Coffee consumption of METU students: a synopsis of student habits, effects of coffee and coffee at campus

İren Su Çelik, Bora Esen, Ata Adanur, Emirhan Kıran

Faculty of Arts and Science, Department of Statistic, METU

2024/2025 Fall

Abstract

Aim – The purpose of this study is to examine the coffee consumption habits of METU students and the effects of drinking coffee on METU students, as well as the status of the coffee sold on the campus, while taking the principles of sustainability into consideration.

Design/methodology/approach – 287 METU students who drink coffee are the respondents to the questionnaire prepared on coffee consumption. The participants answered various questions on daily coffee consumption habits, the effects of coffee consumption, and how well the coffee on campus is.

Findings – METU students consume 1.73 cups of coffee daily and 2.82 cups daily on exam periods, spending 178TL on coffee daily. Most of the students prefer espresso-based coffee with a proportion of 44% and ceramic cups with a proportion of 34%. Students who prefer espresso-based coffee drink more coffee daily than students who prefer their coffee with milk or iced or Turkish coffee. 43% of the students find quality “fair” and price “in the middle”. Most METU students prefer to drink coffee in the morning, and they drink coffee to stay awake or as a social activity. The factor that affects students' buying choices the most is the quality of the coffee.

Originality/value – This study is the first to extensively analyze the general coffee consumption habits of METU students and how they are affected by coffee consumption as well as student remarks regarding to features of coffee sold in METU campus. The findings of this research may clarify future research on coffee consumption habits and effects.

Introduction

Coffee culture grows with around two billion cups consumed daily worldwide and the coffee industry faces pressures to improve its sustainability. The responsibility to improve sustainability does not just imply to the industry but also implies to consumers. Coffee consumers can change their coffee-drinking habits into sustainable coffee practices for the good of our planet and its inhabitants. Limiting coffee intake and avoiding over-consumption makes a good start and so does preparing an adequate amount of coffee to avoid waste. Rejecting single-use plastics, straws, and paper cups and switching from disposable to reusable cups can significantly reduce waste. In take-away situations, usage of thermos can be preferred, if not possible then recycling plastic and paper should be considered.

Coffee is a widely consumed beverage brewed by roasting the seeds of coffee plants in different levels, originating from various countries around the Equator. Different levels of roast and country of origin are what make its flavors unique. Coffee is one of the three

most popular beverages in the world other than tea and water. Although coffee is used as a base for various beverages ranging from lattes to Turkish coffee, its popularity is not just coming from the taste but also its refreshing or awakening effect created by the stimulant caffeine it contains. While this effect can change from one to another, it is what makes coffee popular among students around the world including METU students. METU students buy coffee from various places both inside and outside of the campus, drink it for different reasons including staying awake throughout the day or in exam weeks, and are affected by coffee in diverse ways. Further research on these differences is needed to have a deeper understanding of the effects of drinking coffee, opinions on coffee on campus, and the coffee consumption habits of METU students.

The aim of this research is to determine the general habits of drinking coffee of METU students, distinguish the effects of drinking coffee on the students, and thoughts on the features of coffee sold on METU campus. By examining the responds of randomly selected 287 coffee consumer METU students on the survey conducted on drinking coffee, the consumption of coffee among METU students can be generalized. Also, an inference about how responsible METU students are to improve sustainability can be made. The data obtained and the analysis of these data in this study may provide insight into the general coffee consumption habits of future research on students in Ankara, as well as at METU. Moreover, this study may give a rough idea of the effects of coffee consumption on students in further research.

Methodology

To analyze the coffee consumption of METU students, a questionnaire was prepared initially. The survey includes questions to determine the coffee consumption habits of the students with questions asking when, why, and how much coffee they drink daily on normal days and on exam weeks. Moreover, questions ask what factors influence them to buy coffee from specific places, how much they spend on coffee, which beverage they prefer, and what type of cup they prefer. The questionnaire also includes questions asking how coffee affects the students' academic performance and sleep duration to identify the effects of drinking coffee on students. Also, questions asking how students find the quality and price of coffee sold on campus are included to determine their opinions on the features of coffee at METU. The questionnaire is held short with 12 simple questions.

The initial survey was tested by conducting it on a few peers and changes in answer choices are made after the peer review according to their feedback. The biased answer choices and answers that do not cover all possibilities have been changed and the final version of the questionnaire has been formed with different answer choice types, including an open-ended question at the end, tick all that apply and tick one only choice having other choice.

The questionnaire was conducted on a total of 287 METU students who drink coffee in a time span of 21 days from 25th of November to 15th of December 2024. The students were chosen randomly at random places on METU campus such as the Physics' Building café, YP cafeteria, Devrim, and Çatı. Since all places and all students were chosen randomly, the sample is chosen with simple random sampling method. The survey administered by interviewers (us) is done in person, face-to-face with the respondents and the results are

recorded on smartphones directly by the interviewers. When students were asked whether they would participate, they rarely declined, hence the participation was high. The high participation is possibly due to the topic chosen since nearly every student consumes coffee and they might be interested in the topic. Also, the time spent completing the survey was short, so it might be another factor influencing the students to participate.

The data is transferred into an Excel spreadsheet where every row consists of answers of each student and columns represent the 12 questions. Rows that have null values are excluded from the data set. The null values resulted because 4 students did not want to or could not answer 1 different question each, so a new data set with a sample size equaling 283 is used in the study. Research questions in this study were selected from a group of research questions prepared and analyzed with proper tests. All research questions were determined logically and the ones with significant results, with valid assumptions, were chosen to be included in the study. To make an analysis, some new columns were added where ordinal answers (e.g. very negative, negative, neutral/no opinion, positive, very positive) have transformed into ranks from 1 to 5. Also, preferred beverage answers to the open-ended question included beverage names have been grouped into 4 groups: espresso-based coffee, coffee with milk, iced coffee, and Turkish coffee. The difference in daily coffee consumption during the exam period and the normal period is also typed in a new column. In this study, all tests and their assumptions are conducted via R. Their visualizations are done in Tableau and R and tables are created in Word.

Analysis

To measure the strength of association and the direction of the relationship between daily coffee consumption of students (in cups)(Daily) and daily coffee expenses of students (TL) (Expense), also between daily coffee consumption on exam period (in cups)(Daily on Exam) and change in sleep duration (minutes)(Sleep), correlation is used. Pearson correlation coefficient is appropriate to use since both variables in each association are quantitative, the relationship is linear and there are no outliers.

	Daily	**Expense**	**Daily on Exam**	**Sleep**
Sample Mean	1.73	177.915	2.82	30.96
Sample Standard Deviation	1.001	121.72	1.55	29.966

Table 1. Sample Means and Standard Deviations of the variables used in correlation analysis.

To measure the association between preferred coffee type and preferred cup, chi-square test is used since both variables are categorical. Assumptions are validated since all observations are independent, cells in the contingency table are mutually exclusive, and the expected value of cells should be 5 or greater in at least 80% of cells (in the formed contingency table, Table 4, 90% of the cells are greater than 5). Since chi-square test does not directly give information about the nature of the relationship, estimated conditional probabilities are investigated to get more information about the strength of the relationship.

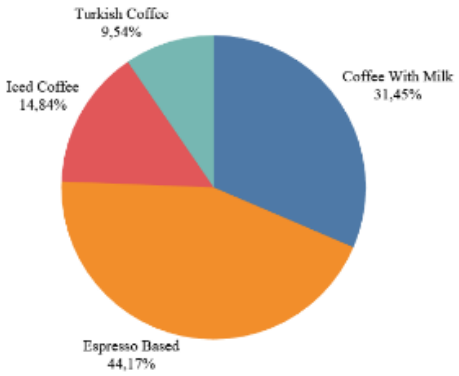

Figure 1. Preference of Coffee Types

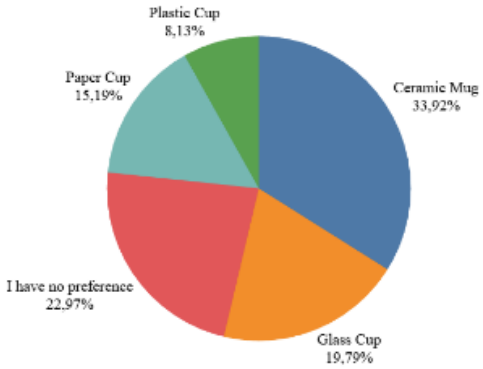

Figure 2. Preference of Cup Types

Sample proportions for preferred coffee types (Figure 1) show that in a sample of 283 METU students, approximately 44% prefer espresso-based coffee and 32% prefer coffee with milk to drink.

Sample proportions for preferred cup types (Figure 2) show that in a sample of 283 METU students, approximately 34% prefer ceramic mugs and 20% prefer glass cups to drink coffee. More than half of the sample uses reusable cups instead of disposable cups which may indicate that METU students feel responsible for improving sustainability. Further interpretation can be made in the analysis results.

To compare the mean number of daily coffee consumption of METU students across groups divided by coffee type preferences, One-Way ANOVA Test is appropriate to use. Individual Shapiro-Wilk Normality Tests were conducted to check the normality assumption and pairwise F-tests were conducted to check the assumption of equal variances ($\square_1^2=\square_2^2=\square_3^2=\square_4^2$). Since the samples are independent and both normality and homogeneity of variance assumptions are valid, One-Way ANOVA is used. To identify which specific groups have significantly different means from each other the posthoc test, Tukey HSD test is used.

	Espresso Based Coffee	**Coffee With Milk**	**Iced Coffee**	**Turkish Coffee**
Sample Mean	1.932	1.736	1.357	1.352
Sample Standard Deviation	0.903	1.113	0.945	0.897

Table 2. Sample Means and Standard Deviations (on daily coffee consumption) of the groups used in One-Way ANOVA Test.

To determine whether there is a linear relationship between daily coffee consumption of METU students on normal period and on exam period, simple linear regression (SLR) model is appropriate, since both independent and dependent variables are quantitative and there is only one independent variable (Table 1 shows sample means and standard deviations of used variables). There are no extreme values and all the assumptions of SLR are valid, so SLR model is used. The assumptions: Linearity, Homoscedasticity, Autocorrelation and Normality are checked by both plots and corresponding tests. Residual-Fitted Plot and RESET Test are used to check Linearity, Residual-Fitted Plot and Breusch–Pagan Test are used to check Homoscedasticity, Residual-Residual Plot and Durbin-Watson Test are used to check Autocorrelation, and Q-Q Plot and Shapiro-Wilk Test are used to check Normality assumption.

To determine whether there is a linear relationship between the dependent variable, daily coffee consumption, and the independent variables, daily coffee expenses of METU students and preferred coffee type, multiple linear regression (MLR) model is appropriate, since there are two independent variables. One of the independent variables is quantitative and the other one is categorical which is used in the model as dummy variable (Table 1 and Table 2 show sample means and standard deviations of used variables). MLR assumptions: Linearity, Homoscedasticity, Autocorrelation and Normality are checked with the same plots and tests used for checking SLR assumptions. No Multicollinearity assumption is checked by Kruskal-Wallis Test since one variable is quantitative and the other one is categorical. Point-Biserial Correlation was not appropriate to check Multicollinearity because the categorical variable is not dichotomous. Since all assumptions are valid and the predictors are linearly independent, no multicollinearity, MLR model is used.

To determine whether there is a linear relationship between academic performance of METU students and daily coffee consumption on exam period, ordinal logistic regression (OLR) model is appropriate, since the independent variable is measured on an ordinal level and the dependent variable is quantitative. Proportional Odds assumption is checked by Brant Test. This assumption basically means that the relationship between each pair of outcome groups must be the same. If the relationship between all pairs of groups is the same, then there is only one set of coefficients, which means that there is only one model. It is concluded that the parallel assumption holds since the probability (p-values) for all variables are greater than 0.05. Therefore, the proportional odds assumption is not violated and the model is a valid.

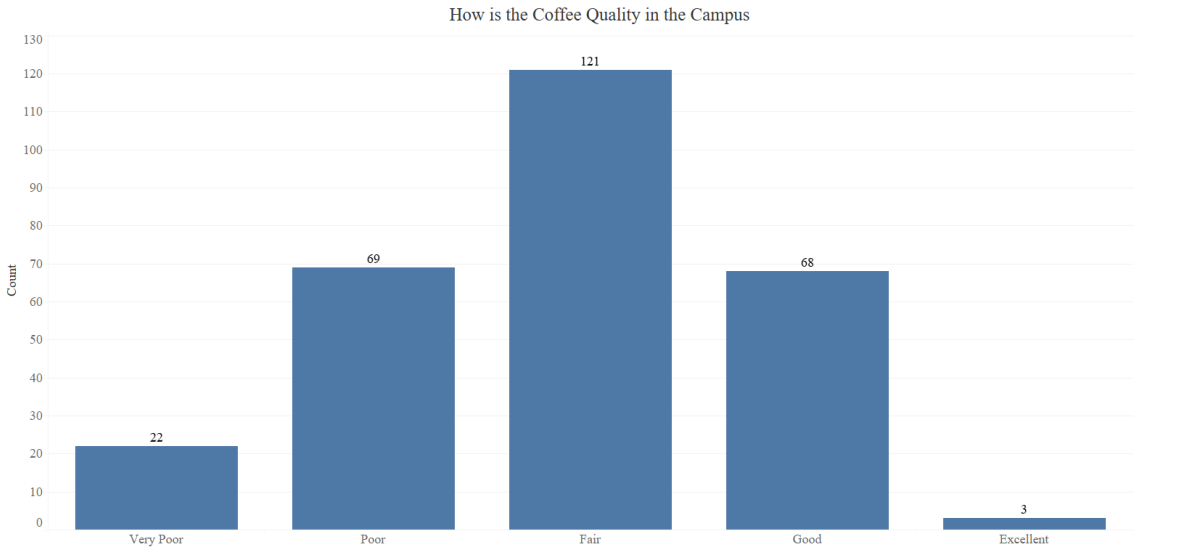

Figure 3. Coffee quality at METU campus

METU students' thoughts on the quality (Figure 3) and the prices (Figure 4) of coffee sold on campus are collected at an ordinal level. These data were not used in any statistical tests in this study, since the assumptions of the tests with these data sets are not valid. The bar charts show that 43% of the students in the sample find quality "fair" and price "in the middle". For quality, the number of students who find it poorly ("poor" and "very poor") is more than those who find it well ("good" and "excellent"). For price, the number of

students in the categories who find it expensive is more than the categories who find it affordable. Further research with more data can be done in order to make regulations of the features, quality and price, of the coffee sold on METU campus.

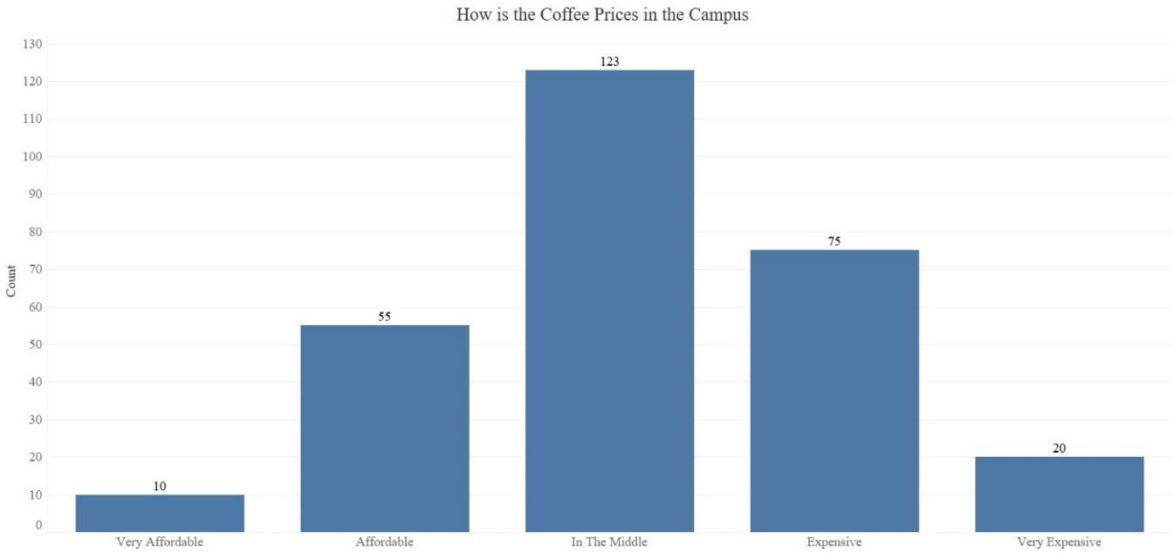

Figure 4. Coffee prices at METU campus

The Frequencies method is used to calculate 3 tick all that apply questions. Since these data are not independent and not mutually exclusive, they are not used in any statistical test. The Frequencies method is used to calculate the total number of votes for each preferred coffee consumption time of the day (Figure 5). The method is also used to calculate the total number of votes for each reason to drink coffee (Figure 6) and for each factor that effects METU students' choice when buying coffee (Figure 7).

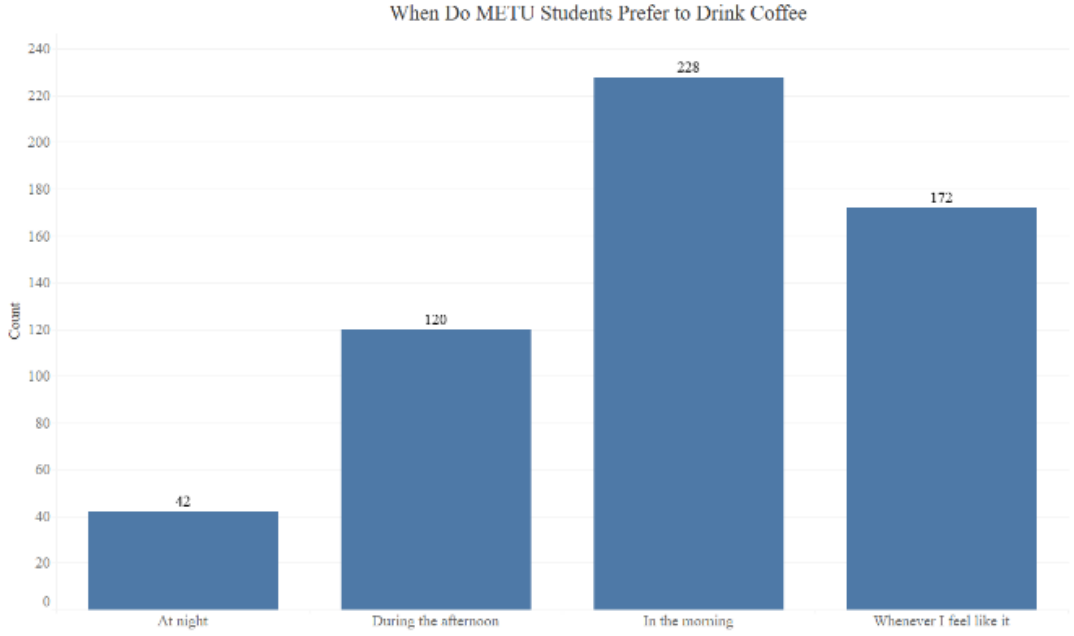

Figure 5. Time of day when METU students drink coffee

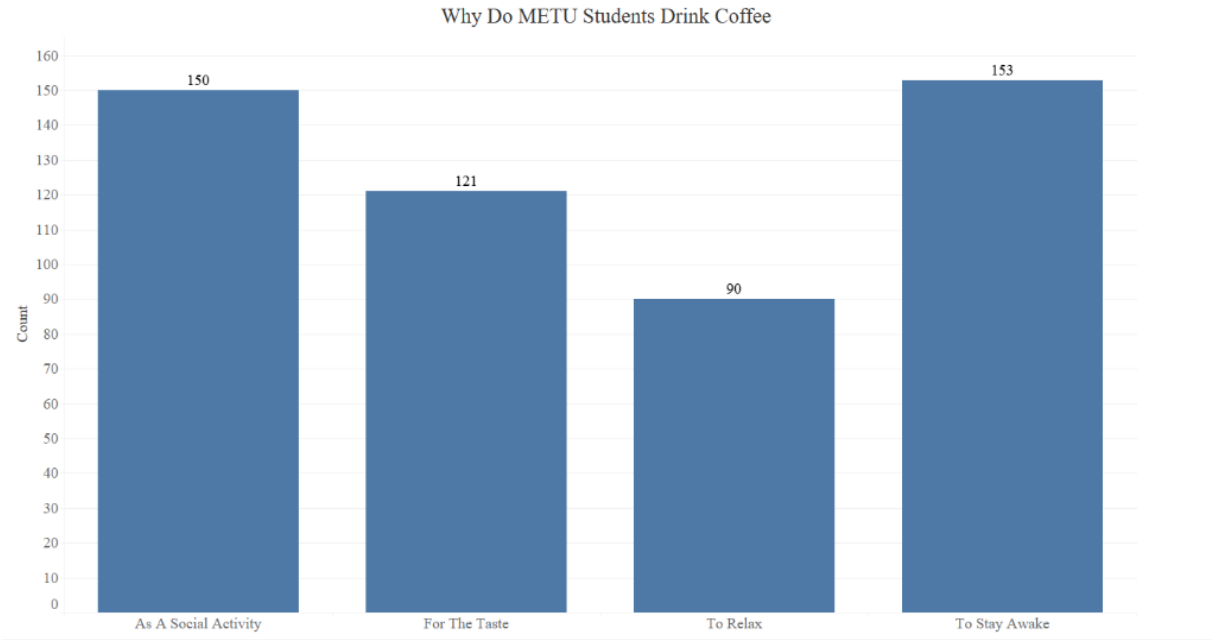

Figure 6. Reasons of METU students to drink coffee

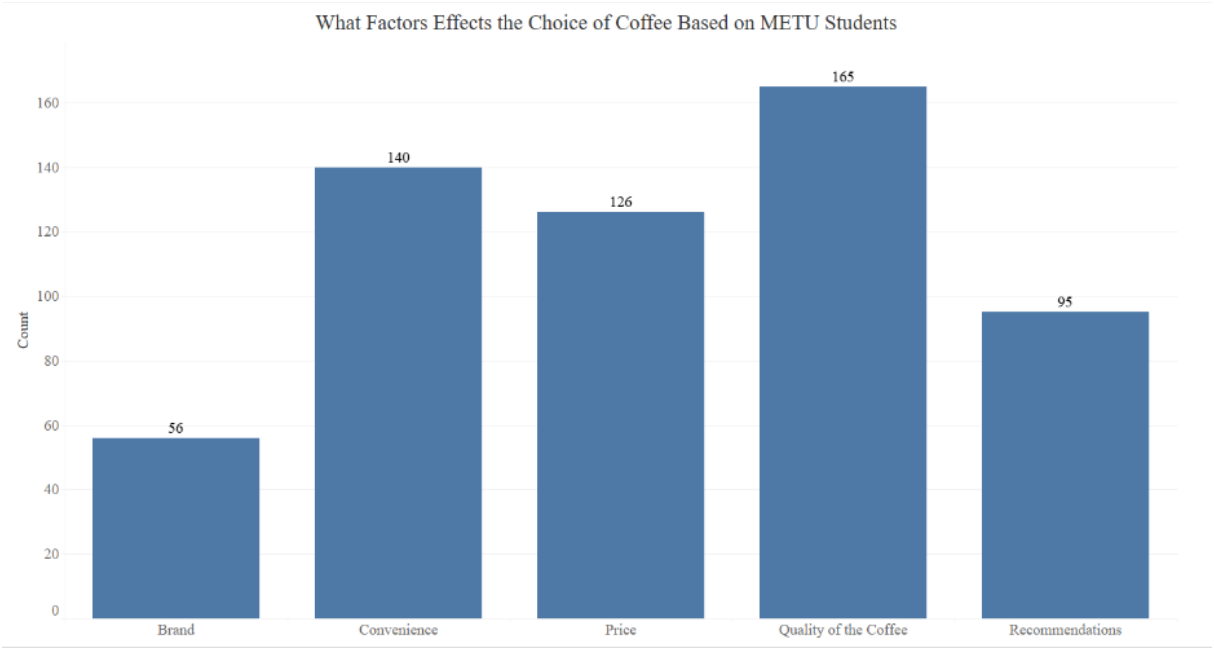

Figure 7. Factors influencing METU students' choice of coffee purchase

The sample shows that most of the METU students might prefer drinking coffee in the morning and their biggest reason to drink coffee might be to stay awake or as a social activity. The factor that affects METU students' choice the most might be the quality of the coffee.

Results

One-Sample Mean Hypothesis Test, t-test, is conducted for 3 variables: daily coffee consumption on normal period (Daily), daily coffee consumption on exam period (Daily on Exam), and daily coffee expenses of METU students (Expense). t-test is used since the sample variances are unknown. Table 3 shows the results of the tests.

	Lower Bound	Upper Bound	Sample Mean
Daily	1.61	1.85	1.73
Daily on Exam	2.64	3.00	2.82
Expense	163.672	192.158	177.915

Table 3. Sample Mean with Lower and Upper Bounds for Confidence Interval of the true mean.

The results of the t-tests show that the daily coffee consumption of METU students on normal period is in between (1.61, 1.85) cups and on exam period is in between (2.64, 3) cups, so the inference of coffee consumption increases on exam period can be made. Also, METU students spend between 164TL and 192TL daily on coffee. These results also show that METU students do not overconsume coffee and be responsible for improving sustainability.

The correlation coefficient between daily coffee consumption of students and daily coffee expenses of students equals 0.5394, so there is a moderate positive relationship between how much the METU students drink coffee and how much they spend on coffee.(Figure 8)

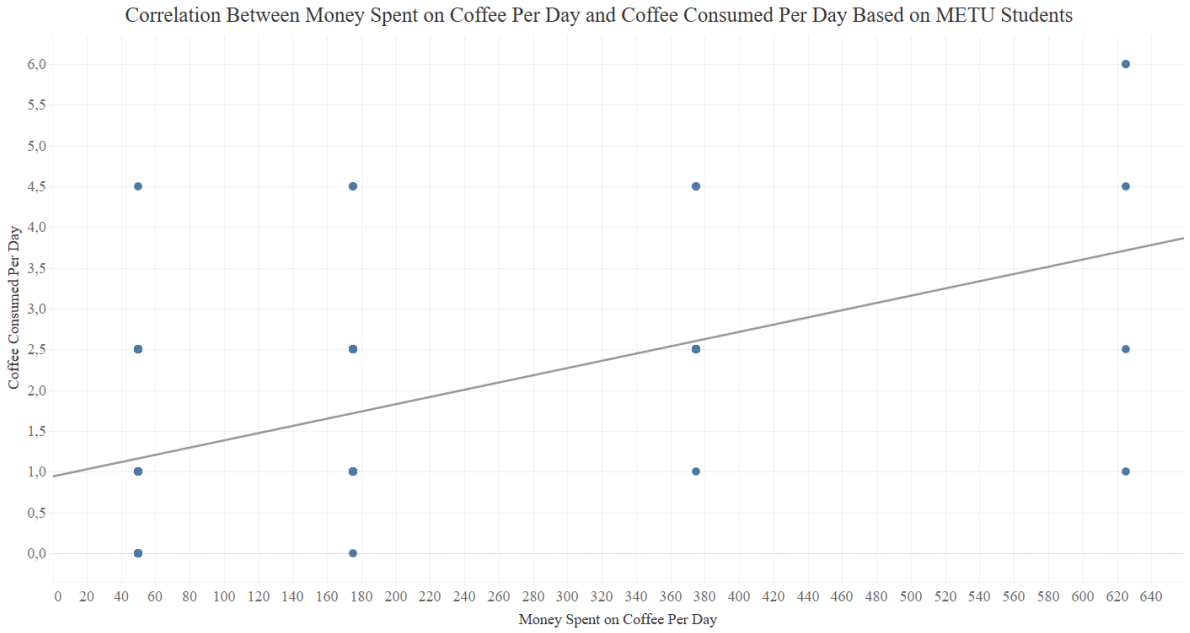

Figure 8. Correlation between daily coffee consumption and daily coffee expenses of students.

The correlation coefficient between daily coffee consumption of METU students on exam periods and the change in their sleep duration equals 0.2644, so there is a low positive relationship between how much the METU students drink coffee on exam periods and how much their sleep is affected by the coffee they consumed. (Figure 9)

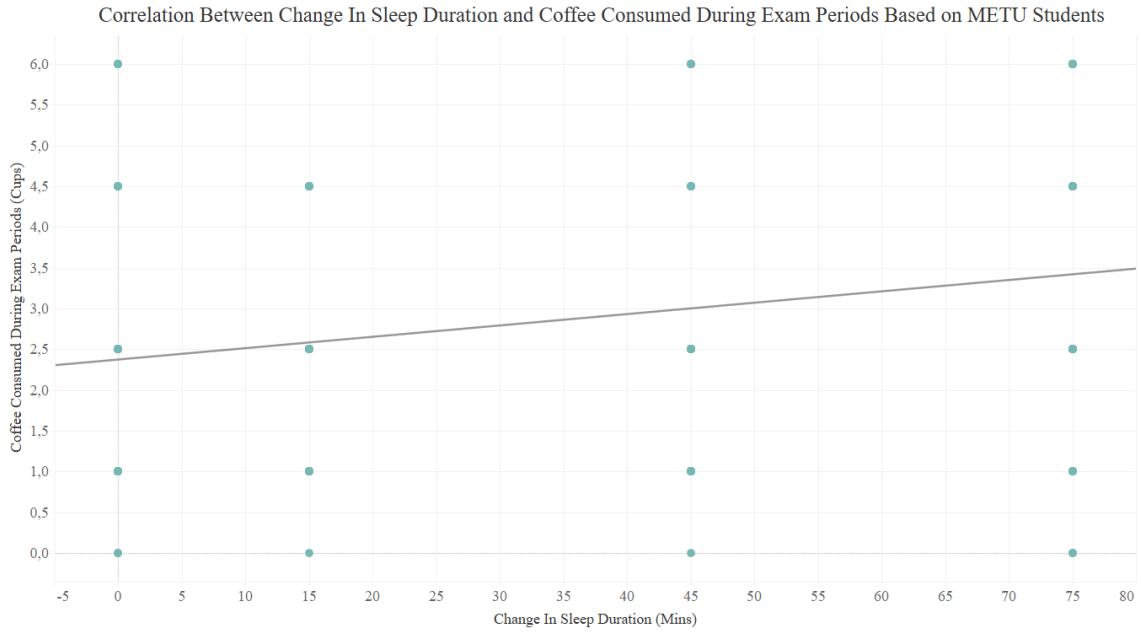

Figure 9. Correlation between daily coffee consumption on exams and change in sleep duration.

Chi-square test results between 2 categorical variables, coffee type preferred and cup type preferred by METU students are shown in the following contingency table (Table 4).

	Type of Coffee				
Type of Cup	Coffee With Milk	Espresso Based Coffee	Iced Coffee	Turkish Coffee	**Total**
Ceramic Mug	29	49	1	17	96
Class Cup	13	11	25	7	56
Paper Cup	17	24	2	0	43
Plastic Cup	5	6	11	1	23
No preference	25	35	3	2	65
Total	89	125	42	27	283

Table 4. Contingency Table of preferred coffee and cup type among METU students.

χ^2 equals 102.01 with p-value smaller than 0.05, so the data supports the conclusion that preferred coffee and cup type are dependent. To get information about the nature of this relationship, estimated conditional probabilities (Table 5) should be investigated.

	Type of Coffee			
Type of Cup	Coffee With Milk	Espresso Based Coffee	Iced Coffee	Turkish Coffee
Ceramic Mug	0.33	0.39	0.03	0.63
Class Cup	0.15	0.09	0.60	0.26
Paper Cup	0.19	0.19	0.04	0
Plastic Cup	0.05	0.05	0.26	0.04
No preference	0.28	0.28	0.07	0.07

Table 5. Estimated conditional probabilities of preferred coffee and cup type.

According to estimated conditional probabilities, reusable cups; ceramic mugs are used more likely by Turkish coffee consumers and glass cups are used more likely by iced coffee consumers. Disposable cups; paper cups are more likely to be used by both espresso based and coffee with milk consumers, plastic cups are more likely to be used by iced coffee

consumers. METU students prefer reusable cups more, showing their responsibility for improving sustainability. Students who prefer paper and plastic cups should re-think their choices for the good of our planet.

According to the results of One-Way ANOVA Test, the mean of daily coffee consumption of METU students across groups divided by coffee type preferences are not same for all, since the p-value is smaller than 0.05. Tukey HSD test is conducted to identify which specific groups have significantly different means from each other. The results show that the means of iced coffee consumers and espresso-based coffee consumers, and the means of Turkish coffee consumers and espresso-based coffee consumers are significantly different, the other means do not differ significantly. The pairwise mean difference between iced and espresso-based equals -0.575 cups, iced coffee consumers drink less coffee, where the difference in population means is between (-1.03, -0.12). The pairwise mean difference between Turkish coffee and espresso-based equals -0.58, Turkish coffee consumers drink less coffee, and the difference in population means is between (-1.18, -0.043). (Sample means and standard deviations are shown in Table 2)

The created simple linear regression model in Figure 10, showing the linear relationship of the variables: daily coffee consumption of METU students on normal period and consumption on exam period, is significant since the p-value of $\hat{B}_1$ is smaller than 0.05. Both $\hat{B}$'s are significant, so the least squares estimation equation is,

$\hat{y} = 1.02249 + (1.03807)*x$

Every 1 unit increase in daily coffee consumption on normal period increases consumption on exam period by 1.03807.

R^2 equals 0.45, so the regression model shows that 45% of the total variation in coffee consumption on exam period is explained by coffee consumption on normal period.

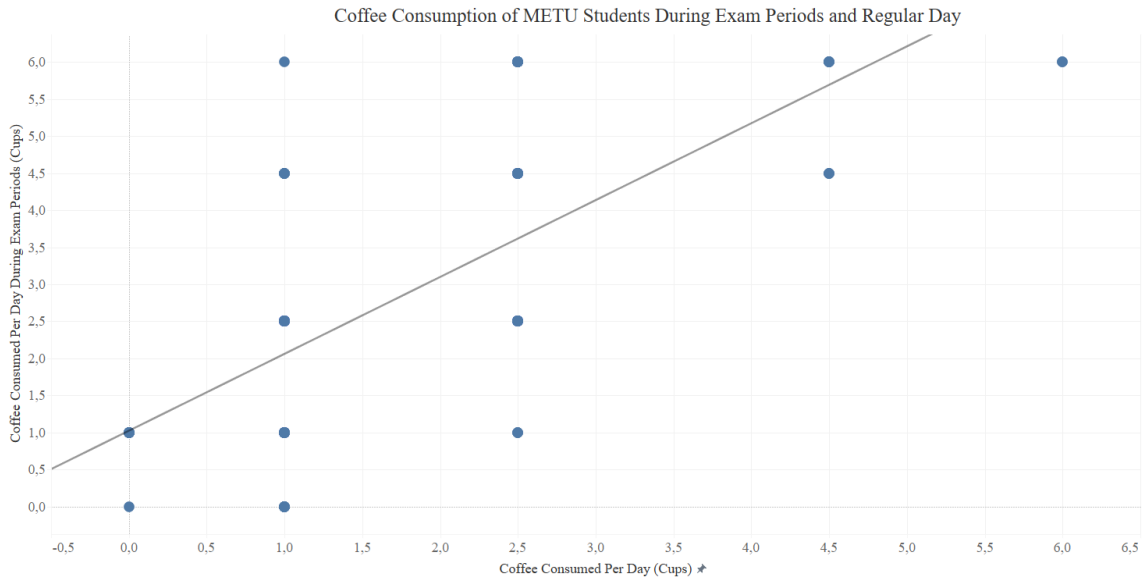

Figure 10. Simple Linear Regression of coffee consumption on regular day and exam periods.

The created multiple linear regression model in Figure 11, showing the linear relationship of the independent variable, daily coffee consumption of METU students on normal period, and the dependent variables, daily coffee expenses of METU students (X1) and preferred coffee type (D), is a significant model since p-value of $\hat{B}_1$ and some p-values of dummy variable are smaller than 0.05. The p-value is larger than 0.05 if the preferred coffee type is Turkish coffee, so it is excluded from the model because of its insignificancy. The least squares estimation equation is,

$\hat{y}$ = 0.9027362 + (0.0045149)*x1 + (0.2355525)*espresso-based – (0.453939)*iced

Expected change in daily coffee consumption on normal period per 1 unit change money spent buying coffee (X1) when D is held constant equals 0.0045149. If the preferred type is espresso-based coffee, then daily coffee consumption increases by 0.2355525 and if the preferred type is iced coffee, then daily coffee consumption decreases by 0.453939, when expense (X1) is held constant.

R^2 equals 0.3473, so the regression model shows that 34.73% of the total variation in coffee consumption on normal period is explained by expense and preferred type of coffee.

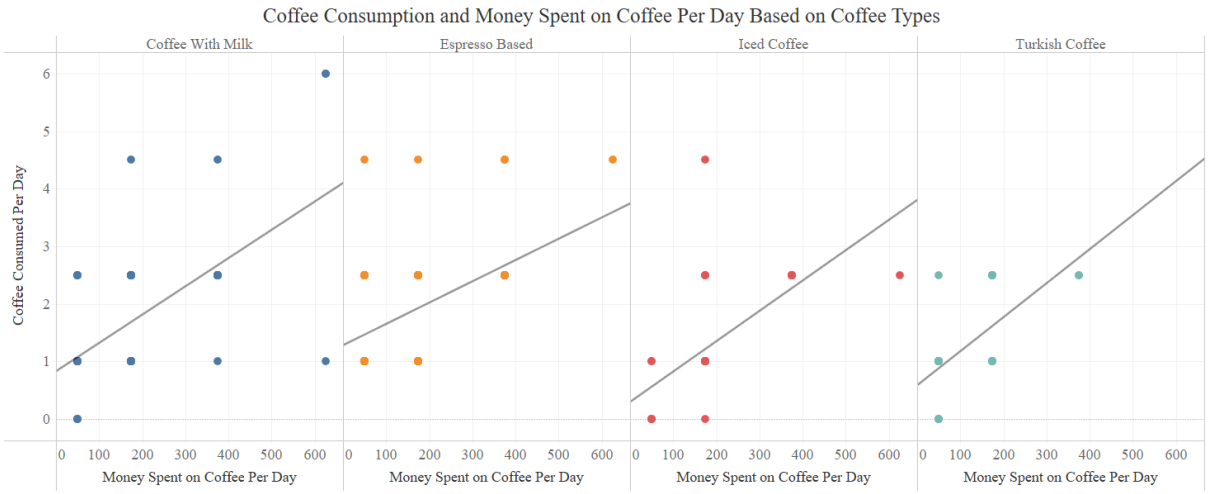

Figure 11. Multiple Linear Regression

The created ordinal logistic regression model in Figure 12 shows the linear relationship of the variables: the ordinal independent variable, effects on academic performance, and the quantitative dependent variable, daily coffee consumption on exam period (X). Since the confidence interval of daily coffee consumption on exam period is (0.444933, 0.7733219), not including 0, the dependent variable and the model is significant. The estimated model can be written as:

logit($\hat{P}(Y \leq 1)$) = -2.6886 – 0.6091 * X
logit($\hat{P}(Y \leq 2)$) = -1.6725 – 0.6091 * X
logit($\hat{P}(Y \leq 3)$) = 1.5095 – 0.6091 * X
logit($\hat{P}(Y \leq 4)$) = 3.7245 – 0.6091 * X

For every one unit increase in student's coffee consumption on exam period, the odds of being more likely to be affected on academic performance (*very positive or positive or no effect or negative versus very negative*) is multiplied by 1.84 (exp(0.6091)) times.

1.84 gives the odds ratio for obtaining a higher j for x_i in:

$$\text{logit}(P(Y \leq j)) = \ln(P(Y \leq j)/1 - P(Y \leq j)) = \alpha j - (\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_5 x_5)$$

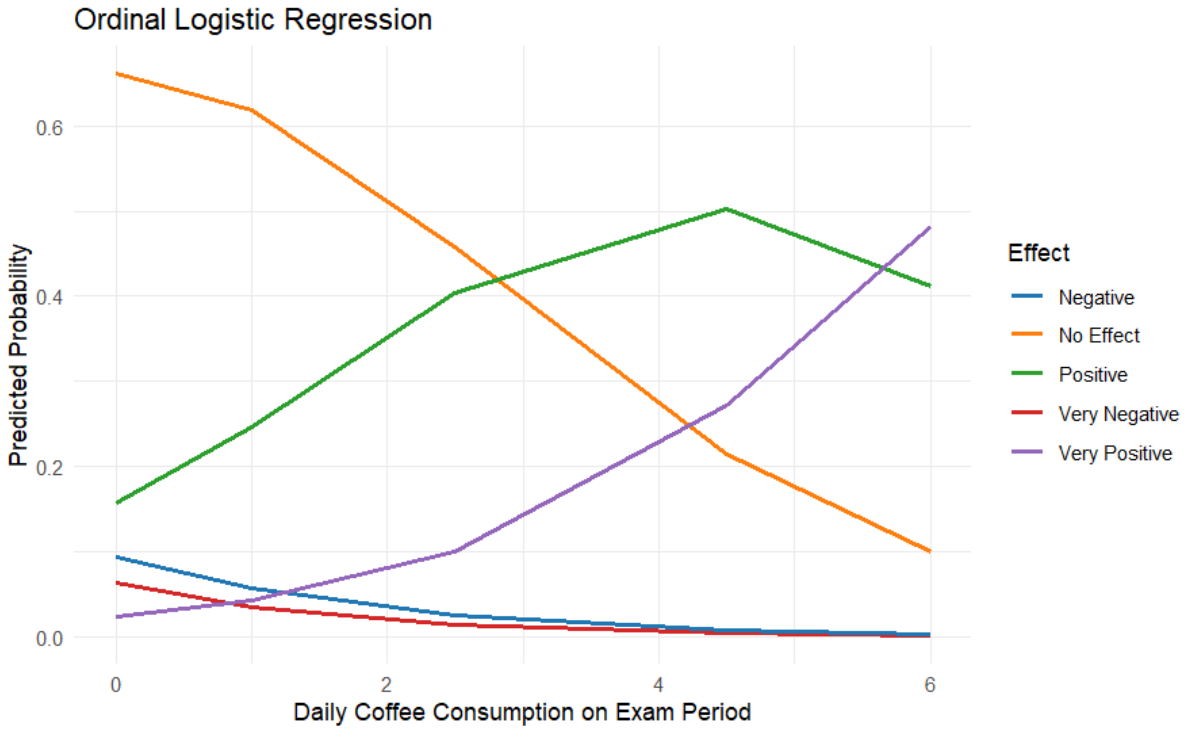

Figure 12. Ordinal Logistic Regression

Conclusion

This study informs about the general coffee consumption habits of METU students as well as the effects of drinking coffee on academic performance and sleep duration. Also, the students' remarks on the quality and the price of the coffee sold on METU campus are given in the study. The generalized habits of the students show that most of the students feel responsible for improving sustainability, by not overconsuming coffee and using reusable cups instead of disposables. The effects on one coffee consumption habit to another are also discussed in this study. Both the data obtained from the surveys conducted and the results of the analysis carried out in this study may provide insight into future research.

References

Barreto Peixoto, J.A. *et al.* (2022) 'Sustainability issues along the coffee chain: From the field to the cup', *Comprehensive Reviews in Food Science and Food Safety*, 22(1), pp. 287–332. Available at: https://doi.org/10.1111/1541-4337.13069.

Bartoloni, S., Ietto, B. and Pascucci, F. (2022) 'Do connoisseur consumers care about sustainability? Exploring coffee consumption practices through netnography', *British Food Journal*, ahead-of-print(ahead-of-print). Available at: https://doi.org/10.1108/bfj-07-2021-0814.

Lee, E. (2019) *Ordinal Logistic Regression on World Happiness Report*, *Medium*. Available at: https://medium.com/evangelinelee/ordinal-logistic-regression-on-world-happiness-report-221372709095.

Lund, B. (2021) 'The questionnaire method in systems research: an overview of sample sizes, response rates and statistical approaches utilized in studies', *VINE Journal of Information and Knowledge Management Systems* [Preprint]. Available at: https://doi.org/10.1108/vjikms-08-2020-0156.

McCombes, S. and George, T. (2019) *Write a Problem Statement for your Research in 3 Steps*, *Scribbr*. Available at: https://www.scribbr.com/research-process/problem-statement/.

Myhrvold, N. (2022) 'Coffee | beverage | Britannica', *Encyclopædia Britannica*. Available at: https://www.britannica.com/topic/coffee.

Ryan, E. (2022) *Research Objectives | Definition & Examples*, *Scribbr*. Available at: https://www.scribbr.com/research-process/research-objectives/.

Thulin, M. (2024) *8 Regression models | Modern Statistics with R*, *Modernstatisticswithr.com*. Available at: https://modernstatisticswithr.com/regression.html#logreg.

UCLA: Statistical Consulting Group (2021) *Ordinal Logistic Regression | R Data Analysis Examples*, *stats.oarc.ucla.edu*. Available at: https://stats.oarc.ucla.edu/r/dae/ordinal-logistic-regression/.

Yeager, K. (no date) *LibGuides: SPSS Tutorials: Working with 'Check All That Apply' Survey Data (Multiple Response Sets)*, *libguides.library.kent.edu*. Available at: https://libguides.library.kent.edu/SPSS/Multiple-Response-Sets.

Zhao, S. *et al.* (2024) 'Reducing carbon footprint of typical coffee consumption from the whole lifecycle viewpoint', *Environmental Impact Assessment Review*, 106, pp. 107476–107476. Available at: https://doi.org/10.1016/j.eiar.2024.107476.