

What Statistical Method is Best for My Project

Christopher J. Malone

cmalone@winona.edu

statsclass.org/isef



DATA
IS THE **NEW** OIL

Source: Clive Humby

DATA
IS THE NEW OIL

but do you have the resource to refine it?

Getting Oil



Getting Oil



Middle School



High School



Getting Oil

**Mentored
Projects**



What does it mean to refine oil?



WIKIPEDIA
The Free Encyclopedia

[Main page](#)

[Contents](#)

[Featured content](#)

[Current events](#)

[Random article](#)

[Donate to Wikipedia](#)

Article [Talk](#)

Oil refinery

From Wikipedia, the free encyclopedia

An **oil refinery** or **petroleum refinery** is an **industrial process plant** where **crude oil** is processed and refined into more useful products such as **petroleum naphtha**, **gasoline**, **diesel fuel**, **asphalt base**, **heating oil**, **kerosene**, and **liquefied petroleum gas**.^{[1][2]}

What does it mean to refine oil?

Crude Oil

Refinery



- Gasoline
- Heating Oil

What does it mean to refine oil?

Crude Oil

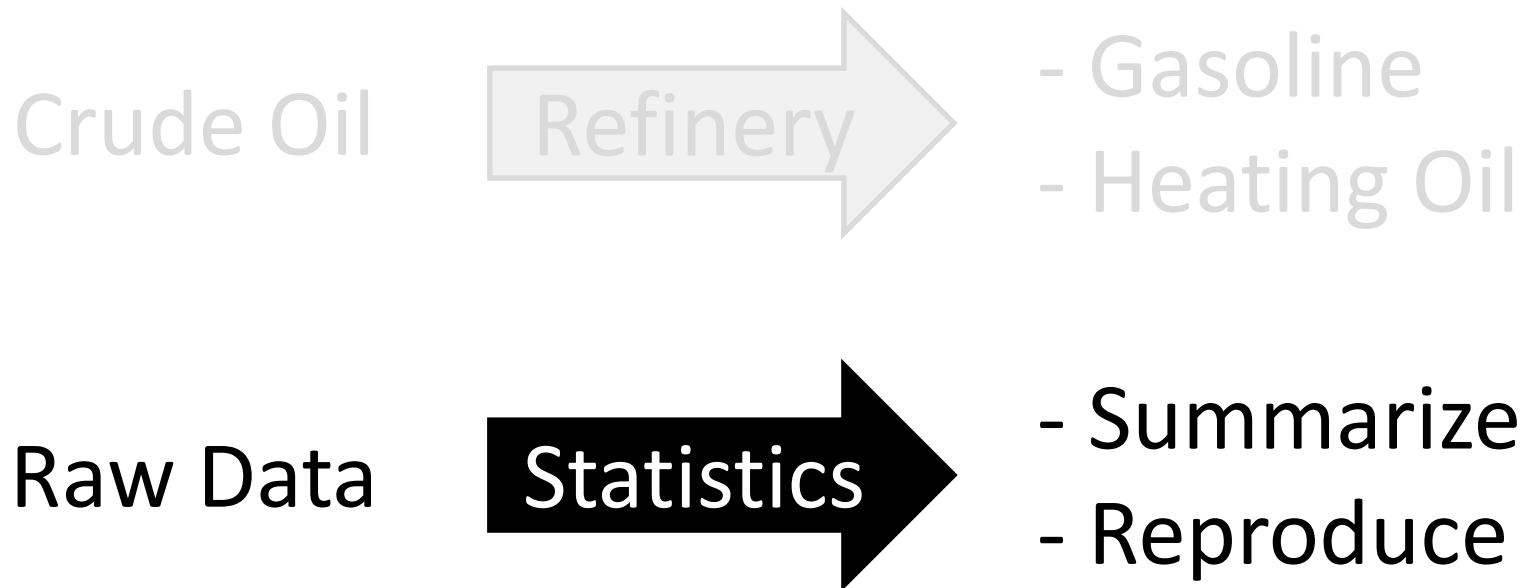
Refinery



- Gasoline
- Heating Oil

What does it mean to refine data?
statsclass.org/isef

What does it mean to refine ~~oil~~?
data?



Minimum Tools for Middle School

- › Methods of Data Collection / Study Design
- › Descriptive Summaries
- › Data Visualization
- › Limitations of Data / Study



Minimum Tools for High School

- › Effective Data Summaries and Visual Displays
- › Methods for ensuring reproducible outcomes, i.e. significance tests, margin-of-error
- › Suggestions for improved study design



Example #1: Energy Efficiency

Roof
vs.
Roof and Walls

Pair	Temperature Change (Change = Inside - Outside)	
	Green Roof	Green Roof and Wall
1	23.5	31
2	25.5	33.5
3	28.25	35.5
4	30	34.5
5	29.5	35
6	22.5	29.5
7	30.25	32
8	26.75	33
9	29.25	34.5

Units of Measurement: °F

Example #1: Initial Thoughts

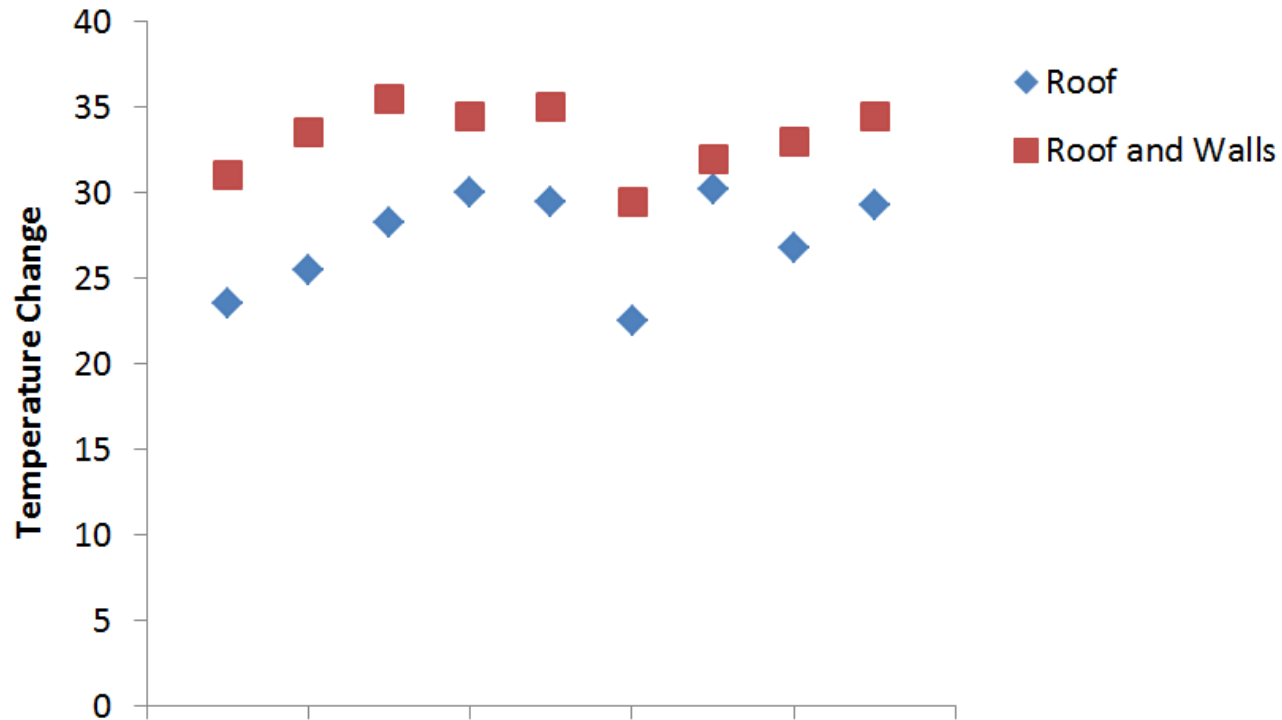
Your Observations

- 1.
- 2.
- 3.

Pair	Temperature Change (Change = Inside - Outside)	
	Green Roof	Green Roof and Wall
1	23.5	31
2	25.5	33.5
3	28.25	35.5
4	30	34.5
5	29.5	35
6	22.5	29.5
7	30.25	32
8	26.75	33
9	29.25	34.5

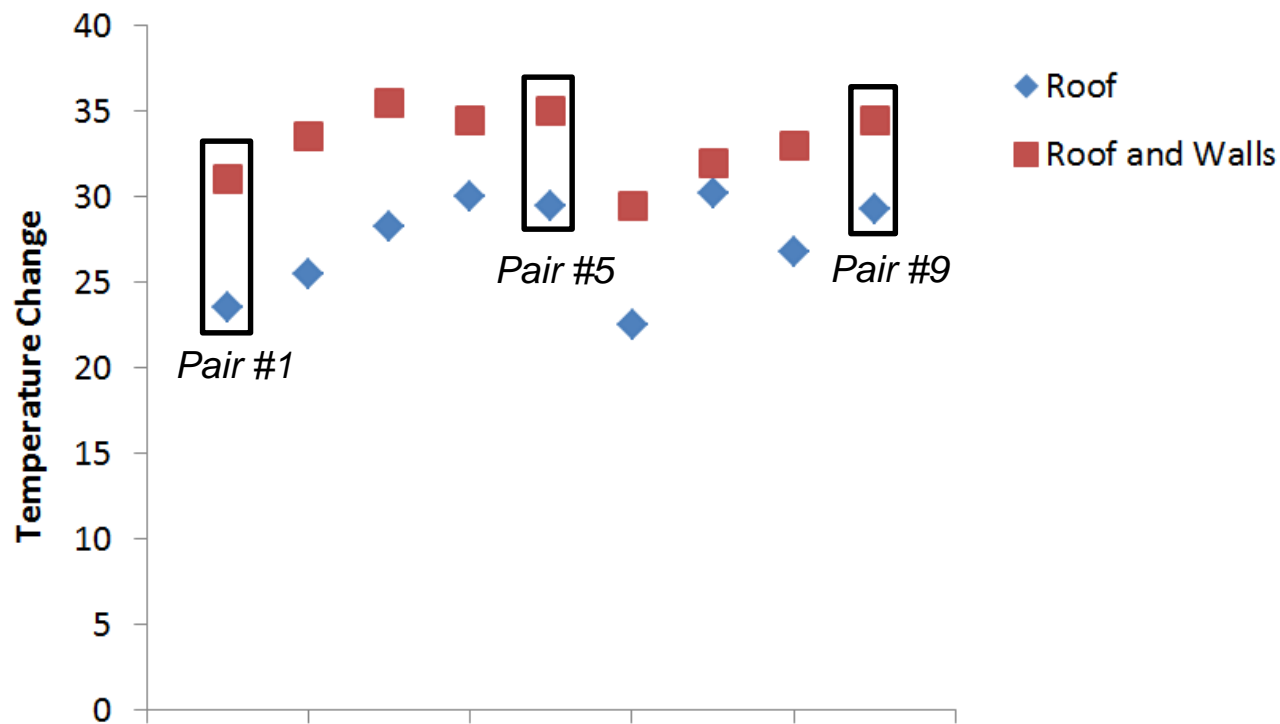
statsclass.org/isef

Example #1: Outcomes Presented



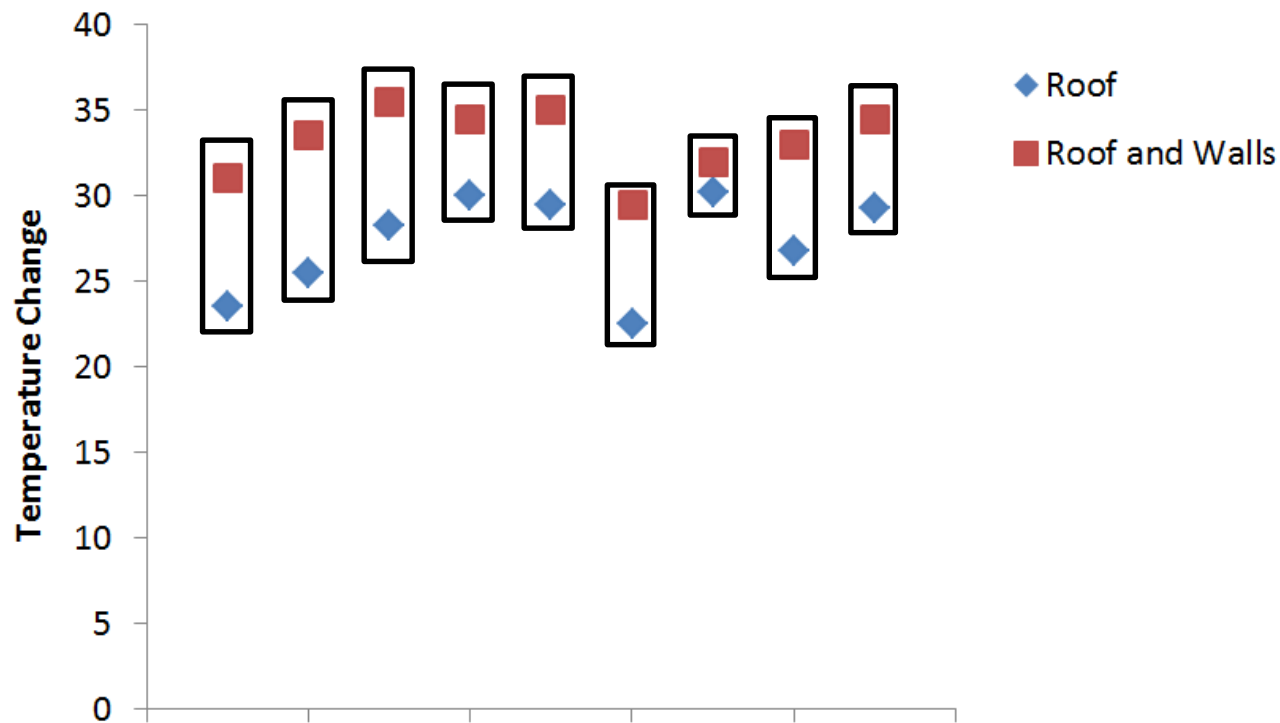
Difference between treatments is statistically significant ($p < 0.01$, paired t-test, $n = 9$ pairs).

Example #1: Outcomes Presented



Difference between treatments is statistically significant ($p < 0.01$, paired t-test, $n = 9$ pairs).

Example #1: Outcomes Presented



Difference between treatments is statistically significant ($p < 0.01$, paired t-test, $n = 9$ pairs).

Example #1: Refined

- › Direct comparisons are of primary interest

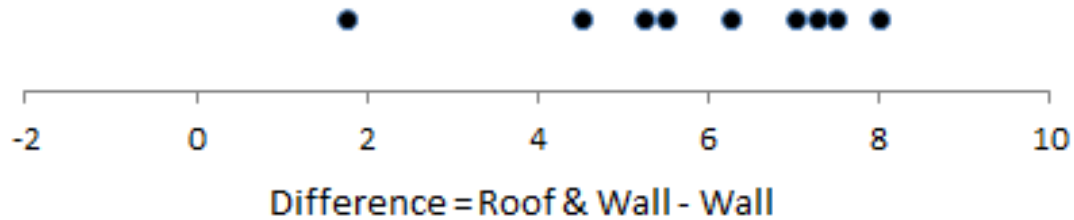
Pair	Temperature Change (Change = Inside - Outside)		Difference
	Green Roof	Green Roof and Wall	
1	23.5	31	7.5
2	25.5	33.5	8
3	28.25	35.5	7.25
4	30	34.5	4.5
5	29.5	35	5.5
6	22.5	29.5	7
7	30.25	32	1.75
8	26.75	33	6.25
9	29.25	34.5	5.25

Questions:

- What information is gained?
- What have we lost?

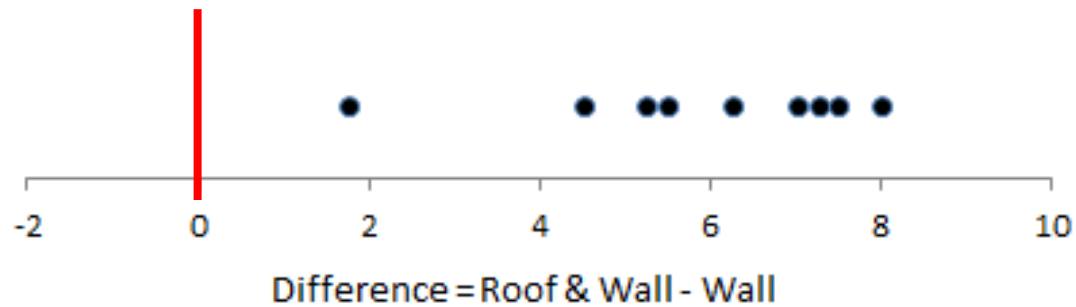
Example #1: Refined

- › Visual Depictions of Differences



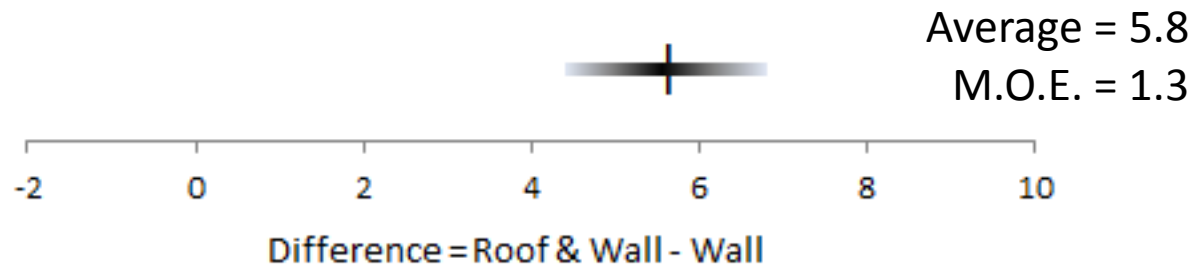
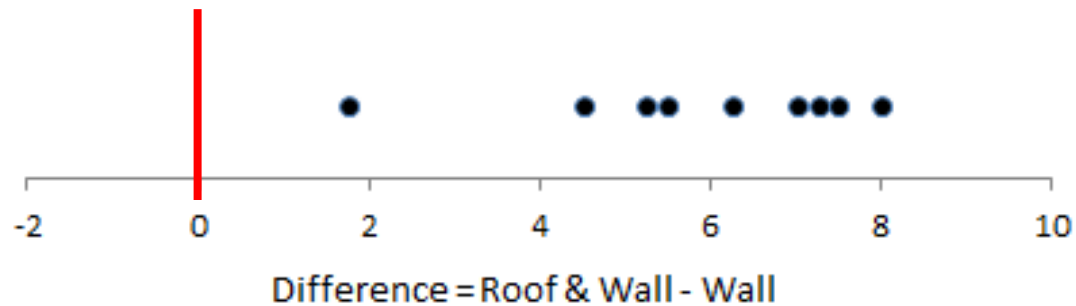
Example #1: Refined

> Visual Depictions of Differences



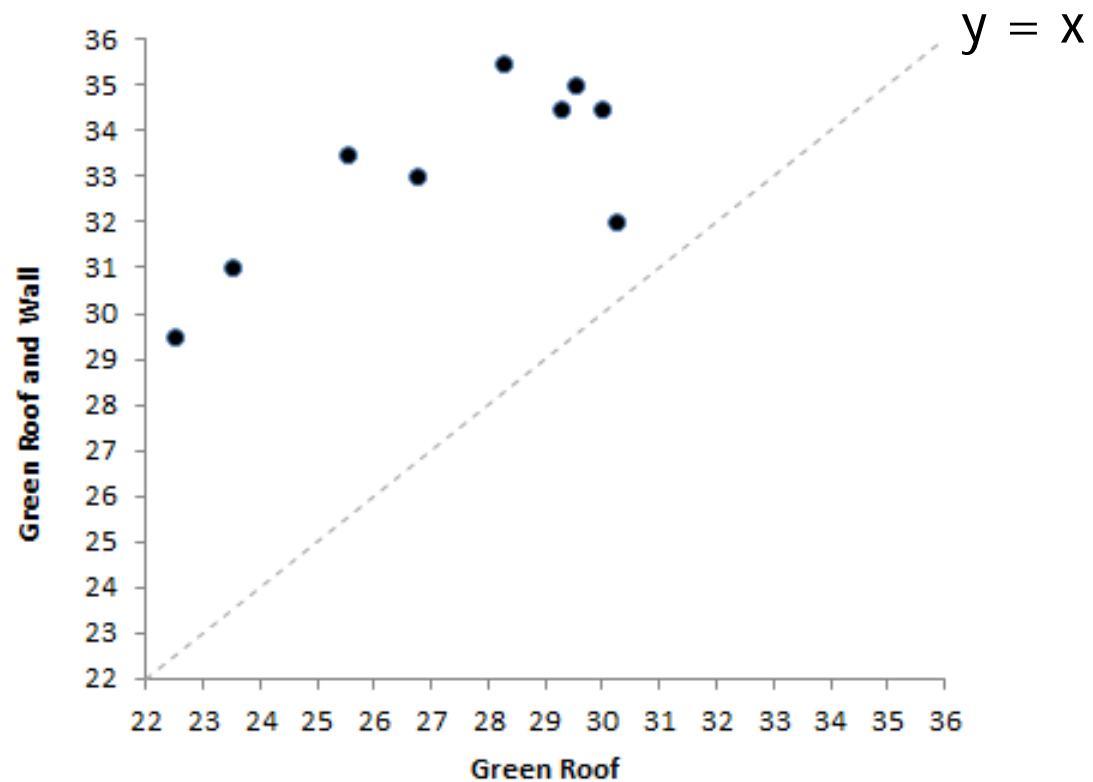
Example #1: Refined +

> Visual Depictions of Differences



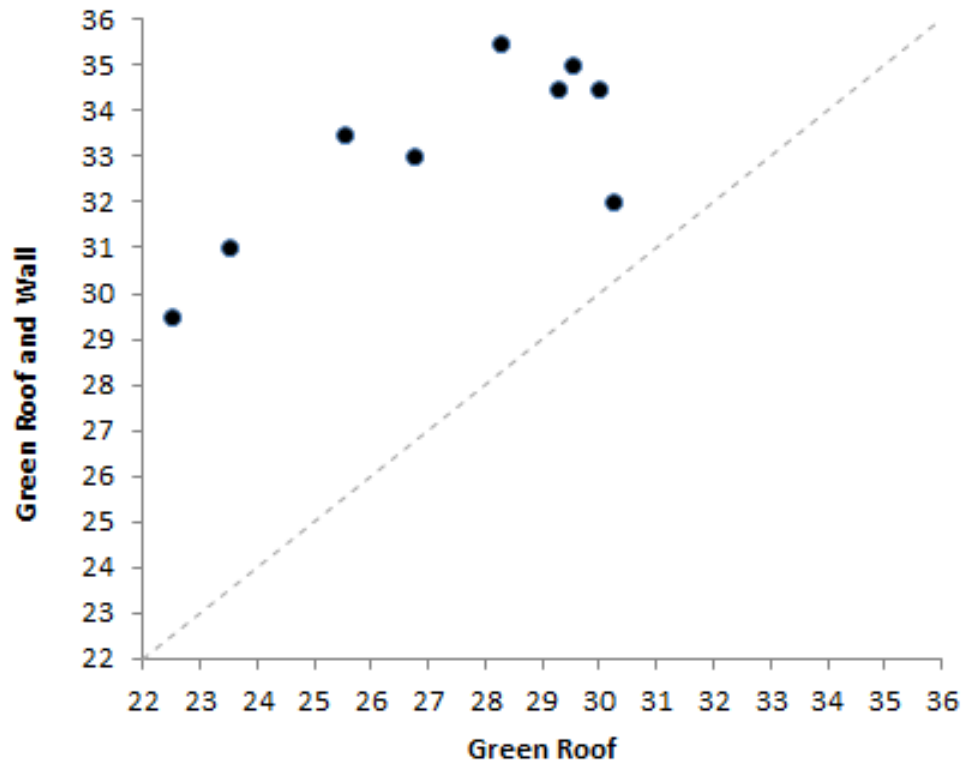
Example #1: Refined +

> Y vs. X Plot



Example #1: Refined +

- Y vs. X Plot (with $Y = X$ line)

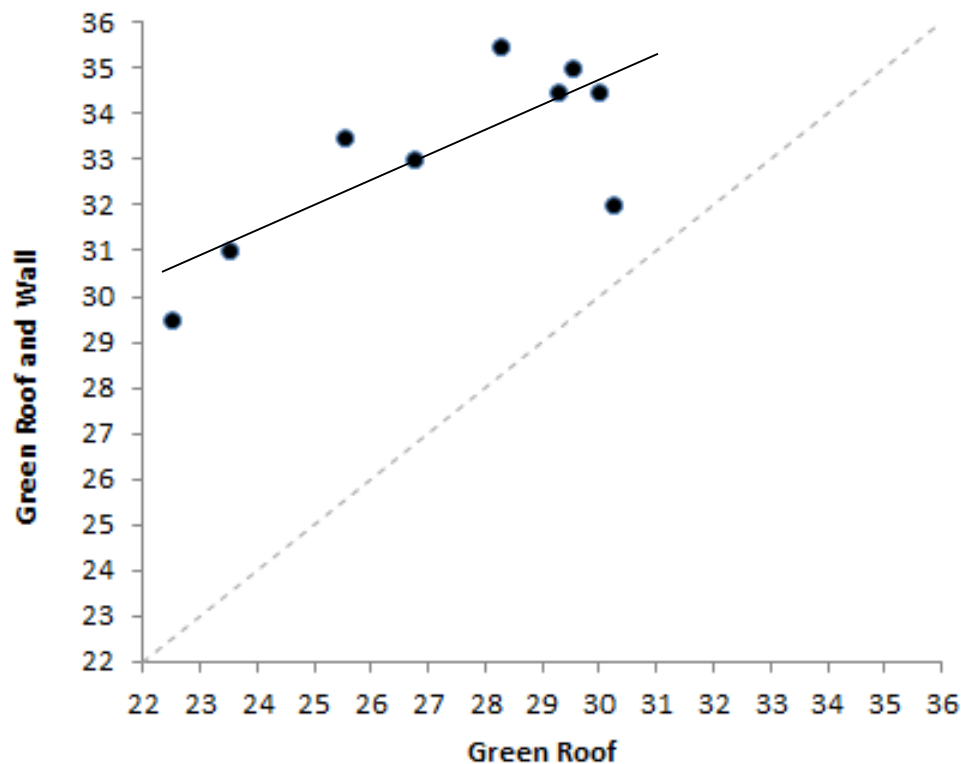


Observations

1. Most points above $y=x$ line, thus walls provide additional benefit (about 6 degrees)

Example #1: Refined +

- Y vs. X Plot (with $Y = X$ line and Trend line)

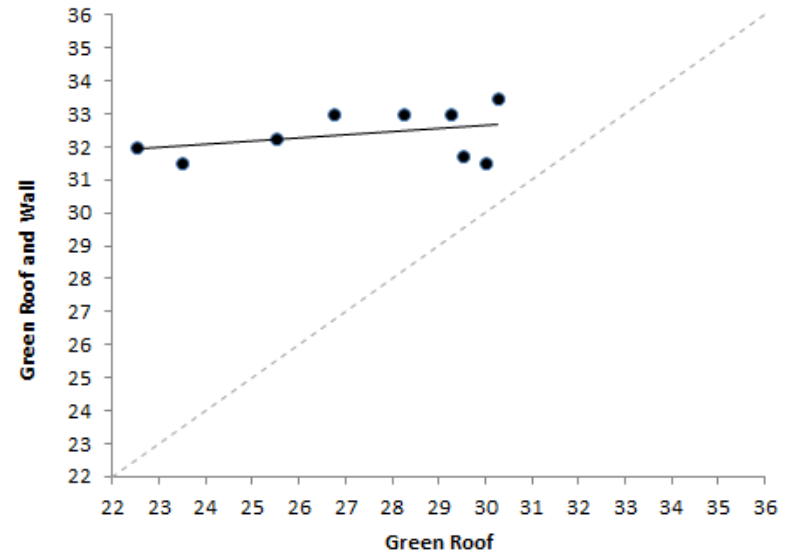
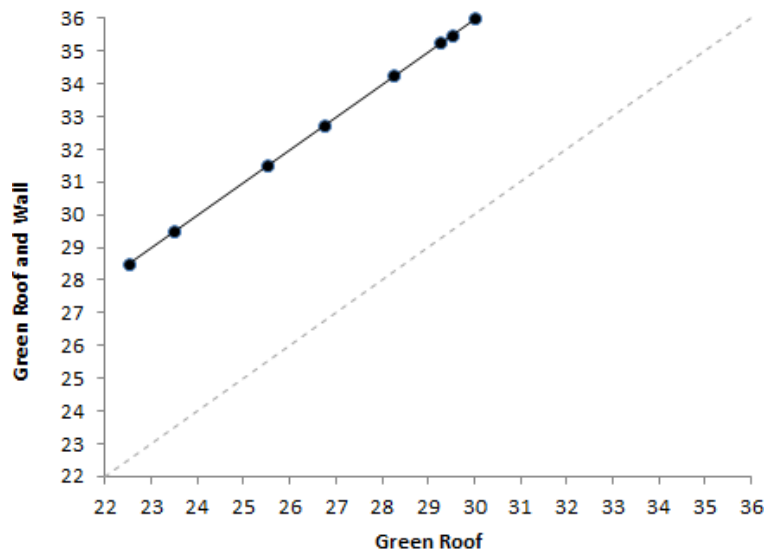


Observations

1. Most points above $y=x$ line, thus walls provide additional benefit (about 6 degrees)
2. This benefit is reduced some when (Inside – Outside) temperature is increased

Example #1: Refined +

› Some What Ifs...



Example #1: Refined +

- > Y vs. X Plot (with $Y = X$ line and Trend line)

Contrast

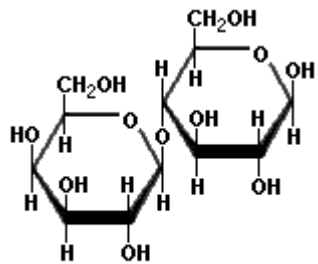
Information gained here to

Difference between treatments is statistically significant ($p < 0.01$, paired t-test, $n = 9$ pairs).

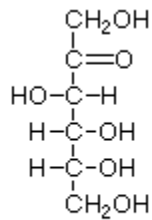
Green Roof and Wall

Green Roof

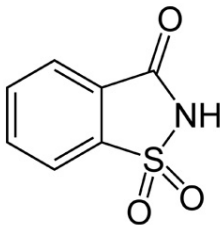
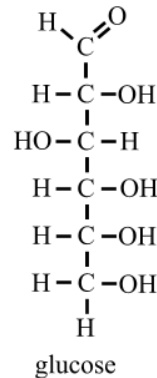
Example #2: Heart Rates of Daphnia



Lactose



D-Fructose

C₇H₅NO₃S Saccharin

Sugar Type	Time Periods				
	Start	5 min	10 min	15 min	20 min
Glucose	300	357	352	358	318
Fructose	384	428	411	429	405
Lactose	358	404	400	396	376
Saccharin (Sweet & Low)	330	376	356	382	386

Average 60 second heart rate of 3 daphnia tested.

Example #2: Initial Observations

Your Observations

1.

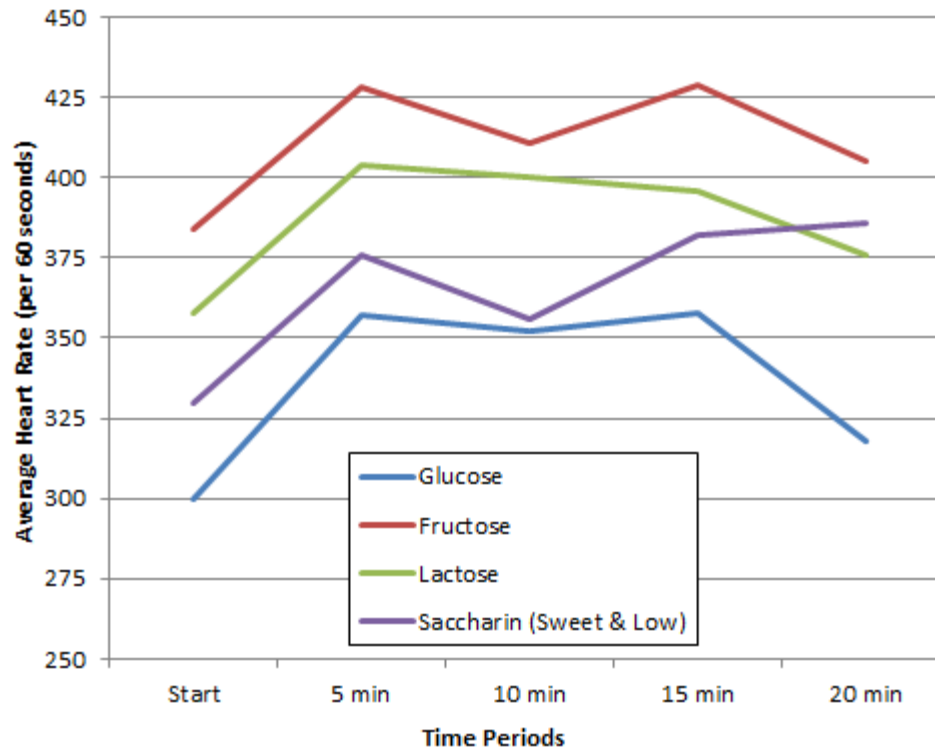
2.

Sugar Type	Time Periods				
	Start	5 min	10 min	15 min	20 min
Glucose	300	357	352	358	318
Fructose	384	428	411	429	405
Lactose	358	404	400	396	376
Saccharin (Sweet & Low)	330	376	356	382	386

Average 60 second heart rate of 3 daphnia tested.

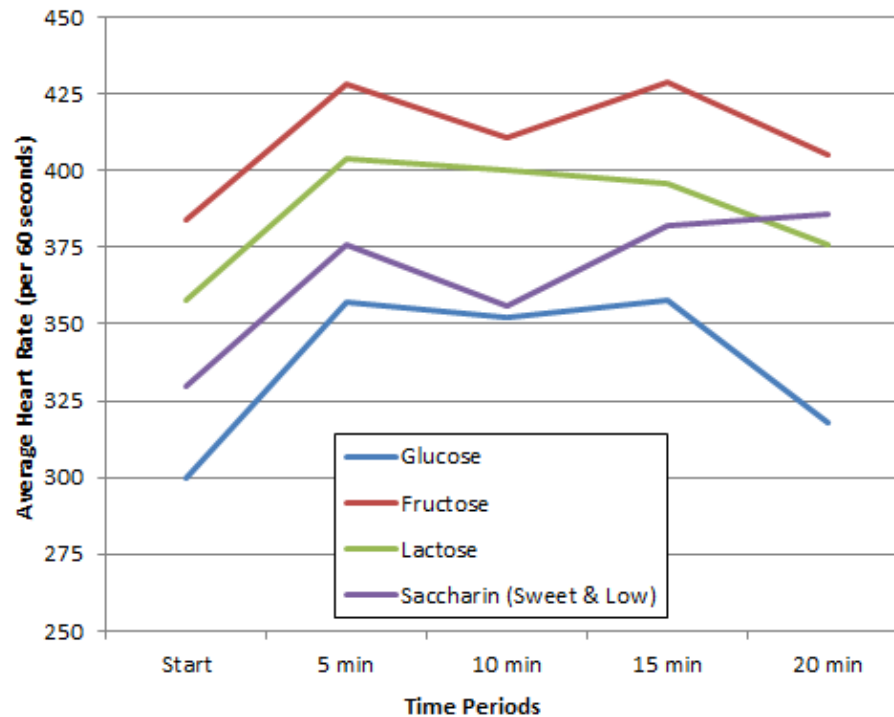
statsclass.org/isef

Example #2: Outcomes Presented



- *Fructose, Lactose, and Glucose caused sustained heart rate increase for about 10 mins*
- *Saccharin caused fluctuated increase in heart rate lasting full 20 mins*

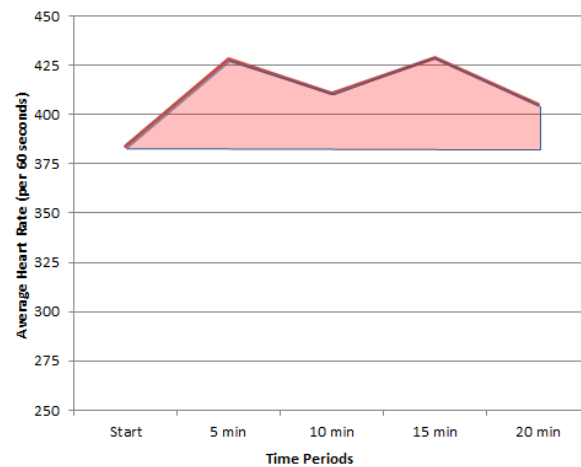
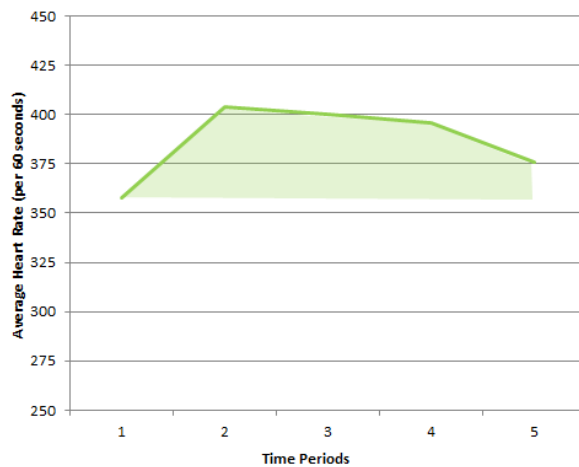
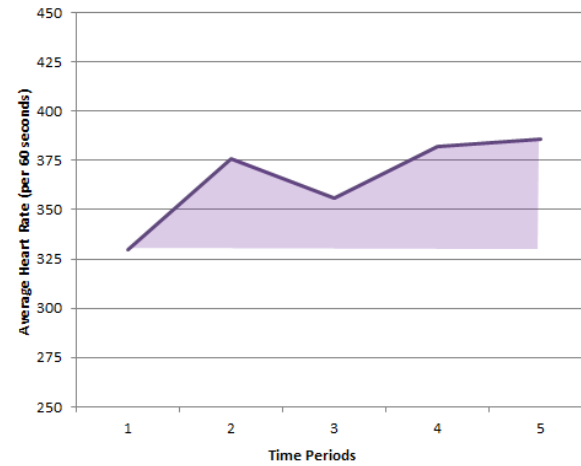
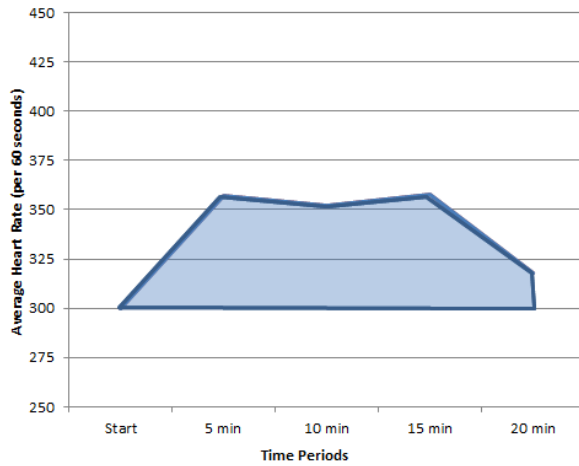
Example #2: Refined



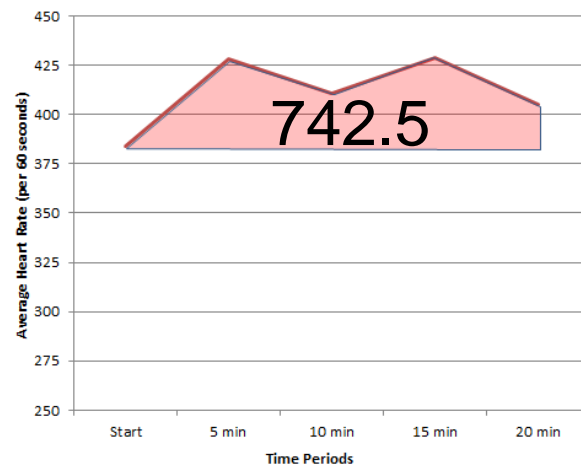
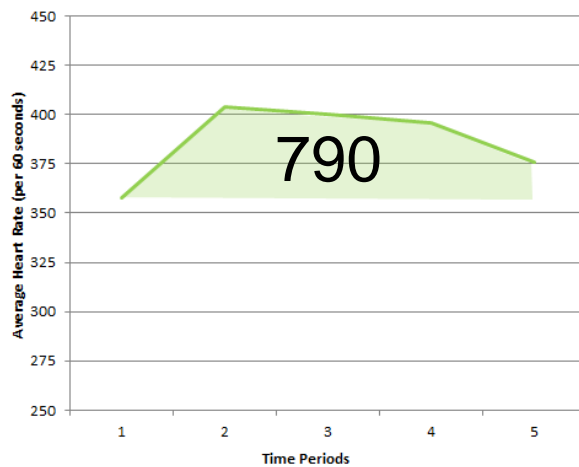
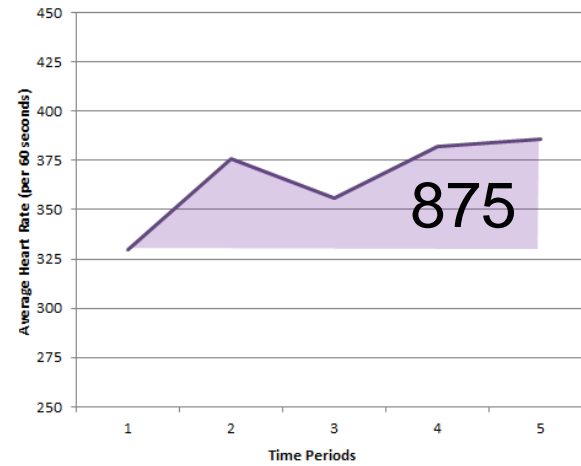
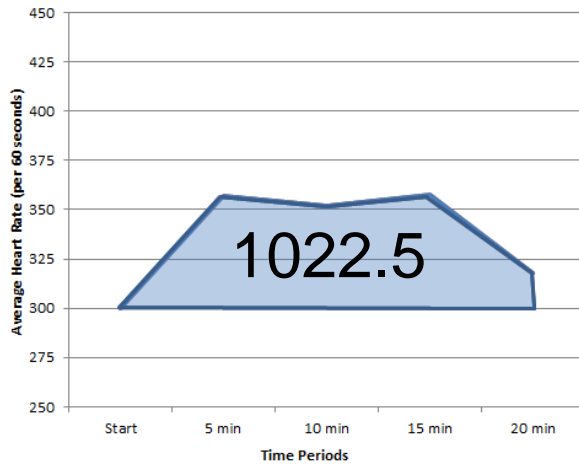
Questions:

- Which sugar had the largest impact on heart rate?
- Your friend suggests Fructose had the least impact, is this correct?

Example #2: Refined



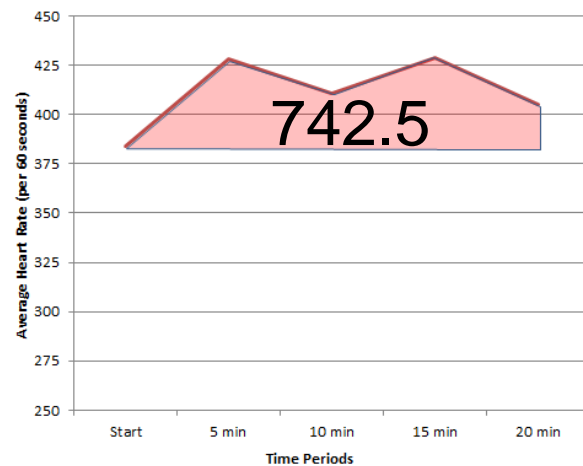
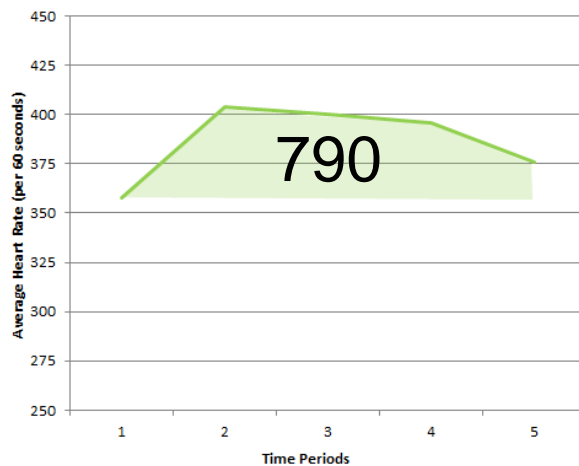
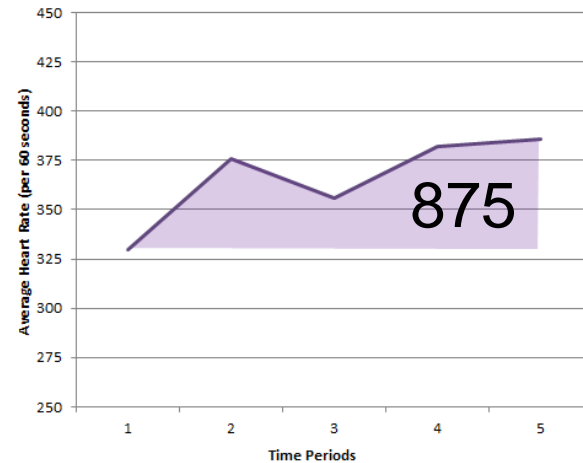
Example #2: Refined



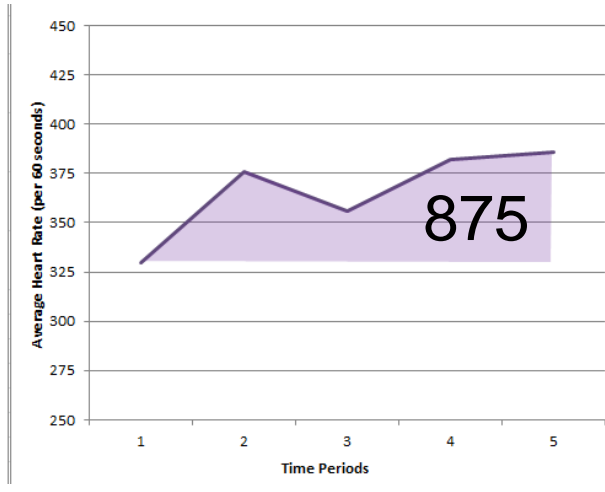
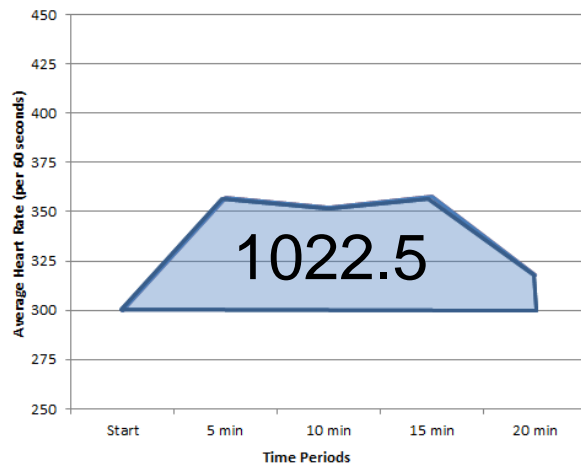
Example #2: Refined

My Observations

1. Effect of Fructose, Lactose and Saccharin is similar
2. Fructose had least impact on heart rate over 20 mins

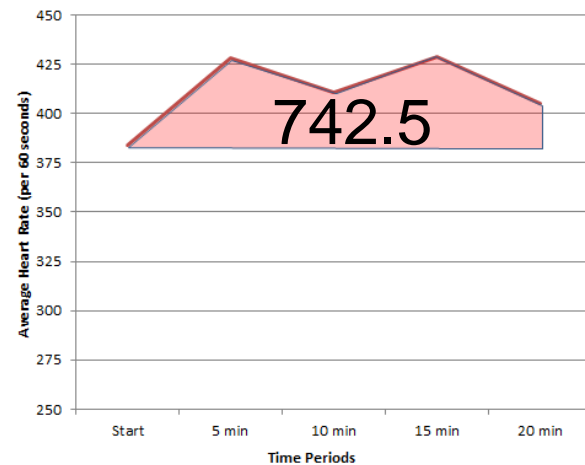


Example #2: Refined



My Observations

- The effect of Glucose is sustainably larger than that of Fructose



Example #2: Refined +

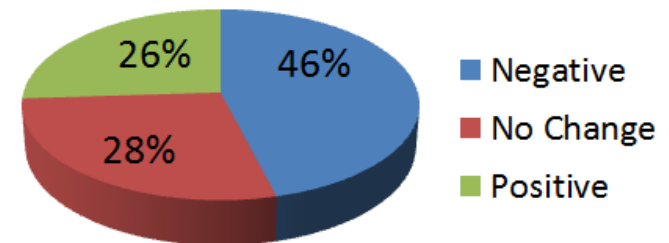
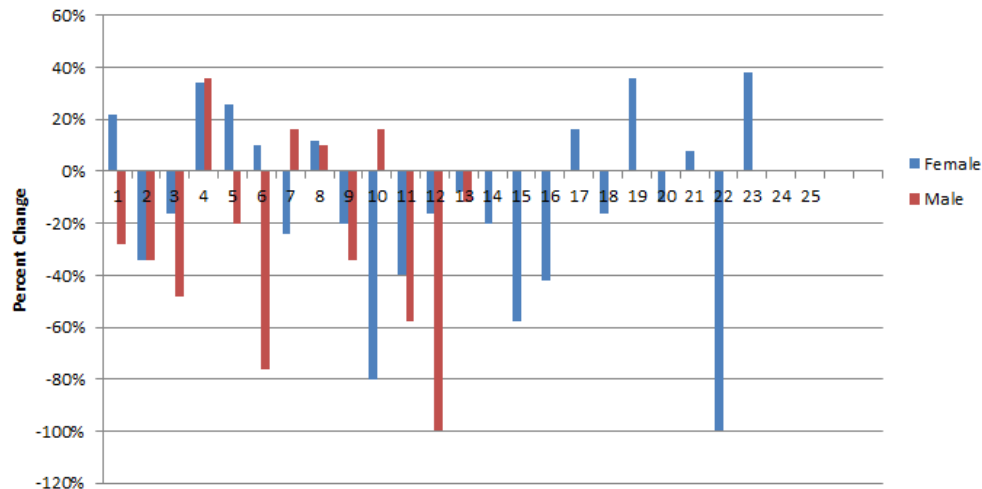
Obtain replicates to understand reproducibility

		Time Periods						
Suga	Time Periods							
Gl	Suga	Time Periods						
Fru	Gl	Suga	Time Periods					
La	Fru	Gl	Sugar Type	Start	5 min	10 min	15 min	20 min
Sac (Sweet	Lac	Fru	Glucose	300	357	352	358	318
	Sac (Sweet	La	Fructose	384	428	411	429	405
	Sac (Sweet	Sac (Sweet	Lactose	358	404	400	395	376
			Saccharin (Sweet & Low)	330	376	366	382	386

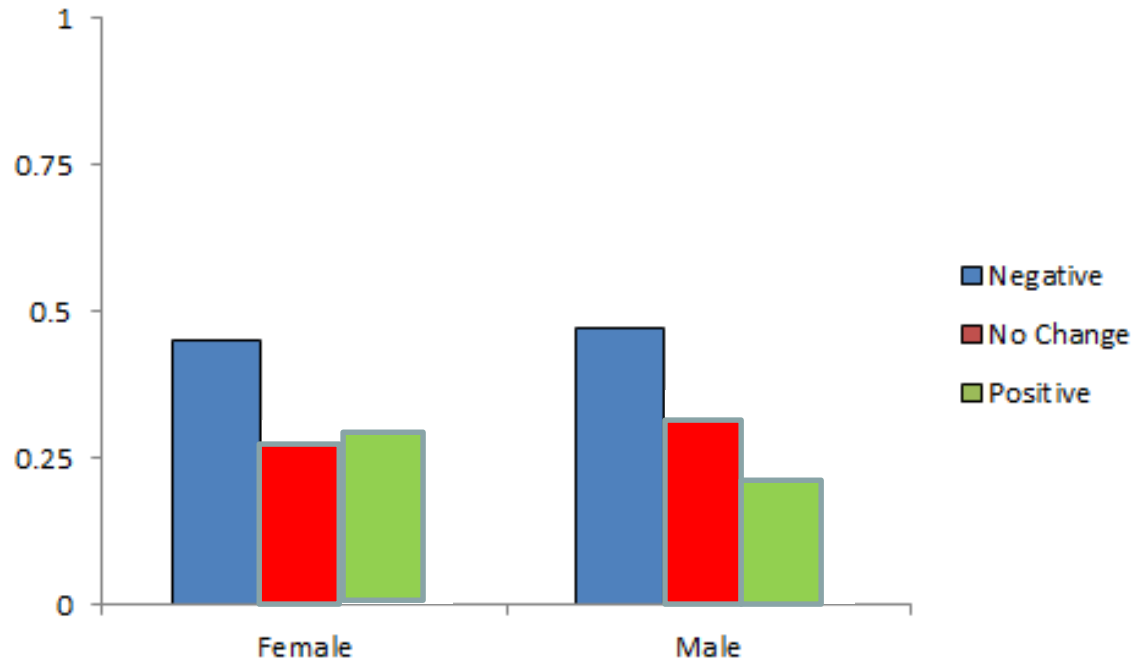
Example #3: Noise and the sense of Smell

Outcomes Presented

1. Almost half had negative effect when exposed to noise
2. Graphs show negative impact greater than positive impact



Example #3: Refined

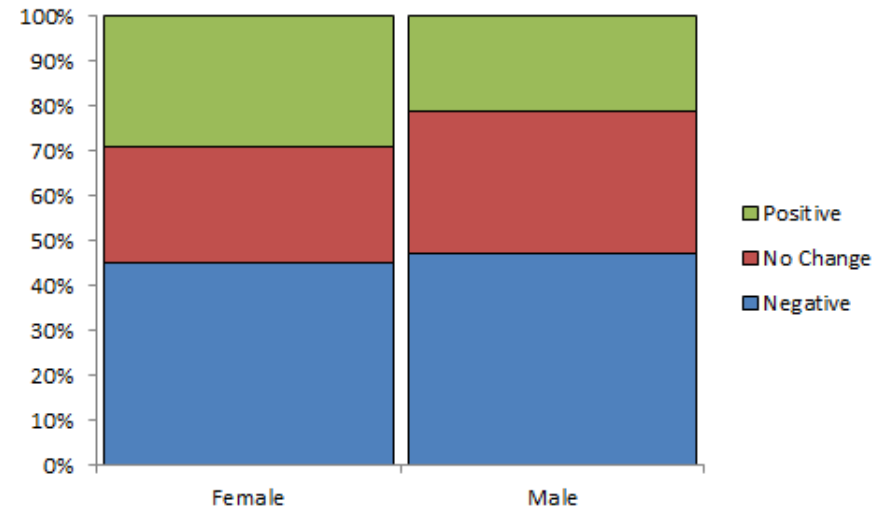


Note: Stack the bars and eliminate space between to obtain 100% Stacked Column graph

Example #3: Refined

My Observations

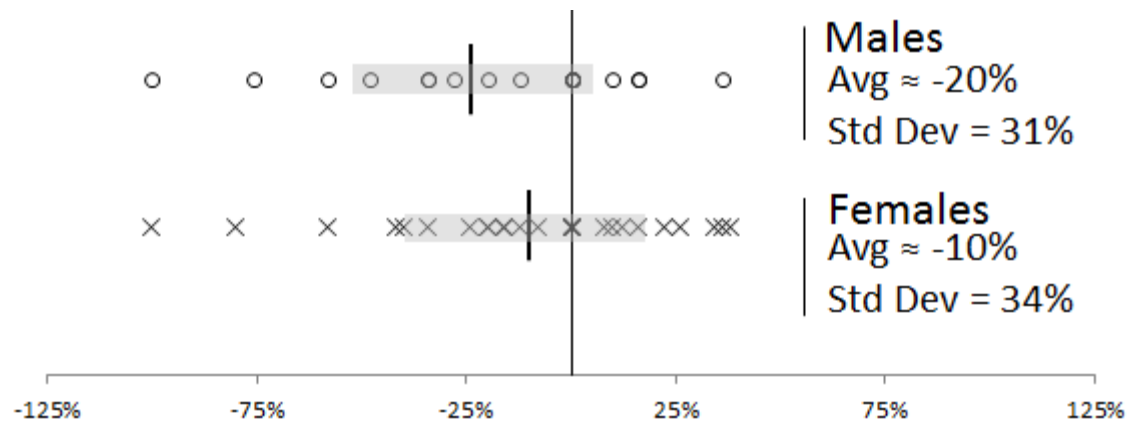
1. About half of the people were negatively affected by the noise
2. This is true regardless of gender
3. Males were slightly less likely to see a positive change



Example #3: Refined +

My Observations

3. A stacked dot chart could be used to contrast numerical measurements



Questions and (hopefully) Answers

Thank you!

Chris Malone
Winona State University
Email: cmalone@winona.edu