Using Yelp to Find Romance in the City: A Case of Restaurants in Four Cities

Sohrab Rahimi
Department of Architecture
The Pennsylvania State University
University Park, PA, USA
sur216@psu.edu

Clio Andris, Xi Liu
Department of Geography
The Pennsylvania State University
University Park, PA, USA
{clio, xiliu}@psu.edu

ABSTRACT

Romantic relationships are an understudied aspect of cities and the built environment. Yet, restaurants continue to attract couples and augment the landscape with visible signs of affection at a table for two—or more. User-generated content (UGC) of restaurant reviews from online review site Yelp (http://yelp.com) provide text on romantic keywords such as "date", "love", "boyfriend", "wife", "anniversary", "family" by geolocated restaurants. We use these to distinguish restaurants and discover features of restaurants associated with various romantic keywords. These features include restaurant ratings and location, as well as comments about the ambiance, food, service, etc.

Using data from the Yelp Dataset Challenge in U.S. cities Charlotte, NC, Las Vegas, NM, Phoenix, AZ, and Pittsburgh, PA, we employ different data mining and correlation tools as well as GIS modeling to learn more about what types of romantic relationships use which parts of the city, and how their choices of restaurants differ by relationship stage. We find that families prefer restaurants that are outside of the central business district (CBD), have good service and high-rated food, while couples—married or dating—prefer hot spots with great ambiance for nightlife. We also find that inexpensive food is not associated with romantic dates, and the quality of service also plays a secondary role to a "classy" and "cozy" atmosphere.

CCS Concepts

• Social and professional topics → User characteristics → Geographic characteristics; 500 • Social and professional topics → User characteristics → Cultural characteristics; 500 • Applied computing → Art and humanities → Architecture (buildings); 300

Keywords

romantic relationships, NLP, Yelp, crowdsourcing, restaurants

1. INTRODUCTION

Romantic relationships are increasingly "hidden" in cities. With online dating, texting and home entertainment options, such as Netflix, it can seem as though finding a couple going to the movies or enjoying a picnic is rare. As a result, affectionate relationships and romantic love dot our landscape more dimly than in the past, which may deprive our urban spaces of visible kindness, dedication and peace that was once so effortlessly on display. Despite a potential decline in couples' use of public spaces, one persistent feature that hosts romantic interpersonal relationships is restaurants. In the U.S., restaurants expenditures exceed spending in higher education, as well as purchases of computers, books, magazines, newspapers, movies and recorded music [1].

Especially at dinnertime, restaurants serve as supportive infrastructure where couples can enjoy time together, and do so not entirely privately, but as a part of city life—they are seen in public, interact with others, and contribute to the local economy [2, 3]. In short, they are a couple in public, and their presence, theoretically, brightens the life of the city. Not all couples are positive influences, and an evening out with friends or alone is still ornamental to city life. Yet, romance should be a more prominently cultivated feature of urban infrastructure and amenity use management, and supporting romance and love should be a goal of the urban planner alongside more traditional concerns [4].

What tools could a planner use for understanding romance in the built environment? If restaurants are a key infrastructure that captures romantic acts and energy, to what extent are they used for romance, which types of restaurants are preferred, what parts of the city are activated, and do these features change based on the type of relationship: dating (boyfriend/girlfriend), marriage, family (presumably a couple with children—although many families are led by single mothers)? In this experiment, we used Yelp data to learn more about the features of restaurants that most appeal to couples: is it the ambiance, the food price, the location? Do certain types of couples (e.g. married, dating, etc.) prefer restaurants in different parts the city?

We used natural language processing (NLP) as well as machine learning techniques to extract restaurant attributes from online review site Yelp (http://yelp.com) and investigated which attributes correlate with the usability of these restaurants for the romantic partners. We asked which keywords are most associated with the "couple" indicators such as dating, love, anniversary, etc. Additionally, we examined the locations of restaurants to find how different romantic partners prefer certain neighborhoods for dinning. We found that couples at early stages of romantic relationships prefer restaurants where people comment positively about their ambience, and prefer to go to concentrated downtown areas. Couples in marriage and family stages place more emphasis on food and service in their choice of restaurants. Restaurants with family and children keywords take up a wider swath of the city, perhaps due to suburban influence and ease of parking. Our findings contribute to urban planning literature, geography, hospitality and tourism as well as an examination on romance and the life cycle.

1.2. Restaurant Features and Customer Expectations

Prior studies have found that different factors affect customers' satisfaction in restaurants. Expectations from restaurants has gone beyond basic nutrition in the past decades [5] to include the dining experience, and accordingly restaurant atmosphere has surpassed food quality [6]. Service quality and interaction between the waiter and the customer affects the customers' chance of return [7]. Expense correlates with expectations for good food and service, and attracting returning customers is harder for expensive restaurants [7].

Researchers have found that decor and design factors can provide a competitive edge in attracting customers [10], comfortable atmosphere affects the perception of good service [11] and sociable and conversational customers help attract other customers to the restaurants [12]. In addition, noise-level and the presence of music in restaurants, i.e. the soundscape, can change the types of interactions—conversational, active or otherwise, between patrons, with both positive and negative outcomes [14-16]. Yelp UGC reviews allow researchers to study these factors simultaneously, and to re-evaluate what is important for patrons and relationships, as respondents can openly choose what factors highlighted for their experience.

2. DATA AND METHOD

The data for this study has been provided by Yelp Inc. online through Round 8 of the Yelp Dataset Challenge [16]. These data include information on just under 86,000 businesses in 10 cities, six of which are in the U.S., amounting to nearly 2.7 million user reviews. Of 86,000 businesses only 20,120 had at least one review. A variety of business types are included in this data such as restaurants, bars, movie theatres, parks, etc.

We first isolated food-related businesses by filtering for business "category", a list of crowdsourced tags provided for each business (e.g. fast food, sandwich, coffee & tea, bar, auto repair, etc.). After filtering, we found 897 unique business categories. Among the food-related businesses, 6139 businesses had a "restaurant" tag, followed by 1280 bars, 463 grocery stores, 357 cafés, 216 bakeries and 183 ice cream shops. 799 places had both "bar" and "restaurant" tags, so that only 481 places were a bar with no restaurant. Given the relevantly small number of other types we only focused on restaurants or bar/restaurants in this study. We used restaurants with more than 50 reviews (N = 1914). The total number of reviews amounts to 361,839 and the maximum number of reviews is 5,042 which belongs to a restaurant in Las Vegas. The average number of reviews for each restaurant is approximately 189.

Although the Yelp dataset includes some binary attributes for these businesses, (e.g. Ambience:: Romantic: False), we relied on NLP techniques to infer features for these restaurants. We "stemmed" keywords so that words such as "eating", "ate" and "eat" are all counted as "eat". Next, we removed commonly-used words by using stop-word lists (e.g. is, are, that, etc.), typical in NLP endeavors. Following, we calculated the single word frequency, focusing on the top 3000 frequent words. The term "place" was the most frequent (90,550 occurrences) and "ricotta" the least (682 occurrences). We filtered these words manually in order to keep small occurrences that were central to the romantic aspect of our research questions such as (i.e. Dating, Boyfriend/Girlfriend, Wife/Husband, Family, Romantic, Intimate, and Anniversary) (Table 1).

Important keywords for restaurant quality and romance, including "Ambience" (which also includes synonyms i.e. "environment", "atmosphere"), "Food", "Service, and "Neighborhood", were determined through bigrams. The main purpose of this step was to find the positive (e.g. "good", "great", "perfect" etc.) and negative (e.g. "bad", "sketchy", etc.) adjectives in order to investigate the impacts of these features on romantic partners' preferences (Table 2). A few interior images are provided to give the reader examples of restaurants with high values in "upscale"," family", and "ambience (+)" (Figure 1). Overall, positive adjectives occur more frequently than negative adjectives. The number of negative adjectives used to describe ambience lower than other variables, therefore, we exclude this variable from the final model.

Figure 1. Photos of restaurants with high upscale value (top row) high family value (middle row), high ambience value, "ambience +", (bottom row) contextualize the type of places that Yelp customers are rating. Images from sourced from Yelp.

Finally, we divided the word frequency (or word group in case of "ambience") by the total number of reviews for each restaurant. Our final dependent and independent variables are ratio variables ranging from 0 to 1 where 1 indicates that that word is found in all comments, and 0 in no comments. We also included the restaurant ratings (1-5) from the business attribute table in our analysis. This dataset sets the foundation for our factor analysis process.

Table 1. Frequency of word counts for different categories of restaurant variables.

Category	Variable	Word	Frequency
Romantic variables (dependent variables)	Dating	Date	3628
	Boyfriend/ Girlfriend	Boyfriend	3107
		Girlfriend	1723
	Wife/ Husband	Wife	5190
		Husband	6450
	Family	Family	7018
	Romantic	Romantic	1173

	Intimate	Intimate	812
	Anniversary	Anniversary	793
Meals and drinks	Breakfast	Breakfast	7048
	Brunch	Brunch	2146
	Lunch	Lunch	12164
	Dinner	Dinner	15781
	Liquor	Beer	24547
		Wine	13483
		Cocktail	3951
		Vodka	2036
		Martini	2030
		Liquor	1375
General features	Cheap	Cheap	5027
	Expensive	Expensive	4466
		Pricy	3059
	Busy	Busy	11970
	Stars (Yelp Rating)	-	-
Acoustics	Quiet	Quiet	1494
	Music	Music	9125
Ambience	Décor	Décor	5812
	Sport	Sport	3660
	Authentic	Authentic	2072
	Traditional	Traditional	1652
	Cozy	Cozy	1512
	Upscale	Upscale	1034
	Touristy	Touristy	964
	Classy	Classy	948
	Modern	Modern	850
	Creative	Creative	755
	Trendy	Trendy	719
	Outdoor	Patio	5923
		Outdoor	2749
	Ambience	Atmosphere	12285
		Ambience	1095
		Environment	1475

Table 2. Variables derived from bigrams. Bigrams helped to distinguish between positive and negative adjective associated with important variables (i.e. food, service, ambience and neighborhood).

Category	**Variable**	**Noun**	**Adjective (ordered by frequency)**	**Count**
Ambience	**Ambience (+)**	**Atmosphere**	nice,lovely,beautiful,great,pretty,cool,cute,amazing,awesome,impressive,interesting,fancy,excellent,wonderful,relaxing,fun	11952
		Ambience	great, nice, lovely, good, enjoyed, perfect, amazing, excellent, fantastic, beautiful, fun, wonderful,cool,relaxing,awedome,friendly,pleasant,decent,impressive,outstanding	
		Environment	great, friendly, nice, relaxing,fun,good,comfortable,lovely,chill,welcoming,cool,pleasant,amazing,peacful,unique	
	Ambience (-)	**Atmosphere**	typical, disappointing, bad, average, bad, mediocre, terrible	761
		Ambience	disappointing, bad, average	
		Environment	-	
Food	**Food (+)**	**Food**	good,great,delicious,amazing,tasty,fresh,awesome,best,decent,fantastic,wonderful,outstanding,yummy,enjoyed,healthy,superb,spectacular,phenomenal,fabulous,unqiue,fancy,exciting	28473
	Food (-)		average,mediocre,bad,terrible,horrible,bland,suck,disgusting,nasty,meh	5204
Service	**Service (+)**	**Service**	great, good, excellent, fast, quick, best, awesome, outstanding, perfect, prompt, incredible, superb	15766
	Service (-)		bad,slow,poor,rude,mediocre,terrible,horrible,worst,disappointing,average,crappy	4319

2.1 Restaurant Attributes

We used seven dependent variables (Dating, Boyfriend/Girlfriend, Wife/Husband, Family, Romantic, Intimate and Anniversary) and 29 independent variables derived from word appearance in a restaurant review. To avoid model overfitting by including all variables in the popular ordinary least squares (OLS) model, we experimented with supervised models and chose the best model with the most reasonable prediction error (i.e. MSE). In all cases, we regressed each dependent variable separately on all other predictor variables in order to find which independent variables correlated with versions of romantic indicators. We also experimented with Pearson correlation of variable co-occurrences.

We started with a least square model. We expected this model to be the least accurate, but could provide a sense of the extent of linearity in the data. We hid 20 percent of the data for validation purposes, and used 80 percent of the data for model training. Among variable selection models, we chose the best subset model with k-fold cross validation, because it has proven one of the most accurate among variable selection methods [17]. For each variable added to the model, the MSE was calculated through a 10-fold cross validation and the model with the lowest MSE was chosen.

Next, we employed the least absolute shrinkage and selection operator (LASSO) regression as this model is commonly used among model shrinkage models. Although the ridge regression and elastic net provide better prediction potentials, LASSO's interpretability aids in result analysis [18] and adds a helpful penalty term for every coefficient added to the model. We cross-validated on a grid of 10,000 lambdas (i.e. shrinkage parameter) ranging from 0.001 to 1010. We selected the best shrinkage parameter through a 10-fold cross validation [19] and then calculated the MSE by applying the best LASSO model to the validation set.

Next, we used principal component regression (PCR) and partial least squares (PLS) models. Both PCR and PLS are widely-used dimension reduction methods. The major advantage of these two models over the models from the previous steps is that they are highly efficient datasets with high co-linearity, whereas the LASSO model which provides only one correlating variable [20]. Yet, PCR and PLS are less interpretable. Between PLS and PCR, the PLS model projects the response variable to a new space while PCR only projects the predictors. We cross-validated PLS and PCR models (k=10), chose a reasonable number of components according to validation results, and calculated the MSE by applying the best models on the validation sets.

Finally, we applied regression trees and random forests on this dataset. Regression trees are robust to outliers and can deal with irrelevant inputs, but tend to be less accurate compared with other models regarding prediction [21, 22]. Random forests, on the other hand, are more accurate learning algorithms that hold the advantages of decision trees [23].

2.2 Spatial Attributes

Yelp data also includes geographical coordinates for each restaurant, which we used to describe the spatial distribution of restaurants and neighborhood preferences of romantic partners at different relationship stages and types. To this end, we conducted a normal weighted kernel density analysis [24] for Las Vegas, Phoenix, Pittsburgh, and Charlotte where there was a higher density of restaurants with more than 50 reviews, due to the large populations of these cities. We found 884 restaurants in Phoenix, 591 restaurants in Las Vegas, 152 restaurants in Charlotte, and 142 restaurants in Pittsburgh garnered more than 50 reviews. For each city, we calculated three separate weighted kernels from the restaurants' coordinates for four key romantic variables: boyfriend/girlfriend, wife/husband, and family. For the ease of visualization, we aggregated "dating" and "boyfriend/girlfriend" into one variable. We then superimposed each of four kernels via contour maps in on maps of each city to compare between differences in neighborhood preferences.

3 RESULTS

3.1.Term Correlations

The results reveal a number of correlations via Pearson correlation between keywords that are found concurrently in the same reviews (Figure 2). The nodes in the correlation network have been ordered based on the Pearson correlation values, that is the higher the absolute Pearson value, the closer the nodes and thicker the tie. This visualization reveals a cluster of highly correlated variables toward the center: "date", "romantic", "upscale", "intimate", "classy", "anniversary", "dinner", "ambience", "ambience (+)" and "liquor". "Wife/husband" is strongly correlated with "service (+)", "dinner", "anniversary", and "food (+)" and negatively associated with "cheap". "Family", on the other hand, falls farther from other romantic variables and is correlated with "food (+)", "service (+)" and "ambience (+)". "Date", "boyfriend and girlfriend" also correlate with "upscale", "romantic", "anniversary" and "dinner". This visualization shows that as romantic partners proceed to different levels of their relationship, factors related to ambience (e.g. ambience, ambience (+), classy, romantic, etc.) become less important and food and service become more important.

3.2.General Correlations

According to the lowest MSE variables used to evaluate each model, random forest, LASSO and best subset performed constantly well for all of the variables (Table 3). Accordingly, we used results from these to interpret results of our research questions. The random forest model revealed the importance of each variable for different response variables as measured by *node purity*: the difference in RSS before and after each split. Importance (node purity) of each variable as derived from the random forest for dating and family showed that ambiance (especially positive ambiance) is important for dating, but that busy and positive food and authenticity is associated with responses that include family. Dinner is the most frequently associated term for boyfriend/girlfriend and husband/wife pairs. For intimate, anniversary and romantic categories, cozy, dinner and ambiance are the top associated keywords, respectively. The findings from the random forests are consistent with our coefficients from the LASSO and best subset models.

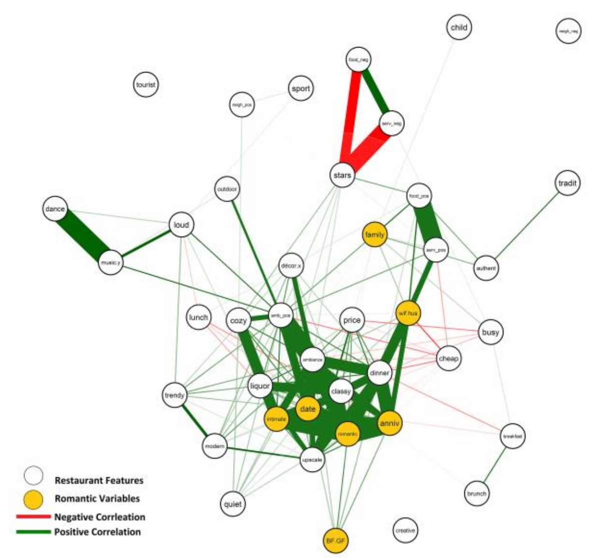

Figure 2. Pearson correlation coefficients of keywords are represented with green lines for a positive correlation between the co-presence of these keywords in one review. Red lines indicate a strong negative correlation. Line width represents the strength of the correlation.

Table 3. Mean squared error (MSE) values for seven romantic dependent variables (columns) based on different tests (rows). The average error is included on the final row. Random forest, best subset and LASSO were the most successful models.

Test	Dating	Boyfriend / Girlfriend	Wife / Husband	Family	Romantic	Intimate	Anniversary
LSR	0.00304	0.00192	0.00747	0.01370	0.00860	0.00030	0.00058
Best Subset	0.00301	0.00191	0.00735	0.01385	0.00085	0.00030	0.00058
Lasso	0.00301	0.00194	0.00730	0.01368	0.00087	0.00029	0.00059
PCR	0.00306	0.01884	0.00722	0.01367	0.00875	0.00031	0.00055
PLS	0.00304	0.00192	0.00746	0.01371	0.00086	0.00030	0.00058
Regression Tree	0.00382	0.00209	0.00898	0.01880	0.00107	0.00034	0.00086
Random Forest	0.00297	0.00195	0.00764	0.01353	0.00077	0.00026	0.00046
Average Error	0.00314	0.00436	0.00763	0.01442	0.00311	0.00030	0.00060

3.3. Restaurant Ambiance

Ambience variables indicated a spectrum of restaurant types that stratify across the range from dating to family, with and boyfriend/girlfriend and wife/husband in between. Ambience-related factors were the most significant factors for the “dating” keyword compared to all other factors including meal and drinks, general features, food, service, location and acoustics. Dating partners prefered “classy”, “trendy”, “cozy” places where Yelpers write positively about ambience. Dating was also negatively correlated with “touristy” and “sport”. On the contrary, families preferred “decorative”, traditional and authentic places. Restaurants chosen by families are negatively associated with “modern”,” classy”,” creative”,” upscale” and “ambience (+)”.

Regarding other romantic variables, “romantic” was positively correlated with “dinner”,” quiet”,” classy”,” cozy” and “ambience” and negatively correlated with “trendy” and “décor”. “Romantic”, “intimate” and “anniversary” were each correlated with “expensive” and in all three Yelpers reported positive service but no clear positive or negative reaction to food.

3.4.Spatial Patterns

Contour maps indicated that generally, restaurants preferred by families are more spread out across the cities while those preferred by married couples or boyfriend/girlfriends and dating partners are more concentrated in more crowded parts of the city. Although these results might be in part biased by the fact that most of our restaurants are located in city centers. Recall that the results from the restaurant feature analysis also showed that people comment positively about the neighborhood of the restaurants that are preferred by families. This issue might be due to the fact that families generally live in less crowded parts of the city and suburbs. Additionally, couples and dating partners are more likely to go to restaurants with positive ambience values (e.g. cozy, classy, etc.) and these restaurants can be more easily found in the downtown areas. This is evident in example of Pittsburgh (Figure 3), Las Vegas (Figure 4) and schematics of other cities (Figure 5). In these images, hotspots represent the distribution of density of all restaurants in that category, so that the sum of colored pixels equals 1.

This information is useful for understanding the different clientele in different neighborhoods—perhaps some neighborhoods have parking lots that are helpful for those who are car-dependent. This analysis can also show where emerging hot-spots are for nightlife, and can potentially attract different consumer groups, such as millennials with disposable income or new retirees who are moving to new, walkable parts of the city. It also gives planners and urbanists a new perspective on providing transportation for patrons and predicting movement.

3.5.Connections

We have also begun to survey volunteer undergraduate and graduate students on the Penn State campus (with over 40,000 students) in rural State College, Pennsylvania (pop. 80,000) through an online Google Form. Students were forwarded the survey via e-mail from department administrators and no compensation is given. All respondents report being in a relationship, a written prerequisite for completing the survey. In our first wave survey of 150 undergraduate and graduate students, respondents chose from the following variables that influence their choice of activities to do with their partners (Table 4).

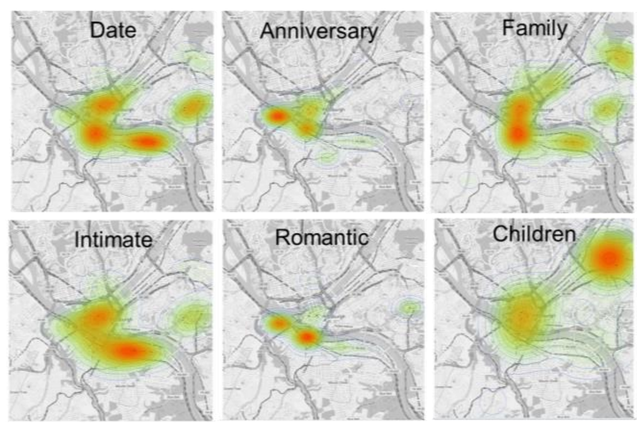

Figure 3. Hotspots of restaurants associated with each of six romantic keywords the Pittsburgh downtown region show that reviews associated with children are given for restaurants in the northeast part of the city, while intimate, date, romantic and anniversary keywords are associated with restaurants on the south shore.

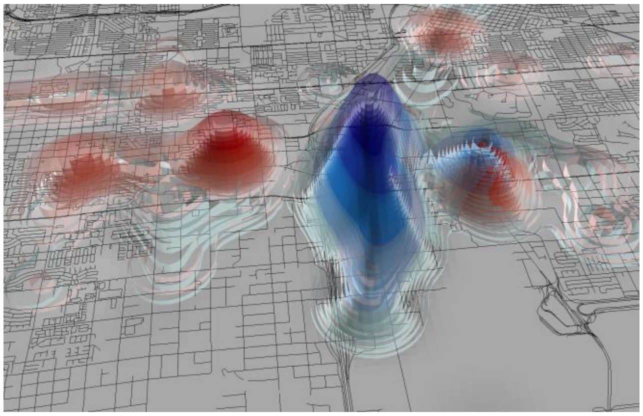

Figure 4. Differences in patron type near the Las Vegas Strip and surrounding area. Red spots indicate a clustering of restaurants with family/children keywords, and blue spots indicate clustering of restaurants with boyfriend/girlfriend/husband/wife/romantic keywords.

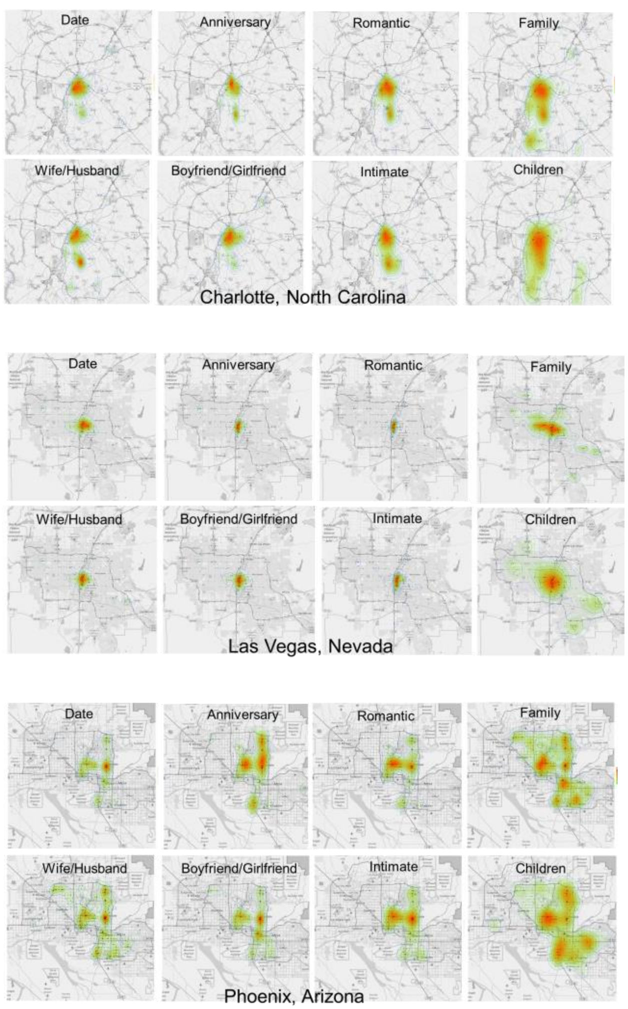

Figure 5. Hotspots of Restaurants. Small multiples of cities Charlotte (top), Las Vegas (middle) and Phoenix (bottom), show kernel hotspots of the restaurants that are associated with each of the eight romantic keywords listed at the top of each small map. In general, family and children have a larger area span.

Table 4. Features that define students' location choices for romantic outings (n = 154).

Feature of Location	**Responses**	**Percentage**
Affordable	**86**	56.6%
Cozy	**77**	50.7%
Nearby	**64**	42.1%
Fun activity (pottery, dancing, bird watching)	**51**	33.6%
Good crowd / friends	**46**	30.3%
Good service	**44**	28.9%
Theme ambiance / music	**43**	28.3%
Different setting (change of pace)	**33**	21.7%
Good hours	**23**	15.1%
Romantic	**21**	13.8%
Tradition	**17**	11.2%
Other	**13**	8.6%
Visibility--to see others and to be seen	**4**	2.6%

The responses of the student population have some overlap with the more general urban Yelp results, and we look forward to understanding more about what influences couples' use of the built environment as this survey progresses and extends to non-students, such as faculty, staff, retirees, high school students and local residents.

4. CONCLUSION AND IMPLICATIONS

We investigated factors that affect the preferences of romantic partners in their choice of restaurants, as one of the most important third places in the post-modern urban life. In this study we used Yelp restaurants comments to analyze factors related to restaurants food and service quality, ambience, affordability, and acoustics and their correlation with the choice of restaurant for romantic partners in different stages of their relationships (i.e. dating, boyfriend/girlfriend, married couples, and family). We experimented with a number of supervised learning methods including the LSR, the best subset model, the LASSO, PCR, PLS, regression trees and random forests and used the best models with the smallest MSE to interpret the final results.

We find too few studies that examine romantic relationships and the city, although there is a natural symbiotic connection between wanting to have memorable experiences with a partner and having the urban infrastructure to do so. We call for more investigations into the dynamics of romantic relationships within the context of transportation, amenities, public space and the servicescape, the latter of which is our contribution here.

In summary, our findings indicate that (a) dating partners prefer restaurants where people talk positively about the ambience and (b) qualities such as being cozy, upscale and classy are among the most significant factors for the dating partners. (c) Wives and husbands and families are more interested to the food and service quality and less concerned with the ambience-related factors. (d) Restaurants preferred by the dating partners are mostly located in dense and crowded areas where as family restaurants are more ubiquitously spread out across the city. (e) Affordability is a less important factor for families and married couples probably due to their age and consequently higher financial status. (f) Families prefer authentic and traditional places whereas the dating partners and boyfriends and girlfriends are more likely to choose trendy and modern places. Our findings are also consistent with the previous studies that also show that food and service is not a determining factors for some groups [5] and that interior design factors [10] and quietness [15] affect one's choice of restaurant.

One limitation of this study was the relatively small number of restaurants with high number of reviews (i.e. 1914 restaurants). This issue limited the geographical location of the restaurants as restaurants with high number of reviews are most likely popular ones located in certain areas. In addition, the NLP method, word frequency and bigrams counts are inherently prone to errors, which we do not discuss in depth there. For example, there is grey area between the meaning of "family", when it comes to romance, as family does not have to include a romantic couple, but could be, for example, a single parent with children, or a grandparent with younger generations in tow, or grown siblings, etc.

Our work shows some promise in the field of consumer research, design, and hospitality literature. Although our results were obtained through experimenting with multiple models and the results from different models were consistent, it is difficult to draw causal inferences from these results. Nevertheless, we suggest that future research consider how cities cater to romance and love in the built environment.

5. REFERENCES

[1] E. Schlosser, *Fast food nation: The dark side of the all-American meal*. Houghton Mifflin Harcourt, 2012.

[2] P. Wardono, H. Hibino, and S. Koyama, "Effects of restaurant interior elements on social dining behavior," *Asian J. Environ. Stud.*, vol. 2, no. 4, pp. 43–53, 2017.

[3] R. Oldenburg, *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community*. Da Capo Press, 1999.

[4] S. Zukin, "Urban Lifestyles: Diversity and Standardisation in Spaces of Consumption," *Urban Stud.*, vol. 35, no. 5–6, pp. 825–839, 1998.

[5] M. Macht, J. Meininger, and J. Roth, "The pleasures of eating: A qualitative analysis," *J. Happiness Stud.*, vol. 6, no. 2, pp. 137–160, 2005.

[6] D. Belman, "Major-league menus," *Restaur. USA Mag. Sept.*, 1996.

[7] C. Ursin, "The art of service.," *Restaur. USA Mag.*, vol. 16, no. 10, 1996.

[8] J. Teboul, *Managing quality dynamics*. Prentice Hall Direct, 1991.

[9] R. Klara, "Please please me," *Restaur. Bus.*, vol. 100, no. 4, p. 22, 2001.

[10] B. J. Pine and J. H. Gilmore, *The experience economy: work is theatre & every business a stage*. Harvard Business Press, 1999.

[11] J. Dulen, "Dazzling by design," *Restaur. Institutions*, vol. 108, no. 20, pp. 40–49, 1998.

[12] W. C. Kim and R. Mauborgne, "Knowing a winning business idea when you see one.," *Harv. Bus. Rev.*, vol. 78, no. 5, pp. 129–138, 200, 2000.

[13] S. Camp, "Café and restaurant acoustic index.," *New Zeal. Acoust. 17*, vol. 17, no. 1, pp. 34–35, 2004.

[14] L. H. Christie, "Psycho-to-building acoustics: are bars, cafes, and restaurants acceptable acoustic environments?," *Unpubl. Indep. Res. Victoria Univ. Wellingt.*, 2004.

[15] J. Rindel, "Acoustical comfort as a design criterion for dwellings in the future," in *16th Biennial Conference of the New Zealand Acoustical Society;"Sound in the Built Environment*, 2002, pp. 1–9.

[16] Yelp, "Yelp Dataset Challenge." [Online]. Available: https://www.yelp.com/dataset_challenge. [Accessed: 26-Apr-2017].

[17] S. Olejnik, J. Mills, and H. Keselman, "Using Wherry's adjusted R 2 and Mallow's Cp for model selection from all possible regressions," *J. Exp. Educ.*, vol. 68, no. 4, pp. 365–380, 2000.

[18] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. R. Stat. Soc. Ser. B (Statistical Methodol.*, vol. 67, no. 2, pp. 301–320, 2005.

[19] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Stat. Soc.*, pp. 267–288, 1996.

[20] S. H. Park, "Collinearity and optimal restrictions on regression parameters for estimating responses," *Technometrics*, vol. 23, no. 3, pp. 289–295, 1981.

[21] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC Press, 1984.

[22] A. Liaw and M. Wiener, "Classification and Regression by randomForest.," *R news*, vol. 2, no. 3, pp. 18–22, 2002.

[23] B. S. Everitt and D. C. Howell, *Encyclopedia of statistics in behavioral science.* John Wiley & Sons Ltd, 2005.

[24] K. Chang, *Introduction to geographic information systems*. McGraw-Hill Higher Education Boston, 2006.