

商管統計資料分析 期末報告

分析康是美高頻率顧客對美妝產品的消費習慣

第13組

工管三 蘇德倫

工管三 林心慈

工管三 張育瑄

工管三 蔡銓驊

工管三 馮依華

工管三 林昱辰

工管三 林翊翎

目錄

一、議題探討原因.....	P.2
二、更換通路原因.....	P.2
三、預期分析目標.....	P.2
四、消費習慣資料分析.....	P.3
五、問題探討.....	P.11
六、銷售策略制定與結論.....	P.13
七、附錄.....	P.14

一、議題探討原因

透過哈佛商業評論的內容 “ Companies can boost profits by almost 100% by retaining just 5% more of their customers.” 得知，留住既有顧客可以有效的增加企業的利潤，因此我們這邊以高頻顧客為我們主要探討的受眾，觀察其高頻顧客對的消費行為，是否會對其客單價有影響。

二、更換通路原因

- 不選小三美日的原因：

由於資料筆數太少，僅管已考慮許多變數與interaction，模型的解釋力仍太低。為了更好地反映出分析的結果，因此決定更換通路。

- 選康是美的原因：

我們組特別想探討折扣對銷售的影響，但資料中屈臣氏的品項欄位幾乎沒有明確記錄折扣的欄位；反觀康是美在品項那欄有明確的名稱，例如：全店88折、不限金額85折、買1送1、點數折現、現金折價券回券、第2件5折等等...。因此我們將鎖定「康是美」作為本次報告主要討論的通路。

三、預期分析目標

- 以康是美客單價為衡量基準，藉由時間、地點、購買數量、購買頻率以及折扣等變因，分析高頻率與非高頻率顧客消費習慣差異，作為康是美銷售策略制定參考，並以提升其業績為最終目標。
- 問題探討：
 1. 是否為高頻率顧客會不會影響客單價
 2. 有無折扣是否影響成為忠實顧客的意願
 3. 將折扣種類區分成六種類型，分析折扣種類的顯著程度以探討對客單價的影響
 4. 了解星期幾的客單價最高，作為後續銷售策略參考
 5. 了解六都中哪個縣市對客單價的貢獻比較顯著，以納入分區銷售策略的考量
 6. 不同產品種類（保養、彩妝、底妝）對客單價的影響是否不同

四、消費習慣資料分析

- 高頻率顧客定義 - 以消費頻率進行篩選

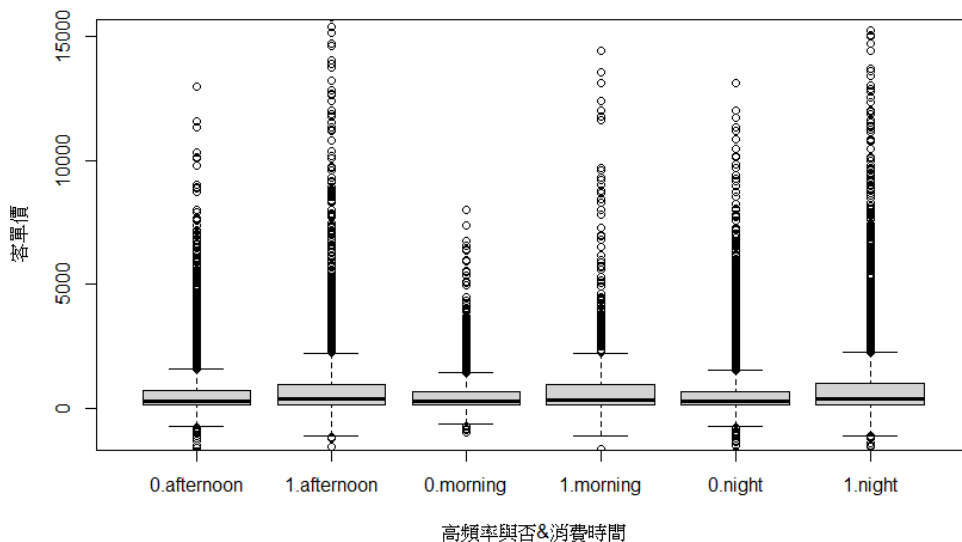
原先我們希望運用RFM模型(註1)來篩選出忠實顧客，但由於若R、F、M任選二或三個篩選，會出現筆數不足的問題。因此，我們僅選擇Frequency (消費頻率) 作為篩選項目，以Frequency > 1為篩選條件式，並將區隔出來的消費者命名為：高頻率顧客。

(*註1：以 Recency (最近一次消費)、Frequency (消費頻率)、Monetary (消費金額) 三個項目來區分顧客類型，以利分析消費者行為及後續策略制定的重要工具)

- 資料EDA (高價值顧客Dataset=1)

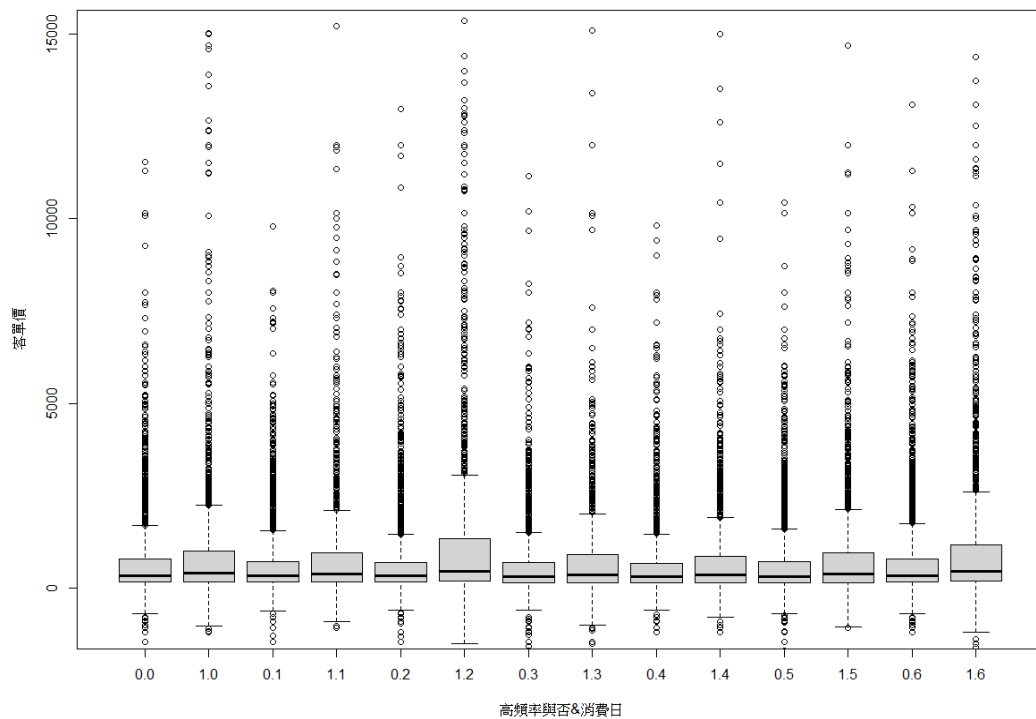
首先我們用消費時間、消費日、消費品項、消費時間四個變因來觀察高頻顧客的是否會影響客單價。

1. 探討「是否為高頻率顧客」及「銷售時間」的不同組合與「客單價」的關係



不論是早上中午晚上，都可以看出高頻顧客的客單價中位數、以及盒鬚圖上界處都高於三個月內只有單次消費的消費者，可以看出高頻顧客會影響客單價高低。但是這邊比較難看出早中晚之間的差異。

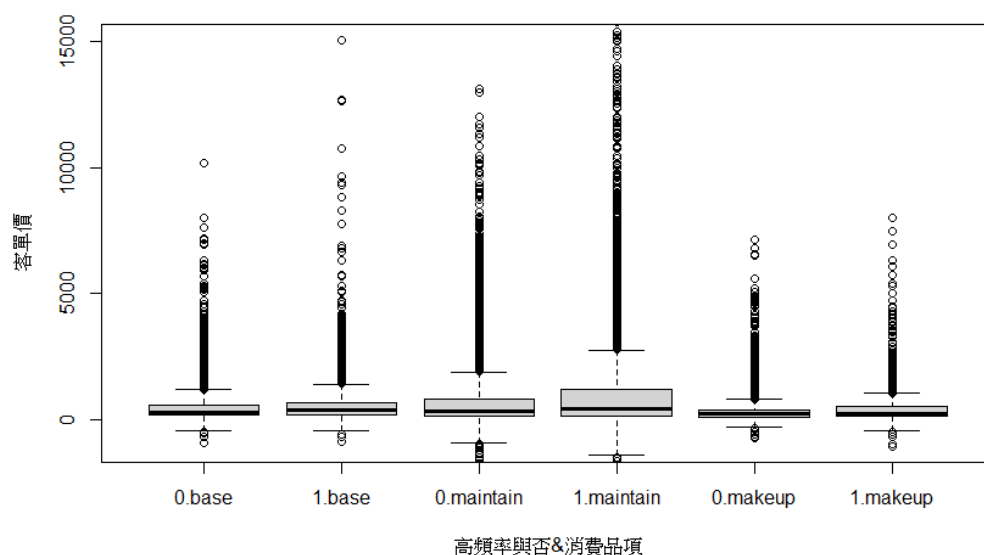
2. 探討「是否為高頻率顧客」及「星期幾」的不同組合與「客單價」的關係



從這張圖也可以看到類似的狀況，高頻顧客的整個盒鬚圖在客單價方面都比非高頻顧客高，可以說明高頻顧客的變異度較單次消費者高，其中又以周二跟週六最高，而週二單次消費者客單價的變異程度也是最小的

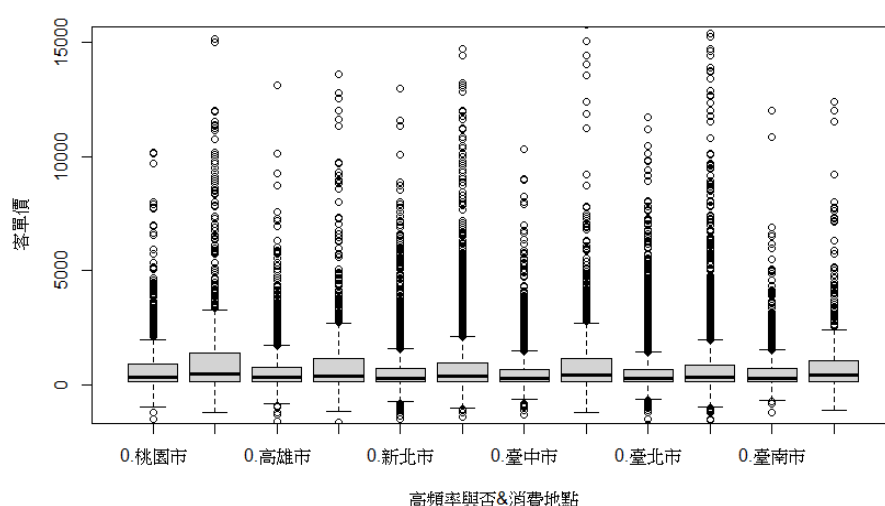
3. 探討「是否為高頻率顧客」及「消費品項」的不同組合與「客單價」的關係

```
boxplot(cosmed$cos_invo_price ~ factor(cosmed$Whether_F)+factor(cosmed$category),xlab = "高頻率與否&消費品項",ylab="客單價",ylim=c(-1000,15000))
```



在這邊可以看到各類化妝品在客單價方面，中位數的差距不大，大多都是高頻顧客略優於單次消費顧客，但這邊可以觀察到底妝跟彩妝的變異程度低，代表消費者傾向在康是美購買一定金額的產品，不容易超出範圍，也不會一次買很多。但是在保養方面，變異程度就相對其他兩個種類大很多，且我們也得知保養的消費金額變異度大。

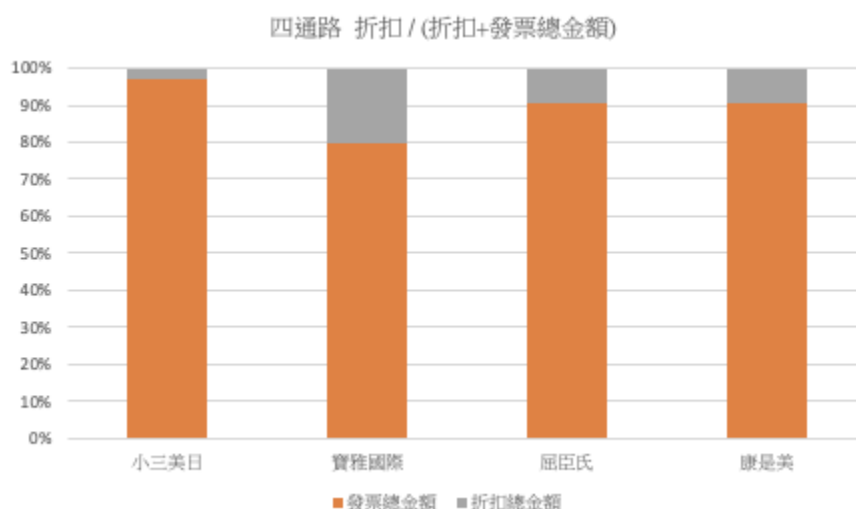
4. 探討「是否為高頻率顧客」及「消費地點」的不同組合與「客單價」的關係



由這張圖，可以各縣市之間的客單價比較差距不大，但是高頻顧客的消費客單價中位數高於單次消費者、變異程度較大。

小結：透過EDA盒鬚圖資料可得知，忠實顧客的客單價中位數較高，且變異程度。

5. 各通路折扣比較



由上圖可知屈臣氏和康是美折扣金額佔折扣與發票總金額比例較為相近，可推測兩通路性質較相似。兩通路與小三美日和寶雅國際有顯著差異。

● 模型建立

(1) Linear Model 1:

Response Variable: 客單價

Explanatory Variables: 美妝三個種類*uniprice + 折扣種類*totprice + 六都各縣市*單次購買量 + 商品購買量*美妝三個種類 + 有無折扣*購買頻率 + 有無折扣*單次購買量

```
lm1 <- lm(cos_invo_price ~ category*uniprice + discount_type*totprice +
city*invo_quant + quant*category + If_Discount*Freq + If_Discount*invo_
quant ,data=cosmed_F)
```

call:

```
lm(formula = cos_invo_price ~ category * uniprice + discount_type *
totprice + city * invo_quant + quant * category + If_Discount *
Freq + If_Discount * invo_quant, data = cosmed_F)
```

Residuals:

Min	1Q	Median	3Q	Max
-3011.56	-72.32	-44.52	25.19	2681.91

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.201e+01	9.348e+00	2.355	0.018524	*
categorymaintain	3.011e+01	8.293e+00	3.631	0.000283	***
categorymakeup	-2.130e+01	1.037e+01	-2.054	0.039965	*
uniprice	2.343e-01	1.557e-02	15.045	< 2e-16	***
discount_type加價購	7.208e+00	2.278e+01	0.316	0.751725	
discount_type品項82折	3.079e+02	7.760e+00	39.681	< 2e-16	***
discount_type買1送1	2.904e+02	1.015e+01	28.611	< 2e-16	***
discount_type滿388現折50元	4.377e+02	2.809e+01	15.583	< 2e-16	***
discount_type滿888現折100元	8.712e+02	2.491e+01	34.969	< 2e-16	***
totprice	9.163e-01	8.439e-03	108.576	< 2e-16	***
city高雄市	1.276e+01	6.166e+00	2.069	0.038525	*
city新北市	1.085e+01	5.074e+00	2.138	0.032481	*
city臺中市	1.461e+01	5.699e+00	2.564	0.010351	*
city臺北市	1.369e+01	4.963e+00	2.759	0.005803	**
city臺南市	6.955e+01	6.463e+00	10.761	< 2e-16	***
invo_quant	1.484e+02	1.664e+00	89.146	< 2e-16	***
quant	-1.420e+02	6.858e+00	-20.710	< 2e-16	***
If_Discount	-3.286e+02	4.968e+00	-66.140	< 2e-16	***
Freq	-1.530e-01	1.101e-01	-1.390	0.164603	
categorymaintain:uniprice	-8.805e-02	1.319e-02	-6.677	2.45e-11	***
categorymakeup:uniprice	5.981e-02	2.286e-02	2.616	0.008897	**
discount_type加價購:totprice	-9.205e-01	1.039e-01	-8.856	< 2e-16	***
discount_type品項82折:totprice	-1.071e+00	3.700e-02	-28.959	< 2e-16	***
discount_type買1送1:totprice	-5.466e-01	2.288e-02	-23.890	< 2e-16	***
discount_type滿388現折50元:totprice	5.265e-01	3.557e-01	1.480	0.138802	
discount_type滿888現折100元:totprice	6.216e-01	1.257e-01	4.947	7.57e-07	***
city高雄市:invo_quant	-6.051e+00	2.085e+00	-2.902	0.003714	**
city新北市:invo_quant	-9.445e+00	1.755e+00	-5.383	7.34e-08	***
city臺中市:invo_quant	-9.511e+00	1.956e+00	-4.863	1.16e-06	***
city臺北市:invo_quant	-1.190e+01	1.729e+00	-6.881	6.00e-12	***
city臺南市:invo_quant	-3.858e+01	2.078e+00	-18.567	< 2e-16	***
categorymaintain:quant	-2.374e+01	6.751e+00	-3.516	0.000438	***
categorymakeup:quant	8.860e+00	8.190e+00	1.082	0.279365	
If_Discount:Freq	-2.194e+00	5.511e-01	-3.981	6.87e-05	***
invo_quant:If_Discount	-1.434e+01	1.013e+00	-14.157	< 2e-16	***

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

Residual standard error: 228.2 on 84921 degrees of freedom
Multiple R-squared: 0.7042, Adjusted R-squared: 0.7041
F-statistic: 5946 on 34 and 84921 DF, p-value: < 2.2e-16

(2) Linear Model 2:

Response Variable: 客單價

Explanatory Variables: 購買頻率 + 是否是高頻率顧客 + 單次購買量*星期幾 + 美妝三個種類*uniprice + 折扣種類*totprice + 六都各縣市*單次購買量 + 商品購買量*美妝三個種類 + 有無折扣*購買頻率 + 有無折扣*單次購買量 + 是否為高頻率顧客*有無折扣

```
lm2<-lm(cos_invo_price~ Freq+Whether_F+invo_quant*day + category*u
niprice + discount_type*totprice + city*invo_quant + quant*category + If_
Discount*Freq + If_Discount*invo_quant + Whether_F*If_Discount ,data=c
osmed_F)
```

Call:

```
lm(formula = cos_invo_price ~ Freq + whether_F + invo_quant *
day + category * uniprice + discount_type * totprice + city *
invo_quant + quant * category + If_Discount * Freq + If_Discount *
invo_quant + whether_F * If_Discount, data = cosmed_F)
```

Residuals:

Min	1Q	Median	3Q	Max
-2996.74	-71.49	-43.84	25.22	2694.33

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.692e+01	9.677e+00	1.748	0.080443 .
Freq	-2.637e-01	1.128e-01	-2.338	0.019399 *
whether_F	9.481e+00	1.815e+00	5.224	1.76e-07 ***
invo_quant	1.483e+02	1.874e+00	79.096	< 2e-16 ***
day1	5.525e+00	4.297e+00	1.286	0.198565
day2	1.907e+00	4.300e+00	0.443	0.657414
day3	8.176e+00	4.447e+00	1.839	0.065980 .
day4	-4.710e+00	4.543e+00	-1.037	0.299795
day5	-1.439e+01	4.173e+00	-3.448	0.000565 ***
day6	1.706e+01	4.092e+00	4.169	3.06e-05 ***
categorymaintain	2.960e+01	8.285e+00	3.573	0.000353 ***
categorymakeup	-2.202e+01	1.036e+01	-2.125	0.033578 *
uniprice	2.344e-01	1.557e-02	15.057	< 2e-16 ***
discount_type加價購	6.375e+00	2.276e+01	0.280	0.779458
discount_type期限內品項82折	3.096e+02	7.788e+00	39.760	< 2e-16 ***
discount_type買1送1	2.932e+02	1.015e+01	28.903	< 2e-16 ***
discount_type滿388現折50元	4.410e+02	2.806e+01	15.716	< 2e-16 ***
discount_type滿888現折100元	8.716e+02	2.489e+01	35.020	< 2e-16 ***
totprice	9.146e-01	8.446e-03	108.288	< 2e-16 ***


```

city高雄市      1.377e+01  6.166e+00  2.233 0.025580 *
city新北市      1.026e+01  5.082e+00  2.018 0.043572 *
city臺中市      1.307e+01  5.697e+00  2.295 0.021738 *
city臺北市      1.387e+01  4.964e+00  2.794 0.005214 ***
city臺南市      6.479e+01  6.518e+00  9.940 < 2e-16 ***
quant           -1.416e+02  6.856e+00 -20.651 < 2e-16 ***
If_Discount     -3.329e+02  5.217e+00 -63.799 < 2e-16 ***
invo_quant:day1 -2.790e+00  1.426e+00 -1.957 0.050371 .
invo_quant:day2 -2.018e+00  1.408e+00 -1.433 0.151974
invo_quant:day3 -3.403e+00  1.462e+00 -2.327 0.019967 *
invo_quant:day4  5.040e+00  1.510e+00  3.337 0.000847 ***
invo_quant:day5  1.036e+01  1.379e+00  7.514 5.77e-14 ***
invo_quant:day6 -5.751e+00  1.328e+00 -4.332 1.48e-05 ***
categorymaintain:uniprice -8.663e-02  1.317e-02 -6.577 4.84e-11 ***
categorymakeup:uniprice  5.858e-02  2.283e-02  2.566 0.010301 *
discount_type加價購:totprice -9.413e-01  1.039e-01 -9.060 < 2e-16 ***
discount_type期限內品項82折:totprice -1.054e+00  3.701e-02 -28.483 < 2e-16 ***
discount_type買1送1:totprice -5.469e-01  2.286e-02 -23.928 < 2e-16 ***
discount_type滿388現折50元:totprice  5.318e-01  3.552e-01  1.497 0.134389
discount_type滿888現折100元:totprice  5.909e-01  1.257e-01  4.701 2.60e-06 ***
invo_quant:city高雄市 -6.671e+00  2.089e+00 -3.193 0.001409 **
invo_quant:city新北市 -9.532e+00  1.764e+00 -5.405 6.51e-08 ***
invo_quant:city臺中市 -8.864e+00  1.958e+00 -4.527 5.98e-06 ***
invo_quant:city臺北市 -1.229e+01  1.732e+00 -7.095 1.30e-12 ***
invo_quant:city臺南市 -3.645e+01  2.122e+00 -17.180 < 2e-16 ***
categorymaintain:quant -2.340e+01  6.745e+00 -3.469 0.000523 ***
categorymakeup:quant  9.891e+00  8.183e+00  1.209 0.226758
Freq:If_Discount -2.723e+00  5.788e-01 -4.705 2.55e-06 ***
invo_quant:If_Discount -1.412e+01  1.024e+00 -13.791 < 2e-16 ***
whether_F:If_Discount  3.008e+00  4.664e+00  0.645 0.518937
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 227.9 on 84907 degrees of freedom
Multiple R-squared:  0.705,    Adjusted R-squared:  0.7048
F-statistic: 4228 on 48 and 84907 DF,  p-value: < 2.2e-16

```

(3) Linear Model 3:

Response Variable: 客單價

Explanatory Variables: 購買頻率 + 是否是高頻率顧客 + 單次購買量*星期幾 + 美妝三個種類*uniprice + 折扣種類*totprice + 六都各縣市*單次購買量 + 商品購買量*美妝三個種類 + 有無折扣*購買頻率 + 有無折扣*單次購買量

```
lm<-lm(cos_invo_price~ Freq+Whether_F+invo_quant*day + category*u
niprice + discount_type*totprice + city*invo_quant + quant*category + If_
```

Discount*Freq + If_Discount*invo_quant,data=cosmed_F

```
call:
lm(formula = cos_invo_price ~ Freq + whether_F + invo_quant *
  day + category * uniprice + discount_type * totprice + city *
  invo_quant + quant * category + If_Discount * Freq + If_Discount *
  invo_quant, data = cosmed_F)

Residuals:
    Min       1Q   Median       3Q      Max
-2996.53   -71.52   -43.79    25.25   2694.47

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.682e+01  9.676e+00   1.738 0.082161 .
Freq          -2.693e-01  1.125e-01  -2.394 0.016653 *
whether_F        9.935e+00  1.673e+00   5.940 2.86e-09 ***
invo_quant     1.482e+02  1.874e+00  79.095 < 2e-16 ***

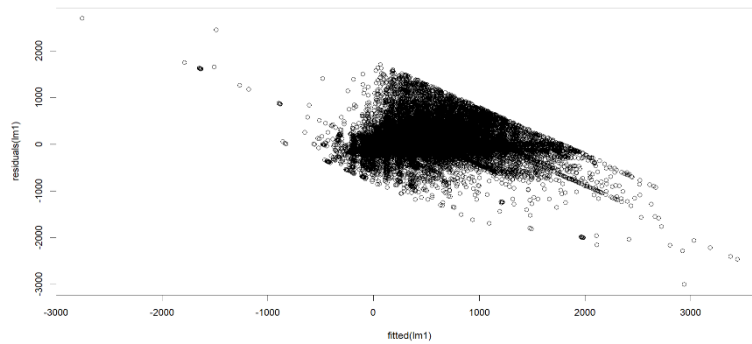
day1           5.538e+00  4.297e+00   1.289 0.197507
day2           1.893e+00  4.300e+00   0.440 0.659867
day3           8.184e+00  4.447e+00   1.840 0.065702 .
day4          -4.718e+00  4.543e+00  -1.039 0.298960
day5          -1.439e+01  4.173e+00  -3.448 0.000565 ***
day6           1.706e+01  4.092e+00   4.168 3.07e-05 ***
categorymaintain 2.958e+01  8.285e+00   3.570 0.000357 ***
categorymakeup  -2.206e+01  1.036e+01  -2.129 0.033228 *
uniprice        2.343e-01  1.557e-02  15.052 < 2e-16 ***
discount_type加價購 6.425e+00  2.276e+01   0.282 0.777747
discount_type期限內品項82折 3.096e+02  7.787e+00  39.755 < 2e-16 ***
discount_type買1送1 2.932e+02  1.015e+01  28.901 < 2e-16 ***
discount_type滿388現折50元 4.409e+02  2.806e+01  15.714 < 2e-16 ***
discount_type滿888現折100元 8.719e+02  2.489e+01  35.032 < 2e-16 ***
totprice       9.147e-01  8.446e-03 108.289 < 2e-16 ***
city高雄市     1.376e+01  6.166e+00   2.231 0.025681 *
city新北市     1.023e+01  5.082e+00   2.013 0.044109 *
city臺中市     1.303e+01  5.697e+00   2.288 0.022138 *
city臺北市     1.385e+01  4.964e+00   2.791 0.005256 **
city臺南市     6.474e+01  6.517e+00   9.933 < 2e-16 ***
quant          -1.416e+02  6.856e+00 -20.651 < 2e-16 ***
If_Discount    -3.319e+02  4.996e+00 -66.433 < 2e-16 ***
invo_quant:day1 -2.792e+00  1.426e+00  -1.959 0.050158 .
invo_quant:day2 -2.023e+00  1.408e+00  -1.437 0.150804
invo_quant:day3 -3.407e+00  1.462e+00  -2.330 0.019821 *
invo_quant:day4 5.050e+00  1.510e+00   3.344 0.000826 ***
invo_quant:day5 1.037e+01  1.379e+00   7.518 5.63e-14 ***
invo_quant:day6 -5.752e+00  1.328e+00  -4.332 1.48e-05 ***
categorymaintain:uniprice -8.658e-02  1.317e-02  -6.572 4.98e-11 ***
categorymakeup:uniprice 5.860e-02  2.283e-02   2.566 0.010281 *
discount_type加價購:totprice -9.408e-01  1.039e-01  -9.056 < 2e-16 ***
discount_type期限內品項82折:totprice -1.055e+00  3.697e-02 -28.538 < 2e-16 ***
discount_type買1送1:totprice -5.469e-01  2.286e-02 -23.926 < 2e-16 ***
discount_type滿388現折50元:totprice 5.317e-01  3.552e-01   1.497 0.134437
discount_type滿888現折100元:totprice 5.945e-01  1.256e-01   4.734 2.21e-06 ***
invo_quant:city高雄市 -6.667e+00  2.089e+00  -3.191 0.001417 **
invo_quant:city新北市 -9.519e+00  1.764e+00  -5.397 6.78e-08 ***
invo_quant:city臺中市 -8.844e+00  1.958e+00  -4.518 6.26e-06 ***
invo_quant:city臺北市 -1.228e+01  1.732e+00  -7.090 1.35e-12 ***
invo_quant:city臺南市 -3.642e+01  2.121e+00 -17.170 < 2e-16 ***
categorymaintain:quant -2.340e+01  6.745e+00  -3.469 0.000523 ***
categorymakeup:quant 9.935e+00  8.183e+00   1.214 0.224720
Freq:If_Discount -2.612e+00  5.528e-01  -4.726 2.29e-06 ***
invo_quant:If_Discount -1.412e+01  1.024e+00 -13.785 < 2e-16 ***

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

Residual standard error: 227.9 on 84908 degrees of freedom
Multiple R-squared:  0.705,    Adjusted R-squared:  0.7049
F-statistic: 4318 on 47 and 84908 DF,  p-value: < 2.2e-16
```

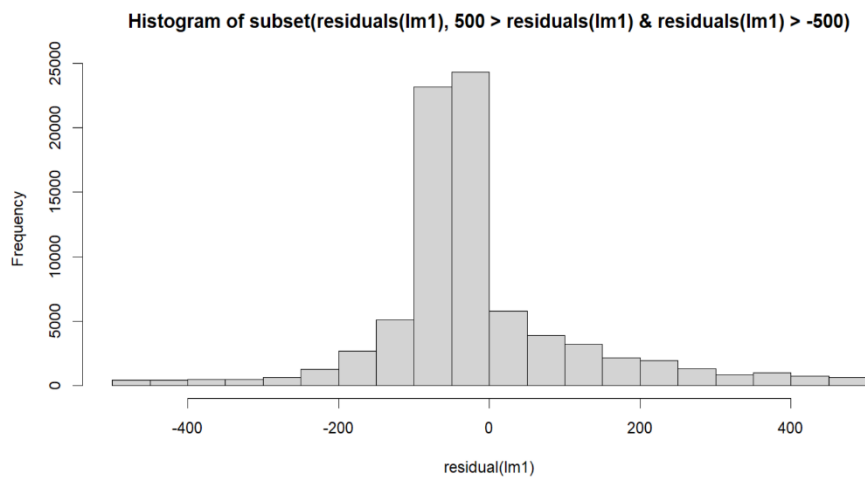
(4) 對Model3進行3+1個假說檢驗

◆ residual plot



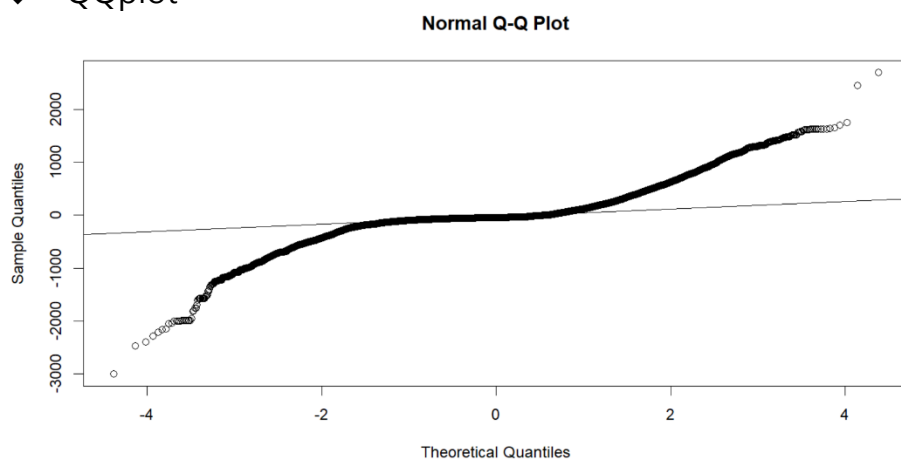
從Linear Model 3的residual plot可以看出資料並沒有均勻散佈。

◆ histogram



由Linear Model 3的histogram發現，資料為正偏分佈，也就是落在左側低分區的人數較多。

◆ QQplot



從Linear Model3的QQplot可以發現資料點往左側內凹，因此資料為正偏分佈。雖然並不貼和虛擬的常態線，但還在可以接受的範圍中。

五、問題探討

Q1: 是否為高頻率顧客會不會影響客單價？

為了得知高頻率顧客是否會影響客單價，我們做了lm3與lm1來比較。

■ anova(lm3,lm1)

Analysis of Variance Table

```
Model 1: cos_invo_price ~ category * uniprice + discount_type * totprice +  
city * invo_quant + quant * category + If_Discount * Freq +  
If_Discount * invo_quant
```

```
Model 2: cos_invo_price ~ Freq + Whether_F + invo_quant * day + category *  
uniprice + discount_type * totprice + city * invo_quant +  
quant * category + If_Discount * Freq + If_Discount * invo_quant
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	84921	442288	522			
2	84908	441036	3173	13	12525348	18.549 < 2.2e-16 ***

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

由ANOVA可得知lm3較lm1佳，此外，也可發現：加入「購買頻率」、「是否為高頻率顧客」以及「單次購買量與星期幾」的interaction能有效提升模型的解釋力。

從lm3中whether_F的係數與p-value可以得知，高頻率顧客對客單價有顯著正向影響，顯示高頻率顧客購買的客單價較非高頻率顧客高。

Q2: 有無折扣是否影響成為高頻顧客的意願

■ anova(lm3,lm2)

Analysis of Variance Table

```
Model 1: cos_invo_price ~ Freq + Whether_F + invo_quant * day + category *  
uniprice + discount_type * totprice + city * invo_quant +  
quant * category + If_Discount * Freq + If_Discount * invo_quant
```

```
Model 2: cos_invo_price ~ Freq + Whether_F + invo_quant * day + category *  
uniprice + discount_type * totprice + city * invo_quant +  
quant * category + If_Discount * Freq + If_Discount * invo_quant +  
whether_F * If_Discount
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	84908	441062	9271			
2	84907	441060	7661	1	21610	0.416 0.5189

由ANOVA可得知lm3較lm2佳。

為了得知有無折扣和高頻率顧客關係，我們做了lm2和lm3做比較，而這兩個model之間唯一的差別在於lm2多了有無折扣和是否為高頻率顧客的interaction，首先從lm2中這有無折扣*是否為高頻率顧客項係數的p value為0.5189左右(也就是兩個模型做anova的結果)，相當不顯著。接著看這兩個model的adjusted R Square，lm2甚至略低於lm3，因此我們選定lm3作為我們的最佳模型，並從lm3可以看出折扣會影響客單價，但不會影響高頻顧客的人數。

Q3: 將折扣種類區分成六種類型，分析折扣種類的顯著程度以探討對客單價的影響

從p值和estimate可看出加價購對於客單價的影響最不顯著，p value為0.777747和0.05相差甚遠，而estimate是e+00，比起其餘的折扣是e+02，對於客單價的幫助相對低很多。而其餘折扣皆有具顯著差異，p value都小於2e-16。

Q4: 了解星期幾的客單價最高，作為後續銷售策略參考？

由Linear Model 3可得知星期六(day6)客單價最高，且具顯著影響。而根據了解實際狀況，康是美確實有在禮拜六促銷，也證實了在星期六做促銷確實有提升客單價的效果。然而其他天數則較不顯著，因此康是美可考慮在其他天制定不同銷售策略，藉此提升整體客單價，增加競爭力。

Q5: 了解六都中哪個縣市對客單價的貢獻比較顯著，以納入分區銷售策略的考量？

六都的p-value皆小於0.05，對客單價的貢獻皆顯著的情況下，又以台南的客單價最高，而其他縣市幾乎沒有很大的差異，因此不太需要針對各地區做分區考量。不過康是美可考慮在台南地區舉辦促銷活動，以提升整體銷售表現。

Q6:不同產品種類（保養、彩妝、底妝）對客單價的影響是否不同？

從Linear Model 3的分析中可以發現，美妝三種類對於客單價的影響程度分別是保養品>底妝>化妝品。且化妝品與銷售數量的interaction並不顯著，反之，保養品與銷售數量的interaction的係數是正向的且呈顯著關係。

因此我們可以確定保養品、底妝、化妝品這三種產品種類對於客單價的影響是不同的。

六、銷售策略制定與結論

1. 由第一個問題我們發現，是否為高頻率顧客對客單價有顯著正向影響，表示高頻率顧客購買的客單價較非高頻率顧客高，然而，購買頻率對客單價有負向影響。因此推測可能是因為這三個月中只購買一次、但客單價高的顧客非常多，造成高頻率顧客購買的客單價比非高頻率顧客高、但購買頻率對客單價是負相關的狀況。
2. 由於自上述問題得知客單價和是否為忠實顧客有顯著正向關係，因此我們可以加強針對會員的活動，增加會員的購買數量和金額。且在模型中我們發現忠實顧客在星期三和星期六的銷售狀況較佳，可以訂定每周這兩天為會員促銷日。
3. 為了吸引更多人成為高頻率顧客，康是美可在一週中的每日推銷不同的策略方案（如每日皆為不同品牌促銷），進行市場區隔，在吸引不同喜好的顧客的同時也能提升整體銷售。

結論：

經模型證實高頻率顧客確實會提升客單價，因此康是美當前應透過不同的促銷方案鞏固既有高頻率顧客，並且吸引更多顧客成為高頻率顧客，藉此提升其業績，並能夠在同業間培養出足夠的競爭力。

七、 附錄

- 程式碼

```
library(readr)
```

```
library(dplyr)
```

```
NTU_Drug <- read_csv("NTU_Drug.csv")
```

#篩選美妝產品

```
filtered_drug <- NTU_Drug %>% filter(grepl('粧前乳|水凝露|蜜粉|唇釉|唇蜜|唇膏|隔離霜|防曬乳|粉餅|眼影盤|眼影筆|腮紅|修容筆|眉粉|化妝水|乳液|粉撲|抗老|凝露|凝凍|凝膠|打亮|BB霜|乳霜|面膜|遮瑕|抗痘|粉底|精華|口紅|護唇|保養|保濕|睫毛|眼妝|底妝|眼睫|眉筆|鼻影', name))
```

#篩選康是美

```
cosmed <- filtered_drug %>% filter(grepl("康是美",channel))
```

#忠實顧客

```
cosmed["Freq"] <- 1
```

```
loyal_customer <- cosmed %>% group_by(deviceid, invo_idx) %>%
```

```
summarise() %>%
```

```
group_by(deviceid) %>%
```

```
summarise(count = n()) %>%
```

```
filter(count > 1)
```

```
cosmed_F <- cosmed[cosmed$deviceid %in% loyal_customer$deviceid, ]
```

Add frequency to cosmed

```
Freq <- replicate(nrow(cosmed), 1)
```

```
for (i in 1:nrow(loyal_customer)){
```

```

  Freq[which(cosmed[, "deviceid"] == loyal_customer$deviceid[i])] <- loyal_customer[i, "count"]
}

```

```

for (i in 1:nrow(cosmed)){
  cosmed[i, "Freq"] <- Freq[[i]]
}

```

```

# Add "whether F" to cosmed
Whether_F <- replicate(nrow(cosmed), 0)
for (i in 1:nrow(loyal_customer)){
  Whether_F[which(cosmed[, "deviceid"] == loyal_customer$deviceid[i])] <- 1
}

```

```

for (i in 1:nrow(cosmed)){
  cosmed[i, "Whether_F"] <- Whether_F[[i]]
}

```

#週幾

```

day <- as.data.frame(format(cosmed$datetime.UTC_8,format = "%w"))
names(day)[1] <- "day"
cosmed["day"] <- day

```

#早中晚

```

library(stringr)
subtime<-as.data.frame(cosmed$datetime.UTC_8)
pos1<-str_locate(subtime[[1]],"2021-")
time_table<-data.frame(time=str_sub(subtime[[1]],pos1[,2]+7,pos1[,2]+8))
time_table$time <- gsub("06|07|08|09|10|11", "morning", time_table$time)

```



```
time_table$time <- gsub("12|13|14|15|16|17", "afternoon", time_table$time)
time_table$time <- gsub("18|19|20|21|22|23|00|01|02|03|04|05", "night", time_table$time)
```

```
cosmed["time"] <- time_table$time
```

品項分類

```
category <- cosmed$name
category[grepl("水凝露|化妝水|乳液|抗老|凝露|凝凍|凝膠|乳霜|面膜|抗痘|精華|保養|保濕",cosmed$name)]<-"maintain"
category[grepl("唇膏|唇釉|唇蜜|口紅|護唇|眉筆|睫毛|眼睫|眼影盤|眼影筆|眼妝|眉粉|修容筆|腮紅|打亮|腮紅|打亮|鼻影",cosmed$name)]<-"makeup"
category[grepl("粉餅|蜜粉|防曬乳|粧前乳|隔離霜|BB霜|底妝|遮瑕|粉底|粉撲",cosmed$name)]<-"base"
category<-as.data.frame(category)
```

```
cosmed["category"] <- category$category
```

折扣分類

```
discount <- replicate(nrow(cosmed), "一般商品")
discount[grepl("一般折扣",cosmed$name)]<-"一般折扣"
discount[grepl("買1送1",cosmed$name)]<-"買1送1"
discount[grepl("加碼-指定面膜", cosmed$name)]<-"加價購"
discount[grepl("開架臉部保養滿888現折100元",cosmed$name)]<-"滿888現折100元"
discount[grepl("開架身體保養滿388現折50元",cosmed$name)]<-"滿388現折50元"
discount[grepl("0224-0226面膜.醫美.彩妝.統藥82折", cosmed$name)]<-"期限內品項82折"
discount<-as.data.frame(discount)
```

```
cosmed["discount_type"] <- discount$discount
```

```
# 是否為折扣
```

```
for (i in 1:nrow(cosmed)){  
  if (cosmed[i, "discount_type"] == "一般商品"){  
    cosmed[i, "Whether_Discount"] = 0  
  }  
  else{  
    cosmed[i, "Whether_Discount"] = 1  
  }  
}
```

```
# 計算cosmetic方面的invo_price
```

```
cos_invo_price <- cosmed %>% group_by(invo_idx) %>%  
  summarise(cos_invo_price = sum(totprice))
```

```
for (i in 1:nrow(cos_invo_price)){  
  cosmed[which(cosmed[, "invo_idx"] == cos_invo_price$invo_idx[i]), "cos_invo_price"] <- cos_invo_price[i, "cos_invo_price"]}
```

```
#城市
```

```
sub<-as.data.frame(cosmed$county_district)  
pos1<-str_locate(sub[[1]],"市|縣")  
city <- data.frame(city=str_sub(sub[[1]],pos1[,1]-2,pos1[,1]))  
cosmed["city"] <- city$city  
cosmed <- cosmed %>% filter(grepl("臺北市|新北市|桃園市|臺中市|臺南市|高雄市",  
city))
```

#是否有折扣

```
discount_invo_idx <- subset(cosmed, Whether_Discount == 1)
for (i in 1:nrow(discount_invo_idx)){
  cosmed[cosmed$invo_idx == discount_invo_idx$invo_idx[i], "If_Discount"] <- 1
}
cosmed[is.na(cosmed$If_Discount), "If_Discount"] = 0
```

#單次購買量

```
tot_quant <- cosmed %>% group_by(invo_idx) %>%
summarise(tot_quant = sum(quant))

for (i in 1:nrow(tot_quant)){
  cosmed[cosmed$invo_idx == tot_quant$invo_idx[i], "invo_quant"] <- tot_quant
  $tot_quant[i]
```

#刪outlier

```
Q <- quantile(cosmed$cos_invo_price, probs=c(.25, .75), na.rm = FALSE)
iqr <- IQR(cosmed$cos_invo_price)
up <- Q[2]+1.5*iqr # Upper Range
low<- Q[1]-1.5*iqr # Lower Range

cosmed_F <- na.omit(subset(cosmed, cosmed$cos_invo_price > (Q[1] - 1.5*iqr)
& cosmed$cos_invo_price < (Q[2]+1.5*iqr)))
```

#Linear Model

```
lm1<-lm(cos_invo_price~ category*uniprice + discount_type*totprice + city*invo_quant + quant*category + If_Discount*Freq + If_Discount*invo_quant ,data=cosmed_F)
```

```
lm2<-lm(cos_invo_price~ Freq+Whether_F+invo_quant*day + category*uniprice + discount_type*totprice + city*invo_quant + quant*category + If_Discount*Freq + If_Discount*invo_quant + Whether_F*If_Discount ,data=cosmed_F)
```

```
lm3<-lm(cos_invo_price~ Freq+Whether_F+invo_quant*day + category*uniprice + discount_type*totprice + city*invo_quant + quant*category + If_Discount*Freq + If_Discount*invo_quant ,data=cosmed_F)
```

ANOVA

```
anova(lm3,lm1)
```

residual plot

```
plot(fitted(lm3), residuals(lm3))
```

```
hist(x = subset(residuals(lm3), 500 > residuals(lm3) & residuals(lm3) > -500),xlab="residual(lm3)")
```

QQplot

```
qqnorm(residuals(lm3))
```

```
qqline(residuals(lm3))
```

EDA

```
boxplot(cosmed$cos_invo_price ~ factor(cosmed$Whether_F)+factor(cosmed$day),xlab = "高頻率與否&消費日",ylab="客單價",ylim=c(-1000,15000))
```

```
boxplot(cosmed$cos_invo_price ~ factor(cosmed$Whether_F)+factor(cosmed$category),xlab = "高頻率與否&消費品項",ylab="客單價",ylim=c(-1000,15000))
```

```
boxplot(cosmed$cos_invo_price ~ factor(cosmed$Whether_F)+factor(cosmed$city),xlab = "高頻率與否&消費地點",ylab="客單價",ylim=c(-1000,15000))
```

```
boxplot(cosmed$cos_invo_price ~ factor(cosmed$Whether_F)+factor(cosmed$time),xlab = "高頻率與否&消費時間",ylab="客單價",ylim=c(-1000,15000))
```

apple電腦顯示中文

```
install.packages("showtext")
```

```
library(showtext)
```

Apple電腦沒有中文要執行下面這四行

```
# font_files() #顯示系統裡擁有的字體檔
```

```
# showtext_auto(enable = TRUE)
```

```
# font_add('Songti', 'Songti.ttc') #選擇要用的字體檔
```

```
# quartz(family='Songti') #接下來便可直接畫圖
```