

Global ABC Company

Group D

Team members: Héctor Orellana, Patrícia Mesquita

Team leader: Héctor Orellana

Introduction	3
Business Problem(s)	3
Business Solution(s)	3
Exploratory Data Analysis	3
Normal Distribution	4
Target Market	5
Conclusion	5

Introduction

After analyzing the data set from the BRFSS, we would like to recommend the CEO of Global ABC Company to invest in a fitness business in North Carolina, as after reading the results, it will definitely be a very profitable business.

The difficulties we had was that there were a lot of questions and the population is really big, so we had to organize the data, to be able to get interesting conclusions and help us make decisions. One of the ways we found to organize the information was to transform the gross data into pivot tables, with the information we found more relevant. After that we did our analysis and used some important charts to validate our problem and solutions.

Below we detail our recommendations and data analysis.

Business Problem

When we first looked at the data variables, what stood out more clearly was the quantity of respondents who are overweight and have health issues. So, we started digging into some specific questions and found out that the problem we want to solve and validate is:

- There are too many overweight respondents who are female and over 45 years old, who apparently do some exercise, although we don't know at what level, and who earn less than \$75k a year.

Business Solution

We will open a **Mature Women only fitness center**. Our Target will be: Women 45 years old or above, who are overweight but have exercised in the last 30 days. They live in Metropolitan areas and have an income of less then \$75k a year, which is the majority, so we will create a low cost gym for this target.

Exploratory Data Analysis

DATA CONSIDERED FOR OUR DECISION:		
People surveyed in North Carolina	5,842	
North Carolina population	10,490,000	
Target Market from Survey	329	

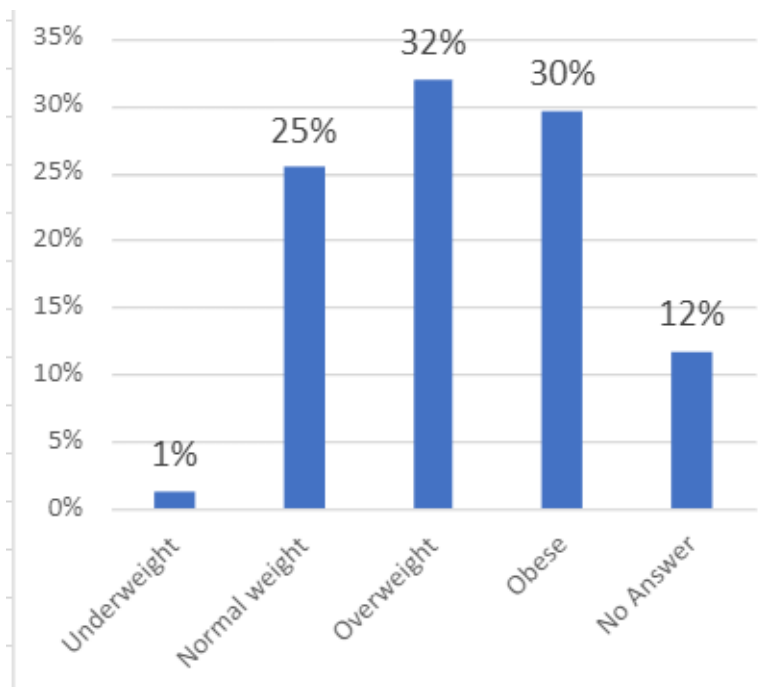
Possible customer %	5.60%	
North Carolina x Target %	590,758	NOTE: possible customers in the state

To get to the above information, we worked the data set with Pivot tables, that can be found here:

<https://onedrive.live.com/view.aspx?resid=6B2EB006038E06B6!1156&ithint=file%2cxlsx&authkey=!AAawbYxf7c-na48>

The most interesting results can be easily understood through the following graphs:

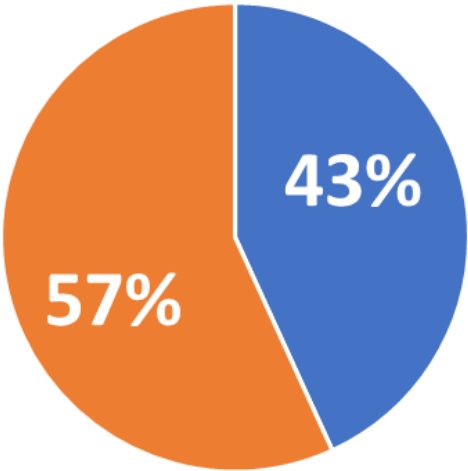
Graph 1: Computed body mass index of women in North Carolina:



Overweight starts at 2500 body mass index, the mean and median are above this limit. People from North Carolina tend to be overweight.

Mean	2850
Median	2744
Standard Deviation	621

Graph 2: People over 45 years old who are female



■ Male ■ Female

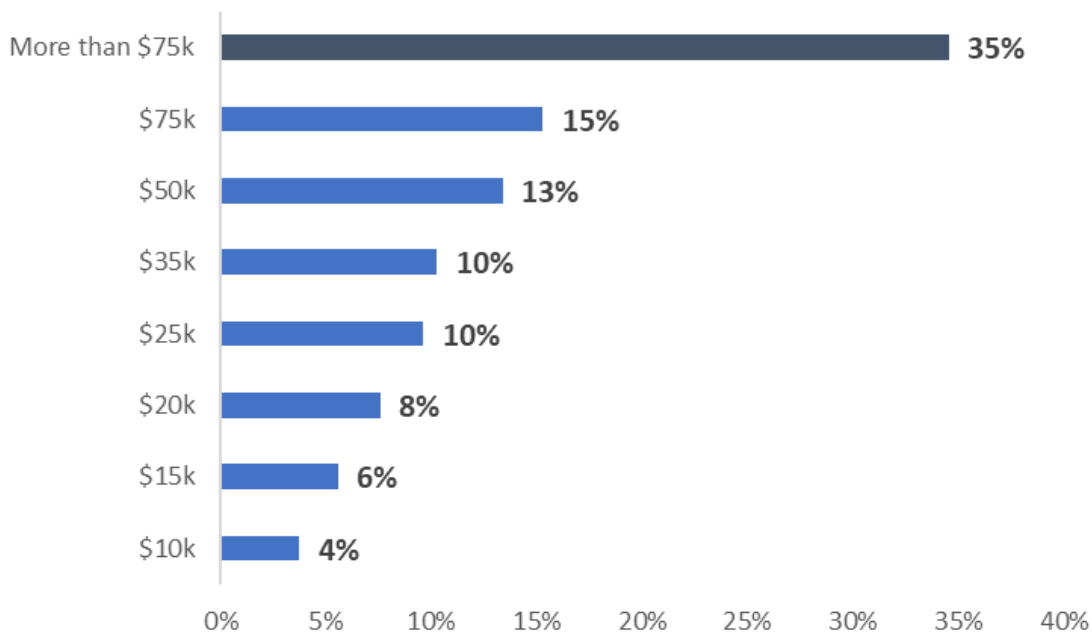
There are more women in North Carolina than there are men, in the ages above 45.

Of all NC, 55% are female.

Age behavior in NC:

Mean	52
Median	54
Standard Deviation	18

Graph 3: Income Level



In the US, most gym members have an income above \$75k per household. In North Carolina, for people above 45, 35% have an income of this level. However the remaining 65% can leave an open market for low subscription price gyms.

Analysis of Frequencies: North Carolina vs. U.S

Imputed Age in Six Groups	Frequency		%	
	North Carolina	US	North Carolina	US
1 (Age 18 to 24)	411	25,652	7.0%	6.4%
2 (Age 25 to 34)	780	44,382	13.4%	11.0%
3 (Age 35 to 44)	806	51,971	13.8%	12.9%
4 (Age 45 to 54)	1,020	62,033	17.5%	15.4%
5 (Age 55 to 64)	1,046	78,089	17.9%	19.4%
6 (Age 65 or older)	1,779	139,831	30.5%	34.8%
Total	5,842	401,958	100.0%	100.0%

Annual household Income Level	Frequency		%	
	North Carolina	US	North Carolina	US
1 (Less than \$10,000)	174	12,903	3.8%	4.0%
2 (\$10,000 - \$14,999)	233	13,705	5.1%	4.3%
3 (\$15,000 - \$19,999)	340	21,048	7.5%	6.5%
4 (\$20,000 - \$24,999)	434	27,719	9.5%	8.6%
5 (\$25,000 - \$34,999)	475	31,410	10.4%	9.8%
6 (\$35,000 - \$49,999)	651	43,851	14.3%	13.6%
7 (\$50,000 - \$74,999)	688	52,486	15.1%	16.3%
8 (\$75,000 or more)	1,565	118,779	34.3%	36.9%
Total	4,560	321,901	100%	100%

Exercise in the past 30 days	Frequency		%	
	North Carolina	US	North Carolina	US
1 (Yes)	4,494	305,883	77.0%	76.2%
2 (No)	1,344	95,393	23.0%	23.8%
Total	5,838	401,276	100%	100%

Computed body mass index in groups	Frequency		%	
	North Carolina	US	North Carolina	US
1 (Underweight : BMI5 < 1850)	78	5,993	1.3%	1.5%
2 (Normal Weight: 1850 <= _BMI5 < 2500)	1,483	110,121	25.4%	27.4%
3 (Overweight: 2500 <= _BMI5 < 3000)	1,868	128,946	32.0%	32.1%
4 (Obese: 3000 <= _BMI5 < 9999)	1,731	115,541	29.6%	28.7%
blank	682	41,357	11.7%	10.3%
Total	5,842	401,958	100%	100%

Gender of Repondent	Frequency		%	
	North Carolina	US	North Carolina	US
1 (Male)	2,640	183,942	45.2%	45.8%
2 (Female)	3,202	218,016	54.8%	54.2%
Total	5,842	401,958	100%	100%

By analyzing these frequencies tables, we can conclude that the North Carolina date set that is important for our business idea formulation, are side by side with the overall results of the

country, which means, that the business may have potential in other states in the country too (may be scalable), which again, validates our business idea potential.

Normal Distribution

We have found a normal distribution with the variable “Computed body mass index”, so that it can help us to better identify our target for the new business.

So, first we have calculated the minimum, maximum, mean and standard deviation of our distribution, as well as the difference of the maximum and the minimum units:

MAX	7115
MIN	1378
DIF	5737

MEAN	2850
ST. DEV	621

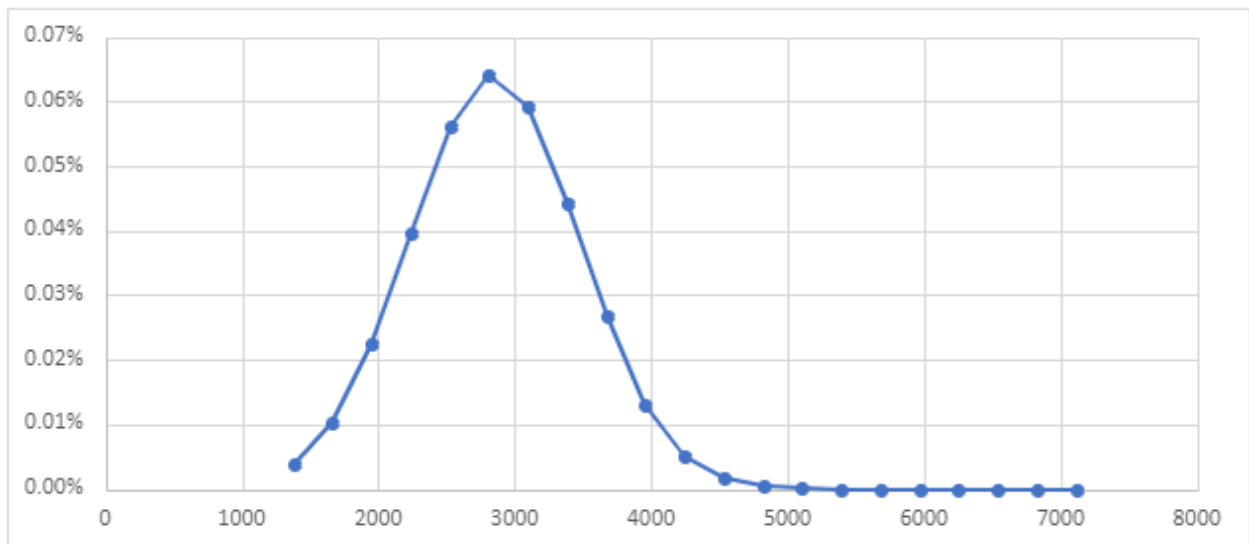
After that, we have split the difference by 20 different columns, and with that we have computed the Probability Norm of each of the units, until the maximum value, which gave us this results:

Variables Unit (BMI)	Probability Norm
1378	0.00%
1665	0.01%
1952	0.02%
2239	0.04%
2525	0.06%
2812	0.06%
3099	0.06%
3386	0.04%

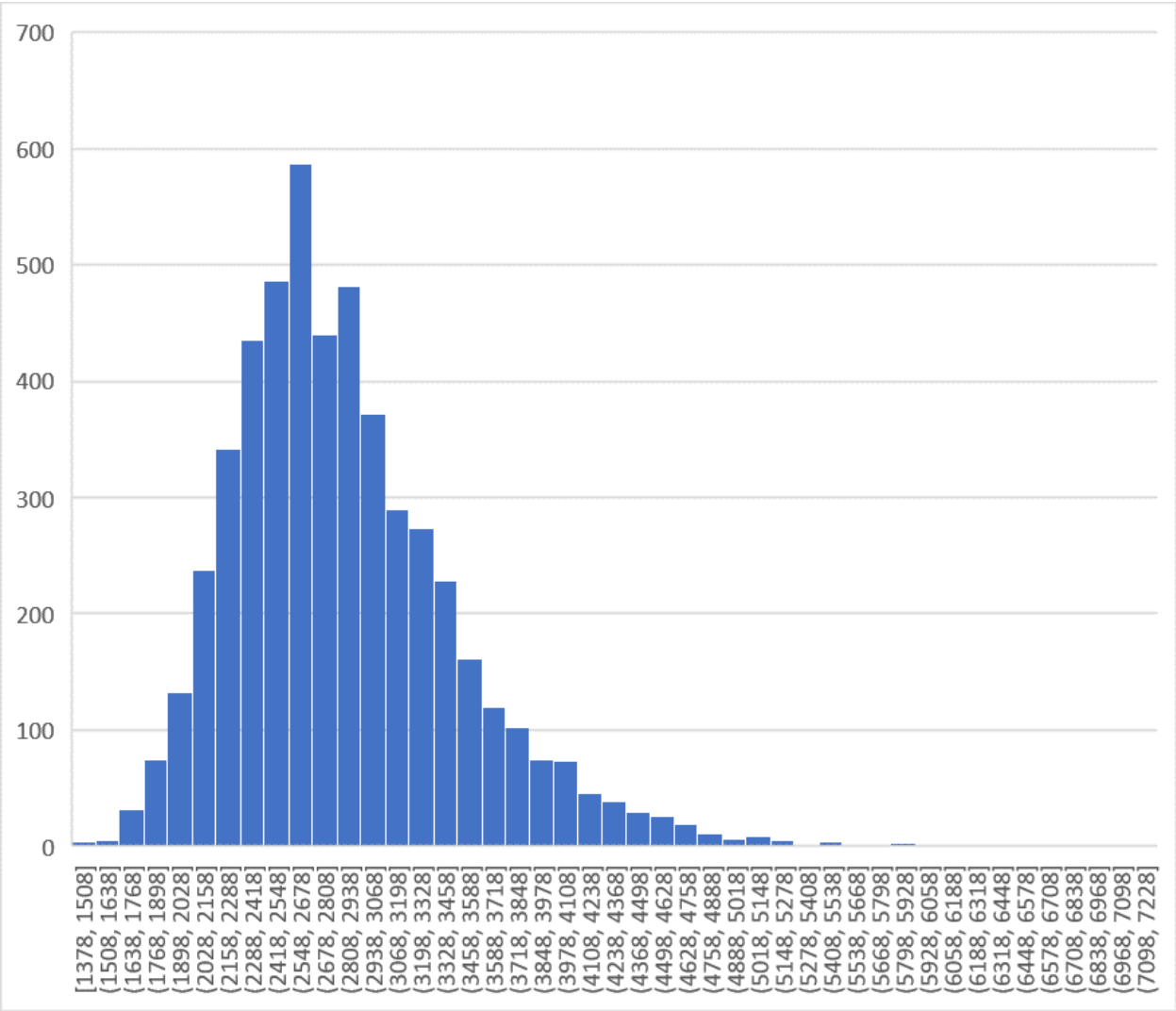
3673	0.03%
3960	0.01%
4247	0.01%
4533	0.00%
4820	0.00%
5107	0.00%
5394	0.00%
5681	0.00%
5968	0.00%
6254	0.00%
6541	0.00%
6828	0.00%
7115	0.00%

So, that we can finally get a normal distribution chart, as well as the histogram:

Graph 4: Normal distribution chart of the Computed body mass index in North Carolina



Graph 5: Histogram representing the Computed body mass index in North Carolina



The insights we get from this normal distribution is that it is very balanced, and that there are more respondents that are above the average body mass (mean), and that the standard deviation of this distribution is not big.

Target Market

After this analysis, we can define with clear evidences, based on the statistic inference on this normal distribution of the body mass index, that our specific target is as stated before, Women

45 years old or above, who live in Metropolitan areas and have an income of less than \$75k a year and that are **overweight or obese** but have exercised in the last 30 days.

So, this statistic inference helped us being more clear on the target, as we understood that costumers we want to reach have a computed body mass index between 2500 (where the overweight category starts), and 5000 (already considered obese).

Below, we detail how did we get to this conclusions:

Calculations to find the probability of having Computed body mass index within 2500-5000 in NC Population:

Mean	2850			
St.Dev	621		Z=	-0.56
X	2500		Probability	29%
Mean	2850		Z=	3.462158
St.Dev	621		Probability	99.97%
X	5000			
Probability from 2500 to 5000				71.20%

	Mathematical way	Excel	
Mean	2849.9	Confidence 16.9	
St.Dev	620.7		
n	5160		
tail	0.025		
tail area	0.975		
Z score	1.96		
Confidence	16.9		
	Confidence interval	2,833	2,867
	Standard error	8.64	

The probability of a person in North Carolina being overweight is high, this gives us an idea of a possible market. Our confidence intervals show us that we can expect people in NC to be well over the lower limit of overweight, with 95% confidence.

Conclusion

Statistical analysis of the North Carolina state population validates that there is a high amount of overweight people. This tells us there is a possibility for a business that tackles the overweight problem. The expected Body Mass Index of the population is above what is considered a normal weight BMI.