The Influence of City Size on Dietary Choices

Hao Cheng
Institut für Kartographie und Geoinformatik
Hannover, Germany
cheng@ikg.uni-hannover.de

Markus Rokicki
L3S Research Center
Hannover, Germany
rokicki@l3s.de

Eelco Herder
L3S Research Center, Germany & Radboud Universiteit Nijmegen, the Netherlands
eelcoherder@acm.org

ABSTRACT

In the past decades, the process of urbanization has shaped general socio-economic aspects of cities with different population sizes. Among them, food consumption is a good indicator to reflect the quality of life. In this paper, we study the impact of city size on food preferences, as shown by users of a large German food sharing community. We quantitatively and qualitatively analyze differences in dietary choices made by users who indicate to live in cities of different sizes, from metropolises and big cities to medium and small towns. Further, we demonstrate that the city size of the creators of online recipes can be predicted with a good accuracy of 86%, using predictors based on recipe authors' profiles, recipe popularity, season, and recipe complexity and contents. The findings indicate that city size is a useful feature to take into account in various other domains.

KEYWORDS

Online food; city size differences; classification; food consumption

1 INTRODUCTION

In the past decades, urbanization has taken place around the world, with increasing numbers of people living in cities. Cities are believed to be focal points for economic growth, innovation, and employment [6]. Researchers have found that socio-economic characteristics are largely shaped by a city's population size [3]. There has been some recent interest in investigating whether people eat differently and show different culinary activities across city sizes. For example, [17] shows differences in Asian cities of different sizes in the process of urbanization, diet change, and transformation of food supply chains.

Various studies on the influence of urbanization on food consumption have been conducted, albeit usually based on questionnaires and interviews [7, 23]. In addition, differences in eating habits between countries, states, and cities have been observed in quantitative studies [2, 8, 13, 19, 22, 24].

One challenge in studying the relation between food consumption and urbanization lies in collecting large amounts of data across cities. Nowadays, recipes and cooking information is readily available and easier to access than before. Online food communities, populated by users with various demographic backgrounds, provide a rich source of information for learning culinary patterns and predicting personal preferences.

In this paper, we investigate how *city size* captures many individual differences in food preferences and thus can serve as a meaningful adddition to more traditional spatial features, such as geographical coordinates and country. To the best of our knowledge, we are the first to use online recipe data for quantitative and qualitative analysis of eating habits and preferences in relation to city size. Further, understanding food preferences across city sizes can be leveraged to improve food recommendation performance, which we confirmed in [5] by comparing different spatio-temporal contexts for context-aware food recommenders.

Contributions. In this paper, we explore the impact of city size in the large German online food community Kochbar[1]. We conduct a two-fold study on differences in food preferences. First, we perform statistical and qualitative analyses to investigate the nature of these differences between different city sizes. Further, we perform a classification experiment to investigate to what extent features related to these differences allow for predicting city size categories for individual recipes and to analyze which of these differences are most meaningful in this context. This way, we aim to provide insights into the nature as well as the impact of city size on differences in the field of cooking and food preferences.

2 RELATED WORK

Influence of City Size. The past decades have witnessed increasing numbers of people moving into cities from rural areas and the expansion of cities of all sizes. Cities provide significant opportunities for economic and social development [6]. City people usually cannot produce their own food, not only due to lack of space, but also due to lack of (spare) time, and therefore are dependent of the city's food chains and food offerings [21]. Bettencourt et al. presented empirical evidence for relations between the population size of cities and a wide range of characteristics, including energy consumption, economic activity, demographics, infrastructure, innovation, employment, patterns of human behavior, using extensive data collected from US metropolitan statistical areas, European larger urban zones, and Chinese urban administrative units [3]. Wealth and prices scale superlinearly with city size, while individual human needs (job, house, household water consumption) scale linearly and material quantities associated with infrastructure scale sublinearly. Sarkar et al. used scaling indicators to analyze income inequality in Australia [20]. They found that a lower-income

This work was partially funded by the German Federal Ministry of Education and Research (BMBF) under project GlycoRec (16SV7172).
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
UMAP'17 Adjunct,, July 09-12, 2017, Bratislava, Slovakia
© 2017 ACM. 978-1-4503-5067-9/17/07…$15.00
DOI: 10.1145/3099023.3099058

[1] http://www.kochbar.de

person is more likely to be found in a small or big city than in a metropole, as the cost of living in metropolitan areas is usually far higher. The level of income, on the other hand, will directly influence the expenditure on food. For example, both Furey et al. [7] and Walker et al. [23] studied the influence of urbanization on local accessibility of affordable, healthy, and nutritious food in Northern Ireland and the US respectively. Lower-income people with limited mobility turned out to suffer most from “food deserts” – areas with limited access to healthy, fresh foods. These observations point to the potential of city size as a proxy to a variety of local differences affecting eating habits and other user preferences.

Regional Differences in Culinary Activities. In addition to studies on urbanization, there are culinary analyses based on location information. In contrast to the former, most of the following works were carried out using online data. Ahn et al. clustered recipes by their flavors and constructed flavor networks to uncover the ingredient preferences of different cuisines worldwide [2]. Similarly, Sajadmanesh used web data to explore worldwide culinary habits [19]. Howell et al. analyzed taste preferences for different countries [8]. Laufer et al. studied Wikipedia data to analyze European food cultures [13]. Wagner et al. used server logs to reveal ingredient preferences in German-speaking countries (Germany, Austria, and Switzerland) [22]. West et al. analyzed web usage logs to discover nutrient patterns of different American states [24].

As indicated above, location features were mostly extracted in terms of individual countries, states, or cities. However, cities with similar sizes are presumed to share several common characteristics: for instance, Readon et al. [17] analyzed diet changes and transformations in food supply chains in Asian cities in the process of urbanization, finding differences over time, as well as between rural and urban areas. Based on these insights, we propose to group cities into subsets according to their population size and to investigate differences in culinary habits and preferences between these groups.

name	#entities	name	#entities
recipes	405864		
categories	246	category classes	7
ingredients	1485		
cooks	18212	≥ 10 recipes	4976
ratings	7796004	5-star-ratings	7724641
raters	19444	≥ 10 ratings	6231
comments	2751820		
commenters	21951	≥ 10 comments	4922

Table 1: Summary of the dataset

3 METHODOLOGY

In this section, we first introduce the dataset and then the ranking and classification algorithms we used for analyzing users and recipes based on city size.

3.1 Dataset

Our study is based on a large-scale crawl from Kochbar.de, provided by Kusmierczyk et al. [12]. Kochbar.de is one of the most popular German online food communities, where users can upload, search, rate, and comment on food recipes. User profiles contain demographic information and the uploaded recipes contain specific information about ingredients, cooking directions, nutritional data, comments, user views, and ratings.

The dataset consists of over 400 thousand recipes in more than 200 categories, published between March 2008 and November 2014 (see Table 1 for an overview). Users provided more than 7 million ratings on the recipes and out of the active raters (those who have rated at least 10 recipes), more than 2 thousand provided location information in their profiles. The ratings are overwhelmingly positive, with over 99% of the ratings being a 5 star rating. Cooks are those who at least uploaded one recipe, over one-forth of them are active ones (those who have uploaded 10 recipes).

City Sizes. City size is not an explicit feature provided by users. Therefore, we use Geonames[2] location data (latitude and longitude) to find the closest city for the users. We adapt a settlement hierarchy[3] to categorize cities based on their populations according to Geonames city population data. In particular, we group cities into five different city sizes: metropolis ($>= 1m$), big city $[500k, 1m)$, medium city $[100k, 500k)$, small city $[50k, 100k)$, and town $[15k, 50k)$.

3.2 Alternative Rank Normalization

In our analysis of differences in recipe content for different city size conditions, recipe title terms, ingredients, and categories that are peculiar or specific to a particular city size are of particular interest. To this end, we make use of techniques for (text) corpus comparison. In their work on termhood extraction via corpus comparison [10], Kit and Liu explored the usefulness of different ranking approaches to describe a given corpus via comparison with a background corpus. In our work, we compare multiple context conditions, such as city sizes, with each other using *Alternative Rank Normalization* (ARN): to characterize recipes in one condition (e.g. to find ingredients used in metropolises but not in others) we construct two ranked lists of the items of interests (e.g. ingredients), one for the corpus of interest (e.g. metropolis) and the background corpus (all other cities). We then calculate the (normalized) rank difference in the item’s rank between both corpuses and sort the items by the rank difference; the most salient items for the corpus are the ones with the largest rank difference (i.e. ranked high in the corpus of interest and ranked low in the background corpus).

3.3 Random Forests

To predict recipe city size, we use Random Forests [4], a state-of-the-art classifier that is resistant to overfitting and that can also be applied to rank importance of features.

Breiman showed that the accuracy of random forests depends on the strength of individual trees and the correlation between the trees [4]. A modified bagging procedure is used to learn the

[2] http://www.geonames.org
[3] https://en.wikipedia.org/wiki/Settlement_hierarchy

tree ensemble. In addition to repeated randomly sampling a new training set to grow a new tree for the ensemble, the tree learner also sub-samples the feature space randomly at each split, thus reducing correlation between individual trees. Breiman pointed out that bagging in this manner enhances the accuracy and can be used to estimate the generalization error, strength, and correlation of combined trees. Although individual trees are sensitive to overfitting, the average of the vote of all the combined trees is not, despite increasing model complexity by incorporating more trees.

4 DATA ANALYSIS

In this section, we analyze users and recipes from different city sizes. In order to reduce noise caused by exotic recipes and eccentric users, we focus on users who have rated at least 20 recipes and recipes with at least 10 ratings. Table 2 shows an overview of the resulting dataset used for our analysis. Recipes are assigned to city sizes based on the location information provided by their cooks.

Table 2: Overview of the city size dataset

sizes	cities	pop.	cooks	recipes	raters[a]	ratings
metro.	3	6425862	339	12735	3632	386194
big	10	6028762	387	12358	3903	399695
med.	74	16352954	778	27014	4250	636914
small	107	8610911	509	14502	3950	433919
town	640	48314158	1876	89705	4574	219259
sum	834	85732647	3889	156314	4718	4675981

[a] sum of distinct raters

4.1 Analysis of User and Recipe Attributes

Figure 1 depicts the percentages of the population, cooks, recipes, and ratings relative to the respective total sums for each city size. In terms of absolute numbers, most users live in towns. Medium-size cities have relatively more cooks than their population (19.1 %). However, this somewhat larger user community provides fewer recipes (17.3 %) and ratings (13.6 %). By contrast, towns (cities of 50,000 and smaller) have relatively fewer cooks (48.2%), but these users are more active in terms of recipes (57.4 %) and ratings (60.3 %). Indeed, the average number of recipes per cook is statistically higher for towns ($M = 47.82$) than for small cities ($M = 28.49$, $W = 337150$, $p < .001$, $r = .07$) and big cities ($M = 31.93$, $W = 449950$, $p < .05$, $r = .05$).

Recipe uploading behavior differs significantly between city sizes during different days of week ($N = 156314$, $\chi^2 = 352.55$, $df = 12$, $p \ll .001$) and seasons ($N = 156314$, $\chi^2 = 63.165$, $df = 4$, $p \ll .001$). For example, cooks from small cities uploaded many more recipes on weekends, whereas cooks from metropolises uploaded fewer recipes on weekdays than expected. In Spring, cooks from medium cities were more motivated to upload recipes; in Autumn, cooks from small cities were less active.

Small but significant differences of cooks' demographics in terms of age and gender are also found between city sizes. Medium city cooks are younger on average ($M = 39.3$) compared to the other city sizes ($M = 40.5$), $W = 266970$, $p \ll .001$, $r = .08$, while town cooks are slightly older ($M = 40.9$). In terms of gender, the distribution is unequal as well. Further, with respect to recipe popularity, differences between genders align with observed differences in the number of ratings between city sizes ($N = 4697785$, $\chi^2 = 145500$, $df = 4$, $p \ll .001$), as suggested by previous work on gender differences in online cooking [18].

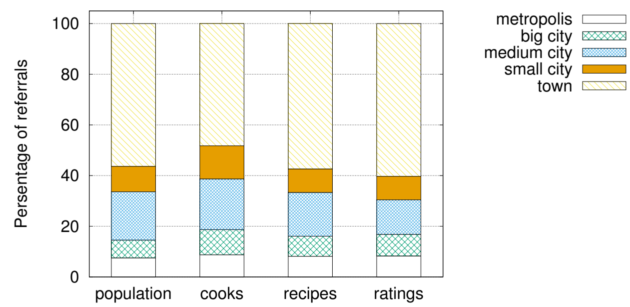

Figure 1: Cook, recipe, and rating fractions across city sizes.

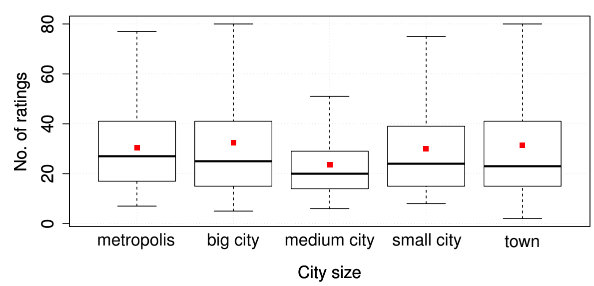

Figure 2: Number of ratings given to recipes that were uploaded by cooks from different city sizes.

In terms of recipe popularity, Figure 2 shows that there are significant differences in the average number of ratings between the city sizes. In particular, big city recipes ($M = 33.8$) and towns ($M = 37.2$) receive far more than medium-size city recipes ($M = 24.4$, town vs. medium city: $W = 47267000$, $p \ll .001$, $r = .32$; big city vs. medium city: $W = 9216300$, $p \ll .001$, $r = .18$). In terms of views, we can observe that small city ($M = 1883$) and big city ($M = 1637$) recipes receive more views than medium city ($M = 1446$) and metropolis recipes ($M = 1165$). The distribution of the comments is very similar to the ratings.

To sum up, significant differences are found for the cooks in the city size division. The majority of cooks are from towns and tend to be more active in uploading recipes. Their demographic attributes – cooks' gender and age – differ depending on the city size category.

4.2 Analysis of Recipe Contents

We continue our analysis by focusing on the features of the recipes in the category "main dishes". Based on 31 identified red-meat ingredients [18], we observe a relationship between the use of red meat and city size ($N = 45497$, $\chi^2 = 20.34$, $df = 4$, $p < .001$). On average, red meat is used more in big cities ($M = 38.4\,\%$) and less in metropolises ($M = 34.7\,\%$) and medium ($M = 36.4\,\%$) cities. In line with the expectation of more exotic dishes in larger cities, spices are more frequently used in big cities ($M = 94.3\,\%$) and metropolises ($M = 93.4\,\%$) than in medium ($91.3\,\%$) cities ($N = 45497, \chi^2 = 35, df = 4, p \ll .001$). This is illustrated nicely by the number of curry dishes decreasing with city size, as shown in Figure 3.

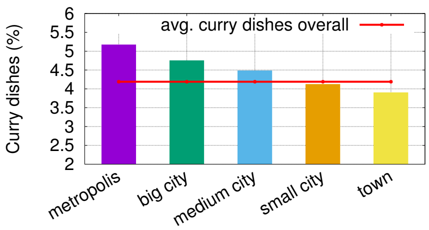

Figure 3: Curries used in main dishes for each city size.

Main dishes from metropolises ($M = 12.6$), big cities ($M = 12.5$), and small cities ($M = 12.3$) use more ingredients than the ones of the medium-cities ($M = 11.52$, e.g. metropolis vs. medium city: $W = 11288000$, $p \ll .001$, $r = .15$). Similarly, metropolis ($M = 45.7min$), small city ($M = 43.9min$), and big city ($M = 43.2min$) main dishes take longer to prepare than medium city dishes ($M = 39.9min$, e.g. metropolis vs. medium city: $W = 4453400$, $p \ll .001$, $r = .55$).

In terms of nutrients, small-city main dishes contain fewer calories ($M = 809kJ$) than dishes from the other city sizes (e.g. small city vs. metropolis $M = 927.7kJ$: $W = 5744900$, $p \ll .001$, $r = .13$), as well as less fat (e.g. small city $M = 15.9g$ vs. metropolis $M = 13.2g$: $W = 5666500$, $p \ll .001$, $r = .12$), but more protein (such as small city $M = 6.7g$ vs. medium city $M = 6.2g$: $W = 8610800$, $p < .001$, $r = .06$). The differences for carbohydrates are not that obvious, but still significant. For example, small-city ($M = 11.7g$) main dishes contain more carbohydrates than big-city ($M = 10.7g$) main dishes ($M = 6.2g$: $W = 4566900$, $p < .001$, $r = .13$).

In summary, metropolis main dishes are more complicated than the medium city main dishes. With regard to the nutrition, the small city main dishes are better in quality than the medium city main dishes.

4.3 Qualitative Analysis of Recipe Contents

Next, we analyze recipe titles[4], categories and ingredients qualitatively by means of ARN (see Section 3.2). Again, we focus on recipes from the category "main dishes", to avoid bias caused by the type of dishes.

Table 3 lists the 20 most peculiar title terms for the different city sizes. There are many foreign dishes in the metropolis and big city categories, some even in the medium city category. By contrast, the small city and town categories contain more traditional recipe titles. For example, *dim sum* and *sticky rice* are typical Asian foods, while *Baden* and *wild* recipes are very local and traditional.

The foreign and traditional pattern is further confirmed by recipe categories (the corresponding table is not shown in this paper due to space limitations). Metropolises and big cities contain more exotic

[4]Recipe titles are preprocessed by applying tokenization and stemming with NLTK (http://www.nltk.org) and manually removing noisy terms.

Table 3: 20 most peculiar terms of the main dish titles for each of the different city sizes.

metropolis	big city	medium city	small city	town
dim sum	dish from oven	main dish	plate dish	from Baden[c]
Alsace	Hessian	**foreign**	small bears	wild recipe
Turkey	**Indonesian**	low-fat	main dish	cooking mama
sticky rice	Franconian	fried	low in calories	poultry dish
buffalo mozzarella	**African**	**Madeira**	mashed potatoes	minced meat
Quinoa	**cabanossi**[b]	menu	noodle dish	summer salad
finger foods	**chutney**	**Moroccan**	meatball	leg slices
sliced potatoes	bacon	artichoke	meat dish	main dish
chakalaka	peanut	oven bag	potato dumplings	witches' cuisine
Pasta sauce	cuisine	Berlin	exotic	made in wok
gremolata	feta cheese	mother	ground pork	herb witch
Harissa	chilli	dish	casserole	lamb steak
fried rice	loin ribs	old	stew	mill's type
fried sausage	fish	fried chicken	Schupfnudeln	roasted
Italy	creamy	**Caribbean**	pan dish	barbecue side
Cajun	dressing	veal meat	casserole dish	carrot salad
deluxe	meat	pie	pork knuckle	spaghetti pan
baked vegetables	cream	snack	grandma's	pasta
dinner	asparagus	roe deer back	potato pancake	crispy
Pleurotus	potatoes	tender	red wine	king prawn

[a] foreign dishes are written in bold
[b] Polish sausage
[c] federal state of Baden-Württemberg of Germany

categories, like Orient, Turkey and Indonesia, whereas location categories for towns include the local East-Frisian, Swabian and Thuringian cuisines, as well as some nearby countries like Denmark and Switzerland. Some unexplainable patterns have been found as well, though, such as the presence of Romania in the town category. This may be caused by a large minority in a particular town or due to a nearby big city or metropolis, where people have far easier access to more exotic ingredients and foreign cuisines.

The same pattern can also be found in the selection of ingredients used. For example, metropolis recipes use *soya*, *quinoa*, *cubeb*, *aioli*, and *tellicherry pepper* distinctively. These ingredients are either from foreign countries, exclusive or expensive. Many of these ingredients are also spice ingredients: we observed a descending use of spices from the metropolis to town categories (see Section 4.2).

We further break down recipes to different genders across city sizes. Women show a preference for vegetarian and sweet dishes in the metropolis and the big city categories. This preference becomes weaker in the smaller city size categories. On the other hand, men show a preference for hearty, spicy, and foreign cuisines.

When we closer inspect the three major metropolises (Berlin, Hamburg, and Munich), we observe that – apart from the foreign cuisine cooked in all three – each city has its own distinct characteristics. Vegetarian and vegan food is most popular in Berlin. The harbour city of Hamburg shows a slight preference for seafood. Munich is the city where you find Bavarian food and southern European cuisines, especially Italian cuisine, and of course a "beer culture": *drink*, *autumn*, and *with alcohol* are among the 20 most peculiar terms or categories list in Munich.

4.4 Summary of Findings

Users from towns are more active than users from other, larger cities. Smaller city sizes are associated with fewer calories and fat, but also with less exotic and more traditional food, as well as with fewer spices. Medium-city recipes use fewer ingredients and take less time - which might indicate that inhabitants of medium cities experience more time pressure - and consequently also receive fewer views and ratings. At several points in our analysis, we found indicators in recipes from medium cities that hint at a lower average income and a larger proportion of immigrants than in both metropolises or towns – which is in line with social-economic literature [11]. The distinctive terms, categories and ingredients further illustrate the nature of these differences.

5 CITY SIZE CLASSIFICATION

In the previous section, pronounced differences in terms of recipe content and users have been shown across city sizes. To find out which of those differences are most discriminative, we perform a classification experiment to predict city size of recipes.

5.1 Setup

Based on our analysis, we select 53 features related to cooks and recipes, as shown in Table 4. For the cook properties, we consider demographic information, as well as some measures of activity and popularity. Recipe content is represented on a high level by recipe categories. In addition, we include features on recipe popularity, recipe structure and complexity, nutrition, and temporal context.

Table 4: Features for recipe city size classification

Cook (6 features)
guest-book messages, age, points, active-days, uploaded recipes, gender.
Recipe popularity (5 features)
comments, ratings, favorites, views, average-rating.
Recipe structure (6 features)
duration, servings, ingredients, price, difficulty, spice-dish.
Recipe nutrients (4 features)
kJ, carbohydrates, protein, fat.
Recipe time (2 features)
season, day-of-week.
Category (30 features)
occasions, special, international, Europe, main-dish, lunch, supper, summer, cheap, milk-products, spring, party, autumn, quick-easy, regional, winter, coffee-cake, vegetarian, meat, cake, intermezzo, snack, healthy-diet, dessert, starter, gluten-free, lactose-free, casserole, no-wheat, allergy.

In order to avoid biases introduced by a small number of very active users who uploaded hundreds or thousands of recipes, we randomly sample ten recipes per user. Furthermore, we perform under-sampling to balance the dataset across classes, which results in 718 recipes per city size. In the following, we first evaluate the discriminative power of features in terms of decrease in Random Forest (RF) accuracy (RF-MDA) [14]. Afterwards, we evaluate classification performance in terms of accuracy using 10-fold cross-validation.

5.2 Results

Feature importance in terms of decrease in Random Forest accuracy is shown in Figure 4. Although we observe clear differences in feature quality between – but also within – feature groups, the feature importance scores show that almost all of the features groups are meaningful. Except for nutrient features, each of the feature groups is evenly represented among the 13 best features. It is worthwhile to note that nutrient values are particularly subject to noise, as some users provide inaccurate values.

Ranked most highly within the eight most useful features are all features that characterize recipe cooks in terms of their demographics, activity and popularity. Surprisingly, the temporal context – recipe season in particular – is also ranked quite highly. The remaining highly ranked feature groups relate to recipe popularity, content, and structure. The importance of popularity features – such as number of comments, ratings, favorites, and views – reinforce the results on differences in terms of recipe popularity shown in the previous section. Further, although the content features are coarse-grained and overall lower-ranked, our findings with regards to differences in food preferences are also reflected in the feature ranking. Categories for special and international dishes are ranked particularly high and categories that encode the other main differences found in the previous section, such as vegetarian and regional dishes, achieve overall similar feature importance scores. In addition, differences in recipe complexity in terms of duration, servings and number of ingredients appear to be useful as well.

Using all of the features, the Random Forest classifier predicts recipe city size with an overall high accuracy of 78%. Restricting the feature set to the top 20 features according to RF-MDA further improves classification accuracy to 86%. The confusion matrix for the latter results, shown in Table 5, further shows that the classification performance is robust across the city sizes, with only minor variations.

Table 5: Confusion Matrix using the top 20 features

classified as →	a	b	c	d	e
a = metropolis	**628**	37	14	21	18
b = big city	28	**610**	24	28	28
c = medium city	25	33	**603**	30	27
d = small city	32	33	22	**617**	14
e = town	19	29	27	15	**628**

6 DISCUSSION AND CONCLUSION

In this paper, we have shown that food preferences depend on the size of city that people are living in and discussed the nature of these differences. Among others, in Germany, people in metropolises eat more foreign food and people in smaller cities and towns eat more traditionally. Medium-city recipes contain less protein but more calories and fat than recipes from other city sizes. These features are sufficient for reliably predicting a user's city size.

The findings of the influence of city size on dietary choices provide meaningful information for food recommendation. In [5], we show that using city size as a feature has a positive impact on context-aware recipe recommendation [1]: among several spatio-temporal contexts (day-of-week, season, and inner-border), city size turned out to give the best performance for food recommendation.

For our analysis, we relied on the user's self-reported location, without taking any nearby larger cities into account. In future work,

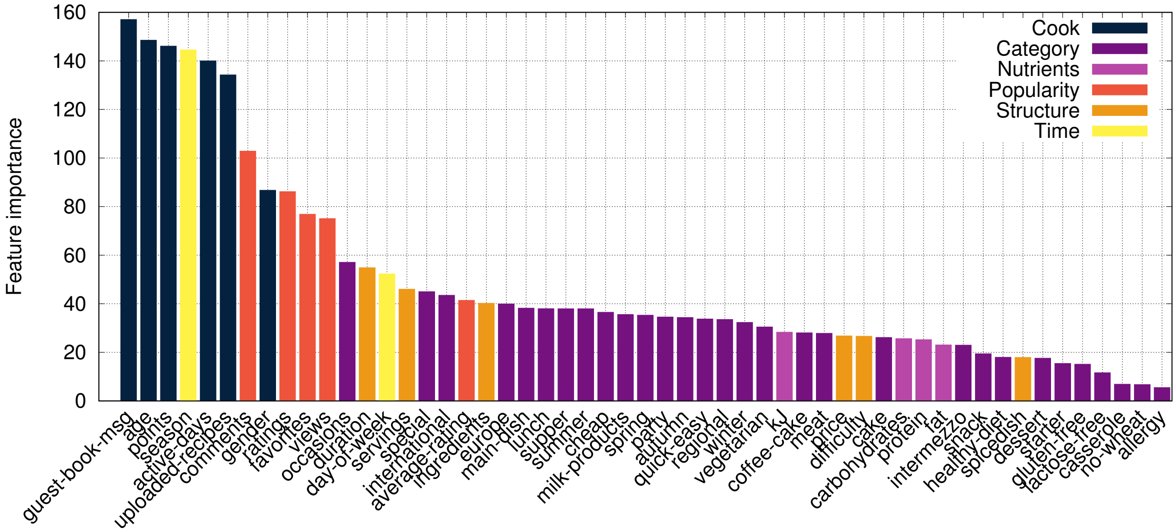

Figure 4: Features ranked by importance in terms of RF-MDA.

it may be useful to take, in addition to the population, distances to adjacent cities into account - in order to distinguish between isolated towns and towns close to big cities or metropolises.

The influence of city size on food preferences, as disccussed in this paper, is in line with the related work as discussed in Section 2. Knowledge on the impact on city size can be effectively translated into measures to reinforce or counteract such effects [16]: for example, it has been found that as city size grows, people tend to be less socially connected and less interested in local politics and affairs [15]; these findings have been used for improving the configuration of metropolitan institutions [9]. In sum, the size of a city has a clear impact on the habits, possibilities, interests and preferences of its inhabitants and therefore is expected to be a useful feature in context-based recommendation in general.

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2011. Context-aware recommender systems. In *Recommender systems handbook*. Springer, 217–253.

[2] Yong-Yeol Ahn, Sebastian E Ahnert, James P Bagrow, and Albert-László Barabási. 2011. Flavor network and the principles of food pairing. *Scientific reports* 1 (2011).

[3] Luís MA Bettencourt, José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B West. 2007. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences* 104, 17 (2007), 7301–7306.

[4] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[5] Hao Cheng, Markus Rokicki, and Eelco Herder. 2017. The Influence of City Size on Dietary Choices and Food Recommendation. In *Proceedings of the 2017 Conference on User Modeling Adaptation and Personalization*. ACM. To appear.

[6] Barney Cohen. 2006. Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in society* 28, 1 (2006), 63–80.

[7] Sinéad Furey, Christopher Strugnell, and Ms Heather McIlveen. 2001. An investigation of the potential existence of"food deserts"in rural and urban areas of Northern Ireland. *Agriculture and Human Values* 18, 4 (2001), 447–457.

[8] Patrick D Howell, Layla D Martin, Hesamoddin Salehian, Chul Lee, Kyler M Eastman, and Joohyun Kim. 2016. Analyzing Taste Preferences From Crowdsourced Food Entries. In *Proceedings of the 6th International Conference on Digital Health Conference*. ACM, 131–140.

[9] Christine A Kelleher and David Lowery. 2009. Central city size, metropolitan institutions and political participation. *British Journal of Political Science* 39, 01 (2009), 59–92.

[10] Chunyu Kit and Xiaoyue Liu. 2008. Measuring mono-word termhood by rank difference via corpus comparison. *Terminology* 14, 2 (2008), 204–229.

[11] Amitabh Kundu and Niranjan Sarangi. 2007. Migration, Employment Status and Poverty: An Analysis across Urban Centres. *Economic and Political Weekly* 42, 4 (2007), 299–306.

[12] Tomasz Kusmierczyk, Christoph Trattner, and Kjetil Nørvåg. 2015. Temporality in online food recipe consumption and production. In *Proceedings of the 24th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 55–56.

[13] Paul Laufer, Claudia Wagner, Fabian Flöck, and Markus Strohmaier. 2015. Mining cross-cultural relations from Wikipedia: A study of 31 European food cultures. In *Proceedings of the ACM Web Science Conference*. ACM, 3.

[14] Gilles Louppe, Louis Wehenkel, Antonio Sutera, and Pierre Geurts. 2013. Understanding variable importances in forests of randomized trees. In *Advances in neural information processing systems*. 431–439.

[15] J Eric Oliver. 2000. City size and civic involvement in metropolitan America. *American Political Science Review* 94, 02 (2000), 361–373.

[16] Michael Parkinson, Richard Meegan, and Jay Karecha. 2015. City size and economic performance: Is bigger better, small more beautiful or middling marvellous? *European Planning Studies* 23, 6 (2015), 1054–1068.

[17] Thomas Reardon, David Tschirley, Michael Dolislager, Jason Snyder, Chaoran Hu, and Stephanie White. 2014. Urbanization, diet change, and transformation of food supply chains in Asia. *Michigan: Global Center for Food Systems Innovation* (2014).

[18] Markus Rokicki, Eelco Herder, Tomasz Kuśmierczyk, and Christoph Trattner. 2016. Plate and prejudice: gender differences in online cooking. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. ACM, 207–215.

[19] Sina Sajadmanesh, Sina Jafarzadeh, Seyed Ali Ossia, Hamid R Rabiee, Hamed Haddadi, Yelena Mejova, Mirco Musolesi, Emiliano De Cristofaro, and Gianluca Stringhini. 2016. Kissing Cuisines: Exploring Worldwide Culinary Habits on the Web. *arXiv preprint arXiv:1610.08469* (2016).

[20] Somwrita Sarkar, Peter Phibbs, Roderick Simpson, and Sachin Wasnik. 2016. The scaling of income distribution in Australia: Possible relationships between urban allometry, city size, and economic inequality. *Environment and Planning B: Planning and Design* (2016), 0265813516676488.

[21] Philipp Stierand. 2008. *Stadt und Lebensmittel*. Ph.D. Dissertation. TU Dortmund.

[22] Claudia Wagner, Philipp Singer, and Markus Strohmaier. 2014. The nature and evolution of online food preferences. *EPJ Data Science* 3, 1 (2014), 1.

[23] Renee E Walker, Christopher R Keane, and Jessica G Burke. 2010. Disparities and access to healthy food in the United States: A review of food deserts literature. *Health & place* 16, 5 (2010), 876–884.

[24] Robert West, Ryen W White, and Eric Horvitz. 2013. From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 1399–1410.