Data Mining 作業一 Association Rules Frequent Itemset
組員：
資科碩一 110753201 曹昱維
資科碩一 110753126 鄭詠儒
資科碩一 110753157 吳仁凱

1.
- 符合條件的 rule 總共有 2099 條，以下為其中 10 條

| | antecedents | consequents | support | confidence |
|---|---|---|---|---|
| 1 | frozenset({'CDR Hot Chocolate', 'Imagine Low Fat French Fries'}) | frozenset({'Quick Extra Lean Hamburger'}) | 0.0001849356688066... | 1.0 |
| 2 | frozenset({'Just Right Canned Yams', 'Hilltop 200 MG Acetominifen'}) | frozenset({'Faux Products HCL Nasal Spray'}) | 0.0001585162875485... | 1.0 |
| 3 | frozenset({'Bravo Fancy Canned Anchovies', 'Booker Low Fat String Cheese'}) | frozenset({'High Top Oranges'}) | 0.0001585162875485... | 1.0 |
| 4 | frozenset({'High Quality Scissors', 'Plato French Roast Coffee'}) | frozenset({'Dollar Monthly Sports Magazine'}) | 0.0001585162875485... | 1.0 |
| 5 | frozenset({'Cormorant Scented Toilet Tissue', 'Hilltop 200 MG Acetominifen'}) | frozenset({'Horatio No Salt Popcorn'}) | 0.0001585162875485... | 1.0 |
| 6 | frozenset({'High Top Summer Squash', 'Even Better Sharp Cheddar Cheese'}) | frozenset({'High Top New Potatos'}) | 0.0001585162875485... | 1.0 |
| 7 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |
| 8 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'High Top New Potatos'}) | 0.0001320969062904... | 1.0 |
| 9 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'High Top New Potatos', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |
| 10 | frozenset({'Just Right Canned Tuna in Water', 'High Top New Potatos'}) | frozenset({'Atomic Tasty Candy Bar', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |

2.
- 符合條件的 rule 總共有 2099 條，以下為其中 10 條

| | antecedents | consequents | support | confidence |
|---|---|---|---|---|
| 1 | frozenset({'CDR Hot Chocolate', 'Imagine Low Fat French Fries'}) | frozenset({'Quick Extra Lean Hamburger'}) | 0.0001849356688066... | 1.0 |
| 2 | frozenset({'Just Right Canned Yams', 'Hilltop 200 MG Acetominifen'}) | frozenset({'Faux Products HCL Nasal Spray'}) | 0.0001585162875485... | 1.0 |
| 3 | frozenset({'Bravo Fancy Canned Anchovies', 'Booker Low Fat String Cheese'}) | frozenset({'High Top Oranges'}) | 0.0001585162875485... | 1.0 |
| 4 | frozenset({'High Quality Scissors', 'Plato French Roast Coffee'}) | frozenset({'Dollar Monthly Sports Magazine'}) | 0.0001585162875485... | 1.0 |
| 5 | frozenset({'Cormorant Scented Toilet Tissue', 'Hilltop 200 MG Acetominifen'}) | frozenset({'Horatio No Salt Popcorn'}) | 0.0001585162875485... | 1.0 |
| 6 | frozenset({'High Top Summer Squash', 'Even Better Sharp Cheddar Cheese'}) | frozenset({'High Top New Potatos'}) | 0.0001585162875485... | 1.0 |
| 7 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |
| 8 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'High Top New Potatos'}) | 0.0001320969062904... | 1.0 |
| 9 | frozenset({'Just Right Canned Tuna in Water', 'High Top Summer Squash'}) | frozenset({'Atomic Tasty Candy Bar', 'High Top New Potatos', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |
| 10 | frozenset({'Just Right Canned Tuna in Water', 'High Top New Potatos'}) | frozenset({'Atomic Tasty Candy Bar', 'Even Better Sharp Cheddar Cheese'}) | 0.0001320969062904... | 1.0 |

3.
- 符合條件的 rule 總共有 100 條，以下為其中 10 條

| | antecedents | consequents | support | confidence |
|---|---|---|---|---|
| 1 | frozenset({'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.29792821709950396 | 1.0 |
| 2 | frozenset({'houseowner_Y', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.17829004960606945 | 1.0 |
| 3 | frozenset({'gender_M', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.1546542165159031 | 1.0 |
| 4 | frozenset({'gender_F', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.14327400058360082 | 1.0 |
| 5 | frozenset({'state_province_CA', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.1246960412411244 | 1.0 |
| 6 | frozenset({'houseowner_N', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.11963816749343449 | 1.0 |
| 7 | frozenset({'total_children_(-0.005, 1.0]', 'occupation_Professional'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.09561326719190741 | 1.0 |
| 8 | frozenset({'yearly_income_$30K - $50K', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.09502966637486626 | 1.0 |
| 9 | frozenset({'num_cars_owned_(1.0, 2.0]', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.09425153195214474 | 1.0 |
| 10 | frozenset({'gender_M', 'houseowner_Y', 'total_children_(-0.005, 1.0]'}) | frozenset({'num_children_at_home_(-0.005, 1.0]'}) | 0.09113899426125863 | 1.0 |

4.
- 將 frequent itemset 的長度限制在 2 以下，minimum confidence 為 0.001，限制 consequents 只包含 product 品項的情況下，找出 435 條 rule，以下為其中 10 條

| | antecedents | consequents | support | confidence |
|---|---|---|---|---|
| 1 | frozenset({'num_children_at_home_(2.0, 3.0]'}) | frozenset({'product_id_Hermanos Peaches'}) | 0.0001038915590842... | 0.0014819436861399... |
| 2 | frozenset({'num_children_at_home_(2.0, 3.0]'}) | frozenset({'product_id_Best Wheat Puffs'}) | 0.0001038915590842... | 0.0014819436861399... |
| 3 | frozenset({'state_province_Yucatan'}) | frozenset({'product_id_Better Noodle Soup'}) | 0.0001038915590842... | 0.0014599661902566... |
| 4 | frozenset({'state_province_Yucatan'}) | frozenset({'product_id_High Top New Potatos'}) | 0.0001038915590842... | 0.0014599661902566... |
| 5 | frozenset({'education_Partial College'}) | frozenset({'product_id_Nationeel Grape Fruit Roll'}) | 0.0001312314430537... | 0.0014058930349715... |
| 6 | frozenset({'state_province_Zacatecas'}) | frozenset({'product_id_Better Canned Tuna in Oil'}) | 0.0001859112109928... | 0.0013626708348362... |
| 7 | frozenset({'education_Partial College'}) | frozenset({'product_id_Bravo Fancy Canned Sardines'}) | 0.0001257634662598... | 0.00134731415851444 |
| 8 | frozenset({'education_Partial College'}) | frozenset({'product_id_Big Time Frozen Corn'}) | 0.0001257634662598... | 0.00134731415851444 |
| 9 | frozenset({'yearly_income_$70K - $90K'}) | frozenset({'product_id_Excellent Orange Juice'}) | 0.0001640393038171... | 0.0013335704125177... |
| 10 | frozenset({'num_children_at_home_(1.0, 2.0]'}) | frozenset({'product_id_Red Wing Copper Pot Scrubber'}) | 0.0001257634662598... | 0.0013232079162351... |

5.

- 將 frequent itemset 的長度限制在 2 以下，minimum confidence 為 0.001，限制 consequents 只包含 product 品項的情況下，12 月以前的 rule 有 551 條，12 月的 rule 有 11101 條
- 12 月以前與 12 月以後的相似 rule 有 129 條(具有相同的 antecedents 與 consequents 視為相似)，以下為其中 10 條

| | antecedents | consequents | support_before | confidence_before | | support_dec | confidence_dec |
|---|---|---|---|---|---|---|---|
| 1 | frozenset({'state_province_Yucatan'}) | frozenset({'product_id_Better Noodle Soup'}) | 0.00010330704... | 0.0014426340801086... | | 0.0001091405184174... | 0.0016260162601626... |
| 2 | frozenset({'state_province_Zacatecas'}) | frozenset({'product_id_Better Canned Tuna in Oil'}) | 0.00018230654... | 0.0013260840737302... | | 0.0002182810368349... | 0.0017182130584192... |
| 3 | frozenset({'education_Partial College'}) | frozenset({'product_id_Big Time Frozen Corn'}) | 0.00012153769... | 0.0013010668748373... | | 0.0001637107776261... | 0.0017657445556209... |
| 4 | frozenset({'education_Partial College'}) | frozenset({'product_id_Consolidated 200 MG Ibuprofen'}) | 0.00012153769... | 0.0013010668748373... | | 0.0001091405184174... | 0.0011771630370806... |
| 5 | frozenset({'num_children_at_home_(1.0, 2.0]'}) | frozenset({'product_id_Booker Low Fat Sour Cream'}) | 0.00012153769... | 0.0012787723785166... | | 0.0001091405184174... | 0.001148105625717566 |
| 6 | frozenset({'occupation_Management'}) | frozenset({'product_id_Hilltop Conditioning Shampoo'}) | 0.00018230654... | 0.0012695725772323... | | 0.0001637107776261... | 0.0011432926829268... |
| 7 | frozenset({'education_Partial College'}) | frozenset({'product_id_Fabulous Cranberry Juice'}) | 0.00011546081... | 0.0012360135310954... | | 0.0001091405184174... | 0.0011771630370806... |
| 8 | frozenset({'education_Partial College'}) | frozenset({'product_id_Big Time Chicken TV Dinner'}) | 0.00011546081... | 0.0012360135310954... | | 0.0001091405184174... | 0.0011771630370806... |
| 9 | frozenset({'education_Partial College'}) | frozenset({'product_id_Black Tie Eyeglass Screwdriver'}) | 0.00011546081... | 0.0012360135310954... | | 0.0001091405184174... | 0.0011771630370806... |
| 10 | frozenset({'education_Partial College'}) | frozenset({'product_id_Nationeel Fudge Brownies'}) | 0.00011546081... | 0.0012360135310954... | | 0.0001091405184174... | 0.0011771630370806... |

- 以相同的 antecedents 為條件聚合 consequents 的操作來整理資料。藉此觀察 12 月以前與 12 月以後的相異之處，換句話說就是觀察 12 月以前類型 A 的顧客會買哪些產品，而 12 月以後類型 A 的顧客又會買哪些產品。

  在 frequent itemset 的長度限制在 2 以下，minimum confidence 為 0.001，限制 consequents 只包含 product 品項的情況下，總共整理出 28 種類型的顧客在 12 月前後買的產品有哪些差異。以職業類型為 skilled manual 的顧客為例，他們在 12 月前購買的內容為左圖，12 月後購買的內容為右圖

```
frozenset({'product_id_ADJ Rosy Sunglasses',
           'product_id_Fabulous Apple Drink',
           'product_id_Fast Fondue Mix',
           'product_id_Nationeel Dried Apples',
           'product_id_Sphinx Wheat Bread',
           'product_id_Steady Whitening Toothpast',
           'product_id_Super Apple Jam',
           'product_id_Super Strawberry Jelly'}),
```

```
frozenset({'product_id_American Chicken Hot Dogs',
           'product_id_American Sliced Ham',
           'product_id_Amigo Scallops',
           'product_id_Atomic Mints',
           'product_id_Atomic Spicy Mints',
           'product_id_BBB Best Brown Sugar',
           'product_id_BBB Best Chunky Peanut Butter',
           'product_id_BBB Best Pepper',
           'product_id_BBB Best Vegetable Oil',
           'product_id_Best Choice Cheese Crackers',
           'product_id_Best Choice Chocolate Donuts',
           'product_id_Best Choice Low Fat Popcorn',
           'product_id_Best Choice Potato Chips',
           'product_id_Best Choice Raspberry Fruit Roll',
           'product_id_Better Canned Tuna in Oil',
           'product_id_Big Time Beef TV Dinner',
           'product_id_Big Time Frozen Chicken Wings',
           'product_id_Big Time Frozen Corn',
           'product_id_Big Time Turkey TV Dinner',
           'product_id_Bird Call 200 MG Acetominifen',
           'product_id_Black Tie Eyeglass Screwdriver',
           'product_id_Blue Label Beef Soup',
           'product_id_Blue Label Canned Tuna in Water',
           'product_id_Blue Label Chicken Ramen Soup',
           'product_id_Blue Label Fancy Canned Anchovies',
           'product_id_Booker Muenster Cheese',
           'product_id_Booker Sharp Cheddar Cheese',
           'product_id_Bravo Canned Tuna in Oil',
           'product_id_Bravo Chicken Noodle Soup',
           'product_id_CDR Apple Jelly',
           'product_id_Carlson Jack Cheese',
           'product_id_Carrington Frozen Cauliflower',
           'product_id_Consolidated 200 MG Acetominifen',
           'product_id_Consolidated Mint Mouthwash',
           'product_id_Cutting Edge Sliced Turkey',
           'product_id_Cutting Edge Turkey Hot Dogs',
           'product_id_Denny Soft Napkins',
```