

進階軟體期末報告

王志明、陳祈廷

ETP

Experience

- OrderGroupCode – 購物車編號
- MemberId – 會員編號
- Quantity – 購買數量
- UnitPrice - 商品單價
- PromotionDiscount – 折扣金額
- PromotionCondition_TotalPrice – 折扣條件
- SalesOrderSlaveDateTime – 下單時間
- PromotionStartDateTime – 優惠活動開始時間
- PromotionEndDateTime – 優惠活動截止時間
- RegisterDateTime – 會員註冊時間
- Birthday – 會員生日
- DiscountType – 優惠方式

- Age – 顧客年齡 (訂單時間-生日)
- MemberTime - 會員時間(訂單時間-註冊時間)

Task

- 研究「滿額折現」此種優惠方式中, 給何種折扣金額可以得到最高的利潤。

Performance

- 透過「假設商品的成本」計算在「滿額折現」此種優惠方式中, 在不同折扣金額的情況下獲得的利潤。利潤最高者即為當然最優選項。

折扣金額：滿足折扣條件的情況下，給予的折現金額
折扣條件：給予優惠的消費門檻

Task

- ▶ 取出Discount Type為「滿額折現」的訂單
- ▶ 設定UnitPrice的7成為利潤，減去折扣金額為實得利潤
- ▶ 將相同折扣金額之實得利潤加總，為該種折扣金額之優惠活動期間所獲得之總利潤
- ▶ 將該總利潤除以該折扣金額之活動之舉行時長，得該折扣金額之每日平均利潤
- ▶ 以折扣金額為橫軸、每日平均利潤為縱軸作圖比較

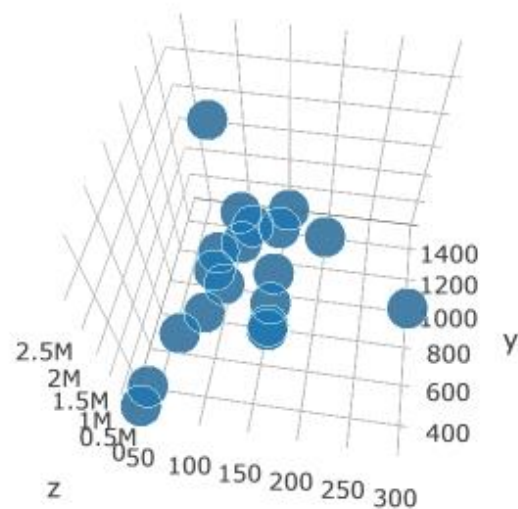
Task



- 由此圖可看出將Discount Price設為100元，可得到最高的平均日利潤

Task

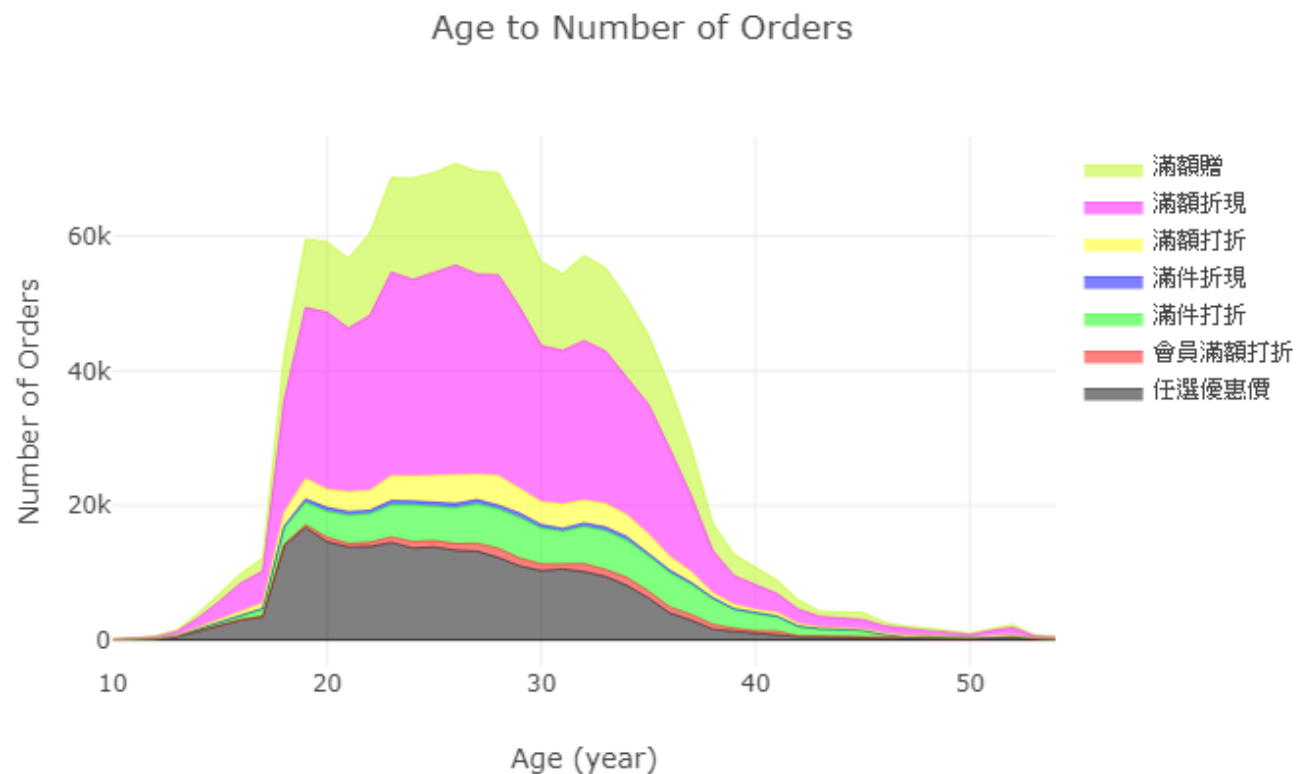
Discount Price and Discount condition versus Daily Profit



進 Jupyter 看

從此圖可發現，折扣金額和折扣條件在特定比率的情況下可以得到座高的日平均收益

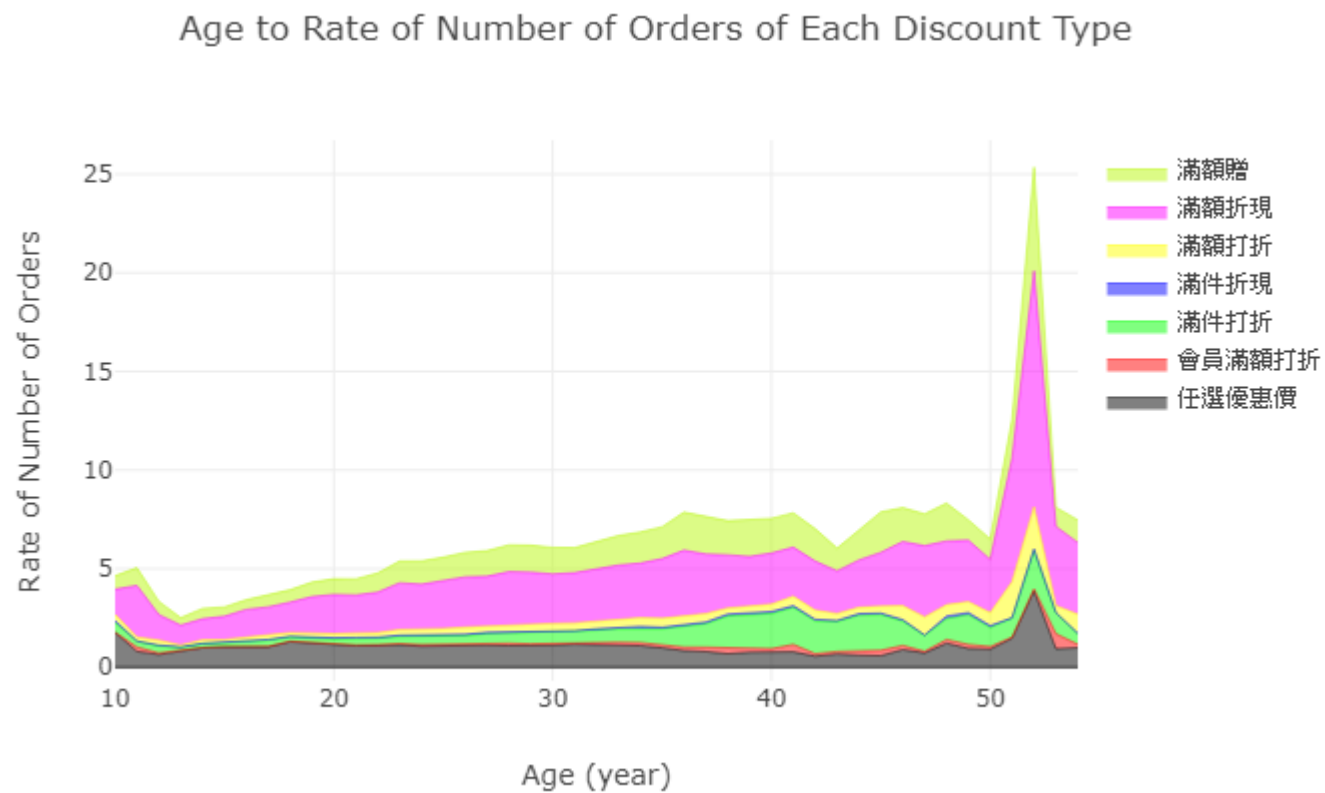
EDA過程



呈現不同年齡層使用不同優惠方式購買商品的訂單數量。

從這張圖發現，使用「滿額折現」的訂單數量最多，最後針對「滿額折現」此種優惠方式進行研究。

EDA過程



另外也想從年齡之間的差異找出值得研究的地方，所以為了減少變因，將縱軸改為計算該種Discount Type訂單占總訂單的比率，另外因為擔心不同Discount Type活動推出的數量不同，也將此因素除去，但最後沒有從這張圖找到可以研究的方向。

EDA過程

- ▶ 取出「滿額折現」此種Discount Type作為研究對象後，我們兵分兩路進行研究。
 1. 利用折扣金額在單價中的比率在做行銷成本的分級。在不同行銷成本中，針對不同的商品單價、年齡、會員時間、上次購買時間作為橫軸的情況下，觀察訂單數量的分布有無特定模式。
 2. 實驗從某折扣金額跨越到另一折扣金額時，多吸引了多少人進行消費，在哪一個區段的效益最高

EDA過程

▶ 並先預設一個ROI來源

1. 同樣的折扣金額，把不恰當的Discount type改成恰當的Discount type，可以吸引更多人
2. 對比區間跨越時，兩區間的Unit price減去Discount value再乘以訂單數量，可獲得總營利

EDA過程

▶ 第一路:

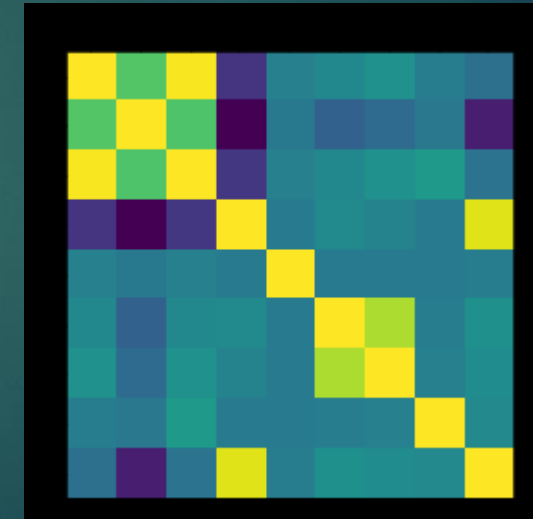
- ▶ 前提:先把被折扣卷吸引而增加購買量的訂單提出。
- ▶ 提出條件：購物車總金額與折扣卷的條件相近

$$ratio = \frac{SlaveTotalPrice - PromotionConditionTotalPrice}{PromotionConditionTotalPrice}$$

- ▶ 假設 $ratio \leq 0.1$ 時，是被優惠卷吸引而進行購買的訂單
- ▶ 為觀察受優惠吸引顧客及未受優惠吸引顧客的變數之間的相關性，製作相關係數表，並畫圖進行觀察比較。

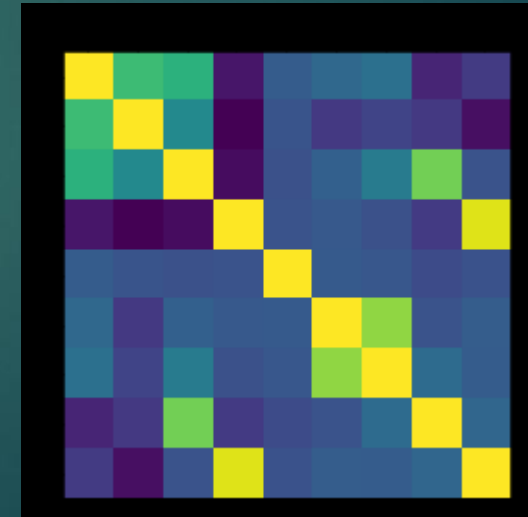
Correlation Coefficient (Attracted)

	Promotion Condition _TotalPric e	Promotion Condition _Discount Price	Price	Promotion Discount	age	memberTi me	delta_day	ratio	discount_ rate
Promotion Condition _TotalPric e	1.000000	0.542414	0.980886	-0.446472	0.035531	0.084669	0.151820	0.015815	-0.081279
Promotion Condition _Discount Price	0.542414	1.000000	0.524931	-0.713585	-0.018401	-0.184299	-0.112319	-0.030431	-0.568505
Price	0.980886	0.524931	1.000000	-0.437425	0.034339	0.084899	0.155216	0.208262	-0.062648
Promotion Discount	-0.446472	-0.713585	-0.437425	1.000000	-0.003581	0.102673	0.050630	-0.000676	0.915841
age	0.035531	-0.018401	0.034339	-0.003581	1.000000	-0.011380	-0.005621	-0.002188	0.007736
memberTi me	0.084669	-0.184299	0.084899	0.102673	-0.011380	1.000000	0.786590	0.009391	0.147053
delta_day	0.151820	-0.112319	0.155216	0.050630	-0.005621	0.786590	1.000000	0.034383	0.120194
ratio	0.015815	-0.030431	0.208262	-0.000676	-0.002188	0.009391	0.034383	1.000000	0.091615
discount_ rate	-0.081279	-0.568505	-0.062648	0.915841	0.007736	0.147053	0.120194	0.091615	1.000000



Correlation Coefficient (Unattracted)

	Promotion Condition _TotalPric e	Promotion Condition _Discount Price	Price	Promotion Discount	age	memberTi me	delta_day	ratio	discount_r ate
Promotion Condition _TotalPric e	1.000000	0.570965	0.509873	-0.282475	0.031817	0.096360	0.140300	-0.223748	-0.127537
Promotion Condition _Discount Price	0.570965	1.000000	0.277832	-0.365244	-0.011446	-0.138906	-0.085851	-0.136669	-0.306891
Price	0.509873	0.277832	1.000000	-0.320556	-0.028142	0.053211	0.200365	0.709063	-0.014404
Promotion Discount	-0.282475	-0.365244	-0.320556	1.000000	-0.013357	0.016472	-0.027365	-0.132729	0.932466
age	0.031817	-0.011446	-0.028142	-0.013357	1.000000	0.019280	0.003332	-0.054880	-0.021227
memberTi me	0.096360	-0.138906	0.053211	0.016472	0.019280	1.000000	0.771889	-0.014095	0.036202
delta_day	0.140300	-0.085851	0.200365	-0.027365	0.003332	0.771889	1.000000	0.110516	0.034391
ratio	-0.223748	-0.136669	0.709063	-0.132729	-0.054880	-0.014095	0.110516	1.000000	0.083902
discount_r ate	-0.127537	-0.306891	-0.014404	0.932466	-0.021227	0.036202	0.034391	0.083902	1.000000



EDA過程

- ▶ 在相關係數表中，受吸引與未受吸引顧客的變數間的相關係數之模態基本相同，沒有得到可以利用的點，因此深入針對被吸引到的群組的各變數做研究

- ▶ 利用Discount Price在Unit Price中的占比定義行銷成本大小(loss)

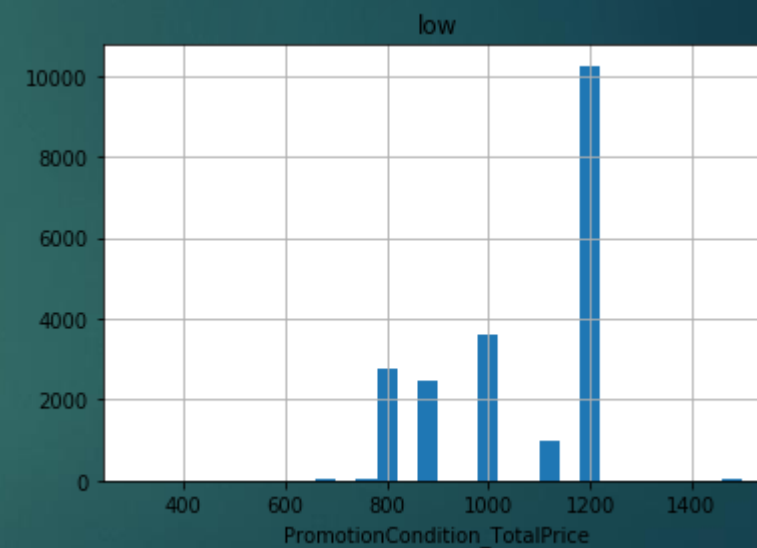
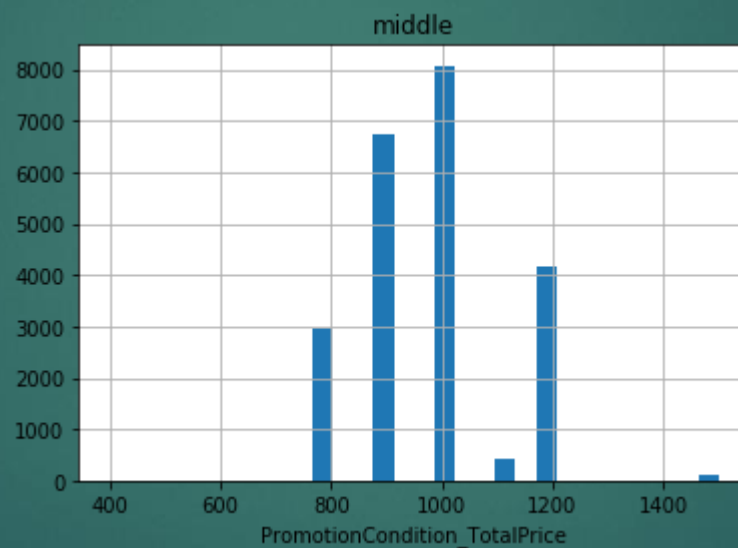
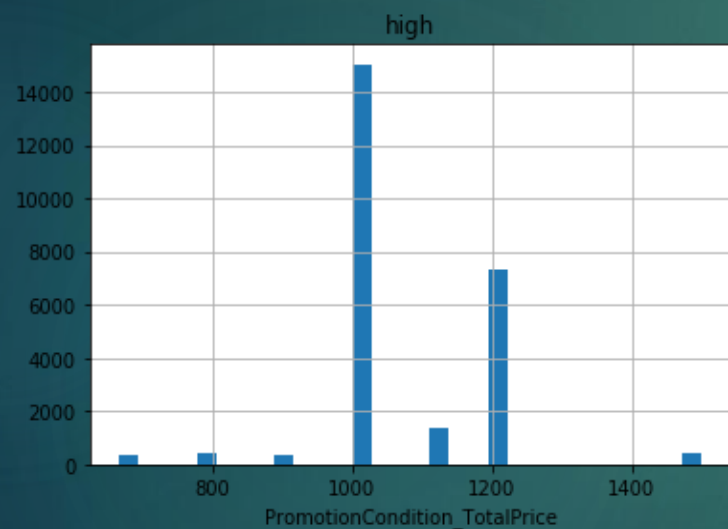
$$loss = \frac{PromotionDiscount}{SlaveTotalPrice}$$

即折扣金額與購物車總金額的，loss還小，成本遇低。

- ▶ 利用loss切分，畫圖觀察3個區域在不同的變數時，有沒有明顯的模態

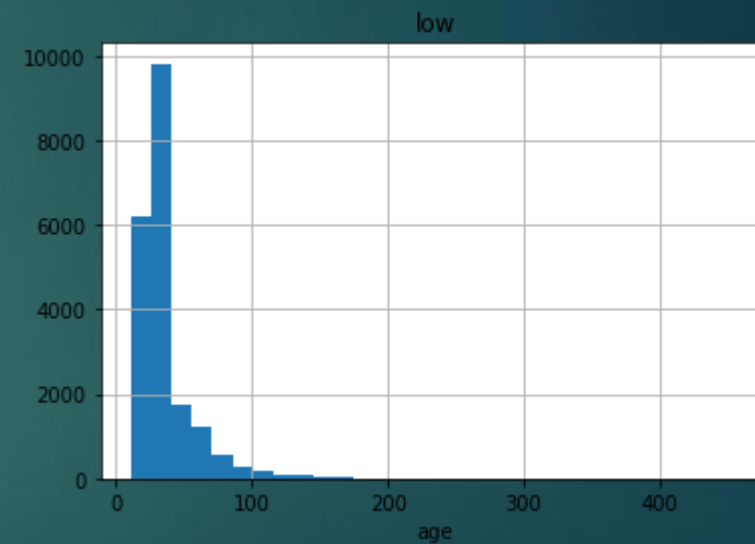
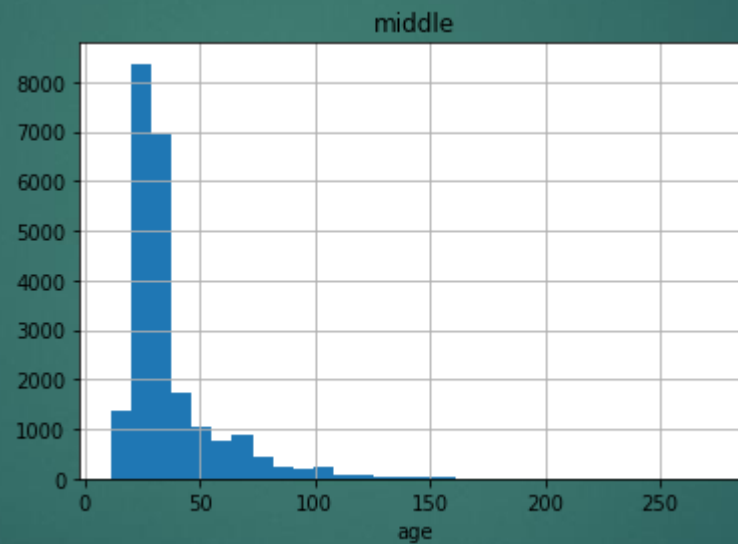
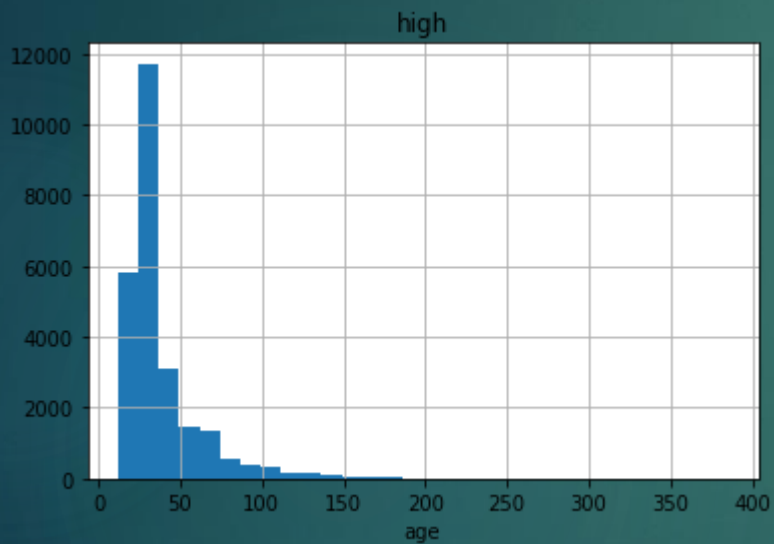
EDA過程

► 變數為 : Promotion_Condition_Total_Price



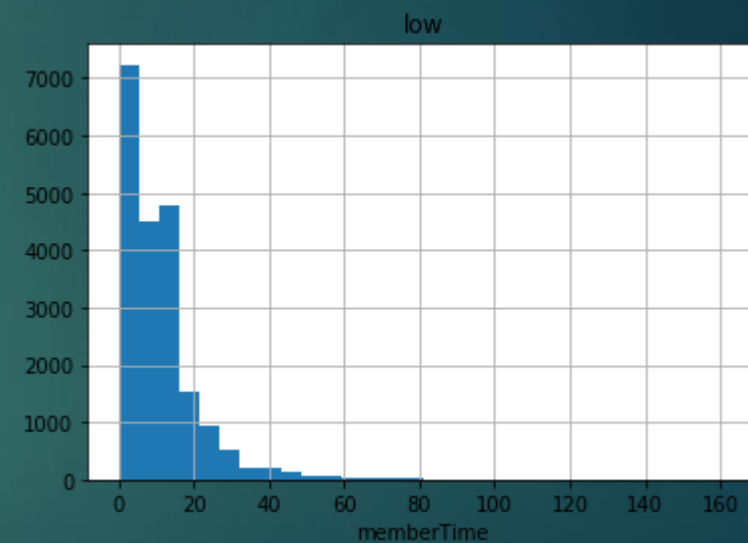
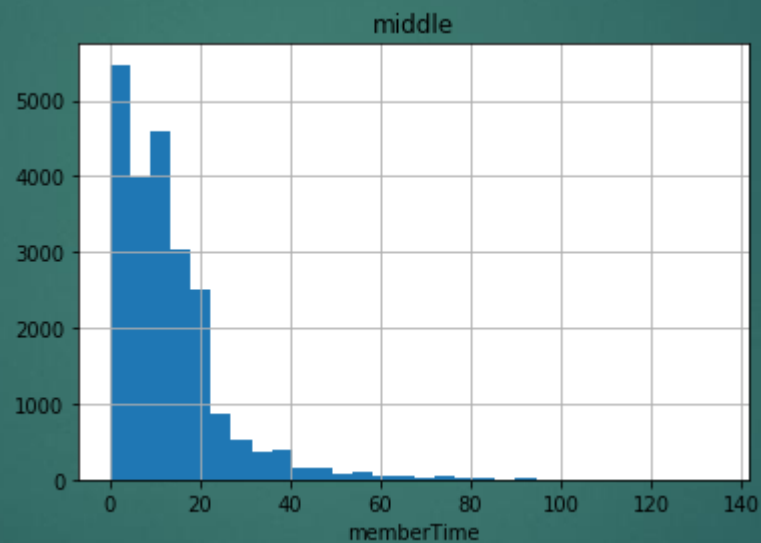
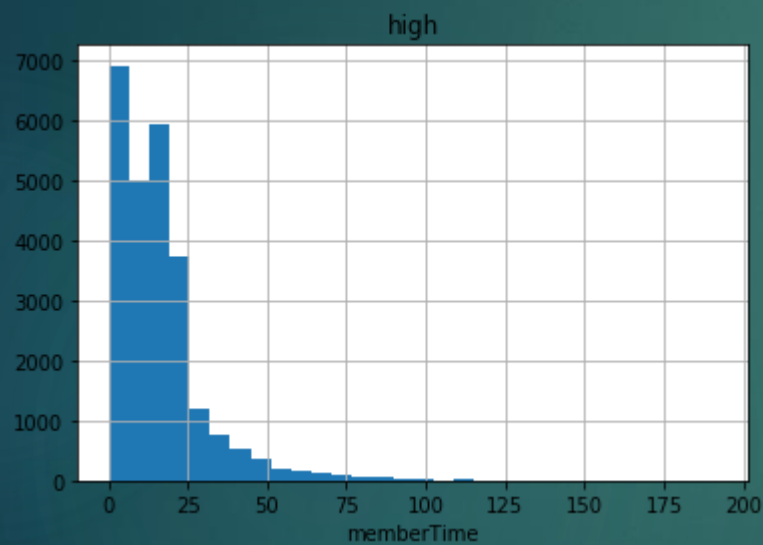
EDA過程

► 變數為 : age



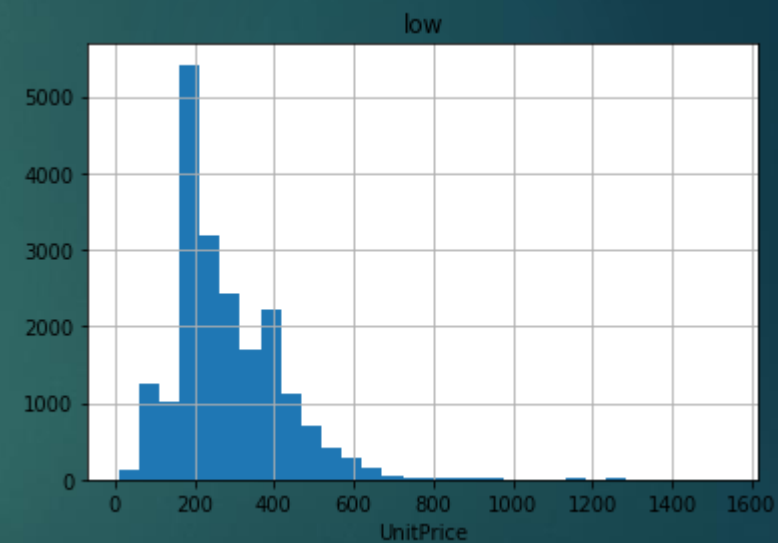
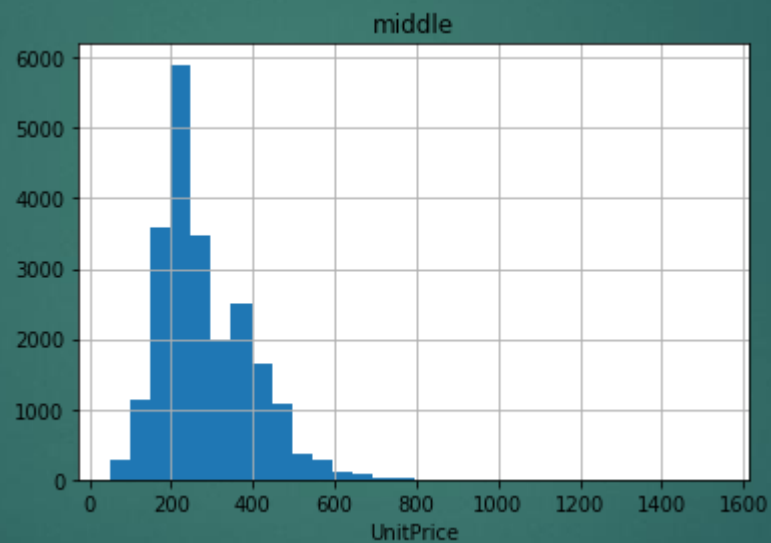
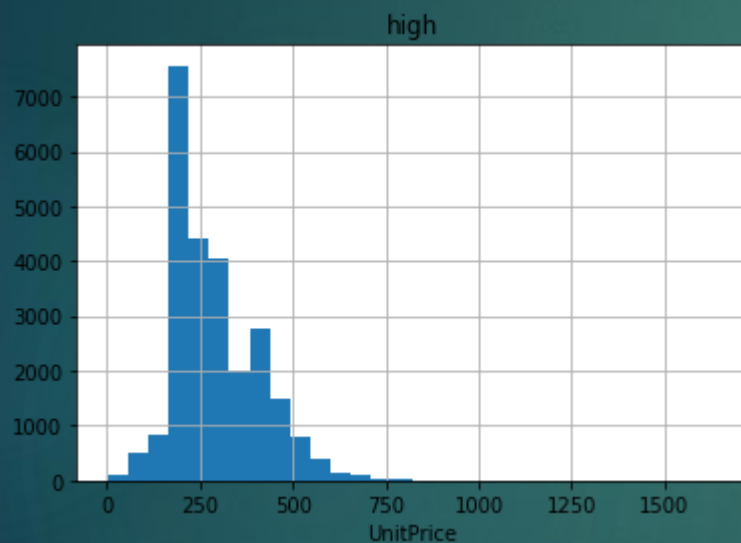
EDA過程

► 變數為 : memberTime



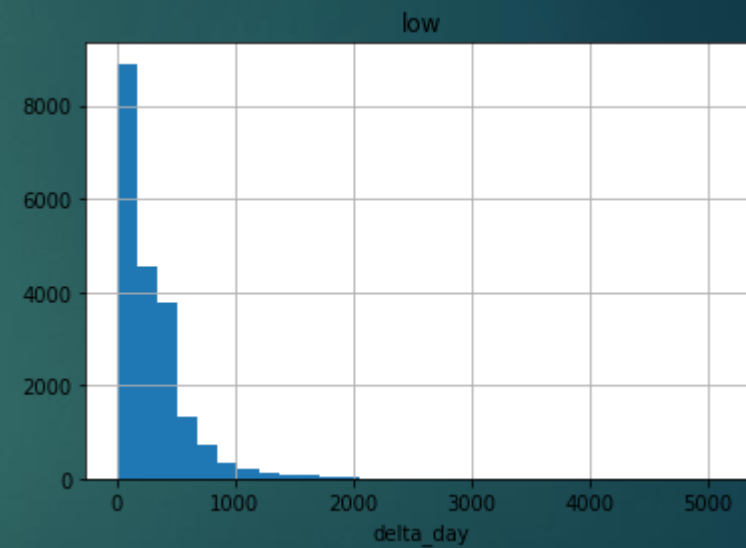
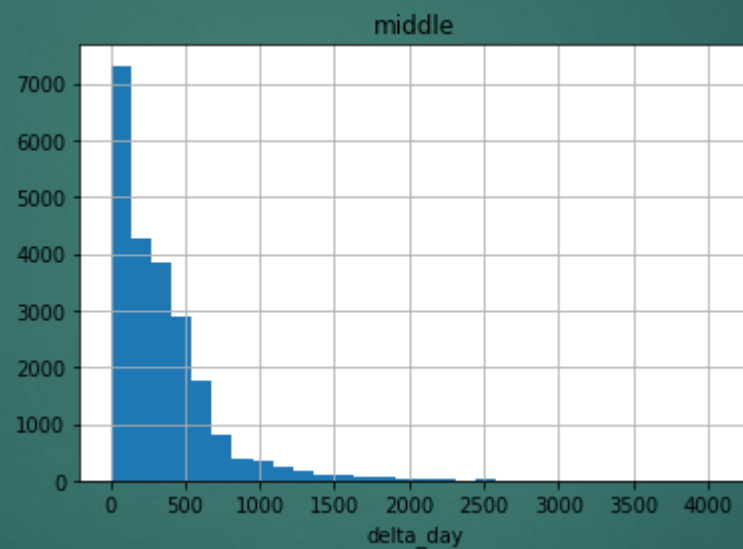
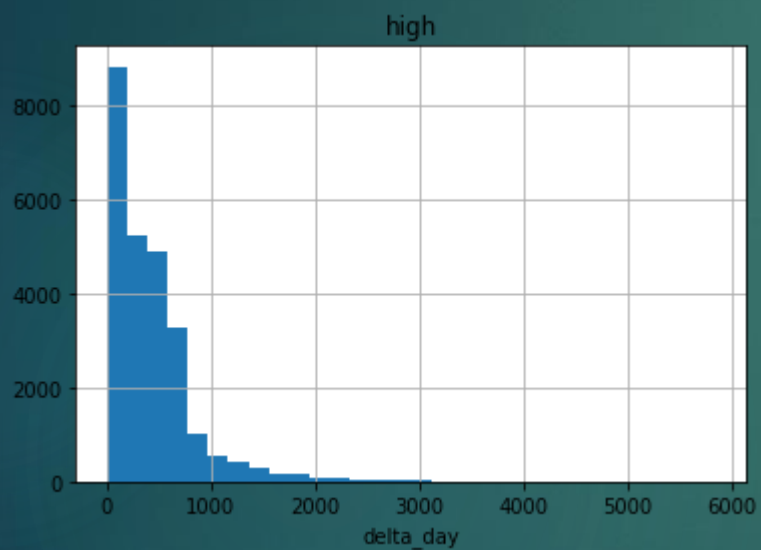
EDA過程

► 變數為 : UnitPrice



EDA過程

► 變數為 : delta_day



第一路結論

- ▶ 在各種的變數中，只有PromotionCondition_totalPrice有結構上的差異。
- ▶ 其他的變數在3個群的 histogram 的形狀結構都十分的相似，因此可以認為，在實驗中使用的變數與loss沒有相關性。

EDA過程

- ▶ 第二路：實驗從某折扣金額跨越到另一折扣金額時，多吸引了多少人進行消費，在哪一個區段的效益最高
- ▶ 為得到此結果須先計算不同折扣金額的折價券的使用量以進行觀察。

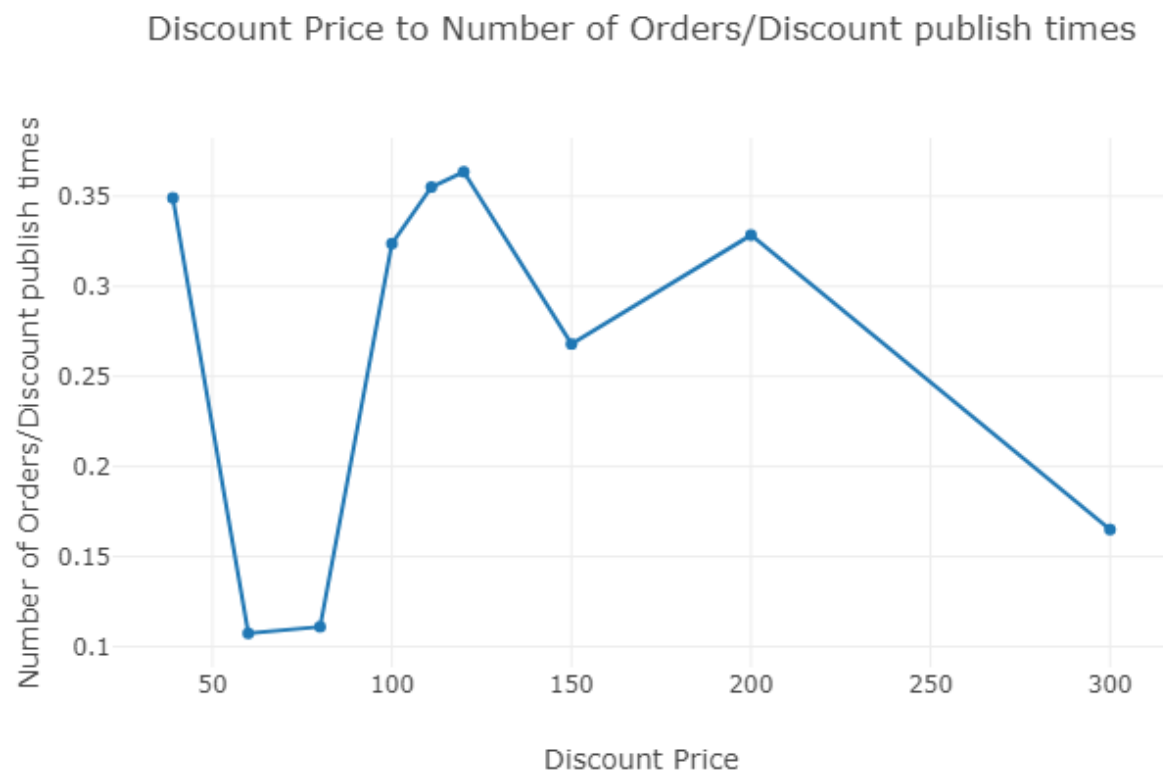
EDA過程



從這張圖觀察到在不同折扣金額中，折扣金額為100的折扣方式的使用量最高。

但思考後認為，不同折扣金額的優惠活動的發行次數會影響到使用另，因此應須計算使用率。

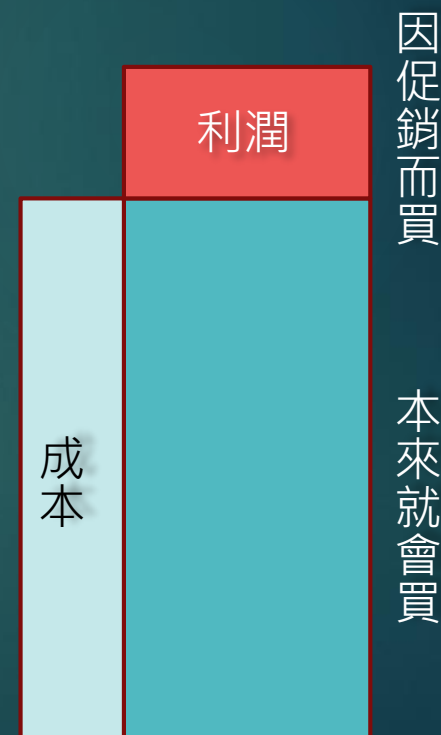
EDA過程



從這張圖觀察到在不同折扣金額中，折扣金額為100的折扣方式的使用率相當高。而更值得注意的是，折扣金額「從80到100」的使用率有相當大的改變，認為相當值得討論。

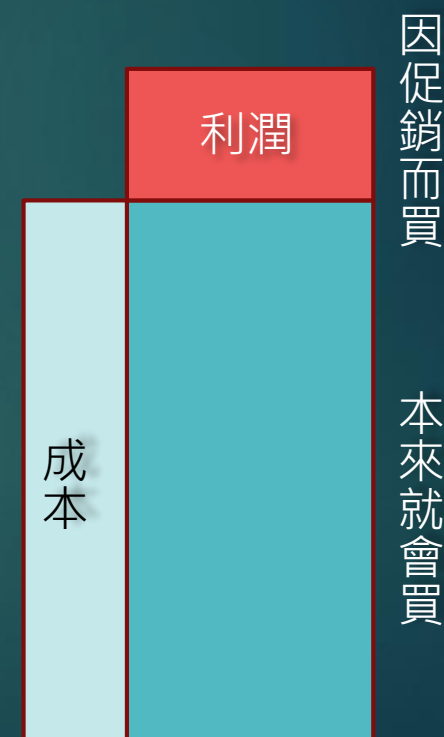
EDA過程

- ▶ 研究ROI計算:
- ▶ 區分沒有促銷也會購買和因為促銷而新增的購買
 - ▶ 購物車金額很接近 Discount conditions 的訂單
 - ▶ 取一段有 Discount 的時間段，和另一段同時間長度但沒有 Discount 的時間段的訂單數量進行比較(考慮季節、時點的影響)
- ▶ 此時成本為因為Discount從沒有促銷也會購買的訂單損失的金額，利潤為因為Discount新增的訂單產生的利潤



EDA過程

- ▶ 購物車金額很接近 Discount conditions 的訂單
 - ▶ 成功：用在第一路，但因為最後找不出模式，因此未計算ROI
- ▶ 取一段有 Discount 的時間段，和另一段同時間長度但沒有 Discount 的時間段的訂單數量進行比較(考慮季節、時點的影響)
 - ▶ 失敗：在Data的三年期間，只有前26天沒有任何優惠活動，因此我們猜測這一段時間可能有特殊意義，因此不能取作比較標準。



EDA過程

- ▶ 因為找不到ROI取用方法，因此回到一開始計算總利潤的方式比較何種Discount Price較為適合，接下來就是一開始介紹的我們採用的Task的研究方式了。

Task



- 由此圖可看出將Discount Price設為100元，可得到最高的平均日利潤