Crime, Business, and Vacant/Abandoned Buildings

Canonical Correlation Analysis

Alejandro Hernandez Segura
Team: **Angry Chickens**

Executive Summary

This report seeks to identify and quantify two relationships between three community indicators: firstly, we are interested in finding the relationship between the characteristics of vacant/abandoned properties with the types of businesses in the city; and secondly, we are interested in measuring the association between the types of businesses in an area with the types of crime that occur in that area.

To analyze these relationships we performed Canonical Correlation Analysis, which evaluates a linear combination of variables between two data sets. In the first model, we used the vacant/abandoned buildings variables as the explanatory variables, and the business types as the response variables. The results of this model were as follows:

- In neighborhoods with a high number of vacant/abandoned properties, there is a strong correlation with children services (day care and children activities facilities), and weak or negatively correlated with all other types of businesses. This suggests two things: first, that there are not enough residents in these neighborhoods to support private businesses (or that they cannot afford to spend their disposable income on luxury items/services), and second, that the only businesses these residents will invest in are those that support their children.

In the second model, we used the business attributes as the explanatory variables and crime attributes as the response variables. The results of this configuration are as follows:

- There is a strong relationship between higher developed, more up-scale neighborhoods and white collar offenses (public indecency, intoxicating compounds & liquor law violations, and deceptive practice crimes).
- The best predictors of the overall measure of crime are manufacturing, motor vehicles services, and home-repair/construction businesses. These explanatory variables are a bit

surprising (the nighttime news is always publicizing gang violence and store robberies), but make sense to some degree; factories, junk yards and construction sites are relatively less densely populated areas, so someone out to commit a crime (regardless of the type of crime) would like to do it in an area where they will not be seen.

- The best predictors of burglary, criminal damage, and motor vehicle theft crimes are the following business types: animal care & services, home-repair/construction, and motor vehicle services. These crimes also make sense; it is conceivable that these places are vandalized and that items are stolen (cars from a lot or materials/equipment from a construction site) from these types of businesses.

Although the results agree for the most part with our intuition on how these community indicators relate to each other, we do have concerns about the statistical method used. Canonical Correlation Analysis requires the number of observations to be significantly higher than the number of attributes in each dataset; however, because we used datasets that were aggregated by zip code, the number of observations in all of the datasets dropped significantly. So we were left with a high number of attributes and a low number of observations in each dataset, which is not ideal for CCA.

Our recommendation for follow up work is to include more cities in the dataset. This would not only raise the number of observations desired for CCA, but could also lead to results that generalize to other large cities (as it stands, the results only apply to Chicago, since we cannot guarantee that other cities have similar neighborhood characteristics).

Introduction

Our group, the Angry Chickens, is very concerned about crime in the City of Chicago, and we would like to identify key factors that are correlated with its increase, decrease or shifts thereby helping to predict areas of concern and possibly mitigating it. Our group has identified five community indicators (building permits, business licenses, vacant and abandoned buildings, city owned land, and buildings violations) that we would like to investigate in relation to crime. This report, in particular, will focus on the vacant and abandoned buildings, business licenses, and crime data sets. Our hope is that our research will help identify meaningful relationships between the indicators that local government can use to improve the welfare of the city and its citizens.

The Data

Although all of the data came from the City of Chicago data portal, the different datasets were provided by a variety of sources within the city: the crime data is supplied by the Chicago Police Department; the business license data is provided by the Department of Business Affairs and Consumer Protection; and the vacant/abandoned buildings data is provided by the city's 311City Services. One of the advantages of having the data provided by the same portal is that, for the most part, all of the data had similar formatting and were available in the same file types. However, because the data is not supplied by the same department of the city, there were a few issues we had to deal with in our attempt to clean and integrate the different data sets.

First, the original data sets spanned different time periods; the crime data set went as far back as 2001 while the business license data set only went back to 2006. As such, our group decided, that for consistency across the different data sets, we would only include observations from 2008 until the end of 2014. Secondly, in order to perform an analysis on the three data sets, we had to find a way to merge the data sets, and though this seems trivial, it proved to be a difficult task. After concluding that it would be best to integrate the data on location (by neighborhood or zip codes, for example); we found that not all of the data sets provided a consistent location value. At best, the only consistent location value came in the form of latitude and longitude coordinates. Here I am grateful to fellow team member, Susan Cavanaugh, for finding and providing a list of latitude/longitude coordinates and how they mapped to their respective zip codes. Through this we were able to convert the coordinates into a tangible locations. Next the question that arose was this: if the data already provided a zip code, do we use the one provided or do we transform the coordinates provided and use the zip code from the mapping? And again we decided that our results would be more consistent if we used the mapping zip codes instead of the ones provided because in this way all of the data was transformed in the same manner. With these issues out of the way, we return our attention to the data.

The data, in its raw form, was quite substantial: the vacant/abandoned buildings data set contained 15 attributes and over 48,000 observations, the building license data set contained over 30 attributes and 400,000 records, and the original crime data set contained over 22 variables and over 5 million records. While inspecting the data, we had to decide what to do with missing coordinate values; and to a large extent, our hand was forced; we could not *manually* substitute a zip code value from an address (that did not include a zip code) or a neighborhood name (like 'Lincoln Park') for the thousands of records, so we decided that we would ignore these observations in our analysis. In the end, after accounting for the missing values and adjusting for the time spans our data sets had 42 thousand, 350 thousand, and 2.5 million records in the vacant buildings, business licenses, and crime data sets respectively.

Next, in order to integrate the data sets, we transformed the data by aggregating by zip codes. The downside to this is that the number of observations would be significantly reduced from tens of thousands or millions (depending on the data set) to only 73 observations in each data set. Conversely, the upside is that our data now has the same amount of observations in each data set, and thus in a form acceptable to perform canonical correlation analysis on (we shall elaborate on this later). Unfortunately, by performing this transformation again we create another problem for ourselves. When we aggregated by zip codes, we did so by vacant building properties, business license types, and crime types; thus, we ended up with 8 vacant building attributes, 31 crime types, and over 100 business license types; this last number, the amount of business types, is a problem because through we only have 73 observations, and we need the number of attributes to be fewer than the number of observations. Therefore, we needed to find a new way to group the business license data. Fortunately, the City of Chicago website groups business licenses by business sector. For example, day care services and after-school programs fall under Children Services business sector. Using the city's business sector groupings we were able to reduce the number of attributes from over 100 to 19.

All in all, the preprocessing step of cleaning, integrating, and transforming the data was an arduous and time-consuming task, but ultimately, also very satisfying. Next, we are pleased to discuss the final form of our data sets. To reiterate, all of the data sets contain the same number of records 73 but vary in attributes. The vacant/abandoned buildings data set contains the following 8 attributes:

1. Occupied – meaning that the building has residents still living in building.
2. Unknown Occupation Status – the residency of the building is unknown
3. Unoccupied – no people reside in the building.

4. Boarded – the building is boarded shut.
5. Unknown Boarding Status
6. Not Boarded
7. People Living In – again, this is different than occupied because here we specify residents that are not living in the building legally (so for example, a homeless person using the building as shelter).
8. No People Living In

Next, we describe the final 19 attributes of the business license data set:

1. Accommodations – hotels, single-room occupancy, bed & breakfasts, and vacation rentals
2. Animal Care and Services – animal boarding, animal grooming, pet shops, veterinarians, etc.
3. Children Services – children's activity facility and day care services
4. Commercial and Business – day labor, dry cleaners, health club, laundry service, firearms dealer, etc. (most business fall under this term, if they did not fall under a different sector)
5. Entertainment – special events (indoor, raffles, or fund-raisers), dance clubs, music halls
6. Food – restaurants, grocery stores, delis, cafés, shared kitchens, etc. Also includes food trucks and ice-cream trucks.
7. Home-Bases Business
8. Home Repair and Construction – board up services, explosives, general contractors, and home repair, hardware stores.
9. Hospital and Commercial Care – adult care centers, adult family care homes, assisted living establishments, hospitals, and long term care facilities.
10. Liquor – liquor stores or retail sales, taverns, and restaurants that serve liquor.

11. Manufacturing – factories and manufacturers (mechanical or chemical).
12. Motor Vehicle Services – gas stations, motor vehicle repair, tire facilities, and junk yards.
13. Other – pharmaceutical or scientific laboratories
14. Outdoor Vending – street performers, street peddlers, seasonal trade shows (indoor or outdoor), and Navy Pier Kiosks
15. Parking – public garages or accessory garage (an accessory garage is a parking lot designed for store access)
16. Pawnshop and Second-hand – pawn shops, thrift stores, vintage clothing stores, etc.
17. Personal Services – massage establishment, nail salons, spas, and tattoo/body piercing shops. Strangely enough also includes tax preparation services and immigration services.
18. Public Vehicles – taxis, public chauffeurs, charter buses, and dispatch services.
19. Tobacco – stores that sell cigarettes and business to business sales.

Finally, we look at the 30 attributes (the crime types) of the crime data set. This time, because of the large number of variables and because most are self-explanatory, we will only explain the crime types that are most vague or unclear.

1. Arson
2. Assault
3. Battery
4. Burglary
5. Criminal Sexual Assault
6. Criminal Damages
7. Criminal Trespass
8. Deceptive Practice
9. Gambling
10. Homicide
11. Human Trafficking
12. Interference with Public Officer
13. Intimidation
14. Kidnapping

15. Liquor Law Violation
16. Motor Vehicle Theft
17. Narcotics
18. Non-Criminal Offense
19. Obscenity
20. Offense Involving Children
21. Other Narcotics Violation
22. Other Offense
23. Prostitution
24. Public Indecency
25. Public Peace Violation
26. Robbery
27. Sex Offense
28. Stalking
29. Theft
30. Weapons Violation

For clarification, let us explain the Non-Criminal Offense, Other-Narcotics Violation, and Other Offenses crime types. A non-criminal offense is crime like a license violation of some kind (selling items on the street without a permit, for example). An Other-Narcotics Violation involves intoxicating compounds that are not drugs (selling or using chemicals for the purpose of intoxication, for example). Lastly, Other Offenses are the crime types that did not fit the other crime categories (example include: parole violations, restraining order violations, and failure to register by a sex offender). Note: we did consider grouping crimes in a similar fashion as we did with the business licensees into business sectors; however, this time, since the City of Chicago's does not provide explicit grouping like it did with businesses, the groups would have to be determined by us. So crime type #30 (weapons violations), was our attempt at this. It is actually a combination of weapons violations and concealed carry violations (which we thought were similar enough that we could put them into a group). However, we had a difficulty trying to group the other crimes together, so we chose not to group the rest of the variables so as to not introduce too much subjectivity, on our behalf, into the data.

Exploratory Data Analysis

Before moving into the breadth of the analysis, we would like to point out a couple of notable points from our data exploration phase. First, the majority of the data is not normally distributed. This is not a problem because for the analysis that we will conduct (CCA) it is not a necessary condition since we will be using it descriptively. Second, there are values that could be considered outliers (for example, there is a particular zip code that contains a disproportionate amount of hotels, so it appears as an outlier in the column). We thought about changing possibly replacing these values; however, it seemed more reasonable to leave the values as they existed (after all, we cannot fault an area for having certain businesses within its boundaries). However, since we left them in, it is possible that these values, which look like mathematical outliers, can greatly influence the results. Lastly, we checked for correlations within the data sets and found several variables related with each other (these correlation matrices can be found in Appendices 3.1, 3.2, and 3.3). This is concerning because we do not want our data to contain a high degree of multicollinearity because this could interfere with the precise effects of each predictor.

Canonical Correlation Analysis

To identify key relationships between the three data sets we will be making use of Canonical Correlation Analysis. In brief, CCA attempts to identify and measure the relationship between two sets of variables by creating new variables, called variates, and maximizing the correlation between the variate pairs. The variates themselves are created through a linear combination of the variables within each data set such that they best explain the variability within and between both data sets. The correlations between the variate pairs and the loadings (the correlations) of the variables to their respective variates are going to explain the majority of the relationships between the two data sets

Although this method sounds ideal for our needs, there are a couple of concerns. In general, and similarly to other statistical methods, CCA requires the data sets to have a high number of observations and a low number of attributes. However, in our data sets, because we aggregated the data by zip code, we ended up with relatively very few observations compared to the amount of attributes. Because of this, we have reservations about making inferences from the model, and instead we opt to use the results in a more descriptive manner; the relationships found from the model are specific to the City of Chicago and thus the results would not generalize well to other large cities.

CCA Models

Now we will delve into the fine points of the results of both models. Although there is a fair amount of technical terminology, none of it should not be considered an obstacle as all terms and technical details are described in a straightforward manner. We begin by examining the results of the vacant/abandoned buildings data against the business data.

Model 1: Vacant/Abandoned Buildings vs. Business Types

In general, the number of variates determined by CCA is equal to the number of variables in the smaller data set. In this case, since the vacant/abandoned buildings data set has fewer variables compared to the business data set, the CCA model will create 8 variate pairs. This is not to say that all 8 variate pairs will be of interest. Canonical Correlation Analysis attempts to find the best model by maximizing the correlation between variate pairs and by keeping the variate pairs uncorrelated with each other. However, it is possible that, while finding the most optimal correlations between variates, the linear combinations of the variables found could have been due to chance and not from meaningful correlations. Using Wilk's Lambda Test we can distinguish the variate correlations that are significant from those that are not (Figure 1.1 below.)

```
Test of H0: The canonical correlations in the
current row and all that follow are zero

     CanR  WilksL      F df1    df2 p.value
1 0.94678 0.00383 2.6108 152 355.14 0.00000
2 0.82373 0.03701 1.6501 126 318.91 0.00024
3 0.75562 0.11514 1.2687 102 280.65 0.06620
4 0.65902 0.26835 0.9435  80 240.21 0.61248
5 0.51941 0.47439 0.6926  60 197.40 0.95175
6 0.47304 0.64966 0.5666  42 152.06 0.98323
7 0.31788 0.83694 0.3723  26 104.00 0.99742
8 0.26266 0.93101 0.3273  12  53.00 0.98084
```

Figure 1.1

As seen in the R output above, only the first two variate pairs are statistically significant at the .05 level. However, just because certain variate pairs are significant does not necessarily mean that the canonical correlations will yield interesting insights. To further clarify which relationships will be of interest, the R software also provides a Scree plot (Figure 1.2 below), which depicts the relative importance of each variate pair. In this case, by locating the knee bend in the plot, we can see that only the first variate pair is likely to provide meaningful insight.

```
    CanR  CanRSQ  Eigen percent    cum                          scree
1 0.9468 0.89640 8.6521 63.1279  63.13 ******************************
2 0.8237 0.67854 2.1108 15.4009  78.53 *******
3 0.7556 0.57096 1.3308  9.7096  88.24 *****
4 0.6590 0.43431 0.7678  5.6018  93.84 ***
5 0.5194 0.26979 0.3695  2.6957  96.54 *
6 0.4730 0.22377 0.2883  2.1033  98.64 *
7 0.3179 0.10105 0.1124  0.8201  99.46
8 0.2627 0.06899 0.0741  0.5407 100.00
```

Figure 1.2

Although, these software tools are very helpful, we will have to examine the loadings, cross-loadings, and variable to variable relationships to find the important relationships that exists between vacant/abandoned buildings variables and business types.

First, note from Figure 1.2 that the correlation between the first variate pair is 0.9468. This tells us that there is a very strong positive linear relationship between the variate pair. Next, we examine the loadings structure of the vacant/abandoned buildings variables with their respective variate.

Loadings

```
$X.xscores
                 Xcan1
Occupied.VB  0.9259351
Unknown.VB   0.5188575
Vacant       0.9509973
VC.Boarded   0.9610463
Unknown      0.8724605
VB.Open      0.9472056
People.False 0.9266584
People.True  0.9620736
```

Figure 1.3

From Figure 1.3, it appears that all of the variables, with the exception of *Unknown.VB*, have a very high correlation with the variate. Thus, the first variate can be thought of as measuring the overall level of vacant/abandoned buildings. On the other hand, looking at the loadings for the business data set (Figure 1.4), we can see that *Children.Services* has the strongest correlation with the business variate. So the business variate is characterized primarily by day care businesses and children's activities facilities.

```
$Y.yscores
                                       Ycan1
accomodations                   -0.002005227
animal.care.and.services        -0.012100144
children.services                0.767298883
commercial.and.business          0.162043003
entertainment                   -0.147485521
food                             0.225455459
home.based.business              0.291797402
home.repair.and.construction     0.399334957
hospital.and.commercial.care     0.167242145
liquor                          -0.086200704
manufacturing                    0.414888562
motor.vehicle                    0.480805980
other                            0.384366905
outdoor.vending                  0.494999399
parking                         -0.180666475
pawnshop.and.secondhand          0.278513091
personal.services               -0.225220029
public.vehicles                 -0.175555140
tobacco                          0.587207688
```

Figure 1.4

Cross-Loadings

Examining the cross-loadings of the vacant buildings variables with the business variate (the software output can be found in Appendix 1.1), we see that all of the variables, with the exception of *Unknown.VB*, have a strong correlation with the business variate. Similarly, the cross-loadings of the business variables with the vacant buildings variate (Appendix 1.2), show that *Children.Services* is mostly correlated with the vacant buildings variates. Therefore, we can think of these findings as reinforcing the loading results we already discussed above (this is the reason why the output was placed in the appendix—it did not highlight or contribute any different information that was not mentioned already).

Variable-to-Variable Relationships

Now that we have examined the loading structures of the variables to the variates, we can paint the whole picture of how the variables relate to one another: the overall measure of vacant/abandoned buildings is associated with kid-centric businesses; If there are a lot of abandoned buildings in a neighborhood, the only shops that thrive are day care services and children activities facilities because parents need these places to keep their children safe while they work (this type of business is seen as a necessity). Additionally, the overall measure of vacant buildings has a very low correlation with businesses that are seen as luxuries (hotels, animal grooming, clubs, etc.). Hence, it appears that if there a lot of empty homes in a zip code, entrepreneurs and would-be business owners do not open a lot of shops in this neighborhood since residents do not appear to have much disposable income (as evidenced by residents' losing their homes).

Model 2: Business Types vs. Crime Types

We just examined how abandoned buildings characteristics, as explanatory variables, affected the business types, as the response variables, of a neighborhood. For the second model, we will be using the business attributes as the explanatory variables and the crime attributes will become the response variables; essentially, we will be trying to use the types of businesses in an area to predict the types of crimes that occur in that area.

For this CCA model, the R software will create 19 variate pairs, and we begin our analysis by inspecting the following software output: the Wilk's Lambda Test (Figure 2.1), the scree plot (Figure 2.2), and the canonical correlations between the variates (can be seen in both Figure 2.1 and Figure 2.2).

```
Test of H0: The canonical correlations in the
current row and all that follow are zero

      CanR  WilksL      F df1    df2 p.value
1  0.99714 0.00000 6.0221 570 471.92 0.00000
2  0.99622 0.00000 4.9001 522 460.34 0.00000
3  0.98957 0.00000 3.9224 476 447.57 0.00000
4  0.97638 0.00000 3.2900 432 433.58 0.00000
5  0.96912 0.00000 2.8802 390 418.35 0.00000
6  0.95710 0.00000 2.5085 350 401.86 0.00000
7  0.94747 0.00001 2.1878 312 384.05 0.00000
8  0.92041 0.00008 1.8769 276 364.89 0.00000
9  0.89098 0.00052 1.6348 242 344.33 0.00001
10 0.86550 0.00253 1.4320 210 322.30 0.00188
11 0.81972 0.01010 1.2373 180 298.74 0.05284
12 0.78241 0.03077 1.0781 152 273.56 0.29452
13 0.71261 0.07935 0.9217 126 246.67 0.69373
14 0.67136 0.16122 0.8062 102 217.94 0.89067
15 0.58862 0.29352 0.6788  80 187.25 0.97531
16 0.56700 0.44913 0.5859  60 154.46 0.99044
17 0.45062 0.66193 0.4243  42 119.42 0.99897
18 0.37969 0.83059 0.3067  26  82.00 0.99941
19 0.17173 0.97051 0.1064  12  42.00 0.99991
```

Figure 2.1

	CanR	CanRSQ	Eigen	percent	cum	scree
1	0.9971	0.99429	174.02992	40.688679	40.69	******************************
2	0.9962	0.99246	131.59954	30.768340	71.46	***********************
3	0.9896	0.97926	47.21022	11.037880	82.49	********
4	0.9764	0.95332	20.42254	4.774846	87.27	****
5	0.9691	0.93920	15.44730	3.611621	90.88	***
6	0.9571	0.91604	10.91082	2.550980	93.43	**
7	0.9475	0.89770	8.77524	2.051675	95.48	**
8	0.9204	0.84715	5.54250	1.295852	96.78	*
9	0.8910	0.79385	3.85082	0.900333	97.68	*
10	0.8655	0.74908	2.98538	0.697991	98.38	*
11	0.8197	0.67194	2.04823	0.478882	98.86	
12	0.7824	0.61217	1.57844	0.369044	99.23	
13	0.7126	0.50781	1.03176	0.241227	99.47	
14	0.6714	0.45073	0.82060	0.191859	99.66	
15	0.5886	0.34647	0.53015	0.123951	99.78	
16	0.5670	0.32149	0.47381	0.110778	99.89	
17	0.4506	0.20306	0.25480	0.059572	99.95	
18	0.3797	0.14417	0.16845	0.039384	99.99	
19	0.1717	0.02949	0.03039	0.007105	100.00	

Figure 2.2

Inspecting the p-values for the different variate pairs in Figure 2.1, the Wilk's Tests tell us that the first 10 variate pairs are statistically significant at the 0.05 level. But, to reiterate, it is unlikely that all 10 will reveal meaningful insights between the business and crime variables (since the model is makes linear combinations of the variables to maximize the correlation between variate, therefore some of the combinations created could be of no practical use). It is more likely that, as the scree plot hints, only the first three variate pairs will reveal important relationships between the two data sets. Next, by examining the canonical correlations, we can see that the first 11 variate pairs have very strong positive linear relationships (correlations above 0.80), but ultimately, only the first three variate pairs were of interest (*the reader may wonder how we determined which variates were of interest, and though it is true the scree plot influenced our choice, we actually examined the loading, cross-loadings, and variable-to variable relationships of the first 6 variate pairs. What we found is that after the third variate pair, the relationships between the business and crime variables did not make much sense. For example, the fourth variate pair made a link between hospitals in a neighborhood and gambling and narcotics crimes, which did not sound reasonable. CCA finds linear combinations of variables that are sound mathematically, but sometimes the results are not practical.). We begin by examining the loadings of these variates.

Loadings

We start by looking at the business variables with their respective variates (Figure 2.3). The variables that have the highest correlation with the first variate, Xcan1, are: liquor, food, entertainment, and home-based businesses. As such we can characterize the first variate, for the most part, as the type of businesses that people visit on a night out (examples: couples going to a restaurant for dinner, or people going out to the bar, or people going to a dance club, etc.). For the second variate, Xcan2, the business attributes most highly correlated with this variate are: motor vehicle services and manufacturing. From these business types, we can describe the second variate as the blue collar businesses. Lastly, the third variate, Xcan3, has the highest positive correlation with home-repair and construction businesses, so this variate can be thought of as representing the handy-man businesses.

```
$X.xscores
                                    Xcan1       Xcan2        Xcan3
accomodations                  -0.7534910 -0.16741553 -0.247126504
animal.care.and.services       -0.6900382  0.13687050  0.376861280
children.services              -0.5017011  0.54741824  0.249306946
commercial.and.business        -0.7623596  0.24655583 -0.058317539
entertainment                  -0.8245262 -0.23507488 -0.080247047
food                           -0.8414951  0.37088524  0.011970686
home.based.business            -0.8225643  0.30366159  0.291537039
home.repair.and.construction   -0.3676890  0.61916777  0.478903497
hospital.and.commercial.care   -0.4388109 -0.02397816 -0.008999437
liquor                         -0.8894642 -0.12618009  0.128951537
manufacturing                  -0.2417707  0.74679156  0.109998720
motor.vehicle                  -0.3115050  0.79634325  0.369120847
other                          -0.2767391  0.37916769 -0.124207738
outdoor.vending                -0.7345055  0.52867074 -0.048043708
parking                        -0.5723821 -0.34089618 -0.181348167
pawnshop.and.secondhand        -0.6505687  0.50188996 -0.050297124
personal.services              -0.7530293 -0.20947713 -0.013201386
public.vehicles                -0.5548762  0.07143721 -0.079090837
tobacco                        -0.7512072  0.55164863  0.110027910
```

Figure 2.3

Next, we examine the loadings of the crime attributes with their variates. The variables most correlated with the first variate, Ycan1, are: public indecency, theft, and deceptive practice. From looking at these crime types, it is difficult to label the variate in a general term, but it will make sense when we discuss the Variable-to-Variable relationships. The second variate, Ycan2, is most correlated with Arson, but it is also highly correlated with 10 other crime types (r = 0.70 and above). So the second variate, since it is correlated with at least half of the crime types, can be thought of as an overall measure of crime. The third variate, Ycan3, is positively correlated with burglary and criminal trespassing crimes, so this variate can be thought of as the variate that captures the crime of breaking and entering.

$Y.yscores

	Ycan1	Ycan2	Ycan3
ASSAULT	-0.25492074	0.6975199	0.06341050
BATTERY	-0.25425035	0.6862428	0.08329644
BURGLARY	-0.35305657	0.6756439	0.25067731
CRIM.SEXUAL.ASSAUL	-0.31517099	0.6618724	0.07356538
CRIMINAL.DAMAGE	-0.36996371	0.7238029	0.20963542
CRIMINAL.TRESPASS	-0.43809125	0.3768500	-0.01655332
DECEPTIVE.PRACTICE	-0.73356122	0.3162475	0.06045651
GAMBLING	0.01779643	0.4414092	-0.07912483
HOMICIDE	-0.10438877	0.6893391	0.01026506
HUMAN.TRAFFICKING	-0.27166536	-0.1191709	0.02108511
INTERFERENCE.WITH	-0.15464963	0.6661116	0.07404394
INTIMIDATION	-0.37567633	0.7431842	-0.03738733
KIDNAPPING	-0.19232270	0.7442549	0.19432740
LIQUOR.LAW.VIOLATION	-0.61772128	0.6010531	0.19286023
MOTOR.VEHICLE.THEFT	-0.37772393	0.7777226	0.22265975
NARCOTICS	-0.10002741	0.5609593	-0.01581233
NON.CRIMINAL	-0.38906042	0.2588171	-0.30113322
OBSCENITY	-0.30255631	0.6609005	0.06394734
OFFENSE.INVOLVING	-0.23566836	0.7465088	0.14467513
OTHER.NARCOTIC.VIOLATION	-0.69435092	-0.1039164	0.04423646
OTHER.OFFENSE	-0.29611055	0.7122716	0.13624707
PROSTITUTION	-0.19787657	0.5100665	-0.05002122
PUBLIC.INDECENCY	-0.76649200	0.1318668	-0.13468568
PUBLIC.PEACE.VIOLATION	-0.26974631	0.7771608	-0.03696115
ROBBERY	-0.28689533	0.6276798	0.08601013
SEX.OFFENSE	-0.50121591	0.7093327	0.09652961
STALKING	-0.49668877	0.5941305	0.08765407
THEFT	-0.75380303	0.4612043	0.08253757
Weapons..Violation..Concealed.Carry.	-0.11235205	0.7052801	0.05205766

Figure 2.4

Cross-Loadings

Similar to the cross-loadings of the first model, the cross-loadings of this model reinforce the results of the loadings above (Appendices 2.1 and 2.2). So, for example, the business variables highly correlated with the crime variates are: liquor, food, entertainment, and home based business. Therefore, since no new insights are added, we will move on the variable relationships.

Variable-to-Variable Relationships

As before, now that we have examined the loading structures, we can try to understand how the business types of a neighborhood can predict the types of crime that occur in that neighborhood. Let us first discuss the variables relationships between the first variate pair. It appears that businesses people visit on a night out (liquor stores, bars, restaurants, and night clubs) are great predictors of public indecency, theft, and deceptive practice crimes. This seems logical—after a night out of drinking it seems plausible that people might engage in thoughtless behavior leading to a public indecency charge. Additionally, since these places are frequented by people with money to spend, it is also plausible that theft crimes (like purse snatching or pocket picking) are likely to occur. Lastly, these business types (which includes home-based businesses) are highly predictive of deceptive practice crimes (examples include: credit card fraud, theft of labor or services, or writing a bad check), which also sounds reasonable.

Moving on to the second variate pair, we can see that motor vehicle services and manufacturing businesses are great predictors of the overall level of crime. This seemed puzzling at first, until we realized that these businesses (junk yards and factories) usually lie together same areas of a neighborhood (the meat packing district, for example), and these are places sparsely populated. Hence, it is not unreasonable to think that criminals might want to hide their shady behavior in places like this, where they are not in the line of sight of most residents.

Lastly, we characterized the third variate of the business dataset as the handy-man businesses (construction businesses and hardware stores), and they are highly predictive of burglary and criminal trespass crimes. This also makes some degree of sense—construction sites, in particular, seem inviting to younger criminals who are probably looking more to have some fun (perhaps vandalizing the site with graffiti) and these places do not usually have security in place until after the construction is complete.

Conclusion and Recommendations

Our group, the Angry Chickens, set out to find the relationship between different community indicators and crime, of which this report focused solely on a subset of those data sets, and we were pleased with results. The CCA models yielded reasonable relationships: overall measure of vacant buildings is predictive of low business in that neighborhood (and the only business type that thrives are day care centers and children activities facilities); businesses that people frequent on a night out are predictive of public indecency, theft, and deceptive practice crimes; the blue collar businesses (those that are in sparsely populated areas) are predictive of the overall general level of crime; and home-repair and construction businesses are predictive of breaking and entering crimes. With that said, we do believe that improvements can be made to gain more knowledge from the data.

We previously mentioned that the number of observations in the data was inadequate for the number of attributes in the data (ideally, we would like to have 20 observations per variable in the data). To correct this, we propose two ideas: One, increase the sample size by adding more data from other cities, or two, decrease the number of attributes. This second option is trickier because either domain expertise would need to further segment the data sets into smaller groups or some feature extraction technique could be used to lower the number of attributes. Either way, these simple enhancements could lead to improved results.

Appendix 1.1 Cross-Loadings: Vacant/Abandoned Buildings variable correlations with Business variate

```
$X.yscores
                 Ycan1
Occupied.VB  0.8766584
Unknown.VB   0.4912448
Vacant       0.9003869
VC.Boarded   0.9099011
Unknown      0.8260296
VB.Open      0.8967969
People.False 0.8773433
People.True  0.9108737
```

Appendix 1.2 Cross-Loadings: Business variable correlations with Vacant Buildings variate

```
$Y.xscores
                                   Xcan1
accomodations                -0.001898512
animal.care.and.services     -0.011456195
children.services             0.726464568
commercial.and.business       0.153419356
entertainment                -0.139636598
food                          0.213457111
home.based.business           0.276268451
home.repair.and.construction  0.378083044
hospital.and.commercial.care  0.158341808
liquor                       -0.081613252
manufacturing                 0.392808912
motor.vehicle                 0.455218320
other                         0.363911565
outdoor.vending               0.468656390
parking                      -0.171051718
pawnshop.and.secondhand       0.263691108
personal.services            -0.213234210
public.vehicles              -0.166212400
tobacco                       0.555957514
```

Appendix 2.1 Cross-Loadings: Vacant/Abandoned Buildings variable correlations with Business variate

```
$X.yscores
                                    Ycan1       Ycan2        Ycan3
accomodations                  -0.7513354 -0.16678306 -0.244550064
animal.care.and.services       -0.6880642  0.13635342  0.372932278
children.services              -0.5002658  0.54535015  0.246707774
commercial.and.business        -0.7601787  0.24562437 -0.057709544
entertainment                  -0.8221674 -0.23418679 -0.079410424
food                           -0.8390878  0.36948408  0.011845885
home.based.business            -0.8202111  0.30251439  0.288497593
home.repair.and.construction -0.3666372  0.61682863  0.473910645
hospital.and.commercial.care -0.4375555 -0.02388757 -0.008905612
liquor                         -0.8869197 -0.12570340  0.127607141
manufacturing                  -0.2410790  0.74397027  0.108851919
motor.vehicle                  -0.3106139  0.79333475  0.365272544
other                          -0.2759474  0.37773524 -0.122912799
outdoor.vending                -0.7324043  0.52667349 -0.047542824
parking                        -0.5707447 -0.33960832 -0.179457505
pawnshop.and.secondhand        -0.6487076  0.49999387 -0.049772746
personal.services              -0.7508750 -0.20868575 -0.013063753
public.vehicles                -0.5532888  0.07116732 -0.078266269
tobacco                        -0.7490582  0.54956456  0.108880804
```

Appendix 2.2 Cross-Loadings: Business variable correlations with Vacant Buildings variate

```
$Y.xscores
                                Xcan1      Xcan2       Xcan3
ARSON                     -0.22552176  0.8513642  0.17617596
ASSAULT                   -0.25419148  0.6948847  0.06274941
BATTERY                   -0.25352300  0.6836502  0.08242802
BURGLARY                  -0.35204657  0.6730913  0.24806385
CRIM.SEXUAL.ASSAUL        -0.31426937  0.6593720  0.07279842
CRIMINAL.DAMAGE           -0.36890534  0.7210684  0.20744985
CRIMINAL.TRESPASS         -0.43683798  0.3754263 -0.01638075
DECEPTIVE.PRACTICE        -0.73146269  0.3150527  0.05982622
GAMBLING                   0.01774552  0.4397416 -0.07829990
HOMICIDE                  -0.10409014  0.6867348  0.01015804
HUMAN.TRAFFICKING         -0.27088819 -0.1187207  0.02086528
INTERFERENCE.WITH         -0.15420722  0.6635951  0.07327199
INTIMIDATION              -0.37460161  0.7403765 -0.03699754
KIDNAPPING                -0.19177251  0.7414432  0.19230142
LIQUOR.LAW.VIOLATION      -0.61595414  0.5987824  0.19084955
MOTOR.VEHICLE.THEFT       -0.37664336  0.7747845  0.22033839
NARCOTICS                 -0.09974126  0.5588400 -0.01564748
NON.CRIMINAL              -0.38794742  0.2578393 -0.29799373
OBSCENITY                 -0.30169077  0.6584037  0.06328066
OFFENSE.INVOLVING         -0.23499417  0.7436885  0.14316681
OTHER.NARCOTIC.VIOLATION  -0.69236455 -0.1035238  0.04377527
OTHER.OFFENSE             -0.29526345  0.7095807  0.13482661
PROSTITUTION              -0.19731049  0.5081395 -0.04949972
```

Appendix 3.1 – Correlations between Vacant/Abandoned Buildings Variables

	Occupied.VB	Unknown.VB	Vacant	VC.Boarded	Unknown	VB.Open	People.False	People.True
Occupied.VB	1	0.399988	0.966578	0.956229	0.828037	0.966514	0.960727	0.959415
Unknown.VB	0.399988	1	0.344412	0.435738	0.500522	0.331916	0.358422	0.326402
Vacant	0.966578	0.344412	1	0.980947	0.886781	0.999674	0.992718	0.991663
VC.Boarded	0.956229	0.435738	0.980947	1	0.912413	0.975824	0.982268	0.963823
Unknown	0.828037	0.500522	0.886781	0.912413	1	0.878895	0.897197	0.860923
VB.Open	0.966514	0.331916	0.999674	0.975824	0.878895	1	0.991191	0.992651
People.False	0.960727	0.358422	0.992718	0.982268	0.897197	0.991191	1	0.969012
People.True	0.959415	0.326402	0.991663	0.963823	0.860923	0.992651	0.969012	1

Appendix 3.2 – Correlations between Business Variables

*Note that the names of the variables were changed in this table to conserve space (Otherwise, the table would take up many pages and be more difficult to read.). The variables here are in the same alphabetical order of the original data set, i.e. *b1* is *Accommodations*, *b2* is *Animal.Care.Services* and so forth.

	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
b1	1	0.321151	0.224753	0.665847	0.853864	0.585689	0.517679	0.021927	0.284735	0.802876
b2	0.321151	1	0.474981	0.474323	0.479496	0.720765	0.80724	0.605563	0.367947	0.684908
b3	0.224753	0.474981	1	0.392097	0.159546	0.546988	0.733843	0.634771	0.389985	0.28609
b4	0.665847	0.474323	0.392097	1	0.823036	0.872255	0.669565	0.461921	0.156212	0.767403
b5	0.853864	0.479496	0.159546	0.823036	1	0.727187	0.567171	0.110514	0.208464	0.925906
b6	0.585689	0.720765	0.546988	0.872255	0.727187	1	0.83672	0.576638	0.348012	0.796561
b7	0.517679	0.80724	0.733843	0.669565	0.567171	0.83672	1	0.747522	0.499916	0.705917
b8	0.021927	0.605563	0.634771	0.461921	0.110514	0.576638	0.747522	1	0.191397	0.271805
b9	0.284735	0.367947	0.389985	0.156212	0.208464	0.348012	0.499916	0.191397	1	0.243259
b10	0.802876	0.684908	0.28609	0.767403	0.925906	0.796561	0.705917	0.271805	0.243259	1
b11	0.05531	0.405919	0.421203	0.463402	0.127149	0.578437	0.504971	0.694907	0.056658	0.208591
b12	0.035594	0.559417	0.643579	0.47759	0.094386	0.63009	0.691168	0.90555	0.215741	0.246597
b13	0.14371	0.297998	0.519887	0.252737	0.14316	0.341965	0.470559	0.360479	0.326033	0.176927
b14	0.563652	0.561489	0.734068	0.671192	0.500757	0.822559	0.837079	0.605395	0.445437	0.588947
b15	0.839442	0.190731	-0.03118	0.622138	0.865684	0.443354	0.277646	-0.13729	0.011143	0.772416
b16	0.295644	0.481095	0.513385	0.718889	0.402996	0.76678	0.683295	0.641433	0.327536	0.434368
b17	0.771137	0.604488	0.164101	0.661786	0.842323	0.668601	0.612323	0.160907	0.164097	0.870621
b18	0.360481	0.452118	0.224656	0.659721	0.540375	0.618527	0.553689	0.434803	0.243912	0.505153
b19	0.460092	0.586595	0.736215	0.773473	0.519725	0.874825	0.811793	0.686499	0.416782	0.611072

Appendix 3.2 (continued) - Correlation between Business Variables

	b11	b12	b13	b14	b15	b16	b17	b18	b19
b1	0.05531	0.035594	0.14371	0.563652	0.839442	0.295644	0.771137	0.360481	0.460092
b2	0.405919	0.559417	0.297998	0.561489	0.190731	0.481095	0.604488	0.452118	0.586595
b3	0.421203	0.643579	0.519887	0.734068	-0.03118	0.513385	0.164101	0.224656	0.736215
b4	0.463402	0.47759	0.252737	0.671192	0.622138	0.718889	0.661786	0.659721	0.773473
b5	0.127149	0.094386	0.14316	0.500757	0.865684	0.402996	0.842323	0.540375	0.519725
b6	0.578437	0.63009	0.341965	0.822559	0.443354	0.76678	0.668601	0.618527	0.874825
b7	0.504971	0.691168	0.470559	0.837079	0.277646	0.683295	0.612323	0.553689	0.811793
b8	0.694907	0.90555	0.360479	0.605395	-0.13729	0.641433	0.160907	0.434803	0.686499
b9	0.056658	0.215741	0.326033	0.445437	0.011143	0.327536	0.164097	0.243912	0.416782
b10	0.208591	0.246597	0.176927	0.588947	0.772416	0.434368	0.870621	0.505153	0.611072
b11	1	0.810891	0.280656	0.532608	-0.01378	0.522637	0.147404	0.195706	0.585057
b12	0.810891	1	0.370658	0.661282	-0.13953	0.681411	0.130734	0.373443	0.743344
b13	0.280656	0.370658	1	0.482186	0.053214	0.284458	0.196068	0.231819	0.416256
b14	0.532608	0.661282	0.482186	1	0.269335	0.678542	0.535945	0.447444	0.854481
b15	-0.01378	-0.13953	0.053214	0.269335	1	0.082231	0.694665	0.277354	0.269818
b16	0.522637	0.681411	0.284458	0.678542	0.082231	1	0.332935	0.602312	0.785552
b17	0.147404	0.130734	0.196068	0.535945	0.694665	0.332935	1	0.450149	0.412612
b18	0.195706	0.373443	0.231819	0.447444	0.277354	0.602312	0.450149	1	0.508812
b19	0.585057	0.743344	0.416256	0.854481	0.269818	0.785552	0.412612	0.508812	1

Appendix 3.3 - Correlations between Crime Variables

	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
c1	1	0.90916	0.915082	0.901098	0.904491	0.938098	0.672212	0.625572	0.613977	0.870458
c2	0.90916	1	0.992249	0.931669	0.981846	0.955414	0.837953	0.720238	0.705844	0.959502
c3	0.915082	0.992249	1	0.917972	0.988159	0.947024	0.834519	0.719923	0.758562	0.968449
c4	0.901098	0.931669	0.917972	1	0.935065	0.978788	0.761048	0.694897	0.511871	0.858794
c5	0.904491	0.981846	0.988159	0.935065	1	0.950556	0.836927	0.743471	0.720573	0.947414
c6	0.938098	0.955414	0.947024	0.978788	0.950556	1	0.783637	0.731301	0.573954	0.887151
c7	0.672212	0.837953	0.834519	0.761048	0.836927	0.783637	1	0.801469	0.567416	0.751863
c8	0.625572	0.720238	0.719923	0.694897	0.743471	0.731301	0.801469	1	0.420129	0.62323
c9	0.613977	0.705844	0.758562	0.511871	0.720573	0.573954	0.567416	0.420129	1	0.787564
c10	0.870458	0.959502	0.968449	0.858794	0.947414	0.887151	0.751863	0.62323	0.787564	1
c11	-0.07275	-0.06705	-0.05416	-0.01665	-0.0266	-0.03958	-0.02801	0.052242	-0.05846	-0.07963
c12	0.877685	0.919844	0.943086	0.802509	0.920668	0.846554	0.764357	0.674466	0.849797	0.922406
c13	0.87641	0.921417	0.902619	0.863154	0.888341	0.91251	0.770695	0.73231	0.556082	0.86123
c14	0.932084	0.921864	0.911789	0.921473	0.907715	0.937396	0.710033	0.658231	0.598065	0.877722
c15	0.719899	0.586084	0.608517	0.645875	0.613638	0.716395	0.492213	0.625401	0.313911	0.511955
c16	0.951342	0.919351	0.918713	0.94322	0.924041	0.970917	0.744848	0.713228	0.550155	0.845437
c17	0.763898	0.830955	0.868862	0.663538	0.831324	0.732513	0.653127	0.577294	0.945532	0.875993
c18	0.303547	0.260083	0.243861	0.30353	0.279939	0.32709	0.209184	0.350877	0.013883	0.205489
c19	0.729424	0.639275	0.643138	0.659534	0.64732	0.721101	0.497754	0.483615	0.395489	0.605924
c20	0.935035	0.967106	0.952178	0.957792	0.945442	0.970323	0.76654	0.669481	0.613155	0.912212
c21	0.12049	0.116268	0.108803	0.184068	0.130429	0.189708	0.297108	0.549545	-0.0666	0.005312
c22	0.915258	0.967849	0.963111	0.940916	0.957776	0.960378	0.807267	0.719705	0.665355	0.908514
c23	0.682803	0.651694	0.687948	0.533856	0.659199	0.606303	0.537857	0.55858	0.605563	0.659364
c24	0.259519	0.280517	0.277476	0.306036	0.329888	0.32043	0.420161	0.713986	0.036458	0.20434
c25	0.903987	0.897256	0.901251	0.851255	0.89315	0.904508	0.719723	0.635834	0.627028	0.887346
c26	0.883174	0.970931	0.980502	0.916452	0.986281	0.925497	0.829739	0.74109	0.736485	0.951317
c27	0.914861	0.888281	0.886076	0.913939	0.896615	0.95414	0.760655	0.755511	0.514758	0.802511
c28	0.799762	0.886713	0.868231	0.880347	0.882539	0.906385	0.792372	0.806684	0.484353	0.81081
c29	0.709032	0.768024	0.761427	0.793239	0.790552	0.817047	0.808352	0.926392	0.404971	0.647518
c30	0.888367	0.970523	0.9672	0.89234	0.953937	0.905047	0.788917	0.635502	0.716945	0.972653

Appendix 3.3 (continued) - Correlations between Crime Variables

	c11	c12	c13	c14	c15	c16	c17	c18	c19	c20
c1	-0.07275	0.877685	0.87641	0.932084	0.719899	0.951342	0.763898	0.303547	0.729424	0.935035
c2	-0.06705	0.919844	0.921417	0.921864	0.586084	0.919351	0.830955	0.260083	0.639275	0.967106
c3	-0.05416	0.943086	0.902619	0.911789	0.608517	0.918713	0.868862	0.243861	0.643138	0.952178
c4	-0.01665	0.802509	0.863154	0.921473	0.645875	0.94322	0.663538	0.30353	0.659534	0.957792
c5	-0.0266	0.920668	0.888341	0.907715	0.613638	0.924041	0.831324	0.279939	0.64732	0.945442
c6	-0.03958	0.846554	0.91251	0.937396	0.716395	0.970917	0.732513	0.32709	0.721101	0.970323
c7	-0.02801	0.764357	0.770695	0.710033	0.492213	0.744848	0.653127	0.209184	0.497754	0.76654
c8	0.052242	0.674466	0.73231	0.658231	0.625401	0.713228	0.577294	0.350877	0.483615	0.669481
c9	-0.05846	0.849797	0.556082	0.598065	0.313911	0.550155	0.945532	0.013883	0.395489	0.613155
c10	-0.07963	0.922406	0.86123	0.877722	0.511955	0.845437	0.875993	0.205489	0.605924	0.912212
c11	1	-0.06663	-0.05102	-0.09426	0.017961	-0.04278	-0.05806	0.292151	-0.02428	-0.08778
c12	-0.06663	1	0.79959	0.864959	0.509442	0.83112	0.941978	0.185734	0.579129	0.866767
c13	-0.05102	0.79959	1	0.86283	0.705176	0.887059	0.713668	0.371622	0.64537	0.909186
c14	-0.09426	0.864959	0.86283	1	0.589175	0.914753	0.747768	0.296473	0.69654	0.960967
c15	0.017961	0.509442	0.705176	0.589175	1	0.758625	0.452204	0.275015	0.619896	0.623088
c16	-0.04278	0.83112	0.887059	0.914753	0.758625	1	0.7203	0.28012	0.753624	0.940698
c17	-0.05806	0.941978	0.713668	0.747768	0.452204	0.7203	1	0.110933	0.512549	0.763166
c18	0.292151	0.185734	0.371622	0.296473	0.275015	0.28012	0.110933	1	0.283401	0.29109
c19	-0.02428	0.579129	0.64537	0.69654	0.619896	0.753624	0.512549	0.283401	1	0.694729
c20	-0.08778	0.866767	0.909186	0.960967	0.623088	0.940698	0.763166	0.29109	0.694729	1
c21	0.330168	0.076509	0.213356	0.087948	0.401173	0.203448	0.048088	0.209864	0.071078	0.108279
c22	-0.07052	0.88719	0.907814	0.928968	0.640873	0.937821	0.796785	0.29227	0.672676	0.972181
c23	-0.04133	0.749226	0.576086	0.635006	0.468388	0.663384	0.750687	0.137599	0.529391	0.602017
c24	0.248612	0.235388	0.386358	0.249228	0.402195	0.334933	0.150222	0.397787	0.309313	0.244784
c25	-0.03966	0.800953	0.920397	0.863396	0.746691	0.88165	0.731895	0.386567	0.725866	0.901376
c26	-0.00814	0.92653	0.859355	0.891244	0.567966	0.90802	0.846859	0.233132	0.622959	0.92747
c27	0.007089	0.782655	0.89335	0.882046	0.802001	0.949121	0.680486	0.399808	0.791293	0.912327
c28	-0.01879	0.733844	0.893664	0.807168	0.686835	0.885705	0.644376	0.309258	0.615382	0.870321
c29	0.054099	0.665378	0.785783	0.699034	0.732538	0.818912	0.551678	0.357162	0.576307	0.733324
c30	-0.06798	0.928724	0.866485	0.90489	0.476123	0.856503	0.821456	0.249078	0.585306	0.926757

Appendix 3.3 (continued) - Correlations between Crime Variables

	c21	c22	c23	c24	c25	c26	c27	c28	c29	c30
c1	0.12049	0.915258	0.682803	0.259519	0.903987	0.883174	0.914861	0.799762	0.709032	0.888367
c2	0.116268	0.967849	0.651694	0.280517	0.897256	0.970931	0.888281	0.886713	0.768024	0.970523
c3	0.108803	0.963111	0.687948	0.277476	0.901251	0.980502	0.886076	0.868231	0.761427	0.9672
c4	0.184068	0.940916	0.533856	0.306036	0.851255	0.916452	0.913939	0.880347	0.793239	0.89234
c5	0.130429	0.957776	0.659199	0.329888	0.89315	0.986281	0.896615	0.882539	0.790552	0.953937
c6	0.189708	0.960378	0.606303	0.32043	0.904508	0.925497	0.95414	0.906385	0.817047	0.905047
c7	0.297108	0.807267	0.537857	0.420161	0.719723	0.829739	0.760655	0.792372	0.808352	0.788917
c8	0.549545	0.719705	0.55858	0.713986	0.635834	0.74109	0.755511	0.806684	0.926392	0.635502
c9	-0.0666	0.665355	0.605563	0.036458	0.627028	0.736485	0.514758	0.484353	0.404971	0.716945
c10	0.005312	0.908514	0.659364	0.20434	0.887346	0.951317	0.802511	0.81081	0.647518	0.972653
c11	0.330168	-0.07052	-0.04133	0.248612	-0.03966	-0.00814	0.007089	-0.01879	0.054099	-0.06798
c12	0.076509	0.88719	0.749226	0.235388	0.800953	0.92653	0.782655	0.733844	0.665378	0.928724
c13	0.213356	0.907814	0.576086	0.386358	0.920397	0.859355	0.89335	0.893664	0.785783	0.866485
c14	0.087948	0.928968	0.635006	0.249228	0.863396	0.891244	0.882046	0.807168	0.699034	0.90489
c15	0.401173	0.640873	0.468388	0.402195	0.746691	0.567966	0.802001	0.686835	0.732538	0.476123
c16	0.203448	0.937821	0.663384	0.334933	0.88165	0.90802	0.949121	0.885705	0.818912	0.856503
c17	0.048088	0.796785	0.750687	0.150222	0.731895	0.846859	0.680486	0.644376	0.551678	0.821456
c18	0.209864	0.29227	0.137599	0.397787	0.386567	0.233132	0.399808	0.309258	0.357162	0.249078
c19	0.071078	0.672676	0.529391	0.309313	0.725866	0.622959	0.791293	0.615382	0.576307	0.585306
c20	0.108279	0.972181	0.602017	0.244784	0.901376	0.92747	0.912327	0.870321	0.733324	0.926757
c21	1	0.134686	0.200684	0.510188	0.094429	0.156786	0.269677	0.303939	0.513641	0.004122
c22	0.134686	1	0.609323	0.2696	0.896732	0.944183	0.909973	0.882149	0.772181	0.914648
c23	0.200684	0.609323	1	0.27933	0.570457	0.672421	0.620346	0.481355	0.531879	0.628754
c24	0.510188	0.2696	0.27933	1	0.287043	0.316046	0.42926	0.422187	0.667097	0.231088
c25	0.094429	0.896732	0.570457	0.287043	1	0.859609	0.910649	0.814326	0.71211	0.861443
c26	0.156786	0.944183	0.672421	0.316046	0.859609	1	0.861215	0.87465	0.775044	0.947083
c27	0.269677	0.909973	0.620346	0.42926	0.910649	0.861215	1	0.862517	0.836832	0.811404
c28	0.303939	0.882149	0.481355	0.422187	0.814326	0.87465	0.862517	1	0.851022	0.808287
c29	0.513641	0.772181	0.531879	0.667097	0.71211	0.775044	0.836832	0.851022	1	0.670836
c30	0.004122	0.914648	0.628754	0.231088	0.861443	0.947083	0.811404	0.808287	0.670836	1