

第34組

統計期末報告 — 網購因素分析

組員： B084012001 李冠融
B084012010 戴子晴



Table of contents

01

建立模型

03

確認模型

02

修正模型

04

管理意涵與
實務應用





01

模型建立

網購行為小調查

*必填

未命名區段

性別 *

☐ 女

☐ 男

☐ 其他: _____

年齡 *

您的回答 _____

月薪/無收入可填每月生活費 *

您的回答 _____

是否會因特殊節慶增加網購金額 *

☐ 會

☐ 不會

是否會為了追趕流行而進行網路購物(如明星代言、親友推薦) *

☐ 會

☐ 不會

天氣會影響您的網購動機嗎 *

☐ 會



☐ 不會

上個月的網購金額 *

您的回答 _____

問卷設計!

- 樣本數: 135
- 預設自變數: 性別、年齡、月收入、特殊節慶、跟隨流行、天氣
- 依變數: 網購金額

 {r}

```
data <- fread("C://Users//Gladice//Desktop//R 下//試做9.csv")

data$gender <- ifelse(data$gender=="male",1,0)
data$holiday <- ifelse(data$holiday=="yes",1,0)
data$trend <- ifelse(data$trend=="yes",1,0)
data$weather <- ifelse(data$weather=="yes",1,0)

agetable <- data.frame(data$age)
dummyage <- dummy.data.frame(agetable)
colnames(dummyage) <- c("agea", "ageb", "agec")
data <- cbind(data, dummyage)

data1m <- lm(spend~gender+income+agea+ageb+holiday+trend+weather, data
=data)
summary(data1m)
```

模型檢定

Residual standard error: 7849 on 127 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6342
F-statistic: 34.19 on 7 and 127 DF, p-value: $< 2.2e-16$

Analysis of Variance Table

Response: spend

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gender	1	1.8216e+08	1.8216e+08	2.9566	0.08796 .
income	1	1.3926e+10	1.3926e+10	226.0379	$< 2e-16$ ***
agea	1	2.8150e+08	2.8150e+08	4.5691	0.03447 *
ageb	1	1.6594e+08	1.6594e+08	2.6933	0.10324 .
holiday	1	1.8812e+08	1.8812e+08	3.0535	0.08298 .
trend	1	1.1534e+06	1.1534e+06	0.0187	0.89138
weather	1	7.5126e+04	7.5126e+04	0.0012	0.97220
Residuals	127	7.8245e+09	6.1610e+07		

- P-value的數值具有顯著程度，足以拒絕

H_0 :全部複迴歸係數皆為0的假設

- 除了trend及weather二自變數以外，其餘自變數皆有相當的顯著程度

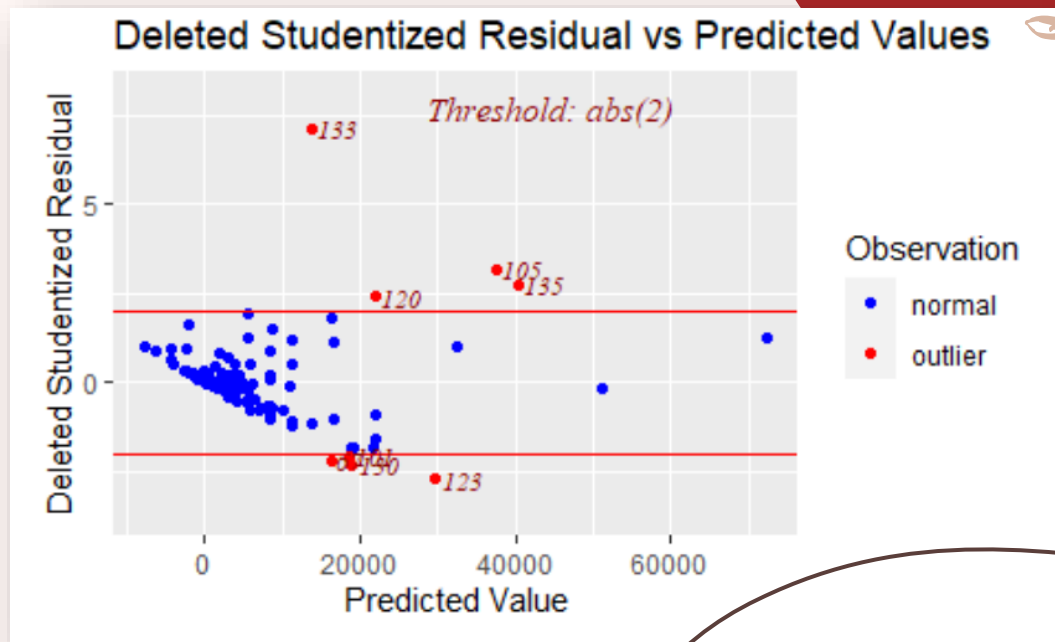
- 多重共線性的檢定部分，所有自變數的數值皆 <10

signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

gender	income	agea	ageb	holiday	trend	weather
1.274881	1.642605	2.555288	1.898849	1.119814	1.091761	1.050130

殘餘項分析

- 使用Deleted Studentized Residual的方式
- 資料的殘差值相當集中於左方中央部分，有越往右越擴散的分布狀況
- 些許離群值





02

修正模型

增刪自變數

```
data_1 <- lm(spend~income,data=data)
```

Residual standard error: 8178 on 133 degrees of freedom
Multiple R-squared: 0.6059, Adjusted R-squared: 0.6029
F-statistic: 204.4 on 1 and 133 DF, p-value: < 2.2e-16

```
data_2 <- lm(spend~gender+income,data=data)
```

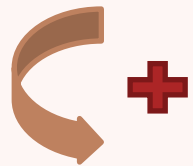
Residual standard error: 8006 on 132 degrees of freedom
Multiple R-squared: 0.6251, Adjusted R-squared: 0.6194
F-statistic: 110 on 2 and 132 DF, p-value: < 2.2e-16

```
data_3 <- lm(spend~gender+income+agea+ageb,data=data)
```

Residual standard error: 7851 on 130 degrees of freedom
Multiple R-squared: 0.6449, Adjusted R-squared: 0.634
F-statistic: 59.03 on 4 and 130 DF, p-value: < 2.2e-16

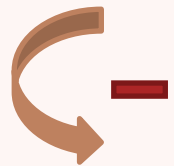
增刪自變數

```
data_4 <- lm(spend~gender+income+agea+ageb+holiday,data=data)
```



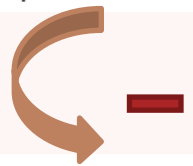
Residual standard error: 7789 on 129 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6398
F-statistic: 48.61 on 5 and 129 DF, p-value: < 2.2e-16

```
data_5 <- lm(spend~gender+income+agea+ageb+holiday+trend,data=data)
```



Residual standard error: 7819 on 128 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6371
F-statistic: 40.2 on 6 and 128 DF, p-value: < 2.2e-16

```
data_6 <- lm(spend~gender+income+agea+ageb+holiday+weather,data=data)
```



Residual standard error: 7819 on 128 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.637
F-statistic: 40.19 on 6 and 128 DF, p-value: < 2.2e-16



03

確認模型

修正後模型

```
data1m_new <- lm(spend~gender+income+agea+ageb+holiday,data=data)
summary(data1m_new)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-7.896e+03	2.502e+03	-3.157	0.00199	**
gender	-2.500e+03	1.497e+03	-1.670	0.09744	.
income	2.660e-01	1.973e-02	13.481	< 2e-16	***
agea	6.194e+03	2.138e+03	2.897	0.00443	**
ageb	2.900e+03	2.017e+03	1.438	0.15291	
holiday	2.751e+03	1.562e+03	1.761	0.08061	.

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

Residual standard error: 7789 on 129 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6398
F-statistic: 48.61 on 5 and 129 DF, p-value: < 2.2e-16

模型檢定比較

Residual standard error: 7849 on 127 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6342
F-statistic: 34.19 on 7 and 127 DF, p-value: < 2.2e-16

Analysis of Variance Table

Response: spend

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gender	1	1.8216e+08	1.8216e+08	2.9566	0.08796 .
income	1	1.3926e+10	1.3926e+10	226.0379	< 2e-16 ***
agea	1	2.8150e+08	2.8150e+08	4.5691	0.03447 *
ageb	1	1.6594e+08	1.6594e+08	2.6933	0.10324
holiday	1	1.8812e+08	1.8812e+08	3.0535	0.08298 .
trend	1	1.1534e+06	1.1534e+06	0.0187	0.89138
weather	1	7.5126e+04	7.5126e+04	0.0012	0.97220
Residuals	127	7.8245e+09	6.1610e+07		

修正前

Residual standard error: 7789 on 129 degrees of freedom
Multiple R-squared: 0.6533, Adjusted R-squared: 0.6398
F-statistic: 48.61 on 5 and 129 DF, p-value: < 2.2e-16

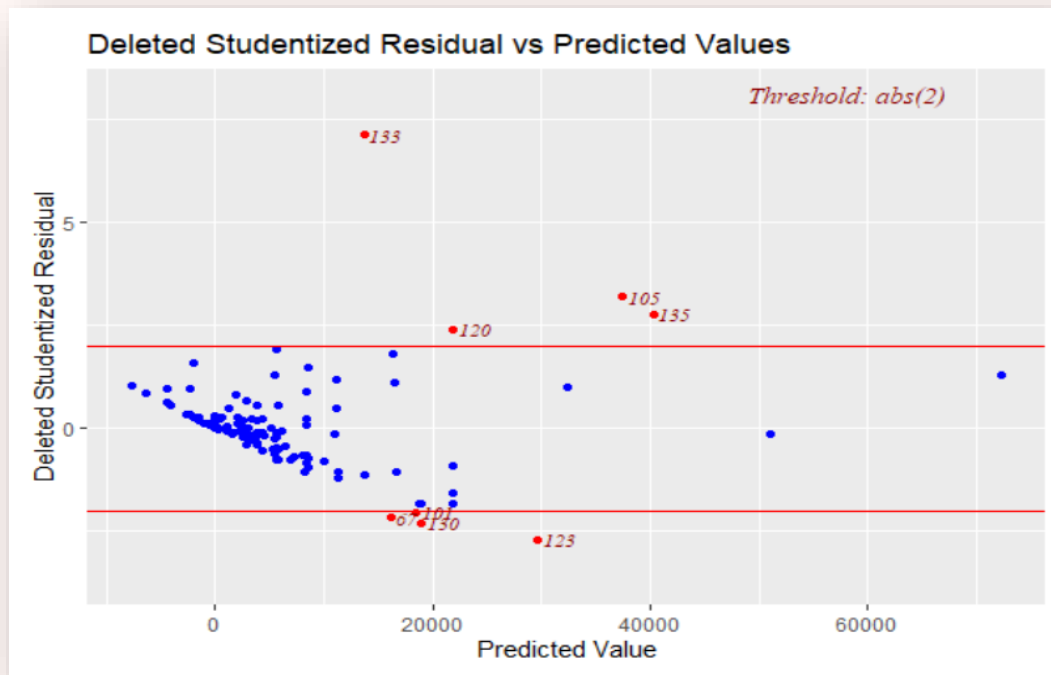
Analysis of Variance Table

Response: spend

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gender	1	1.8216e+08	1.8216e+08	3.0027	0.08551
income	1	1.3926e+10	1.3926e+10	229.5615	< 2e-16
agea	1	2.8150e+08	2.8150e+08	4.6403	0.03309
ageb	1	1.6594e+08	1.6594e+08	2.7353	0.10058
holiday	1	1.8812e+08	1.8812e+08	3.1011	0.08061
Residuals	129	7.8257e+09	6.0664e+07		

修正後

模型修正後的 殘餘項分析



離群值

樣本編號	性別 (gender)	月收入 (income)	節慶 (holiday)	年齡 (age)	網購金額 (spend)
67	男	100000	不會	50+	0
101	男	98000	會	50+	3000
105	女	130000	會	25~49	10000
120	女	100000	會	50+	2000

樣本編號	性別 (gender)	月收入 (income)	節慶 (holiday)	年齡 (age)	網購金額 (spend)
105	女	150000	會	25~49	60000
120	女	100000	會	25~49	40000
133	女	80000	會	50+	60000
135	女	180000	會	50+	60000

自變數相關性

	gender	age	income	holiday	trend	weather	spend
gender	1.00000000	0.31820612	0.286217080	0.07052260	-0.18723149	-0.060740765	0.08983864
age	0.31820612	1.00000000	0.584223008	0.26166127	-0.15510713	-0.017241657	0.33696895
income	0.28621708	0.58422301	1.000000000	0.17546261	-0.10497096	0.002728857	0.77836565
holiday	0.07052260	0.26166127	0.175462605	1.00000000	0.08917149	0.086375124	0.20635152
trend	-0.18723149	-0.15510713	-0.104970960	0.08917149	1.00000000	0.177340439	-0.04677184
weather	-0.06074076	-0.01724166	0.002728857	0.08637512	0.17734044	1.00000000	0.01353639
spend	0.08983864	0.33696895	0.778365650	0.20635152	-0.04677184	0.013536389	1.00000000

- Income <-> Spend



04

管理意涵與 實務應用

商品行銷

01

Income

消費者的購物
金額會隨著收
入的上升而增
加

02

Age

消費者的網路
購物金額會隨
著年齡增長而
下降

03

Holiday

逢臨特殊節
慶時，消費者
的網路購物金
額上升

04

Gender

女性消費者的
網路購物金額
較高

The background features several decorative elements: a black line-art branch with small buds in the top-left; a large, solid maroon shape in the top-right with a small leafy branch extending from it; a large, solid tan shape in the bottom-left with two long, thin leaf outlines; and a simple black line forming a large, open curve in the bottom-right.

Thanks for Your
Listening