

# Huohua Siwei Mock Project (Applied Economist Position)

Yangzhuoran Fin Yang (Monash University)

20 March, 2021

## 1 Introduction

In online education, trial classes have been offered to potential customers (parents and kids) in an attempt to increase sales. The probability of making a sale after such a trial class has been known as the conversion rate. In the study of conversion, likelihood labels have been assigned to students based on their estimated conversion rate. The action of putting video feeds of students who are most likely to make purchases in the middle of the screen can be made by the presenter. The presenter shifts focus to and make possible eye contact with the middle of the screen.

This project attempts to study the effect of two variables on conversion: showing labels of the likelihood of purchase to the presenter (`is_show`) and sorting students in such an order that the most likely purchasers are put in the centre of attention (`is_sort`). Possible explanations of the mechanism are discussed. Further application of labelling and sorting, and experimental design to assess the effectiveness are proposed.

The rest of the report is structured as follows. Section 2 introduces the data and experimental design. Section 3 analyses the relationships between conversion and label showing and sorting. Section 4 discusses possible rationales of the relationship. Section 5 proposes further applications of label-showing and sorting as well as possible assessments of the effectiveness of such application. Section 6 makes a brief summary.

The following packages have been used in this report:

```
# For data analysis
library(tidyverse)
library(DescTools)
library(margins)
library(boot)
# For compilation and presentation
library(bookdown)
library(knitr)
library(kableExtra)
library(gridExtra)
```

## 2 Data

```
unzip("dataset.zip")
data_a <- read_csv("data_set_a.csv")
data_b <- read_csv("data_set_b.csv")
```

The data has been provided in two different sets, A and B. There are 31460 observations in data set A where the students were not selected to participate in the experiment. Thus no label-showing or sorting has been made to observations in set A, whereas in data set B of 2964 observations, students have been assigned to four groups with equal probability: with labelling and with sorting, with labelling but with no sorting, with sorting but no labelling, and no labelling or sorting.

Table 1: Variable definition

Variable name	definition	Possible values
uid	Student ID	Integer
cr_id	Class ID	Integer
seat_id	Index of position of video feeds	Row index and column index of seats, e.g. first row first seat 1-1
is_king_kong	Whether the video feeds are in the middle of the screen, known as king kong seats	1 for Yes, 0 for No
prob	Estimated conversion rate	Numerical value between 0 and 1
tag	Likelihood label of conversion	P1 to P4 from least likely to most likely
is_pay	Whether conversion is made	1 for Yes, 0 for No
is_sort	Whether video feeds have been sorted	1 for Yes, 0 for No
is_show	Whether the label has been shown to the presenter	1 for Yes, 0 for No

The definition of variables can be found in Table 1. The middle portion of all the seats (video feeds) is known as king kong seats as the presenter puts more attention to this part of students. The likelihood labels are assigned to students based on their estimated conversion rate from most likely to least likely. For classes where seats have been sorted, students with the most likely label (P4) have been put in the middle, surrounded by students with the second most likely label (P3), then third likely (P2) and least likely (P1).

### 3 Effects of Labelling and Sorting

#### 3.1 Quick Independence Tests

Before we consider the influences from other control variables, let us examine the simple correlation between the treatments (labelling and sorting) and conversion. We will be focusing on data set B for this part of the analysis. The frequency of having a conversion with labelling or sorting is shown in Figure 1. Visually, on the left of Figure 1, having the label shown to the presenter makes little difference to whether the customer converts or not. On the right, while the numbers of conversion with sorting and without sorting are similar, the number of non-conversion with sorting is smaller than that without sorting, which indicates a possible correlation.

Two tests of independence between labelling and conversion and between sorting and conversion are conducted. Results are shown below. With a huge p-value, we cannot reject the null hypothesis that both labelling and sorting are independent with conversion, so this is another indication that they may be not correlated.

```
## Test of independence between labelling and conversion
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data: .
## X-squared = 0.33071, df = 1, p-value = 0.5652
## -----
## Test of independence between sorting and conversion
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data: .
## X-squared = 0.0041717, df = 1, p-value = 0.9485
```

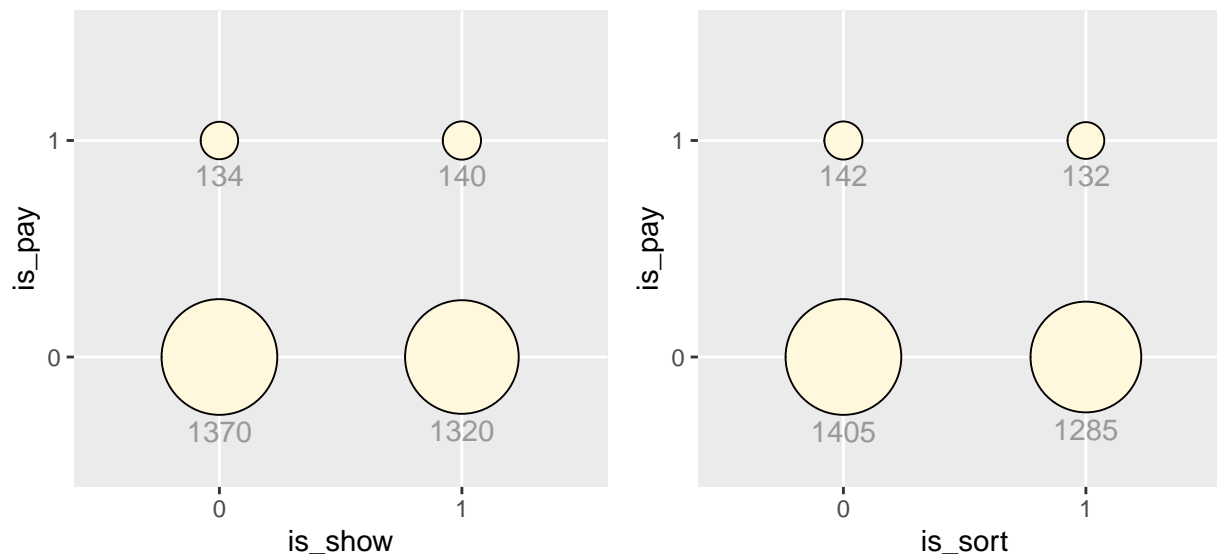


Figure 1: Balloon plot of frequencies of labelling and sorting with respect to conversion.

### 3.2 Direct effect

We can examine the direct effect of labelling and sorting by applying a logistic (logit) regression with conversion being the regressand, and two treatment variables and all other control variables being regressors. The result has been shown below. For the control variables I have included the estimated conversion rate (prob) and combinations of interaction between being in king kong seats (is\_king\_kong) and likelihood label (tag).

```
data_b %>%
  glm(is_pay ~ is_show + is_sort + prob + is_king_kong*tag,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()
```

```
##
## Call:
## glm(formula = is_pay ~ is_show + is_sort + prob + is_king_kong *
##      tag, family = binomial(link = "logit"), data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9865  -0.4535  -0.3294  -0.2680   3.1667
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.46981    0.29124  -11.914  < 2e-16 ***
## is_show         0.08683    0.13088    0.663  0.507046
## is_sort         0.03826    0.13204    0.290  0.771983
## prob           2.82545    1.50281    1.880  0.060092 .
## is_king_kong   -1.59240    1.03970   -1.532  0.125621
## tagp2           0.33486    0.31714    1.056  0.291029
## tagp3           0.86861    0.33908    2.562  0.010417 *
```

```
## tagp4          1.48719    0.39406    3.774 0.000161 ***
## is_king_kong:tagp2 2.23906    1.06738    2.098 0.035931 *
## is_king_kong:tagp3 2.31993    1.06743    2.173 0.029752 *
## is_king_kong:tagp4 1.45076    1.06319    1.365 0.172398
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1826.7 on 2963 degrees of freedom
## Residual deviance: 1696.7 on 2953 degrees of freedom
## AIC: 1718.7
##
## Number of Fisher Scoring iterations: 7
```

In this logit model, the main contribution of variation in conversion comes from having a label indicating a high conversion rate (`tagp4` or `tagp3`), and having a label with a medium conversion rate and being in the king kong seats (interaction between `is_king_kong` and `tagp2` or `tagp3`). This can be seen by a small p-value indicating significant coefficients of corresponding regressors. The treatment variables of interest (`is_show` and `is_sort`), however, are insignificant in the determination of conversion with a large p-value. Although the coefficient is positive, we do not have enough evidence to say labelling or sorting has a positive effect on conversion. There is no significant direct effect of labelling on conversion. This conclusion is consistent with preliminary analysis using plots and tests of independence.

Applications with a probit model and a probability linear regression yield the same conclusion.

### 3.3 Interactions

With no direct effects implied by the above analysis, I also considered possible indirect effects of labelling and sorting. Will labelling increase the positive effects created by being in the king kong seats? The presenter of the class may put different attention to students in king kong seats with different labels. Another possibility is a high likelihood label may benefit from being shown to the presenter, but not so much for a low likelihood label. This idea is examined using a logit model below with interactions between being in king kong seats (`is_king_kong`), labels (`tag`), and showing labels to the presenter (`is_show`). Unfortunately, none of the terms involving label-showing has any statistical significance. The assumption is not supported.

```
data_b %>%
  glm(is_pay ~ is_king_kong*tag*is_show + is_sort + prob,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()

##
## Call:
## glm(formula = is_pay ~ is_king_kong * tag * is_show + is_sort +
##      prob, family = binomial(link = "logit"), data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0744  -0.5030  -0.3285  -0.2994   2.9267
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -3.78685    0.46035  -8.226 < 2e-16 ***
## is_king_kong    -0.53820    1.10434  -0.487 0.626008
## tagp2           0.64897    0.50647   1.281 0.200068
```

```
## tagp3          0.81864    0.55137    1.485 0.137613
## tagp4          1.99237    0.55212    3.609 0.000308 ***
## is_show        0.62616    0.56707    1.104 0.269509
## is_sort        0.04680    0.13242    0.353 0.723773
## prob          2.88455    1.51154    1.908 0.056347 .
## is_king_kong:tagp2 1.37210    1.15246    1.191 0.233820
## is_king_kong:tagp3 1.34659    1.17189    1.149 0.250524
## is_king_kong:tagp4 0.33493    1.14414    0.293 0.769725
## is_king_kong:is_show -12.93818 284.75834 -0.045 0.963760
## tagp2:is_show    -0.54762    0.64919   -0.844 0.398928
## tagp3:is_show     0.06368    0.67630    0.094 0.924983
## tagp4:is_show    -0.95520    0.64031   -1.492 0.135761
## is_king_kong:tagp2:is_show 12.54198 284.75876 0.044 0.964869
## is_king_kong:tagp3:is_show 12.85328 284.75878 0.045 0.963998
## is_king_kong:tagp4:is_show 13.04986 284.75869 0.046 0.963447
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1826.7 on 2963 degrees of freedom
## Residual deviance: 1685.0 on 2946 degrees of freedom
## AIC: 1721
##
## Number of Fisher Scoring iterations: 15
```

A similar analysis is conducted on the interaction of sorting, king kong seats, and likelihood labels using a logit regression below. The hypothesis in such a scenario is a certain level of likelihood may benefit more from having the seats sorted. Again, no term involving sorting appears to have any significance.

```
data_b %>%
  glm(is_pay ~ is_king_kong*tag*is_sort + is_show + prob,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()

##
## Call:
## glm(formula = is_pay ~ is_king_kong * tag * is_sort + is_show +
##      prob, family = binomial(link = "logit"), data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0073  -0.5047  -0.3418  -0.2681   2.9951
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.45980    0.39334  -8.796 < 2e-16 ***
## is_king_kong   -1.06798    1.07694  -0.992  0.32135
## tagp2          0.45223    0.44216   1.023  0.30642
## tagp3          0.97607    0.45450   2.148  0.03175 *
## tagp4          1.55971    0.49201   3.170  0.00152 **
## is_sort        0.02519    0.54422   0.046  0.96309
## is_show        0.08220    0.13112   0.627  0.53071
## prob          2.77010    1.50377   1.842  0.06546 .
```

```

## is_king_kong:tagp2          1.27330    1.13503    1.122    0.26194
## is_king_kong:tagp3          1.49645    1.13194    1.322    0.18616
## is_king_kong:tagp4          0.86059    1.11794    0.770    0.44142
## is_king_kong:is_sort       -12.15404   317.78416   -0.038    0.96949
## tagp2:is_sort               -0.24029    0.63002   -0.381    0.70290
## tagp3:is_sort               -0.26678    0.65396   -0.408    0.68332
## tagp4:is_sort               -0.17538    0.62546   -0.280    0.77917
## is_king_kong:tagp2:is_sort  13.00736   317.78454    0.041    0.96735
## is_king_kong:tagp3:is_sort  12.74493   317.78455    0.040    0.96801
## is_king_kong:tagp4:is_sort  12.33468   317.78448    0.039    0.96904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1826.7  on 2963  degrees of freedom
## Residual deviance: 1690.6  on 2946  degrees of freedom
## AIC: 1726.6
##
## Number of Fisher Scoring iterations: 15

```

### 3.4 Conclusion

After examination of both direct and indirect effects from labelling and sorting, no correlation between labelling and sorting, and conversion has been discovered.

## 4 Possible Mechanism

Because the data is experimental from a randomised block design where sorting and labelling have an equal chance of being assigned to class without influences from other factors, they are considered as exogenous. That means no unobserved factors in the random components have an impact on both treatment and conversion. Therefore, the above regression analysis is valid with a consistent estimation of the true effects of treatment.

I consider three possible reasons for the inconsequential roles of labelling and sorting:

1. In the estimation of conversion rate, which can be represented by variable **prob** and consequently **tag**, elements that contribute to the majority of the variation of conversion have been included. Some of those elements that contribute to the variation of conversion are also what drives the effects of labelling and sorting. From this perspective, the insignificance of labelling and sorting is a result of underlying factors being covered in the estimated conversion rate.
2. The majority of impacts from labelling comes from that of sorting, which then is included in the effect of king kong seats. In many classes, being in the king kong seats is a result of sorting by likelihood labels. Sorting contributes to conversion indirectly by putting students in king kong seats.
3. Little variation in conversion can be explained or predicted by observed variables.

Starting with the first possible reason, because the algorithm producing the estimated conversion rate is unknown, no theoretical or analytical analysis can be made on the underlying factors. Instead, the explanatory power of the estimated conversion rate represented by likelihood labels is examined. The average marginal effect (AME) of likelihood labels (**tag**) in a logit model using data set A with only the labels as regressor is calculated in Table 2. The labels rather than the estimated probability itself is used for better interpretability. With a baseline of P1 tag, having a tag of P4 increases the probability of conversion by 0.15, which is not small in probability terms. This is an indication of the strong explanatory power of the estimated conversion rate.

In the second possible reason, king kong seats serve as a proxy of labelling and sorting. I have regressed king kong seats on estimated conversion rate and interaction between likelihood labels, sorting, and label showing

Table 2: Average marginal effects of estimated conversion rate represented using labels.

tagp2	0.02395
tagp3	0.07453
tagp4	0.1514

using a logit model. The result is shown below. In this model, whether the student is in a king kong seat can be affected by sorting, especially when the student has a high likelihood label (P3 or P4), which can be seen from the significant coefficients of interaction between P3 or P4 tag and sorting. This is expected: this is exactly what sorting is meant to do.

```
data_b %>%
  glm(is_king_kong ~ tag*is_sort*is_show + prob,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()

##
## Call:
## glm(formula = is_king_kong ~ tag * is_sort * is_show + prob,
##      family = binomial(link = "logit"), data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4177  -0.8764  -0.8015   1.2465   1.7885
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.81850    0.19225  -4.257 2.07e-05 ***
## tagp2          -0.00696    0.23127  -0.030 0.975992
## tagp3           0.09708    0.27138   0.358 0.720552
## tagp4           0.14210    0.32771   0.434 0.664575
## is_sort        -0.56454    0.27741  -2.035 0.041848 *
## is_show        -0.19202    0.25869  -0.742 0.457920
## prob           1.53933    1.25153   1.230 0.218712
## tagp2:is_sort    0.34574    0.32900   1.051 0.293320
## tagp3:is_sort    1.01086    0.35974   2.810 0.004954 **
## tagp4:is_sort    1.24083    0.36833   3.369 0.000755 ***
## tagp2:is_show    0.18277    0.31194   0.586 0.557939
## tagp3:is_show   -0.06426    0.34880  -0.184 0.853838
## tagp4:is_show    0.24813    0.34906   0.711 0.477173
## is_sort:is_show  0.35050    0.39699   0.883 0.377293
## tagp2:is_sort:is_show -0.25089  0.47087  -0.533 0.594153
## tagp3:is_sort:is_show  0.03174  0.51463   0.062 0.950817
## tagp4:is_sort:is_show -0.58790  0.52907  -1.111 0.266482
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 3821.2  on 2963  degrees of freedom
## Residual deviance: 3707.3  on 2947  degrees of freedom
## AIC: 3741.3
```

```
##
## Number of Fisher Scoring iterations: 4
```

With the effect of sorting on king kong seats being established, I would like to test the (indirect) effect of sorting on conversion by only regressing the conversion on labelling and sorting in the model below. After removing king kong seats from the regressors along with the estimated conversion rate, any impact made by those variables is now included in label-showing and sorting. While labelling and sorting are still not significant on the 5% level, labelling is significant on the 10% level and the coefficient of sorting has a p-value of 0.126: close enough to the 10% level. This is, to some extent, supportive of reason 2 and 1: the effects of labelling and sorting are covered by king kong seats and the model estimating conversion rate.

```
data_b %>%
  glm(is_king_kong ~ is_show + is_sort,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()

##
## Call:
## glm(formula = is_king_kong ~ is_show + is_sort, family = binomial(link = "logit"),
##      data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9484  -0.9395  -0.8948   1.4251   1.4892
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.68511     0.06892  -9.940  <2e-16 ***
## is_show      -0.02340     0.07791  -0.300    0.764
## is_sort       0.11936     0.07791   1.532    0.126
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 3821.2  on 2963  degrees of freedom
## Residual deviance: 3818.6  on 2961  degrees of freedom
## AIC: 3824.6
##
## Number of Fisher Scoring iterations: 4
```

Now it comes down to the third possible reason: maybe our attempt of using those observed data to explain conversion is futile after all. There may be many factors making a difference to conversion that can not be captured or recorded, for example, the wealth of a family and the occupations of the parents. It is likely that the observed data can only account for a minimal portion of the variation in conversion. Sadly, in the model analysing the direct effect (Section 3.2), this is suggested by the pseudo goodness of fit, a mere 0.07. A value this small indicates the model is not so useful in explaining the variation in conversion.

```
data_b %>%
  glm(is_pay ~ is_show + is_sort + prob + is_king_kong*tag,
      family = binomial(link = "logit"),
      data = .) %>%
  PseudoR2()

##      McFadden
```



```
## 0.07116617
```

A 10-fold cross-validation of the same model is also indicative of such a conclusion with an error rate that is as huge as 0.91. Although a logit model is not known for its accuracy, an error rate of 0.91 is still rather unacceptable: randomly guessing the outcome of conversion can do better than this model.

```
# 10 fold cross validationn
error_rate <- data_b %>%
  glm(is_pay ~ is_show + is_sort + prob + is_king_kong*tag,
      family = binomial(link = "logit"),
      data = .) %>%
  cv.glm(data_b,
        glmfit = .,
        cost = function(obs, fitted){
          sum(obs== as.numeric(fitted>=0.5))/length(obs)
        },
        K=10)
{cat("10 Fold Cross Validation Error Rate\n")
cat(error_rate$delta[[1]])}
```

```
## 10 Fold Cross Validation Error Rate
```

```
## 0.9075574
```

## 5 Further application

### 5.1 New Sorting Method

The idea is from an observation of the logit model in Section 3.2. The output of the model is again printed here.

```
data_b %>%
  glm(is_pay ~ is_show + is_sort + prob + is_king_kong*tag,
      family = binomial(link = "logit"),
      data = .) %>%
  summary()

##
## Call:
## glm(formula = is_pay ~ is_show + is_sort + prob + is_king_kong *
##      tag, family = binomial(link = "logit"), data = .)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9865  -0.4535  -0.3294  -0.2680   3.1667
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -3.46981    0.29124 -11.914  < 2e-16 ***
## is_show         0.08683    0.13088   0.663  0.507046
## is_sort         0.03826    0.13204   0.290  0.771983
## prob           2.82545    1.50281   1.880  0.060092 .
## is_king_kong   -1.59240    1.03970  -1.532  0.125621
## tagp2          0.33486    0.31714   1.056  0.291029
## tagp3          0.86861    0.33908   2.562  0.010417 *
## tagp4          1.48719    0.39406   3.774  0.000161 ***
## is_king_kong:tagp2 2.23906    1.06738   2.098  0.035931 *
```

```
## is_king_kong:tagp3  2.31993    1.06743    2.173 0.029752 *
## is_king_kong:tagp4  1.45076    1.06319    1.365 0.172398
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1826.7  on 2963  degrees of freedom
## Residual deviance: 1696.7  on 2953  degrees of freedom
## AIC: 1718.7
##
## Number of Fisher Scoring iterations: 7
```

Notice the interaction terms of king kong seats and tag P2 and P3 have quite significant positive coefficients. With the baseline being tag P1, this suggests being in the king kong seats is rather beneficial to increase the chance of conversion for students with medium likelihood tags (P2 and P3), compared to P1. This is not true for students who have a high likelihood tag of P4: the coefficient of interaction between king kong seats and tag P4 is not significant, so being in the king kong seats will be the same to P1 and P4 students.

The original method sorts students from the middle to the outer circle by their likelihood label. The more likely they make a purchase, the closer they are to the middle of the screen. But the observation on the behaviour of the interaction terms suggests a different method of sorting compared to what has been done up to this point. Since having a P4 tag may be indifferent to being in the king kong seats or not, putting students with a P2 or P3 tag to the middle of the screen may result in a better conversion rate overall. The order of tags based on this sorting method would be P3, P2, P4, and P1, from the middle to the outer circle.

Under this new definition of sorting, the analysis in this report can be done similarly and randomised block design can be produced the same way:

1. Showing labels in half of the classes, where the classes are selected randomly;
2. Sorting using this new method in half of the classes which are selected randomly.

The analysis of the explanatory power of this new sorting method on conversion can be evidential to the usefulness of new sorting. Conducting the same experiment with sorting of different orders of labels can help to find the best sorting order.

## 5.2 Coloured Labels

The purpose of the likelihood label is to attract the presenter's attention so they will focus on students who are more likely to make purchases in order to increase the conversion rate. The appearance of labels implemented for now is not different from other indicative information in an online class interface, which can only attract a limited amount of attention, especially when the presenter is focused mainly on teaching and delivery of the materials.

It is a general rule in visualisation that colours are pre-attentive: they attract attention before the reader even realise it (see Cleveland and McGill (1984) and Bostock and Heer (2009) among others). The same might be applied to user interface design. This feature of colour can be applied to draw the presenter's attention to the likelihood label. Colours with high saturation and high temperature (e.g. red and yellow) can be set to labels that need more attention to increase the chance of having a conversion. Under the assumption that more attention on students with a high likelihood to purchase can help with conversion rate, we can highlight tag P4 to P1 with colours from most saturated to least saturated. If it is the case discovered from interaction term in regression in Section 5.1, i.e. attention resulted from colours works better with medium likelihood tags (P3 and P2) the same way king kong seats works with them, then we can set P3 to have the most saturated colours, followed by P2, P4, and P1.

The power of colour on labels can be examined by adding another sampling group to label-showing and sorting in the randomised block design: adding colours in one of two ways mentioned above to half of the

classes which are chosen randomly. The same analysis in this paper then can be applied similarly to test the effect of adding colours and to determine what is the best format of colouring.

## 6 Summary

In this paper, I conclude that showing likelihood labels of purchase to the presenter and sorting students from the most likely purchaser to least likely have little impact to conversion, either directly or indirectly. Possible reasons have been discussed.

Further application of using a different sorting method to shift the presenter's attention to more needed students and adding colours to labels are proposed. Experiments testing the power of such applications are suggested.

## Appendix: Comments on the Inconsistency

### Variable description

In the project instruction, the user ID has been indicated as `user_id`, while in the data set it is actually `uid`. The same goes for classroom ID of `classroom_id` in the instruction and `cr_id` in the data set.

### Imbalanced Randomised Block Design

In data set B, the number of observations in different blocks is shown in Table 3. It does not strictly follow a randomised block design with an equal probability of being in each block.

Table 3: Number of observations in different block of randomised block design

	Not Sorted	Sorted
Label Not Shown	693	811
Label shown	854	606

### Number of Tags

In the project instruction, three tags are specified whereas in the data set there are four: P1, P2, P3, and P4.

## References

Bostock, Michael, and Jeffrey Heer. 2009. “Protovis: A Graphical Toolkit for Visualization.” IEEE.

Cleveland, William S, and Robert McGill. 1984. “Graphical Perception: Theory Experimentation and Application to the Development of Graphical Methods.” Taylor & Francis Group.