Process Improvement Project Storyboard: Monthly Saving Improvement

Define 7/3

- Monthly Saving was $1231.52 in average.
- The goal was to save $518.4 more, Improvement Rare = 42%.
- In order to extract more detail insight, the data set was prepared by week. Mean of Weekly saving is $307.88.

Weekly Expenses	
Mean	307.8791667
Standard Error	53.40803329
Median	358.515
Mode	#N/A
Standard Deviation	320.4481997

- Determined the defect by the Mean. SQL of current saving was 1.85.

Original Process (Nov ~ Jul)	
Defect opportunities per unit (D) =	4
How many units in the data set (U)	9
Total actual defets (D X U)	36
Total actual defects (A)	13
Defect-per-opportunity rate (A ÷ DU = DPO)	36%
Defects per million opportunities (DPMO)	361111
SQL value =	1.85

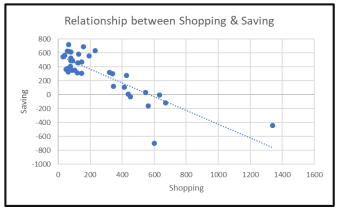
Measure 7/10

- The purpose was to **find the major expense.** Figured out the factor with highest correlation coefficient to leverage the affect of actions we were going to take.

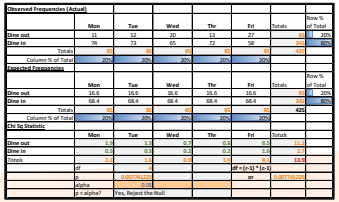
- According to the plot, I decided to do some further Analyses on **Grocery, Restaurant and Shopping.**

Analyze 7/17

- **Shopping** is the major expense category which has a strong relationship with **Saving** because **R = 0.82.**



- **p = 0.0077 < α = 0.05,** Chi-Square Test showed a strong evidence that I dined out more on **Wednesday and Friday.**



Improve 8/15

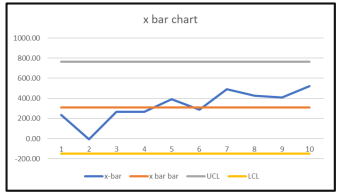
- Stop browsing Amazon and Facebook coupon sharing group. Starting bringing lunch box on Wednesday and Friday.
- Mean of Weekly Saving on August was $520.44.
- Hypothesis Test showed that I made a significant improvement and SQL of saving is 6 now.

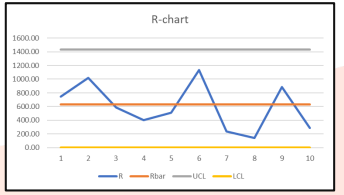
Original	
Mean	307.88
Standard Error	320.45
Count	36
New	
Mean	520.44
Standard Error	126.10
Count	4
α	0.05
Z-Value	-2.57
P-Value	0.005049
Reject Ho	

Original Process (Aug)	
Defect opportunities per unit (D) =	4
How many units in the data set (U)	1
Total actual defets (D X U)	4
Total actual defects (A)	0
Defect-per-opportunity rate (A ÷ DU = DPO)	0%
Defects per million opportunities (DPMO)	0
SQL value =	6

Control 8/29

- The improvement rate was actually 69% and beyond my original goal.
- The trend of saving goes up but still within the limit. It makes me confident with this result. **I'm capable of saving this much money.**





- Will take further actions when the saving become stable.

Data Collection Method

Resources

- Credit card & Bank account statement
- Venmo, PayPal statement

Data Wrangling

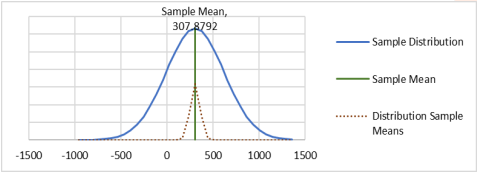
- All resources cover almost 99% information of the family expense.
- Mostly we use **continuous data (money) for analyzing**, and only **use discrete data for chi-square analysis.**
- Pulled up the record from **last November to this July (9 months)** and break them down into weeks (36 weeks) for analysis. The record of **August (4 weeks)** is for verifying the improvement.
- Some expenses like cash spending in an Asian restaurant or medical bill were excluded because they are out of scope and the affect is minor (< 2%), I'll treat these expenses as investment lost of stock and **won't become the measurement error.**

Descriptive Statistics of Weekly Saving

Current Situation

- My monthly discretionary income is $3854 and I thought I can easily save money for 1 credit ($1782) per month. My goal is to save 30% more which is $518.4.
- Obviously I was wrong. The Mean of Weekly Saving is $307.88 according to the data which means I only save $1231.52 in average.
- I realized I spent too much money before and need to take actions immediately. The goal still sticks with $518.4 because I think it's a reasonable amount. And the improvement rate will go up to 42% (518.4 / 1231.52 = 0.42)

Weekly Expenses	
Mean	307.8791667
Standard Error	53.40803329
Median	358.515
Mode	#N/A
Standard Deviation	320.4481997
Sample Variance	102687.0487
Kurtosis	1.807994357
Skewness	-1.296030887
Range	1421.87
Minimum	-698.7
Maximum	723.17
Sum	11083.65
Count	36



Sample Size Calculation

Margin Error

- Monthly Saving goal is $518.4, so I take $130 as Weekly Saving goal which is the Margin Error.

Result

- The result shows that I need 23 samples at least. It can be satisfied because I've already had 36 samples.

Sample Size Calculation	
Define α =	0.05
μ_o =	307.8792
Margin Error =	130
Standard Deviation =	315.9662
Lookup or Calculate $Z_{\alpha/2}$ =	1.96
Calculate sammple size (n) =	22.69377226

Box & Whisky Plot of Expenses Categories

Study of Expenses Categories

- All expenses can be categorized into 4 categories – Grocery, Restaurant, Shopping, Gasoline.
- It looks like Shopping is the major expense, but I also found the mean of Grocery is similar to the mean of Shopping.

Actions

- Will do hypothesis test to compare Grocery and Shopping to see if Shopping is really larger than Grocery.

Hypothesis Test – Grocery & Shopping

Test Result

- Ho: Mean of Grocery ≥ Mean of Shopping
  Ha: Mean of Grocery < Mean of Shopping
- We can reject the Null because the P-value = 4.46139E-07 is much smaller than α = 0.05. We have a strong evidence to prove that the Shopping is the major expense.

Actions

- Will do Linear Regression to see how much that Shopping affect the Saving.

grocery	
Mean	217.08
Standard Error	26.36
Count	36
shopping	
Mean	259.80
Standard Error	45.00
Count	36
α	0.05
Z-Value	-4.91
P-Value	4.46139E-07
Reject Ho	

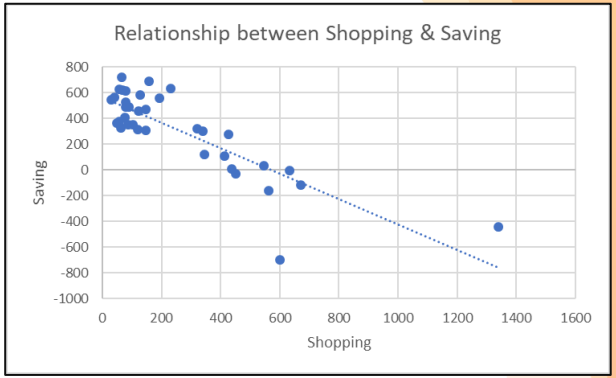
Linear Regression – Shopping & Weekly Saving

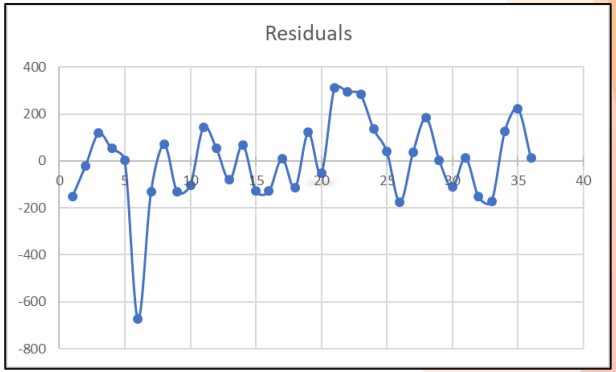
Why & How

- Shopping is the major expense. I would like to check how strong that Shopping can affect the Weekly Saving.
- R = 0.83, Residuals are randomly distributed. There is a **strong relationship between shopping and saving.** The more I shop, the less I save.

Actions

- Now we are pretty sure that **Shopping is the key factor for saving more money.** The action we take on Shopping will efficiently increase the most on Saving.
- Stop browsing Amazon and notices from Facebook coupon sharing group.





Regression Statistics	
Multiple R	0.829894589
R Square	0.68872503
Adjusted R Square	0.679569883
Standard Error	181.3946609
Observations	36

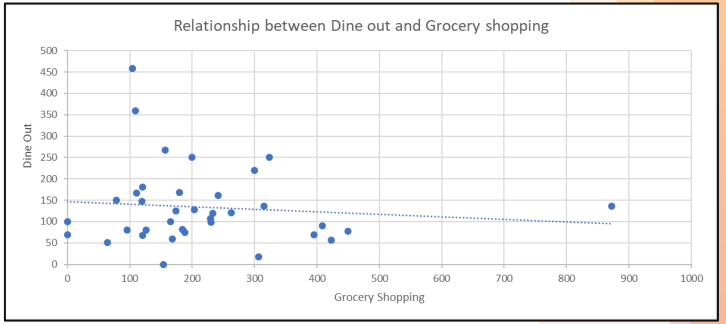
Linear Regression – Dine Out & Grocery Shopping

Why & How

- Basically these 2 categories are for getting food. I would like to see if there is trade-off relationship.
- R is only 0.1, There is almost no relationship between these 2 category

Actions

- No significant relationship means the expense on grocery shopping won't go up if I reduce the times of dinning out.
- Will do Chi-Square to see is there any way to reduce the times.



Regression Statistics	
Multiple R	0.103861944
R Square	0.010787303
Adjusted R Square	-0.018307188
Standard Error	93.26939002
Observations	36

Chi-Square – Work Day & Dine Out - 1

Why & How

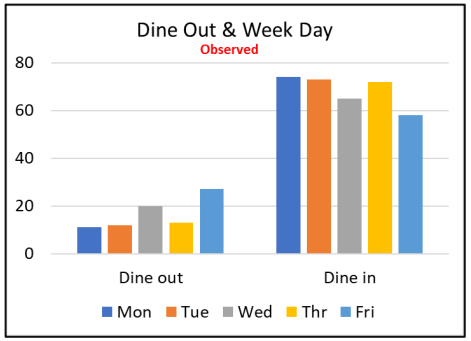
- This test is to figure out if I especially ate out with coworkers on certain days.
- Ho: Dine out and Day of a week are independent.
  Ha: Dine out and Day of a week are dependent.
- p = 0.0077 and **α = 0.05, we can reject the Ho because p < α.**
- **It's obvious that I dinned out more on Wednesday and Friday.**

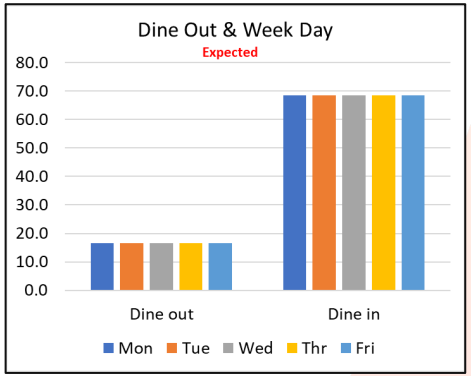
Why & How

- Middle of a week and Friday may make me especially feel casual and would like to hang out with coworkers more. Do prepare lunch boxes on these 2 days for staying at my seat and studying statistics.

Chi-Square – Work Day & Dine Out - 2

Observed Frequencies (Actual)							
	Mon	**Tue**	**Wed**	**Thr**	**Fri**	Totals	Row % of Total
Dine out	11	12	20	13	27	83	20%
Dine in	74	73	65	72	58	342	80%
Totals	85	85	85	85	85	425	
Column % of Total	20%	20%	20%	20%	20%		
Expected Frequencies							
	Mon	**Tue**	**Wed**	**Thr**	**Fri**	Totals	Row % of Total
Dine out	16.6	16.6	16.6	16.6	16.6	83	20%
Dine in	68.4	68.4	68.4	68.4	68.4	342	80%
Totals	85	85	85	85	85	**425**	
Column % of Total	20%	20%	20%	20%	20%		
Chi Sq Statistic							
	Mon	**Tue**	**Wed**	**Thr**	**Fri**	*Totals*	
Dine out	1.9	1.3	0.7	0.8	6.5	11.2	
Dine in	0.5	0.3	0.2	0.2	1.6	2.7	
Totals	2.3	1.6	0.9	1.0	8.1	**13.9**	
	df	4			**df = (r-1) * (c-1)**		
	p	0.007741229			**or**	0.007741229	
	alpha	0.05					
	p < alpha?	**Yes, Reject the Null**					





SQL

Defects Definition

- The Mean of weekly saving is $307.88. I choose this as the threshold to define defect because I hope the higher standard to force me to force me to save more.

Result

- The original process has an SQL = 1.85 and the new process has an SQL = 6.
- The new sample size is not large enough to be convincing but it's still a good sign as a new beginning.

Original Process (Nov ~ Jul)	
Defect opportunities per unit (D) =	4
How many units in the data set (U)	9
Total actual defets (D X U)	36
Total actual defects (A)	13
Defect-per-opportunity rate (A ÷ DU = DPO)	36%
Defects per million opportunities (DPMO)	361111
SQL value =	1.85

Original Process (Aug)	
Defect opportunities per unit (D) =	4
How many units in the data set (U)	1
Total actual defets (D X U)	4
Total actual defects (A)	0
Defect-per-opportunity rate (A ÷ DU = DPO)	0%
Defects per million opportunities (DPMO)	0
SQL value =	6

Hypothesis Test – Original & New (Process)

Test Result

- Ho: Mean of Original ≥ Mean of New
  Ha: Mean of Original < Mean of New
- We can reject the Null because the P-value = 0.005 is smaller than α = 0.05. We have a strong evidence to prove that the Improvement is significant.

Actions

- Keep on doing it and keep checking the progress by control chart.

Original	
Mean	307.88
Standard Error	320.45
Count	36
New	
Mean	520.44
Standard Error	126.10
Count	4
α	0.05
Z-Value	-2.57
P-Value	0.005049
Reject Ho	

Control Chart – Saving - 1

Why & How

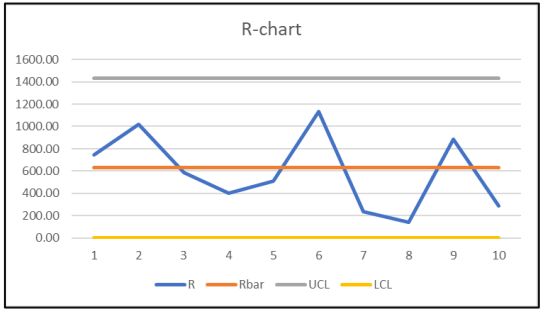
- The saving last month is $2082.76, the **improvement rate was 69%.** Which is much more than my original goal 42%.
- The control chart can help me to see how the trend goes.

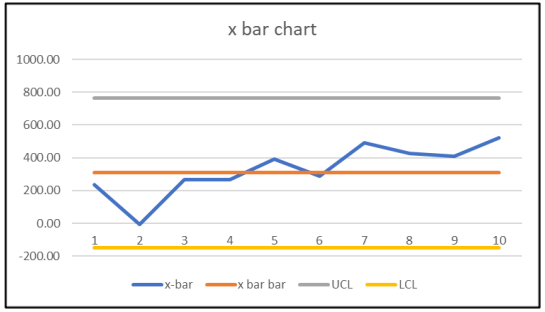
Result

- The trend of saving is going up but still within the limit. I believe it means that I'm totally capable of doing it. I just wasted too much money before.
- Will find out other way to save more when the trend reaches a plateau.

Control Chart – Saving - 2

n=4							R-chart					x bar chart			
	Sample				Calculations>>>			Centerline					Centerline		
Sample	**1**	**2**	**3**	**4**			**R**	**Rbar**	**UCL**	**LCL**		x-bar	**x bar bar**	**UCL**	**LCL**
Month 1	-31.64	-117.63	625.47	471.87			743.10	628.29	1432.50	0		237.02	307.88	766.5292	-150.771
Month 2	30.16	-698.70	315.25	321.05			1019.75	628.29	1432.50	0		-8.06	307.88	766.5292	-150.771
Month 3	375.27	121.73	579.78	-5.14			584.92	628.29	1432.50	0		267.91	307.88	766.5292	-150.771
Month 4	408.15	298.79	5.69	349.86			402.46	628.29	1432.50	0		265.62	307.88	766.5292	-150.771
Month 5	545.60	305.76	617.30	106.15			511.15	628.29	1432.50	0		393.70	307.88	766.5292	-150.771
Month 6	-444.06	632.48	690.44	279.04			1134.50	628.29	1432.50	0		289.48	307.88	766.5292	-150.771
Month 7	526.98	326.53	561.58	555.78			235.05	628.29	1432.50	0		492.72	307.88	766.5292	-150.771
Month 8	491.38	352.20	489.84	364.83			139.18	628.29	1432.50	0		424.56	307.88	766.5292	-150.771
Month 9	-161.31	613.20	723.17	456.80			884.48	628.29	1432.50	0		407.97	307.88	766.5292	-150.771
Month 10	558.49	622.70	564.18	336.37			286.33	628.29	1432.50	0		520.44	307.88	766.5292	-150.771
						Rbar =	628.29				x bar bar=	307.88			
						UCL = D4*Rbar		**1432.496**			UCL = xbarbar+A2*Rbar		**766.52924**		
						LCL = D3*Rbar		**0**			LCL = xbarbar-A2*Rbar		**-150.77091**		





Conclusion

- The saving last month is $2081.76, I actually saved $850.24 more than the original goal $518.4
- The improvement rate was 69%, more than my original goal 42%.
- SQL has been improved from 1.85 to 6.
- Since there is no out-of-control signal in the control chart. I believe the goal saving is achievable and will keep these actions as my current rule of saving.
- Shopping is the major expense. Will take some further actions on it to save more money.