동의보감

신승균, 이성주, 신익규



Outline

- Overview
- Unit Conversion
- Logic [Filter & Clustering]
- Visualizations Shiny

Unit Conversion

Nutrient	Current DV Unit*	Current Conversion*	New DV Unit (required by 2019- 2020)	New Conversion (required by 2019-2020)
Vitamin A	IU	1 IU = 0.3 mcg retinol 1 IU = 0.6 mcg beta-carotene	mcg RAE	1 mcg RAE = 1 mcg retinol 1 mcg RAE = 2 mcg supplemental beta-carotene 1 mcg RAE = 12 mcg beta-carotene 1 mcg RAE = 24 mcg alpha-carotene 1 mcg RAE = 24 mcg beta-cryptoxanthin
Vitamin E	IU	1 IU = 0.67 mg for <i>d</i> -alpha- tocopherol (natural) 1 IU = 0.9 mg for <i>dl</i> -alpha- tocopherol (synthetic)	mg alpha-tocopherol	1 mg vitamin E (as alpha-tocopherol) label claim = 1 mg of natural α -tocopherol 1 mg vitamin E (as alpha-tocopherol) label claim = 2 mg of synthetic α -tocopherol
Vitamin D	IU	1 IU = 0.025 mcg	mcg	1 IU = 0.025 mcg
Folate	mcg		mcg DFE	1 mcg DFE = 1 mcg folates 1 mcg DFE = 0.6 mcg folic acid
Niacin	mg		mg NE	1 mg NE = 1 mg niacinamide 1 mg NE = 1 mg inositol hexanicotinate 1 mg NE = 1 mg niacin 1 mg NE = 60 mg tryptophan

IU = International Unit

RAE = Retinol Activity Equivalents

DFE = Dietary Folate Equivalents

NE = Niacin Equivalents

Source: https://dsid.od.nih.gov/Conversions.php

Literature Review: United States

Dietary Reference Intakes (DRIs):
Recommended Dietary Allowances and Adequate Intakes, Vitamins

Life Stage Group	Vitamin A (μg/d) ^{<u>a</u>}	Vitamin C (mg/d)	$\frac{Vitamin\ D}{(\mu g/d)^{\underline{b},\underline{c}}}$	Vitamin E (mg/d) ^d	Vitamin K (μg/d)	Thiamin (mg/d)
Infants						
0–6 <u>mo</u>	400*	40*	10*	4*	2.0*	0.2*
6–12 <u>mo</u>	500*	50*	10*	5*	2.5*	0.3*
Children						
1-3 <u>y</u>	300	15	15	6	30*	0.5
4–8 <u>y</u>	400	25	15	7	55*	0.6
Males						
9–13 <u>y</u>	600	45	15	11	60*	0.9
14–18 у	900	75	15	15	75*	1.2
19–30 у	900	90	15	15	120*	1.2
31–50 <u>y</u>	900	90	15	15	120*	1.2
51–70 <u>y</u>	900	90	15	15	120*	1.2
> 70 <u>y</u>	900	90	20	15	120*	1.2

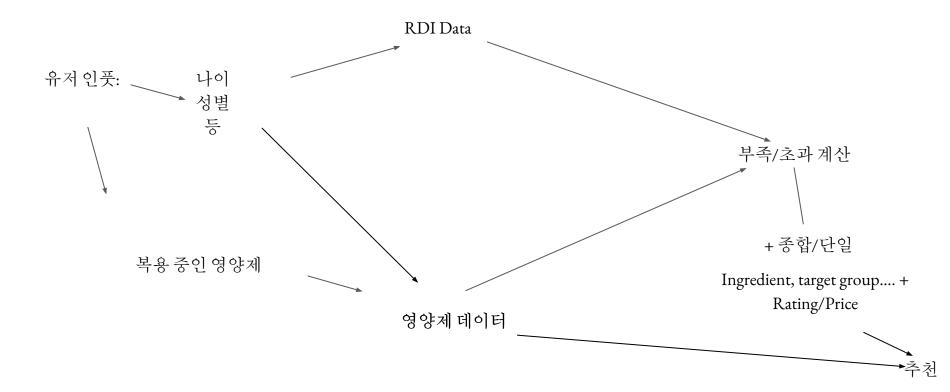
15/0.025=600 IU

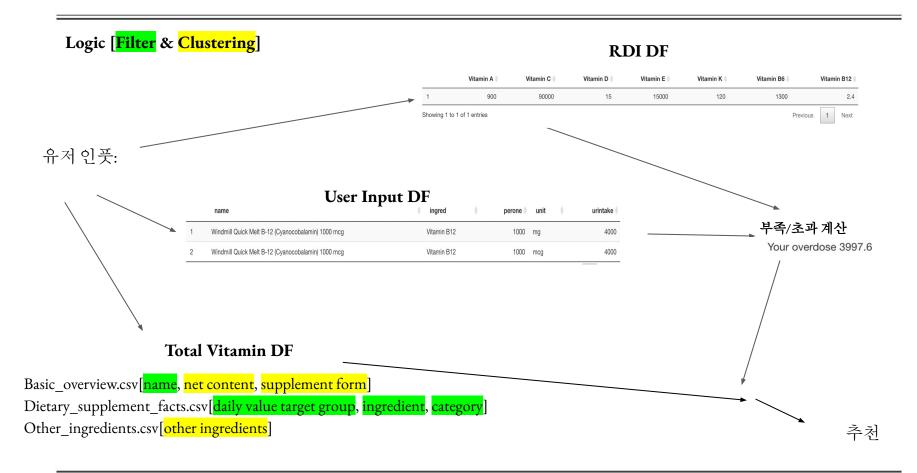
Dietary Reference Intakes (DRIs): Tolerable Upper Intake Levels, Vitamins

Life Stage Group	Vitamin A (μg/d) ^a	Vitamin C (mg/d)	Vitamin D (μg/d)	Vitamin E (mg/d) ^{b,c}	Vitamin K	Thiamin
Infants				1		
0–6 <u>mo</u>	600	<u>ND</u> e	25	<u>ND</u>	ND	ND
6–12 <u>mo</u>	600	ND	38	<u>ND</u>	ND	ND
Children						
1-3 у	600	400	63	200	ND	ND
4–8 <u>y</u>	900	650	75	300	<u>ND</u>	ND
Males						
9–13 у	1,700	1,200	100	600	ND	ND
14–18 <u>y</u>	2,800	1,800	100	800	<u>ND</u>	ND
19–30 <u>y</u>	3,000	2,000	100	1,000	ND	<u>ND</u>
31–50 <u>y</u>	3,000	2,000	100	1,000	ND	ND
51–70 <u>y</u>	3,000	2,000	100	1,000	ND	ND
> 70 y	3,000	2,000	100	1,000	ND	<u>ND</u>

100/0.025 = 4000 IU

Logic [Filter & Clustering]





Logic [Filter & Clustering]

- Duplicate 제거, 중복 개수 파악 (종합/ 싱글 비타민)

```
## # A tibble: 10 x 6
## # Groups:
              Name [9]
                                   'Serving Size' Target Categories Ingrednum multi
      Name
      <chr>>
                                                  <chr> <chr>
                                   <chr>>
                                                                        <int> <chr>
   1 Windmill Quick Melt B-12 (C- 1.0 Melt(s)
                                                  Adult- vitamin
                                                                            2 Multi
   2 Windmill B-1 (Thiamine Hydr- 1.0 Tablet(s) Adult- vitamin
                                                                            1 Sing-
   3 Windmill B-12 (Cyanocobalam- 1.0 Tablet(s)
                                                  Adult- vitamin
                                                                            2 Multi
                                  1.0 Capsule(s) Adult- vitamin
   4 Thorne Methylcobalamin
                                                                            1 Sing-
   5 Thorne Pyridoxal 5'- Phosph- 1.0 Capsule(s) Adult- vitamin
                                                                            1 Sing-
   6 Thorne Vitamin K2
                                   15.0 Drop(s)
                                                  Adult- vitamin
                                                                            1 Sing-
   7 Allergy Research Group Pure- 2.0 Capsule(s) Adult- vitamin
                                                                            4 Multi
   8 The Vitamin Shoppe Dry A No- 1.0 Capsule(s) Adult- vitamin
                                                                            2 Multi
## 9 The Vitamin Shoppe Vitamin ~ 1.0 Lozenge(s) Adult- vitamin
                                                                            2 Multi
## 10 The Vitamin Shoppe Vitamin ~ 1.0 Lozenge(s) Adult- vitamin
                                                                            2 Multi
```

Logic [Filter & Clustering]

A tibble: 10 x 6 ## # Groups: Name [9] 'Serving Size' Target Categories Ingrednum multi Name <chr>> <chr>> <chr> <chr> <int> <chr> 1 Windmill Quick Melt B-12 (C- 1.0 Melt(s) Adult- vitamin 2 Multi 2 Windmill B-1 (Thiamine Hydr- 1.0 Tablet(s) Adult- vitamin 1 Sing-3 Windmill B-12 (Cyanocobalam~ 1.0 Tablet(s) Adult~ vitamin 2 Multi 4 Thorne Methylcobalamin 1.0 Capsule(s) Adult- vitamin 1 Sing-5 Thorne Pyridoxal 5'- Phosph- 1.0 Capsule(s) Adult- vitamin 1 Sing-6 Thorne Vitamin K2 15.0 Drop(s) Adult- vitamin 1 Sing-7 Allergy Research Group Pure- 2.0 Capsule(s) Adult- vitamin 4 Multi 8 The Vitamin Shoppe Dry A No- 1.0 Capsule(s) Adult- vitamin 2 Multi ## 9 The Vitamin Shoppe Vitamin ~ 1.0 Lozenge(s) Adult- vitamin 2 Multi ## 10 The Vitamin Shoppe Vitamin ~ 1.0 Lozenge(s) Adult- vitamin 2 Multi

Basic overview

name
Brand
Net contents
Supplement form

Dietary Supplement Facts (Clustering)

/ Daily value Target group/ Ingredient/ Category

크롤링

Rating, 가격, 리뷰개수 = > 최종 추천

Python 클러스터링/크롤링?

Basic_overview.csv[net content, supplement form]

Dietary_supplement_facts.csv[daily value target group, ingredient, category]

Other_ingredients.csv[other ingredients]

```
: chr [1:20000] "https://dsld.od.nih.gov/labe
  $ URL
  $ DSLD ID
                                 : num [1:20000] 542 543 544 545 546 547 548 5
  $ Product Name
                                 : chr [1:20000] "Vitamin World B-2 100 mg" "\
                                 : chr [1:20000] "Vitamin World" "Vitamin World"
  $ Brand Name
                                 : chr [1:20000] "100.0 Easy To Swallow Coated
  $ Net Contents
  $ Serving Size
                                 : chr [1:20000] "2.0 Tablet(s)" "1.0 Tablet(s
  $ Product Type [LanguaL]
                                 : chr [1:20000] "DIETARY SUPPLEMENT, VITAMIN
  $ Supplement Form [LanguaL]: chr [1:20000] "TABLET [E0155]" "TABLET [E01
                                    Other Ingredients
Other_Ingredients_1
                          20216 obs
                      : chr [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme...
 $ URL
                      : num [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme... 5
 $ DSLD ID
                      : chr [1:2021 Ascorbic Acid, Vegetable Cellulose, Citrus Bioflavonoi...
 $ Product Name
  $ Other Ingredients: chr [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme... ]
```

20000 obs. of 11 variables

Vegetable Cellulose, Dicalcium Phosphate, Stearic Aci...

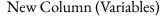
Basic Overview 1

Python 클러스터링/크롤링?

Basic overview.csv[net content, supplement form]

Dietary_supplement_facts.csv[daily value target group, ingredient, category]

Other_ingredients.csv[other ingredients]



- 'Form of Supplement' \rightarrow Categorical Var (i.e. tablet \rightarrow "1")
- 'Serving Size' → Categorical Var (i.e. 500 pills \rightarrow "5", 100 pills \rightarrow "1")
- 'Ingredients' \rightarrow Logical Var (i.e. more than $5 \rightarrow$ "1" == 종합)
- Possibly 'Other Ingredients'

```
Basic Overview 1
                            20000 obs. of 11 variables
  $ URL
                                 : chr [1:20000] "https://dsld.od.nih.gov/labe
  $ DSLD ID
                                 : num [1:20000] 542 543 544 545 546 547 548 5
  $ Product Name
                                 : chr [1:20000] "Vitamin World B-2 100 mg" "\
                                 : chr [1:20000] "Vitamin World" "Vitamin World"
  $ Brand Name
                                 : chr [1:20000] "100.0 Easy To Swallow Coated
  $ Net Contents
  $ Serving Size
                                 : chr [1:20000] "2.0 Tablet(s)" "1.0 Tablet(s
                                 : chr [1:20000] "DIETARY SUPPLEMENT, VITAMIN
  $ Product Type [LanguaL]
  $ Supplement Form [LanguaL]: chr [1:20000] "TABLET [E0155]" "TABLET [E01
                                     Other Ingredients
Other_Ingredients_1
                          20216 obs
                       : chr [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme...
 $ URL
                       : num [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme... 5
 $ DSLD ID
                       : chr [1:2021 Ascorbic Acid, Vegetable Cellulose, Citrus Bioflavonoi...
 $ Product Name
  $ Other Ingredients: chr [1:2021 Dicalcium Phosphate, Vegetable Cellulose, Croscarme... ]
                                     Vegetable Cellulose, Dicalcium Phosphate, Stearic Aci...
```

```
Dietary Supplement Fac... 193938 obs. of 10 variables
  $ URL
                               : chr [1:193938] "https://dsld.od.nih.gov/la
```

: num [1:193938] 542 543 544 544 544 545 546 \$ DSLD ID : chr [1:193938] "Vitamin World B-2 100 mg" \$ Product Name

\$ Serving Size : chr [1:193938] "2.0 Tablet(s)" "1.0 Tablet \$ Daily Value Target Group : chr [1:193938] "Adults and children 4 or m

\$ Ingredient

: chr [1:193938] "Riboflavin" "Vitamin B6" " \$ DSLD Ingredient Categories: chr [1:193938] "vitamin" "vitamin" "vitamin" Visualizations - Shiny

Next Step

Unit / Data Scaling

Finalize dataframes and try clustering

Web Scraping if possible

Shiny (Stack input & output as dataframe)

Thank you