



GoodVitamins

An Interactive Vitamin Product Review
Summarization App From iHerb.



Skills:

Web Scraping, Data Cleaning, Data Wrangling, Text Preprocessing,
Visualization, Machine Learning, NLP.



Tools:

Selenium, Pandas, Matplotlib, Universal Sentence Encoder, K-means
clustering, TextBlob.



[linkedin.com/in/andyphamto/](https://www.linkedin.com/in/andyphamto/)



github.com/andy-pham-72/

Introduction:

- Supplements are regulated by the FDA; However, the framework is different from drugs. The **FDA does not review dietary supplement products for safety and effectiveness** before they are marketed.
- Question: **Are all the supplements effective?**



Photo by Benedetto Cristofani

Why iHerb ?

- iHerb is an **online global supplements retailer**.
- iHerb has a **very active community** where users keep **posting** and **updating detailed review contents** about the products.



iHerb Brands

(115)

 Include Discontinued

Categories

Search by Category

Vitamin D (274)

Vitamin C (406)

Multivitamins (797)

Vitamin B (487)

Vitamin K (70)

Vitamin E (76)

Vitamin A (20)

Vitamin A & D (6)

Brands

Search by Brand

 Now Foods (135) California Gold Nutrition (29) Solgar (107) 21st Century (74) Life Extension (48)[+ Show More](#)

Ratings

 ★★★★★ (460) ★★★★★☆ (1614) ★★★★☆ (14) ★★★☆☆ (1) ★☆☆☆☆ (1)

Formulated For

 Women (222) Men (143)

iHerb Brands



California Gold Nutrition, Gold C, Vitamin C, 1,000 mg, 60 Veggie Capsules

★★★★★ 29764

\$5.00



Now Foods, High Potency Vitamin D-3, 125 mcg (5,000 IU), 120 Softgels

★★★★★ 31065

\$6.58



Now Foods, High Potency Vitamin D-3, 5,000 IU, 240 Softgels

★★★★★ 31065

\$10.99



Now Foods, High Potency Vitamin D-3, 50 mcg (2,000 IU), 120 Softgels

★★★★★ 20688

\$4.55

Total numbers of review



Now Foods, Highest Potency Vitamin D-3, 250 mcg (10,000 IU), 120 Softgels

★★★★★ 10014

\$8.46



Now Foods, Vitamin D-3 High Potency, 50 mcg (2,000 IU), 240 Softgels

★★★★★ 20688

\$8.15



Solaray, Vitamin D3 + K2, Soy-Free, 125 mcg (5000 IU), 60 VegCaps

★★★★★ 6969

\$14.99



California Gold Nutrition, Vitamin D3, 125 mcg (5,000 IU), 90 Fish Gelatin Softgels

★★★★★ 6064

\$4.00

iHerb Brands



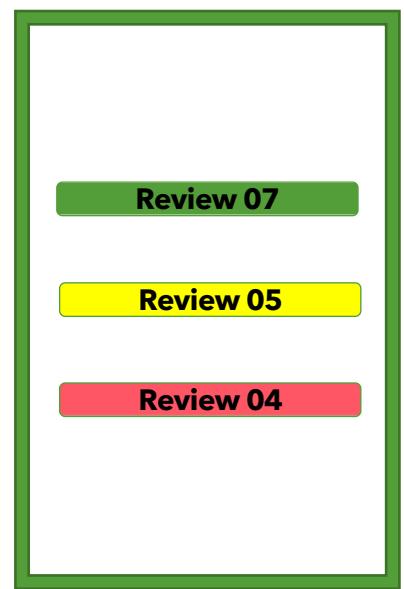
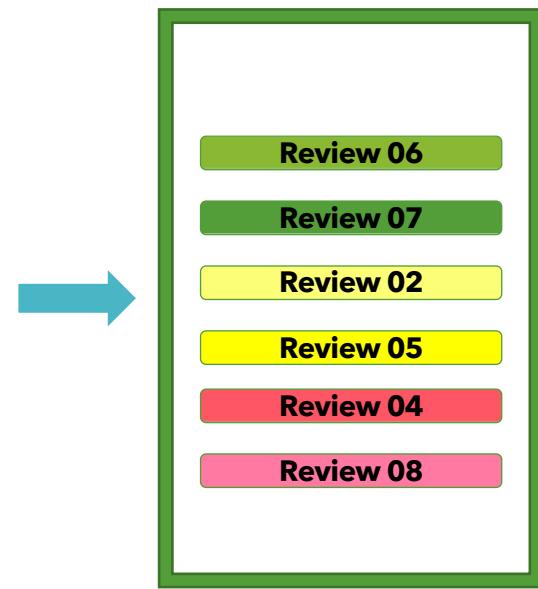
How the algorithm works



Keeps all the review contents as data points.

Universal Sentence Encoder flow:

- + Tokenizes all the reviews.
- + Each review is converted to a 512-dimentional vector.
- + These vectors are processed based on Semantic Textual Similarity (similar meaning).



Locates **the most common perspectives**.

Using **k-means clustering** to **find the cluster centers**

Finds the **top representative review** for each perspective

Using **inner product calculation** to find the review vector which is closest to the cluster center.

<https://goodvitamins.herokuapp.com>

THANK YOU !



Try it out on:
Goodvitamins.herokuapp.com



[Linkedin.com/in/andyphamto](https://www.linkedin.com/in/andyphamto)



Aqpham02@gmail.com



[Github.com/andy-pham-72](https://github.com/andy-pham-72)