

**ĐỀ TÀI PHÂN TÍCH
DINH DƯỠNG THỰC
THỰC PHẨM VIỆT
NAM**

HCMUS - VNU Students

Mục Lục

Tổng quan

- Đề tài
- Dataset

Tiền Xử Lý

- Xử lý kiểu

Phân Tích Dữ Liệu

- Trực quan hóa
- Phân tích thống kê

Tổng kết

- Kiến thức chuyên môn
- Kiến thức môn học

Về Nhóm

Là nhóm sinh viên đến từ trường Đại Học Khoa Học Tự Nhiên

18120184 Nguyễn Nguyên Khang

18120210 Phạm Tổng Bình Minh

18120237 Bạch Tăng Thắng

18120264 Nguyễn Duy Vũ

18120283 Nguyễn Chiêu Bản

Giáo Viên: Bùi Tiến Lên

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

Đề Tài Trực Quan Hóa Dữ Liệu

Nhóm 6 - Lớp Trực Quan Hóa Dữ Liệu 18_21

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

> **Tổng Quan**

- > Tiền Xử Lý
- > Phân Tích Dữ Liệu
- > Tổng kết

Tổng Quan

Nhóm muốn nghiên cứu, tìm hiểu các thông tin thú vị về thực phẩm Việt Nam. Đây là vấn đề về **khoa học thực phẩm**, gắn chặt với đời sống hàng ngày của chúng ta, ảnh hưởng trực tiếp đến sức khỏe, và đặc biệt là trong thời kỳ dịch bệnh, ta đã để tâm hơn đến sức khỏe, và có nhiều thời gian hơn để quan tâm đến bữa ăn

Tổng Quan

Từ suy nghĩ đó, nhóm đã sử dụng bộ dữ liệu được trích xuất từ sách "Bảng Thành Phần Dinh Dưỡng Việt Nam", cung cấp bởi Bộ Y Tế - Viện Dinh Dưỡng

Tổng Quan

	TÊN THỰC PHẨM	Calories (kcal)	Protein (g)	Fat (g)	Carbonhydrates (g)	Chất xơ (g)	Cholesterol (mg)	Canxi (mg)	Photpho (mg)	Sắt (mg)	Natri (mg)	Kali (mg)	Beta Caroten (mcg)	Vitamin A (mcg)	Vitamin B1 (mg)	Vitamin C (mg)	Loại
0	Gạo nếp cái	346	8,6	1,5	74,9	0,6	0	32	98	1,2	3	282	0	0	0,14	0,0	Ngũ cốc và sản phẩm chế biến từ chúng
1	Gạo tẻ	344	7,9	1,0	76,2	0,4	0	30	104,0	1,3	5	241	0	0	0,10	0,0	Ngũ cốc và sản phẩm chế biến từ chúng
2	Bắp tươi	196	4,1	2,3	39,6	1,2	0	20	187,0	1,5	0	0	170	0	0,21	0,0	Ngũ cốc và sản phẩm chế biến từ chúng

Dữ liệu có 162 dòng, 17 cột, mỗi dòng đại diện cho một một loại thực phẩm Việt Nam và dinh dưỡng của nó.

Tổng Quan

Lưu ý rằng trong tập dữ liệu của ta, carbohydrate là glucid, không bao gồm fiber/chất xơ và sugar/đường tổng số như Total carbohydrates trong các bảng dinh dưỡng thường được in trên bao bì.

Tổng Quan

Nutrition Facts

Serving size 121 g/mL

Servings Per Container --

Amount Per Serving

Calories 200 Calories from Fat 78

% Daily Value

Total Fat 9g **13%**

Saturated Fat 3g **13%**

Cholesterol 20mg **7%**

Sodium 358mg **15%**

Total Carbohydrate 21g **7%**

Dietary Fiber 1g **4%**

Sugars 8g

Protein 8g

Vitamin A 10% • Vitamin C 7%

Calcium 9% • Iron 6%

Nguồn:

https://www.researchgate.net/figure/Nutrition-Facts-for-general-audience-foods-Average-nutritional-content-of-foods-in_fig1_7643116

Tổng Quan

Theo tài liệu Bộ Y Tế - Viện Dinh Dưỡng Việt Nam 2007[1]:

Năng lượng và ba chất dinh dưỡng chính sinh năng lượng:

- Protid/protein/chất đạm
- Lipid/chất béo
- Glucid/chất bột đường và chất xơ

[1]Bộ Y Tế - Viện Dinh Dưỡng Việt Nam. Nhu cầu dinh dưỡng khuyến nghị cho người Việt Nam, 2007.

http://chuyentrang.viendinhduong.vn/FileUpload/Documents/240312_031923bang_nhu_cau_dd_cho_nguoi_viet_nam_2006.pdf?fbclid=IwAR1sC1qhluzhJUdRxK6Hi-F2XboQs5b86wVLHsLCi_WE9GvhbFBmtjxxGg4

Tổng Quan

Theo tài liệu Bộ Y Tế - Viện Dinh Dưỡng Việt Nam 2007[1]:

Các chất khoáng và vi khoáng:

- Calci
- Sắt
- Iod
- Kẽm
- Selen

[1]Bộ Y Tế - Viện Dinh Dưỡng Việt Nam. Nhu cầu dinh dưỡng khuyến nghị cho người Việt Nam, 2007.

http://chuyentrang.viendinhduong.vn/FileUpload/Documents/240312_031923bang_nhu_cau_dd_cho_nguoi_viet_nam_2006.pdf?fbclid=IwAR1sC1qhluzhJUdRxK6Hi-F2XboQs5b86wVLHsLCi_WE9GvhbFBmtjxxGg4

Tổng Quan

Theo tài liệu Bộ Y Tế - Viện Dinh Dưỡng Việt Nam 2007[1]:

Các vitamin:

- Vitamin tan trong dầu: Vitamin A, D, E, K
- Vitamin tan trong nước: C, B1, B2, B3 (Niacin), B9 (Folat), B6 và B12

[1]Bộ Y Tế - Viện Dinh Dưỡng Việt Nam. Nhu cầu dinh dưỡng khuyến nghị cho người Việt Nam, 2007.

http://chuyentrang.viendinhduong.vn/FileUpload/Documents/240312_031923bang_nhu_cau_dd_cho_nguoi_viet_nam_2006.pdf?fbclid=IwAR1sC1qhluzhJUdRxK6Hi-F2XboQs5b86wVLHsLCi_WE9GvnbFBmtjxxGg4

Tổng Quan

Theo tài liệu Bộ Y Tế - Viện Dinh Dưỡng Việt Nam 2007[1]:

Nước và các chất điện giải:

- Sodium (Na)
- Potassium (K)
- Chloride (Cl)

[1]Bộ Y Tế - Viện Dinh Dưỡng Việt Nam. Nhu cầu dinh dưỡng khuyến nghị cho người Việt Nam, 2007.

http://chuyentrang.viendinhduong.vn/FileUpload/Documents/240312_031923bang_nhu_cau_dd_cho_nguoi_viet_nam_2006.pdf?fbclid=IwAR1sC1qhluzhJUdRxK6Hi-F2XboQs5b86wVLHsLCi_WE9GvnbFBmtjxxGg4

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

Đề Tài Trực Quan Hóa Dữ Liệu

Nhóm 6 - Lớp Trực Quan Hóa Dữ Liệu 18_21

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

> Tổng Quan

> **Tiền Xử Lý**

> Phân Tích Dữ Liệu

> Tổng kết

Tiền Xử Lý

- Chỉnh sửa kiểu thuộc tính: một số cột đang ở dạng categorical thay vì numerical do cách viết “,” thay vì “.”
- Không có giá trị thiếu

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

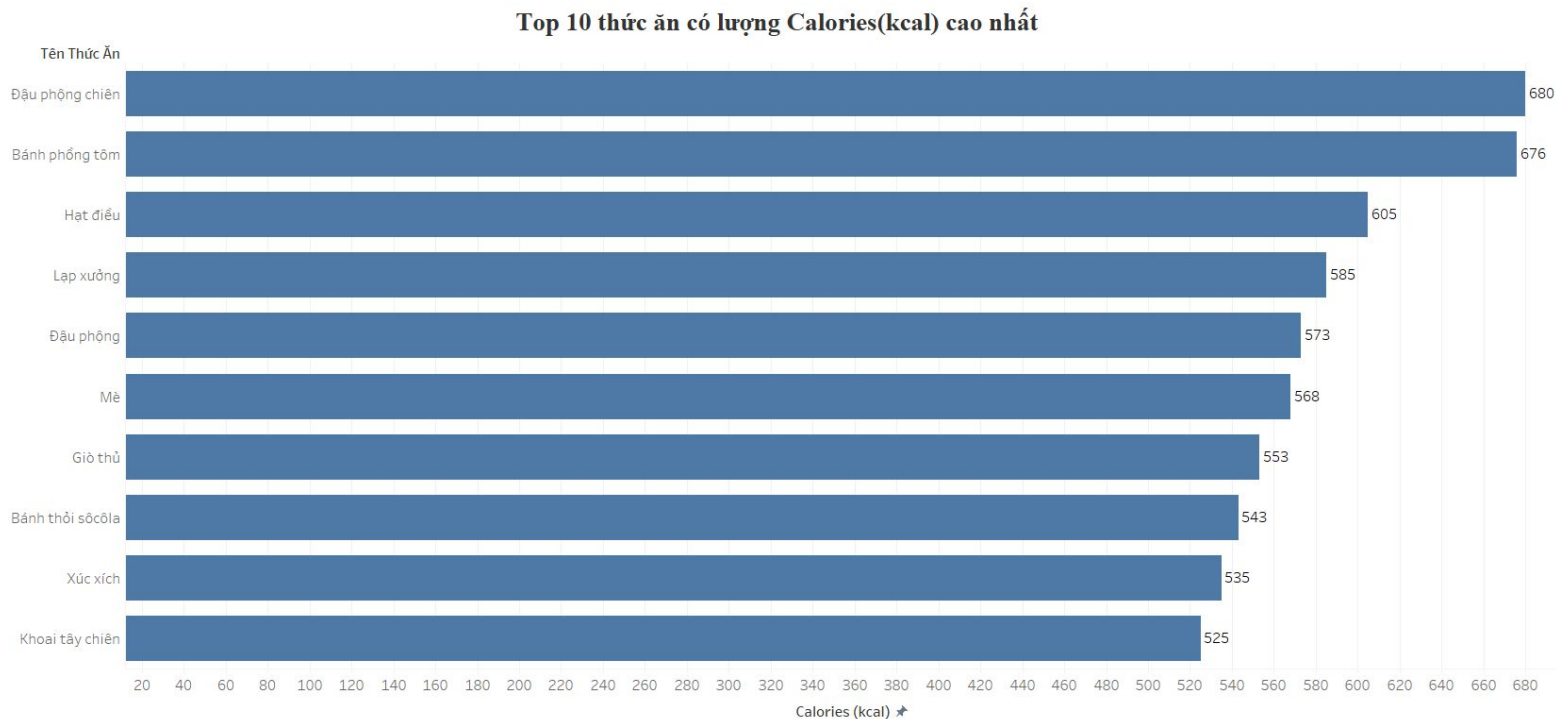
Đề Tài Trực Quan Hóa Dữ Liệu

Nhóm 6 - Lớp Trực Quan Hóa Dữ Liệu 18_21

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

- > Tổng Quan
- > Tiền Xử Lý
- > **Phân Tích Dữ Liệu**
- > Tổng kết

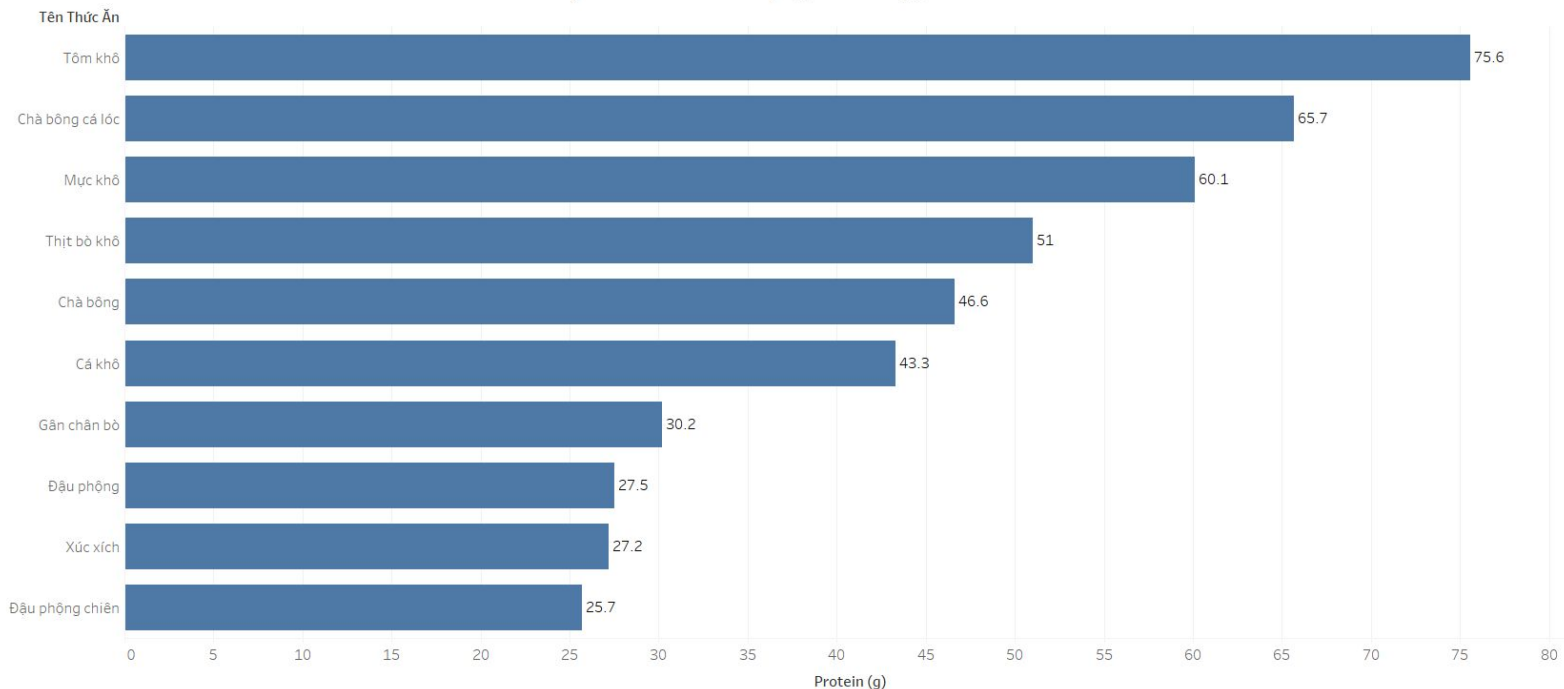
Top 10 Thực Phẩm Có Dinh Dưỡng Cao



“đậu phộng chiên” và “bánh phồng tôm” là các thức ăn có lượng calories cao nhất.

Top 10 Thực Phẩm Có Dinh Dưỡng Cao

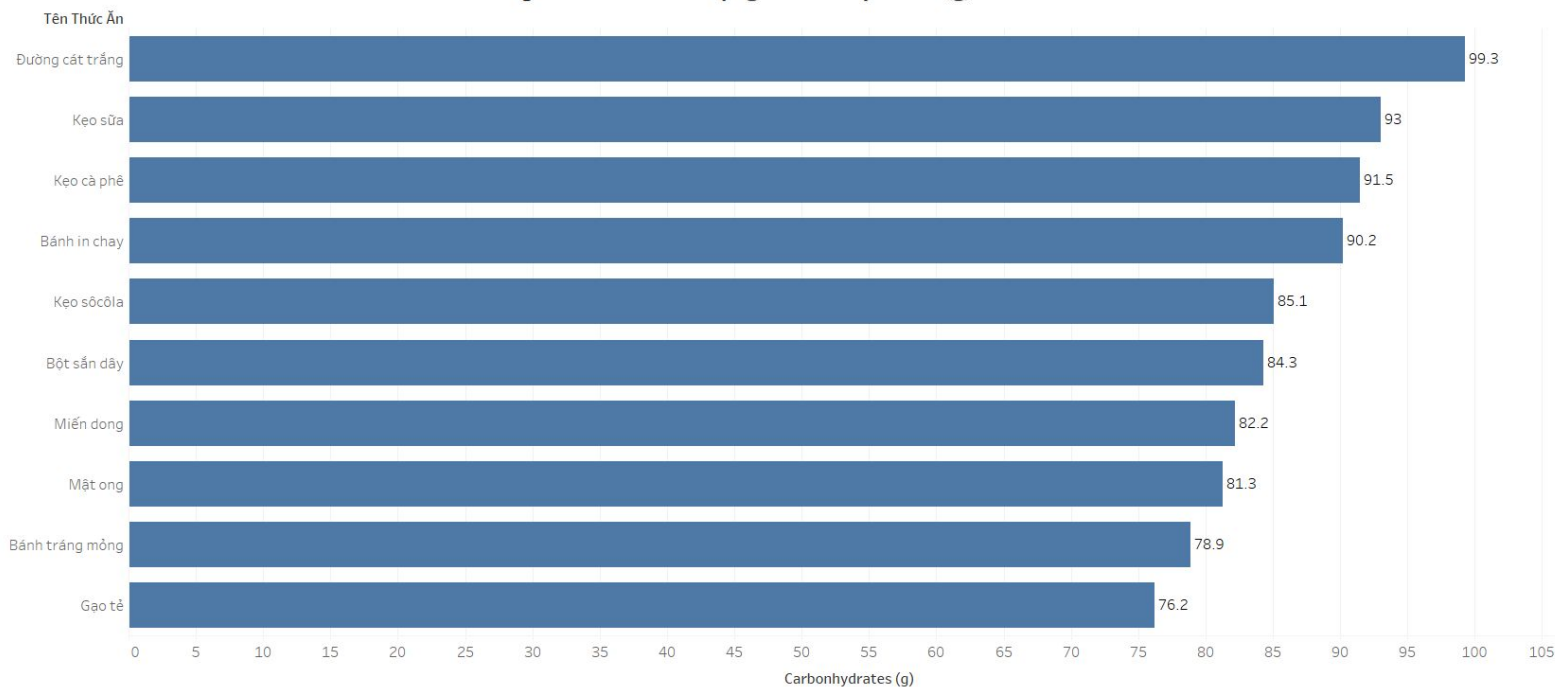
Top 10 thức ăn có lượng Protein(g) cao nhất



“Tôm khô” và “chà bông cá lóc” là các thức ăn có lượng protein cao nhất.

Top 10 Thực Phẩm Có Dinh Dưỡng Cao

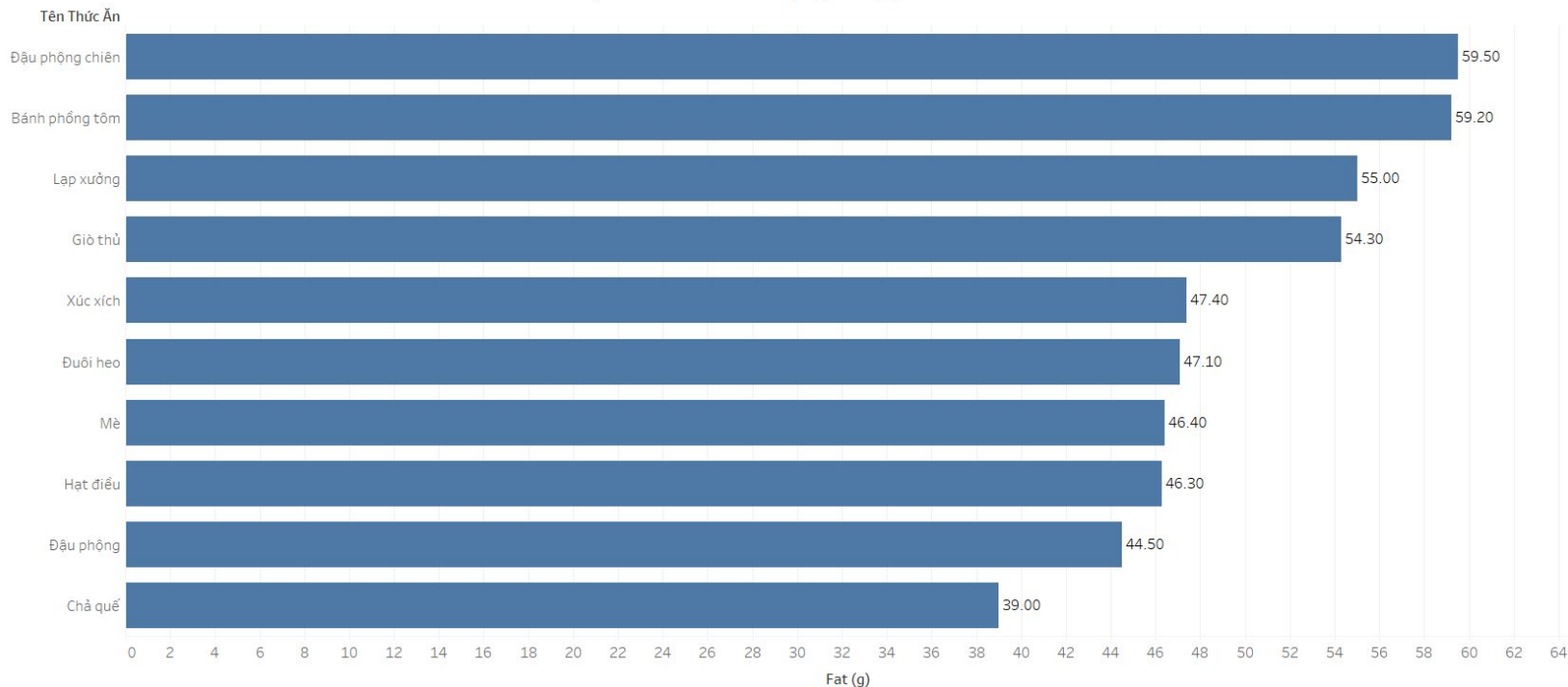
Top 10 thức ăn có lượng Carbohydrates(g) cao nhất



Đường cát trắng, bánh kẹo, mật ong, và các chế phẩm từ sắn, gạo như miến dong, bột sắn dây là loại thức ăn có lượng carbohydrates(g) cao.

Top 10 Thực Phẩm Có Dinh Dưỡng Cao

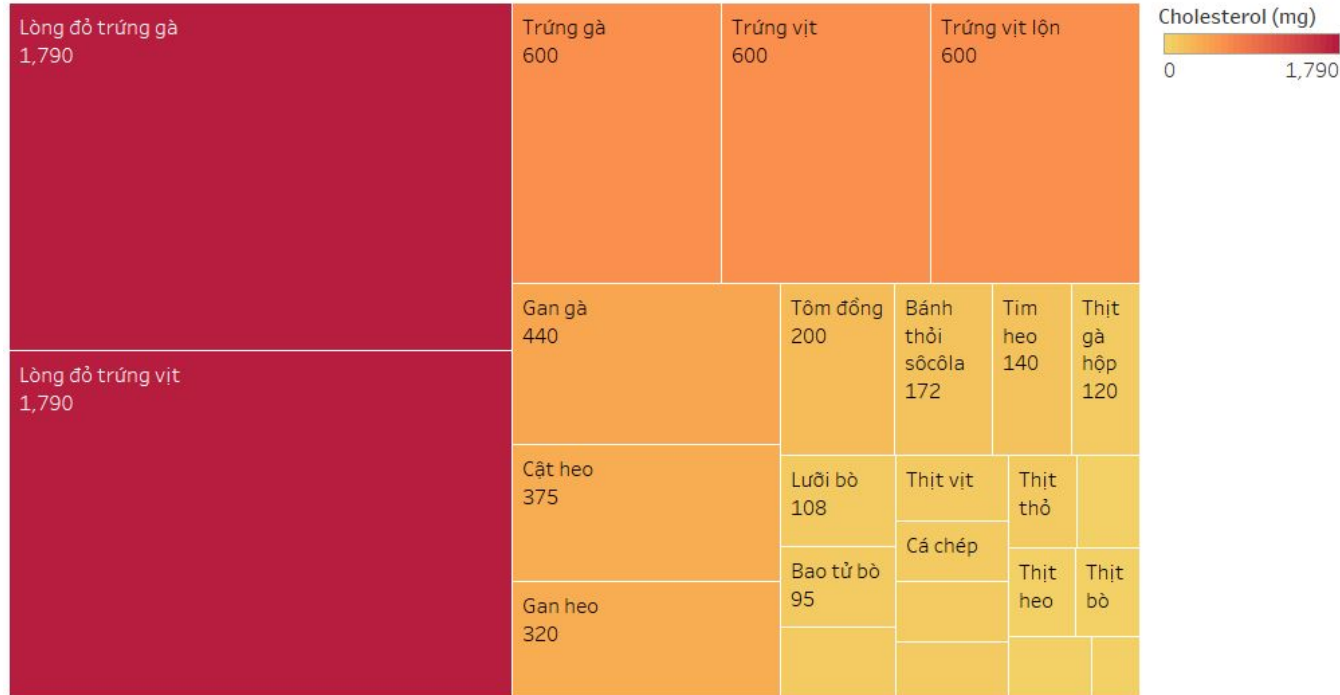
Top 10 thức ăn có lượng Fat(g) cao nhất



Các loại đồ chiên như **đậu phộng chiên**, **bánh phồng tôm**, **lạp xường**, các loại mỡ động vật như **giò thủ**, **xúc xích**, **đuôi heo**, mà các loại hạt như **mè**, **hạt điều**, **đậu phộng** là các loại thực phẩm giàu **chất béo**.

Biểu đồ

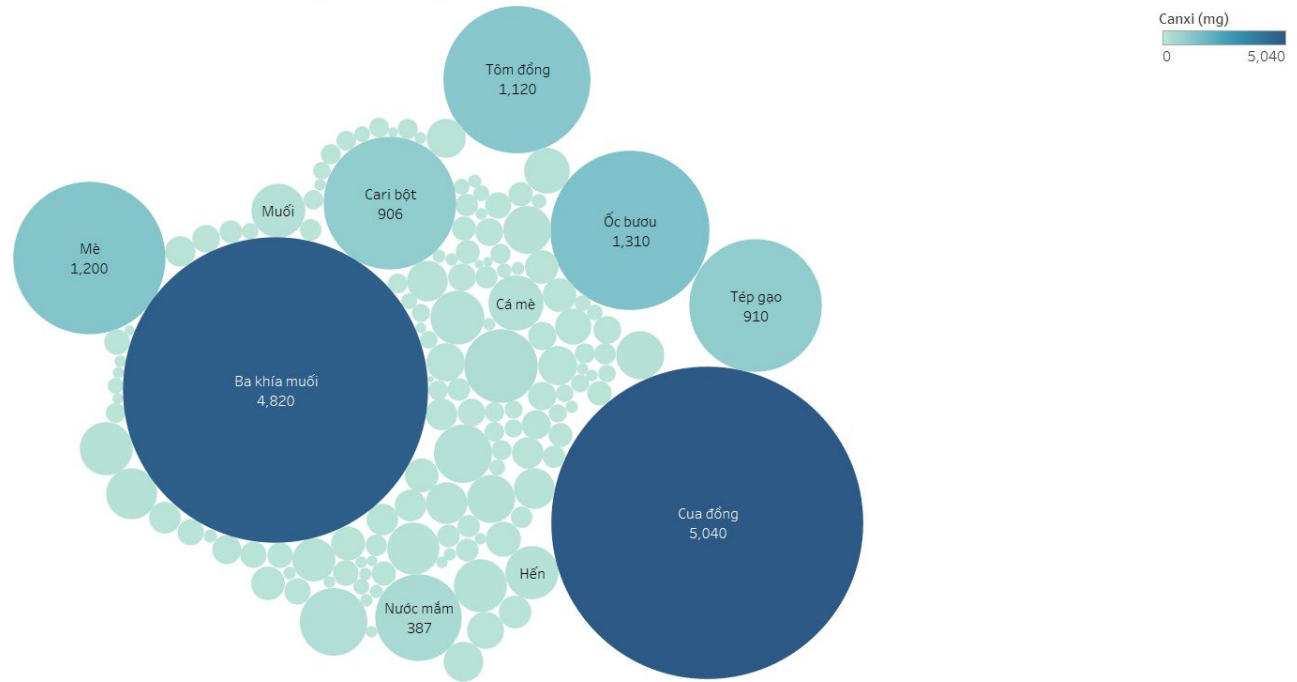
Biểu đồ thể hiện lượng Cholesterol(mg) trong thức ăn



Các loại **trứng** đặc biệt là **lòng đỏ trứng** có chứa lượng cholesterol cao, tiếp đó nội tạng động vật như **gan, cật, tim**. Vì vậy người cao tuổi, người bị xơ vữa động mạch nên hạn chế các thức ăn trên để **ngừa mỡ trong máu**.

Biểu đồ

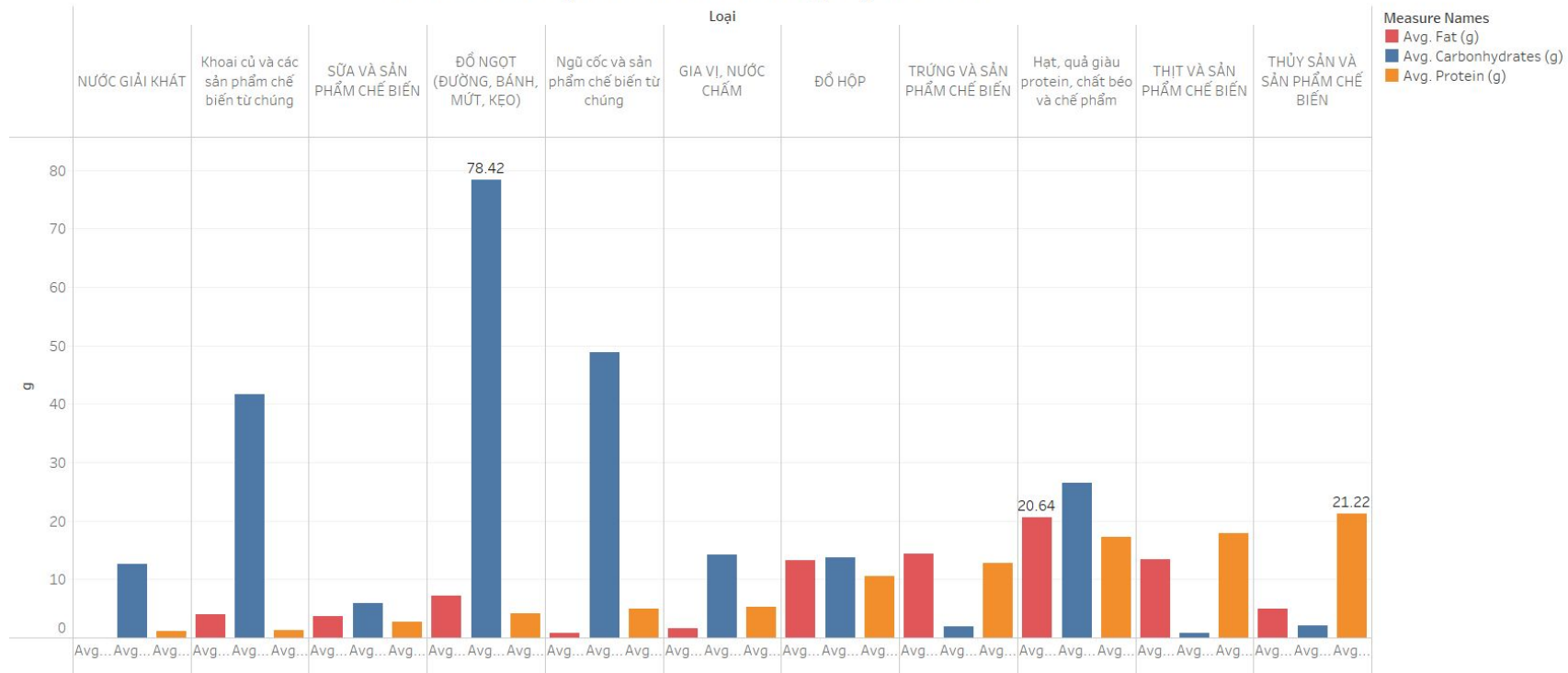
Biểu đồ lượng canxi(mg) trong mỗi loại thức ăn



Để bổ sung canxi ta cần ăn các loại thực phẩm giàu canxi như **Ba khía muối, cua đồng, tôm đồng.**

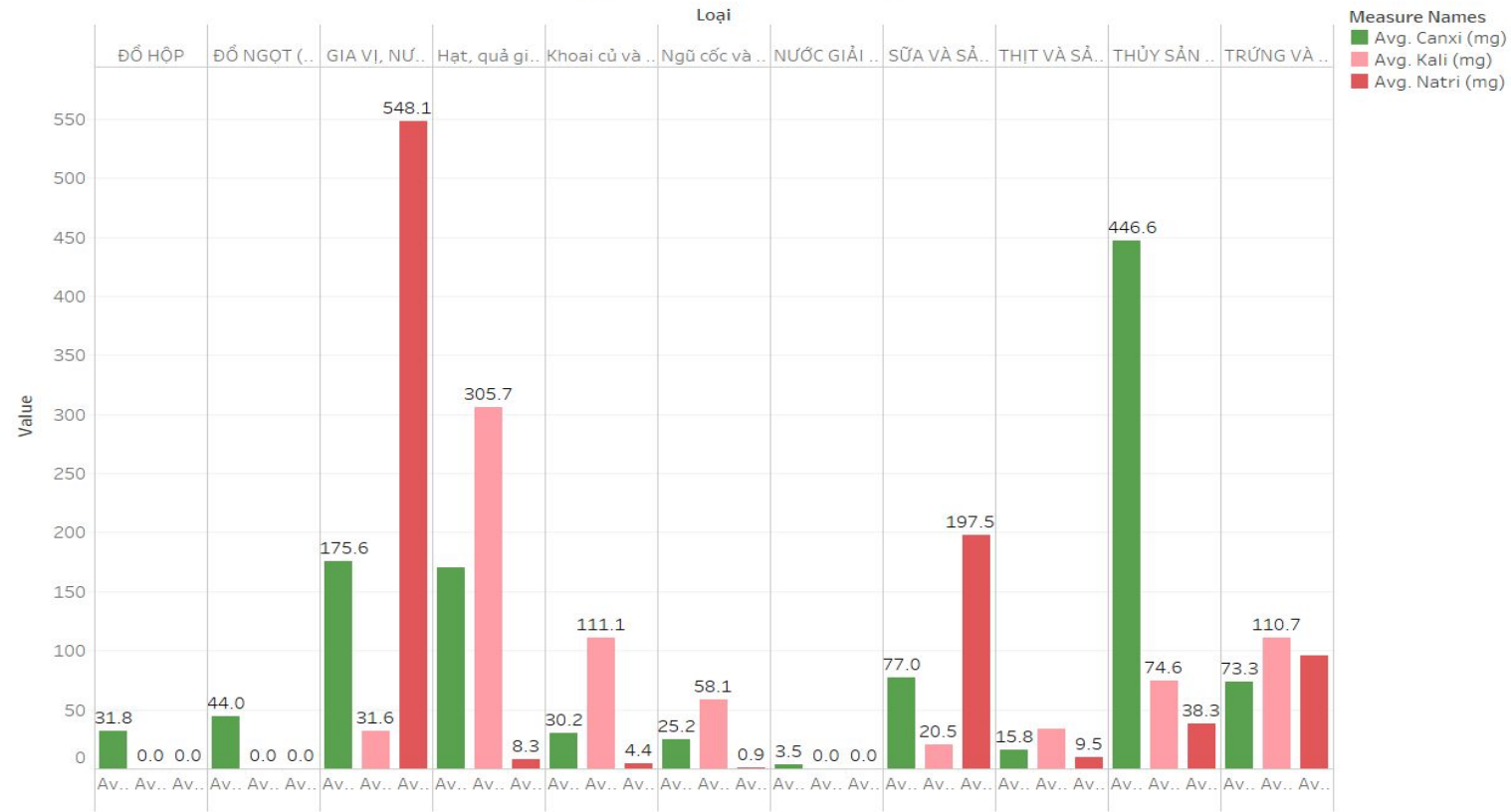
Biểu đồ

Biểu đồ carb,protein,fat của từng loại thức ăn



Biểu đồ

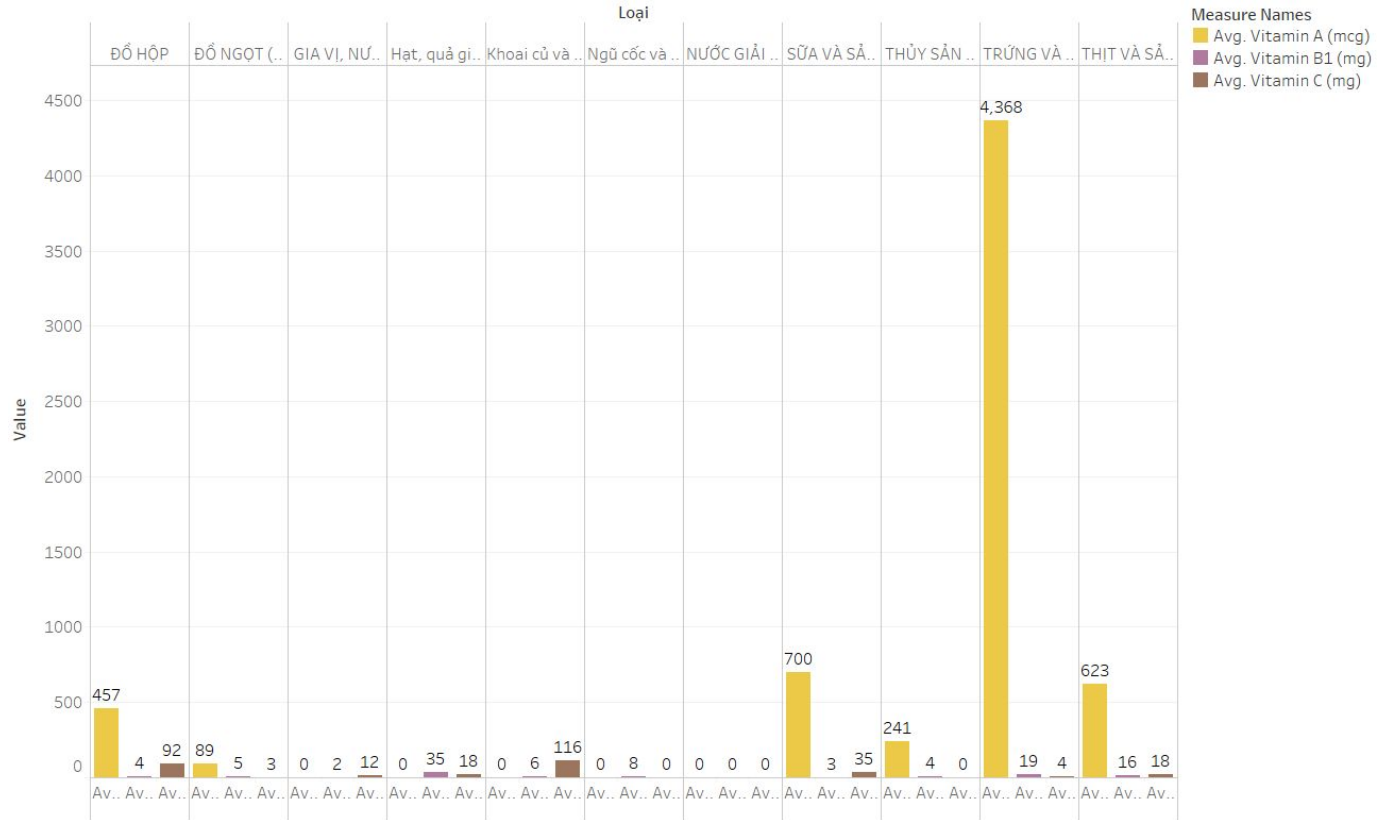
Biểu đồ Canxi, Kali, Natri có trong từng loại thức ăn



Avg. Canxi (mg), Avg. Kali (mg) and Avg. Natri (mg) for each Loại. Color shows details about Avg. Canxi (mg), Avg. Kali (mg) and Avg. Natri (mg).

Biểu đồ

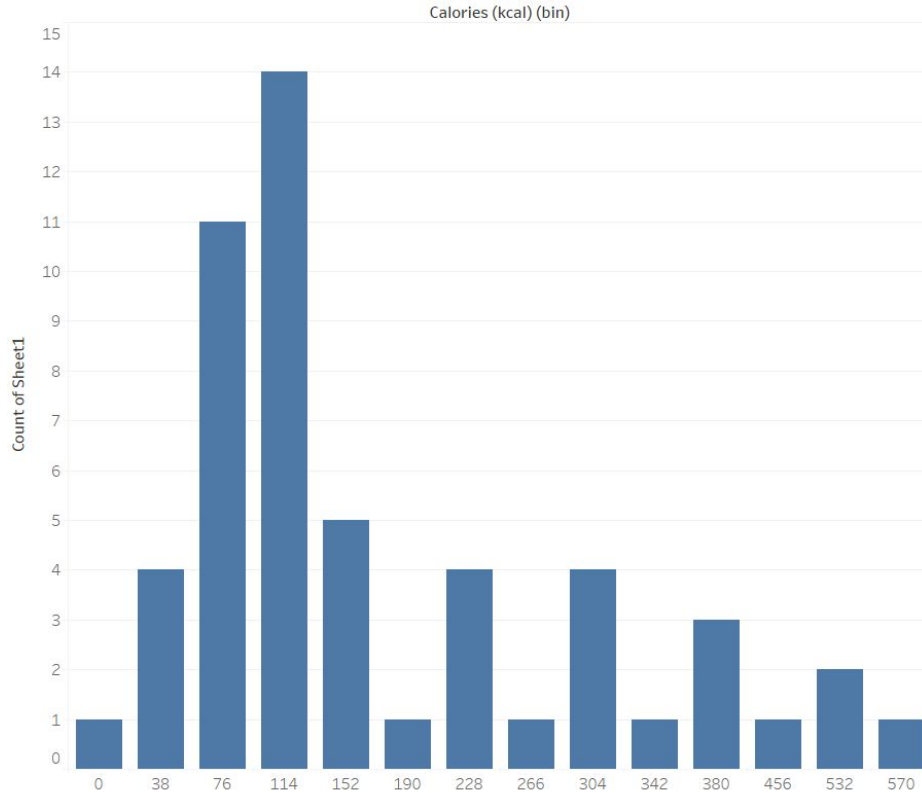
Biểu đồ vitamin có trong từng loại thức ăn



Avg. Vitamin A (mcg), Avg. Vitamin B1 (mg) and Avg. Vitamin C (mg) for each Loại. Color shows details about Avg. Vitamin A (mcg), Avg. Vitamin B1 (mg) and Avg. Vitamin C (mg). The view is filtered on Loại, which keeps 11 of 11 members.

Biểu đồ

Biểu đồ phân bố Calories của Thịt và sản phẩm từ thịt

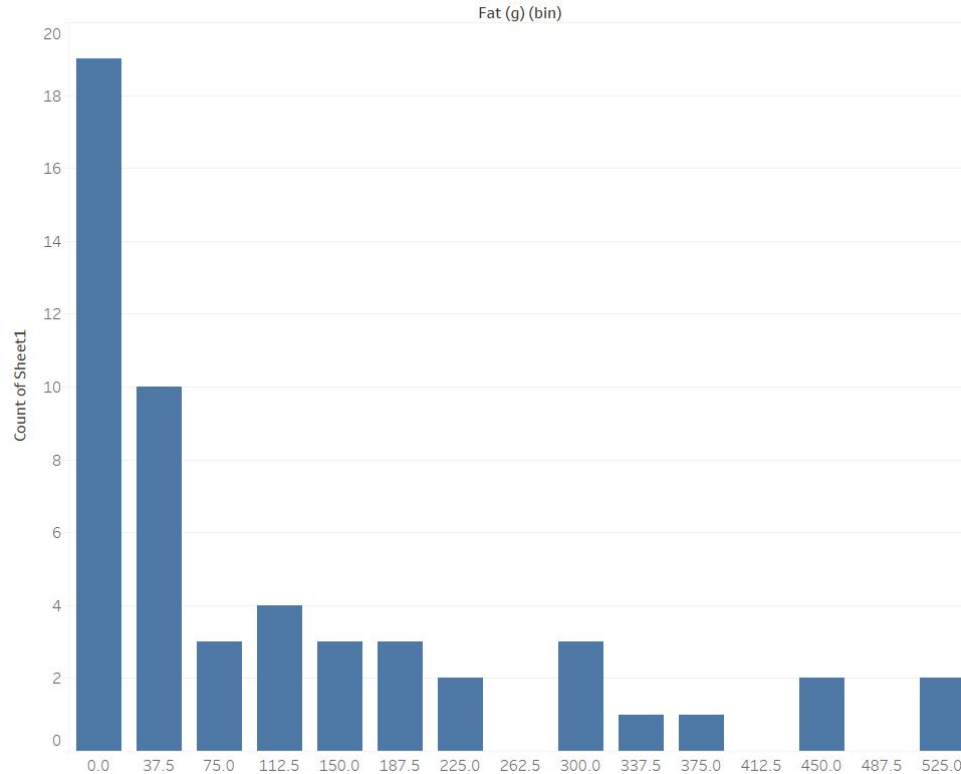


Count of Sheet1 for each Calories (kcal) (bin). The data is filtered on Loại and Tên Thức Ăn. The Loại filter keeps THỊT VÀ SẢN PHẨM CHẾ BIẾN. The Tên Thức Ăn filter keeps 162 of 162 members.

- Lượng Calories phân bố ở mức 76 đến 152 calories mỗi món (30 món), một mức phân bố khá thấp cho những món chứa nhiều Calories như thịt
- Các món ăn có sự rõ ràng về mặt năng lượng cung cấp
- Khẩu phần ăn đầy đủ lượng Calories cần phải được tạo từ sự kết hợp của nhiều món ăn

Biểu đồ

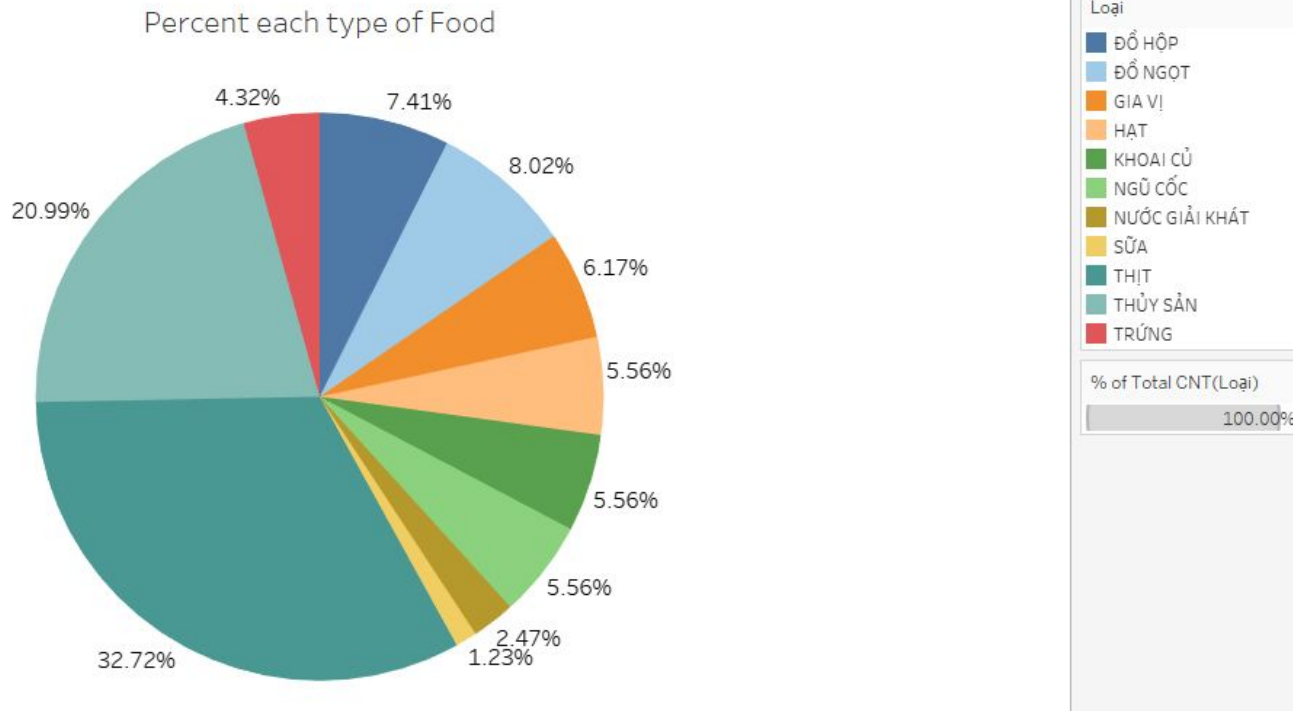
Biểu đồ phân bố Fat trong Thịt và các sản phẩm từ thịt



Count of Sheet1 for each Fat (g) (bin). The data is filtered on Loại, which keeps THỊT VÀ SẢN PHẨM CHẾ BIẾN.

- Lượng Fats phân bố tương đối đồng đều
- Các sản phẩm từ thịt chứa tương đối ít Fats, hơn một nửa có chứa lượng Fats là bằng 0
- Còn lại rải rác từ 37.5g đến 525g Fat tuy nhiên ở số lượng rất ít món ăn

Biểu đồ



- “Thịt” và “Thủy sản” là 2 loại thực phẩm xuất hiện nhiều nhất trong tập dữ liệu. Tổng của 2 loại thực phẩm này chiếm 1 nửa của tập dữ liệu
- “Sữa” và “Nước giải khát” là 2 loại thực phẩm xuất hiện ít nhất

So Sánh Các Thực Phẩm Quen Thuộc

Các thực phẩm quen thuộc do nhóm tự chọn như sau:

'Bánh mì', 'Gạo tẻ', 'Bún', 'Đậu phụ', 'Thịt bò', 'Thịt gà ta', 'Thịt heo ba chỉ',
'Chả lụa', 'Xúc xích', 'Cá ngừ', 'Cá rô phi', 'Cá thu', 'Tôm biển', 'Trứng gà',
'Sữa bò tươi', 'Đường cát trắng', 'Muối', 'Nước mắm', 'Xì dầu', 'Bia'

So Sánh Các Thực Phẩm Quen Thuộc

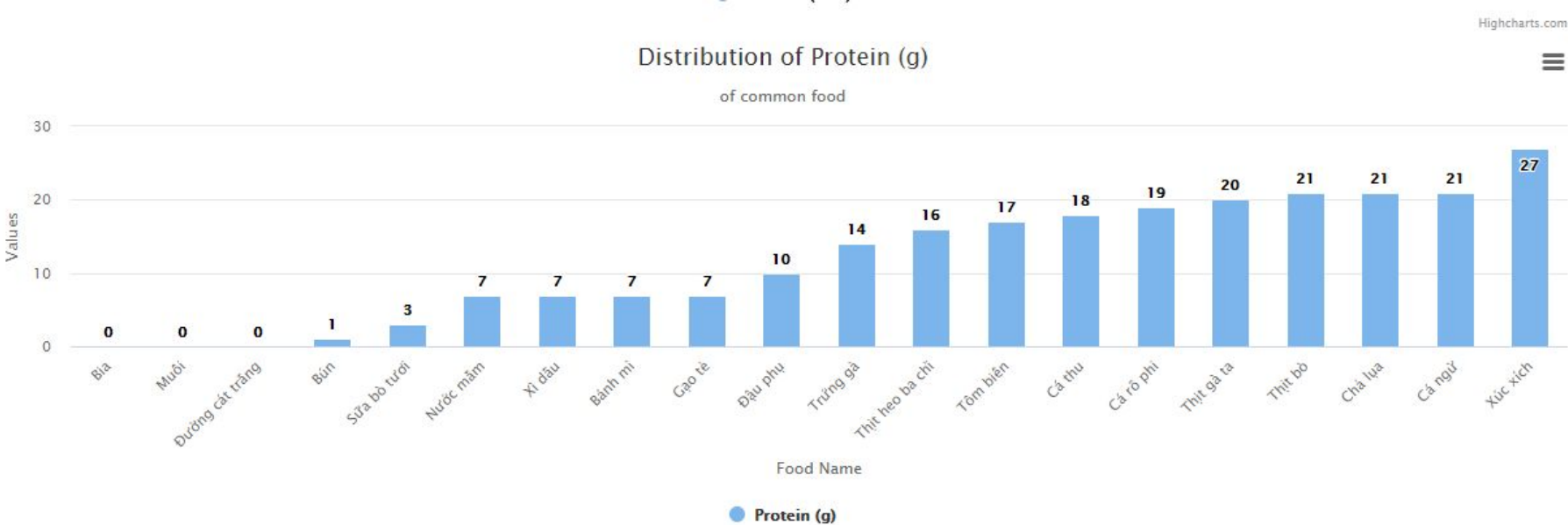


Distribution of Calories (kcal)

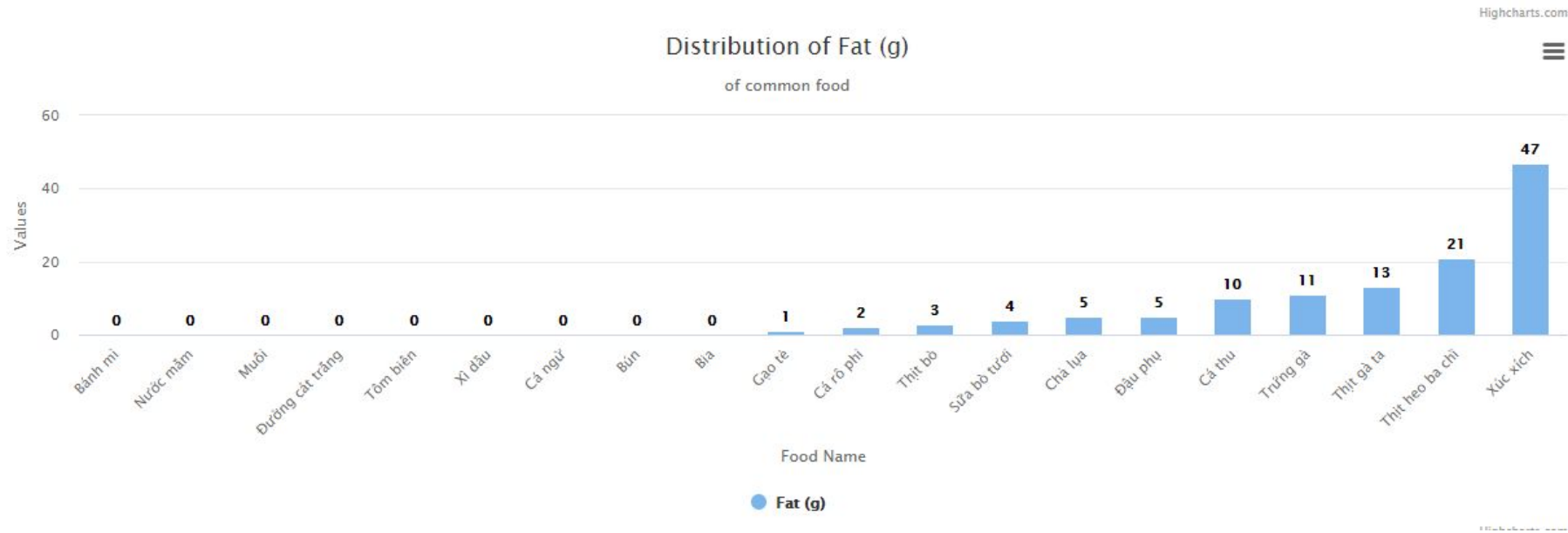
of common food



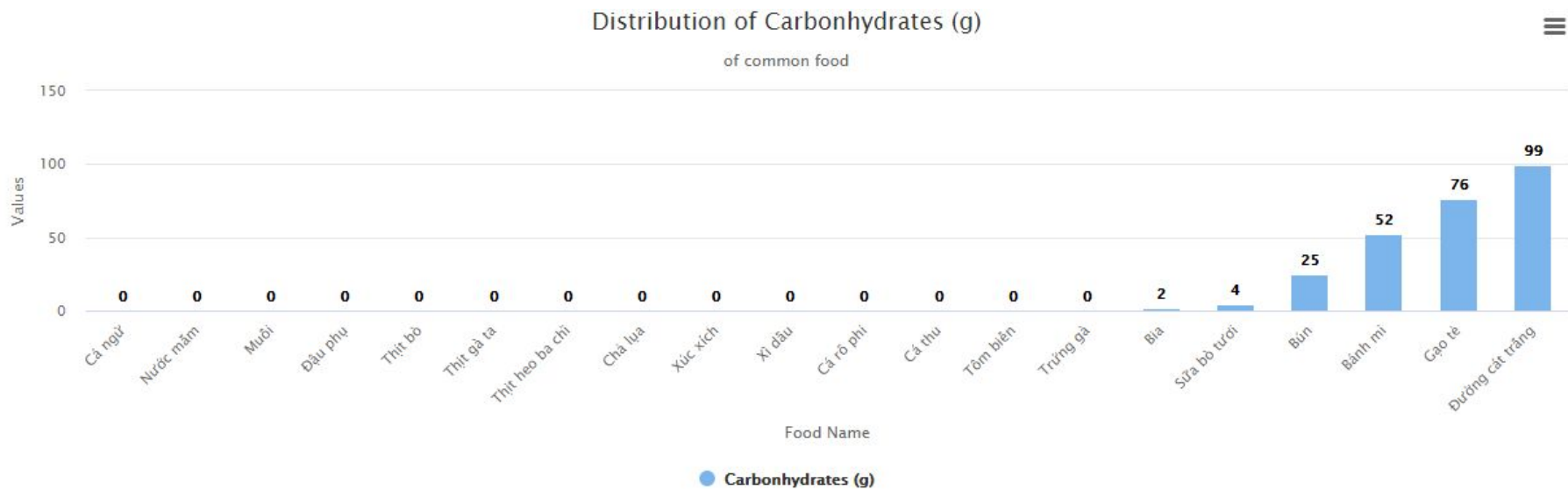
So Sánh Các Thực Phẩm Quen Thuộc



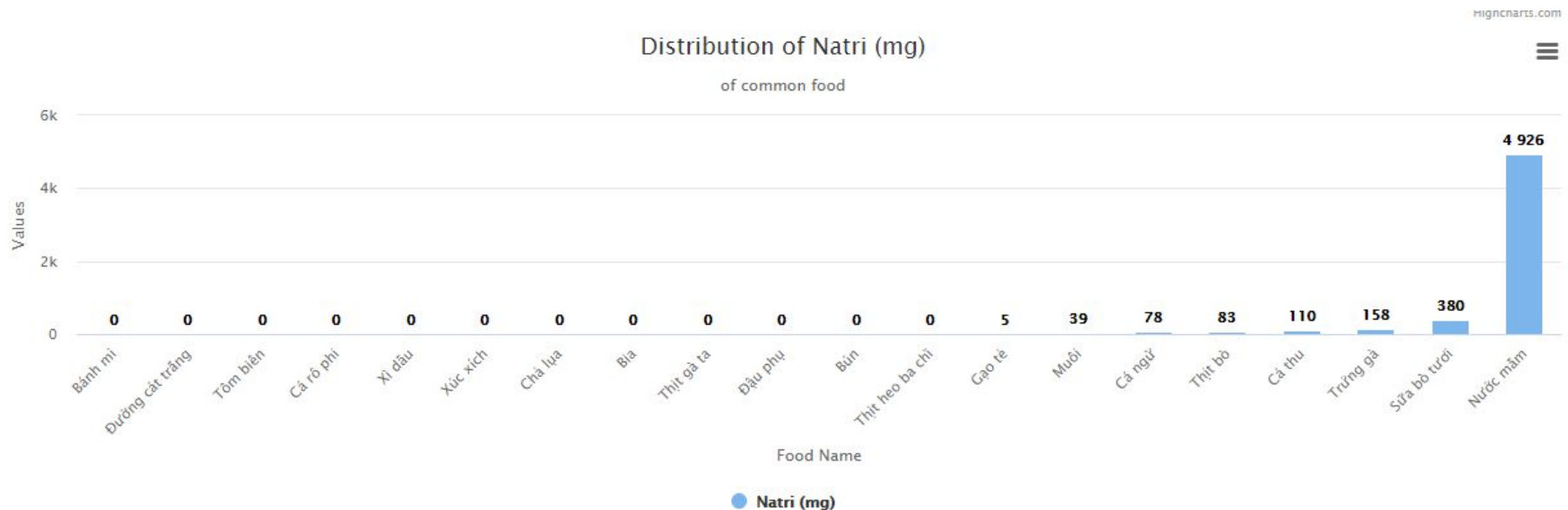
So Sánh Các Thực Phẩm Quen Thuộc



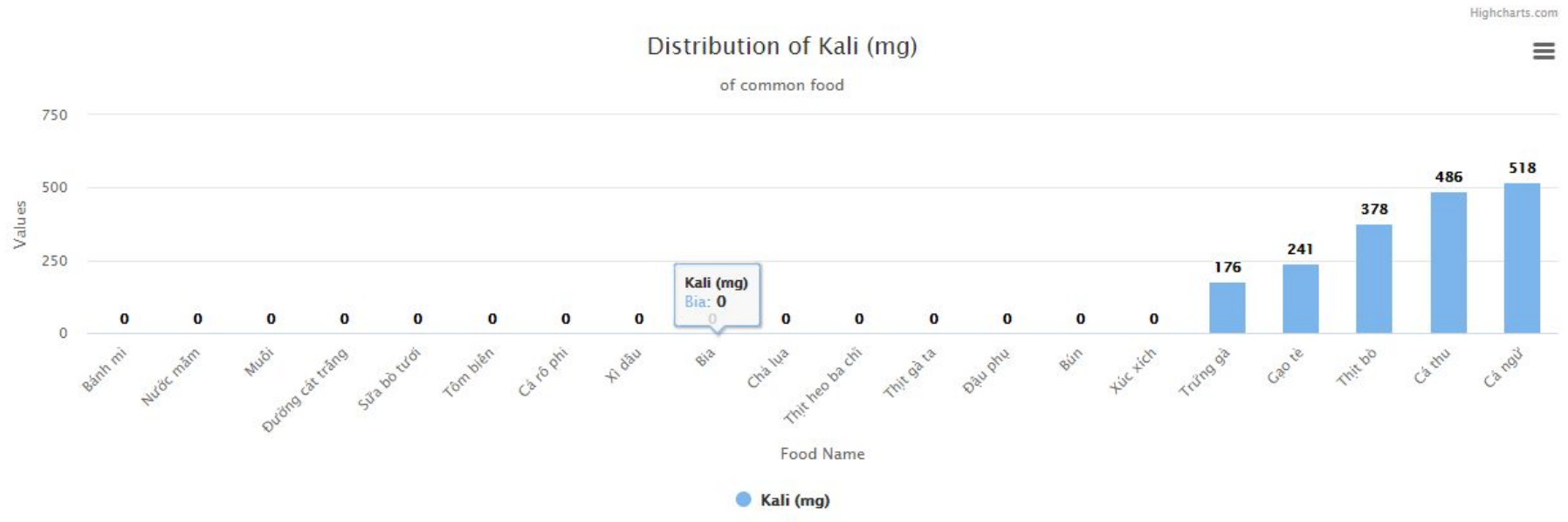
So Sánh Các Thực Phẩm Quen Thuộc



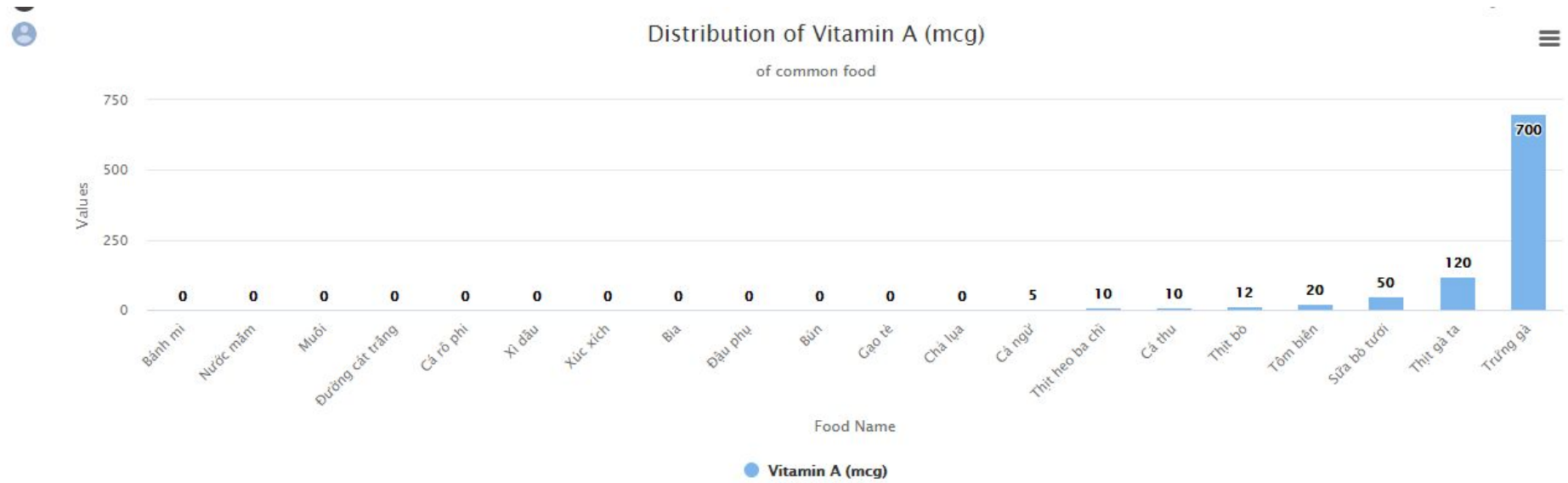
So Sánh Các Thực Phẩm Quen Thuộc



So Sánh Các Thực Phẩm Quen Thuộc

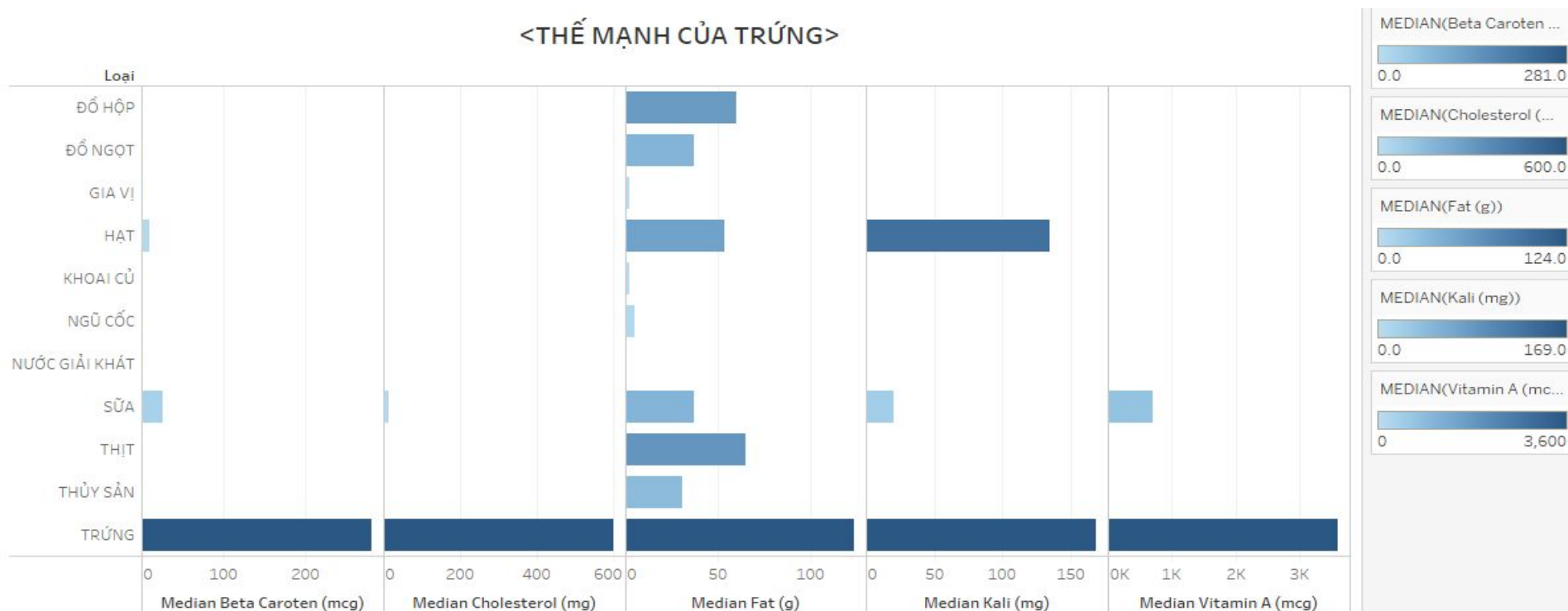


So Sánh Các Thực Phẩm Quen Thuộc



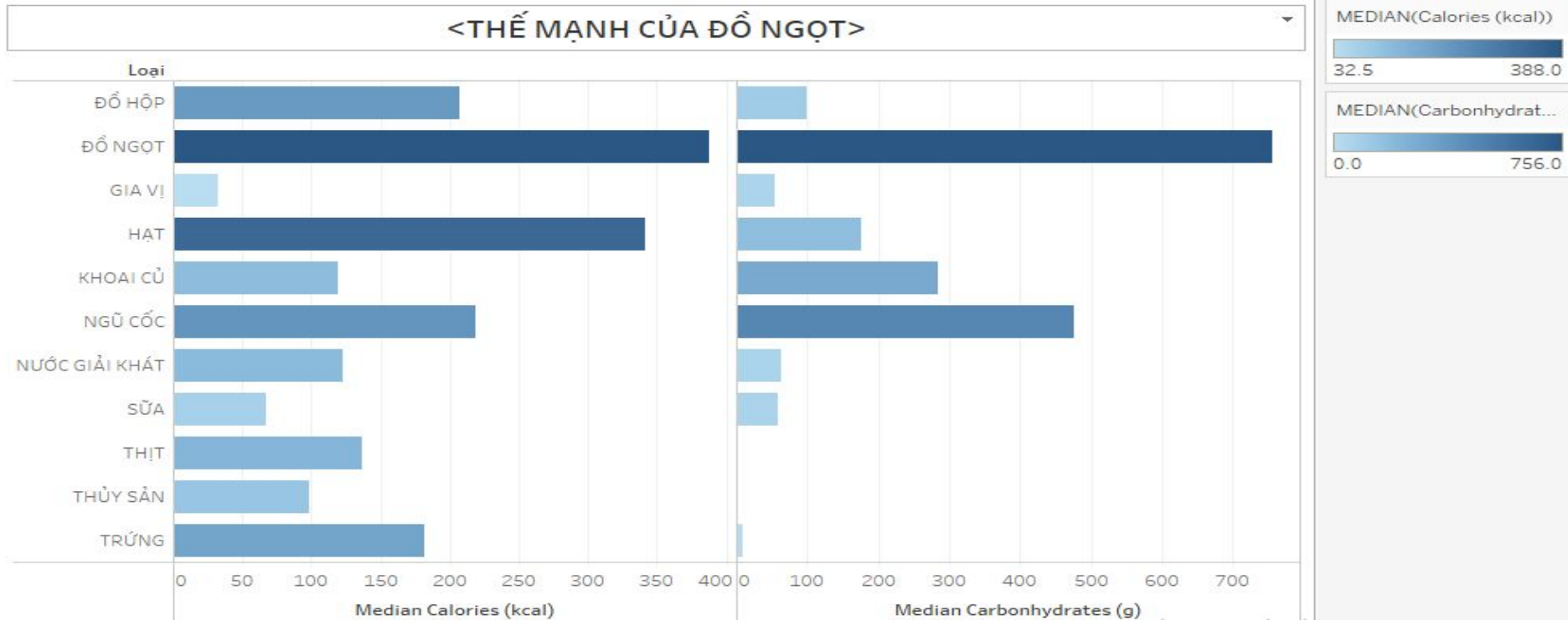
Loại Thực Phẩm

<THỂ MẠNH CỦA TRỨNG>



- Có thể thấy “Trứng” là loại thực phẩm trung bình cung cấp nhiều “Beta Caroten”, “Cholesterol”, “Fat”, “Kali”, “Vitamin A”

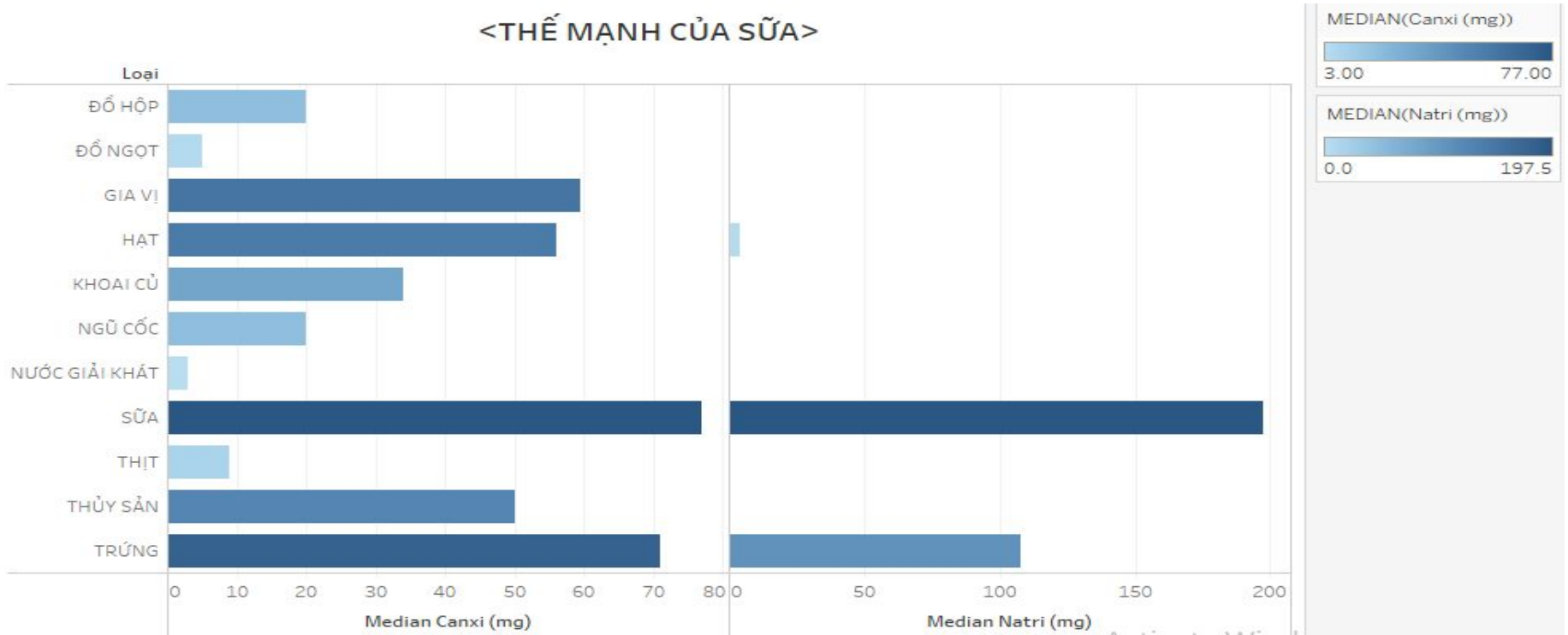
Loại Thực Phẩm



- Có thể thấy “Đồ ngọt” là loại thực phẩm trung bình cung cấp nhiều “Calories” và “Carbonhydrates”

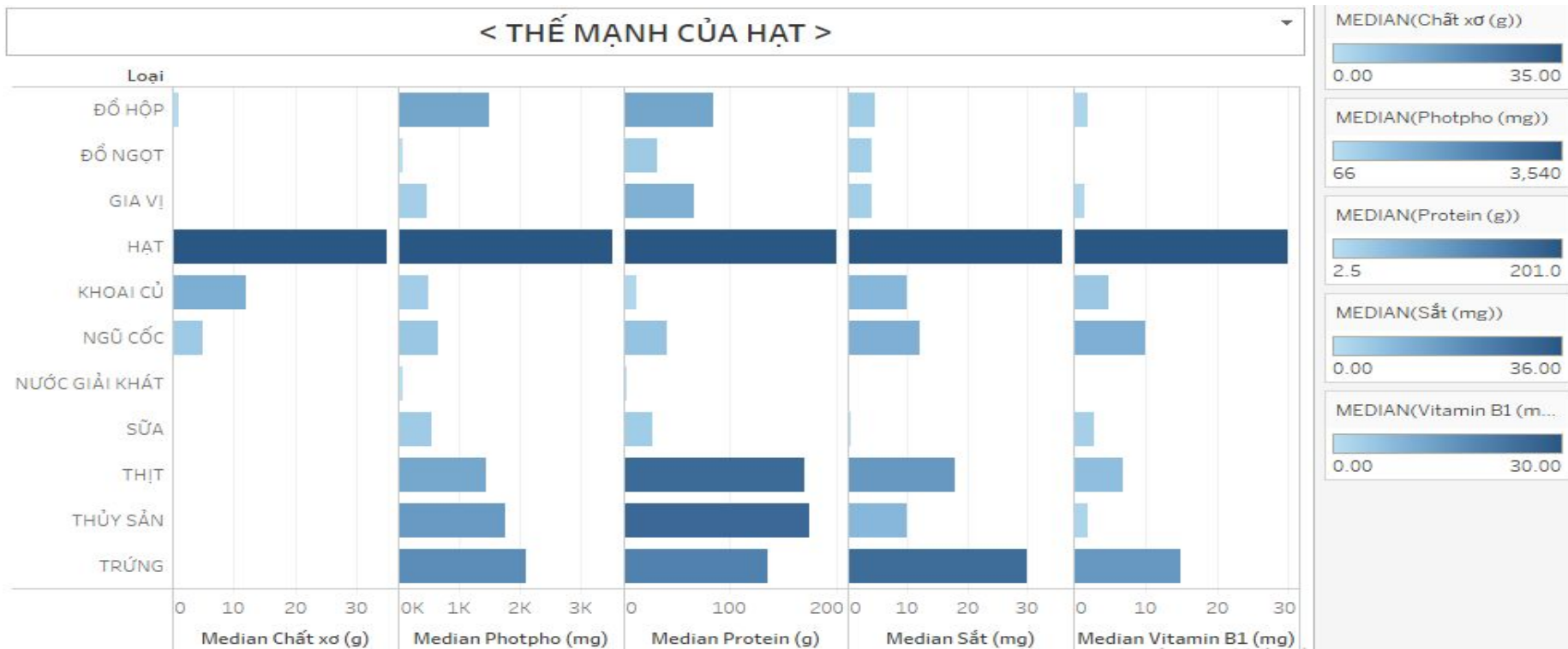
Loại Thực Phẩm

<THỂ MẠNH CỦA SỮA>



- Có thể thấy “Sữa” là loại thực phẩm cung cấp nhiều “Canxi” và “Natri” nhất

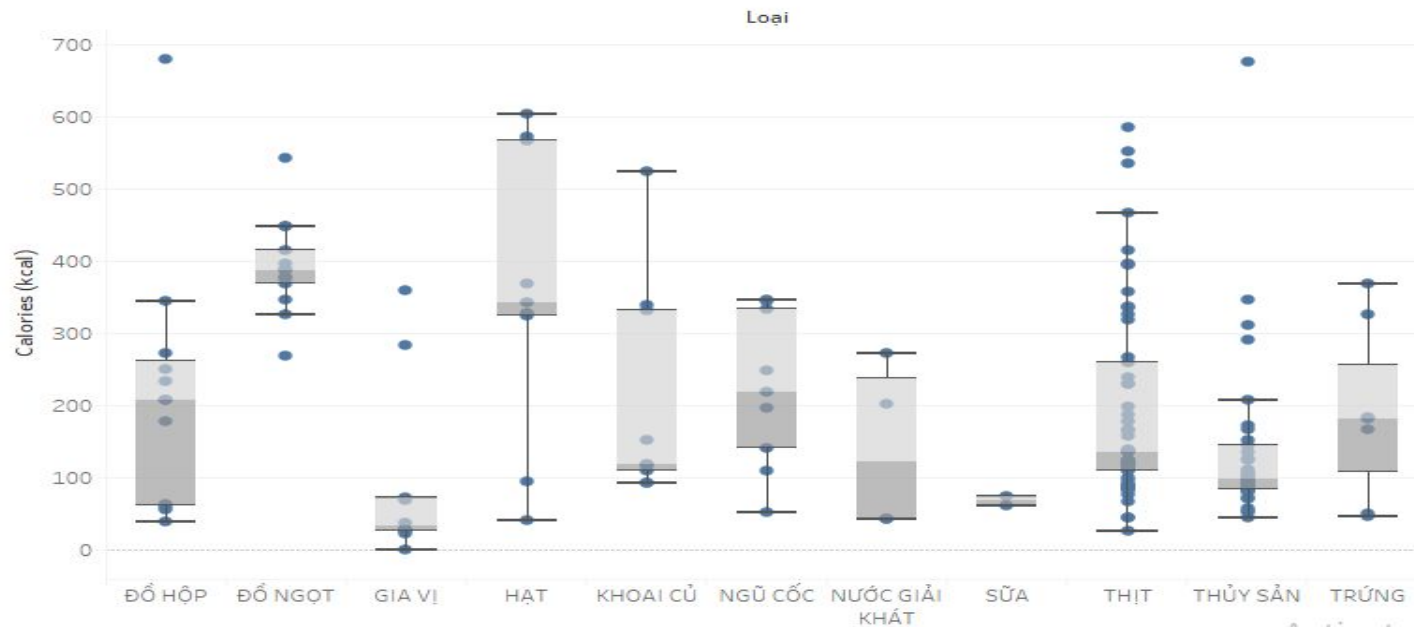
Loại Thực Phẩm



- Có thể thấy “Hạt” là loại thực phẩm cung cấp nhiều “Chất xơ”, “Photpho”, “Protein”, “Sắt” và “Vitamin B1” nhất

Loại Thực Phẩm

<PHÂN BỐ CALORIES CỦA CÁC LOẠI THỰC PHẨM>



Summary

Count:	162
SUM(Calories (kcal))	
Sum:	32,792
Average:	202.42
Minimum:	0
Maximum:	680
Median:	138.00
Standard deviation:	152.6
First quartile:	89.00
Third quartile:	326.75

- Như đã phân tích ở trên thì “Đồ ngọt” là loại thực phẩm trung bình có chỉ số “Calories” cao nhất tuy nhiên thực phẩm có lượng Calories cao nhất lại là “Đậu phộng chiên” của loại “Đồ hộp”. Thấp nhất là “Muối” của “Gia vị”.
- Mặt khác có thể thấy “Đồ ngọt” có dao động về mức Calories khá thấp và các loại thực phẩm hầu như đều nằm tập trung ở mức Q3 của “Calories” vì vậy kết luận “Đồ ngọt” là loại thực phẩm cung cấp nhiều “Calories” là không hề sai

Trung bình giữa 'Natri (mg)', 'Kali (mg)' giữa
'Thịt và sản phẩm chế biến' và 'Thủy sản và sản
phẩm chế biến' có bằng nhau?

Ho: trung bình `Natri (mg)` ở Thịt và sản phẩm chế biến = trung bình `Natri (mg)` ở Thủy sản và sản phẩm chế biến

Ha: trung bình `Natri (mg)` ở Thịt và sản phẩm chế biến khác trung bình `Natri (mg)` ở Thủy sản và sản phẩm chế biến

stat=-1.794, p=0.073

KẾT LUẬN:

Với mức ý nghĩa 0.05, ta chấp nhận Ho, bác bỏ Ha.

Trung bình `Natri (mg)` ở Thịt và sản phẩm chế biến = trung bình `Natri (mg)` ở Thủy sản và sản phẩm chế biến

Ho: trung bình `Kali (mg)` ở Thịt và sản phẩm chế biến = trung bình `Kali (mg)` ở Thủy sản và sản phẩm chế biến

Ha: trung bình `Kali (mg)` ở Thịt và sản phẩm chế biến khác trung bình `Kali (mg)` ở Thủy sản và sản phẩm chế biến

stat=-1.467, p=0.142

KẾT LUẬN:

Với mức ý nghĩa 0.05, ta chấp nhận Ho, bác bỏ Ha.

Trung bình `Kali (mg)` ở Thịt và sản phẩm chế biến = trung bình `Kali (mg)` ở Thủy sản và sản phẩm chế biến

Mối Quan Hệ Giữa Các Biến

OLS Regression Results						
=====						
Dep. Variable:	Calories (kcal)	R-squared:	0.980			
Model:	OLS	Adj. R-squared:	0.978			
Method:	Least Squares	F-statistic:	563.8			
Date:	Sun, 18 Jul 2021	Prob (F-statistic):	6.60e-119			
Time:	17:41:33	Log-Likelihood:	-726.21			
No. Observations:	162	AIC:	1480.			
Df Residuals:	148	BIC:	1524.			
Df Model:	13					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	14.5964	3.909	3.734	0.000	6.871	22.322
Protein (g)	3.7971	0.227	16.714	0.000	3.348	4.246
Fat (g)	8.9035	0.135	65.971	0.000	8.637	9.170
Carbohydrates (g)	3.8868	0.076	50.918	0.000	3.736	4.038
Chất xơ (g)	-0.0463	1.605	-0.029	0.977	-3.218	3.125
Cholesterol (mg)	-0.0004	0.011	-0.037	0.971	-0.023	0.022
Canxi (mg)	-0.0009	0.003	-0.287	0.775	-0.007	0.006
Photpho (mg)	-0.0013	0.018	-0.071	0.943	-0.038	0.035
Sắt (mg)	-0.1898	0.349	-0.544	0.587	-0.879	0.499
Natri (mg)	-0.0027	0.005	-0.576	0.565	-0.012	0.006
Kali (mg)	-0.0019	0.013	-0.151	0.881	-0.027	0.023
Beta Caroten (mcg)	0.0023	0.016	0.149	0.882	-0.029	0.033
Vitamin A (mcg)	0.0005	0.002	0.211	0.834	-0.004	0.005
Vitamin B1 (mg)	0	0	nan	nan	0	0
Vitamin C (mg)	-0.3207	0.329	-0.975	0.331	-0.970	0.329
=====						
Omnibus:	328.643	Durbin-Watson:	0.894			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	113947.817			
Skew:	10.844	Prob(JB):	0.00			
Kurtosis:	131.105	Cond. No.	inf			
=====						

Mối Quan Hệ Giữa Các Biến

OLS Regression Results

Dep. Variable:	Calories (kcal)	R-squared:	0.980			
Model:	OLS	Adj. R-squared:	0.980			
Method:	Least Squares	F-statistic:	2574.			
Date:	Sun, 18 Jul 2021	Prob (F-statistic):	7.27e-134			
Time:	10:52:10	Log-Likelihood:	-727.24			
No. Observations:	162	AIC:	1462.			
Df Residuals:	158	BIC:	1475.			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	12.2154	3.366	3.629	0.000	5.566	18.864
Protein (g)	3.8258	0.160	23.916	0.000	3.510	4.142
Fat (g)	8.9346	0.126	70.715	0.000	8.685	9.184
Carbonhydrates (g)	3.8981	0.070	55.524	0.000	3.759	4.037
Omnibus:	330.172	Durbin-Watson:	0.870			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	116571.152			
Skew:	10.944	Prob(JB):	0.00			
Kurtosis:	132.579	Cond. No.	61.5			

Mối Quan Hệ Giữa Các Biến

OLS Regression Results

```
=====
Dep. Variable:          Calories (kcal)    R-squared (uncentered):          0.992
Model:                  OLS               Adj. R-squared (uncentered):      0.992
Method:                 Least Squares     F-statistic:                     6705.
Date:                  Sun, 18 Jul 2021   Prob (F-statistic):              4.09e-167
Time:                  12:20:20          Log-Likelihood:                  -733.73
No. Observations:      162              AIC:                             1473.
Df Residuals:          159              BIC:                             1483.
Df Model:              3
Covariance Type:       nonrobust
=====
```

```
=====
              coef      std err          t      P>|t|      [0.025      0.975]
-----
Protein (g)      4.2439      0.115     36.856      0.000       4.017      4.471
Fat (g)          9.0681      0.125     72.306      0.000       8.820      9.316
Carbohydrates (g) 4.0474      0.059     68.580      0.000       3.931      4.164
=====
```

```
=====
Omnibus:          332.081    Durbin-Watson:           0.841
Prob(Omnibus):    0.000     Jarque-Bera (JB):        120630.217
Skew:             11.066     Prob(JB):                0.00
Kurtosis:         134.838     Cond. No.                2.57
=====
```


PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

Đề Tài Trực Quan Hóa Dữ Liệu

Nhóm 6 - Lớp Trực Quan Hóa Dữ Liệu 18_21

PHÂN TÍCH DINH DƯỠNG THỰC PHẨM VIỆT NAM

- > Tổng Quan
- > Tiền Xử Lý
- > Phân Tích Dữ Liệu
- > **Tổng kết**

Tổng Kết

Yếu tố dinh dưỡng là một vấn đề quan trọng, cần được lưu tâm hơn trong đời sống, các bữa ăn cần được cân bằng, đa dạng vì mỗi món ăn đều có dinh dưỡng riêng khác nhau.

Qua dự án lần này, nhóm đã ôn tập lại được các kiến thức đã học: trực quan hóa bằng các thư viện, phần mềm, sử dụng thống kê để phân tích dữ liệu, đưa ra nhận xét, kể lại câu chuyện từ các biểu đồ.

Thanks for watching