Process Owner: Ray LaForte

Decrease Sugar Intake by 50%

Key Dates: Define 1/20/20 → Measure 1/27/20 → Analyze 2/9/20 → Improve 2/19/20 → Control 3/3/20

DEFINE

Problem Statement:

The goal of this study is to decrease my daily added or processed sugar intake by 50% or consuming less than 37.5 g of sugar so I can live a more healthy lifestyle, which can have positive effects like improving quality of life, reducing the probability of chronic disease like CVD, thus increase life expectancy, and prevent taxpayers for having to pay for medical bills caused by high sugar intake.

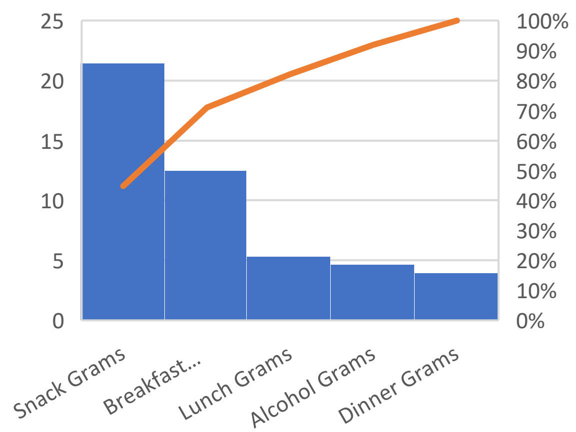
Business Impact:

The 2 week baseline measurement of the amount of added or processed sugar I consume a day is about 48 grams per day. That is above the average of about 37.5 grams per day. Reducing my added sugar consumption by 50% will have my intake well below the 37.5 grams per day threshold that the American Heart Association recommends and will reduce my risk of chronic disease like cardiovascular disease (CVD). The cost to treat a patient with CVD is about $19,000 per year and if I plan on living to 85 years old, this will save taxpayers about $1.2 million dollars if I do not develop the disease.

SQL= .4 or 14% yield

MEASURE

Pareto of Sugar Consumption



Sugar Category	Mean Weekly Sugar in Grams Consumed	Contribution in %	Cumulative %
Snack Grams	21.43	44.78%	44.78%
Breakfast Grams	12.5	26.12%	70.90%
Lunch Grams	5.36	11.20%	82.10%
Alcohol Grams	4.64	9.69%	91.79%
Dinner Grams	3.93	8.21%	100.00%
Total	47.86	100.00%	

Snack grams, Breakfast grams, and Lunch grams make up 80% of sugar consumed

ANALYZE

Hypothesis Test

Ho : u $\geq$ 37.5 versus Ha : u < 37.5

Reject H0 if tdata$\leq$ -tcrit

*t*data= 47.5-37.5/(21/(14^.5)) = 1.7817

1.7817 is not $\leq$ -1.771 so we do not reject the null hypothesis.

There is evidence at level of significance α=0.05/confidence level 95% that the sample mean of sugar consumption per day has exceeded 37.5g or more.

Multiple Linear Regression Analysis

- Sleep Time (negatively correlated)
- Miles Driven (positively correlated)

These two variables together increased my adjusted R squared value by 4% from 72.5% to 76.7% after I eliminated work time and social time from the regression analysis

SUMMARY OUTPUT	
Miles Driven & Sleep Time	
Regression Statistics	
Multiple R	0.896147682
R Square	0.803080669
Adjusted R Square	0.767277154
Standard Error	9.911402208
Observations	14

SUMMARY OUTPUT	
Miles Driven, Sleep Time, Work Time, Social Time	
Regression Statistics	
Multiple R	0.899827509
R Square	0.809689546
Adjusted R Square	0.725107122
Standard Error	10.77202455
Observations	14

IMPROVE

- SQL for Baseline: .4 or 14% yield
- SQL after Improvements: 2.3 or 78.8% yield

Mean For Baseline	Mean For Improvements
47.5 G of sugar	21.86 G of sugar

***Sugar intake decreased by more than 50%

CONTROL

- Limit Breakfast and Snack Grams of Sugar
- Aim to sleep at least 7.68 hours a night
- Limit driving to less than 25 miles per day

Define

Problem Statement- The goal of this study is to decrease my daily added or processed sugar intake by at least 50% so I can live a more healthy lifestyle, which can have positive effects like improving quality of life, reducing the probability of chronic disease like Cardiovascular Disease CVD, thus increase life expectancy, and prevent taxpayers for having to pay for medical bills caused by high sugar intake.

Business Process- The baseline measurement of the amount of added or processed sugar I consume a day is about 48 grams per day. That is above the average that is recommended by the American Heart Association to consume for an American male at about 37.5 grams per day, or 150 calories. That turns out to be 7.5% of calories for the day. Reducing my added sugar consumption by 50% will have my intake well below the 37.5 grams per day threshold that the American Heart Association recommends and will reduce my risk of chronic disease like cardiovascular disease (CVD). The cost to treat a patient with CVD is about $19,000 per year and if I plan on living to 85 years old, this will save taxpayers about $1.2 million dollars if I do not develop the disease.

SQL

Date	Y
Baseline Data	Total grams of sugar per day
Day 1 (Jan 27)	40
Day 2 (Jan 28)	90
Day 3 (Jan 29)	15
Day 4 (Jan 30)	40
Day 5 (Jan 31)	50
Day 6 (Feb 1)	40
Day 7 (Feb 2)	50
Day 8 (Feb 3)	40
Day 9 (Feb 4)	65
Day 10 (Feb 5)	10
Day 11 (Feb 6)	60
Day 12 (Feb 7)	70
Day 13 (Feb 8)	45
Day 14 (Feb 9)	50
SQL for Baseline Data	
1. Defects per day (Not eating less than 37.5g sugar)	D=1
2. Number of days	14
3. Total possible defects per day	14
4. Total actual defects	12
5. Defects per opportunity rate	12/14 = 85.71%
6. Defects per million opportunities	857,142.85
7. SQL Value	0.4 14% yield

Descriptive Statistics

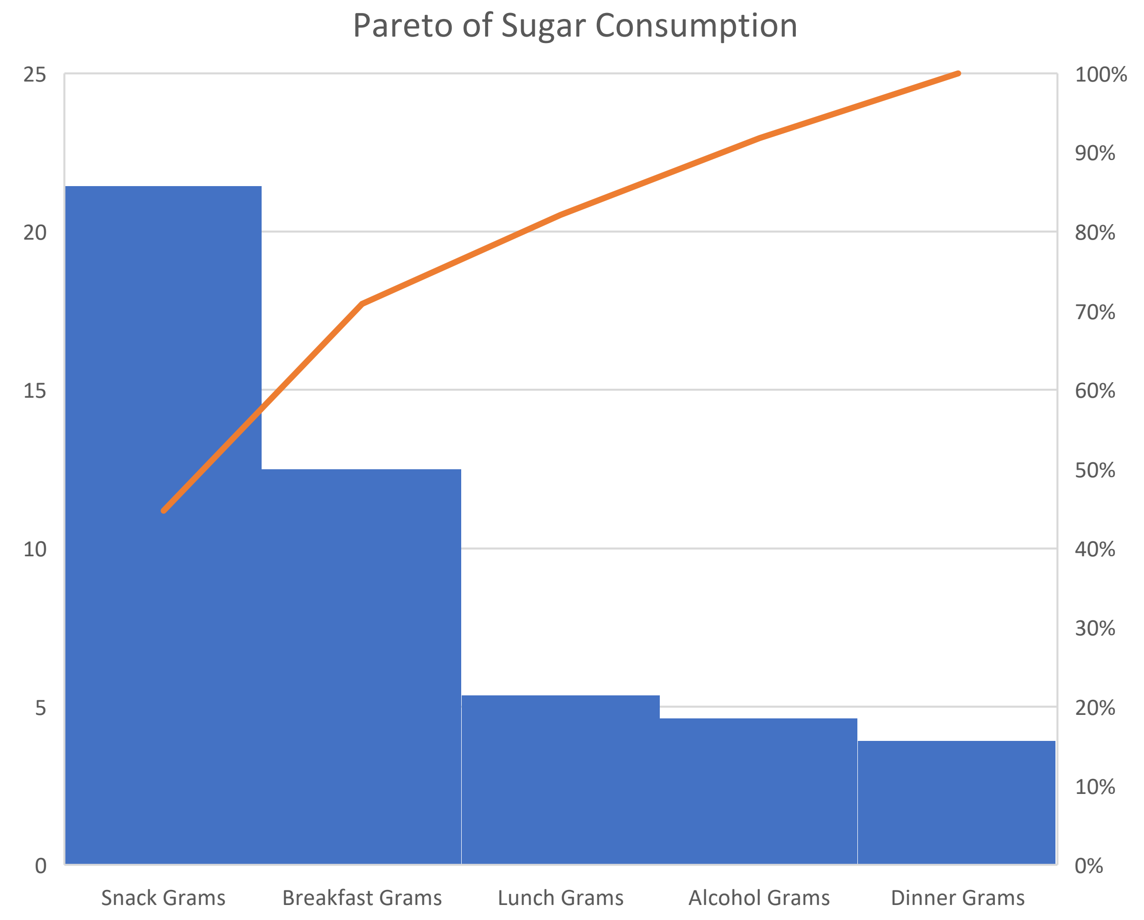
Descriptive Statistics for Baseline Data Sugar Consumption	
Mean	47.5
Standard Error	5.49100163
Median	47.5
Mode	40
Standard Deviation	20.54544681
Sample Variance	422.1153846
Kurtosis	0.790315345
Skewness	0.06984904
Range	80
Minimum	10
Maximum	90
Sum	665
Count	14

Measure- Data Measurement Plan

Performance Measure	Data Source	Type of Data	Data Collection Method	Frequency of Collection	Collection Dates
Breakfast grams of sugar	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Lunch grams of sugar	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Dinner grams of sugar	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Snack grams of sugar	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Miles Driven	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/2-3/3/20
Social Time in Mins	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Work Time in Mins	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Workout (yes/no)	Personally monitor and put into notes app	Discrete	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Weekday or Weekend	Calendar	Discrete	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Alcohol grams of sugar	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20
Sleep Time in Mins	Personally monitor and put into notes app	Continuous	Translate notes from iphone to excel	Daily collection/weekly upload	Baseline Data 1/27/20-2/9/20 Data after Improvements 2/19/20-3/3/20

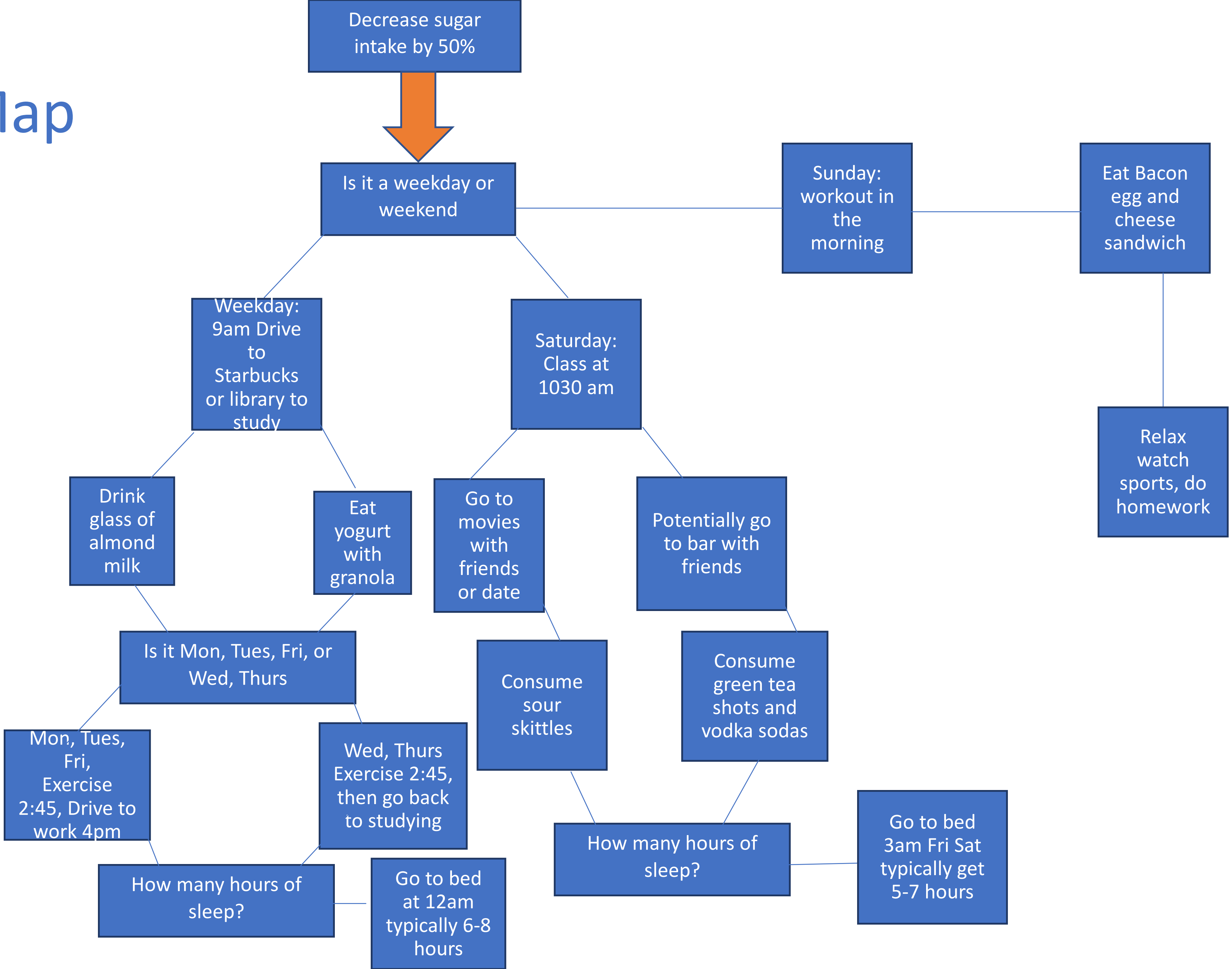
Pareto Histogram of Sugar Consumption

Sugar Category	Mean Weekly Sugar in Grams Consumed	Contribution in %	Cumulative %
Snack Grams	21.43	44.78%	44.78%
Breakfast Grams	12.5	26.12%	70.90%
Lunch Grams	5.36	11.20%	82.10%
Alcohol Grams	4.64	9.69%	91.79%
Dinner Grams	3.93	8.21%	100.00%
Total	**47.86**	**100.00%**	



Snack, Breakfast, and Lunch grams make up about 80% of sugar consumption

Process Map



Decrease sugar intake by 50%
Is it a weekday or weekend
Weekday: 9am Drive to Starbucks or library to study
Saturday: Class at 1030 am
Sunday: workout in the morning
Eat Bacon egg and cheese sandwich
Relax watch sports, do homework
Drink glass of almond milk
Eat yogurt with granola
Go to movies with friends or date
Potentially go to bar with friends
Is it Mon, Tues, Fri, or Wed, Thurs
Consume sour skittles
Consume green tea shots and vodka sodas
Mon, Tues, Fri, Exercise 2:45, Drive to work 4pm
Wed, Thurs Exercise 2:45, then go back to studying
How many hours of sleep?
Go to bed at 12am typically 6-8 hours
How many hours of sleep?
Go to bed 3am Fri Sat typically get 5-7 hours

Analysis- Hypothesis Test

- **Hypothesis Test for Baseline Data**

- A sample of 14 days of data was taken for the total sugar consumption of each one of these days. The sample mean was 47.5 grams consumed a day. Assume the sample standard deviation equals 21. I've set a goal to achieve a sample mean of less than or equal to 37.5g of sugar consumed a day

- Ho : u $\geq$ 37.5 versus Ha : u < 37.5
- Reject Ho if tdata$\leq$ -tcrit

- where u represents the sample mean total sugar consumed per day.

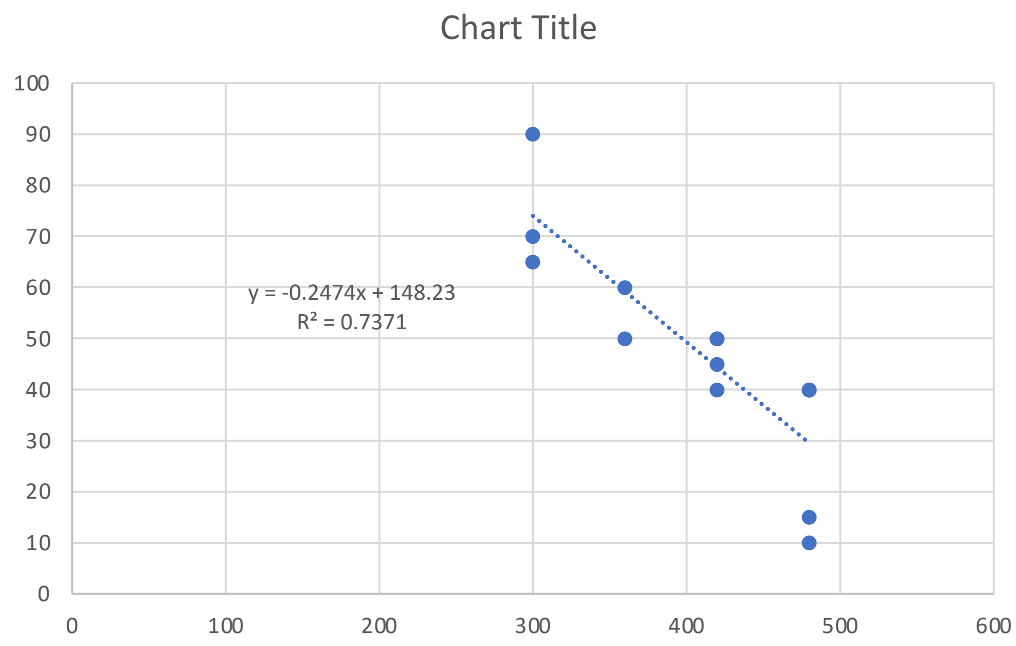
- The sample size is n = 14 so we need to verify normality. Perform t test for the mean and find t- critical. with a sample mean of x bar = 47.5 and s = 21

- To find tcrit we turn to the t table. Select the column with our α value 0.05. Then choose the row with our df=n−1=14−1=13 tcrit=1.771. The test is left tailed so the rejection rule is "Reject Ho if tdata $\leq$ -tcrit ; we will reject Ho if tdata$\leq$ -1.771

- *t*data= 47.5-37.5/(21/(14^.5)) = 1.7817
- 1.7817 is not $\leq$ -1.771 so we do not reject the null hypothesis.
- There is evidence at level of significance α=0.05 that the sample mean of sugar consumption per day has exceeded 37.5g or more.

Simple Linear Regression Analysis

Observations	Total grams of sugar Y	Sleep Time X
1	40	480
2	90	300
3	15	480
4	40	480
5	50	420
6	40	420
7	50	360
8	40	480
9	65	300
10	10	480
11	60	360
12	70	300
13	45	420
14	50	420



Running simple linear regression analysis on the continuous x variable, Sleep Time, with my increased sugar intake, Y, showed a negative correlation. The less I slept, the more sugar I consumed. The p-value was negative, extremely low, and very close to 0. Based on the regression statistics, the Correlation Coefficient value of 0.8586 indicates a strong correlation between the two variables. The Adjusted R Square value being above .7 also shows a strong model and indicates that decreased sleep time was a key factor.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.858571951
R Square	0.737145795
Adjusted R Square	0.715241278
Standard Error	38.04701905
Observations	14

ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	4045.087549	4045.087549	33.65268411	8.45885E-05			
Residual	12	1442.412451	120.2010376					
Total	13	5487.5						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	148.229572	17.60938995	8.41764379	2.22475E-06	109.8620072	186.5971367	109.8620072	186.5971367
Sleep Time X	-0.247405966	0.042648162	-5.801093355	8.45885E-05	-0.340328328	-0.154483605	-0.340328328	-0.154483605

Multiple Linear Regression Analysis

Total grams of sugar per day Y	Sleep Time in Mins X	Miles Driven X
40	480	20
90	300	40
15	480	0
40	480	60
50	420	15
40	420	0
50	360	0
40	480	20
65	300	30
10	480	15
60	360	20
70	300	30
45	420	60
50	420	30

I ran multiple linear regression with all my continuous x variables after the simple linear regression and only saw a slight increase in my adjusted R square value from 0.7152 to 0.7251. I analyzed which variables were not helping my adjusted R square and those were the ones with the highest P values. I eliminated work time and social and time and then ran another multiple regression summary and saw a much bigger increase in adjusted R square to 0.7673, which indicates a better and more improved model. I concluded that the variables more miles driven and decreased sleep time had the biggest impact together at increasing my sugar intake.

SUMMARY OUTPUT	
Miles Driven, Sleep Time, Work Time, Social Time	
Regression Statistics	
Multiple R	0.899827509
R Square	0.809689546
Adjusted R Square	0.725107122
Standard Error	10.77202455
Observations	14

SUMMARY OUTPUT	
Miles Driven & Sleep Time	
Regression Statistics	
Multiple R	0.896147682
R Square	0.803080669
Adjusted R Square	0.767277154
Standard Error	9.911402208
Observations	14

ANOVA									
	df	SS	MS	F	Significance F				
Regression	2	4406.905169	2203.452584	22.43021874	0.000131397				
Residual	11	1080.594831	98.23589373						
Total	13	5487.5							
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
Intercept	138.3373894	16.73302216	8.267328402	4.77242E-06	101.5082559	175.1665228	101.5082559	175.1665228	
Sleep Time X	-0.239424108	0.038778729	-6.174109212	6.96584E-05	-0.324775516	-0.154072701	-0.324775516	-0.154072701	
Miles Driven	0.273511665	0.142516812	1.919153693	0.081269386	-0.040165722	0.587189053	-0.040165722	0.587189053	

Residual Output for Multiple Linear Regression

RESIDUAL OUTPUT

Observation	*Predicted Total grams of sugar per day*	*Residuals*
1	28.88405068	11.11594932
2	77.45062349	12.54937651
3	23.41381738	-8.413817377
4	39.8245173	0.175482703
5	41.88193886	8.118061142
6	37.77926388	2.220736122
7	52.14471038	-2.144710378
8	28.88405068	11.11594932
9	74.71550684	-9.715506838
10	27.51649236	-17.51649236
11	57.61494368	2.385056315
12	74.71550684	-4.715506838
13	54.1899638	-9.189963797
14	45.98461384	4.015386163

Sample Size

Sample Size for Sugar Consumed (95% confidence level)

n= ((z*STDEV)/E)^2 n= 14, STDEV=21, z= 1.96, Find E

14= ((1.96*21)/E)^2

E= 11

- I went with 14 days for my baseline sample size because I had roughly a 6 week timeline to complete the project by the time I figured out my topic. I also collected another 14 days worth of data after I implemented my improvements so I'd match the exact days for the baseline data. I thought the 2 weeks (14 days) was suitable and within the scope of the project timeline. I utilized the remaining 2 weeks to construct the full project together.

Improve- SQL

SQL for Baseline Data	
1. Defects per day (Not eating less than 37.5g sugar)	D=1
2. Number of days	14
3. Total possible defects per day	14
4. Total actual defects	12
5. Defects per opportunity rate	12/14 = 85.71%
6. Defects per million opportunities	857,142.85
7. SQL Value	0.4 14% yield

SQL for Improved Data	
1. Defects per day (Not eating less than 37.5g sugar)	D=1
2. Number of days	14
3. Total possible defects per day	14
4. Total actual defects	3
5. Defects per opportunity rate	3/14 = 21.4%
6. Defects per million opportunities	214,285.71
7. SQL Value	2.3 or 78.8%

The SQL value for my process improved dramatically from a 14% yield to a 78.8% yield. I concluded after the analysis step that I had to **decrease my miles driven and increase my sleep time**. These were the two variables that had the biggest impact on my sugar consumption.

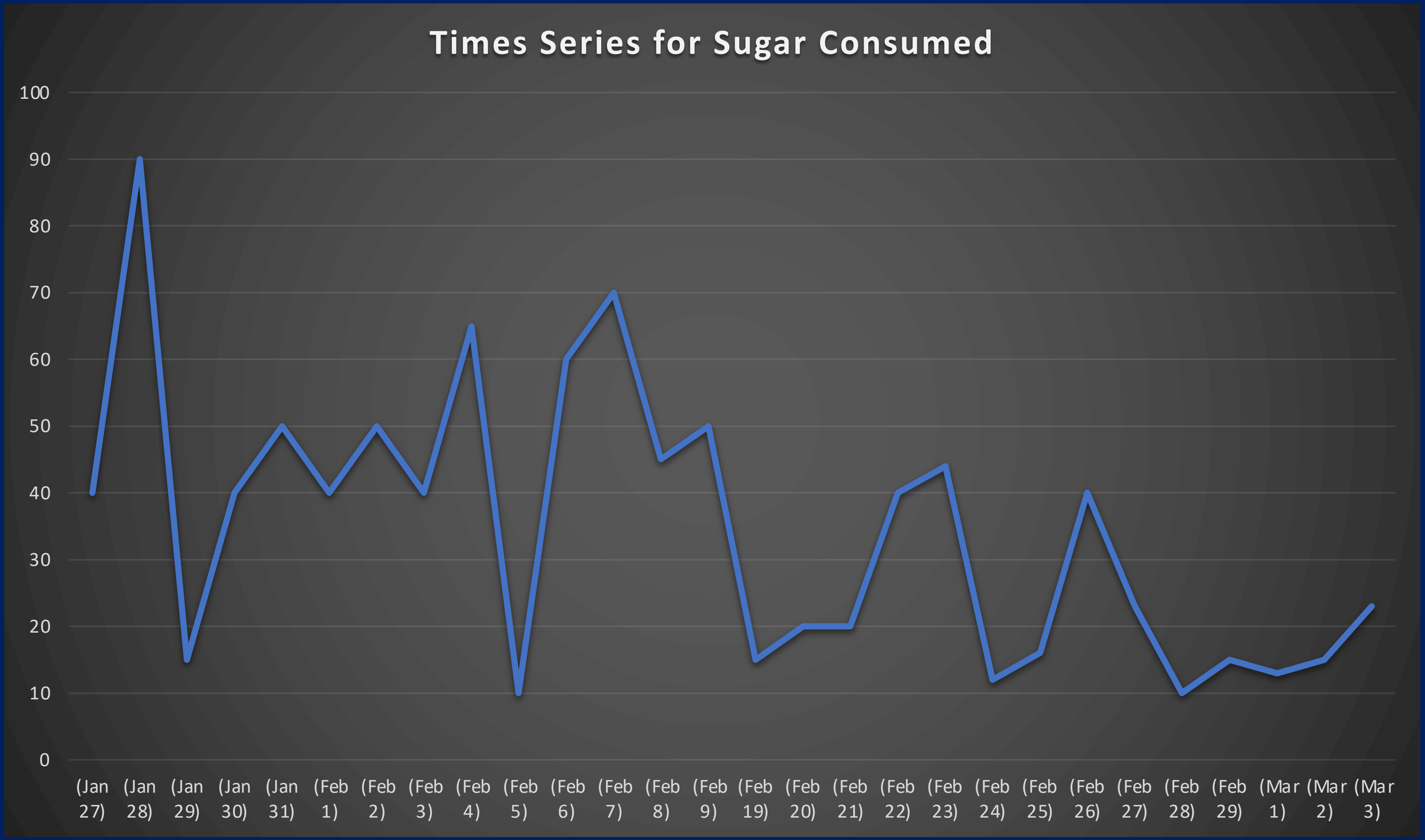
Descriptive Statistics: Before and After

Descriptive Statistics for Baseline Data	
Mean	47.5
Standard Error	5.49100163
Median	47.5
Mode	40
Standard Deviation	20.54544681
Sample Variance	422.1153846
Kurtosis	0.790315345
Skewness	0.06984904
Range	80
Minimum	10
Maximum	90
Sum	665
Count	14

Descriptive Statistics for Data after Improvements	
Mean	21.85714286
Standard Error	3.010708257
Median	18
Mode	15
Standard Deviation	11.26503879
Sample Variance	126.9010989
Kurtosis	-0.039038727
Skewness	1.149023988
Range	34
Minimum	10
Maximum	44
Sum	306
Count	14

Based on descriptive statistics, my average (mean) sugar consumption did decrease by more than 50% after I implemented my improvements from 47.5 to 21.857. I achieved my goal and if I continue to keep my sugar intake down like this, I’m on track to live a healthy lifestyle and not develop chronic diseases like CVD that could cost tax payers up to 1.2 million dollars if I live to 85 years old.

Times Series Graph



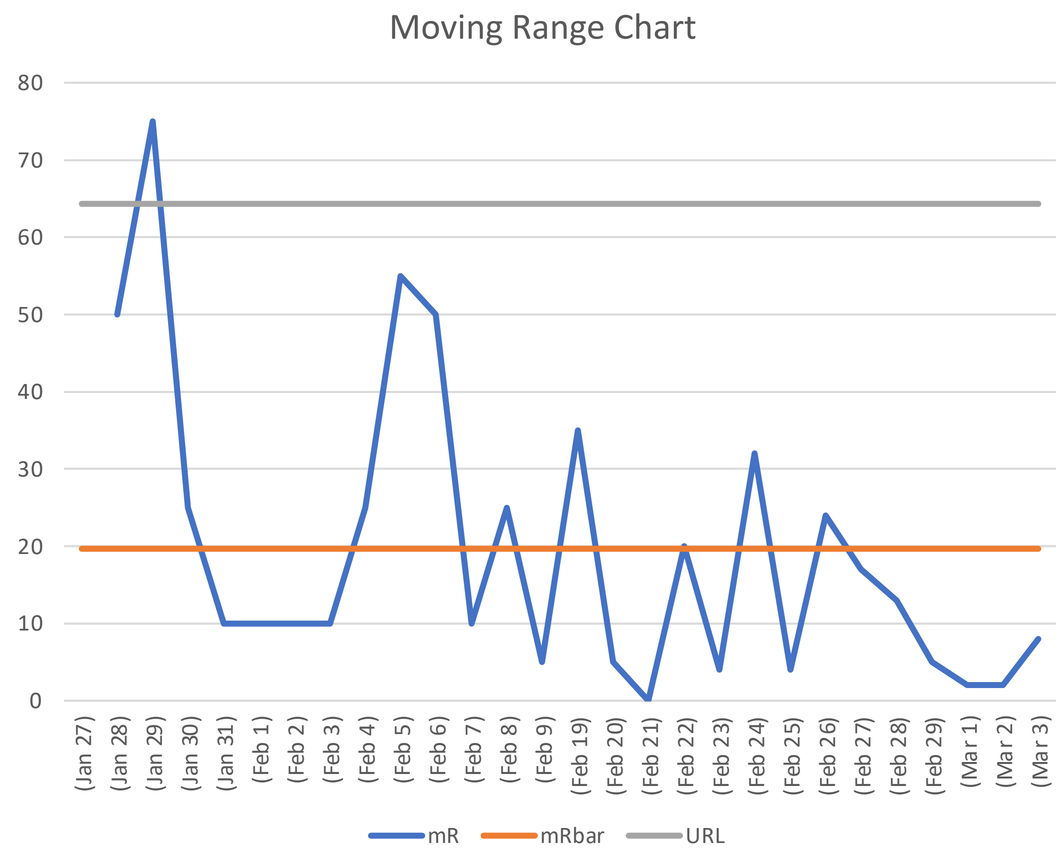
***Jan 27-Feb 9 is baseline data

***Feb 19-Mar 3 is the data after improvements

The time series graph shows lower sugar numbers from Feb 19th thereafter because of the improvements to my process. Also, the up and down spikes are less severe compared to the baseline data numbers.

Moving Range Chart with Data

Baseline Data and Improved Data	X	mR	mRbar	URL
(Jan 27)	40		19.67	64.3209
(Jan 28)	90	50	19.67	64.3209
(Jan 29)	15	75	19.67	64.3209
(Jan 30)	40	25	19.67	64.3209
(Jan 31)	50	10	19.67	64.3209
(Feb 1)	40	10	19.67	64.3209
(Feb 2)	50	10	19.67	64.3209
(Feb 3)	40	10	19.67	64.3209
(Feb 4)	65	25	19.67	64.3209
(Feb 5)	10	55	19.67	64.3209
(Feb 6)	60	50	19.67	64.3209
(Feb 7)	70	10	19.67	64.3209
(Feb 8)	45	25	19.67	64.3209
(Feb 9)	50	5	19.67	64.3209
(Feb 19)	15	35	19.67	64.3209
(Feb 20)	20	5	19.67	64.3209
(Feb 21)	20	0	19.67	64.3209
(Feb 22)	40	20	19.67	64.3209
(Feb 23)	44	4	19.67	64.3209
(Feb 24)	12	32	19.67	64.3209
(Feb 25)	16	4	19.67	64.3209
(Feb 26)	40	24	19.67	64.3209
(Feb 27)	23	17	19.67	64.3209
(Feb 28)	10	13	19.67	64.3209
(Feb 29)	15	5	19.67	64.3209
(Mar 1)	13	2	19.67	64.3209
(Mar 2)	15	2	19.67	64.3209
(Mar 3)	23	8	19.67	64.3209



Based on the moving range chart, the values were much higher and had bigger spikes with the baseline data. The values for the data after improvements were mostly below the mRbar line with less variation with the up and down movement.

Control

Limit Breakfast and Snack Grams of Sugar

Aim to sleep at least 7.68 hours a night

Limit driving to less than 25 miles per day

I've concluded that these three factors are the most important for me to control my sugar intake to keep me healthy.