# DietR-

# A dietary analysis tool for ASA24 and NHANES in R



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





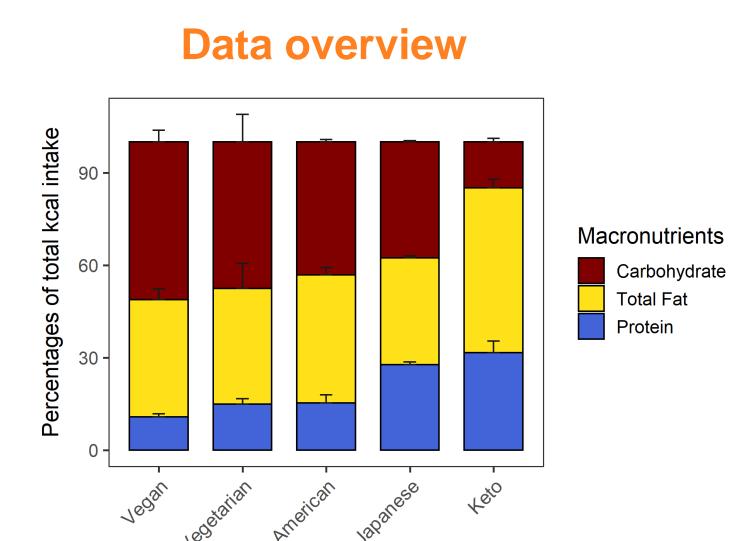
#### Background

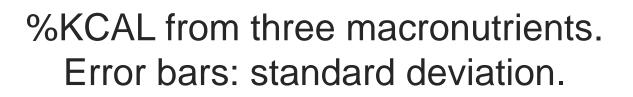
- Analysis of 24-hour recall data can be complicated and difficult.
- Many dietary datasets and dietary analysis tools are written in SAS.
- R is open-source and customizable with packages.
- We developed a package "DietR" to analyze NHANES and ASA24 data with R.

## Functionality of DietR

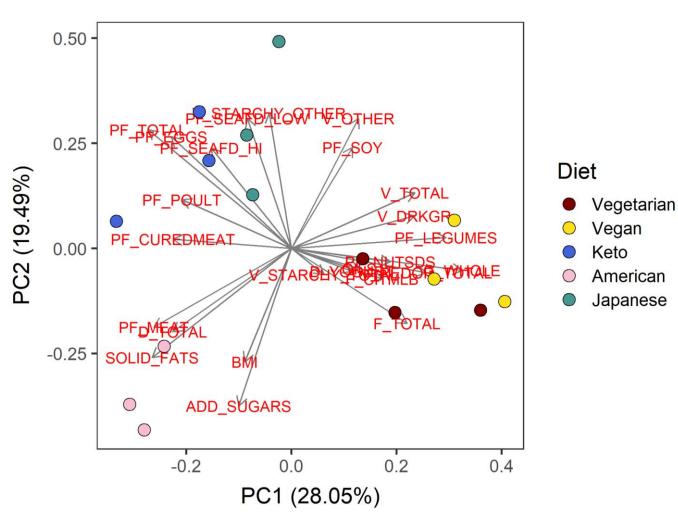
Data preparation	Load, filter, compute total food intake for each participant, compute means of food intake across days/groups, filter the total data for outliers.
Data overview	Data summary, % KCAL by macronutrients in barcharts.
Diversity	Compute $\alpha$ -diversity indices for dietary records, participants, 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



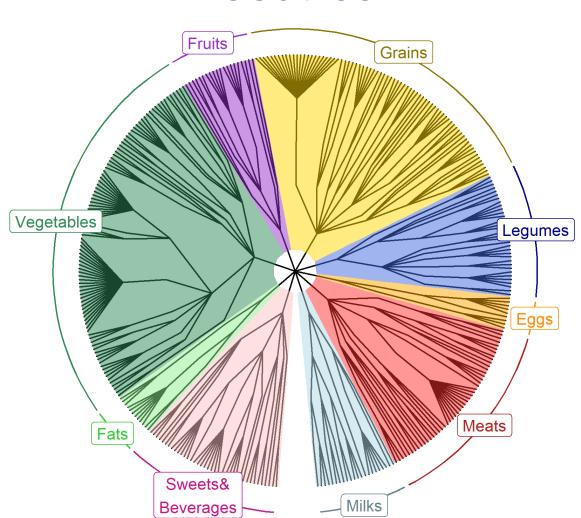


# Clustering



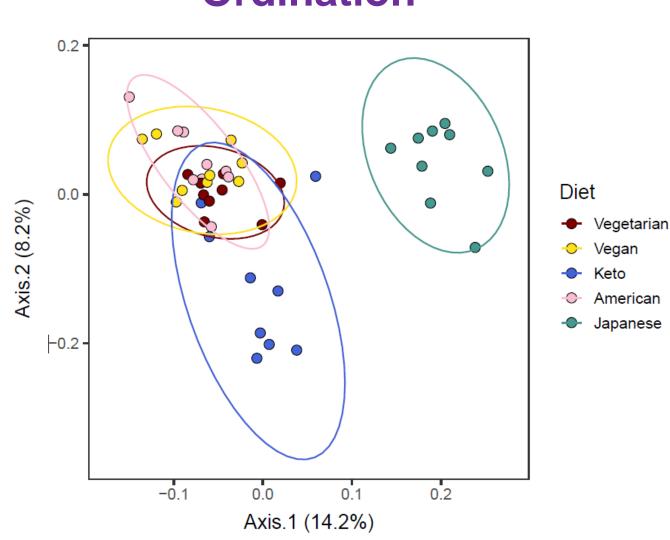
PCA based on food categories averaged across 3 days.

#### 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. waist size or BMI?

# 

https://github.com/computational-nutrition-lab/DietR

Website with tutorials

GitHub repo

Availability

References

sampling

2019;25(6):789-802.



https://computational-nutrition-lab.github.io/DietR/

[1] Johnson AJ, Vangay P, Al-Ghalith GA, et al. Daily

[2] Simpson GL, Minchin PR, De Caceres M, et al. vegan:

[3] Karlsen MC, Ellmore GS, McKeown N. Seeds—Health

[4] Mitchell DC, Marinangeli CPF, Pigat S, et al. Pulse

benefits, barriers to incorporation, and strategies for

practitioners in supporting consumption among

intake improves nutrient density among US adult

personalized

Preprint on medRxiv

reveals

associations in humans. Cell

Community Ecology Package. 2022.

consumers. Nutr Today. 2016;51(1):50-59.

consumers. Nutrients. 2021;13(8):2668.



https://doi.org/10.1101/2023.07.07.23292390



diet-microbiome

Host Microbe.

#### Calculated nuts/seeds/legumes α-diversity.

nuts/seeds/legumes) from 2 days of recalls.

Extracted reported food items with their

- Defined diversity groups (Table 1).
- Analysis of covariance (ANCOVA) with Age, Sex, Race, Income, Education, KCAL as covariates

NHANES 2015–16, n=3,641, 18+ yo, with waist

foodcodes starting with 4 (Foodcode 4xxxxxxxxx:

#### nuts/seeds/legumes Diversity DivGroup consumed index DivNA 1,819 NA Div0 1,105 0.027 - 0.66360 Div1 >1 Div2 357 0.66 - 1.95>1

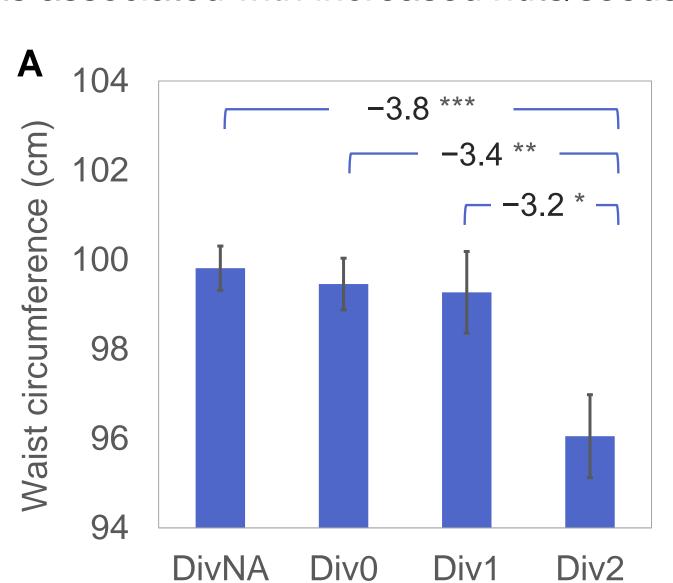
**Table 1:** α-diversity groups. DivNA represents no intake of nuts/seeds/legumes. Div0 are individuals who consumed 1 type of nuts/seeds/legumes. Div1 and Div2 consumed more than 1 type of nuts/seeds/legumes.

#### Results & Discussion

circumference & BMI.

Methods

- 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.



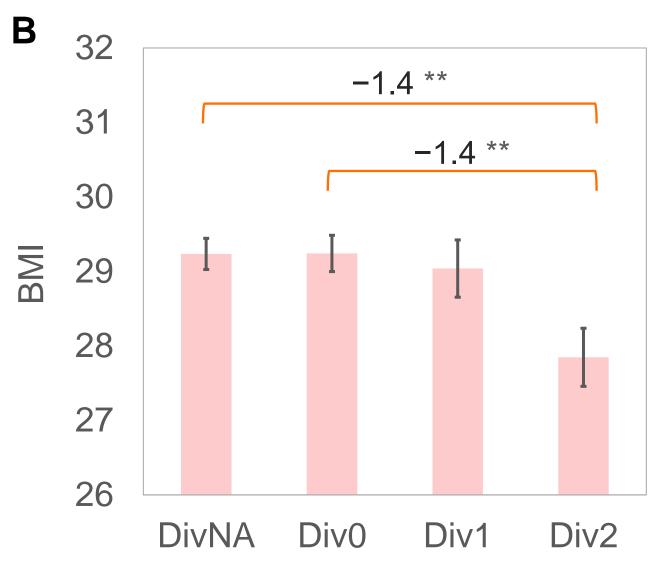


Figure 2: Bar charts showing emmeans ± SE for ANCOVA models for (A) waist circumference and (B) BMI; pairwise differences shown are significantly different. \*\*\*: *p*<0.001, \*\*: *p*<0.01, \*: *p*<0.05.

## 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.