IC24026 USDA Data Analytics

A general analysis of Foundation Foods and SR Legacy food components

Team members: Anthony Cheng, David Chu, Eric Huang, Ethan Chang

The problem

USDA

- Organization Name: U.S. Department of Agriculture Food Data Central (FDC)
- Challenge Name: USDA FDC
- Difficulty: Level 2: Participants with some data analysis background.

Context

- The FDC integrates nutrient and food component data into one system.
- Lack of Resources/Staff lead to new ways to incorporate data

Findings

- Food Components Changes
- Comparison
- Mean Value Comparison
- Nutrients/Components Ranges

Challenges deep-dive

Challenge 1

Cleaning/Organizing Data

Aside from dividing the data into more easily manageable categories, we organized the information into a more consumable format. This enabled us to more effectively analyze the data set.

Challenge 2

Analysis

Took cleaned data to describe and illustrate, condense and recap, and evaluate the data. Pattern Analysis.

Challenge 3

Conclusion

Formed conclusions about the analysis.

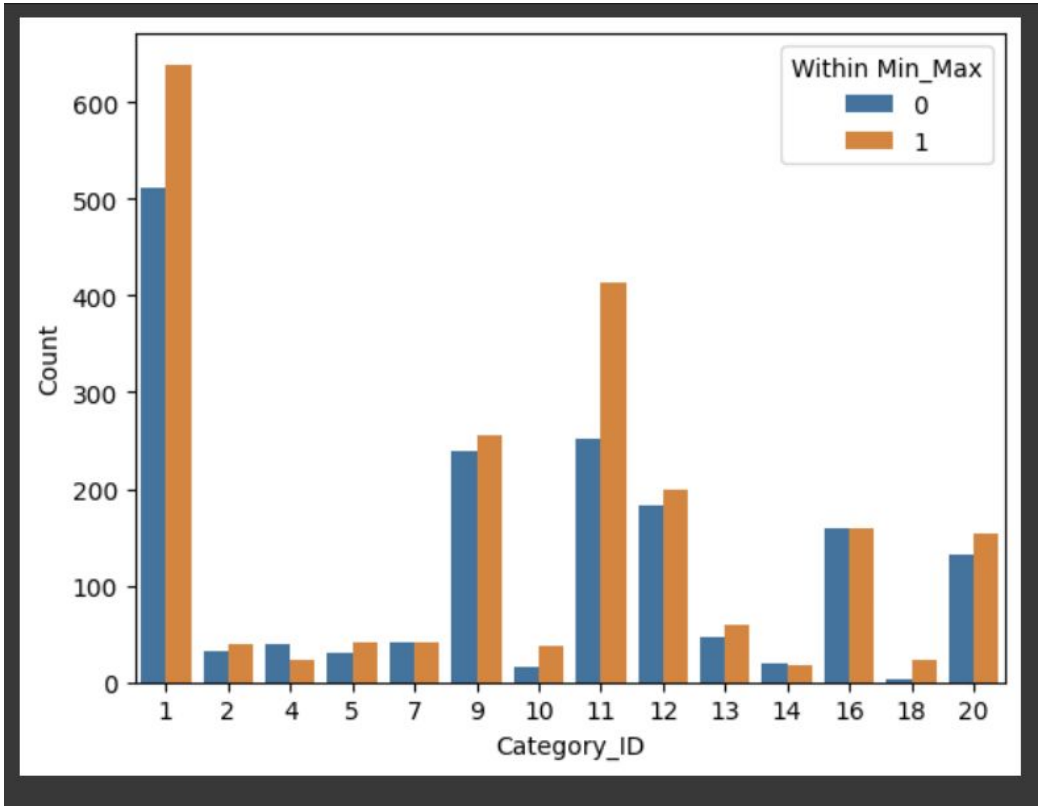
Data Cleaning/Organization

1. New Category: Percentage Change
    a. Three Outcomes
        i. Infinite
        ii. 0
        iii. Finite Pos/Neg
2. Cleaning
    a. Duplicates Rows
    b. Missing Values
    c. Similar Food Items measured with different units

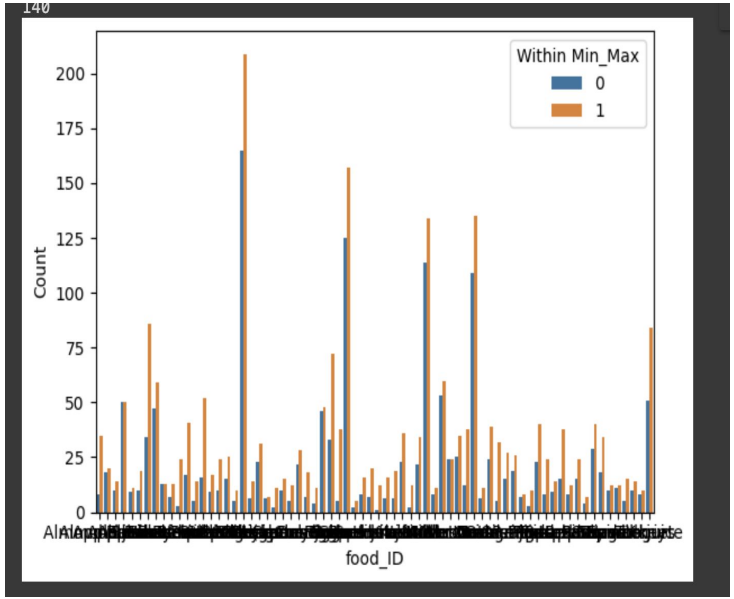
Analyzing Food Category, and Overall Number of Change

	Category	Food Description	Net_Zero	Net_Inf	Pos_Finite	Neg_One	Neg_Finite
0	1	Dairy and Egg Products	48.033126	39.130435	29.544109	40.000000	27.928342
1	2	Spices and Herbs	3.519669	4.347826	2.249852	2.000000	1.056500
2	4	Fats and Oils	0.414079	10.869565	1.243339	2.500000	1.607717
3	5	Poultry Products	0.621118	0.000000	1.598579	1.500000	2.113000
4	7	Sausages and Luncheon Meats	1.863354	8.695652	3.374778	2.000000	3.950390
5	9	Fruits and Fruit Juices	18.219462	8.695652	10.834813	15.500000	14.561323
6	10	Pork Products	0.621118	0.000000	1.065719	1.000000	1.607717
7	11	Vegetables and Vegetable Products	17.805383	2.173913	16.873890	18.500000	17.087735
8	12	Nut and Seed Products	4.554865	8.695652	11.071640	7.500000	9.416628
9	13	Beef Products	0.828157	2.173913	2.013025	2.000000	3.353238
10	14	Beverages	2.070393	10.869565	0.473653	2.500000	0.918695
11	16	Legumes and Legume Products	0.414079	0.000000	10.716400	0.000000	7.579237
12	18	Baked Products	0.414079	2.173913	1.065719	0.000000	0.643087
13	20	Cereal Grains and Pasta	0.621118	2.173913	7.874482	3.500000	8.176390

Category: Analyzing SR to FF Min/Max



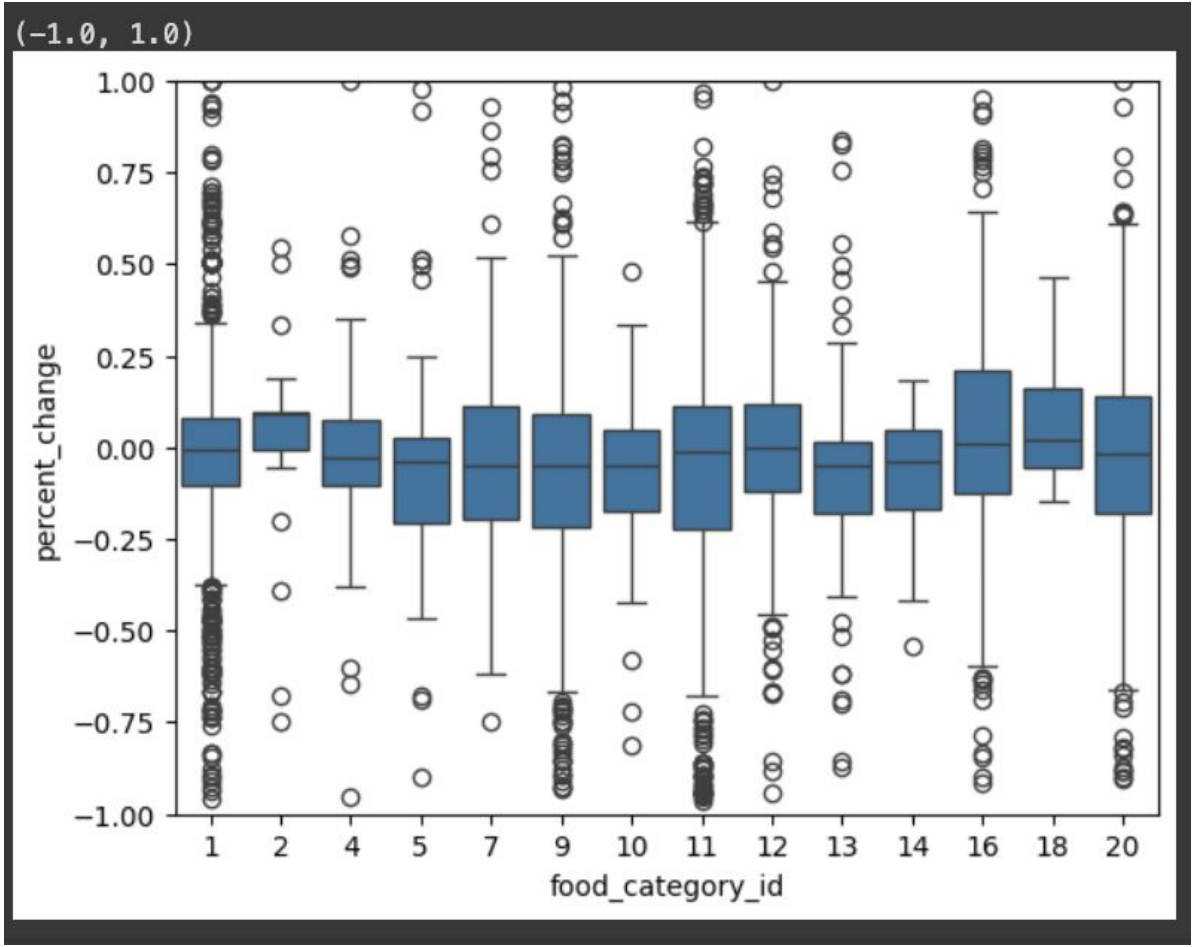
Food Description: Analyzing SR to FF Min/Max



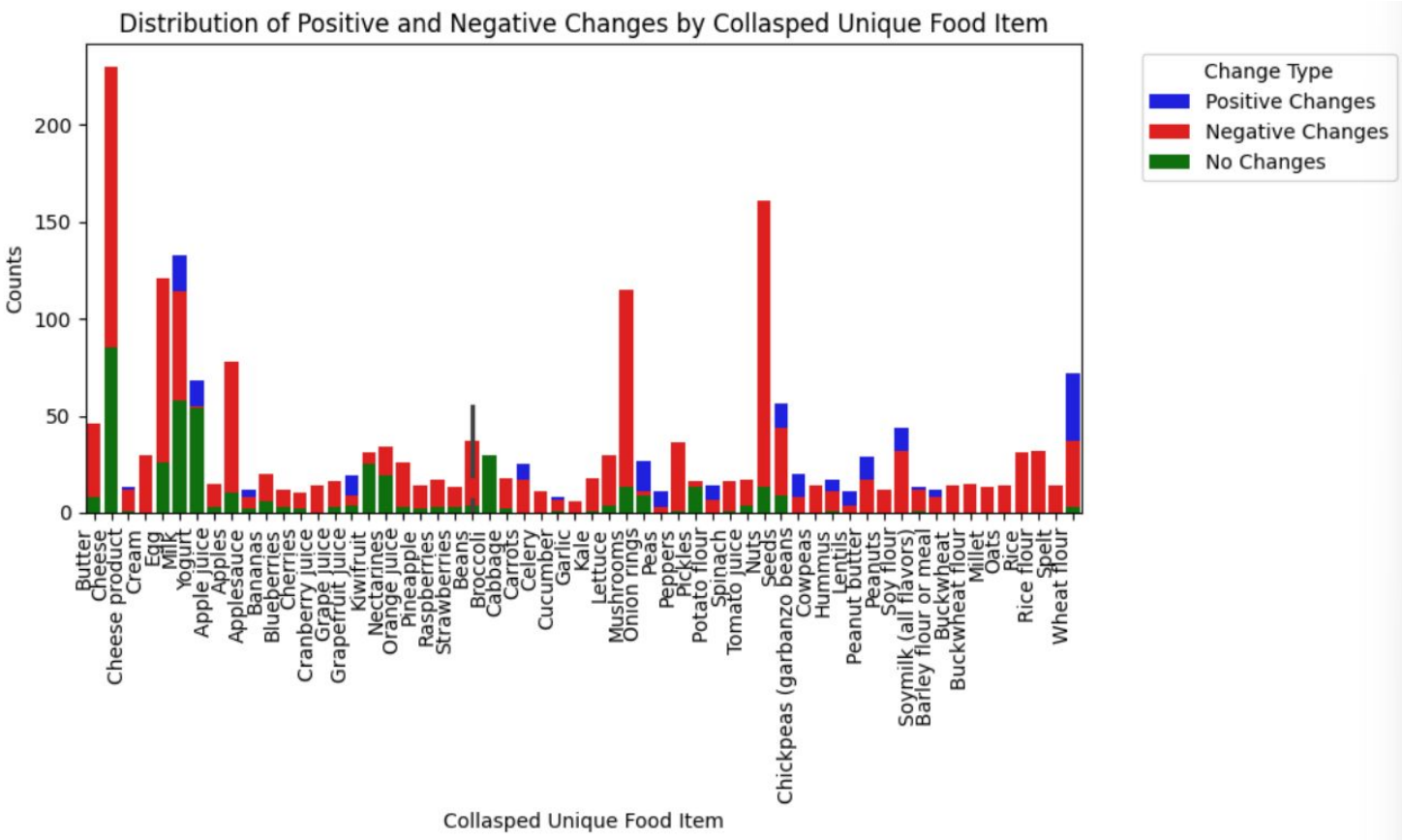
food_ID	Within Min_Max	Count
Cheese	1	209
Cheese	0	165
Flour	1	157
Nuts	1	135
Milk	1	134
Flour	0	125
Milk	0	114
Nuts	0	109
Beans	1	86
Yogurt	1	84
Eggs	1	72
Mushroom	1	60
Beef	1	59
Mushroom	0	53

Cherries	0	6
Chickpeas	0	6
Celery	0	5
Cranberry juice	0	5
Strawberries	0	5
Onion rings	0	5
Buckwheat	0	5
Flaxseed	0	5
Garlic	1	5
Cucumber	0	4
Salt	0	4
Bread	0	3
Peas	0	3
as (garbanzo beans	0	2
Garlic	0	2
Lentils	0	2
Ham	0	1

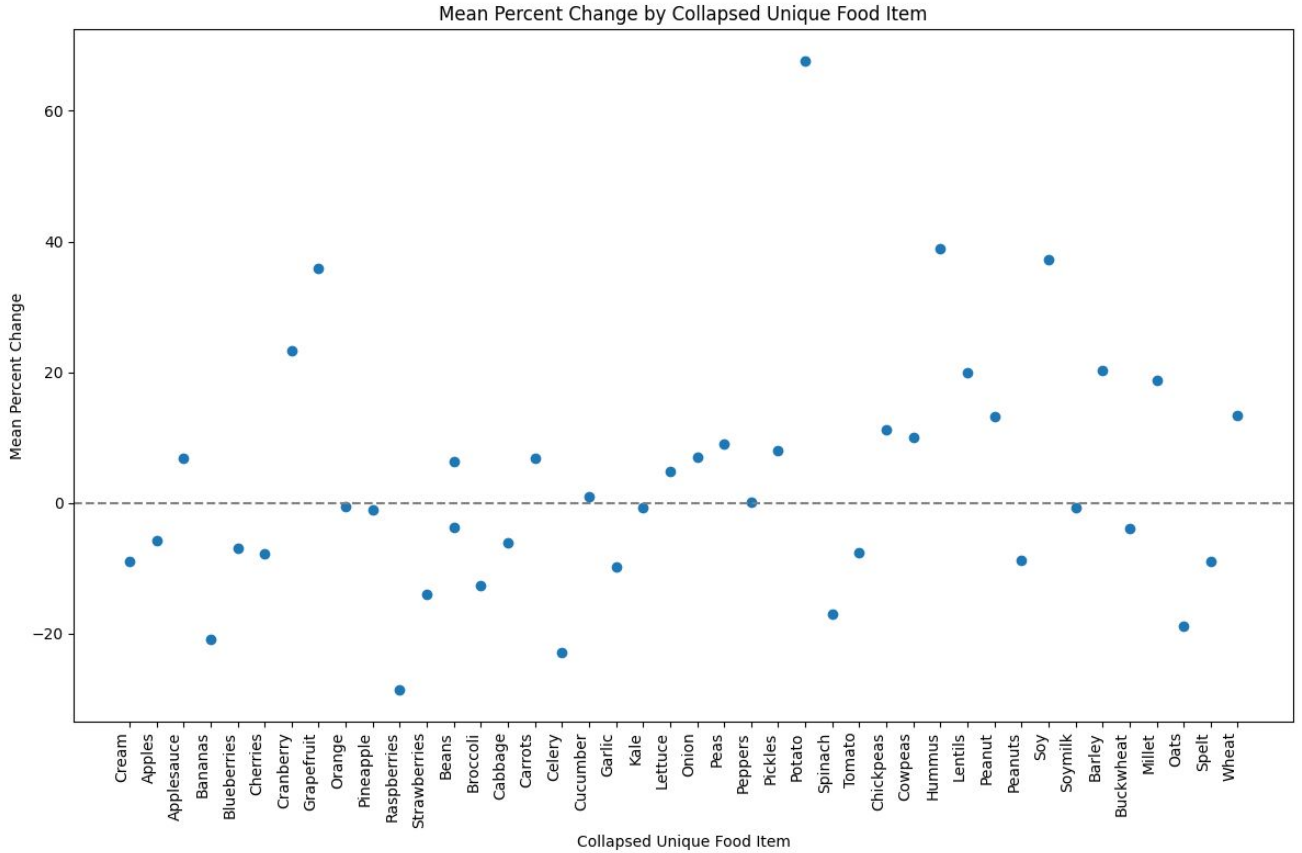
Analysis


(−1.0, 1.0)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
percent_change
1
2
4
5
7
9
10
11
12
13
14
16
18
20
food_category_id

Analysis


Distribution of Positive and Negative Changes by Collasped Unique Food Item
Counts
0
50
100
150
200
Change Type
Positive Changes
Negative Changes
No Changes
Butter
Cheese
Cheese product
Cream
Egg
Milk
Yogurt
Apple juice
Apples
Applesauce
Bananas
Blueberries
Cherries
Cranberry juice
Grape juice
Grapefruit juice
Kiwifruit
Nectarines
Orange juice
Pineapple
Raspberries
Strawberries
Beans
Broccoli
Cabbage
Carrots
Celery
Cucumber
Garlic
Kale
Lettuce
Mushrooms
Onion rings
Peas
Peppers
Pickles
Potato flour
Spinach
Tomato juice
Nuts
Seeds
Chickpeas (garbanzo beans
Cowpeas
Hummus
Lentils
Peanut butter
Peanuts
Soy flour
Soymilk (all flavors)
Barley flour or meal
Buckwheat
Buckwheat flour
Millet
Oats
Rice
Rice flour
Spelt
Wheat flour
Collasped Unique Food Item

Analysis

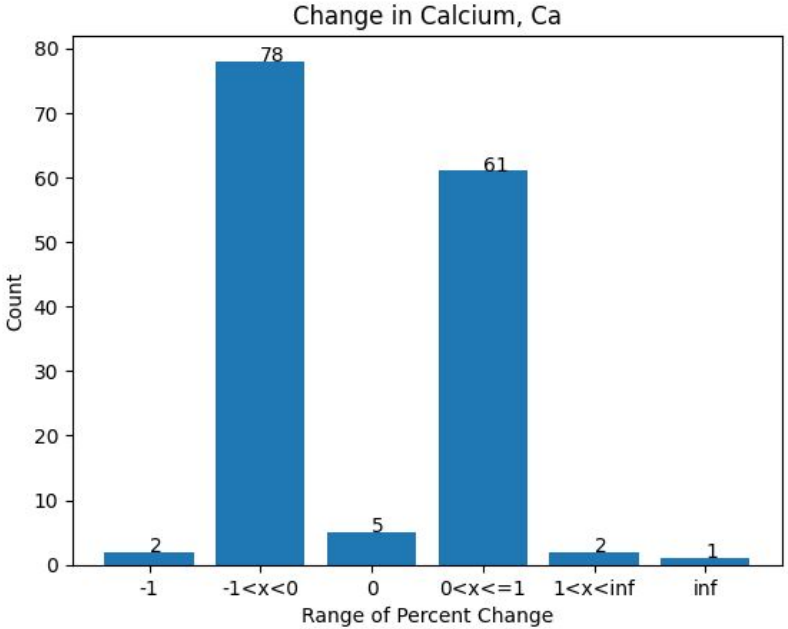

Mean Percent Change by Collapsed Unique Food Item
Mean Percent Change
60
40
20
0
−20
Cream
Apples
Applesauce
Bananas
Blueberries
Cherries
Cranberry
Grapefruit
Orange
Pineapple
Raspberries
Strawberries
Beans
Broccoli
Cabbage
Carrots
Celery
Cucumber
Garlic
Kale
Lettuce
Onion
Peas
Peppers
Pickles
Potato
Spinach
Tomato
Chickpeas
Cowpeas
Hummus
Lentils
Peanut
Peanuts
Soy
Soymilk
Barley
Buckwheat
Millet
Oats
Spelt
Wheat
Collapsed Unique Food Item

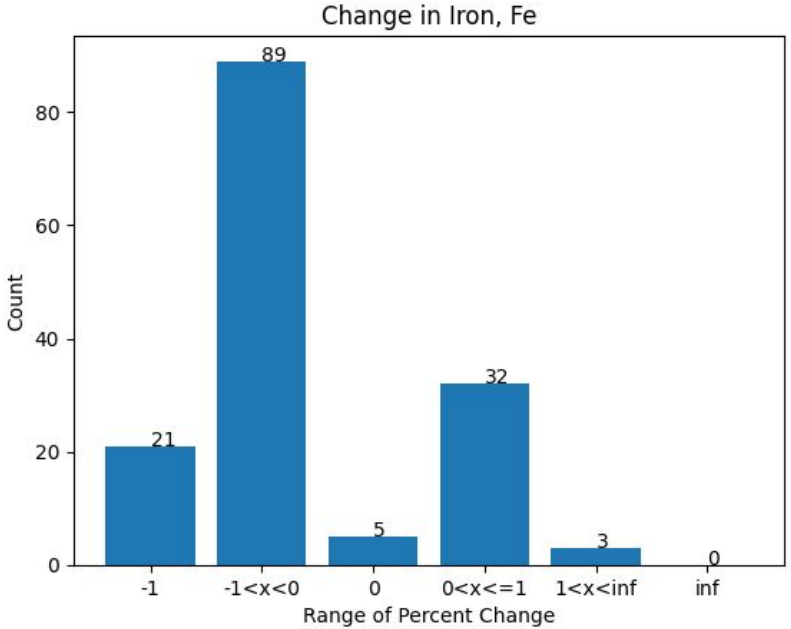
Components that People in the U.S. Commonly have a Deficiency of

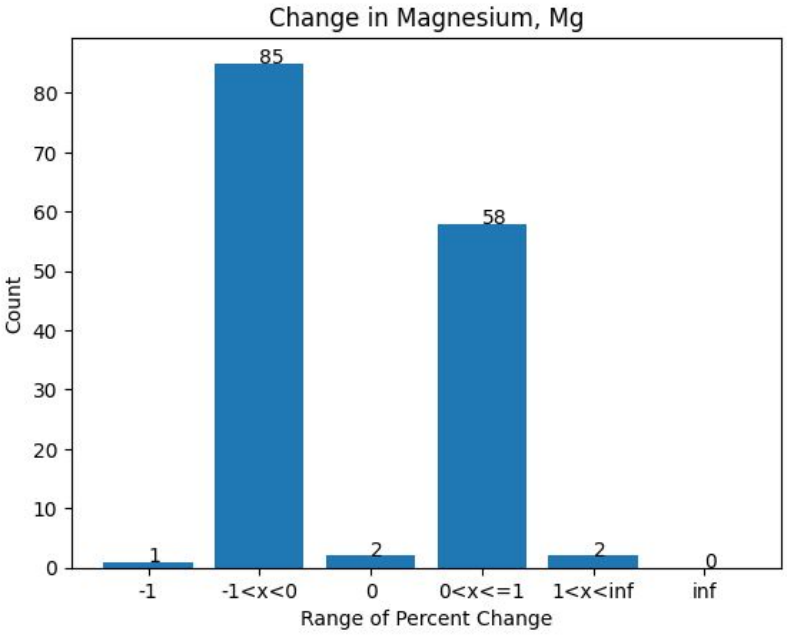
1. Calcium
2. Iron
3. Iodine
4. Magnesium
5. Vitamin A
6. Vitamin B12
7. Vitamin D

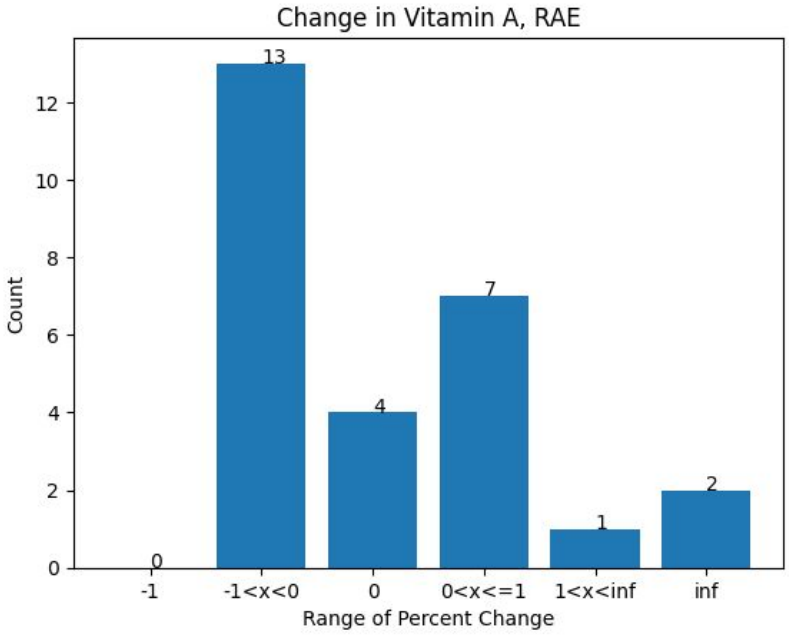
FF_Component	Neg 1 Count	Neg Count	Zero Count	Positive Count	Inf Count	Nonzero Count
Calcium, Ca	2.0	80.0	5.0	64.0	1.0	144.0
Iron, Fe	21.0	110.0	5.0	35.0	0.0	145.0
Magnesium, Mg	1.0	86.0	2.0	60.0	0.0	146.0
Vitamin A, RAE	0.0	13.0	4.0	10.0	2.0	23.0
Vitamin B-12	3.0	14.0	0.0	9.0	1.0	23.0
Vitamin D (D2 + D3)	8.0	21.0	6.0	7.0	3.0	28.0

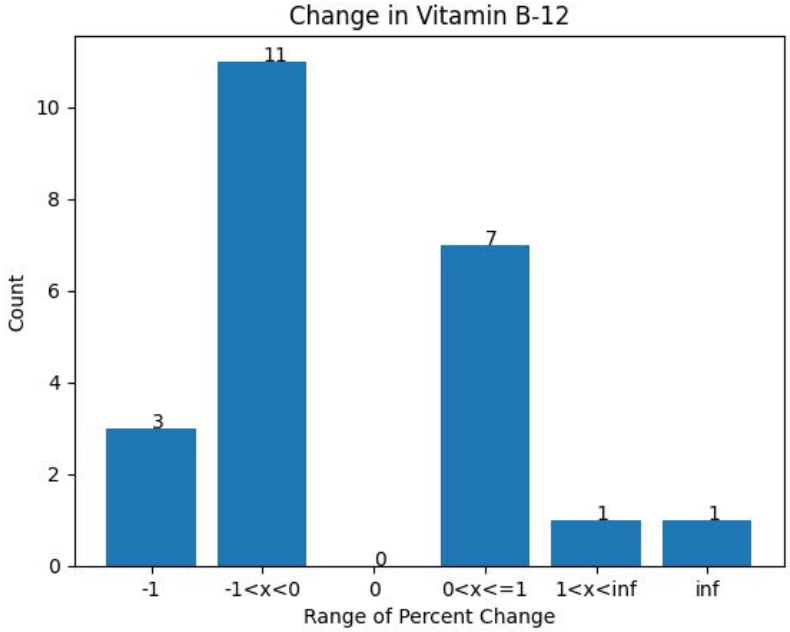
Percent Change in Components that People in the U.S. Commonly have a Deficiency of

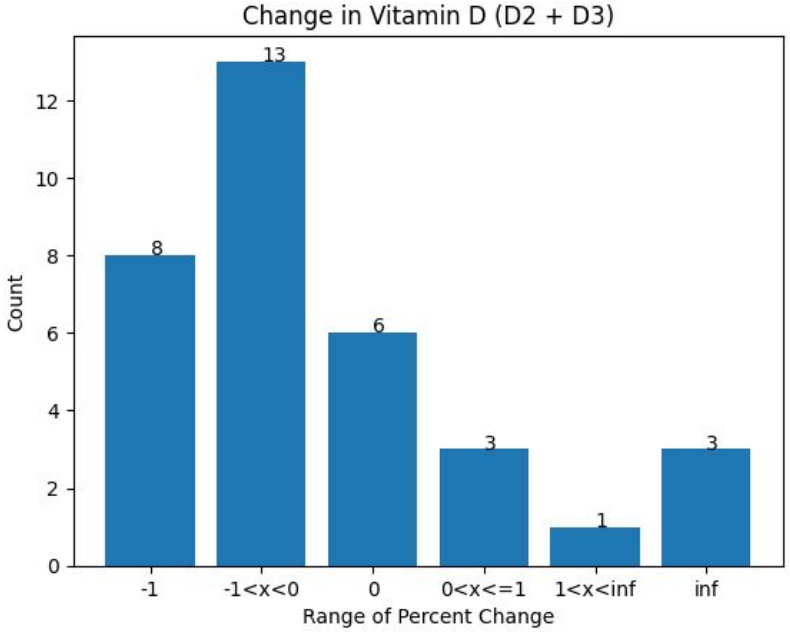





Change in Magnesium, Mg
Count
Range of Percent Change
-1
-1<x<0
0
0<x<=1
1<x<inf
inf
1
85
2
58
2
0


Change in Vitamin A, RAE
Count
Range of Percent Change
-1
-1<x<0
0
0<x<=1
1<x<inf
inf
0
13
4
7
1
2


Change in Vitamin B-12
Count
Range of Percent Change
-1
-1<x<0
0
0<x<=1
1<x<inf
inf
3
11
0
7
1
1


Change in Vitamin D (D2 + D3)
Count
Range of Percent Change
-1
-1<x<0
0
0<x<=1
1<x<inf
inf
8
13
6
3
1
3

Foods that Experienced a Decrease in the Components

1. **Calcium**: apples, beans, beef, chicken, pork, dairy products, eggs, flour, mushrooms, nuts, rice.
2. **Iron**: apples, beans, beef, chicken, cabbage, dairy products, eggs, flour, lettuce, mushrooms, nuts, rice.
3. **Magnesium**: apples, beans, beef, chicken, pork, dairy products, eggs, flour, mushrooms, nuts.
4. **Vitamin A**: bananas, broccoli, cheese.
5. **Vitamin B12**: dairy products, sausages.
6. **Vitamin D**: cheese, eggs, milk, mushrooms.

Conclusion

There was change

Focus on the changed stuff