
UNIVERSIDAD DE GRANADA

MASTER PROFESIONAL EN INGENIERÍA INFORMÁTICA

PRÁCTICA 2

MongoDB

Autor:

Manuel Jesús García Manday
(nickter@correo.ugr.es)

Master en Ingeniería Informática

5 de mayo de 2017

Índice

1. Objetivo.	3
2. Consulta de Documentos.	3
2.1. Visualiza la colección pedidos y familiarízate con ella. Observa los distintos tipos de datos y sus estructuras dispares.	3
2.2. Visualiza sólo el primer documento de la colección. Utiliza los métodos <code>.limit()</code> y <code>.findOne()</code> . . .	9
2.3. Visualiza el cliente con el <code>id_cliente = 2222</code>	11
2.4. Visualiza los clientes que hayan pedido algún producto de más de 94 euros.	13
2.5. Visualiza los clientes de Jaén o Salamanca (excluye los datos de los pedidos). Utiliza los operadores <code>\$or</code> e <code>\$in</code>	13
2.6. Visualiza los clientes que no tienen campo pedido.	14
2.7. Visualiza los clientes que hayan nacido en 1963.	14
2.8. Visualiza los clientes que hayan pedido algún producto fabricado por Canon y algún producto cuyo precio sea inferior a 15 euros.	15
2.9. Datos personales (<code>id_cliente</code> , Nombre, Dirección, Localidad y <code>Fnacimiento</code>) de los clientes cuyo nombre empieza por la cadena <code>ç</code> (No distinguir entre mayúsculas y minúsculas).	19
2.10. Visualiza los datos personales de los clientes (excluyendo <code>_id</code>). Limita los documentos a 4.	20
2.11. Ídem anterior pero ordenando los documentos por Localidad (ascendente) e <code>id_cliente</code> (descendente).	21
3. Agregación.	22
3.1. N° total de clientes.	22
3.2. N° total de clientes de Jaén	22
3.3. Facturación total clientes por localidad.	23
3.4. Facturación media de clientes por localidad para las localidades distintas a Jaén con facturación media mayor de 5000. Ordenación por Localidad descendente. Eliminar el <code>_id</code> y poner el nombre en mayúsculas.	23
3.5. Calcula la cantidad total facturada por cada cliente (uso de "unwind").	23
4. MapReduce.	23
4.1. Encontrar las ciudades más cercanas sobre la colección recién creada mediante un enfoque MapReduce conforme a los pasos que se ilustran en el tutorial práctico.	24
4.2. ¿Cómo podríamos obtener las ciudades más distantes en cada país?.	27
4.3. ¿Qué ocurre si en un país hay dos parejas de ciudades que están a la misma distancia mínima? ¿Cómo harías para que aparecieran todas?.	31
4.4. ¿Cómo podríamos obtener adicionalmente la cantidad de parejas de ciudades evaluadas para cada país consultado?.	37
4.5. ¿Cómo podríamos calcular la distancia media entre las ciudades de cada país?.	40
4.6. ¿Mejoraría el rendimiento si creamos un índice? ¿Sobre qué campo? Comprobadlo.	44

1. Objetivo.

El objetivo de esta práctica es familiarizarse con el uso de un sistema de gestión de bases de datos en entornos Big Data. Para ello haremos uso de la aplicación más conocida como es MongoDB.

2. Consulta de Documentos.

Crear la colección pedidos sobre la que se realizarán diversas operaciones CRUD.

2.1. Visualiza la colección pedidos y familiarízate con ella. Observa los distintos tipos de datos y sus estructuras dispares.



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.find().pretty();
{
  "_id" : ObjectId("59087944fd111cefb7a5ec8"),
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romeros 14",
  "Localidad" : "Sevilla",
  "Fnacimiento" : ISODate("1963-04-03T00:00:00Z"),
  "Facturacion" : 5000,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 390,
          "Cantidad" : 1
        },
        {
          "id_producto" : 2,
          "Nombre" : "Tablet 8 pulgadas",
          "Precio_unidad" : 95,
          "Cantidad" : 1
        }
      ]
    }
  ]
},
```

Figura 1: Visualizando la colección (I).

```

    {
      "id_pedido" : 2,
      "Productos" : [
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 115,
          "Cantidad" : 3
        }
      ]
    }
  ]
}
{
  "_id" : ObjectId("59087945fd111cefbd7a5ec9"),
  "id_cliente" : 2222,
  "Nombre" : "Juan Gomez",
  "Direccion" : "Perpetuo Socorro 9",
  "Localidad" : "Salamanca",
  "Fnacimiento" : ISODate("1960-08-17T00:00:00Z"),
  "Facturacion" : 6500,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 100,
          "Cantidad" : 1
        }
      ],
    }
  ]
}

```

Figura 2: Visualizando la colección (II).

```
        {
            "id_producto" : 42,
            "Nombre" : "Portatil ASM Mod. 254",
            "Fabricante" : "Intel",
            "Precio_unidad" : 455,
            "Cantidad" : 2
        },
        {
            "id_producto" : 27,
            "Nombre" : "Cable USB",
            "Precio_unidad" : 11,
            "Cantidad" : 12
        }
    ]
},
{
    "id_pedido" : 2,
    "Productos" : [
        {
            "id_producto" : 77,
            "Nombre" : "Impresora Laser",
            "Fabricante" : "Canon",
            "Precio_unidad" : 128,
            "Cantidad" : 3
        },
        {
            "id_producto" : 42,
            "Nombre" : "Portatil ASM Mod. 254",
            "Fabricante" : "Intel",
            "Precio_unidad" : 451,
            "Cantidad" : 5
        },
    ],
}
```

Figura 3: Visualizando la colección (III).

```
    {
      "id_producto" : 42,
      "Nombre" : "Portatil ASM Mod. 254",
      "Fabricante" : "Intel",
      "Precio_unidad" : 451,
      "Cantidad" : 5
    },
    {
      "id_producto" : 21,
      "Nombre" : "Disco Duro 500GB",
      "Precio_unidad" : 99,
      "Cantidad" : 10
    }
  ]
},
{
  "id_pedido" : 3,
  "Productos" : [
    {
      "id_producto" : 1,
      "Nombre" : "Pentium IV",
      "Fabricante" : "Intel",
      "Precio_unidad" : 94,
      "Cantidad" : 5
    },
    {
      "id_producto" : 95,
      "Nombre" : "SAI 5H Mod. 258",
      "Precio_unidad" : 213,
      "Cantidad" : 2
    },
  ],
}
```

Figura 4: Visualizando la colección (IV).

```
        {
            "id_producto" : 21,
            "Precio_unidad" : 66,
            "Nombre" : "Disco Duro 500GB",
            "Cantidad" : 10
        }
    ]
}
}
{
    "_id" : ObjectId("59087945fd111cefb7a5eca"),
    "id_cliente" : 3333,
    "Nombre" : "Carlos Montes",
    "Direccion" : "Salsipuedes 13",
    "Localidad" : "Jaen",
    "Fnacimiento" : ISODate("1967-11-25T00:00:00Z"),
    "Facturacion" : 8000
}
{
    "_id" : ObjectId("59087945fd111cefb7a5ecb"),
    "id_cliente" : 4444,
    "Nombre" : "Carmelo Coton",
    "Direccion" : "La Luna 103",
    "Localidad" : "Jaen",
    "Fnacimiento" : ISODate("1969-01-06T00:00:00Z"),
    "Facturacion" : 12300
}
```

Figura 5: Visualizando la colección (V).

```

{
  "_id" : ObjectId("59087945fd111cefb7a5ecc"),
  "id_cliente" : 5555,
  "Nombre" : "Cristina Miralles",
  "Direccion" : "San Fernando 28",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1970-07-12T00:00:00Z"),
  "Facturacion" : 16500,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 95,
          "Nombre" : "SAI 5H Mod. 258",
          "Precio_unidad" : 211,
          "Cantidad" : 2
        },
        {
          "id_producto" : 42,
          "Nombre" : "Portatil ASM Mod. 254",
          "Precio_unidad" : 460,
          "Fabricante" : "Intel",
          "Cantidad" : 2
        },
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 119,
          "Cantidad" : 2
        }
      ]
    }
  ]
}

```

Figura 6: Visualizando la colección (VI).

```

    }
  ]
}
{
  "_id" : ObjectId("59087945fd111cefb7a5ecd"),
  "id_cliente" : 6666,
  "Nombre" : "Chema Pamundi",
  "Direccion" : "Recogidas 54",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1969-02-04T00:00:00Z"),
  "Facturacion" : 5000
}
{
  "_id" : ObjectId("59087945fd111cefb7a5ece"),
  "id_cliente" : 777,
  "Nombre" : "Alberto Matero",
  "Direccion" : "Pelayo 4",
  "Localidad" : "Sevilla",
  "Facturacion" : 2500,
  "Pedidos" : null
}

```

Figura 7: Visualizando la colección (VII).

2.2. Visualiza sólo el primer documento de la colección. Utiliza los métodos `.limit()` y `.findOne()` .



```

CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
> db.pedidos.findOne();
{
  "_id" : ObjectId("59087944fd111cefb7a5ec8"),
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romeros 14",
  "Localidad" : "Sevilla",
  "Fnacimiento" : ISODate("1963-04-03T00:00:00Z"),
  "Facturacion" : 5000,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 390,
          "Cantidad" : 1
        },
        {
          "id_producto" : 2,
          "Nombre" : "Tablet 8 pulgadas",
          "Precio_unidad" : 95,
          "Cantidad" : 1
        }
      ]
    },
    {
      "id_pedido" : 2,
      "Productos" : [
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 115,
          "Cantidad" : 3
        }
      ]
    }
  ]
}

```

Figura 8: Visualizando el primer documento (I).



A terminal window titled "CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es" displays the following command and its output:

```
> db.pedidos.find().limit(1).pretty();
{
  "_id" : ObjectId("59087944fd111cefb7a5ec8"),
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romero 14",
  "Localidad" : "Sevilla",
  "Fnacimiento" : ISODate("1963-04-03T00:00:00Z"),
  "Facturacion" : 5000,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 390,
          "Cantidad" : 1
        },
        {
          "id_producto" : 2,
          "Nombre" : "Tablet 8 pulgadas",
          "Precio_unidad" : 95,
          "Cantidad" : 1
        }
      ]
    },
    {
      "id_pedido" : 2,
      "Productos" : [
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 115,
          "Cantidad" : 3
        }
      ]
    }
  ]
}
```

Figura 9: Visualizando el primer documento (II).

2.3. Visualiza el cliente con el id_cliente = 2222 .



A terminal window titled "CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es" displays the output of a MongoDB query. The query is `db.pedidos.find({id_cliente: 2222}).pretty();`. The output is a JSON document representing a client and their orders.

```
[> db.pedidos.find({id_cliente: 2222}).pretty();
{
  "_id" : ObjectId("59087945fd111cefb7a5ec9"),
  "id_cliente" : 2222,
  "Nombre" : "Juan Gomez",
  "Direccion" : "Perpetuo Socorro 9",
  "Localidad" : "Salamanca",
  "Fnacimiento" : ISODate("1960-08-17T00:00:00Z"),
  "Facturacion" : 6500,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 100,
          "Cantidad" : 1
        },
        {
          "id_producto" : 42,
          "Nombre" : "Portatil ASM Mod. 254",
          "Fabricante" : "Intel",
          "Precio_unidad" : 455,
          "Cantidad" : 2
        },
        {
          "id_producto" : 27,
          "Nombre" : "Cable USB",
          "Precio_unidad" : 11,
          "Cantidad" : 12
        }
      ]
    }
  ]
},
{
  ...
}
```

Figura 10: Visualizando los datos del cliente (I).

```

{
  "id_pedido" : 2,
  "Productos" : [
    {
      "id_producto" : 77,
      "Nombre" : "Impresora Laser",
      "Fabricante" : "Canon",
      "Precio_unidad" : 128,
      "Cantidad" : 3
    },
    {
      "id_producto" : 42,
      "Nombre" : "Portatil ASM Mod. 254",
      "Fabricante" : "Intel",
      "Precio_unidad" : 451,
      "Cantidad" : 5
    },
    {
      "id_producto" : 21,
      "Nombre" : "Disco Duro 500GB",
      "Precio_unidad" : 99,
      "Cantidad" : 10
    }
  ]
},

```

Figura 11: Visualizando los datos del cliente (II).

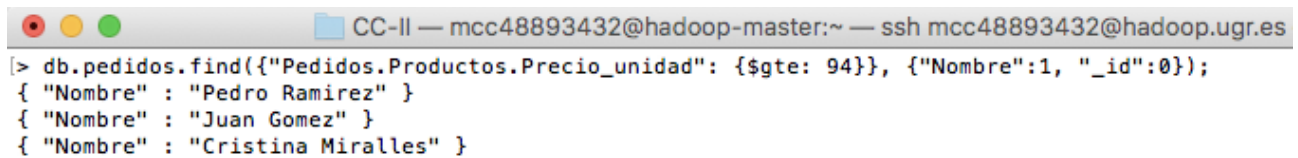
```

{
  "id_pedido" : 3,
  "Productos" : [
    {
      "id_producto" : 1,
      "Nombre" : "Pentium IV",
      "Fabricante" : "Intel",
      "Precio_unidad" : 94,
      "Cantidad" : 5
    },
    {
      "id_producto" : 95,
      "Nombre" : "SAI 5H Mod. 258",
      "Precio_unidad" : 213,
      "Cantidad" : 2
    },
    {
      "id_producto" : 21,
      "Precio_unidad" : 66,
      "Nombre" : "Disco Duro 500GB",
      "Cantidad" : 10
    }
  ]
}

```

Figura 12: Visualizando los datos del cliente (III).

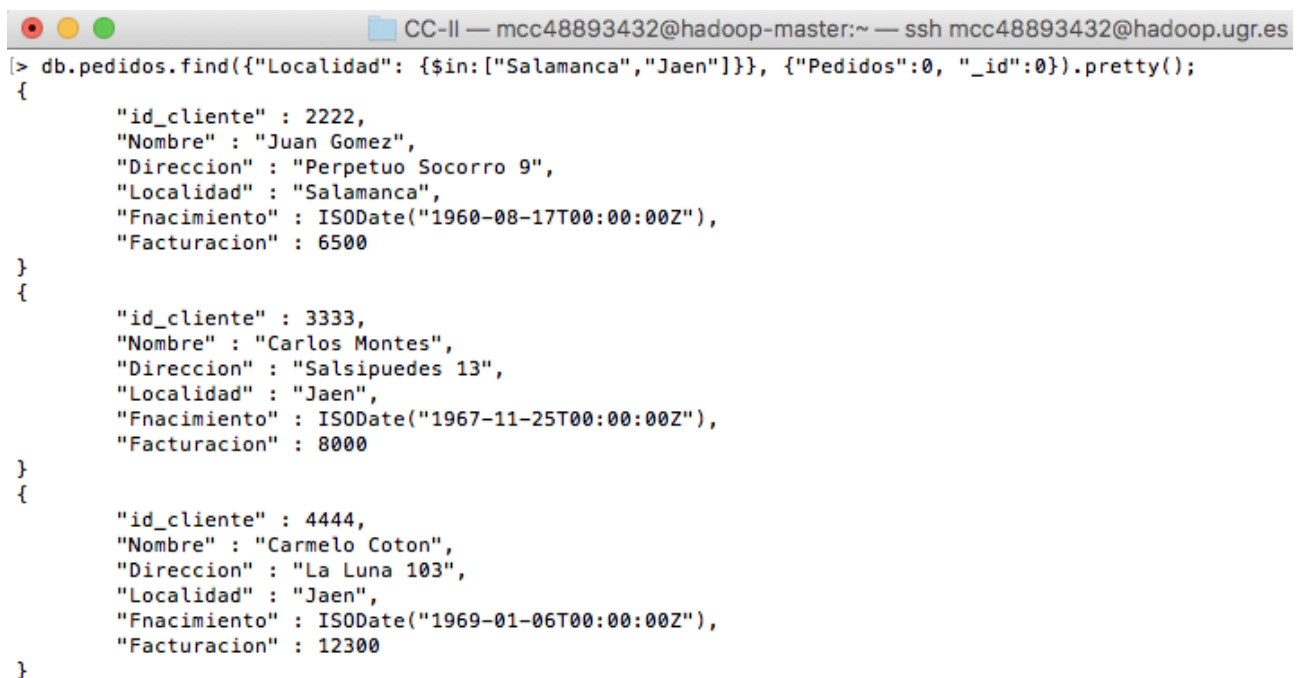
2.4. Visualiza los clientes que hayan pedido algún producto de más de 94 euros.



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.find({"Pedidos.Productos.Precio_unidad": {$gte: 94}}, {"Nombre":1, "_id":0});
{ "Nombre" : "Pedro Ramirez" }
{ "Nombre" : "Juan Gomez" }
{ "Nombre" : "Cristina Miralles" }
```

Figura 13: Visualizando los clientes.

2.5. Visualiza los clientes de Jaén o Salamanca (excluye los datos de los pedidos). Utiliza los operadores \$or e \$in .



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.find({"Localidad": {$in:["Salamanca","Jaen"]}}, {"Pedidos":0, "_id":0}).pretty();
{
  "id_cliente" : 2222,
  "Nombre" : "Juan Gomez",
  "Direccion" : "Perpetuo Socorro 9",
  "Localidad" : "Salamanca",
  "Fncimiento" : ISODate("1960-08-17T00:00:00Z"),
  "Facturacion" : 6500
}
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fncimiento" : ISODate("1967-11-25T00:00:00Z"),
  "Facturacion" : 8000
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fncimiento" : ISODate("1969-01-06T00:00:00Z"),
  "Facturacion" : 12300
}
```

Figura 14: Visualizando los clientes.

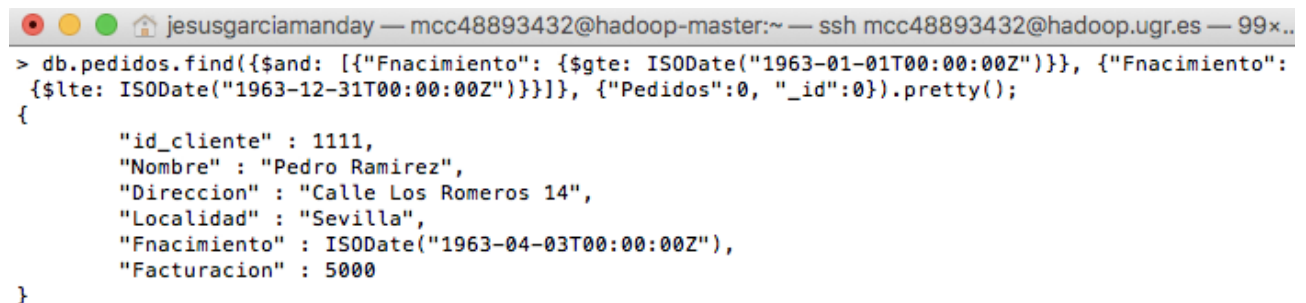
2.6. Visualiza los clientes que no tienen campo pedido.



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
> db.pedidos.find({Pedidos: {$exists: false}}, {"_id":0}).pretty();
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1967-11-25T00:00:00Z"),
  "Facturacion" : 8000
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1969-01-06T00:00:00Z"),
  "Facturacion" : 12300
}
{
  "id_cliente" : 6666,
  "Nombre" : "Chema Pamundi",
  "Direccion" : "Recogidas 54",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1969-02-04T00:00:00Z"),
  "Facturacion" : 5000
}
```

Figura 15: Visualizando los clientes.

2.7. Visualiza los clientes que hayan nacido en 1963.



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es — 99x..
> db.pedidos.find({$and: [{"Fnacimiento": {$gte: ISODate("1963-01-01T00:00:00Z")}}, {"Fnacimiento":
{$lte: ISODate("1963-12-31T00:00:00Z")}]}, {"Pedidos":0, "_id":0}).pretty();
{
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romero 14",
  "Localidad" : "Sevilla",
  "Fnacimiento" : ISODate("1963-04-03T00:00:00Z"),
  "Facturacion" : 5000
}
```

Figura 16: Visualizando los clientes.

2.8. Visualiza los clientes que hayan pedido algún producto fabricado por Canon y algún producto cuyo precio sea inferior a 15 euros.

```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es — 120x43
> db.pedidos.find({$or: [{"Pedidos.Productos.Fabricante": "Canon"}, {"Pedidos.Productos.Precio_unidad": 15}]}).pretty();
{
  "_id" : ObjectId("59087944fd111cefd7a5ec8"),
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romeros 14",
  "Localidad" : "Sevilla",
  "Fnacimiento" : ISODate("1963-04-03T00:00:00Z"),
  "Facturacion" : 5000,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 390,
          "Cantidad" : 1
        },
        {
          "id_producto" : 2,
          "Nombre" : "Tablet 8 pulgadas",
          "Precio_unidad" : 95,
          "Cantidad" : 1
        }
      ]
    },
    {
      "id_pedido" : 2,
      "Productos" : [
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 115,
          "Cantidad" : 3
        }
      ]
    }
  ]
}
```

Figura 17: Visualizando los clientes (I).

```

{
  "_id" : ObjectId("59087945fd111cefb7a5ec9"),
  "id_cliente" : 2222,
  "Nombre" : "Juan Gomez",
  "Direccion" : "Perpetuo Socorro 9",
  "Localidad" : "Salamanca",
  "Fnacimiento" : ISODate("1960-08-17T00:00:00Z"),
  "Facturacion" : 6500,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 1,
          "Nombre" : "Pentium IV",
          "Fabricante" : "Intel",
          "Precio_unidad" : 100,
          "Cantidad" : 1
        },
        {
          "id_producto" : 42,
          "Nombre" : "Portatil ASM Mod. 254",
          "Fabricante" : "Intel",
          "Precio_unidad" : 455,
          "Cantidad" : 2
        },
        {
          "id_producto" : 27,
          "Nombre" : "Cable USB",
          "Precio_unidad" : 11,
          "Cantidad" : 12
        }
      ]
    }
  ],
}

```

Figura 18: Visualizando los clientes (II).

```

{
  "id_pedido" : 2,
  "Productos" : [
    {
      "id_producto" : 77,
      "Nombre" : "Impresora Laser",
      "Fabricante" : "Canon",
      "Precio_unidad" : 128,
      "Cantidad" : 3
    },
    {
      "id_producto" : 42,
      "Nombre" : "Portatil ASM Mod. 254",
      "Fabricante" : "Intel",
      "Precio_unidad" : 451,
      "Cantidad" : 5
    },
    {
      "id_producto" : 21,
      "Nombre" : "Disco Duro 500GB",
      "Precio_unidad" : 99,
      "Cantidad" : 10
    }
  ]
},

```

Figura 19: Visualizando los clientes (III).


```
{
  "id_pedido" : 3,
  "Productos" : [
    {
      "id_producto" : 1,
      "Nombre" : "Pentium IV",
      "Fabricante" : "Intel",
      "Precio_unidad" : 94,
      "Cantidad" : 5
    },
    {
      "id_producto" : 95,
      "Nombre" : "SAI 5H Mod. 258",
      "Precio_unidad" : 213,
      "Cantidad" : 2
    },
    {
      "id_producto" : 21,
      "Precio_unidad" : 66,
      "Nombre" : "Disco Duro 500GB",
      "Cantidad" : 10
    }
  ]
}
```

Figura 20: Visualizando los clientes (IV).

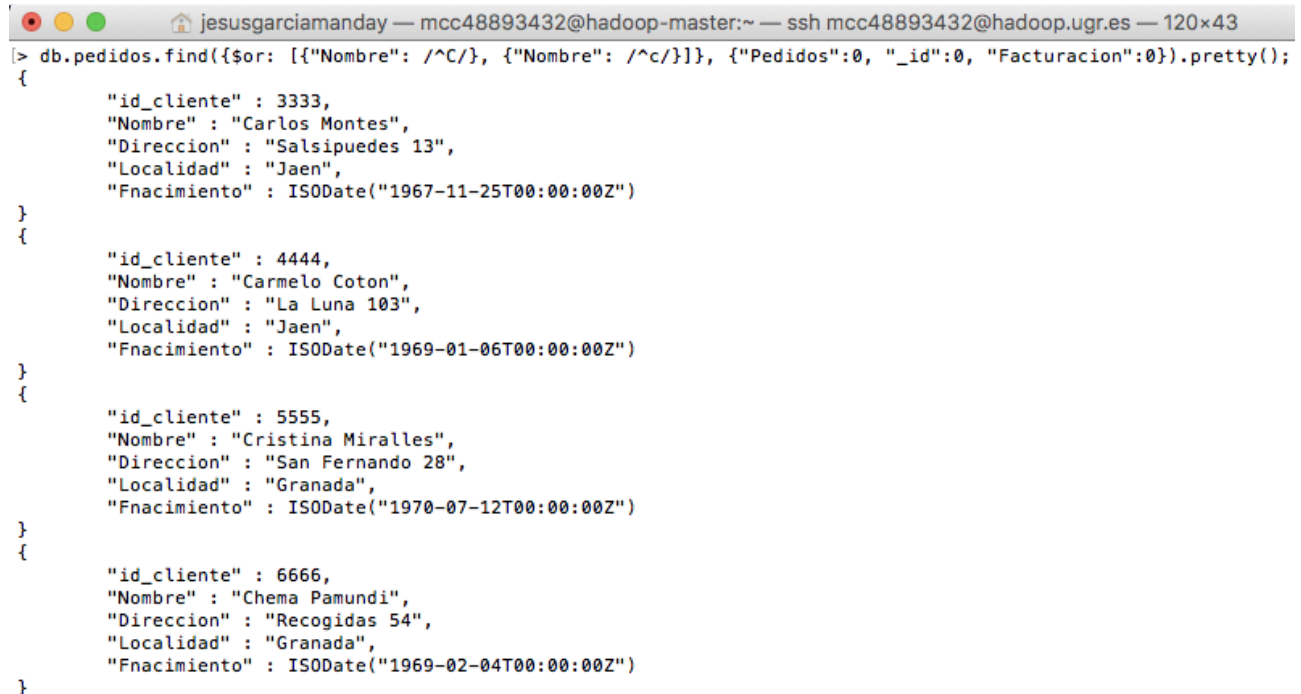
```

{
  "_id" : ObjectId("59087945fd111cefb7a5ecc"),
  "id_cliente" : 5555,
  "Nombre" : "Cristina Miralles",
  "Direccion" : "San Fernando 28",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1970-07-12T00:00:00Z"),
  "Facturacion" : 16500,
  "Pedidos" : [
    {
      "id_pedido" : 1,
      "Productos" : [
        {
          "id_producto" : 95,
          "Nombre" : "SAI 5H Mod. 258",
          "Precio_unidad" : 211,
          "Cantidad" : 2
        },
        {
          "id_producto" : 42,
          "Nombre" : "Portatil ASM Mod. 254",
          "Precio_unidad" : 460,
          "Fabricante" : "Intel",
          "Cantidad" : 2
        },
        {
          "id_producto" : 77,
          "Nombre" : "Impresora Laser",
          "Fabricante" : "Canon",
          "Precio_unidad" : 119,
          "Cantidad" : 2
        }
      ]
    }
  ]
}

```

Figura 21: Visualizando los clientes (V).

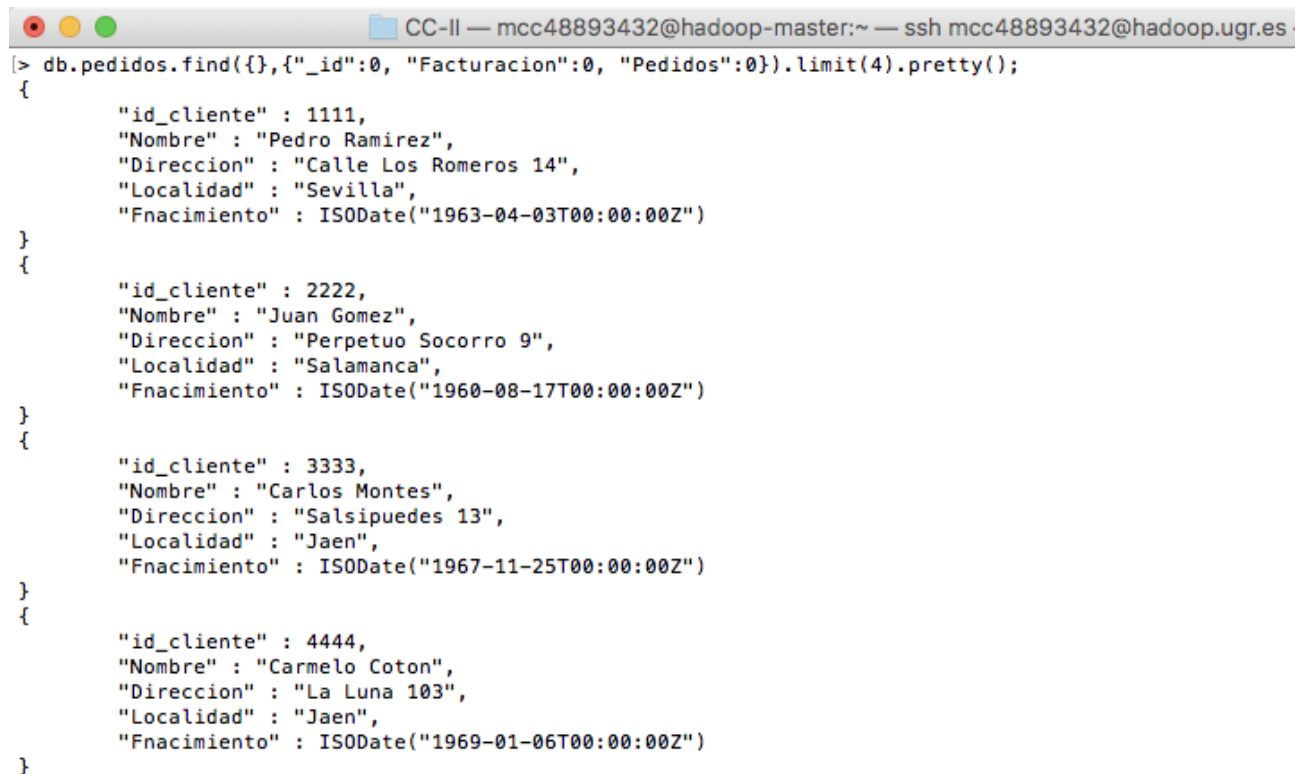
2.9. Datos personales (id_cliente, Nombre, Dirección, Localidad y Fnacimiento) de los clientes cuyo nombre empieza por la cadena ζ (No distinguir entre mayúsculas y minúsculas).



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es — 120x43
[> db.pedidos.find({$or: [{"Nombre": /^C/}, {"Nombre": /^c/}]}, {"Pedidos":0, "_id":0, "Facturacion":0}).pretty();
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1967-11-25T00:00:00Z")
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1969-01-06T00:00:00Z")
}
{
  "id_cliente" : 5555,
  "Nombre" : "Cristina Miralles",
  "Direccion" : "San Fernando 28",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1970-07-12T00:00:00Z")
}
{
  "id_cliente" : 6666,
  "Nombre" : "Chema Pamundi",
  "Direccion" : "Recogidas 54",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1969-02-04T00:00:00Z")
}
```

Figura 22: Visualizando los clientes.

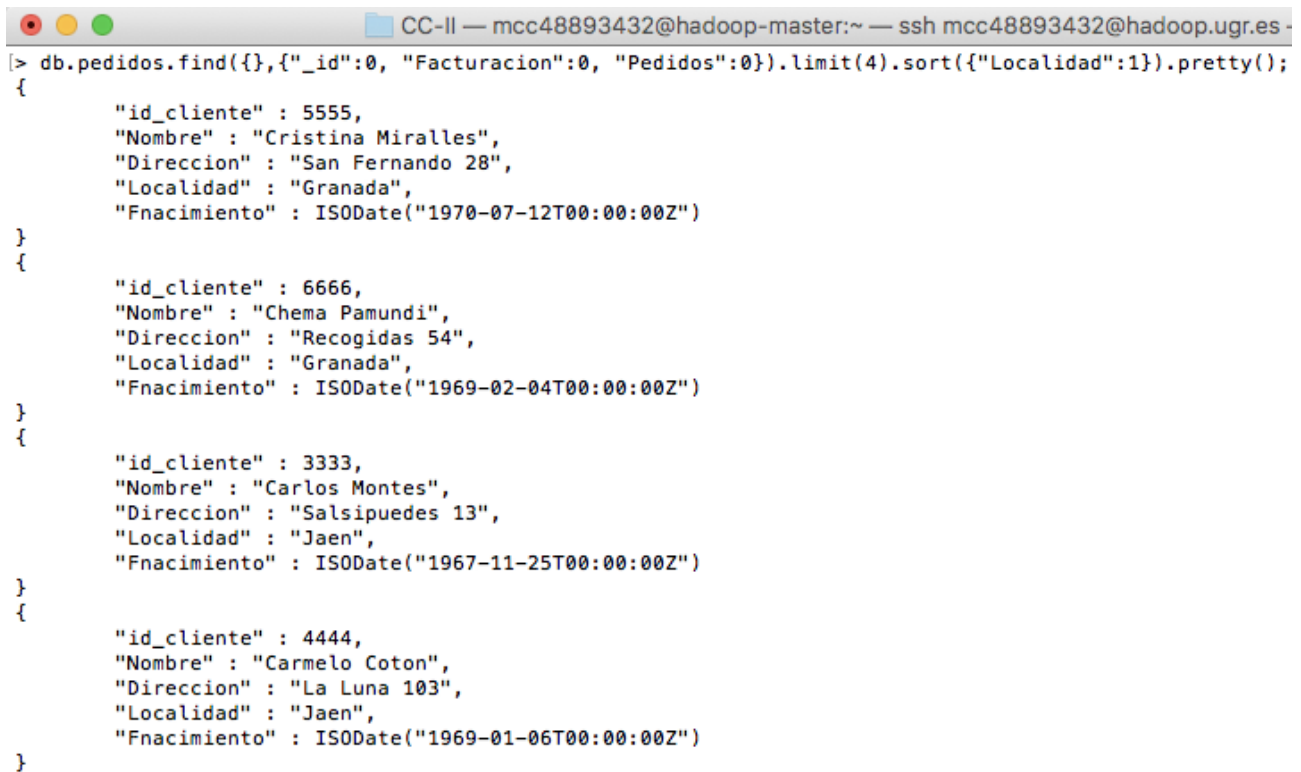
2.10. Visualiza los datos personales de los clientes (excluyendo `_id`). Limita los documentos a 4.



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.find({}, {"_id":0, "Facturacion":0, "Pedidos":0}).limit(4).pretty();
{
  "id_cliente" : 1111,
  "Nombre" : "Pedro Ramirez",
  "Direccion" : "Calle Los Romero 14",
  "Localidad" : "Sevilla",
  "Fncimiento" : ISODate("1963-04-03T00:00:00Z")
}
{
  "id_cliente" : 2222,
  "Nombre" : "Juan Gomez",
  "Direccion" : "Perpetuo Socorro 9",
  "Localidad" : "Salamanca",
  "Fncimiento" : ISODate("1960-08-17T00:00:00Z")
}
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fncimiento" : ISODate("1967-11-25T00:00:00Z")
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fncimiento" : ISODate("1969-01-06T00:00:00Z")
}
```

Figura 23: Visualizando los datos personales de los clientes.

2.11. Ídem anterior pero ordenando los documentos por Localidad (ascendente) e id_cliente (descendente).



The terminal window shows a MongoDB query executed in a shell. The query filters documents where '_id' is 0, 'Facturacion' is 0, and 'Pedidos' is 0, limits the results to 4, and sorts them by 'Localidad' in ascending order. The output shows four JSON documents, each representing a client's personal data. The documents are sorted by 'Localidad' (Granada, Granada, Jaen, Jaen) and then by 'id_cliente' in descending order (5555, 6666, 3333, 4444).

```
[> db.pedidos.find({},{"_id":0, "Facturacion":0, "Pedidos":0}).limit(4).sort({"Localidad":1}).pretty();
{
  "id_cliente" : 5555,
  "Nombre" : "Cristina Miralles",
  "Direccion" : "San Fernando 28",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1970-07-12T00:00:00Z")
}
{
  "id_cliente" : 6666,
  "Nombre" : "Chema Pamundi",
  "Direccion" : "Recogidas 54",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1969-02-04T00:00:00Z")
}
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1967-11-25T00:00:00Z")
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1969-01-06T00:00:00Z")
}
```

Figura 24: Visualizando los datos personales de los clientes ascendentemente por **Localidad**.



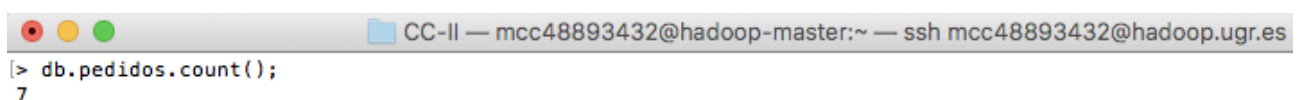
```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es —
[> db.pedidos.find({}, {"_id":0, "Facturacion":0, "Pedidos":0}).limit(4).sort({"id_cliente":-1}).pretty();
{
  "id_cliente" : 6666,
  "Nombre" : "Chema Pamundi",
  "Direccion" : "Recogidas 54",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1969-02-04T00:00:00Z")
}
{
  "id_cliente" : 5555,
  "Nombre" : "Cristina Miralles",
  "Direccion" : "San Fernando 28",
  "Localidad" : "Granada",
  "Fnacimiento" : ISODate("1970-07-12T00:00:00Z")
}
{
  "id_cliente" : 4444,
  "Nombre" : "Carmelo Coton",
  "Direccion" : "La Luna 103",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1969-01-06T00:00:00Z")
}
{
  "id_cliente" : 3333,
  "Nombre" : "Carlos Montes",
  "Direccion" : "Salsipuedes 13",
  "Localidad" : "Jaen",
  "Fnacimiento" : ISODate("1967-11-25T00:00:00Z")
}
```

Figura 25: Visualizando los datos personales de los descendientes por **id_cliente**.

3. Agregación.

A partir de la colección **pedidos** utilizaremos consultas más complejas por medio de los operadores de agregación (pipeline).


3.1. N° total de clientes.



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.count();
7
```

Figura 26: Visualizando el numero de clientes.

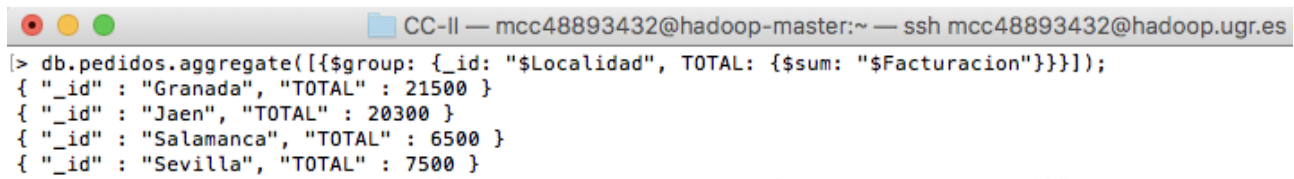
3.2. N° total de clientes de Jaén



```
CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.find({"Localidad": "Jaen"}).count();
2
```

Figura 27: Visualizando el numero de clientes.

3.3. Facturación total clientes por localidad.



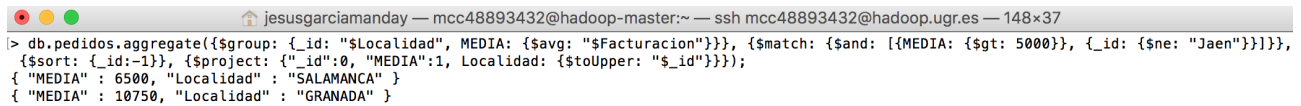
```

CC-II — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
[> db.pedidos.aggregate([{$group: {_id: "$Localidad", TOTAL: {$sum: "$Facturacion"}}}]);
{ "_id" : "Granada", "TOTAL" : 21500 }
{ "_id" : "Jaen", "TOTAL" : 20300 }
{ "_id" : "Salamanca", "TOTAL" : 6500 }
{ "_id" : "Sevilla", "TOTAL" : 7500 }

```

Figura 28: Visualizando las localidades y el total de facturación.

3.4. Facturación media de clientes por localidad para las localidades distintas a Jaén con facturación media mayor de 5000. Ordenación por Localidad descendente. Eliminar el _id y poner el nombre en mayúsculas.



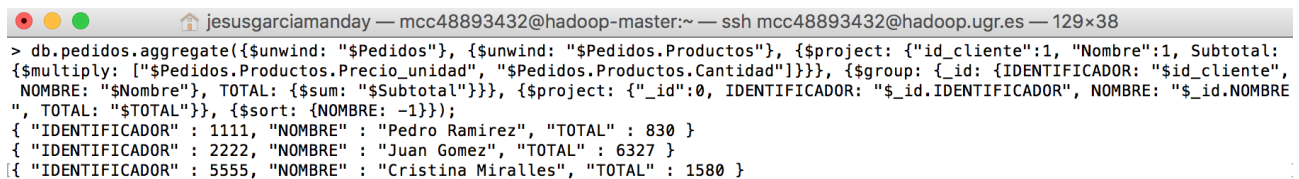
```

jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es — 148x37
[> db.pedidos.aggregate({$group: {_id: "$Localidad", MEDIA: {$avg: "$Facturacion"}}}, {$match: {$and: [{MEDIA: {$gt: 5000}}, {_id: {$ne: "Jaen"}}]}},
{$sort: {_id:-1}}, {$project: {"_id":0, "MEDIA":1, Localidad: {$toUpper: "$_id"}}});
{ "MEDIA" : 6500, "Localidad" : "SALAMANCA" }
{ "MEDIA" : 10750, "Localidad" : "GRANADA" }

```

Figura 29: Visualizando las localidades y facturación media.

3.5. Calcula la cantidad total facturada por cada cliente (uso de "unwind").



```

jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es — 129x38
[> db.pedidos.aggregate({$unwind: "$Pedidos"}, {$unwind: "$Pedidos.Productos"}, {$project: {"id_cliente":1, "Nombre":1, Subtotal:
{$multiply: ["$Pedidos.Productos.Precio_unidad", "$Pedidos.Productos.Cantidad"]}}, {$group: {_id: {IDENTIFICADOR: "$id_cliente",
NOMBRE: "$Nombre"}, TOTAL: {$sum: "$Subtotal"}}}, {$project: {"_id":0, IDENTIFICADOR: "$_id.IDENTIFICADOR", NOMBRE: "$_id.NOMBRE",
TOTAL: "$TOTAL"}}, {$sort: {NOMBRE: -1}});
{ "IDENTIFICADOR" : 1111, "NOMBRE" : "Pedro Ramirez", "TOTAL" : 830 }
{ "IDENTIFICADOR" : 2222, "NOMBRE" : "Juan Gomez", "TOTAL" : 6327 }
{ "IDENTIFICADOR" : 5555, "NOMBRE" : "Cristina Miralles", "TOTAL" : 1580 }

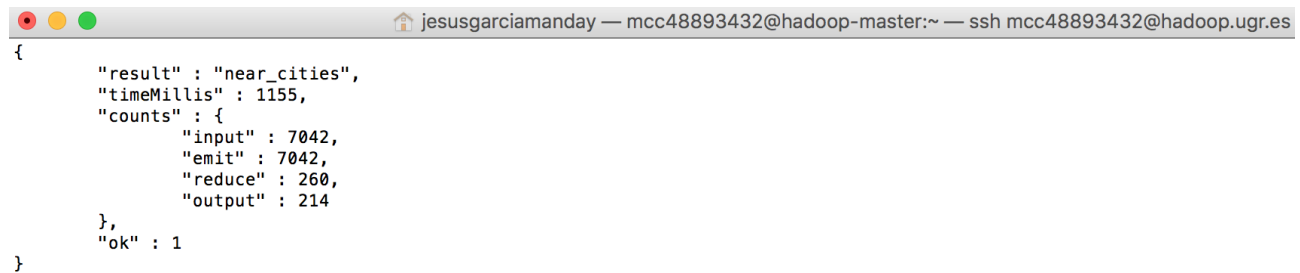
```

Figura 30: Visualizando los clientes y el total.

4. MapReduce.

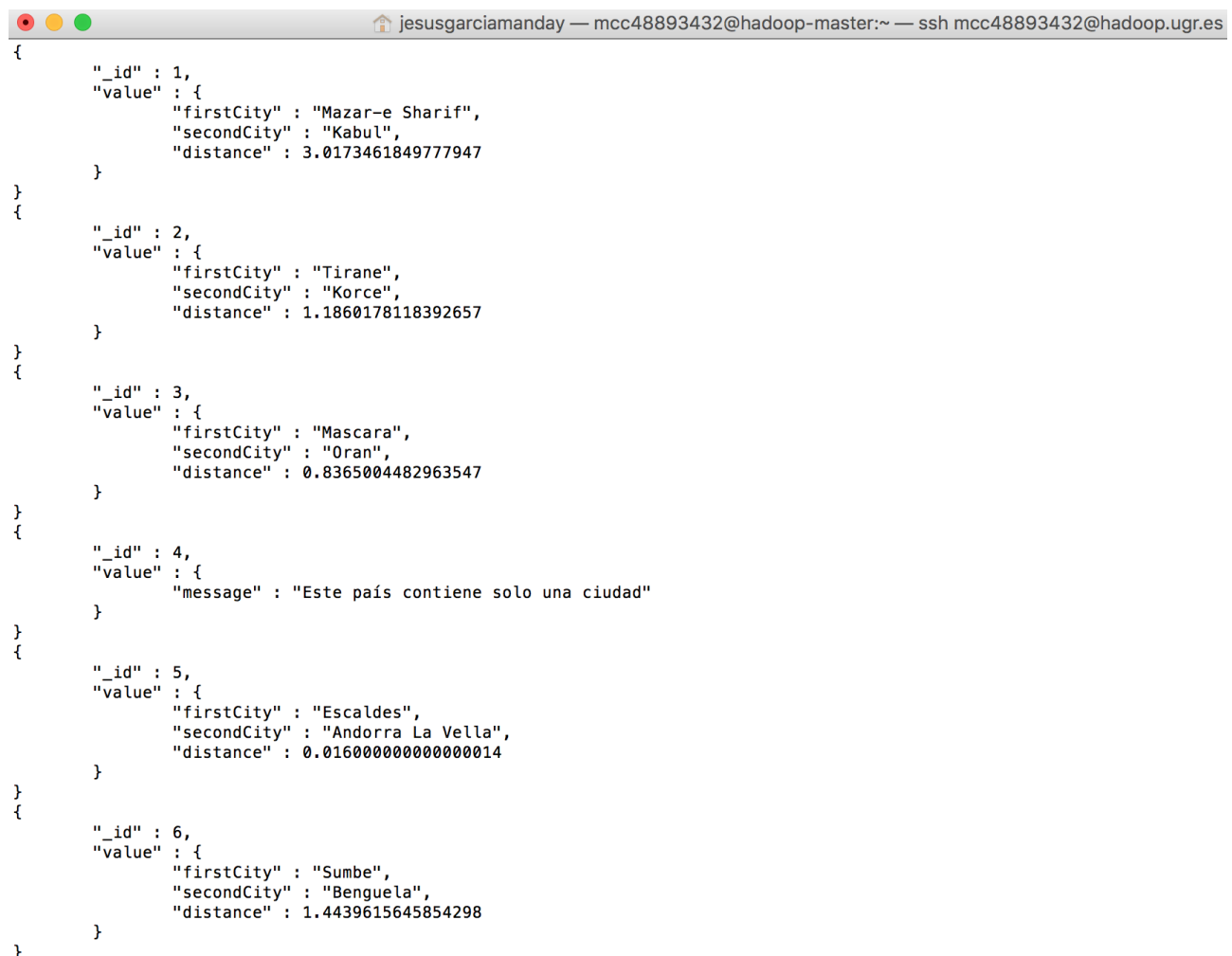
Vamos a utilizar la base de datos libre **GeoWorldMap** de **GeoBytes**. Es una base de datos de países, con sus estados/regiones y ciudades importantes.

4.1. Encontrar las ciudades más cercanas sobre la colección recién creada mediante un enfoque MapReduce conforme a los pasos que se ilustran en el tutorial práctico.



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
{
  "result" : "near_cities",
  "timeMillis" : 1155,
  "counts" : {
    "input" : 7042,
    "emit" : 7042,
    "reduce" : 260,
    "output" : 214
  },
  "ok" : 1
}
```

Figura 31: Visualizando los documentos (I).



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
{
  "_id" : 1,
  "value" : {
    "firstCity" : "Mazar-e Sharif",
    "secondCity" : "Kabul",
    "distance" : 3.0173461849777947
  }
}
{
  "_id" : 2,
  "value" : {
    "firstCity" : "Tirane",
    "secondCity" : "Korce",
    "distance" : 1.1860178118392657
  }
}
{
  "_id" : 3,
  "value" : {
    "firstCity" : "Mascara",
    "secondCity" : "Oran",
    "distance" : 0.8365004482963547
  }
}
{
  "_id" : 4,
  "value" : {
    "message" : "Este país contiene solo una ciudad"
  }
}
{
  "_id" : 5,
  "value" : {
    "firstCity" : "Escaldes",
    "secondCity" : "Andorra La Vella",
    "distance" : 0.016000000000000014
  }
}
{
  "_id" : 6,
  "value" : {
    "firstCity" : "Sumbe",
    "secondCity" : "Benguela",
    "distance" : 1.4439615645854298
  }
}
```

Figura 32: Visualizando los documentos (II).


```
{
  "_id" : 7,
  "value" : {
    "message" : "Este país contiene solo una ciudad"
  }
}
{
  "_id" : 8,
  "value" : {
    "firstCity" : "McMurdo Station",
    "secondCity" : "Molodesjnaja",
    "distance" : 183.962470629202
  }
}
{
  "_id" : 9,
  "value" : {
    "firstCity" : "Falmouth",
    "secondCity" : "Saint Johns",
    "distance" : 0.12037026210821469
  }
}
{
  "_id" : 10,
  "value" : {
    "firstCity" : "Turdera",
    "secondCity" : "Lomas De Zamora",
    "distance" : 0.015999999999998238
  }
}
{
  "_id" : 11,
  "value" : {
    "firstCity" : "Spitak",
    "secondCity" : "Vanadzor",
    "distance" : 0.2200460179144342
  }
}
{
  "_id" : 12,
  "value" : {
    "message" : "Este país contiene solo una ciudad"
  }
}
}
```

Figura 33: Visualizando los documentos (III).

```
{
  "_id" : 14,
  "value" : {
    "firstCity" : "Williamstown",
    "secondCity" : "Kalgoorlie",
    "distance" : 0.01600000000005343
  }
}
{
  "_id" : 15,
  "value" : {
    "firstCity" : "Wiener Neustadt",
    "secondCity" : "Neudörfl",
    "distance" : 0.03712142238654041
  }
}
{
  "_id" : 16,
  "value" : {
    "message" : "Este país contiene solo una ciudad"
  }
}
{
  "_id" : 17,
  "value" : {
    "firstCity" : "Old Bight",
    "secondCity" : "Freetown",
    "distance" : 0.07468600939935845
  }
}
{
  "_id" : 18,
  "value" : {
    "firstCity" : "Muharraq",
    "secondCity" : "Al Manamah",
    "distance" : 0.035805027579939586
  }
}
{
  "_id" : 20,
  "value" : {
    "firstCity" : "Dhaka",
    "secondCity" : "Comilla",
    "distance" : 0.8367855161270389
  }
}
```

Figura 34: Visualizando los documentos (IV).

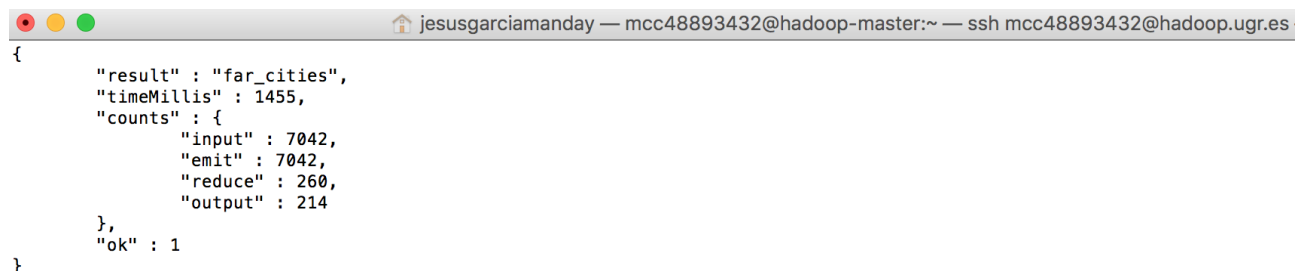
```

{
  "_id" : 21,
  "value" : {
    "firstCity" : "Warrens",
    "secondCity" : "Christchurch",
    "distance" : 0.024041630560339342
  }
}
{
  "_id" : 23,
  "value" : {
    "firstCity" : "Molodechno",
    "secondCity" : "Minsk",
    "distance" : 0.8294443923494809
  }
}
Type "it" for more
|> it
{
  "_id" : 24,
  "value" : {
    "firstCity" : "Woluwe-Saint-Pierre",
    "secondCity" : "Woluwe-Saint-Lambert",
    "distance" : 0.024041630560344366
  }
}
{
  "_id" : 25,
  "value" : {
    "firstCity" : "San Ignacio",
    "secondCity" : "Belmopan",
    "distance" : 0.31820119421523463
  }
}
{
  "_id" : 26,
  "value" : {
    "firstCity" : "Porto-Novo",
    "secondCity" : "Cotonou",
    "distance" : 0.22703523955544888
  }
}
}

```

Figura 35: Visualizando los documentos (V).

4.2. ¿Cómo podríamos obtener las ciudades más distantes en cada país?.



```

jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
{
  "result" : "far_cities",
  "timeMillis" : 1455,
  "counts" : {
    "input" : 7042,
    "emit" : 7042,
    "reduce" : 260,
    "output" : 214
  },
  "ok" : 1
}

```

Figura 36: Visualizando los documentos (I).

```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
{
  "_id" : 1,
  "value" : {
    "city1" : "Herat",
    "city2" : "Kabul",
    "dist" : 6.98542375235748
  }
}
{
  "_id" : 2,
  "value" : {
    "city1" : "Korce",
    "city2" : "Tirane",
    "dist" : 1.1860178118392657
  }
}
{
  "_id" : 3,
  "value" : {
    "city1" : "Oran",
    "city2" : "Annaba",
    "dist" : 8.495467144307016
  }
}
{
  "_id" : 4,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 5,
  "value" : {
    "city1" : "Andorra La Vella",
    "city2" : "Escaldes",
    "dist" : 0.016000000000000014
  }
}
{
  "_id" : 6,
  "value" : {
    "city1" : "Lubango",
    "city2" : "Luanda",
    "dist" : 7.976476540428111
  }
}
```

Figura 37: Visualizando los documentos (II).

```
{
  "_id" : 7,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 8,
  "value" : {
    "city1" : "Molodesjnaja",
    "city2" : "McMurdo Station",
    "dist" : 183.962470629202
  }
}
{
  "_id" : 9,
  "value" : {
    "city1" : "Saint Johns",
    "city2" : "Falmouth",
    "dist" : 0.12037026210821469
  }
}
{
  "_id" : 10,
  "value" : {
    "city1" : "San Rafael",
    "city2" : "Ushuaia",
    "dist" : 30.929864031385584
  }
}
{
  "_id" : 11,
  "value" : {
    "city1" : "Gyumri",
    "city2" : "Yerevan",
    "dist" : 0.9049468989946328
  }
}
{
  "_id" : 12,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
}
```

Figura 38: Visualizando los documentos (III).

```
{
  "_id" : 14,
  "value" : {
    "city1" : "Byron Bay",
    "city2" : "Shark Bay",
    "dist" : 40.16699782159477
  }
}
{
  "_id" : 15,
  "value" : {
    "city1" : "Gänserndorf",
    "city2" : "Göttzisz",
    "dist" : 7.172467427601189
  }
}
{
  "_id" : 16,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 17,
  "value" : {
    "city1" : "Matthew Town",
    "city2" : "Freeport",
    "dist" : 7.516713244497227
  }
}
{
  "_id" : 18,
  "value" : {
    "city1" : "Muharraq",
    "city2" : "Jasra",
    "dist" : 0.18866372200293755
  }
}
{
  "_id" : 20,
  "value" : {
    "city1" : "Comilla",
    "city2" : "Saginaw",
    "dist" : 94.16656784655582
  }
}
```

Figura 39: Visualizando los documentos (IV).

```

{
  "_id" : 21,
  "value" : {
    "city1" : "Bridgetown",
    "city2" : "Speightstown",
    "dist" : 0.15056227947264958
  }
}
{
  "_id" : 23,
  "value" : {
    "city1" : "Brest",
    "city2" : "Mahilyow",
    "dist" : 6.8794688748478245
  }
}
Type "it" for more
|> it
{
  "_id" : 24,
  "value" : {
    "city1" : "Arlon",
    "city2" : "Veurne",
    "dist" : 3.4406330812802466
  }
}
{
  "_id" : 25,
  "value" : {
    "city1" : "Orange Walk",
    "city2" : "Punta Gorda",
    "dist" : 1.9828234918923053
  }
}
{
  "_id" : 26,
  "value" : {
    "city1" : "Natitingou",
    "city2" : "Cotonou",
    "dist" : 4.10772990835571
  }
}
}

```

Figura 40: Visualizando los documentos (V).

4.3. ¿Qué ocurre si en un país hay dos parejas de ciudades que están a la misma distancia mínima? ¿Cómo harías para que aparecieran todas?.

```

jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es
{
  "result" : "near_cities2",
  "timeMillis" : 1437,
  "counts" : {
    "input" : 7042,
    "emit" : 7042,
    "reduce" : 260,
    "output" : 214
  },
  "ok" : 1
}

```

Figura 41: Visualizando los documentos (I).



A terminal window with a title bar showing a home icon, the username 'jesusgarciamanday', and the command 'ssh mcc48893432@hadoop-master:~'. The terminal displays four JSON documents, each with an '_id' field and a 'value' field. The first three documents have a 'cities' array and a 'dist' field. The fourth document has a 'message' field.

```
{
  "_id" : 1,
  "value" : {
    "cities" : [
      "Kabul",
      "Mazar-e Sharif"
    ],
    "dist" : 3.0173461849777947
  }
}
{
  "_id" : 2,
  "value" : {
    "cities" : [
      "Korce",
      "Tirane"
    ],
    "dist" : 1.1860178118392657
  }
}
{
  "_id" : 3,
  "value" : {
    "cities" : [
      "Oran",
      "Mascara"
    ],
    "dist" : 0.8365004482963547
  }
}
{
  "_id" : 4,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
```

Figura 42: Visualizando los documentos (II).


```
{
  "_id" : 5,
  "value" : {
    "cities" : [
      "Andorra La Vella",
      "Escaldes"
    ],
    "dist" : 0.016000000000000014
  }
}
{
  "_id" : 6,
  "value" : {
    "cities" : [
      "Benguela",
      "Sumbe"
    ],
    "dist" : 1.4439615645854298
  }
}
{
  "_id" : 7,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 8,
  "value" : {
    "cities" : [
      "Molodesjnaja",
      "McMurdo Station"
    ],
    "Reading",
    "McMurdo Station"
  ],
  "dist" : 183.962470629202
}
}
```

Figura 43: Visualizando los documentos (III).

```
{
  "_id" : 9,
  "value" : {
    "cities" : [
      "Saint Johns",
      "Falmouth"
    ],
    "dist" : 0.12037026210821469
  }
}
{
  "_id" : 10,
  "value" : {
    "cities" : [
      "Turdera",
      "Lomas De Zamora"
    ],
    "dist" : 0.015999999999998238
  }
}
{
  "_id" : 11,
  "value" : {
    "cities" : [
      "Vanadzor",
      "Spitak"
    ],
    "dist" : 0.2200460179144342
  }
}
{
  "_id" : 12,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
```

Figura 44: Visualizando los documentos (IV).

```
{
  "_id" : 14,
  "value" : {
    "cities" : [
      [
        "Kalgoorlie",
        "Williamstown"
      ]
    ],
    "dist" : 0.0160000000000005343
  }
}
{
  "_id" : 15,
  "value" : {
    "cities" : [
      [
        "Neudörfl",
        "Wiener Neustadt"
      ]
    ],
    "dist" : 0.03712142238654041
  }
}
{
  "_id" : 16,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 17,
  "value" : {
    "cities" : [
      [
        "Freetown",
        "Old Bight"
      ]
    ],
    "dist" : 0.07468600939935845
  }
}
```

Figura 45: Visualizando los documentos (V).

```

{
  "_id" : 18,
  "value" : {
    "cities" : [
      [
        "Al Manamah",
        "Muharraq"
      ],
      [
        "Manama",
        "Muharraq"
      ]
    ],
    "dist" : 0.035805027579939586
  }
}
{
  "_id" : 20,
  "value" : {
    "cities" : [
      [
        "Comilla",
        "Dhaka"
      ]
    ],
    "dist" : 0.8367855161270389
  }
}
{
  "_id" : 21,
  "value" : {
    "cities" : [
      [
        "Christchurch",
        "Warrens"
      ]
    ],
    "dist" : 0.024041630560339342
  }
}

```

Figura 46: Visualizando los documentos (VI).

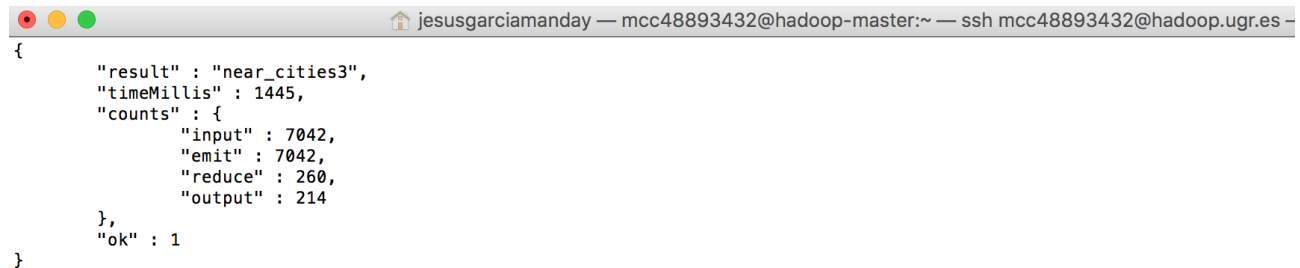
```

{
  "_id" : 23,
  "value" : {
    "cities" : [
      [
        "Minsk",
        "Molodechno"
      ],
      [
        "Minsk",
        "Molodechno"
      ]
    ],
    "dist" : 0.8294443923494809
  }
}
Type "it" for more
> █

```

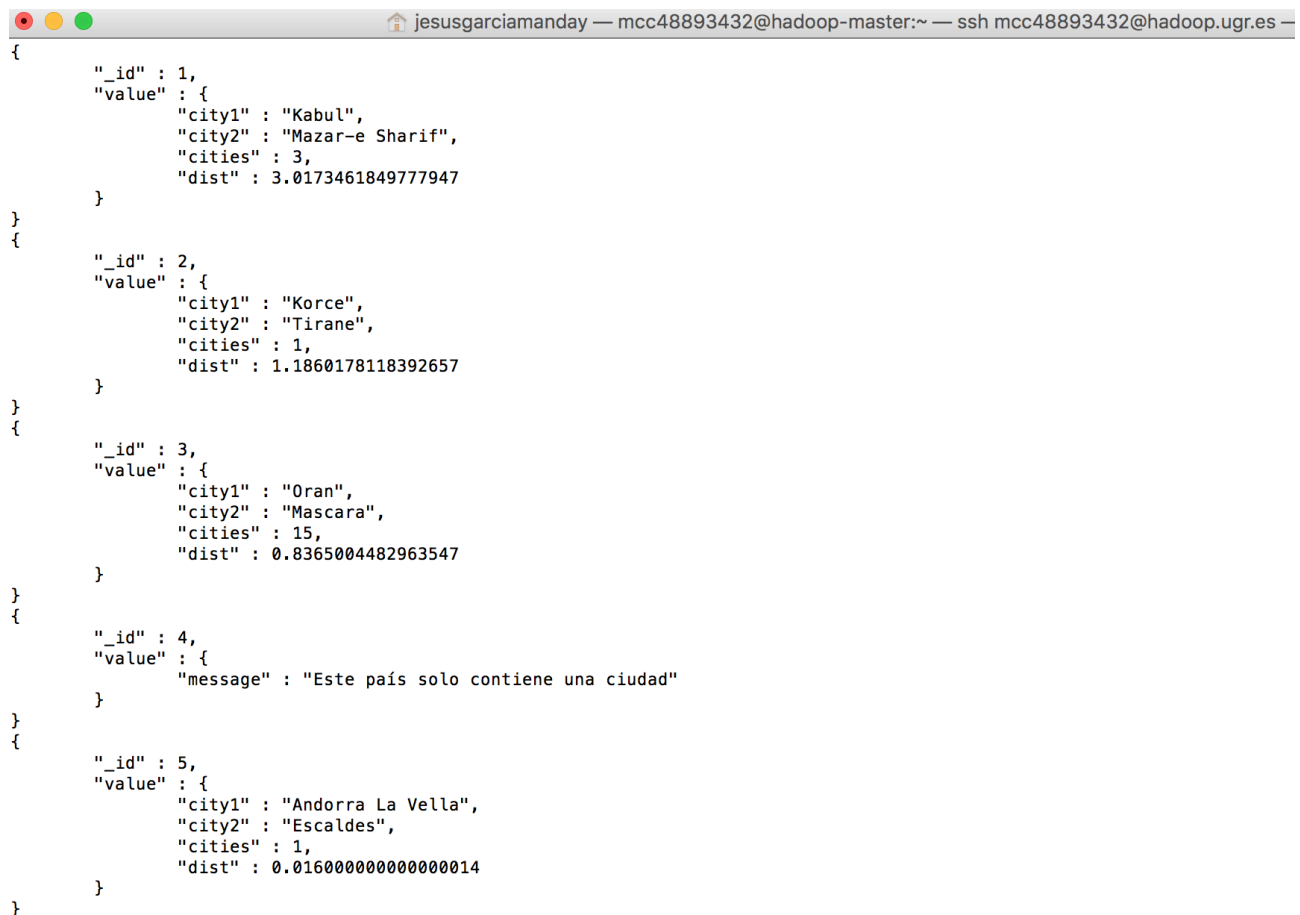
Figura 47: Visualizando los documentos (VII).

4.4. ¿Cómo podríamos obtener adicionalmente la cantidad de parejas de ciudades evaluadas para cada país consultado?.



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es —  
{  
  "result" : "near_cities3",  
  "timeMillis" : 1445,  
  "counts" : {  
    "input" : 7042,  
    "emit" : 7042,  
    "reduce" : 260,  
    "output" : 214  
  },  
  "ok" : 1  
}
```

Figura 48: Visualizando los documentos (I).



```
jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es —  
{  
  "_id" : 1,  
  "value" : {  
    "city1" : "Kabul",  
    "city2" : "Mazar-e Sharif",  
    "cities" : 3,  
    "dist" : 3.0173461849777947  
  }  
}  
{  
  "_id" : 2,  
  "value" : {  
    "city1" : "Korce",  
    "city2" : "Tirane",  
    "cities" : 1,  
    "dist" : 1.1860178118392657  
  }  
}  
{  
  "_id" : 3,  
  "value" : {  
    "city1" : "Oran",  
    "city2" : "Mascara",  
    "cities" : 15,  
    "dist" : 0.8365004482963547  
  }  
}  
{  
  "_id" : 4,  
  "value" : {  
    "message" : "Este país solo contiene una ciudad"  
  }  
}  
{  
  "_id" : 5,  
  "value" : {  
    "city1" : "Andorra La Vella",  
    "city2" : "Escaldes",  
    "cities" : 1,  
    "dist" : 0.016000000000000014  
  }  
}
```

Figura 49: Visualizando los documentos (II).

```
{
  "_id" : 6,
  "value" : {
    "city1" : "Benguela",
    "city2" : "Sumbe",
    "cities" : 10,
    "dist" : 1.4439615645854298
  }
}
{
  "_id" : 7,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 8,
  "value" : {
    "city1" : "Molodesjnaja",
    "city2" : "McMurdo Station",
    "cities" : 2,
    "dist" : 183.962470629202
  }
}
{
  "_id" : 9,
  "value" : {
    "city1" : "Saint Johns",
    "city2" : "Falmouth",
    "cities" : 1,
    "dist" : 0.12037026210821469
  }
}
{
  "_id" : 10,
  "value" : {
    "city1" : "Turdera",
    "city2" : "Lomas De Zamora",
    "cities" : 1431,
    "dist" : 0.015999999999998238
  }
}
```

Figura 50: Visualizando los documentos (III).

```
{
  "_id" : 11,
  "value" : {
    "city1" : "Vanadzor",
    "city2" : "Spitak",
    "cities" : 6,
    "dist" : 0.2200460179144342
  }
}
{
  "_id" : 12,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 14,
  "value" : {
    "city1" : "Kalgoorlie",
    "city2" : "Williamstown",
    "cities" : 26105,
    "dist" : 0.0160000000000005343
  }
}
{
  "_id" : 15,
  "value" : {
    "city1" : "Neudörfl",
    "city2" : "Wiener Neustadt",
    "cities" : 1377,
    "dist" : 0.03712142238654041
  }
}
{
  "_id" : 16,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
}
```

Figura 51: Visualizando los documentos (IV).

```

{
  "_id" : 17,
  "value" : {
    "city1" : "Freetown",
    "city2" : "Old Bight",
    "cities" : 170,
    "dist" : 0.07468600939935845
  }
}
{
  "_id" : 18,
  "value" : {
    "city1" : "Al Manamah",
    "city2" : "Muharraq",
    "cities" : 20,
    "dist" : 0.035805027579939586
  }
}
{
  "_id" : 20,
  "value" : {
    "city1" : "Comilla",
    "city2" : "Dhaka",
    "cities" : 10,
    "dist" : 0.8367855161270389
  }
}
{
  "_id" : 21,
  "value" : {
    "city1" : "Christchurch",
    "city2" : "Warrens",
    "cities" : 6,
    "dist" : 0.024041630560339342
  }
}
{
  "_id" : 23,
  "value" : {
    "city1" : "Minsk",
    "city2" : "Molodechno",
    "cities" : 9,
    "dist" : 0.8294443923494809
  }
}
Type "it" for more
>

```

Figura 52: Visualizando los documentos (V).


4.5. ¿Cómo podríamos calcular la distancia media entre las ciudades de cada país?.

```

jesusgarciamanday — mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es -
{
  "result" : "average_distance",
  "timeMillis" : 1583,
  "counts" : {
    "input" : 7042,
    "emit" : 7042,
    "reduce" : 260,
    "output" : 214
  },
  "ok" : 1
}

```

Figura 53: Visualizando los documentos (I).



A terminal window with a title bar showing a home icon, the username 'jesusgarciamanday', and the connection details 'mcc48893432@hadoop-master:~ — ssh mcc48893432@hadoop.ugr.es'. The terminal displays five JSON documents, each with an '_id' and a 'value' field. The first three documents contain city names and distances. The fourth document contains a message. The fifth document contains city names and distances.

```
{
  "_id" : 1,
  "value" : {
    "city1" : "Kabul",
    "city2" : "Mazar-e Sharif",
    "dist_avg" : 5.148174719390839,
    "dist" : 3.0173461849777947
  }
}
{
  "_id" : 2,
  "value" : {
    "city1" : "Korce",
    "city2" : "Tirane",
    "dist_avg" : 1.1860178118392657,
    "dist" : 1.1860178118392657
  }
}
{
  "_id" : 3,
  "value" : {
    "city1" : "Oran",
    "city2" : "Mascara",
    "dist_avg" : 4.592117250388198,
    "dist" : 0.8365004482963547
  }
}
{
  "_id" : 4,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 5,
  "value" : {
    "city1" : "Andorra La Vella",
    "city2" : "Escaldes",
    "dist_avg" : 0.016000000000000014,
    "dist" : 0.016000000000000014
  }
}
```

Figura 54: Visualizando los documentos (II).

```
{
  "_id" : 6,
  "value" : {
    "city1" : "Benguela",
    "city2" : "Sumbe",
    "dist_avg" : 4.011856586626525,
    "dist" : 1.4439615645854298
  }
}
{
  "_id" : 7,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 8,
  "value" : {
    "city1" : "Molodesjnaja",
    "city2" : "McMurdo Station",
    "dist_avg" : 183.962470629202,
    "dist" : 183.962470629202
  }
}
{
  "_id" : 9,
  "value" : {
    "city1" : "Saint Johns",
    "city2" : "Falmouth",
    "dist_avg" : 0.12037026210821469,
    "dist" : 0.12037026210821469
  }
}
{
  "_id" : 10,
  "value" : {
    "city1" : "Turdera",
    "city2" : "Lomas De Zamora",
    "dist_avg" : 7.516723478942258,
    "dist" : 0.0159999999999998238
  }
}
```

Figura 55: Visualizando los documentos (III).

```
{
  "_id" : 11,
  "value" : {
    "city1" : "Vanadzor",
    "city2" : "Spitak",
    "dist_avg" : 0.5826373265925909,
    "dist" : 0.2200460179144342
  }
}
{
  "_id" : 12,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
{
  "_id" : 14,
  "value" : {
    "city1" : "Kalgoorlie",
    "city2" : "Williamstown",
    "dist_avg" : 13.523198216104323,
    "dist" : 0.016000000000005343
  }
}
{
  "_id" : 15,
  "value" : {
    "city1" : "Neudörfl",
    "city2" : "Wiener Neustadt",
    "dist_avg" : 2.48556192868551,
    "dist" : 0.03712142238654041
  }
}
{
  "_id" : 16,
  "value" : {
    "message" : "Este país solo contiene una ciudad"
  }
}
}
```

Figura 56: Visualizando los documentos (IV).

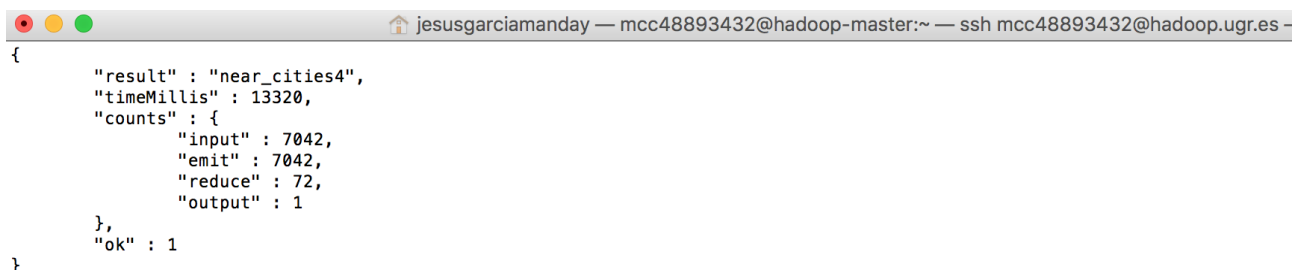
```

{
  "_id" : 17,
  "value" : {
    "city1" : "Freetown",
    "city2" : "Old Bight",
    "dist_avg" : 2.8268589972830935,
    "dist" : 0.07468600939935845
  }
}
{
  "_id" : 18,
  "value" : {
    "city1" : "Al Manamah",
    "city2" : "Muharraq",
    "dist_avg" : 0.10929194178280277,
    "dist" : 0.035805027579939586
  }
}
{
  "_id" : 20,
  "value" : {
    "city1" : "Comilla",
    "city2" : "Dhaka",
    "dist_avg" : 38.22592916882991,
    "dist" : 0.8367855161270389
  }
}
{
  "_id" : 21,
  "value" : {
    "city1" : "Christchurch",
    "city2" : "Warrens",
    "dist_avg" : 0.08047164379109523,
    "dist" : 0.024041630560339342
  }
}
{
  "_id" : 23,
  "value" : {
    "city1" : "Minsk",
    "city2" : "Molodechno",
    "dist_avg" : 3.3298240952743003,
    "dist" : 0.8294443923494809
  }
}
Type "it" for more

```

Figura 57: Visualizando los documentos (V).

4.6. ¿Mejoraría el rendimiento si creamos un índice? ¿Sobre qué campo? Comprobadlo.



```

{
  "result" : "near_cities4",
  "timeMillis" : 13320,
  "counts" : {
    "input" : 7042,
    "emit" : 7042,
    "reduce" : 72,
    "output" : 1
  },
  "ok" : 1
}

```

Figura 58: Visualizando los documentos.

Se puede ver en la imagen como el rendimiento en tiempo no mejora, pero disminuye el número de entradas y salidas.