Paper AA14-2014

Understanding Change through Different Methodological Lens

Jie Liao, M.S. Statistics

ABSTRACT

The search for the best methods to study change using longitudinal data has been a major concern for analysts in most industries, in particularly the retail industry. This article offers different examples of available techniques to analyze the process of change. First, the article offers examples of how conventional models based on generalized linear models (GLM) can be fitted to longitudinal data using a variety of different procedures, from paired t-test to repeated-measures analysis of variance (rANOVA). Second, the article will discuss similarities and differences between these techniques. The general purpose of this article is to provide a demonstration that the study of change can be analyzed through different methodological lens. The most immediate goal is to help analysts implement these techniques as a useful way to examine change in their area of interest.

1. INTRODUCTION

Change is the only constant. And in no industry is that more apparent than retail. With higher expectations, new kinds of customers, new kinds of competition, you might be hard pressed to find an industry facing more change. As retailing continues to evolve, there is no doubt that the analysis of change has also become a frequent and often necessary sophistication for retailers to remain competitive.

While the study of change is of obvious importance to the retail industry and despite the fact that longitudinal methods have enjoyed a remarkable period of growth in the past three decades, it is somewhat surprising to find the change methodology most often used in the retail industry by analysts is usually limited to pre-post comparisons. This chapter is intended to introduce some not so novel change models. Numerical examples are used to convey the main assumptions and procedures used, and all analyses are carried out using procedures from the SAS/STAT® module. We first begin with some perspective on change.

PERSPECTIVE ON CHANGE

Longitudinal data, also known as repeated measures, is collected by taking time-sequential records of responses on the same subject. With the collection of longitudinal data, analysts are interested in investigating the process of change for whatever topic of interest. For example, in social science, the objective might be to understand better whether juvenile recidivism drops after incarceration. In the retail industry, it could be an attempt to understand better customer shopping behavior as a function of receiving rewards coupons. Given these types of questions, some analysts are quick to employ pre/post, two occasion change analysis to address the question of change. Still, even fewer analysts would venture beyond simple pre/post methods to provide solutions to the change questions of interest. So what happens if the question of change extends beyond explaining the process of change beyond two occasions? What then?

The way we explain change from a statistical perspective is directly related to the methodology we used. Because the questions we ask about change vary so widely, so too must the change methodology. Using one or two conventional techniques may not be sufficient to address the topic of interest adequately. In the sections below, we describe a number of different change methodologies and the data used in our analyses.

2. METHODS

2.1 DATA DESCRIPTION

To demonstrate the selected change methodologies, we generated sample data of a hypothetical retail company that has a customer loyalty program. The data used in this paper is generated from a rewards campaign which is conducted by a retailer to stimulate consumption. During the campaign, the retailer sends out coupons every month to customers who purchased over a certain amount in the prior month. The more these customers spent, the larger the coupon amount they would earn.

For analysis purpose, the retailer tracked down every customer who received the coupon (known as fulfiller) -- among which some redeemed the coupon (known as redeemer), while others did not (known as non-redeemer) -- to see if these redeemers are more likely to be better customers than expected. The expected level is approximately by the non-redeemer group. Total number of tracked customers reaches approximately 40,000.

There are two waves of data, each in a 12-month timeframe where repeated measures are taken on a monthly basis. One wave is before the campaign launch, that is, in pre-launch phase. It mainly consists of customers' monthly transaction data, e.g. monthly spending/ monthly number of trips. The other wave is after the campaign launch, that is, in post-launch phase. It contains not only customers' transaction data, but also their fulfillment and redemption information.

2.2 VARIABLES

Variables for analysis fall into two categories: demographic and financial (Table 2.1).

Table 2.1 List of used variables and the corresponding categories

CATEGORY	TYPE	VARIABLE NAME	VARIABLE DESCRIPTION
Demographic	Independent Variable	redeem_ind	redeemer indicator, with 1 for redeemer, -1 for non-redeemer
Financial	Dependent Variable	spend1, spend2, . . . , spend24	spending in the corresponding months;
	Independent Variable	pre_multi_chnnl_ind	multi-channel shopper indicator in pre-launch phase, with 1 for multi-channel shopper, 0 otherwise;
		pre_trips	number of purchase trips made in pre-launch phase, standardized with mean 0 and std 1;
		pre_ads	average dollars spent per trip in pre-launch phase, standardized with mean 0 and std 1;
		pre_recency	time since last purchase (referred to as recency) in pre-launch phase, standardized with mean 0 and std 1;
		pre_3M_store_spend	most recent 3-month in-store spending in pre-launch phase, standardized with mean 0 and std 1;
		pre_credit_delta	credit limit change by the end of pre-launch phase, standardized with mean 0 and std 1
		pre_credit_limit_original	original credit limit, standardized with mean 0 and std 1;

We want to study the change of customer spending over pre- and post-launch phases. Specifically, we want to test whether redeemers outperform non-redeemers while holding certain variables constant. In addition, we would like to know whether redeemers would spend more in the post-launch period than non-redeemers.

3. MODELS: PAIRED T-TEST

3.1 OVERVIEW

A paired t-test is conducted when data sample consists of matched pairs of similar subjects, and the difference between paired means is of interest. It has several assumptions: (1) dependent variables should be continuous, (2) subjects are independent, (3) no evident outliers of paired difference exist, (4) paired differences should approximately follow normal distribution. Paired t-test is quite robust to violation of the normality assumption as sample size increases. The degree of freedom for paired t-test equals (n/2 – 1), where n denotes the total number of observations (see equation (1)).

$$t = \frac{d}{S_d/\sqrt{\mathrm{n}/2}} \sim t(\mathrm{n}/2 - 1) \tag{1}$$

where $d = \sum_j^{n/2}(Y_{1j} - Y_{0j})\,/(n/2)$, $S_d = \sqrt{\sum_j^{n/2}(Y_{1j} - Y_{0j})^2\,/(n/2 - 1)}.$

We use t-test to see how redeemer is different from non-redeemer in terms of spending change from pre-launch phase to post-launch phase, which could be divided into three sub-questions

a. Does redeemer have a higher spending in post-launch phase?
b. Does non-redeemer has a higher spending in post-launch phase?
c. Is redeemer superior to non-redeemer as to spending increase in post-launch phase?

We can employ three t-tests to address them respectively.

To prepare the data set for paired t-test analysis, we calculated the total spending for both the pre- and post-launch phase and derive their differences. Table 3.1 and Table 3.2 provide a descriptive summary of the sample data.

Table 3.1 Summary of Sample Data for Paired T-Test

	Total			Spend		
Customer Redemption Groups	**N**	**Percentage (%)**		**pre-launch**	**post-launch**	**Delta ***
Overall	45,684	100	**Mean**	96.13	112.87	16.74
			S.D.	77.58	70.23	67.71
Sub-groups						
Redeemer	10,028	22	**Mean**	119.15	150.24	31.09
			S.D.	93.27	94.60	79.51
Non-Redeemer	35,656	78	**Mean**	89.66	102.36	12.70
			S.D.	71.24	57.44	63.41

Note: * Delta = (post-launch) - (pre-launch);

Table 3.2 Profile Summary for Redeemers and Non-Redeemers

	PRE_MULTI_CHNNL_IND	PRE_TRIPS	PRE_ADS	PRE_RECENCY	PRE_3M_STORE_SPEND	PRE_CREDIT_DELTA	PRE_CREDIT_LIMIT_ORIGINAL
	Overall Means and Standard Deviations						
Mean	0.168	0.000	0.000	0.000	0.000	0.000	0.000
S.D.	0.374	1.000	1.000	1.000	1.000	1.000	1.000
	Means and Standard Deviations for Redeemer						
Mean	0.163	0.479	-0.194	-0.183	0.233	0.030	0.178
S.D.	0.369	1.326	0.777	0.799	1.129	1.048	1.033
	Means and Standard Deviations for Non-Redeemer						
Mean	0.169	-0.136	0.053	0.052	-0.068	-0.009	-0.050
S.D.	0.375	0.832	1.036	1.044	0.936	0.985	0.985

Note: All variables are standardized with mean 0 and std 1, except PRE_MULTI_CHNNL_IND

We used the TTEST procedure to answer the above questions. Below are sample codes we used for this part of the analysis.

```
PROC TTEST DATA = dataset_name;
/* sum_pre_spend is the sum of spend in pre-launch period,
   Sum_post_spend is the sum of spend in post-launch period */
       PAIRED sum_post_spend* sum_pre_spend;
       TITLE "Testing the Equality of Paired Difference";
RUN;
```

3.2 RESULTS

Table 3.3 T-test results

	Mean	SD	DF	t Value	Pr > |t|
			Paired T-test for Redeemer		
Q1	31	79.51	10,027	39	<0.001
			Paired T-test for Non-Redeemer		
Q2	13	63.41	35,655	38	<.0001
			Two Sample T-test for Redeemer and Non-Redeemer's Spending Delta		
Q3	18	67.28	45,682	24	<.0001

The last two columns in Table 3.3 show results from the three t-tests where the t values are positive and their associated p values are smaller than 0.01. These results suggest that both redeemers and non-redeemers would spend more in post-launch period, in addition, customers who redeemed tended to spend even more in post-launch period than those who did not redeemer.

3.3 LIMITATION

Even though paired t-test is easy to apply and interpret, there are some assumptions when applying the test. It assumes that redeemers and non-redeemers are sampled from the same population. But if you look at redeemers and non-redeemers' profiles, you will find out that redeemers are naturally better than non-redeemers even before the launch of the program. As shown in Table 3.2, redeemers in general have a shorter recency and greater number of trips than that of non-redeemers. In other words, most redeemers made their last purchase not long before the launch of the program and visited the store more frequently. So the question becomes: is the difference in spending behavior a function of the rewards program and/or factors not currently considered in the model. We first turn to univariate repeated measures analysis to help us understand better factors that might adjust and make more clear these findings.

4. MODEL: MULTIPLE REGRESSION

4.1 OVERVIEW

We next use multiple regression to address some of the limitations associated with paired t-test. One of the many advantages multiple regression enjoys over paired t-test is that it allows for the examination of multiple predictors. The typical multiple regression model is usually written as

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \tag{2}$$

Using regression models to examine longitudinal data will most likely violate the assumption of independence of error since measures taken over time usually correlate within subjects. One way to correct for this non-independence is to split the total variance into its separate between-subjects and within-subjects components. Since we are interested in how customers change over time, we will focus only on the within-subjects part. Judd and McClelland (1989) offer a straightforward way to analyze repeated measures data within a regression framework. (The interested reader should consult Judd & McClelland (1989), pp 248-340 for details).

One of the corrective ways to get rid of non-independence is to collapse those sequential data into one composite score for each subject through some sort of transformation. Types of transformation rely on what you expect from the data. You can simply sum up all observations per subject if it makes sense as to reflect the variables in question. Or you could design an analogous contrast code for dependent variables to test specific hypothesis. For example, if the spending difference between 1st month and 2nd month is what you're looking for, you could assign a weight of 1 to spend1, a weight of -1 to spend2, and weights of 0 to all the other observations within subjects. For each subject, we would calculate the weighted sum of observations and then divide by the square root of the sum of the squared weights. The composite score for the ith subject is calculated by the formula below.

$$W_i = \frac{\sum_j \pi_j Y_{ji}}{\sqrt{\sum_j \pi_j^2}} \qquad (3)$$

Where π_j represents weight for the j^{th} observation and Y_{ji} represents the j^{th} observation for the i^{th} subject.

Note that we're dividing the weighted sum of observations by root squared sum of weights, rather than by sum of weights. There're two reasons behind it. First, if we use contrast code as weight, the squared sum guarantees a legitimate denominator, otherwise, simple sum of contrast code renders a value of 0. Second, such scaling of response combination helps create sums of squares in reasonable scales for each factor, which can be put directly into ANOVA source table (Judd and McClelland, 1989).

We are interested in answering the following questions:

a. Is there a difference between redeemers and non-redeemers in their pre/post program launch behavior controlling for a number of covariates, including recent 3-month in-store spending, recency, credit limit, etc.?
b. Is this difference in spending between the two groups evident in the first 3 months, 6 months, and 12 months post program launch?

Below is a sample data step showing how the within-subject variables are created.

```
DATA ds_reg;
SET sample_data;
   /* short term contrast */
   W1 = (spend13+spend14+spend15) - (spend1+spend2+spend3);
   W1 = W1/ 6**0.5;
RUN;
```

After the within-subject variables are created, we examine the within-subject variance using the Regression procedure. Please note any interaction term needs to be created in DATA step beforehand if we use PROC REG instead of other linear regression procedures like PROC GLM. Below is a snippet of code used for this analysis.

```
PROC REG DATA=ds_reg;
      CLASS redeem_ind  pre_multi_chnnl_ind;
      /*'pre_:' denotes all covariates and their selected interaction terms*/
      MODEL W1 = redeem_ind pre_: ;
RUN;
```

4.2 RESULTS

The results for the univariate repeated measures analysis is presented in Table 4.1.

Note that the parameter estimates reported below are back-transformed to their original unit to make the results more meaningful. (Please see notes sections for details about back-transformation process).

Table 4.1 ANOVA Source Table (Short Term : 3 Months)

	b	SS	df	MS	F	p
Within Subjects Effect						
W1	1.51	333,424	1	333,424	1,766	<0.001
W1 * REDEEM_IND	**0.45**	32,258	1	32,258	171	<0.001
W1 * PRE_MULTI_CHNNL_IND	0.19	1,228	1	1,228	7	0.0107
W1 * PRE_TRIPS	-2.06	552,482	1	552,482	2,926	<0.001
W1 * PRE_ADS	-1.33	177,281	1	177,281	939	<0.001
W1 * PRE_RECENCY	-0.30	9,386	1	9,386	50	<0.001
W1 * PRE_3M_STORE_SPEND	0.98	134,317	1	134,317	711	<0.001
W1 * PRE_CREDIT_DELTA	-0.10	1,555	1	1,555	8	0.004

	b	SS	df	MS	F	p
W1 * PRE_CREDIT_LIMIT_ORIGINAL	-0.32	19,416	1	19,416	103	<0.001
W1 * PRE_RECENCY * REDEEM_IND	-0.19	4,048	1	4,048	21	<0.001
W1* PRE_3M_STORE_SPEND * REDEEM_IND	0.31	16,650	1	16,650	88	<0.001
W1 * PRE_ADS * REDEEM_IND	-0.33	12,451	1	12,451	66	<0.001
Total W1 Within-Subject Error			45,672	189		

Note:
(1) W1 is calculated by 3-month spending in post-launch phase minus 3-month spending in pre-launch phase;
(2) b = unstandardized coefficients; SS = Sums of Squares; df = degree of freedom; MSS = Mean Sums of Squares; F = F test values;
(3) Back-transformation: b is derived from coefficient estimate divided by root squared sum of weights, F is derived from squared t value, SS is derived from F multiplied by mean squared error.

Table 4.1 reflects the effect size and contribution of variance for each factor. Take the highlighted coefficient for interaction of W1 with redeemer indicator as an example. Regarding short term spending growth over the two phases, redeemer seems to spend additional $0.90 (considering the coding scheme of redeemer indicator, spending discrepancy is obtained by taking difference between $0.45 and -$0.45) more than non-redeemer in each of the 3 months irrespective of anything else. Its minute p value indicates the reliable impact of coupon redemption on short-term spending growth.

Likewise, we can produce ANOVA source tables for W2 as mid-term spending difference (6 months) and W3 as long-term spending difference (12 months) in the same way. Table 4.2 provides a partial summary that emphasizes on the effect of response contrast and its interaction with REDEEM_IND in short-term, mid-term and long-term scenarios.

Table 4.2 ANOVA Source Table (Partial Summary)

	b	SS	df	MS	F	p
Within Subjects Effect for Short-Term Contrast						
W1	1.51	333,424	1	333,424	1,766	<0.001
W1 * REDEEM_IND	**0.45**	32,258	1	32,258	171	<0.001
Within Subjects Effect for Mid-Term Contrast						
W2	1.11	362,503	1	362,503	2,417	<0.001
W1 * REDEEM_IND	**0.52**	85,896	1	85,896	573	<0.001
Within Subjects Effect for Long-Term Contrast						
W3	1.06	657,381	1	657,381	4,762	<0.001
W1 * REDEEM_IND	**0.72**	332,644	1	332,644	2,410	<0.001

Note: W1 equals 3-month spending in post-launch phase minus 3-month spending in pre-launch phase; b = unstandardized coefficients; SS = Sums of Squares; df = degree of freedom; MS = Mean Sums of Squares; F = F test values;

4.3 LIMITATION

Univariate multiple regression allows us to control some covariates while testing the within-subject effect of pre-/post-launch phase, as well as the interaction of program phases with redeemer indicator. However, it doesn't take the correlation structure of repeated measures into consideration. Thus, it's impossible to perform an overall test on significance of within-subject effect if within-subject factor has more than two levels. Not only that, if you have multiple repeated-measure factors, you can't use multiple regression to test their interaction effect. For example, our 24 longitudinal observations can be separated into pre-launch and post-launch phase, and each phase would have a year-long monthly seasonality. Hence we have two repeated-measure factors, pre/post phase and monthly seasonality, where monthly seasonality is nested under pre/post phase. If we stick with univariate multiple regression, we won't be able to know whether seasonality differs over pre/post phase and whether redeem status is augmenting the difference. To address the questions here, we introduce multivariate repeated-measures ANOVA in next session.

5. MODEL: MULTIVARIATE REPEATED-MEASURES ANOVA

5.1 OVERVIEW

Multivariate repeated-measures ANOVA originates from univariate ANOVA, with multiple repeated measures as dependent variables. In essence, ANOVA uses variance in testing the overall statistical significance of the group means differences, whereas multivariate repeated-measures ANOVA uses variance-covariance matrix (Gregory Carey, 1998). In other words, repeated-measures ANOVA accounts for intra-individual correlation. Analogous to univariate ANOVA, repeated-measures ANOVA has counterpart assumptions of multivariate normality, independent subjects and homogeneity of variance between groups. In addition, it asks for sphericity, which refers to the equality of variance of difference between any pair of repeated-measures factor levels. If sphericity assumption does not hold, multivariate test is preferable to univariate test for within-subject variance unless some correction has been made for univariate test according to how far along sphericity is violated. Wilks' lambda, the Hotelling-Lawley trace, and Roy's greatest root are examples of multivariate test statistics PROC GLM will output.

Both PROC GLM and the ANOVA procedure can perform repeated-measures ANOVA. In contrast to PROC ANOVA, PROC GLM can deal with unbalanced data. Table 5.1 explains basic syntax of PROC GLM for repeated measures analysis of variance.

Table 5.1 Conduct Repeated-Measure ANOVA with SAS

INSTRUCTION	SAS CODE
	PROC GLM DATA = ds_rANOVA;
CLASS statement : redeem_ind is a categorical variable	**CLASS** redeem_ind;
MODEL statement : spend1, spend2, . . . , spend24 are on the left side of the model equation, representing response variables; redeem_ind and other covariates are on the right side of the equation, representing predictors. statement option *solution offers* regression estimation results for each single repeated measures listed in MODEL statement; statement option *ss3* provides Type III sum of squares in Analysis of Variance output	**MODEL** spend1 - spend24 = redeem_ind &covariate. / solution ss3;
REPEATED statement : we specify two repeated-measure factors here --- phase and seasonality; phase has 2 factor levels and seasonality has 12 levels; *contrast* transformation of repeated measures is conducted for phase, and a set of orthogonal *polynomial* transformation is for seasonality; *contrast(1)* means to generate contrasts between the first level of a factor and the other levels. statement option *summary* offers ANOVA tables for each contrasts defined by repeated-measure factors, along with their interaction with independent variables specified in MODEL statement; statement option *printm* displays the transformation matrics that define the contrast statement option *printe* not only shows error covariance matrix for each combination of repeated-measure factors, but also provides sphericity tests for each set of transformed response variables	**REPEATED** phase 2 contrast(1) , seasonality 12 polynomial / summary printm printe;
	RUN;

5.2 RESULTS

Both multivariate test and adjusted univariate test of hypotheses for within-subject effects demonstrate significant results, meaning that program phases, polynomial seasonality, and their interactions all exert certain impact on customer spending. Moreover, the interactions of these within-subject factors with redeemer indicator also play some roles.

5.3 LIMITATION

Apparently, multivariate repeated-measures ANOVA outperforms univariate multiple regression by focusing on a bigger picture. But it shares the problem with univariate multiple regression that it can only tell you whether there is a difference among groups, shedding no light on where the difference lies. With regard to the question that which pairs of within-subject groups result in the difference, we should accompany repeated-measures ANOVA with further multiple comparisons to solve the puzzle.

6. DISCUSSION

6.1 RESULTS SUMMARY

All paired t-tests have extremely large t-values along with p values under 0.01, which means all null hypotheses are rejected. In other words, redeemer and non-redeemer both spend more in post-launch phase, where redeemer's spend increases even more. But we don't know if the increase is due to other factors, so we switch to univariate multiple regression.

By examining p values from univariate multiple regression analysis, we could conclude that all within-subject factors and their interaction with between-subject covariates would exert influence on spending. As we move from short term, midterm to long term, we could see the highlighted within-subject effect coefficients in Table 4.2, representing redeemers' additional average incremental spending per month over non-redeemers, are actually climbing up. With that being said, as the rewards program moves forward, redeemers are turning into better consumers, and their advantage against non-redeemers has become increasingly prominent. However, when making allowance for another within-subject factor, seasonality, within each phase, there's no way to figure out whether change is still significant if we're committed to univariate multiple regression. Therefore, we turn to repeated-measures ANOVA.

The sphericity test for repeated-measures ANOVA exhibits a p value under 0.01, indicating the assumption of sphericity fails. So we should go with multivariate test or adjusted univariate test for within-subject effect. The multivariate tests in SAS involving within-subject effects report significant results with four separate multivariate test statistics (Wilks' lambda, Pillai's Trace, the Hotelling-Lawley trace, and Roy's greatest root), while univariate tests also report small adjusted p values after Greenhouse-Geisser corrections and Huynh-Feldt-Lecoutre corrections. Both multivariate test and univariate test proved the existence of within-subject effects.

6.2 SIMILARITY

From data perspective, paired t-test, univariate multiple regression and multivariate repeated-measures ANOVA are all sensitive to atypical outliers, non-independence between subjects and heterogeneous variance among groups. And large sample size can attenuate the effect of deviation from normality for these tests.

For application, univariate multiple regression and repeated-measures ANOVA can both test if any change exists over time. But neither of these techniques can tell us where the change, if any, comes from.

6.3 DIFFERENCES

From data perspective, repeated-measures ANOVA can take care of the non-independence within subjects and handle multiple repeated measures, whereas paired t-test and univariate multiple regression can't except for some specific questions.

For application, paired t-test is beaten by univariate multiple regression, since univariate multiple regression can explain more variance by controlling covariates. But univariate multiple regression is not as good as repeated-measures ANOVA when it comes to testing overall effect of change over time, or interaction of within-subject effects.

6.4 CONCLUSION

Even though the simple pre-post change comparison prevails currently in the retail sector, it reveals itself to not be the most appropriate technique to address some of the more complex questions in the ever evolving retail industry. Analysts in retail sector may consider broadening their methodological tool belt in anticipation of this trending.

In this paper, we demonstrated the applications of different longitudinal methods to analyze change over time. As we shift from paired t-test to univariate multiple regression, and finally to repeated-measures ANOVA, more dependable analysis results are delivered, and more sophisticated questions are answered. We are not suggesting that one method is better than another. Instead, we are suggesting the application of the right technique to a given questions. Some techniques are more appropriate to address some questions, meanwhile, they all call for an in-depth understanding of the data to ensure its consistency with their assumptions. Without fully understanding your data and crystalizing the key questions you are trying to address, it becomes easy to mismatch statistical methods to questions. In addition, more time has to be devoted if opting for advanced methods when we could have solved the problem in a much easier way.

Obviously, we omitted a number of alternative change methods in this paper, including latent growth curve model with time-varying covariates. We encourage you to explore these and many other useful change models if of interest to you.

REFERENCES

[1]. Todd D Little, Kai U. Schnabel, JÜrgen Baumert. 2000. *Modeling Longitudinal and Multilevel Data - Practical Issues, Applied Approaches and Specific Examples*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[2]. Charles M. Judd, Gary H. McClelland, Carey S. Ryan. 1989. *Data Analysis: A Model-Comparison Approach*. San Diego, CA: Harcourt Brace Jovanovich (HBJ).

[3]. Gregory Carey. "Multivariate Analysis of Variance (MANOVA): I. Theory". Retrieved from http://ibgwww.colorado.edu/~carey/p7291dir/handouts/manova1.pdf. (1998)

[4]. David C. Howell. 2012. *Statistical Methods for Psychology, 8th Edition*. Independence, KY: Wadsworth Cengage Learning.

ACKNOWLEDGMENTS

We would like to profusely thank Alphonse Damas and Amy Yu for their guide and mentorship during the course of this work. We also thank Marketing Insights Department at Alliance Data for providing us the opportunity to attend the conference.

CONTACT INFORMATION

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

Name: Jie Liao
Enterprise: Alliance Data
Address: 3100 Easton Square Pl
City, State, ZIP: Columbus, Ohio, 43219
Work Phone: 614-944-5331
E-mail: jie.liao@alliancedata.com

APPENDIX

```
*========================================================*
** Paired T-Test
** (1) Create Data Set ds_paired_ttest From sample_data
*========================================================*;
DATA  ds_paired_ttest;
       SET sample_data;
       sum_pre_spend = SUM(OF spend1-spend12);
       sum_post_spend = SUM(OF spend13-spend24);
       spend_delta = sum_post_spend - sum_pre_spend;
RUN;

*** (2) Paired T-Test with PROC GLM ***;

PROC TTEST DATA = ds_paired_ttest (WHERE = (redeem_ind = 1));
       PAIRED sum_post_spend*sum_pre_spend;
       TITLE "Testing the Equality of Paired Difference for Redeemers";
RUN;

PROC TTEST DATA = ds_paired_ttest (WHERE = (redeem_ind = -1));
       PAIRED sum_post_spend*sum_pre_spend;
       TITLE "Testing the Equality of Paired Difference for Non-Redeemers";
RUN;

PROC TTEST DATA = ds_paired_ttest;
       CLASS redeem_ind;
       VAR spend_delta;
       TITLE "Testing the Equality of Delta Means for Redeemers and Non-Redeemers";
RUN;


*===================================================*
** Univariate Multiple Regression
** (1) Create Data Set ds_reg From sample_data
*===================================================*;
/* Transform Dependent Variables */
Data ds_reg;
       SET sample_data;

       ** between subject **;
       W0=0;
       W0 = SUM(of spend:);
       W0=W0/24**0.5;

       ** within subject : short-term (first 3 months difference from pre to post) **;
       W1=0;
       W1 = SUM(OF spend13-spend15) - SUM(OF spend1-spend3);
       W1=W1/6**0.5;

       ** within subject : mid-term (first 6 months difference from pre to post) **;
       W2=0;
       W2 = SUM(OF spend13-spend18) - SUM(OF spend1-spend6);
       W2=W2/12**0.5;

       ** within subject : long-term (total 12 months difference from pre to post) **;
       W3=0;
       W3 = SUM(OF spend13-spend24) - SUM(OF spend1-spend12);
       W3=W3/24**0.5;
RUN;

/* Create Interaction Terms in DATA step */
DATA ds_reg;
       SET ds_reg;
       pre_recency_by_redeem = pre_recency * redeem_ind;
       pre_3M_store_spend_by_redeem = pre_3M_store_spend * redeem_ind;
```

```
        pre_ads_by_redeem = pre_ads * redeem_ind;
RUN;



*** (2) Univariate Multiple Regression Analysis with PROC REG ***;
%LET pre =
        redeem_ind
        pre_multi_chnnl_ind
        pre_trips
        pre_ads
        pre_recency
        pre_3m_store_spend
        pre_credit_delta
        pre_credit_limit_original
        pre_recency_by_redeem
        pre_3M_store_spend_by_redeem
        pre_ads_by_redeem
;

/* Between Subject */
/* Detect Outliers */
PROC REG DATA = ds_reg PLOTS(MAXPOINTS=NONE)=DIAGNOSTICS;
        MODEL W1 = redeem_ind;
        MODEL W1 = redeem_ind &pre.;
        OUTPUT OUT = output P=yhat R=yresid COOKD=cookd PRESS=press;
RUN;

/* Short Term Within Subject: First 3 months in each phase */
PROC REG DATA = ds_reg;
        MODEL W1 = redeem_ind;
        MODEL W1 = redeem_ind &pre.;
RUN;

/* Midterm Within Subject: First 6 months in each phase */
PROC REG DATA = ds_reg;
        MODEL W2 = redeem_ind;
        MODEL W2 = redeem_ind &pre.;
RUN;

/* Long Term Within Subject: First 12 months in each phase */
PROC REG DATA = ds_reg;
        MODEL W3 = redeem_ind;
        MODEL W3 = redeem_ind &pre.;
RUN;


*=================================================*
** Repeated-Measures ANOVA
** (1) Create Data Set ds_rANOVA From sample_data
*=================================================*;
DATA ds_rANOVA;
        SET sample_data;
RUN;

*** (2) Repeated-Measures ANOVA with PROC GLM ***;
%LET covariate =
        redeem_ind
        pre_multi_chnnl_ind
        pre_trips
        pre_ads
        pre_recency
        pre_3m_store_spend
        pre_credit_delta
        pre_credit_limit_original
        pre_recency*redeem_ind
```

```
        pre_3M_store_spend*redeem_ind
        pre_ads*redeem_ind
;

ODS GRAPHICS ON;
PROC GLM DATA = ds_rANOVA PLOTS(MAXPOINTS=NONE)=(DIAGNOSTICS RESIDUALS) /* Detect
Outliers */ ;
        CLASS redeem_ind pre_multi_chnnl_ind;
        MODEL spend: = redeem_ind &covariate.
              / solution ss3;
        REPEATED phase 2 contrast(1), seasonality 12 polynomial
              / summary printm printe;
        LSMEANS redeem_ind /* Least squares & adjusted means for ANCOVA */
                         / STDERR      /*  .. and std errors          */
                           PDIFF ;     /*  ... and p-values for diff */
        CONTRAST 'Redeemer vs Non-Redeemer'  redeem_ind 1 -1/E;
        ESTIMATE "Redeemer - Non-Redeemer" redeem_ind 1 -1/divisor = 2;
        OUTPUT OUT=RANOVA_out
              P = fitvar /* Predicted values */
              R = residvar  /* Residuals */
              COOKD = cookd /* cooks' D influence statistic */
              H = h /* leverage */
              STDR = std_r/*standard error of the residual */;
RUN;
ODS GRAPHICS OFF;
```