FIT5125 Assignment 3 Week 8

Telling a data story

1. Research Question

What is the correlation between visa status and student nutrition?

2. Descriptive Metrics of Choice

Median values of nutritional intake

The best average of our dataset, as the variance is high

Standard deviation of nutritional intake

Measures how wide the variance is in the dataset

Results

Median

Visa vs Nutrition - Median	Dataset Size	Portion size (g/ml)	Energy (kJ)	Carbohydrates	Protein	Fat	Satd FA	Cholesterol	Added sugar	Total sugars	Dietary fibre (g)	Starch	Caffeine	Total trans fatty acids (g)	Polyunsaturated fatty acids (g)	Monounsaturated fatty acids (g)
Overall	600	1534.35	5706.18	154.935	64.225	47.155	16.19	223.255	10.86	41.765	12.315	108.145	27.605	668.115	7.04	18.18
Citizen	58	1612.35	4731.035	119.69	59.99	44.78	14.68	236.68	7.09	30.45	10.88	95.735	49.545	631.93	7.125	17.83
Student Visa	521	1533.37	5824.71	159.96	64.53	46.71	16.06	222.31	11.22	42.56	12.33	109.73	26.63	667.21	7.04	18.03

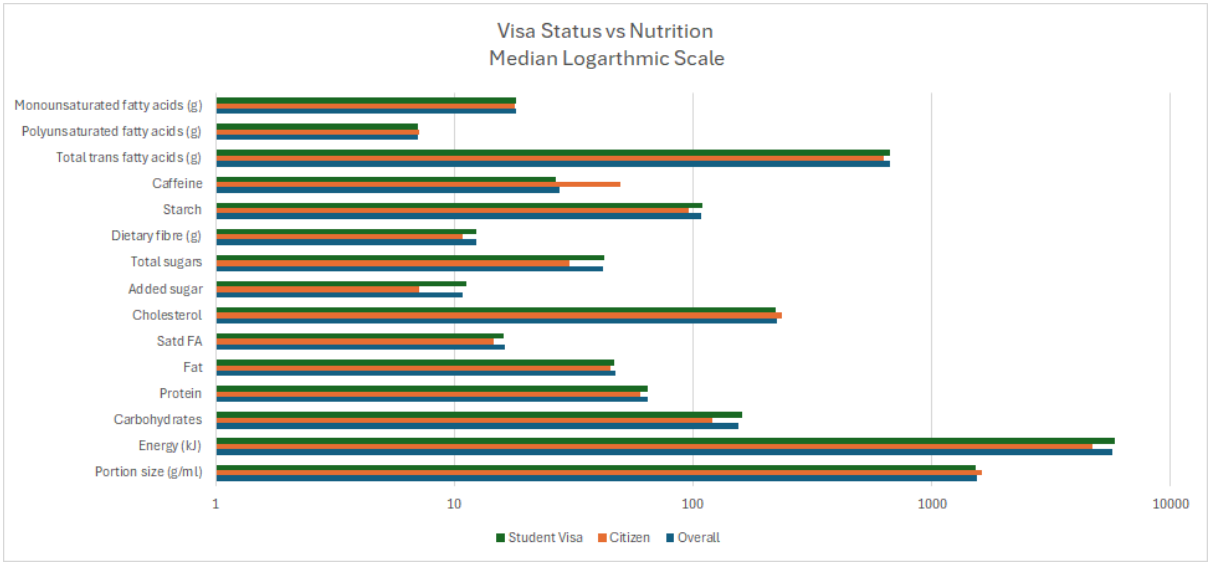
Standard deviation

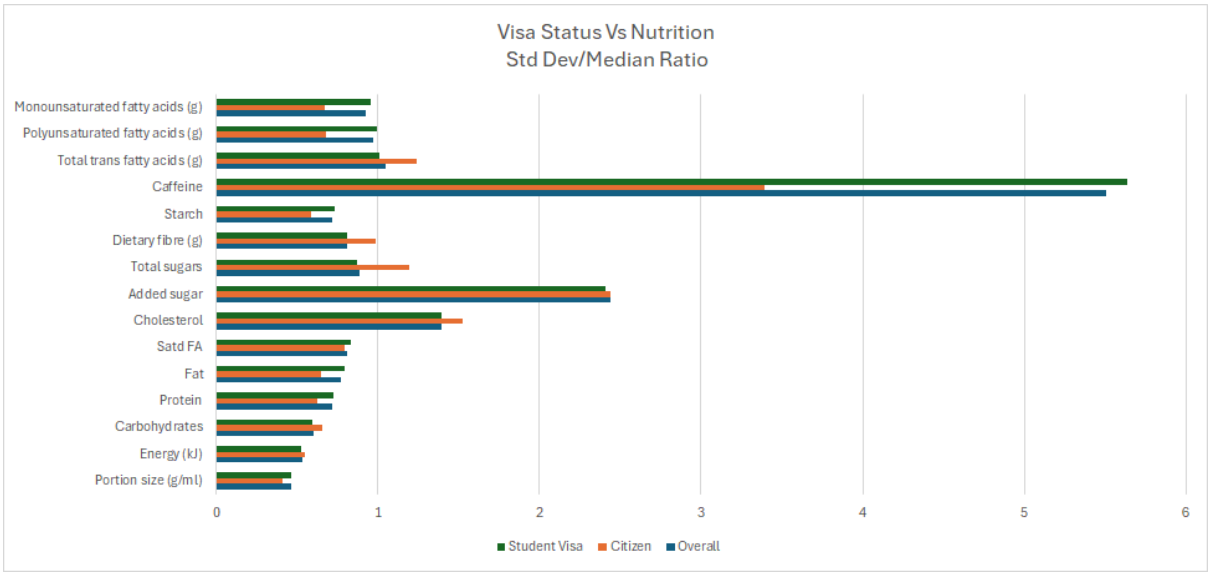
Visa vs Nutrition - StdDev	Dataset Size	Portion size (g/ml)	Energy (kJ)	Carbohydrates	Protein	Fat	Satd FA	Cholesterol	Added sugar	Total sugars	Dietary fibre (g)	Starch	Caffeine	Total trans fatty acids (g)	Polyunsaturated fatty acids (g)	Monounsaturated fatty acids (g)
Overall	600	704.8861276	3020.9636	93.54550946	45.933	36.195	13.087	311.933626	26.45657447	37.00294808	9.960991511	77.7891	152.085	702.3215343	6.809434957	16.75032635
Citizen	58	660.3951682	2575.6078	78.39142849	37.532	29.094	11.672	360.835719	17.27878708	36.27200197	10.7417013	56.3369	167.981	784.6734053	4.844639578	12.01827435
Student Visa	521	714.9477296	3067.624	95.28599006	46.761	37.038	13.315	309.099328	27.02505233	36.92430027	9.938523412	80.3424	150.178	672.448401	6.972707961	17.26680284

Standard deviation/Median%

StdDev/Median%	Dataset Size	Portion size (g/ml)	Energy (kJ)	Carbohydrates	Protein	Fat	Satd FA	Cholesterol	Added sugar	Total sugars	Dietary fibre (g)	Starch	Caffeine	Total trans fatty acids (g)	Polyunsaturated fatty acids (g)	Monounsaturated fatty acids (g)
Overall	600	45.94%	52.94%	60.38%	71.52%	76.76%	80.83%	139.72%	243.61%	88.60%	80.89%	71.93%	550.93%	105.12%	96.72%	92.14%
Citizen	58	40.96%	54.44%	65.50%	62.56%	64.97%	79.51%	152.46%	243.71%	119.12%	98.73%	58.85%	339.05%	124.17%	67.99%	67.40%
Student Visa	521	46.63%	52.67%	59.57%	72.46%	79.29%	82.91%	139.04%	240.86%	86.76%	80.60%	73.22%	563.94%	100.79%	99.04%	95.77%

3. Visualisation





4. Narrative Description of Findings

From the data gathered, it clearly shows that students who are citizens eat considerably less calories than international students on a student visa, at a median of 1093 kJ (19%) less than international students. Yet, student citizens actually eat slightly larger portions at 78.78 g/mL (5.15%) more than international students, and somehow also eat more cholesterol at 14.37 more. Interestingly, student citizens also consume vastly more caffeine at 49.545 vs 26.63 (86% more).

While not part of the original research question, it is difficult to ignore the vastly larger numbers of international students participating in the survey than citizens.

The survey had a very large variance in data collected, so the top and bottom 5% of the data was filtered out to remove the most extreme outliers. Even after filtering, the standard deviations are frequently more than half the mean, and caffeine intake in particular has an extremely high standard deviation:median ratio of up to 563.94%, indicating the population is heavily right-skewed on many of these metrics.

The most extreme examples before filtering were someone eating only 148.99 kJ on the low end, and another eating 89836 kJ on the high end - concerning, but likely and hopefully simply fake.