Using SAS® Marketing Optimization to improve Campaigning at Africa's largest Pay TV service provider

Jean Murray de Villiers, MultiChoice South Africa

ABSTRACT

Traditionally, the entertainment industry incurs significant cost in the engagement of a diverse customer base with appropriate messaging and slow marketing turnaround, leading to reduced customer satisfaction levels and customer retention rates. MultiChoice South Africa, Africa's largest Digital Satellite Television service provider leverages a rich set of historical and personal customer information, in both structured and unstructured format, in order to understand and predict customer behavior and marketing campaign propensity. These predictive models are utilized to improve campaign targeting, generating significant incremental revenue, whilst reducing costs of communication. This paper describes MultiChoice's data science journey, using SAS Marketing Optimization, to campaign to the right customers at the right time, through the right communication channel to ensure that the company maximize profit, improve retention rates, and improve overall customer satisfaction.

INTRODUCTION

MultiChoice South Africa is a leading video entertainment and internet company in the African continent. The MultiChoice Group operates under various brands, including: DStv, M-Net, SuperSport, GOtv, DStv Digital Media and DStv Media Sales. MultiChoice provide a broad range of products and services that together produce entertainment to millions in Africa. In order to enjoy MultiChoice entertainment requires a customer to subscribe to an entertainment package and acquire a physical electronic decoder device that, in combination, provide access to content in the traditional viewing way. MultiChoice campaign to their broad customer base via a number of communication channels. Campaign types include standard operational messaging, educational or informative messaging, and marketing campaigns that proactively address marketing concerns such as to support upsell, cross sell and retention strategies. Propensity models are generated to provide the marketing team with proactive intelligence around their customers' likely future behavior and serves as input into the marketing optimization process. The complete data science journey is described in a high-level overview with the focus of this paper on the model generation and campaign optimization phases and all the procedural steps followed.

OBJECTIVE

The overall objective of the data science journey is to provide the marketing optimization engine that is utilized to achieve significant increases in campaign revenue, and reductions in communication volumes and fulfilment costs of campaigns using SAS® Marketing Optimization.

CHALLENGES ON THE ROAD TO MARKETING OPTIMIZATION

Challenges faced on our journey include working with Big Data (Volume, Velocity, Variety, Variability), Integration across different platforms, Complex Business Requirements, and Analytical Challenges.

BIG DATA

Volume – A huge amount of customer and product data is collected across various platforms.

Velocity – Data is captured daily and is stored, processed, and transformed into actionable information.

Variety – Data is source in a structured as well as unstructured format and requires the appropriate data pre-processing.

Variability – The introduction of new information systems and product offerings over time produces data that vary over time.

INTEGRATION ACROSS DIFFERENT PLATFORMS

Integrating between various systems pose its own challenges, including working with multiple project teams, each with their specific project goals and success criteria. There are various integration points between the SQL Server, the 3rd Party Campaign Management platform and the SAS® stack of solutions. The 3rd Party Campaign Management platform is utilized for campaign generation and campaign execution. The SQL Server environment facilitate the data processes and data storage. The SAS stack of solutions is used to generate propensity models, score new data, and optimization campaign lists.
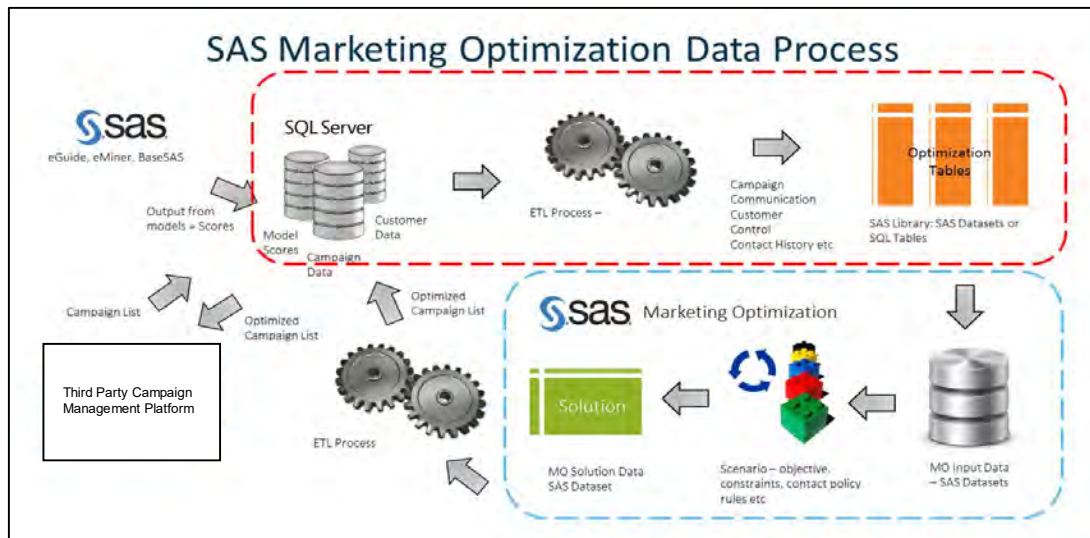
Figure 1. SAS Marketing Optimization Data Process Utilized

COMPLEX BUSINESS REQUIREMENTS

The business dynamics in a modern entertainment vendor includes complex business requirements that govern campaign processes, contact policies, as well as quarantine rules that define and limit the amount of contact sessions allowed over a defined time period. Customers migrate between packages and states *(Active / Inactive)* throughout their tenure. In addition, operational and budgetary constraints need to be included in the planning and execution of marketing campaigns.

ANALYTICAL CHALLENGES

Opportunistic Data – Data that is used for the predictive models is opportunistic by nature. It is gathered from operational & billing systems and is unrelated to typical statistical analysis, massive in nature, contains errors and outliers, missing values, etc. and therefore needs thorough data preparation prior to model building and optimization (Patetta, 2008).

Mixed Measurement Scales – Input variables have a variety of measurement scales and need to be standardized prior to model building.

High Dimensionality – The data used for model building has a large amount of input variables. This limits the ability to explore and model the relationships among variables and is known as the *curse of dimensionality*. This is addressed through the application of proper variable reduction and selection techniques (Patetta, 2008).

Rare Target Event – Rare target events are observed in initial samples drawn from the population data and pose risks to reliable model building. Separate Sampling is used to draw samples that disproportionally over-represent the target events. The priors produced when scoring new data is adjusted for over-sampling to ensure that there are no biases.

Model Selection – The champion model is selected by selecting the most parsimonious model from the set of models build for each propensity model type.

THE DATA SCIENCE JOURNEY

In this section the Data Science journey will be discussed from a high level view. Figure 2 below, provides a high level overview of the data science journey.
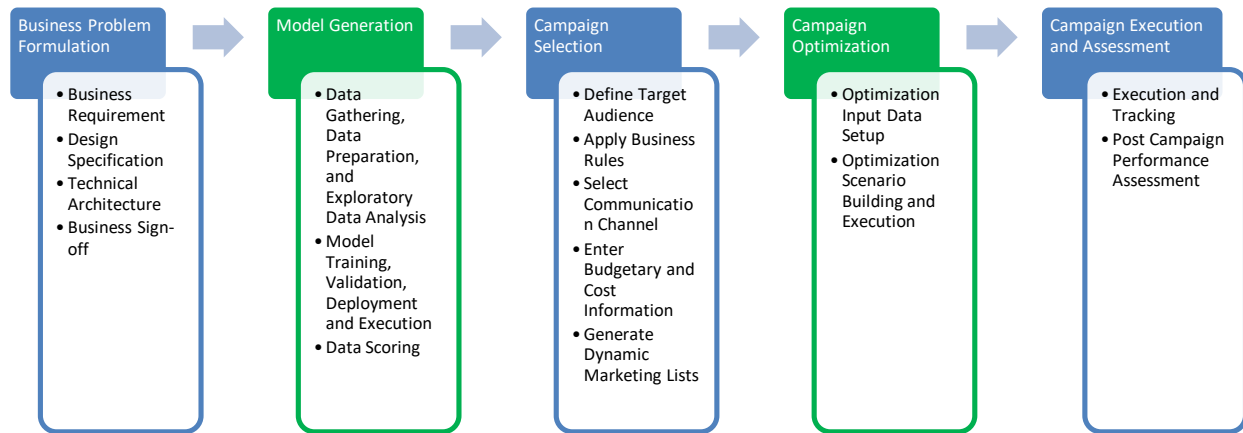
HIGH LEVEL OVERVIEW

Business Problem Formulation	Model Generation	Campaign Selection	Campaign Optimization	Campaign Execution and Assessment
• Business Requirement • Design Specification • Technical Architecture • Business Sign-off	• Data Gathering, Data Preparation, and Exploratory Data Analysis • Model Training, Validation, Deployment and Execution • Data Scoring	• Define Target Audience • Apply Business Rules • Select Communication Channel • Enter Budgetary and Cost Information • Generate Dynamic Marketing Lists	• Optimization Input Data Setup • Optimization Scenario Building and Execution	• Execution and Tracking • Post Campaign Performance Assessment

Figure 2. Data Science Journey Overview

PREPARATION AND PLANNING

The first step covers preparation and planning for the journey. This step is important to ensure that we defined the project correctly and plan for all this steps in this journey accordingly. This is done by understanding business needs and documenting it in a formal process. There is also a need to generate the necessary support documentation to support the plan and its related verification checks. In the present case, this step was done in conjunction with SAS Institute South Africa in a formal project proposal process.

MODEL GENERATION

SAS® Enterprise Guide is used for model generation and the following steps are followed:

1. Data preparation steps include data access, explorative data analysis, data cleaning, data manipulation, and feature engineering.

2. Model building and deployment include model training, validation, deployment, and execution.

3. Data scoring include the scoring of new data on regular time intervals to support the campaign management and marketing optimization processes.

CAMPAIGN SELECTION

The 3rd Party campaign management platform defines the lists of customers for each campaign by following these steps:

1. Campaign target selection that filters campaigning data based on customer demographics and package-device combinations.

2. Quarantine rules are applied to check offer eligibility based on previous communications.

3. Business rules are applied that adhere to governance requirements, as well as definitive customer and product rules that covers customer segmentation filtering.

4. Communication channels are selected.

5. Campaign cost and budgetary information is captured.

6. Dynamic marketing lists are generated for each campaign to be included in the optimization period and passed on to SAS® Marketing Optimization for optimization.

CAMPAIGN OPTIMIZATION

SAS® Marketing Optimization is used to for campaign list optimization. SAS® Marketing Optimization is a component within the greater SAS® Customer Intelligence solution. On a customer level, SAS® Marketing Optimization is fed propensity scores and the revenue and cost associated information related to campaigning for optimization. This information along with key optimization metrics are used to build optimization scenarios.

CAMPAIGN EXECUTION AND ASSESSMENT

At this stage of the journey the optimized marketing campaign lists are passed back to the 3rd party campaign management platform where the campaigns are executed and tracked to assess campaign performance. The campaign performance information is passed back to the SQL server and made available to the SAS platform for post optimization reporting and future modelling.

PROPENSITY MODEL GENERATION

Propensity models are generated using SAS Base code within SAS® Enterprise Guide 6.1. The following four phase explain the modelling process that generate the campaign associated propensity scores that stored in the SQL database and used during the optimization stage of the journey. The creation journey for the models is discussed in the below section. One of the models, the Premium Voluntary Disconnect model, is further dissected to provide more details around the process.

OVERVIEW OF THE PROPENSITY MODELS

The following propensity models are generated to support marketing efforts. These propensity models estimate the likelihood for a customer to move from one state to another as a propensity score for a defined period (typically a calendar month).

The following package offerings are included:

- Premium
- Compact Plus (Previously known as Extra)
- Compact
- Family
- Access

The following campaign type associated propensity models exist for each of the above packages. If a customer is on the top package, Premium, he does not have a propensity score to move to a higher package and visa-a-versa for the lowest package, Access. Similar logic follows for customers already on the EXPLORA™ decoder and customers that have already moved into an inactive state.

Campaign Type	Model Type	Description
Stay Connected	Retain	This model estimates the propensity of a customer to remain a customer on the same package for the upcoming month.
Package Upgrade	Upgrade	This model estimates the propensity of a customer to upgrade to a package of higher financial value for the upcoming month.
Package Rightsizing	Right-Sizing	This model estimates the propensity of a customer to downgrade to a package of lower financial value for the upcoming month.

| Device Upgrade to the EXPLORA Decoder	Upgrade to EXPLORA	This model estimates the propensity of a customer to upgrade their decoder device to the EXPLORA for the upcoming month.

Table 1. Campaign Types and their Associated Model Type

DATA GATHERING AND PREPARATION

In the sections below describe the data gathering and preparation stage in further details.

Data Gathering and Defining Sources

Data is gathered from multiple sources systems and collated in a single repository that helps support analytical models generated, CRM processes, and ultimately the marketing optimization process. The data resides in a SQL database environment and the SAS environments access these data via an ODBC connection defined within SAS® Management Console for each environment. The display below illustrates the four stages with their associated SAS programs.
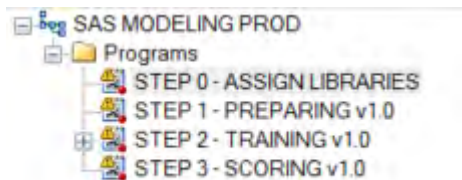
Display 1. Modeling Stages

In the first step, STEP 0, the necessary SAS libraries are assigned to ensure that the right library definitions and database connections are active. The following SAS macro programs are used to complete the second step in the model generation process:

Display 2. Data Preparation Macro Program

Preparing the Tables and Filter

The following customer information is sourced from five distinct tables:

- Demographic
- Geographic
- Tenure
- Interactions
- Subscriptions type
- Decoder type
- Payment status
- Package and device migration made during the last 30 days

The last 12 months' worth of information is used for the model training window. The data is further filtered to only contain information for the five package offerings of interest.

Creating Target Variables

The target variables are derived from migration data from one month to another for each of the models. The targets are derived in such a way to indicate a migration that happened in the month to follow.

Merging Tables

The five fact tables are merged by the row identifiers information date and customer numbers to form one flat table.

Collapsing Categorical Levels

Categorical variables with levels that have a frequency of less than 5 percent are collapsed with other smaller or larger groups. Business rules are considered with the collapsing of categorical groupings to ensure that the newly formed groups still adhere to the business logic. This reduces the dimensions of the model and allows the model to better classify more distinctly.

Sampling

Rare target events are observed in initial samples drawn from the population data and pose risks to reliable model building. Therefore, the data samples are oversampled by means of separate sampling. This sampling technique is standard practice in supervised classification and draws samples that disproportionally over-represent the events relative to non-events when compared to the population dataset and is believed to lead to better predictions (Scott and Wild, 1986). The priors produced when scoring new data is later adjusted to correct for over-sampling and ensure that there are no biases. The SAS code below is used to create separate samples:

```
proc sql;
    create table customer_P_V_C_E  as
    select * from &lib..APC_Active
    where Flag_Vol_Disconnect = 1 and product_current = "01Premium";

    create table customer_P_V_C_N as
    select * from &lib..APC_Active
    where Flag_Vol_Disconnect = 0 and product_current = "01Premium";
quit;
```

For each of the models, the samples were divided into training and validation *(holdout)* datasets for honest assessment purposes. The SAS code below is used to create the training and validation samples:

```
proc surveyselect data=Premium_Sample samprate=0.70 seed=49201
out=Premium_Sample outall
        method=srs noprint;
run;
Proc SQL;
    Create Table &lib..Premium_Sample_T as
    Select *
    From Premium_Sample
    where selected = 1;

    Create table &lib..Premium_Sample_V as
    Select *
    From Premium_Sample
    where selected = 0;
Quit;
```

EXPLORATORY DATA ANALYSIS

Assumption Validation

The assumption that the logit has a linear relationship with the predictor variables were verified visually by graph. The logits by the predictor variables where plotted and a linear relationship between the logit and predictor variables where observed.

MODEL METHODOLOGY

Model performance and interpretability are the two main criteria that is taken in consideration when deciding on the modelling method used. Decision trees, logistic regression, and neural networks were explored in the training phase for all models and although decision trees and neural networks in some instances slightly out-performed their logistic regression counterpart, for the sake of interpretability of the parameters by means of odds ratios, logistic regression was selected for the first iteration of the models.

MODEL TRAINING, VALIDATION, DEPLOYMENT, AND EXECUTION

This phase of the Data Science journey is an iterative process and the models are regularly revised and improved to ensure that they remain relevant to the current business rules and ever-changing business environment.

Model Training and Variable Selection

The following SAS code is used for model training and variable selection:

```
proc logistic data=&lib..premium_sample_T
outmodel=&lib..Premium_VC_Betas;
    class device_current PrimaryInteractionType MethodOf
    (param=ref ref='Pay_3mAvg_Retailer') AssumedLanguage
    (param=ref ref='english') SurnameIndicator
    CustomerSegmentationDesc information_month (param=ref
    ref='4')month_movement;
    model flag_vol_disconnect (event="1") = device_current
    revenue_current tenure_months information_month
    PrimaryInteractionType MethodOf SurnameIndicator
    CustomerSegmentationDesc month_movement/
    selection=backward;
    output out=PVC_Train_out p=ppred;
```

```
        score data=&lib..premium_sample_V out=PVC_Val_out
        outroc=pvc_vroc fitstat;
    run;
```

The backward stepwise variable selection method is used for variable selection within logistic regression procedure by stating the selection equals backward option in the model statement. Output 1 shows the remaining variables in the model after backward variables selection was conducted.

Note: No further effects can be removed from the model without violating the hierarchy requirements.

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Device_Current	5	5596.9617	<.0001
Revenue_Current	1	11895.5758	<.0001
Tenure_Months	1	8455.3519	<.0001
Information_Month	11	20744.0367	<.0001
PrimaryInteractionTy	8	62647.7542	<.0001
Methodof	6	14486.6369	<.0001
SurnameIndicator	2	5619.6241	<.0001
CustomerSegmentation	2	62344.6961	<.0001
Month_Movement	2	2219.1036	<.0001

Output 1. Variable Selection Result

Model Validation

The holdout sample for each model is used for model validate. Model discrimination and model calibration ability is measured using the following statistics.

Discrimination

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	86.5	Somers' D	0.729
Percent Discordant	13.5	Gamma	0.729
Percent Tied	0.0	Tau-a	0.365
Pairs	87364165889	c	0.865

Output 2. Association of Predicted Probabilities and Observed Responses

The following measures are used to determine how well the model discriminates between events and non-events. The area under the ROC curve (AUC / C statistic) plots true positive rate (sensitivity) against the false positive rate (1-specificity). For the model at hand the observed C-statistic (AUC) is 0.865 which is deemed a favorable score. The Somer's D or better known as Gini statistic measures the degree to which the model has better discrimination power than the model with random scores. It is calculated as twice the AUC minus 1 and the general rule of thumb is that it should be greater than 0.4. A score of 0.729 was observed for the model and therefore indicate a model that has strong discrimination power (Bhalla, 2015).

Calibration

Fit Statistics for SCORE Data						
Data Set	Total Frequency	Log Likelihood	Error Rate	AIC	AICC	BIC
PDATA.PREMIUM_SAMPLE_V	253365	-117447	0.2111	234971.1	234971.1	235378.3

Data Set	SC	R-Square	Max-Rescaled R-Square	AUC	Brier Score
PDATA.PREMIUM_SAMPLE_V	235378.3	0.368204	0.490941	0.863039	0.14942

Output 3. Fit Statistics

The brier score is an important measure of calibration and looks at the mean squared difference between the predicted probability and the actual outcome. The lower the Brier score is for a set of predictions, the better the predictions are calibrated (Bhalla, 2015). The model has a Brier Score of 0.14942 which indicate that the model is well calibrated.

Model Deployment and Execution

The champion models are promoted from the Development environment to the Quality Assurance (QA) environment for final testing. Once they have passed all final quality assurance tests they are promoted to the Production Environment. The propensity models are executed in SAS® Enterprise Guide manually at this stage by running the SAS scoring script and final scores are written back to the SQL Environment to be made available for Marketing execution processes.

DATA SCORING

At the beginning of every month, the previous months' data is used for scoring. The data used for scoring go through the similar data preparation process as explained in the data preparation section, without creating target events. The SAS® Logistics regression procedure is used to score new data and the scores are corrected for oversampling using the *prior event* option in the score statement. Finally, the scores from all the models are collated into a single flat file and is written back to the SQL Database. The SAS code below is used to score the Premium Voluntary Disconnect model:

```
%let pi1 = 0.0049;
proc logistic inmodel=&lib..Premium_VC_Betas;
     score data=&lib..P_Score_&YM. out=PVC_Score_&YM._out
     priorevent=&pi1;
run;
```

CAMPAIGN SELECTION

This stage of the marketing campaign process is covered by the 3rd party campaign management solution and produces dynamic campaign marketing lists. If the marketing lists are eligible for optimization, they go through an ETL process that structure the data in the correct table format before making them available to SAS® Marketing Optimization.

CAMPAIGN OPTIMIZATION

The SAS® Marketing Optimization component within the SAS® Customer Intelligence stack is used to optimize the marketing lists. On a customer level, SAS® Marketing Optimization is fed propensity scores and campaign associated information. This information along with key optimization metrics are used for the creation of optimization scenarios.

OPTIMIZATION OVERVIEW

Typically, the campaign process starts with a brief from a marketer, defining a campaign that has been planned, to support an aspect of a broader marketing strategy. This step is covered in the campaign selection section. If only a single campaign is planned and there are associated propensity model scores, then a customer list is selected by means of prioritizing customers with higher scores relative to the rest as part of the campaign selection criteria. If customers are eligible for competing marketing efforts such as: high likelihood to churn, likely to upgrade their decoder to that latest version, or likely to upgrade their packages and constraints need to be included such as contact policies, communication constraints, quarantine rules, etc. then a more sophisticated approach such as mathematical optimization need to be employed for campaign selection.

OPTIMIZATION ENGINE

Qualifying customers are identified for each of the campaigns for a defined period, and assuming that there is an overlap of customers between these campaigns, then these campaigns are promoted into the optimization process. A central optimization process, which involves both functional and business processes, is conducted using the scenario analysis. The set of campaigns is then optimized based on the business rules, contact policy rules and various constraints included in the selected scenario.

SAS® Marketing Optimization present marketers with the expected outcome of each set of objectives, rules and constraints, that allow them to make an informed decision as to the effects of each constraint or strategic decision applied. This allow the marketer to choose the most ideal scenario which maximizes the objective while also considering the constraints applied. The outputs from this process are sets of optimized campaign lists, which is submitted for execution to the channels via the third party campaign managers.
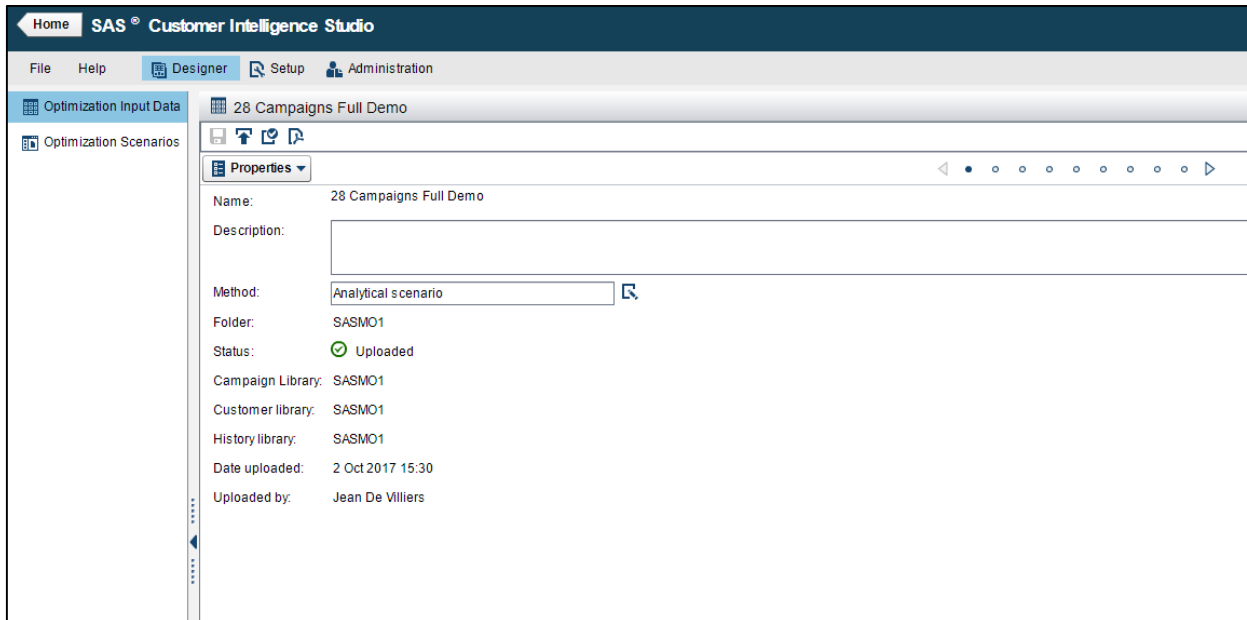
OPTIMIZATION PROBLEM

During the time of completion of this paper, the project was in final testing. The following describes the optimization problem addressed to stress-test the system:

- Twenty-eight campaigns to be included. The twenty-eight individual campaigns comprised of the four campaign types and the five packages (Premium, Compact Plus, Compact, Family, Access) within. The breakdown of the individual campaigns is further explained in the propensity model generation section.

- All of the active customers in the five packages (5,422,501 eligible offers).

- Each customer is only allowed to be assigned to one campaign.

- Communication channels include Email and SMS. Each campaign has a single communication channel so that there is a 1-1 relationship between campaigns and communications. This means that customers will either be placed in an email or SMS campaign but not both.

- Model scores for propensity models are on a customer level and linked to each of the campaigns included for optimization.

- Optimization objective is to maximize propensity scores.

- Optimization period covers a 7-day window.

- Optimization contact policy allows customers to be contacted once every seven days.

- Quarantine rules are applied in the 3rd party campaign management solution and the lists that are passed on to SAS® Marketing Optimization has already passed quarantine.

- Stay Connected campaigns should have a minimum of 10 000 offers per campaign as a business rule constraint.

Optimization Input Data

The first step is to configure and structure the input data before scenarios can be built. This part includes loading data from predefined SAS Libraries, configuring the tables needed for optimization, selecting the right customer attributes and table settings, etc.


Display 3. Optimization Input Data

Optimization Scenarios

In the second step optimization scenarios are built and executed. The optimization problem is translated into an optimization scenario and covers the following steps:

Input Data Measures

In this step the input data measures are viewed. This measures page allows users to create new calculated measures as well as indicator variables. The propensity scores and value measures are used to define the optimization objective function.
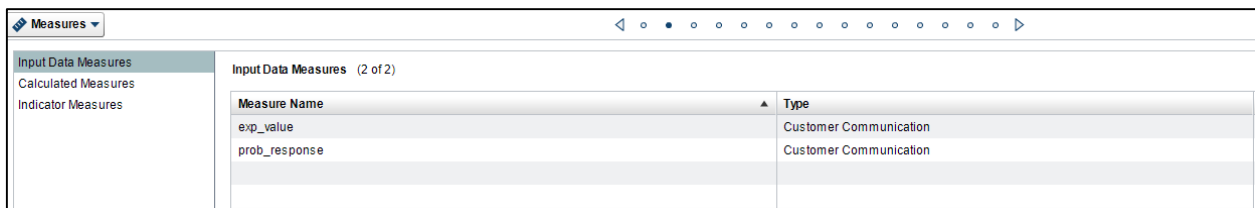
Display 4. Input Data Measures

Suppression Rules

No suppression rules are applied for the communication type and customer attributes in the suppression rules page. The quarantine rules are applied in the 3rd party campaign management solution.

Optimization Objective

Display 5. Optimization Objective

For this optimization scenario the overall objective is to maximize the propensity scores over all customers and all campaigns.

Optimization Constraints

Type	Name	Operator	Limit	Value	Opportunity Cost
Cellsize	Limit Stay Connected - Access	At most	10,000.00	10,002.00	0.19
Cellsize	Limit Stay Connected - Compact	At most	10,000.00	9,985.00	0.14
Cellsize	Limit Stay Connected - Extra	At most	10,000.00	9,999.00	0.10
Cellsize	Limit Stay Connected - Family	At most	10,000.00	10,001.00	0.06
Cellsize	Limit Stay Connected - Premium	At most	10,000.00	10,003.00	0.09

Display 6. Stay Connected Constraints

A limit of 10,000 customers per Stay Connected campaign is created in the Constraints Page. Note that this is a soft constraint that allow some flexibility around the final numbers allocated to each Stay Connected campaign.

Max/Min Contact Policies

✓ Use max/min contact policies in optimization					
Name	**Prepopulated**	**Operator** ▲	**Limit**	**Communication Filter**	**Customer Filter**
New Max/Min Contact Policy	No	At most	1		

Display 7. Max/Min Contact Policies

Each customer is only allowed to be contacted once every 7-days. Therefore, each customer is only allowed to be assigned to one of the twenty-eight campaigns.

Blocking Contact Policies

No blocking policies are applied.

Solution Setup

No solution setup is created and therefore the default solution tables is produced. This page allows users to customize the final fields they want to be included in the output tables.

Execute Optimization Scenario

Status	Method	Start Time	End Time	Elapsed Time	Log	Promoted Solution	Published Library	Published LASR Solution	Modified By	Objective
⊘ Optimized	Manual	23 Feb 2018 12:47	23 Feb 2018 12:48	0:53	View	None	None	None	Jean De Villiers	212,782.87
⊘ Optimized	Manual	21 Nov 2017 15:41	21 Nov 2017 15:42	1:29	View	None	None	None	Jean De Villiers	70,890,537.67
⊘ Optimized	Manual	5 Oct 2017 09:51	5 Oct 2017 09:52	1:04	View	None	None	None	Jean De Villiers	70,890,537.67
⊘ Optimized	Manual	4 Oct 2017 15:20	4 Oct 2017 15:22	1:05	View	None	None	None	Jean De Villiers	70,890,537.67
⊘ Optimized	Manual	3 Oct 2017 08:19	3 Oct 2017 08:20	1:06	View	None	None	None	Jean De Villiers	70,890,537.67
⊘ Optimized	Manual	2 Oct 2017 15:47	2 Oct 2017 15:47	0:53	View	None	None	None	Jean De Villiers	212,782.87
⊘ Optimized	Manual	2 Oct 2017 15:35	2 Oct 2017 15:36	1:06	View	None	None	None	Jean De Villiers	70,890,537.67

Display 8. Optimization Page

The scenario is optimized manually by selecting the optimize scenario button. Optimizations can be scheduled and the history tab within the optimization page allow users to track the status of the optimization and view other optimization execution information. The below display shows that the total elapsed time for running the optimization is 0:53 seconds. The green check in the status field indicate that the optimization ran successfully.

Optimization Results

Display 9. Summary Report

SAS® Marketing Optimization produce a summary report that includes the objective summary and constraint summary. It also produces summary reports by offer category, channel description, cost contact, cost incentive, offer revenue, sub category, campaign code, campaign description, channel code, and communication code. The following optimization results for the optimized offer by channel, sub category, and campaign description is further discussed below. Note that the campaign reduction calculations are based on the assumption that all eligible offers would been campaigned to.

Channel_Description	Eligible Offers	Offers
EMAIL	13,471,259	1,570,700
SMS	16,209,233	3,851,801

Output 4. Optimized Channels

The following is the optimization results viewed by channel description. The output shows that before the optimization, there were 13,471,259 eligible Email offers and 16,209,233 eligible SMS offers. After the optimization that number was reduced to 1,570,700 Email offers and 3,851,801 SMS offers. This implies an 88.3% reduction in Email offers and a 76.3% reduction in SMS offers.

Sub_Category	Eligible Offers	Offers
Right-sizing	10,636,376	2,205,867
Stay Connected	5,422,454	49,990
Upgrade Package	10,845,784	1,874,685
Upgrade To Explora	2,775,878	1,291,959

Output 5. Optimized Sub Category

The following is the optimization results viewed by sub category. Right-sizing is reduced from 10,636,376 eligible offers to 2,205,867 offers, a 79.3% reduction. Stay connected is reduced from 5,422,454 eligible offers to 49,990 offers, a 99.1% reduction. Upgrade Package is reduced from 10,845,784 eligible offers to 1,874,685 offers, an 82.7% reduction, and Upgrade to EXPLORA™ is reduced from 2,775,878 eligible offers to 1,291,959 offers, a 53.5% reduction. These reductions in offers suggest great cost savings.

campaign_desc	Eligible Offers	Offers
Right-sizing – Compact To Access	2,349,607	1,035,028
Right-sizing – Compact To Family	2,349,607	41,885
Right-sizing – Extra To Access	170,542	234
Right-sizing – Extra To Compact	170,554	116,545
Right-sizing – Extra To Family	170,542	415
Right-sizing – Family To Access	246,860	210,995
Right-sizing – Premium To Access	1,294,597	145
Right-sizing – Premium To Compact	1,294,689	234,729
Right-sizing – Premium To Extra	1,294,689	564,218
Right-sizing – Premium To Family	1,294,689	1,673
Stay Connected – Access	1,360,791	10,002
Stay Connected – Compact	2,349,607	9,985
Stay Connected – Extra	170,554	9,999
Stay Connected – Family	246,813	10,001
Stay Connected – Premium	1,294,689	10,003

Output 6. Optimized Campaign Offers Part I

Upgrade Package – Access To Compact	1,360,791	1,228,017
Upgrade Package – Access To Extra	1,360,791	102
Upgrade Package – Access To Family	1,360,791	115,517
Upgrade Package – Access To Premium	1,360,650	7,153
Upgrade Package – Compact To Extra	2,248,871	38,998
Upgrade Package – Compact To Premium	2,248,871	423,030
Upgrade Package – Extra To Premium	164,484	36,004
Upgrade Package – Family To Compact	246,860	15,839
Upgrade Package – Family To Extra	246,836	60
Upgrade Package – Family To Premium	246,839	9,965
Upgrade To Explora - Compact	2,047,912	800,681
Upgrade To Explora - Extra	130,816	7,357
Upgrade To Explora - Premium	597,150	483,921

Output 7. Optimized Campaign Offers Part II

Output 6 and 7 shows all of the eligible offers within each of the campaign prior to optimization and the optimized offers at the right. Due to the soft constraint of at least 10,000 offers placed on each of the Stay Connected campaigns, all of their offers are either more than 10,000 or very close to that number such as 9,999 offers for the Stay Connected – Extra campaign.

CAMPAIGN EXECUTION AND ASSESSMENT

The optimized marketing campaign lists are passed back to the 3rd Party campaign management via the platform where the campaigns are executed and tracked to assess campaign performance. Campaign performance information is passed back to the SQL server and made available to the SAS platform for post-optimization reporting and future modelling.

CONCLUSION

In this paper a Data Science journey is covered using SAS software to generate the necessary propensity models to support the marketing campaigning process, and the optimization of campaigns to ensure that the right campaign is selected for each customer, utilizing the right communication channel to ensure that the company maximize profit, improve retention rates, and improve overall customer satisfaction.

The different stages of the journey are viewed from a high level and the challenges faced are also discussed briefly. The propensity models generated provides the necessary scores that is utilized by SAS® Marketing Optimization along with the necessary business rules and constraints to optimize campaigns. The results from the 28 campaign stress test promise great cost reduction through cleverly picking the most suited campaign for each customer. The stress test further concludes that a basic marketing optimization scenario that includes 28 campaigns with some basic constraints applied on a customer list of over 5 million can run within minutes. This stress-test shows the immense power of SAS® Marketing Optimization that we believe will ensure that the company maximize profit, improve retention rates, and improve overall customer satisfaction in the months to come.

REFERENCES

Bhalla, D. Check Predictive Model Performance. Listen data. Accessed on 7 November 2017. https://www.listendata.com/2015/01/model-performance-in-logistic-regression.html

Patetta, M. 2008. Predictive Modeling Using Logistic Regression Course Notes. Cary, NC: SAS Institute Inc.

Scott, A. J. and Wild, C. J. 1986. "Fitting Logistic Regression Models under Case-Control or Choice

Based Sampling", Journal of the Royal Statistical Society B, 48, 170-182.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Jean Murray de Villiers
Jean.devilliers@multichoice.co.za or jmdevilliers@gmail.com