



# Wine Retailer Case

MSMA Group II

Qianyu Dong | Sichun Li | Xinbei Jin | Congrong Shen | Xiao Shi

# Key Insights

- The average causal effect of **emails** is **highly significant** and its size is **\$1.26**.
- Randomized check shows our email effect is **randomly assigned**, hence no bias would be in experiment.
- Sending email to consumers who **purchased wine before** would have stronger effectiveness. Emails are especially effective to **Sauvignon Blanc buyers**.
- According to scoring results, **55%** of training data should be targeted.
- An average customer from **targeted group** did his/her last purchase about **2.5 months ago**, spent an average **\$94 on Chardonnay**, **\$28 on Sauvignon Blanc** and **\$31 on Cabernet Sauvignon**.

# Methodology

- Conducted **randomization check** to test whether experimental manipulation is assigned randomly to all baseline variables.
- Used **slice and dice analysis** to illustrate the potential for targeting on responses for the email campaign.
- Built **causal forest model** to obtain best estimates for causal effect and make individual-level predictions. (X: "last\_purchase", "visits", "chard", "sav\_blanc", "syrah", "cab"; Y: "purchase"; w: "group")
- **Score = 30%(Margin rate)\*Lift - 0.1(Cost)**, where Lift is the prediction of causal forest, which equals to the variance between purchase with email and purchase without email.
- **Targeting via Scoring**: target consumers whose **score > 0**.

# Average Causal Effects

```
lm(formula = purch ~ group + chard + sav_blanc + syrah + cab +  
  last_purch + visits, data = d)
```

Residuals:

Min	1Q	Median	3Q	Max
-420.37	-14.57	-10.31	-1.72	1798.77

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	14.5269957	0.4363336	33.293	< 2e-16 ***
groupemail	1.2603997	0.3101382	4.064	4.83e-05 ***
chard	0.0346117	0.0007959	43.489	< 2e-16 ***
sav_blanc	0.0433309	0.0020630	21.004	< 2e-16 ***
syrah	0.0240070	0.0149648	1.604	0.109
cab	0.0489413	0.0020948	23.363	< 2e-16 ***
last_purch	-0.0718125	0.0017235	-41.667	< 2e-16 ***
visits	-0.0627548	0.0655217	-0.958	0.338

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 43.39 on 78304 degrees of freedom

Multiple R-squared: 0.05836, Adjusted R-squared: 0.05827

F-statistic: 693.3 on 7 and 78304 DF, p-value: < 2.2e-16

---

\* Variable “past\_purch” is dropped because it is perfectly collinear with the purchases by categories

- The experimental manipulation (“group”) **passed the randomization check**
- Customers spent \$14.53 without receiving the emails
- The effect of **emails** is **highly significant** and the effect size is **\$1.26**
- Past purchases of chard / sav\_blanc / cab and the amount of last purchase are also significant

# Slice and Dice Analysis

```
Call:
lm(formula = purch ~ purch_sav_blanc + group:purch_sav_blanc,
    data = d)
```

Residuals:

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

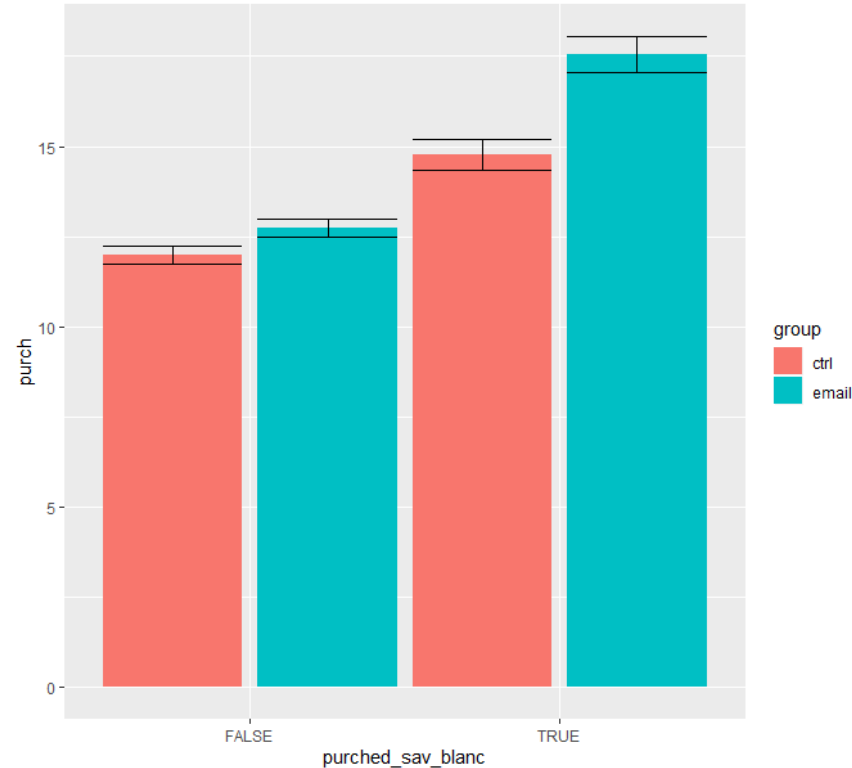
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	11.9785	0.2672	44.824	< 2e-16	***
purch_sav_blancTRUE	2.7750	0.4995	5.555	2.78e-08	***
purch_sav_blancFALSE:groupemail	0.7541	0.3781	1.995	0.0461	*
purch_sav_blancTRUE:groupemail	2.8045	0.5963	4.704	2.56e-06	***

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Effect of email is **significant** to both “have purchased Sauvignon Blanc” group and “haven’t purchased Sauvignon Blanc” group.

Emails are **more effective** for **consumers who purchased Sauvignon Blanc**, bringing a purchase lift of **\$2.80**, while for those who **haven’t purchased Sauvignon Blanc**, emails can only raise the purchase by **\$0.75**.



# Summary of Target Customers

- According to scoring results, **55%** of training data should be targeted.

targeted	43,103	55%
non_targeted	35,209	45%

- An average customer from targeted group** did his/her last purchase about 2.5 months ago, spent an average \$94 on Chardonnay, \$28 on Sauvignon Blanc and \$31 on Cabernet Sauvignon.

istarget	last_purch	vistis	chard	sav_blanc	syrah	cab
targeted	72.90017	5.659142	93.69655	28.16207	2.298140	30.50920
non_targeted	111.06251	5.631969	49.91804	24.94670	3.504276	22.76584

---

\* “last\_purch”, “visits”, “chard”, “sav\_chard”, “syrah”, “cab” are average values.