Quantitative Effects of Chat Expansion on Chat Volume: For Finance & Planning

10/9/2014

Formal Analysis Authored by: Carole Bonner

TABLE OF CONTENTS

QUANTITATIVE EFFECTS OF CHAT EXPANSION ON CHAT VOLUME: FOR FINANCE & PLANNING

Carole B. Bonner, MSET, MSAS

ABSTRACT

The Company website's "Contact Us" page was redesigned with a prominent Chat Option September 29, 2014. This Expansion was an initiated to promote Chat and decrease Call Volume. Although chat volume increased after the September 29th expansion, call volume remained the same. A multivariate regression analysis was conducted to examine the quantitative effects of the chat expansion on call volume and reasons that the call volume remained steady. The resulting model accurately predicts peaks and troughs in chat volume at the 99.99% confidence level (F=71.40, p-value<.0001). Likewise, the model is a good representation of the variation in the process with an Adjusted R-squared of 72%. Holding the other regressors constant, the Chat Volume increased by about 4563 or about 50% per day since the expansion. This was almost twice the increase that resulted from the secondary iPhone release. Not surprisingly, each call that is answered within 20 seconds, decreases Chat Volume by 3415.15. Extrapolations cannot be made outside of the date parameter used in this particular analysis.

RECOMMENDATIONS:

It is conclusive that the Chat Expansion has effectively decreased Chat Volume and this decrease can safely be reflected by finance and planning.

QUANTITATIVE EFFECTS OF CHAT EXPANSION ON CHAT VOLUME: FOR FINANCE & PLANNING

INTRODUCTION

The Company website's "Contact Us" page was redesigned with a prominent Chat Option September 29, 2014. This Expansion was an initiated to promote Chat and decrease Call Volume. Notable dates are as follows:

- September 18th –19th   Primary iPhone Release
- September 29th Chat Expansion
- October 2nd–3rd Secondary iPhone Release
- October 3rd Louisville Weather Event caused 26,000 iPhone delays

Although chat volume increased after the September 29th expansion, call volume remained the same. This analysis explores the effects of the chat expansion on call volume and reasons that the call volume remained steady.

METHODOLOGY

A multivariate regression analysis was conducted to examine the quantitative effect of a new Chat Expansion for the company's website.

Data Cleaning and Checking Assumptions:

Before conducting the regression, the data was screened for missing data, outliers, and assumption violations. Multiple imputation procedure was used to replace each missing value with a set of plausible values that represent the uncertainty about the right value to impute, and a univariate analysis was performed.

Table 1: Variable Description and Coding

Coding	Form	Description
CallVol	Count	Call Volume
ServLvl20secs	Count	2014 Adequate Service Level
ServLvlLY	Count	2013 Adequate Service Level
AvgDurLY	Count	2013 Average Chat Duration
AHTTY	Count	2014 Average Handle Time
AHTLY	Count	2013 Average Handle Time
DOW	M-F	Weekday (M – F only)
iphonerel1	Yes=1 No=0	1st iphone release
iphonerel2		2nd iphone release
ChatExp		Chat Expansion Implemented

The residuals plots were generated and are shown in Figure1. The normal probability plot was relatively good with a slight rise on the left tail and a slight drop on the right end; this may be due to outliers but it is no cause for worry since the ANOVA is robust. Leverage plot has some points that are high and some which are low, so no action needs to be taken. Cook's D plot shows some observations that are spiked; however none of the values are greater than 1.0, so there is no need to standardize the data. The residuals versus fitted plot showed linear relationships for all variables. Box Cox Analysis indicated that a square root transformation would be appropriate, if necessary and beneficial. (α=0.5) The transformation did not significantly improve homoscedasticity.

Table 2: Basic Statistical Measures for ChatVol

Location		Variability	
Mean	8000.276	Std Deviation	1551
Median	7611.500	Variance	2406007
Mode	7322.000	Range	9691
		Interquartile Range	1355

Extreme values listed in Table 3 were observed but are in keeping with the exceptions in the model and should be included in the model.

MODEL SELECTION

Stepwise Regression based on Adjusted R-squared, a combination of forward and backward selection, was used as a model selection tool ($\alpha = 0.15$ entry and exclusion from the model). The main effects as well as the interaction terms were examined. The variables that were found to contribute significantly to the model are list in Table 4. (0<- p-values <- 0.015)

The second order interactions of the significant regressors were examined for multicollinearity using the Variation Inflation Factor (VIF_k which quantifies how much the variance is inflated when examining how variables interact with one another or between main effects. The formula for VIF is given by $VIF_k = \frac{Var(b_k)}{Var(b_k)_{min}} = \frac{1}{1-R_k^2}$ where R_k^2 is   the R^2-value obtained by regressing the kth predictor on the remaining predictors. Of course, the greater the

linear dependence among the predictor x_k and the other predictors, the larger the R_k^2 value. The larger the R_k^2 value, the larger the variance of b_k. So, if VIF_k is larger than 10, the variance is very large or inflated. Of the thirty interactions, the variables with no significant multicollinearity ($VIF < 10$) were not found to be predictors for the model. Therefore, no interactions were be added to the model.

According to the coefficients shown in Table 4, Chat Volume is became higher at the Secondary iPhone release as well as with the Chat expansion. It is also not surprising that the Chat Volume decreases when the 20 second Service Level is achieved.

The regression model is given by

$$\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 - \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7$$

The resulting regression model was found to be

$$\begin{aligned} ChatVol = {} & 1312.26 + 0.05(CallVol) \\ & - 3415.15(ServLvl20secs) \\ & + 2008.11(ServLvlLY) \\ & + 7.09(AvgDurLY) \\ & - 11.18(AHTLY) \\ & + 2973.03(iphonerel2) \\ & + 4563.41(ChatExp) \end{aligned}$$

MODEL ADEQUACY

An overall F test was computed (Table 5). The model has a good fit at the 99.99% confidence level ($F =$ 71.40, $p - value < .0001$). Adjusted R-squared is an adjustment of the R-squared that penalizes the addition of extraneous predictors to the model. Adjusted R-squared is computed using the formula $R^2 = 1 - \left(\frac{(1 - Rsq)(N - 1)}{N - k - 1} \right)$ where k is the number of predictors. Likewise, the model is a good representation of the variation in the process with an Adjusted R-squared of 72%.

The coefficients for each predictor **is** the UNIQUE effect of that predictor on the response variable. It's not the full effect unless all predictors are independent. It's the effect after controlling for other variables in the model. The coefficients in the selected model seem to be reasonable in both signs and magnitudes.

CONCLUSIONS AND RECOMMENDATIONS

Holding the other regressors constant, the Chat Volume increased by about 4563 or about 50% per day since the expansion. This was almost twice the increase that resulted from the secondary iPhone release. Not surprisingly, each call that is answered within 20 seconds, decreases Chat Volume by 3415.15. Extrapolations cannot be made outside of the date parameter used in this particular analysis.

RECOMMENDATIONS

It is conclusive that the Chat Expansion has decreased Chat Volume and this decrease can safely be reflected in the new budget.

APPENDICES

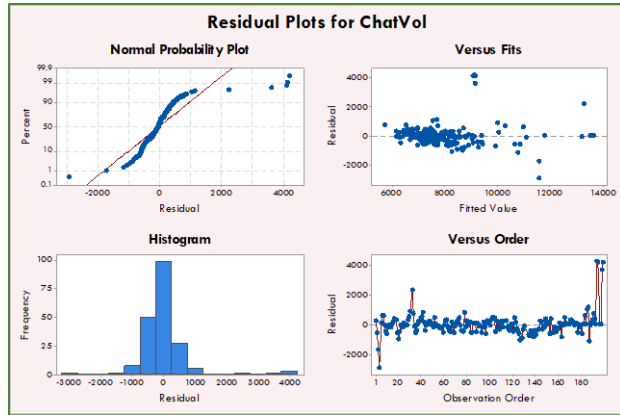## APPENDIX A: CHARTS AND TABLES

Figure 1: Residual Plots

Table 3: Extreme Observations

Lowest		Highest	
Value	Obs	Value	Obs
5821	77	13266	199
6118	13	13352	194
6128	82	13514	193
6329	12	13648	196
6337	87	15512	32

Table 4: Coefficients

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	1312.26	1626.39	0.81	0.4208
CallVol	1	0.05	0.01	9.01	<.0001
ServLvl20secs	1	-3415.15	388.18	-8.80	<.0001
ServLvlLY	1	2008.11	647.09	3.10	0.002
AvgDurLY	1	7.09	2.18	3.25	0.001
AHTLY	1	-11.18	4.28	-2.61	0.010
iphonerel2	1	2973.03	615.75	4.83	<.0001
ChatExp	1	4563.41	840.72	5.43	<.0001

Table 5: Analysis of Variance

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	341366625	48766661	71.40	<.0001
Error	187	127722364	683007		
Corrected Total	194	469088989			

Root MSE	826.44255	R-Square	0.7277
Dependent Mean	8001.74359	Adj R-Sq	0.7175
Coeff Var	10.32828		

APPENDIX B: SAS CODE

```
ODS HTML CLOSE;
DM log "OUT;CLEAR;LOG;CLEAR;" log continue ;
DM log 'next results; clear; cancel;' whostedit continue ;
ODS HTML NEWFILE=NONE;
TITLE 'CHAT ANALYSIS - REGRESSION';

LIBNAME GBS "C:\Users\cbonner4\Documents";  RUN;

PROC IMPORT OUT= GBS.CHATa
    DATAFILE= "C:\Users\cbonner4\Documents\CHAT.csv"
    DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

DATA GBS.chat;
SET GBS.chata;

********************************************
DEFINING TRANSFORMATIONS and INTERACTIONS
********************************************;
    sqchat = sqrt(chatvol);
    lchat = log(chatvol);
    callvol_DOW=callvol*DOW;
    callvol_iphonerel2=callvol*iphonerel2;
    callvol_ChatExp=callvol*ChatExp;
    DOW_iphonerel2=DOW*iphonerel2;
    DOW_ChatExp=DOW*ChatExp;
    iphonerel2_ChatExp=iphonerel2*ChatExp;

**********************************
RECODING DAY OF WEEK
**********************************;
    IF DOW=2 THEN DOW2=1; DOW3=0; DOW4=0;
DOW5=0; DOW6=0;
    IF DOW=3 THEN DOW2=0; DOW3=1; DOW4=0;
DOW5=0; DOW6=0;
    IF DOW=4 THEN DOW2=0; DOW3=0; DOW4=1;
DOW5=0; DOW6=0;
    IF DOW=5 THEN DOW2=0; DOW3=0; DOW4=0;
DOW5=1; DOW6=0;
    IF DOW=6 THEN DOW2=0; DOW3=0; DOW4=0;
DOW5=0; DOW6=0;
RUN;

%LET ylist = ChatVol;
%LET time = day;
%LET grlist = DOW2 DOW3 DOW4 DOW5 DOW6
iphonerel1 iphonerel2 WeatherEvent ChatExp;
%LET grlist2 = DOW2 DOW3 DOW4 DOW5 DOW6
iphonerel2 ChatExp;
%LET grlist3 = iphonerel2 ChatExp;
%LET xlist = CallVol ServLvl20secs    ServLvlLY DurationTY
AvgDurLY AHTTY AHTLY;
%LET lylist = lchat;
%LET sqylist = sqchat;
```

```
*********************************************
  UNIVARIATE ANALYSIS AND SCATTERPLOT
*********************************************;
PROC UNIVARIATE DATA=GBS.chat;
VAR &xlist;
RUN;

PROC FREQ DATA=GBS.chata;
BY DOW iphonerel1 iphonerel2 WeatherEvent ChatExp;
RUN;

ODS GRAPHICS ON;
PROC SGSCATTER DATA=GBS.chat;
  LABEL CallVol='CallVol' ServLvl20secs='ServLvl20secs'
ServLvlLY='ServLvlLY'    DurationTY='DurationTY'
AvgDurLY='AvgDurLY' AHTTY='AHTTY'    AHTLY='AHTLY';
  LABEL CallVol='CallVol' ServLvl20secs='ServLvl20secs'
ServLvlLY='ServLvlLY'    DurationTY='DurationTY'
AvgDurLY='AvgDurLY' AHTTY='AHTTY'    AHTLY='AHTLY';
  MATRIX CallVol ServLvl20secs    ServLvlLY    DurationTY
AvgDurLY    AHTTY    AHTLY/ DIAGONAL=(HISTOGRAM);
RUN;

**********************************
BOXCOX and TRANFORMATION
**********************************;
PROC TRANSREG DATA = GBS.chat TEST;
    MODEL BOXCOX(&YLIST) = IDENTITY(&xlist &grlist);  RUN;

PROC REG DATA = GBS.chat;
    MODEL &ylist = &xlist &grlist /PARTIAL XPX SS1 CORRB
SEQB P R INFLUENCE; RUN;

PROC GLM DATA = GBS.chat PLOTS=all ;
    MODEL &ylist = &xlist &grlist3/ P ;
    OUTPUT OUT = GBS.CHATex
    P=pred R=Residual RSTUDENT=r1 DFFITS=dffits
    COOKD=cookd H=hatvalue PRESS=res_del ; RUN;

*********************************
Selection/Validation
*********************************;
PROC GLMSELECT Data=GBS.chat Plots=All Seed=523654;
    CLASS &grlist ;
    Partition fraction(Test=0.4);
    MODEL &ylist = &xlist &grlist/Selection=Stepwise(Select=SL
SLE=0.15 SLS=0.15 Choose=AdjRSQ)Details=ALL
Hierarchy=Single Stats=ALL showpvalues;
    OUTPUT Out=GBS.glm;  RUN;

PROC MIXED DATA=GBS.chat;
    CLASS &grlist;
    MODEL &lylist = callvol | DOW | iphonerel2 | ChatExp/
HTYPE=1;
    REPEATED / TYPE=un SUB= DOW;  RUN;

ODS RTF CLOSE;
ODS GRAPHICS OFF;
```