**University of Connecticut**

**MS in Business Analytics and Project Management**

****

**OPIM 5510 Web Analytics**

**Professor Xinxin Li**

**Ad Display Analysis on Taobao**

**Team 5**

**Lei Cao**

**Mrudula Kadam**

**Yukuan Li**

**Jieying Qin**

**Si Zhu**

**Table of Contents**

* Company Background ………………………………………...………………………….1
* Data Description ……………………………………………....………………………….1
* Metric Analysis ……………………………………….………………………………….1
* Suggestions ………………………………………………………………....…………….2
* Experiment Evaluation ……………………………………………………………...…….3
* Appendix ………………………………………………………………………………….4

**Company Background**

Taobao is the Asia-Pacific region’s largest network of retail business, founded by the Alibaba group in May 2003. It is a popular online shopping and retail platform in China, with nearly 500 million registered users, more than 60 million regular visitors a day, and more than 800 million products online each day, an average of 48,000 items are sold every minute.

In 2009, it became the largest general store in China, with annual turnover reaching 208.3 billion yuan. By the end of 2011, the peak daily turnover in Taobao was 4.38 billion yuan, creating 2.708 million direct and full employment opportunities. With the expansion of the scale of Taobao and the increase of the number of users, Taobao has also changed from a single C2C network bazaar to a comprehensive retail circle including C2C, group purchase, distribution, auction and other e-commerce models. At present, it has become one of the electronic commerce transaction platforms in the world. March 15,2016,315 party exposure, Taobao merchants brush single and other deceptive consumer phenomenon.

Taobao website has a massive amount of traffic and customers. Each customer has their own consuming habit. Besides, different users are interested in different kinds of goods when visiting an e-commerce site. So how to choose a more attractive ad is very important to increase CTR. Our goal is to use historical behavior data to optimize the online advertising system for Taobao website.

**Data Description**

The dataset sampled 1140000 users from the website of Taobao for 8 days of ad display / click logs (26 million records) to form the original sample skeleton. The original dataset includes three separate data sets, raw\_sample, ad\_feature, and user\_profile.

***raw\_sample***: Users from Taobao website for 8 days of ad display or click logs (26 million records) with 26557961 rows and 6 columns, where the variables and their descriptions are:

***ad\_feature***: Basic information of all ads in raw\_sample with 846811 rows and 6 columns, where the variables and their descriptions are:

***user\_profile***: Basic information of 1060000 users in raw\_sample with 1061768 rows and 9 columns, where the variables and their descriptions are:

For data processing, we merged the three datasets and removed some redundant and missing values. We also checked the value consistency for dummy variables and did some recode to make them consistent. For field description, see Appendix 2.1.

**Metric Analysis**

Taobao’s objective is increasing their users as well as increasing the brand’s reach. In order to understand the effectiveness of the campaigns, the two key metrics that we are using are: CTR (primary) and CPM (secondary). CTR helps us in understanding the effectiveness of campaigns and also helps us identify the campaigns that should be scaled and the ones that need to be pruned. CPM is a secondary metric as it helps us increase the awareness of the brand and also track the ROI for the customers. Considering CTR as the primary metric for analysis, we decided to perform regression analysis to determine the impact of various factors contributing to CTR. As CTR is affected by clicks and impressions, we constructed CTR as our target variable by aggregating records as impressions and summing clicks of 1 as clicks. Then we built the regression model and applied the causality concept to gain insights on what variables would contribute to CTR so that the company can take corresponding methods to improve it.

As we intended to explore how males and females behave differently towards ad displays, we decided to focus mainly on gender and create interactions mostly between it and other predictor variables. The formulation of the linear regression model we constructed is attached at Appendix 3.1.

From the estimation results (Appendix 3.2), we learned that males contribute less to CTR, because the final\_gender\_code-1is significant according to its p-value and the estimate is negative. Moreover, p-values of interactions between final\_gender\_code-1 and period\_code-1, and interactions between final\_gender\_code-1 and period\_code-2 are less than 0.05. And the coefficients are negative, representing males in the afternoon and evening contribute less to CTR. Regarding the effect of afternoon and evening on males, they are statistically the same according to the hypothesis testing result (Appendix 3.3.5) because the p-value of the F test is greater than 0.05. However, the estimates of period\_code-1 and period\_code-2 are positive, meanwhile the p-values are less than 0.05. So these two terms are also significant but they indicate that females in the afternoon and evening contribute more to CTR. According to the result of the hypothesis testing, afternoon and evening have statistically different effects on females because the p-value is less than 0.05 (Appendix 3.3.6). Finally, p-values of the pvalue\_level.Q and interactions between final\_gender\_code-1 and pvalue\_level.Q are both less than 0.05, standing for their significance to CTR. However, the estimates are opposite, where it is positive for the pvalue\_level.Q but negative for the interactions between final\_gender\_code-1 and pvalue\_level.Q. So, it can be interpreted that females with consumption grade of high contribute more to CTR while males with consumption grade of high contribute less. For more hypothesis testing, see the entire Appendix 3.3.

**Suggestions**

CTR is important to advertisers, it affects revenue cause conversion happens after clicking from customers. It also affects the cost in this case. In CPC advertising systems like the one in Alibaba, advertisements are ranked based on effective Cost Per Mille, which is a product of bid price and CTR. Hence how to improve CTR is one main concern for advertisers.

Gender plays an important role in this case. Based on the result from Metric Analysis before, we can see that gender is statistically significant(p-value of final\_gender\_code is less than 0.05) when we are predicting the CTR. From the experience in real life, we may notice that the reason for purchasing the same product is different. When buying a car, a man might think 'it will help to get me to work and save more time'. Whereas the woman is thinking, 'it will help us to spend more time together and have travel on the weekend'. In other words, the motivations for men and women are different. The quote "Men are from Mars women are from Venus" from the 1991 book of the same name by John Gray is probably one of the world's most well known. Also from the same book is this fantastic quote "Men are motivated when they feel needed while women are motivated when they feel cherished." If we convert this to the advertisements on Taobao, we may need to develop different kinds of Ad-version for males and females.

Based on these insights, we offer a suggestion for a gender targeting strategy. For the same product, we develop two kinds of Ad-version to deliver different emotions and motivations to men and women. We deliver a sense of being cherished to women, feeling comfortable about buying, and feels appreciated by the vendor. We deliver a sense of problem-solving to men to let them feel they get what exactly they need. The features are as follows (see Appendix 4.1)

**Experiment Evaluation**

Since we have put forward the gender targeting strategy and developed different AD-version for males and females. We want to know whether our suggestion does bring a positive effect on our CTR. We would develop an experiment to make a causal analysis. Since the experiment on male and females is exactly the same, so we just choose Ad-version 1 to illustrate. The assumption is that we have whatever we want to conduct the experiments. The factors are as follows:

**Objective**: To check the effect of Ad-version 1 on CTR

**Treatment**: For the same product, use the Ad-version 1 advertisement on treatment group

**Outcome Variable**: CTR

**Time to run**: 1 + 1 month

**Method**:

(1) PSM, cus no random assignment so PSM help to make our subjects comparable

(2) DID method, to exclude the difference between two groups by nature

(3) Significance testing on β3, to decide the causal effect on CTR

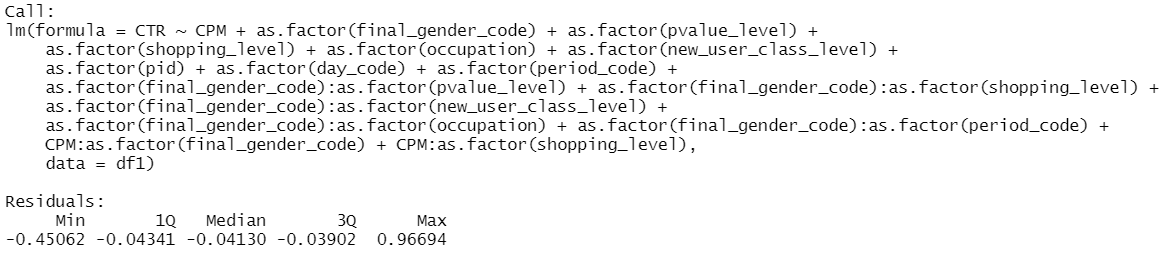
The experience pipeline is as follows, to make it easier to follow, we randomly assign values in the chart(see Appendix 5.1)

**Appendix**

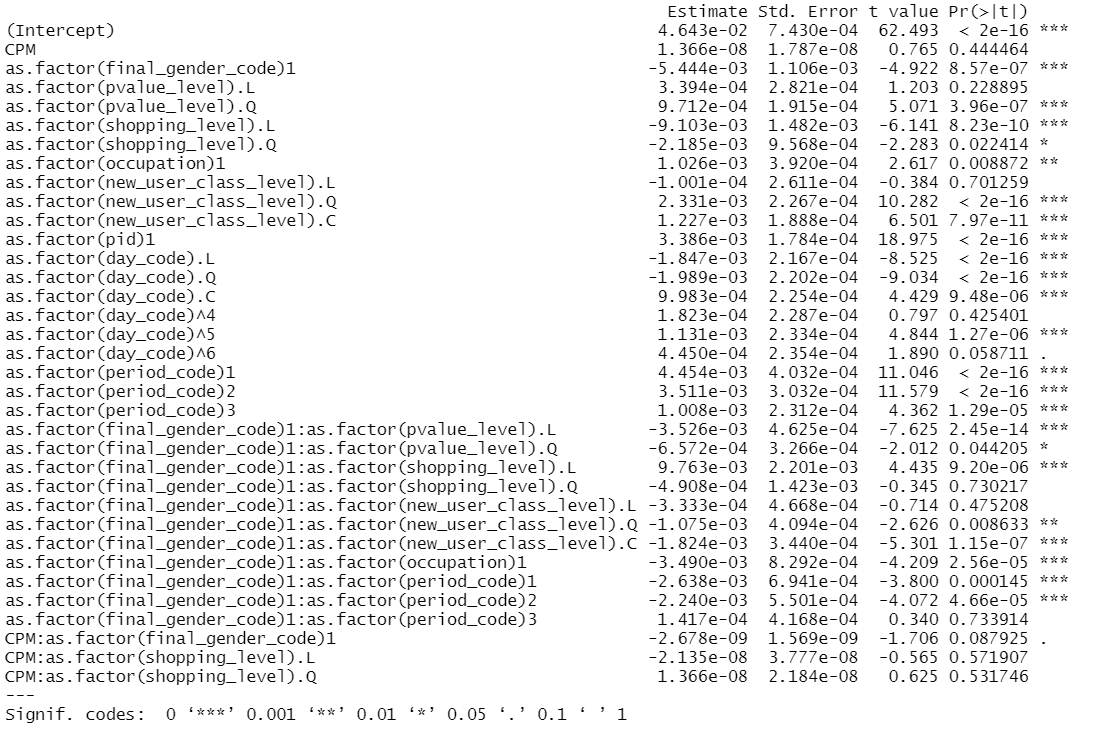
2.1 Field Description

|  |  |
| --- | --- |
| Field | Description |
| user | User ID |
| timestamp | time stamp (eg. Bigint, 1494032110 stands for 2017-05-06 08:55:10 |
| adgroupid | advertising group ID |
| pid | scenario |
| noclk | 1—not click; 0—click |
| clk | 1—click; 0—not click |
| adgroup\_id | Ad ID |
| cate\_id | category ID |
| campaign\_id | campaign ID |
| customer | advertiser ID |
| brand | brand ID |
| price | advertising price |
| cms\_segid | micro group ID |
| cms\_group\_id | cms group ID |
| final\_gender\_code | 1—male; 2—female |
| age\_level | customer segment based on age |
| pvalue\_level | consumption grade, 1—low, 2—mid, 3—high |
| shopping\_level | 1— shallow user; 2—moderate user; 3—heavy user |
| occupation | 1— college students, 2 — non-college students |
| new\_user\_class\_level | City level |

3.1 Regression Model Formulation

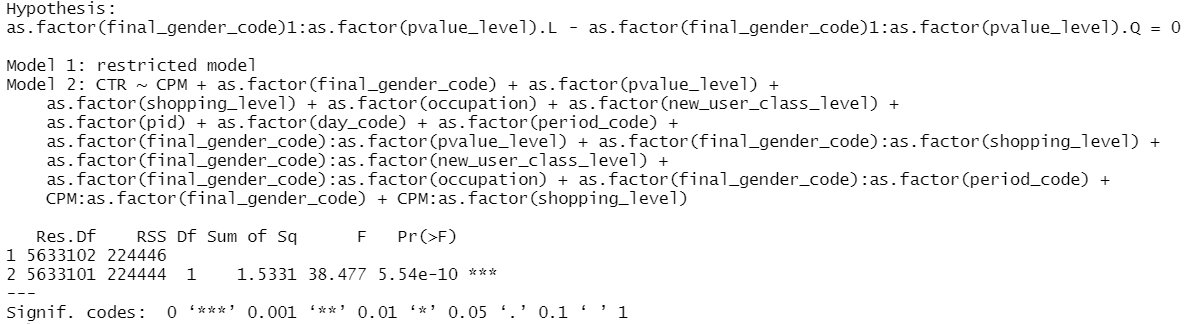


3.2 Regression Estimation Results



3.3 Hypothesis Testing

1. *compare effect of consumption grade of mid and high on males*

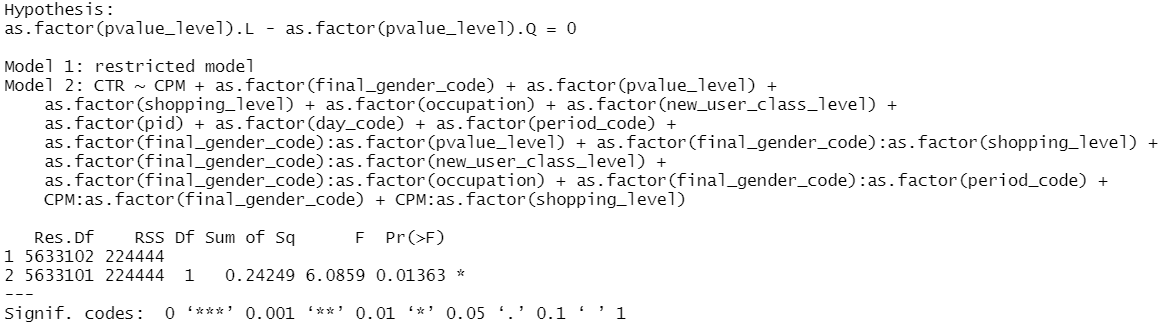


H0: Consumption grades of mid and high have the same effect on males.

H1: Consumption grades of mid and high have different effects on males.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say the effects from consumption grades of mid and high on males are statistically **different**.

1. *compare effect of consumption grade of mid and high on females*

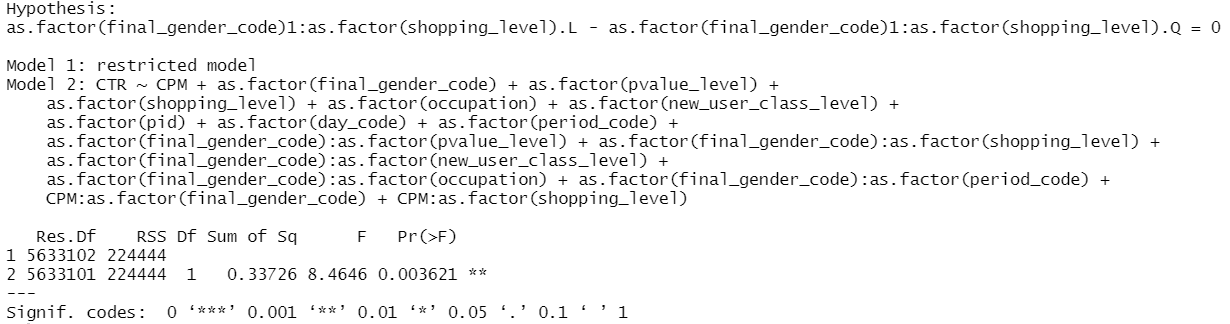


H0: Consumption grades of mid and high have the same effect on females.

H1: Consumption grades of mid and high have different effects on females.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say the effects from consumption grades of mid and high on females are statistically **different**.

1. *compare effect of shopping level of moderate and deep on males*

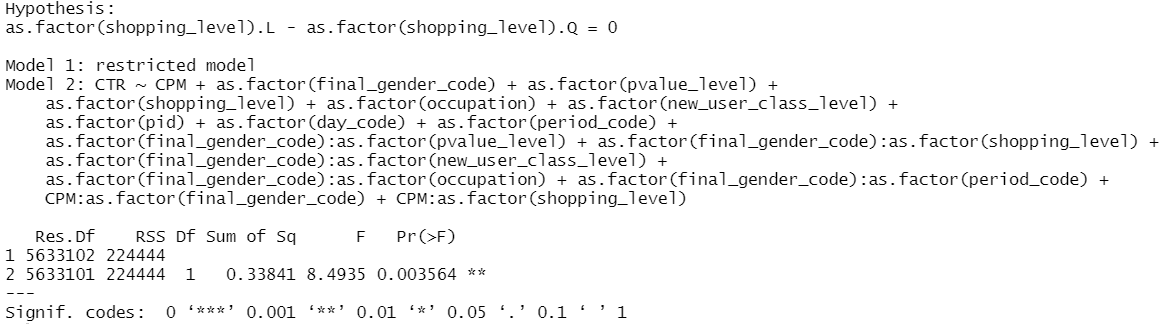


H0: Shopping levels of moderate and deep have the same effect on males.

H1: Shopping levels of moderate and deep have different effects on males.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say the effects from shopping levels of moderate and deep on males are statistically **different**.

1. *compare effect of shopping level of moderate and deep on females*

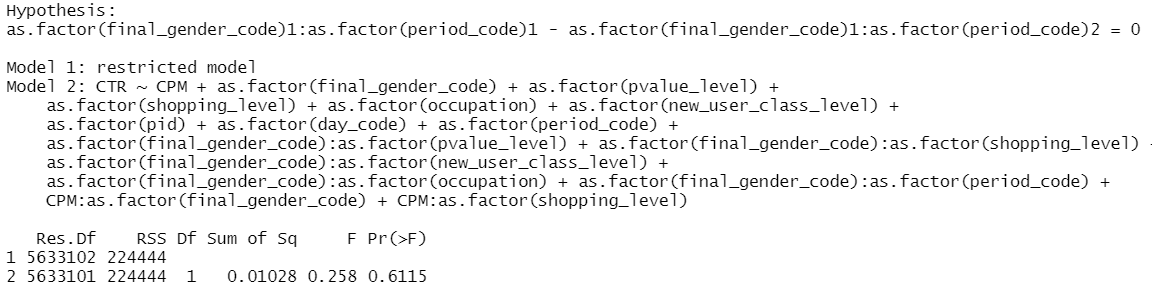


H0: Shopping levels of moderate and deep have the same effect on females.

H1: Shopping levels of moderate and deep have different effects on females.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say the effects from shopping levels of moderate and deep on females are statistically **different**.

1. *compare effect of afternoon and evening on males*

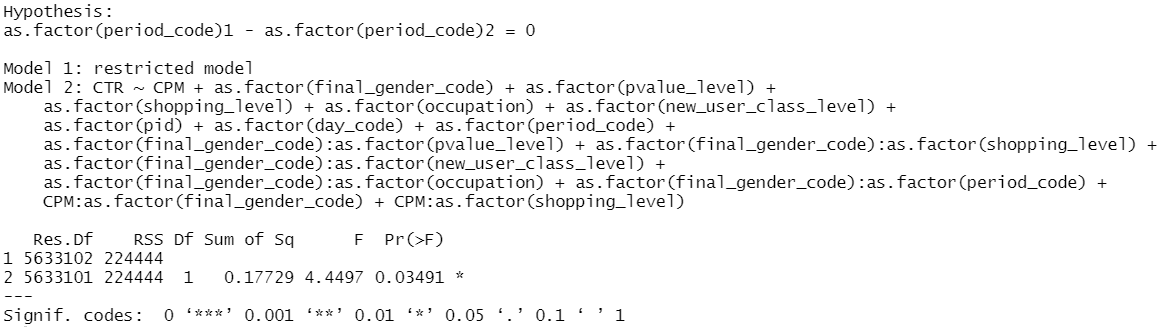


H0: Afternoon and evening have the same effect on males.

H1: Afternoon and evening have different effects on males.

Since the p-value of the F test on the hypothesis H0 is **greater than 0.05**, we fail to reject H0. Hence effects from afternoon and evening on males are statistically **the same**.

1. *compare effect of afternoon and evening on females*

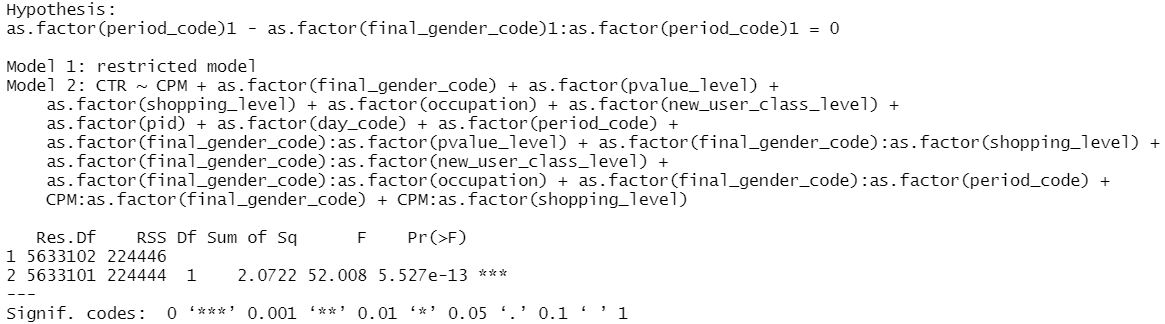


H0: Afternoon and evening have the same effect on females.

H1: Afternoon and evening have different effects on females.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say the effects from afternoon and evening on females are statistically **different**.

1. *compare afternoon effect between genders*

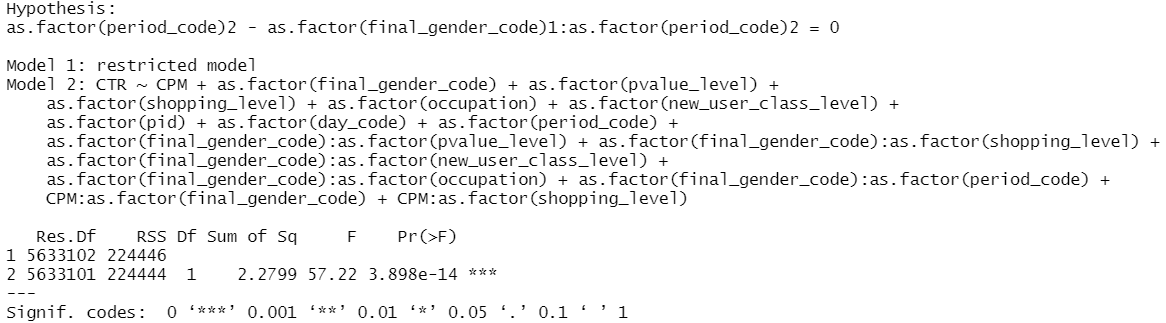


H0: Afternoon has the same effect between males and females.

H1: Afternoon has different effects between males and females.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say afternoon has statistically **different** effects between males and females.

1. *compare evening effect between genders*



H0: Evening has the same effect between males and females.

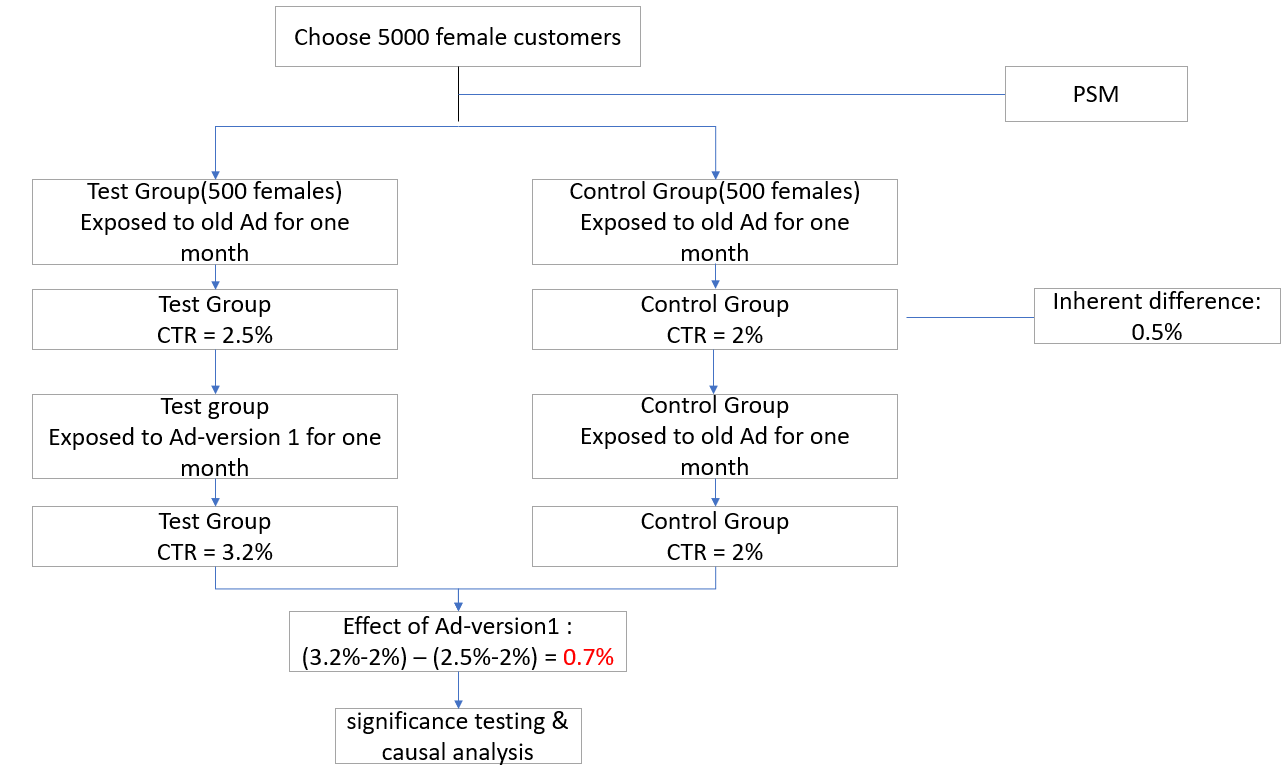
H1: Evening has different effects between males and females.

Since the p-value of the F test on the hypothesis H0 is **less than 0.05**, we reject H0 and say evening has statistically **different** effects between males and females.

4.1 New ad version design and features for men and women

|  |  |
| --- | --- |
| **Gender Targeting Strategy** | |
| Ad-version 1(for female) | Ad-version 2(for male) |
| * Great customer service and good feedback * Great value, great offers, rewards * Discount and coupons on Birthday, Women’s day | * It can solve the problem in front of you * How-to videos * Addressing the ‘Pain Points’ |
|
|

5.1 Experiment design and pipeline

****