

Andrews University Digital Commons @ Andrews University

Master's Theses

Graduate Research

2015


Ranking Interesting Changes in Correlation Coefficient Matrix Results from Varying Data Partitions in Causal Graphic Modeling

Yesica Daniela Bravo Gonzalez

Andrews University

This research is a product of the graduate program in [Engineering & Computer Science MS](#) at Andrews University. [Find out more](#) about the program.

Follow this and additional works at: <https://digitalcommons.andrews.edu/theses>

 Part of the [Applied Statistics Commons](#), and the [Longitudinal Data Analysis and Time Series Commons](#)

Recommended Citation

Bravo Gonzalez, Yesica Daniela, "Ranking Interesting Changes in Correlation Coefficient Matrix Results from Varying Data Partitions in Causal Graphic Modeling" (2015). *Master's Theses*. 66.
<https://digitalcommons.andrews.edu/theses/66>

This Thesis is brought to you for free and open access by the Graduate Research at Digital Commons @ Andrews University. It has been accepted for inclusion in Master's Theses by an authorized administrator of Digital Commons @ Andrews University. For more information, please contact repository@andrews.edu.



Seek Knowledge. Affirm Faith. Change the World.

Thank you for your interest in the

**Andrews University Digital Library
of Dissertations and Theses.**

*Please honor the copyright of this document by
not duplicating or distributing additional copies
in any form without the author's express written
permission. Thanks for your cooperation.*

ABSTRACT

RANKING INTERESTING CHANGES IN CORRELATION
COEFFICIENT MATRIX RESULTS FROM
VARYING DATA PARTITIONS IN
CAUSAL GRAPHIC MODELING

by

Yesica Daniela Bravo Gonzalez

Chair: Roy Villafane, Ph.D.

ABSTRACT OF GRADUATE STUDENT RESEARCH

Thesis

Andrews University

College of Arts and Sciences

Title: RANKING INTERESTING CHANGES IN CORRELATION COEFFICIENT MATRIX RESULTS FROM VARYING DATA PARTITIONS IN CAUSAL GRAPHIC MODELING

Name of researcher: Yesica Daniela Bravo Gonzalez

Name and degree of faculty adviser: Roy Villafane, Ph.D.

Date completed: August 2015

Problem

In life we need to compare situations in order to select the best solution. The study in this paper is about analyzing data (variables), which is also called data mining. There are situations where it is not enough to compare variables among themselves at one specific moment. Sometimes it is necessary to compare the behavior of variables at different periods of time and know how they behave at different times in order to select the best arrangements for any situation.

Method

To find correlation among variables, traffic intersections were simulated so they could be compared, since the correlation coefficient matrix is normalized. This type of

matrix was used to compare intersections in different time variances to find the most interesting information. By comparing each point from the first matrix with each point to the second matrix one can find the intersections that are busier and have a larger difference from the others. Also, two formulas were found to help find the most interesting correlations; in one of those I modified the harmonic mean formula to obtain a balance between two important details.

Results

By using these two new formulas the most interesting information between variables may be found, such as those that are the most popular or least popular (average value) and those that are very different from or very similar to each other (difference value) at different times. “Rank 1” is the value of the balance between the average and the difference, with values ranging between 0 and 0.6. A 0 means that those intersections have very low values in averages and differences, and 0.6 means the opposite. The formula “Rank 2” is based on assigning weight into the average and the difference categories. Once the formula is applied, the values would be between 0 and 1, where 0 will mean that their average or their difference is low, depending on which one was assigned more weight. A value of 1 would mean the opposite. The weight depends on what is needed for a specific situation.

Conclusions

By comparing two correlation coefficient matrices from any type of data in different time periods (since this type of matrix is already normalized) anybody can find out very interesting information for any situations where we need to know how different and popular any types of variables may be. Finally, the most interesting information may

be identified by calculating the average or the difference between variables. As an example, these formulas were used to compare traffic intersections, and the result obtained was a rank with the most popular intersections to the less important intersections, which confirmed previously observed traffic patterns.

Andrews University
College of Arts and Sciences

RANKING INTERESTING CHANGES IN CORRELATION
COEFFICIENT MATRIX RESULTS FROM
VARYING DATA PARTITIONS IN
CAUSAL GRAPHIC MODELING

A Thesis

Presented in Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by

Yesica Daniela Bravo Gonzalez

2015

© Copyright by Yesica Daniela Bravo Gonzalez 2015
All Rights Reserved

RANKING INTERESTING CHANGES IN CORRELATION
COEFFICIENT MATRIX RESULTS FROM
VARYING DATA PARTITIONS IN
CAUSAL GRAPHIC MODELING

A thesis
presented in partial fulfillment
of the requirements for the degree
Master of Science

by

Yesica Daniela Bravo Gonzalez

APPROVAL BY THE COMMITTEE:

Roy Villafane, Ph.D., Chair

Committee Member Name, Degree

Committee Member Name, Degree

Date approved

To: Families Bravo, Gonzalez, and Schooley

TABLE OF CONTENTS

LIST OF TABLES.....	iv
ACKNOWLEDGMENTS.....	vi
Chapter	
INTRODUCTION	1
Problem Statement.....	1
Purpose of Study.....	2
Justification.....	2
Methodology.....	3
Thesis Organization	3
PREVIOUS RELATED WORKS	4
METHODOLOGY	8
Purpose of Study.....	12
Methodology.....	14
RESULTS	26
SUMMARY, CONCLUSION AND FUTURE WORK.....	50
Summary.....	50
Conclusion and Future Work.....	51
REFERENCE LIST	52

LIST OF TABLES

1. Variance Components Example.....	11
2. Covariance Matrix Example	11
3. Correlation Coefficient Example	12
4. Week Data (Simulated).....	14
5. Week Correlation Coefficient Matrix	15
6. Weekend Data (Simulated).....	16
7. Weekend Correlation Coefficient Matrix	17
8. Table Of Two Intersections Data Comparison To Be Compared.....	18
9. Two Intersections Comparison Rank 1	19
10. Stoplights with a Higher Average.....	21
11. Stoplights with a Higher Difference	23
12. Intersections with Balanced Average and Distance Weights	25
13. Complete Comparison with All the Intersections Using Rank 1	26
14. Complete Results from Stoplights with a Higher Average using Rank 2.....	30
15. Complete Results from Stoplights with a Higher Difference using Rank 2	33
16. Complete Results from Intersections with Balanced Average and Distance Weights Using Rank 2	36
17. Items Sold Morning (Data Simulated).....	40
18. Items Sold Evening (Data Simulated).....	41
19. Morning Correlation Coefficient Matrix	42
20. Evening Correlation Coefficient Matrix	42

21.	Complete comparison results from the store items using Rank 1.....	43
22.	Complete comparison results from the store items using Rank 2.....	45

ACKNOWLEDGMENTS

I thank God for giving me the opportunity of taking this challenge, but overall for being with me and helping me to accomplish it. Also, I thank my family and the friends that He provided for me; they were a big support to finish my goal, and without them this would not be possible. Thank you Daniel, Rogelia, Sergio, Laura, Dani, Kari, Heather, Greg, Jim, Rowen, Noemi, Elisa, Ogbonnaya, and Lici; you have been amazing.

Also I want to thank to my professors for their support and effort to make my study successful. Dr. Villafane and Mr. Bill Wolfer: I couldn't have better professors; thank you for watching out for me in class and outside of class. Finally, I thank Andrews University for its generous support. I am thankful to you all.

CHAPTER 1

INTRODUCTION

Comparing situations and data is extremely necessary in order to select the best option and to make the necessary arrangements to obtain an optimal result. To find the best solutions it is necessary to compare variables with each other to find their correlation and to find those that are the most popular or least popular. This research is focused on comparing variables on different time periods and finding the most interesting information among them. Since comparing variables is a very extensive field that involves precise work, this study uses Data Analysis, one of the areas related to data mining, statistics, and causality graphical modeling, to find the variables' correlation. According to Rouse (2015), "Data analytics (DA) is the science of examining raw data with the purpose of drawing conclusions about that information."

Problem Statement

To analyze data, variables are usually compared with each other at one period of time to see how they behave and to find the necessary information from it, but the need may arise to compare variables from different time periods. In order to do this, it was necessary to find the best solution and the best way of comparing them. Also in real-life applications, not only is it necessary to compare one type of data, but any type of data (situation); that is why this research topic was chosen.

Out of the infinite options for data analysis, the research paper reviewed talked exclusively about analyzing traffic lights in a city in a specific time period. The authors of the article proposed to add time-varying observations so that the evolution of those causalities may be seen.

Purpose of Study

To add time-vary observations, it is necessary to decide what periods of time are to be compared. Once those matrices are calculated, the most interesting and least interesting correlations among those time periods may be found. To do this, the average and difference between those two time periods was calculated, and then a formula was used to find the perfect mean between the average and the difference among all intersections. Finally they were arranged in ascending order according to the highest values in average and difference. A second formula was incorporated to add weight to the average or to the difference, depending on which one is most needed.

Justification

Since different ways were found to select the necessary and best information, they may be applied to a range of everyday situations. Many people may need to know different things about the stoplights in a city for multiple reasons. These people include construction workers, policemen taking care of the city, big companies that transport material, and even the government. Because of these situations, some people would like to know which stoplights have a larger correlation, and others need to know which ones don't have much correlation during different periods of time. Therefore, a formula is needed that can give us the flexibility to select what we need for a specific situation.

Methodology

Even though the reviewed article provided information about how to find correlation among stoplights in a city, the purpose of this study was to find out what the most interesting correlation was among any type of variable. To find the most interesting correlation between variables, the difference and the average between the correlation measures of those variables (stoplights) needed to be obtained. With those two attributes, important information between those variables may be found, taking the time variance into consideration. Here are the important questions to consider before we analyze data:

1. How different are the correlations between intersections?
2. What is the average between those two points?
3. Considering those previous questions, which variables have the best balance?

These are the main points that will be analyzed in order to find the most interesting information.

Thesis Organization

This thesis is organized with five chapters: Introduction, Previous Related Work, Methodology, Results and Summary, Conclusion and Future Work.

Chapter 1 is the summary of each section of the thesis.

Chapter 2 talks about previous research about comparison between matrices.

Chapter 3 talks about the method used to compare variables in different time periods.

Chapter 4 finally shows the results found using these formulas.

Chapter 5 is a summary of the conclusions of this study and ideas for future work.

CHAPTER 2

PREVIOUS RELATED WORKS

The article, *Overlapping Decomposition for Causal Graphical Modeling*, suggests that there will be future work in the field to find the correlation among variables at different time periods. Further information was sought about several ways in which this has been done in the past. All articles found focused on the comparison of the covariance matrix or in the comparison of vectors, but they did not focus on identifying the correlation between each variable and its highest correlation among others.

In the past, several researchers have compared matrices. Among them are Lovsfold (1986), Shaw (1995), and Cheverud (1988, 1995, 1996). According to Cheverud & Marroig (2007), the approaches used included Random Skewers, T-statistics (T-Test), and Common Principal Components (CPC).

The random skewers test is used on genetic or phenotypic matrices. Roff et al. commented about the research of Calsbeek & Goodnight (2009), who questioned whether populations would respond similarly to selection and random skewers test,

which is given by the multivariate breeder's equation $R_A = G_A R_A^{-1} S$, $R_B = G_B P_B^{-1} S$ where S is the selection vector. The random skewers approach can be applied in this case by using the product GP^{-1} in place of either G or P . The selection skewers can be constructed to apply selection in a particular direction if such is a focus of interest (Calsbeek & Goodnight, 2009) or, as in this case, by randomly selecting skewers as previously. (p. 1120)

The Common Principal Components (CPC) uses the decomposition into eigenvalues and eigenvectors (principal components) to share more complex information

than simply being similar or not. It is not a very detailed comparison, but it can be a useful method because it allows a diversity of hypotheses of the relationships between matrices.

Cheverud & Marroig also make reference to Roff and colleagues (2001, 2003, 2004), who introduced other methods of matrix comparison, such as the T-method. The T-method is based on summation of all diagonal and off-diagonal elements in the two matrices (a sum of the number of diagonal elements plus elements above and below the diagonal). Another method that is similar to this is the T^2 Statistics, which can be calculated by the sum of the squared difference between corresponding matrix elements. Also, these statistic comparisons can be extended to three or more populations by computing all pairwise elements, as stated in Cheverud & Marroig (2007):

To test the null hypothesis that the elements in one matrix from one population do not differ from the corresponding elements in the other population matrix, we estimate the probability associated with each T value observed via randomizations. Hypothesis testing is conducted by permuting individuals between data arrays and calculating pseudo values of T or T^2 from the randomization procedure (T_r and T_r^2). Significance is evaluated as the fraction of times $T_{obs} > T_r$ (or $T_{obs}^2 > T_r^2$). (p. 462)

Another method, called Bartlett's test, is based on comparing the determinants of the matrix, which "are measured by the 'volumes' occupied" (Roff, Prokkola, Krams & Rantala, 2012, p.1117). These researchers also refer to the Modified Mantel test, which removes the differences in sizes between matrices and compares differences in shape. This uses the null hypothesis; in this case, the opposite is used, namely, that the two matrices are equal.

The Flury hierarchy method is another way of comparing covariance matrices, which uses a ranking arrangement. This method, "which is itself part of a larger hierarchical structure (Boik, 2002), extends the dimensionality approach of rank analysis

by analyzing the differences between the eigenvectors” (ibid. p.1118). Roff et al. talked about the research by Phillips & Arnold (1999), which found that this method recognizes similarities that can be arranged ascendingly:

(i) Unrelated structure, matrices share no principal components in common, (ii) Partial Common Principal Components, matrices share some of the principal components, (iii) Common Principal Components, matrices share all of the principal components but the eigenvalues differ, (iv) Proportionality, matrices have the same set of principal components but the eigenvalues of one matrix differ from another by a constant proportion, (v) Equality, matrices have identical principal components and eigenvalues. (p. 1118)

Roff et al. also mentioned two methods that are very similar: the Jackknife-MANOVA and the Jackknife-eigenvalues tests. Jackknife-MANOVA is a statistical procedure to produce a set of pseudo values that then uses MANOVA, which compares the differences between two or more vectors of means. The Jackknife-eigenvalues test basically compares the trace (the sum of the eigenvalues) of two matrices; then it is also compared with the MANOVA or with the T-test. The T-test is also known as T-statistics; it compares whether two groups have different average values.

All of these are ways to compare matrices. Some of them found a formula to compare matrices just for animals and plants, but not to compare any type of data. Since the purpose of this study was to compare any type of data, the formula is not very useful. Other formulas use the eigenvalues and eigenvectors to compare an entire matrix with another matrix. A different method was to add values from the number of diagonal elements plus elements above and below the diagonal to compare with the results from other matrices using the same procedure. Most of the observed comparisons were comparing the whole matrix with just one value, but this study was about comparing the values among themselves. Also, none of those formulas compared both parameters

(difference and average) to find out those matrices' values with the most interesting and different results.

CHAPTER 3

METHODOLOGY

This research can help to obtain the most interesting information about how variables behave by comparing any type of data at multiple times. Since this allows one to compare different types of data at different periods of time, it may be used to analyze practically anything.

Overlapping *Decomposition for Causal Graphical Modeling* talked about how the stoplights in a city were correlated; the main research was focused on finding correlations between those graphs that were already correlated with each other and putting them together in new sub-graphs. The article suggests adding time variances as a future work (Han, Song, Cong & Xie, 2012). The researchers suggest that correlations may be found not just among lights at one period of time, but also among different days or times. Doing so can prove useful for different situations, such as determining on which days or times a city's stoplights are very busy, in order to avoid them.

To analyze data at one specific moment, they used the causal graphical model and statistics as the main tools. The rationale for the use of graphical models is explained by Borgelt, Steinbrecher & Kruse (2009): "Since high-dimensional domains tend to be unmanageable as a whole (and the more so if imprecision and uncertainty are involved), it is necessary to decompose the available information. In graphical modeling [Whittaker 1990, Kruse et al. 1991, Lauritzen 1996] such a decomposition exploits (conditional)

dependence and independence relations between the attributes used to describe the domain under consideration.” (p. 10)

Overlapping Decomposition for Causal Graphical Modeling explains causal graphical models as follows: “Causal graphical models are established to meaningfully characterize causal or statistical relationships that exist among variables of interest and quantify them” (Han, Song, Cong, & Xie, 2012, p. 114). Another article comments: “The undirected graphical model... referred to as the Markov network... [has] been applied in a wide range of scientific and engineering problems to infer the local conditional dependency of the variables” (Cheng, 2013, p. 1). It further states that “A pairwise Markov network can represent an undirected graph $G = (V, E)$, where V is the node set representing the collection of random variables, and E is the edge set where the existence of an edge is equivalent to the conditional dependency between the corresponding pair of variables, with the rest of the graph” (ibid).

Han, Song, Cong, & Xie decomposed a large graphical model into multiple graphs to find their overlapping decomposition from those sub-graphs; this is called overlapping decomposition. In the reviewed article they use a Gaussian graphical model (GGM) to find their causalities. GGM is also called covariance selection or a concentration graph: “GGMs have been widely used for modeling the dependency structure among a set of variables (Whittaker, 1990). Such models use undirected graphs to specify the conditional independence structures among the variables” (Sun & Li, 2012, p.1). It is one of the most popular causal graphical modeling methods, to identify causalities between variables.

In the reviewed article, the researchers consider p random variables $X = (x_1, \dots, x_p)$, each variable x_i having n observations $x_i = (x_i^1, \dots, x_i^n)^T$, where we usually have $n \gg p$. Without loss of generality, we assume X follows a multivariate Gaussian distribution $N(\mu, \Sigma)$, also called Normal Distribution in statistics, where the mean vector μ is p -dimensional and each element in the covariance matrix Σ is the expected value $\Sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$. The causality matrix Ω is the inverse of the covariance matrix, i.e., $\Omega = \Sigma^{-1}$. There exists a causal relationship between variables x_i and x_j iff $\Omega_{ij} \neq 0$. (Han et al, p.115)

To use the covariance matrix, the variables of the matrix must be on the same scale to have better results. Covariance matrices are not strictly compatible with each other: “Columns and rows of variance/covariance matrices usually differ in scale because larger measurements tend to have larger variances and the reverse is true for smaller measurements. This scaling problem means that the randomized columns and associated rows of the variance/covariance matrix are not strictly comparable to one another” (Cheverud & Marroig, 2007, p.464).

The inverse of this matrix (Σ^{-1}) is the inverse covariance matrix, also known as the concentration matrix or the precision matrix (Hsieh, Sustik, Dhillon & Ravikumar, 2011, p.1). Each variance theoretically is divided into three parts:

Table 1

Variance Components Example:

Unique variance: 0.29	Shared variance among other variables: 0.75	Error: 0.03
-----------------------	---	-------------

Table 2

Covariance Matrix Example:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	33.54													
2	12.37	5.86												
3	14.29	5.47	7.49											
4	18.46	7.13	8.46	12.14										
5	14.96	6.23	6.76	8.64	8.79									
6	12.83	5.04	6.25	7.37	6.295	6.14								
7	27.78	10.59	12.88	16.57	13.47	11.43	26.61							
8	29.27	11.66	13.39	16.83	14.45	12.42	26.46	28.94						
9	26.94	10.32	12.16	15.66	13.23	10.84	24.15	24.83	23.92					
10	18.7	7.3	8.6	11.4	8.7	7.35	16.05	16.65	16	13.1				
11	14.07	5.01	6.89	8.83	6.25	5.50	13.21	13.22	12.50	8.7	7.84			
12	26.51	10.38	11.63	14.59	13.21	10.70	22.89	25.66	22.22	14.9	11.54	25.12		
13	14.3	5.9	6.42	7.85	7.87	6.08	12.77	13.78	12.46	8	5.7	12.53	7.88	
14	19.54	7.17	7.74	10.96	8.96	7.40	16.68	17.49	15.76	10.5	8.19	16.78	8.27	13.39

As known, the covariance matrix by itself finds the correlation among all the variables. If only the correlation between two variables is needed while leaving out the correlation among other variables, it is necessary to use the inverse of the covariance matrix; this is called the partial covariance matrix. The covariance formula is not needed in detail; Excel already has it, and it is called covariance. If the variables are not in the same scale, it is necessary to use the correlation coefficient tool to find their exact correlation. This way the data is normalized using the correlation to have a better result.

Additionally, when only the correlation between two variables needs to be found, partial correlation must be used to obtain just this type of correlation. To obtain the correlation coefficient Excel can be used; the function is called correlation.

Table 3

Correlation Coefficient Example:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2	0.88	1												
3	0.90	0.82	1											
4	0.91	0.84	0.88	1										
5	0.87	0.86	0.83	0.83	1									
6	0.89	0.83	0.92	0.85	0.85	1								
7	0.92	0.84	0.91	0.92	0.88	0.89	1							
8	0.93	0.89	0.90	0.89	0.90	0.93	0.95	1						
9	0.95	0.87	0.90	0.91	0.91	0.89	0.95	0.94	1					
10	0.89	0.83	0.86	0.90	0.81	0.81	0.85	0.85	0.90	1				
11	0.86	0.73	0.89	0.90	0.75	0.79	0.91	0.87	0.91	0.85	1			
12	0.91	0.85	0.84	0.83	0.88	0.86	0.88	0.95	0.90	0.82	0.82	1		
13	0.87	0.86	0.83	0.80	0.94	0.87	0.88	0.91	0.90	0.78	0.73	0.89	1	
14	0.92	0.80	0.77	0.85	0.82	0.81	0.88	0.88	0.88	0.79	0.79	0.91	0.80	1

Purpose of Study

Sometimes the need arises to compare all kinds of data in everyday situations, and those are not always in the same scale. Since correlation coefficients normalize data in the same scale, anything can be compared, in any type of scale.

In *Introductory Statistics: A Conceptual Approach Using R*, an example is presented to illustrate the difficulty of comparing measurements made in inches and ounces versus those made in centimeters and grams. The following is suggested: “If

metric is the problem, however, the situation is easy to remedy. We can solve the problem by putting all variables on the same scale” (Ware, W., Ferron, J. & Miller, B., 2013, p.57).

This article was making reference to a “standardized” covariance, which was in fact the correlation coefficient. In this way not just stoplights in a city can be compared, but also, for example, ages with salaries (things that are in a different scale) or any other data for which it is necessary to find out whether a correlation exists.

To compare time variances from stoplights in a city, either two correlation coefficient matrices or two partial-correlation coefficient matrices must be compared, since correlation coefficient finds the correlation among all the variables and the partial correlation coefficient finds the correlation between only two variables. Which type of calculation is needed is determined according to the specific situation. In order to standardize the scales and account for all the stoplights in a city, this study compares correlation matrices.

In my research, data was recreated to resemble the stoplights in a city, how the lights of a city behave, and especially how to distinguish those that behave very differently from each other. This could be useful in different situations; for example, it can be used by people that want to coordinate city traffic to keep it more fluent, or to designate the location of policemen in key points around the city. The purpose of analyzing data is to find the most different or interesting behavior between those two variables at different periods of time.

Methodology

Working with time variance requires the comparison of two matrices; to be more precise, each point of a matrix with another matrix's point has to be compared. The table below contains simulated data from Berrien Springs' traffic intersections during the week. It provides some points to compare:

Table 4

Week Data (Simulated)

Number of cars passing through intersection from 8:45 a.m. to 9:00 a.m. Data simulated over 20 days.														
Day	1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir	3) Old U.S. 31 & College Ave	4) E Campus Cir & Timberland Dr	5) International Ct & Timberland Dr	6) Timberland Dr & Kephart Ln	7) Kephart Ln & N Main St	8) Kephart Ln & Old U.S. 31	9) W Ferry St & Old U.S. 31	10) E Snow Rd & Old U.S. 31	11) George Ave & Old U.S. 31	12) Garland Ave & Old U.S. 31	13) International Ct & N Main St	14) Old U.S. 31 & Griggs Ave
1	39	39	25	15	14	20	29	24	45	46	23	12	9	10
2	40	40	28	14	12	19	30	20	36	38	24	16	11	5
3	42	42	26	11	13	22	34	19	40	43	23	12	15	13
4	45	45	30	12	11	23	30	24	46	44	20	11	15	8
5	36	36	27	13	15	24	29	23	38	40	21	15	11	16
6	40	40	25	16	10	23	28	21	43	39	24	15	16	11
7	46	46	21	18	13	20	25	25	44	37	20	11	13	7
8	38	38	24	19	14	21	26	19	39	40	19	16	14	6
9	43	43	22	20	12	19	24	17	37	42	24	13	12	9
10	44	44	25	21	16	20	26	26	42	45	23	14	16	5
11	40	40	30	26	12	22	25	20	41	36	20	13	12	12
12	39	39	29	22	11	19	29	23	40	40	19	14	11	11
13	37	37	21	19	15	20	24	19	37	44	22	12	15	9
14	42	42	23	23	16	18	22	16	44	40	23	16	10	10
15	41	41	22	24	15	14	25	21	40	39	24	12	13	7
16	36	36	24	25	11	13	26	25	46	37	20	14	14	14
17	40	40	25	21	16	15	28	20	38	42	19	12	12	13
18	37	37	28	22	13	18	30	16	43	39	22	13	16	9
19	44	44	29	20	14	21	26	22	44	40	23	11	11	6
20	42	42	30	24	15	20	28	19	28	42	24	15	14	7

Table 5

Week Correlation Coefficient Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2	1	1												
3	0.01	0.01	1											
4	-0.11	-0.11	-0.06	1										
5	0.03	0.03	-0.27	0.18	1									
6	0.19	0.19	0.37	-0.59	-0.17	1								
7	-0.11	-0.11	0.53	-0.62	-0.29	0.29	1							
8	0.17	0.17	0.05	-0.17	-0.15	0.05	0.10	1						
9	0.11	0.11	-0.12	-0.10	-0.29	-0.04	-0.06	0.40	1					
10	0.17	0.17	-0.05	-0.38	0.37	0.20	0.21	0.09	-0.09	1				
11	0.22	0.22	-0.10	-0.10	0.13	0.04	-0.01	-0.28	-0.24	0.21	1			
12	-0.39	-0.39	0.08	0.11	0.01	0.04	-0.08	-0.27	-0.34	-0.24	0.12	1		
13	0.03	0.03	-0.04	-0.02	-0.17	0.09	0.18	-0.00	-0.00	0.08	0.02	-0.07	1	
14	-0.51	-0.51	0.03	-0.07	-0.14	-0.00	0.22	0.00	0.12	-0.12	-0.33	0.01	-0.14	1

Table 6

Weekend Data (Simulated)

Number of cars passing through intersection from 9:00 a.m. to 9:15 a.m. Simulated over 20 days during weekend.														
Day	1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	3) Old U.S. 31 & College Ave	4) E Campus Cir Dr & Timberland Dr	5) International Ct & Timberland Dr	6) Timberland Dr & Kephart Ln	7) Kephart Ln & N Main St	8) Kephart Ln & Old U.S. 31	9) W Ferry St & Old U.S. 31	10) E Snow Rd & Old U.S. 31	11) George Ave & Main St	12) Garland Ave & Old U.S. 31	13) International Ct & N Main St	14) Old U.S. 31 & Griggs Ave
1	20	10	10	8	11	12	19	21	25	24	15	26	10	11
2	10	5	3	3	5	6	9	10	15	16	10	15	5	4
3	19	9	9	9	10	11	20	19	24	25	16	24	9	10
4	9	6	2	2	6	5	10	9	14	15	9	14	4	7
5	21	11	7	10	9	10	18	20	23	23	14	25	8	12
6	8	4	4	3	4	4	8	8	16	18	13	17	2	5
7	18	7	8	11	8	9	21	18	22	22	17	23	7	13
8	6	5	5	5	5	7	11	11	14	19	11	13	3	2
9	19	10	10	9	10	12	20	21	24	21	15	26	10	11
10	5	6	3	4	6	6	10	9	15	14	10	15	5	4
11	20	9	9	12	12	10	19	19	25	24	16	24	9	10
12	8	4	2	2	4	5	9	10	13	16	9	19	4	6
13	17	11	7	10	10	9	18	18	23	25	14	25	8	12
14	7	5	4	3	3	7	8	11	12	15	11	17	2	5
15	20	8	8	11	8	11	21	20	26	23	17	23	7	13
16	9	6	5	4	2	5	11	8	15	18	12	12	3	2
17	19	9	9	9	11	10	20	21	24	22	15	26	11	11
18	6	4	2	2	5	6	10	9	16	17	10	15	5	4
19	17	10	7	10	9	9	17	20	23	24	16	24	9	10
20	10	5	4	5	4	7	7	7	14	19	9	14	4	6

Table 7

Weekend Correlation Coefficient Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2	0.88	1												
3	0.90	0.82	1											
4	0.91	0.84	0.88	1										
5	0.87	0.86	0.83	0.83	1									
6	0.89	0.83	0.92	0.85	0.85	1								
7	0.92	0.84	0.91	0.92	0.88	0.89	1							
8	0.93	0.89	0.90	0.89	0.90	0.93	0.95	1						
9	0.95	0.87	0.90	0.91	0.91	0.89	0.95	0.94	1					
10	0.89	0.83	0.86	0.90	0.81	0.81	0.85	0.85	0.90	1				
11	0.86	0.73	0.89	0.90	0.75	0.79	0.91	0.87	0.91	0.85	1			
12	0.91	0.85	0.84	0.83	0.88	0.86	0.88	0.95	0.90	0.82	0.82	1		
13	0.87	0.86	0.83	0.80	0.94	0.87	0.88	0.91	0.90	0.78	0.73	0.89	1	
14	0.92	0.80	0.77	0.85	0.82	0.81	0.88	0.88	0.88	0.79	0.79	0.91	0.80	1

To work with two correlation coefficients in this case, each point from each matrix was compared with each point from the other matrix; for example, the first point in the first vector (intersection) in the week matrix was compared with each point of the first vector (intersection) from the weekend matrix. This way the goal of identifying the most interesting relations between variables is reached. Here is an example to demonstrate how these formulas work:

Table 8

Table of Two Intersections – Data to Be Compared

A Intersection	B Intersection	Week Correlation Coefficient	Weekend Correlation Coefficient
1) Old U.S. 31 & Andrews Blvd	9) W Ferry St & Old U.S. 31	0.11	0.95
1) Old U.S. 31 & Andrews Blvd	8) Kephart Ln & Old U.S. 31	0.17	0.93
1) Old U.S. 31 & Andrews Blvd	7) Kephart Ln & N Main St	-0.11	0.92
1) Old U.S. 31 & Andrews Blvd	12) Garland Ave & Old U.S. 31	-0.39	0.91
1) Old U.S. 31 & Andrews Blvd	14) Old U.S. 31 & Griggs Ave	-0.51	0.92
1) Old U.S. 31 & Andrews Blvd	4) E Campus Cir Dr & Timberland Dr	-0.11	0.91
1) Old U.S. 31 & Andrews Blvd	3) Old U.S. 31 & College Ave	0.01	0.90
1) Old U.S. 31 & Andrews Blvd	6) Timberland Dr & Kephart Ln	0.19	0.89
1) Old U.S. 31 & Andrews Blvd	10) E Snow Rd & Old U.S. 31	0.17	0.89
1) Old U.S. 31 & Andrews Blvd	13) International Ct & N Main St	0.03	0.87
1) Old U.S. 31 & Andrews Blvd	11) George Ave & Main St	0.22	0.86
1) Old U.S. 31 & Andrews Blvd	5) International Ct & Timberland Dr	0.03	0.87
1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	1	0.88

First of all, to start a comparison of matrices, they must contain positive numbers.

To do so, calculate the square of every number so there are no negative numbers. After this the following formulas may be used to define the most interesting information. These are the formulas to obtain the difference and the average between these two points:

$$\text{Difference } d = |V1 - V2|.$$

$$\text{Average } a = 0.5 (V1+V2).$$

In the area of information retrieval, the harmonic mean is used to compare two values and obtain the best combination of those values. Since this harmonic mean finds the best value between two numbers, with a little change and using part of this formula, the mean from those values can be found. This is the modified harmonic mean formula used to compare two correlation coefficients:

$$\text{Rank 1} = 2 \frac{da}{d+a}$$

The intersections that are the most and the least interesting can be found by arranging the values of the Rank 1 formula in ascending order, because they have the largest differences and the highest averages between each other, which was identified as the most interesting information. Here is an example of the correlation using this formula:

Table 9

Two Intersections Comparison Rank 1.

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r ² – Week Correlation Coefficient	r ² – Weekend Correlation Coefficient	Difference	Average	Rank 1 Formula
1) Old U.S. 31 & Andrews Blvd	9) W Ferry St & Old U.S. 31	0.11	0.95	0.01	0.90	0.89	0.45	0.60
1) Old U.S. 31 & Andrews Blvd	8) Kephart Ln & Old U.S. 31	0.17	0.93	0.03	0.88	0.85	0.45	0.59
1) Old U.S. 31 & Andrews Blvd	7) Kephart Ln & N Main St	-0.11	0.92	0.01	0.86	0.85	0.43	0.57
1) Old U.S. 31 & Andrews Blvd	12) Garland Ave & Old U.S. 31	-0.39	0.91	0.15	0.83	0.68	0.49	0.57
1) Old U.S. 31 & Andrews Blvd	14) Old U.S. 31 & Griggs Ave	-0.51	0.92	0.26	0.85	0.58	0.55	0.57
1) Old U.S. 31 & Andrews Blvd	4) E Campus Cir Dr & Timberland Dr	-0.11	0.91	0.01	0.83	0.82	0.42	0.56
1) Old U.S. 31 & Andrews Blvd	3) Old U.S. 31 & College Ave	0.01	0.90	0.00	0.81	0.81	0.40	0.54
1) Old U.S. 31 & Andrews Blvd	6) Timberland Dr & Kephart Ln	0.19	0.89	0.03	0.79	0.75	0.41	0.53
1) Old U.S. 31 & Andrews Blvd	10) E Snow Rd & Old U.S. 31	0.17	0.89	0.02	0.79	0.76	0.41	0.53
1) Old U.S. 31 & Andrews Blvd	13) International Ct & N Main St	0.03	0.87	0.00	0.77	0.77	0.38	0.51
1) Old U.S. 31 & Andrews Blvd	11) George Ave & Main St	0.22	0.86	0.04	0.75	0.70	0.40	0.51
1) Old U.S. 31 & Andrews Blvd	5) International Ct & Timberland Dr	0.03	0.87	0.00	0.75	0.75	0.38	0.50
1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	1	0.88	1	0.77	0.22	0.88	0.35

This is part of the simulated data where the intersections are arranged based on their highest difference and highest average values. If one looks at the raw data (week and weekend) from the correlation coefficients, one can see that the values are arranged from those that are farther from each other to the ones that are closer. This way the best balance can be found between those two characteristics and they can be arranged in the order needed at a specific point in time.

While working with this formula I realized that sometimes it may be desirable to know how far apart geometrically they are, while other times it's more important to know how highly correlated they are based on their mean, but not both. In this case I was looking for other formulas and found out that if weight is assigned to the data elements, one can figure out either how different they are or what are the highest averages among them. Trying to find a formula for it I realized that this can be done:

$$\text{Distance weight} = w_d [0, 1]$$

$$\text{Average weight} = w_a [0, 1]$$

To find out either their distance or average, more weight was put into the element that is needed according to the specific situation, using the following formula:

$$\text{Rank2} = dw_d + aw_a.$$

To see how this formula works, here are some examples from Excel. First, the weight values are set, making sure that their sum equals 1. For example:

$$a \text{ in range } [0, 1]$$

$$w_a = a \text{ (weight of average)}$$

$$w_d = 1 - a \text{ (weight of difference)}$$

$$w_a + w_d = 1$$

Sample equation with assigned weights:

$$a = 0.87452 \rightarrow w_a = 0.87452$$

$$w_d = 0.12548$$

Arranging these values indicates that there's a need to know which ones have a larger average than the others. Here are some examples.

Table 10

Stoplights with a Higher Average

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r ² Week Correlation Coefficient	r ² Weekend Correlation Coefficient	Difference	Average	w _d	w _a	d w _d	a w _a	d w _d + a w _a
1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	1	0.88	1	0.77	0.22	0.88	0.12	0.87	0.02	0.77	0.80
1) Old U.S. 31 & Andrews Blvd	14) Old U.S. 31 & Griggs Ave	-0.51	0.92	0.26	0.85	0.58	0.55	0.12	0.87	0.07	0.48	0.56
1) Old U.S. 31 & Andrews Blvd	9) W Ferry St & Old U.S. 31	0.11	0.95	0.01	0.90	0.89	0.45	0.12	0.87	0.11	0.40	0.51
1) Old U.S. 31 & Andrews Blvd	8) Kephart Ln & Old U.S. 31	0.17	0.93	0.03	0.88	0.85	0.45	0.12	0.87	0.10	0.39	0.50
1) Old U.S. 31 & Andrews Blvd	12) Garland Ave & Old U.S. 31	-0.39	0.91	0.15	0.83	0.68	0.49	0.12	0.87	0.08	0.43	0.51
1) Old U.S. 31 & Andrews Blvd	7) Kephart Ln & N Main St	-0.11	0.92	0.01	0.86	0.85	0.43	0.12	0.87	0.10	0.38	0.49
1) Old U.S. 31 & Andrews Blvd	4) E Campus Cir Dr & Timberland Dr	-0.11	0.91	0.01	0.83	0.82	0.42	0.12	0.87	0.10	0.37	0.47
1) Old U.S. 31 & Andrews Blvd	3) Old U.S. 31 & College Ave	0.01	0.90	0.00	0.81	0.81	0.40	0.12	0.87	0.10	0.35	0.45
1) Old U.S. 31 & Andrews Blvd	6) Timberland Dr & Kephart Ln	0.19	0.89	0.03	0.79	0.75	0.41	0.12	0.87	0.09	0.36	0.46
1) Old U.S. 31 & Andrews Blvd	10) E Snow Rd & Old U.S. 31	0.17	0.89	0.02	0.79	0.76	0.41	0.12	0.87	0.09	0.36	0.45
1) Old U.S. 31 & Andrews Blvd	13) International Ct & N Main St	0.03	0.87	0.00	0.77	0.77	0.38	0.12	0.87	0.09	0.33	0.43
1) Old U.S. 31 & Andrews Blvd	11) George Ave & Main St	0.22	0.86	0.04	0.75	0.70	0.40	0.12	0.87	0.08	0.35	0.43
1) Old U.S. 31 & Andrews Blvd	5) International Ct & Timberland Dr	0.03	0.87	0.00	0.75	0.75	0.38	0.12	0.87	0.09	0.33	0.42

This previous table shows how their average is arranged in ascending order and how those intersections are arranged based on their average. The following examples are similar, changing their weight to obtain the higher difference between stoplights.

$$a = 0.17 \rightarrow w_a = 0.17$$

$$w_d = 0.82$$

Table 11

Stoplights with a Higher Difference

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r ² Week Correlation Coefficient	r ² Weekend Correlation Coefficient	Difference	Average	w _d	w _a	d w _d	a w _a	d w _d + a w _a
1) Old U.S. 31 & Andrews Blvd	9) W Ferry St & Old U.S. 31	0.11	0.95	0.01	0.90	0.89	0.45	0.82	0.17	0.73	0.08	0.81
1) Old U.S. 31 & Andrews Blvd	8) Kephart Ln & Old U.S. 31	0.17	0.93	0.03	0.88	0.85	0.45	0.82	0.17	0.70	0.07	0.78
1) Old U.S. 31 & Andrews Blvd	7) Kephart Ln & N Main St	-0.11	0.92	0.01	0.86	0.85	0.43	0.82	0.17	0.70	0.07	0.77
1) Old U.S. 31 & Andrews Blvd	4) E Campus Cir Dr & Timberland Dr	-0.11	0.91	0.01	0.83	0.82	0.42	0.82	0.17	0.67	0.07	0.75
1) Old U.S. 31 & Andrews Blvd	3) Old U.S. 31 & College Ave	0.01	0.90	0.00	0.81	0.81	0.40	0.82	0.17	0.67	0.07	0.74
1) Old U.S. 31 & Andrews Blvd	13) International Ct & N Main St	0.03	0.87	0.00	0.77	0.77	0.38	0.82	0.17	0.63	0.06	0.70
1) Old U.S. 31 & Andrews Blvd	10) E Snow Rd & Old U.S. 31	0.17	0.89	0.02	0.79	0.76	0.41	0.82	0.17	0.63	0.07	0.70
1) Old U.S. 31 & Andrews Blvd	6) Timberland Dr & Kephart Ln	0.19	0.89	0.03	0.79	0.75	0.41	0.82	0.17	0.62	0.07	0.70
1) Old U.S. 31 & Andrews Blvd	5) International Ct & Timberland Dr	0.03	0.87	0.00	0.75	0.75	0.38	0.82	0.17	0.62	0.06	0.69
1) Old U.S. 31 & Andrews Blvd	11) George Ave & Main St	0.22	0.86	0.04	0.75	0.70	0.40	0.82	0.17	0.58	0.06	0.64
1) Old U.S. 31 & Andrews Blvd	12) Garland Ave & Old U.S. 31	-0.39	0.91	0.15	0.83	0.68	0.49	0.82	0.17	0.56	0.08	0.64
1) Old U.S. 31 & Andrews Blvd	14) Old U.S. 31 & Griggs Ave	-0.51	0.92	0.26	0.85	0.58	0.55	0.82	0.17	0.48	0.09	0.57
1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	1	0.88	1	0.77	0.22	0.88	0.82	0.17	0.18	0.15	0.33

In this other example their difference is arranged in ascending order and the values of those intersections from both the week and the weekend are arranged based on their difference. The following example is similar, but the weight placed on the difference value is only a little bit higher than the average.

$$a = 0.47452 \rightarrow w_a = 0.47$$

$$w_d = 0.52$$

Table 12

Intersections with Balanced Average and Distance Weights

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r ² Week Correlation Coefficient	r ² Weekend Correlation Coefficient	Difference	Average	w _d	w _a	dw _d	aw _a	dw _d + aw _a
1) Old U.S. 31 & Andrews Blvd	9) W Ferry St & Old U.S. 31	0.11	0.95	0.01	0.90	0.89	0.45	0.52	0.47	0.46	0.21	0.68
1) Old U.S. 31 & Andrews Blvd	8) Kephart Ln & Old U.S. 31	0.17	0.93	0.03	0.88	0.85	0.45	0.52	0.47	0.44	0.21	0.66
1) Old U.S. 31 & Andrews Blvd	7) Kephart Ln & N Main St	-0.11	0.92	0.01	0.86	0.85	0.43	0.52	0.47	0.44	0.20	0.65
1) Old U.S. 31 & Andrews Blvd	4) E Campus Cir Dr & Timberland Dr	-0.11	0.91	0.01	0.83	0.82	0.42	0.52	0.47	0.43	0.20	0.63
1) Old U.S. 31 & Andrews Blvd	3) Old U.S. 31 & college Ave	0.01	0.90	0.00	0.81	0.81	0.40	0.52	0.47	0.42	0.19	0.61
1) Old U.S. 31 & Andrews Blvd	10) E Snow Rd & Old U.S. 31	0.17	0.89	0.02	0.79	0.76	0.41	0.52	0.47	0.40	0.19	0.59
1) Old U.S. 31 & Andrews Blvd	6) Timberland Dr & Kephart Ln	0.19	0.89	0.03	0.79	0.75	0.41	0.52	0.47	0.39	0.19	0.59
1) Old U.S. 31 & Andrews Blvd	12) Garland Ave & Old U.S. 31	-0.39	0.91	0.15	0.83	0.68	0.49	0.52	0.47	0.35	0.23	0.59
1) Old U.S. 31 & Andrews Blvd	13) International Ct & N Main St	0.03	0.87	0.00	0.77	0.77	0.38	0.52	0.47	0.40	0.18	0.58
1) Old U.S. 31 & Andrews Blvd	5) International Ct & Timberland Dr	0.03	0.87	0.00	0.75	0.75	0.38	0.52	0.47	0.39	0.18	0.57
1) Old U.S. 31 & Andrews Blvd	14) Old U.S. 31 & Griggs Ave	-0.51	0.92	0.26	0.85	0.58	0.55	0.52	0.47	0.30	0.26	0.57
1) Old U.S. 31 & Andrews Blvd	11) George Ave & Main St	0.22	0.86	0.04	0.75	0.70	0.40	0.52	0.47	0.36	0.19	0.55
1) Old U.S. 31 & Andrews Blvd	2) Andrews Blvd & E Campus Cir Dr	1	0.88	1	0.77	0.22	0.88	0.52	0.47	0.11	0.42	0.53

CHAPTER 4

RESULTS

In the previous chapter one can see how the formulas were applied and a short result from those examples, but here are the complete results from those formulas. The previous comparison was only between the first intersection and the rest of the other intersections; what follows is the comparison of all of the other intersections from the formula:

$$\text{Rank 1} = 2 \frac{da}{d + a}$$

Table 13

Complete Comparison with All the Intersections Using Rank 1

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	Rank 1 Formula
8	12	-0.27	0.95	0.07	0.91	0.83	0.49	0.62
7	9	-0.07	0.96	0.00	0.92	0.91	0.46	0.61
8	9	0.40	0.94	0.16	0.89	0.73	0.53	0.61
7	8	0.11	0.95	0.01	0.91	0.90	0.46	0.61
1	9	0.12	0.95	0.01	0.90	0.89	0.46	0.61
5	13	-0.17	0.95	0.03	0.89	0.87	0.46	0.60
1	8	0.18	0.94	0.03	0.88	0.85	0.46	0.59
3	6	0.38	0.92	0.14	0.85	0.71	0.50	0.58
6	8	0.06	0.93	0.00	0.87	0.86	0.44	0.58

Table 13—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	Rank 1 Formula
1	7	-0.12	0.93	0.01	0.86	0.85	0.44	0.58
1	12	-0.39	0.91	0.15	0.83	0.68	0.49	0.57
1	14	-0.52	0.92	0.27	0.85	0.58	0.56	0.57
5	9	-0.30	0.91	0.09	0.83	0.74	0.46	0.57
9	11	-0.25	0.91	0.06	0.83	0.77	0.45	0.57
4	9	-0.10	0.92	0.01	0.84	0.83	0.43	0.57
9	12	-0.35	0.91	0.12	0.82	0.70	0.47	0.56
1	4	-0.12	0.91	0.01	0.84	0.82	0.43	0.56
4	10	-0.39	0.90	0.15	0.82	0.67	0.48	0.56
12	14	0.01	0.92	0.00	0.84	0.84	0.42	0.56
7	11	-0.01	0.91	0.00	0.84	0.84	0.42	0.56
8	13	-0.01	0.91	0.00	0.83	0.83	0.42	0.56
3	9	-0.13	0.91	0.02	0.83	0.81	0.42	0.55
5	8	-0.16	0.91	0.02	0.82	0.80	0.42	0.55
3	8	0.05	0.91	0.00	0.83	0.83	0.42	0.55
3	7	0.54	0.91	0.29	0.83	0.54	0.56	0.55
9	13	-0.01	0.91	0.00	0.82	0.82	0.41	0.55
4	11	-0.11	0.90	0.01	0.82	0.81	0.41	0.55
6	7	0.30	0.89	0.09	0.80	0.71	0.44	0.55
9	10	-0.09	0.90	0.01	0.82	0.81	0.41	0.55
4	8	-0.18	0.90	0.03	0.81	0.77	0.42	0.54
1	3	0.01	0.90	0.00	0.81	0.81	0.41	0.54
3	11	-0.11	0.90	0.01	0.81	0.80	0.41	0.54
2	8	0.18	0.90	0.03	0.80	0.77	0.42	0.54
1	6	0.20	0.89	0.04	0.80	0.76	0.42	0.54
1	10	0.17	0.89	0.03	0.80	0.77	0.41	0.54
6	9	-0.04	0.89	0.00	0.80	0.80	0.40	0.53
5	7	-0.30	0.88	0.09	0.78	0.69	0.43	0.53
12	13	-0.07	0.89	0.01	0.79	0.79	0.40	0.53
7	14	0.23	0.88	0.05	0.78	0.73	0.42	0.53
4	7	-0.62	0.92	0.39	0.85	0.46	0.62	0.53
5	12	0.02	0.89	0.00	0.79	0.79	0.40	0.53

Table 13—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	Rank 1 Formula
8	14	0.01	0.89	0.00	0.79	0.79	0.39	0.53
8	11	-0.29	0.88	0.08	0.77	0.69	0.43	0.53
3	4	-0.06	0.89	0.00	0.79	0.78	0.40	0.53
7	13	0.19	0.88	0.03	0.78	0.74	0.41	0.53
7	12	-0.08	0.89	0.01	0.78	0.78	0.40	0.52
9	14	0.12	0.88	0.01	0.78	0.76	0.40	0.52
1	13	0.03	0.88	0.00	0.77	0.77	0.39	0.52
6	13	0.09	0.87	0.01	0.76	0.76	0.39	0.51
1	11	0.22	0.87	0.05	0.75	0.70	0.40	0.51
2	9	0.12	0.87	0.01	0.76	0.75	0.39	0.51
1	5	0.03	0.87	0.00	0.76	0.76	0.38	0.51
3	10	-0.06	0.87	0.00	0.75	0.75	0.38	0.50
2	5	0.03	0.87	0.00	0.75	0.75	0.38	0.50
2	13	0.03	0.87	0.00	0.75	0.75	0.38	0.50
7	10	0.22	0.86	0.05	0.74	0.69	0.39	0.50
2	12	-0.39	0.86	0.15	0.73	0.58	0.44	0.50
10	11	0.22	0.86	0.05	0.74	0.69	0.39	0.50
6	12	0.04	0.86	0.00	0.74	0.74	0.37	0.50
5	6	-0.18	0.86	0.03	0.73	0.70	0.38	0.50
4	14	-0.07	0.86	0.01	0.74	0.73	0.37	0.49
8	10	0.09	0.86	0.01	0.73	0.72	0.37	0.49
2	7	-0.12	0.85	0.01	0.72	0.71	0.37	0.48
3	12	0.08	0.85	0.01	0.72	0.71	0.36	0.48
2	4	-0.12	0.85	0.01	0.71	0.70	0.36	0.48
2	6	0.20	0.84	0.04	0.71	0.67	0.37	0.48
3	5	-0.27	0.83	0.07	0.69	0.62	0.38	0.47
4	5	0.18	0.84	0.03	0.70	0.67	0.37	0.47
2	10	0.17	0.83	0.03	0.69	0.66	0.36	0.47
4	12	0.12	0.84	0.01	0.70	0.68	0.36	0.47
3	13	-0.05	0.84	0.00	0.70	0.70	0.35	0.47
10	12	-0.25	0.82	0.06	0.67	0.61	0.37	0.46
5	14	-0.15	0.83	0.02	0.68	0.66	0.35	0.46

Table 13—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	Rank 1 Formula
6	10	0.21	0.82	0.04	0.67	0.63	0.36	0.46
2	3	0.01	0.83	0.00	0.68	0.68	0.34	0.45
11	12	0.13	0.82	0.02	0.68	0.66	0.35	0.45
5	10	0.37	0.81	0.14	0.66	0.52	0.40	0.45
6	14	0.00	0.82	0.00	0.67	0.67	0.33	0.44
4	6	-0.60	0.85	0.36	0.73	0.37	0.54	0.44
11	14	-0.34	0.80	0.11	0.64	0.53	0.38	0.44
13	14	-0.14	0.81	0.02	0.65	0.63	0.33	0.44
4	13	-0.03	0.80	0.00	0.64	0.64	0.32	0.43
10	14	-0.12	0.79	0.02	0.63	0.61	0.32	0.42
2	14	-0.52	0.81	0.27	0.66	0.39	0.46	0.42
6	11	0.04	0.79	0.00	0.63	0.63	0.31	0.42
10	13	0.09	0.79	0.01	0.62	0.61	0.31	0.41
3	14	0.04	0.77	0.00	0.60	0.60	0.30	0.40
5	11	0.14	0.75	0.02	0.57	0.55	0.29	0.38
2	11	0.22	0.74	0.05	0.55	0.50	0.30	0.37
11	13	0.03	0.73	0.00	0.54	0.54	0.27	0.36
1	2	1.00	0.88	1.00	0.78	0.22	0.89	0.35

The following table contains the complete results from the comparisons among all the lights in the city using formula Rank 2.

$$Rank2 = dw_d + aw_a$$

$$a = 0.87452 \Rightarrow w_a = 0.87452 (0.87)$$

$$w_d = 0.12548 (0.13)$$

Table 14

Complete Results from Stoplights with a Higher Average Using Rank 2

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
1	2	1.00	0.88	1.00	0.78	0.22	0.89	0.13	0.87	0.03	0.78	0.81
4	7	-0.62	0.92	0.39	0.85	0.46	0.62	0.13	0.87	0.06	0.54	0.60
1	14	-0.52	0.92	0.27	0.85	0.58	0.56	0.13	0.87	0.07	0.49	0.56
3	7	0.54	0.91	0.29	0.83	0.54	0.56	0.13	0.87	0.07	0.49	0.56
8	9	0.40	0.94	0.16	0.89	0.73	0.53	0.13	0.87	0.09	0.46	0.55
8	12	-0.27	0.95	0.07	0.91	0.83	0.49	0.13	0.87	0.10	0.43	0.53
3	6	0.38	0.92	0.14	0.85	0.71	0.50	0.13	0.87	0.09	0.43	0.52
4	6	-0.60	0.85	0.36	0.73	0.37	0.54	0.13	0.87	0.05	0.48	0.52
1	12	-0.39	0.91	0.15	0.83	0.68	0.49	0.13	0.87	0.09	0.43	0.52
7	9	-0.07	0.96	0.00	0.92	0.91	0.46	0.13	0.87	0.11	0.40	0.52
7	8	0.11	0.95	0.01	0.91	0.90	0.46	0.13	0.87	0.11	0.40	0.52
1	9	0.12	0.95	0.01	0.90	0.89	0.46	0.13	0.87	0.11	0.40	0.51
5	13	-0.17	0.95	0.03	0.89	0.87	0.46	0.13	0.87	0.11	0.40	0.51
4	10	-0.39	0.90	0.15	0.82	0.67	0.48	0.13	0.87	0.08	0.42	0.51
1	8	0.18	0.94	0.03	0.88	0.85	0.46	0.13	0.87	0.11	0.40	0.51
9	12	-0.35	0.91	0.12	0.82	0.70	0.47	0.13	0.87	0.09	0.41	0.50
5	9	-0.30	0.91	0.09	0.83	0.74	0.46	0.13	0.87	0.09	0.40	0.50
1	7	-0.12	0.93	0.01	0.86	0.85	0.44	0.13	0.87	0.11	0.38	0.49
6	8	0.06	0.93	0.00	0.87	0.86	0.44	0.13	0.87	0.11	0.38	0.49
9	11	-0.25	0.91	0.06	0.83	0.77	0.45	0.13	0.87	0.10	0.39	0.49
4	9	-0.10	0.92	0.01	0.84	0.83	0.43	0.13	0.87	0.10	0.37	0.48
6	7	0.30	0.89	0.09	0.80	0.71	0.44	0.13	0.87	0.09	0.39	0.48
1	4	-0.12	0.91	0.01	0.84	0.82	0.43	0.13	0.87	0.10	0.37	0.48
12	14	0.01	0.92	0.00	0.84	0.84	0.42	0.13	0.87	0.11	0.37	0.47
7	11	-0.01	0.91	0.00	0.84	0.84	0.42	0.13	0.87	0.10	0.37	0.47
3	9	-0.13	0.91	0.02	0.83	0.81	0.42	0.13	0.87	0.10	0.37	0.47
5	8	-0.16	0.91	0.02	0.82	0.80	0.42	0.13	0.87	0.10	0.37	0.47
8	13	-0.01	0.91	0.00	0.83	0.83	0.42	0.13	0.87	0.10	0.36	0.47
3	8	0.05	0.91	0.00	0.83	0.83	0.42	0.13	0.87	0.10	0.36	0.47

Table 14—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
4	11	-0.11	0.90	0.01	0.82	0.81	0.41	0.13	0.87	0.10	0.36	0.46
5	7	-0.30	0.88	0.09	0.78	0.69	0.43	0.13	0.87	0.09	0.38	0.46
4	8	-0.18	0.90	0.03	0.81	0.77	0.42	0.13	0.87	0.10	0.37	0.46
9	13	-0.01	0.91	0.00	0.82	0.82	0.41	0.13	0.87	0.10	0.36	0.46
9	10	-0.09	0.90	0.01	0.82	0.81	0.41	0.13	0.87	0.10	0.36	0.46
1	6	0.20	0.89	0.04	0.80	0.76	0.42	0.13	0.87	0.10	0.37	0.46
2	8	0.18	0.90	0.03	0.80	0.77	0.42	0.13	0.87	0.10	0.36	0.46
2	12	-0.39	0.86	0.15	0.73	0.58	0.44	0.13	0.87	0.07	0.39	0.46
8	11	-0.29	0.88	0.08	0.77	0.69	0.43	0.13	0.87	0.09	0.37	0.46
3	11	-0.11	0.90	0.01	0.81	0.80	0.41	0.13	0.87	0.10	0.36	0.46
1	3	0.01	0.90	0.00	0.81	0.81	0.41	0.13	0.87	0.10	0.36	0.46
1	10	0.17	0.89	0.03	0.80	0.77	0.41	0.13	0.87	0.10	0.36	0.46
7	14	0.23	0.88	0.05	0.78	0.73	0.42	0.13	0.87	0.09	0.36	0.46
2	14	-0.52	0.81	0.27	0.66	0.39	0.46	0.13	0.87	0.05	0.40	0.45
6	9	-0.04	0.89	0.00	0.80	0.80	0.40	0.13	0.87	0.10	0.35	0.45
7	13	0.19	0.88	0.03	0.78	0.74	0.41	0.13	0.87	0.09	0.36	0.45
12	13	-0.07	0.89	0.01	0.79	0.79	0.40	0.13	0.87	0.10	0.35	0.45
5	12	0.02	0.89	0.00	0.79	0.79	0.40	0.13	0.87	0.10	0.35	0.45
8	14	0.01	0.89	0.00	0.79	0.79	0.39	0.13	0.87	0.10	0.35	0.44
3	4	-0.06	0.89	0.00	0.79	0.78	0.40	0.13	0.87	0.10	0.35	0.44
7	12	-0.08	0.89	0.01	0.78	0.78	0.40	0.13	0.87	0.10	0.35	0.44
9	14	0.12	0.88	0.01	0.78	0.76	0.40	0.13	0.87	0.10	0.35	0.44
1	11	0.22	0.87	0.05	0.75	0.70	0.40	0.13	0.87	0.09	0.35	0.44
1	13	0.03	0.88	0.00	0.77	0.77	0.39	0.13	0.87	0.10	0.34	0.44
6	13	0.09	0.87	0.01	0.76	0.76	0.39	0.13	0.87	0.09	0.34	0.43
2	9	0.12	0.87	0.01	0.76	0.75	0.39	0.13	0.87	0.09	0.34	0.43
7	10	0.22	0.86	0.05	0.74	0.69	0.39	0.13	0.87	0.09	0.34	0.43
10	11	0.22	0.86	0.05	0.74	0.69	0.39	0.13	0.87	0.09	0.34	0.43
1	5	0.03	0.87	0.00	0.76	0.76	0.38	0.13	0.87	0.10	0.33	0.43
3	10	-0.06	0.87	0.00	0.75	0.75	0.38	0.13	0.87	0.09	0.33	0.43
2	5	0.03	0.87	0.00	0.75	0.75	0.38	0.13	0.87	0.09	0.33	0.42

Table 14—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
2	13	0.03	0.87	0.00	0.75	0.75	0.38	0.13	0.87	0.09	0.33	0.42
5	6	-0.18	0.86	0.03	0.73	0.70	0.38	0.13	0.87	0.09	0.33	0.42
6	12	0.04	0.86	0.00	0.74	0.74	0.37	0.13	0.87	0.09	0.33	0.42
4	14	-0.07	0.86	0.01	0.74	0.73	0.37	0.13	0.87	0.09	0.33	0.42
8	10	0.09	0.86	0.01	0.73	0.72	0.37	0.13	0.87	0.09	0.32	0.41
3	5	-0.27	0.83	0.07	0.69	0.62	0.38	0.13	0.87	0.08	0.34	0.41
5	10	0.37	0.81	0.14	0.66	0.52	0.40	0.13	0.87	0.07	0.35	0.41
2	7	-0.12	0.85	0.01	0.72	0.71	0.37	0.13	0.87	0.09	0.32	0.41
2	6	0.20	0.84	0.04	0.71	0.67	0.37	0.13	0.87	0.08	0.33	0.41
3	12	0.08	0.85	0.01	0.72	0.71	0.36	0.13	0.87	0.09	0.32	0.41
2	4	-0.12	0.85	0.01	0.71	0.70	0.36	0.13	0.87	0.09	0.32	0.41
4	5	0.18	0.84	0.03	0.70	0.67	0.37	0.13	0.87	0.08	0.32	0.40
2	10	0.17	0.83	0.03	0.69	0.66	0.36	0.13	0.87	0.08	0.32	0.40
10	12	-0.25	0.82	0.06	0.67	0.61	0.37	0.13	0.87	0.08	0.32	0.40
4	12	0.12	0.84	0.01	0.70	0.68	0.36	0.13	0.87	0.09	0.31	0.40
11	14	-0.34	0.80	0.11	0.64	0.53	0.38	0.13	0.87	0.07	0.33	0.39
3	13	-0.05	0.84	0.00	0.70	0.70	0.35	0.13	0.87	0.09	0.31	0.39
6	10	0.21	0.82	0.04	0.67	0.63	0.36	0.13	0.87	0.08	0.31	0.39
5	14	-0.15	0.83	0.02	0.68	0.66	0.35	0.13	0.87	0.08	0.31	0.39
11	12	0.13	0.82	0.02	0.68	0.66	0.35	0.13	0.87	0.08	0.30	0.39
2	3	0.01	0.83	0.00	0.68	0.68	0.34	0.13	0.87	0.09	0.30	0.38
6	14	0.00	0.82	0.00	0.67	0.67	0.33	0.13	0.87	0.08	0.29	0.37
13	14	-0.14	0.81	0.02	0.65	0.63	0.33	0.13	0.87	0.08	0.29	0.37
4	13	-0.03	0.80	0.00	0.64	0.64	0.32	0.13	0.87	0.08	0.28	0.36
10	14	-0.12	0.79	0.02	0.63	0.61	0.32	0.13	0.87	0.08	0.28	0.36
6	11	0.04	0.79	0.00	0.63	0.63	0.31	0.13	0.87	0.08	0.28	0.35
10	13	0.09	0.79	0.01	0.62	0.61	0.31	0.13	0.87	0.08	0.27	0.35
3	14	0.04	0.77	0.00	0.60	0.60	0.30	0.13	0.87	0.07	0.26	0.34
5	11	0.14	0.75	0.02	0.57	0.55	0.29	0.13	0.87	0.07	0.26	0.32
2	11	0.22	0.74	0.05	0.55	0.50	0.30	0.13	0.87	0.06	0.26	0.32
11	13	0.03	0.73	0.00	0.54	0.54	0.27	0.13	0.87	0.07	0.23	0.30

Following are the complete results, obtained by assigning different weight values with formula Rank 2:

$$a = 0.17452 \rightarrow w_a = 0.17452 (0.17)$$

$$w_d = 0.82548 (0.83)$$

Table 15

Complete Results from Stoplights with a Higher Difference Using Rank 2

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	$d w_d$	$a w_a$	$d w_d + a w_a$
7	9	-0.07	0.96	0.00	0.92	0.91	0.46	0.83	0.17	0.75	0.08	0.83
7	8	0.11	0.95	0.01	0.91	0.90	0.46	0.83	0.17	0.74	0.08	0.82
1	9	0.12	0.95	0.01	0.90	0.89	0.46	0.83	0.17	0.74	0.08	0.82
5	13	-0.17	0.95	0.03	0.89	0.87	0.46	0.83	0.17	0.71	0.08	0.79
6	8	0.06	0.93	0.00	0.87	0.86	0.44	0.83	0.17	0.71	0.08	0.79
1	8	0.18	0.94	0.03	0.88	0.85	0.46	0.83	0.17	0.70	0.08	0.78
1	7	-0.12	0.93	0.01	0.86	0.85	0.44	0.83	0.17	0.70	0.08	0.78
8	12	-0.27	0.95	0.07	0.91	0.83	0.49	0.83	0.17	0.69	0.09	0.77
12	14	0.01	0.92	0.00	0.84	0.84	0.42	0.83	0.17	0.69	0.07	0.76
7	11	-0.01	0.91	0.00	0.84	0.84	0.42	0.83	0.17	0.69	0.07	0.76
4	9	-0.10	0.92	0.01	0.84	0.83	0.43	0.83	0.17	0.69	0.07	0.76
8	13	-0.01	0.91	0.00	0.83	0.83	0.42	0.83	0.17	0.69	0.07	0.76
3	8	0.05	0.91	0.00	0.83	0.83	0.42	0.83	0.17	0.68	0.07	0.75
1	4	-0.12	0.91	0.01	0.84	0.82	0.43	0.83	0.17	0.68	0.07	0.75
9	13	-0.01	0.91	0.00	0.82	0.82	0.41	0.83	0.17	0.68	0.07	0.75
1	3	0.01	0.90	0.00	0.81	0.81	0.41	0.83	0.17	0.67	0.07	0.74
3	9	-0.13	0.91	0.02	0.83	0.81	0.42	0.83	0.17	0.67	0.07	0.74
9	10	-0.09	0.90	0.01	0.82	0.81	0.41	0.83	0.17	0.67	0.07	0.74
4	11	-0.11	0.90	0.01	0.82	0.81	0.41	0.83	0.17	0.67	0.07	0.74
5	8	-0.16	0.91	0.02	0.82	0.80	0.42	0.83	0.17	0.66	0.07	0.73
3	11	-0.11	0.90	0.01	0.81	0.80	0.41	0.83	0.17	0.66	0.07	0.73

Table 15—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
6	9	-0.04	0.89	0.00	0.80	0.80	0.40	0.83	0.17	0.66	0.07	0.73
5	12	0.02	0.89	0.00	0.79	0.79	0.40	0.83	0.17	0.65	0.07	0.72
8	14	0.01	0.89	0.00	0.79	0.79	0.39	0.83	0.17	0.65	0.07	0.72
12	13	-0.07	0.89	0.01	0.79	0.79	0.40	0.83	0.17	0.65	0.07	0.72
3	4	-0.06	0.89	0.00	0.79	0.78	0.40	0.83	0.17	0.65	0.07	0.72
9	11	-0.25	0.91	0.06	0.83	0.77	0.45	0.83	0.17	0.64	0.08	0.72
4	8	-0.18	0.90	0.03	0.81	0.77	0.42	0.83	0.17	0.64	0.07	0.71
7	12	-0.08	0.89	0.01	0.78	0.78	0.40	0.83	0.17	0.64	0.07	0.71
2	8	0.18	0.90	0.03	0.80	0.77	0.42	0.83	0.17	0.64	0.07	0.71
1	13	0.03	0.88	0.00	0.77	0.77	0.39	0.83	0.17	0.64	0.07	0.70
1	10	0.17	0.89	0.03	0.80	0.77	0.41	0.83	0.17	0.63	0.07	0.70
1	6	0.20	0.89	0.04	0.80	0.76	0.42	0.83	0.17	0.63	0.07	0.70
9	14	0.12	0.88	0.01	0.78	0.76	0.40	0.83	0.17	0.63	0.07	0.70
5	9	-0.30	0.91	0.09	0.83	0.74	0.46	0.83	0.17	0.61	0.08	0.69
8	9	0.40	0.94	0.16	0.89	0.73	0.53	0.83	0.17	0.60	0.09	0.69
1	5	0.03	0.87	0.00	0.76	0.76	0.38	0.83	0.17	0.63	0.07	0.69
6	13	0.09	0.87	0.01	0.76	0.76	0.39	0.83	0.17	0.62	0.07	0.69
2	5	0.03	0.87	0.00	0.75	0.75	0.38	0.83	0.17	0.62	0.07	0.69
2	13	0.03	0.87	0.00	0.75	0.75	0.38	0.83	0.17	0.62	0.07	0.69
3	10	-0.06	0.87	0.00	0.75	0.75	0.38	0.83	0.17	0.62	0.07	0.69
7	13	0.19	0.88	0.03	0.78	0.74	0.41	0.83	0.17	0.61	0.07	0.68
2	9	0.12	0.87	0.01	0.76	0.75	0.39	0.83	0.17	0.62	0.07	0.68
6	12	0.04	0.86	0.00	0.74	0.74	0.37	0.83	0.17	0.61	0.06	0.68
7	14	0.23	0.88	0.05	0.78	0.73	0.42	0.83	0.17	0.60	0.07	0.68
3	6	0.38	0.92	0.14	0.85	0.71	0.50	0.83	0.17	0.58	0.09	0.67
4	14	-0.07	0.86	0.01	0.74	0.73	0.37	0.83	0.17	0.61	0.06	0.67
6	7	0.30	0.89	0.09	0.80	0.71	0.44	0.83	0.17	0.59	0.08	0.67
8	10	0.09	0.86	0.01	0.73	0.72	0.37	0.83	0.17	0.60	0.06	0.66
9	12	-0.35	0.91	0.12	0.82	0.70	0.47	0.83	0.17	0.58	0.08	0.66
3	12	0.08	0.85	0.01	0.72	0.71	0.36	0.83	0.17	0.59	0.06	0.65
1	11	0.22	0.87	0.05	0.75	0.70	0.40	0.83	0.17	0.58	0.07	0.65
1	12	-0.39	0.91	0.15	0.83	0.68	0.49	0.83	0.17	0.56	0.09	0.65

Table 15—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
2	7	-0.12	0.85	0.01	0.72	0.71	0.37	0.83	0.17	0.58	0.06	0.65
5	6	-0.18	0.86	0.03	0.73	0.70	0.38	0.83	0.17	0.58	0.07	0.65
5	7	-0.30	0.88	0.09	0.78	0.69	0.43	0.83	0.17	0.57	0.08	0.64
2	4	-0.12	0.85	0.01	0.71	0.70	0.36	0.83	0.17	0.58	0.06	0.64
8	11	-0.29	0.88	0.08	0.77	0.69	0.43	0.83	0.17	0.57	0.07	0.64
7	10	0.22	0.86	0.05	0.74	0.69	0.39	0.83	0.17	0.57	0.07	0.64
10	11	0.22	0.86	0.05	0.74	0.69	0.39	0.83	0.17	0.57	0.07	0.64
3	13	-0.05	0.84	0.00	0.70	0.70	0.35	0.83	0.17	0.57	0.06	0.64
4	10	-0.39	0.90	0.15	0.82	0.67	0.48	0.83	0.17	0.55	0.08	0.63
4	12	0.12	0.84	0.01	0.70	0.68	0.36	0.83	0.17	0.56	0.06	0.63
2	3	0.01	0.83	0.00	0.68	0.68	0.34	0.83	0.17	0.56	0.06	0.62
2	6	0.20	0.84	0.04	0.71	0.67	0.37	0.83	0.17	0.55	0.06	0.62
4	5	0.18	0.84	0.03	0.70	0.67	0.37	0.83	0.17	0.55	0.06	0.61
2	10	0.17	0.83	0.03	0.69	0.66	0.36	0.83	0.17	0.55	0.06	0.61
6	14	0.00	0.82	0.00	0.67	0.67	0.33	0.83	0.17	0.55	0.06	0.61
5	14	-0.15	0.83	0.02	0.68	0.66	0.35	0.83	0.17	0.55	0.06	0.61
11	12	0.13	0.82	0.02	0.68	0.66	0.35	0.83	0.17	0.54	0.06	0.60
4	13	-0.03	0.80	0.00	0.64	0.64	0.32	0.83	0.17	0.53	0.06	0.59
6	10	0.21	0.82	0.04	0.67	0.63	0.36	0.83	0.17	0.52	0.06	0.58
3	5	-0.27	0.83	0.07	0.69	0.62	0.38	0.83	0.17	0.51	0.07	0.58
1	14	-0.52	0.92	0.27	0.85	0.58	0.56	0.83	0.17	0.48	0.10	0.58
13	14	-0.14	0.81	0.02	0.65	0.63	0.33	0.83	0.17	0.52	0.06	0.58
6	11	0.04	0.79	0.00	0.63	0.63	0.31	0.83	0.17	0.52	0.05	0.57
10	12	-0.25	0.82	0.06	0.67	0.61	0.37	0.83	0.17	0.51	0.06	0.57
10	14	-0.12	0.79	0.02	0.63	0.61	0.32	0.83	0.17	0.51	0.06	0.56
10	13	0.09	0.79	0.01	0.62	0.61	0.31	0.83	0.17	0.50	0.05	0.56
2	12	-0.39	0.86	0.15	0.73	0.58	0.44	0.83	0.17	0.48	0.08	0.55
3	7	0.54	0.91	0.29	0.83	0.54	0.56	0.83	0.17	0.45	0.10	0.55
3	14	0.04	0.77	0.00	0.60	0.60	0.30	0.83	0.17	0.49	0.05	0.54
5	11	0.14	0.75	0.02	0.57	0.55	0.29	0.83	0.17	0.45	0.05	0.50
11	14	-0.34	0.80	0.11	0.64	0.53	0.38	0.83	0.17	0.43	0.07	0.50
5	10	0.37	0.81	0.14	0.66	0.52	0.40	0.83	0.17	0.43	0.07	0.50

Table 15—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 Week Correlation Coefficient	r^2 Weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
11	13	0.03	0.73	0.00	0.54	0.54	0.27	0.83	0.17	0.44	0.05	0.49
4	7	-0.62	0.92	0.39	0.85	0.46	0.62	0.83	0.17	0.38	0.11	0.49
2	11	0.22	0.74	0.05	0.55	0.50	0.30	0.83	0.17	0.41	0.05	0.46
2	14	-0.52	0.81	0.27	0.66	0.39	0.46	0.83	0.17	0.32	0.08	0.40
4	6	-0.60	0.85	0.36	0.73	0.37	0.54	0.83	0.17	0.30	0.09	0.40
1	2	1.00	0.88	1.00	0.78	0.22	0.89	0.83	0.17	0.18	0.16	0.34

Here is a third result with different weight values, again using Formula Rank 2:

$$a = 0.47452 \rightarrow w_a = 0.47452 (0.47)$$

$$w_d = 0.52548 (0.53)$$

Table 16

Complete Results from Intersections with Balanced Average and Distance Weights Using Rank 2

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 week Correlation Coefficient	r^2 weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
7	9	-0.07	0.96	0.00	0.92	0.91	0.46	0.53	0.47	0.48	0.22	0.70
7	8	0.11	0.95	0.01	0.91	0.90	0.46	0.53	0.47	0.47	0.22	0.69
1	9	0.12	0.95	0.01	0.90	0.89	0.46	0.53	0.47	0.47	0.22	0.69
5	13	-0.17	0.95	0.03	0.89	0.87	0.46	0.53	0.47	0.45	0.22	0.67
8	12	-0.27	0.95	0.07	0.91	0.83	0.49	0.53	0.47	0.44	0.23	0.67
1	8	0.18	0.94	0.03	0.88	0.85	0.46	0.53	0.47	0.45	0.22	0.66
6	8	0.06	0.93	0.00	0.87	0.86	0.44	0.53	0.47	0.45	0.21	0.66

Table 16—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 week Correlation Coefficient	r^2 weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
1	7	-0.12	0.93	0.01	0.86	0.85	0.44	0.53	0.47	0.45	0.21	0.66
4	9	-0.10	0.92	0.01	0.84	0.83	0.43	0.53	0.47	0.44	0.20	0.64
12	14	0.01	0.92	0.00	0.84	0.84	0.42	0.53	0.47	0.44	0.20	0.64
7	11	-0.01	0.91	0.00	0.84	0.84	0.42	0.53	0.47	0.44	0.20	0.64
8	13	-0.01	0.91	0.00	0.83	0.83	0.42	0.53	0.47	0.44	0.20	0.64
1	4	-0.12	0.91	0.01	0.84	0.82	0.43	0.53	0.47	0.43	0.20	0.63
8	9	0.40	0.94	0.16	0.89	0.73	0.53	0.53	0.47	0.38	0.25	0.63
3	8	0.05	0.91	0.00	0.83	0.83	0.42	0.53	0.47	0.43	0.20	0.63
9	13	-0.01	0.91	0.00	0.82	0.82	0.41	0.53	0.47	0.43	0.20	0.63
3	9	-0.13	0.91	0.02	0.83	0.81	0.42	0.53	0.47	0.43	0.20	0.63
4	11	-0.11	0.90	0.01	0.82	0.81	0.41	0.53	0.47	0.42	0.20	0.62
9	10	-0.09	0.90	0.01	0.82	0.81	0.41	0.53	0.47	0.42	0.20	0.62
1	3	0.01	0.90	0.00	0.81	0.81	0.41	0.53	0.47	0.43	0.19	0.62
5	8	-0.16	0.91	0.02	0.82	0.80	0.42	0.53	0.47	0.42	0.20	0.62
9	11	-0.25	0.91	0.06	0.83	0.77	0.45	0.53	0.47	0.41	0.21	0.62
3	11	-0.11	0.90	0.01	0.81	0.80	0.41	0.53	0.47	0.42	0.19	0.61
5	9	-0.30	0.91	0.09	0.83	0.74	0.46	0.53	0.47	0.39	0.22	0.61
6	9	-0.04	0.89	0.00	0.80	0.80	0.40	0.53	0.47	0.42	0.19	0.61
3	6	0.38	0.92	0.14	0.85	0.71	0.50	0.53	0.47	0.37	0.24	0.61
4	8	-0.18	0.90	0.03	0.81	0.77	0.42	0.53	0.47	0.41	0.20	0.61
12	13	-0.07	0.89	0.01	0.79	0.79	0.40	0.53	0.47	0.41	0.19	0.60
5	12	0.02	0.89	0.00	0.79	0.79	0.40	0.53	0.47	0.42	0.19	0.60
2	8	0.18	0.90	0.03	0.80	0.77	0.42	0.53	0.47	0.40	0.20	0.60
8	14	0.01	0.89	0.00	0.79	0.79	0.39	0.53	0.47	0.41	0.19	0.60
3	4	-0.06	0.89	0.00	0.79	0.78	0.40	0.53	0.47	0.41	0.19	0.60
1	10	0.17	0.89	0.03	0.80	0.77	0.41	0.53	0.47	0.40	0.20	0.60
1	6	0.20	0.89	0.04	0.80	0.76	0.42	0.53	0.47	0.40	0.20	0.60
7	12	-0.08	0.89	0.01	0.78	0.78	0.40	0.53	0.47	0.41	0.19	0.60
1	12	-0.39	0.91	0.15	0.83	0.68	0.49	0.53	0.47	0.36	0.23	0.59
9	12	-0.35	0.91	0.12	0.82	0.70	0.47	0.53	0.47	0.37	0.22	0.59
1	13	0.03	0.88	0.00	0.77	0.77	0.39	0.53	0.47	0.41	0.18	0.59
9	14	0.12	0.88	0.01	0.78	0.76	0.40	0.53	0.47	0.40	0.19	0.59

Table 16—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 week Correlation Coefficient	r^2 weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
6	7	0.30	0.89	0.09	0.80	0.71	0.44	0.53	0.47	0.37	0.21	0.58
7	13	0.19	0.88	0.03	0.78	0.74	0.41	0.53	0.47	0.39	0.19	0.58
7	14	0.23	0.88	0.05	0.78	0.73	0.42	0.53	0.47	0.38	0.20	0.58
6	13	0.09	0.87	0.01	0.76	0.76	0.39	0.53	0.47	0.40	0.18	0.58
4	10	-0.39	0.90	0.15	0.82	0.67	0.48	0.53	0.47	0.35	0.23	0.58
1	5	0.03	0.87	0.00	0.76	0.76	0.38	0.53	0.47	0.40	0.18	0.58
2	9	0.12	0.87	0.01	0.76	0.75	0.39	0.53	0.47	0.39	0.18	0.58
2	5	0.03	0.87	0.00	0.75	0.75	0.38	0.53	0.47	0.40	0.18	0.57
2	13	0.03	0.87	0.00	0.75	0.75	0.38	0.53	0.47	0.40	0.18	0.57
3	10	-0.06	0.87	0.00	0.75	0.75	0.38	0.53	0.47	0.39	0.18	0.57
1	14	-0.52	0.92	0.27	0.85	0.58	0.56	0.53	0.47	0.31	0.27	0.57
5	7	-0.30	0.88	0.09	0.78	0.69	0.43	0.53	0.47	0.36	0.20	0.57
6	12	0.04	0.86	0.00	0.74	0.74	0.37	0.53	0.47	0.39	0.18	0.57
8	11	-0.29	0.88	0.08	0.77	0.69	0.43	0.53	0.47	0.36	0.20	0.56
4	14	-0.07	0.86	0.01	0.74	0.73	0.37	0.53	0.47	0.39	0.18	0.56
1	11	0.22	0.87	0.05	0.75	0.70	0.40	0.53	0.47	0.37	0.19	0.56
8	10	0.09	0.86	0.01	0.73	0.72	0.37	0.53	0.47	0.38	0.18	0.56
3	7	0.54	0.91	0.29	0.83	0.54	0.56	0.53	0.47	0.29	0.27	0.55
5	6	-0.18	0.86	0.03	0.73	0.70	0.38	0.53	0.47	0.37	0.18	0.55
7	10	0.22	0.86	0.05	0.74	0.69	0.39	0.53	0.47	0.36	0.19	0.55
10	11	0.22	0.86	0.05	0.74	0.69	0.39	0.53	0.47	0.36	0.19	0.55
3	12	0.08	0.85	0.01	0.72	0.71	0.36	0.53	0.47	0.37	0.17	0.55
2	7	-0.12	0.85	0.01	0.72	0.71	0.37	0.53	0.47	0.37	0.17	0.54
2	4	-0.12	0.85	0.01	0.71	0.70	0.36	0.53	0.47	0.37	0.17	0.54
1	2	1.00	0.88	1.00	0.78	0.22	0.89	0.53	0.47	0.12	0.42	0.54
4	7	-0.62	0.92	0.39	0.85	0.46	0.62	0.53	0.47	0.24	0.29	0.54
3	13	-0.05	0.84	0.00	0.70	0.70	0.35	0.53	0.47	0.37	0.17	0.53
4	12	0.12	0.84	0.01	0.70	0.68	0.36	0.53	0.47	0.36	0.17	0.53
2	6	0.20	0.84	0.04	0.71	0.67	0.37	0.53	0.47	0.35	0.18	0.53
4	5	0.18	0.84	0.03	0.70	0.67	0.37	0.53	0.47	0.35	0.17	0.52
2	10	0.17	0.83	0.03	0.69	0.66	0.36	0.53	0.47	0.35	0.17	0.52
2	3	0.01	0.83	0.00	0.68	0.68	0.34	0.53	0.47	0.36	0.16	0.52

Table 16—Continued

Intersection A	Intersection B	Week Correlation Coefficient	Weekend Correlation Coefficient	r^2 week Correlation Coefficient	r^2 weekend Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
5	14	-0.15	0.83	0.02	0.68	0.66	0.35	0.53	0.47	0.35	0.17	0.51
2	12	-0.39	0.86	0.15	0.73	0.58	0.44	0.53	0.47	0.30	0.21	0.51
11	12	0.13	0.82	0.02	0.68	0.66	0.35	0.53	0.47	0.35	0.16	0.51
6	14	0.00	0.82	0.00	0.67	0.67	0.33	0.53	0.47	0.35	0.16	0.51
3	5	-0.27	0.83	0.07	0.69	0.62	0.38	0.53	0.47	0.33	0.18	0.51
6	10	0.21	0.82	0.04	0.67	0.63	0.36	0.53	0.47	0.33	0.17	0.50
10	12	-0.25	0.82	0.06	0.67	0.61	0.37	0.53	0.47	0.32	0.17	0.50
4	13	-0.03	0.80	0.00	0.64	0.64	0.32	0.53	0.47	0.34	0.15	0.49
13	14	-0.14	0.81	0.02	0.65	0.63	0.33	0.53	0.47	0.33	0.16	0.49
6	11	0.04	0.79	0.00	0.63	0.63	0.31	0.53	0.47	0.33	0.15	0.48
10	14	-0.12	0.79	0.02	0.63	0.61	0.32	0.53	0.47	0.32	0.15	0.48
10	13	0.09	0.79	0.01	0.62	0.61	0.31	0.53	0.47	0.32	0.15	0.47
5	10	0.37	0.81	0.14	0.66	0.52	0.40	0.53	0.47	0.27	0.19	0.46
3	14	0.04	0.77	0.00	0.60	0.60	0.30	0.53	0.47	0.31	0.14	0.46
11	14	-0.34	0.80	0.11	0.64	0.53	0.38	0.53	0.47	0.28	0.18	0.45
4	6	-0.60	0.85	0.36	0.73	0.37	0.54	0.53	0.47	0.19	0.26	0.45
5	11	0.14	0.75	0.02	0.57	0.55	0.29	0.53	0.47	0.29	0.14	0.43
2	14	-0.52	0.81	0.27	0.66	0.39	0.46	0.53	0.47	0.20	0.22	0.42
11	13	0.03	0.73	0.00	0.54	0.54	0.27	0.53	0.47	0.28	0.13	0.41
2	11	0.22	0.74	0.05	0.55	0.50	0.30	0.53	0.47	0.26	0.14	0.40

A specific subset of results may be extracted through this formula by considering the weight as a percentage and with each weight set according to its importance in a specific situation.

Here is an additional example of how to use these formulas to compare items being sold in a store and how to determine the most interesting information, such as

which items are sold the most, whether an item sells more in the morning or afternoon, etc. First of all, three things must be decided:

1. For which items would the store owners like to know the correlation?
2. How long would they like to collect information?
3. For what time periods will they compare the information?

Let's say that they would like to collect information about ten items during the morning and afternoon for ten days. Here is the raw data:

Table 17

Items Sold Morning (Data Simulated)

Day	1) Shampoo	2) Toothpaste	3) cereal	4) milk	5) beef	6) chicken	7) strawberries	8) oranges	9) jewelry	10) TV
1	15	10	20	25	13	10	6	9	5	3
2	14	13	15	20	11	9	4	8	4	2
3	16	15	18	19	9	11	9	10	8	5
4	14	11	14	17	14	12	10	12	9	7
5	18	10	19	13	15	14	8	11	7	9
6	19	13	17	21	12	10	5	7	6	8
7	15	10	18	20	13	9	7	9	5	3
8	17	16	15	18	10	8	4	8	4	2
9	13	12	19	13	17	13	5	13	9	4
10	12	10	19	17	18	15	11	10	7	7

Table18

Items Sold Evening (Data Simulated)

	1) Shampoo	2) Toothpaste	3) Cereal	4) Milk	5) Beef	6) Chicken	7) Strawberries	8) Oranges	9) Jewelry	10) TV
1	20	15	25	30	18	15	11	10	9	2
2	19	18	20	25	16	14	9	6	5	1
3	21	20	23	24	14	16	6	8	6	5
4	19	21	19	23	19	17	10	4	10	7
5	23	15	24	28	20	19	7	7	8	8
6	24	18	22	26	18	15	5	11	6	3
7	20	15	23	25	18	14	8	9	11	6
8	22	21	20	23	10	13	4	5	7	4
9	18	17	24	18	22	18	8	7	4	9
10	17	15	24	22	23	20	6	13	5	10

Once the data is obtained, calculate its correlation coefficient, which can be done on an Excel spreadsheet:

Table 19

Morning Correlation Coefficient Matrix

Items	1) Shampoo	2) Toothpaste	3) Cereal	4) Milk	5) Beef	6) Chicken	7) Strawberries	8) Oranges	9) Jewelry	10) TV
1) Shampoo	1									
2) Toothpaste	0.34	1								
3) Cereal	-0.08	-0.49	1							
4) Milk	0.13	0.04	0.00	1						
5) Beef	-0.53	-0.73	0.45	-0.50	1					
6) Chicken	-0.35	-0.52	0.45	-0.64	0.80	1				
7) Strawberries	-0.33	-0.44	0.18	-0.23	0.40	0.68	1			
8) Oranges	-0.48	-0.35	0.21	-0.70	0.60	0.69	0.46	1		
9) Jewelry	-0.27	-0.19	0.15	-0.59	0.47	0.72	0.59	0.84	1	
10) TV	0.29	-0.35	0.17	-0.40	0.42	0.70	0.58	0.27	0.57	1

Table 20

Evening Correlation Coefficient Matrix

Items	1) Shampoo	2) Toothpaste	3) Cereal	4) Milk	5) Beef	6) Chicken	7) Strawberries	8) Oranges	9) Jewelry	10) TV
1) Shampoo	1									
2) Toothpaste	0.17	1								
3) Cereal	-0.08	-0.79	1							
4) Milk	0.53	-0.34	0.20	1						
5) Beef	-0.47	-0.64	0.51	-0.18	1					
6) Chicken	-0.35	-0.37	0.45	-0.27	0.79	1				
7) Strawberries	-0.46	-0.26	0.08	0.26	0.35	0.03	1			
8) Oranges	-0.07	-0.65	0.68	0.19	0.45	0.29	-0.18	1		
9) Jewelry	0.19	-0.09	-0.10	0.50	-0.11	-0.27	0.40	-0.19	1	
10) TV	-0.39	-0.24	0.32	-0.55	0.64	0.83	-0.15	0.13	-0.09	1

Applying the Rank 1 formula (New harmonic mean) and arranging those values in ascending order, here is the interesting information:

Table 21

Complete Comparison Result from the Store Items Using Rank 1

Items A	Item B	Morning Correlation Coefficient	Evening Correlation Coefficient	r^2 Morning Correlation Coefficient	r^2 Evening Correlation Coefficient	Difference	Average	Rank 1 Formula
8) Oranges	9) Jewelry	0.84	-0.19	0.71	0.04	0.68	0.37	0.48
2) Toothpaste	3) Cereal	-0.49	-0.79	0.24	0.63	0.39	0.43	0.41
6) Chicken	9) Jewelry	0.72	-0.27	0.52	0.07	0.44	0.29	0.35
4) Milk	8) Oranges	-0.70	0.19	0.49	0.04	0.45	0.26	0.33
6) Chicken	8) Oranges	0.69	0.29	0.47	0.08	0.39	0.28	0.33
6) Chicken	7) Strawberries	0.68	0.03	0.47	0.00	0.47	0.23	0.31
3) Cereal	8) Oranges	0.21	0.68	0.04	0.46	0.41	0.25	0.31
6) Chicken	10) TV	0.70	0.83	0.49	0.68	0.19	0.59	0.29
2) Toothpaste	8) Oranges	-0.35	-0.65	0.12	0.43	0.31	0.27	0.29
4) Milk	6) Chicken	-0.64	-0.27	0.40	0.07	0.33	0.24	0.28
5) Beef	10) TV	0.42	0.64	0.17	0.41	0.23	0.29	0.26
7) Strawberries	10) TV	0.58	-0.15	0.34	0.02	0.32	0.18	0.23
7) Strawberries	9) Jewelry	0.59	0.40	0.35	0.16	0.19	0.25	0.22
9) Jewelry	10) TV	0.57	-0.09	0.32	0.01	0.31	0.16	0.22
5) Beef	8) Oranges	0.60	0.45	0.36	0.20	0.16	0.28	0.20
1) Shampoo	4) Milk	0.13	0.53	0.02	0.28	0.26	0.15	0.19
2) Toothpaste	5) Beef	-0.73	-0.64	0.53	0.41	0.12	0.47	0.19
4) Milk	10) TV	-0.40	-0.55	0.16	0.31	0.14	0.24	0.18
4) Milk	5) Beef	-0.50	-0.18	0.25	0.03	0.22	0.14	0.17
2) Toothpaste	6) Chicken	-0.52	-0.37	0.27	0.14	0.13	0.20	0.16
1) Shampoo	8) Oranges	-0.48	-0.07	0.23	0.01	0.23	0.12	0.16
5) Beef	9) Jewelry	0.47	-0.11	0.22	0.01	0.21	0.12	0.15
7) Strawberries	8) Oranges	0.46	-0.18	0.21	0.03	0.18	0.12	0.15
4) Milk	9) Jewelry	-0.59	0.50	0.34	0.25	0.09	0.30	0.14
2) Toothpaste	7) Strawberries	-0.44	-0.26	0.19	0.07	0.13	0.13	0.13
1) Shampoo	7) Strawberries	-0.33	-0.46	0.11	0.21	0.10	0.16	0.12

Table 21—Continued

Items A	Item B	Morning Correlation Coefficient	Evening Correlation Coefficient	r^2 Morning Correlation Coefficient	r^2 Evening Correlation Coefficient	Difference	Average	Rank 1 Formula
1) Shampoo	5) Beef	-0.53	-0.47	0.28	0.22	0.06	0.25	0.10
3) Cereal	5) Beef	0.45	0.51	0.20	0.26	0.06	0.23	0.09
1) Shampoo	10) TV	0.29	-0.39	0.09	0.15	0.07	0.12	0.09
1) Shampoo	2) Toothpaste	0.34	0.17	0.12	0.03	0.09	0.07	0.08
2) Toothpaste	4) Milk	0.04	-0.34	0.00	0.11	0.11	0.06	0.08
2) Toothpaste	10) TV	-0.35	-0.24	0.12	0.06	0.06	0.09	0.08
3) Cereal	10) TV	0.17	0.32	0.03	0.10	0.07	0.07	0.07
5) Beef	7) Strawberries	0.40	0.35	0.16	0.12	0.03	0.14	0.06
8) Oranges	10) TV	0.27	0.13	0.07	0.02	0.06	0.05	0.05
1) Shampoo	9) Jewelry	-0.27	0.19	0.07	0.04	0.04	0.05	0.04
5) Beef	6) Chicken	0.80	0.79	0.64	0.63	0.01	0.64	0.03
3) Cereal	4) Milk	0.00	0.20	8.82	0.04	0.04	0.02	0.03
4) Milk	7) Strawberries	-0.23	0.26	0.05	0.07	0.02	0.06	0.03
2) Toothpaste	9) Jewelry	-0.19	-0.09	0.03	0.01	0.03	0.02	0.02
3) Cereal	7) Strawberries	0.18	0.08	0.03	0.01	0.03	0.02	0.02
3) Cereal	9) Jewelry	0.15	-0.10	0.02	0.01	0.01	0.02	0.01
1) Shampoo	6) Chicken	-0.35	-0.35	0.12	0.12	8.32	0.12	1.67
3) Cereal	6) Chicken	0.45	0.45	0.20	0.20	5.55	0.20	1.11
1) Shampoo	3) Cereal	-0.08	-0.08	0.01	0.01	0.00	0.01	0.00

Applying the Rank 2 formula, where a higher weight was assigned to the average, here is the information:

Table 22

Complete Comparison Result from the Store Items Using Rank 2

Item A	Item B	Morning Correlation Coefficient	Evening Correlation Coefficient	r^2 Morning Correlation Coefficient	r^2 Evening Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
5) Beef	6) Chicken	0.80	0.79	0.64	0.63	0.01	0.64	0.20	0.80	0.00	0.51	0.52
6) Chicken	10) TV	0.70	0.83	0.49	0.68	0.19	0.59	0.20	0.80	0.04	0.47	0.51
8) Oranges	9) Jewelry	0.84	-0.19	0.71	0.04	0.68	0.37	0.20	0.80	0.13	0.30	0.43
2) Toothpaste	3) Cereal	-0.49	-0.79	0.24	0.63	0.39	0.43	0.20	0.80	0.08	0.35	0.43
2) Toothpaste	5) Beef	-0.73	-0.64	0.53	0.41	0.12	0.47	0.20	0.80	0.02	0.38	0.40
6) Chicken	9) Jewelry	0.72	-0.27	0.52	0.07	0.44	0.29	0.20	0.80	0.09	0.24	0.32
6) Chicken	8) Oranges	0.69	0.29	0.47	0.08	0.39	0.28	0.20	0.80	0.08	0.22	0.30
4) Milk	8) Oranges	-0.70	0.19	0.49	0.04	0.45	0.26	0.20	0.80	0.09	0.21	0.30
3) Cereal	8) Oranges	0.21	0.68	0.04	0.46	0.41	0.25	0.20	0.80	0.08	0.20	0.28
6) Chicken	7) Strawberries	0.68	0.03	0.47	0.00	0.47	0.23	0.20	0.80	0.09	0.19	0.28
5) Beef	10) TV	0.42	0.64	0.17	0.41	0.23	0.29	0.20	0.80	0.05	0.23	0.28
2) Toothpaste	8) Oranges	-0.35	-0.65	0.12	0.43	0.31	0.27	0.20	0.80	0.06	0.22	0.28
5) Beef	8) Oranges	0.60	0.45	0.36	0.20	0.16	0.28	0.20	0.80	0.03	0.23	0.26
4) Milk	9) Jewelry	-0.59	0.50	0.34	0.25	0.09	0.30	0.20	0.80	0.02	0.24	0.26
4) Milk	6) Chicken	-0.64	-0.27	0.40	0.07	0.33	0.24	0.20	0.80	0.07	0.19	0.26
7) Strawberries	9) Jewelry	0.59	0.40	0.35	0.16	0.19	0.25	0.20	0.80	0.04	0.20	0.24
4) Milk	10) TV	-0.40	-0.55	0.16	0.31	0.14	0.24	0.20	0.80	0.03	0.19	0.22
1) Shampoo	5) Beef	-0.53	-0.47	0.28	0.22	0.06	0.25	0.20	0.80	0.01	0.20	0.21
7) Strawberries	10) TV	0.58	-0.15	0.34	0.02	0.32	0.18	0.20	0.80	0.06	0.15	0.21
3) Cereal	5) Beef	0.45	0.51	0.20	0.26	0.06	0.23	0.20	0.80	0.01	0.18	0.20
9) Jewelry	10) TV	0.57	-0.09	0.32	0.01	0.31	0.16	0.20	0.80	0.06	0.13	0.19
2) Toothpaste	6) Chicken	-0.52	-0.37	0.27	0.14	0.13	0.20	0.20	0.80	0.03	0.16	0.19
1) Shampoo	4) Milk	0.13	0.53	0.02	0.28	0.26	0.15	0.20	0.80	0.05	0.12	0.17
3) Cereal	6) Chicken	0.45	0.45	0.20	0.20	5.55	0.20	0.20	0.80	1.10	0.16	0.16
4) Milk	5) Beef	-0.50	-0.18	0.25	0.03	0.22	0.14	0.20	0.80	0.04	0.11	0.16
1) Shampoo	7) Strawberries	-0.33	-0.46	0.11	0.21	0.10	0.16	0.20	0.80	0.02	0.13	0.15
1) Shampoo	8) Oranges	-0.48	-0.07	0.23	0.01	0.23	0.12	0.20	0.80	0.04	0.10	0.14
5) Beef	9) Jewelry	0.47	-0.11	0.22	0.01	0.21	0.12	0.20	0.80	0.04	0.09	0.13
7) Strawberries	8) Oranges	0.46	-0.18	0.21	0.03	0.18	0.12	0.20	0.80	0.04	0.10	0.13
2) Toothpaste	7) Strawberries	-0.44	-0.26	0.19	0.07	0.13	0.13	0.20	0.80	0.02	0.10	0.13
5) Beef	7) Strawberries	0.40	0.35	0.16	0.12	0.03	0.14	0.20	0.80	0.01	0.11	0.12

Table 22—Continued

Item A	Item B	Morning Correlation Coefficient	Evening Correlation Coefficient	r^2 Morning Correlation Coefficient	r^2 Evening Correlation Coefficient	Difference	Average	w_d	w_a	dw_d	aw_a	$dw_d + aw_a$
1) Shampoo	10) TV	0.29	-0.39	0.09	0.15	0.07	0.12	0.20	0.80	0.01	0.10	0.11
1) Shampoo	6) Chicken	-0.35	-0.35	0.12	0.12	8.32	0.12	0.20	0.80	1.63	0.10	0.10
2) Toothpaste	10) TV	-0.35	-0.24	0.12	0.06	0.06	0.09	0.20	0.80	0.01	0.07	0.09
1) Shampoo	2) Toothpaste	0.34	0.17	0.12	0.03	0.09	0.07	0.20	0.80	0.02	0.06	0.08
2) Toothpaste	4) Milk	0.04	-0.34	0.00	0.11	0.11	0.06	0.20	0.80	0.02	0.05	0.07
3) Cereal	10) TV	0.17	0.32	0.03	0.10	0.07	0.07	0.20	0.80	0.01	0.05	0.07
4) Milk	7) Strawberries	-0.23	0.26	0.05	0.07	0.02	0.06	0.20	0.80	0.00	0.05	0.05
1) Shampoo	9) Jewelry	-0.27	0.19	0.07	0.04	0.04	0.05	0.20	0.80	0.01	0.04	0.05
8) Oranges	10) TV	0.27	0.13	0.07	0.02	0.06	0.05	0.20	0.80	0.01	0.04	0.05
3) Cereal	4) Milk	0.00	0.20	8.82	0.04	0.04	0.02	0.20	0.80	0.01	0.02	0.02
2) Toothpaste	9) Jewelry	-0.19	-0.09	0.03	0.01	0.03	0.02	0.20	0.80	0.01	0.02	0.02
3) Cereal	7) Strawberries	0.18	0.08	0.03	0.01	0.03	0.02	0.20	0.80	0.00	0.02	0.02
3) Cereal	9) Jewelry	0.15	-0.10	0.02	0.01	0.01	0.02	0.20	0.80	0.00	0.01	0.02
1) Shampoo	3) Cereal	-0.08	-0.08	0.01	0.01	0.00	0.01	0.20	0.80	0.00	0.00	0.00

Once the values in the column formula are arranged in descending order, what is obtained is the maximum values of the difference and the average. If arranged in ascending order, the minimum values are obtained. The following are explanations of the meaning of the average and the difference in this study, and the use of Rank 1 and Rank 2 formulas:

Meaning of the average: A higher value means a busy intersection or an important variable, a low value means a less busy intersection or a less important variable.

Meaning of the difference: A higher value means that those variables are very different at those two time periods; a low value means that those variables are not very different at those two time periods.

Rank 1 formula values: By arranging those values in descending order, the variables are arranged according to which ones are busier and have a larger difference at both time periods.

Rank 2 formula values: By assigning more or less weight to the difference or average, the maximum value is obtained depending on which one is assigned more weight. By arranging the values in the column formula in descending order, those intersections that have a maximum value are obtained; in other words, those that have a larger difference to those that have a smaller difference, or busier to less busy (important), depending on what is necessary in a specific case.

Here is how to use these formulas for any situation:

1. Decide which items need to be compared.
2. Select the observations or the time to collect the data.
3. Choose which periods of time are going to be compared (days, weeks, months).
4. Once the data is ready, calculate the correlation coefficient of each of those matrices.
5. Decide what information needs obtaining:

The difference and the average (Rank 1 formula)

Just the difference (Rank 2 formula)

Just the average (Rank 2 formula)

6. For both formulas, make sure that the values are positive. To do this, square the correlation coefficient values.
7. Obtain their differences and their averages with these formulas:

Difference $d = |V1 - V2|$.

Average $a = 0.5 (V1+V2)$.

8. After deciding what information is needed, use one of the following formulas:

Rank 1: Here it is necessary to use the previous data from the averages and differences that were already calculated

$$Rank\ 1 = 2 \frac{da}{d + a}$$

Rank 2: To use this formula, add two parameters:

Distance weight = w_d [0, 1]

Average weight = w_a [0, 1]

For the formula to work, the values (w_d and w_a) must add up to 1. After that, use the already-calculated averages and differences to add them to this formula:

$$Rank2 = dw_d + aw_a$$

The final thing to do for both formulas is to arrange those values in descending order to have the data shown from the most interesting (important) information to the least important information, depending on what needs to be calculated. The Rank 1 formula will obtain the intersections that are higher in difference and higher in average to the lower ones. The Rank 2 formula will obtain the intersections that are higher in difference or higher in average to the lower ones.

The maximum values to obtain in each formula will be as follows (these maximum values are for both formulas):

Difference $d = |V1 - V2|$.

$d = |1 - 0| = 1$ which will be the highest difference

Average $a = 0.5 (V1+V2)$.

$a = 0.5 (1+1) = 1$ which will be the highest average

Since: $d \neq a$ (they will never be equal)

$$a > d \rightarrow a = 0.5$$

$$d > a \rightarrow d = 1$$

or

$a = 1$ and $d = 0$ or vice versa.

Rank 1 formula values would be between 0 and .6, based on this:

Example of the maximum value:

$$\text{Rank 1} = 2 \frac{(1)(0.5)}{1+0.5} = 2 \frac{0.5}{1.5} = 0.6$$

Those values that are in the top, with the value closer to 0.6, are the ones that have a perfect mean (higher) between the difference and the average, compared to those values that are closer to zero, which means that they do not have a perfect balance between the difference and the average.

Rank 2 formula values would range between 0 and 1, with 1 being the highest.

See the following example: $1(1) + 0.5(0) = 1$ or

$$1(0) + 0.5(1) = 0.5 \text{ or}$$

$$1(0.5) + 0.5(0.5) = 0.75$$

When the weight on average has a higher value (closer to 1), it means that those two variables on the top, having a higher value, have a higher correlation (are busier or more interesting) than the lower variables, which have a value closer to 0. If more weight is placed on the difference, a value closer to 1 will mean that those two intersections are very different, as compared to intersections with a lesser value. The value in the intersections that are the most similar will be 0 or close to 0.

CHAPTER 5

SUMMARY, CONCLUSION AND FUTURE WORK

Summary

This thesis is focused on finding correlations among variables on different time periods and comparing several types of variables with each other. The article that inspired this research was focused on finding overlapping correlations among variables, but in just one period of time. The future work recommended by this article was to compare variables on different time periods. To find the correlation among variables on different time periods, a correlation coefficient matrix was used in this document, since this matrix is already normalized and it is simple to compare the matrices to each other. Since the correlation coefficient results are the values of how much those variables are correlated with each other, comparing two matrices value by value to obtain their average and difference shows how each set of values correlates with the other, whether it be high or low, and how different each set of values correlates with each other.

Also, the harmonic mean was modified to find out the best or the worse combination of variables between the average and difference among each option. While working with the previous formula, I realized that not everybody would like to know the average and the difference of each pairing of variables at the same time; so I then found a way to choose which one I wanted to know based on how much weight I assigned to one of the categories. Simulated data was used to prove my theory.

Conclusion and Future Work

In this research the most interesting correlation among variables was found by comparing two correlation coefficients from two different time periods. With the first formula (“Rank 1”), a perfect mean between the average and difference between two intersections can be found. The other formula (“Rank 2”) produces the highest value only in their difference or only in their average by assigning a weight value to each, depending on what needs to be known in a specific situation. Also, with the causal graphical model, how their correlation changes at different times can be observed. In reality these formulas can be used to compare any other types of data from a wide variety of situations. To take advantage of the most interesting situations, a way was needed to compare information (data) at different time periods. With these formulas, the correlation between several data elements at different times may be found and arranged as needed.

In future work, the results of two correlation coefficients and two partial correlation coefficients could be compared to know how much of a difference would be obtained as a result. Maybe by comparing these two matrices future researchers can identify from which one they can obtain the most correct correlation among variables.

REFERENCE LIST

- Begin, M., Roff, D., & Debat, V. (2004). *The effect of temperature and wing morphology on quantitative genetic variation in the cricket *Gryllus firmus*, with an appendix examining the statistical properties of the Jackknife–MANOVA method of matrix comparison*. Retrieved May 1, 2015 from <http://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2004.00772.x/epdf>
- Borgelt, C., Steinbrecher, M., & Kruse, R. (2009). *Graphical models representations for learning, reasoning and data mining*. Retrieved May 1, 2015 from United Kingdom: Wiley & Sons. http://samples.sainsburysebooks.co.uk/9780470749562_sample_379060.pdf
- Cheng, J. (2013). *Mix and covariance dependent graphical models*. Retrieved May 1, 2015 from http://deepblue.lib.umich.edu/bitstream/handle/2027.42/99862/jjeche_1.pdf?sequence=1
- Cheverud, J., & Marroig, G. (2007). *Research article comparing covariance matrices: Random Skewers method compared to the Rouse common principal components model*. Retrieved May 1, 2015 from <http://www.scielo.br/pdf/gmb/v30n2/a27v30n2.pdf>
- Han, L. Song, G., Cong, G., & Xie, K. (2012). *Overlapping decomposition for causal graphical modeling*. Retrieved May 1, 2015 from <http://wan.poly.edu/KDD2012/docs/p114.pdf>
- Hsieh, Sustik, Dhillon, & Ravikumar. (2011). *Sparse inverse covariance matrix estimation using quadratic approximation*. Retrieved May 1, 2015 from <https://www.cs.utexas.edu/~pradeepr/paperz/invcov.pdf>
- Roff, Prokkola, Krams, & Rantala. (2012) *There is more than one way to skin a Gmatrix*. Retrieved May 1, 2015 from <http://users.utu.fi/mjranta/reprints/Roff%20et%20al%202012.pdf>
- Rouse, M. (2005–2015). *Search data management*. Retrieved from <http://searchdatamanagement.techtarget.com/definition/data-management>
- Sun, H & Li, H. (2012). *Robust Gaussian graphical modeling via l_1 penalization*. Retrieved May 1, 2015 from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3535542/pdf/nihms378767.pdf>
- Ware, W., Ferron, J., & Miller, B. (2013). *Introductory statistics: A conceptual approach using R*. New York, NY: Routledge.