What is a sandwich?

A data analysis in $\mathbb{R}^{43}$

Vincent Macri David White Matthew Stuart

Group D

June 6, 2018

Abstract

In this paper we set out to examine the sandwich views of students and teachers at William Lyon Mackenzie C.I.

Our study consisted of a survey conducted in person with 140 participants. We asked participants basic demographic information about where they fit into the William Lyon Mackenzie Collegiate Institute community, and examined correlations between sandwich views and demographics.

We propose the creation of a two-dimensional sandwich alignment chart, inspired by two-dimensional political axes. The sandwich alignment chart has a dimension for "sandwich purity", and "sandwich orthodoxy".

We hypothesize that:

1. Sandwich purity and sandwich orthodoxy will be positively correlated.
2. Students in the MaCS or Gifted program are more likely to have a low orthodoxy score.

Contents

Part I

Introduction

Chapter 1

Terminology and notation

As this paper is dealing with sandwich mathematics, which is a developing field, we will define the terms we have coined in this chapter.

Food question A food question is any question asked about sandwiches or their ingredients. We asked 43 food questions as a part of this study. Participants answered each question on a 0 to 10 scale. We purposely started the scale at 0 so as to make 5 exactly in the middle of the range of possible responses.

Demographic question A demographic question is any of the non-food questions asked in our survey. These questions asked about participants demographics, in order to examine any correlations between demographics and sandwich views.

Chapter 2

Methodology

2.1 Survey type

We conducted our survey as mix of a stratified, voluntary, and random sample. We surveyed roughly 10% of the William Lyon Mackenzie population.

Unfortunately, we failed to collect a perfectly stratified sample. However, an analysis of the data shows that to be inconsequential.

We made manual corrections to categorical to correct for similar, blank, and inappropriate responses. As part of this, we grouped ethnicity into the following 11 categories: Caucasian, Chinese, East Asian, Filipino, Jewish, Korean, Middle Eastern, Mixed, Other, South Asian, and Vietnamese.

Chapter 3

Metrics

3.1 Axes metrics

We have two metrics used in our calculations: purity and orthodoxy.

Purity is how pure a respondent's definition of a sandwich is. The less things a respondent considers a sandwich, the greater their purity score will be. Similarly, the more things a respondent considers a sandwich, the lower their purity score will be.

Orthodoxy is a measure of how much a respondent differs from the mean set of responses. The less a respondent's answers differ from the mean set of answers, the greater their orthodoxy score will be. Similarly, the more a respondent's answers differ from the mean set of answers, the lower their orthodoxy score will be.

Both purity and orthodoxy are bound in the range $[-1, 1]$.

We describe the general scoring system in subsection 3.1.1 on the following page. This scoring system is used to calculate the purity metric described in subsection 3.1.2 on the next page and the orthodoxy metric described in subsection 3.1.3 on page 6.

3.1.1 Scoring

While participants answered each food question on a 0 to 10 scale, it is more convenient to perform calculations using a -5 to 5 scale. We converted responses from the 0 to 10 scale to the -5 to 5 scale by subtracting each response from 5.

Formally, for each response to a food question, we calculate the score for the response by the passing the response through the sandwich spectrum function, defined in 1.

Definition 1 (Sandwich spectrum function)**.** The sandwich spectrum function is defined as:

$$s : \{x \in \mathbb{R} \mid 0 \leq x \leq 10\} \rightarrow \{x \in \mathbb{R} \mid -5 \leq x \leq 5\} \text{ by } s(x) = 5 - x \tag{3.1}$$

Due to the format of our survey, all responses are integers, and are mapped to another integer by the sandwich spectrum function. Although, in principle, the sandwich spectrum function works for real numbers as well.

We can create a table for s:

x	0	1	2	3	4	5	6	7	8	9	10
$s(x)$	5	4	3	2	1	0	-1	-2	-3	-4	-5

One will note that this gives responses that were originally high a lower score. This is intentional. Subsection 3.1.2 will show it to be useful for calculating the purity metric, and subsection 3.1.3 on the next page will show it to be irrelevant for calculating the orthodoxy metric.

3.1.2 Purity

The purity score for a respondent is defined as the sum of a respondent's scores divided by the maximum possible score.

The maximum score for a question is 5, and there are 43 food questions. This means that the maximum possible score is $43 \times 5 = 215$.

Definition 2 (Sadwich purity function)**.** For a given response with a set of 43 food answers, A, we define the sandwich purity function as:

$$p : \mathbb{R}^{43} \rightarrow \mathbb{R} \text{ by } p(A) = \frac{\sum_{i=1}^{43} A_i}{215} \tag{3.2}$$

This definition illustrates why we subtract each response from 5 to get the score. The sandwich spectrum function will assign higher purity values for lower responses. Since a lower response to a question implies a more pure definition of a sandwich, 1 on the preceding page is a valid metric.

3.1.3 Orthodoxy

To calculate orthodoxy for each respondent, we take the score for each question as a dimension of a vector, which creates a vector in 43-dimensional Euclidean space.

We also calculate the mean response for each question, and create an additional $\mathbb{R}^{43}$ vector from that. This vector is referred to as the mean vector, and denoted as $\vec{m}$.

The orthodoxy score for a respondent is defined as the cosine similarity between the respondent's $\mathbb{R}^{43}$ vector and $\vec{m}$.
Definition 3 (Mean vector)**.** To calculate the value of the mean vector, $\vec{m}$:

Let $A \in \mathbb{R}^{43}$ be the set of response vectors. Then, $n(A)$ is the cardinality of the set A.

Then:

$$\vec{m} = \frac{\sum_{i=1}^{n(A)} A_i}{n(A)} \tag{3.3}$$

Definition 4 (Sandwich orthodoxy function)**.** To calculate the orthodoxy score for a respondent:

Let $\vec{r}$ be a vector in $\mathbb{R}^{43}$ defined as having its each of its components equal to the score for each food question.

$$o : \mathbb{R}^{43} \to \mathbb{R} \text{ by } o(\vec{r}) = \frac{\vec{r} \cdot \vec{m}}{\|\vec{r}\|\|\vec{m}\|} \tag{3.4}$$

Since we are calculating orthodoxy as the cosine of the angle between two vectors, it is useful to have some vector components be negative, as that allows respondents to have a negative orthodoxy score if they answer opposite to the mean response. It does not matter what direction the vectors are in, as we are only looking at the angle between them. This means that the sandwich spectrum function could have been defined as subtracting 5 instead of subtracting from 5 for the purposes of the orthodoxy function. Since both definitions would have worked for orthodoxy, we stick with 1 on page 5 for the sake of consistency with the purity metric.

Part II

Results Summary

Part III

Analysis

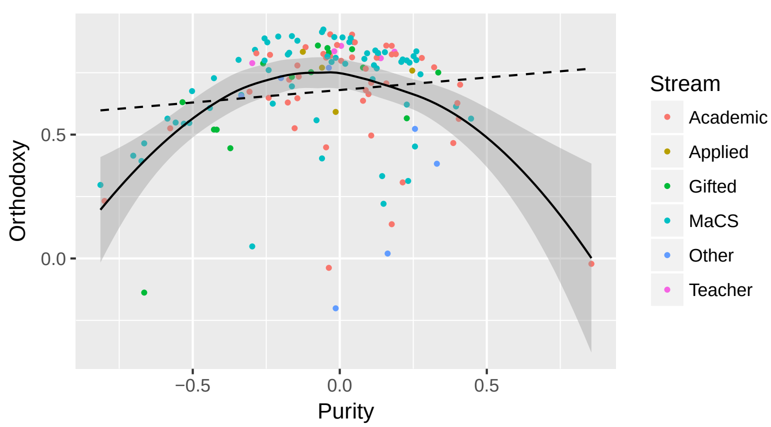
Chapter 4

Outliers

We begin our analysis by plotting the purity vs orthodoxy and performing both a linear and locally weighted regression analysis in figure 4.1.

4.1 With outliers

Figure 4.1: A plot of purity vs orthodoxy with data points coloured based on stream.



Here, the solid curve is the result of the locally weighted regression analysis, and the shaded area is the 95% confidence interval for that analysis. The dashed line is the result of the linear regression.

The linear regression has a positive slope, which proves the first hypothesis that purity and orthodoxy are positively correlated.

4.2 Identifying outliers

From the locally weighted regression analysis, it is clear that some points are extremely deviate. We next create residual plots using both the linear model and the locally weighted regression model to search for outliers in figure 4.2 and figure 4.3 on the following page.

Figure 4.2: A residual plot using the linear regression analysis method.

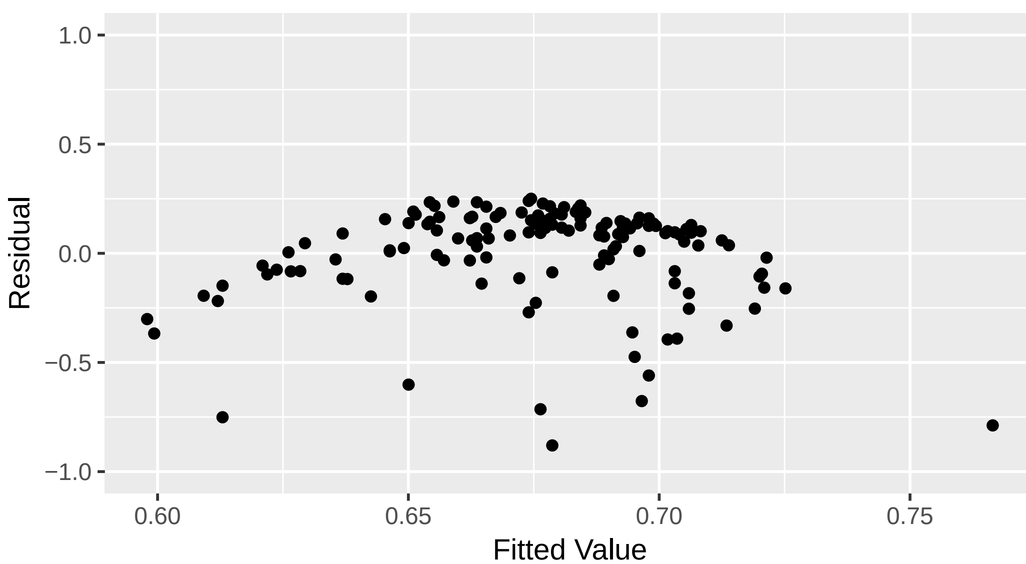
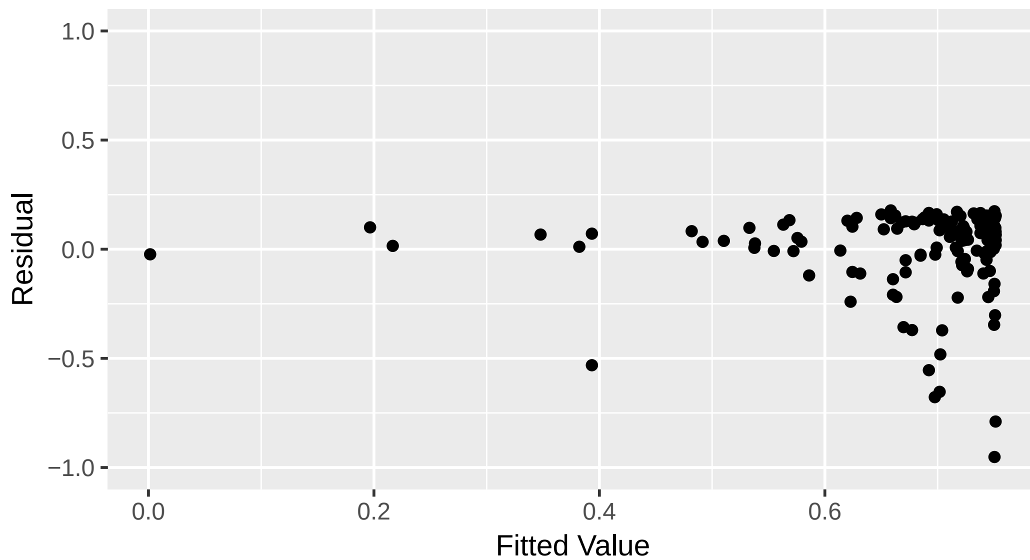


Figure 4.3: A residual plot using the locally weighted regression analysis method.



We will eliminate any points with an absolute residual exceeding a certain value.

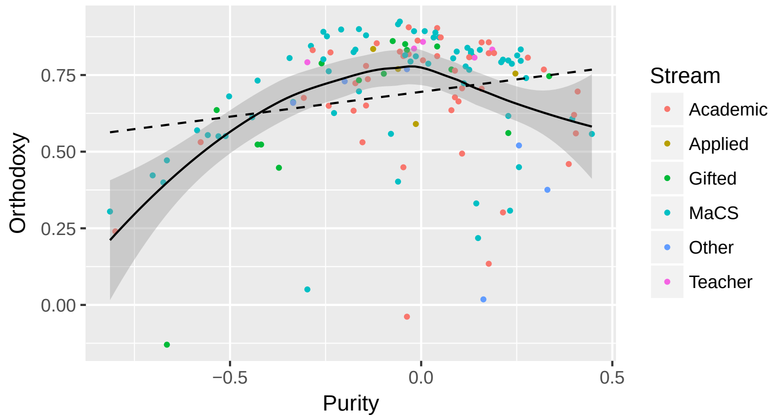
For the linear analysis, we exclude any points with an absolute residual above 0.76. These are respondents #81, and #111. For the locally weighted analysis, we exclude any points with an absolute residual above 0.80. This is just respondent #81.

So, respondents #81 and #111 are the only major outliers. We will continue the rest of the analysis as if they did not do the survey at all.

4.3 Outliers removed

We once again plot orthodoxy vs purity by stream in figure 4.4 on the next page, this time with the outliers removed.

Figure 4.4: A plot of purity vs orthodoxy with data points coloured based on stream. Outliers have been removed.



Here, the solid curve is the result of the locally weighted regression analysis, and the shaded area is the 95% confidence interval for that analysis. The dashed line is the result of the linear regression.

Part IV

Appendices

Appendix A

Questions

A.1 Sandwich Survey

What defines a sandwich?

A.2 Demographic Information

We need to study correlations and demographics.

A.2.1 What grade are you in?

Respondents were asked to select only one option.

○ Grade 9 ○ Grade 10 ○ Grade 11 ○ Grade 12 ○ Teacher

A.2.2 What is/are your favourite subjects?

Answer based on what department teaches the subject at Mackenzie.
Respondents were asked to check all that apply.

- ☐ Arts
- ☐ Business
- ☐ Computer Science and Technology
- ☐ Co-op and Leadership
- ☐ English
- ☐ Languages
- ☐ Library
- ☐ Math
- ☐ Physical Education
- ☐ Science
- ☐ Social Sciences
- ☐ Special Education
- ☐ Other: ______

A.2.3 What is your ethnic background?

Respondents were asked to provide a free response.

A.3 Student Questions

Questions in this section were asked only of respondents who identified as being in grade 9, 10, 11, or 12. It was not asked of respondents who identified as a teacher.

A.3.1 What is your academic stream?

Select whatever option you most closely identify with.
Respondents were asked to select only one option.

- ○ Applied
- ○ Academic
- ○ MaCS
- ○ Gifted
- ○ Other: ______

A.4 Teacher Questions

Questions in this section were asked only of respondents who identified as being a teacher. It was not asked of respondents who identified as being in grade 9, 10, 11, or 12.

A.4.1 What departments do you work in at Mackenzie?

Answer based on what department teaches the subject at Mackenzie.
Respondents were asked to check all that apply.

- ☐ Arts
- ☐ Business
- ☐ Computer Science and Technology
- ☐ Co-op and Leadership
- ☐ English
- ☐ Languages
- ☐ Library
- ☐ Math
- ☐ Physical Education
- ☐ Science
- ☐ Social Sciences
- ☐ Special Education
- ☐ Student Services
- ☐ Other: ______

A.5 Sandwich Questions

For each of the following, answer on a scale from 0 to 10, where 0 is not at all a sandwich, and 10 is definitely a sandwich.

A.5.1 Ham between two slices of bread

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [13].

A.5.2 Bacon, lettuce, and tomato between two slices of bread

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [32].

A.5.3 Grilled cheese

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [4].

A.5.4 Panini

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [37].

A.5.5 Sub

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image cropped from [9].

A.5.6 Hamburger

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [6].

A.5.7 Hotdog

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [34].

A.5.8 Ice cream between two cookies

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [8].

A.5.9 Burrito

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [39].

A.5.10 Taco

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [7].

A.5.11 Calzone

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [20].

A.5.12 Quesadilla

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [16].

A.5.13 Open-faced sandwich

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [30].

A.5.14 Pizza

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [31].

A.5.15 Two slices of pizza, face down, on top of each other

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [40].

A.5.16 Slice of bread between two more slices of bread

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [27].

A.5.17 Oreo cookie

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [12].

A.5.18 Ritz cracker

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [22].

A.5.19 Sushi burrito

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [15].

A.5.20 Nigiri sushi

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [41].

A.5.21 Slice of cheese

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image cropped from [25].

A.5.22 Salad with croutons

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [14].

A.5.23 Brick between two slices of bread

Respondents were asked to select only one option, with 0 being "Not at all a sandwich" and 10 being "Definitely a sandwich".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image created in GIMP using [42] and [23].

A.6 Ingredient Questions

For each of the following, answer on a scale from 0 to 10, where 0 is not at all a sandwich ingredient, and 10 is definitely a sandwich ingredient.

A.6.1 Cold cuts

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [33].

A.6.2 Tomato

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [1].

A.6.3 Cheese

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image cropped from [25].

A.6.4 Lettuce

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [43].

A.6.5 Mayonnaise

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [17].

A.6.6 Mustard

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [2].

A.6.7 Ketchup

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [38].

A.6.8 Carrot

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [18].

A.6.9 Peanut butter

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [28].

A.6.10 Jelly

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [35].

A.6.11 Olives

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [5].

A.6.12 Fish

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [24].

A.6.13 Egg

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image cropped from [26].

A.6.14 Fruit

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [36].

A.6.15 Chicken foot

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [21].

A.6.16 Rice

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [11].

A.6.17 Potato

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [10].

A.6.18 Soup

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [29].

A.6.19 Pasta

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [3].

A.6.20 Grass

Respondents were asked to select only one option, with 0 being "Not at all a sandwich ingredient" and 10 being "Definitely a sandwich ingredient".

○	○	○	○	○	○	○	○	○	○	○
0	1	2	3	4	5	6	7	8	9	10

Image from [19].

Appendix B

Source Code

B.1 Functions.R

```
# Calculate the cosine similarity between two vectors.
cosineSimilarity <- function(a, b) {
        return (sum(a * b) / (sqrt(sum(a^2)) * sqrt(sum(b^2))))
}
```

B.2 Calculate.R

```
# This is a program to calculate the orthodoxy and purity scores on the sandwich spectrum for all respondents.

source("Functions.R")

# Define the constants for what columns data is in inside the CSV file.
GRADE_COLUMN <- 1 # What column the grade data is in.
SUBJECTS_COLUMN <- 2 # What column the favourite subjects data is in.
BACKGROUND_COLUMN <- 3 # What column the ethnic background data is in.
STREAM_COLUMN <- 4 # What column the stream data is in.
```

```
QUESTIONS_START <- 6 # What column the food questions start at.
QUESTIONS_END <- 48 # What column the food questions end at.
lockBinding("GRADE_COLUMN", globalenv())
lockBinding("SUBJECTS_COLUMN", globalenv())
lockBinding("BACKGROUND_COLUMN", globalenv())
lockBinding("STREAM_COLUMN", globalenv())
lockBinding("QUESTIONS_START", globalenv())
lockBinding("QUESTIONS_END", globalenv())

respondents <- read.csv("NoOutliers.csv") # Read in the CSV.
NUM_RESPONDENTS <- nrow(respondents)
lockBinding("NUM_RESPONDENTS", globalenv())

foodResponses <- as.matrix(respondents)[,QUESTIONS_START:QUESTIONS_END] # Convert responses to a matrix.
foodResponses <- apply(foodResponses, 1, as.numeric) # Make the matrix numeric.
foodResponses <- 5 - foodResponses

NUM_QUESTIONS <- nrow(foodResponses) # The number of sandwich questions.
lockBinding("NUM_QUESTIONS", globalenv())

totalResponse <- numeric(NUM_QUESTIONS)
averageResponse <- numeric(NUM_QUESTIONS)

# Sum up the total score for each question by respondent.
for (i in 1 : NUM_RESPONDENTS) {
        totalResponse <- totalResponse + (foodResponses[,i])
}
averageResponse <- totalResponse / NUM_RESPONDENTS # Divide by the number of respondents to find the mean.

orthodoxyScores <- numeric(NUM_RESPONDENTS)
purityScores <- numeric(NUM_RESPONDENTS)

for (i in 1 : NUM_RESPONDENTS) {
        orthodoxyScores[i] <- cosineSimilarity(foodResponses[,i], averageResponse)
        purityScores[i] <- sum(foodResponses[,i])
}
purityScores <- purityScores / (5 * NUM_QUESTIONS)

# Put data into frame.
```

```
data <- data.frame(purity = purityScores, orthodoxy = orthodoxyScores, grade = respondents[,GRADE_COLUMN], subjects = respondents[,SUBJECTS_COLUMN], background = respondents[,BACKGROUND_COLUMN], stream = respondents[,STREAM_COLUMN])
```

B.3 BoxPlots.R

```
# This program creates box plots of the respondents.

library(ggplot2)
source("Calculate.R")

makeBoxPlot <- function(categoryData, categoryName, categoryTitle) {
    boxPlotData <- data[categoryData %in% names(table(categoryData))[table(categoryData) > 1],] # Remove all categorical data points only occurring once, as these data are not helpful for a box plot.

    # Create the purity plot.
    dataPlot <- ggplot(boxPlotData, aes_string(x = categoryName, y = "purity", fill=categoryName)) # Setup the plot.
    dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
    dataPlot <- dataPlot + geom_boxplot() # Add the data points.
    dataPlot <- dataPlot + labs(x = categoryTitle, y = "Purity") # Give axes proper labels.
    dataPlot <- dataPlot + theme(legend.position = "none") # Remove the legend.
    ggsave(paste(categoryTitle, "Purity.pdf"), plot=dataPlot, width=9, height=8)

    # Create the orthodoxy plot.
    dataPlot <- ggplot(boxPlotData, aes_string(x = categoryName, y = "orthodoxy", fill=categoryName)) # Setup the plot.
    dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
```

```
        dataPlot <- dataPlot + geom_boxplot() # Add the data points.
        dataPlot <- dataPlot + labs(x = categoryTitle, y = "Orthodoxy") # Give axes proper labels.
        dataPlot <- dataPlot + theme(legend.position = "none") # Remove the legend.
        ggsave(paste(categoryTitle, "Orthodoxy.pdf"), plot=dataPlot, width=9, height=8)
}

makeBoxPlot(data$grade, "grade", "Grade")
makeBoxPlot(data$background, "background", "Ethnic Background")
makeBoxPlot(data$stream, "stream", "Stream")

boxPlotData <- data

# Create the purity plot.
dataPlot <- ggplot(boxPlotData, aes(x = 1, y = purity)) # Setup the plot.
dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
dataPlot <- dataPlot + geom_boxplot() # Add the data points.
dataPlot <- dataPlot + labs(x = "All respondents", y = "Purity") # Give axes proper labels.
dataPlot <- dataPlot + theme(axis.text.x = element_blank(), axis.ticks.x = element_blank()) # Remove the x axis.
ggsave("PurityBoxPlot.pdf", plot=dataPlot, width=2, height=8)

# Create the orthodoxy plot.
dataPlot <- ggplot(boxPlotData, aes(x = 1, y = orthodoxy)) # Setup the plot.
dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
dataPlot <- dataPlot + geom_boxplot() # Add the data points.
dataPlot <- dataPlot + labs(x = "All respondents", y = "Orthodoxy") # Give axes proper labels.
dataPlot <- dataPlot + theme(axis.text.x = element_blank(), axis.ticks.x = element_blank()) # Remove the x axis.
ggsave("OrthodoxyBoxPlot.pdf", plot=dataPlot, width=2, height=8)
```

B.4 PurityOrthodoxyPlot.R

```
# This is a program to plot orthodoxy vs purity on the sandwich spectrum for all respondents.

library(ggplot2)
source("Calculate.R")

makeScatterPlot <- function(categoryName, categoryTitle) {
        # Create the plot.
        dataPlot <- ggplot(data, aes(purity, orthodoxy)) # Setup the plot.
        #dataPlot <- dataPlot + coord_cartesian(xlim = c(-1, 1), ylim = c(-1, 1)) # Set the graph limits.
        dataPlot <- dataPlot + geom_point(aes_string(colour = categoryName), size = 0.75) # Add the data points.
        dataPlot <- dataPlot + geom_smooth(method = loess, level = 0.95, colour = "black", size = 0.5) # Add the confidence curve.
        dataPlot <- dataPlot + geom_smooth(method = lm, se = FALSE, colour = "black", size = 0.5, linetype="dashed") # Add the line of best fit.
        dataPlot <- dataPlot + labs(x = "Purity", y = "Orthodoxy", colour = categoryTitle) # Give axes and legend proper labels.
        ggsave(paste(categoryTitle, "PurityVsOrthodoxy.pdf", sep = ""), plot=dataPlot, width=5.5, height=3)
}

makeScatterPlot("stream", "Stream")
makeScatterPlot("grade", "Grade")
makeScatterPlot("background", "Ethnic Background")
```

B.5 Residuals.R

```
# This program creates a residual plot of the respondents.

library(ggplot2)
library(broom)

source("Calculate.R")

```

```r
residualPlotData <- data

# Create the linear model.
mod <- lm(orthodoxy ~ purity, data = residualPlotData)
df <- augment(mod)
#print(df[which(df$.resid < -0.76),])
which(df$.resid < -0.76)

# Create the residual plot.
dataPlot <- ggplot(df, aes(.fitted, .resid)) + geom_point()
dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
dataPlot <- dataPlot + labs(x = "Fitted Value", y = "Residual") # Give axes proper labels.
ggsave("LinearResidualPlot.pdf", plot=dataPlot, width=5.5, height=3)

# Create the loess model.
mod <- loess(orthodoxy ~ purity, data = residualPlotData)
df <- augment(mod)
#print(df[which(df$.resid < -0.8),])
which(df$.resid < -0.80)

# Create the residual plot.
dataPlot <- ggplot(df, aes(.fitted, .resid)) + geom_point()
dataPlot <- dataPlot + coord_cartesian(ylim = c(-1, 1)) # Set the graph limits.
dataPlot <- dataPlot + labs(x = "Fitted Value", y = "Residual") # Give axes proper labels.
ggsave("LoessResidualPlot.pdf", plot=dataPlot, width=5.5, height=3)
```

B.6 Levels.R

```r
# Output the levels for certain columns of interest in the input.

# Define the constants for what columns data is in inside the CSV file.
GRADE_COLUMN <- 1 # What column the grade data is in.
SUBJECTS_COLUMN <- 2 # What column the favourite subjects data is in.
```

```r
BACKGROUND_COLUMN <- 3 # What column the ethnic background data is in.
STREAM_COLUMN <- 4 # What column the stream data is in.
QUESTIONS_START <- 6 # What column the food questions start at.
QUESTIONS_END <- 48 # What column the food questions end at.
lockBinding("GRADE_COLUMN", globalenv())
lockBinding("SUBJECTS_COLUMN", globalenv())
lockBinding("BACKGROUND_COLUMN", globalenv())
lockBinding("STREAM_COLUMN", globalenv())
lockBinding("QUESTIONS_START", globalenv())
lockBinding("QUESTIONS_END", globalenv())

respondents <- read.csv("CleanedData.csv") # Read in the CSV.

print(paste(nrow(respondents), "respondents"))

summary(respondents[,GRADE_COLUMN])
print(paste(length(levels(respondents[,GRADE_COLUMN])), " grades"))

#summary(respondents[,SUBJECTS_COLUMN])
#print(paste(length(levels(respondents[,SUBJECTS_COLUMN])), " subjects."))

summary(respondents[,BACKGROUND_COLUMN])
print(paste(length(levels(respondents[,BACKGROUND_COLUMN])), "ethnic backgrounds"))

summary(respondents[,STREAM_COLUMN])
print(paste(length(levels(respondents[,STREAM_COLUMN])), " streams"))
```

Bibliography

[1] Adityamadhav83. *Sliced Round Tomato at Madhurawada, Visakhapatnam*. Oct. 10, 2012. URL: https://commons.wikimedia.org/wiki/File:Sliced_Round_Tomato_at_Madhurawada.JPG (visited on 2018-05-21).

[2] Aeæ commonswiki. *A squeeze bottle of French's Classic Yellow Mustard*. May 21, 2006. URL: https://commons.wikimedia.org/wiki/File:French%27s_classic_yellow_mustard.jpg (visited on 2018-05-21).

[3] Tim Bartel. *Spaghetti*. Dec. 20, 2004. URL: https://commons.wikimedia.org/wiki/File:Spaghetti.jpg (visited on 2018-05-21).

[4] Mack Male from Edmonton Canada AB. *Grilled Cheese Olympics*. Mar. 14, 2010. URL: https://commons.wikimedia.org/wiki/File:Simple_grilled_cheese.jpg (visited on 2018-05-19).

[5] Renee Comet. *green olives*. 1994. URL: https://commons.wikimedia.org/wiki/File:NCI_2_green_olives.jpg (visited on 2018-05-21).

[6] Renee Comet. *Hamburger*. 1994. URL: https://commons.wikimedia.org/wiki/File:Cheeseburger_(2).jpg (visited on 2018-05-19).

[7] Renee Comet. *Hard-shell taco with meat, cheese, lettuce, tomatoes, and onions*. 1994. URL: https://commons.wikimedia.org/wiki/File:NCI_Visuals_Food_Taco.jpg (visited on 2018-05-19).

[8] Renee Comet. *Ice Cream Sandwich*. 1994. URL: https://commons.wikimedia.org/wiki/File:Ice_cream_sandwich_(1).jpg (visited on 2018-05-19).

[9] Crisco 1492. *A Subway pizza sub in Hong Kong (near the Grand Kowloon)*. Aug. 10, 2013. URL: https://commons.wikimedia.org/wiki/File:Subway_Pizza_Sub,_Hong_Kong.JPG (visited on 2018-05-19).

[10] Doudoulolita. *3 potatoes*. Mar. 21, 2012. URL: https://commons.wikimedia.org/wiki/File:Pommes_de_terre.png (visited on 2018-05-21).

[11] Earth100. *The four, main, common forms of consumed rice*. Nov. 16, 2012. URL: https://commons.wikimedia.org/wiki/File:White,_Brown,_Red_%26_Wild_rice.jpg (visited on 2018-05-21).

[12] Evan-Amos. *Two Double Stuf Oreos, by Nabisco. Bought in a New York supermarket.* Mar. 31, 2017. URL: https://commons.wikimedia.org/wiki/File:Nabisco-Double-Stuf-Oreos.jpg (visited on 2018-05-20).

[13] Fancy steve. *A plain ham sandwich*. Nov. 11, 2011. URL: https://commons.wikimedia.org/wiki/File:Ham_sandwich.jpg (visited on 2018-05-19).

[14] jeffreyw. *Romaine lettuce salad*. Jan. 28, 2011. URL: https://commons.wikimedia.org/wiki/File:Romaine_lettuce_salad.jpg (visited on 2018-05-19).

[15] jen. *sushirrito*. July 9, 2013. URL: https://commons.wikimedia.org/wiki/File:Sushirrito_(9310623804).jpg (visited on 2018-05-21).

[16] Jennifer. *Lamb quesadillas made with leftover grilled lamb, cheddar cheese, red onion, and peppers from the garden*. Feb. 7, 2005. URL: https://commons.wikimedia.org/wiki/File:Lamb_quesadilla.jpg (visited on 2018-05-20).

[17] Joey 424. *mayonaise on da hot afternoon [sic]*. June 1, 1933. URL: https://commons.wikimedia.org/wiki/File:Jorgedamayo.jpg (visited on 2018-05-21).

[18] Kander. *Carrots on display at local greengrocer*. Oct. 27, 2007. URL: https://commons.wikimedia.org/wiki/File:Carrots.JPG (visited on 2018-05-21).

[19] Ubaid kareem. *Beauty of grass*. Oct. 28, 2013. URL: https://commons.wikimedia.org/wiki/File:Beauty_of_grass.jpg (visited on 2018-05-21).

[20] Kobako. *Calzone*. Aug. 7, 2013. URL: https://commons.wikimedia.org/wiki/File:Calzone04.jpg (visited on 2018-05-20).

[21] Krista. *A chicken foot*. Dec. 24, 2007. URL: https://commons.wikimedia.org/wiki/File:A_chicken_foot.jpg (visited on 2018-05-21).

[22] Lance. *Lance Gluten Free Sandwich Crackers*. URL: https://images-na.ssl-images-amazon.com/images/I/71bduHPOL9L._SL1500_.jpg (visited on 2018-05-21).

[23] Lowes. *Bennett is brick*. Apr. 16, 2017. URL: https://commons.wikimedia.org/wiki/File:Brick_Bennett.jpg (visited on 2018-05-19).

[24] Tiia Monto. *Cod fillet baked in oven.* Mar. 7, 2018. URL: https://commons.wikimedia.org/wiki/File:Baked_cod_fillet.jpg (visited on 2018-05-21).

[25] Tiia Monto. *Emmental cheese slices.* Mar. 26, 2016. URL: https://commons.wikimedia.org/wiki/File:Emmental_slices.jpg (visited on 2018-05-21).

[26] Porsche997SBS. *A picture of an egg being fried*. Nov. 10, 2006. URL: https://commons.wikimedia.org/wiki/File:Fried_egg,_sunny_side_up_(black_background).PNG (visited on 2018-05-21).

[27] Qwantz. *This is an image of a toast sandwich, shot from the side. Sandwich was made and photographed by me, Ryan. It was okay.* Mar. 26, 2014. URL: https://commons.wikimedia.org/wiki/File:An_image_of_a_toast_sandwich,_shot_from_the_side.jpg (visited on 2018-05-20).

[28] Susan Slater. *peanut butter*. Aug. 4, 2015. URL: https://commons.wikimedia.org/wiki/File:96_-_IMG_20150804_111725.jpg (visited on 2018-05-21).

[29] Stu Spivack. *Creamed corn soup*. June 14, 2006. URL: https://commons.wikimedia.org/wiki/File:Creamed_corn.jpg (visited on 2018-05-21).

[30] stef yau stef. *tuna, guacamole sauce and cherry tomatoes*. Aug. 5, 2007. URL: https://commons.wikimedia.org/wiki/File:Open_faced_tuna_sandwich_with_cherry_tomato_and_guacamole_spread.jpg (visited on 2018-05-20).

[31] Becky Stern. *Pizza at State of the Fruit meeting*. June 19, 2015. URL: https://commons.wikimedia.org/wiki/File:Pizza_at_State_of_the_Fruit_meeting_(18611225834).jpg (visited on 2018-05-20).

[32] Groves Steven. *A bacon, lettuce and tomato sandwich on toasted bread.* Feb. 14, 2011. URL: https://commons.wikimedia.org/wiki/File:BLT_sandwich_on_toast.jpg (visited on 2018-05-19).

[33] Stilfehler. *Pastrami*. May 8, 2009. URL: https://commons.wikimedia.org/wiki/File:Pastrami.jpg (visited on 2018-05-21).

[34] Evan Swigart. *Hot Dog*. Sept. 8, 2010. URL: https://commons.wikimedia.org/wiki/File:Hotdog_-_Evan_Swigart.jpg (visited on 2018-05-19).

[35] terri_bateman. *blackberry jelly*. July 21, 2014. URL: https://commons.wikimedia.org/wiki/File:Blackberry_jelly_(20248988916).jpg (visited on 2018-05-21).

[36] Tomwsulcer. *Fruit platter*. Mar. 30, 2013. URL: https://commons.wikimedia.org/wiki/File:Fruit_platter_at_a_party_with_kiwi_melons_strawberries_etc.JPG (visited on 2018-05-21).

[37] Tomwsulcer. *Panini sandwich at a restaurant*. Feb. 27, 2013. URL: https://commons.wikimedia.org/wiki/File:Panini_sandwich_in_a_restaurant.JPG (visited on 2018-05-19).

[38] Tt341. *Heinz Organic Ketchup*. Oct. 2, 2016. URL: https://commons.wikimedia.org/wiki/File:Heinz_USDA_Organic_Tomato_Ketchup.jpg (visited on 2018-05-21).

[39] Elliot Volkman. *A basic burrito*. June 24, 2009. URL: https://commons.wikimedia.org/wiki/File:Basic_burrito.jpg (visited on 2018-05-19).

[40] Gabby Wang. *Pizza sandwich*. URL: https://s3.amazonaws.com/spoonuniversi-wpengine/spoonuniversi/wp-content/uploads/sites/71/2015/02/pizza-sandwich-1.jpg (visited on 2018-05-20).

[41] www.bluewaikiki.com. *Salmon Nigiri Sushi with chopsticks, 2008*. Sept. 5, 2008. URL: https://commons.wikimedia.org/wiki/File:Salmon_Nigiri_Sushi_with_chopsticks,_2008.jpg (visited on 2018-05-21).

[42] yamada taro. *Slices Of Bread Free Stock Photo - Public Domain Pictures*. URL: https://www.publicdomainpictures.net/en/view-image.php?image=26198&picture=slices-of-bread (visited on 2018-05-19).

[43] Rainer Zenz. *Romaine lettuce*. 2006. URL: https://commons.wikimedia.org/wiki/File:Romaine.jpg (visited on 2018-05-21).