DietR –
A dietary analysis tool for ASA24 and NHANES in R

Rie Sadohara[1], David Jacobs[2], Mark A. Pereira[2], Abigail Johnson[2]
[1]Kyoto-city, Kyoto, Japan [2]Division of Epidemiology & Community Health, University of Minnesota

Background

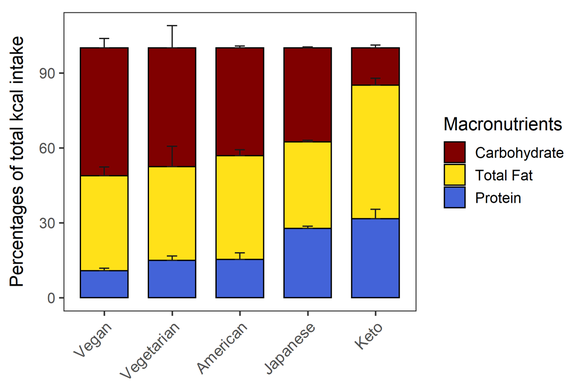
- Analysis of 24-hour recall data can be complicated and difficult.
- Many dietary datasets and dietary analysis tools are written in SAS.
- R is open-source and customizable with packages.
- We developed a package “DietR” to analyze NHANES and ASA24 data with R.

Functionality of DietR

Data preparation	Load, filter, compute total food intake for each participant, compute means of food intake across days/groups, filter the total data for outliers.
Data overview	Data summary, % KCAL by macronutrients in barcharts.
Diversity	Compute α−diversity indices for dietary records, participants, or food groups.
Clustering	Principal component analysis (PCA), *k*-means, select the optimal *k*
Foodtree	Build foodtrees [1] where foods in FNDDS are hierarchically grouped, visualize foodtrees, generate individual food consumption tables (“vegan” package [2]).
Ordination	Principal Coordinate Analysis (PCoA) based on their food consumption amount and the similarity of foods taken into account (“vegan” package [2]).

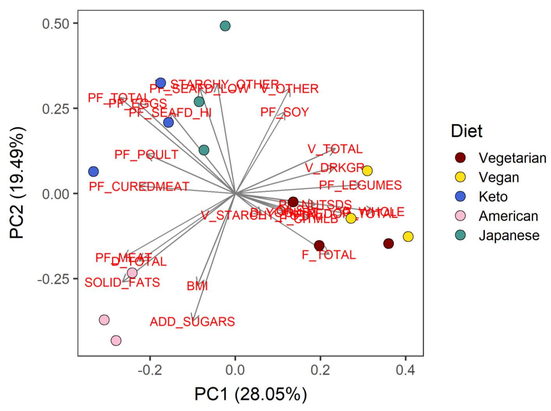
Demonstration

Data overview



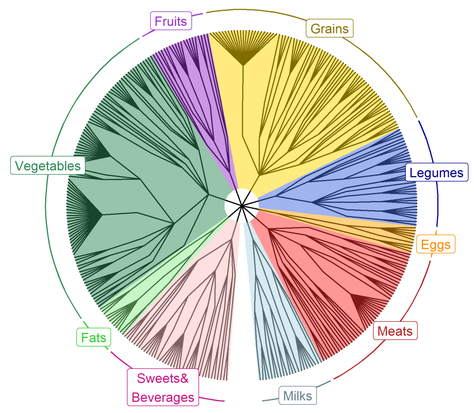
%KCAL from three macronutrients. Error bars: standard deviation.

Clustering



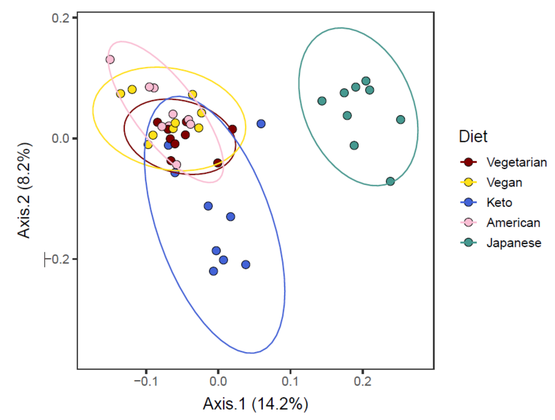
PCA based on food categories averaged across 3 days.

Foodtree



4-level hierarchical grouping of all reported food items.

Ordination



PCoA with consumption and food hierarchy.

Figure 1: Examples of plots created with DietR using a set of simulated ASA24 dietary records designed to show differences in eating patterns. The example dataset includes 15 imagined people with 5 different diets: Vegetarian, Vegan, Keto, American, and Japanese.

Use case vignette: nuts/seeds/legumes diversity & body measures

Background and Research question

- Previous studies suggest nuts/seeds/legumes have positive impacts on health [3].
- Is diversity of nuts/seeds/legumes consumption related to body measures, e.g. BMI or waist size?

Methods

- NHANES 2015−16, n=3,641, 18+ yo, with waist circumference & BMI.
- Extracted reported food items with their foodcodes starting with 4 (Foodcode 4xxxxxxxx: nuts/seeds/legumes) from two days of recalls.
- Calculated nuts/seeds/legumes α−diversity.
- Defined diversity groups (**Table 1**).
- Analysis of covariance (ANCOVA) with Age, Gender, Income, Education, KCAL as covariates

DivGroup	n	nuts/seeds/legumes consumed	Diversity index
DivNA	1,819	0	NA
Div0	1,105	1	0
Div1	360	>1	0.027 − 0.66
Div2	357	>1	0.66 − 1.95

Table 1: α−diversity groups. DivNA represents no intake of nuts/seeds/legumes. Div0 are individuals who consumed 1 type of nuts/seeds/legumes. Div1 and Div2 consumed more than 1 type of nuts/seeds/legumes.

Results & Discussion

- More diverse nuts/seeds/legumes consumption is associated with lower waist circumference.
- Div2 had 3.8 cm lower waist circumference than DivNA (p<0.001) and 3.4 cm lower than Div0 (p<0.01).
- Div2 had 1.4 lower BMI than DivNA and Div0 (p<0.01 for both).
- In contrast, higher KCAL intake was associated with increased nuts/seeds/legume diversity.
- Pulse intake in NHANES is associated with better quality diets [4]. Thus, nuts/seeds/legumes diversity could be a useful index to explore the health-promoting effects of this food group.
- Physical exercise, drinking, and smoking habits may be confounders.
- 2 day-data may have been insufficient to capture nuts/seeds/legumes diversity.

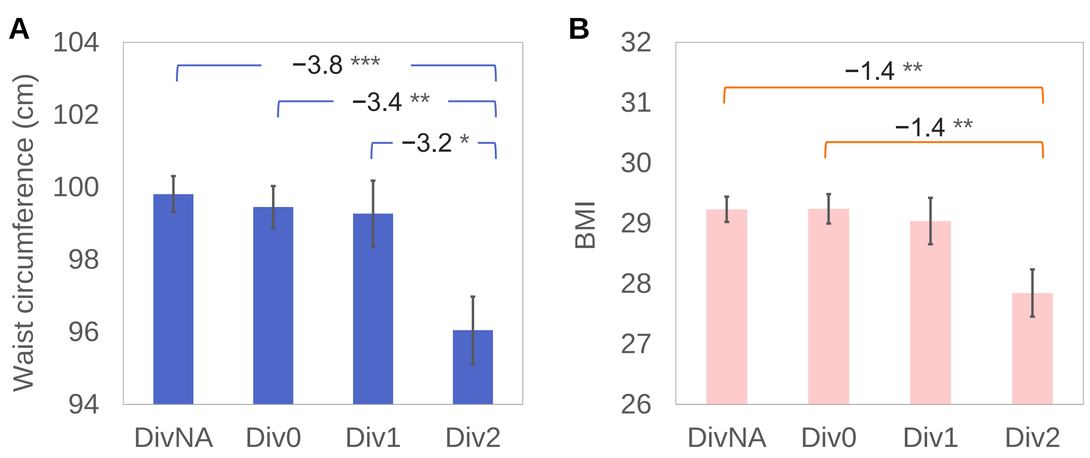


Figure 2: Bar charts showing emmeans ± SE for ANCOVA models for (A) waist circumference and (B) BMI; pairwise differences shown are significantly different.
***: p<0.001, **: p<0.01, *: p<0.05.

Availability

- GitHub repo
  https://github.com/computational-nutrition-lab/DietR
- Website with tutorials
  https://computational-nutrition-lab.github.io/DietR/
- Preprint on medRxiv
  https://doi.org/10.1101/2023.07.07.23292390

References

[1] Johnson AJ, Vangay P, Al-Ghalith GA, et al. Daily sampling reveals personalized diet-microbiome associations in humans. Cell Host Microbe. 2019;25(6):789-802.
[2] Simpson GL, Minchin PR, De Caceres M, et al. vegan: Community Ecology Package. 2022.
[3] Karlsen MC, Ellmore GS, McKeown N. Seeds—Health benefits, barriers to incorporation, and strategies for practitioners in supporting consumption among consumers. Nutr Today. 2016;51(1):50-59.
[4] Mitchell DC, Marinangeli CPF, Pigat S, et al. Pulse intake improves nutrient density among US adult consumers. Nutrients. 2021;13(8):2668.

Acknowledgements

This project was supported by internal institutional start-up funds from the University of Minnesota. The authors would like to thank **Mo Hutti** for the create_corr_frame function which generates a correlation table with ordination axes and variables; **Pajau Vangay** for the collapse_by_correlation function which removes correlated variables; and **Suzie Hoops** for the matrix multiplication operation and her insights into statistical analyses.