Project Proposal

In this research project, considering that I do not have previous econometric research experience, but I would like to have some flexibility over the methodology, I will be choosing **the regular track** and completing this project **individually**.

The UBC food service has recently changed its service format at the cafeterias to an all-you-can-eat mode with a single payment. Its concern is whether such a shift has led to more waste than before. After reviewing the preliminary data provided by UBC food service, it is noticed that when comparing the dining data and the waste data, both the consumption and the waste during lunch hours and dinner hours seem to be relatively higher than those during breakfast hours, with the dinner hours being the highest in general. Therefore, this research project hopes to answer the following question: **within each day, across three different facilities, how does each dining hour contribute to the final waste, or more specifically, what proportion of waste does each dining hour contribute relative to their consumption?** The result of this research can allow all the facilities to adjust the size of each cuisine serving according to the dining hours to reduce food waste.

The datasets used in this project include two main parts: the food consumption data, and the food waste data. The datasets are first categorized by their facilities, then by the dining hours. Therefore, the key variables will involve **the consumption data and the waste data for each dining hour within each day**. The consumption data can be used to determine the proportion of each dining hour's consumption within a day; while the waste data can be added together to find the total waste each day. With these two general data, a model that relates the waste and the proportion of consumption each day can be built for analysis.

Though preliminary, the following table includes a general glance at the key statistics involved:

		Consumption			Waste		
		Mean	SD	Max	Mean	SD	Max
Gather @ Place Vanier	Breakfast	736.6	342.95	1372.0	28.73	41.69	219.54
	Lunch	1004.8	328.47	1513.0	49.75	57.65	362.98

	Dinner	907	347.17	1468.0	288.5	202.94	1170.7
Feast @ Totem Park	Breakfast	693.2	386.63	1797.0	45.97	91.90	909.71
	Lunch	1019	248.69	1806.0	112.58	305.65	2980.00
	Dinner	1110	388.39	1884.0	235.31	91.99	672.53
Open Kitchen	Breakfast	753.4	279.68	1298.0	33.35	20.57	110.70
	Lunch	1292.1	329.59	1794.0	86.28	69.26	600.99
	Dinner	1456	342.01	2093.0	259.77	699.22	7322.18

Table 1: Preliminary Summary Statistics (Can be modified in deeper analysis)

Since the purpose of the project is to highlight the temporal pattern of consumption and waste within each day, therefore, if the waste of the ith day in the data is labelled to be W_i, and the breakfast, lunch, and dinner consumption of the ith day are labelled C_{bi}, C_{li}, C_{di} respectively, then, it is possible to set up a linear regression equation as the following:

$$W_i = \beta_0 + \beta_b C_{bi} + \beta_l C_{li} + \beta_d C_{di} + \epsilon_i$$

In this equation, β_0 indicates an unavoidable waste each day, β_b represents the factor of how much breakfast consumption contributes to the total waste, β_l represents the factor of how much lunch consumption contributes to the total waste, and β_d represents the factor of how much dinner consumption contributes to the total waste. The error, in this case, represents general fluctuation across days. The key assumptions made in this model include these points:

1. It is assumed that the source of waste comes primarily from the food taken by the students.
2. It is assumed that the consumption of each dining hour across days is relatively stable and does not suffer from dramatic increases and drops.

Note that the variables within this regression model may be subject to change, as standardization is not performed yet as a preliminary analysis.