# Chopsticks!

A few researchers set out to determine the optimal length of chopsticks for children and adults. They came up with a measure of how effective a pair of chopsticks performed, called the "Food Pinching Performance." The "Food Pinching Performance" was determined by counting the number of peanuts picked and placed in a cup (PPPC).

## An investigation for determining the optimum length of chopsticks.

Link to Abstract and Paper (http://www.ncbi.nlm.nih.gov/pubmed/15676839)
*the abstract below was adapted from the link*

Chopsticks are one of the most simple and popular hand tools ever invented by humans, but have not previously been investigated by ergonomists (https://www.google.com/search?q=ergonomists). Two laboratory studies were conducted in this research, using a randomised complete block design (http://dawg.utk.edu/glossary/whatis_rcbd.htm), to evaluate the effects of the length of the chopsticks on the food-serving performance of adults and children. Thirty-one male junior college students and 21 primary school pupils served as subjects for the experiment to test chopsticks lengths of 180, 210, 240, 270, 300, and 330 mm. The results showed that the food-pinching performance was significantly affected by the length of the chopsticks, and that chopsticks of about 240 and 180 mm long were optimal for adults and pupils, respectively. Based on these findings, the researchers suggested that families with children should provide both 240 and 180 mm long chopsticks. In addition, restaurants could provide 210 mm long chopsticks, considering the trade-offs between ergonomics and cost.

## For the rest of this project, answer all questions based only on the part of the experiment analyzing the thirty-one adult male college students.

Download the data set for the adults (https://www.udacity.com/api/nodes/4576183932/supplemental_media/chopstick-effectivenesscsv/download), then answer the following questions based on the abstract and the data set.

**If you double click on this cell**, you will see the text change so that all of the formatting is removed. This allows you to edit this block of text. This block of text is written using Markdown (http://daringfireball.net/projects/markdown/syntax), which is a way to format text using headers, links, italics, and many other options. You will learn more about Markdown later in the Nanodegree Program. Hit shift + enter or shift + return to show the formatted text.

**1. What is the independent variable in the experiment?**

The indepedent variable for this experiment is lenght of chopsticks.

**2. What is the dependent variable in the experiment?**

The dependent variable is the food pinching efficienty.

**3. How is the dependent variable operationally defined?**

The dependent variable (chopstick lenght in mm) is operationally defined as food pinching performance determined by counting the number of peanuts picked and placed in a cup (PPPC).

**4. Based on the description of the experiment and the data set, list at least two variables that you know were controlled.**

Based on description of the experiment and the data set number of controlled variables can be identidfied:

•sex of participants (only men were chosen for this experiment)

•education status of the participants (only junior college students were chosen for this experiment)

•age of the participants (only 31 years old were chosen for the experiment)

One great advantage of ipython notebooks is that you can document your data analysis using code, add comments to the code, or even add blocks of text using Markdown. These notebooks allow you to collaborate with others and share your work. For now, let's see some code for doing statistics.

```
In [2]:  import pandas as pd

         # pandas is a software library for data manipulation and analysis
         # We commonly use shorter nicknames for certain packages. Pandas is often abbr
         eviated to pd.
         # hit shift + enter to run this cell or block of code
```

In [3]:
```
path = r'~/Downloads/chopstick-effectiveness.csv'
# Change the path to the location where the chopstick-effectiveness.csv file i
s located on your computer.
# If you get an error when running this block of code, be sure the chopstick-e
ffectiveness.csv is located at the path on your computer.

dataFrame = pd.read_csv(path)
dataFrame
```

Out[3]:

|     | Food.Pinching.Efficiency | Individual | Chopstick.Length |
| --- | --- | --- | --- |
| **0** | 19.55 | 1 | 180 |
| **1** | 27.24 | 2 | 180 |
| **2** | 28.76 | 3 | 180 |
| **3** | 31.19 | 4 | 180 |
| **4** | 21.91 | 5 | 180 |
| **5** | 27.62 | 6 | 180 |
| **6** | 29.46 | 7 | 180 |
| **7** | 26.35 | 8 | 180 |
| **8** | 26.69 | 9 | 180 |
| **9** | 30.22 | 10 | 180 |
| **10** | 27.81 | 11 | 180 |
| **11** | 23.46 | 12 | 180 |
| **12** | 23.64 | 13 | 180 |
| **13** | 27.85 | 14 | 180 |
| **14** | 20.62 | 15 | 180 |
| **15** | 25.35 | 16 | 180 |
| **16** | 28.00 | 17 | 180 |
| **17** | 23.49 | 18 | 180 |
| **18** | 27.77 | 19 | 180 |
| **19** | 18.48 | 20 | 180 |
| **20** | 23.01 | 21 | 180 |
| **21** | 22.66 | 22 | 180 |
| **22** | 23.24 | 23 | 180 |
| **23** | 22.82 | 24 | 180 |
| **24** | 17.94 | 25 | 180 |
| **25** | 26.67 | 26 | 180 |
| **26** | 28.98 | 27 | 180 |
| **27** | 21.48 | 28 | 180 |
| **28** | 14.47 | 29 | 180 |
| **29** | 28.29 | 30 | 180 |
| **...** | ... | ... | ... |
| **156** | 26.18 | 2 | 330 |

|     | Food.Pinching.Efficiency | Individual | Chopstick.Length |
| --- | --- | --- | --- |
| **157** | 25.93 | 3 | 330 |
| **158** | 28.61 | 4 | 330 |
| **159** | 20.54 | 5 | 330 |
| **160** | 26.44 | 6 | 330 |
| **161** | 29.36 | 7 | 330 |
| **162** | 19.77 | 8 | 330 |
| **163** | 31.69 | 9 | 330 |
| **164** | 24.64 | 10 | 330 |
| **165** | 22.09 | 11 | 330 |
| **166** | 23.42 | 12 | 330 |
| **167** | 28.63 | 13 | 330 |
| **168** | 26.30 | 14 | 330 |
| **169** | 22.89 | 15 | 330 |
| **170** | 22.68 | 16 | 330 |
| **171** | 30.92 | 17 | 330 |
| **172** | 20.74 | 18 | 330 |
| **173** | 27.24 | 19 | 330 |
| **174** | 17.12 | 20 | 330 |
| **175** | 23.63 | 21 | 330 |
| **176** | 20.91 | 22 | 330 |
| **177** | 23.49 | 23 | 330 |
| **178** | 24.86 | 24 | 330 |
| **179** | 16.28 | 25 | 330 |
| **180** | 21.52 | 26 | 330 |
| **181** | 27.22 | 27 | 330 |
| **182** | 17.41 | 28 | 330 |
| **183** | 16.42 | 29 | 330 |
| **184** | 28.22 | 30 | 330 |
| **185** | 27.52 | 31 | 330 |

186 rows × 3 columns

Let's do a basic statistical calculation on the data using code! Run the block of code below to calculate the average "Food Pinching Efficiency" for all 31 participants and all chopstick lengths.

```
In [4]: dataFrame['Food.Pinching.Efficiency'].mean()
```

```
Out[4]: 25.00559139784947
```

This number is helpful, but the number doesn't let us know which of the chopstick lengths performed best for the thirty-one male junior college students. Let's break down the data by chopstick length. The next block of code will generate the average "Food Pinching Effeciency" for each chopstick length. Run the block of code below.

```
In [5]: meansByChopstickLength = dataFrame.groupby('Chopstick.Length')['Food.Pinching.
        Efficiency'].mean().reset_index()
        meansByChopstickLength

        # reset_index() changes Chopstick.Length from an index to column. Instead of t
        he index being the length of the chopsticks, the index is the row numbers 0,
         1, 2, 3, 4, 5.
```

Out[5]:

|   | Chopstick.Length | Food.Pinching.Efficiency |
|---|---|---|
| 0 | 180 | 24.935161 |
| 1 | 210 | 25.483871 |
| 2 | 240 | 26.322903 |
| 3 | 270 | 24.323871 |
| 4 | 300 | 24.968065 |
| 5 | 330 | 23.999677 |

**5. Which chopstick length performed the best for the group of thirty-one male junior college students?**

240mm chopstick performed the best

In [6]:
```
# Causes plots to display within the notebook rather than in a new window
%pylab inline

import matplotlib.pyplot as plt

plt.scatter(x=meansByChopstickLength['Chopstick.Length'], y=meansByChopstickLe
ngth['Food.Pinching.Efficiency'])
            # title="")
plt.xlabel("Length in mm")
plt.ylabel("Efficiency in PPPC")
plt.title("Average Food Pinching Efficiency by Chopstick Length")
plt.show()
```

Populating the interactive namespace from numpy and matplotlib
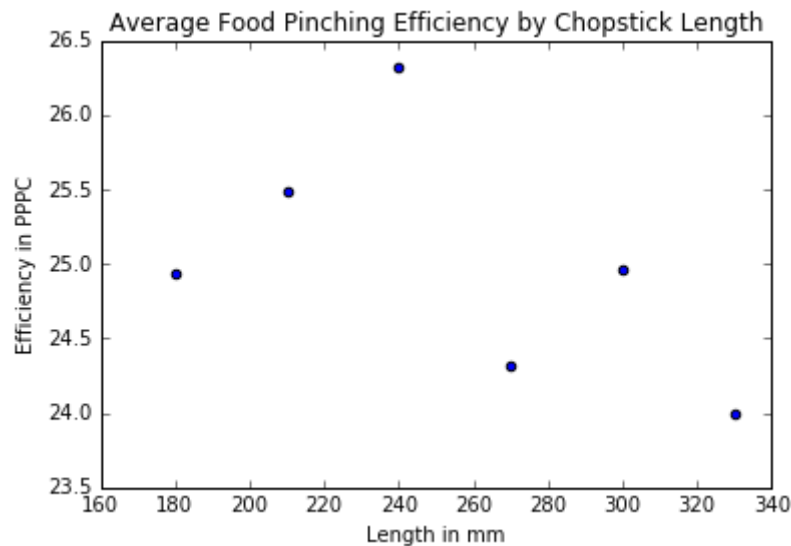


**6. Based on the scatterplot created from the code above, interpret the relationship you see. What do you notice?**

We can observe that the average efficenty of food pinching is increasing starting with 180 to 240mm chopstick. We see the liniar relationship. We can see 270mm and 330mm chopsticks have the lowest food pinching effcienty. The average food pinching efficiency of 300mm is higher than of 270mm and 330mm chopsticks. This can be interpreted as an outlier or possibily a preference by a minority of the people who took a part in this experiment.

**In the abstract the researchers stated that their results showed food-pinching performance was significantly affected by the length of the chopsticks, and that chopsticks of about 240 mm long were optimal for adults.**

**7a. Based on the data you have analyzed, do you agree with the claim?**

**7b. Why?**

Yes i agree with the claim as we can see from the data set that the 240mm showed to have the highest (maximum) average of food pinching efficeny in above described experiment.