STATISTICAL AND MULTIVARIATE ANALYSIS OF DIETARY HABITS AND LIFESTYLE BEHAVIORS IN RELATION TO BODY WEIGHT

Valbona SALLAI MAZLLAMI[1], Vesna KNIGHTS[1], Vezirka JANKULOSKA, Zora UZUNOVSKA[1] , Tatjana BLAZHEVSKA[1], Valentina VELKOVSKI[1], Kristina GACOSKA[1]

[1]University "St. Kliment Ohridski" Bitola, Faculty of Technology and Technical Science, Republic of North Macedonia, (valbona.mazlami@uklo.edu.mk , vesna.knights@uklo.edu.mk)

Abstract

Dietary habits and lifestyle behaviors are important factors associated with body weight and metabolic health. Nut and seed consumption, cooking oil preference, sleep duration, and meal frequency may influence nutritional status and body-weight-related outcomes.
This cross-sectional study presents a statistical analysis of dietary habits based on questionnaire responses collected from 22 participants. Descriptive statistics, Bayesian correlation, Bayesian regression, crosstab analysis, and Multiple Correspondence Analysis (MCA) were applied to investigate relationships among body weight, age, nut and seed consumption, bread intake, cooking oil usage, and lifestyle variables.
The results showed an average body weight of 96.14 kg and an average age of 50.14 years among participants. Most participants reported irregular meal frequency, short sleep duration, and predominant use of sunflower oil. Bayesian analyses did not identify a reliable association between age and body weight or between meal frequency and body weight. MCA suggested that body weight, bread consumption, nut and seed intake, and cooking oil preference contributed most strongly to the latent variability structure of the dataset.
The findings provide exploratory insight into dietary and lifestyle behaviors associated with body weight and may support future nutritional and public health investigations using larger population samples.

Keywords: dietary habits, lifestyle, Bayesian analysis, cross-sectional study, nutrition, Multiple Correspondence Analysis

Introduction

Dietary patterns, sleep duration, and socio-economic conditions are recognized correlates of body weight regulation and metabolic disease. Nutritional behaviors such as nut and seed consumption, bread intake, meal frequency, and cooking oil preference may influence obesity risk, insulin resistance, and long-term cardiometabolic health (Campos and Egea 2025). In recent years, increased attention has been directed toward the role of healthy dietary fats and lifestyle habits in preventing obesity and type 2 diabetes.
Nuts and seeds are rich in unsaturated fatty acids, dietary fiber, antioxidants, vitamins, and minerals, and their regular consumption has been associated with improved lipid metabolism and lower obesity prevalence. Sabaté and Ang (2009) reported that frequent nut consumption may contribute to lower body weight and reduced risk of metabolic disorders. Similarly, systematic reviews suggest that nut-enriched diets generally do not promote weight gain and may support satiety regulation and metabolic balance (Nishi et al. 2021).
Lifestyle-related variables such as short sleep duration and irregular meal patterns have also been associated with increased BMI and impaired metabolic homeostasis. Sleep deprivation may alter appetite-regulating hormones and increase energy intake, while irregular eating behavior may contribute to unhealthy dietary patterns (Campos and Egea 2025). However, the strength and direction of these associations often vary depending on sample characteristics, population structure, and statistical methodology.
From a computational and statistical perspective, nutritional datasets frequently contain both continuous and categorical variables with potentially complex interactions. Classical

descriptive statistics are commonly used to summarize distributions and identify dominant dietary patterns, while Bayesian statistical methods provide probabilistic estimation of associations under uncertainty, particularly in small samples. Bayesian correlation and Bayesian regression approaches allow estimation of posterior parameter distributions and credible intervals instead of relying exclusively on frequentist significance testing.
In addition, exploratory multivariate techniques such as Multiple Correspondence Analysis (MCA) enable dimensionality reduction and visualization of latent relationships among categorical variables. MCA transforms categorical data into multidimensional geometric structures using inertia decomposition and eigenvalue analysis, allowing identification of variables contributing most strongly to the overall variance structure of the dataset (Balakrishna et al. 2022).
Therefore, the aim of this study was to evaluate dietary habits related to nut and seed consumption, bread intake, cooking oil preference, and lifestyle behaviors in relation to body weight using descriptive statistics, Bayesian computational analyses, and exploratory multivariate methods in a cross-sectional sample of adult participants.

Materials and Methods

2.1 Study Design and Sample: This study represents a cross-sectional secondary statistical analysis of questionnaire-based nutritional and lifestyle data collected from a convenience sample of 22 participants (N = 22). The analysis aimed to investigate associations between body weight, dietary habits, and selected lifestyle-related variables using both classical and Bayesian statistical approaches.

2.2 Data Source and Preparation: The dataset was imported from an Excel worksheet (Book1.xlsx, Sheet1) into IBM SPSS Statistics software for preprocessing and statistical analysis. The active dataset was subsequently stored as *data.sav*. User-defined missing values were treated as missing during all analyses.
Because several questionnaire responses contained slight textual inconsistencies (e.g., variations in category formatting), SPSS treated some responses as separate categorical levels during frequency and multivariate analyses.

2.3 Variables: The study included both continuous and categorical variables. Continuous variables: Age (years), Body weight (kg), height (m), BMI (kg/m^2). Categorical variables: Religion Employment status Socio-economic condition (income categories) Sleep duration Diabetes diagnosis (yes/no) Meal frequency Frequency of nut consumption Amount of nut consumption Frequency of seed consumption Daily bread intake Bread type Primary cooking oil used for cooking

2.4 Statistical Analysis: Descriptive and inferential statistical analyses were performed using IBM SPSS Statistics.

2.4.1 Descriptive Statistics and Frequency Analysis: Continuous variables were summarized using:
arithmetic mean $\tilde{x}$
standard deviation (SD),
minimum values,
maximum values.

The arithmetic mean was calculated according to: $\tilde{x} = \Sigma_{i=1}^{n} x_i$ (1), while the standard deviation was calculated as: $S = \sqrt{\frac{\Sigma_{i=1}^{n}(x_i - \tilde{x})^2}{n-1}}$ (2).

Categorical variables were analyzed using frequency distributions and percentage representation .

$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$ (3)

where:

$P(\theta|D)$ represents the posterior distribution,
$P(D|\theta)$ represents the likelihood,
$P(\theta)$ represents the prior distribution,
P(D) represents the marginal likelihood.
Posterior means, posterior modes, variances, and 95% credible intervals were reported.

2.4.2 Bayesian Regression Analysis

Bayesian regression analysis was conducted to examine the relationship between body weight and meal frequency categories. The linear regression model can be expressed as:

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ (4)

where:

yi - dependent variable (the outcome for observation *i*)
x_i − independent variable (the predictor for observation *i*)
β_0 - intercept (the value of y when x=0)
β_1 - regression coefficient (the change in y for a one-unit increase in x)
ε_i- error term (the unexplained variation for observation *i*)

Reference priors were applied and posterior parameter estimation was obtained through Bayesian inference. ANOVA summaries generated by SPSS were additionally reported.

2.4.3 Crosstab Analysis: Crosstabulations were performed descriptively to evaluate distributions between:

body weight and seed-consumption frequency,
body weight and daily bread intake,
body weight and employment status,
body weight and sleep duration,
body weight and meal frequency.

Due to sparse cell frequencies and numerous single-observation categories, inferential interpretation of contingency associations was limited.

2.4.4 Nonparametric Analysis

Nonparametric tests including one-sample, related-sample, and independent-sample procedures were performed primarily for diagnostic and exploratory purposes because the small sample size and sparse categorical structure reduced the robustness of classical parametric assumptions.

2.4.5 Multiple Correspondence Analysis (MCA)

Multiple Correspondence Analysis (MCA) was applied to investigate latent multidimensional relationships among categorical dietary and lifestyle variables. Variable-principal normalization was used with two extracted dimensions and a maximum of 100 iterations. Convergence was successfully achieved.

MCA decomposes categorical variance using inertia-based dimensional representation:

$I = \sum \frac{(f_{ij} - f_i f_j)^2}{f_i f_j}$ (5),

where:

f_{ij} – represent observed joint frequencies,

$f_i f_j$ – represents expected frequencies under indepedence

SPSS automatically discretized string variables using ranking-based transformation procedures.

All statistical analyses used two-sided inference procedures, while Bayesian analyses reported 95% credible intervals. Given the relatively small sample size and sparse contingency tables, all findings were interpreted cautiously and considered exploratory rather than confirmatory.

Results

3.1. Sample Characteristics: A total of 22 participants were included in the statistical analysis (N = 22). The average age of participants was 50.14 years (SD = 12.40), with values ranging from 29 to 68 years. The average body weight was 96.14 kg (SD = 26.55), with a minimum value of 60 kg and a maximum value of 191 kg. These descriptive indicators suggest that the analyzed sample consisted predominantly of middle-aged individuals with relatively elevated body weight values.

Table 1. Descriptive statistics of continuous variables

Variable	N	Minimum	Maximum	Mean	SD
Age (years)	22	29	68	50.14	12.40
Body weight (kg)	22	60	191	96.14	26.55
BMI (kg/m^2)	22	22.6	58.95	41.86	10.56

The arithmetic mean was calculated using equation $\bar{x} = \frac{\sum x}{N}$ (6), while the standard deviation was estimated using equation $S = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$ (7), where xi represents individual observations and n is the sample size.

3.2. Distribution of Categorical Variables: The categorical analysis demonstrated that the majority of participants identified as Orthodox Christian (68.2%), followed by Muslim participants (27.3%). According to employment status, office-based work was the most represented occupational category (40.9%), while unemployed individuals and pensioners each represented 22.7% of the sample.

Regarding socio-economic condition, the most frequent income category was 300 € (27.3%), followed by 500 € and above 500 € (22.7% each). Most participants reported sleeping between 5–6 hours per night (45.5%), while 22.7% reported sleeping 7–8 hours.

Half of the participants (50%) reported having diabetes mellitus. Meal frequency analysis showed that irregular meal consumption dominated within the sample (54.5%), whereas only 22.7% reported regular meal frequency.

Dietary indicators demonstrated that nut consumption was most commonly reported as 1–2 times weekly (36.4%), while seed consumption was frequently occasional or irregular. White bread (40.9%) and integral bread (36.4%) represented the dominant bread categories. Sunflower oil was the predominant cooking oil, used by 72.7% of participants.

Table 2. Distribution of main categorical variables

Variable	Dominant category	Percentage (%)
Religion	Orthodox Christian	68.2
Employment status	Office work	40.9
Sleep duration	5–6 h	45.5

Diabetes diagnosis	Yes	50.0
Meal frequency	Irregular	54.5
Bread type	White bread	40.9
Cooking oil	Sunflower oil	72.7

The categorical frequency distributions indicate relatively heterogeneous nutritional and lifestyle patterns within the analyzed sample.

3.3. Descriptive and Diagnostic Statistical Testing: One-sample t-tests were performed for the continuous variables age and body weight. As expected, both variables produced statistically significant results ($p < 0.001$), primarily because the hypothesized test value was fixed at zero. Consequently, these analyses were considered diagnostic rather than substantively informative for nutritional interpretation.
The t-statistic is defined with equation:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \quad (8),$$

where:
$\bar{x}$ – the sample mean
μ_0 – the hipothesiyed population mean
s – the sample standard deviation
n – the sample size

In addition, one-sample chi-square goodness-of-fit tests were performed to evaluate whether categorical responses were uniformly distributed across categories. Significant deviations from equal category distribution were observed for religion ($p = 0.001$), sleep duration ($p = 0.006$), daily nut-consumption amount ($p = 0.013$), seed-consumption frequency ($p = 0.045$), bread type ($p = 0.002$), and cooking oil preference ($p < 0.001$). These findings indicate the presence of dominant behavioral and dietary patterns within the analyzed sample. Due to the relatively small sample size and sparse categorical frequencies, the chi-square results should be interpreted cautiously.
The chi-square goodness-of-fit statistic is defined with equation :

$$x^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (9),$$

where:
O_i- observed frequency in category i
E_i – expected frequency in category i
$\sum$ - sum over all categories

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories of Пациенти occur with equal probabilities.	One-Sample Chi-Square Test	1,000[1]	Retain the null hypothesis.
2	The categories of религија occur with equal probabilities.	One-Sample Chi-Square Test	,001[1]	Reject the null hypothesis.
3	The categories of работно место occur with equal probabilities.	One-Sample Chi-Square Test	,327[1]	Retain the null hypothesis.
4	The categories of социо-економск условн occur with equal probabilities.	One-Sample Chi-Square Test	,802[1]	Retain the null hypothesis.
5	The categories of времетрајно на спиење occur with equal probabilities.	One-Sample Chi-Square Test	,006[1]	Reject the null hypothesis.
6	The categories defined by дали е диагостициран со диабет = не an да occur with probabilities 0,5 and 0,5.	One-Sample Binomial Test	1,000[1,1]	Retain the null hypothesis.
7	The categories of Фрекфенција на оброци occur with equal probabilities.	One-Sample Chi-Square Test	,108[1]	Retain the null hypothesis.
8	The categories of колку често консумирате јаткасти плодови occur with equal probabilities.	One-Sample Chi-Square Test	,054[1]	Retain the null hypothesis.
9	The categories of Колку консумирате дневно occur with equal probabilities.	One-Sample Chi-Square Test	,013[1]	Reject the null hypothesis.
10	The categories of Колку често консумирате семиња occur with equal probabilities.	One-Sample Chi-Square Test	,045[1]	Reject the null hypothesis.
11	The categories of Колку парчиња леб јадете дневно occur with equal probabilities.	One-Sample Chi-Square Test	,074[1]	Retain the null hypothesis.
12	The categories of Какoв леб користите occur with equal probabilities.	One-Sample Chi-Square Test	,002[1]	Reject the null hypothesis.
13	The categories of Кое масло превенственого користете за готвење occur with equal probabilities.	One-Sample Chi-Square Test	,000[1]	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

[1] Exact significance is displayed for this test.

[1] Lilliefors Corrected

[2] This is a lower bound of the true significance.

Figure 1. Summary of one-sample chi-square goodness-of-fit tests for categorical variables.

3.4. Associations Between Variables

3.4.1. Bayesian Correlation Between Age and Body Weight

Bayesian correlation analysis was applied to evaluate the probabilistic relationship between age and body weight. The posterior mean correlation coefficient was approximately:

$r \approx -0.304$

with a 95% credible interval ranging from −0.648 to 0.059. The posterior mode was approximately −0.344. Since the credible interval crosses zero, the Bayesian evidence is insufficient to support a reliable association between age and body weight within this sample.

Table 3. Bayesian correlation between age and body weight

Parameter	Value
Posterior mean	−0.304
Posterior mode	−0.344
95% credible interval	[−0.648 ; 0.059]

Bayesian inference was estimated according to Bayes' theorem equation:

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)} \qquad (10),$$

P(A|B) - posterior probability

P(B | A) – likelihood

P(A) - prior probability

P(B) - *evidence* (marginal probability)

3.4.2. Bayesian Regression: Body Weight vs. Meal Frequency: Bayesian regression analysis was conducted to investigate whether meal frequency predicts body weight. The obtained model was not statistically significant:
$F=1.246, p=0.310$
indicating that meal frequency was not a reliable predictor of body weight in this sample. Posterior intervals for meal-frequency categories were broad and overlapped zero, reflecting substantial uncertainty associated with the small sample size.

Table 4. Bayesian regression analysis

Parameter	Value
F statistic	1.246
p-value	0.310

The regression model is represented with equation :
$Y = a + bX$ (11),
where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept (the value of y when x = 0).

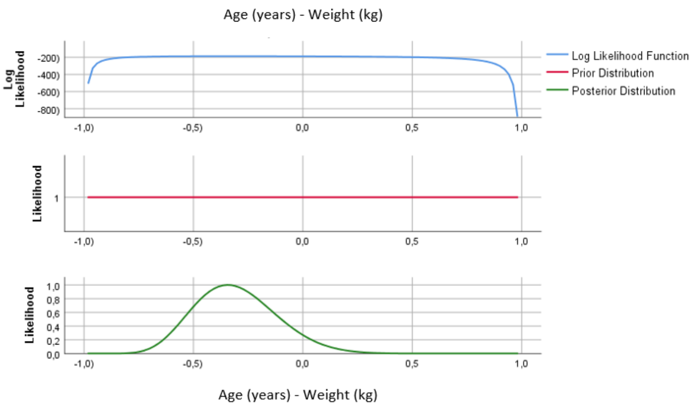

Figure 2. Bayesian posterior distribution and likelihood estimation for the relationship between age and body weight.

The posterior distribution demonstrated a weak negative tendency between age and body weight, although the posterior credible interval crossed zero, indicating insufficient Bayesian evidence for a reliable linear association.

3.4.3. Crosstab Analysis
Descriptive crosstab analyses were performed between body weight and:
seed consumption frequency,
bread-slice intake,
employment status,
sleep duration,
meal frequency.
The contingency tables contained numerous sparse cells with single observations and zero frequencies, limiting the reliability of classical inferential χ^2 association testing. Consequently, the crosstab results were interpreted descriptively rather than inferentially.
The distributions suggest heterogeneous behavioral and nutritional patterns across body-weight categories without clearly defined clustering tendencies.

3.5. Multiple Correspondence Analysis (MCA)
Multiple Correspondence Analysis (MCA) was applied as an exploratory multivariate method for identifying latent relationships among categorical variables. Convergence was achieved

after 36 iterations using variable-principal normalization. The first two dimensions produced eigenvalues of approximately 7.57 and 7.49, respectively, accounting for approximately 54% of the variance per dimension.

Table 5. MCA dimensional structure

Dimension	Eigenvalue	Explained variance (%)
Dimension 1	7.57	54.07
Dimension 2	7.49	53.47

MCA decomposes categorical variance through inertia estimation with equation:

$$I = \sum i \sum j \frac{(pij - ricj)2}{ricj}$$

where:

I - total inertia (total variance explained in MCA)
P_{ij} - joint proportion of category i and category j
R_i - marginal proportion (row mass) for category i
C_j - marginal proportion (column mass) for category j
$\sum$ - summation over all categories (rows and columns)

Discrimination measures indicated that age, body weight, daily bread consumption, nut frequency, and cooking oil usage contributed most strongly to the latent variance structure of the dataset. Given the relatively small sample size and ranking-based discretization of categorical variables in SPSS, MCA findings should be interpreted as exploratory and hypothesis-generating rather than confirmatory.

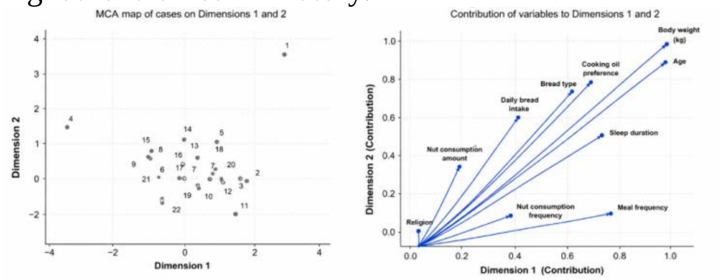

Figure 3. Multiple Correspondence Analysis (MCA) results showing participant distribution and variable contribution across the first two dimensions.

The MCA results presented in Figure 3 provide a multidimensional representation of both participant distribution and variable contribution within the analyzed nutritional and lifestyle dataset. The left panel illustrates the object-point map of participants projected onto the first two MCA dimensions, while the right panel presents the discrimination measures showing the contribution strength of individual variables to the latent multidimensional structure.

In the left panel, each numbered point represents an individual participant positioned according to the similarity of categorical dietary and lifestyle characteristics. Participants located closer to each other exhibit more similar response patterns, whereas participants positioned farther apart demonstrate greater heterogeneity in nutritional and behavioral profiles. Most participants are clustered around the central region of the multidimensional space, indicating moderate similarity among dietary habits and lifestyle behaviors within the sample. However, several observations, particularly cases 1 and 4, appear relatively distant from the primary cluster, suggesting more atypical participant profiles compared to the remaining sample population. These isolated cases may reflect differences in body weight, sleep duration, bread consumption, cooking oil preference, or other analyzed lifestyle variables.

The right panel presents the discrimination measures for the first two MCA dimensions. The length and direction of each vector indicate the relative contribution of variables to the explained multidimensional variance. Variables positioned farther from the origin demonstrate stronger discriminatory contribution to the latent structure of the dataset. The strongest contributions were observed for body weight and age, indicating that these variables

contributed most substantially to the separation of participants across the MCA dimensions. Cooking oil preference, bread type, daily bread intake, and sleep duration also demonstrated relatively strong contributions, suggesting that these nutritional and lifestyle characteristics play an important role in defining variability within the analyzed sample. In contrast, variables such as religion, nut-consumption frequency, and meal frequency exhibited weaker contributions, as reflected by their shorter vectors and closer proximity to the origin.
Overall, the MCA findings suggest that anthropometric characteristics together with selected dietary and lifestyle variables contribute substantially to the multidimensional variability of the analyzed population. Although the observed clustering tendencies are exploratory and limited by the relatively small sample size, the analysis provides useful insight into the latent structure of nutritional and behavioral patterns within the dataset.

Discussion

The present study explored dietary habits and lifestyle-related variables in relation to body weight using descriptive statistics, Bayesian inference, and exploratory multivariate analysis. The findings demonstrated that irregular meal frequency, short sleep duration, predominant sunflower oil consumption, and frequent white-bread intake were common characteristics within the analyzed sample. Although Bayesian analyses did not identify statistically reliable associations between age, meal frequency, and body weight, the multivariate MCA results suggested that anthropometric characteristics together with selected dietary variables contributed substantially to the latent variability structure of the dataset.
The observed predominance of sunflower oil and white-bread consumption reflects dietary behaviors commonly reported in transitional and middle-income populations, where refined carbohydrates and vegetable oils remain highly represented in everyday nutrition. Similar nutritional tendencies have been discussed in studies investigating obesity-related dietary patterns and metabolic health (Campos and Egea 2025). In addition, irregular meal frequency and short sleep duration observed in the present sample are consistent with previous findings suggesting that lifestyle-related behaviors may influence metabolic regulation and appetite control. Similar observations regarding obesity-related nutritional behaviors and mathematical analysis of obesity predictors were reported by Knights et al. (2024), where statistical and machine learning methods identified BMI and dietary variables as important contributors to obesity-related outcomes.
The Bayesian correlation analysis demonstrated only a weak negative tendency between age and body weight, with posterior credible intervals crossing zero. This finding indicates substantial uncertainty in the estimated relationship and suggests that age alone may not sufficiently explain body-weight variability within the analyzed sample. The broad credible intervals likely reflect the relatively small sample size and heterogeneous participant characteristics. Bayesian approaches are particularly useful in small exploratory datasets because they allow probabilistic interpretation of parameter uncertainty rather than relying exclusively on classical significance testing. Similar applications of mathematical and machine learning approaches in metabolic-disease prediction were presented by Mazlami and Antonska Knights (2025), where logistic regression and statistical modeling successfully identified important predictors of type 2 diabetes mellitus.
The MCA results provided additional insight into the multidimensional structure of the dataset. Body weight and age demonstrated the strongest discriminatory contribution to the first two MCA dimensions, followed by cooking oil preference, bread type, and sleep duration. These findings suggest that anthropometric and nutritional characteristics jointly contribute to latent behavioral structure within the analyzed population. Similar applications of exploratory multivariate approaches in nutritional epidemiology have shown that dietary variables

frequently cluster together and may reflect broader lifestyle patterns rather than isolated nutritional effects (Balakrishna et al. 2022). Furthermore, Markovikj and Knights (2022) demonstrated that optimization approaches in nutritional modeling may support more sustainable dietary planning and healthier nutritional structures.

The present findings are partially consistent with previous studies reporting beneficial effects of nuts and seeds on metabolic health and body-weight regulation. Nuts and seeds are rich in unsaturated fatty acids, antioxidants, dietary fiber, and micronutrients, which may contribute to improved satiety and lipid metabolism (Sabaté and Ang 2009). Systematic reviews further indicate that moderate nut consumption generally does not promote weight gain and may support long-term metabolic balance (Nishi et al. 2021). However, in the present study, nut-consumption frequency demonstrated relatively limited discriminatory contribution within the MCA analysis, potentially due to the small sample size and heterogeneous consumption patterns among participants.

The observed variability in body weight and dietary habits may additionally reflect differences in dietary adherence and behavioral responses among individuals. Similar findings were reported by Markovikj, Knights, and Gajdoš Kljusurić (2023), who demonstrated variability in body-weight reduction efficiency among overweight and obese adults following ketogenic dietary interventions. Their results further support the importance of individualized nutritional approaches and continuous monitoring of dietary behavior.

From a methodological perspective, the integration of Bayesian statistics, chi-square testing, and multivariate exploratory methods demonstrates the growing importance of computational approaches in nutritional and biomedical research. Experimental and statistical design methodologies similar to those used in the present study were discussed by Antoska Knights and Millaku (2023), emphasizing the applicability of mathematical and statistical frameworks in complex health-related datasets.

Several limitations should be considered when interpreting the findings. First, the study included a relatively small convenience sample, limiting statistical power and generalizability. Second, several categorical variables contained sparse frequencies and minor textual inconsistencies in questionnaire responses, which increased category fragmentation during SPSS processing. Third, the cross-sectional design does not permit causal interpretation of relationships between dietary habits and body weight. Consequently, the findings should be considered exploratory and hypothesis-generating rather than confirmatory.

Despite these limitations, the study demonstrates the applicability of Bayesian statistics and exploratory multivariate techniques such as MCA in nutritional and lifestyle research. The integration of descriptive statistics, Bayesian inference, chi-square testing, and multidimensional analysis provides a computational framework for investigating complex behavioral datasets containing both continuous and categorical variables. Future studies should include larger and more balanced population samples, standardized questionnaire structures, and additional nutritional biomarkers to improve statistical robustness and biological interpretation.

Conclusions

The present study investigated dietary and lifestyle behaviors in relation to body weight using descriptive statistics, Bayesian computational analysis, and Multiple Correspondence Analysis (MCA). The results demonstrated that irregular meal frequency, short sleep duration, sunflower-oil consumption, and white-bread intake were common characteristics within the analyzed sample. Bayesian analyses did not identify reliable associations between age, meal frequency, and body weight, while MCA indicated that body weight, bread consumption, and

cooking-oil preference contributed most strongly to the latent variability structure of the dataset.
Despite the relatively small sample size, the study demonstrates the applicability of Bayesian statistics and exploratory multivariate methods in nutritional and lifestyle research. The integration of mathematical and statistical approaches provides a useful computational framework for investigating complex behavioral datasets containing both continuous and categorical variables. Future studies should include larger population samples and additional nutritional and metabolic indicators to improve statistical robustness and predictive interpretation.

References

[1]. Antoska Knights V. and Millaku J., 2023. Three-factor experimental design as a tool in applied statistics. *International Journal of Statistics and Applied Mathematics*, Vol.8, No.1, Part A, pp.46-49. DOI: https://doi.org/10.22271/maths.2023.v8.i1a.929

[2]. Balakrishna R., et al., 2022. Consumption of nuts and seeds and health outcomes: A systematic review. *Nutrients*, Vol.14, Article 5049. Available at: Nutrients Journal Article

[3]. Campos S.B. and Egea M.B., 2025. Ingesting nuts can regulate adipokines expression in individuals living with overweight and obesity. *Nutrients*, Vol.17, No.13, Article 2138. Available at: MDPI Nutrients Article

[4]. Knights V., Blazevska T., Markovic G. and Gajdoš Kljusurić J., 2024. Mathematical analysis of statistical design of experiment and machine learning methods in identifying factors influencing obesity. *Medical Research Archives*, Vol.12, No.9. DOI: https://doi.org/10.18103/mra.v12i9.5790

[5]. Markovikj G. and Knights V., 2022. Model of optimization of the sustainable diet indicators. *Journal of Hygienic Engineering and Design*, Vol.39, pp.1-8. Available at: Key Publishing

[6]. Markovikj G., Knights V. and Gajdoš Kljusurić J., 2023. Body weight loss efficiency in overweight and obese adults in the ketogenic reduction diet program—Case study. *Applied Sciences*, Vol.13, No.19, Article 10704. DOI: https://doi.org/10.3390/app131910704

[7]. Mazlami V. and Antonska Knights V., 2025. Applied mathematical techniques and machine learning for identifying predictors of type 2 diabetes. *Journal of Natural Sciences and Mathematics*, Vol.10, No.19-20, pp.89-99. DOI: https://doi.org/10.62792/ut.jnsm.v10.i19-20.p3096

[8]. Mazllami V. and Antoska-Knights V., 2025. Примена на математички модели и техники за предикција на дијабетес тип 2. Годишна конференција (летна сесија) на студентите на трет циклус студии за презентација на докторски проекти, Универзитет „Св. Климент Охридски" Битола. Available at: UKLO PDF

[9]. Mazllami V. and Antoska Knights V., 2026. Навики на исхрана кај популација со дијабетес тип 2 и математичко моделирање и оптимизација на режим на исхрана. Втор докторски проект, Технолошко-технички факултет – Велес, Универзитет „Св. Климент Охридски" – Битола, Република Северна Македонија.

[10] Nishi S.K., Viguiliouk V., Blanco Mejia S., et al., 2021. Are fatty nuts a weighty concern? A systematic review and meta-analysis. *Obesity Reviews*, Vol.22, Article e13330. Available at: Obesity Reviews Article

[11] Sabaté J. and Ang Y., 2009. Nuts and health outcomes: New epidemiologic evidence. *American Journal of Clinical Nutrition*, Vol.89, No.5, pp.1643S-1648S. Available at: American Journal of Clinical Nutrition Article

[12] Simovska V., Antoska-Knights V., Kalevska T., Nedelkovska-Nikoloska D., Blazevska T., Damjanovski D. and Simovska A., 2016. Development of academic education of nutritionists in the Republic of Macedonia. *Advances in Nutrition*, Vol.7, No.1, p.24A. DOI: https://doi.org/10.1093/advances/7.1.24A