

Saulo Gabriel López Antonio

data-analysis-gdl-20-04

PROYECTO

“Captación y cuentas de ahorro de la banca mexicana”

1.- Definición del proyecto (fuentes y problemas a resolver)

A) Fuente Base de datos

<https://portafoliodeinformacion.cnbv.gob.mx/bm1/Paginas/infooper.aspx>

Reporte 040-4A-R10 y 040-4A-R11 del 2020

B) Problema a resolver

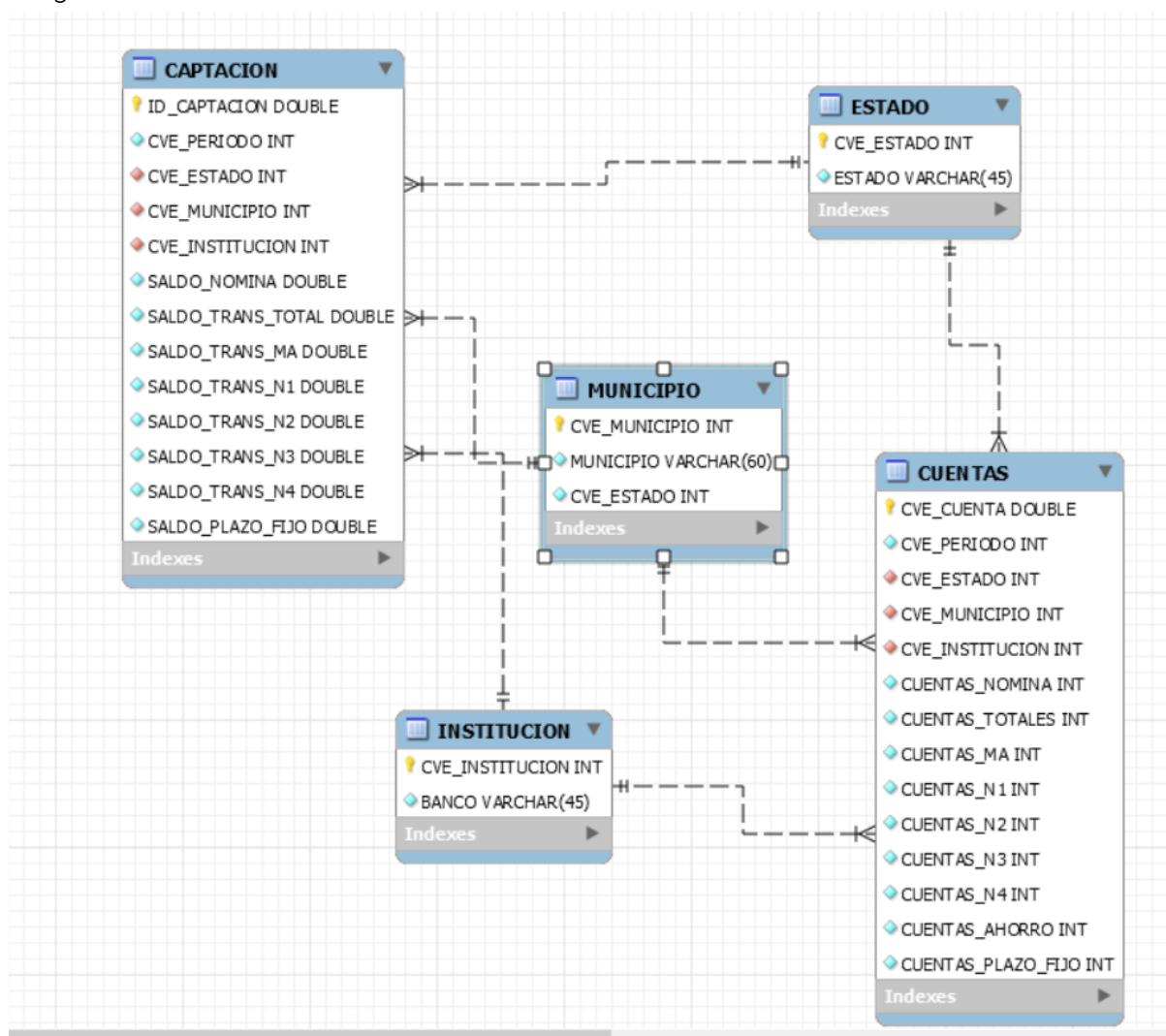
Cierta institución bancaria requiere conocer el mercado de las cuentas de captación y el saldo que hay por estado y/o municipio en México, además de saber cómo es la situación del resto de las instituciones bancarias

2.- Definición de la base de datos

Archivos utilizados:

- a) Cuentas.csv (Base con la información histórica de las cuentas de ahorro)
- b) Captación.csv (Base con la información histórica del saldo de las cuentas de ahorro)
- c) Municipio.csv (Base con nombres de los municipios)
- d) Estado.csv (Base con los nombres de los Estados)
- e) Institucion.csv (Base con los nombres de los bancos de México)

Diagrama entidad relación



Glosario

Base Cuentas	
CVE_CUENTA	Clave única de la base de cuentas
CVE_PERIODO	Clave del periodo que tiene el registro
CVE_ESTADO	Clave del estado al que pertenece el registro
CVE_MUNICIPIO	Clave del municipio al que pertenece el registro
CVE_INSTITUCION	Clave de la institución al que pertenece el registro
CUENTAS_NOMINA	Numero de cuentas de nomina
CUENTAS_TOTALES	Cuentas de ahorro transaccionales (N1+N2+N3+N4)
CUENTAS_MA	Cuentas de ahorro transaccionales que no tienen algún crédito con la institución
CUENTAS_NIVEL1	Cuentas transaccionales de nivel 1 (N1)
CUENTAS_NIVEL2	Cuentas transaccionales de nivel 2 (N2)
CUENTAS_NIVEL3	Cuentas transaccionales de nivel 3 (N3)
CUENTAS_NIVEL4	Cuentas transaccionales de nivel 4 (N4)
CUENTAS_AHORRO	Cuentas nivel 4 pertenecientes al grupo de la institución (subuniverso N4)
CUENTAS_PLAZOFIJO	Cuentas de plazo fijo

Base Captacion	
ID_CAPTACION	Clave única de la base de captación
CVE_PERIODO	Clave del periodo que tiene el registro
CVE_ESTADO	Clave del estado al que pertenece el registro
CVE_MUNICIPIO	Clave del municipio al que pertenece el registro
CVE_INSTITUCION	Clave de la institución al que pertenece el registro
SALDO_NOMINA	Saldo de cuentas de nomina
SALDO_TRANS_TOTAL	Saldo de ahorro transaccionales (N1+N2+N3+N4)
SALDO_TRANS_MA	Saldo cuentas de ahorro transaccionales que no tienen algún crédito con la institución
SALDO_TRANS_N1	Saldo de cuentas transaccionales de nivel 1 (N1)
SALDO_TRANS_N2	Saldo de cuentas transaccionales de nivel 2 (N2)
SALDO_TRANS_N3	Saldo de cuentas transaccionales de nivel 3 (N3)
SALDO_TRANS_N4	Saldo de cuentas transaccionales de nivel 4 (N4)
SALDO_AHORRO	Saldo nivel 4 pertenecientes al grupo de la institución (subuniverso N4)
SALDO_PLAZO_FIJO	Saldo de cuentas de plazo fijo

Base Estado	
CVE_ESTADO	Clave única de la base de Estado
ESTADO	Nombre del Estado de la República Mexicana

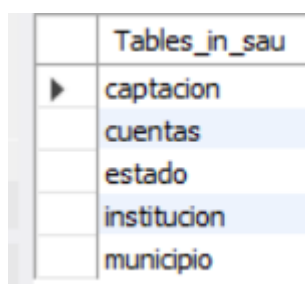
Base Institución	
CVE_INSTITUCION	Clave única de la base Institución
BANCO	Nombre de la Institución bancaria

Base Municipio	
CVE_MUNICIPIO	Clave única de la base Municipio
MUNICIPIO	Nombre del Municipio
CVE_ESTADO	Clave del estado al que pertenece el Municipio



Sesión 1: Fundamentos de SQL

1.1 Usando la base de datos SAU, mostrar las tablas que contienen y la estructura de la tabla "cuenta"

```
USE SA;
SHOW TABLES;
DESCRIBE CUENTAS;
```



	Tables_in_sau
▶	captacion
	cuentas
	estado
	institucion
	municipio

Result Grid						
Filter Rows:		Export:  Wrap Cell Content: 				
	Field	Type	Null	Key	Default	Extra
►	CVE_CUENTA	double	NO	PRI	NULL	
	CVE_PERIODO	int	NO		NULL	
	CVE_ESTADO	int	NO	MUL	NULL	
	CVE_MUNICIPIO	int	NO	MUL	NULL	
	CVE_INSTITUCION	int	NO	MUL	NULL	
	CUENTAS_NOMINA	int	NO		NULL	
	CUENTAS_TOTALES	int	NO		NULL	
	CUENTAS_MA	int	NO		NULL	
	CUENTAS_NIVEL1	int	NO		NULL	
	CUENTAS_NIVEL2	int	NO		NULL	
	CUENTAS_NIVEL3	int	NO		NULL	
	CUENTAS_NIVEL4	int	NO		NULL	


Result 19 x

Output

Action Output

1.2 ¿Cuáles son los municipios del estado 4?

```
SELECT *
FROM MUNICIPIO
WHERE num_estado = 4
```




Result Grid			
Filter Rows:		Edit: 	
	CVE_MUNICIPIO	MUNICIPIO	NUM_ESTADO
►	4001	Calkini	4
	4002	Campeche	4
	4003	Ciudad Del Carmen	4
	4004	Champoton	4
	4005	Hecelchakan	4
	4006	Hopelchen	4
	4007	Palizada	4
	4008	Tenabo	4
	4009	Escarcega	4
	4010	Calakmul	4
	4011	Candelaria	4
▲	NULL	NULL	NULL

MUNICIPIO 17 x

Output

1.3 ¿Cuáles son los bancos y que saldo total de cuentas transaccionales se tienen en mayo del 2020 (202005), que no tengan saldo de plazo fijo y que sean del municipio 2002?

```
SELECT
CVE_INSTITUCION,
SALDO_TRANS_TOTAL
from CAPTACION
WHERE CVE_PERIODO = 202005
AND CVE_MUNICIPIO = 2002
AND SALDO_PLAZO_FIJO > 0
```

Result Grid   Filter Rows: <input type="text"/> Export: 		
	CVE_INSTITUCION	SALDO_TRANS_TOTAL
▶	40002	7164755915
	40012	10844727208
	40014	4872665553
	40021	4002187483
	40030	555996400
	40036	449849133
	40044	2276675580
	40058	1241892451
	40072	7175954720
	40112	125885538
	40127	424413812
	40128	371677

CAPTACION 16 x

1.4 ¿Cuál es el top 9 de bancos que tienen más cuentas nivel 2 en abril y que sean del municipio 09014?

```
SELECT
CVE_INSTITUCION,
CUENTAS_NIVEL2
FROM CUENTAS
WHERE CVE_MUNICIPIO = 9014
AND CVE_PERIODO = 202004
order by CUENTAS_NIVEL2 desc limit 9
```

	CVE_INSTITUCION	CUENTAS_NIVEL2
▶	40012	2374759
	40072	29932
	40002	15393
	40130	11550
	40044	7398
	40014	3455
	40030	1607
	40128	352
	40036	243

CUENTAS 15 x

1.5 ¿Cuáles son los 20 municipios con menos saldo de plazo fijo del banco 40012 en enero del 2020?

```
select
CVE_MUNICIPIO,
SALDO_PLAZO_FIJO
from CAPTACION
WHERE CVE_INSTITUCION = 40012
AND CVE_PERIODO = 202001
AND SALDO_PLAZO_FIJO>0
ORDER BY SALDO_PLAZO_FIJO
limit 20
```

	CVE_MUNICIPIO	SALDO_PLAZO_FIJO
▶	12045	674000
	19018	1200000
	8062	1403971
	19041	3432294
	4007	3990588
	12040	4428884
	30085	5237457
	23009	6652807
	8052	6694567
	30116	7232000
	15023	8480232
	10031	8360006

CAPTACION 14 x

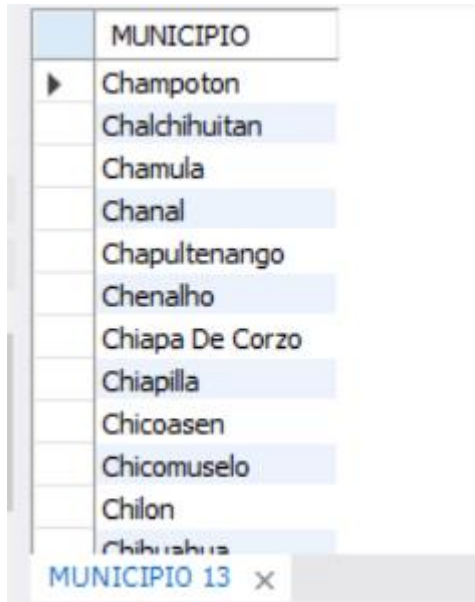
Output

Action Output

Sesión 2: Agrupaciones y subconsultas

2.1 ¿Cuáles son todos los municipios que empiezan con 'Ch'?

```
SELECT MUNICIPIO
FROM MUNICIPIO
WHERE MUNICIPIO LIKE 'Ch%'
```

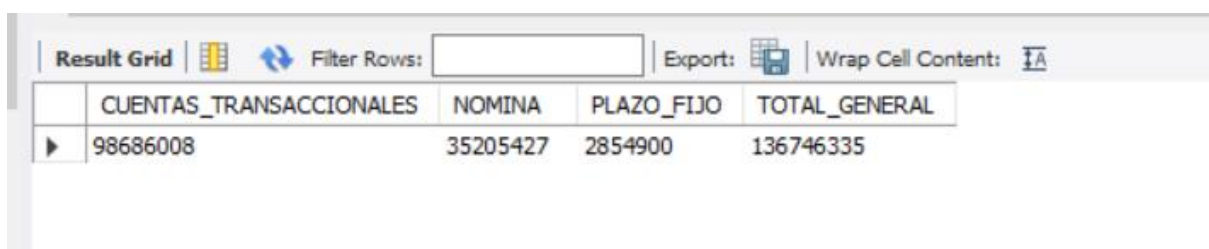


A screenshot of a database query result. It shows a table with one column labeled 'MUNICIPIO'. The table contains 13 rows of data, all starting with 'Ch'. The rows are: Champoton, Chalchihuitan, Chamula, Chanal, Chapultenango, Chenalho, Chiapa De Corzo, Chiapilla, Chicoasen, Chicomuselo, Chilon, Chihualpa, and Chihuahua. The table is displayed in a grid format with alternating light blue and white rows. At the bottom, there is a label 'MUNICIPIO 13' with a close button (X).

MUNICIPIO
Champoton
Chalchihuitan
Chamula
Chanal
Chapultenango
Chenalho
Chiapa De Corzo
Chiapilla
Chicoasen
Chicomuselo
Chilon
Chihualpa
Chihuahua

2.2 Obtener el total de cuentas (Nomina, plazo fijo y transaccionales) en mayo del 2020 y el total general de estas

```
select
SUM(CUENTAS_TOTALES) AS CUENTAS_TRANSACCIONALES,
SUM(CUENTAS_NOMINA) AS NOMINA,
SUM(CUENTAS_PLAZOFIJO) AS PLAZO_FIJO,
SUM(CUENTAS_TOTALES) + SUM(CUENTAS_NOMINA) + SUM(CUENTAS_PLAZOFIJO)
AS TOTAL_GENERAL
FROM
CUENTAS
WHERE CVE_PERIODO = 202005
```



A screenshot of a database query result. It shows a table with four columns: CUENTAS_TRANSACCIONALES, NOMINA, PLAZO_FIJO, and TOTAL_GENERAL. The table has one row of data. The values are: 98686008, 35205427, 2854900, and 136746335. The table is displayed in a grid format with alternating light blue and white rows. Above the table, there is a toolbar with buttons for 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content'.

	CUENTAS_TRANSACCIONALES	NOMINA	PLAZO_FIJO	TOTAL_GENERAL
▶	98686008	35205427	2854900	136746335

2.3 Obtener el total de saldo de plazo fijo que hay por institución en cada periodo y el numero de municipios en donde se encuentran

```

Select
CVE_PERIODO,
CVE_INSTITUCION,
SUM(SALDO_PLAZO_FIJO) AS SALDO_PLAZO_FIJO,
COUNT(CVE_MUNICIPIO) AS MUNICIPIO
from CAPTACION
WHERE SALDO_PLAZO_FIJO > 0
group by
CVE_PERIODO,
CVE_INSTITUCION

```

Result Grid				
Filter Rows:		Export:		
Wrap Cell Content:				
	CVE_PERIODO	CVE_INSTITUCION	SALDO_PLAZO_FIJO	MUNICIPIO
▶	202005	40002	166023400366	509
	202005	40012	295956010928	572
	202005	40014	286377777868	340
	202005	40021	174759957555	310
	202005	40030	87123335232	120
	202005	40036	23167715496	178
	202005	40042	21959984896	32
	202005	40044	195069318602	196
	202005	40058	54369422715	60
	202005	40072	298491385936	331
	202005	40112	29060051148	36
	202005	40127	19549540854	771

Result 40 x

Output

Action Output

2.4 ¿Cuál es el nombre las instituciones bancarias que no tienen saldo plazo fijo en el municipio 2002 o tienen cuenta nivel 3 en cualquiera de los periodos?

```

SELECT BANCO
FROM INSTITUCION
WHERE CVE_INSTITUCION IN
(select cve_institucion
FROM
CAPTACION
where SALDO_PLAZO_FIJO > 0
AND CVE_MUNICIPIO = 2002)
OR CVE_INSTITUCION IN
(select
cve_institucion
FROM CUENTAS
WHERE CUENTAS_NIVEL3 > 0)

```


BANCO
Banamex
BBVA Bancomer
Santander
HSBC
Banco del Bajío
Inbursa
Scotiabank
Banregio
Banorte
Monex
Banco Azteca
Autofin
INSTITUCION 10 x

2.5 ¿Cuál es el total de cuentas de Compartamos que se tienen por estado y en cuantos municipios tienen cuentas durante mayo?

```

SELECT
  CVE_ESTADO,
  SUM(CUENTAS_TOTALES) AS CUENTAS,
  COUNT(CVE_MUNICIPIO) AS MUNICIPIOS
FROM CUENTAS
WHERE CVE_PERIODO = 202005
AND CUENTAS_TOTALES > 0
AND CVE_INSTITUCION = (Select CVE_INSTITUCION
from INSTITUCION
WHERE BANCO LIKE '%compartamos%')
GROUP BY CVE_ESTADO

```

	CVE_ESTADO	CUENTAS	MUNICIPIOS
▶	1	532	1
	2	12703	5
	3	4425	3
	4	6311	5
	5	13997	12
	6	765	3
	7	20821	20
	8	4204	7
	9	22895	12
	10	11011	7
	11	7837	16
	12	19545	14

Result 59 ×

Sesión 3: Joins y Vistas

3.1 ¿Cuáles son los 3 bancos con más saldo de nómina en febrero 2020?

```
select
  B.BANCO,
  C.NOMINA AS SALDO
from
  (
    SELECT
      CVE_INSTITUCION,
      SUM(SALDO_NOMINA) AS NOMINA
    FROM
      CAPTACION
    WHERE CVE_PERIODO = 202002
    GROUP BY CVE_INSTITUCION
  ) AS C
LEFT JOIN INSTITUCION B
ON C.CVE_INSTITUCION = B.CVE_INSTITUCION
ORDER BY C.NOMINA DESC LIMIT 3
```

	BANCO	SALDO
▶	BBVA Bancomer	167539708646
	Banorte	76432707031
	Banamex	73524842186

3.2 ¿Cuántos bancos tienen al menos una cuenta N4 por municipio y estado al cierre de mayo?

```

SELECT
E.ESTADO,
M.MUNICIPIO,
BANCOS
FROM
(
SELECT
COUNT(CVE_INSTITUCION) AS BANCOS,
CVE_MUNICIPIO,
CVE_ESTADO
from CUENTAS
WHERE CUENTAS_TOTALES > 0
AND CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
) AS C
LEFT JOIN
MUNICIPIO AS M
ON M.CVE_MUNICIPIO = C.CVE_MUNICIPIO
LEFT JOIN
ESTADO AS E
ON E.CVE_ESTADO = C.CVE_ESTADO

```

	ESTADO	MUNICIPIO	BANCOS
▶	Aguascalientes	Aguascalientes	20
	Aguascalientes	Asientos	1
	Aguascalientes	Calvillo	5
	Aguascalientes	Cosio	1
	Aguascalientes	Jesus Maria (Aguascalientes)	4
	Aguascalientes	Pabellon De Arteaga	5
	Aguascalientes	Rincon De Romos	5
	Aguascalientes	San Jose De Gracia	1
	Aguascalientes	Tepezala	1
	Aguascalientes	El Llano	1
	Aguascalientes	San Francisco De Los Romo	2
	Baja California	Ensenada	18

Result 24 x

Output

3.3 Usando la consulta anterior obtener el saldo total que hay por cada municipio

```

SELECT
E.ESTADO,
M.MUNICIPIO,

```

```
BANCOS,
S.SALDO
FROM
(
SELECT
COUNT(CVE_INSTITUCION) AS BANCOS,
CVE_MUNICIPIO,
CVE_ESTADO
from CUENTAS
WHERE CUENTAS_TOTALES > 0
AND CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
) AS C
LEFT JOIN
MUNICIPIO AS M
ON M.CVE_MUNICIPIO = C.CVE_MUNICIPIO
LEFT JOIN
ESTADO AS E
ON E.CVE_ESTADO = C.CVE_ESTADO
LEFT JOIN
(SELECT
CVE_MUNICIPIO,
SUM(SALDO_TRANS_TOTAL) AS SALDO
from CAPTACION
WHERE CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
) AS S
ON S.CVE_MUNICIPIO = C.CVE_MUNICIPIO
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
ESTADO	MUNICIPIO	BANCOS	SALDO
Aguascalientes	Aguascalientes	20	33207134998
Aguascalientes	Asientos	1	605215
Aguascalientes	Calvillo	5	653178835
Aguascalientes	Cosio	1	71972
Aguascalientes	Jesus Maria (Aguascalientes)	4	187072481
Aguascalientes	Pabellon De Arteaga	5	466851427
Aguascalientes	Rincon De Romos	5	600247625
Aguascalientes	San Jose De Gracia	1	79014
Aguascalientes	Tepezala	1	78923018
Aguascalientes	El Llano	1	227337
Aguascalientes	San Francisco De Los Romo	2	39855319
Baja California	Ensenada	18	13060031500

Result 30 x

Output

Action Output

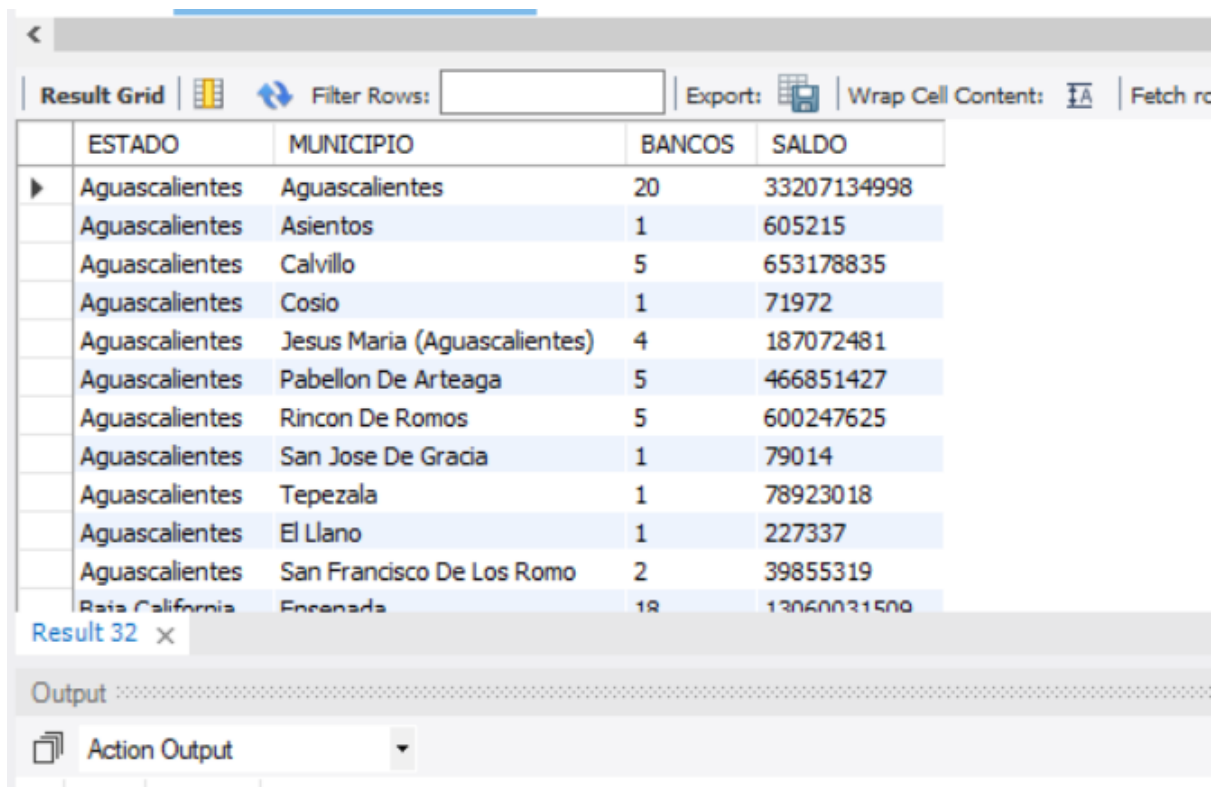
3.4 Hacer la consulta anterior usando vistas

```
CREATE VIEW TBANCOS_MAYO AS
(
SELECT
COUNT(CVE_INSTITUCION) AS BANCOS,
CVE_MUNICIPIO,
CVE_ESTADO
from CUENTAS
WHERE CUENTAS_TOTALES > 0
AND CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
);
```

```
CREATE VIEW TSALDO_MAYO AS
(
SELECT
CVE_MUNICIPIO,
SUM(SALDO_TRANS_TOTAL) AS SALDO
from CAPTACION
WHERE CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
```

);

```
SELECT
E.ESTADO,
M.MUNICIPIO,
BANCOS,
S.SALDO
FROM TBANCOS_MAYO AS C
LEFT JOIN
MUNICIPIO AS M
ON M.CVE_MUNICIPIO = C.CVE_MUNICIPIO
LEFT JOIN
ESTADO AS E
ON E.CVE_ESTADO = C.CVE_ESTADO
LEFT JOIN
TSALDO_MAYO
AS S
ON S.CVE_MUNICIPIO = C.CVE_MUNICIPIO;
```



The screenshot shows a database query result grid with the following data:

	ESTADO	MUNICIPIO	BANCOS	SALDO
▶	Aguascalientes	Aguascalientes	20	33207134998
	Aguascalientes	Asientos	1	605215
	Aguascalientes	Calvillo	5	653178835
	Aguascalientes	Cosio	1	71972
	Aguascalientes	Jesus Maria (Aguascalientes)	4	187072481
	Aguascalientes	Pabellon De Arteaga	5	466851427
	Aguascalientes	Rincon De Romos	5	600247625
	Aguascalientes	San Jose De Gracia	1	79014
	Aguascalientes	Tepezala	1	78923018
	Aguascalientes	El Llano	1	227337
	Aguascalientes	San Francisco De Los Romo	2	39855319
	Baja California	Ensenada	18	13060031500

Result 32 x

Output

Action Output

3.5 Con el ejercicio anterior, agregar el total de cuentas que había en mayo

```
CREATE VIEW TBANCOS_MAYO AS
(
```

```

SELECT
COUNT(CVE_INSTITUCION) AS BANCOS,
CVE_MUNICIPIO,
CVE_ESTADO
from CUENTAS
WHERE CUENTAS_TOTALES > 0
AND CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
);

```

```

CREATE VIEW TSALDO_MAYO AS
(
SELECT
CVE_MUNICIPIO,
SUM(SALDO_TRANS_TOTAL) AS SALDO
from CAPTACION
WHERE CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
);

```

```

CREATE VIEW TCUENTAS_MAYO AS
(
SELECT
CVE_MUNICIPIO,
SUM(CUENTAS_TOTALES) AS CUENTAS
from CUENTAS
WHERE CVE_PERIODO = 202005
GROUP BY
CVE_MUNICIPIO,
CVE_ESTADO
);

```

```

SELECT
E.ESTADO,
M.MUNICIPIO,
B.BANCOS,
S.SALDO,
C.CUENTAS
FROM TBANCOS_MAYO AS B
LEFT JOIN
MUNICIPIO AS M
ON M.CVE_MUNICIPIO = B.CVE_MUNICIPIO
LEFT JOIN

```

```

ESTADO AS E
ON E.CVE_ESTADO = B.CVE_ESTADO
LEFT JOIN
TSALDO_MAYO
AS S
ON S.CVE_MUNICIPIO = B.CVE_MUNICIPIO
LEFT JOIN
TCUENTAS_MAYO
AS C
ON C.CVE_MUNICIPIO = B.CVE_MUNICIPIO

```

	ESTADO	MUNICIPIO	BANCOS	SALDO	CUENTAS
▶	Aguascalientes	Aguascalientes	20	33207134998	1038046
	Aguascalientes	Asientos	1	605215	1322
	Aguascalientes	Calvillo	5	653178835	19593
	Aguascalientes	Cosio	1	71972	326
	Aguascalientes	Jesus Maria (Aguascalientes)	4	187072481	17899
	Aguascalientes	Pabellon De Arteaga	5	466851427	17837
	Aguascalientes	Rincon De Romos	5	600247625	23695
	Aguascalientes	San Jose De Gracia	1	79014	175
	Aguascalientes	Tepezala	1	78923018	4127
	Aguascalientes	El Llano	1	227337	420
	Aguascalientes	San Francisco De Los Romo	2	39855319	4584
	Baja California	Ensenada	18	13060031500	471517

Result 33 x

Output

Action Output

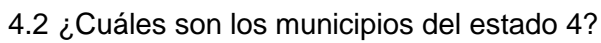
Sesión 4: Fundamentos de MongoDB

4.1 Mostrar los datos de cuentas totales y de instituciones

```

{
  CVE_INSTITUCION: 1,
  CUENTAS_TOTALES: 1
}

```

```
{
  CVE_ESTADO: 4
}
```

PROJECT

```
{
  _id: 0,
  MUNICIPIO: 1
}
```


The screenshot shows the MongoDB Compass interface. The top navigation bar includes 'Documents', 'Aggregations', 'Schema', 'Explain Plan', and 'Indexes'. The 'Documents' tab is active, displaying a list of documents. The query builder on the left includes the following stages:

- FILTER**: `{CVE_MUNICIPIO: 2002, SALDO_PLAZO_FIJO: { $gt: 0 }, CVE_PERIODO: 202005}`
- PROJECT**: `{_id: 0, CVE_INSTITUCION: 1, SALDO_TRANS_TOTAL: 1}`
- SORT**: `{SALDO_TRANS_TOTAL: -1}`
- AGGREGATION**: `{ $limit: 9 }`

The main panel displays the following documents (showing 1-18 of 18):

CVE_INSTITUCION	SALDO_TRANS_TOTAL
40002	7164755915
40012	10844727208
40014	4872665553
40021	4002187483
40030	555996400
40036	449849133

The right sidebar shows the query stages and a timestamp: 'Sun Jul 26 2020 13:16:25 GMT-0500 (ho..)'.

4.4 ¿Cuál es el top 9 de bancos que tienen más cuentas nivel 2 en abril y que sean del municipio 09014?

FILTER

```
{
  CVE_PERIODO: 202004,
  CVE_MUNICIPIO: 9014
}
```

PROJECT

```
{
  _id: 0,
  CVE_INSTITUCION: 1
}
```

SORT

```
{
  CUENTAS_NIVEL2: -1
}
```

LIMIT

9

Documents
Aggregations
Schema
Explain Plan
Indexes

FILTER
{CVE_PERIODO: 202004, CVE_MUNICIPIO: 09014}
OPTIONS

PROJECT
{ _id: 0, CVE_INSTITUCION: 1 }

SORT
{ CUENTAS_NIVEL2: -1 }
MAXTIMEMS
5000

COLLATION
SKIP
0
LIMIT
9

FIND
RESET
...

VIEW
{}

Displaying documents 1 - 9 of 9
REFRESH

CVE_INSTITUCION: 40012

CVE_INSTITUCION: 40072

CVE_INSTITUCION: 40002

CVE_INSTITUCION: 40130

CVE_INSTITUCION: 40044

CVE_INSTITUCION: 40014

CVE_INSTITUCION: 40030

CVE_INSTITUCION: 40128

Sun Jul 26 2020 13:21:03 GMT-0500 (ho...)

FILTER
{
CVE_PERIODO: 202004,
CVE_MUNICIPIO: 9014
}

PROJECT
{
_id: 0,
CVE_INSTITUCION: 1
}

SORT
{
CUENTAS_NIVEL2: -1
}

LIMIT
9

Sun Jul 26 2020 13:20:46 GMT-0500 (ho...)

FILTER
{
CVE_PERIODO: 202004,
CVE_MUNICIPIO: 9014
}

PROJECT
{
CVE_INSTITUCION: 1
}

SORT

4.5 ¿Cuáles son los 20 municipios con menor saldo de plazo fijo del banco 40012 en enero del 2020?

```

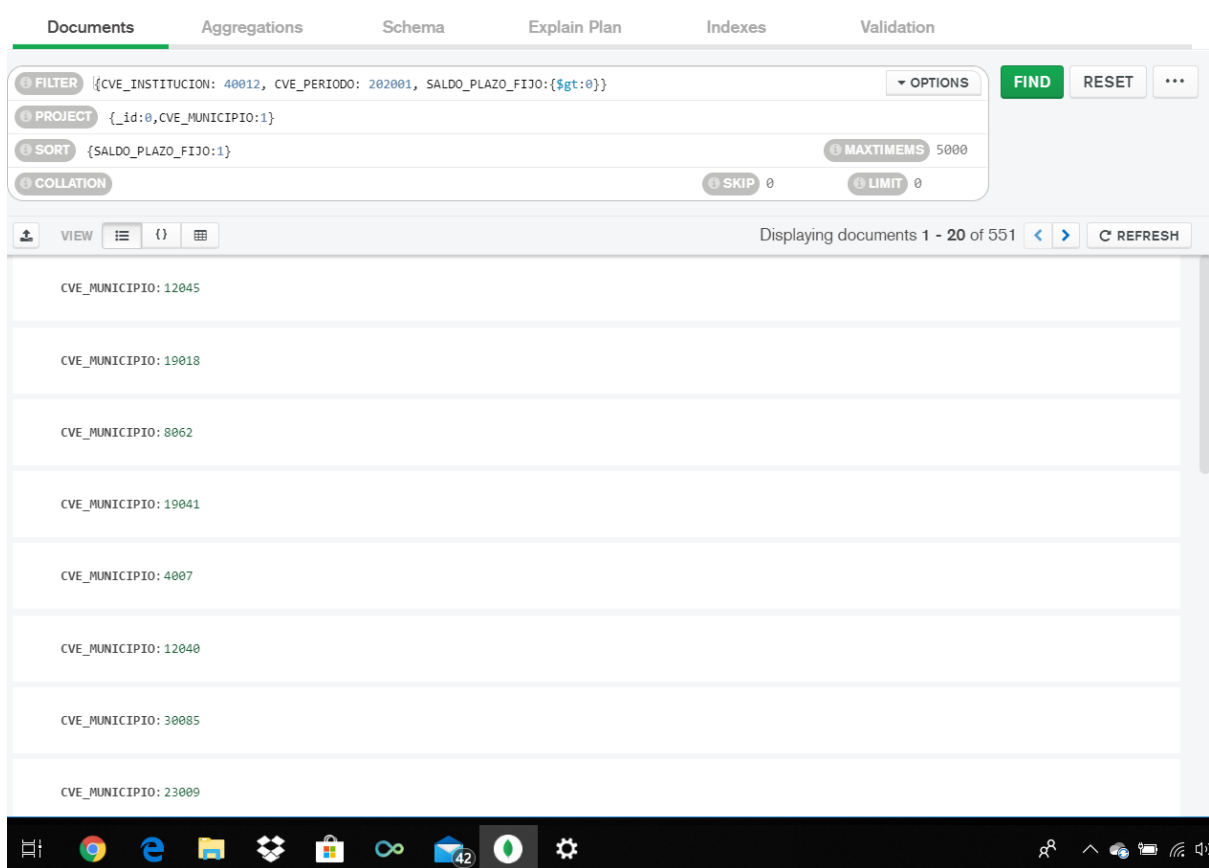
FILTER
{
  CVE_PERIODO: 202004,
  CVE_MUNICIPIO: 9014
}

PROJECT
{
  _id: 0,
  CVE_INSTITUCION: 1
}

SORT
{
  CUENTAS_NIVEL2: -1
}

LIMIT
9

```



Sesión 5: Consultas en Mongo DB

5.1 Obtener todos los municipios que empiezen con 'Ch'

FILTER

```
{
  MUNICIPIO: RegExp('^Ch', i)
}
```

PROJECT

```
{
  _id: 0,
  MUNICIPIO: 1
}
```

The screenshot shows the MongoDB Compass interface. The main window displays a list of documents with the following 'MUNICIPIO' values: "Champoton", "Chalchihuitan", "Chamula", "Chanal", "Chapultenango", "Chenalho", and "Chiapa De Corzo". The right sidebar shows the query stages:

```

FILTER
{
  MUNICIPIO: RegExp('^Ch', 1)
}

PROJECT
{
  _id: 0,
  MUNICIPIO: 1
}

```

Below the project stage, the execution time is shown: "Sun Jul 26 2020 13:56:50 GMT-0500 (ho...".

5.2 Mostrar la institución, estado y municipio de aquellos bancos que tengan cuentas n1 o n2 en enero 2020

```

[{$match: {
  $or: [
    {CUNETAS_NIVEL1: {$gt: 0}}
  ],
  {CUENTAS_NIVEL2: {$gt: 0}}
}
],
  CVE_PERIODO: 202005
}}, {$project: {_id: 0,
  CVE_ESTADO: 1,
  CVE_MUNICIPIO: 1,
  CVE_INSTITUCION: 1
}}]

```


Output after **\$project** stage ⓘ (Sample of 1 document)

```

1 /**
2  * specifications: The fields to
3  * include or exclude.
4  */
5 {
6   _id: 0,
7   NOMINA: 1,
8   CUENTAS_TRANS: 1,
9   CUENTAS_PLAZO_FIJO: 1,
10  TOTAL: { $add: [ '$NOMINA', '$CUENTAS_TRANS', '$CUENTAS'
11 }

```

```

NOMINA: 35205427
CUENTAS_TRANS: 98686006
CUENTAS_PLAZO_FIJO: 2854900
TOTAL: 136746333

```

5.4 ¿Cuál es el total de saldo de plazo fijo que hay por institución en cada periodo y en cuantos municipios se encuentran?

```

[{$match: {
  SALDO_PLAZO_FIJO: {$gt:0}
}}, {$group: {
  _id: {PERIODO: '$CVE_PERIODO', BANCO: '$CVE_INSTITUCION'},
  PLAZO_FIJO: { $sum: '$SALDO_PLAZO_FIJO'
  },
  MUNICIPIO: {$sum: 1}
}}]

```

Output after **\$group** stage ⓘ (Sample of 20 documents)

```

1 /**
2  * id: The id of the group.
3  * fieldN: The first field name.
4  */
5 {
6   id: {PERIODO: '$CVE_PERIODO', BANCO: '$CVE_INSTITUCION'},
7   PLAZO_FIJO: { $sum: '$SALDO_PLAZO_FIJO'
8   },
9   MUNICIPIO: {$sum: 1}
10 }

```

```

_id: Object
PLAZO_FIJO: 12649788
MUNICIPIO: 1

```

```

_id: Object
PLAZO_FIJO: 903220
MUNICIPIO: 1

```

ADD STAGE

5.5 Obtener el saldo de las cuentas N3 que hay por estado de cada institución en enero del 2020

```

[ {
  $match: {
    CVE_PERIODO: 202001
  }
}, {
  $group: {
    _id: {
      ESTADO: '$CVE_ESTADO',
      INSTITUCION: '$CVE_INSTITUCION'

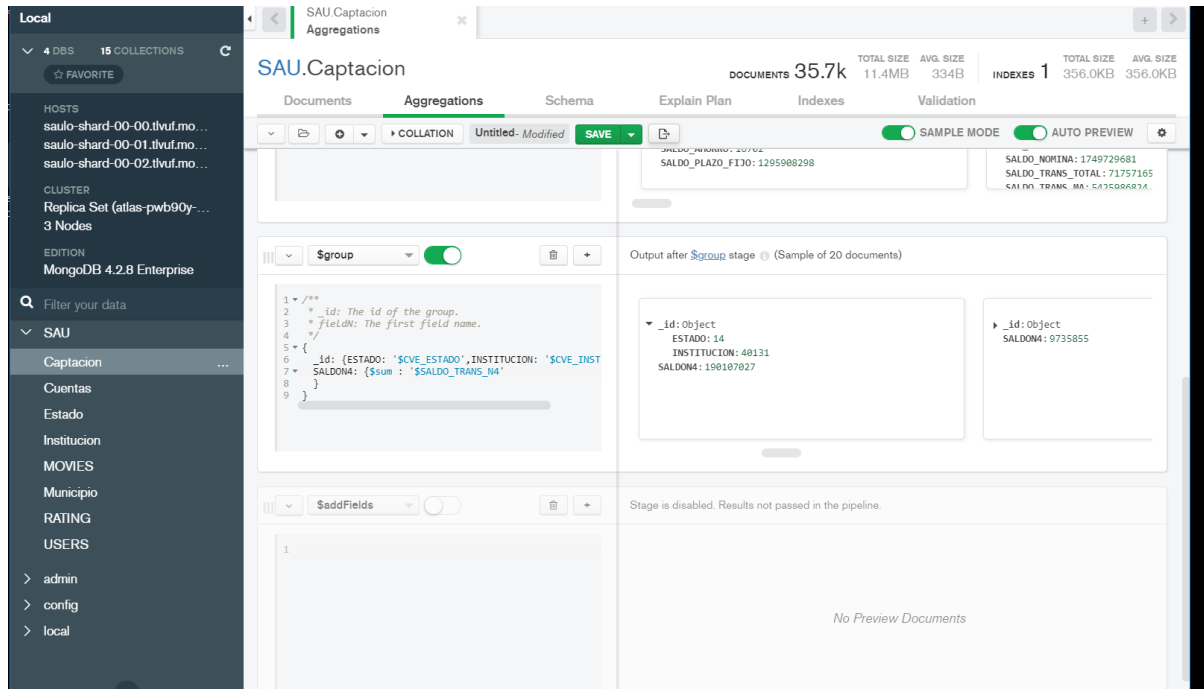
```



```

    },
    SALDON4: {
      $sum: '$SALDO_TRANS_N4'
    }
  }
}
}}

```



Sesión 6: Agregaciones

6.1 Obtener el promedio de cuentas por institución que hay por cada estado, en mayo

```

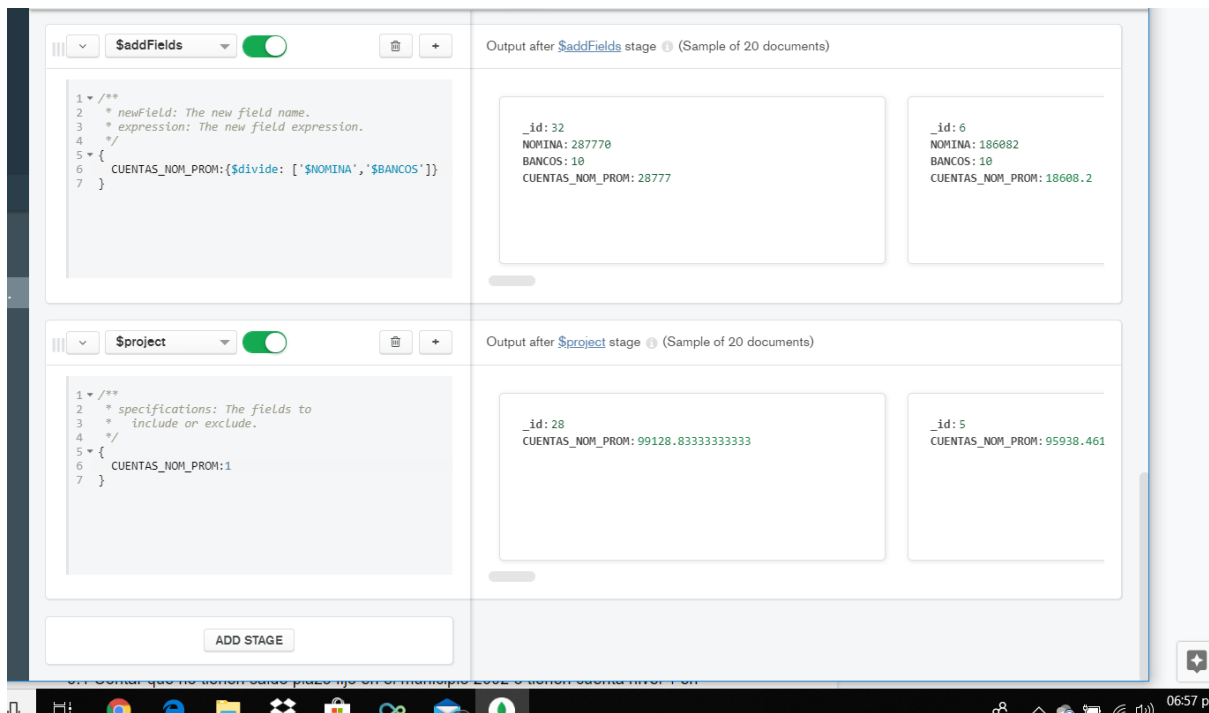
[{$match: {
  CVE_PERIODO: 202005,
  CUENTAS_NOMINA: {$gt:0}
}}, {$group: {
  _id: {ESTADO: '$CVE_ESTADO', INSTITUCION: '$CVE_INSTITUCION'},
  NOMINA: {
    $sum: '$CUENTAS_NOMINA'
  }
}}, {$addFields: {
  CVE_ESTADO: '$_id.ESTADO'
}}, {$group: {
  _id: '$CVE_ESTADO' ,
  NOMINA: {
    $sum: '$NOMINA'
  },
  BANCOS: {$sum: 1}
}}

```

```

}}, {$addFields: {
  CUENTAS_NOM_PROM: {$divide: ['$NOMINA', '$BANCOS']}
}}, {$project: {
  CUENTAS_NOM_PROM: 1
}}

```

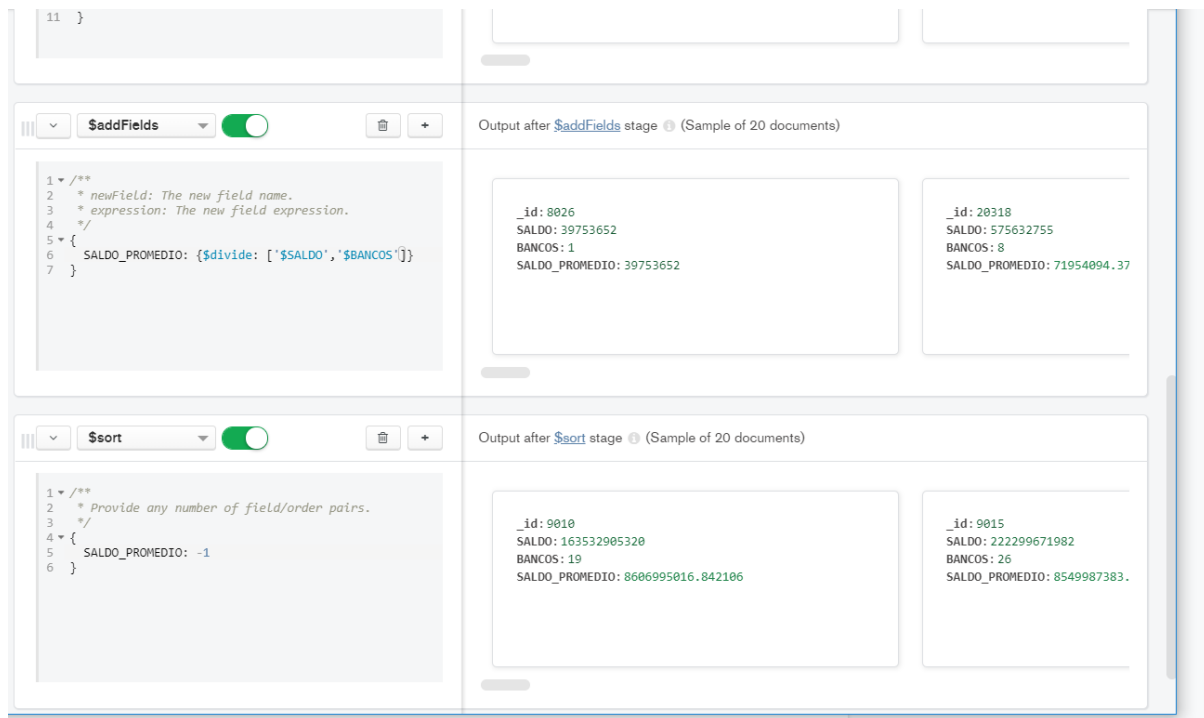


6.2 ¿Cuál es el ranking de municipios por el saldo promedio de las instituciones de plazo fijo en febrero 2020?

```

[{$match: {
  SALDO_PLAZO_FIJO: {$gt: 0},
  CVE_PERIODO: 202002
}}, {$group: {
  _id: '$CVE_MUNICIPIO',
  SALDO: {
    $sum: '$SALDO_PLAZO_FIJO'
  },
  BANCOS: {$sum: 1}
}}, {$addFields: {
  SALDO_PROMEDIO: {$divide: ['$SALDO', '$BANCOS']}
}}, {$sort: {
  SALDO_PROMEDIO: -1
}}

```



6.3 ¿Cuál es el total de cuentas totales (sin nomina ni plazo fijo) de Compartamos en mayo que se tienen por estado y en cuantos municipios esta?

```
[{$lookup: {
  from: 'Institucion',
  localField: 'CVE_INSTITUCION',
  foreignField: 'CVE_INSTITUCION',
  as: 'BANCO_A'
}}, {$addFields: {
  BANCO_OBJETO: {$arrayElemAt: ["$BANCO_A", 0]}
}}, {$addFields: {
  BANCO: '$BANCO_OBJETO.BANCO'
}}, {$match: {
  BANCO: /compartamos/i,
  CVE_PERIODO: 202005
}}, {$group: {
  _id: '$CVE_ESTADO',
  CUENTAS: {
    $sum: '$CUENTAS_TOTALES'
  },
  MUNICIPIO: {$sum: 1}
}}]
```

```

1 ▾ /**
2   * query: The query in MQL.
3   */
4 ▾ {
5   BANCO: /compartamos/i,
6   CVE_PERIODO: 202005
7 }

```

```

CVE_ESTADO: 1
CVE_MUNICIPIO: 1001
CVE_INSTITUCION: 40130
CUENTAS_NOMINA: 0
CUENTAS_TOTALES: 532
CUENTAS_MA: 532
CUENTAS_NIVEL1: 0
CUENTAS_NIVEL2: 7
CUENTAS_NIVEL3: 0
CUENTAS_NIVEL4: 525

```

```

_id: ObjectId("5f1dc06a26c8...")
CVE_CUENTA: 1
CVE_PERIODO: 202005
CVE_ESTADO: 2
CVE_MUNICIPIO: 2001
CVE_INSTITUCION: 40130
CUENTAS_NOMINA: 0
CUENTAS_TOTALES: 2254
CUENTAS_MA: 2254

```

▾

\$group

🔴

🗑️

+

Output after \$group stage ⓘ (Sample of 20 documents)

```

1 ▾ /**
2   * id: The id of the group.
3   * fieldN: The first field name.
4   */
5 ▾ {
6   _id: '$CVE_ESTADO',
7   CUENTAS: {
8     $sum: '$CUENTAS_TOTALES'
9   },
10  MUNICIPIO: {$sum: 1}
11 }

```

```

_id: 11
CUENTAS: 7837
MUNICIPIO: 16

```

```

_id: 25
CUENTAS: 6415
MUNICIPIO: 6

```

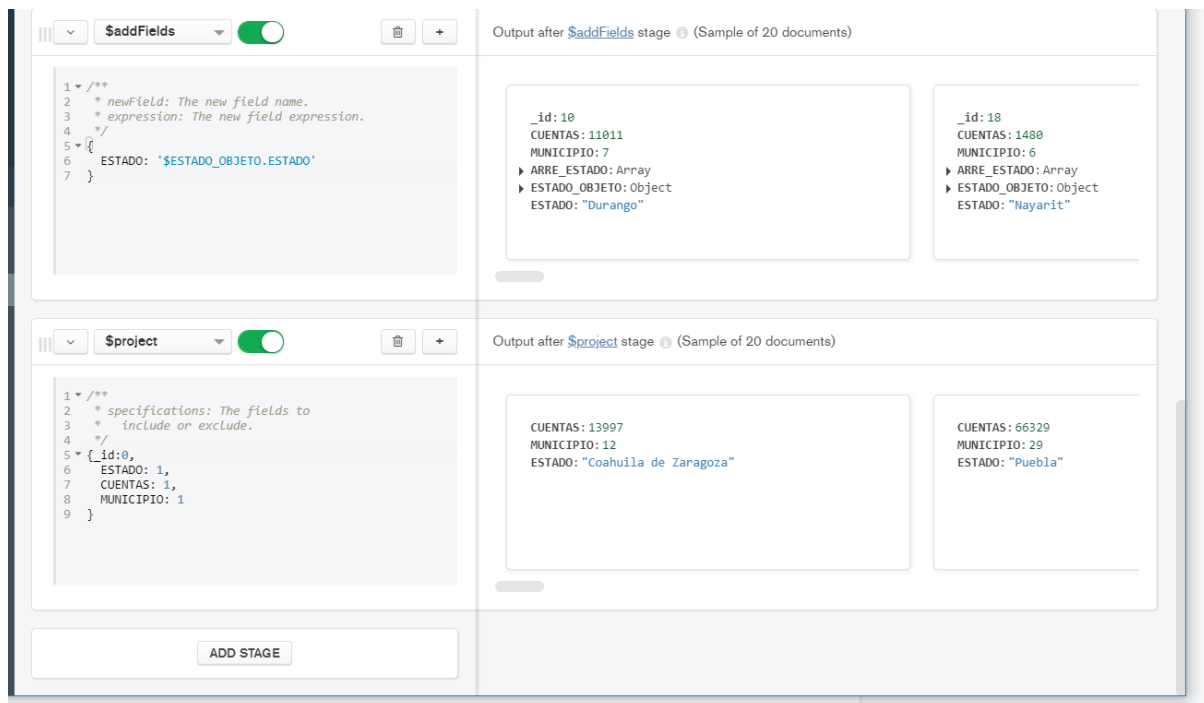
ADD STAGE

6.4 Usando el ejercicio anterior, crear una vista y poner el nombre del estado

```

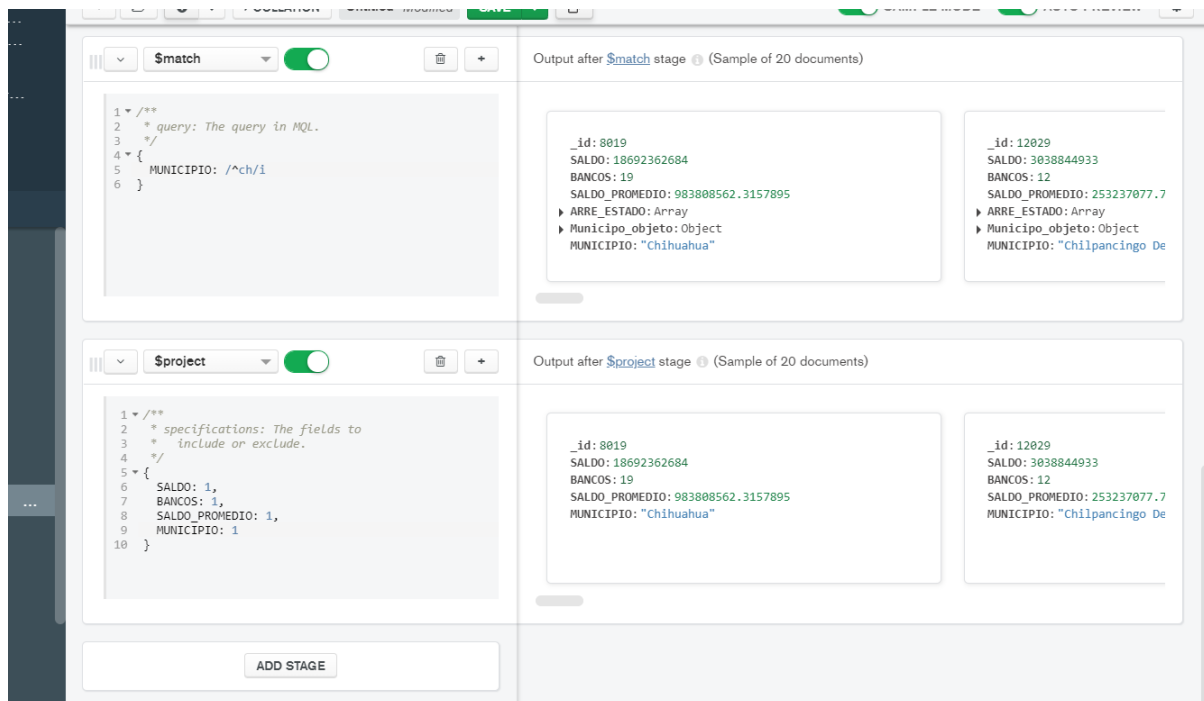
[{$lookup: {
  from: 'Estado',
  localField: '_id',
  foreignField: 'CVE_ESTADO',
  as: 'ARRE_ESTADO'
}}, {$addFields: {
  ESTADO_OBJETO: {$arrayElemAt: ['$ARRE_ESTADO',0]}
}}, {$addFields: {
  ESTADO: '$ESTADO_OBJETO.ESTADO'
}}, {$project: {_id:0,
  ESTADO: 1,
  CUENTAS: 1,
  MUNICIPIO: 1
}}]

```



6.5 Hacer una vista del ejercicio 6.2 y poner el nombre del municipio y filtrar para mostrar los municipios que empiezan con ch

```
[{$lookup: {
  from: 'Municipio',
  localField: '_id',
  foreignField: 'CVE_MUNICIPIO',
  as: 'ARRE_ESTADO'
}}, {$addFields: {
  Municipio_objeto: {$arrayElemAt: ['$ARRE_ESTADO', 0]}
}}, {$addFields: {
  MUNICIPIO: '$Municipio_objeto.MUNICIPIO'
}}, {$match: {
  MUNICIPIO: /^ch/i
}}, {$project: {
  SALDO: 1,
  BANCOS: 1,
  SALDO_PROMEDIO: 1,
  MUNICIPIO: 1
}}]
```



Conclusiones

- 1) SQL workbench se vuelve lenta en la carga de información, además de que se tiene que hacer las relaciones entre tablas y definir las llaves foráneas, pero una vez terminado este proceso, es muy fácil el hacer consultas que tengan varias relaciones
- 2) Mongo es una herramienta que puede funcionar muy bien para la carga de información y si se tiene todos los registros en una sola tabla es muy fácil poder consultar, sin embargo, tener diferentes tablas se vuelve complicado el hacer el lookup entre ellas, debido a que se generan objetos con arreglos de datos, por lo que se tienen que hacer más procesos para convertir estos en campos de la tabla
- 3) Debido a que esta información de la banca, ya está estructurada para ser usada como una base de datos relacional, optaría por hacer todos los procesos en SQL