IDENTIFYING KYIV RESTAURANT
FEATURES VIA SENTIMENT ANALYSIS
USING TRIPADVISOR REVIEWS

by

Kateryna Moisiienko

A thesis submitted in partial fulfillment of the
requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2020

Thesis Supervisor: _____ Professor Maksym Obrizan

Approved by _____
        Head of the KSE Defense Committee, Professor Olena Besedina

        _____

        _____

        _____

        Date _____

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Polarized terms Words associated with positive or negative emotion

Negators Words that invert polarized meaning, e.g. "not good"

Amplifiers Words that increase emotional intent

De-amplifiers Words that decrease emotional intent

Association Correlation with every other word in the review. Scores range from 0 to 1. A score of 1 means that two words always appear together in documents, while a score approaching 0 means the terms seldom appear in the same document

CHAPTER 1. INTRODUCTION

"Although a great restaurant experience must include great food, a bad restaurant experience can be achieved through bad service alone. Ideally, service is invisible. You notice only when something goes wrong."

– Dana Spiotta

The food service industry has grown significantly over the last years in Ukraine. The pace of modern life is forcing people to delegate cooking to professionals, order food or eat out of their homes. On the other hand, Ukraine becomes a very interesting place to visit for foreign tourists due to low prices, historical heritage and its close neighborhood to European countries. The demand for restaurants and cafés has grown, but does it mean the increase in restaurants success and profit? A growing market means more customers look to dine out. However, it also means that there is increased competition — especially as more and more restaurants open in major megalopolis areas. What makes one restaurant more successful than another does? What are the major concerns of customers for a great meal? We could say that tasty food, good location and great service may guarantee customer satisfaction, but it might not be true depending on cuisine type, price level, trends and other factors.

Nowadays food service industry faces with unseen before crisis – COVID19 pandemic. The tight quarantine restrictions, introduced in spring, caused the closure about 30% of catering establishments[1] and among those there are many previously very successful and recognizable ones – café chain Zheltok, Rakhat Lukum restaurant, Lost&Found bar etc. At the same time, many new restaurants are opening, for example in October more

---

[1] https://www.dsnews.ua/society/30-kafe-i-restoranov-zakrylis-posle-karantina-v-ukraine-13062020142000

than 40 restaurants were opened only in Kyiv[2]. Thus, for some factors there are restaurants for which the crisis opens up new opportunities.

Being motivated to know more about food service in Ukraine, its development in time and how restaurants survive the coronavirus crisis, the author faces the problem that the majority of studies about food service in Ukraine are quite general and contain only descriptive statistics or not freely available, because surveys are provided for private organizations or restaurants.

Since we have no access to restaurants private information: number of employers, profit, costs, customers bill details, - we are going to provide a survey based on data from the TripAdvisor, which nowadays is leading website offering business listings and reviews[3] for restaurants, cafes, and any other businesses which cater to travelers and local citizens. TripAdvisor is a perfect service for tourists and citizens in order to find a hotel, café or restaurant. It makes search easier, suggesting filters by price, cuisine, location and finally – ranking. After making a decision about cuisine type, price level and location, final choice may depend on customers` reviews. Nevertheless, most of the reviews often contain very less details about that particular product and have more of other sentences, which are not useful to the potential customer. Hence, we need to extract only information required by the customer and trim out other undesired information. Talking about not only customers, but also owners, investors, if they could see the big picture about food industry, their rivals, trends and update actual information more frequent, probably their decisions could be more successful and precise.

The purpose of this research is to investigate which restaurant characters or determinants make significant impact on customers satisfaction and then on its popularity (ranking) in Ukraine, check trends and assumptions, find out which dishes and drinks are

---

[2] https://horeca-ukraine.com/novye-zavedenija-kieva-v-oktjabre/

[3] https://tripadvisor.mediaroom.com/us-about-us

most often mentioned in the reviews, how the emotion intensity in the reviews correlates with the reviews length, etc. The rank is very important for restaurants because TripAdvisor shows restaurants in ascending ranking order and the lower the rank the further the restaurant will be in the list and the lower probability to be seeing by seeker. The data collected for restaurants could be conditionally divided into two parts. First contains general description about restaurants listed on TripAdvisor in eight large cities (Kyiv, Odessa, Lviv, Kharkiv, Vinnytsia, Zaporizhia, Dnipro and Kherson) such as ranking in the total list of restaurants in the particular city, number of reviews, average rating, which is defined by customers from one to five, where "one" means "terrible" and "five" – "excellent", total number of reviews, cuisine type and price level. This data is necessary to understand the general picture of restaurants and cafes in Ukraine, what the proportion of cheap, middle are and expensive restaurants, what are their average ratings, etc. Second – data with reviews details for Kyiv restaurants. Since Kyiv restaurants types cover variety cuisines, the number of restaurants in Kiev almost reaches the number of restaurants in seven remaining targeted cities taken together and on the other hand, it has as much reviews as all mentioned above cities we decided to concentrate sentiment analysis especially on Kyiv restaurants reviews. After cleansing reviews from stop words, punctuation, stemming them we are going to provide data mining analysis and visualize the results. Then we will conduct sentiment analysis for them, in particular to find the polarity scores to each review, identify positive and negative words in the reviews as well as negators, amplifiers, de-amplifiers that has also impact on the contextual sentiment. We expect the number of positive reviews be higher than negative; positive relationship between the polarity score and reviews length.

We suppose this research to be very useful for investors and executives in restaurants who are going to beat competitors, increase revenue streams, and grow their organization.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

Understanding how external factors influence restaurant[4] industry in Ukraine is essential in making smart decisions and guiding the restaurant to greater heights. Government stability is an important factor in the success of restaurant industry. In 2018, "Political stability and absence of violence/terrorism" index in Ukraine has the value 20.89, which took 123rd place among 144 countries[5]. When people are not confident about their future wealth due to war, anti-government protests, conflicts and frequent government changes, they buy essentials firstly and postpone entertainments. Restaurant dishes are more likely to be elastic goods, consumption of which will fall much faster if income declines. Inflation rate, unemployment rate, exchange rate, consumer disposable income in Ukraine become major obstacles for restaurant owners. Inflation and exchange rate affects prices for ingredients and as a result, stronger restaurants lose profits, while small newcomers may shut down or substitute ingredients to cheaper, local and probably with lower quality.

However, according to dynamic food service development in Ukraine, we could assume that even having less money and fears due to inflation, low salaries, war and instability people still demand dining out. We cannot make our judgement based on restaurants overall turnout in Ukraine, but the number of reviews and its dynamic through time signals that the review number per week increases when people start to visit restaurants more. On the Figure 1 we may observe how weekly reviews number changes for Kyiv restaurants on TripAdvisor from January 2008 to September 2020. Being at the point of 240 reviews per week in early March 2020 and then fall drastically in April to less than 10 reviews per week – the count, which was last time in late 2011. This graph shows a huge impact of quarantine restrictions on this industry. Important to mention that this graph does not include the reviews for food delivery.

---

[4] Under the word "restaurants" author means "restaurants, cafes, bars and pubs"

[5] https://tcdata360.worldbank.org/

Figure 1. The number of reviews for Kyiv restaurants per week in TripAdvisor
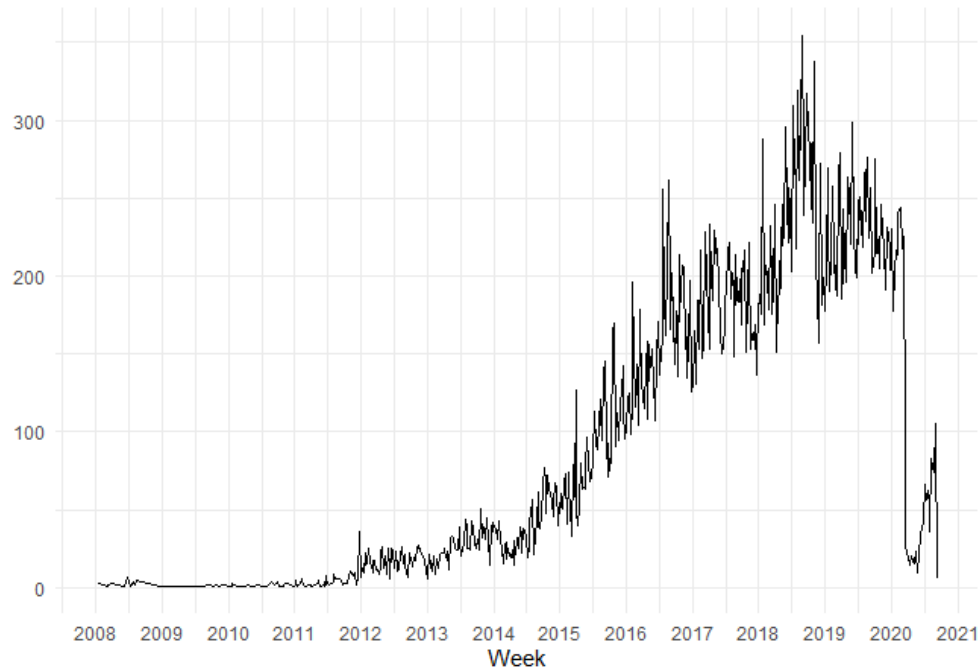
Like many industries, the restaurant industry faces a variety of challenges keeping up with the rapid pace of change driven by the consumer trends. Growing preferences for healthier food options, concerns over environmental sustainability, increased competition from grocery stores, heightened consumer expectations, and the rising spending power of the millennial generation of consumers is accelerating the industry's response to such trends (Appetite for change, KPMG 2016). In 2017 more than a half of Ukrainians in age of 18-30 years old visited restaurants regularly[6], for the range 30-39 years 40% approved that they dine out regularly and only 33% of 40-49 years old. It means that young people from 18 to 40 years old drive restaurant industry in Ukraine the most. Born between 1980 to 2000, these 20 to 40-years-old people, also called "Millennials", make up 29,3% of the total population in Ukraine[7]. Millennials' food choices are much different from those of

---

[6] According to Kiev International Institute of Sociology

[7] Estimated manually according to statistical publication "Resident population of Ukraine by sex and age on January 1, 2020" from State Statistics Service of Ukraine

their parents. Eve Turow (2015) made interesting and deep research about Millenials: they are more passionate about food than any other in history is, nearly 50% of Gen Y-ers (or Millennials) self-identify as "foodies". She also claimed that Millennials more than ever tend to share a feedback in social media, encourage or discourage others to purchase, what they purchased.

One more interesting survey about Millennials was «What Millennials Eat: How social data pushes innovation» (Linkfluence 2018). This report looks at the way that Millennials talk about food online. What motivates them to buy, their sentiment towards certain foods, and what do they want to see more in the food market. The methodology of the research were the following: first authors collected texts of public conversations from Facebook, twitter and Instagram from January 2018 to November 2018, keep only two languages – English in the United states and French in France and analyzed the content of it. They found out that nearly 38% of millennials wright about food in The United States and 48% Millennials in France are about food. This and previous related studies confirm that nowadays dinning in the restaurant is more than only physiological needs satisfaction, but the way to relax, share experience, enjoy etc.

Among Ukrainians scientists, Hirniak L.I. and Glagola V.A. (2018) provided research about current state, perspectives and trends of restaurant business development in Ukraine. They discussed how crisis impact on the restaurants` number and turnover in Ukraine and concluded which trends will dominate next years: development trends of certain national cuisines, restaurants offering craft dishes, the implementation of healthy diet, cooking in the presence of the customer from farm products, opening restaurants in the format casual, original establishments of factories and markets. We assume these trends be reflected in increasing word frequencies among reviews.

Digitalization improves service and quality. Years ago restaurants collected handwritten customers reviews in books of complaints and suggestions, which had almost no effect on the restaurants` reputation. Today every single positive or negative review may

be seeing by thousands of people immediately after the meal. Traveling websites forces restaurants to remain competitive, because they have made it easier for people to compare restaurants by cuisine, price level or look through reviews. In this context, consumers are increasingly seeking information and recommendations ahead of a restaurant booking. Online reviews are very decisive when searching for a restaurant, both while travelling and in one's home. Nowadays, if two neighborhood restaurants offers quite similar products, customer will choose the cheaper one. The solution is to give a unique service or exceptional quality. Higher customer satisfaction should lead to increased probability of repeated purchase, which in turn should result in greater restaurant sales.

Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis of customers' opinion would cover much more approaches than only average rating on some particular travel website. These days Sentiment analysis used in social media monitoring and Voice of Customer to track customer reviews, survey responses, competitors, etc. However, it is also practical for use in business analytics and situations in which text needs to be analyzed.

Plenty of works of sentiment analysis classifies reviews (or documents) by their overall sentiment, defining positive, negative words, negators, amplifiers and de-amplifiers and determine whether whole review is negative or positive (Pang and Lee, 2004). A number or researchers choose the method of dividing the whole text into single words and then assign to each word a polarity score (Turney, 2002). In this work author apply both methods in order to analyze reviews from different point of view.

Many studies related to sentiment analysis were conducted based on Yelp reviews. Raja Irfan Sabir and Muhammad Irfan (2014) explored customer satisfaction of Pakistan local restaurants industry.

New-York scientists Boya Yu, Jiaxu Zhou, Yi Zhang, Yunong Cao (2015) identified different features for restaurants of different cuisines. Their method was based on a high-accuracy support vector machine (SVM) model, calculating word scores and measuring the polarity. The importance of this work is that the essential features discovered might not only help customers to choose their favorite cuisine, but also provide restaurants with their advantages and shortages.

Theresa Wilson, Janyce Wiebe and Paul Hoffmann (2004) suggested a new approach to phrase-level sentiment analysis that first determines whether an expression is neutral or polar and then disambiguates the polarity of the polar expression. Sometimes the review may not include sentiment words, but overall have a positive or negative meaning; this method motivates to consider this part of the reviews separately to avoid wrong polarity.

Victoria Svensson in Stockholm University provided interesting research about The Effect of Demand Density on Firm Productivity in the Restaurant Industry. Her findings suggest that firm productivity dispersion is lager in markets with higher demand density. She also found evidence that suggests that the minimum productivity level is lower in more dense markets. We can adapt her idea taking rating or ranking as productivity and compare dispersion of restaurants rating depending on the demand density. For the restaurant, industry demand density is estimated as the number of inhabitants per square kilometer.

We did not succeed in finding natural language processing, text mining and sentiment analysis on reviews of restaurants or cafes in Ukraine. However, the topicality is clear due to necessity to provide restaurants policy in order to improve customers experience more accurate.

CHAPTER 3. METHODOLOGY

Many people use TripAdvisor to find a good restaurant. Nonetheless, overall rating for each restaurant, cuisine type or location is not enough for independently judging other important aspects such as flavor, atmosphere, service quality etc. In this case, we need to apply for natural language processing, data mining and sentiments analysis of the reviews.

We are going to check the several assumptions about food trends among customers of Ukrainian restaurants. First, we would like to look through time if there any changes in people attitude to service comparing to food. We suppose increasing of mentioning "service", "staff" in reviews. No doubts that "food" remains a defining characteristic of restaurant, but requirements and attitude to service could change.

Recent years we observe the development of coffee culture in Ukraine: not only restaurants and cafes offer coffee, but almost every street or park has a coffee kiosk in it. In 2018 among all types of coffee imported by Ukraine, including instant and ground coffee, more than half are green coffee beans[8]. Restaurants are gradually moving away from dark Italian roast to fresh Ukrainian roast. They are forced to implement the trend, because now people eat in restaurants, but drink coffee in the coffee shop opposite. People already know that coffee doesn't have to be so bitter that it has to be "quenched" with sugar and milk. Coffee shops have already shown them that a balanced coffee without milk and sugar can be achieved with high quality beans and fresh roasts.[9] Taking into account possible seasonality of this drink we are going to analyze presence of "coffee" in Kyiv reviews.

Hirniak L.I. and Glagola V.A. (2018) mentioned several trends in their work: development trends of certain national cuisines, restaurants offering craft dishes. We

---

[8] http://www.ukrstat.gov.ua/

[9] https://delo.ua/business/kak-restorannyj-biznes-perezhil-pjat-mesjacev-ka-372150/

assume these trends be reflected in increasing word frequencies, associations and bigrams among reviews.

Often authors will use more words in the review when they are more passionate. We supposed that the more strongly client's feeling, the more they may wright and with more opportunities for polarized word usage. Conversely a less impassioned customer may not feel compelled to spend a lot of time writing a review. We are going to check this for our review data.

3.1. Data collection

We gathered data using web-scraping tools from TripAdvisor[10] website for restaurants from January 2008 to September 2020. Data collection was divided in two parts. First – restaurant characteristics in eight large cities: Kyiv, Odessa, Lviv, Kharkiv, Dnipro, Vinnytsia, Kherson, and Zaporizhia. Each restaurant from above cities includes following information: ranking in particular city, name, address, review numbers, cuisine, price level, rating, dispersion among customers grades "terrible", "poor", "average", "very good", "excellent" and existence of "Certificate of excellence", which is an option from TripAdvisor and dispersion of reviews languages. Second part - the base of the sentiment analysis on TripAdvisor - reviews for Kyiv restaurants. In addition, we collected dates, device type from which review was made (phone or laptop).

3.2. Data description

Basic statistics about our data in general will help us to understand the global picture and scales. With time series data we can provide an analysis for better understanding how restaurant industry changed over time and how did consumers` subjective opinions were changed. First check comparison between cities could be provided: which city has more

---

[10] https://www.tripadvisor.com/

luxury restaurants relative to middle prices and cheap, how those are ranked by reviews, does exist a difference among number of restaurants belonging to different cuisines, etc.

3.3. Data preparation and text mining

Since people of different nationalities write reviews in different languages, we need to translate all texts to English for further analysis. Text mining is a process of distilling actionable insights from text. All reviews for Kyiv restaurants is a very unstructured data, which is impossible to evaluate manually without text mining approaches. Using *tm* package we will change class of reviews to volatile corpus. Unnecessary information reduction (making all text lowercase, removing stop words, punctuation, stemming words, removing sparse terms and then excessive white spaces) will help us to navigate and draw important features. We are going to unify dishes names for all the reviews, because transliteration of Ukrainian words can be done differently. In this research, we will concentrate on "bag of words" approach to divide text into separated words regardless word type or order in sentence. We can also define most frequent meet words for restaurants, which will help to identify the common concern of customers and to understand what our customers like or dislike about dishes, atmosphere, service, prices etc.

After data preprocessing we can step to extracting important insights, build bigrams, trigrams to see the most frequent combination of words and visualizations, e.g. word clouds, word networks, histograms, graphs etc.

3.4. Sentiment analysis

Sentiment analysis helps you extract an author's feelings from the review. While "bag of words" method aims to perform analyses without regards to the author's emotion, sentiment analysis uses techniques to match the text to know expressive words like "awful" or "wonderful". Many sentiment analysis methods use a subjectivity lexicons – predefined list of words, associated with a specific emotion or positive or negative feelings. In this

work we are going to use a *qdap* polarity function and work with the tidytext sentiments tibble made of 2 different lexicons – "bing", "afinn". Depending on the type of lexicon sentiment outputs can be as "positive" or "negative", be classified to some of 8 emotions ("trust", "surprise", "sadness", "fear", "anger", "joy", "disgust" and "anticipation"). Bing lexicon contains 6 786 words and afinn lexicon – 2477 words, despite that these counts does not cover the lexicon of people, there are several reasons why it works. First, a linguist named George Zipf created a "Zipf's law", second the principle of least effort also helps support using small lexicons. Zipf's law states that given some text, the frequency of any single word is inversely proportional to its rank in the frequency table. We are going to check Zipf's law on our review data. Sentiment analysis is much more complex than it could appear.

Polarity() function identifies words from the subjectivity lexicon, once a polarized words are "tagged" the function creates a context cluster around them, which includes the four words before and two words after the identified words. So, removing stop words in the previous step will impact the polarity scores, because it affects the words in the cluster. Since we removed all the punctuations from the text, we have to do the same for negators (e.g. "can`t" become "cant") to be properly identified. Then each of the individual words is classified as polarized, neutral, negator, amplifier or de-amplifier. Basically a negative term has a value of -1 and positive has a value of +1, negator reverses the sign, amplifier is valued as 0.8 and de-amplifier is valued at -0.8. In the end all the polarity values are summed and divided by the square root of the word count.

CHAPTER 4. DATA

The restaurant dataset was collected from TripAdvisor in eight Ukrainian Cities: Kyiv, Odessa, Lviv, Dnipro, Vinnytsia, Kharkiv, Kherson, Zaporizhia using ParseHub[11] webscraping program. The data contains details for 5034 restaurants with the following variables: restaurant name, cuisine type, price level, webpage on TripAdvisor, restaurant webpage, review numbers, numbers of travelers rating "excellent", "very good", "average", "poor" and "terrible", average traveler rating, city, existence of Certificate of excellence.

Table 1. Sample of restaurant details table

City	Restaurant name	Total rest. number in city	Traveler rating					Rating	Price
			Exce-llent	Very good	Ave-rage	Poor	Terrible		
Kyiv	4. Mama Manana	2089	271	57	12	7	9	4,61	$$-$$$
Dnipro	1. Mishi Blyahera	398	93	26	2	2	2	4,65	$$-$$$
Kharkiv	1. Italyanskaya Redaktsiya№2	731	13	1	0	0	0	4,93	$
Lviv	1. 4friends Whiskey Pub	540	100	15	4	1	0	4,78	$$-$$$
Odessa	1. Restaurant About U	932	20	3	2	0	1	4,58	$
…	…	…	…	…	…	…	…	…	…

As expected, the highest number of restaurants are in Kyiv, then in Odessa and Kharkiv. TripAdvisor awards the best restaurants with the Travelers' Choice award to restaurants that consistently earn great reviews from travelers and are ranked within the top 10% of properties on TripAdvisor.

---

[11] https://www.parsehub.com/

Figure 2. Number of restaurants across cities. TripAdvisor travelers' choice.
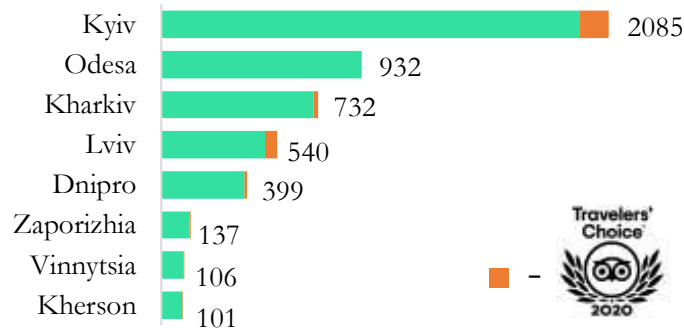
Talking about the distribution of restaurants with different price levels, we may observe the tendency that Lviv has the highest percentage of cheap restaurants comparing to another cities. Important to mention that restaurant defines on his own its price level and the price difference between cheap and middle can be minor.

Figure 3. Distribution of cheap, middle prices and expensive restaurants across cities
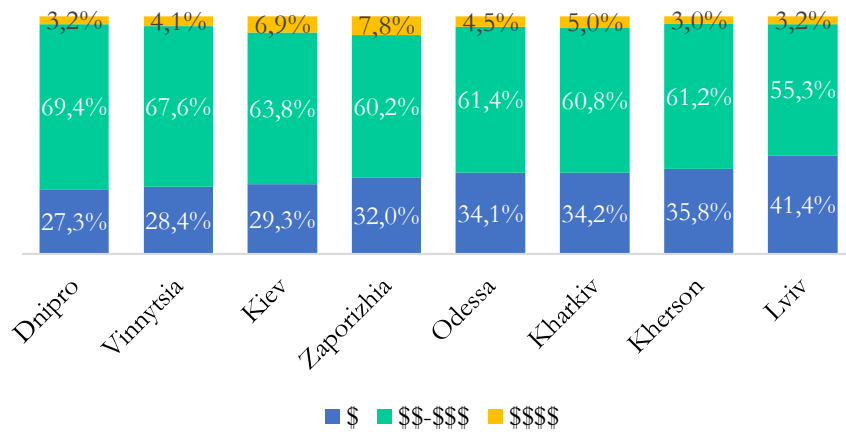
Now let us look, how proportion of reviews for cheap, middle priced and expensive restaurants changed in time. In 2011-2013 30-40% reviews was for expensive restaurants then it started to decrease with several turn back in summer 2014 and starting from 2015 more than 90% of the reviews is written to cheap and middle priced restaurants.

We are not judging the number of expensive restaurants, but only the fact that the reviews proportion decreased. According to Figure 6. the highest proportion non-Ukrainian and non-Russian reviews refers to expensive restaurants, one possible reason is that after the revolution in 2013, after war starting in Donbass and Crimea annexation foreigners could decline their trips to Ukraine due to instability, danger. On the other hand with the crisis start some expensive restaurants could shut down and be replaced by more affordable, so with more middle priced and cheap restaurants more possibility to visit and write the reviews.

Figure 4. The number of reviews for Kyiv restaurants per week across cheap, middle priced and expensive restaurants
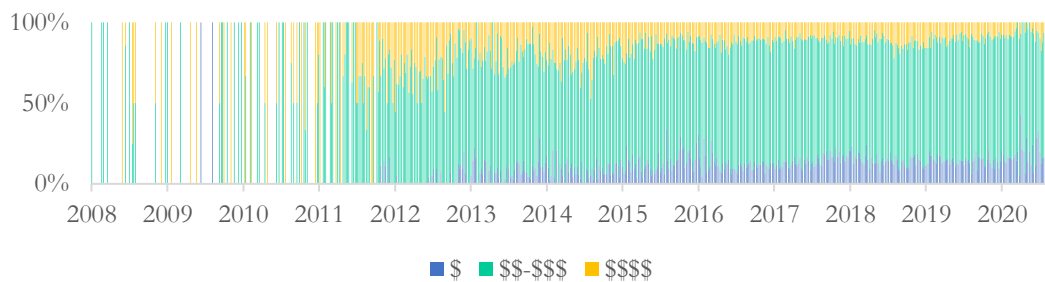
We classified reviews for Kyiv restaurants into two groups – Ukrainian and Russian reviews and all other languages. Notice that, we are not talking about nationalities, but only review language. The share of "other languages" could be even higher due to wide spread of Russian language in other countries. It turns out that first years of TripAdvisor in Ukraine mostly foreigners used it due to absence of Ukrainian or Russian language among reviews. Since people from abroad are more familiar with TripAdvisor, we could say that they rely on restaurant ranking on TripAdvisor. Steep decline of "other languages" reviews are expected due to closed country borders.

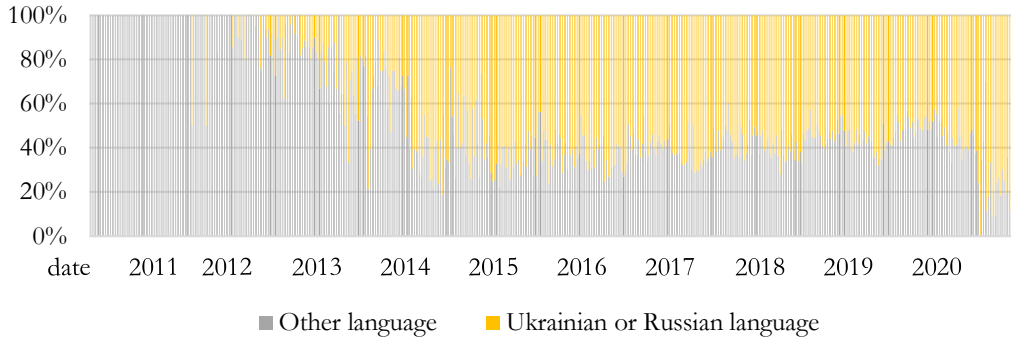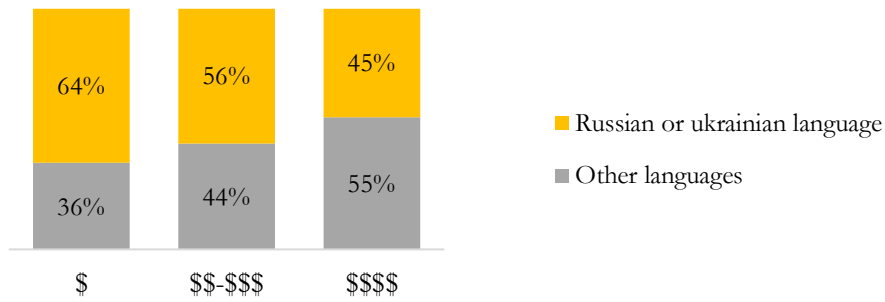Figure 5. Distribution of reviews languages across time

Figure 6. Proportion of reviews for Kyiv restaurants across languages

Important TripAdvisor distinction from Yelp[12] is that TripAdvisor does not have a widely used "star system", but its unique bubble rating system. The biggest disadvantage for us, as researchers, is that these bubbles cannot be collected as a numbers, because actually they are pictures. One more complicating detail is that the cuisine type is written by a restaurant but is not chosen from the existed sole list. The possible types of cuisine are so diverse that we are not going to rely on manually classified data by cuisine type.

In order to make some steps to understand the logic of TripAdvisor ranking we built regressions to find patterns. Firstly, we wanted to check the hypothesis that restaurants that are more expensive have more chances to have better ranking number and

---

[12] https://www.yelp.com, analogue of TripAdvisor.

for this reason, we started from the simple linear regression for the entire sample. In general (and ceteris paribus) price review number, rating, existence of certificate of excellence positively affect the ranking number. Then we get the increase in price level from cheap to middle lead to increase the probability of better ranking and this coefficient is significant. Nevertheless, switching from middle price restaurant to expensive will not increase the probability to have a better ranking number because the coefficient is not significant.

Review data was collected separately using Octoparse[13] web scraper tool and it includes reviews only for Kyiv due to convincing diversity and number of restaurants, which almost the same as seven remained restaurants. We collected 54 384 reviews, date of the reviews for further retrospective analysis, name or the reviewer, type of the device (mobile phone or laptop).

On Figure 7. we could see how mobile phones almost displaced laptops or stationery computers in the review writing, which is not surprising due to technology development. During the quarantine, people wrote reviews from their laptops more often, because of easiness to access them at home.

Figure 7. Mobile versus Laptop in reviews' writing
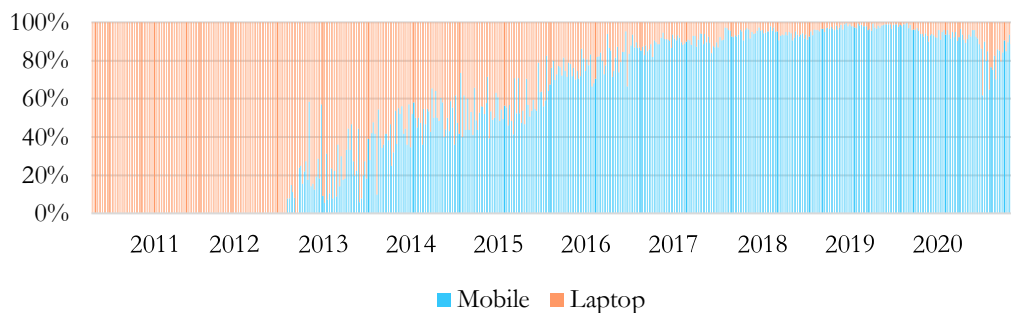
Our review data follows Zipf's law, but not as steep as expected. The graph is shown in appendices. We have a very steep fall in word frequency.

---

[13] https://www.octoparse.com/

CHAPTER 5. RESULTS

5.1. The most popular drinks and dishes in the reviews for Kyiv

Making a general picture of customer preferences goes in line with the names of dishes and drink, which they comment. On the word clouds below we may observe these products, bigger size of the word means higher frequency. Table 2. Shows top 7 dishes and drinks together with associations for them and most frequent related adjectives.

Figure 8. Most frequent dishes and drink mentioned in reviews for Kyiv restaurants.

The leaders in times mentioned in the reviews in "beer" and then "wine". This can be explained in the next way: if person order several dishes, he or she can use word "food" to express the attitude, but we cannot replace "beer" by some more general word. One of the most frequent adjective for "salad" is "Caesar" as food experts claimed about recent trends: old food ideas are also in demand, Ukrainian and Italian cuisine and of course, the everlasting "Caesar"[14]. Cocktails are associated with alcohol and with hookans.

---

[14] https://www.epravda.com.ua/rus/columns/2019/06/18/648777/

Table 2. Top seven most frequent dishes and drinks in reviews for Kyiv restaurants

Word	Freq.	Association (correlation)			Adjectives	
Beer	7578	pub (0.38)	bar (0.31)	enjoy (0.21)	cfart	delicious
Wine	5891	list (0.29)	glass(0.28)	bottle (0.26)	bottle	Georgian
Meat	5215	steak (0.17)	beef (0.15)	piece (0.15)	delicious	grill
Salad	4931	order (0.15)	sauce (0.14)	tomato (0.14)	Caesar	delicious
Pizza	4293	italian (0.16)	delivery (0.14)	pasta (0.14)	delicious	order
Dessert	3378	cream (0.11)	cake (0.1)	coffee(0.1)	delicious	-
Cocktail	3231	bar (0.19)	alcohol (0.11)	hookah (0.07)	delicious	-

5.2. Georgian, Asian and other national cuisines as a trend in Kyiv

One of the most frequent adjective for wine is "Georgian". What is interesting, Georgian dishes – khinkali and khachapuri take 13th and 21th positions respectively among dishes, being more popular than borshch! Georgian chacha is 14th among drinks. Sushi are 11th in this list. This facts proves the assumption of a trended developing different national cuisines.

5.3. Ukrainians become more service oriented

What is more important "food" or "service" for the restaurant success? This question is one of the most discussed on retail market. For sure, both vectors are important, but let have a look how visitors of Ukrainian restaurants attend these statements. According to Figure 9. Starting from 2014 service and food go almost in line, comparing to the previous period when domination of "food" felt. It is clear that people start to talk more about staff, waiters. Restaurant owners should take into account the fact that staff quality becomes more important. "Prices" become less interesting to mention in the reviews, perhaps customers shift their concerns.

Figure 9. "Food", "service", "staff" and "price" in terms of reviewer interests

5.4. Coffee culture development in Kyiv restaurants

Drinking coffee becomes a daily ritual for many Ukrainians. The coffee quality becomes higher with make this more attractive. To prove this we grouped words about coffee, e.g. cappuccino, latte, Americano, removed seasonality and find out how its presence in reviews changed in time. On Figure 10. Except of increase from 2014 we observe its domination comparing to "tea" mentioning in the reviews. The rise of coffee consumption is supposed to be more dynamic due to high penetration of small coffee shops, the data of which we do not have.

Figure 10. "Coffee" and "tea" in terms of reviewer interests
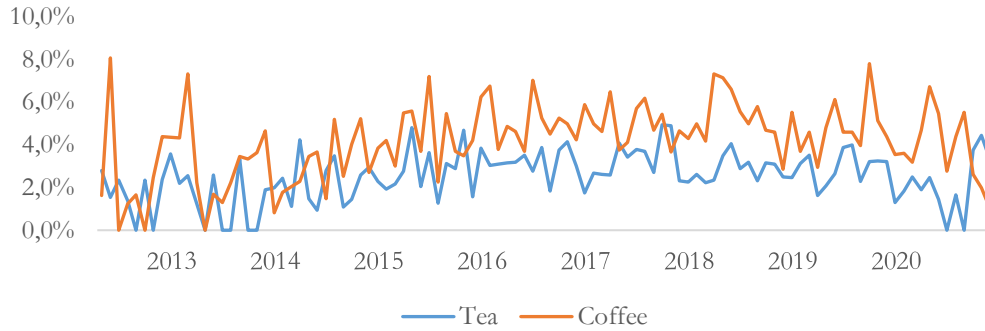
5.5. The more impressions – the longer the review

This statement could be expected, but we have also checked it based on our data. The stronger impressions will be, the higher change that customer will describe his attitude in more details. The longer the review – the more information could the restaurant extract and the more possibility for better understanding what is good and what is bad to pretend negative experience in the future. Important to mention that trend line for negative polarity is steeper.

Figure 11. Relationship between the review length and polarity.
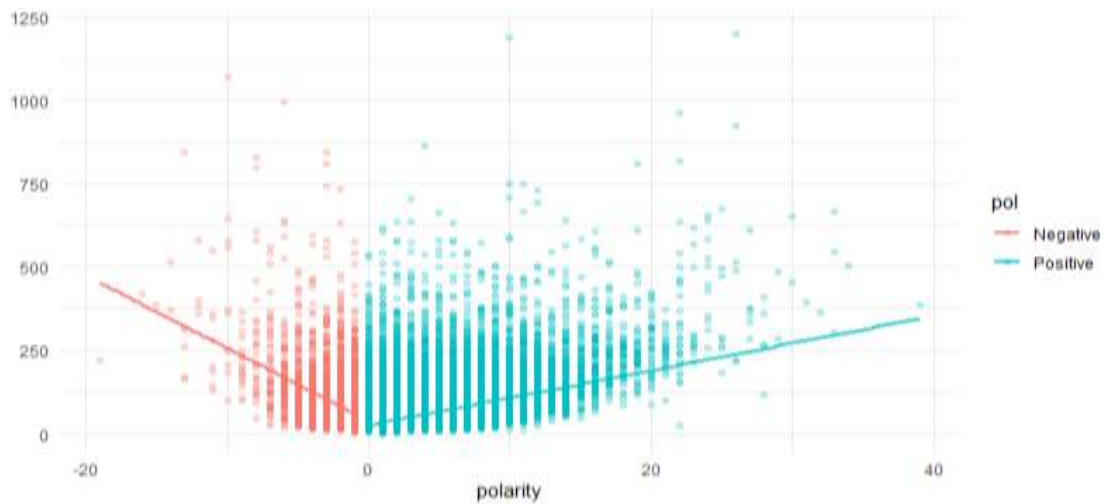
5.6. Cheap restaurants are also good restaurants

There is a stereotype that cheap cafes have low food quality and bad service due to lower investment and lower turnover. Using "bing" lexicon lexicon for polarity analysis we build a density plots comparing polarity scores for cheap, middle priced and expensive restaurants. We see that they possibility to have the same polarity score for these three groups is almost equal. Cheap restaurants have a very little more dense negative polarity, but it is very close to middle priced restaurants.

Figure 12. Bing score Densities for different price levels restaurants
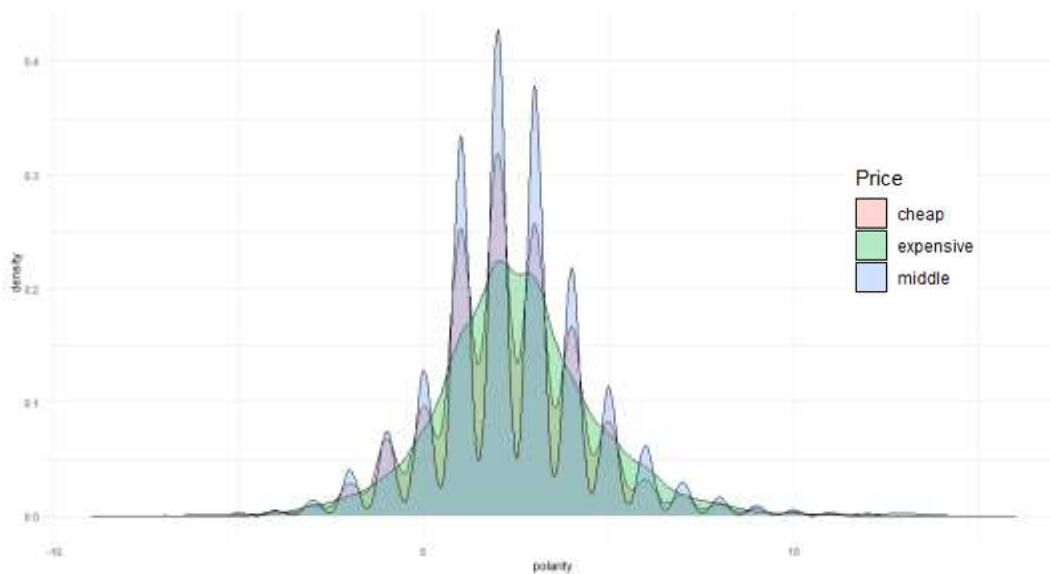
5.7. Most popular combinations of words

Fortunatelly "delicious food" and "friendly staff" are the most frequent pairs of words. Hight frequency also have description about atmosphere – "cozy atmosphere", "pleasant atmosphere", "nice atmosphere". The good signal that "highly recommend" is also a common phrase. Quite often people mentioned that table should be booked. Visitors of Ukrainian restaurans value their time, due to hight frequency of mentioning time

duration. What was unexpected is that more than 1000 reviews customers talks about live music. It may be a considerable motivation to write a review on TripAdvisor.

Figure 13. Word Network for most frequent pairs of words.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

Nowadays restaurants experiencing difficult times of change and adaptation to new realities due to coronavirus pandemic. First, it means that foreign tourists cannot visit Ukrainian restaurants, but even for Ukrainians the access is restricted. Now, more than ever, it's key for restaurants to pay close attention to Voice of Customer to improve the customer experience.

Since we don't have an access to restaurants financial data, we provided our research based on the leading website for travelers and citizens, who need to find a hotel or restaurant – TripAdvisor. The target for our analysis was restaurants details of eight large cities – Kyiv, Odessa, Lviv, Dnipro, Kharkiv, Vinnytsia, Kherson and Zaporizhia. Text mining and sentiment analysis has been done on the reviews for Kyiv restaurants.

We provided a descriptive statistics for 5034 restaurants in Ukraine, Find out the tendency that Lviv has the highest percentage of cheap restaurants comparing to another cities. Then we compared how proportion of reviews for cheap, middle priced and expensive restaurants changed in time: In 2011 about 30-40% reviews on TripAdvisor was about expensive restaurants and then this share started to decrease and from 2015 more than 90% of the reviews is written to cheap and middle priced restaurants.

We collected 54 384 reviews about Kyiv restaurants and translated all of them to English. For convenience we classified all the reviews to "Russian-Ukrainian language" and "other languages". From 2008 to 2012 almost all reviews were written in another from ukrainian and russian languages. It could mean that first years TripAdvisor were used mostly by foreigners, because they were already familiar with this website.

Starting from 2013 we observe how mobile phones almost displaced laptops or stationery computers in the review writing, which is not surprising due to technology development.

After data cleansing and preparation we provided a text mining and then a sentiment analysis on all reviews. Top most frequent mentioned in the reviews dishes and drinks were describing together with their related adjectives and associations.

Beer and wine mentioned the most frequent times not only among drinks but dishes at all. "Cocktails", which takes the 7th place among top dishes and drinks, is quite often in the reviews with "hookah". The most popular adjective for "salad" turns out a "Caesar".

Starting from 2014 service and food go almost in line, comparing to the previous period when domination of "food" felt. It is clear that people start to talk more about staff, waiters. Restaurant owners should take into account the fact that staff quality becomes more important. "Prices" become less interesting to mention in the reviews, perhaps customers shift their concerns.

Coffee trend is not the new one, but it still develops very dynamically. Recent years "coffee" is mentioned more often than a "tea.

We checked the assumption that with the higher level of impressions customer will write a longer review. In our review data the speed of increasing the negative polarity with the same review length is higher relative to positive polarity. We found the reviews with the highest positive polarity and highest negative polarity. Positive ones is really long and includes a detailed information about the positive experience starting from the staff and ending with the food taste and quality. But as for the most negative review, it is also long enough but it is negative not due to the service of food, but because of the group of rude people, who almost started a fight with him and his mother. The only one comment witch was concerned to the restaurant responsibility was the fact that the fact did not react on

this situation. The main lesson of this review is that bad situations may happen independently to the primary restaurants obligations, but the restaurant is responsible for everything, what happens inside the restaurant.

By analyzing and getting insights from customer reviews of Kyiv restaurants on TripAdvisor we come to several insights, which could help restaurant owners to see a big picture and real estate on Ukrainian restaurants.

The nearest perspective strategy is to concentrate on Ukrainians preferences. Since nobody know, how long the pandemic and closed country borders could be we advise to apply all forces to local citizens' satisfaction.

As we know since March about 30% of restaurants had to shut down. But in line with this many restaurants are opened every month. It means that with a drastic change in economic situation, trends and preferences also change very dynamically. There are so many negative news around; people want to get positive emotions from somewhere. That's why interesting, fresh concepts with especial care about customer win.

The key segment of customers, which drives the growth and help to retain turnover is Millennials. As discussed earlier Millennials are young people from 18 to 40 years old, who has already started to earn their own money and perceive dinning out not only as a physiological needs satisfaction, but the way to relax, share experience, enjoy etc.

We showed that in reviews people pay more and more attention to impressions about staff, atmosphere than before. Which approves that giving the better service will finally bring great results. In general, before the crisis started people gradually decreased the attention to prices in the review text. During the quarantine percentage of reviews with prices description increased, so they become more sensitive to price changes.

We summarized several scenarios, which are trended now and can be successfully embodied in life. Some of them are easier to implement, because they do not entail great

investment, which is very good in the unstable economic situation and some are about starting new trendy food project.

6.1. Live music during the meal

We were surprised that so many people shared in the reviews their emotions about live music. It means that it can be very impressive and inspire to share about this option with friend or in the social media. As we remember, Millennials are exactly the people, who feels the necessity to be recognized in the social media.

6.2. Implementing weeks of national cuisines

If opening a new restaurant of trending cuisines, as Georgian or Asian is not the option, but people still demand Khachapuri, Khinkali or sushi, several most popular dishes can be implemented as "chef menu" or "seasonal menu". "Khachapuri" and "Khinkalli" mentioned more frequent than Ukrainian dumplings in the reviews.

6.3. Old food ideas are also in demand

Text mining shows that the most frequent mentioned dishes are well known and everlasting Caesar salad, meat stakes, cakes, which is preferred by people despite the new, untasted before dishes. Fair to order a new dishes which will turn out to be not tasty motivates people rely on already known dishes.

In Ukraine there is a huge lack of open researches about restaurant industry. There are several researches, but they are not freely available, so we cannot judge its expertise. The field for further work is very wide. On one hand, deep sentiment analysis should be provided for all Ukrainian cities, every review is a single story witch is better to analyze together and separately from the whole data. In this work we provided a quite generalized research, but it could be much deeper. We suppose this work to be a first motivation for the further analysis of Ukrainian restaurants.

REFERENCES

An appetite for change. Key trends driving innovation in the restaurant industry. 2016. KPMG.

Linkfluence. 2018. What Millennials eat: How social data pushes innovation. https://www.linkfluence.com/hubfs/Ebooks/What%20Millennials%20Eat%20(1).pdf?hsLang=en

Worls Bank. Statistical information: https://datacatalog.worldbank.org/

Eve Turov. 2015. A Taste of Generation Yum: How the Millennial Generation's Love for Organic Fare, Celebrity Chefs and Microbrews Will Make or Break the Future of Food. *The Atlantic* (August).

Hirniak L.I., Glagola V.A. 2018.Current state, perspectives and trends of restaurant Business development in Ukraine. *Economy and enterprise management 16*

Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. *In Association for Computational Linguistics* (July)

Peter D. Turney. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *In Association for Computational Linguistics 40* (July): 417-424. https://www.researchgate.net/publication/248832100_Thumbs_Up_or_Thumbs_Down_Semantic_Orientation_Applied_to_Unsupervised_Classification_of_Reviews

Raja Irfan Sabir, Muhammad Irfan, Naeem Akhtar, Muhammad Abbas Pervez and Asad ur Rehman. 2014. Customer Satisfaction in the Restaurant Industry; Examining the Model in Local Industry Perspective. *Journal of Asian Business Strategy*, vol. 4, issue 1 (June): 18-31. http://www.aessweb.com/pdf-files/2-104-4(1)2014-JABS-18-31.pdf

Boya Yu, Jiaxu Zhou, Yi Zhang, Yunong Cao. 2015. Identifying Restaurant Features via Sentiment Analysis on Yelp Reviews. https://arxiv.org/ftp/arxiv/papers/1709/1709.08698.pdf

Theresa Wilson, Janyce Wiebe, Paul Hoffmann. 2004. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. *Computational Linguistics* 35 (September)

State Statistics Service of Ukraine. Statistical information. http://www.ukrstat.gov.ua.

Victoria Svensson. The Effect of Demand Density on Firm Productivity in the
    Restaurant Industry. *Stockholm School of Economics.*
    http://arc.hhs.se/download.aspx?MediumId=4524

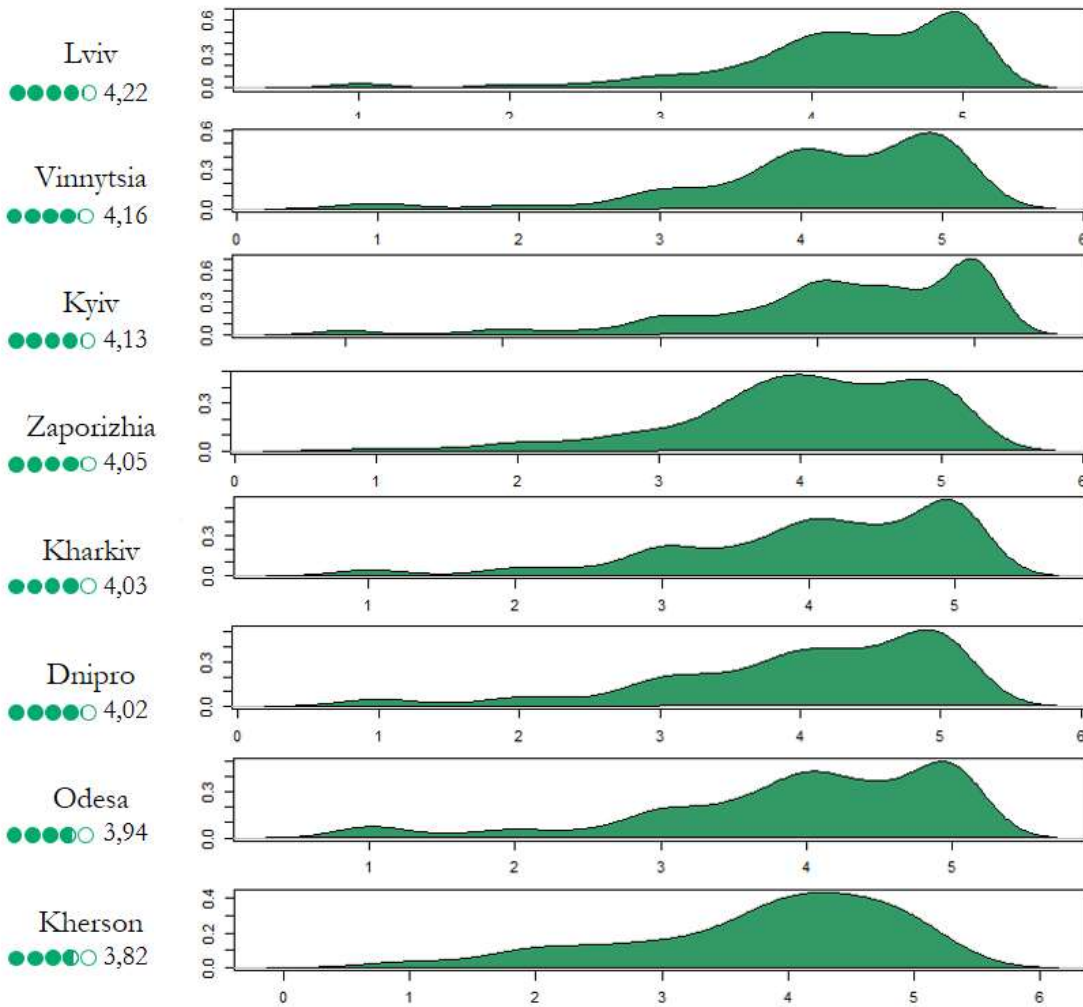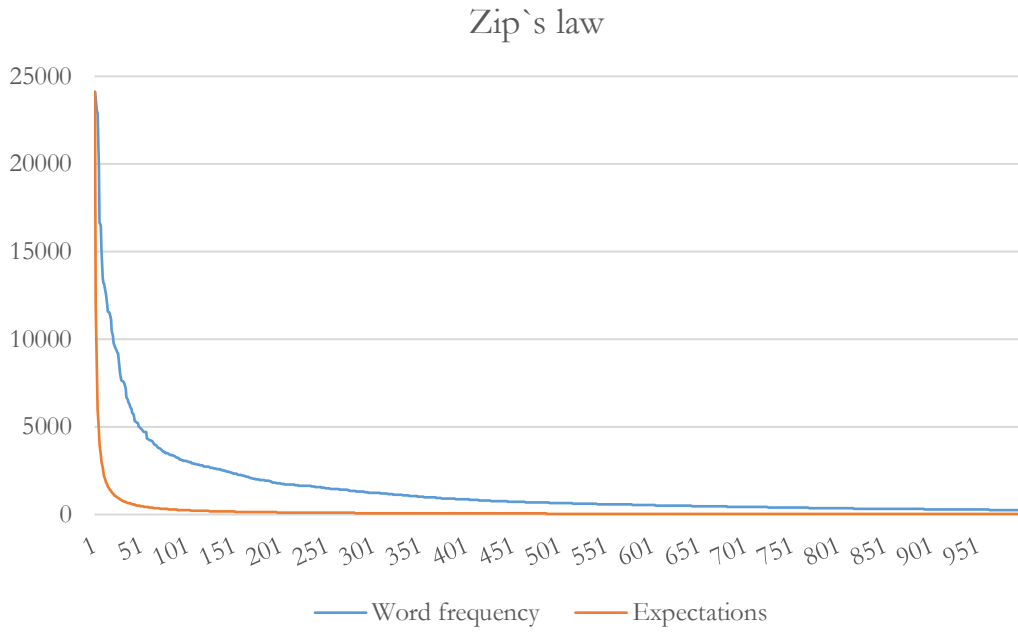Figure 1. Density plots of Travelers rating across cities

Figure 2. Zipf's law

Zip`s law

Figure 3. Sentiment word frequency