Lean Six Sigma Project

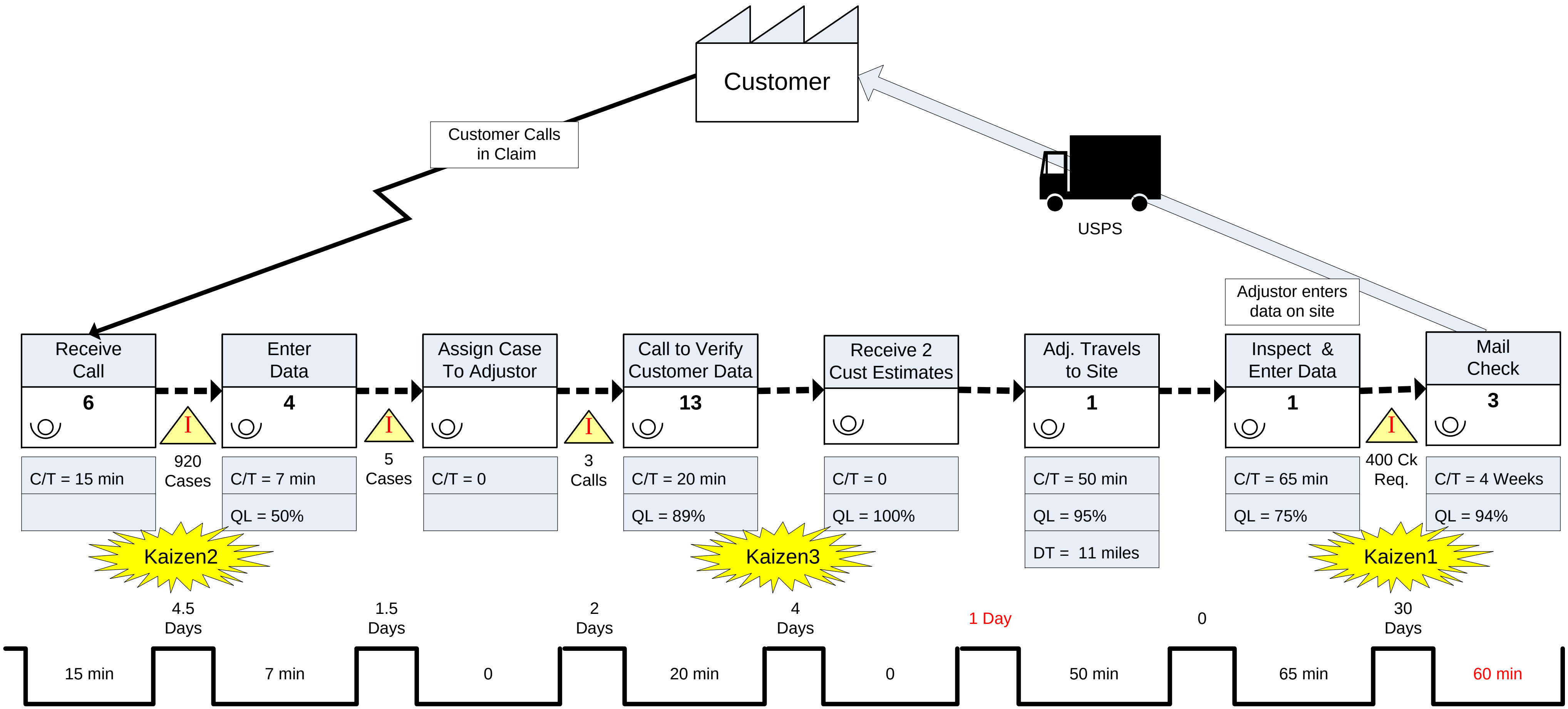
Acme Insurance Company

Dr. Michael S. Jordan

Georgia State University Teaching Case

Acme Insurance Company

Claims Processing VSM



Total Lead Time = 43 Days
Processing Time = 3.6 Hours

Takt Time = $\frac{\text{Available Work Time per Shift}}{\text{Customer Demand per Shift}}$

= $\frac{\text{450 Min}}{\text{200 Claims}}$

= 2.25 Claims (Cases) per Min

Define

<u>The Business Context</u>

The Acme Insurance Company is in the business of insuring vehicles. The company has three processing centers in three different geographic regions that handle two types of insurance claims – A and B.

The company processes on the average 200 claims per day, where a day is 7.5 hours. Customers have been complaining that the company is too slow in processing accident claims. On the average, the company takes 43 days to mail a claim check to a customer from the time the customer first calls in the claim. The actual "touch time" to process a claim within the 43 days is 3.6 hours (processing time), indicating a lot of delays and wait time between process steps. Acme management has also noted that they are not happy with the process either claiming that the cost to process claims is too high. In fact, the company actually looses money on claims processing.

Upon cursory inspection of the available data, it has been determined that the primary outcome of interest (big Y) is to reduce the cost per claim. It has been determined that the primary factor driving the cost per claim is total lead-time for processing a claim (the Big X). When we view total lead time as an outcome variable (small y relative to the Big Y), we also recognize that the small x's or factors driving total lead time are the time delays in the process (see Acme Value Stream Map). These delays are due to wasteful activities in the value stream.

We developed a value stream map of the claims process. We have identified three areas in the value stream where delays are particularly high -- 1) the time it takes to enter a customer claim in the computer after receipt of a claim call, 2) the time it takes to assign a claim to an available adjustor once the claim has been entered, and 3) the time it takes to actually process a claims check at the end of the process. Removing the wasteful activities driving these delays will dramatically reduce the total lead-time of the process, thus significantly reducing the cost per claim.

By reducing the total lead time of the claims process, the customer will be more satisfied while, at the same time, the total cost per claim will decrease since cost per claim is a function of lead time. This is true because most of the cost of processing a claim is fixed cost in the form of human resources, offices, computers, etc. By reducing the total lead-time for processing a claim, Acme can process more claims per day – increasing claims throughput. When waste is eliminated from the process, capacity is freed up in the form of human resources. This freed up capacity can be converted to financial benefits by processing more claims per day. Thus, the focus of this Lean Six Sigma project is to determine and eliminate the drivers (the x's) of the delays in the process, which will reduce total lead-time. This will ultimately reduce the cost per claim by increasing claims throughput. Secondary effects will be happier customers and brand lift, which could produce an increase in demand for Acme insurance products over time via word of mouth.

Data was collected (150 observations) from all three claims processing centers. There appears to be significant variation in the cost per claim between the centers. Preliminary statistical analysis was performed on the dataset (see Table 1). Descriptive statistics indicate that the mean of all observations is 285.95 with a standard deviation of 33.35. Further, the cost per claim ranges between 236.42 and 347.61. Descriptive statistics confirm that the data is normal.

Table 1

Cost/claim	
Mean	285.95
Standard Error	2.72
Median	279.79
Mode	-
Standard Deviation	33.35
Sample Variance	1112.16
Kurtosis	-1.42
Skewness	0.31
Range	111.19
Minimum	236.42
Maximum	347.61
Sum	42892.66
Count	150.00
Largest (1)	347.61
Smallest (1)	236.42
Confidence Level (95.0%)	5.38

A multi-histogram graph was created segmenting the data by region and product type (see Figure 1). The graph indicates that the highest variation and cost per claim is occurring in region 1 for product B with the next highest occurring in region 2 for product B. Also, region 3 shows moderately high cost for product B as well. The lowest cost is occurring in regions 1, 2, and 3 for product A. Figure 2 shows the same information in another format. Variation is lowest in region 3 for product B. Based on the business case for this project, we see cost per claim as a function of total claim lead-time. Variation in the cost per claim is symptomatic of uneven lead-time due to undetermined factors (the x's). Also, Figure 1 suggests that that products A and B may require different processing times. Specifically, product B appears to require more time than product A, possibly due to more requirements.

Figure 1

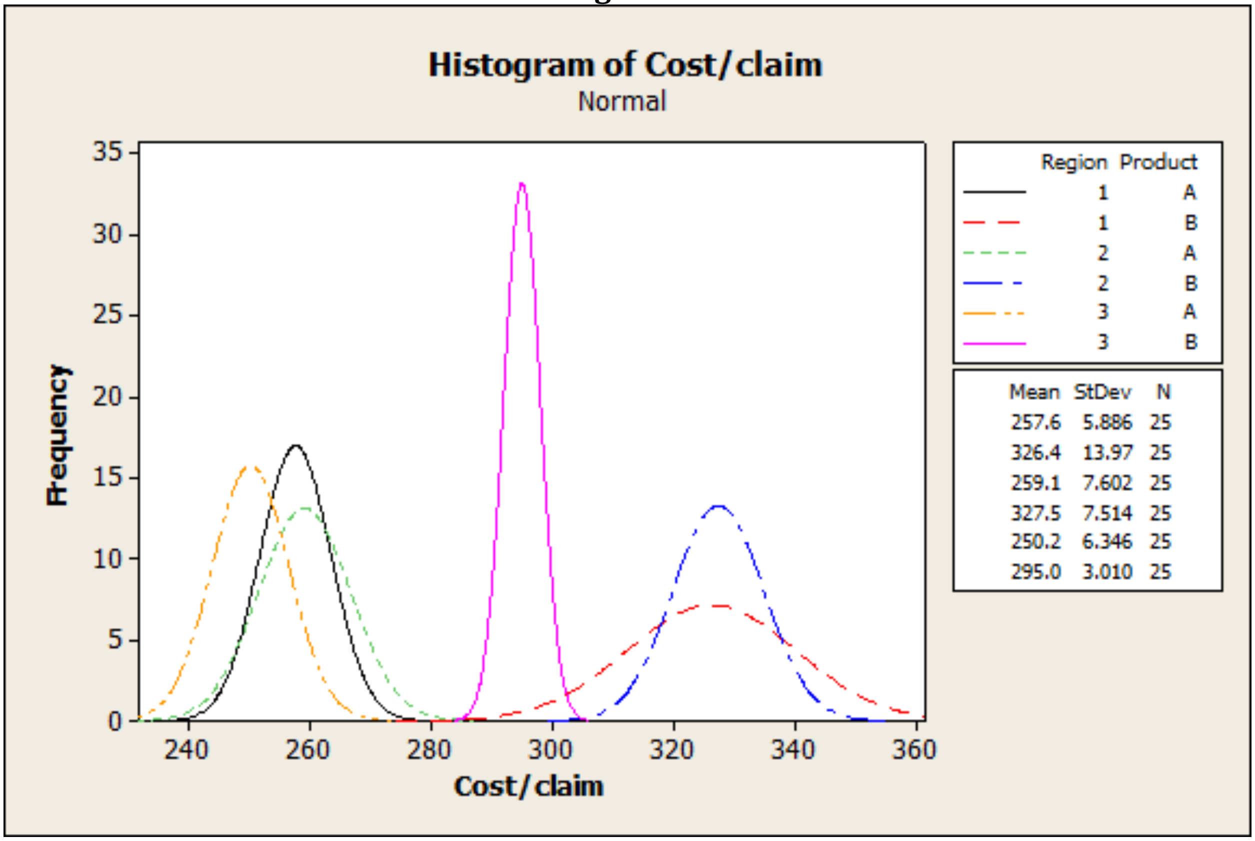
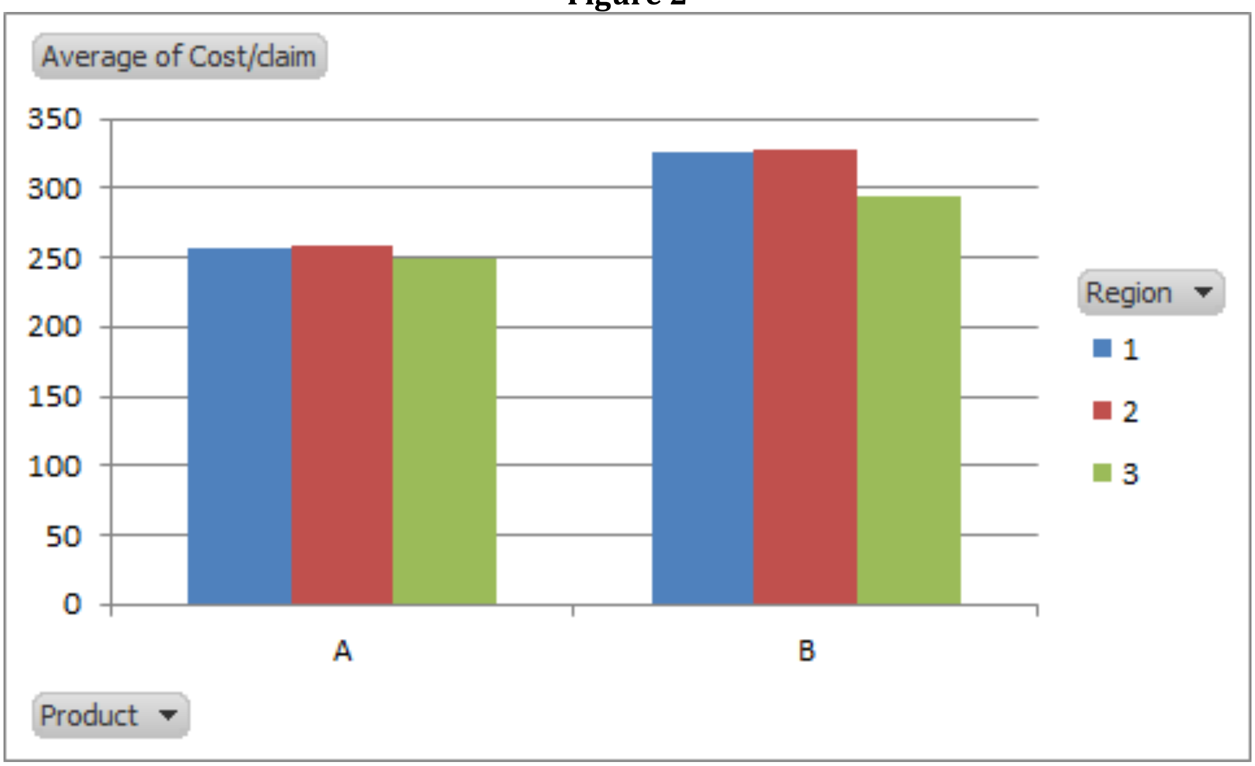
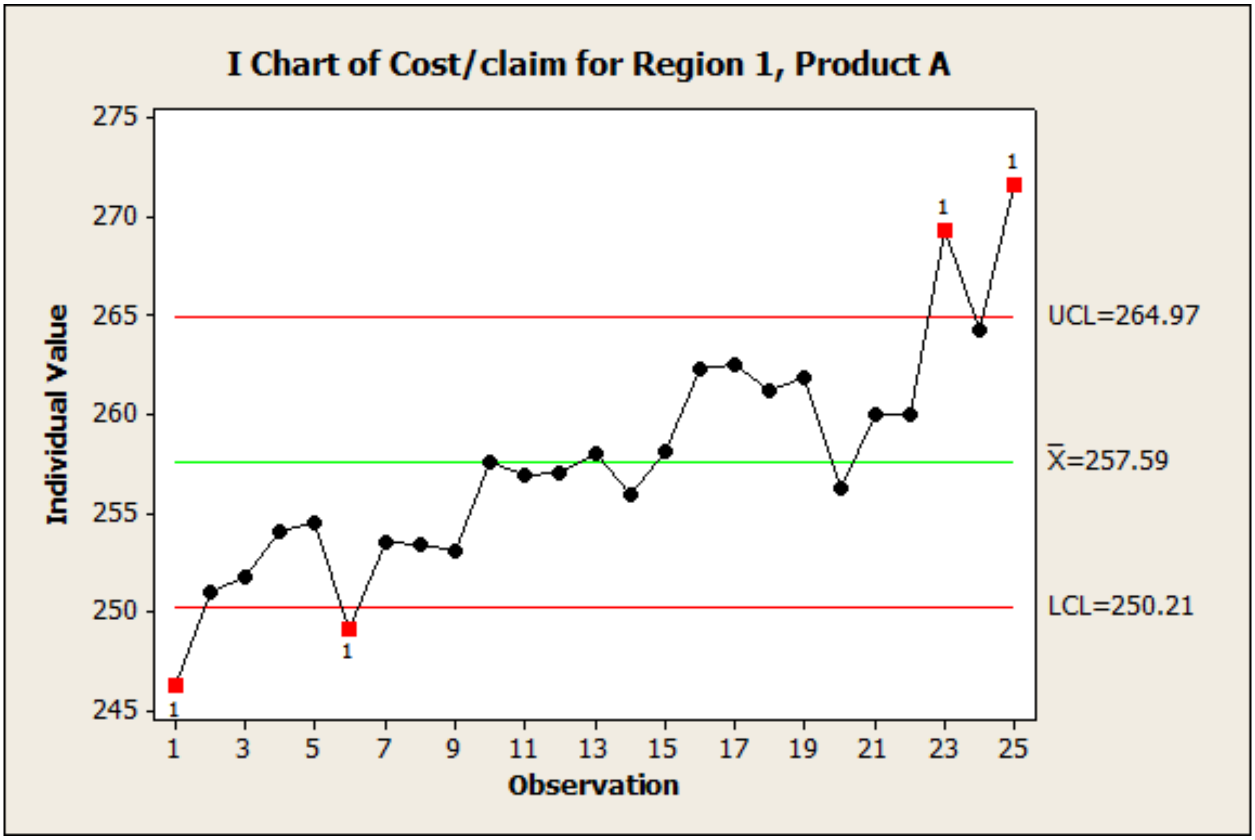


Figure 2



Measure

Figure 3



We created a control chart for cost per claim for region 1, product A (Figure 3), which indicates that this process is not within statistical control and therefore not predictable. The two data points below the lower control limit shows a very low cost exceeding three standard deviations from the mean in the negative (good) direction. The two higher points may indicate a special cause of variation. However, assignable cause will be hard to determine without additional time series data. The remaining observations that are within the control limits suggest an upward trend which may indicate that the underlying cause(s) or the x's that are driving delays in the process are getting progressively worse in region 1 for product A. Since Time and cost are related, we consider the cost per claim to be a proxy for process time. Higher cost per claim indicates increase delays due to x factors.

Table 2

Product Type	Region 1	2	3	Grand Total
A	258	259	250	256
B	326	327	295	316
Grand Total	**292**	**293**	**273**	**286**

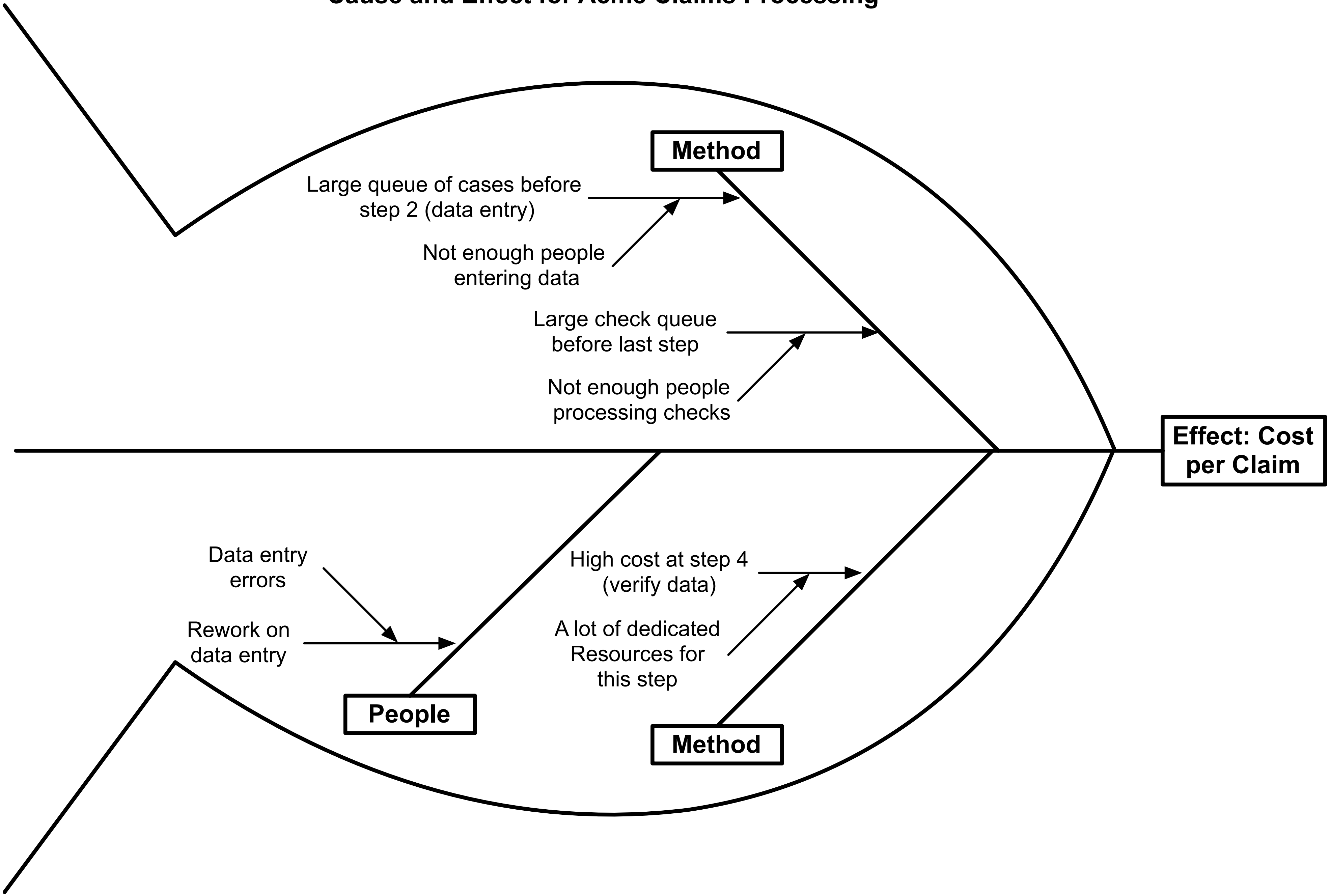
Next, we created a pivot table from the available data. The pivot table shows a cross sectional view of the data -- the average cost per claim for each region and product type. The pivot table reveals significant variation in the cost per claim for product B in regions 1 and 2. These two regions have a much higher cost per claim for product B than does region 3. Further, regions 1 and 2 have a moderately higher cost for product A than does region 3. This pivot table is graphically depicted in Figure 2. Based on the business case logic for the project that Time is driving cost, the biggest opportunity for improving the Big Y (reducing the cost per claim) is to reduce the delays in the process in three areas where these delays are particularly high (prioritized in the order presented):

- Kaizen 1 – Step 8
- Kaizen 2 – Step 2
- Kaizen 3 – Step 5

These areas are indicated by Kaizen bursts on the value stream map.

Next, we developed a cause and effect diagram to brainstorm the possible causes of delays in the process that drive up the cost per claim. We have concluded that rework due to errors in processing claims is one of the primary drivers of cost per claim. To investigate this further, we ran descriptive statistics and histogram charts for each of the x's of interest which may be contributing to process errors.

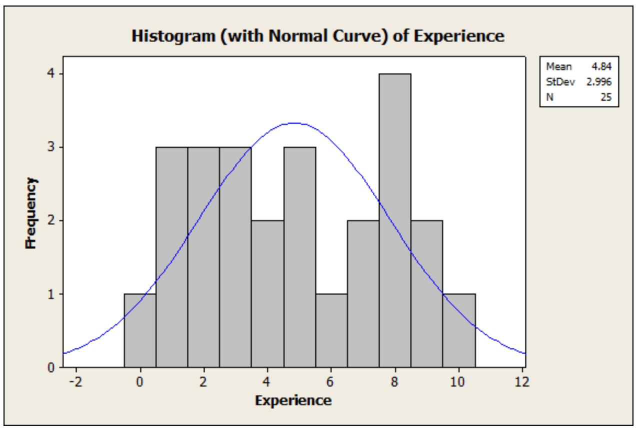
Cause and Effect for Acme Claims Processing

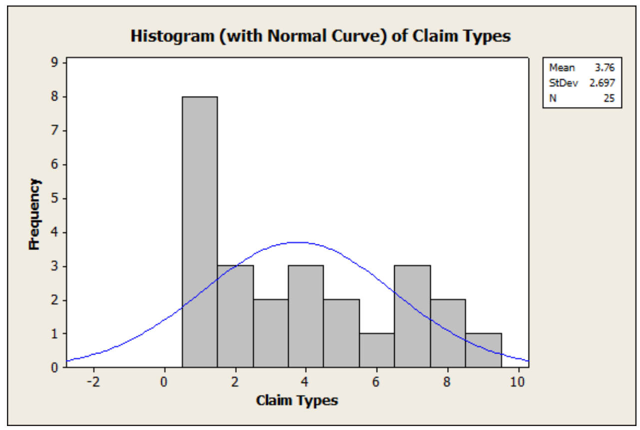


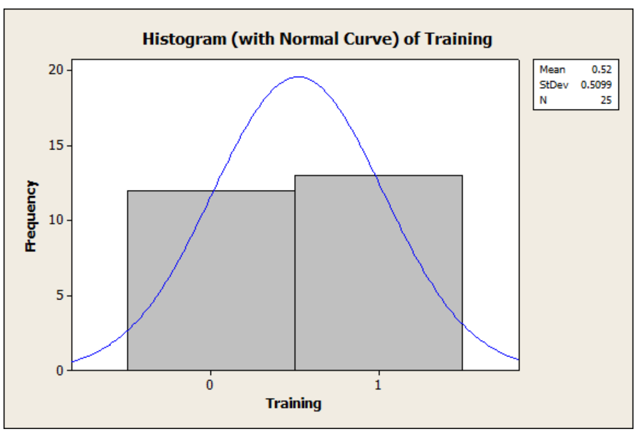
Descriptive Statistics and Distribution Histograms for X's on Errors Worksheet

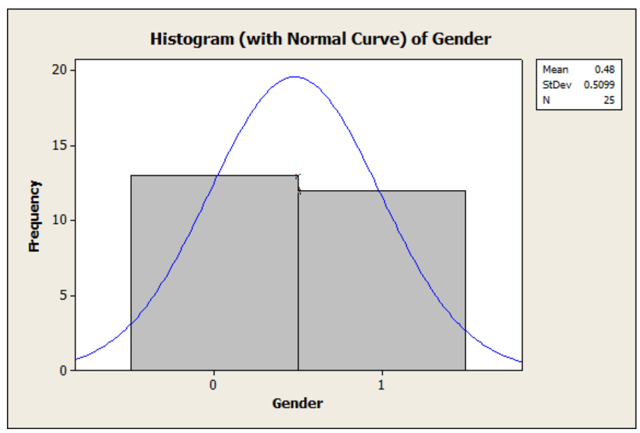
Variable	N	N*	Mean	SE Mean	StDev	Variance	Range
Experience	25	0	4.840	0.599	2.996	8.973	10.000
Claim Types	25	0	3.760	0.539	2.697	7.273	8.000
Training	25	0	0.520	0.102	0.510	0.260	1.000
Gender	25	0	0.480	0.102	0.510	0.260	1.000
No. Claims	25	0	295.2	25.4	126.8	16082.3	365.0

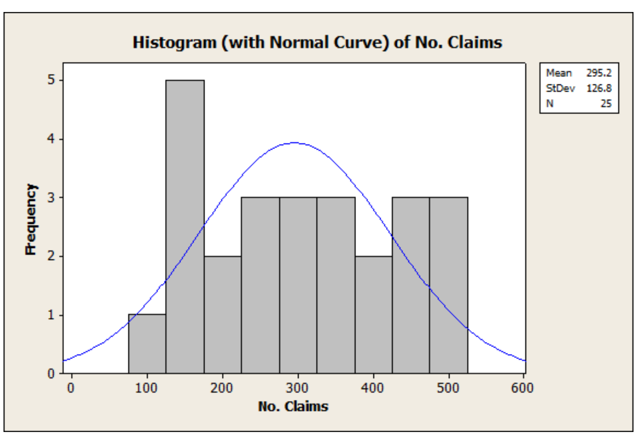
Figure 4











Based on this output (Figure 4), we would eliminate Training and Gender as factors (x's) that lead to errors since the binary values for each factor are evenly distributed. We would say that that the potential contributors of processing error would be the Claims Type, Number of Claims Handled, and Experience. It may be the case that Claim Type may have a larger influence on processing errors than the other two.

Possible Measurement Errors

The number of errors per processor could introduce measurement error. First, the data would not include errors that were not detected. Second, errors can be misclassified. Additionally, since there is a 4.5-day lead-time before a customer claim is entered into the computer, the total time to process a claim may not include this delay. That is, the total time to process a claim may start from when the claim is entered into the computer and not when the customer call is first received. This would have the effect of understating the cost per claim.

Benchmarking

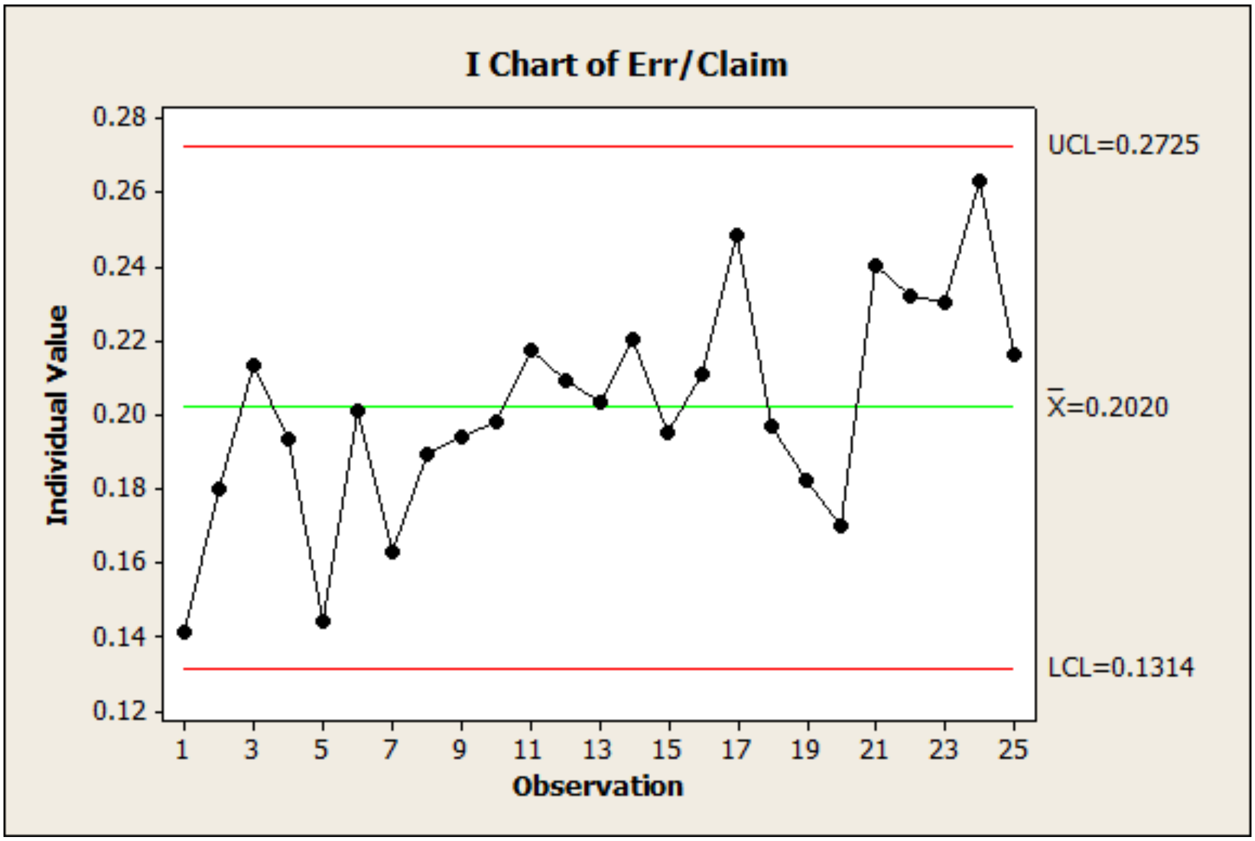
It may be necessary to benchmark internally and externally to affirm that certain project goals are reasonable. Based on the business case logic for the project, the cost per claim (the Big Y) is a lagging indicator of total lead-time for claims processing. Therefore, the cost per claim will continue to do down as total lead-time goes down because increased claims throughput will convert fixed capacity released from the elimination of waste resulting in more claims being processed per day. More claims processed per day with the same fixed capacity will lower the cost per claim. The cost per claim will eventually level off as the claims process is optimized.

Based on the data, we surmise that product B takes longer to process than product A, possibly because product B is more complex or has more requirements than product A. The goal is to reduce total lead-time for processing product A from 43 days to 10 days. The goal is also to reduce total lead-time for processing product B from 43 days to 15 days. The improvement team can benchmark other insurance companies to determine what their total lead-time is for processing claims. The team can also perform a process time study that tracks the cycle times of randomly chosen claims for both products A and B to determine what the process can accommodate.

Analyze

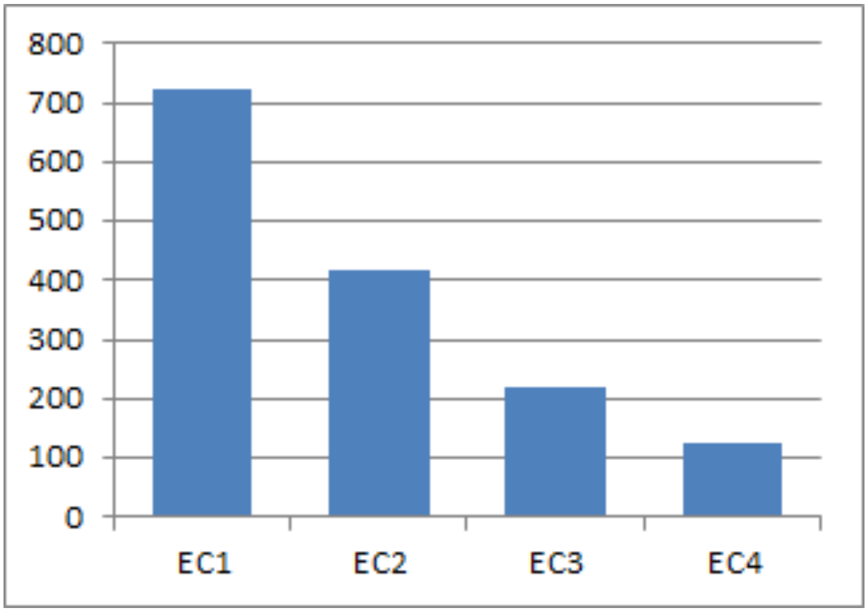
Next, the Error per Claim ratio was plotted on an I Chart (Figure 5) to graphically view the variance pattern of the errors over a six-month period. The I Chart indicates a discernable upward trend of errors, which suggests that the errors are getting worse over time.

Figure 5



Next, we created a sorted histogram of the Error Codes by frequency (Figure 6) over the six-month period and performed a Pareto analysis. The Pareto analysis reveals that EC1 and EC2 together account for 77% of all the errors suggesting that the improvement team should focus on these two error types for maximum leverage.

Figure 6



We then computed the correlation coefficients (Table 3) and created scatter plots (Figure 7) for each of the 'x' factors against the Error Ratio (the 'y' variable) to determine their relationships.

Table 3: Correlations

Pearson correlation of Ratio and Experience = -0.870
P-Value = 0.000

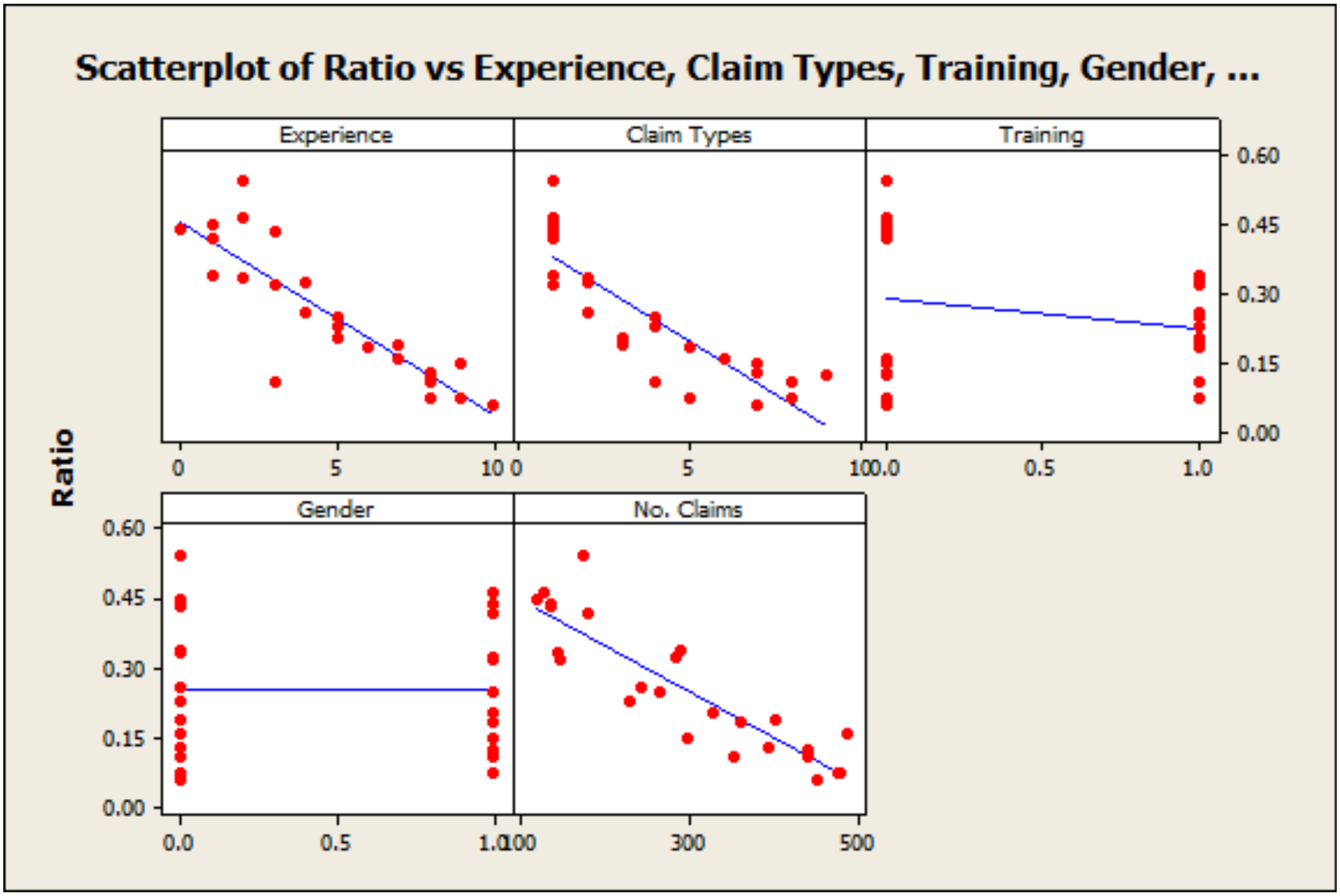
Pearson correlation of Ratio and Claim Types = -0.860
P-Value = 0.000

Pearson correlation of Ratio and Training = -0.223
P-Value = 0.284

Pearson correlation of Ratio and Gender = 0.005
P-Value = 0.981

Pearson correlation of Ratio and No. Claims = -0.892
P-Value = 0.000

Figure 7



Experience, Claim Types, and No. Claims all have a high negative correlation with the Error Ratio as indicated in Table 3 and Figure 7. Further, all three correlation coefficients are significant, $p < .001$. Specifically, this means that 1) as Experience increases, the Error Rate decreases; 2) as claim types increase (from 1 to 9), the Error Rate decreases, 3) as the No. Claims increases, the Error Rate decreases. Training and Gender do not have a significant correlation to Error Ratio.

Next, a pivot table was created to compare the Error Ratios by the categories of Training and Gender to determine any relationships (Table 4). The pivot table reveals no significant differences in the number of females versus males that received training. That is, females received about the same training as males did. Further, as previous graphs indicate, both Gender and Training values are distributed roughly evenly within the Error Ratio. This suggests that Gender and Training do not a have a significant relationship to the Error Ratio.

Table 4

Categories	Count of Ratio
Female	**12**
No	6
Yes	6
Male	**13**
No	6
Yes	7
Grand Total	**25**

Next, a step-wise regression was performed (Table 5) to determine the most parsimonious regression equation. As earlier analyses revealed, Gender does not have a significant relationship to the Error Ratio. We had concluded earlier that Training was a weak predictor as well. However, the step-wise regression suggests that the most parsimonious regression model will include all variables except for Gender. This may be due to interaction effects between Training and other variables. The No. of Claims variable just barely makes the $p < .05$ cutoff and therefore is still included in the regression equation. We then performed another regression that includes these four variables to obtain the final regression equation (Table 6).

Table 5

Step	1	2	3	4
Constant	0.5505	0.5354	0.5585	0.5392
No. Claims	-0.00101	-0.00062	-0.00054	-0.00025
T-Value	-9.48	-4.18	-4.22	-2.06
P-Value	0.000	0.000	0.000	0.053
Experience		-0.0207	-0.0239	-0.0177
T-Value		-3.33	-4.46	-3.95
P-Value		0.003	0.000	0.001
Training			-0.059	-0.080
T-Value			-3.13	-5.08
P-Value			0.005	0.000
Claim Types				-0.0221
T-Value				-3.81
P-Value				0.001
S	0.0659	0.0550	0.0465	0.0363
R-Sq	79.62	86.44	90.74	94.63
R-Sq(adj)	78.73	85.21	89.42	93.56

Table 6

Regression Statistics	
Multiple R	0.973
R Square	0.946
Adjusted R Square	0.936
Standard Error	0.036
Observations	25

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	0.464	0.116	88.166	2.07396E-12
Residual	20	0.026	0.001		
Total	24	0.490			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	0.5392	0.0204	26.3994	5.09468E-17	0.4966
Experience	-0.0177	0.0045	-3.9508	0.001	-0.0271
No. Claims	-0.0003	0.0001	-2.0553	0.053	-0.0005
Training	-0.0803	0.0158	-5.0752	5.78282E-05	-0.1133
Claim Types	-0.0221	0.0058	-3.8072	0.001	-0.0342

Based on the regression output, we are now able to create a regression equation that predicts the Error Ratio as a function of Experience, No. Claims, Training, and Claim Types. The coefficients for each variable represent the slope of the best fitting regression line through the data points between predictor and outcome variable.

Regression Equation:

Error Ratio = -.0177 (Experience) - 0.0003 (No. Claims) - 0.0803 (Training) – 0.0221 (Claim Type) + 0.5392

The multivariate regression (Table 6) indicates an R square of 0.946 where the model as a whole is significant, $p < .001$. The interpretation is that the model explains 94.6% of the variance in the outcome (y) variable Error Ratio. Regression output shows a standard error of 0.036 for the model. The standard error indicates how well the model will be able to predict Error Rates, which in this case means the model is a good predictor.

The Business Case for Lean Six Sigma

Question: What is Lean Six Sigma and why should we use it at Acme Insurance Company?

Response to Acme CEO:

Lean Six Sigma is a process improvement system that will enable us to identify and eliminate variation and waste in our processes that drive up costs and reduce the quality of our services. When variation and waste are removed, we will be able to reclaim the capacity that is currently tied up by wasteful activities. The capacity can then be used to grow our company without the need to invest more capital. At the same time, profitability on current business will increase and improved quality may grow top line revenues due to customer loyalty and word of mouth. For example, you know that problem in processing that we have been trying to figure out that makes a lot of our customers angry and costs the company lots of money? Well, we can eliminate this problem using Lean Six Sigma by using data driven methods to identify the root cause of this problem and then to eliminate it permanently. This costs us nothing except our time, but costs us plenty if we continue to flounder around trying to fix the problem with trial and error reasoning.