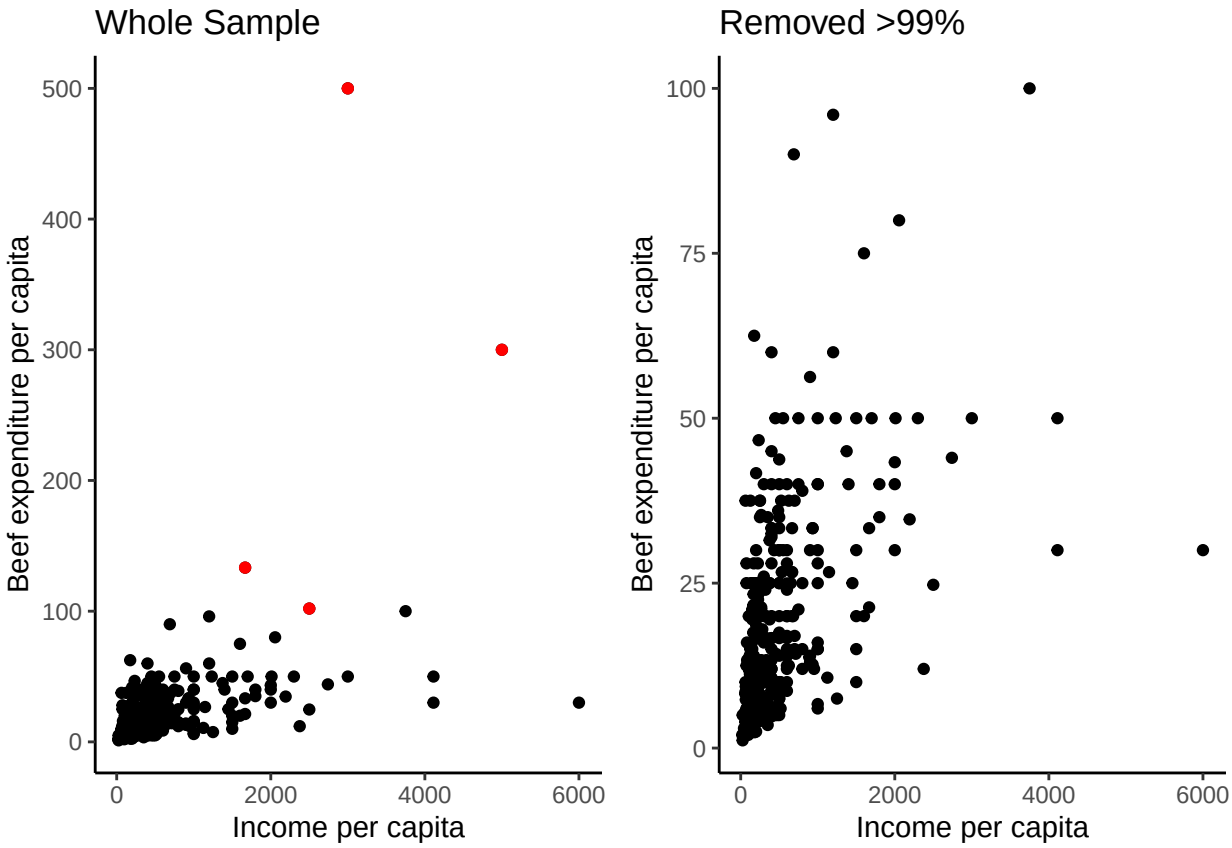
"Beef Survey Data Analysis"

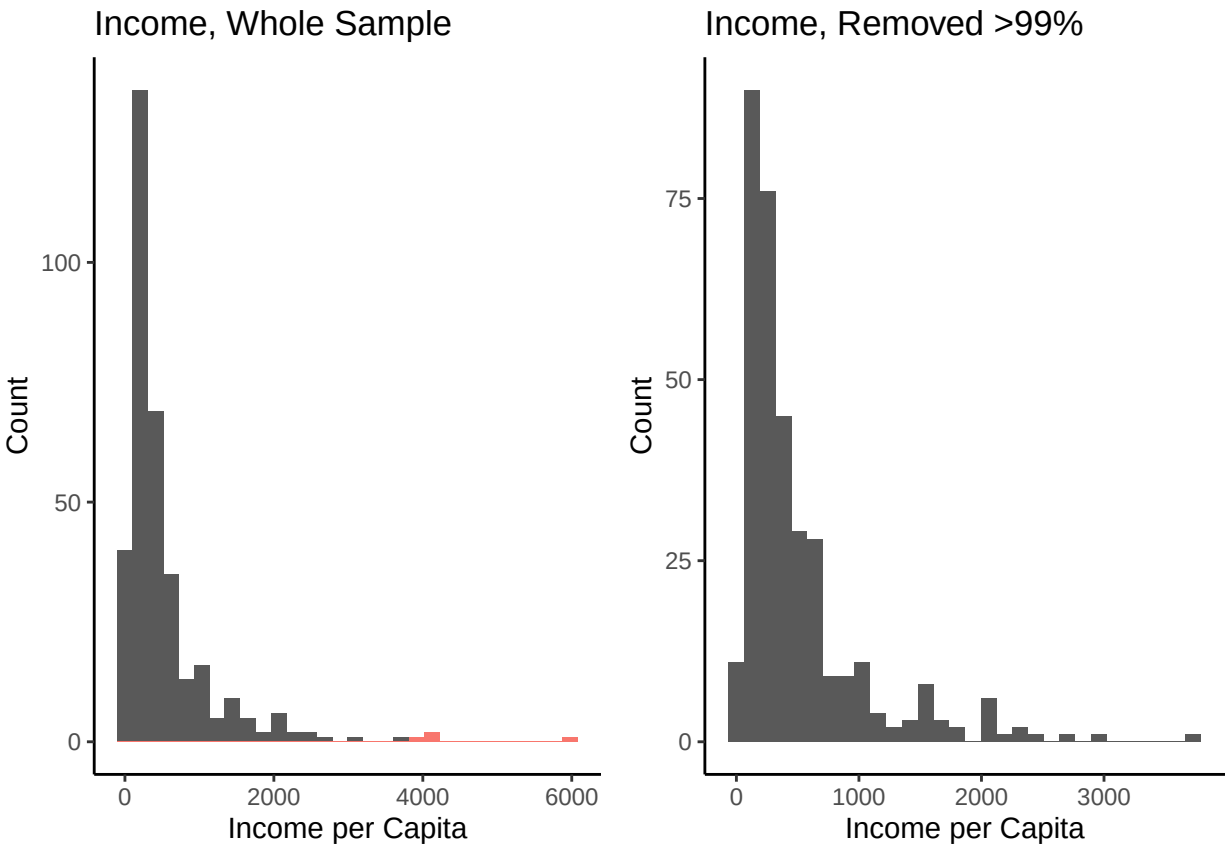
FRE518 (January 2022)

1 Data Cleaning and Descriptives

- Tabulated every single variable and recoded those do not match the data dictionary
- Converted categorical variables to factor variables for use in analysis later on
- Removed observations if beef expenditure $> 99^{th}$ percentile.



- Removed observations if income is $> 99^{th}$ percentile



- Removed observations if WTP amount is $> 99^{th}$ percentile

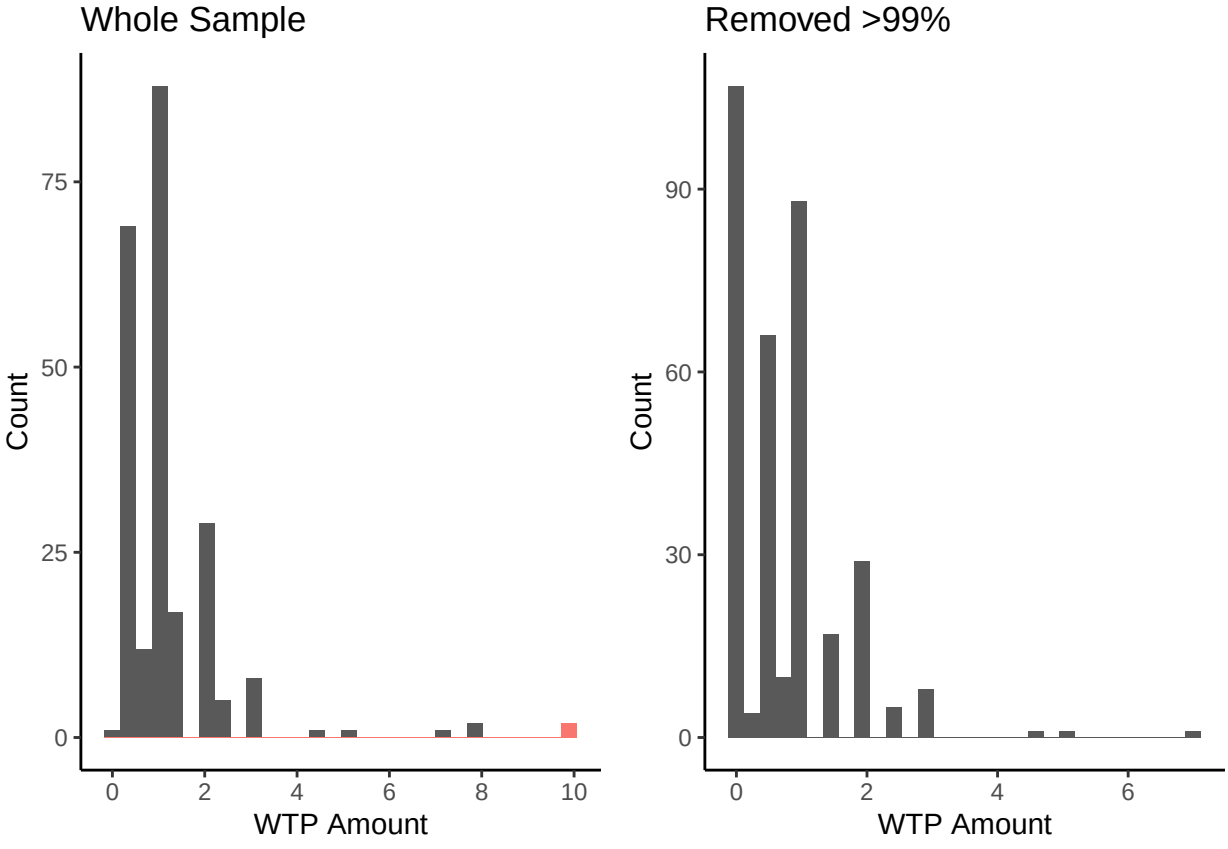


Table 1: Descriptive Statistics of Categorical Variables (Whole Sample)

		N	%
City	Cali	160	46.8
	Popayan	101	29.5
	Mercaderes	39	11.4
	Bordo	42	12.3
Gen	Female	104	30.4
	Male	237	69.3
Study	No education	3	0.9
	Incomplete primary	23	6.7
	Full primary	34	9.9
	Incomplete secondary	43	12.6
	Full secondary	116	33.9
	Technical	49	14.3
	Incomplete university	27	7.9
	University	38	11.1
	Postgraduate	8	2.3
Strata	Strata 1 (lowest)	102	29.8
	Strata 2	87	25.4
	Strata 3	104	30.4
	Strata 4	27	7.9
	Strata 5	17	5.0
	Strata 6 (highest)	5	1.5
Min12	No children under 12	195	57.0
	With children under 12	147	43.0
Preference1	Beef	105	30.7
	Chicken	132	38.6
	Fish	51	14.9
	Pork	43	12.6
	Other	1	0.3
Knowledge	No knowledge	187	54.7
	Very little knowledge	52	15.2
	Some knowledge	58	17.0
	Good knowledge	25	7.3
	Much knowledge	17	5.0
PayMore	No	106	31.0
	Yes	233	68.1

What matters is the selection criteria/sampling method. We don't really need an equal split if that isn't representative of society

Most are not knowledgeable about environmental impacts of beef production, yet they are willing to pay more

Table 2: Descriptive Statisics of Contiuous Variables (Whole Sample)

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Age	62	1	44.1	15.0	18.0	43.0	81.0
HhSize	12	0	3.7	1.7	1.0	4.0	13.0
BeefConsumption	9	1	2.8	1.6	0.0	2.0	7.0
ChickenConsumption	19	5	3.0	1.7	0.0	3.0	7.0
PorkConsumption	15	22	1.3	1.1	0.0	1.0	7.0
FishConsumption	16	29	1.2	1.3	0.0	1.0	7.0
BeefExpenditurePerCapita	106	4	19.2	15.6	1.2	14.0	100.0
BeefComsumptionPerCapita	16	1	3.1	2.6	0.0	2.0	16.0
WtpAmount	19	1	0.8	0.9	0.0	0.5	7.0
IncomePerCapita	113	1	495.5	534.3	20.0	300.0	3750.0

- Frequency of protein consumption in the household (times per week)
- BeefExpenditurePerCapita = Per capita expenditure of beef consumption ($)
- BeefConsumptionPerCapita = Per capita consumption of beef (lb)

Table 3: Descriptive Statistics of Categorical Variables, Stratified by Gender

		Female (N=104)		Male (N=237)	
		N	Pct.	N	Pct.
City	Cali	54	51.9	106	44.7
	Popayan	32	30.8	69	29.1
	Mercaderes	10	9.6	29	12.2
	Bordo	8	7.7	33	13.9
Study	No education	1	1.0	2	0.8
	Incomplete primary	9	8.7	14	5.9
	Full primary	6	5.8	28	11.8
	Incomplete secondary	11	10.6	32	13.5
	Full secondary	30	28.8	85	35.9
	Technical	14	13.5	35	14.8
	Incomplete university	16	15.4	11	4.6
	University	15	14.4	23	9.7
	Postgraduate	1	1.0	7	3.0
Strata	Strata 1 (lowest)	24	23.1	77	32.5
	Strata 2	22	21.2	65	27.4
	Strata 3	41	39.4	63	26.6
	Strata 4	10	9.6	17	7.2
	Strata 5	5	4.8	12	5.1
	Strata 6 (highest)	2	1.9	3	1.3
Min12	No children under 12	69	66.3	126	53.2
	With children under 12	35	33.7	111	46.8
Preference1	Beef	32	30.8	73	30.8
	Chicken	37	35.6	94	39.7
	Fish	17	16.3	34	14.3
	Pork	14	13.5	29	12.2
	Other	1	1.0	0	0.0
Knowledge	No knowledge	39	37.5	147	62.0
	Very little knowledge	17	16.3	35	14.8
	Some knowledge	30	28.8	28	11.8
	Good knowledge	7	6.7	18	7.6
	Much knowledge	11	10.6	6	2.5
PayMore	No	31	29.8	74	31.2
	Yes	73	70.2	160	67.5

Table 4: Descriptive Statistics of Contiuous Variables, Stratified by Gender

	Female (N=104)		Male (N=237)	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	47.4	15.6	42.7	14.5
HhSize	3.6	1.3	3.8	1.8
BeefConsumption	2.9	1.7	2.7	1.6
ChickenConsumption	3.1	1.6	3.0	1.7
PorkConsumption	1.4	1.3	1.2	1.1
FishConsumption	1.2	1.0	1.3	1.4
WtpAmount	0.7	0.7	0.8	0.9
IncomePerCapita	587.6	578.9	457.1	510.0
BeefExpenditurePerCapita	22.7	14.8	17.7	15.7
BeefComsumptionPerCapita	3.7	2.8	2.9	2.5

Table 5: Descriptive Statistics of Cateogrical Variables, Stratified by City

		Cali (N=160)		Popayan (N=101)		Mercaderes (N=39)		Bordo (N=42)	
		N	Pct.	N	Pct.	N	Pct.	N	Pct.
Gen	Female	54	33.8	32	31.7	10	25.6	8	19.0
	Male	106	66.2	69	68.3	29	74.4	33	78.6
Study	No education	0	0.0	0	0.0	3	7.7	0	0.0
	Incomplete primary	3	1.9	6	5.9	7	17.9	7	16.7
	Full primary	10	6.2	13	12.9	4	10.3	7	16.7
	Incomplete secondary	22	13.8	12	11.9	6	15.4	3	7.1
	Full secondary	45	28.1	40	39.6	11	28.2	20	47.6
	Technical	25	15.6	15	14.9	6	15.4	3	7.1
	Incomplete university	21	13.1	4	4.0	1	2.6	1	2.4
	University	27	16.9	9	8.9	1	2.6	1	2.4
	Postgraduate	6	3.8	2	2.0	0	0.0	0	0.0
Strata	Strata 1 (lowest)	12	7.5	32	31.7	32	82.1	26	61.9
	Strata 2	31	19.4	37	36.6	5	12.8	14	33.3
	Strata 3	78	48.8	22	21.8	2	5.1	2	4.8
	Strata 4	19	11.9	8	7.9	0	0.0	0	0.0
	Strata 5	17	10.6	0	0.0	0	0.0	0	0.0
	Strata 6 (highest)	3	1.9	2	2.0	0	0.0	0	0.0
Min12	No children under 12	111	69.4	47	46.5	21	53.8	16	38.1
	With children under 12	49	30.6	54	53.5	18	46.2	26	61.9
Preference1	Beef	49	30.6	30	29.7	16	41.0	10	23.8
	Chicken	61	38.1	41	40.6	10	25.6	20	47.6
	Fish	16	10.0	18	17.8	10	25.6	7	16.7
	Pork	25	15.6	11	10.9	3	7.7	4	9.5
	Other	1	0.6	0	0.0	0	0.0	0	0.0
Knowledge	No knowledge	71	44.4	63	62.4	25	64.1	28	66.7
	Very little knowledge	25	15.6	16	15.8	5	12.8	6	14.3
	Some knowledge	35	21.9	12	11.9	5	12.8	6	14.3
	Good knowledge	14	8.8	5	5.0	4	10.3	2	4.8
	Much knowledge	12	7.5	5	5.0	0	0.0	0	0.0
PayMore	No	48	30.0	24	23.8	14	35.9	20	47.6
	Yes	110	68.8	77	76.2	24	61.5	22	52.4

Table 6: Descriptive Statistics of Contiuous Variables, Stratified by City

	Cali (N=160)		Popayan (N=101)		Mercaderes (N=39)		Bordo (N=42)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age	46.5	15.9	42.2	13.6	42.4	14.7	40.6	13.9
HhSize	3.4	1.6	4.0	1.3	4.2	2.2	3.9	2.0
BeefConsumption	2.6	1.4	3.1	1.9	2.9	1.8	2.7	1.5
ChickenConsumption	3.0	1.5	3.4	1.9	2.3	1.6	2.9	1.7
PorkConsumption	1.4	1.0	1.0	1.1	1.0	1.2	1.4	1.5
FishConsumption	1.3	1.4	1.1	1.3	1.0	1.0	1.3	1.3
WtpAmount	0.8	1.0	0.9	0.8	0.7	0.7	0.6	0.7
IncomePerCapita	729.7	637.3	322.6	358.9	217.0	184.2	269.7	188.3
BeefExpenditurePerCapita	24.7	18.9	16.0	11.1	11.0	7.5	14.7	9.6
BeefComsumptionPerCapita	3.6	2.8	2.9	2.4	2.2	2.4	2.8	2.1

The absolute value of WTP for beef produced in a more environmentally-friendly manner is similar across cities, but quite different if compared relative to income per capita and relative to beef expenditure per capita. For example 0.8/24.7 = 3% in Cali compared to 0.7/11 = 6% in Mercaderes

Table 7: Descriptive Statistics of Cateogrical Variables, Stratified by WTP more

		No (N=106)		Yes (N=233)	
		N	Pct.	N	Pct.
City	Cali	48	45.3	110	47.2
	Popayan	24	22.6	77	33.0
	Mercaderes	14	13.2	24	10.3
	Bordo	20	18.9	22	9.4
Gen	Female	31	29.2	73	31.3
	Male	74	69.8	160	68.7
Study	No education	3	2.8	0	0.0
	Incomplete primary	14	13.2	9	3.9
	Full primary	15	14.2	18	7.7
	Incomplete secondary	17	16.0	26	11.2
	Full secondary	32	30.2	84	36.1
	Technical	11	10.4	37	15.9
	Incomplete university	6	5.7	21	9.0
	University	6	5.7	31	13.3
	Postgraduate	1	0.9	7	3.0
Strata	Strata 1 (lowest)	44	41.5	56	24.0
	Strata 2	25	23.6	62	26.6
	Strata 3	27	25.5	77	33.0
	Strata 4	6	5.7	21	9.0
	Strata 5	2	1.9	14	6.0
	Strata 6 (highest)	2	1.9	3	1.3
Min12	No children under 12	63	59.4	130	55.8
	With children under 12	43	40.6	103	44.2
Preference1	Beef	32	30.2	73	31.3
	Chicken	39	36.8	92	39.5
	Fish	13	12.3	37	15.9
	Pork	18	17.0	25	10.7
	Other	0	0.0	1	0.4
Knowledge	No knowledge	77	72.6	109	46.8
	Very little knowledge	11	10.4	41	17.6
	Some knowledge	12	11.3	46	19.7
	Good knowledge	2	1.9	23	9.9
	Much knowledge	3	2.8	14	6.0

Table 8: Descriptive Statistics of Contiuous Variables, Stratified by WTP more

	No (N=106)		Yes (N=233)	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	46.4	15.6	43.0	14.7
HhSize	3.7	1.6	3.8	1.7
BeefConsumption	2.8	1.6	2.8	1.7
ChickenConsumption	2.9	1.6	3.1	1.7
PorkConsumption	1.2	1.2	1.3	1.1
FishConsumption	1.2	1.6	1.3	1.2
WtpAmount	0.0	0.0	1.2	0.8
IncomePerCapita	408.2	501.9	532.8	543.5
BeefExpenditurePerCapita	20.2	19.1	18.9	13.7
BeefComsumptionPerCapita	3.0	2.6	3.2	2.6

Those that are WTP more are slightly younger.

Table 9: Descriptive Statistics of Contiuous Variables, Stratified by Knowledge of Environmental Impact

	None (N=187)		Very Little (N=52)		Little (N=58)		Some (N=25)		A lot (N=17)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age	44.0	14.2	43.2	15.3	45.8	16.6	42.6	17.4	45.3	15.4
HhSize	3.9	1.7	3.4	1.4	3.6	1.6	4.1	2.5	3.1	1.1
BeefConsumption	2.8	1.7	3.1	1.7	2.6	1.4	2.8	1.6	2.7	1.8
ChickenConsumption	3.1	1.8	2.9	1.6	3.1	1.5	2.7	1.2	2.7	1.6
PorkConsumption	1.2	1.1	1.5	1.4	1.1	1.0	1.7	1.6	1.2	0.9
FishConsumption	1.2	1.5	1.1	1.0	1.4	1.1	1.2	1.4	1.2	0.6
WtpAmount	0.6	0.8	0.9	0.8	1.0	0.9	1.1	0.6	1.3	1.6
IncomePerCapita	357.2	369.8	599.5	656.1	703.9	697.0	666.8	591.7	697.5	565.2
BeefExpenditurePerCapita	16.2	13.7	22.5	17.3	23.3	18.2	19.2	12.5	28.4	17.8
BeefComsumptionPerCapita	2.7	2.2	3.6	3.1	3.5	2.6	3.8	2.9	4.4	3.1

Can see that average WTP amount goes up as knowledge of environmental impacts of beef production goes up

Table 10: Descriptive Statistics of Categorical Variables, Stratified by Knowledge of Environmental Impact

		None (N=187)		Very Little (N=52)		Little (N=58)		Some (N=25)		A lot (N=17)	
		N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
City	Cali	71	38.0	25	48.1	35	60.3	14	56.0	12	70.6
	Popayan	63	33.7	16	30.8	12	20.7	5	20.0	5	29.4
	Mercaderes	25	13.4	5	9.6	5	8.6	4	16.0	0	0.0
	Bordo	28	15.0	6	11.5	6	10.3	2	8.0	0	0.0
Gen	Female	39	20.9	17	32.7	30	51.7	7	28.0	11	64.7
	Male	147	78.6	35	67.3	28	48.3	18	72.0	6	35.3
Study	No education	3	1.6	0	0.0	0	0.0	0	0.0	0	0.0
	Incomplete primary	17	9.1	3	5.8	2	3.4	1	4.0	0	0.0
	Full primary	26	13.9	4	7.7	3	5.2	0	0.0	1	5.9
	Incomplete secondary	29	15.5	9	17.3	3	5.2	2	8.0	0	0.0
	Full secondary	67	35.8	20	38.5	16	27.6	6	24.0	6	35.3
	Technical	23	12.3	5	9.6	14	24.1	3	12.0	3	17.6
	Incomplete university	13	7.0	4	7.7	5	8.6	3	12.0	2	11.8
	University	6	3.2	7	13.5	13	22.4	8	32.0	3	17.6
	Postgraduate	3	1.6	0	0.0	1	1.7	2	8.0	2	11.8
Strata	Strata 1 (lowest)	69	36.9	13	25.0	12	20.7	5	20.0	2	11.8
	Strata 2	56	29.9	13	25.0	11	19.0	4	16.0	3	17.6
	Strata 3	46	24.6	17	32.7	24	41.4	9	36.0	7	41.2
	Strata 4	10	5.3	7	13.5	5	8.6	3	12.0	2	11.8
	Strata 5	5	2.7	1	1.9	5	8.6	2	8.0	3	17.6
	Strata 6 (highest)	1	0.5	1	1.9	1	1.7	2	8.0	0	0.0
Min12	No children under 12	105	56.1	30	57.7	33	56.9	16	64.0	9	52.9
	With children under 12	82	43.9	22	42.3	25	43.1	9	36.0	8	47.1
Preference1	Beef	53	28.3	17	32.7	19	32.8	8	32.0	7	41.2
	Chicken	72	38.5	23	44.2	21	36.2	8	32.0	7	41.2
	Fish	35	18.7	5	9.6	5	8.6	5	20.0	1	5.9
	Pork	19	10.2	5	9.6	13	22.4	4	16.0	2	11.8
	Other	0	0.0	1	1.9	0	0.0	0	0.0	0	0.0
PayMore	No	77	41.2	11	21.2	12	20.7	2	8.0	3	17.6
	Yes	109	58.3	41	78.8	46	79.3	23	92.0	14	82.4

education seems to be related to knowledge of environmental impacts of beef production

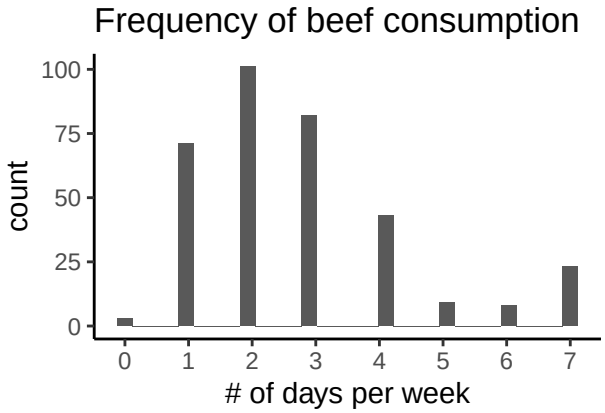
Table 11: Descriptive Statistics of Numerical Variables, Stratified by Protein Preference

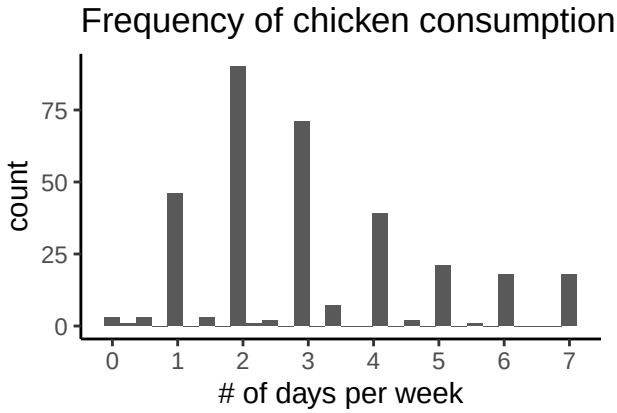
	Beef (N=105)		Chicken (N=132)		Fish (N=51)		Pork (N=43)		Other (N=1)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Age	44.1	15.5	43.1	14.4	44.2	16.6	44.2	13.0	64.0	
HhSize	3.7	1.5	3.7	1.6	4.0	2.2	3.5	1.7	4.0	
BeefConsumption	3.7	1.7	2.3	1.4	2.5	1.7	2.5	1.4	3.0	
ChickenConsumption	2.5	1.4	3.8	1.6	2.5	1.7	2.5	1.5	2.0	
PorkConsumption	1.2	0.8	1.0	1.0	1.6	1.4	1.6	1.6	2.0	
FishConsumption	1.0	1.1	1.0	1.0	1.5	1.4	2.1	1.9	1.0	
WtpAmount	0.7	0.7	0.8	0.8	0.8	0.8	0.9	1.4	0.7	
IncomePerCapita	545.0	599.7	452.3	477.1	415.0	463.5	568.4	531.9	2741.2	
BeefExpenditurePerCapita	23.5	19.1	15.6	11.4	17.0	12.9	21.8	18.1	44.0	
BeefComsumptionPerCapita	4.1	3.1	2.6	2.2	2.8	2.3	2.8	2.2	4.0	

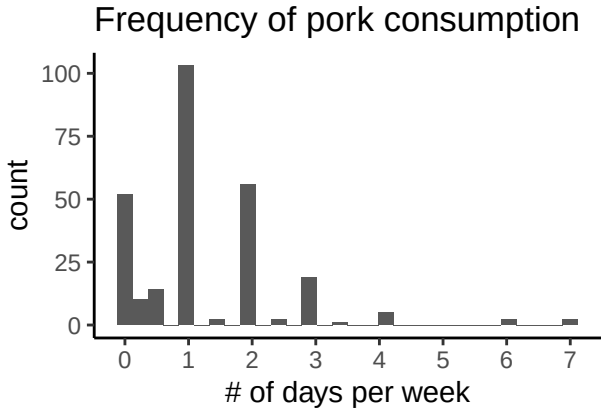
Table 12: Descriptive Statistics of Categorical Variables, Stratified by Protein Preference

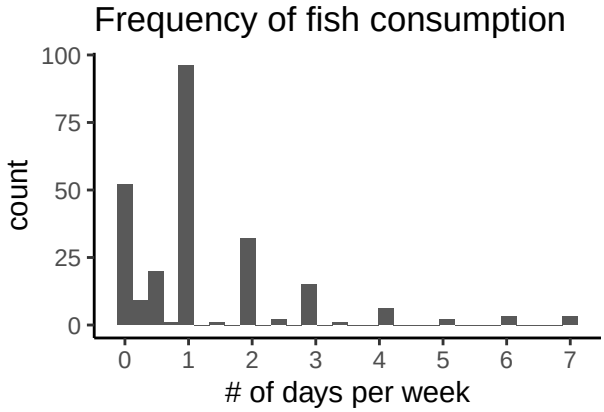
		Beef (N=105)		Chicken (N=132)		Fish (N=51)		Pork (N=43)		Other (N=1)	
		N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
City	Cali	49	46.7	61	46.2	16	31.4	25	58.1	1	100.0
	Popayan	30	28.6	41	31.1	18	35.3	11	25.6	0	0.0
	Mercaderes	16	15.2	10	7.6	10	19.6	3	7.0	0	0.0
	Bordo	10	9.5	20	15.2	7	13.7	4	9.3	0	0.0
Gen	Female	32	30.5	37	28.0	17	33.3	14	32.6	1	100.0
	Male	73	69.5	94	71.2	34	66.7	29	67.4	0	0.0
Study	No education	1	1.0	0	0.0	1	2.0	1	2.3	0	0.0
	Incomplete primary	7	6.7	8	6.1	5	9.8	3	7.0	0	0.0
	Full primary	14	13.3	13	9.8	5	9.8	2	4.7	0	0.0
	Incomplete secondary	11	10.5	12	9.1	8	15.7	5	11.6	1	100.0
	Full secondary	34	32.4	57	43.2	11	21.6	12	27.9	0	0.0
	Technical	11	10.5	15	11.4	10	19.6	12	27.9	0	0.0
	Incomplete university	11	10.5	8	6.1	4	7.8	4	9.3	0	0.0
	University	12	11.4	16	12.1	6	11.8	3	7.0	0	0.0
	Postgraduate	4	3.8	3	2.3	1	2.0	0	0.0	0	0.0
Strata	Strata 1 (lowest)	27	25.7	41	31.1	20	39.2	13	30.2	0	0.0
	Strata 2	26	24.8	35	26.5	14	27.5	11	25.6	0	0.0
	Strata 3	35	33.3	40	30.3	10	19.6	12	27.9	0	0.0
	Strata 4	12	11.4	9	6.8	3	5.9	3	7.0	0	0.0
	Strata 5	3	2.9	5	3.8	3	5.9	4	9.3	1	100.0
	Strata 6 (highest)	2	1.9	2	1.5	1	2.0	0	0.0	0	0.0
Min12	No children under 12	54	51.4	80	60.6	29	56.9	24	55.8	0	0.0
	With children under 12	51	48.6	52	39.4	22	43.1	19	44.2	1	100.0
Knowledge	No knowledge	53	50.5	72	54.5	35	68.6	19	44.2	0	0.0
	Very little knowledge	17	16.2	23	17.4	5	9.8	5	11.6	1	100.0
	Some knowledge	19	18.1	21	15.9	5	9.8	13	30.2	0	0.0
	Good knowledge	8	7.6	8	6.1	5	9.8	4	9.3	0	0.0
	Much knowledge	7	6.7	7	5.3	1	2.0	2	4.7	0	0.0
PayMore	No	32	30.5	39	29.5	13	25.5	18	41.9	0	0.0
	Yes	73	69.5	92	69.7	37	72.5	25	58.1	1	100.0

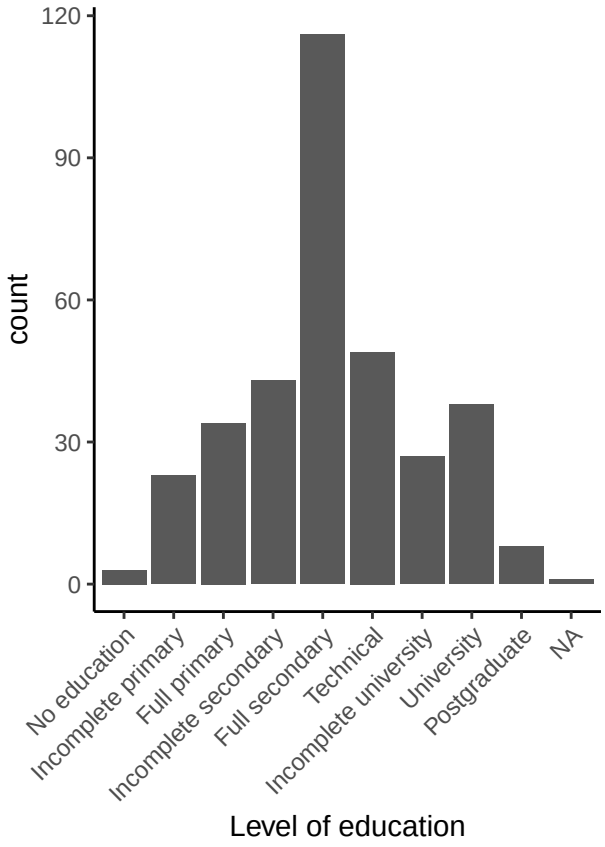
2 Some more visualizations

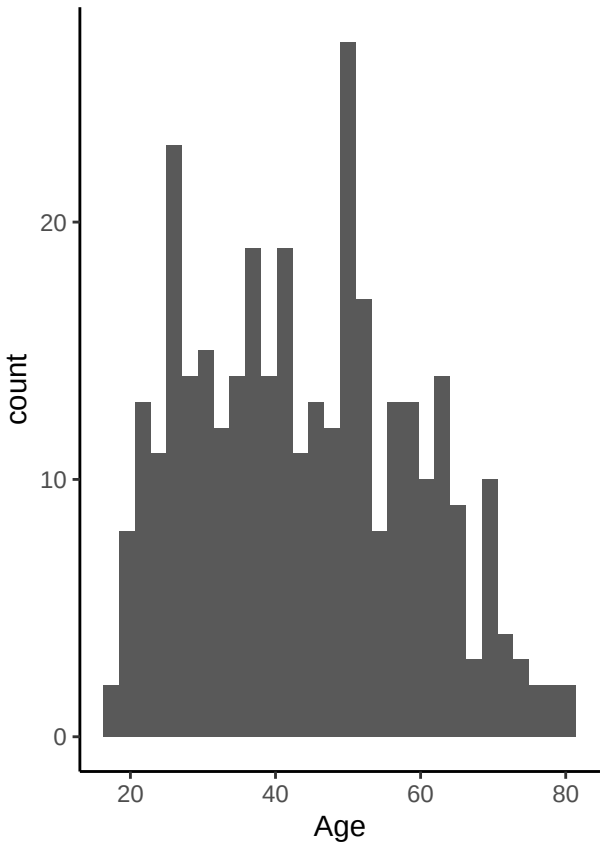









120
90
60
30
0
count
No education
Incomplete primary
Full primary
Incomplete secondary
Full secondary
Technical
Incomplete university
University
Postgraduate
NA
Level of education


20
10
0
count
20
40
60
80
Age

3 Statistical Tests

3.1 Correlations

- WTP Amount and Age: -0.1506 (p-val: 0.006)
- WTP Amount and Household Size: -0.0363 (p-val: 0.836)
- WTP Amount and Education: 0.3309 (p-val: 0)
- WTP Amount and Income per Capita: 0.2175 (p-val: 0)
- WTP Amount and Socioeconomic Strata: 0.2032 (p-val: 2×10^{-4})
- WTP Amount and Knowledge: 0.2695 (p-val: 0)
- WTP Amount and Beef Expenditure per Capita: 0.0616 (p-val: 0)
- WTP Amount and Beef Consumption: -0.0237 (p-val: 0.0167)
- WTP Amount and Chicken Consumption: 0.0353 (p-val: 0.0167)
- WTP Amount and Fish Consumption: 0.0989 (p-val: 0.3995)
- WTP Amount and Pork Consumption: 0.0671 (p-val: 0.6683)

most strongly correlated: education, income, socioeconomic strata, knowledge -- all makes sense

negatively correlated with age - can make sense as younger people may care more about the environment

odd to see that there is a positive correlation with beef expenditure per capita but negative correlation with frequency of beef consumption

3.2 T-tests

- Had to convert relevant variables to numeric, because the `t.test()` function does not accept factor variables
- Due to the way the questions were asked, we can take a look if their answers "make sense". The table below shows the frequency of beef (vertical) and chicken (horizontal) consumption for those who answered beef is their preferred protein. The values in the upper right part of the matrix indicates people who eat more chicken than beef, even though their protein preference is beef. A t-test shows that the difference in means between these two groups is statistically significant.

```
# observe chicken and beef consumption are different for those who prefer beef
table(beef$BeefConsumption[beef$Preference1 == "Beef"], beef$ChickenConsumptionRound[beef$Preference1 == "Beef"])
```

freq of chicken consumption

freq of beef consumption

	0	1	2	3	4	5	6	7
1	0	2	2	1	0	0	0	0
2	0	6	6	5	3	1	0	0
3	1	8	11	3	2	1	1	0
4	0	1	12	4	3	0	0	0
5	0	1	4	1	0	1	0	0
6	0	2	2	0	0	0	0	0
7	0	2	2	1	2	2	0	3

The highlighted values are the number of people who indicated they prefer beef as protein but actually consume chicken more frequently than beef. For example, the highlighted 2 means that there are 2 people in the data where they indicated beef as their preferred protein but eat chicken twice per week and eat beef only once per week.

- A number of t-tests are run to compare means of different groups in the data.

```
# t.test if beef and chicken consumption are different for those who prefer beef
# summary(beef$BeefConsumption[beef$Preference1 == "Beef"])
# summary(beef$ChickenConsumption[beef$Preference1 == "Beef"])
t.test(beef$BeefConsumption[beef$Preference1 == "Beef"], beef$ChickenConsumption[beef$Preference1 == "Beef"])
```

```
##
##  Welch Two Sample t-test
##
## data:  beef$BeefConsumption[beef$Preference1 == "Beef"] and beef$ChickenConsumption[beef$Preference1 == "Beef"
## t = 5.5858, df = 195.71, p-value = 7.711e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.8051239 1.6839253
## sample estimates:
## mean of x mean of y
##  3.711538  2.467014
```

Average beef consumption if prefers beef = 3.7
Average chicken consumption if prefers chicken = 2.5

People are "consistent" in their answers that if they prefer beef they actually consume more beef (of course there are caveats in terms of availability and preferences)

```
# two sided t-test
# t.test if WTP is different for those who prefer beef than those who prefer other protein (Prefers Beef = 1)
t.test(beef$WtpAmount ~ beef$PreferBeef)
```

```
##
##  Welch Two Sample t-test
##
## data:  beef$WtpAmount by beef$PreferBeef
## t = 1.0618, df = 264.66, p-value = 0.2893
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.08422065  0.28137037
## sample estimates:
## mean in group 0 mean in group 1
##       0.8341518       0.7355769
```

Average WTP if male = 0.834
Average WTP if female = 0.736

p-value is not statistically different

```
# t.test if WTP is different for gender (Male = 1)
t.test(beef$WtpAmount ~ beef$GenNumeric)
```

```
##
##  Welch Two Sample t-test
##
## data:  beef$WtpAmount by beef$GenNumeric
## t = -1.4207, df = 270.59, p-value = 0.1566
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.30420008  0.04919363
## sample estimates:
## mean in group 0 mean in group 1
##       0.7033981       0.8309013
```

```
# t.test if WTP is different if hh has children (Yes = 1)
t.test(beef$WtpAmount ~ beef$ChildNumeric)
```

```
##
##  Welch Two Sample t-test
##
## data:  beef$WtpAmount by beef$ChildNumeric
## t = 0.66226, df = 327.56, p-value = 0.5083
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.1223508  0.2465350
## sample estimates:
## mean in group 0 mean in group 1
##       0.8246575       0.7625654
```

```
# t.test if WTP is different for those with income is above average (~1200usd) (Yes = 1)
beef$HighIncome <- ifelse(beef$IncomePerCapita >= 1200, 1, 0)
t.test(beef$WtpAmount ~ beef$HighIncome)
```

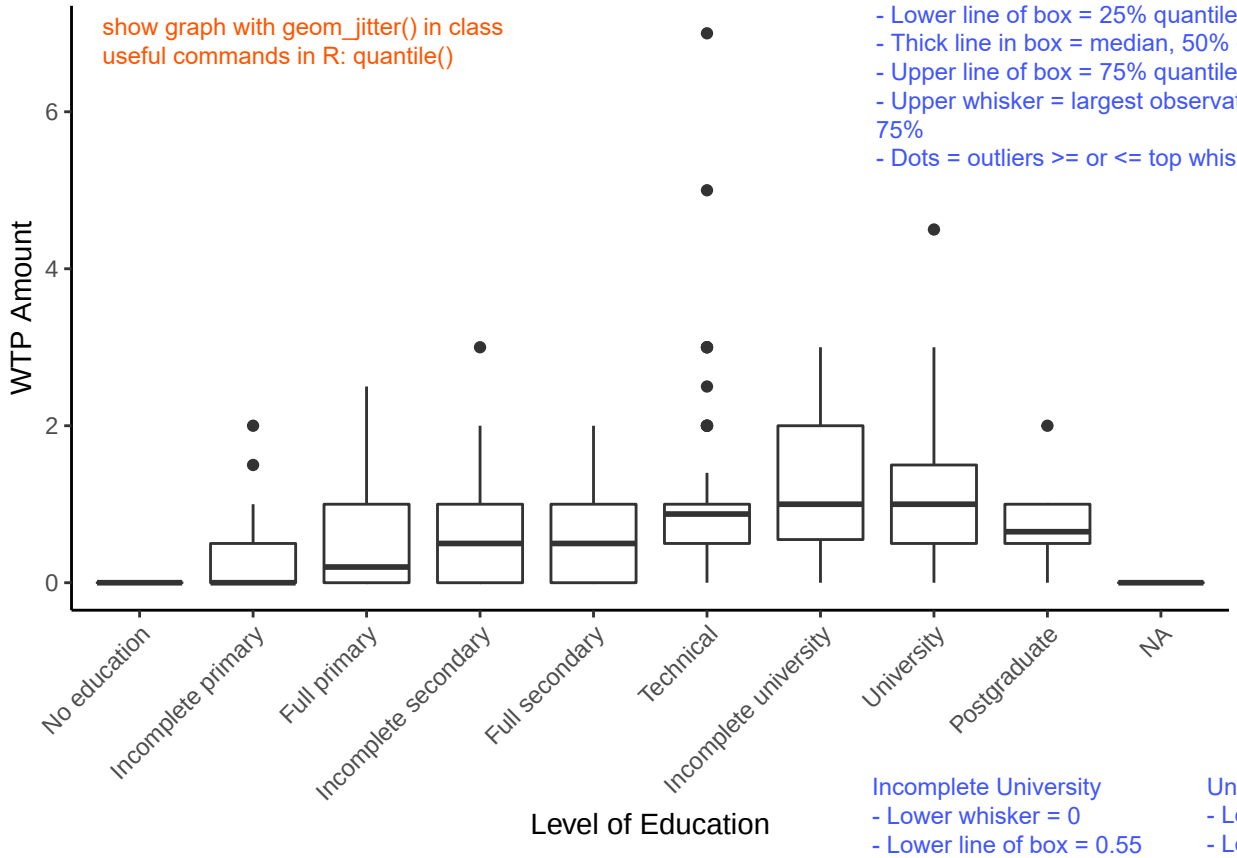
For this one, I looked up average income per month in Colombia and created a dummy variable = 1 if their income >= 1200.

```
##
##  Welch Two Sample t-test
##
## data:  beef$WtpAmount by beef$HighIncome
## t = -1.647, df = 32.234, p-value = 0.1093
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.89743385  0.09486811
## sample estimates:
## mean in group 0 mean in group 1
##        0.750330        1.151613
```

3.3 ANOVA

- IQR = interquartile range
- Lower whisker = smallest observation within 1.5*IQR below 25th percentile
- Lower line of box = 25% quantile
- Thick line in box = median, 50% quantile
- Upper line of box = 75% quantile
- Upper whisker = largest observation within 1.5*IQR above 75%
- Dots = outliers >= or <= top whisker + 1.5 * IQR

show graph with geom_jitter() in class
useful commands in R: quantile()



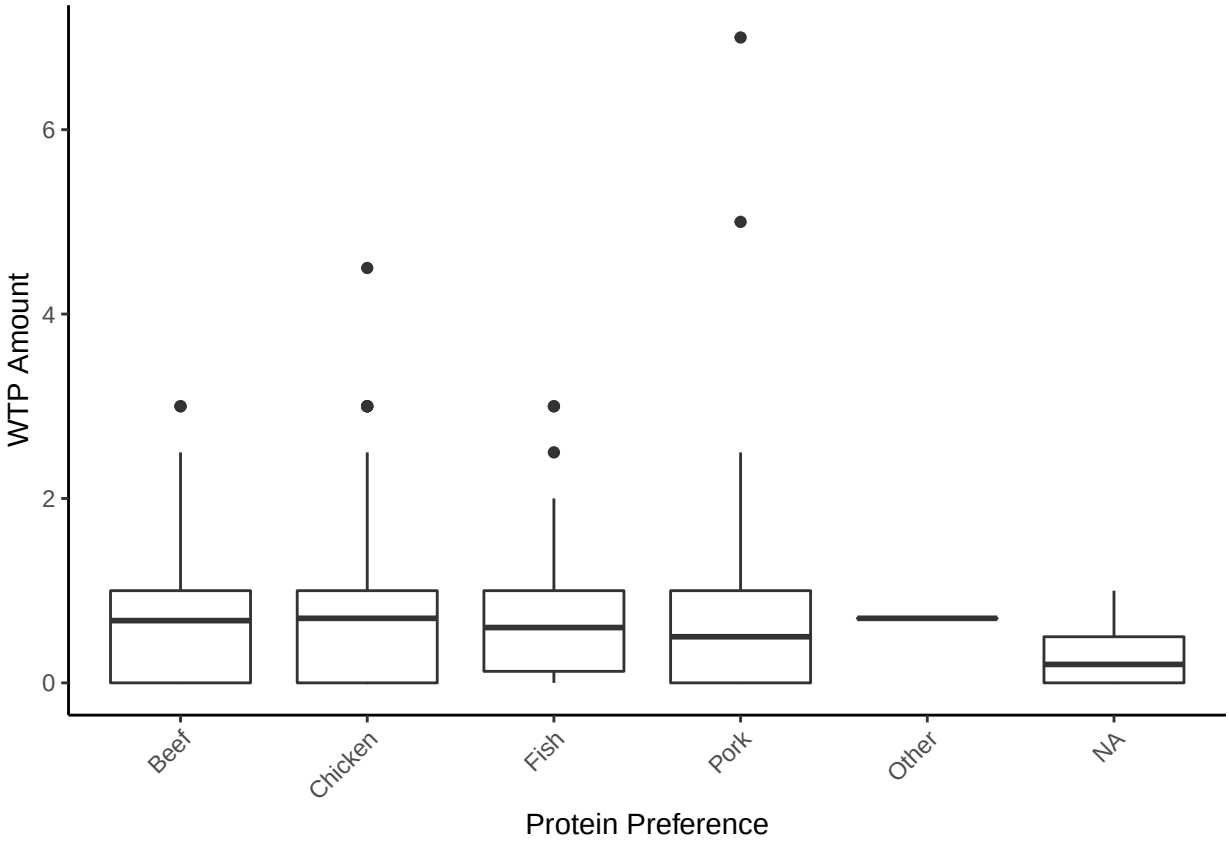
Incomplete University
- Lower whisker = 0
- Lower line of box = 0.55
- Median = 1
- Upper line of box = 2
- Upper whisker = 3
- IQR = 1.45

University
- Lower whisker = 0
- Lower line of box = 0.5
- Median = 1
- Upper line of box = 1.5
- Upper whisker = 3
- IQR = 1

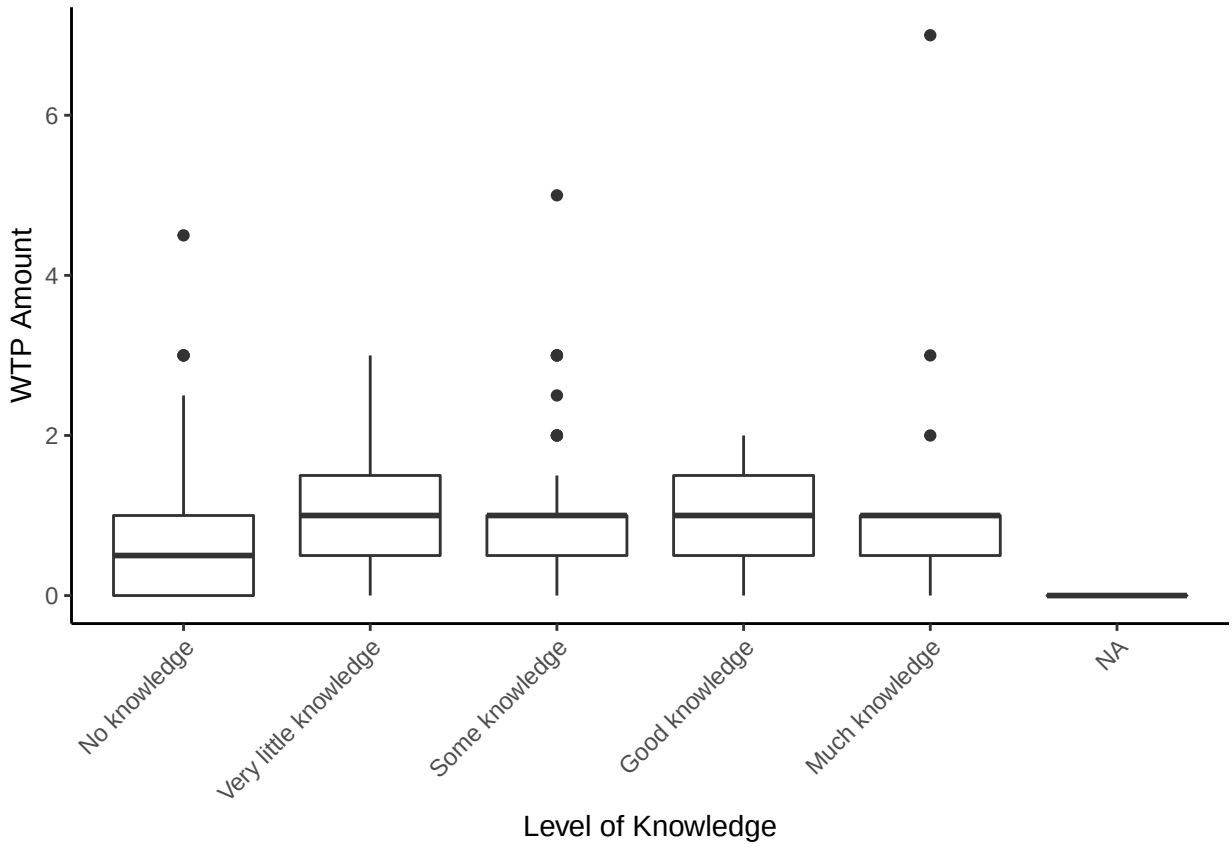
Technical
- Lower whisker = 0
- Lower line of box = 0.25
- Median =0.875
- Upper line of box = 1
'- Upper whisker = ?
- IQR = 0.5

```
## Analysis of Variance Table
##
## Response: WtpAmount
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Study       8  28.141  3.5176  5.1557 4.521e-06 ***
## Residuals 327 223.104  0.6823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##  Kruskal-Wallis rank sum test
##
## data:  WtpAmount by Study
## Kruskal-Wallis chi-squared = 40.569, df = 8, p-value = 2.509e-06

##                                                 diff       p adj
## Incomplete primary-No education             0.36956522 0.998353372
## Full primary-No education                   0.44218750 0.993555903
## Incomplete secondary-No education           0.54418605 0.973484534
## Full secondary-No education                 0.73318966 0.846588330
## Technical-No education                      1.10729167 0.374421550
## Incomplete university-No education          1.21923077 0.276189514
## University-No education                     1.18378378 0.294389136
## Postgraduate-No education                   0.78750000 0.894180688
## Full primary-Incomplete primary             0.07262228 0.999996644
## Incomplete secondary-Incomplete primary     0.17462083 0.996277066
## Full secondary-Incomplete primary           0.36362444 0.594331636
## Technical-Incomplete primary                0.73772645 0.014267765
## Incomplete university-Incomplete primary    0.84966555 0.011171932
## University-Incomplete primary               0.81421857 0.007355915
## Postgraduate-Incomplete primary             0.41793478 0.948745542
## Incomplete secondary-Full primary           0.10199855 0.999844732
```

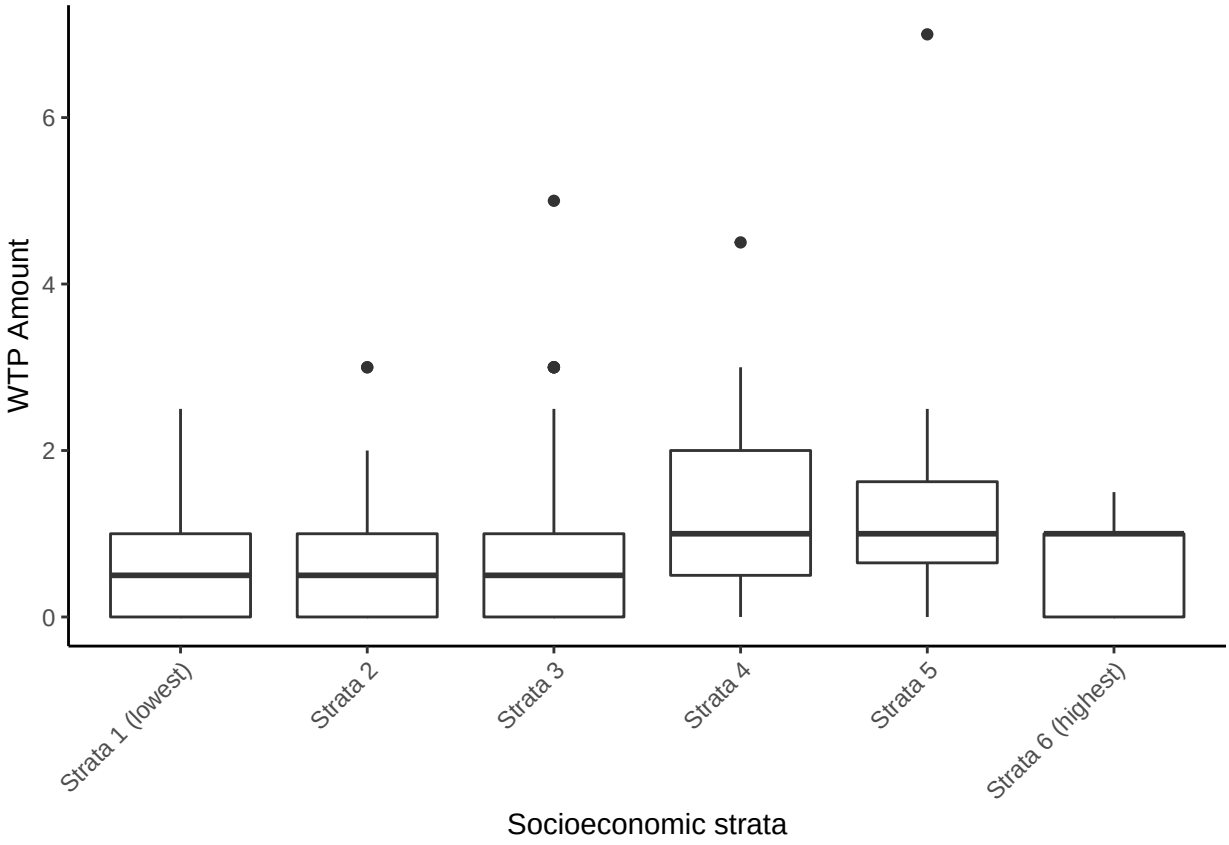
```
## Full secondary-Full primary                  0.29100216 0.705614696
## Technical-Full primary                       0.66510417 0.013962351
## Incomplete university-Full primary           0.77704327 0.012408084
## University-Full primary                      0.74159628 0.007180149
## Postgraduate-Full primary                    0.34531250 0.979607812
## Full secondary-Incomplete secondary          0.18900361 0.936140827
## Technical-Incomplete secondary               0.56310562 0.034558947
## Incomplete university-Incomplete secondary   0.67504472 0.030278998
## University-Incomplete secondary              0.63959774 0.017939761
## Postgraduate-Incomplete secondary            0.24331395 0.997677036
## Technical-Full secondary                     0.37410201 0.174730601
## Incomplete university-Full secondary         0.48604111 0.147686007
## University-Full secondary                    0.45059413 0.095085430
## Postgraduate-Full secondary                  0.05431034 0.999999966
## Incomplete university-Technical              0.11193910 0.999772359
## University-Technical                         0.07649212 0.999971621
## Postgraduate-Technical                      -0.31979167 0.984405869
## University-Incomplete university            -0.03544699 0.999999980
## Postgraduate-Incomplete university          -0.43173077 0.932987869
## Postgraduate-University                     -0.39628378 0.949277866
```




```
## Analysis of Variance Table
##
## Response: WtpAmount
##              Df Sum Sq Mean Sq F value Pr(>F)
## Preference1   4   0.78 0.19501  0.2541 0.9071
## Residuals   323 247.90 0.76748

##
##  Kruskal-Wallis rank sum test
##
## data:  WtpAmount by Preference1
## Kruskal-Wallis chi-squared = 2.2694, df = 4, p-value = 0.6863

##                       diff     p adj
## Chicken-Beef   0.106407810 0.8871742
## Fish-Beef      0.066423077 0.9921576
## Pork-Beef      0.115613553 0.9513731
## Other-Beef    -0.035576923 0.9999994
## Fish-Chicken  -0.039984733 0.9987497
## Pork-Chicken   0.009205743 0.9999972
## Other-Chicken -0.141984733 0.9998470
## Pork-Fish      0.049190476 0.9988587
## Other-Fish    -0.102000000 0.9999600
## Other-Pork    -0.151190476 0.9998097
```




```
## Analysis of Variance Table
##
## Response: WtpAmount
##            Df Sum Sq Mean Sq F value    Pr(>F)
## Knowledge   4  14.49  3.6226  5.0647 0.0005666 ***
## Residuals 331 236.75  0.7153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##  Kruskal-Wallis rank sum test
##
## data:  WtpAmount by Knowledge
## Kruskal-Wallis chi-squared = 25.841, df = 4, p-value = 3.407e-05

##                                          diff      p adj
## Very little knowledge-No knowledge   0.32240803 0.11061911
## Some knowledge-No knowledge          0.35052474 0.04874041
## Good knowledge-No knowledge          0.45786957 0.08444014
## Much knowledge-No knowledge          0.65645780 0.01999743
## Some knowledge-Very little knowledge 0.02811671 0.99979366
## Good knowledge-Very little knowledge 0.13546154 0.96503168
## Much knowledge-Very little knowledge 0.33404977 0.61922064
## Good knowledge-Some knowledge        0.10734483 0.98417057
## Much knowledge-Some knowledge        0.30593306 0.68420593
## Much knowledge-Good knowledge        0.19858824 0.94517496
```




```
## Analysis of Variance Table
##
## Response: WtpAmount
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Strata      5  18.217  3.6434  5.1613 0.0001417 ***
## Residuals 331 233.653  0.7059
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
##  Kruskal-Wallis rank sum test
##
## data:  WtpAmount by Strata
## Kruskal-Wallis chi-squared = 16.686, df = 5, p-value = 0.005135

##                                           diff       p adj
## Strata 2-Strata 1 (lowest)            0.13336816 0.889057336
## Strata 3-Strata 1 (lowest)            0.26989311 0.204215785
## Strata 4-Strata 1 (lowest)            0.66548822 0.004122159
## Strata 5-Strata 1 (lowest)            0.88724747 0.001499463
## Strata 6 (highest)-Strata 1 (lowest)  0.12474747 0.999523863
## Strata 3-Strata 2                     0.13652494 0.874608160
## Strata 4-Strata 2                     0.53212005 0.048846524
## Strata 5-Strata 2                     0.75387931 0.013641290
## Strata 6 (highest)-Strata 2          -0.00862069 0.999999999
## Strata 4-Strata 3                     0.39559511 0.250974950
## Strata 5-Strata 3                     0.61735437 0.071281319
## Strata 6 (highest)-Strata 3          -0.14514563 0.999001037
## Strata 5-Strata 4                     0.22175926 0.960490490
## Strata 6 (highest)-Strata 4          -0.54074074 0.772739352
## Strata 6 (highest)-Strata 5          -0.76250000 0.485760692
```

4 Key Insights

- After removing some outliers, only 342 observations were used in the analysis.
- 68% indicated they were willing to pay more for beef produced in an environmentally friendly manner, even though majority said they have no knowledge about the environmental impacts of beef production.
- The average amount people are willing to pay for beef produced in an environmentally-friendly manner is US$0.79 , which represents about 4% of their beef expenditure per capita.
- There is a significant and positive correlation between willingness to pay amount and education, income and knowledge of environmental impact of beef production. The positive correlation between WTP amount and beef expenditure but negative correlation with frequency of beef consumption is a bit puzzling.
- Based on ANOVA results, there are some differences in willingness to pay based on education levels, level of knowledge of environmental impact of beef production, socioeconomic strata, but not protein preference