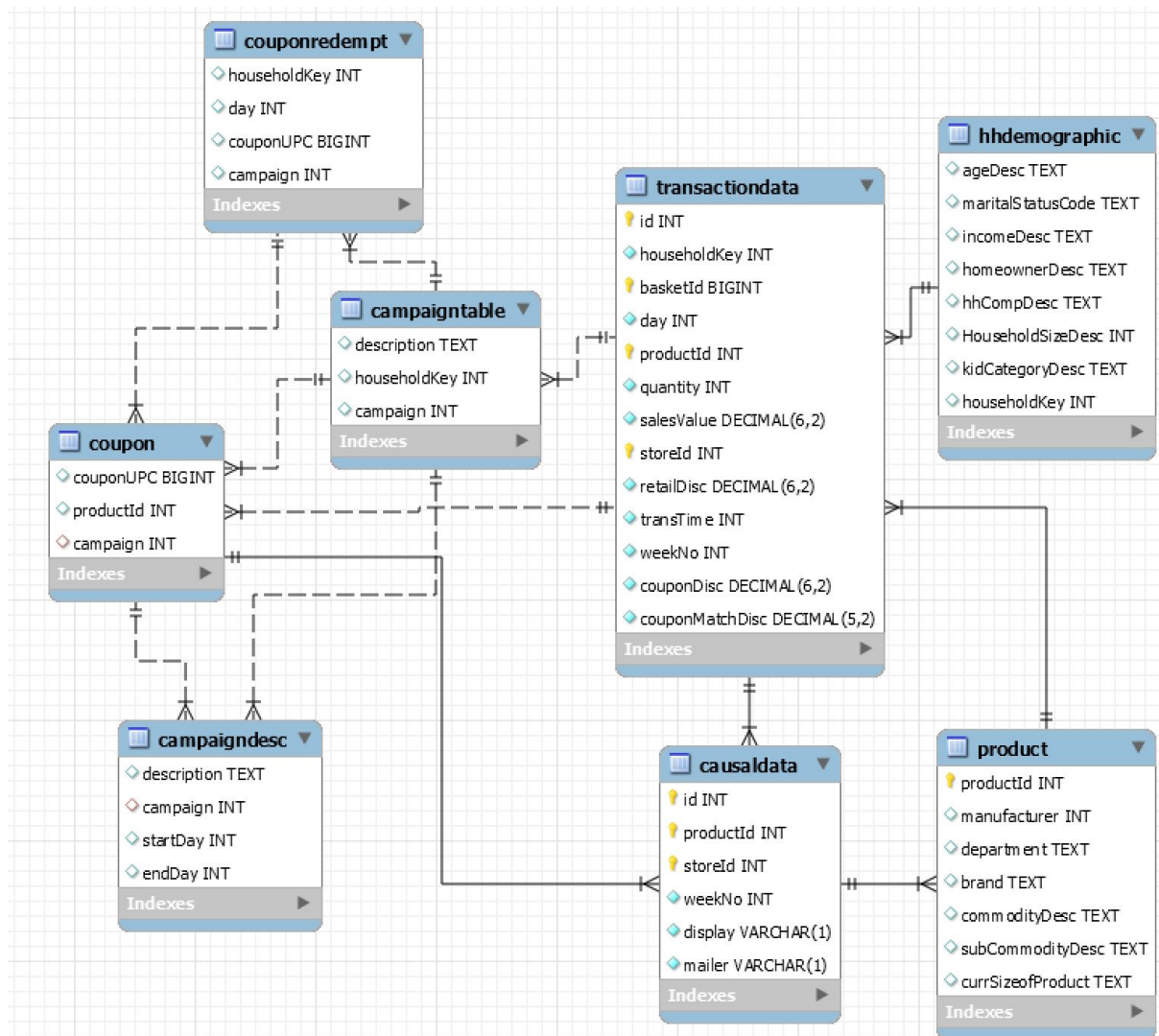


Proyecto Bedu MySQL y MongoDB

Las tablas que estaré analizando son transacciones por 2 años de un grupo de 2,500 dueños de casas que atienden seguido a la tienda. Se va analizar las ventas que se tuvieron a través de este tiempo, el impacto de las campañas, los clientes y los productos que se vendieron.

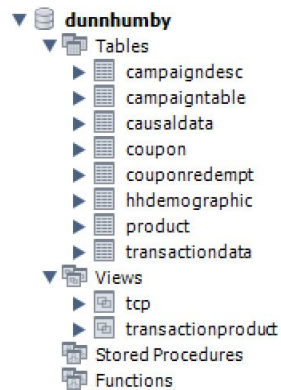
La base de datos es muy amigable. Los datos que se utilizan a través del análisis no tienen ningún significado. Las tiendas, productos, y campañas se identifican por números.El tiempo se encuentra en minutos.

Schema



Para este proyecto utilice mongoDB y MySQL con las mismas bases de datos.

BBDD en MySQL



BBDD en MongoDB

campaignDesc	30	97.7 B	2.9 KB	1	20.0 KB	
campaignTable	7,208	85.3 B	600.4 KB	1	92.0 KB	
casualData	99,999	121.1 B	11.5 MB	1	900.0 KB	
coupon	124,548	90.7 B	10.8 MB	1	1.1 MB	
couponRedempt	2,318	103.3 B	233.9 KB	1	48.0 KB	
hhDemographic	801	236.2 B	184.8 KB	1	24.0 KB	
product	92,353	222.6 B	19.6 MB	1	832.0 KB	
tp (view on: transactionData)	0	-	0.0 B		0.0 B	
transactionData	99,999	278.8 B	26.6 MB	1	900.0 KB	

Consultas

1. Muestra el rango de edades de las personas que gastan más a través del tiempo.

```
SELECT      AVG(t.transTime)/60 AS tiempotrans, SUM(t.salesValue + t.retailDisc
+ t.couponDisc + t.couponMatchDisc) AS ventaTot, h.ageDesc
FROM transactiondata t
JOIN hhdemographic h
      ON t.householdKey = h.householdKey
GROUP BY h.ageDesc
ORDER BY ventaTot DESC;
```

tiempotrans	ventaTot	ageDesc
26.02376731	50521.93	45-54
26.45564669	37852.55	35-44
26.97496226	26619.36	25-34
24.89805317	12985.01	65+
25.78810493	10728.95	55-64
27.14412854	8297.51	19-24

2. Qué tipo de clientes gastan menos?

```
SELECT SUM(t.salesValue + t.retailDisc + t.couponDisc + t.couponMatchDisc) AS
ventaTot, h.ageDesc
FROM transactiondata t
JOIN hhdemographic h
      ON t.householdKey = h.householdKey
GROUP BY h.ageDesc
ORDER BY ventaTot
LIMIT 1;
```

Los clientes que gastan menos son los de 19-24 años.

ventaTot	ageDesc
8297.51	19-24

3. Qué tipo de clientes gastan más?

```
SELECT SUM(t.salesValue + t.retailDisc + t.couponDisc + t.couponMatchDisc) AS
ventaTot, h.ageDesc
FROM transactiondata t
JOIN hhdemographic h
      ON t.householdKey = h.householdKey
GROUP BY h.ageDesc
ORDER BY ventaTot DESC
LIMIT 1;
```

Los clientes que gastan más son los de 45-54 años.

ventaTot	ageDesc
50521.93	45-54

4. Cuántas campañas hicieron?

MySQL:

```
SELECT count(*), campaign
FROM campaigntable
group by campaign;
```

Se hicieron 30 campañas en total.

MongoDB:

```
[{$group: {
  _id: "$CAMPAIGN"
}}, {$sort: {
  _id: 1
```

```
}}, {$count: 'campañas'})
```

```
campañas: 30
```

5. Cuánto duró cada campaña?

MySQL:

```
SELECT (endDay - startDay) AS dias, campaign  
FROM campaigndesc  
ORDER BY dias DESC;
```

dias	campaign
161	15
70	20
65	14
63	27
61	28
60	24
56	3
55	18
53	29
48	8
47	13
46	11
46	30
40	26
38	23
37	1
34	7
34	5
32	25
32	21
32	22
32	19
32	17
32	16
32	12
32	10
32	9
32	6
32	4
32	2

MongoDB:

```
[{$addFields: {  
  startday: {$convert: {input: "$START_DAY", to: "double"}},  
  endday: {$convert: {input: "$END_DAY", to: "double"}}  
}}, {$addFields: {
```

```
  duracion: {  
    $subtract: ["$endday", "$startday"]  
  }  
}}, {$project: {  
  duracion: 1,  
  CAMPAIGN: 1
```

```
}}, {$sort: {
  duracion: -1
}}]
```

```
_id: ObjectId("5f18fd4b70f188223430ec66")
CAMPAIGN: "15"
duracion: 161
```

```
_id: ObjectId("5f18fd4b70f188223430ec68")
CAMPAIGN: "20"
duracion: 70
```

6. Ordena las campañas de mayor a menor de acuerdo al éxito que se obtuvo

```
SELECT count(*) AS cantidad, campaign
FROM campaigntable
group by campaign
ORDER BY cantidad DESC;
```

cantidad	campaign
1133	18
1077	13
1076	8
361	30
332	26
276	22
244	20
224	14
214	11
202	17
198	7
188	16
187	25
183	23
176	9
170	12
166	5
130	19
123	10
118	29
100	24
81	4
65	6
65	21
48	2
17	15
17	28
13	1
12	3
12	27

7. Qué semana fue en la que se ganó más?

MySQL:

```
SELECT SUM(salesValue + retailDisc + couponDisc + couponMatchDisc) AS
ventaTot, weekNo
FROM transactiondata
GROUP BY weekNo
ORDER BY ventaTot DESC
LIMIT 1;
```

En la semana 11 se ganó más.

ventaTot	weekNo
41971.15	11

MongoDB:

```
{ $addFields: {
  retailDisc: { $convert: { input: "$RETAIL_DISC", to: "double" } },
  salesValue: { $convert: { input: "$SALES_VALUE", to: "double" } },
  couponDisc: { $convert: { input: "$COUPON_DISC", to: "double" } },
  couponMatchDisc: { $convert: { input: "$COUPON_MATCH_DISC", to: "double" } }
}, { $addFields: {
  ventasTot: { $sum: ["$retailDisc", "$salesValue", "$couponDisc",
"$couponMatchDisc"]
}
}}, { $group: {
  _id: "$WEEK_NO",
  ventas: {
    $sum: "$ventasTot"
  }
}}, { $sort: {
  ventas: -1
}}, { $limit: 1}]
```

_id: "11"
ventas: 41971.15

8. Qué semana fue en la que se ganó menos?

MySQL:

```
SELECT SUM(salesValue + retailDisc + couponDisc + couponMatchDisc) AS
ventaTot, weekNo
FROM transactiondata
GROUP BY weekNo
ORDER BY ventaTot
LIMIT 1;
```

En la primera semana se ganó menos.

ventaTot	weekNo
4154.42	1

MongoDB:

```
{ $addFields: {
  retailDisc: { $convert: { input: "$RETAIL_DISC", to: "double" } },
  salesValue: { $convert: { input: "$SALES_VALUE", to: "double" } },
  couponDisc: { $convert: { input: "$COUPON_DISC", to: "double" } },
  couponMatchDisc: { $convert: { input: "$COUPON_MATCH_DISC", to: "double" } }
}, { $addFields: {
  ventasTot: { $sum: ["$retailDisc", "$salesValue", "$couponDisc",
"$couponMatchDisc"]
}
}}
```

```

}}, {$group: {
  _id: "$WEEK_NO",
  ventas: {
    $sum: "$ventasTot"
  }
}}, {$sort: {
  ventas: 1
}}, {$limit: 1}]

```

```

_id: "1"
ventas: 4154.42

```

9. Qué edad compran más a través del tiempo?

```

SELECT      sum(t.transTime) AS tiempotrans, h.ageDesc
FROM transactiondata t
JOIN hhdemographic h
      ON t.householdKey = h.householdKey
GROUP BY h.ageDesc
ORDER BY tiempotrans DESC
LIMIT 1;

```

Los clientes que compran más son los de 45-54 años.

tiempotrans	ageDesc
22103547	45-54

10. Qué edad compran menos a través del tiempo?

```

SELECT      sum(t.transTime) AS tiempotrans, h.ageDesc
FROM transactiondata t
JOIN hhdemographic h
      ON t.householdKey = h.householdKey
GROUP BY h.ageDesc
ORDER BY tiempotrans
LIMIT 1;

```

Los clientes que compran menos a través del tiempo son los de 19-24 años.

tiempotrans	ageDesc
3666086	19-24

11.Cuál es el departamento que vende más?

```

CREATE VIEW transactionproduct AS
(SELECT t.day, t.productId, t.weekNo, p.department, p.brand, t.householdKey,
t.storeId
FROM transactiondata t
JOIN product p
      ON t.productId = p.productId);

```

```

SELECT COUNT(*) AS cantventas, department
FROM transactionproduct

```

```
GROUP BY department
ORDER BY cantventas DESC
LIMIT 1;
```

El departamento que vende más es grocery.

cantventas	department
62814	GROCERY

12. Cuál es el departamento que vende menos?

```
SELECT COUNT(*) AS cantventas, department
FROM transactionproduct
GROUP BY department
ORDER BY cantventas
LIMIT 1;
```

El departamento que vende menos es pork.

cantventas	department
1	PORK

13. Cuántas campañas estuvieron relacionadas al departamento que vende menos?

```
CREATE VIEW tcp AS
(SELECT c.campaign, t.productId, p.department, p.brand
FROM transactiondata t
JOIN coupon c
ON t.productId = c.productId
JOIN product p
ON t.productId = p.productId);
```

```
SELECT COUNT(*) AS cantcampanas, department
FROM tcp
WHERE department = 'PORK';
```

Se hicieron 3 campañas al departamento de PORK.

cantcampanas	department
3	PORK

14. Cuántas campañas estuvieron relacionadas al departamento que vende más?

```
SELECT COUNT(*) AS cantcampanas, department
FROM tcp
WHERE department = 'PORK';
CREATE VIEW tcp2 AS
(SELECT COUNT(campaign), campaign, department
FROM tcp
WHERE department = 'GROCERY'
GROUP BY campaign);
```

```
SELECT COUNT(*)
FROM tcp2;
```

Estuvieron 26 campañas relacionadas al departamento de GROCERY.

COUNT(*)
26

15. Qué campañas se utilizaron más?

```
SELECT c.campaign, COUNT(c.campaign) AS cantidad
FROM transactiondata t
JOIN coupon c
      ON t.productId = c.productId
WHERE t.couponDisc < 0
GROUP BY c.campaign
ORDER BY cantidad DESC;
```

campaign	cantidad
13	498
18	442
8	294
27	107
26	100
30	85
24	32
17	28
2	26
23	23
14	23
25	22

16. Cuántos departamentos hay?

```
{{ $group: {
  _id: '$DEPARTMENT'
}}, { $sort: {
  _id: 1
}}, { $match: {
  _id: { $ne: "" }
}}, { $count: 'dpto' }}
```

dpto: 43

17. Cuántos productos hay?

```
{{ $group: {
  _id: '$PRODUCT_ID'
}}, { $sort: {
  _id: 1
}}, { $match: {
  _id: { $ne: "" }
}}, { $count: 'productos' }}
```

productos: 92353

18. Cuántos tiendas hay?

```
[$group: {
  _id: "$STORE_ID"
}], {$match: {
  _id: { $ne: " " }
}}, {$count: 'tiendas'}
```

```
tiendas: 115
```

19. Cuánto se ganó en total?

```
[$addFields: {
  retailDisc: {$convert: {input: "$RETAIL_DISC", to: "double"}},
  salesValue: {$convert: {input: "$SALES_VALUE", to: "double"}},
  couponDisc: {$convert: {input: "$COUPON_DISC", to: "double"}},
  couponMatchDisc: {$convert: {input: "$COUPON_MATCH_DISC", to: "double"}}
}], {$addFields: {
  ventaTot: {
    $sum: ["$retailDisc", "$salesValue", "$couponDisc", "$couponMatchDisc"]
  }
}}, {$group: {
  _id: null,
  ventas: {
    $sum: "$ventaTot"
  }
}}]
```

```
_id: null
ventas: 241747.78
```

20. Crea una vista donde se vea ventas, departamento, household_key, PRODUCT_ID, y STORE_ID.

```
[$lookup: {
  from: 'product',
  localField: 'PRODUCT_ID',
  foreignField: 'PRODUCT_ID',
  as: 'PRODUCT'
}], {$addFields: {
  retailDisc: {$convert: {input: "$RETAIL_DISC", to: "double"}},
  salesValue: {$convert: {input: "$SALES_VALUE", to: "double"}},
  couponDisc: {$convert: {input: "$COUPON_DISC", to: "double"}},
```

```

couponMatchDisc: {$convert: {input: "$COUPON_MATCH_DISC", to: "double"}}
}}, {$unwind: {
  path: "$PRODUCT"
}}, {$addFields: {
  departamento: "$PRODUCT.DEPARTMENT",
  ventasTot: { $sum: ["$retailDisc", "$salesValue", "$couponDisc",
"$couponMatchDisc"]
}
}}, {$project: {
  ventasTot:1,
  departamento:1,
  PRODUCT_ID:1,
  STORE_ID:1,
  household_key:1
}}]

```

AVH.tp (view on: AVH.transactionData)

MODIFY SOURCE

Read Only

Documents Aggregations Explain Plan Indexes

COLLATION Untitled SAVE SAMPLE MODE AUTO PREVIEW

N/A Documents in the Collection

Select an operator to construct expressions used in the aggregation pipeline stages. [Learn more](#)

Preview of Documents in the Collection

```

_id: ObjectId("5f18fc4d70f18822342ddf27")
household_key: "2375"
PRODUCT_ID: "1004906"
STORE_ID: "364"
departamento: "PRODUCE"
ventasTot: 0.7899999999999999

```

```

_id: ObjectId("5f18fc4d70f18822342ddf2c")
household_key: "2375"
PRODUCT_ID: "826249"
STORE_ID: "364"
departamento: "GROCERY"
ventasTot: 1.38

```

21. Menciona los rangos de edad de las campañas

```

[{$group: {
  _id: "$AGE_DESC"
}}, {$sort: {
  _id: 1
}}]

```

```

_id: "19-24"

```

```

_id: "25-34"

```

22. Cuales productos están relacionados a grocery?

```

[{$match: {
  DEPARTMENT: "GROCERY"
}}, {$project: {
  PRODUCT_ID: 1,
  DEPARTMENT: 1,
  MANUFACTURER:1
}}, {$sort: {

```

PRODUCT_ID: 1

}}

```
_id: ObjectId("5f18fafd70f1882234296708")  
PRODUCT_ID: "1000050"  
MANUFACTURER: "1046"  
DEPARTMENT: "GROCERY"
```

```
_id: ObjectId("5f18fafd70f1882234296709")  
PRODUCT_ID: "1000057"  
MANUFACTURER: "1046"  
DEPARTMENT: "GROCERY"
```

23. Cuántos productos están relacionados a grocery?

```
{{ $match: {  
  DEPARTMENT: "GROCERY"  
}}, { $project: {  
  PRODUCT_ID: 1,  
  DEPARTMENT: 1,  
  MANUFACTURER: 1  
}}, { $sort: {  
  PRODUCT_ID: 1  
}}, { $count: 'productos' ]}
```

Hay 39,021 productos relacionados a grocery.

```
productos: 39021
```

24. Cuales productos están relacionados a pork?

```
{{ $match: {  
  DEPARTMENT: "PORK"  
}}, { $project: {  
  PRODUCT_ID: 1,  
  DEPARTMENT: 1,  
  MANUFACTURER: 1  
}}]
```

```
_id: ObjectId("5f18fafa70f1882234297183")  
PRODUCT_ID: "1023792"  
MANUFACTURER: "3030"  
DEPARTMENT: "PORK"
```

25. Cuántos productos están relacionados a pork?

```
{{ $match: {  
  DEPARTMENT: "PORK"  
}}, { $project: {  
  PRODUCT_ID: 1,  
  DEPARTMENT: 1,  
  MANUFACTURER: 1  
}}]
```

```
}}, {$count: 'productos'}]
```

Hay un producto relacionado a pork.

```
productos: 1
```

26. Cuántos productos se vendieron en total?

```
[$addFields: {  
  cantidad: {$convert: {input: "$QUANTITY", to: "double"}}  
}], {$group: {  
  _id: null,  
  cantproductosvendidos: {  
    $sum: "$cantidad"  
  }  
}}]
```

Se vendieron 4,939,052 productos.

```
_id: null  
cantproductosvendidos: 4939052
```

27. Cuánto se ganó en el departamento de grocery?

```
SELECT SUM(t.salesValue + t.retailDisc + t.couponDisc + t.couponMatchDisc) AS  
ventaTot, p.department  
FROM transactiondata t  
JOIN product p  
  ON t.productId = p.productId  
WHERE p.department = "GROCERY"  
GROUP BY p.department;
```

El departamento de grocery ganó \$119,679.01.

ventaTot	department
119679.01	GROCERY

28. Cuánto se ganó en el departamento de pork?

```
SELECT SUM(t.salesValue + t.retailDisc + t.couponDisc + t.couponMatchDisc) AS  
ventaTot, p.department  
FROM transactiondata t  
JOIN product p  
  ON t.productId = p.productId  
WHERE p.department = "PORK"  
GROUP BY p.department;
```

El departamento de pork ganó \$7.96.

ventaTot	department
7.96	PORK

29. Muestra las ventas totales y cantidad de productos vendidos por tienda

```
SELECT SUM(salesValue + retailDisc + couponDisc + couponMatchDisc) AS  
ventaTot, storeId, SUM(quantity) AS cantProd  
FROM transactiondata  
GROUP BY storeId  
ORDER BY ventaTot DESC;
```

ventaTot	storeId	cantProd	ventaTot	storeId	cantProd
12685.51	367	306680	5.00	3168	2
5817.22	375	246661	4.44	21	4
5291.11	32004	2754	4.02	2602	4
5245.68	361	346996	3.88	3076	2
4829.34	356	2765	3.42	2748	4
4767.19	327	201368	3.29	2915	1
4642.42	315	2169	2.80	77	2
4176.06	406	61128	2.58	588	2
3659.54	333	1696	1.99	2996	1
3603.76	372	61368	1.62	3124	2
3521.67	292	2054	-1.45	673	5

Conclusiones

El mercado que genera más del 20% está en el rango de edad de 45-54. En las próximas campañas de mercado me enfocaría más en ese mercado para poder crecer. Sin embargo, analizaría más a fondo cuáles campañas y cuales productos son los que compran menos para generar más compras y rotación de esos productos. La próxima vez tomar en cuenta que la duración de las campañas no afecta sino más bien al área que está enfocado y también el mercado. Cerraría las tiendas con ventasTot negativas y también las que venden muy pocos productos.