Workout frequency and effectiveness

Jennifer Mead

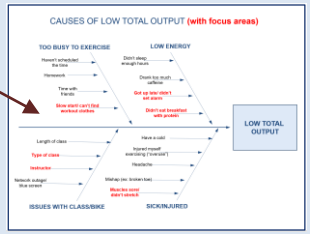
Key dates	Define 10/30/17	Measure 9/29/17	Analyze 11/20/17	Improve 11/26/17	Control 12/7/17

DEFINE

Business problems

- Can I use my new stationary bike enough to justify the expense?
- I bought the bike to improve my fitness level. Can I do that? Which behaviors and classes have the most impact on my fitness level?

Many causes, but some aren't under our control.



Fishbone diagram

Goals

- Increase **Ride Frequency**.
- Increase **Total Output** (a combination of cadence, resistance, and ride length).

MEASURE

Collected automatically by the bike

Ride Frequency

- Baseline: **17/30 (57%)**. Goal: **28/30 (93%)**.

Total Output

- Baseline: Avg of **85**. Goal: Avg of **95**.

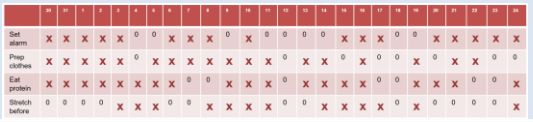
Other data

- Instructor, class type.

Collected in a checksheet

- Frequency of setting my alarm, prepping workout clothes, eating protein for breakfast, and stretching before my ride.

Checksheet



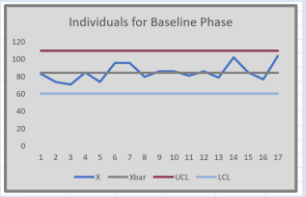
ANALYZE

Average of Total Output during Baseline	Easy	Medium	Hard	Grand Total
Female	79.0	87.4		86.4
Male	83.7	78.8	94.0	83.8
Grand Total	82.5	84.3	94.0	85.0

Early analysis suggests **"hard" classes and classes led by women** may result in higher output.

Riding every other day gives consistent results, but not much improvement. **Would riding daily help?**

Control chart



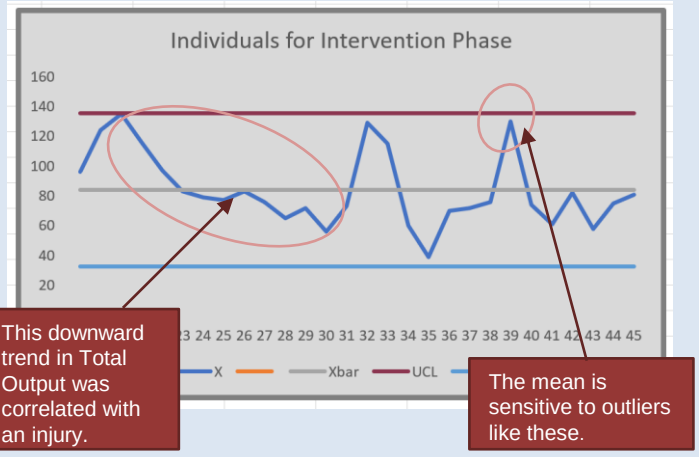
IMPROVE

Neither goal was met. The ride goal was very close, but the output goal was not.

# of Rides	Baseline	Intervention	Goal
Count	17/30	28/31	28/30
Percentage	57%	90%	93%
Goal achieved?	NO		

Total Output	Baseline	Intervention	Goal
Range	104	135	n/a
Median	74	76	n/a
Mean	85	84	95
Goal achieved?	NO		

The control chart shows that output has a downward trend with **special cause**.



CONTROL

Average of Total Output for all days	Easy	Medium	Hard	Grand Total
Female	73.7	83.1	113.0	84.1
Male	86.6	79.6	94.5	84.8
Grand Total	80.1	81.7	102.4	84.4

Next steps

- If I want to increase output, a hard class with a female instructor is my best bet.
- The injury appeared to be a confounding variable in the experiment. After recovering, consider running the experiment again. Track injuries to see if the correlation persists.

TEAM MEMBERS

Jennifer Mead

TECHNIQUES AND TESTS USED

Process map, Ishikawa diagram, Checksheet, Data stratification tree, Confidence intervals, Control charts, Hyphothesis tests (z-test), Mean/Median/Range, Pie charts, Multiple linear regression.

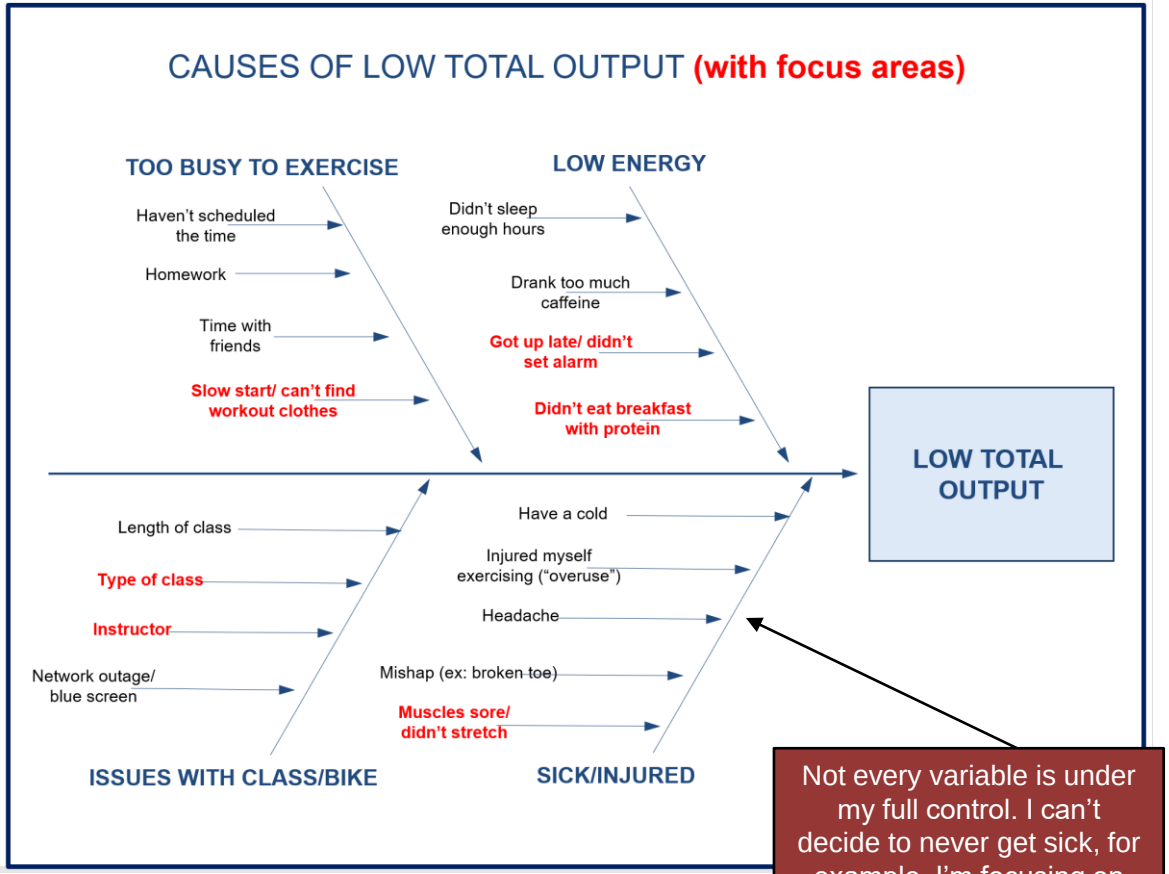
DEFINE

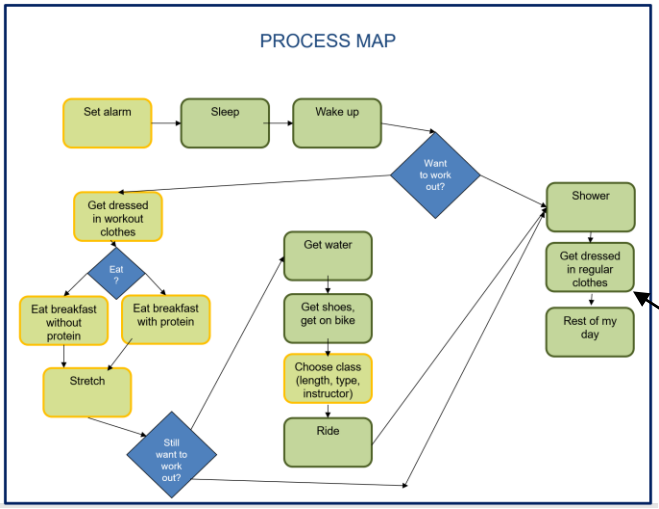
What problems do we want to solve?

I recently bought a stationary bike that includes a subscription to cycling classes, which I watch on its built-in screen. The bike and subscription are expensive. **I want to improve my level of fitness…and use the bike often enough to justify the expense.**

There are many ways I could modify my behavior to increase the chance that I would ride each day and have enough energy to ride effectively.

There are also choices for instructor and class type that might influence my performance during classes.





In the process map, some steps are consistent (wake up, shower) and others have more variation (set alarm, eat breakfast with protein).
The ones with more variation are better targets for improvement.

Not every variable is under my full control. I can't decide to never get sick, for example. I'm focusing on areas that I have the best chance of influencing.

MEASURE

Data measurement plan

Questions about the process	Performance measure	Collected by	When collected	Possible measure-ment errors	Target sample size
Can I increase how often I ride?	Proportion of days with rides, baseline compared to intervention	Bike, then downloaded in Excel	Automatically when I ride	Bike or network failure	30
Would setting my alarm help me have enough time to workout?	Count of alarm sets	I record in checksheet and transfer to summary slide	By hand, after my ride	User failure	30
Would setting aside my workout clothes before I go to bed help?	Count of prepped clothes	I record in checksheet and transfer to summary slide	By hand, after my ride	User failure	30
Can I increase my average Total Output?	Average Total Output for baseline and intervention phases	Bike, then downloaded in Excel	Automatically when I ride	Bike or network failure	30/ each phase
Would eating protein at breakfast give me more energy before/during my ride?	Count of breakfasts with protein	I record in checksheet and transfer to summary slide	By hand, after my ride	User failure	30
Would stretching before my ride help me avoid injury and increase stamina?	Count of stretching sessions before rides	I record in checksheet and transfer to summary slide	By hand, after my ride	User failure	30
Are some instructors more motivating/demanding in a way that impacts my performance?	Average Total Output per instructor Average for male/female	Bike, then downloaded in Excel	Automatically when I ride	Bike or network failure	60
Are some class types better for getting higher outputs?	Average Total Output per class, grouped by class type	Bike, then downloaded in Excel	Automatically when I ride	Bike or network failure	60

Measurement errors are possible! To reduce them, I have a recording system set up in advance, and double-check values before calculations.

If the bike doesn't record data correctly, or if it changes how it calculates output, there is little I can do.

Thirty is a large sample size, but since we are trying to detect a small change, we may not have a big enough sample to support the margin of error we want to use.

ANALYZE

Rides: Starting state & definition of success

- **Goal:** During the baseline period, I rode 17/30 days. **My goal is to ride 28/30 days during the intervention period (93%).**
- **Operational definition:** The bike records rides automatically. A ride must be recorded by the bike; I must be the rider; the ride must be at least 15 minutes.
- **Data type:** Discrete data; each day I ride or don't ride.
- **S.Q.L.:** Sigma Quality Level for 17/30 days is about 1.6. If I achieve my goal of 28/30, the S.Q.L. will be 3. (See at right.)
- **Sample size:**

Sigma Quality Level for the current process

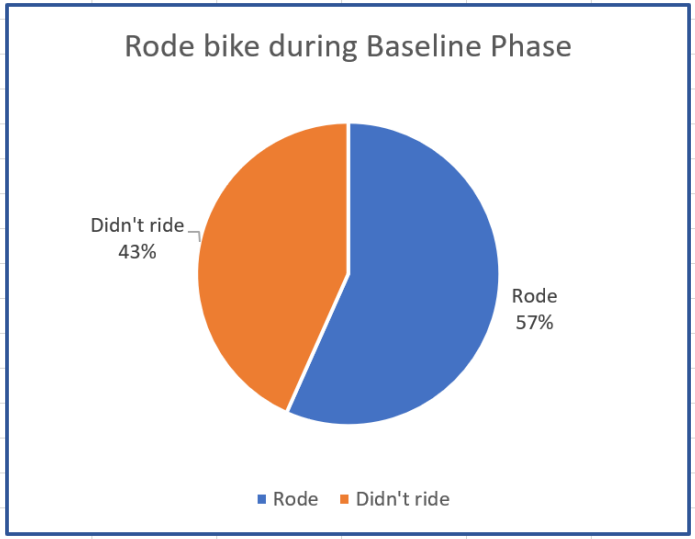
Baseline: 17 rides in 30 days		
Defect opportunity per unit	D	1
Units made per period	U	30
Total possible defects per period	D x U	30
Total actual defects in this period	A	13
Defect opportunity rate (DPO)	A/DU * 100	0.4
Defects per million opportunity	DPO * 1000000	433,333
SQL value from SQL table for DPMO	<look up>	between 1.6 and 1.7

The goal is an S.Q.L. of 3.

For discrete data (days rode bike)	
zstar confidence level	1.96
p (yes)	17
1-p (no)	13
E margin of error (choose)	3
sample size for ideal conf level (.95) and ideal margin of error (3)	
n = ((zstar^2)*p*(1-p)/E^2	94
at conf level 0.95, margin of error for actual sample size of n = 30	
E = sqrt(((z^2)*p*(1-p))/n)	5

Ideally, we'd have a sample size of at least 94 days. That's not possible in our pilot, so we'll need to make some compromises on margin or error.

With a sample of 30, and a confidence level of 95%, the best margin of error we can get is 5.



ANALYZE

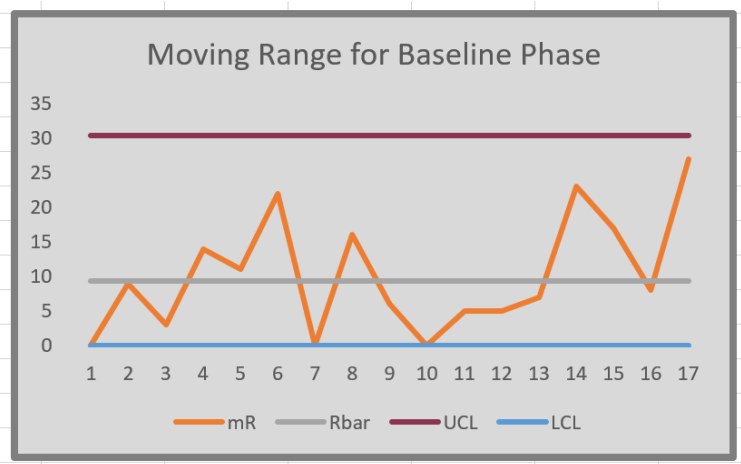
Total Output: Starting state & definition of success

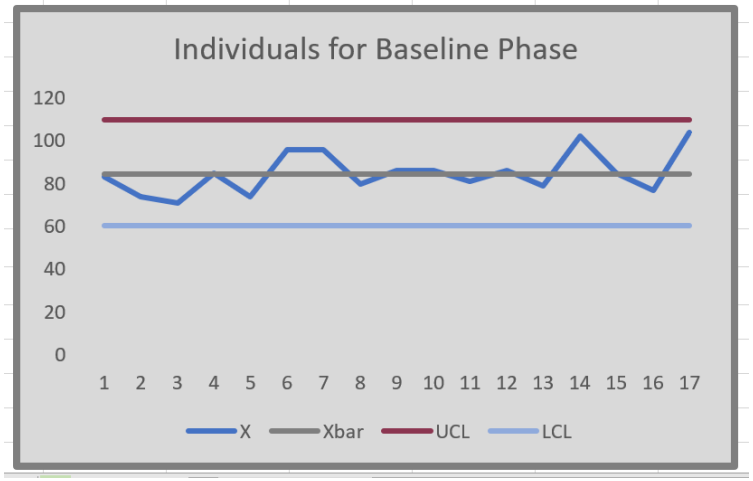
- **Goal:** During the baseline period, my average Total Output was 85. **My goal is to have an average Total Output of 95 during the intervention period.**
- **Operational definition:** The bike calculates Total Output automatically for each ride. A ride must be recorded by the bike; I must be the rider; the ride must be at least 15 minutes.
- **Data type:** Continuous data. Total Output is calculated from a combination of cadence, resistance, and ride length. I'm comparing average Total Output across two 30-day periods.
- **S.Q.L.:** Sigma Quality Level doesn't apply here, because there is no "right" goal for output. It varies by rider and over time.
- **Sample size:**

For continuous data (total output)		
zstar confidence level	1.96	
sigma hat std dev	40.93	
E margin of error (choose)	3	
sample size for ideal conf level (.95) and ideal margin of error (3)		
n = ((z*sigma)/E)^2	715	
at conf level 0.95, margin of error for actual sample size of n =	60	30
E = (z * sigma)/sqrt(n)	10	15
E if use 90% conf	9	12
E if use 80% conf	7	10

Ideally, we'd have a sample size of at least 715 days. That's not possible in our pilot, so we'll have to make trade-offs with confidence level and margin of error.

This is a challenge. With our sample size and a confidence level of 95%, the best margin of error we can get is 10. But since our start (85) and end (95) are within 10 of each other, the results won't tell us as much as we'd hoped





IMPROVE

Did the proportion of rides increase?

Yes, the proportion did increase, and quite dramatically. It did not increase as much as we intended, so we did not achieve the goal.
In practice, the results may be close enough.

# of Rides	Baseline	Intervention	Goal
Count	17/30	28/31	28/30
Percentage	57%	90%	93%
Goal achieved?	NO		

Goal: Increase proportion of rides to 28/30 (0.93)

Hypotheses

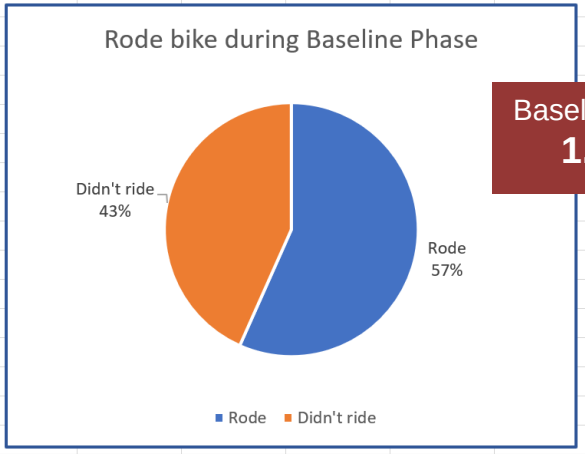
- H0: The proportion of rides is < 0.93
- HA: The proportion of rides is >= 0.93
- Alpha is 0.05
- Confidence level is 95%, margin of error is 5

Test

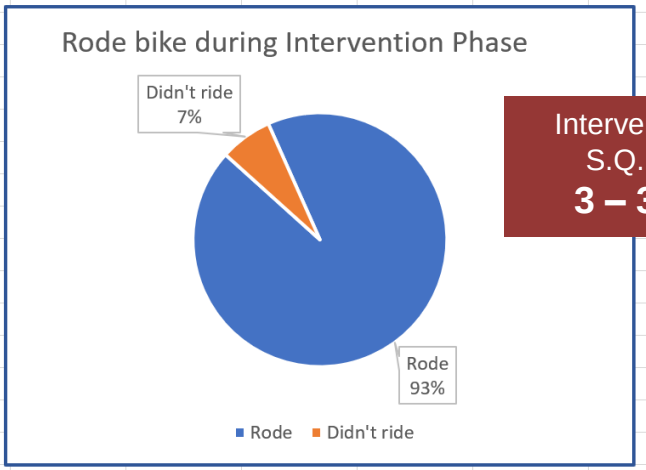
- Data is discrete, one sample against a standard
- It's an upper-right-tailed test, using the formula
  =1-(NORM.S.DIST(z, TRUE))
- Sample size is large (31)
- Population Mu ("the standard") is .93
- P number of successes is 28
- 1-p number of failures is 3
- Z = -0.67
- P-value = 0.75

Conclusion

The p-value of 0.75 is greater than the alpha of 0.05. We do not reject the null hypothesis. At a confidence level of 95% and a margin of error of 5, there is not convincing evidence that the proportion of rides is >= 0.93.



Baseline S.Q.L.:
1.6-1.7



Intervention
S.Q.L.:
3 – 3.1

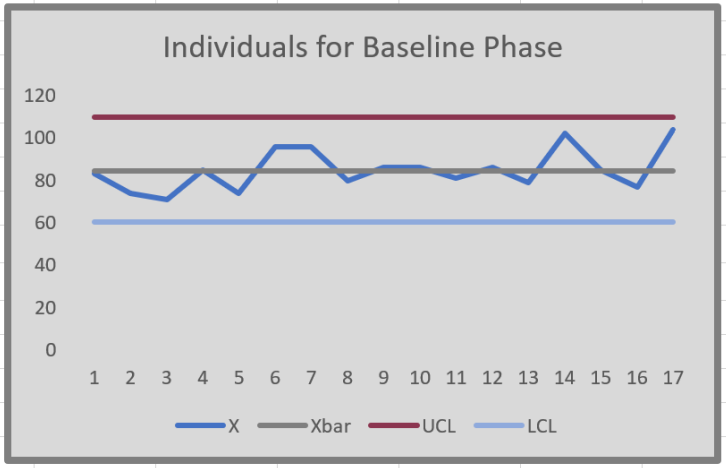
IMPROVE

Did the Total Output increase?

No, it did not increase. It went down.

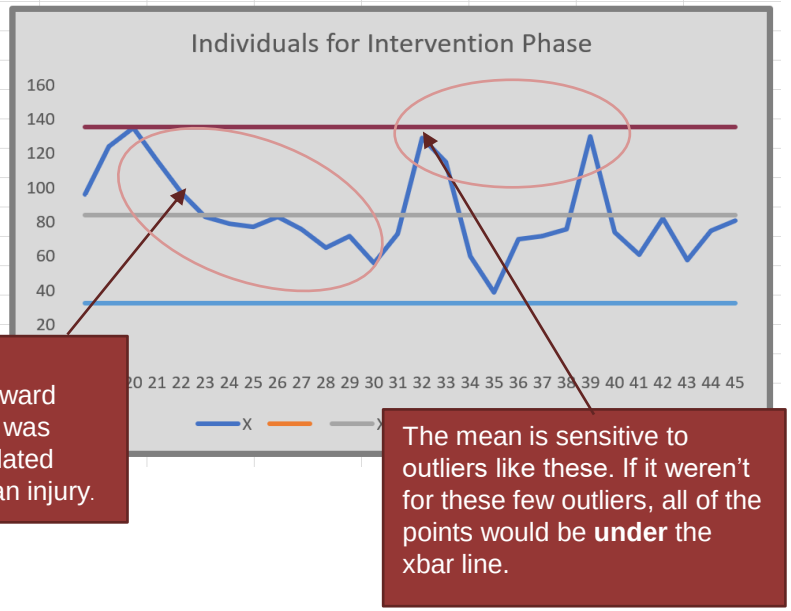
Total Output	Baseline	Intervention	Goal
Range	104	135	n/a
Median	74	76	n/a
Mean	85	84	95
Goal achieved?	NO		

Goal: Increase Total Output from 85 to 95



Hypotheses

- H0: The Total Output in the Intervention phase is <95.
- HA: The Total Output in the Intervention phase is >= 95.
- Alpha is 0.05.



This downward trend was correlated with an injury.

The mean is sensitive to outliers like these. If it weren't for these few outliers, all of the points would be **under** the xbar line.

Test

- Data is continuous, one sample against a standard
- It's an upper-right-tailed test, using the formula
  =1-(NORM.S.DIST(z, TRUE))
- Sample size is large (31)
- Population Mu ("the standard") is 95
- Sample mean xbar is 76
- Standard deviation sigma is 34
- Z = -3.11
- P-value > 0.99

Conclusion

With a confidence level of 95% and a margin of error of 15, we fail to reject H0. There is not convincing evidence that the mean is greater than or equal to 95.

CONTROL

Can we predict future performance?

For Number of rides: **NO**

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.446181
R Square	0.199078
Adjusted R Square	0.141869
Standard Error	0.410868
Observations	61

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	2.34977	0.587442	3.479849	0.013133235
Residual	56	9.453509	0.168813		
Total	60	11.80328			

	Coefficients	*andard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	0.579941	0.071716	8.086684	5.57E-11	0.436277527	0.723605	0.436278	0.723605
AlarmCount	0.276249	0.13598	2.031536	0.046956	0.003847599	0.548651	0.003848	0.548651
ClothesCount	0.046221	0.150419	0.307279	0.759771	-0.255105027	0.347546	-0.25511	0.347546
ProteinCount	-0.06475	0.159324	-0.40638	0.686012	-0.383910785	0.254418	-0.38391	0.254418
StretchCount	0.301107	0.144452	2.084484	0.041689	0.011735676	0.590478	0.011736	0.590478

The low adjusted R Square (.14) tells us that this multiple linear regression model isn't a good model for number of rides.

But this high R Squared (.82) indicates that this multiple linear regression model is useful for predicting Total Output.

For Total Output: **YES**

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.921757076
R Square	0.849636108
Adjusted R Square	0.821038687
Standard Error	16.56579451
Observations	61

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	86836.43	17367.29	79.10745956	7.32853E-24
Residual	56	15367.83	274.4255		
Total	61	102204.3			

	Coefficients	*andard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	-7.1054273576E-15	4.141449	-1.7E-15	1	-8.29631852	8.296319	-8.29632	8.296319
Easy	-22.33343596	7.676732	-2.90924	0.00518784	-37.7117776	-6.95509	-37.7118	-6.95509
Medium	-20.86530172	7.157723	-2.91508	0.005104497	-35.2039444	-6.52666	-35.2039	-6.52666
Hard	0	0	65535	1.7892E-222	0	0	0	0
Male	102.1422414	7.804982	13.0868	1.12541E-18	86.50698413	117.7775	86.50698	117.7775
Female	102.8103448	8.029094	12.80473	2.8234E-18	86.72613731	118.8946	86.72614	118.8946

Regression equation: y = 0 - 22.2(easy) - 20.9(medium) + 102.1(male) + 102.8(female)

Low P-values indicate that these variables contribute to the regression equation. Include their co-efficients to build the equation.

Regression equation:
Y = 0 – 22.2(easy) – 20.9 (medium) + 102.1 (male) + 102.8(female)

A "hard" class led by a female instructor gets the highest result.

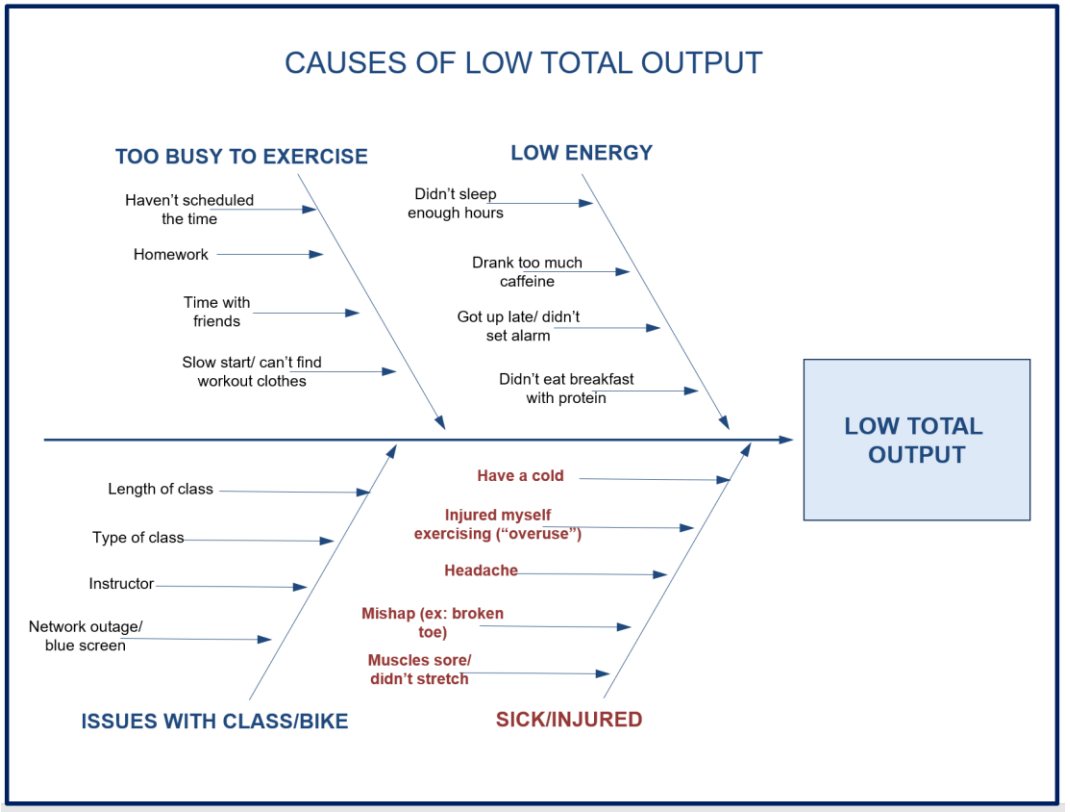
CONTROL

Next steps

- To increase my Total Output, my best bet is to **take "hard" classes with female instructors**.
- We ran into issues with sample size in the pilot (even though we had 30 or 60 samples for each test). **In the next phase, we should aim to collect data for at least three months.** This would decrease our margin of error without sacrificing confidence level.
- Since we think that injury was a confounding variable in the process, in the next phase we should **track the various forms of injuries** from the fishbone diagram.
- Before starting the next testing phase, it's smart to **let the current injury heal**.

Average of Total Output for all days

	Easy	Medium	Hard	Grand Total
Female	73.7	83.1	113.0	84.1
Male	86.6	79.6	94.5	84.8
Grand Total	80.1	81.7	102.4	84.4



This is the same fishbone diagram from the Define phase.