Selcuk Journal of Agriculture and Food Sciences SJAFS

http://sjafs.selcuk.edu.tr/sjafs/index

(2022) 36 (2), 146-165

e-ISSN: 2458-8377

DOI:10.15316/SJAFS.2022.021

Research Article

Defining and Predicting Consumers' Bread Choices based on Socio-Demographic Characteristics and Healthy Living Orientations

Ferit COBANOGLU[1,*], Ayse Demet KARAMAN[2], Serdal OGUT[3]

[1]Aydin Adnan Menderes University, Faculty of Agriculture, Department of Agricultural Economics, Aydın, Türkiye
[2]Aydin Adnan Menderes University, Engineering Faculty, Department of Food Engineering, Aydin, Türkiye
[3]Aydin Adnan Menderes University, Graduate School of Health Sciences, Aydin, Türkiye

ARTICLE INFO

Article history:
Received date: 21.04.2022
Accepted date: 27.05.2022

Keywords:
Behavioral Attitudes
Bread Consumption
Bread Types
Consumer Awareness
Healthy Live Style
Whole Grain Bread

ABSTRACT

The bread still constitutes an essential part of meeting the daily nutritional requirement. This study, it is tried to define and predict the behavioral attitudes of consumers toward bread consumption. The primary purpose of this study is to examine the socio-demographic characteristics of individuals and their bread consumption preferences and perceptions with a focus on quality and healthy life. The data are obtained digitally with 9583 consumers from survey questions created on google form throughout Turkey throughout the year 2020. First, descriptive statistics of the answers to the comprehensive survey questions are revealed. Then, correlation analysis was conducted to show the relationships between the variables, in line with each sub-objective set up under the primary purpose. Finally, by using some data mining and machine learning algorithms, it is tried to reveal future expectations based on some main determinant parameters such as health and quality expectation for bread consumption of consumers. All statistical and inferential evaluations were made using the Konstanz Information Miner (KNIME) analytics platform. The findings discover that increasing income and education level in bread consumption improves the level of awareness about health and quality of life in people, and these attitudes and perceptions bring about a proportional decrease in bread consumption of individuals. On the other hand, as a vicious circle, it reveals that the sedentary life brought by work-life in modern society causes an increase in body weight and an unhealthy living ecosystem. The findings can offer meaningful advice to all actors and political decision-makers involved in the bread market and supply chain and consumers.

1. Introduction

Bread is a historical and traditional food that has been proven to be used more than 5000 years ago in Mesopotamia and Ancient Egypt (De Boni et al., 2019). Today, with the increase in population, bread consumption also shows steady growth. On the other hand, it is seen that there has been a significant decrease in bread consumption in recent years, especially in developed countries (Eglite and Kunkulberga, 2017). While the average per capita bread consumption is 70 kg globally, this amount is 59 kg in European countries. While the lowest levels are in question in Great Britain and Italy with approximately 31 kg, the consumption of bread per capita is 95 kg in Bulgaria and 104 kg in Turkey (Eglite and Kunkulberga, 2017; De Boni et al., 2019). While bread consumption per capita is only 10 kg in African countries, it is 118 kg in Middle Eastern countries (Cagri, 2016).

In addition to the socio-demographic characteristics and consumption habits of the consumers, the product characteristics and variety, the type of grain from which it is obtained, the product processing and production system, the possible effects of bread on human health are very important (Sandvik et al., 2014; Sajdakowska et al., 2019).

The share of bread (about 11%) in consumers' food and beverage expenditures is still significant (WB, 2019). However, almost half of the daily average calorie need of Turkish people is still met by bread. Turkey entered the Guinness World Records with the title of "highest bread consumption per capita" in 2000 (GWR, 2020). Due to the reasons that are tried to be explained, it is observed that bread is one of the most popular food

* Corresponding author email: *ferit.cobanoglu@adu.edu.tr*

products in Turkey. On the other hand, it has been established by most of society, not only as a beloved product but also as an essential sacred value (Sarica et al., 2021).

This study tried to define and predict the behavioral attitudes of consumers toward bread consumption. Therefore, consumers' most distinctive socio-demographic characteristics and the most specific attitudes and behaviors related to bread consumption are selected. To achieve as clear results as possible, descriptive statistics such as arithmetic mean and standard deviation and classical statistical methods such as correlation analysis are used. In addition, one of the crucial facts that make the study superior is the use of decision trees and naive Bayes algorithms, which are data mining methods. Thus, a more straightforward definition and estimation are possible.

After the introduction section, the methodology used is represented. In the following stage, descriptive statistics of the variables, correlation analyses, and the results of various machine learning algorithms are presented. In each process, the nodes used in the Konstanz Information Miner (KNIME) (2021) are first introduced, and then the main findings are presented. Finally, the conclusion is presented. Therefore, as far as we know, this research is the first to examine the socio-demographic characteristics of the individual and their bread consumption preferences and perceptions using this method.

2. Materials and Methods

Sample and Data Collection Process

The data are obtained in the digital environment with 9583 consumers from survey questions created on google form throughout Turkey throughout the year 2020. A survey was conducted with 9583 consumers in 81 provinces, and most consumers are concentrated in Aydın, Antalya, Hatay, Muğla, Kocaeli, Mersin, Eskişehir, Van, Tekirdağ, Gaziantep (Figure 1).

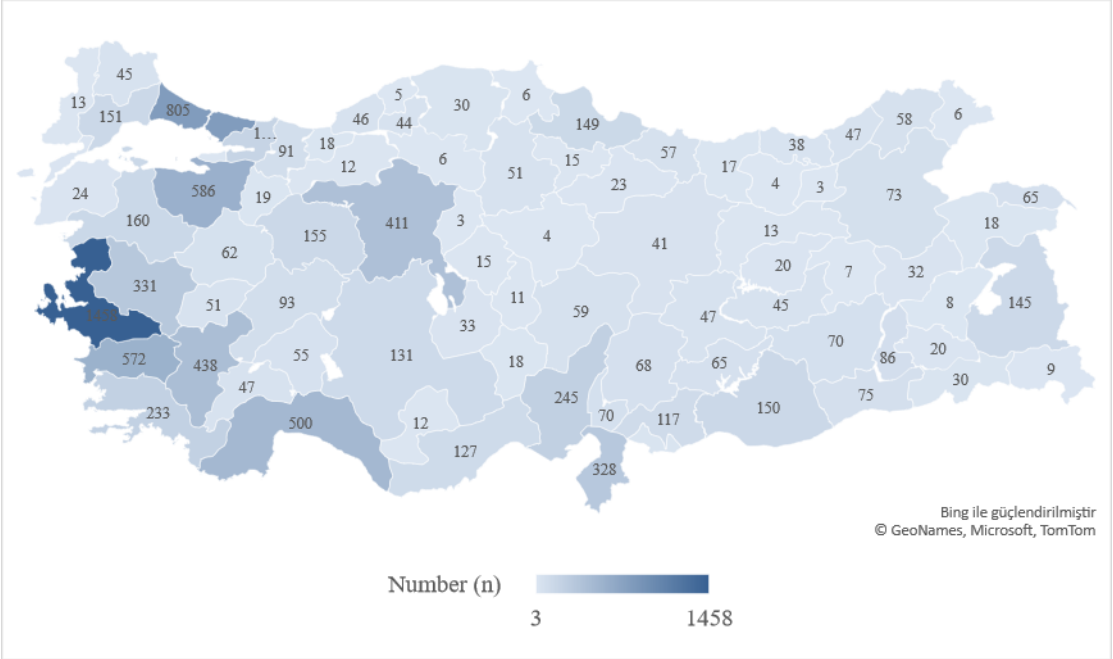


Figure 1
Research provinces in Turkey

Description of Questionnaire

The subjects covered in the study are built on certain main bodies, and the questions in the survey study are formed within this framework.

(1) Some personal and socio-demographic characteristics of the participants.

(i) Age (kg)

(ii) Gender (1 = male, 0 = female)

(iii) Marital status (1 = married, 0 = otherwise)

(iv) Height (cm)

(v) Body weight (kg)

(vi) Education (0 = illiterate, 1 = primary education, 2 = high school, 3 = Bachelor's degree, 4 = postgraduate)

(vii) Working status (1 = yes, 0 = otherwise)

(viii) a question: Do you have a disease diagnosed by a doctor? (1 = yes, 0 = otherwise)

(ix) a question: Do you exercise regularly? (1 = yes, 0 = otherwise)

(x) Income rate (₺/month) (1 = up to 750, 2 = 751-1500, 3 = 1501-3000, 4 = 3001-5000, 5 = 5001-10000, 6: over 10000) [Note: According to the data of the Organization for Economic Cooperation and Development (OECD), 1 $ has been calculated as approximately 7 ₺ in 2020 (OECD, 2021)].

(2) Various characteristics, perceptions, and attitudes of consumers regarding their bread consumption preferences.

(i) a question: Do you think bread is a portion of healthy food? (1: yes, 0: otherwise)

(ii) a question: Which of the following types of bread do you think is the healthiest? (1: yes, 0: otherwise)

- Grain bread (mixed, grain, wholewheat, rye, oat, corn)
- White bread, pita, lavash (wheat flour)
- Sourdough, dough bread (wheat, wholewheat, grain)
- Wholewheat bread, pita, lavash (wholewheat flour)

(iii) a question: Why do you think additives are used in bread making? (1: yes, 0: otherwise)

- Increase the durability of bread
- Increasing the nutritional value
- To give aroma and flavor
- Correcting and improving the appearance
- Delay staling

(iv) a question: What kind of health effects do you think the additives used in bread making have? (1: yes, 0: otherwise)

- Causes allergies
- Causes obesity
- Causes other health problems
- Causes an increase in cholesterol
- Does not cause health problems

(v) a question: How do you know when the bread is stale (too much to be consumed)? (1: yes, 0: otherwise)

- If the bread is crumbly
- If the bread is moldy
- If the bread is hard
- If the bread is hard and crumbles easily
- If the appearance of the bread is distorted
- If the appearance of the bread has changed
- If it is not consumed on the day, it is taken
- If the smell of the bread has changed
- If the taste of the bread has changed

(vi) Attitudes, ideas between bread, and some health-related features. (1 = I totally disagree and 5 = I totally agree)

- Wholewheat bread should be preferred instead of white bread to reduce the risk of obesity, heart, and chronic diseases
- Bread strengthens the immune system
- Bread regulates the digestive system
- Wholewheat bread is healthier
- White bread is unhealthy
- Eating bread makes you gain weight
- Overeating bread makes you gain weight
- Bread is a source of energy
- Consuming bread keeps you full
- Additives used in bread making are harmful to our health

(vii) a question: Where do you often buy your bread? (1: yes, 0: otherwise)

- Bakery
- Grocer
- Buffet
- I do it myself at home
- Municipal outlet
- Market

(viii) a question: Which type of bread do you often consume? (1: yes, 0: otherwise)

- White bread
- Pita
- Lavash (with wheat flour)
- Sourdough/dough bread (wheat, wholewheat grain)
- Grain bread (mixed grain, wholewheat, rye, oat, corn)
- Wholewheat bread

(ix) a question: What are the points you pay attention to when buying bread? (1: yes, 0: otherwise)

- Being close to reach
- Size
- Packaging
- Company name
- View of bread
- Producing quality bread
- Sellers are smiling
- Additive in bread
- The calorie level of bread
- Easy to find
- To be salt-free
- Being hot and fresh
- To be well cooked
- Being cooked in a wood oven
- Produced/sold in a hygienic environment

(x) a question: What do you think about the quality of the bread you consume? (1: very bad, 2: bad, 3: average, 4: good, 5: very good)

(xi) a question: What are the reasons why you do not find the quality of the bread satisfactory? (1: yes, 0: otherwise)

- The manufacturer's disregard for quality
- Uneducated bakers
- Presence of excessive and unconscious additives in bread
- Bread not cooked well
- Bakers not paying attention to cleanliness

- Poor quality of flour used in bread production
- Using bleach in bread making
- Bread is tasteless

(xii) a question: How many slices of bread do you eat per day? (No = 0, 1 = one or two slices in a day, 2 = three to four slices in a day, 3 = five to six slices in a day, 4 = more than six slices in a day)

(xiii) a question: Do you eat bread with meals? (1: yes, 0: otherwise)

(xiv) a question: Which food or meal would you like to consume bread with? (1: yes, 0: otherwise)

- Soup
- Slops
- Meat, chicken, and rice
- Rice and pasta
- I never consume

(xv) a question: In which meals do you mostly consume bread? (1: yes, 0: otherwise)

- At breakfast
- At lunch
- At snacks
- At the dinner

(xvi) a question: How do you determine the staling of bread? (1: yes, 0: otherwise)

- Bread hardening and crumbling
- Changing the taste of bread
- Bread molding
- If not consumed that day
- If the appearance of the bread has changed

(xvii) a question: How do you evaluate stale bread? (1: yes, 0: otherwise)

- I make breadcrumbs, crusty bread, etc.
- I use it dry
- I feed the animals
- I use it as a meatball paste
- I give the bread to the milkman and the porter
- I do not evaluate it in any way. I throw it away
- I use it in soup and food
- I give to what is needed

(xviii) a question: How do you store the bread? (1: yes, 0: otherwise)

- I keep it in the freezer
- I consume it by heating
- I keep it in the fridge

3. Results and Discussion

Descriptive Statistics of the Variables

First, descriptive statistics of the variables used in the study are presented (Table 1). Then, the workflow used for this is stated below (Figure 2).

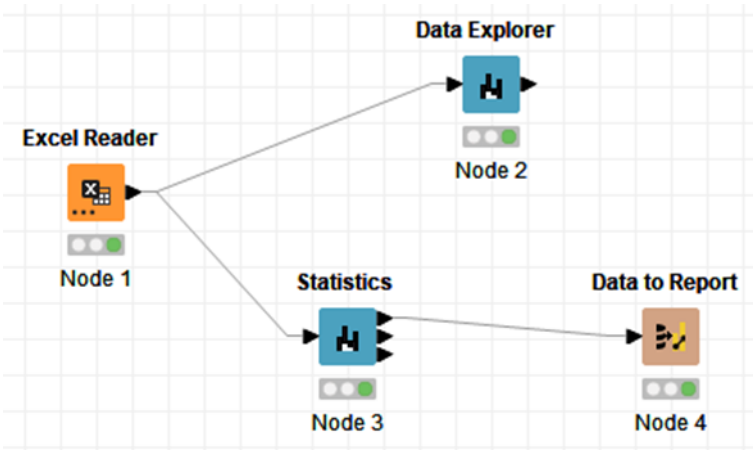


Figure 2
KNIME workflow for descriptive statistics

Table 1
Descriptive statistics of the variables/statements

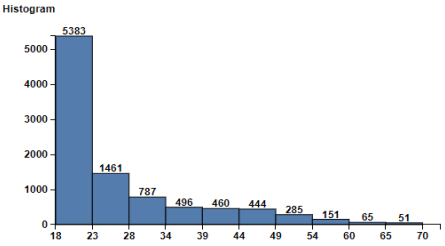
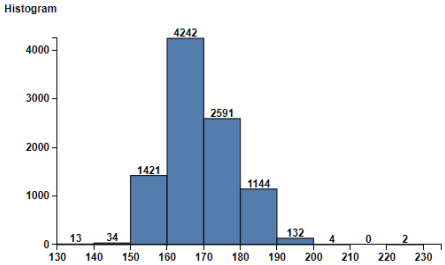
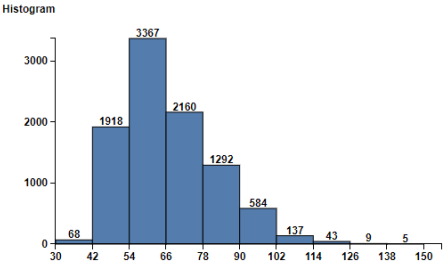
Variables/statements	Mean	Standard deviation		Frequency (n)
Age	27. 06	10.38		
Gender	0.32	0.46		0=6568, 1=3015
Marital status	0.27	0.44		0=6995, 1=2588
Height	168.03	9.13		
Bodyweight	66.21	14.81		
Education	2.75	0.64		0=34, 1=528, 2=1670, 3=6947, 4=404
Working status	0.35	0.48		0=6240, 1=33431
Do you have a disease diagnosed by a doctor?	0.14	0.35		0=8235, 1=1348
Do you exercise regularly?	0.42	0.49		0=5551, 1=4032
Income rate	3.01	1.54		1=2557, 2=1072, 3=1898, 4=2182, 5=1547, 6=327
Do you think bread is a portion of healthy food?	0.49	0.50		0=4874, 1=4709
Grain bread (mixed, grain, wholewheat, rye, oat, corn)	0.45	0.50		0=5311, 1=4272
White bread, pita, lavash (wheat flour)	0.14	0.35		0=8219, 1=1364
Sourdough, dough bread (wheat, wholewheat, grain)	0.20	0.40		0=7709, 1=1874
Wholewheat bread, pita, lavash (wholewheat flour)		0.22	0.41	0=7510, 1=2073
Increase the durability of bread		0.66	0.48	0=3307, 1=6276
Increasing the nutritional value		0.25	0.44	0=7152, 1=2431
To give aroma and flavor		0.31	0.46	0=6580, 1=3003
Correcting and improving the appearance		0.36	0.48	0=6167, 1=3416
Delay staling		0.70	0.46	0=2911, 1=6672
Causes allergies		0.03	0.17	0=9294, 1=289
Causes obesity		0.54	0.50	0=4421, 1=5162
Causes other health problems		0.21	0.41	0=7607, 1=1976
Causes an increase in cholesterol		0.10	0.29	0=8667, 1=916
Does not cause health problems		0.05	0.21	0=9148, 1=435
If the bread is crumbly		0.39	0.49	0=5865, 1=3718
If the bread is moldy		0.70	0.46	0=2898, 1=6685
If the bread is hard		0.49	0.50	0=4848, 1=4735
If the bread is hard and crumbles easily		0.45	0.50	0=5234, 1=4349
If the appearance of the bread is distorted		0.31	0.46	0=6593, 1=2990
If the appearance of the bread has changed		0.39	0.49	0=5844, 1=3739
If it is not consumed on the day, it is taken		0.09	0.29	0=8698, 1=885
If the smell of the bread has changed		0.48	0.50	0=5003, 1=4580
If the taste of the bread has changed		0.45	0.50	0=5289, 1=4294
Wholewheat bread should be preferred instead of white bread to reduce the risk of obesity, heart, and chronic diseases		3.86	0.85	1=165, 2=447, 3=1880 4=5178, 5=1913
Bread strengthens the immune system		2.69	1.01	1=1274, 2=2761, 3=3465 4=1833, 5=250
Bread regulates the digestive system		2.76	1.03	1=1157, 2=2752, 3=3202 4=2209, 5=263
Wholewheat bread is healthier		3.88	0.78	1=122, 2=485, 3=1406 4=6030, 5=1540
White bread is unhealthy		3.43	1.06	1=293, 2=2023, 3=1786 4=4197, 5=1284

Table 1 (continued)
Descriptive statistics of the variables/statements

Eating bread makes you gain weight	3.28	1.20	1=451, 2=3164, 3=644 4=3894, 5=1430
Eating too much bread makes you gain weight	4.35	0.71	1=73, 2=219, 3=228 4=4843, 5=4220
Bread is a source of energy	3.39	1.06	1=522, 2=1797, 3=1575 4=4809, 5=880
Consuming bread keeps you full	3.56	1.07	1=484, 2=1658, 3=656 4=5573, 5=1212
Additives used in bread making are harmful to our health	3.84	0.87	1=117, 2=618, 3=1988 4=4855, 5=2005
Bakery	0.55	0.50	0=4314, 1=5269
Grocer	0.15	0.36	0=8155, 1=1428
Buffet	0.01	0.10	0=9491, 1=92
I do it myself at home	0.07	0.25	0=8932, 1=651
Municipal outlet	0.02	0.15	0=9362, 1=221
Market	0.20	0.40	0=7661, 1=1922
White bread	0.46	0.50	0=5179, 1=4404
Pita	0.67	0.47	0=3147, 1=6436
Lavash (with wheat flour)	0.46	0.50	0=5178, 1=4405
Sourdough/dough bread (wheat, wholewheat grain)	0.13	0.34	0=8317, 1=1266
Grain bread (mixed grain, wholewheat, rye, oat, corn)	0.20	0.40	0=7702, 1=1881
Wholewheat bread	0.21	0.41	0=7551, 1=2032
Being close to reach	0.51	0.50	0=4654, 1=4929
Size	0.03	0.18	0=9258, 1=325
Packaging	0.04	0.19	0=9244, 1=339
Company name	0.01	0.09	0=9505, 1=78
View of bread	0.05	0.22	0=9107, 1=476
Producing quality bread	0.10	0.30	0=8645, 1=938
Sellers are smiling	0.00	0.06	0=9549, 1=34
Additive in bread	0.03	0.16	0=9319, 1=264
The calorie level of bread	0.02	0.14	0=9391, 1=192
Easy to find	0.04	0.19	0=9212, 1=371
To be salt-free	0.01	0.07	0=9533, 1=50
Being hot and fresh	0.09	0.29	0=8719, 1=864
To be well cooked	0.06	0.24	0=8977, 1=606
Being cooked in a wood oven	0.03	0.17	0=9296, 1=287
Produced/sold in a hygienic environment	0.05	0.22	0=9096, 1=487
What do you think about the quality of the bread you consume?	3.60	0.79	1=68, 2=314, 3=4330 4=3573, 5=1298
The manufacturer's disregard for quality	0.45	0.50	0=5283, 1=4300
Uneducated bakers	0.06	0.24	0=8986, 1=597
Presence of excessive and unconscious additives in bread	0.27	0.45	0=6961, 1=2622
Bread not cooked well	0.05	0.22	0=9103, 1=480
Bakers not paying attention to cleanliness	0.08	0.26	0=8863, 1=720
Poor quality of flour used in bread production	0.07	0.26	0=8870, 1=713
Using bleach in bread making	0.06	0.25	0=8966, 1=617
Bread is tasteless	0.03	0.18	0=9261, 1=322
How many slices of bread do you eat per day?	2.05	1.12	0=386, 1=3228, 2=2832 3=1792, 4=1345
Do you eat bread with meals?	0.81	0.39	0=1838, 1=7745
Soup	0.69	0.46	0=2988, 1=6595
Slops	0.39	0.49	0=5896, 1=3687
Meat, chicken, and rice	0.17	0.38	0=7934, 1=1649
Rice and pasta	0.02	0.13	0=9420, 1=163
I never consume	0.03	0.17	0=9285, 1=298
At breakfast	0.60	0.49	0=3875, 1=5708
At lunch	0.08	0.28	0=8793, 1=790
At snacks	0.01	0.11	0=9456, 1=127
At the dinner	0.28	0.45	0=6923, 1=2660
Bread hardening and crumbling	0.45	0.50	0=5241, 1=4342
Changing the taste of bread	0.06	0.23	0=9053, 1=530
Bread molding	0.28	0.45	0=6918, 1=2665
If not consumed that day	0.09	0.28	0=8753, 1=830
If the appearance of the bread has changed	0.07	0.25	0=8944, 1=639
I make breadcrumbs, crusty bread, etc.	0.39	0.49	0=5895, 1=3688
I use it dry	0.29	0.45	0=6799, 1=2784
I feed the animals	0.60	0.49	0=3877, 1=5706
I use it as a meatball paste	0.46	0.50	0=5135, 1=4448
I give the bread to the milkman and the porter	0.09	0.28	0=8748, 1=835
I do not evaluate it in any way, and I throw it away	0.10	0.31	0=8588, 1=995
I use it in soup and food	0.37	0.48	0=6009, 1=3574
I give to what is needed	0.06	0.24	0=8996, 1=587
I keep it in the freezer	0.17	0.38	0=7964, 1=1619
I consume it by heating	0.29	0.45	0=6806, 1=2777
I keep it in the fridge	0.19	0.40	0=7725, 1=1858

Correlation Analysis

In statistical terminology, Spearman's rank correlation coefficient or Spearman's ρ, defined after Charles Spearman and often presented as {rho} in Greek, refers to the non-parametric measure of a rank correlation (statistical dependence between the rank of two variables). It gives the degree of relationship between two variables using a monotonic function. The Spearman correlation

between the two variables is equal to the Pearson correlation between these two variables; While Pearson's correlation reveals linear relationships, Spearman's correlation presents monotonic relationships (whether linear or not). In the absence of repeated data values, a perfect Spearman correlation of +1 or -1 occurs when each variable is a perfectly monotonic function. Conceptually, the Spearman correlation between two variables is high when the observations have similar (or identical for a correlation of 1) order between the two variables (i.e., the relative position label of the observations within the variable: 1st, 2nd, 3rd, etc.). is happening. Conversely, variables and observations are low when the two variables are different (or the opposite for a -1 correlation). The Spearman coefficient is suitable for both continuous and discrete ordinal variables. The p-value (two-sided) for these columns expresses the probability that an uncorrelated system will produce at least an extreme degree of correlation if the mean of the correlation is zero (Anonymous, 2021). At this stage, the nodes used in the KNIME platform are listed below (Figure 3).

First, it was aimed to reveal whether there is a relationship between the level of individuals related to bread and their demographic characteristics (Table 2).

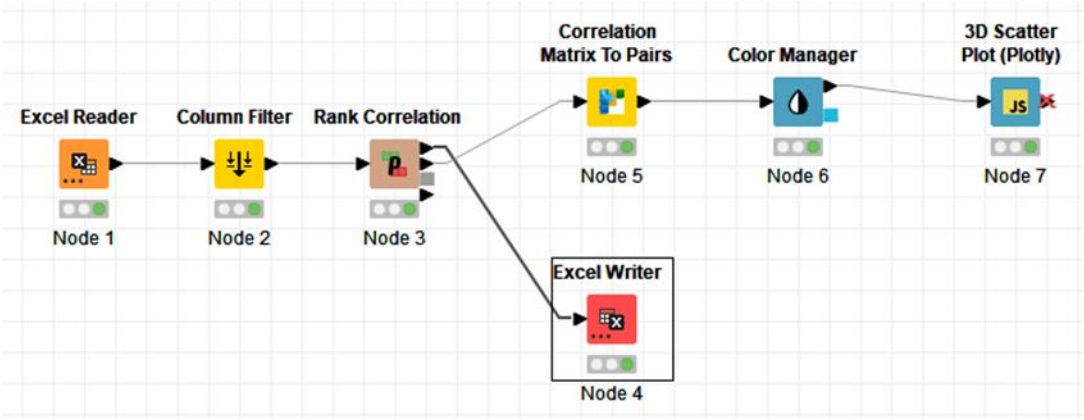


Figure 3

Workflow used in the KNIME platform for correlation analysis

Table 2

The relationship between the knowledge level of individuals about the healthiest of bread and their socio-demographic features

	Consumer opinions about the healthiest type of bread			
	Grain bread	White bread, pita, lavash	Sourdough, dough bread	Wholewheat bread, pita, lavash
Age	-0.10***	-0,09***	0,17	0,03***
Gender	-0,09***	0.11***	-0.01	0.02**
Marital status	-0.09***	-0.06***	0.14	0.04***
Height	-0.04***	0.08***	-0.02**	-0.00
Body weight	-0.07***	0.00	0.04***	0.04***
Education	0.05***	-0.05***	0.03***	-0.04***
Working status	-0.06***	0.00	0.07***	-0.00
Income rate	-0.05***	-0.04***	0.09	0.00

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

There is usually a negative relationship with a low correlation coefficient between the socio-demographic characteristics of the consumers and the bread types they think are the healthiest. It turns out that the socio-demographic characteristics of the consumers and the characteristics of the health status of various bread types differ. Adebayo et al. (2017) stated that the demands and preferences of individuals belonging to different ethnic groups residing in Finland for healthy bread and food products also differ significantly. Sandvik et al. (2017) comparatively analyzed the situation between rye bread consumption and the expectation of consuming healthier products among young and old consumers in Sweden. They revealed that the structure and color (light and dark) of rye bread effectively form health properties. Demirtaş et al. (2018) stated that white bread is widely produced and consumed in Turkey, and regional consumption is also typical. Sarica et al. (2021), in a study they conducted in Turkey, found that insufficient control of the mother and father's occupation, monthly income, and bakery products were important in bread consumption. Therefore, it can be stated that the amount of bread consumption decreases with the increase in income. Sarica et al. (2021) also confirm this inference.

The relationships between the socio-demographic characteristics of the individuals and the additive they use in bread making are also examined (Table 3). The socio-demographic characteristics and various expressions of individuals about the additives used in bread

making are mostly positive, and statistically significant, but this relationship had a low ρ coefficient. Simmons et al. (2014) and Bacarea et al. (2021) declared that highly processed foods, which are calorie-dense and rich in carbohydrates, lipids, flavor enhancers, and food additives, are becoming popular among children and adolescents. It is emphasized that this popularity also continued to increase insidiously due to the lack of targeted education of the population. Loloei et al. (2019) emphasized that individuals with a relatively higher level of education can catch the opportunity to learn at a higher level, their awareness increases, and they can better understand the harmful effects of additives that are not allowed in bread.

Table 3
Correlation relationship between socio-demographic characteristics and bread additive

Variables	Opinions on food additives used in bread making				
	Increase the durability of bread	Increasing the nutritional value	To give aroma and flavor	Correcting and improving the appearance	Delay staling
Age	0.07***	-0.09***	-0.03***	-0.01	-0.02*
Gender	-0.06***	0.02*	-0.02**	0.04***	-0.02**
Marital status	0.05***	-0.06***	-0.05***	-0.01	-0.02**
Height	-0.04***	-0.01	-0.01	0.01	0.00
Body weight	-0.01	-0.03***	-0.03***	0.02**	-0.01
Education	0.08***	-0.04***	0.07***	-0.00	0.07***
Working status	0.03***	-0.05***	-0.01	-0.01	-0.02**
Income rate	0.05***	-0.07***	-0.01	-0.01	0.03***

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

The possible effects of additives used in bread making on individuals' health and socio-demographic characteristics are examined (Table 4). It is clarified that the correlation relationship between the variables in these groups is mostly negative, and they had low correlation coefficients. It can be stated that there is an awareness among individuals that there is a relationship between the additives used in bread and various health problems, albeit at a certain level. It can be deduced that the additives in question contribute to the formation of allergies and obesity and that they are one of the crucial reasons for the appearance of cholesterol and other health problems. It is underlined that awareness of the negative effects of some synthetic food additives has fueled high consumer demand for clean-label bread products (Cappelli et al., 2020). This is also confirmed in similar studies (Montagnese et al., 2015; Vandevijvere et al., 2017).

Table 4
Interaction of bread additives with possible health effects and socio-demographic characteristics

Variables	Opinions on bread additives with possible health effects				
	It causes allergies	It causes obesity	It causes other health problems	It causes an increase in cholesterol	It does not cause health problems
Age	0.05***	-0.01*	0.03***	-0.10***	-0.00
Gender	-0.03***	-0.01	-0.02	0.02**	0.06***
Marital status	0.03***	0.02**	0.01	-0.08***	-0.01*
Height	-0.01	-0.03***	0.00	0.01	0.04***
Body weight	-0.02**	0.03***	-0.03***	-0.03***	0.03
Education	0.03***	-0.05***	0.03***	0.01	-0.03***
Working status	0.01	0.00	-0.01	-0.04***	0.02*
Income rate	0.02*	-0.01	-0.00	-0.03***	0.00

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

Stale aging occurs with the hardening of breadcrumbs, which starts during the cooling process of bread products and gradually hardens, dries, and crumbles during storage, and the crust becomes soft and leathery (Dong and Karboune, 2021). The correlation between understanding that bread is too stale for consumption and some socio-demographic characteristics is also evaluated (Table 5). Some variables are positive, some are negative, and there is no correlation between the remaining few variables. The correlation coefficients are also low among the variables analyzed at this stage. Although there is a negative correlation between age and bread being hard enough, it is defined that the correlation coefficient (ρ) was relatively higher than the other coefficients.

Table 5
The correlation between understanding that bread is too stale for consumption and socio-demographic features

Variables	Perceptions of bread's stale for consumption								
	If the bread is crumbly	If the bread is moldy	If the bread is hard	If the bread is hard and crumbles easily	If the appearance of the bread is distorted	If the appearance of the bread has changed	If it is not consumed on the day, it is taken	If the smell of the bread has changed	If the taste of the bread has changed
Age	0.02**	0.07***	-0.08***	-0.05***	-0.02**	-0.04***	-0.01	0.02**	0.02
Gender	-0.02**	-0.05**	0.03***	-0.03***	-0.03***	-0.04***	0.03***	-0.06***	-0.04***
Mrt. St.	0.02**	0.07***	-0.08***	-0.05***	-0.03***	-0.05***	-0.01	0.02*	0.02
Height	-0.03***	-0.03***	0.03***	-0.02**	-0.03***	-0.02**	0.02**	-0.04***	-0.03***
Body weight	-0.01	0.00	-0.01	-0.05***	-0.03***	-0.04***	-0.00	-0.04***	-0.02*
Education	-0.04***	0.01	0.01	0.04***	0.05***	0.06***	-0.02**	0.05***	0.03***
Working status	0.02**	0.04***	-0.02**	-0.01	0.00	-0.00	0.01	0.01	0.01
Income rate	-0.01	0.05***	-0.03***	-0.02*	-0.00	-0.01	0.01	0.01	0.01

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

As the working status and income level increase, the correlation coefficient between the perceptions of bread consumption seems to have a positive relationship. This reveals a positive relationship between the statements about the staleness of bread as the income level of the individual's increases and the opportunity to work improves. The correlation between various socio-demographic characteristics and some bread and health-related characteristics is also analyzed (Table 6).

Table 6
Correlation between various socio-demographic characteristics and some bread and health-related attributes

Variables	Bread and health-related characteristics									
	Wholewheat bread should be preferred instead of white bread to reduce the risk of obesity, heart, and chronic diseases	Bread strengthens the immune system	Bread regulates the digestive system	Wholewheat bread is healthier	White bread is unhealthy	Eating bread makes you gain weight	Eating too much bread makes you gain weight	Bread is a source of energy	Consuming bread keeps you full	Additives used in bread making are harmful to our health
Age	0.05***	-0.00	-0.01	0.04***	0.15	0.13	-0.01	-0.13***	-0.08***	0.11
Gender	-0.11***	0.03***	0.03***	-0.05***	-0.06***	-0.06***	0.08***	-0.08***	0.11	-0.08***
Marital status	0.04***	-0.02**	-0.04**	0.03***	0.12	0.13	0.00	-0.14***	-0.06***	0.10
Height	-0.07***	0.03***	0.03***	-0.03***	-0.04***	0.02**	-0.04***	0.04***	0.06***	-0.04***
Body weight	0.01	-0.00	-0.00	0.03***	0.06***	0.12	0.05***	-0.03***	-0.02**	0.02**
Education	0.05***	-0.02*	0.01	0.04***	0.03***	-0.11***	0.06***	0.07***	-0.02	0.04***
Working status	0.01	-0.03***	-0.04***	0.03***	0.08***	0.09	-0.01	-0.10***	-0.04***	0.03***
Income rate	0.04***	-0.03***	-0.01	0.04***	0.11	0.07***	0.04***	-0.06***	-0.07***	0.07***

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

Sandvik et al. (2018) state that most consumers find the bread healthy. A quarter of the participants stated that they had no idea about the relation of bread to health and that most of the consumers in this group had a low level of education. It is found that consumers' cues are more important than labels in their perception of healthy bread. The study tried to define the consumers' perceptions of bread related to health by discovering which quality attributes related to bread and health and whether there were differences in terms of age, gender, and education level. Three-quarters of consumers rank bread as healthy among the course, like whole grain, fibre-rich, sourdough, crunchy, less sugar, dark, rye, and seed. They stated that such terms are in interaction. Concepts such as trademark, homemade, and cores are also discussed at this point. Bread is perceived as healthy mainly because they contain fiber, is good for the stomach, saturates well, and has beneficial glycemic properties. One of the difficulties in defining healthy bread, especially among consumers with a low level of education, is a preference for consumption. It is suggested that many of the health effects that are important to consumers cannot be conveyed on food packages, and therefore consumers should use their clues to identify these characteristics. This is especially true if bread is labeled, for example, as sourdough bread or rye bread despite its low content. They are summarized as it may cause misdirection of consumers.

In this study, there is mostly a positive correlation between the variables discussed here, whereas the correlation coefficient is small, as in the other sections. It is revealed that there is a relatively high negative correlation of 0.11 only between gender and the variable that whole wheat bread should be used instead of white bread to reduce obesity, heart, and chronic diseases. It can be stated that this perception is especially higher in women. As the income level of individuals increases and some symptoms of diseases that differ according to various health reasons begin to appear, the perception of consumption towards bread varieties with low calories and relatively high nutritional value develops. While most consumers buy bread from bakers, grocers, and greengrocers, it is observed that the number of consumers who mark the option to make it myself at home remains very low. In 2020, when a heavy closure period was passed due to the covid-19 epidemic, it is represented that although bread making at home has increased to a certain extent, most consumers still buy bread from outside the home in Turkey.

It is defined that most individuals prefer to consume pita, white bread, lavash (with wheat flour) (Table 1). It is discovered that there are statistically significant relationships between the socio-demographic characteristics of individuals and the correlation relationship between the types of bread they frequently consume, especially between some variables (Table 7). It is observed that as age, body weight, education, and income level increase, consumption of white bread, pita, and lavash decreases, while preference for grain bread and wholewheat increases. It is observed that there is a similar development to the above trend as the working opportunities of individuals improve.

Table 7
Correlation between bread types consumed and socio-demographic characteristics

Variables	Bread types consumed					
	White bread	Pita	Lavash (with wheat flour)	Sourdough/dough bread (wheat, wholewheat grain)	Grain bread (mixed grain, wholewheat, rye, oat, corn)	Whole-wheat bread
Age	-0.15***	-0.14***	-0.15***	0.15	0.04***	0.01
Gender	0.08***	0.07***	0.08***	-0.02	-0.07***	-0.02
Marital status	-0.09***	-0.08***	-0.09***	0.10	0.01	0.02**
Height	0.03***	0.02**	0.03***	-0.03***	-0.00	-0.01
Body weight	-0.06***	-0.04***	-0.06***	0.03***	0.02**	0.03***
Education	-0.07***	-0.07***	-0.07***	0.03***	0.06***	0.00
Working status	-0.07***	-0.06***	-0.07***	0.04***	0.04***	0.01
Income rate	-0.10***	-0.09***	-0.10***	0.07***	0.05***	0.01

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

At this stage, although it is defined that there are statistically significant, both negative and positive correlation relationships among many variables, the high correlation coefficients between age and income level and the types of bread consumed reveal that the difference is more visible. It is explored that there is a negative correlation between the age of the consumer and the types of bread consumed. This is also an expected situation. As the age of the individual increases, it appears that there is a significant decrease in the amount of consumption of bread types (Figure 4).

These inferences, Demirtaş et al. (2018), Sarica et al. (2021), and Dong and Karboune (2021), are also like the findings and literature reports.

The interaction between the socio-demographic characteristics of consumers and the points they pay attention to when buying bread is examined below (Table 8).

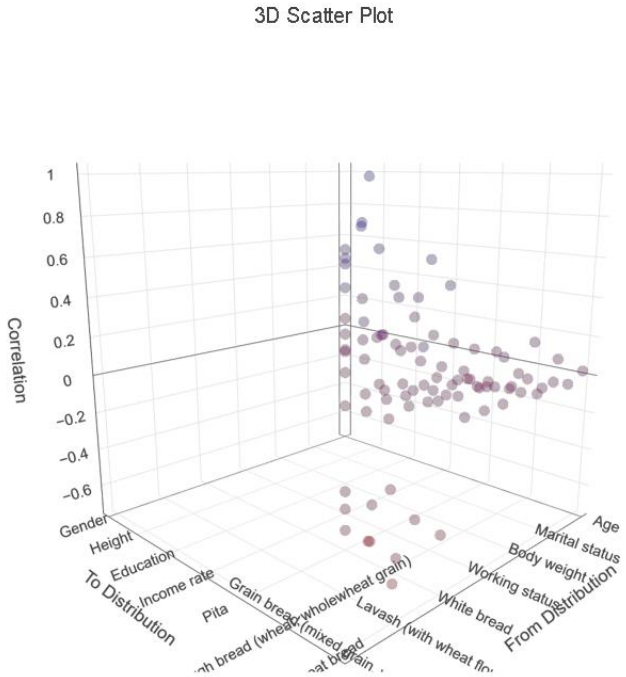


Figure 4

Scatter plot graph for correlation values intended for bread types consumed and socio-demographic features

Table 8

Relationship between considerations when buying bread and socio-demographic characteristics of the consumers

Variables	Socio-demographic characteristics							
	Age	Gender	Marital status	Height	Bodyweight	Education	Working status	Income rate
Being close to reach	-0.09***	0.05***	-0.06***	0.03***	0.06	0.00	-0.04***	-0.05***
Size	-0.02**	-0.01	-0.03***	-0.01	-0.01	-0.01	-0.01	-0.02***
Packaging	0.04***	-0.02**	0.03***	-0.02**	0.01	-0.03***	0.02**	0.00
Company name	0.00	0.00	-0.01	0.00	0.00	0.02*	0.02*	0.01
View of bread	-0.03***	-0.01	-0.02**	-0.01	-0.02***	0.02**	-0.02	-0.01
Producing quality bread	0.01	-0.01	-0.00	-0.01	0.00	0.02**	-0.01	0.01
Sellers are smiling	-0.01	0.02**	-0.00	-0.00	0.00	-0.00	0.03***	0.01
Additive in bread	-0.01	-0.01	-0.01	0.00	0.00	0.01	-0.01	-0.01
The calorie level of bread	-0.02***	0.03***	-0.02	0.03***	0.03***	-0.01	-0.02**	-0.02
Easy to find	-0.03***	0.02**	-0.03***	0.02	0.01	-0.00	-0.02**	-0.02*
To be salt-free	0.04***	-0.00	0.02***	0.00	0.01	-0.02**	-0.01	0.00
Being hot and fresh	-0.05***	0.00	-0.03***	-0.01	-0.01*	0.00	-0.04***	-0.03***
To be well cooked	-0.01	-0.00	-0.02**	-0.02	-0.01	-0.01	-0.04***	-0.01
Being cooked in a wood oven	0.02	0.00	0.01	-0.01	-0.00	-0.02**	0.01	0.00
Produced/sold in a hygienic environment	-0.00	-0.03***	-0.01	-0.03***	-0.02*	-0.01	-0.04***	-0.01

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

When the correlation coefficients between the socio-demographic characteristics of the consumers and the subjects taken into consideration when purchasing bread are examined, it is observed that although there are negative and positive correlation coefficients, there is no significant correlation coefficient. It is clarified that the correlation coefficient values are quite small. In this study, various socio-demographic characteristics of individuals, access to places where bread is purchased, packaging, calorie level of bread, easy availability, salt-free, hot, fresh, well-cooked, and hygienically produced and sold found to be more related than other characteristics. Akdemir et al. (2020) emphasizes that the factors of being close to home, freshness, hygiene, and being of good quality come to the fore in choosing a place where bread is purchased. Sandvik et al. (2017) determined that young consumers generally prefer soft, juicy texture (low moisture absorption) and sweet flavored bread.

On the other hand, bread with this content was also perceived as the least healthy bread type, not rye bread. It was defined that this situation is associated with eating mainly white bread or a sweet loaf of bread during childhood and finding convenience and familiarity in food selection, especially among males. The high level of education of the individuals, being a woman and expressing the importance of health, ethical awareness, and natural

content in food selection was associated with the preference for healthier, chewy resistant, and sour-flavored rye bread, especially among young consumers. It is clarified that older consumers often like whole-grain rye-rich bread slightly more. On the other hand, individuals in the elderly group liked soft white bread more, had lower education levels, were born, and raised in Sweden, were associated with less consumption of dark, chewy bread in childhood, and the importance of natural ingredients and food choice were negatively related. The combination of sensory properties such as light color and soft texture led to the perception that the bread is less healthy than rye bread, while the opposite was found for healthier bread and rye bread.

The connection between the socio-economic features of the consumers and the quality of the bread they consume is evaluated (Table 9).

Table 9
The connection between socio-demographic features of the consumers and their opinions on the quality of the bread

Variables	Socio-demographic characteristics							
	Age	Gender	Marital status	Height	Bodyweight	Education	Working status	Income rate
Opinions on quality of bread	-0.01	0.04***	-0.01	0.05***	0.03***	-0.03***	0.01	0.01

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

Although there is a statistically significant correlation between some socio-demographic characteristics of the consumers and the quality of the bread they consume and a negative correlation between some of their characteristics, it is observed that these correlation coefficients are quite small, as in most of the previous section. On the other hand, there is a positive correlation between consumers' gender, height, body weight, and bread quality thoughts. Amini-Rarani et al. (2021) examined the role of social health and demographic factors on bread quality in Iran. They found that low social characteristics of individuals increase the probability of low-quality bread production, while illiteracy or low education level increases the risk of low-quality bread production. It is reported that social health and responsibility, awareness, and empathy are associated with bread quality. Health policies have been referred to the social health of bakeries to increase the quality of bread. In addition, this study covers low and medium social health, 7-12 years of education; it was emphasized that lower work experience, rental property status, and living in the south increased the likelihood of baking poor quality bread. Only one factor was determined to be in a protective position, including the central location. The research conducted by Eglite and Kunkulberga (2017) aims to determine the trends in bread selection and consumption of consumers in Latvia. It was discovered that the reasons for the consumer preference related to the quality and price of the bread and the trust and behavior of the consumers are different for different types of bread. While the determining factor in the selection of wheat bread was the price, the choice of rye bread was determined by the previous experience, namely the producer of the consumed bread. It is thought that consumers can increase their bread consumption if more delicious and high-quality bread is produced.

In the following stage, the correlation between the thought of why the quality of the consumed bread is not satisfactory and the socio-demographic characteristics of the consumers are also analyzed (Table 10).

Table 10
Links between socio-demographic attributes and the notion that the quality of bread is not sufficient

Variables	Socio-demographic attributes							
	Age	Gender	Marital status	Height	Bodyweight	Education	Working status	Income rate
The manufacturer's disregard for quality	0.01	0.04***	-0.00	0.02**	0.02**	-0.01	0.01	0.00
Uneducated bakers	0.05***	0.04***	0.04***	0.01	0.03***	-0.01	0.02*	0.02*
Presence of excessive and unconscious additives in bread	0.02	-0.01	-0.00	0.01	0.00	0.02	0.00	0.01
Bread not cooked well	0.01	0.01	0.01	-0.01	0.02	-0.04***	-0.00	-0.00
Bakers not paying attention to cleanliness	0.02*	-0.05***	0.02**	-0.04***	-0.01	-0.02**	-0.03***	0.00
Poor quality of flour used in bread production	-0.00	0.02	-0.02*	-0.00	0.00	0.00	0.00	0.00
Using bleach in bread making	0.05***	-0.03***	0.03***	-0.03***	-0.01	0.02**	0.03***	0.02**
Bread is tasteless	-0.03***	-0.01	-0.03***	-0.01	-0.01	0.02*	-0.03***	-0.02*

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

At this stage, although it is discovered that there are negative correlations between some variables and positive correlations between others, the correlation coefficients between statistically significant variables are also found to be at a very low level. While there is a positive correlation between most of the socio-demographic characteristics of the consumers and the bread producers' disregard for quality, untrained bakers, the presence

of unknown or excessive additives in the bread content, the bread not baked well, the low quality of the flour used in the bread production, the bakers' lack of cleanliness. There is a negative correlation between the lack of importance and the poor quality of the bread. Dong and Karboune (2021) emphasized a continuing interest in optimizing post-production quality and extending the shelf life in relation to the preservation and enhancement of the economic and social values, flavors, and textural properties of bread a staple food product. Second, it has been the subject of numerous studies and reviews, in which different approaches and views are discussed. However, it was emphasized that the differences in freshness, flavor, and textural quality of bread are still a cause for concern in bread making. At the same time, the industry expects bread products with high-quality qualities and without synthetic ingredients that satisfy consumers' tastes and sustainable lifestyles. It was defined that it is very important to focus mainly on the quality profiles of bread, including flavor, rheological, textural, and sensory aspects. The linkage between consumers' daily eaten slices of bread and bread consumption in meals and some socio-demographic characteristics are also examined (Table 11). Although it is clarified that there is a statistically significant, negative, and positive correlation between these two groups of variables and the socio-demographic characteristics of consumers, the correlation coefficients are found to be quite small.

As the education level and income of the individual increase, the slice of bread eaten daily decreases, while the bread consumption increases relatively as the age increases, being married and working in a job because there is a positive correlation between them. Interestingly, as the education level and income increase, the consumption of bread with food decreases. However, with increasing age, being married, and increasing height and body weight, bread consumption with food increases.

This part of the study examines the connection between which food meals consumers consume with bread and which meals they mostly consume (Table 12).

Table 11
The connection between socio-demographic features of the consumers and their bread consumption in meals

Variables	Socio-demographic attributes							
	Age	Gender	Marital status	Height	Bodyweight	Education	Working status	Income rate
Number of slices of bread eaten per day	0.07***	0.25	0.08***	0.15	0.16	-0.13***	0.05***	-0.03***
Eating bread with meals	0.04***	0.11	0.04***	0.05***	0.07***	-0.07***	0.01	-0.02*

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

Table 12
The correlation between bread and the food or meals consumed meals the customers mostly consume

Variables	Bread and food or meals consumed by the consumers				
	Soap	Slops	Meat, chicken, and rice	Rice and pasta	I never consume
At breakfast	0.01	-0.02**	-0.10***	-0.07***	-0.21***
At lunch	-0.01	0.02*	0.03***	0.01	-0.05***
At snacks	-0.03***	-0.03***	-0.02**	-0.02	-0.02**
At the dinner	0.02**	0.06***	0.13	0.08***	-0.11***

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

As expected, it is observed that bread is consumed more intensively at lunch and dinner along with slops, meat, chicken, and rice, rice, and pasta, and the correlation between these variables is therefore statistically significant. The most interesting and remarkable result is that the correlation between the variables that do not consume any bread at breakfast is negative, whereas the correlation coefficient is relatively high. On the other hand, a negative but statistically significant yet low correlation coefficient is determined between lunch and snack meals and the expression I never consume. In this process, it can be predicted that the rapidly changing living and working conditions and the active participation of women in working life have significant effects. Although there are important structural changes and breaks in all areas of working life and life in this covid 19 process, it is an inevitable reality that there are irreversible changes and developments in eating habits. Schwedhelm et al. (2018) found a high correlation between food products such as bread, margarine, butter, and cheese that are often eaten together.

On the other hand, they determined that there is a strong negative correlation between potatoes, pasta, rice, tea, and breakfast, which are easily substituted for each other. Charlebois et al. (2020) conducted a research study examining the eating habits of Canadians and highlighting important trends in food consumption patterns and showing the socio-economic motivations that affect food management practices today. It has been suggested that contemporary Canadians experience a disruption in mealtimes, an increase in snacking frequency, and an erosion of the desire or ability to cook. As these fragmented eating habits and disintegration of traditional food patterns represent a challenge for public health nutrition in Canada and are particularly relevant to changes in the provision of dietary guidance nationally, further research on this topic is recommended.

This last part of the study focuses on how the stale bread is determined, how stale bread is evaluated, and how the bread is stored (Table 13).

Table 13
Correlation between understanding and evaluation of stale bread and storage of bread

Variables	Evaluating of stale bread							
	I make bread-crumbs, crusty bread, etc.	I use it dry	I feed the animals	I use it as a meatball paste	I give the bread to the milkman and porter	I do not evaluate it in any way. I throw it away	I use it in soup and food	I give to what is needed
Bread hardening and crumbling	0.04***	0.02**	-0.00	0.07***	-0.03***	-0.05***	0.04***	-0.00
Changing the taste of bread	-0.01	0.01	0.03***	-0.03***	0.01	-0.02	-0.00	0.01
Bread molding	-0.03***	-0.02**	-0.02**	-0.04***	0.00	0.06***	-0.04***	-0.03***
If not consumed that day	-0.01	-0.01	0.02**	0.00	0.01	-0.01	-0.00	0.03***
If the appearance of the bread has changed	0.05***	0.05***	0.08***	0.06***	0.03***	0.04***	0.06***	0.04***
I keep it in the freezer	0.18	0.20	0.05***	0.22	0.03***	-0.09***	0.21	0.11
I consume it by heating	0.12	0.16	0.03***	0.18	0.02**	-0.13***	0.28	0.12
I keep it in the fridge	0.15	0.20	-0.02*	0.19	0.05***	-0.10***	0.24	0.12

*, **, *** denotes statistically significant at p<0.10, p<0.05, and p<0.01, respectively.

It is determined that the correlation between most of the variables related to the evaluation of stale bread and most of the variables related to how the stale bread is understood and how the bread is stored is statistically significant, both positive and negative. On the other hand, the correlation coefficients calculated in this section are also quite small. Concerning the staling of the bread; if the appearance of the bread is changed, it is defined that there is a positive correlation between the expressions of storing the bread in the refrigerator, consumption by heating and keeping the bread in the freezer, and all evaluation attitudes of stale bread. However, again about the staleness of the bread, it is determined that there is a negative correlation relationship between the change in taste, molding of bread, and all evaluation attitudes toward stale bread. In this case, it can be stated that consumers make as much effort as possible to consume bread in the absence of a negative phenomenon related to human health, especially in the perception of the staleness of bread. At this point, it is thought that a great effort is made to prevent the bread from being thrown away as garbage, and there is an important positive perception in this regard. However, in the staleness of bread, evaluation forms may affect human health, especially when it comes to microbial activities, consumption of bread, feeding to animals rather than human consumption, etc. These interpretations are approved by Demirtaş et al. (2018), Sarica et al. (2021), and Dong and Karboune (2021).

Supervised Classification Models

One of the most preferred data mining methods is organizing various data into predefined categories by applying classification (Patil and Shinde, 2020). Nguyen and Do (2020) compared many data mining techniques for training a large dataset for classification optimization. According to their study, Naïve Bayes algorithms could give the best prediction findings. The classification task defines the estimation of the target variable using the input variables. If a target is provided as part of the dataset, classification is a supervised task. Therefore, it is important to analyze the performance of supervised classification models before using them in a classification task. In the study, the performance of supervised classification models such as Decision Tree and Naive Bayes is evaluated using the KNIME Analytics platform. In both machine learning algorithms, firstly, consumers consider whether bread is healthy food, and secondly, how many slices of bread do you eat per day? The focus is on approaches for estimating expressions. In general, creating a classification model consists of two steps. The first is the training phase, where the model is created and the parameters of this data, called training data, are adjusted as it helps to build a model. The second is the testing phase, in which the learned model is applied to new data. The model is then evaluated according to its performance on the test data (Basha et al., 2018). In all models, 80% of the data is used for training and 20% for testing.

Decision Tree Modeling

A decision tree is a classification model that uses a tree-like structure to present multiple decision paths. Crossing each path and classifying an input sample leads to a different path. A Decision Tree is an algorithm that uses a supervised machine learning approach by repeatedly splitting data to solve regression and classification problems based on a given variable. The data is first divided into nodes, and then the tree's leaf represents the final decisions. The main purpose of the decision tree is to create a model that can be used to predict the target

variable by learning the simple decision rules obtained from the training data (Chauhan, 2020). The tree is created during a training process using training data. Leaf nodes express the name of the class, while the decision node is a non-leaf node (Eesa et al., 2015a; Eesa et al., 2015b). The decision tree handles categorical and numerical data. The nonlinear relationship between the arguments does not affect the efficiency of the tree (Tian et al., 2019). No preprocessing of data is required. However, the possibility of over-learning may arise when the tree is repeatedly rebuilt (Ibrahim and Abdulazeez, 2021).

Naïve Bayes Modeling

A Naïve Bayes model uses a probabilistic approach to classification. Bayes Theorem is used to understand the relationship between the input and the output class. Naïve Bayes is a probabilistic and statistical method based on the classification algorithm. It is a standard algorithm in machine learning applications as it offers a simple approach to all features contributing equally to the final decision. Computational efficiency equals this simplicity, making the Naive Bayes approach exciting and suitable for different fields. The basic component of the Naive Bayes classification is the prior, next, and class conditional probability (Wibawa et al., 2019). This method offers many benefits, such as being easy and very useful for large data sets. It can be used for binary and multiclass classification problems. Relatively less training data is required and can be used for both discrete and continuous data (Uddin et al., 2019). For example, this algorithm can easily filter spam and classify documents (Yaswanth and Riyazuddin, 2020).

Performance Metrics

Key performance metrics for result validation; Precision, Recall, F-Measure, True Positive (TP) Ratio, False Positive (FP) Ratio, ROC Area, Kappa Stats and Accuracy, etc. These performance measures help to understand the performance of a model better and clearly. Brief explanations of these performance estimators are presented below (Gupta et al., 2021). When a model is created, or existing models are used for a classification approach, that model's success is considered the number of correct predictions from all predictions made. However, this information only gives the accuracy of the classification. The confusion matrix alone is often not enough information to decide whether a model is good enough. True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN). True Positive (TP): Correctly predicted positive values. This indicates that the value of the actual class and the estimated class are the same. True Negative (TN): These are correctly predicted negative values. This indicates that the value of the actual class and the estimated class are the same. False Positive (FP): This value appears when your actual class and the predicted class conflict. False Negative (FN): This value appears when your actual class conflicts with the predicted class. While it is desired to increase the true positive and true negative areas, decreasing the false positive and false negative areas shows that the classification performance is good. Thanks to the confusion matrix, the following metrics can be calculated. Accuracy: Accuracy is the most intuitive performance measure and is the ratio of correctly predicted observations to total observations. If the model used has high accuracy, it can be considered that the model is the best. However, in cases where the number of false-positive and false-negative values is quite different, other parameters should be considered to evaluate the model's performance. Precision: Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. This is also called Positive Predictive Value. Precision can be thought of as a measure of the precision of classifiers. Low sensitivity can also indicate many false positives. Recall, Sensitivity: The ratio of correctly predicted results to the total number of positives. It is the proportion of correctly predicted positive observations for all observations in the classification. It can be thought of as a measure of the integrity of the classifiers. Low sensitivity indicates a lot of false negatives. Specificity: In fact, it is an indicator of how many of the negative values are predicted as negative. F-score (F-measure): It is the harmonic mean of recall and precision. Therefore, it considers both false positives and false negatives. Looking at the F-score is more useful than looking at the accuracy value, especially in cases where there is an uneven class distribution. If false positives and false negatives occur in similar numbers, looking at the accuracy value give the best results for classification success. If the numbers of false positives and false negatives are very different, it is necessary to look at sensitivity and sensitivity. ROC area: Also known as the Receiver Operating Characteristic, it plots the performance measurement graph. The true-positive ratio is plotted against the false-positive ratio. ROC curves are widely adopted in machine learning to significantly visualize, organize, and measure performance. The area under the ROC curve is called the AUC (Area under the curve). Cohen's Kappa statistics: A measure that relates Observed Accuracy to Expected Accuracy. Kappa statistic is also used to evaluate classifiers among themselves.

Comparison of Two Machine Learning Algorithms Findings

First, the results of the two algorithms are compared to predict consumers' opinion of whether bread is healthy food. The numbers 1 and 0 represent yes and otherwise, respectively. Workflows of Decision Tree (Figure 5) and Naïve Bayes algorithms (Figure 6) in KNIME are represented. The confusion matrices and performance metrics of both algorithms are represented together in Table 14.

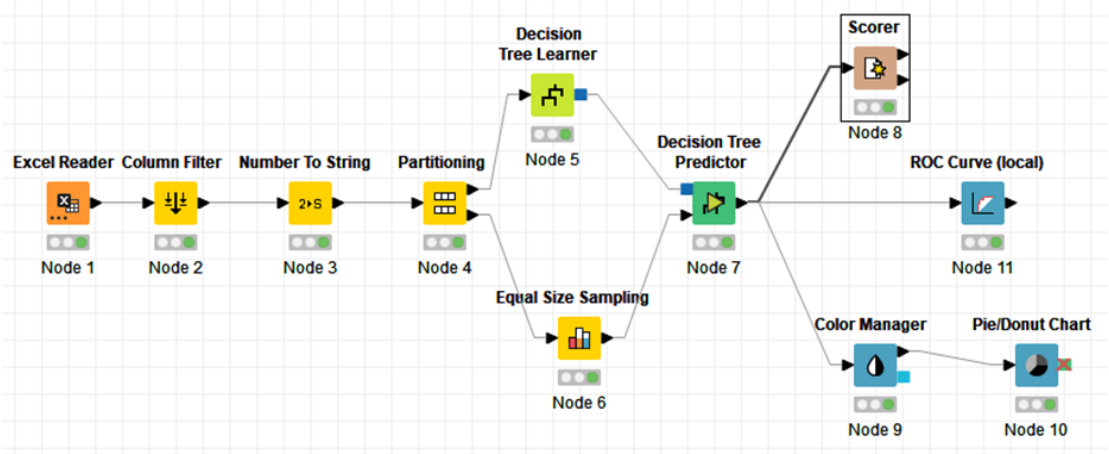


Figure 5
Workflow of Decision Tree defining consumers' perception of whether bread is a healthy food

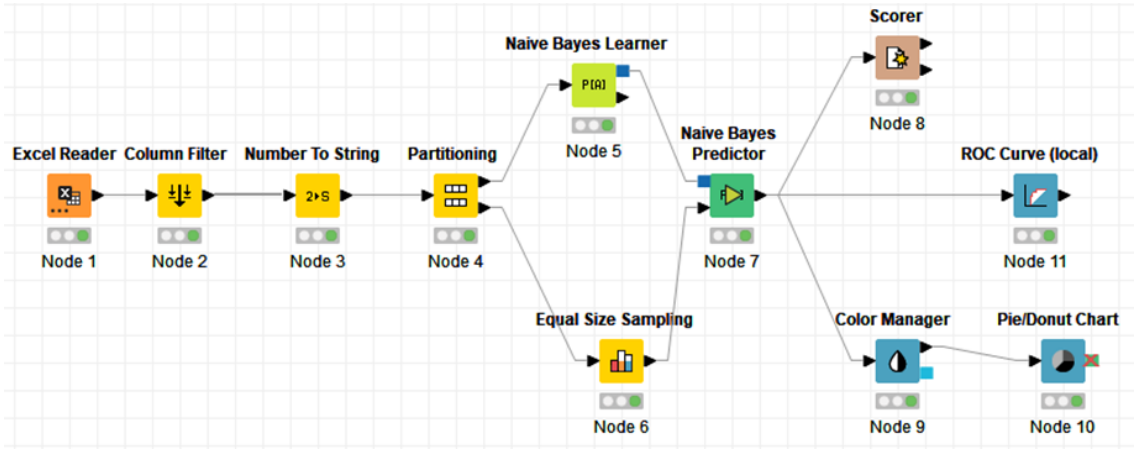


Figure 6
Workflow of Naïve Bayes describes consumers' perception of whether bread is a healthy food

Table 14
Confusion matrices and performance metrics for bread healthiness' perceptions

Decision Tree		Predicted value		Accuracy	Performance metrics				
		Yes	Otherwise		Sensitivity	Specificity	Precision	F-score	Cohen's Kappa
Actual value	Yes	649	228	0.738	0.733	0.742	0.740	0.737	0.476
	Otherwise	236	657		0.742	0.733	0.736	0.739	
Naïve Bayes		Predicted value		Accuracy	Performance metrics				
		Yes	Otherwise		Sensitivity	Specificity	Precision	F-score	Cohen's Kappa
Actual value	Yes	621	155	0.744	0.651	0.838	0.800	0.718	0.488
	Otherwise	333	799		0.838	0.651	0.706	0.766	

It can be stated that the findings obtained from the Naive Bayes algorithm in terms of many performance metrics, especially the accuracy measurement, are more consistent than the Decision Tree algorithm, and the predictive power is higher. In the Decision Tree algorithm, among 1306 correct predictions, the predictions of yes and other cases are approximately equal (Figure 7). In the Naive Bayes algorithm, within 1420 correct predictions, the yes option is estimated with an accuracy of 41% and the other case option with an accuracy of 59% (Figure 8). Furthermore, the ROC curves of the Decision Tree and Naive Bayes algorithms are quite similar (Figure 9 and Figure 10). The area value under the ROC curve is 0.7524 and 0.8383 for the Decision Tree and Naive Bayes algorithms, respectively. Therefore, it can be said that the Naive Bayes algorithm is slightly more successful than the Decision Tree modeling.

Second, the results of the two algorithms are compared to the statement, how many slices of bread do you eat per day? The numbers 0, 1, 2, 3, and 4 refer to no, one or two slices, three to four slices, five to six slices, or more than six slices in a day. Finally, workflows of Decision Tree (Figure 11) and Naïve Bayes algorithms (Figure 12) in KNIME are represented.

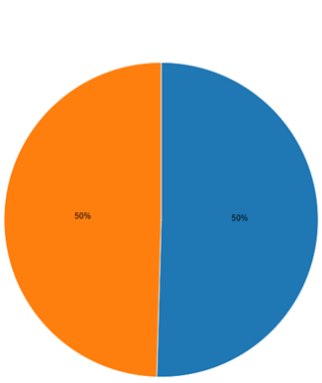


Figure 7
The prediction level of bread's perception of being healthy with the Decision Tree algorithm

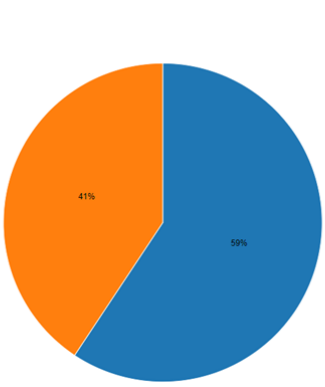


Figure 8
The prediction level of bread's perception of being healthy with the Naïve Bayes algorithm

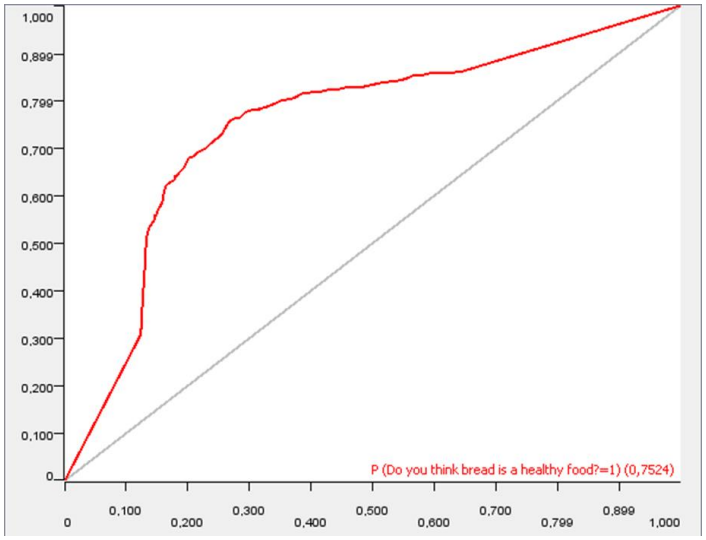


Figure 9
ROC curve for the perception of being bread healthiness with Decision Tree algorithm

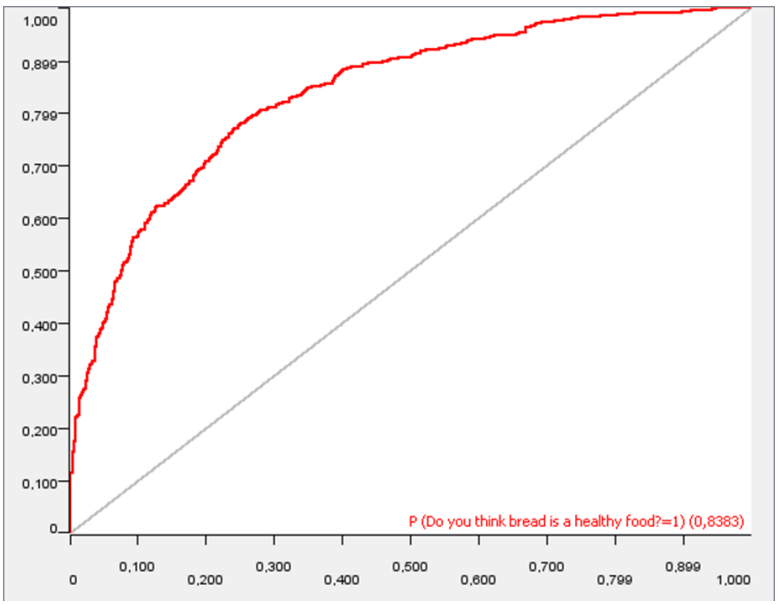


Figure 10
ROC curve for the perception of being bread healthiness with Naïve Bayes algorithm

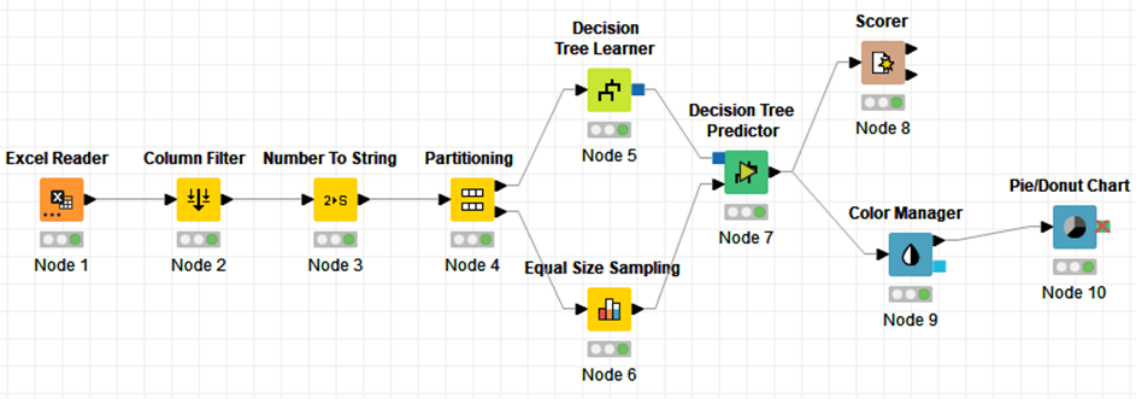


Figure 11
Workflow of Decision Tree algorithm revealing consumers' habit for eating bread per day

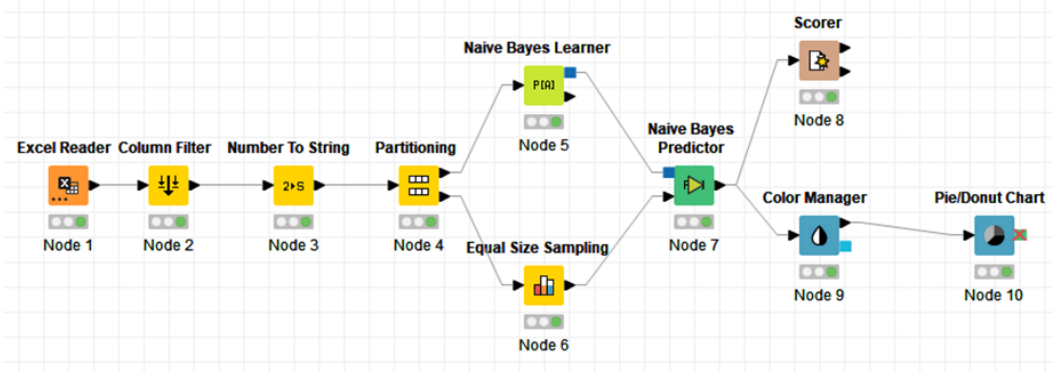


Figure 12
Workflow of Naïve Bayes algorithm representing consumers' habit for eating bread per day

The confusion matrices and performance metrics of both algorithms are represented together in Table 15 for this stage of the study. Although the performance metric values obtained from the Naive Bayes algorithm are higher than those obtained from the Decision Trees algorithm, it can still be stated that they are relatively low. A total of 181 correct predictions were made in the Decision Trees algorithm, and a total of 187 correct predictions were made in the Naive Bayes algorithm. The percentages of correct predictions made in the decision tree algorithm are 12%, 31%, 27%, 16% and 14% for the values 0, 1, 2, 3 and 4, respectively (Figure 13). In the Naive Bayes algorithm, the percentages for 0, 1, 2, 3 and 4 values are 39%, 10%, 17%, 9% and 25%, respectively (Figure 14).

Table 15
Confusion matrices and performance metrics of habit for eating bread

Decision Tree		Predicted value					Accuracy	Performance metrics				
		0	1	2	3	4		Sensitivity	Specificity	Precision	F-score	Cohen's Kappa
Actual value	0	49	24	10	3	0		0.570	0.991	0.942	0.710	
	1	3	44	21	9	9		0.512	0.735	0.326	0.398	
	2	0	34	33	8	11	0.421	0.384	0.765	0.289	0.330	0.276
	3	0	16	30	28	12		0.326	0.878	0.400	0.359	
	4	0	17	20	22	27		0.314	0.907	0.458	0.372	
Naïve Bayes		Predicted value					Accuracy	Performance metrics				
		0	1	2	3	4		Sensitivity	Specificity	Precision	F-score	Cohen's Kappa
Actual value	0	79	3	0	0	2		0.940	0.750	0.485	0.640	
	1	48	12	12	3	9		0.143	0.908	0.279	0.189	
	2	18	14	24	7	21	0.445	0.286	0.854	0.329	0.306	0.307
	3	11	10	22	20	21		0.238	0.952	0.556	0.333	
	4	7	4	15	6	52		0.143	0.842	0.495	0.550	

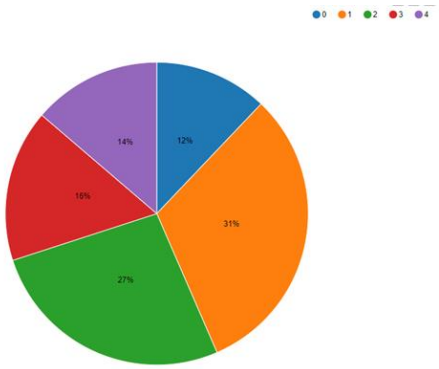


Figure 13
The prediction level of eating bread habit with the Decision Tree algorithm

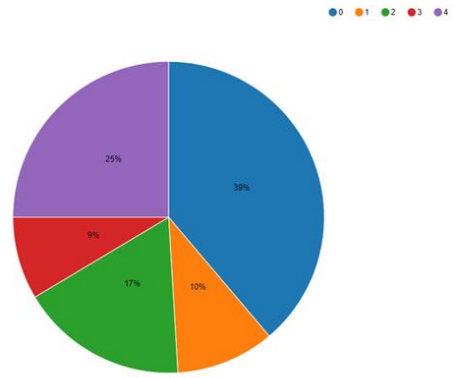


Figure 14
The prediction level of eating bread habit with the Naïve Bayes algorithm

Conclusion

This study is conducted at the national level with a large sample size of consumers. With this structure, it can be stated that the study is unique, and the results obtained are suitable for generalization. Although bread is sacred to Turkish society, it still constitutes an important component of Turkish food culture. Although bread consumption in Turkey seems to be high compared to many western countries, individuals are trying to reduce their bread consumption due to obesity and various health problems, which show a significant prevalence rate. While the results obtained from this study support the inferences, the characteristics such as bread consumption, taste, aroma, appearance, nutritiveness, and low calories are tried to be reduced by the influence of factors such as income level and living conditions, gender, are considered. Bread consumption preferences are directed towards whole wheat bread and rye bread. Descriptive statistics and correlation analysis can reveal the change in bread consumption preferences seen in society in recent years. The interaction between consumer attitudes and perceptions and socio-demographic characteristics that affect bread consumption preferences is evident. Key and structural concepts such as healthy life, high-quality bread consumption, low calories, changing living and working conditions, and women's participation in business life at least as much as men come to the fore. In addition, in this study, using Decision Trees and Naïve Bayes algorithms, classification approaches within the supervised data mining methods were used in the bread consumption perception and preferences of consumers.

It is focused on the attitudes, perceptions, and thoughts of consumers on whether bread is healthy food and about daily consumed bread slices. As a result of both algorithms, it is defined that the estimation results of bread being a healthy food are more consistent and valid. In future studies, it is thought that it would be appropriate to use other machine learning algorithms in bread consumption preferences by expanding and customizing the subjects and purposes.

4. Acknowledgements

We would like to thank the individuals who participated in this study for their participation, especially during the period of the Covid-19 epidemic.

5. References

Adebayo FA, Itkonen ST, Koponen P, Prattala R, Harkanen T, Lamberg-Allardt C, Erkkola M (2017). Consumption of healthy foods and associated socio-demographic factors among Russian, Somali and Kurdish immigrants in Finland. *Scandinavian Journal of Public Health*, 45: 277-287.

Akdemir Ş, Keskin F, Ünal H, Miassi Y (2020). Analysis of behaviors and attitudes of bread consumption in Turkey: Case of the province of Adana. *European Journal of Science and Technology*, 19: 659-667.

Amini-Rarani M, Abutoraabi SH, Nosratabadi M (2021). The role of social health and demographic factors in bread quality: An ecological study in Isfahan, Iran. *Journal of Food Quality, 2021*, 1-6.

Anonymous (2021). Spearman's rank correlation coefficient. https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient/ Accessed 19 July 2021.

Bacarea A, Bacarea VC, Cinpeaunu C, Teodorescu C, Seni AG, Guine RPF, Tarcea M (2021). Demographic, anthropometric and food behavior data towards healthy eating in Romania. *Foods*, 10: 487, 1-16.

Basha SM, Rajput DS, Poluru RK, Brushan SB, Basha, SAK (2018). Evaluating the performance of supervised classification models: Decision Tree and Naïve Bayes using KNIME. *International Journal of Engineering & Technology*, 7: 248-253.

Cappelli A, Oliva N, Cini E (2020). A systematic review of gluten-free dough and bread: Dough rheology, bread characteristics, and improvement strategies. *Applied Sciences*, 10: 6559.

Cagri H (2016). Wastage Problem: Bread Wastage Problem in Turkey and its Effects on the Turkish Economy, Master's thesis, Nevsehir Haci Bektas Veli University, Nevsehir.

Charlebois S, Somogyi S, Kirk SFL (2020). Fragmented food habits and the disintegration of traditional meal patterns: A challenge to public health nutrition in Canada? *Journal of International Food & Agribusiness Marketing*, 32: 69-78.

Chauhan YJ (2020). Cardiovascular disease prediction using classification algorithms of machine learning. *International Journal of Science and Research*, 9: 194-200.

De Boni A, Pasqualone A, Roma R, Acciani C (2019). Traditions, health, and environment as bread purchase drivers: A choice experiment on high-quality artisanal Italian bread. *Journal of Cleaner Production,* 221: 249-260.

Demirtaş B, Kaya A, Dağıstan E (2018). Consumers' bread consumption habits and waste status: Hatay/Turkey example. *Turkish Journal of Agriculture-Food Science and Technology,* 6*:* 1653-1661.

Dong Y, Karboune S (2021). A review of bread qualities and current strategies for bread bioprotection: Flavor, sensory, rheological, and textural attributes. *Comprehensive Reviews in Food Science and Food Safety*, 20*:* 1937-1981.

Eesa AS, Orman Z, Brifcani AMA (2015a). A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. *Expert Systems with Applications,* 42: 2670-2679.

Eesa AS, Orman Z, Brifcani AMA (2015b). A new feature selection model based on ID3 and bees algorithm for intrusion detection system. *Turkish Journal of Electrical Engineering & Computer Sciences*, 23: 615-622.

Eglite A, Kunkulberga D (2017). Bread choice and consumption trends. *Foodbalt,* 2017, 178-182.

Gupta A, Suri B, Kumar V, Jain P (2021). Extracting rules for vulnerabilities detection with static metrics using machine learning. *International Journal of System Assurance Engineering and Management,* 12: 65-76.

GWR (2020). Guinness World Records 2020. Largest Bread Consumption per capita. https://www.guinnessworldrecords.com/world-records/largest-bread-consumption-per-capita / Accessed 23 June 2021.

Ibrahim IM, Abdulazeez AM (2021). The role of machine learning algorithms for diagnosing diseases. *Journal of Applied Science and Technology Trends*, 02: 10-19

KNIME (2021). KNIME Analytics Platform. https://www.knime.com/knime-analytics-platform / Accessed 01 May 2021.

Loloei S, Pouraram H, Majdzadeh R, Takian A, Goshtaei M, Djazayery A (2019). Policy analysis of salt reduction in bread in Iran. *AIMS Public Health,* 6: 534-545.

Montagnese C, Santarpia L, Buonifacio M, Nardelli A, Caldara AR, Silvestri E, Contaldo F, Pasanisi F (2015). European food-based dietary guidelines: A comparison and update. *Nutrition,* 31: 908-15,

Nguyen TTS, Do PMT (2020). Classification optimization for training a large dataset with Naïve Bayes. *Journal of Combinatorial Optimization,* 40: 141–169.

OECD (2021). Organization for Economic Cooperation and Development. https://data.oecd.org/conversion/exchange-rates.htm/ Accessed July 2021.

Patil P, Shinde S. (2020). Performance analysis of different classification algorithms: Naïve Bayes, Decision Tree and K-Star. *Journal of Critical Reviews,* 7: 1160-1164.

Sandvik P, Kihlberg I, Lindroos AK, Marklinder I, Nydahl M (2014). Bread consumption patterns in Swedish national dietary survey focusing particularly on whole-grain and rye bread. *Food & Nutrition Research 2014,* 58, 24024.

Sandvik P, Nydahl M, Marklinder I, Naes T, Kihlberg I (2017). Different liking but similar healthiness perceptions of rye bread among younger and older consumers. *Food Quality and Preference,* 61: 26-37.

Sandvik P, Nydahl M, Kihlberg I, Marklinder I (2018). Consumers' health-related perceptions of bread – Implications for labeling and health communication. *Appetite,* 121: 285-293.

Sajdakowska M, Gebski J, Zakowska-Biemans S (2019). Willingness to eat bread with health benefits: habits, taste and health in bread choice. *Public Health,* 167: 78-87,

Sarica D, Demircan V, Erturk A, Arslantas N (2021). An econometric analysis of the factors affecting consumers' bread waste and consumption behaviour: a case study of Isparta province, Turkey. *British Food Journal 123,* 4: 1449-1464.

Schwedhelm C, Iqbal K, Knüppel S, Schwingshackl L, Boeing G (2018). Contribution to the understanding of how principal component analysis–derived dietary patterns emerge from habitual data on food consumption. *The American Journal of Clinical Nutrition,* 1: 227-235.

Simmons AL, Schlezinger JJ, Corkey BE (2014). What are we putting in our food that is making us fat? Food additives, contaminants, and other putative contributors to obesity. *Current Obesity Reports,* 1: 273-85.

Tian X, Chong Y, Huang Y, Guo P, Li M, Zhang W, Du Z, Li X, Hao Y (2019). Using machine learning algorithms to predict hepatitis B surface antigen seroclearance. *Computational and Mathematical Methods in Medicine, 6915850,*

Uddin S, Khan A, Hossain ME, Moni MA (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making,* 19: 281.

Vandevijvere S, MacKenzie T, Ni Mhurchu C (2017). Indicators of the relative availability of healthy versus unhealthy foods in supermarkets: A validation study. *International Journal of Behavioral Nutrition and Physical Activity 14, 53,*

WB (2019). Global Consumption Database, The World Bank. http://datatopics.worldbank.org/consumption/home/ Accessed 28 June 2021.

Wibawa AP, Kurniawan AC, Murti DMP, Adiperkasa RP, Putra SM, Kurniawan SA, Nugraha YR (2019). Naïve Bayes classifier for journal quartile classification. *International Journal of Recent Contributions from Engineering, Science & IT,* 7: 91-99,

Yaswanth R, Riyazuddin YM (2020). Heart disease prediction using machine learning techniques. *International Journal of Innovative Technology and Exploring Engineering,* 9: 1456-1460.