ANALYSIS - TOP 98 HEALTHY FOOD

By: Nizam Jaafar

|	Calories	Protein (g)	Fiber (g)	Vitamin C (mg)	Antioxidant Score
Calories	1				
Protein (g)	0.662552151	1			
Fiber (g)	0.578018217	0.258662946	1		
Vitamin C (mg)	-0.212027512	-0.223645397	0.011237358	1	
Antioxidant Score	-0.125098697	-0.102985671	-0.042942676	-0.126943081	1

From the heatmap, we can conclude that there is moderate correlation for Calories with Protein and Fiber where the results are above 0.663 and 0.578 respectively.

Protein (g) Line Fit Plot

Coefficient = 8.470169
P-Value = 1.08E-13
R^2 = 0.439

Protein (g) Residual Plot
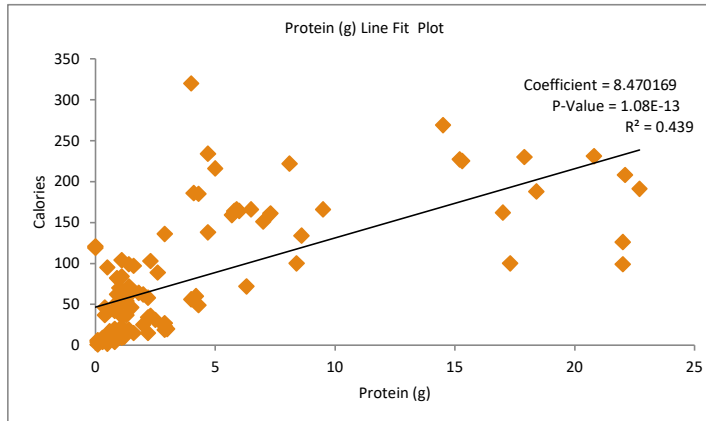
We did a regression analysis on Calories and Protein where Calories as dependable variable and Protein as independent variable. From the outcome, as can see in the line fit plot, R square resulting reading of 0.439 or 43.9% which the variation or changes of Calorie explained by Protein content. This a decent fit but there are also other variable or factors influencing Calories.

Meanwhile, on the P-Value result suggesting that we can accept to reject the null relationship which to conclude the amount of Protein in food has a meaningful and statistically significant impact on the number of Calories.

As per coefficient result, the expected value of Calories to increase for each additional gram of Protein is approximately around 8.47 units. This relationship is highly significant.

Fiber (g) Line Fit Plot

Coefficient = 12.948700
P-Value = 4.60E-10
R^2 = 0.3341

Fiber (g) Residual Plot
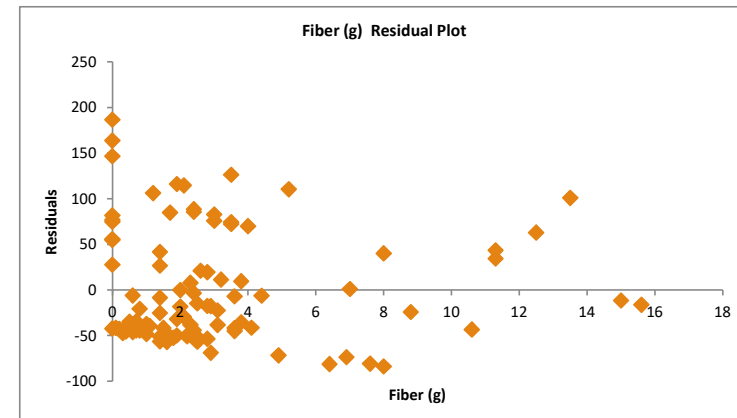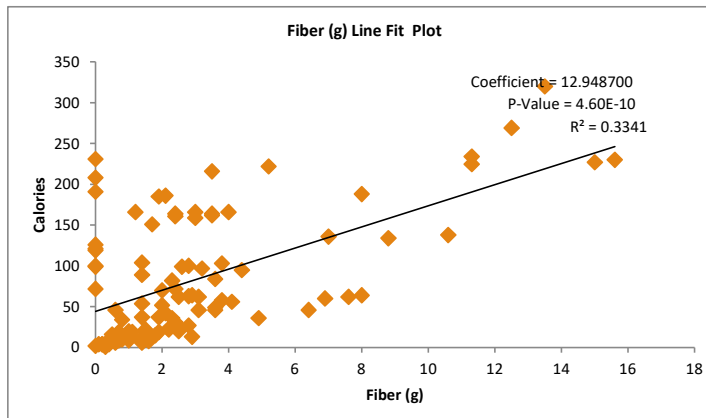
We did a regression analysis on Calories and Fiber where Calories as dependable variable and Fiber as independent variable. From the outcome, as can see in the line fit plot, R square resulting reading of 0.3341 or 33.41% which the variation or changes of Calories explained by Fiber content. This a decent fit but there are also other variable or factors influencing Calories.

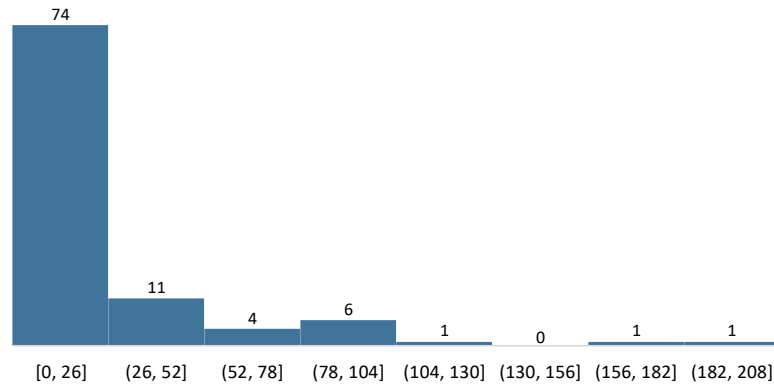Meanwhile, on the P-Value result suggesting that we can accept to reject the null relationship which to conclude the amount of Fiber in food has a meaningful and statistically significant impact on the number of Calories.

As per coefficient result, the expected value of Calories to increase for each additional gram of Fiber is approximately around 12.95 units. This relationship is highly significant.

ANALYSIS - TOP 98 HEALTHY FOOD

By: Nizam Jaafar

Distribution by Food: Vitamin C

74 | 11 | 4 | 6 | 1 | 0 | 1 | 1
[0, 26] | (26, 52] | (52, 78] | (78, 104] | (104, 130] | (130, 156] | (156, 182] | (182, 208]

This is distribution plot for Vitamin C for each food listed. Most of the food ranging 0 - 26 mg of Vitamin C with total count of 74.

Red Bell Peppers represent the high reading of Vitamin C with value of 190 mg.

Distribution by Food: Fiber

61 | 23 | 4 | 3 | 4 | 3
[0, 3] | (3, 5] | (5, 8] | (8, 10] | (10, 13] | (13, 16]
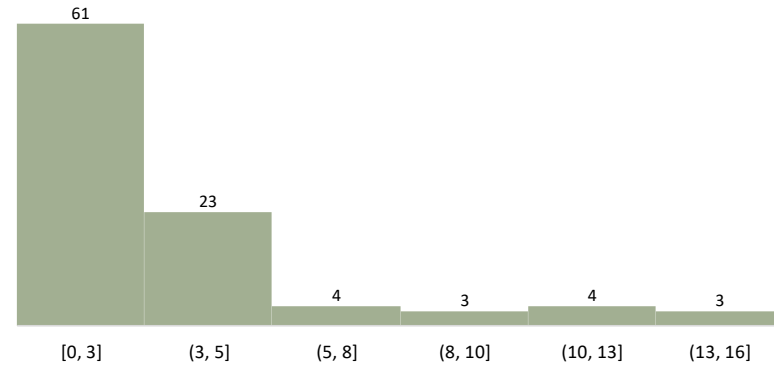
This is distribution plot for Fiber for each food listed. Most of the food ranging around 0 - 3 g of Fiber with total count of 61.

Lentils, Black Beans and Avocado represent the high reading of Fiber with value ranging 13 - 16 g.

Distribution by Food: Calories

48 | 21 | 15 | 9 | 4 | 1
[1, 58] | (58, 115] | (115, 172] | (172, 229] | (229, 286] | (286, 343]
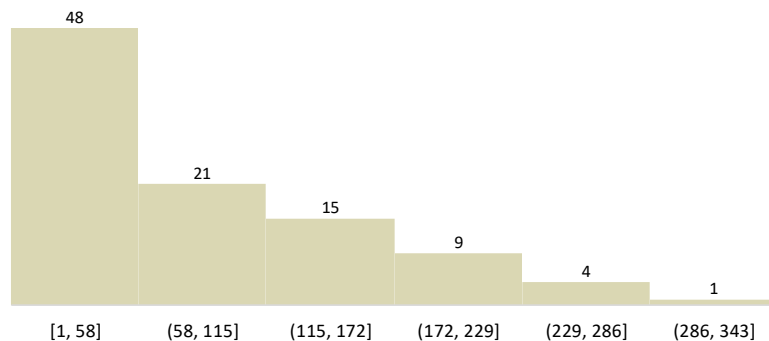
This is distribution plot for Calories for each food listed. Most of the food ranging 1 - 58 Calories with total count of 48.

Avocado represent the high reading of Calories with value of 320.

Distribution by Food: Protein

71 | 14 | 1 | 6 | 5 | 1
[0, 5] | (5, 9] | (9, 14] | (14, 18] | (18, 23] | (23, 27]

This is distribution plot for Protein for each food listed. Most of the food ranging 0 - 5 g of Protein with total count of 71.

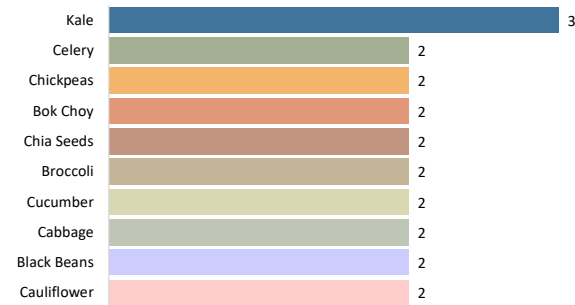Sardines represent the high reading of Protein with value ranging 23 - 27 g.

ANALYSIS - TOP 98 HEALTHY FOOD

By: Nizam Jaafar

Hierarchy by Origin: Most Healthiest Food

Legend:
- Mediterranean Region
- South America
- Worldwide
- Middle East
- North America
- Southeast Asia
- Central Asia
- Central America
- China
- Japan
- India
- Europe
- Central/South America
- Italy
- Ancient Greece
- Ancient Persia
- Ancient Rome

Treemap labels: Mediterranean Region, 10; South America, 8; Worldwide, 7; Middle East, 6; North America, 6; Southeast Asia, 6; Central Asia, 5; Central America, 5; Japan, 4; India, 4; China, 4; Europe, 3; Central/South America, 3; Ancient Persia, 2; Italy, 2; Ancient Greece, 2; Ancient Rome, 1; Europe/North...; Cari... 1; Mex... 1; Ind... 1; Mid... East...; East Asia...; Gree... 1; North Paci...; East... Me... 1; South Asia, 1; Sri Lan... 1; Nort... Euro...; Nort... Euro...; Asia, 1; Asia Mi... 1; Cy... 1; Africa, 1; Euro... West...; West... Asia, 1; Mediterra...

Out of 37 origins, it can be seen that Mediterranean region emerge as the leader for the most number of healthiest food out of top 98 list, with total of 10 count of foods or 10.20%.

2nd in place is South America with 8 food counts, and third with 7 is the food origin is worldwide as per list.

Other regions, distribute evenly from 1 - 6 counts of food.

An extensive research can be done to find the reason behind why this two regions (Mediterranean and South America) contribute most out of top 98 healthiest food in the world.

We have to exclude region "Worldwide" from the category as it not specified the origin of the food deeply.

Is it an environmental factor or maybe geological?

Is there any additonal food out of 98 foods that belong to these three regions? What are the numbers?

From the pie chart on the right-hand side, after analyzing the nutritional content correlates with the number of healthy foods per origin. Specifically, 10.40% of the 125 nutrional contents found in the 98 healthiest food are present in Mediterranean food, followed by South America at 7.20%.

Additionally, food of worldwide origin accounts for approximately 6.40% of the nutrial contents.

By Origin: Top 5 % Distribution of Nutritional Content

- Central America, 5.60%
- Mediterranean Region, 10.40%
- Middle East, 6.40%
- South America, 7.20%
- Worldwide, 6.40%

ANALYSIS - TOP 98 HEALTHY FOOD

By: Nizam Jaafar

By Food: Top 10 Count of Nutritional Content per 100g

Food	Count
Kale	3
Celery	2
Chickpeas	2
Bok Choy	2
Chia Seeds	2
Broccoli	2
Cucumber	2
Cabbage	2
Black Beans	2
Cauliflower	2

By Nutritional Content: No. of Food

Nutrient	No. of Food
Vitamin C	14
Vitamin K	13
Calories	11
Protein	9
Fiber	7
Omega-3	7
Vitamin A	6
Antioxidants	4
Probiotics	4
Selenium	3

Go into analysis of nutrional content found in a top 98 healthy food - as shown in the chart above, Kale is a leader of food that contain the highest distinct nutrients per 100 gram serving with 8 distinct nutrient.

The other top 9 with count of 2 distinct nutrients per serving of 100g, showcasing a diverse range of benifical nutrients.

Lastly, for other 88 foods ranging from 1 to 2 contents.

By nutritional content, the type of nutrition that found the most in a single plate of food is coming from Vitamin C with total of 14 foods, follow by Vitamin K at 13 of counts and at third, Calories with 11 foods.

As we go deeper into the analysis, by category of nutrition in this top 10 list, we can conclude that Vitamin accountable for most of the nutrition can be found in the top 98 healthy foods.

By Weigthage: Top 10 Healthiest Food

Legend: Antioxidant, Protein, Fiber, Vitamin C, Calories
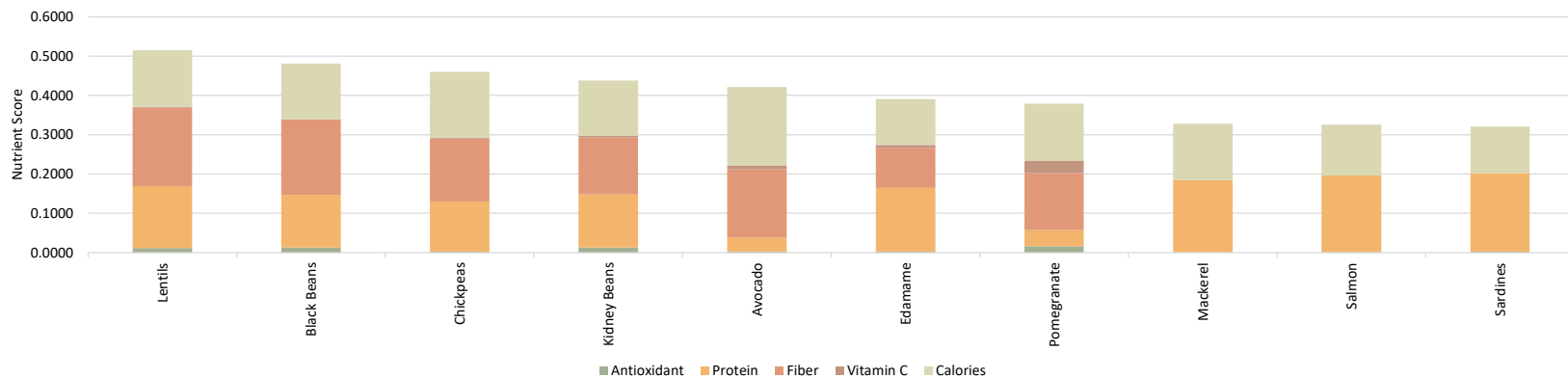
To find the food that can provide most benefit to health, we did a weightage scoring based on the essential nutrients as presented in the dataset - Antioxidant, Protein, Fiber, Vitamin C and Calories. The numbers for each nutrient than convert into normalized value with formula of:

Normalized Value = Actual Number / Max Number

Weightage score than being calculated by segregated the weightage of each nutrient to 0.2 per nutrient with return of value 1 for all 5 nutrients. To conclude the scoring, the formula being use is (n is the type of nutrient:

Weightage Score = (n1 * 0.2) + (n2 * 0.2) + (n3 * 0.2) + (n4 * 0.2) + (n5 * 0.2)

The outcome after calculate the weigthage scoring after ranked is as what being showed in above chart for top 10 foods - Lentils is the most healthiest food based on the final scoring due to its balanced content of essential nutrients. Black Beans at number 2 and Chickpeas at number 3 which these two foods have high protein and fiber content.

Pomegranate have the most nutrient contents of Vitamin C and also Antioxidant. This due to Pomegranate is a fruit which in general, fruit are often rich with these two nutrients.

Fish like Mackerel, Salmon and Sardines, are high in Protein and Calorie but lack of Antioxidant, Fiber and Vitamin C, which this directly resulting on lower overall scores based on the weightage set.

Conclusion to have a broad spectrum of nutritional contents, a person must commit to balancing different type of food from fish, fruit & legumes (plant based foods), etc.