

Contenido

INSTALACIÓN DE MONGODB:	2
IMPORTACION DE BD TWEETS:	5
CONSULTAS POR MONGOSH:	6
3.1.....	7
3.1.1.....	7
3.1.2.....	8
3.1.3.....	8
3.1.4.....	8
3.2.....	9
3.2.1.....	10
3.2.2.....	12
3.2.3.....	12
3.3.....	15
3.3.1.....	15
3.3.2.....	16
3.4.....	16

LIBRERIAS PYTHON:

Estas son necesarias para el punto 3.4:

```
import pymongo
import psycpg2
import re
import pandas as pd
import geopandas as gp
import matplotlib.pyplot as plt
from collections import Counter
from wordcloud import WordCloud
```

Wordcloud: <https://pypi.org/project/wordcloud/>

pip install wordcloud

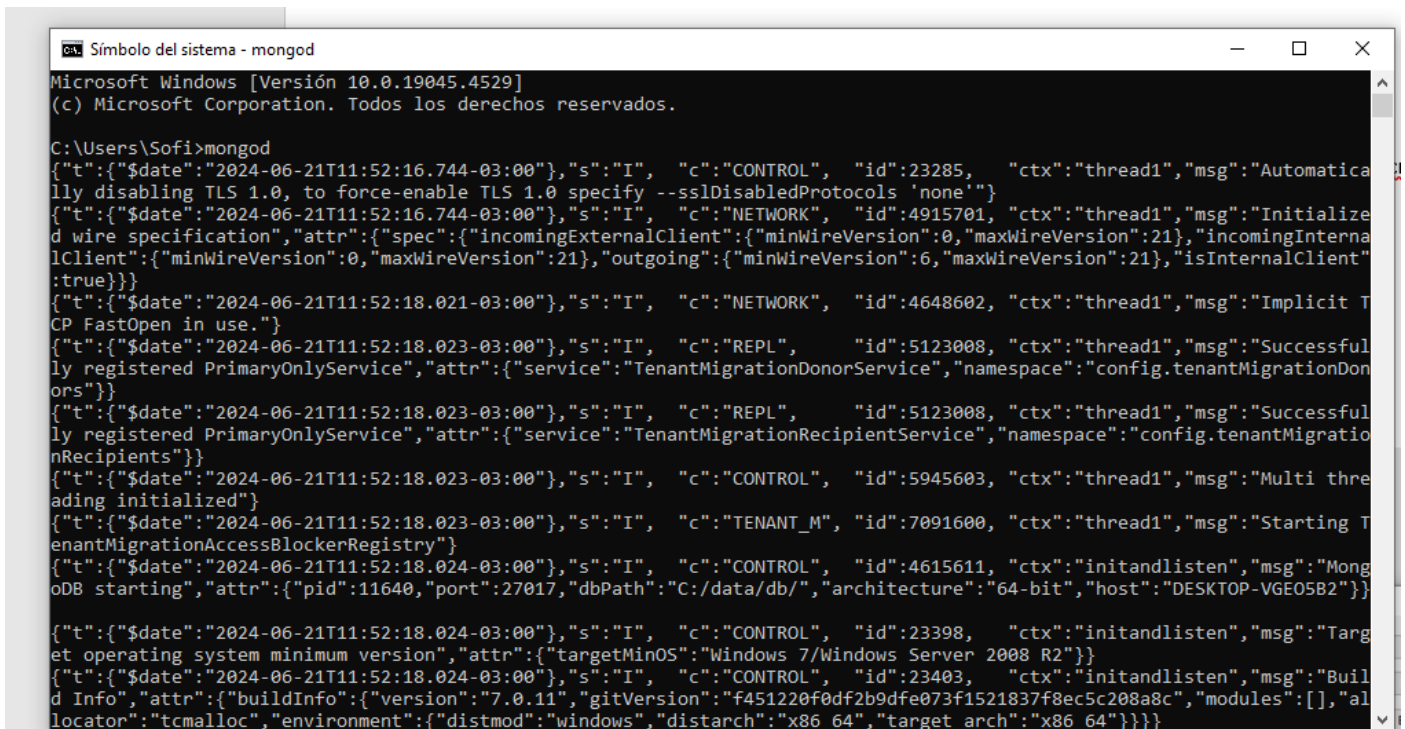
INSTALACIÓN DE MONGODB:

<https://www.youtube.com/watch?v=Lf6KqCbSwyg>

Debemos descargarla e instalarla del sitio oficial MongoDB Community Server.

Si decido NO instalar MongoD como un servicio, cada vez que quiero usarlo debo correr el comando por cmd para correr el servicio manualmente:

> mongod



```
Símbolo del sistema - mongod
Microsoft Windows [Versión 10.0.19045.4529]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Users\Sofi>mongod
{"t":{"$date":"2024-06-21T11:52:16.744-03:00"},"s":"I",  "c":"CONTROL",  "id":23285,   "ctx":"thread1","msg":"Automatica
lly disabling TLS 1.0, to force-enable TLS 1.0 specify --sslDisabledProtocols 'none'"}
{"t":{"$date":"2024-06-21T11:52:16.744-03:00"},"s":"I",  "c":"NETWORK",  "id":4915701, "ctx":"thread1","msg":"Initialize
d wire specification", "attr":{"spec":{"incomingExternalClient":{"minWireVersion":0,"maxWireVersion":21},"incomingIntern
alClient":{"minWireVersion":0,"maxWireVersion":21},"outgoing":{"minWireVersion":6,"maxWireVersion":21},"isInternalClient"
:true}}}}
{"t":{"$date":"2024-06-21T11:52:18.021-03:00"},"s":"I",  "c":"NETWORK",  "id":4648602, "ctx":"thread1","msg":"Implicit T
CP FastOpen in use."}
{"t":{"$date":"2024-06-21T11:52:18.023-03:00"},"s":"I",  "c":"REPL",       "id":5123008, "ctx":"thread1","msg":"Successful
ly registered PrimaryOnlyService", "attr":{"service":"TenantMigrationDonorService","namespace":"config.tenantMigrationDon
ors"}}
{"t":{"$date":"2024-06-21T11:52:18.023-03:00"},"s":"I",  "c":"REPL",       "id":5123008, "ctx":"thread1","msg":"Successful
ly registered PrimaryOnlyService", "attr":{"service":"TenantMigrationRecipientService","namespace":"config.tenantMigratio
nRecipients"}}
{"t":{"$date":"2024-06-21T11:52:18.023-03:00"},"s":"I",  "c":"CONTROL",  "id":5945603, "ctx":"thread1","msg":"Multi thre
ading initialized"}
{"t":{"$date":"2024-06-21T11:52:18.023-03:00"},"s":"I",  "c":"TENANT_M",  "id":7091600, "ctx":"thread1","msg":"Starting T
enantMigrationAccessBlockerRegistry"}
{"t":{"$date":"2024-06-21T11:52:18.024-03:00"},"s":"I",  "c":"CONTROL",  "id":4615611, "ctx":"initandlisten","msg":"Mong
oDB starting", "attr":{"pid":11640,"port":27017,"dbPath":"C:/data/db/", "architecture":"64-bit", "host":"DESKTOP-VGE05B2"}}

{"t":{"$date":"2024-06-21T11:52:18.024-03:00"},"s":"I",  "c":"CONTROL",  "id":23398,   "ctx":"initandlisten","msg":"Targ
et operating system minimum version", "attr":{"targetMinOS":"Windows 7/Windows Server 2008 R2"}}
{"t":{"$date":"2024-06-21T11:52:18.024-03:00"},"s":"I",  "c":"CONTROL",  "id":23403,   "ctx":"initandlisten","msg":"Buil
d Info", "attr":{"buildInfo":{"version":"7.0.11", "gitVersion":"f451220f0df2b9dfe073f1521837f8ec5c208a8c", "modules":[], "al
locator":"tcmalloc", "environment":{"distmod":"windows", "distarch":"x86_64", "target_arch":"x86_64"}}}}

```

Luego agregar en las variables de entorno – Path la ruta a la carpeta bin de la instalación.

Variables de entorno

Variables de usuario para Sofi

Variable	Valor
OneDrive	C:\Users\Sofi\OneDrive
OneDriveConsumer	C:\Users\Sofi\OneDrive
Path	C:\Users\Sofi\AppData\Local\Programs\Python\Python312\Scripts
TEMP	C:\Users\Sofi\AppData\Local\Temp
TMP	C:\Users\Sofi\AppData\Local\Temp

Nueva... Editar... Eliminar

Variables del sistema

Variable	Valor
OS	Windows_NT
Path	C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbem
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64
PROCESSOR_IDENTIFIER	Intel64 Family 6 Model 158 Stepping 9, GenuineIntel
PROCESSOR_LEVEL	6
PROCESSOR_REVISION	9e09

Nueva... Editar... Eliminar

Aceptar Cancelar

Editar variable de entorno

%SystemRoot%\system32

%SystemRoot%

%SystemRoot%\System32\Wbem

%SYSTEMROOT%\System32\WindowsPowerShell\v1.0\

%SYSTEMROOT%\System32\OpenSSH\

C:\Program Files\NVIDIA Corporation\NVIDIA NvDLISR

C:\Program Files (x86)\NVIDIA Corporation\PhysX\Common

C:\Program Files\MongoDB\Server\7.0\bin

C:\Program Files\MongoDB\mongosh-2.2.9-win32-x64\bin

C:\Program Files\PostgreSQL\16\bin

Nuevo

Modificar

Examinar...

Eliminar

Subir

Bajar

Editar texto...

Aceptar Cancelar

También elegir la opción de instalar [pgAdmin 4](#) que es el motor gráfico.

las variables de entorno – Path la ruta a la carpeta [bin](#) de la instala

IMPORTACION DE BD WORLD:
CONFIGURACIONES DE IDLE:

3.2

3.4

3.5

3.6

3.7

3.8

3.9

4.0

4.1

4.2

4.3

4.4

4.5

4.6

4.7

4.8

4.9

5.0

5.1

5.2

5.3

5.4

5.5

5.6

5.7

5.8

5.9

6.0

6.1

6.2

6.3

6.4

6.5

6.6

6.7

6.8

6.9

7.0

7.1

7.2

7.3

7.4

7.5

7.6

7.7

7.8

7.9

8.0

8.1

8.2

8.3

8.4

8.5

8.6

8.7

8.8

8.9

9.0

9.1

9.2

9.3

9.4

9.5

9.6

9.7

9.8

9.9

10.0

10.1

10.2

10.3

10.4

10.5

10.6

10.7

10.8

10.9

11.0

11.1

11.2

11.3

11.4

11.5

11.6

11.7

11.8

11.9

12.0

12.1

12.2

12.3

12.4

12.5

12.6

12.7

12.8

12.9

13.0

13.1

13.2

13.3

13.4

13.5

13.6

13.7

13.8

13.9

14.0

14.1

14.2

14.3

14.4

14.5

14.6

14.7

14.8

14.9

15.0

15.1

15.2

15.3

15.4

15.5

15.6

15.7

15.8

15.9

16.0

16.1

16.2

16.3

16.4

16.5

16.6

16.7

16.8

16.9

17.0

17.1

17.2

17.3

17.4

17.5

17.6

17.7

17.8

17.9

18.0

18.1

18.2

18.3

18.4

18.5

18.6

18.7

18.8

18.9

19.0

19.1

19.2

19.3

19.4

19.5

19.6

19.7

19.8

19.9

20.0

20.1

20.2

20.3

20.4

20.5

20.6

20.7

20.8

20.9

21.0

21.1

21.2

21.3

21.4

21.5

21.6

21.7

21.8

21.9

22.0

22.1

22.2

22.3

22.4

22.5

22.6

22.7

22.8

22.9

23.0

23.1

23.2

23.3

23.4

23.5

23.6

23.7

23.8

23.9

24.0

24.1

24.2

24.3

24.4

24.5

24.6

24.7

24.8

24.9

25.0

25.1

25.2

25.3

25.4

25.5

25.6

25.7

25.8

25.9

26.0

26.1

26.2

26.3

26.4

26.5

26.6

26.7

26.8

26.9

27.0

27.1

27.2

27.3

27.4

27.5

27.6

27.7

27.8

27.9

28.0

28.1

28.2

28.3

28.4

28.5

28.6

28.7

28.8

28.9

29.0

29.1

29.2

29.3

29.4

29.5

29.6

29.7

29.8

29.9

30.0

30.1

30.2

30.3

30.4

30.5

30.6

30.7

30.8

30.9

31.0

31.1

31.2

31.3

31.4

31.5

31.6

31.7

31.8

31.9

32.0

32.1

32.2

32.3

32.4

32.5

32.6

32.7

32.8

32.9

33.0

33.1

33.2

33.3

33.4

33.5

33.6

33.7

33.8

33.9

34.0

34.1

34.2

34.3

34.4

34.5

34.6

34.7

34.8

34.9

35.0

35.1

35.2

35.3

35.4

35.5

35.6

35.7

35.8

35.9

36.0

36.1

36.2

36.3

36.4

36.5

36.6

36.7

36.8

36.9

37.0

37.1

37.2

37.3

37.4

37.5

37.6

37.7

37.8

37.9

38.0

38.1

38.2

38.3

38.4

38.5

38.6

38.7

38.8

38.9

39.0

39.1

39.2

39.3

39.4

39.5

39.6

39.7

39.8

39.9

40.0

40.1

40.2

40.3

40.4

40.5

40.6

40.7

40.8

40.9

41.0

41.1

41.2

41.3

41.4

41.5

41.6

41.7

41.8

41.9

42.0

42.1

42.2

42.3

42.4

42.5

42.6

42.7

42.8

42.9

43.0

43.1

43.2

43.3

43.4

43.5

43.6

43.7

43.8

43.9

44.0

44.1

44.2

44.3

44.4

44.5

44.6

44.7

44.8

44.9

45.0

45.1

45.2

45.3

45.4

45.5

45.6

45.7

45.8

45.9

46.0

46.1

46.2

46.3

46.4

46.5

46.6

46.7

46.8

46.9

47.0

47.1

47.2

47.3

47.4

47.5

47.6

47.7

47.8

47.9

48.0

48.1

48.2

48.3

48.4

48.5

48.6

48.7

48.8

48.9

49.0

49.1

49.2

49.3

49.4

49.5

49.6

49.7

49.8

49.9

50.0

50.1

50.2

50.3

50.4

50.5

50.6

50.7

50.8

50.9

51.0

51.1

51.2

51.3

51.4

51.5

51.6

51.7

51.8

51.9

52.0

52.1

52.2

52.3

52.4

52.5

52.6

52.7

52.8

52.9

53.0

53.1

53.2

53.3

53.4

53.5

53.6

53.7

53.8

53.9

54.0

54.1

54.2

54.3

54.4

54.5

54.6

54.7

54.8

54.9

55.0

55.1

55.2

55.3

55.4

55.5

55.6

55.7

55.8

55.9

56.0

56.1

56.2

56.3

56.4

56.5

56.6

56.7

56.8

56.9

57.0

57.1

57.2

57.3

57.4

57.5

57.6

57.7

57.8

57.9

58.0

58.1

58.2

58.3

58.4

58.5

58.6

58.7

58.8

58.9

59.0

59.1

59.2

59.3

59.4

59.5

59.6

59.7

59.8

59.9

60.0

60.1

60.2

60.3

60.4

60.5

60.6

60.7

60.8

60.9

61.0

61.1

61.2

61.3

61.4

61.5

61.6

61.7

61.8

61.9

62.0

62.1

62.2

62.3

62.4

62.5

62.6

62.7

62.8

62.9

63.0

63.1

63.2

63.3

63.4

63.5

63.6

63.7

63.8

63.9

64.0

64.1

64.2

64.3

64.4

64.5

64.6

64.7

64.8

64.9

65.0

65.1

65.2

65.3

65.4

65.5

65.6

65.7

65.8

65.9

66.0

66.1

66.2

66.3

66.4

66.5

66.6

66.7

66.8

66.9

67.0

67.1

67.2

67.3

67.4

67.5

67.6

67.7

67.8

67.9

68.0

68.1

68.2

68.3

68.4

68.5

68.6

68.7

68.8

68.9

69.0

69.1

69.2

69.3

69.4

69.5

69.6

69.7

69.8

69.9

70.0

70.1

70.2

70.3

70.4

70.5

70.6

70.7

70.8

70.9

71.0

71.1

71.2

71.3

71.4

71.5

71.6

71.7

71.8

71.9

72.0

72.1

72.2

72.3

72.4

72.5

72.6

72.7

72.8

72.9

73.0

73.1

73.2

73.3

73.4

73.5

73.6

73.7

73.8

73.9

74.0

74.1

74.2

74.3

74.4

74.5

74.6

74.7

74.8

74.9

75.0

75.1

75.2

75.3

75.4

75.5

75.6

75.7

75.8

75.9

76.0

76.1

76.2

76.3

76.4

76.5

76.6

76.7

76.8

76.9

77.0

77.1

77.2

77.3

77.4

77.5

77.6

77.7

77.8

77.9

78.0

78.1

78.2

78.3

78.4

78.5

78.6

78.7

78.8

78.9

79.0

79.1

79.2

79.3

79.4

79.5

79.6

79.7

79.8

79.9

80.0

80.1

80.2

80.3

80.4

80.5

80.6

80.7

80.8

80.9

81.0

81.1

81.2

81.3

81.4

81.5

81.6

81.7

81.8

81.9

82.0

82.1

82.2

82.3

82.4

82.5

82.6

82.7

82.8

82.9

83.0

83.1

83.2

83.3

83.4

83.5

83.6

83.7

83.8

83.9

84.0

84.1

84.2

84.3

84.4

84.5

84.6

84.7

84.8

84.9

85.0

85.1

85.2

85.3

85.4

85.5

85.6

85.7

85.8

85.9

86.0

86.1

86.2

86.3

86.4

86.5

86.6

86.7

86.8

86.9

87.0

87.1

87.2

87.3

87.4

87.5

87.6

87.7

87.8

87.9

88.0

88.1

88.2

88.3

88.4

88.5

88.6

88.7

88.8

88.9

89.0

89.1

89.2

89.3

89.4

89.5

89.6

89.7

89.8

89.9

90.0

90.1

90.2

90.3

90.4

90.5

90.6

90.7

90.8

90.9

91.0

91.1

91.2

91.3

91.4

91.5

91.6

91.7

91.8

91.9

92.0

92.1

92.2

92.3

92.4

92.5

92.6

92.7

92.8

92.9

93.0

93.1

93.2

93.3

93.4

93.5

93.6

93.7

93.8

93.9

94.0

94.1

94.2

94.3

94.4

94.5

94.6

94.7

94.8

94.9

95.0

95.1

95.2

95.3

95.4

95.5

95.6

95.7

95.8

95.9

96.0

96.1

96.2

96.3

96.4

96.5

96.6

96.7

96.8

96.9

97.0

97.1

97.2

97.3

97.4

97.5

97.6

97.7

97.8

97.9

98.0

98.1

98.2

98.3

98.4

98.5

98.6

98.7

98.8

98.9

99.0

99.1

99.2

99.3

99.4

99.5

99.6

99.7

99.8

99.9

100.0

100.1

100.2

100.3

100.4

100.5

100.6

100.7

100.8

100.9

101.0

101.1

101.2

101.3

101.4

101.5

101.6

101.7

101.8

101.9

102.0

102.1

102.2

102.3

102.4

102.5

102.6

102.7

102.8

102.9

103.0

103.1

103.2

103.3

103.4

103.5

103.6

103.7

103.8

103.9

104.0

104.1

104.2

104.3

104.4

104.5

104.6

104.7

104.8

104.9

105.0

105.1

105.2

105.3

105.4

105.5

105.6

105.7

105.8

105.9

106.0

106.1

106.2

106.3

106.4

106.5

106.6

106.7

106.8

106.9

107.0

107.1

107.2

107.3

107.4

107.5

107.6

107.7

107.8

107.9

108.0

108.1

108.2

108.3

108.4

108.5

108.6

108.7

108.8

108.9

109.0

109.1

109.2

109.3

109.4

109.5

109.6

109.7

109.8

109.9

110.0

110.1

110.2

110.3

110.4

110.5

110.6

110.7

110.8

110.9

111.0

111.1

111.2

111.3

111.4

111.5

111.6

111.7

111.8

111.9

112.0

112.1

112.2

112.3

112.4

112.5

112.6

112.7

112.8

112.9

113.0

113.1

113.2

113.3

113.4

113.5

113.6

113.7

113.8

113.9

114.0

114.1

114.2

114.3

114.4

114.5

114.6

114.7

114.8

114.9

115.0

115.1

115.2

115.3

115.4

115.5

115.6

115.7

115.8

115.9

116

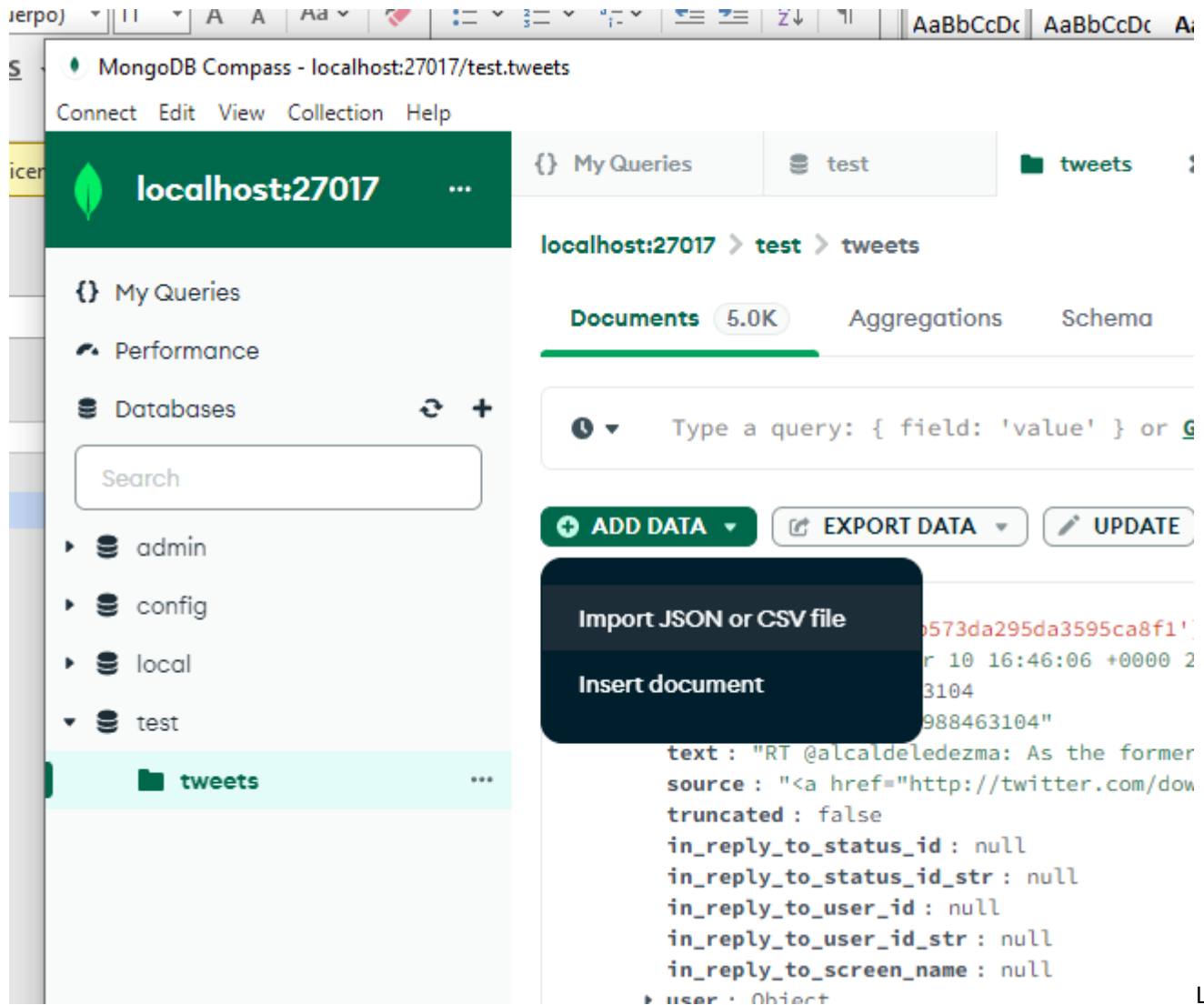
IMPORTACION DE BD TWEETS:

Abro el MongoDB Compass. Me conecto. Creo la base de datos "test" y la colección "tweets"

ADD DATA > Import JSON or CSV file > tweets.json

O

ADD DATA > Import JSON or CSV file > crisis.20190410.json



Luego

CONSULTAS POR MONGOSH:

Corro el comando mongosh para conectar:

```
❏ mongosh mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000
Microsoft Windows [Versión 10.0.19045.4529]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Users\Sofi>mongosh
Current Mongosh Log ID: 66759522a769ab2e6290defd
Connecting to:      mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.2.9
Using MongoDB:      7.0.11
Using Mongosh:      2.2.9

For mongosh info see: https://docs.mongodb.com/mongosh-shell/

-----
  The server generated these startup warnings when booting
  2024-06-21T11:52:18.337-03:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
  2024-06-21T11:52:18.337-03:00: This server is bound to localhost. Remote systems will be unable to connect to this server. Start the server with --bind_ip <address> to specify other interfaces.
  --bind_ip_all to bind to all interfaces. If this behavior is desired, start the server with --bind_ip 127.0.0.1 to disable this warning
-----

test>
```

> show dbs

Me muestra las bases de datos existentes

```
test> show dbs
admin      40.00 KiB
config     72.00 KiB
local      72.00 KiB
test       9.20 MiB
test>
```

> use test

Entro a la base de datos test

```
test> use test
already on db test
test>
```

> show collections

```
test> show collections
tweets
test>
```

<https://medium.com/@xagustin93/comandos-b%C3%A1sicos-de-mongodb-59978f2a81b7>

3.1

3.1.1

Seleccionar el id y el texto de 10 "documents"

db.collection.find(query, projection, options)

Yo necesito trabajar con el campo projection. Por lo cual si ingreso la query de esta forma:

```
> db.tweets.find({id:1,text:1}).limit(10)
```

```
test> db.tweets.find({id:1,text:1}).limit(10)
```

No me devuelve nada porque interpreta que {id:1, text:1} es una condición.

Para que eso no suceda en el primer parametro debo poner {}, ya que no hay condicion alguna y ahora si, en el segundo parametro poner la proyección, es decir, los campos que quiero que me traiga con un 1, o los campos que NO quiero que me traiga con un 0:

```
> db.tweets.find({}, {id:1, text:1}).limit(10)
```

```
test> db.tweets.find({}, {id:1, text:1}).limit(10)
[
  {
    _id: ObjectId('6674b573da295da3595ca8f1'),
    id: Long('1116019536988463104'),
    text: 'RT @alcaldeledezma: As the former Mayor of Caracas, on behalf of millions of Venezuelans, I express our heartfelt gratitude to President @r...'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f2'),
    id: Long('1116019536988463104'),
    text: 'Costa de Marfil: Observamos con pena el estancamiento político e institucional. Cuya repercusión en la crisis hum... https://t.co/kXKc5KHjPH'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f3'),
    id: Long('1116019539798646784'),
    text: '@CamilaFocareta Te vendo los míos. Tengo una crisis financiera. Tengo mas cosas.'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f4'),
    id: Long('1116019539840466944'),
    text: 'RT @amoration: We now fail at both disaster response & displacement policy with thousands of new homeless families adding to #exponential #'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f5'),
    id: Long('1116019541795061760'),
    text: 'RT @Volvei0Riqueime: Racing es campeón cada vez que en Argentina hay una crisis económica'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f6'),
    id: Long('1116019542394920960'),
    text: 'RT @elcosodelapizza: tengo que aprobar todas las materias, no engordar , hacer ejercicio, tener vida social y no tener ninguna crisis exist...'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f7'),
    id: Long('1116019542671659008'),
    text: '@UN @antonioquterres @AlAzhar Silence in the middle of the Venezuelan. Holocaust? Silence is approval of humanitari... https://t.co/sbEgMiuLQd'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f8'),
    id: Long('1116019543539879936'),
    text: '#Brexit Remain MP's made sure they got a remain PM by making sure that the decision never went to Tory membership... https://t.co/ZDihhnJdbp'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f9'),
    id: Long('1116019544173240320'),
    text: 'RT @Jenn198523: The Rio Grande Valley sector alone takes in 38% of all illegal immigration.\n' +
      '\n' +
      'This BP agent says they're taking in MIGRANT S...'
  },
  {
    _id: ObjectId('6674b573da295da3595ca8fa'),
    id: Long('1116019546387668992'),
    text: 'RT @sunchick8: @LegoBeni @Nailic15_4 @EW @Avengers @Breznican 1) Bullshit\n' +
      '2) Existential identity crisis\n' +
      '3) Her name is literally MARVEL \n' +
      '4...'
  }
]
```

3.1.2

Seleccionar los lenguajes distintos de los tweets

```
> db.tweets.distinct('lang')
```

```
test> db.tweets.distinct('lang')
[
  'ar', 'ca', 'de', 'en',
  'es', 'et', 'fr', 'hi',
  'ht', 'in', 'it', 'ja',
  'nl', 'pl', 'th', 'tl',
  'tr', 'und'
]
test>
```

<https://www.mongodb.com/docs/manual/reference/method/db.collection.distinct/>

3.1.3

Seleccionar el id, el nombre, la descripcion y la cantidad de followers de aquellos usuarios que tengan mas de 100000 followers

```
> db.tweets.find({"user.followers_count":{"$gt":10000}},{id:1, "user.name":1, "user.description":1,
"user.followers count":1})
```

```
test> db.tweets.find({"user.followers_count":{"$gt":10000}},{_id:1, "user.name":1, "user.description":1, "user.followers_count":1})
[
  {
    _id: ObjectId('6674b573da295da3595ca8f2'),
    id: Long('1116019536690544640'),
    user: {
      name: 'Jessica Carrillo',
      description: 'Periodista venezolana, fiebrúa e impertinente l RT are not endorsements l English, Spanish and Spanglish l Aquí hay de todo.',
      followers_count: 19508
    }
  },
  {
    _id: ObjectId('6674b573da295da3595ca8f4'),
    id: Long('1116019539840466944'),
    user: {
      name: 'Mark Horvath',
      description: 'Do-Gooder | Founder of @InvisiblePeople | Digital Storyteller | Marketing Consultant | Keynote Speaker https://www.patreon.com/invisiblepeople',
      followers_count: 27721
    }
  },
  {
    _id: ObjectId('6674b573da295da3595ca8fb'),
    id: Long('1116019547419623424'),
    user: {
      name: 'Carpe Covfefe ☹️',

```

<https://www.mongodb.com/docs/manual/reference/operator/query/gt/>

3.1.4

Seleccionar el id, el nombre y la cantidad de followers de los 10 usuarios con más followers ordenado de manera descendente

Para devolver 10, debo usar `sort({"atributo":tipoDeOrden})` y `limit(num):`

En tipoDeOrden puedo poner 1-> Ascendente; -1-> Descendente

```
> db.tweets.find({"user.followers_count":{"$gt":10000}},{id:1,"user.name":1,"user.followers_count":1}).sort({"user.followers_count":-1}).limit(10)
```



```

test> db.tweets.find({"user.followers_count":{"$gt":10000}},{id:1, "user.name":1, "user.followers_count":1}).sort({"user.followers_count":-1}).limit(10)
[
  {
    _id: ObjectId('6674b577da295da3595cb73a'),
    id: Long('1116022507820142592'),
    user: { name: 'Bloomberg', followers_count: 5272857 }
  },
  {
    _id: ObjectId('6674b575da295da3595cb0a0'),
    id: Long('1116021164705484800'),
    user: { name: 'El Espectador', followers_count: 4789126 }
  },
  {
    _id: ObjectId('6674b579da295da3595cbb7d'),
    id: Long('1116023550756958208'),
    user: { name: 'El Nacional', followers_count: 4756177 }
  },
  {
    _id: ObjectId('6674b573da295da3595ca9a3'),
    id: Long('1116019673336832000'),
    user: { name: 'El Nacional', followers_count: 4756164 }
  },
  {
    _id: ObjectId('6674b573da295da3595cab28'),
    id: Long('1116019989838991360'),
    user: { name: 'El Nacional', followers_count: 4756164 }
  },
  {
    _id: ObjectId('6674b573da295da3595caa13'),
    id: Long('1116019762751049728'),
    user: { name: 'AnimalPolitico.com', followers_count: 2006679 }
  },
  {
    _id: ObjectId('6674b574da295da3595cac52'),
    id: Long('1116020237298753536'),
    user: { name: 'Venezuela al Dia', followers_count: 1423039 }
  },
  {
    _id: ObjectId('6674b579da295da3595cbb86'),
    id: Long('1116023560722624512'),
    user: { name: 'CaraotaDigital', followers_count: 1301404 }
  },
  {
    _id: ObjectId('6674b579da295da3595cbb39'),
    id: Long('1116023479546056704'),
    user: { name: 'CaraotaDigital', followers_count: 1301404 }
  },
  {
    _id: ObjectId('6674b579da295da3595cbb14'),
    id: Long('1116023452924895232'),
    user: { name: 'CaraotaDigital', followers_count: 1301404 }
  }
]
test>

```

<https://www.mongodb.com/docs/manual/reference/method/cursor.sort/>

<https://www.mongodb.com/docs/manual/reference/method/cursor.limit/>

3.2

db.runCommand(

```

{
  mapReduce: <input-collection>,
  map: function() { ... },
  reduce: function(key, values) { ... },
  out: { merge: <output-collection> },
  query: <query>
}
)

```

Collection
↓

```

db.orders.mapReduce(
  map    → function() { emit( this.cust_id, this.amount ); },
  reduce → function(key, values) { return Array.sum( values ) },
  {
    query: { status: "A" },
    out: "order_totals"
  }
)

```

<https://www.mongodb.com/docs/manual/core/map-reduce/>

<https://www.mongodb.com/docs/manual/reference/command/mapReduce/#mongodb-dbcommand-dbcmd.mapReduce>

3.2.1

En base al campo \source", determinar la cantidad de usuarios que hay por cada canal.

```

db.tweets.mapReduce(
  function(){ emit( this.source, 1 ); },
  function(source, value) { return Array.sum(value) },
  {
    out: { inline : 1 }
  }
)

```

```

> db.tweets.mapReduce(function(){emit(this.source,1);},function(source,value){return
Array.sum(value)},{out:{inline:1}})

```

```

}
test> db.tweets.mapReduce(function(){emit(this.source,1);},function(source,value){return Array.sum(value)}},{out:{inline:1}})
{
  results: [
    {
      _id: '<a href="http://construyenpais.com" rel="nofollow">ConstruyenPais.com</a>',
      value: 1
    },
    {
      _id: '<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>',
      value: 818
    },
    {
      _id: '<a href="http://tapbots.com/tweetbot" rel="nofollow">Tweetbot for iOS</a>',
      value: 5
    },
    {
      _id: '<a href="http://www.nextscripts.com" rel="nofollow">NextScriptsPR02</a>',
      value: 1
    },
    {
      _id: '<a href="http://www.sendurtweet.com" rel="nofollow">SendUrTweet</a>',
      value: 1
    },
    {
      _id: '<a href="http://www.dontspreadmywealth.com" rel="nofollow">PostTwee</a>',
      value: 1
    },
    {
      _id: '<a href="http://discontents.com.au/trovenewsbot" rel="nofollow">TroveNewsBot</a>',
      value: 1
    },
    {
      _id: '<a href="http://khalilhumam.com" rel="nofollow">KhalilHumam.com</a>',
      value: 1
    },
    {
      _id: '<a href="http://www.nextscripts.com" rel="nofollow">NextScriptsPR04</a>',
      value: 1
    },
    {
      _id: '<a href="http://www.nextscripts.com" rel="nofollow">NextScriptsPR05</a>',
      value: 1
    }
  ]
}

```

Antes intente hacer lo mismo pero poniendo un “\$” antes del source:

```

db.tweets.mapReduce(function(){emit(this.$source,1);},function(source,value){return
Array.sum(value)}},{out:{inline:1}})

```

Pero eso me rompió la lógica, me devolvía la suma total de documentos (5000). No los distinguía por source.

El operador \$ en MongoDB tiene un significado especial y no puede ser utilizado como parte de los nombres de los campos en las consultas o en las operaciones de Map-Reduce. El motivo principal es que el operador \$ se utiliza para identificar operadores de MongoDB, como los operadores de consulta (\$eq, \$gt, \$in, etc.) o los operadores de actualización en las operaciones de escritura (\$set, \$push, \$addToSet, etc.).

Aquí hay algunos puntos clave sobre el operador \$ en MongoDB:

1. **Operadores de MongoDB:** Los operadores en MongoDB comienzan con el símbolo \$. Por ejemplo, \$eq se usa para realizar comparaciones de igualdad, \$gt para comparaciones de mayor que, etc. Estos operadores son fundamentales para construir consultas complejas y para actualizar documentos en MongoDB.
2. **Sintaxis de los campos:** Los nombres de los campos en MongoDB no pueden comenzar con el símbolo \$. Esto es una limitación de la sintaxis y está diseñado así para evitar confusiones entre los nombres de los campos y los operadores.
3. **Ejemplo de uso incorrecto:** Si intentas usar \$source como nombre de campo en MongoDB, MongoDB lo interpretará como si estuvieras intentando usar un operador \$source, lo cual no existe. Esto provocará errores o resultados inesperados en tus consultas.

Para ilustrar esto en el contexto de tu pregunta:

- **Uso incorrecto:** Si intentas usar db.usuarios.mapReduce con this.\$source en la función de map, MongoDB interpretará \$source como un operador especial, lo cual no es válido como nombre de campo. Esto causará errores de sintaxis o resultados incorrectos.

- **Uso correcto:** Debes utilizar nombres de campos válidos que no comiencen con \$. Por ejemplo, puedes usar `this.source` para referirte al campo `source` dentro de la función de map, como se mostró en el ejemplo anterior.

En resumen, el operador \$ en MongoDB no puede ser utilizado como parte de los nombres de los campos en tus consultas o en las operaciones de Map-Reduce. Debes utilizar nombres de campos válidos que no comiencen con este símbolo para evitar errores y garantizar el funcionamiento correcto de tus consultas y operaciones en MongoDB.

3.2.2

Determinar la cantidad de tweets por cada lenguaje.

```
db.tweets.mapReduce(
    function() { emit( this.lang, 1 ); },
    function( lang, value ) { return Array.sum(value) },
    {
        out: { inline : 1 }
    }
)
```

```
> db.tweets.mapReduce(function(){emit(this.lang,1);},function(lang,value){return Array.sum(value)},{out:{inline:1}})
```

```
test> db.tweets.mapReduce(function(){emit(this.lang,1);},function(lang,value){return Array.sum(value)},{out:{inline:1}})
DeprecationWarning: Collection.mapReduce() is deprecated. Use an aggregation instead.
See https://docs.mongodb.com/manual/core/map-reduce for details.
{
  results: [
    { _id: 'pl', value: 1 },
    { _id: 'ja', value: 4 },
    { _id: 'pl', value: 1 },
    { _id: 'tr', value: 1 },
    { _id: 'hi', value: 1 },
    { _id: 'en', value: 2448 },
    { _id: 'ht', value: 1 },
    { _id: 'nl', value: 8 },
    { _id: 'de', value: 2 },
    { _id: 'in', value: 2 },
    { _id: 'und', value: 69 },
    { _id: 'et', value: 1 },
    { _id: 'es', value: 2389 },
    { _id: 'fr', value: 25 },
    { _id: 'ar', value: 3 },
    { _id: 'tl', value: 2 },
    { _id: 'it', value: 5 },
    { _id: 'th', value: 36 },
    { _id: 'ca', value: 2 },
    { _id: 'ja', value: 4 }
  ],
  ok: 1
}
test>
```

3.2.3

*Clasificar los textos de los tweets en corto (< 10 palabras), mediano (>= 10, < 20 palabras) y largo (>= 20 palabras).
Devolver la cantidad de tweets en cada una de las clases.*

Solución final:

```

db.tweets.mapReduce(
  function() {
    var palabras = this.text.split(" ");
    var cantidadPalabras = palabras.length;

    if (cantidadPalabras < 10) {
      emit("textoCorto", 1);
    } else if (cantidadPalabras >= 20) {
      emit("textoLargo", 1);
    } else {
      emit("textoMediano", 1);
    }
  },
  function(key, values) {
    return Array.sum(values);
  },
  {
    out: { inline: 1 }
  }
)

```

Intentos previos:

Primer Intento:

Funciona correctamente pero no me hace una sumatoria total de los resultados sino que hace la clasificación individual por tweet:

```

db.tweets.mapReduce(
  function() { emit( this.text, 1 ); },
  function(text, value ){
    var palabras = text.split(" ");
    var cantidadPalabras = palabras.length;
    var clasificacion = {
      textoCorto:0,
      textoMediano:0,

```

```

        textoLargo:0
    };
    if(cantidadPalabras < 10){
        clasificacion.textoCorto += 1;
    }else if(cantidadPalabras >= 20){
        clasificacion.textoLargo +=1;
    }else{
        clasificacion.textoMediano +=1;
    };
    return clasificacion;
},
{
    out: { inline : 1 }
}
)

```

Segundo intento:

Intente usar un scope para declarar una variable global pero no puede usarse esto en mongodb directamente porque this.clasificacion no lo reconoce:

```

db.tweets.mapReduce(
    function() { emit( this.text, 1 ); },
    function(text, value ){
        var palabras = text.split(" ");
        var cantidadPalabras = palabras.length;
        if(cantidadPalabras < 10){
            this.clasificacion.textoCorto += 1;
        }else if(cantidadPalabras >= 20){
            this.clasificacion.textoLargo +=1;
        }else{
            this.clasificacion.textoMediano +=1;
        };
        return this.clasificacion;
    },

```

```

{
  scope: {
    clasificacion : {
      textoCorto:0,
      textoMediano:0,
      textoLargo:0
    }
  },
  out: { inline : 1 }
}
)

```

3.3

3.3.1

1. Listar los 10 usuarios que publicaron más tweets (ordenarlos de manera descendente por cantidad de tweets).

```

db.tweets.aggregate(
  {"$group": {"_id": "$user.name", "count": {"$sum": 1}}},
  {"$sort": {"count": -1}},
  {"$limit": 10}
)

```

```

test> db.tweets.aggregate(
... {"$group": {"_id": "$user.name", "count": {"$sum": 1}}},
...   {"$sort": {"count": -1}},
...   {"$limit": 10}
... )
[
  { _id: 'chavez vive', count: 13 },
  { _id: 'CaraotaDigital', count: 11 },
  { _id: 'ana bertha malave ', count: 11 },
  { _id: 'Aura Jimenez', count: 9 },
  { _id: 'Danielle RN 🇩🇪 🇩🇪 🇩🇪 🇩🇪 ', count: 9 },
  { _id: 'kay_12000', count: 8 },
  { _id: 'JORGE C. C. M.', count: 7 },
  { _id: 'Jubilee USA Network', count: 7 },
  { _id: 'Centauro', count: 6 },
  { _id: 'Arlex Limongi Lárez', count: 6 }
]
test>

```

3.3.2

2. Listar por lenguaje la cantidad de followers del usuario con mayor cantidad de followers que publica en ese lenguaje

```
db.tweets.aggregate(  
  {"$group": {"_id": "$lang", "maxFollowersCount": {"$max": "$user.followers_count"} }},  
  {"$sort": {"maxFollowersCount": -1}}  
)
```

```
test> db.tweets.aggregate( { "$group": { "_id": "$lang", "maxFollowersCount": { "$max": "$user.followers_count" } } }, { "$sort": { "maxFollowersCount": -1 } } )  
[  
  { _id: 'en', maxFollowersCount: 5272857 },  
  { _id: 'es', maxFollowersCount: 4780126 },  
  { _id: 'und', maxFollowersCount: 09035 },  
  { _id: 'fr', maxFollowersCount: 10424 },  
  { _id: 'it', maxFollowersCount: 10278 },  
  { _id: 'ca', maxFollowersCount: 6095 },  
  { _id: 'nl', maxFollowersCount: 1483 },  
  { _id: 'ja', maxFollowersCount: 1391 },  
  { _id: 'th', maxFollowersCount: 1362 },  
  { _id: 'in', maxFollowersCount: 1107 },  
  { _id: 'tl', maxFollowersCount: 1077 },  
  { _id: 'pl', maxFollowersCount: 879 },  
  { _id: 'et', maxFollowersCount: 493 },  
  { _id: 'de', maxFollowersCount: 331 },  
  { _id: 'ar', maxFollowersCount: 77 },  
  { _id: 'hi', maxFollowersCount: 44 },  
  { _id: 'ht', maxFollowersCount: 12 },  
  { _id: 'tr', maxFollowersCount: 0 }  
]  
test>
```

3.4