2023.05.19 Waist2 method, after talking with Mark and David via email for the 2nd time.

ANCOVA Model

Response = μ + DivGroup + Age + Gender + Ethnicity + Income + Education + kcal\*  + ε,

1 2 3 4 5 6 7

where DivGroup is the treatment effect, and Age, Gender, Ethnicity, Income, Education, and kcal intake were covariates, and ε is the error term.

QC-ed Total ()

|

No missing data in waist and BMI, and variables 1-7

|

Remove the 361 missing data in Income.

|

Make an OTU table

|

Remove outliers in nuts/seeds/legumes consumption amount in the OTU table.

|

Calc divGroup

|

Plot a boxplot of amount x DivGroup.

|

Make table of Demographic (var. 2-7) x DivGroup table. Percentages +- SD.

|

Make table of foods: PF\_LEG, PF\_TOTAL, V\_TOTAL, KCAL etc. x DivGroup.

There may be outliers that are only visible when I made divGroup table, and averages may be affected. If there is an outlier in those food variables, just remove them individually, I guess.

|

Kcal is different among DivGroups…

|

Make a table of nutrients: CARB, PUFA, etc./ 2000 kcal x DivGroup.

|

Test Waist vs. DivGroup and BMI vs. DivGroup using the ancova model shown above.

Done!!

Variable list <https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/DEMO_I.htm#RIAGENDR>

|  |  |
| --- | --- |
| BMI | BMXBMI |
| WAIST | BMXWAIST |
| Age | RIDAGEYR |
| Gender | RIAGENDR |
| Ethnicity | RIDRETH3 |
| Family IPR | INDFMPIR |
| Edu (<20 yo) | DMDEDUC3 |
| Edu (>=20) | DMDEDUC2 |
| kcal | KCAL |