

ENTREGA NOSQL



UNIVERSIDAD COMPLUTENSE MADRID

Rafael Vargas Vargas
UNIVERSIDAD COMPLUTENSE DE MADRID |

ENTREGA MÓDULO NOSQL

INTRODUCCIÓN

Desde el departamento estratégico de la empresa COVAP están ideando un plan de internacionalización y puesto que el mercado asiático está en auge se han decidido por hacer una investigación en las principales ciudades de Birmania dado que al ser un mercado desconocido es importante conocer los patrones de compra de Yangon, Mandalay y Naypyitaw ya que son las que cuentan con una mayor concentración de población dentro del territorio y tienen un mayor potencial para nuestros intereses.

El dataset se compone de los siguientes campos:

ID de factura: Identificador único para cada transacción.

Sucursal: la ubicación de la sucursal del supermercado (por ejemplo, Yangon, Naypyitaw, Mandalay).

Ciudad: La ciudad en la que se encuentra ubicada la sucursal del supermercado.

Tipo de Cliente: Indica si el cliente es 'Miembro' o 'Normal'.

Género: Género del cliente.

Línea de productos: la categoría del producto vendido (por ejemplo, salud y belleza, accesorios electrónicos, hogar y estilo de vida).

Precio Unitario: Precio por unidad del producto.

Cantidad: Número de artículos comprados.

Impuesto del 5%: Monto del impuesto calculado sobre la transacción a una tasa del 5%.

Total: Monto total de la transacción incluyendo impuestos.

Fecha: Fecha de la transacción.

Hora: Hora de la transacción.

Pago: método de pago utilizado (por ejemplo, efectivo, billetera electrónica, tarjeta de crédito).

COGS: Costo de bienes vendidos, que representa el costo bruto de los productos.

Porcentaje de Margen Bruto: Porcentaje fijo de ganancia por cada venta (4,7619%).

Ingreso Bruto: Beneficio obtenido de la transacción.

Calificación: Calificación de satisfacción del cliente (sobre 10).

ANÁLISIS DEL DATASET

Comenzamos con una visualización del dataset para tener una visión global de los datos que tenemos disponibles para realizar el estudio:

```
db.Supermercados.find()
```

| Supermercados 0.019 s 1000 Docs | | | | | | | | | 50 | | | | p. 1 | 1 - 50 | Table | |
|---------------------------------|--------------------------|-------------|------|---------------|--------|--------------|------------------------|----------|--------|-----------|--|--|------|--------|-------|--|
| _id * | Invoice ID | Branch | City | Customer type | Gender | Product line | Unit price | Quantity | Tax 5% | | | | | | | |
| 1 | 67aa6c92d9d11e0d1bda6113 | 750-67-8428 | A | Yangon | Member | Female | Health and beauty | 74.69 0 | 7 | 26.1415 0 | | | | | | |
| 2 | 67aa6c92d9d11e0d1bda6114 | 226-31-3081 | C | Naypyitaw | Normal | Female | Electronic accessories | 15.28 0 | 5 | 3.82 0 | | | | | | |
| 3 | 67aa6c92d9d11e0d1bda6115 | 631-41-3108 | A | Yangon | Normal | Male | Home and lifestyle | 46.33 0 | 7 | 16.2155 0 | | | | | | |
| 4 | 67aa6c92d9d11e0d1bda6116 | 123-19-1176 | A | Yangon | Member | Male | Health and beauty | 58.22 0 | 8 | 23.288 0 | | | | | | |
| 5 | 67aa6c92d9d11e0d1bda6117 | 373-73-7910 | A | Yangon | Normal | Male | Sports and travel | 86.31 0 | 7 | 30.2085 0 | | | | | | |
| 6 | 67aa6c92d9d11e0d1bda6118 | 699-14-3026 | C | Naypyitaw | Normal | Male | Electronic accessories | 85.39 0 | 7 | 29.8865 0 | | | | | | |
| 7 | 67aa6c92d9d11e0d1bda6119 | 355-53-5943 | A | Yangon | Member | Female | Electronic accessories | 68.84 0 | 6 | 20.652 0 | | | | | | |
| 8 | 67aa6c92d9d11e0d1bda611a | 315-22-5665 | C | Naypyitaw | Normal | Female | Home and lifestyle | 73.56 0 | 10 | 36.78 0 | | | | | | |
| 9 | 67aa6c92d9d11e0d1bda611b | 665-32-9167 | A | Yangon | Member | Female | Health and beauty | 36.26 0 | 2 | 3.626 0 | | | | | | |
| 10 | 67aa6c92d9d11e0d1bda611c | 692-92-5582 | B | Mandalay | Member | Female | Food and beverages | 54.84 0 | 3 | 8.226 0 | | | | | | |
| 11 | 67aa6c92d9d11e0d1bda611d | 351-62-0822 | B | Mandalay | Member | Female | Fashion accessories | 14.48 0 | 4 | 2.896 0 | | | | | | |
| 12 | 67aa6c92d9d11e0d1bda611e | 529-56-3974 | B | Mandalay | Member | Male | Electronic accessories | 25.51 0 | 4 | 5.102 0 | | | | | | |
| 13 | 67aa6c92d9d11e0d1bda611f | 365-64-0515 | A | Yangon | Normal | Female | Electronic accessories | 46.95 0 | 5 | 11.7375 0 | | | | | | |
| 14 | 67aa6c92d9d11e0d1bda6120 | 252-56-2699 | A | Yangon | Normal | Male | Food and beverages | 43.19 0 | 10 | 21.595 0 | | | | | | |
| 15 | 67aa6c92d9d11e0d1bda6121 | 829-34-3910 | A | Yangon | Normal | Female | Health and beauty | 71.38 0 | 10 | 35.69 0 | | | | | | |
| 16 | 67aa6c92d9d11e0d1bda6122 | 299-46-1805 | B | Mandalay | Member | Female | Sports and travel | 93.72 0 | 6 | 28.116 0 | | | | | | |
| 17 | 67aa6c92d9d11e0d1bda6123 | 656-95-9349 | A | Yangon | Member | Female | Health and beauty | 68.93 0 | 7 | 24.1255 0 | | | | | | |
| 18 | 67aa6c92d9d11e0d1bda6124 | 765-26-6951 | A | Yangon | Normal | Male | Sports and travel | 72.61 0 | 6 | 21.783 0 | | | | | | |
| 19 | 67aa6c92d9d11e0d1bda6125 | 329-62-1586 | A | Yangon | Normal | Male | Food and beverages | 54.67 0 | 3 | 8.2005 0 | | | | | | |
| 20 | 67aa6c92d9d11e0d1bda6126 | 319-50-3348 | B | Mandalay | Normal | Female | Home and lifestyle | 40.3 0 | 2 | 4.03 0 | | | | | | |
| 21 | 67aa6c92d9d11e0d1bda6127 | 300-71-4605 | C | Naypyitaw | Member | Male | Electronic accessories | 86.04 0 | 5 | 21.51 0 | | | | | | |
| 22 | 67aa6c92d9d11e0d1bda6128 | 371-85-5789 | B | Mandalay | Normal | Male | Health and beauty | 87.98 0 | 3 | 13.197 0 | | | | | | |
| 23 | 67aa6c92d9d11e0d1bda6129 | 273-16-6619 | B | Mandalay | Normal | Male | Home and lifestyle | 33.2 0 | 2 | 3.32 0 | | | | | | |

Podemos observar si hay una representación equitativa de las ciudades en las cuales se va a realizar el estudio, y para ello agrupamos el numero de transacciones por ciudad al igual que el importe de esas ventas de manera que si es similar podemos dar comienzo al estudio.

```
db.Supermercados.aggregate([
{
  $group: {
    _id: "$City",
    totalSales: {$sum: "$Total"},
    totalTransactions: {$sum: 1},
  }
}
]);
```

| | _id * | Ventas_totales | Operaciones_totales |
|---|-----------|----------------------|---------------------|
| 1 | Yangon | 106.200,3705 (0.11M) | 340 |
| 2 | Naypyitaw | 110.568,7065 (0.11M) | 328 |
| 3 | Mandalay | 106.197,672 (0.11M) | 332 |

Ahora que certificamos que la representación de las ciudades es similar podemos empezar a investigar a través de esta clasificación en que ciudad se realiza un gasto medio mayor por compra para empezar a comparar entre las diferentes opciones.

```
db.Supermercados.aggregate([
{
  $group: {
    _id: "$City",
    Gasto_medio: {$avg: "$Total"}
  }
}
]);
```

| Go to line 3 Supermercados 3 Docs | | |
|-----------------------------------|-----------|---------------|
| | _id * ◆ | Gasto_medio ◆ |
| 1 | Mandalay | 319,8725.0 |
| 2 | Yangon | 312,354.0 |
| 3 | Naypyitaw | 337,0997.0 |

De primera instancia podemos observar rápidamente que se produce un mayor gasto por compra en Naypyitaw con una diferencia bastante clara sobre el resto que se encuentran más igualadas.

A continuación, podemos ver una comparativa de valoración de los usuarios entre las distintas ciudades donde realizan sus compras de manera que podemos ver que sucursal es más rentable y cuál de ellas tiene mayor valoración.

```
db.Supermercados.aggregate([
{
  $group: {
    _id: {Ciudad: "$City", sucursal: "$Branch"},
    total_ventas: {$sum: "$Total"},
    rating_promedio: {$avg: "$Rating"}
  }
},
{$sort: { total_ventas: -1}}
]);
```

| | _id | | total_ventas ⬇ | rating_promedio ⬇ |
|---|-----------|------------|----------------------|-------------------|
| | Ciudad ⬇ | sucursal ⬇ | | |
| 1 | Naypyitaw | C | 110.568,7065 (0.11M) | 7,0729.0 |
| 2 | Yangon | A | 106.200,3705 (0.11M) | 7,0271.0 |
| 3 | Mandalay | B | 106.197,672 (0.11M) | 6,8181.0 |

En este caso como podemos observar fácilmente, la sucursal “C” que corresponde a la situada en la ciudad de Naypyitaw es la que mayor facturación de ventas obtiene como ya mencionamos anteriormente pero también cuenta con un mayor promedio de valoración de cara a sus clientes de esta manera y por el momento, La sucursal “C” se posiciona por encima a la hora de tomar la decisión de donde situar nuestra sucursal.

Otra cuestión fundamental que tenemos que resolver a la hora de realizar ventas es investigar acerca de las preferencias de los clientes y si varían en los distintos puntos de venta que tenemos en nuestra investigación.

```
db.Supermercados.aggregate([
```

```
  {$group: {_id: {ciudad: "$City", categoria: "$Product line"}, total_vendido: {$sum: "$Total"}}},
```

```
  {$sort: {total_vendido: -1}}]
```

| | | |
|-----------|------------------------|---------------------|
| Naypyitaw | Electronic accessories | 18.968,9745 (19.0K) |
| Yangon | Electronic accessories | 18.317,1135 (18.3K) |
| Mandalay | Electronic accessories | 17.051,4435 (17.1K) |
| Naypyitaw | Fashion accessories | 21.560,07 (21.6K) |
| Mandalay | Fashion accessories | 16.413,3165 (16.4K) |
| Yangon | Fashion accessories | 16.332,5085 (16.3K) |
| Naypyitaw | Food and beverages | 23.766,855 (23.8K) |
| Yangon | Food and beverages | 17.163,1005 (17.2K) |
| Mandalay | Food and beverages | 15.214,8885 (15.2K) |
| Mandalay | Health and beauty | 19.980,66 (20.0K) |
| Naypyitaw | Health and beauty | 16.615,326 (16.6K) |
| Yangon | Health and beauty | 12.597,753 (12.6K) |
| Yangon | Home and lifestyle | 22.417,1955 (22.4K) |
| Mandalay | Home and lifestyle | 17.549,1645 (17.5K) |
| Naypyitaw | Home and lifestyle | 13.895,553 (13.9K) |
| Mandalay | Sports and travel | 19.988,199 (20.0K) |
| Yangon | Sports and travel | 19.372,6995 (19.4K) |
| Naypyitaw | Sports and travel | 15.761,928 (15.8K) |

Ahora contamos con los datos clasificados por las diferentes categorías de productos que se venden principalmente en las distintas ciudades y de esta manera podemos comparar las preferencias dependiendo de la ciudad donde se realiza la compra.

Otra información interesante que podría resultar de utilidad es si los productos más caros se compran en menor cantidad y que diferencia hay entre los distintos grupos.

```
db.Supermercados.aggregate([
  {$group: {_id: "$Product line", precio_promedio: {$avg: "$Unit price"}, cantidad_promedio:
  {$avg: "$Quantity"}}},
  {$sort: { precio_promedio: -1}}
])
```

| | _id * ⇅ | precio_promedio ⇅ | cantidad_promedio ⇅ |
|---|------------------------|-------------------|---------------------|
| 1 | Fashion accessories | 57,1537.0 | 5,0674.0 |
| 2 | Sports and travel | 56,9933.0 | 5,5422.0 |
| 3 | Food and beverages | 56,0089.0 | 5,4713.0 |
| 4 | Home and lifestyle | 55,3169.0 | 5,6938.0 |
| 5 | Health and beauty | 54,8545.0 | 5,6184.0 |
| 6 | Electronic accessories | 53,5516.0 | 5,7118.0 |

En este caso, no hay grandes diferencias entre los grupos y no nos aporta información demasiado relevante.

Otro aspecto relevante es saber el comportamiento de distintos clientes (Member vs Normal) si tienen diferentes gastos.

```
db.Supermercados.aggregate([
  {$group: {_id: "$Customer type", gasto_total: {$sum: "$Total"}, gasto_promedio: {$avg:
  "$Total"}}}
])
```

| _id * ⇅ | gasto_total ⇅ | gasto_promedio ⇅ |
|---------|---------------------|------------------|
| Member | 164.223,444 (0.16M) | 327,7913.0 |
| Normal | 158.743,305 (0.16M) | 318,1229.0 |

Fácilmente entendible que aquellos clientes que gozan de descuentos en su tienda de confianza tienen más motivos para realizar un mayor gasto pero cierto es que no hay una diferencia demasiado significativa, pero ¿sucede igual en cada ciudad?

```

db.Supermercados.aggregate([
  {
    $group: {
      _id: {tipo_cliente:"$Customer type", ciudad: "$City"},
      gasto_total: {$sum: "$Total"},
      gasto_promedio: {$avg: "$Total"}
    }
  },
  {$sort: {"_id.ciudad": 1, "_id.tipo_cliente": 1}}
])

```

| | _id | | gasto_total ↕ | gasto_promedio ↕ |
|---|----------------|-----------|---------------------|------------------|
| | tipo_cliente ↕ | ciudad ↕ | | |
| 1 | Member | Mandalay | 53.704,686 (53.7K) | 325,4829.0 |
| 2 | Normal | Mandalay | 52.492,986 (52.5K) | 314,3293.0 |
| 3 | Member | Naypyitaw | 56.881,2825 (56.9K) | 336,5756.0 |
| 4 | Normal | Naypyitaw | 53.687,424 (53.7K) | 337,6568.0 |
| 5 | Member | Yangon | 53.637,4755 (53.6K) | 321,1825.0 |
| 6 | Normal | Yangon | 52.562,895 (52.6K) | 303,8318.0 |

En este caso, tanto las ciudades de Yangon como Mandalay siguen el patrón del resultado general, pero en Naypyitaw se iguala tanto entre ambos tipos que incluso la categoría “Member” es un poco superior, esto puede ser a que en esta sucursal se hagan más compras de manera esporádica que no de manera recurrente y puede ser que la fidelidad del cliente sea menor que en el resto.

Un aspecto para el apartado comercial es ver el método de pago que usan los clientes en las distintas ciudades para anticipar la actuación el método de pago.

```

db.Supermercados.aggregate([
  {$group: {_id: {ciudad: "$City", pago: "$Payment"}, cantidad: {$sum: 1}}},
  {$sort: {cantidad: -1}}
])

```

| | _id | | cantidad ↕ |
|---|-----------|-------------|------------|
| | ciudad ▲ | pago ⇅ | |
| 1 | Mandalay | Ewallet | 113 |
| 2 | Mandalay | Cash | 110 |
| 3 | Mandalay | Credit card | 109 |
| 4 | Naypyitaw | Cash | 124 |
| 5 | Naypyitaw | Ewallet | 106 |
| 6 | Naypyitaw | Credit card | 98 |
| 7 | Yangon | Ewallet | 126 |
| 8 | Yangon | Cash | 110 |
| 9 | Yangon | Credit card | 104 |

Aquí podemos observar que en Mandalay y Yangon el método de pago más utilizado es la Ewallet aunque cabe destacar que en Mandalay la distribución de uso del método de pago entre las diferentes compras es muy igualitario. Por otro lado, en Naypyitaw sobresale el método de utilizar efectivo para realizar el pago, seguido por Ewallet.

CONCLUSIONES

Cabe destacar que tenemos que situarnos desde la perspectiva que nuestra entidad en este caso COVAP tiene como valor de negocio que es la producción y comercialización de alimentos y bebidas entonces, todas nuestras conclusiones están condicionadas por este factor que debemos tener en cuenta.

Si visualizamos los primeros indicadores que tenemos en los dos primeros informes podemos ver que el mayor gasto general y promedio por compra se concentra en la ciudad de Naypyitaw, en primera instancia podemos intuir que puede ser una gran ubicación para situar la nueva tienda, pero no debemos precipitarnos y mirar un poco más con detalle otros factores.

Una relación directa que tiene entre sucursal y ventas es la valoración que el cliente en la cual a mayor valoración se obtiene una mayor facturación por lo que es importante tener una imagen y un servicio cuidado si queremos adentrarnos dentro del mercado birmano.

Clasificando las preferencias de cada ciudad por categoría de producto, debemos centrarnos en “Food and beverages” que es la categoría que corresponde a nuestro modelo de negocio en el cual una vez más Naypyitaw tiene una gran diferencia respecto al resto de ciudades en términos de facturación.

La comparación global por categoría no nos transmite mucha información muy relevante, pero podemos ver que la categoría de “Food and beverages” se sitúa en tercera posición a nivel de precio y a nivel de unidades de compra es la segunda menor solo por encima de la sección de “Fashion accesorios” pero recalcar que las diferencias son mínimas y no debemos centrar este aspecto como uno de los principales para tomar la decisión.

El aspecto general a nivel global de la relación entre gasto de cada tipo de cliente muestra una diferencia entre aquellos clientes “Members” donde tienen un gasto superior a los “Normales”, este dato extrapolado a cada ciudad hace que este patrón se repita en las ciudades de Yangon y

Mandalay pero en Naypyitaw no se cumple este patrón ya que ambas categorías se encuentran muy igualadas, podemos deducir que en esta región no es tan importante el aspecto de fidelización del cliente ya que ambos tipos tienen un gasto medio muy similar.

Respecto al método de pago el cliente suele preferir utilizar la Ewallet a la hora de realizar sus compras menos en la ciudad de Naypyitaw donde el cash es la preferencia de la mayoría de los compradores de la ciudad. Esto puede ser debido a que es una ciudad más turística y de compra casual ya que unido al aspecto de fidelización del cliente nos ayuda a dar por buenas esta conclusión.

Una vez analizada la información con la que contamos, parece ser que la ciudad de Naypyitaw es una buena opción para iniciar la primera tienda de este mercado debido a que es la ciudad donde más facturación existe en relación con el resto de las comparadas y los clientes tienen preferencia por la categoría de los productos que ofrecemos. Aparte de eso hay que tener en cuenta que no sería muy útil centrarse en una estrategia de fidelización de clientes y utilizar esta inversión en optimizar costes para competir por precios con la competencia o buscar una diferenciación de nuestro producto.

Aparte de esto tendremos que ofrecer un buen servicio con unas buenas instalaciones para que el rating del establecimiento sea positivo ya que existe una relación directa entre la valoración del cliente y la facturación de los establecimientos.