Analysis and Recommendation of Wine Retailer Business

2020.02.17

Highlines

1. Email marketing will increase the sales by **$1.35** on average
2. Best Slicing and dicing criterion: **Sav Blanc purchase history**
    - Will generate $8,341 profits
3. **55%** of the customers in experiment should be targeted for email marketing
    - Method: Casual Forest
    - Logic: Predicted Purchase Amount * Margin - Cost

Methodology

➢ Definition

- ***Causal Effect: τ = y1-y0***

  (y1: money spent with email, y0: money spent without email)

➢ Average Causal Effect

- Regression : lm(purch ~ group)
- ***β1 : Average Causal Effect*** of Email on Purchase

➢ Slice & Dice

- Baseline Variables
- Conditional causal effect
- Regression : lm(purch ~ group * havepurchased)
- Coefficient of interaction: ***difference in the effect*** of email between two groups

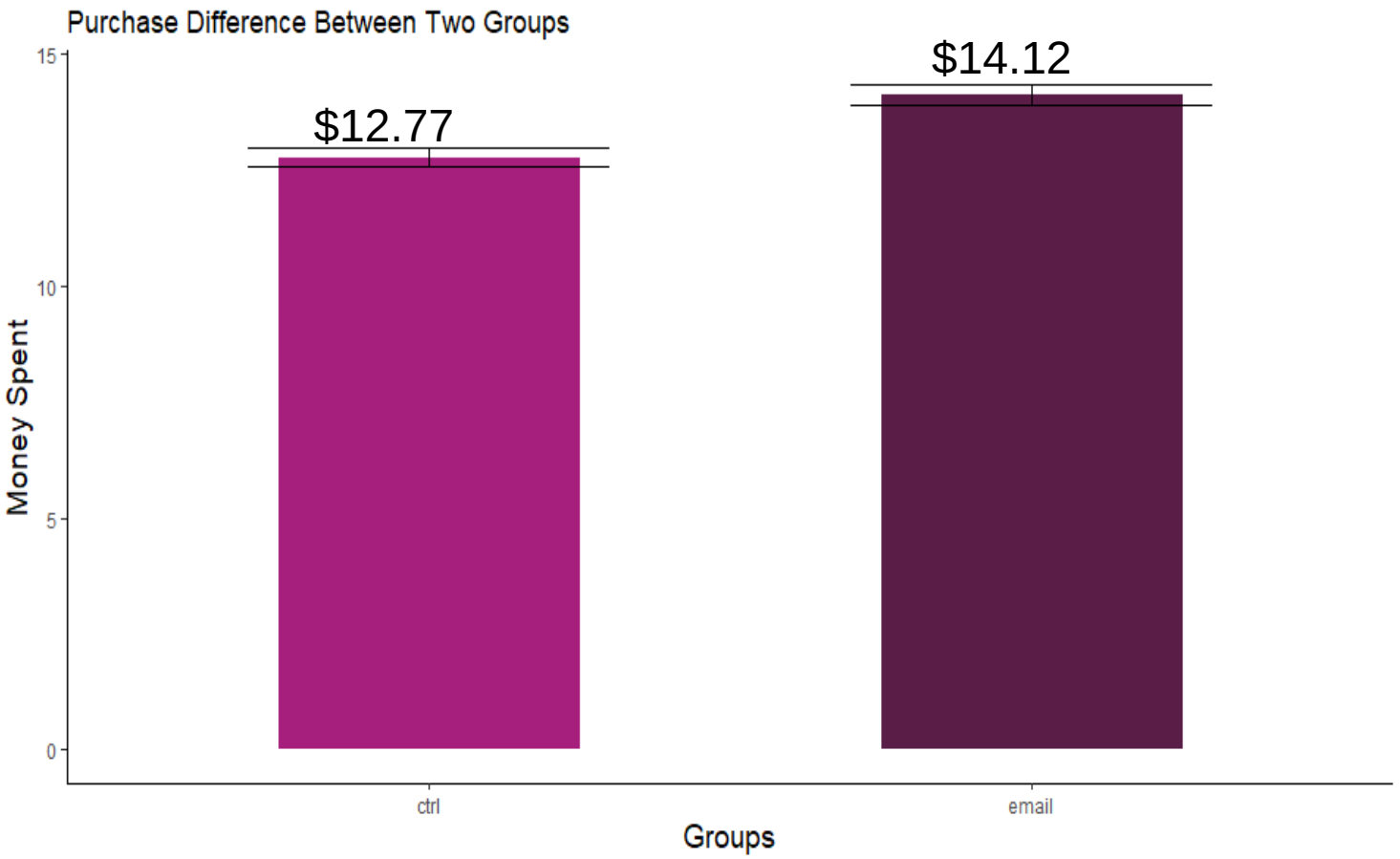
➢ Causal forest

- Models estimating conditional causal effects when causal variable is assigned randomly
- ***Predicting τ |X*** (causal effects using baseline variables)

➢ Scoring and targeting rules

- ***Score: τ * Margin-cost*** = τ * 0.3-0.1
- Targeting rules: ***Target when Score > 0***

Average Causal Effect



Email has potential to promote sales:

- The effect is highly significant, and the effect size is $1.35 for GroupEmail.
- It is just a general observation. To identify which groups of customers can purchase more, applying slicing and dicing analysis is necessary.

Coefficients	Esitimate	Std Error	t Value	Pr(>|t|)	Significance
(Intercept)	12.77	0.23	56.53	< 2e-16	***
groupemail	1.35	0.32	4.21	2.52E-05	***

Slicing and Dicing Analysis

```
Call:
lm(formula = purch ~ group * recentPurch, data = d)

Residuals:
    Min      1Q  Median      3Q     Max
 -21.03  -19.29   -7.66   -6.74 1791.47

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                6.7399     0.3102  21.727   <2e-16 ***
groupemail                 0.9158     0.4393   2.085   0.0371 *
recentPurchTRUE           12.5536     0.4475  28.054   <2e-16 ***
groupemail:recentPurchTRUE  0.8229     0.6328   1.301   0.1934
```

- Slicing and dicing by recent purchase(last purchase<60)
- There is no difference between recent group and less recent purchase group
- Not a good example of slicing and dicing.

⇦ **No Significant**

```
Call:
lm(formula = purch ~ group * havepurchased, data = d)

Residuals:
    Min      1Q  Median      3Q     Max
 -35.85  -11.08  -10.06  -10.06 1801.42

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                  10.0597     0.2373  42.387  < 2e-16 ***
groupemail                    1.0230     0.3358   3.046  0.00232 **
havepurchasedTRUE            22.4822     0.6832  32.907  < 2e-16 ***
groupemail:havepurchasedTRUE  2.2825     0.9629   2.370  0.01777 *
```

- Slicing and dicing by past purchase amount(past purchase>300)
- There is *significant difference* between the groups with past purchase over $300 and not
- Profit: ((1.0230+2.2825)*0.3-0.1)*4801=4,280.81

⇦ **All Significant**

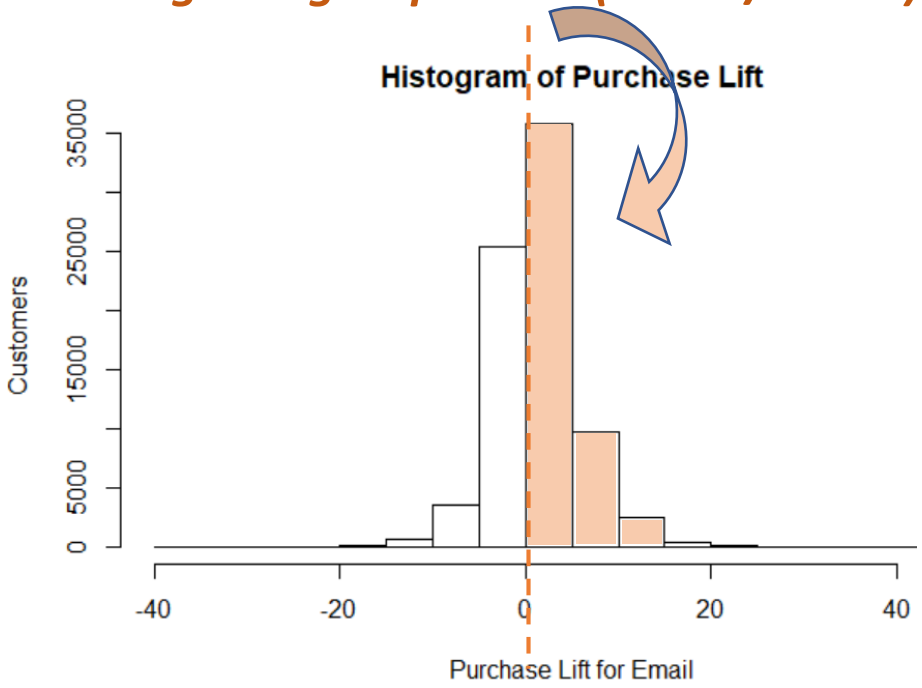
```
Call:
lm(formula = purch ~ group * boughtsavblanc, data = d)

Residuals:
    Min      1Q  Median      3Q     Max
 -17.56  -12.73  -11.98  -11.98 1794.94

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                   11.9785     0.2672  44.824  < 2e-16 ***
groupemail                     0.7541     0.3781   1.995  0.04609 *
boughtsavblancTRUE             2.7750     0.4995   5.555 2.78e-08 ***
groupemail:boughtsavblancTRUE  2.0504     0.7060   2.904  0.00368 **
```

- Slicing and dicing by purchasing Sauvignon blanc (sav_blanc>0)
- There is significant difference between the groups with Sauvignon blanc purchase history and not
- Profit: (2.8045*0.3-0.1)*11251= 8,340.93 > 4,280.81
- The best example of slicing and dicing.

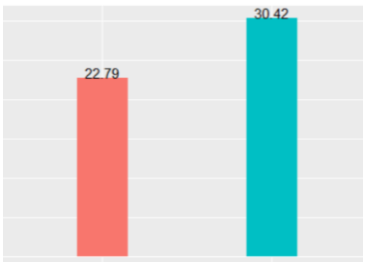
⇦ **All Significant**

Causal Forest Prediction

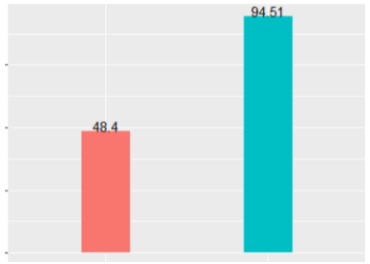
- Scoring and targeting rules
  - ***Score: τ * Margin-cost*** = τ * 0.3-0.1
  - Targeting rules: ***Target when Score > 0***
  - ***Targeted group: 55% (43416/78312)***



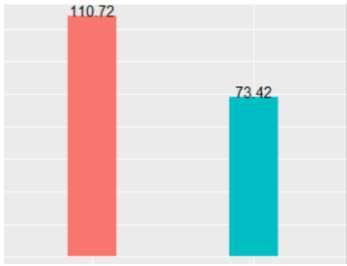
- **Summary of targeted group**
  - *Mean Value* of baseline variables (recent purchase, money spent on Chardonnay, Cabernet Sauvignon, Sauvignon blanc, Syrah wine and times of visit.
  - *Consumers we should target*:
  - Purchasing more Cab, more Chard, more Sav_blanc
  - Purchasing less Syrah
  - With closer last purchase date



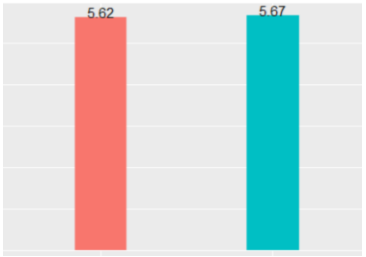
Cab($)



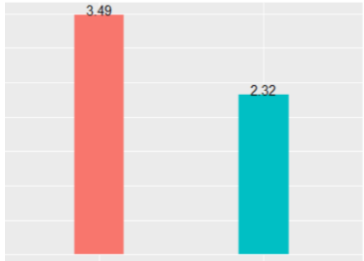
Chard ($)



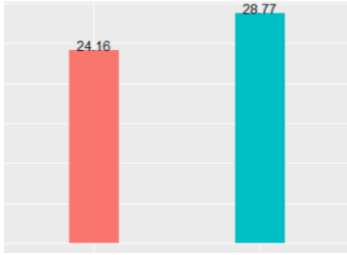
Last-purchase(days)



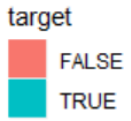
Visits (times)



Syrah ($)

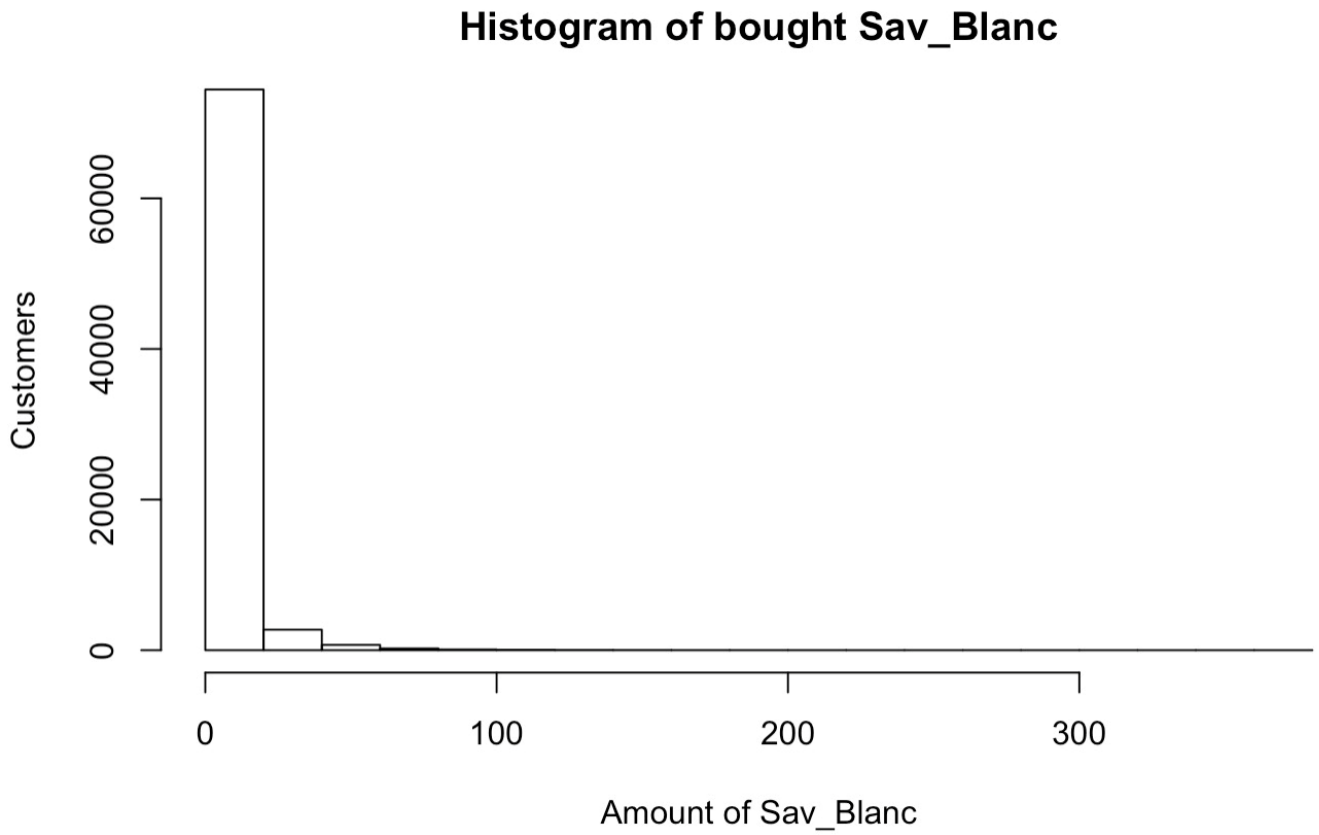


Sav_blanc ($)



Q&A

Appendix



- **Summaries of the baseline variables for targeted and non-targeted groups.**

target	last purchase	chard	sav blanc
TRUE	72.94	94.51	28.77
FALSE	111.45	48.4	24.16
target	syrah	cab	visits
TRUE	2.32	30.42	5.67
FALSE	3.49	22.79	5.62