

# STAT 579 Final Project

Chelsey Legacy, Lindong Zhou, Evan Pete Walsh\*

November 23, 2014

## **Abstract**

Analysis of the Yelp Academic Dataset.

---

\*Statistics graduate students at Iowa State University of Science and Technology

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data</b>	<b>1</b>
<b>3</b>	<b>Restaurants</b>	<b>1</b>
3.1	Breakfast . . . . .	1
3.2	Lunch . . . . .	1
3.3	Dinner . . . . .	1
<b>4</b>	<b>Bars</b>	<b>1</b>
<b>5</b>	<b>Hotels</b>	<b>1</b>
<b>6</b>	<b>Fitness</b>	<b>2</b>
<b>7</b>	<b>Conclusion</b>	<b>2</b>

## 1 Introduction

Lindong

## 2 Data

The Yelp Academic Dataset<sup>1</sup> provides data enthusiasts with the exciting opportunity to explore an incredible collection of information regarding the characteristics and quality of hundreds of businesses across the United States, Canada, and UK. Specifically, the data includes details and reviews on 250 of the closest business to 30 large universities, including the Arizona State, UNLV, the University of Edinburgh, the University of Wisconsin, and the University of Waterloo, to name a few. The raw data is in json format and contains three different types of json objects: 'business,' 'review,' and 'user.'

Describe each object....

For the scope of this study, we will only be using the 'business' dataset. Our first concern was manipulating the data into a more usable format.

Go through method of cleaning....

## 3 Restaurants

### 3.1 Breakfast

### 3.2 Lunch

### 3.3 Dinner

## 4 Bars

Chelsey

## 5 Hotels

Lindong

---

<sup>1</sup>[https://www.yelp.com/academic\\_dataset](https://www.yelp.com/academic_dataset)

## **6 Fitness**

Lindong

## **7 Conclusion**

Chelsey