

Who-is-who of food.docx

by Turnitin LLC

Submission date: 24-May-2024 02:03AM (UTC+0300)

Submission ID: 2386729685

File name: AgADxBcAAvFbgFI (140.56K)

Word count: 731

Character count: 3772

Who-is-who of food

Question one

a) To explain the underlying association between brand and category, statistical frequency distribution tables and statistical bar charts. The frequency distribution table have provided count of each sandwich category within each brand. The numbers obtained from the frequency table, indicate the frequency of each category within each brand. The rightmost column shows the total number of products for each brand while the bottom row shows number of products for each category across brands.

The analysis shows that products in brand A are distributed among the different categories. The frequency table also highlights the brands with limited categories (Brand B, Brand D and Brand F). Dominant categories in some of the brands are also noted (Brand E and Brand L). Sparse categories include Ham and Tuna, offered by few brands (Brand A and Brand G for Ham and Tuna, respectively), and Veggie, offered only by Brands G and H, with Brand G having more variety.

Table 1.0

Frequency distribution between Brand and Category

The FREQ Procedure

Frequency

Table of Brand by Category									
Brand	Category								
	Beef	Chicken	Fish	Frozen	Ham	Tuna	Turkey	Veggie	Total
A	5	4	1	0	1	1	3	0	15
B	0	2	0	0	0	0	0	0	2
C	1	3	1	0	0	0	1	0	6
D	1	1	0	0	0	0	0	0	2
E	0	7	1	0	0	0	0	0	8
F	0	1	0	0	0	0	0	0	1
G	0	2	0	0	1	1	2	2	8
H	0	3	0	0	0	0	0	0	3
I	0	0	2	0	0	0	0	0	2
J	0	0	0	3	0	0	0	0	3
K	0	0	0	4	0	0	0	0	4
L	0	0	0	11	0	0	0	0	11
Total	7	23	5	18	2	2	6	2	65

Chi square value of <.0001 indicated that there is statistical difference between sandwich brand and category

Table 2.0

Statistics from the frequency distribution table

1

Statistics for Table of Brand by Category

Statistic	DF	Value	Prob
WARNING: 99% of the cells have expected counts less than 5. Chi-Square may not be a valid test.			
Chi-Square	77	139.1016	<.0001
Likelihood Ratio Chi-Square	77	125.0995	0.0004
Mantel-Haenszel Chi-Square	1	3.7673	0.0523
Phi Coefficient		1.4629	
Contingency Coefficient		0.8255	
Cramer's V		0.5529	

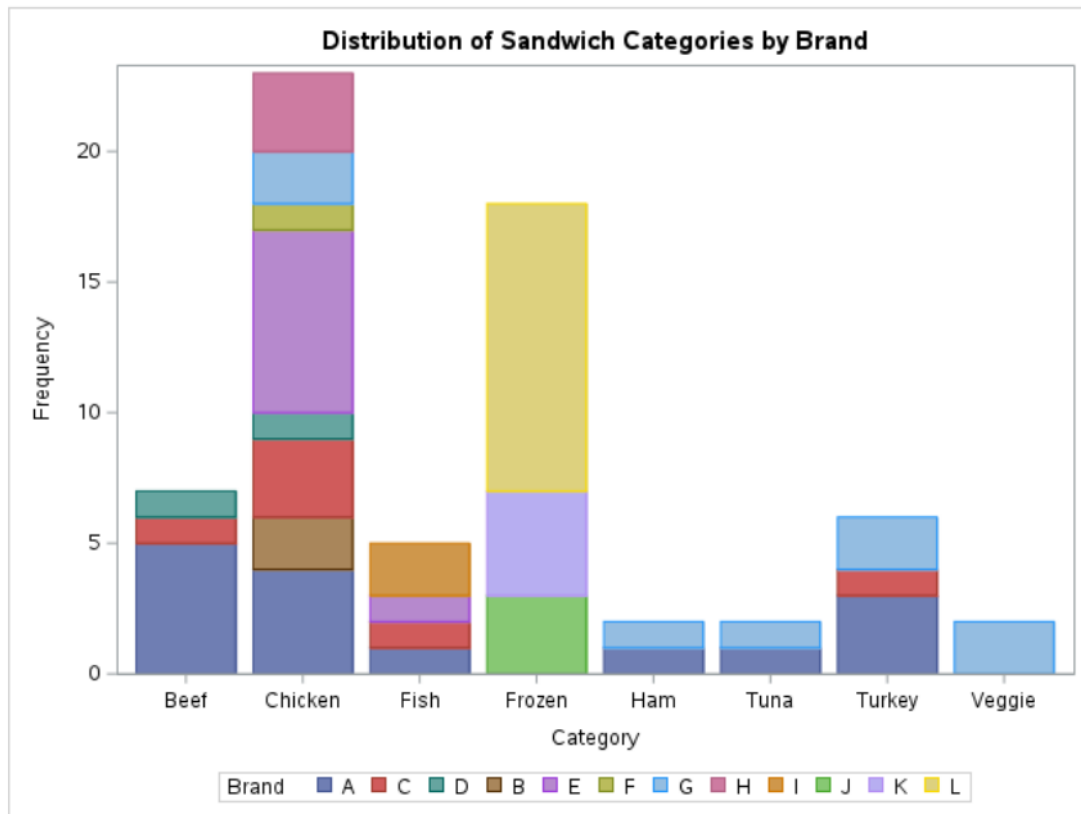
Sample Size = 65

A stacked bar chart was generated to investigate the association between brand and category.

Stacked bar chart is more advance in data visualization than a simple bar chart. The use of a staked bar chart was necessitated by the structure of the variable brand (multiple brands).

Figure 1.0

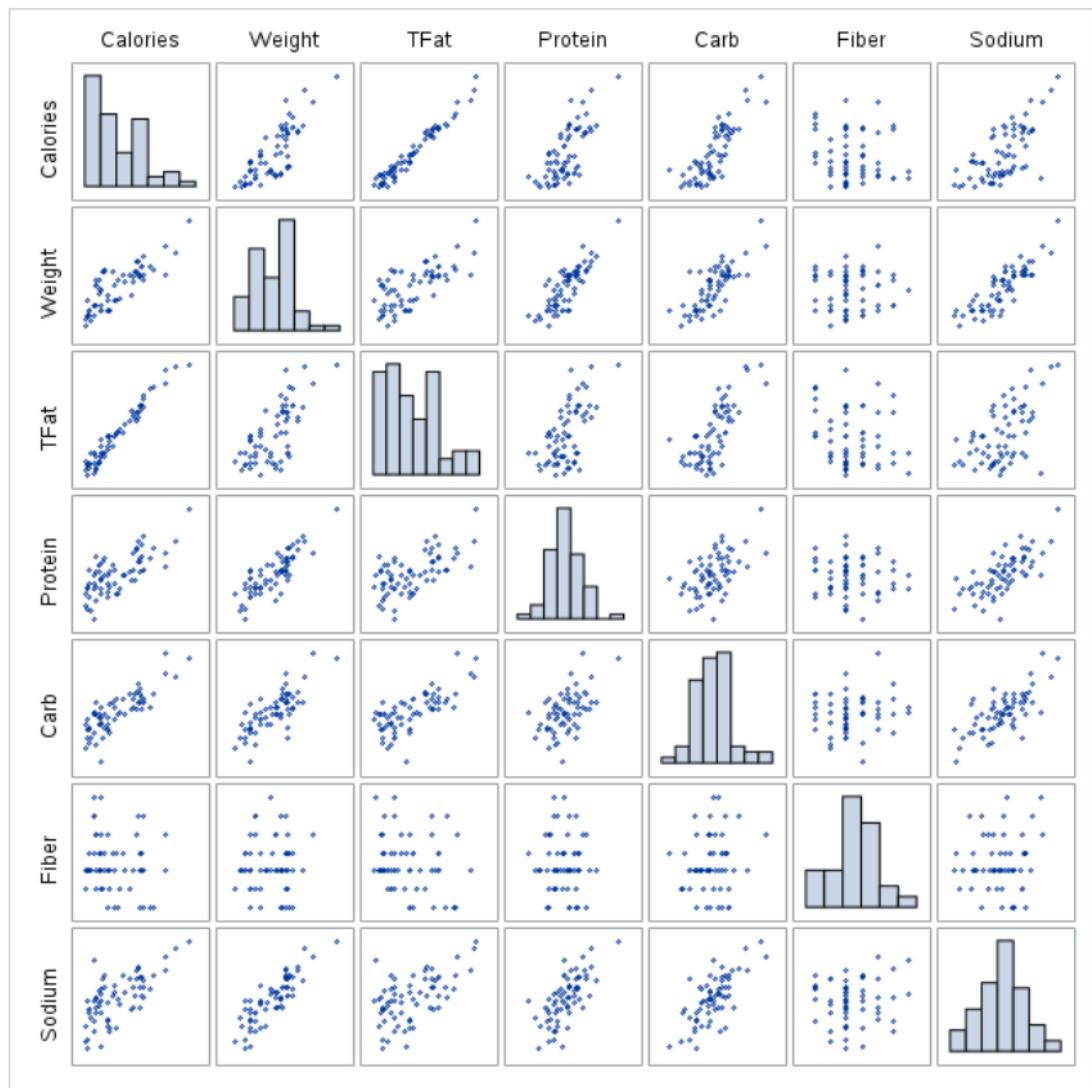
Distribution of Sandwich Categories by Brand



a) To investigate the relationship between variable, a statistical regression model was adopted and fitted for the variables of interest. Additionally, correlation analysis also depicts the existing relationship between the variables. The use of statistical charts offers a visual representation of the relationship between the variables. Visual present data in graphical form making it easier for the consumers to comprehend. The existing relationship among the variables can easily be identified.

Figure 2.0:

Scatter plot matrix

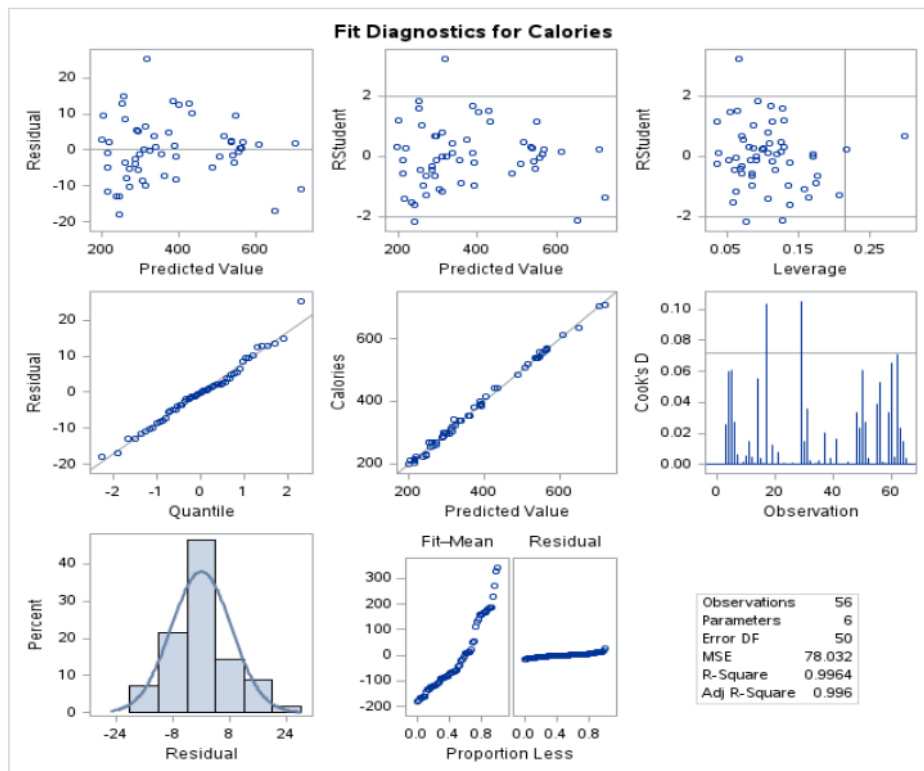


The figure above shows pairwise relationships between nutritional variables (Total Fat, Protein, Carbs, Fiber, Sodium) and physical variables (Calories, Weight). Positive correlations are evident between Calories, Weight, and several nutritional variables. Histograms on the diagonal provide the distribution of each variable. The relationships help identify which nutritional factors most strongly influence calories and weight.

Regression analysis is one of the methods used to investigate variable relationship in a data. R squared score measures the goodness of the model fitted. To investigate the exist relationship between nutritional variables physical variables, a regression model was fit on calories where the physical variables were treated as the response variable and the nutritional variables acted as he explanatory variables. The model on calories had an R squared score of 99.64%, suggesting that about 99.64% of variability in calories could be explained by the nutritional variables used.

Figure 2.0

Diagnostic fit for calories



References

Bhalla, D. (2020). *SAS: PROC FREQ with Examples*. ListenData.

<https://www.listendata.com/2016/02/sas-proc-freq-tutorial.html>

Frequency Distribution Node. (2021).

<https://support.sas.com/documentation/onlinedoc/dfdmstudio/2.5/dmpdmsug/Content/df>

[U_PFProf_FreqDist.html](#)

Gunst, R. F., & Mason, R. L. (2018). *Regression analysis and its application: a data-oriented approach*. CRC Press.

In, J., & Lee, S. (2017). Statistical data presentation. *Korean journal of anesthesiology*, 70(3), 267.

SAS OnDemand for academics. (2018). SAS. https://www.sas.com/el_gr/software/on-demand-for-academics.html

Who-is-who of food.docx

ORIGINALITY REPORT

8%

SIMILARITY INDEX

8%

INTERNET SOURCES

6%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to University of the Western Cape

Student Paper

5%

2

hdl.handle.net

Internet Source

3%

Exclude quotes On

Exclude bibliography On

Exclude matches Off