

This readme file was generated on 2023-03-29 by Anna-Stiina Suur-Uski

## GENERAL INFORMATION

**Title of Dataset:** LoCard Food Classification

**Authors:** Satu Kinnunen<sup>1</sup> (0000-0001-7673-0243), Noora Kanerva<sup>1</sup> (0000-0001-6776-9357), Jaakko Nevalainen<sup>2</sup> (0000-0001-6295-0245), Anna-Stiina Suur-Uski<sup>3</sup> (0000-0001-5029-2863), Maijaliisa Erkkola<sup>1</sup> (0000-0002-6966-1523), Mikael Fogelholm<sup>1</sup> (0000-0001-8110-102X), Henna Vepsäläinen<sup>1</sup> (0000-0002-0177-3609), Jelena Meinilä<sup>1</sup> (0000-0002-6377-1377), Hannu Saarijärvi<sup>2</sup> (0000-0001-5803-9037)

<sup>1</sup> University of Helsinki, Finland

<sup>2</sup> Tampere University, Finland

<sup>3</sup> Tampere University Library, Finland

### Contacts:

Name: Kinnunen, Satu

ORCID: 0000-0001-7673-0243

Institution: University of Helsinki, Finland

Address: PL 66 (Agnes Sjöbergin katu 2), 00014 University of Helsinki, Finland

Email: satu.kinnunen@helsinki.fi

Name: Kanerva, Noora

ORCID: 0000-0001-6776-9357

Institution: University of Helsinki, Finland

Address: PL 66 (Agnes Sjöbergin katu 2), 00014 University of Helsinki, Finland

Email: noora.kanerva@helsinki.fi

**Date of data collection:** from 2016-09-01 to 2018-12-31

### Information about funding sources that supported the collection of the data:

This research was funded by the Academy of Finland (#350862 JN; #350863 MF), the Juho Vainio Foundation (#202200480 ME; #202100202 JM), the Emil Aaltonen Foundation (JM), the Yrjö Jahnesson Foundation (JM), and the Finnish Food Research Foundation (HV)

## SHARING/ACCESS INFORMATION

### Licenses/restrictions placed on the data:

[Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

### Links to publications that cite or use the data:

N/A

**Recommended citation for this dataset:**

Kinnunen, S., Kanerva, N., Nevalainen, J., et al. (2023) LoCard Food Classification. Zenodo.  
DOI: 10.5281/zenodo.7781352

**DATA & FILE OVERVIEW****File List:**

- locard\_classification\_nutrients\_28DEC2022.csv
- LoCard\_classes\_fin-eng\_2020-06-09.xlsx
- README.pdf

**METHODOLOGICAL INFORMATION**

Description of methods used for collection/generation of data:

The grocery purchase data including 3574 product groups was received from the retailer for research purposes. The data was reclassified into appropriate categories suitable for the use of nutrition and health research.

Methods for processing the data:

A four-level hierarchical classification of product groups was used. Each class on the broadest level of hierarchy (Class 1) was subsequently divided into a reasonable number of finer sub-classes starting with Class level 2, followed by Class 3, and finally, Class 4, which was the most detailed level of hierarchy. The main ingredient of the product group, type of the food and purpose of use, nutritional content, and carbon footprint were considered in the reclassification process. The classified food groups were linked with Finnish food composition database.

The authors strongly recommend reading the detailed description of the classification is published in <https://doi.org/10.21203/rs.3.rs-2826970/v1>

DATA-SPECIFIC INFORMATION FOR: locard\_classification\_nutrients\_28DEC2022.csv

Number of variables: 17

Number of cases/rows: 1029

Variable List:

| Variable name | Explanation  | More information   |
|---------------|--|--|
| Class_1       | highest level of hierarchy including 35 main food groups based on <i>healthiness (Nordic Nutrition Recommendations)</i> and <i>main ingredients (grouping of the Finnish Food Composition Database Fineli)</i> .   | Nordic Council of Ministers. Nordic nutrition recommendations. 5th ed. Vol. Nord 2013/009. Denmark: Norden; 2013.<br><a href="http://www.fineli.fi">www.fineli.fi</a>  |
| Class_2       | 114 food groups.   |  |
| Class_3       | 159 food groups unique to Class_3, and 170 unique combinations of Class_1 to Class_3.  |  |
| Class_4       | Finest level of hierarchy. Includes 164 unique food groups, and 198 unique combinations of Class_1 to Class_4.   |  |
| fineli_foodid | Identification code of the food in the Finnish Food Composition Database Fineli (version 20).<br><br>Some values are taken from US database ("usda"), from product label ("producer"), or analysed in laboratory as part of an intervention study ("beanman"). | <a href="http://www.fineli.fi">www.fineli.fi</a><br><br>Reinivuo H, Hirvonen T, Ovaskainen ML, Korhonen T, Valsta LM. Dietary survey methodology of FINDIET 2007 with a risk assessment perspective. Public health nutrition. 2010 Jun;13(6A):915–9.<br><br>Values used from USDA:<br>75224023 Green peas, canned, cooked, no added fat<br>75510050 Olive tapenade<br>91708000 Fruit peel, candied |

|                          |   |  |
|--------------------------|---|--|
|                          |   | <p>59003000 Meat substitute, cereal- and vegetable protein-based, fried</p> <p>91718200 Chocolate flavored sprinkles</p> <p>19202 Puddings, vanilla, dry mix, instant</p> <p>Values used from producer: Deliciest, Maisku quark bar, vanilla, 38g<br/>Jalostaja, Italian stew, 300g<br/>Colman's, mustard powder 100g</p> <p>Values used from the Bean Man study:<br/>vegetable cheese<br/>(more information: anne-maria.pajari@helsinki.fi)</p> |
| protein_g                | protein content in grams  |  |
| saccharose_g             | saccharose content in grams   |  |
| fibre_g                  | fibre content in grams  |  |
| pufa_g                   | polyunsaturated fat in grams  |  |
| safa_g                   | saturated fat in grams  |  |
| ca_mg                    | calcium in milligrams   |  |
| fe_mg                    | iron in milligrams  |  |
| Salt                     | salt in grams   |  |
| vit_c_mg                 | vitamin C in milligrams   |  |
| vit_d_ug                 | vitamin D in micrograms   |  |
| folate_mg                | folate in milligrams  |  |
| nutrient_rich_food_index | <p>Nutrient rich food index was calculated per 100 g of product using 11 nutrients, of which eight were regarded positive (protein, fibre, polyunsaturated fatty acids (PUFA), calcium, iron, vitamin D, vitamin C and folate) and three negative (saturated fatty acids (SFA), saccharose and salt) in terms of anticipated health effects. First, positive and negative scores were</p> | <p>Drewnowski A, Fulgoni VL. Nutrient density: principles and evaluation tools. Am J Clin Nutr. 2014 May;99(5 Suppl):1223S-8S.</p> <p>Recommended values are from Finnish Nutrition Recommendations which are the same as in the Nordic Nutrition recommendations, except for salt which is 5000mg in the Finnish</p>  |

|  |  |  |
|--|--|--|
|  | <p>calculated as an average of percentages of daily recommendation (DR%):</p> <p>Positive score:<br/>(DR% protein + DR% fibre + DR% PUFA + DR% Ca + DR% Fe + DR% Vit D + DR% Vit C + DR% folate) / 8</p> <p>Negative score:<br/>(DR% sucrose + DR% SFA + DR% salt) / 3</p> <p>Last, NRFI was calculated by subtracting the negative score from the positive score. The scores vary from -6.538 to 1.261. Positive score reflects higher nutritional quality and negative score reflect lower nutritional quality of the grocery product group.</p> | <p>recommendations (6000 mg in the Nordic nutrition recommendations).</p> <p>Recommended values for calculating the percentage of daily recommendation (DR%) are:</p> <p>Protein = 90g<br/>(corresponding 15% of energy in 2400 kcal diet)</p> <p>Fibre = 25g</p> <p>Polyunsaturated fat (PUFA)=20g (corresponding 7.5% of energy in 2400 kcal diet)</p> <p>Calcium (Ca) = 800mg</p> <p>Iron (Fe) = 9mg</p> <p>Vitamin D = 10µg</p> <p>Vitamin C =75mg</p> <p>Folate = 300mg</p> <p>Sucrose = 60g<br/>(corresponding 10% of energy in 2400 kcal diet)</p> <p>Saturated fat = 26.7g<br/>(corresponding 10% of energy in 2400 kcal diet)</p> <p>Salt = 5000mg</p> <p>Nordic Council of Ministers.<br/>Nordic nutrition recommendations. 5th ed.<br/>Vol. Nord 2013/009.<br/>Denmark: Norden; 2013.</p> |
|--|--|--|

Missing data codes: Does not apply for the data in question

Specialized formats or other abbreviations used:

Ancillary data file (LoCard\_classes\_fin-eng\_2020-06-09.xlsx) lists the abbreviated food group names used in the classification data with full names both in English and in Finnish.

DATA-SPECIFIC INFORMATION FOR: LoCard\_classes\_fin-eng\_2020-06-09.xlsx

Number of variables: 3

Number of cases/rows: 241

Variable List:

| Variable name    | Description   |
|------------------|---|
| Tuoteryhmä       | Food group names in Finnish.                                  |
| Food group       | Food group names in English.                                  |
| Abbreviates name | Abbreviated food group names used in the classification data. |

Missing data codes: Does not apply for the data in question.

Specialized formats or other abbreviations used: For alcoholic beverages numbers OXY refer to the alcohol content (X.Y%) of the products.