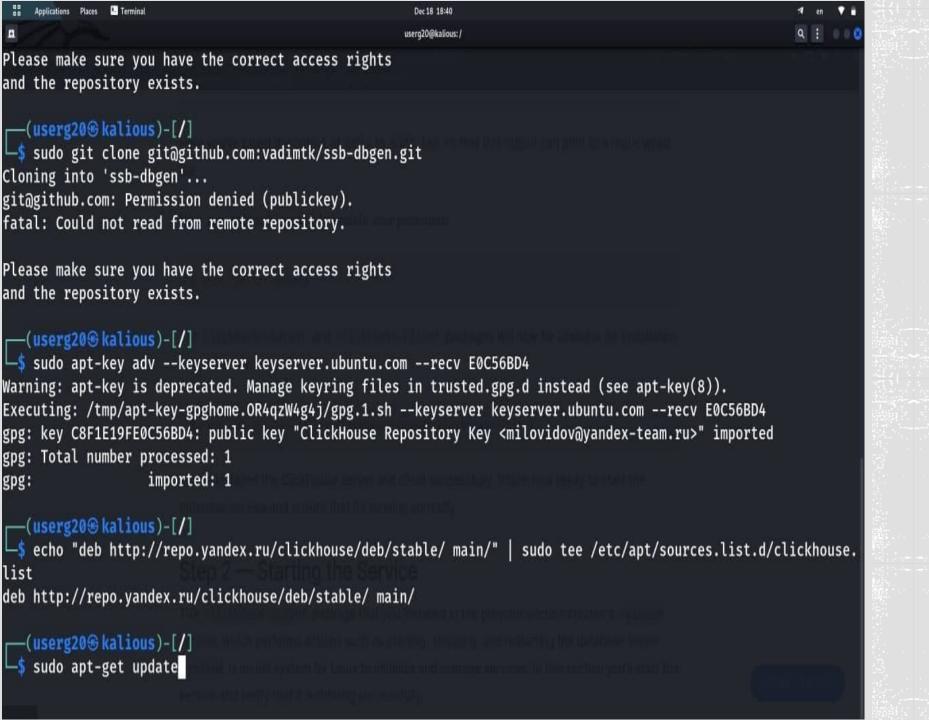
پروژه انبار داده با شمای سناره ای

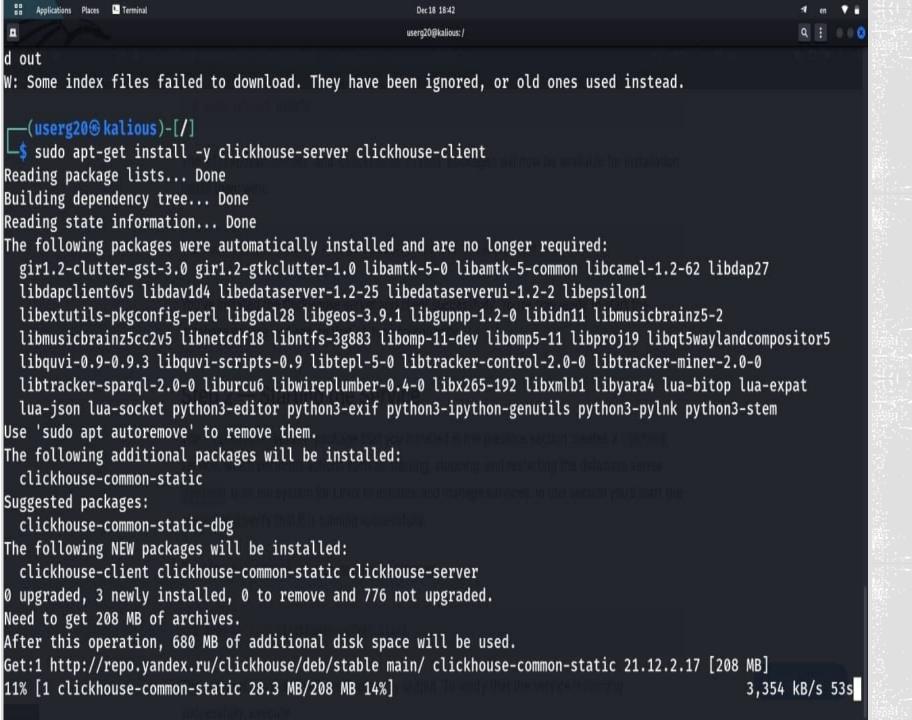


علیرضا عباسی نوخندان , یونس مجمودی (Group 20)



کلید را تعریف و در مسیر etc/apt/source.list.d/clickhouse ریپاسیتوری را ادد می کنیم.





در اینجا کلیک هاوس را نصب می کنیم.



```
Applications Places - Terminal
                                                      Dec 18 18:45
Processing triggers for kali-menu (2021.4.1) ...
  —(userg20⊛ kalious)-[/]
sudo service clickhouse-server start
 —(userg20⊕kalious)-[/]
sudo service clickhouse-server status
clickhouse-server.service - ClickHouse Server (analytic DBMS for big data)
     Loaded: loaded (/etc/systemd/system/clickhouse-server.service; enabled; vendor preset: disabled)
    Active: active (running) since Sat 2021-12-18 18:44:46 GMT; 9s ago
   Main PID: 4693 (clckhouse-watch)
      Tasks: 207 (limit: 14122)
    Memory: 134.8M
        CPU: 826ms
    CGroup: /system.slice/clickhouse-server.service
             ├─4693 clickhouse-watchdog "" "" "" "" "" ""
                                                            --config=/etc/clickhouse-server/config.xml --pid-f>
             └─4694 /usr/bin/clickhouse-server --config=/etc/clickhouse-server/config.xml --pid-file=/run/click>
Dec 18 18:44:46 kalious systemd[1]: Started ClickHouse Server (analytic DBMS for big data).
Dec 18 18:44:46 kalious clickhouse-server[4693]: Processing configuration file '/etc/clickhouse-server/config.x
Dec 18 18:44:46 kalious clickhouse-server[4693]: Logging trace to /var/log/clickhouse-server/clickhouse-server.>
Dec 18 18:44:46 kalious clickhouse-server[4693]: Logging errors to /var/log/clickhouse-server/clickhouse-server
Dec 18 18:44:46 kalious clickhouse-server[4694]: Processing configuration file '/etc/clickhouse-server/config.x
Dec 18 18:44:46 kalious clickhouse-server[4694]: Saved preprocessed configuration to '/var/lib/clickhouse/prepr
Dec 18 18:44:46 kalious clickhouse-server[4694]: Processing configuration file '/etc/clickhouse-server/users.xm>
Dec 18 18:44:46 kalious clickhouse-server[4694]: Merging configuration file '/etc/clickhouse-server/users.d/def
Dec 18 18:44:46 kalious clickhouse-server[4694]: Saved preprocessed configuration to '/var/lib/clickhouse/prepr>
lines 1-20/20 (END)
```

سرویس کلیک هاوس را استارت و برای اطمینان حاصل شدن از نصب کامل و صحیح clickhouse باید از دستور زیر استفاده کنیم:

sudo service
clickhouse-server
status

تا با پیغام زیر روبرو شویم که نشان دهنده نصب درست clickhouse است:

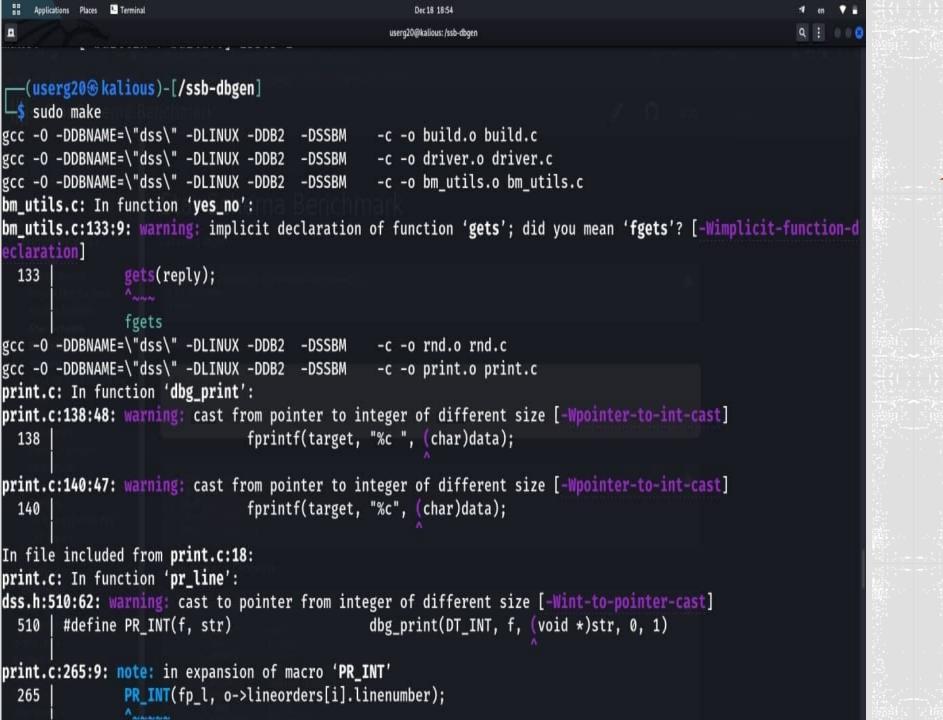
Active (running)



```
(userg20% kalious)-[/]
sudo git clone https://github.com/vadimtk/ssb-dbgen.git
Cloning into 'ssb-dbgen'...
remote: Enumerating objects: 50, done.
remote: Total 50 (delta 0), reused 0 (delta 0), pack-reused 50
Receiving objects: 100% (50/50), 96.51 KiB | 457.00 KiB/s, done.
Resolving deltas: 100% (12/12), done.
__(userg20⊕ kalious)-[/]
_$
```

پس از نصب کلیک هاوس اکنون ریپاسیتوری مربوطه را از گیت کلون می کنیم





پس از مراحل فوق باید با استفاده از دستور **make** در دایر کتوری dbgen فایل را ایجاد کنیم.

```
/usr/bin/ld: bm_utils.o: in function `yes_no':
bm_utils.c:(.text+0x4c): warning: the `gets' function is dangerous and should not be used.
gcc -O -DDBNAME=\"dss\" -DLINUX -DDB2 -DSSBM -c -o qgen.o qgen.c
qgen.c: In function 'process_options':
qgen.c:275:19: warning: implicit declaration of function 'getopt' [-Wimplicit-function-declaration]
           while((flag = getopt(cnt, args, "ab:cdhi:n:Nl:o:p:r:s:t:vx")) != -1)
 275
agen.c: At top level:
qgen.c:388:1: warning: return type defaults to 'int' [-Wimplicit-int]
      main(int ac, char **av)
In file included from dss.h:48,
                from qgen.c:20:
ggen.c: In function 'main':
config.h:165:25: warning: implicit declaration of function 'getpid' [-Wimplicit-function-declaration]
       #define DSS PROC
 165
ggen.c:409:54: note: in expansion of macro 'DSS_PROC'
                       rndm = (long)((unsigned)time(NULL) * DSS_PROC);
 409
gcc -O -DDBNAME=\"dss\" -DLINUX -DDB2 -DSSBM
                                                -c -o varsub.o varsub.c
gcc -O -DDBNAME=\"dss\" -DLINUX -DDB2 -DSSBM -O -o ggen build.o bm utils.o ggen.o rnd.o varsub.o text.o bcd2.o
permute.o speed seed.o -lm
/usr/bin/ld: bm_utils.o: in function `yes_no':
bm utils.c:(.text+0x4c): warning: the `gets' function is dangerous and should not be used.
__(userg20⊕ kalious)-[/ssb-dbgen]
```

Dec 18 18:53

Applications Places - Terminal

سپس پس از مراحل فوق باید با استفاده از دستور make در دایر کتوری dbgen فایل را ایجاد کنیم.



برای وارد کردن داده ها در جداول در خط فرمان دستورات زیر را وارد میکنیم:

\$./dbgen -s 1000 -T c

\$./dbgen -s 1000 -T 1

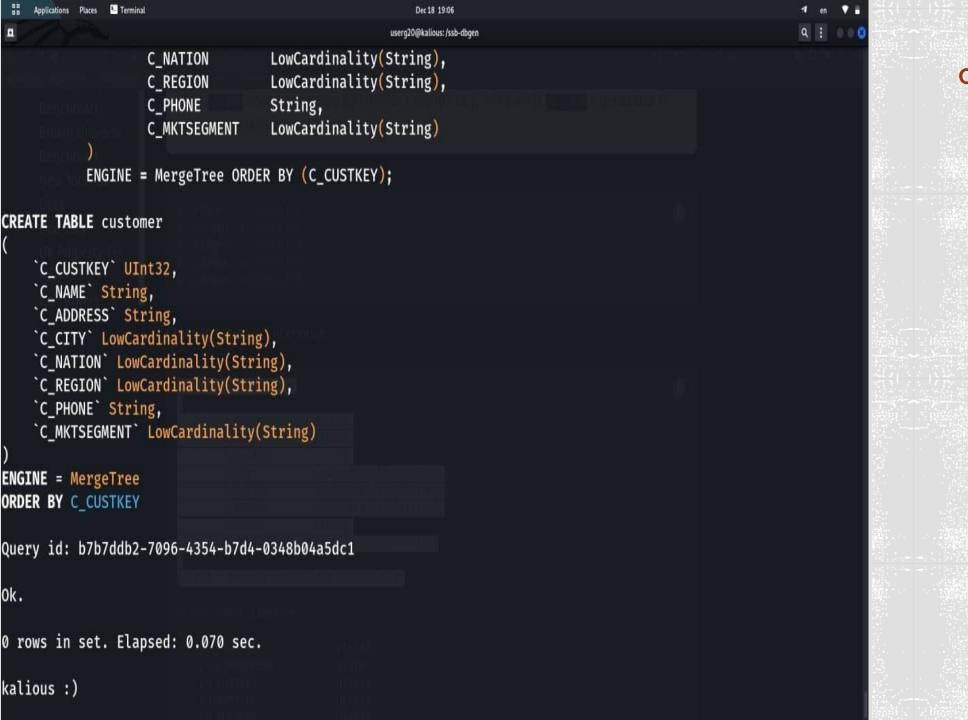
\$./dbgen -s 1000 -T p

\$./dbgen -s 1000 -T s

این کدها برای قرار دادن مقادیر و داده ها در جدول است.

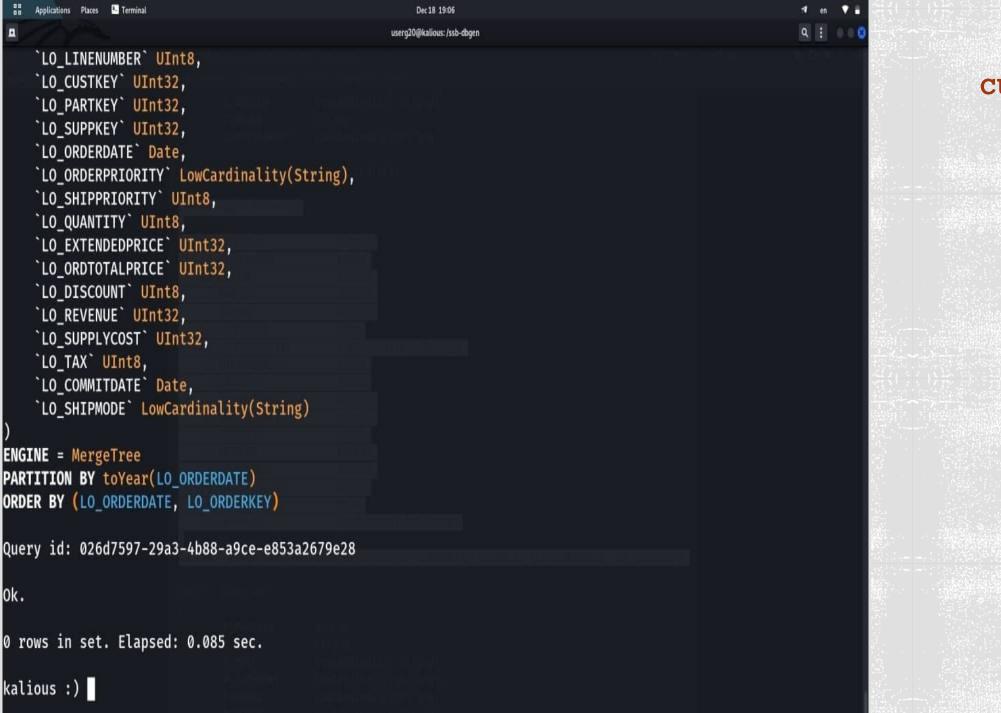
عدد 1000 ع- نشان دهنده آن است که جدول ایجاد شده دارای شش میلیارد تاپل است؛ که میتوان برای جلوگیری از اتلاف فضا از اعداد کوچکتری استفاده میکنیم. (ما از عدد ۱ عاستفاده کردیم که میتواند جداولی با حداکثر شش میلیون تاپل ایجاد کند).

حروف لاتین آخرهر لاین نشان دهنده نام هر جدول است که با حرف اول نام آن جدول مشخص شده اند.



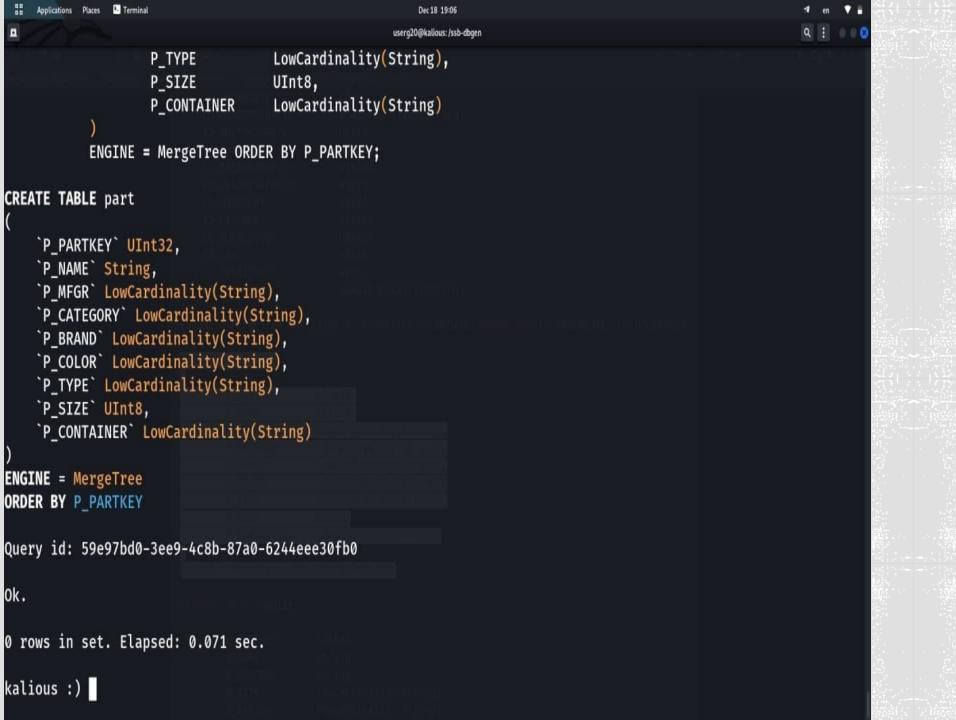
lipse ایجاد جدول customer



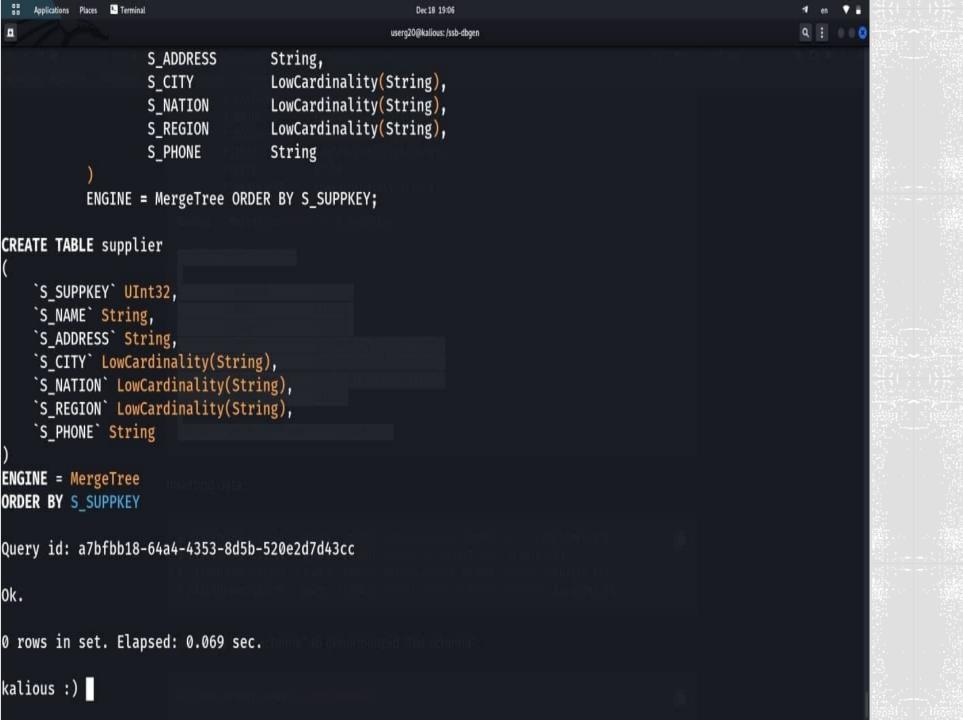


customer ایجاد جدول



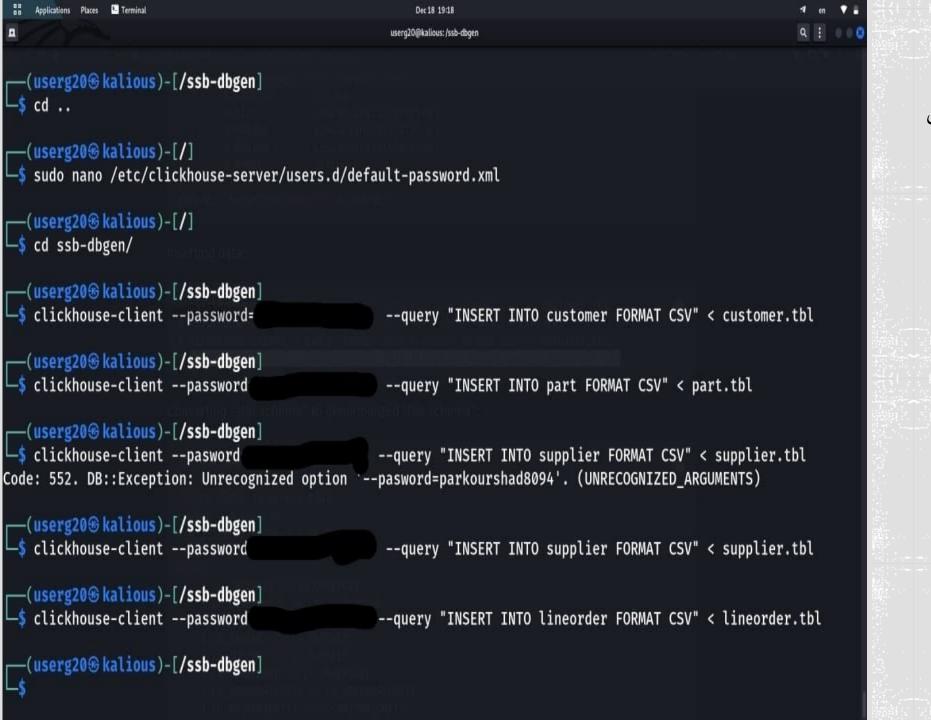


part ایجاد جدول



supplier ایجاد جدول





در این مرحله داده را از فایل هایی با فرمت **۷۵ا**خوانده و در جداول مربوطه درج می کنیم.



شمای جداول را از استار به فلت کانورت می کنیم.



شمای جداول را از استار به فلت کانورت می کنیم.



```
Applications Places - Terminal
                                              Dec 21 11:47
   s.S_PHONE AS S_PHONE,
   p.P_NAME AS P_NAME,
   p.P_MFGR AS P_MFGR,
   p.P_CATEGORY AS P_CATEGORY,
   p.P_BRAND AS P_BRAND,
   p.P_COLOR AS P_COLOR,
   p.P_TYPE AS P TYPE,
   p.P_SIZE AS P_SIZE,
   p.P_CONTAINER AS P_CONTAINER
FROM lineorder AS l
INNER JOIN customer AS c ON c.C_CUSTKEY = l.LO_CUSTKEY
INNER JOIN supplier AS s ON s.S_SUPPKEY = l.LO_SUPPKEY
INNER JOIN part AS p ON p.P_PARTKEY = 1.LO_PARTKEY
Query id: 41b14dd5-09a9-4839-878b-452feb5e16ef
Ok.
0 rows in set. Elapsed: 52.263 sec. Processed 6.23 million rows, 267.75 MB (119.27 tho
usand rows/s., 5.12 MB/s.)
```

شمای جداول را از استار به فلت کانورت می کنیم.



```
Applications Places Terminal
                                                                                                                ∢ fa ♥ ...
                                                         Dec 18 19:47
                                                                                                                Q : 008
                                                      userg20@kalious:/ssb-dbgen
               S CITY ASC.
               P BRAND ASC;
SELECT
   toYear(LO ORDERDATE) AS year,
   S_CITY,
   P_BRAND,
   sum(LO_REVENUE - LO_SUPPLYCOST) AS profit
FROM lineorder_flat
WHERE (S_NATION = 'UNITED STATES') AND ((year = 1997) OR (year = 1998)) AND (P_CATEGORY = 'MFGR#14')
GROUP BY
   year,
   S_CITY,
   P_BRAND
ORDER BY
   year ASC,
   S CITY ASC,
   P BRAND ASC
Query id: c7b5e7ee-2865-400a-b8c8-208b98373bc1
—year——S_CITY—
                     --P_BRAND-
                                       -profit—
 1997
         UNITED ST0
                       MFGR#141
                                     46368864
         UNITED STØ
 1997
                       MFGR#1410
                                     98883339
         UNITED STØ
                       MFGR#1411
 1997
                                    108499200
         UNITED ST0
 1997
                       MFGR#1412
                                     54712353
 1997
         UNITED ST0
                       MFGR#1413
                                    122258439
 1997
         UNITED STØ
                       MFGR#1414
                                     72712764
 1997
                                     19257003
         UNITED STØ
                       MFGR#1415
```

کویری ۴.۳: این کوئری سال.شهر و مجموع سود ها را به شرط این که اهل ایالات متحده باشند و سال ۱۹۹۸ یا ۱۹۹۷ باشد و در دسته بندی MFGR#14 قرار بگیرند بر اساس سال و شهر و برند گروهبندی کرده و به ترتیب صعودی سال سپس در صورت تساوی سال به ترتیب صعودی شهر و سپس صعودی برند نشان می دهد.

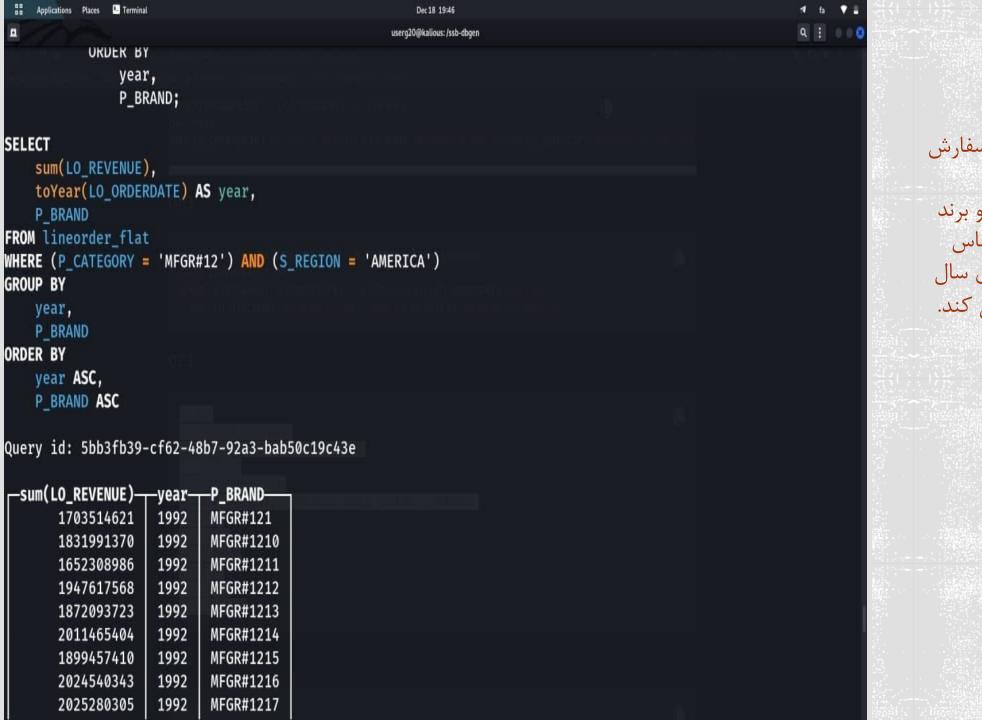


694 rows in set. Elapsed: 0.049 sec. Processed 3.59 million rows, 55.62 MB (72.84 million rows/s., 1.13 GB/s.)

kalious :)

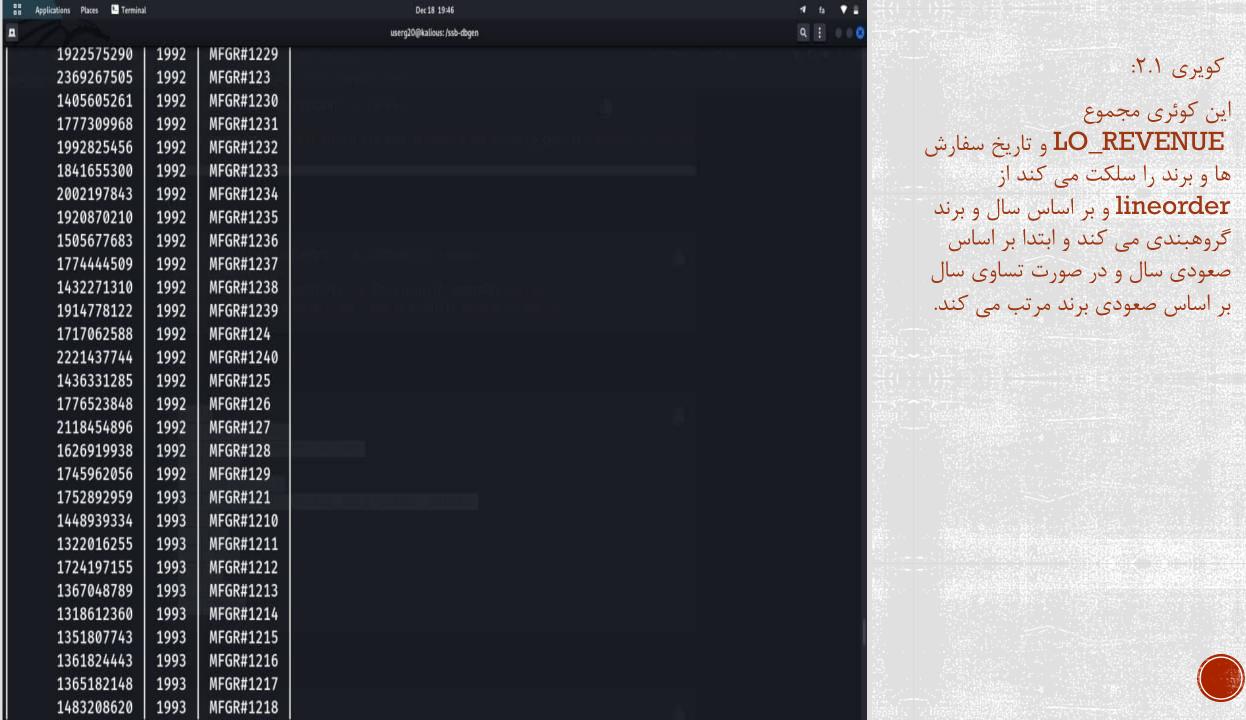
کویری ۴.۳: این کوئری سال.شهر و مجموع سود ها را به شرط این که اهل ابالات متحده باشند و سال ۱۹۹۸ با ۱۹۹۷ باشد و در دسته بندی MFGR#14 قرار بگیرند بر اساس سال و شهر و برند گروهبندی کرده و به ترتیب صعودی سال سپس در صورت تساوی سال به ترتیب صعودی شهر و سیس صعودی برند نشان می دهد.

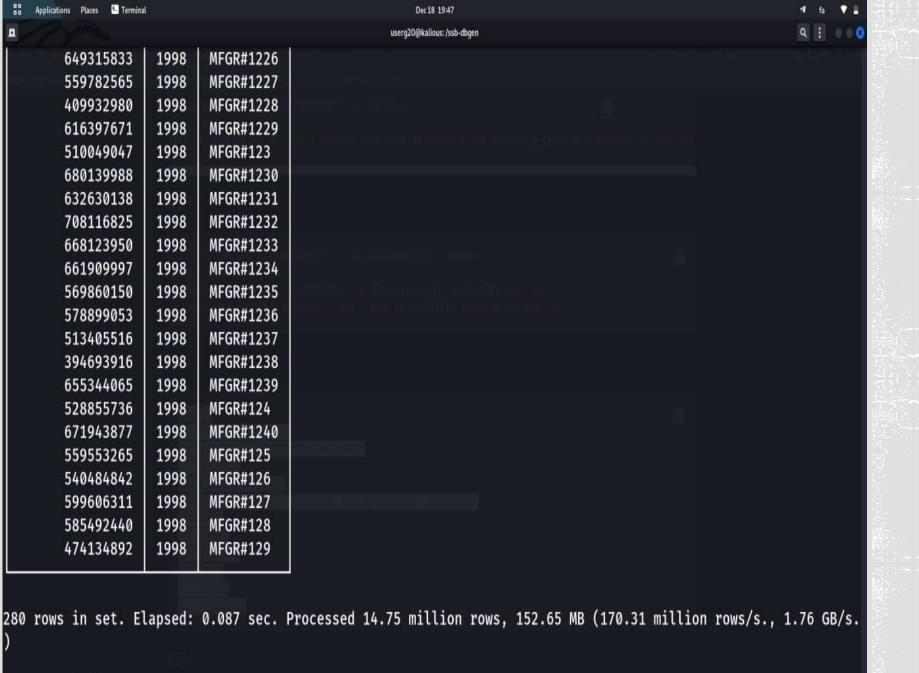




کویری ۲.۱:

این کوئری مجموع LO_REVENUE و تاریخ سفارش ها و برند را سلکت می کند از lineorder و بر اساس سال و برند گروهبندی می کند و ابتدا بر اساس صعودی سال و در صورت تساوی سال بر اساس صعودی برند مرتب می کند.





kalious :)

کویری ۲.۱:

این کوئری مجموع LO_REVENUE و تاریخ سفارش ها و برند را سلکت می کند از lineorder و بر اساس سال و برند گروهبندی می کند و ابتدا بر اساس صعودی سال و در صورت تساوی سال بر اساس صعودی برند مرتب می کند.