Introduction

When one of the world's largest lenders discovered that revenues for its Latin American auto financing operations were flat, it embarked on a mission to collect feedback from first-tier customers, in this case, auto dealers, and second-tier customers, borrowers, or car buyer. Overwhelmingly, the dealers reported that *credit decision response time* was the number one issue that the lender needed to improve. A quick initial credit decision keeps the customer from shopping elsewhere, both for a car and for the financing.

A team of *Six Sigma Black Belts* was formed at the lender's company then began to quantify the effects of slow response time and learned that 40% of credit applications weren't moving past the initial application stage, either because the application was rejected or the customer went elsewhere for financing. This 40% figure translated into $110 million in lost revenue annually for the lender.

Given the *voice of the customer* (VOC) data, which highlighted the importance of credit decision response times and the tremendous opportunity to capture lost revenue, the lender's Six Sigma team embarked on an ambitious improvement journey to accomplish three goals:

- Identify the root causes of uncompetitive response times.
- Reduce response times to surpass the competition.
- Increase revenue by converting 5% of "lost" applications to purchased contracts

Since the team is aware that statistical inferential techniques provide solid foundation to achieve the above goals, they hired you as Analytic Consultant to lead their statistical analysis and help them draw reasonable solutions.

Lending Process

In the United States, the initial credit decision typically takes just a few minutes, but in Latin American countries the time is much greater—sometimes as long as 24 hours. As a leading analyst of the Six Sigma team, you had your team investigate lending process first. Figure 1 shows a process map detailing the initial activities included in the auto loan application process.

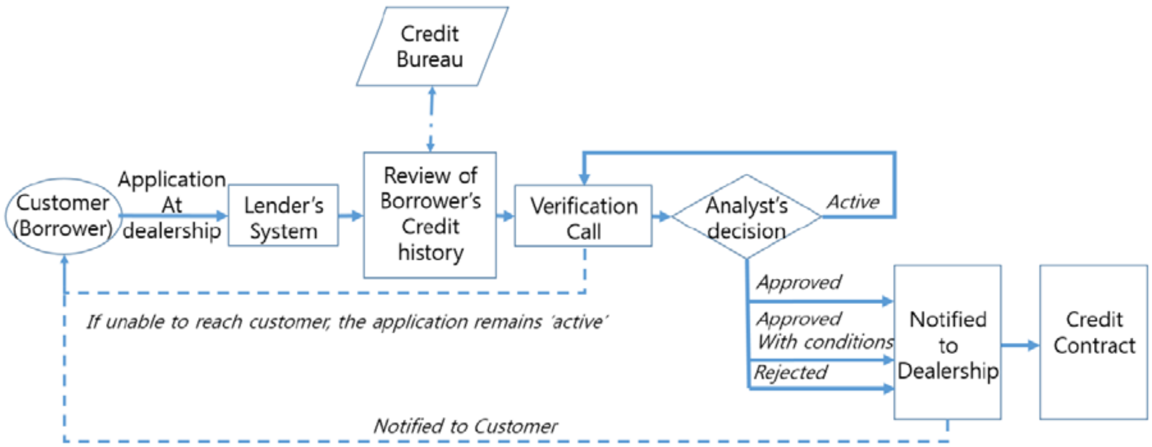

Figure 1 Process Map of Credit Contract Process

There are major milestones in the credit-contract process:

- A customer (borrower) fills out the application form at dealership, and a dealer enters key credit parameters and transcribe the form into the lender's system.

- An initial review on borrower's credit history is performed with the aid of credit bureau. Credit bureau provides personal credit history when requested by lender's company. Then lender's company verifies the credit history by calling a customer. If a call is not reached to customer, the application remains "active", and the call is continued to be attempted.

- An initial credit decision is made by analyst in the lender's company to accept the application as-is (approved), accept with conditions (approved with conditions), reject the application, or classify the application as "active." The application remains active if verification call is not reached to customer.

- If the application is approved, the lending company purchases the credit contract.

Uncovering the root causes of uncompetitive response time

Then, you had led your team to develop and implement the required data collection plans for the organization's processing centers in several Latin American countries. The team has collected data for three months by recording sales operation from auto dealers. The sales data includes basic information about the car that was sold (maker, model, price, etc.), as well as the dealer's information (monthly sales, inventories, dealer skills). The sales data is recorded in the first Excel Sheet, "I. Sales Related Variables". Note that the sales data is uniquely identified by "Sales ID" variable in the first column.

At the same time, the Six Sigma team turned to process mapping to follow the life cycle of a loan application through the lender's process, i.e. for each sales instance, the team kept track of time spent for the process elements identified in Figure 1. The time breakdown for the process is recorded in the second Excel Sheet, "II Process Time Breakdown". Note that "Sales ID" variable in the first column links the life cycle of a loan application to the related sales variables in the previous sheet.

At the draft review of collected data, the team discovered a new twist: No matter how fast the lender responded to the car dealers with credit decisions, customers weren't notified by the car dealers until significant time had elapsed. Two reasons for these delays were uncovered:

- The dealers' salespeople were busy selling cars and didn't take the time to pass along the initial credit decisions to the buyers.

- Low inventory levels caused some salespeople to delay financing notification until the vehicles became available.

With this information in hand, your team also gathered survey responses from the lending company's staffs (analysts, system engineers, support personnel, and dealers in partnership) and from the second-tier customers, borrowers, to identify potential factors that slow response time and to discover how staffs and customers differ in their perspectives. Figure 2 depicts *Fish Bone diagram* with potential factors identified by brainstorming, and these factors were surveyed to evaluate their relative impact on slowing response time.

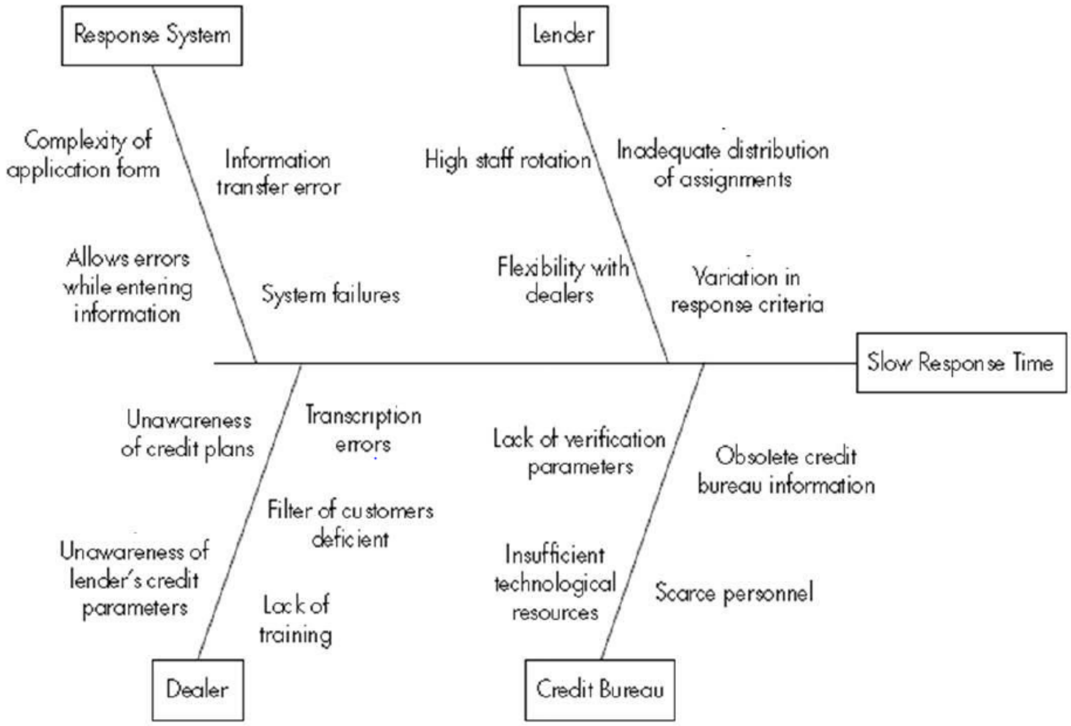

Figure 2 Fish Bone Diagram of Potential Factors Attributable to Slow Response Time

Survey responses are recorded in the third Excel Sheet, "III Survey Responses". A respondent is uniquely identified by personal ID and affiliation, either as staff or customer. For each factor, a respondent was asked to evaluate its impact on slow response time, from 1 (no impact) to 5 (extremely high impact).

Data Appraisal and Analytic Plans

Now the team of Six Sigma Black Belts is ready to perform statistical analysis using the Excel sheets, (I) Sales related variables, (II) Process Time Breakdown, and (III) Survey Responses. Since (I) and (II) have objective variables from sales operation and variable types are all quantitative, you can use them to build a *Regression model* to predict a part of, or entire response time, with time in (II) as dependent variable, and sales related variables in (I) as independent variables. Depending on the model performance, this model can predict if how much of the response time can be reduced by controlling the independent variables. On the other hand, (III) has subjective variables on potential factors in four categories. It first helps to prioritize among factors-a higher value on a factor indicates higher priority. However, you are aware that relying on a mere average is not as reliable as considering standard deviation altogether. Therefore, you decide to use *Confidence Interval* to compare the relative importance of factors to the response time. By comparing Confidence Interval of all 17 factors, you expect to determine which factors have statistically greater impact than others. In addition, comparing the mean responses between staffs and customers can reveal different perspectives from both groups of stakeholders. *Hypothesis testing* is useful to statistically diagnose if the response between staff and customer is different or not. If necessary, you can make assumptions on the parameters (e.g. alpha).

Procedural Steps of Analysis

Follow these steps to complete your analysis. Analysis outcomes can vary depending on the parameters you define, the assumption you set up, etc. Always make sure that your analysis contains the requirements stressed in bold characters.

(1) You want to determine which category in the square in Figure 2 has the greatest impact on response time, based on survey response. By comparing the confidence intervals of four categories in the dataset (III), **determine which category** among 'response system', 'dealer', 'lender', and 'credit bureau', **are the most important** to response time. You can select multiple categories if they are not significantly different in their confidence interval. Clearly define parameters if needed.

(2) Next, you want to determine which factor <u>within</u> the category in (1) has the greatest impact on response time, based on survey response. Construct confidence intervals for each factor <u>within</u> the category you identified in (1) from the dataset (III). By comparing these confidence intervals, **rank relative importance of the factors**. If any two factors are not significantly different in confidence interval, you may assign an equal rank. Clearly define your own parameters and assumptions if needed. Also assume that the survey response is normally distributed.

(3) For the dataset (II), draw a **graphical chart** that best illustrates which process element in Figure 1 is a bottle neck to total response time. You may choose a graphical chart that is suitable to visualize a bottle neck. Think

of the types of charts you learned in class. To which category among 'response system', 'dealer', 'lender', and 'credit bureau' does the bottle neck belong to?

(4) By combining the analytic results from subjective survey in (III) and from objective observation in (II), you want to identify the category that is the most critical to response time. Integrate your findings in (1) and (3), and tell if **which one among** 'response system', 'dealer', 'lender', and 'credit bureau' is **the most attributable** to slow response time. If you have conflicting results from subjective survey (III) and objective data (II), select one with a reasonable ground and justify your ground.

(5) Next, you want to build a quantitative model to predict response time. First, combine (I) and (II) in terms of unique Sales ID, which is a key to link both datasets. Then, build a regression model to predict a part or the whole response time using any sales related variables, i.e. your dependent variable can be either total response time, or partial time for process elements in Figure 1. Explore variety of candidate Y's and X's, and propose only **one regression model** with the highest fit (Adjusted-R^2) you can obtain.

(6) Include **ANOVA table** for your best fit model in (5), and **diagnose model significance**, and **test the significance of all the coefficients**. In addition, attach the **residual plots** together with your **diagnosis**.

(7) Using your regression model in (5), you want to predict the shortest response time you can achieve. **Predict the lowest time**, Y, you can achieve from the set of your independent variables, X's. Assume that your X's do not take on values outside the values given in the dataset, i.e. do not extrapolate.

(8) You want to estimate Total response time that can be saved by controlling independent variables, X's. This step is necessary to justify your model. Based on your prediction in (7), estimate **how much % of Total response time can be saved** on average if you adopt desirable X's. If needed in the course of estimation, make assumptions.

(9) Integrating your findings from (4) and (5)-(8), **propose the best set of solutions** to reduce response time. There is no limit in the number of solutions. The solution can be at the management level, e.g. hire more people, or at the operational level, e.g. set up approval criteria. Note that all solutions are logically derived from the analytic findings in (4) and (5)-(8).

(10) (Optional) For gauging the difficulty of the exam, the estimated amount of time you put into this exam?

Reporting Guidelines

- Type all your answers concisely and neatly following the steps with numbers (1) to (9). Do not forget to put the numbers to each answer. Lengthy report does not really count here. Just include all the answers that are required.
- For your output, pdf file format is highly recommended. You will submit the file through Angel Dropbox by Friday (4/9/2014) at 11:59pm. No hard copy. No email. No late work accepted.