Partiendo de los datos de la discografía de álbumes de estudio de Pink Floyd wget http://es.wikipedia.org/wiki/Anexo:Discograf%C3%ADa de Pink Floyd

```
Año Album Ranking Certificación EEUU Ranking Certificación UK
1967 The Piper at the Gates of Dawn 131 6
1968 A Saucerful of Secrets - 9
1969 Music from the Film More 153 9
1969 Ummagumma (Disco 2) 74 5
1970 Atom Heart Mother 55 1
1971 Meddle 70 3
1972 Obscured by Clouds 46 6
1973 The Dark Side of the Moon 1 1
1975 Wish You Were Here 1.1
1977 Animals 3 2
1979 The Wall 1 3
1983 The Final Cut 6 1
1987 A Momentary Lapse of Reason 3 3
1994 The Division Bell 1 1
2014 The Endless River 3 1
```

Resolver las siguientes preguntas:

1. Crear un namespace llamado discografía

hbase shell

create namespace 'discografia'

```
bigdata@bigdata:~
bigdata@bigdata:~$ hbase shell
2016-09-24 13:38:08,225 WARN [main] util.NativeCodeLoader: Unable to load nativ
e-hadoop library for your platform... using builtin-java classes where applicable
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.1, r8d8a7107dc4ccbf36a92f64675dc60392f85c015, Wed Mar 30 11:19:21 CD
T 2016
hbase(main):001:0> create_namespace 'discografia'
0 row(s) in 0.3540 seconds
hbase(main):002:0>
```

describe_namespace 'discografia'

vemos creado el namespace

```
hbase(main):002:0> describe_namespace 'discografia'
DESCRIPTION
{NAME => 'discografia'}
1 row(s) in 0.2370 seconds
```

List_namespace

```
hbase(main):003:0> list_namespace
NAMESPACE
default
discografia
ejercicios
hbase
masterHadoop
5 row(s) in 0.0770 seconds
hbase(main):004:0>
```

2. Crear una tabla llamada discos que contenga sólo las columnas: grupo, nombre_disco, anio, ranking (a su vez tendrá identificadores para EEUU y UK)

aquí creammos una tabla con las columnas/familias de columnas 'id_nombre_disco', 'nombre_grupo', 'anio', 'ranking'

create 'discografia:discos', 'id_nombre_disco', 'nombre_grupo', 'anio', 'ranking'

```
hbase(main):002:0> create 'discografia:discos', 'id_nombre_disco', 'nombre_grupo', 'anio', 'ranking' 0 row(s) in 1.3550 seconds

=> Hbase::Table - discografia:discos hbase(main):003:0>
```

3. Cargar los datos en HBase (se pueden emplear los comandos put o una importación desde PIG). Nota: algunos años hay más de un disco por lo que la clave será el año y un secuencial, y no se disponen de todos los datos de rankings por lo que esos datos nulos no será necesario cargarlos

Aqui vemos que tenemmos la tabla sin entradas scan 'discografia:discos'

```
hbase(main):001:0> scan 'discografia:discos'
ROW COLUMN+CELL
0 row(s) in 0.3110 seconds
hbase(main):002:0>
```

Esta es la tabla que vamos a cargar:

			ranking	
id_Nombre_Disco	Grupo	Año	EEUU	UK
The Piper at the Gates of Dawn	Pink Floyd	1967	31	6
A Saucerful of Secrets	Pink Floyd	1968	-	9
Music from the Film More	Pink Floyd	1969	153	9
Ummagumma (Disco 2)	Pink Floyd	1969	74	5
Atom Heart Mother	Pink Floyd	1970	55	1
Meddle	Pink Floyd	1971	70	3
Obscured by Clouds	Pink Floyd	1972	46	6
The Dark Side of the Moon	Pink Floyd	1973	1	1
Wish You Were Here	Pink Floyd	1975	1	1
Animals	Pink Floyd	1977	3	2
The Wall	Pink Floyd	1979	1	3
The Final Cut	Pink Floyd	1983	6	1
A Momentary Lapse of Reason	Pink Floyd	1987	3	3
The Division Bell	Pink Floyd	1994	1	1
The Endless River	Pink Floyd	2014	3	1
39/Smooth	Green Day	1990	-	-
Kerplunk	Green Day	1992	-	-
Dookie	Green Day	1994	2	13
Insomniac	Green Day	1995	2	8
Nimrod	Green Day	1997	10	11
Warning	Green Day	2000	4	4
American Idiot	Green Day	2004	1	1
21st Century Breakdown	Green Day	2009	1	1
¡Uno!	Green Day	2012	2	2
¡Dos!	Green Day	2012	9	10
¡Tré!	Green Day	2012	13	31
Revolution Radio	Green Day	2016	-	-
Bullet in a Bible	Green Day	2005	8	6
Awesome as Fuck	Green Day	2011	14	14
On the Radio	Green Day	2012	-	-

Vamos a meter la primera entrada con clave r1

put 'discografia:discos', 'r1', 'id_nombre_disco', 'The Piper at the Gates of Dawn'

```
hbase(main):002:0> put 'discografia:discos', 'r1', 'id_nombre_disco', 'The Piper at the Gates of Dawn'
0 row(s) in 0.0400 seconds

hbase(main):003:0> scan'discografia:discos'
ROW COLUMN+CELL
r1 column=id_nombre_disco:, timestamp=1475524003413, value=The Piper at the Gates of Da
wn
1 row(s) in 0.0470 seconds

hbase(main):004:0>
```

put 'discografia:discos', 'r1', 'nombre_grupo', 'Pink Floyd'

```
hbase(main):004:0> put 'discografia:discos', 'r1', 'nombre_grupo', 'Pink Floyd'
0 row(s) in 0.0080 seconds
hbase(main):005:0> scan'discografia:discos'
ROW COLUMN+CELL
r1 column=id_nombre_disco:, timestamp=1475524003413, value=The Piper at the Gates of Da
wn
r1 column=nombre_grupo:, timestamp=1475524199024, value=Pink Floyd
1 row(s) in 0.0400 seconds
hbase(main):006:0>
```

put 'discografia: discos', 'r1', 'anio', '1967'

```
hbase(main):006:0> put 'discografia:discos', 'r1', 'anio', '1967'
0 row(s) in 0.0160 seconds

hbase(main):007:0> scan'discografia:discos'
ROW COLUMN+CELL
r1 column=anio:, timestamp=1475524354664, value=1967
r1 column=id_nombre_disco:, timestamp=1475524003413, value=The Piper at the Gates of Da
wn
r1 column=nombre_grupo:, timestamp=1475524199024, value=Pink Floyd
1 row(s) in 0.0430 seconds

hbase(main):008:0>
```

put 'discografia: discos', 'r1', 'ranking: eeuu', '31'

```
hbase(main):008:0> put 'discografia:discos', 'r1', 'ranking:eeuu', '31'
0 row(s) in 0.0220 seconds

hbase(main):009:0> scan'discografia:discos'
ROW COLUMN+CELL
r1 column=anio:, timestamp=1475524354664, value=1967
r1 column=id_nombre_disco:, timestamp=1475524003413, value=The Piper at the Gates of Da
wn
r1 column=nombre_grupo:, timestamp=1475524199024, value=Pink Floyd
r1 column=ranking:eeuu, timestamp=1475524597305, value=31

1 row(s) in 0.0380 seconds

hbase(main):010:0>
```

put 'discografia:discos', 'r1', 'ranking:uk', '6'

Ya tenemos la primera linea. Vamos a por las demás

```
put 'discografia: discos', 'r1', 'id_nombre_disco', ' The Piper at the Gates of Dawn' put 'discografia: discos', 'r2', 'id_nombre_disco', ' A Saucerful of Secrets' put 'discografia: discos', 'r3', 'id_nombre_disco', ' Music from the Film More' put 'discografia: discos', 'r4', 'id_nombre_disco', ' Atom Heart Mother' put 'discografia: discos', 'r6', 'id_nombre_disco', ' Atom Heart Mother' put 'discografia: discos', 'r6', 'id_nombre_disco', ' Meddle' put 'discografia: discos', 'r8', 'id_nombre_disco', ' Obscured by Clouds' put 'discografia: discos', 'r8', 'id_nombre_disco', ' The Dark Side of the Moon' put 'discografia: discos', 'r10', 'id_nombre_disco', ' Wish You Were Here' put 'discografia: discos', 'r11', 'id_nombre_disco', ' The Wall' put 'discografia: discos', 'r11', 'id_nombre_disco', ' The Final Cut' put 'discografia: discos', 'r12', 'id_nombre_disco', ' The Final Cut' put 'discografia: discos', 'r13', 'id_nombre_disco', ' The Division Bell' put 'discografia: discos', 'r16', 'id_nombre_disco', ' The Indless River' put 'discografia: discos', 'r16', 'id_nombre_disco', ' The Indless River' put 'discografia: discos', 'r18', 'id_nombre_disco', 'Bookie' put 'discografia: discos', 'r18', 'id_nombre_disco', 'Nerplunk' put 'discografia: discos', 'r18', 'id_nombre_disco', 'Nookie' put 'discografia: discos', 'r20', 'id_nombre_disco', 'Nimrod' put 'discografia: discos', 'r21', 'id_nombre_disco', 'Nimrod' put 'discografia: discos', 'r22', 'id_nombre_disco', 'Warning' put 'discografia: discos', 'r22', 'id_nombre_disco', 'American Idiot' put 'discografia: discos', 'r22', 'id_nombre_disco', 'Plos!' put 'discografia: discos', 'r26', 'id_nombre_disco', 'Ipos!' put 'discografia: discos', 'r26', 'id_nombre_disco', 'Revolution Radio' put 'discografia: discos', 'r26', 'id_nombre_disco', 'Revolution Radio' put 'discografia: discos', 'r28', 'id_nombre_disco', 'Revolution Radio' put 'discografia: discos', 'r28', 'id_nombre_disco', 'Awesome as Fuck' put 'discografia: discos', 'r28', 'id_nombre_disco', 'On the Radio'
```

```
put 'discografia:discos', 'r1', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r2', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r4', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r4', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r6', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r6', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r8', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r8', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r8', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r10', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r10', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r11', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r12', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r13', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r14', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r16', 'nombre_grupo', 'Pink Floyd'
put 'discografia:discos', 'r16', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r18', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r18', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r20', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r21', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r21', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r22', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r22', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r24', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r26', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r27', 'nombre_grupo', 'Green Day'
put 'discografia:discos', 'r28', 'nombre_grupo', 'Green Day'
put 'discografia:discos',
```

```
put 'discografia:discos', 'r1', 'anio', 1967
put 'discografia:discos', 'r2', 'anio', 1968
put 'discografia: discos', 'r3', 'anio', 1969
put 'discografia:discos', 'r4', 'anio', 1969
put 'discografia:discos', 'r5', 'anio', 1970
put 'discografia:discos', 'r6', 'anio', 1971
put 'discografia:discos', 'r7', 'anio', 1972
put 'discografia:discos', 'r8', 'anio', 1973
put 'discografia:discos', 'r9', 'anio', 1975
put 'discografia:discos', 'r10', 'anio', 1977
put 'discografia:discos', 'r11', 'anio', 1979
put 'discografia:discos', 'r12', 'anio', 1983
put 'discografia:discos', 'r13', 'anio', 1987
put 'discografia:discos', 'r14', 'anio', 1994
put 'discografia: discos', 'r15', 'anio', 2014
put 'discografia:discos', 'r16', 'anio', 1990
put 'discografia:discos', 'r17', 'anio', 1992
put 'discografia:discos', 'r18', 'anio', 1994
put 'discografia: discos', 'r19', 'anio', 1995
put 'discografia:discos', 'r20', 'anio', 1997
put 'discografia:discos', 'r21', 'anio', 2000
put 'discografia:discos', 'r22', 'anio', 2004
put 'discografia: discos', 'r23', 'anio', 2009
put 'discografia: discos', 'r24', 'anio', 2012
put 'discografia:discos', 'r25', 'anio', 2012
put 'discografia:discos', 'r26', 'anio', 2012
put 'discografia:discos', 'r27', 'anio', 2016
put 'discografia: discos', 'r28', 'anio', 2005
put 'discografia:discos', 'r29', 'anio', 2011
put 'discografia:discos', 'r30', 'anio', 2012
```

```
put 'discografia:discos', 'r1', 'ranking:eeuu', 31
put 'discografia:discos', 'r3', 'ranking:eeuu', 153
put 'discografia:discos', 'r4', 'ranking:eeuu', 74
put 'discografia:discos', 'r5', 'ranking:eeuu', 55
put 'discografia:discos', 'r6', 'ranking:eeuu', 70
put 'discografia:discos', 'r7', 'ranking:eeuu', 46
put 'discografia:discos', 'r8', 'ranking:eeuu', 1
put 'discografia:discos', 'r9', 'ranking:eeuu', 1
put 'discografia:discos', 'r10', 'ranking:eeuu', 3
put 'discografia:discos', 'r11', 'ranking:eeuu', 1
put 'discografia:discos', 'r12', 'ranking:eeuu', 6
put 'discografia:discos', 'r13', 'ranking:eeuu', 3
put 'discografia:discos', 'r14', 'ranking:eeuu', 1
put 'discografia:discos', 'r15', 'ranking:eedd', '
put 'discografia:discos', 'r18', 'ranking:eeuu', 2
put 'discografia:discos', 'r19', 'ranking:eeuu', 2
put 'discografia:discos', 'r20', 'ranking:eeuu', 10
put 'discografia:discos', 'r21', 'ranking:eeuu', 4
put 'discografia:discos', 'r22', 'ranking:eeuu', 1
put 'discografia:discos', 'r23', 'ranking:eeuu', 1
put 'discografia:discos', 'r24', 'ranking:eeuu', 2
put 'discografia:discos', 'r25', 'ranking:eeuu', 9
put 'discografia:discos', 'r26', 'ranking:eeuu', 9
put 'discografia:discos', 'r26', 'ranking:eeuu', 13
put 'discografia:discos', 'r28', 'ranking:eeuu', 8
put 'discografia:discos', 'r29', 'ranking:eeuu', 14
put 'discografia:discos', 'r1', 'ranking:uk', 6
put 'discografia:discos', 'r2', 'ranking:uk', 9
put 'discografia:discos', 'r3', 'ranking:uk', 9
put 'discografia:discos', 'r4', 'ranking:uk', 5
put 'discografia:discos', 'r5', 'ranking:uk', 1
put 'discografia:discos', 'r6', 'ranking:uk', 3
put 'discografia:discos', 'r7', 'ranking:uk', 6
put 'discografia:discos', 'r8', 'ranking:uk', 1
put 'discografia:discos', 'r9', 'ranking:uk', 1
put 'discografia:discos', 'r10', 'ranking:uk', 2
put 'discografia:discos', 'r11', 'ranking:uk', 3
put 'discografia:discos', 'r12', 'ranking:uk', 1
put 'discografia:discos', 'r13', 'ranking:uk', 3
put 'discografia:discos', 'r14', 'ranking:uk', 1
put 'discografia:discos', 'r15', 'ranking:uk', 1
put 'discografia:discos', 'r18', 'ranking:uk', 13
put 'discografia:discos', 'r19', 'ranking:uk', 8
put 'discografia:discos', 'r20', 'ranking:uk', 11
put 'discografia:discos', 'r21', 'ranking:uk', 4
put 'discografia:discos', 'r22', 'ranking:uk', 1
put 'discografia: discos', 'r23', 'ranking: uk', 1
put 'discografia:discos', 'r24', 'ranking:uk', 2
put 'discografia:discos', 'r25', 'ranking:uk', 10
put 'discografia:discos', 'r26', 'ranking:uk', 31
put 'discografia:discos', 'r28', 'ranking:uk', 6
put 'discografia:discos', 'r29', 'ranking:uk', 14
```

```
blase(main):022:0> put 'discografia:discos', 'r19', 'id_nombre_disco', 'Insomniac' or ow(s) in 0.0120 seconds

hbase(main):023:0> put 'discografia:discos', 'r20', 'id_nombre_disco', 'Nimrod' orow(s) in 0.0180 seconds

hbase(main):023:0> put 'discografia:discos', 'r21', 'id_nombre_disco', 'Warning' or ow(s) in 0.0270 seconds

hbase(main):025:0> put 'discografia:discos', 'r22', 'id_nombre_disco', 'American Idiot' orow(s) in 0.0100 seconds

hbase(main):026:0> put 'discografia:discos', 'r22', 'id_nombre_disco', '21st Century Breakdown' orow(s) in 0.0170 seconds

hbase(main):027:0> put 'discografia:discos', 'r24', 'id_nombre_disco', ';Uno!' orow(s) in 0.0270 seconds

hbase(main):027:0> put 'discografia:discos', 'r24', 'id_nombre_disco', ';Uno!' orow(s) in 0.0200 seconds

hbase(main):028:0> put 'discografia:discos', 'r25', 'id_nombre_disco', ';Dos!' orow(s) in 0.0040 seconds

hbase(main):029:0> put 'discografia:discos', 'r26', 'id_nombre_disco', ';Tré!' orow(s) in 0.0050 seconds

hbase(main):030:0> put 'discografia:discos', 'r27', 'id_nombre_disco', 'Revolution Radio' orow(s) in 0.0210 seconds

hbase(main):031:0> put 'discografia:discos', 'r28', 'id_nombre_disco', 'Bullet in a Bible ' orow(s) in 0.0210 seconds

hbase(main):032:0> put 'discografia:discos', 'r29', 'id_nombre_disco', 'Awesome as Fuck ' orow(s) in 0.0210 seconds
```

4. Escribir la instrucción y el resultado de consultar el álbum del año 1968

Primero vamos a ver que entrada, que clave tiene en año 1968.

Para ello salimos de hbase (exit) y arrancamos:

start-hvbase.sh

```
bigdata@bigdata:~$ start-hbase.sh
starting master, logging to /home/bigdata/hbase/logs/hbase-bigdata-master-bigdat
a.out
bigdata@bigdata:~$ hbase shell
2016-10-04 00:56:06,362 WARN [main] util.NativeCodeLoader: Unable to load nativ
e-hadoop library for your platform... using builtin-java classes where applicabl
e
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.1, r8d8a7107dc4ccbf36a92f64675dc60392f85c015, Wed Mar 30 11:19:21 CD
T 2016
```

volvemos a entrar en hbase (hbase shell)

y ejecutamos busqueda:

scan 'discografia:discos', { COLUMNS => 'anio', FILTER => "ValueFilter(=, 'binary:1968')" }

YAasabemos que es r2, entonces ejecutamos:

get 'discografia:discos','r2'

5. Añadir la información de los álbumes de estudio de GreenDay wget https://en.wikipedia.org/wiki/Green Day discography

Ya lo hemos hecho en los puntos anteriores.

6. Ejecuta varias instrucciones (describiéndolas) para analizar datos anteriores

Vamos a verr el ranking en EEUUde la entrada con clave r1

get 'discografia:discos','r1', 'ranking:eeuu'

```
hbase(main):001:0> get 'discografia:discos','r1', 'ranking:eeuu'
COLUMN CELL
ranking:eeuu timestamp=1475527861393, value=31
1 row(s) in 0.3820 seconds
```

Vamos a ver el ranking en UK de la entrada con clave r15

get 'discografia:discos','r15', 'ranking:uk'

```
hbase(main):002:0> get 'discografia:discos','r15', 'ranking:uk'
COLUMN CELL
ranking:uk timestamp=1475527962232, value=1
1 row(s) in 0.0200 seconds
```

Vamos a ver los datos de la entrada con clave r25 get 'discografia:discos','r25'

```
hbase(main):004:0> get 'discografia:discos','r25'

COLUMN

anio:

id_nombre_disco:

nombre_grupo:

ranking:eeuu

ranking:uk

5 row(s) in 0.0300 seconds

CELL

timestamp=1475527600338, value=2012

timestamp=1475527514735, value=2012

timestamp=1475527514735, value=2012

timestamp=1475527223181, value=Green Day

timestamp=1475527862177, value=9

timestamp=1475527962496, value=10
```

Ver que filas son del grupo Green Day

scan 'discografia:discos', { COLUMNS => 'nombre grupo', FILTER => "ValueFilter(=, 'binary:Green Day')" }

```
hbase(main):005:0> scan 'discografia:discos', {    COLUMNS => 'nombre_grupo',    FILTER => "ValueFilter( =, 'binary:Green Da
                                      COLUMN+CELL
 г16
                                      column=nombre_grupo:, timestamp=1475527155945, value=Green Day
г17
г18
                                      column=nombre_grupo:, timestamp=1475527176205, value=Green Day column=nombre_grupo:, timestamp=1475527183661, value=Green Day
 г19
г20
                                      column=nombre_grupo:, timestamp=1475527189884, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527196332, value=Green Day
 r21
r22
r23
r24
r25
                                      column=nombre_grupo:, timestamp=1475527200820, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527208314, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527213457, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527217752, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527223181, value=Green Day
 г26
                                      column=nombre_grupo:, timestamp=1475527229025, value=Green Day
                                      column=nombre_grupo:, timestamp=1475527244673, value=Green Day
 г27
                                      column=nombre_grupo:, timestamp=1475527244704, value=Green Day column=nombre_grupo:, timestamp=1475527244717, value=Green Day column=nombre_grupo:, timestamp=1475527244736, value=Green Day
 г28
 5 row(s) in 0.1360 seconds
```

Ver que filas llegaron al nº 1 en eeuu

scan 'discografia:discos', { COLUMNS => 'ranking:eeuu', FILTER => "ValueFilter(=, 'binary:1')" }

Ver que filas no llegaron al nº 20 en uk

scan 'discografia:discos', { COLUMNS => 'ranking:uk', FILTER => "ValueFilter(<, 'binary:20')" }