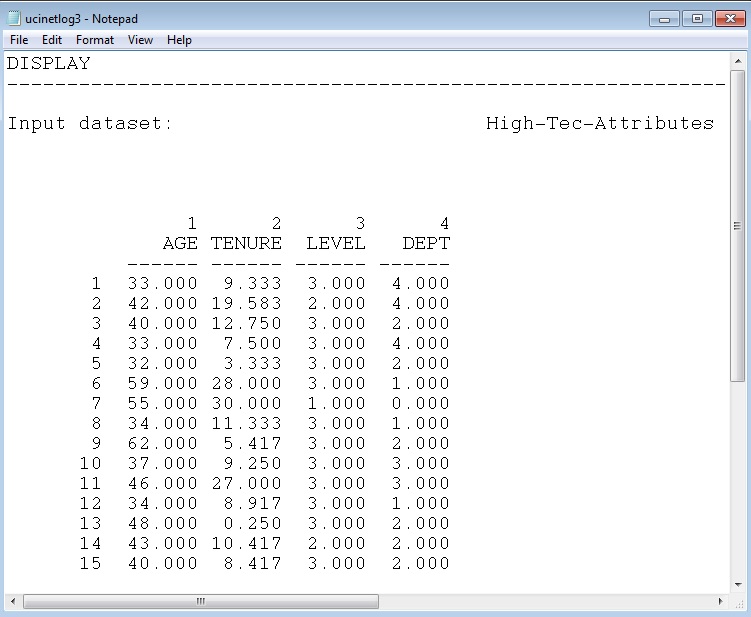
1. Testing monadic hypotheses.
   1. You should have already unpacked the KRACK-HIGH-TEC dataset, but if not, do so now. You will get three datasets (REPORTS\_TO, ADVICE, FRIENDSHIP). We are going to use the ADVICE dataset. Run Network | Centrality & Power| Degree on this dataset, specifying that the data are directed. By default, it will name the node-level output **ADVICE-Deg**.
   2. We are particularly interested in who is sought after for advice, which is captured by InDegree centrality. Display the output dataset (ADVICE-Deg) using the big “D” icon on the tool bar, and note what column has the measure “InDegree” (the number of people who sought advice from each actor).

See Advice InDegree output below:



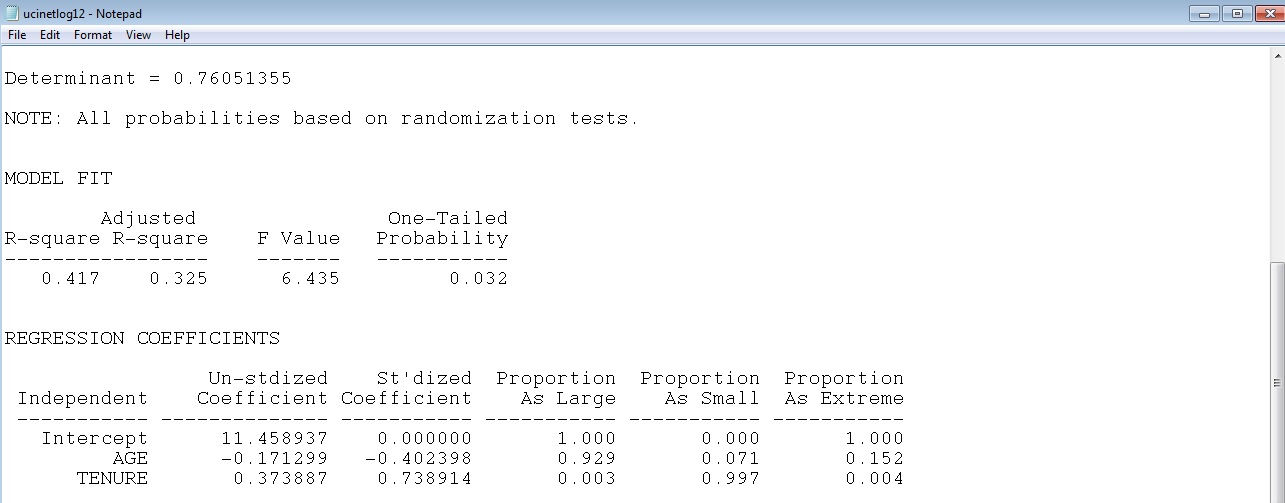
* 1. Display (Big D) the HIGH-TEC-ATTRIBUTES dataset to determine which columns the AGE and TENURE attributes are in.

See HIGH-TEC-ATTRIBUTES file below:



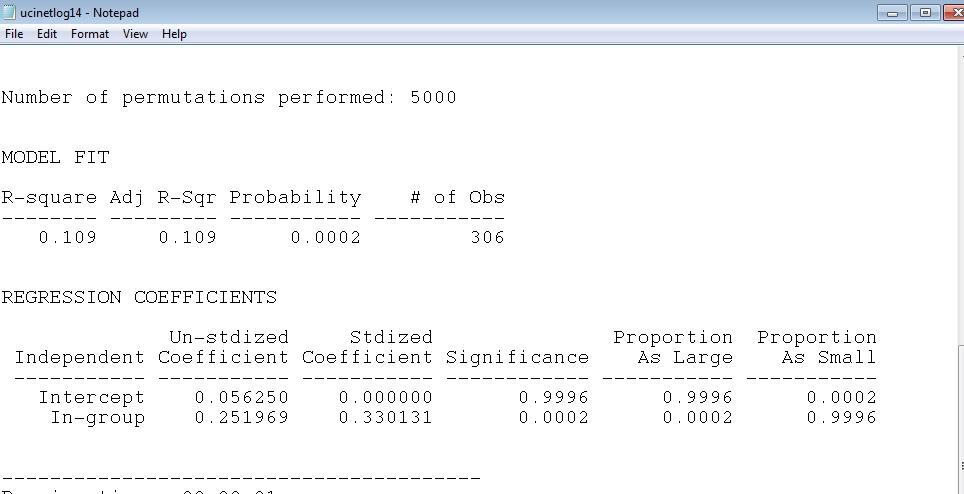
* 1. Now, it is common wisdom that people look to the “senior” people for advice, but is unclear in an organizational context whether senior is “older” or “longer tenured”. You will test if either of these is supported by the data. Run Tools | Testing Hypotheses | Node-Level | Regression specifying ADVISE for your dependent dataset with the appropriate column for InDegree and HIGH-TEC-ATTRIBUTES and the appropriate columns for your independent dataset (i.e., the columns for Age and Tenure separated by a space), and set the number of permutations to 10000. Which meaning of “senior” do the data support?

Based on the node-level regression, the coefficient for TENURE is significant at the p <. 01 level (b = .373, p = .004), while AGE is not (b = -.171, ns). Thus, the data indicate that TENURE is a better predictor of incoming advice than AGE. It is possible that AGE is conflated with TENURE, and AGE is negative because it is “over-fitting” the data. See results below:



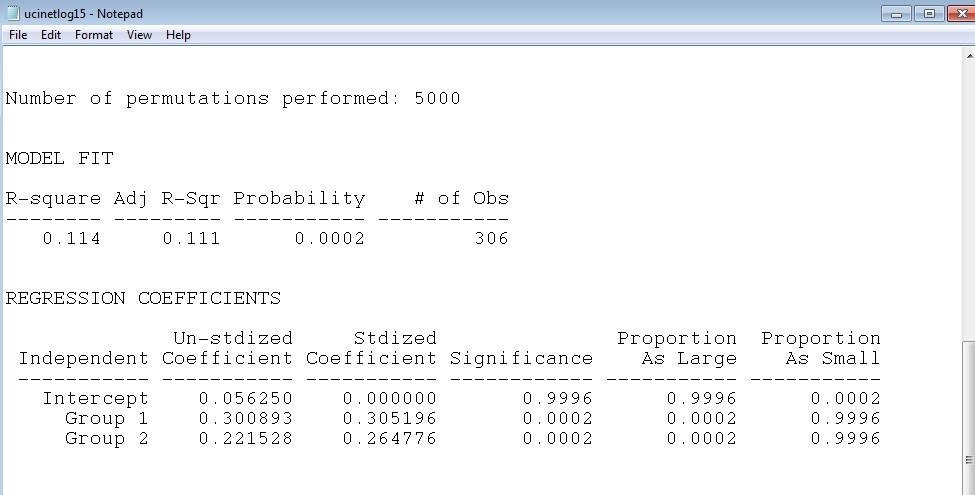
1. Testing Mixed-Dyadic Monadic hypotheses
   1. Since it is only fitting that we end where we started, we shall use the campnet data for these final exercises.
   2. You will run Tools | Testing Hypotheses | Mixed Dyadic/Nodal | Categorical attributes | Anova Density twice. For both, specify CAMPNET as the network matrix, and the gender column of the CAMPATTR matrix as the Actor Attribute. For the first run, choose “Constant Homophily” for your model, and for the second, choose “Variable Homophily”. Interpret both sets of results. What do they mean? Is there homophily? Which gender tends to be more homophilous?

Results of CAMPNET ANOVA with “Constant Homophily”:



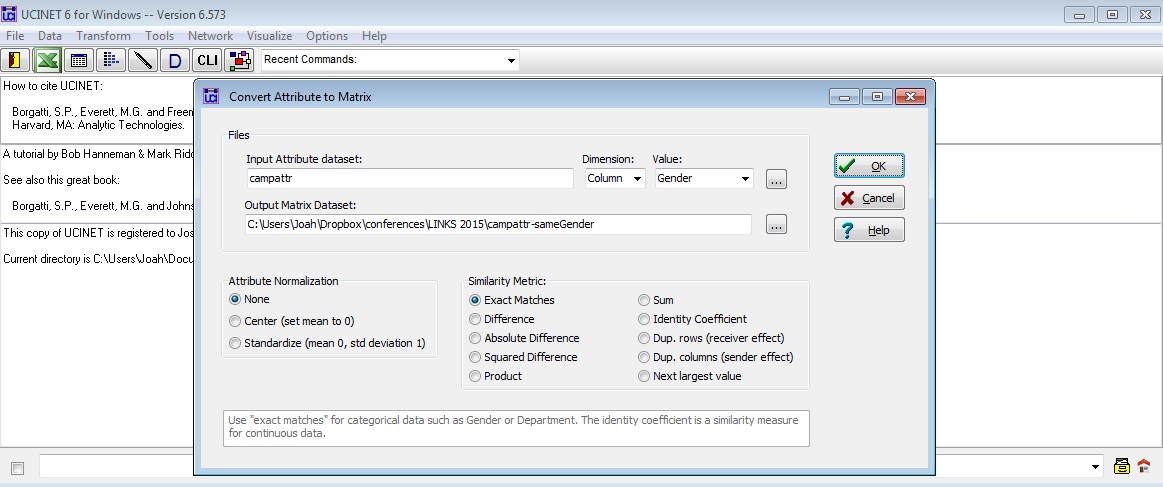
The coefficient for “in-group” is significant; this means that individuals were likely to associate with other who shared the same gender. Thus, homophily was a factor related to network ties.

Results of CAMPNET ANOVA with “Variable Homophily”:



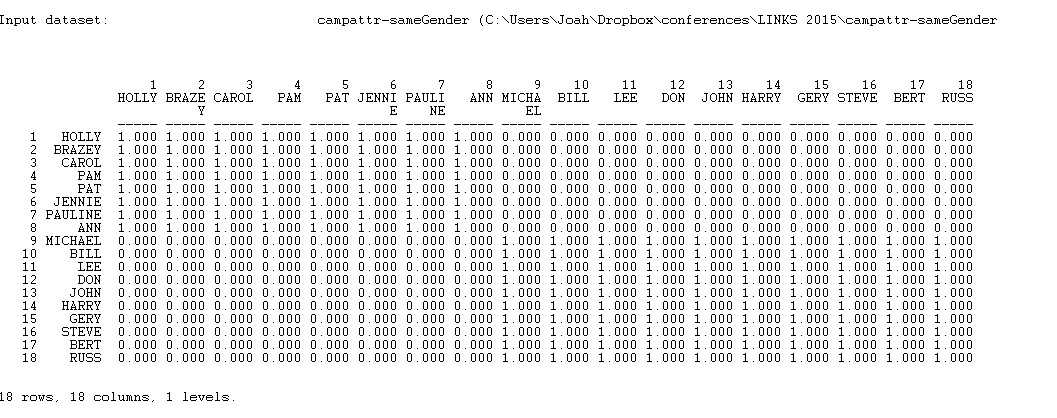
Results indicate that members of group 1 (female) associated with each other. Also, members of group 2 (male) associated with each other as well. Thus, homophily is working within specific groups (men to men and women to women). Based on these results, men (group 2) had a higher coefficient than women (group 1), but this difference is not that meaningful.

1. Using QAP for Mixed Monadic/Dyadic Hypotheses testing.
   1. Use Data | Attribute to matrix, create a matrix of exact matches among the actors in Campnet based on gender.

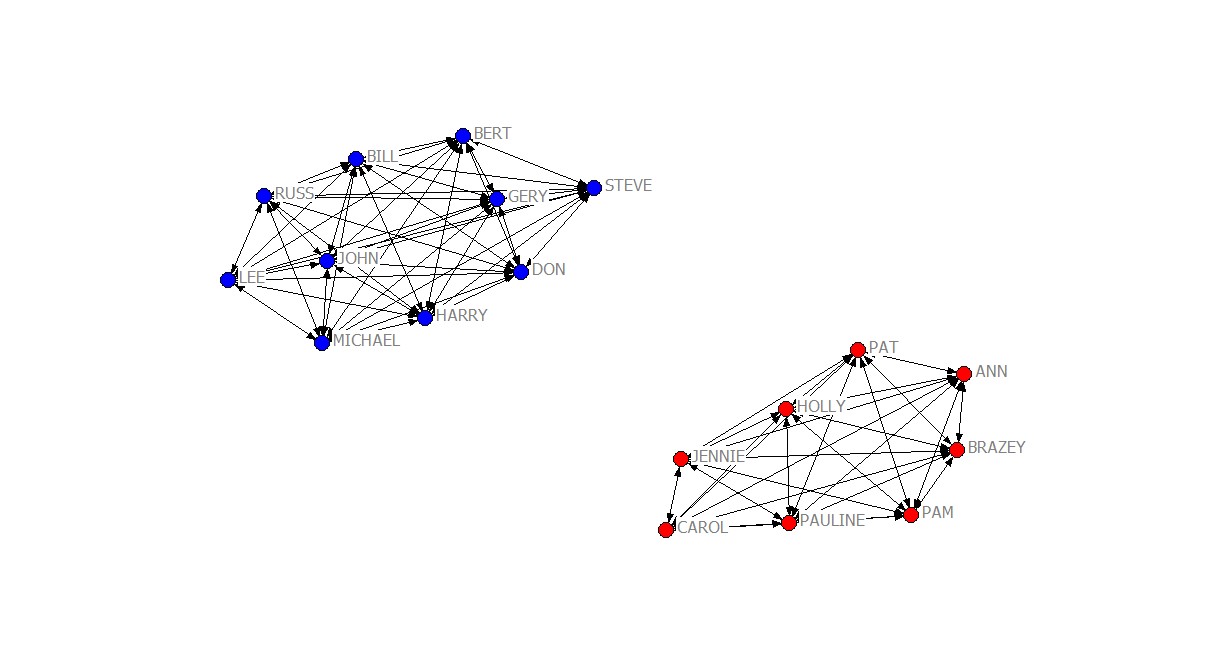


* 1. View this new matrix (named CAMPATTR-SameGender by default) in Netdraw. What does the diagram show?

The display of the UCINET file for “sameGender”:

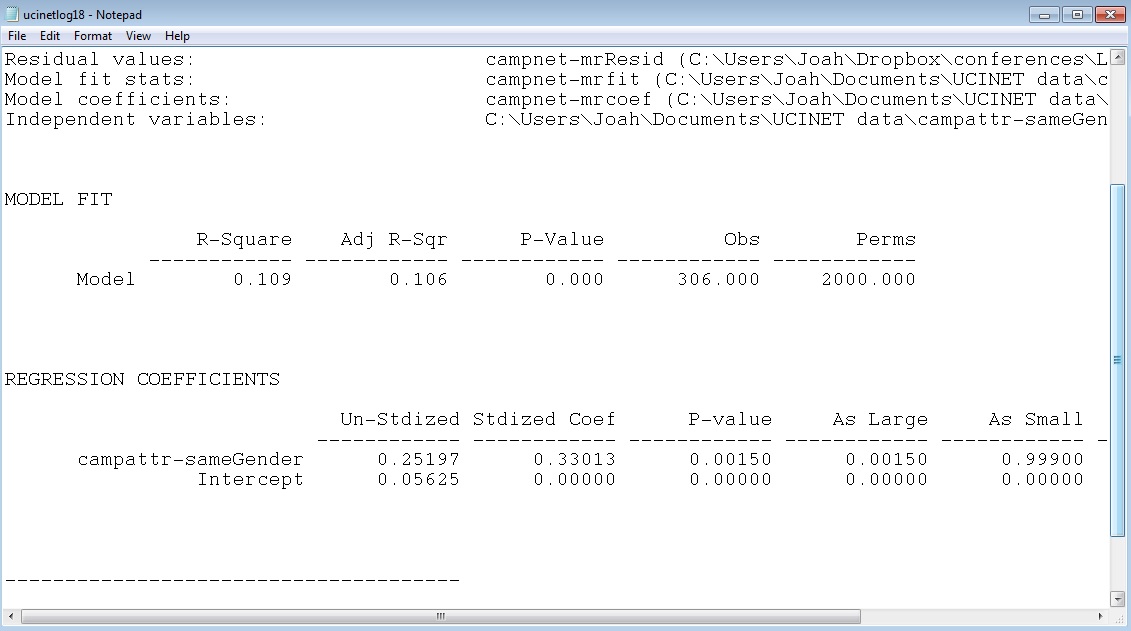


See the diagram of the above UCINET file in Netdraw below. Blue indicates male, red indicates female:



* 1. Use Tools | Testing Hypotheses | QAP Regression to regress the Campnet network on this new matrix of gender similarity, CAMPATTR-SameGender. What do the results show?

The MRQAP shows the same coefficients and r-squared as the “constant homophily” model ANOVA above (5b). Same gender is related to network ties.



* 1. Do you prefer this approach of the ANOVA Density Tables? When might you use each of these separate techniques?

ANOVA is useful when there is only one predictor variable (independent variable). However, with MRQAP, additional predictors (and controls) can be added to the model. Otherwise, the methods are equivalent.

* 1. When would you choose Moran’s I (or Geary’s C) instead of the ANOVA Density Tables? How would you use QAP to test for Autocorrelation in those cases?

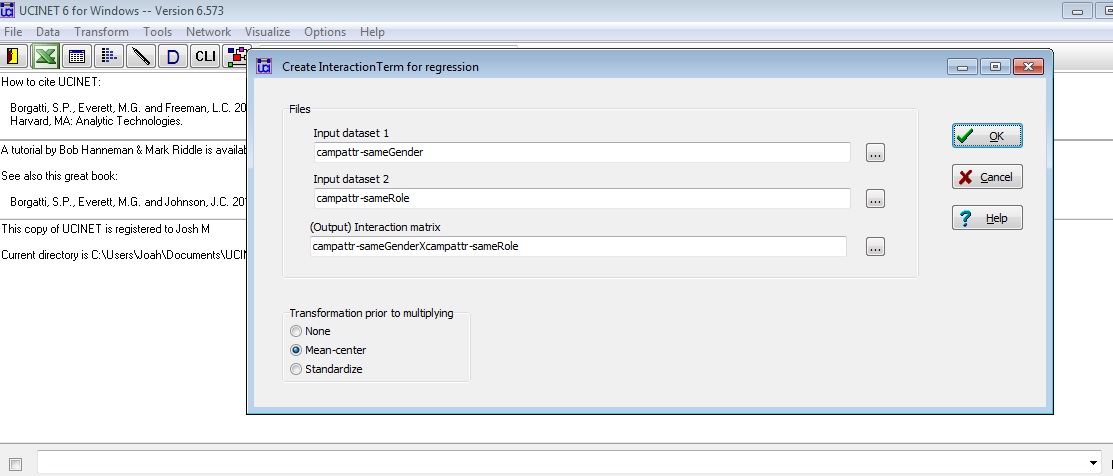
You would use Moran’s I (or Geary’s C) if the attribute (predictor) is continuous or interval/ratio. ANOVA works with categorical variables (binary predictors). For autocorrelation, you would use the same technique, but use something other than “exact match” when creating the “sameGender” variable. For example, you could use absolute value difference, or even product to construct the sameGender variable from the attribute data.

**ADVANCED [OPTIONAL]**

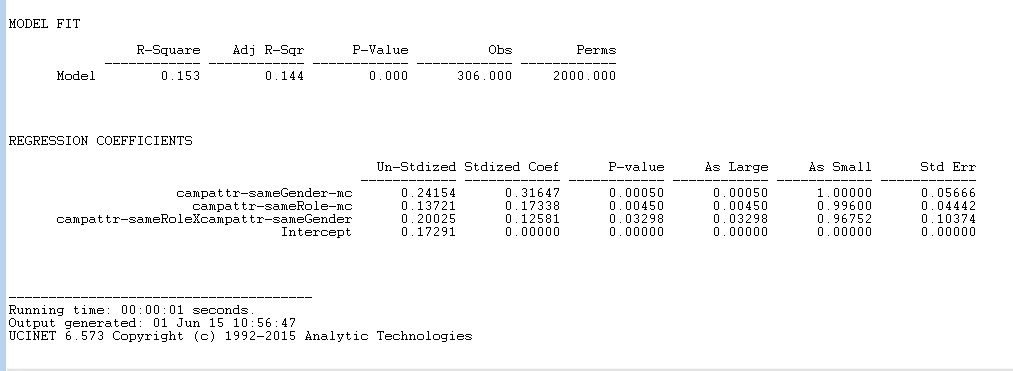
One reason you might be forced to use the QAP method for testing autocorrelation is you want to test two variables at once, and possibly even in interaction term. So, using the UCINET Data | Attribute to Matrix function, create a “SameRole” matrix as well.

Now, go to Transform | Make Interaction Term for Regression, and specify the two matrices you created (Campattr-SameGender and Campattr-SameRole by default), as an interaction term, and call the output file **Interaction** (the default name is too long to type).

Create interaction terms in UCINET. You can designate “Mean-center” or leave it as “None”-- it will not affect the results of the model. If you mean-center UCINET will create three new files: campattr-sameGenderXcampattr-sameRole, campattr-sameGender-mc, and campattr- sameRole-mc (the mc designates a mean-centered variable to use in the regression model).



Now, run the regression as you typically would to test for interaction. What are the results?



Based on the MRQAP above, the interaction term is significant at the p < .05 level (b = .200, p = .032). Thus, individuals with the same gender are more likely to associate if they also have the same role. Note that the r-squared has increased from .109 in the ANOVA (same gender) model to .153 in the MRQAP model with sameRole and the interaction term added. This seems to be a small increase.

Thank you all for a great workshop! Hope you got a lot out of it. We will post answers to Labs to the shared dropbox folder sometime soon!