

CASSANDRA

Code to create streaming keyspace:

```
CREATE KEYSPACE streaming  
WITH REPLICATION = { 'class' : 'SimpleStrategy', 'replication_factor' : 1 };
```

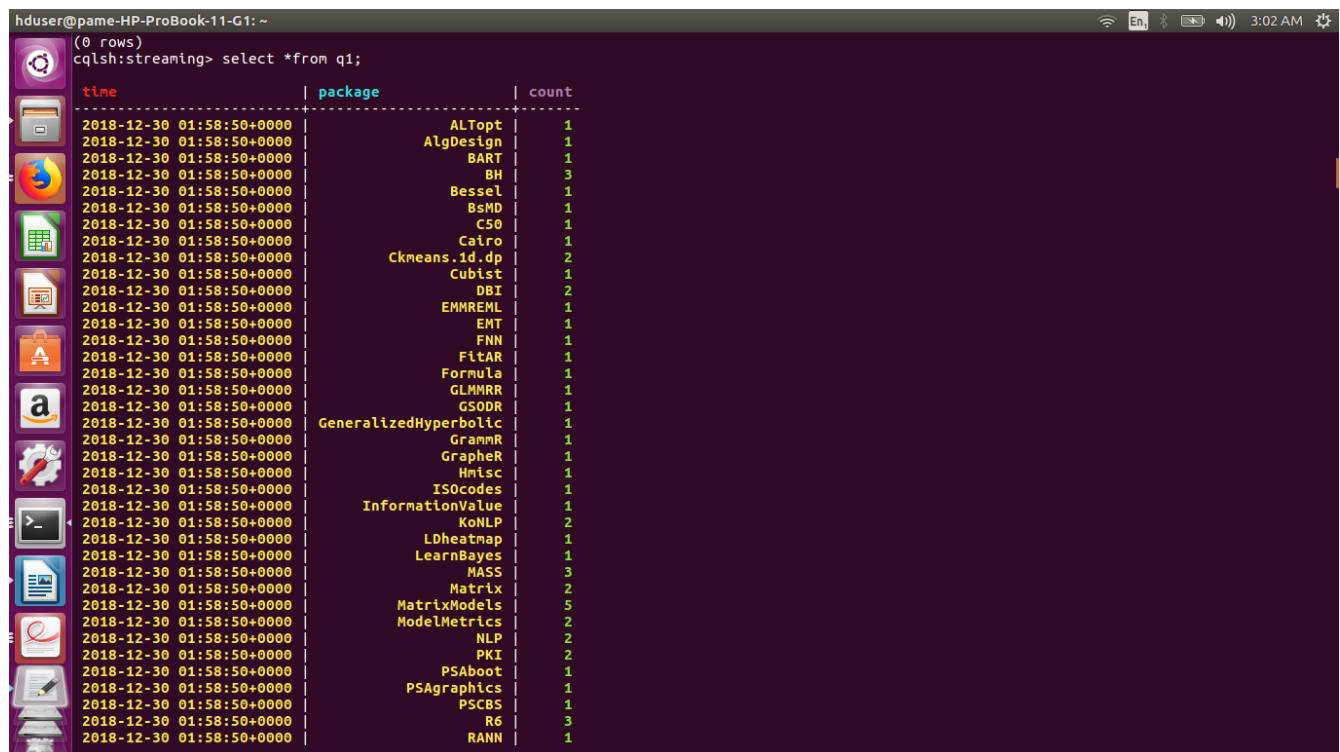
Streaming Computations:

Q1. To calculate the number of downloads of each package.

CQL Code Table:

```
:cqlsh:streaming> CREATE TABLE q1 (time text, package text, count int, PRIMARY KEY(time,  
package));
```

ScreenShots:



hduser@pame-HP-ProBook-11-G1: ~
(0 rows)
cqlsh:streaming> select * from q1;

time	package	count
2018-12-30 01:58:50+0000	ALTopt	1
2018-12-30 01:58:50+0000	AlgDesign	1
2018-12-30 01:58:50+0000	BART	1
2018-12-30 01:58:50+0000	BH	3
2018-12-30 01:58:50+0000	Bessel	1
2018-12-30 01:58:50+0000	BsMD	1
2018-12-30 01:58:50+0000	C50	1
2018-12-30 01:58:50+0000	Cairo	1
2018-12-30 01:58:50+0000	Ckmeans.1d.dp	2
2018-12-30 01:58:50+0000	Cubist	1
2018-12-30 01:58:50+0000	DBI	2
2018-12-30 01:58:50+0000	EMMREML	1
2018-12-30 01:58:50+0000	EMT	1
2018-12-30 01:58:50+0000	FNN	1
2018-12-30 01:58:50+0000	FitAR	1
2018-12-30 01:58:50+0000	Formula	1
2018-12-30 01:58:50+0000	GLMMRR	1
2018-12-30 01:58:50+0000	GSODR	1
2018-12-30 01:58:50+0000	GeneralizedHyperbolic	1
2018-12-30 01:58:50+0000	GramMR	1
2018-12-30 01:58:50+0000	GraphER	1
2018-12-30 01:58:50+0000	HmLsc	1
2018-12-30 01:58:50+0000	ISOCodes	1
2018-12-30 01:58:50+0000	InformationValue	1
2018-12-30 01:58:50+0000	KoNLP	2
2018-12-30 01:58:50+0000	LDheatmap	1
2018-12-30 01:58:50+0000	LearnBayes	1
2018-12-30 01:58:50+0000	MASS	3
2018-12-30 01:58:50+0000	Matrix	2
2018-12-30 01:58:50+0000	MatrixModels	5
2018-12-30 01:58:50+0000	ModelMetrics	2
2018-12-30 01:58:50+0000	NLP	2
2018-12-30 01:58:50+0000	PKI	2
2018-12-30 01:58:50+0000	PSAboot	1
2018-12-30 01:58:50+0000	PSAgraphics	1
2018-12-30 01:58:50+0000	PSCBS	1
2018-12-30 01:58:50+0000	R6	3
2018-12-30 01:58:50+0000	RANN	1

```
hduser@pame-HP-ProBook-11-G1: ~  
2018-12-30 03:23:55+0000 tidyhydat 1  
2018-12-30 03:23:55+0000 tidyimpute 1  
2018-12-30 03:23:55+0000 tidymodels 1  
2018-12-30 03:23:55+0000 tidyposterior 1  
2018-12-30 03:23:55+0000 tidypredict 1  
2018-12-30 03:23:55+0000 tidyquant 1  
2018-12-30 03:23:55+0000 tidyr 5  
2018-12-30 03:23:55+0000 tidyselect 5  
2018-12-30 03:23:55+0000 tidyverse 3  
2018-12-30 03:23:55+0000 timeDate 1  
2018-12-30 03:23:55+0000 timetk 1  
2018-12-30 03:23:55+0000 tinytex 4  
2018-12-30 03:23:55+0000 tnaptools 1  
2018-12-30 03:23:55+0000 tseries 3  
2018-12-30 03:23:55+0000 urca 1  
2018-12-30 03:23:55+0000 uroot 1  
2018-12-30 03:23:55+0000 usethis 6  
2018-12-30 03:23:55+0000 utf8 6  
2018-12-30 03:23:55+0000 uuid 1  
2018-12-30 03:23:55+0000 vegan 1  
2018-12-30 03:23:55+0000 vlaplot 1  
2018-12-30 03:23:55+0000 viridis 1  
2018-12-30 03:23:55+0000 viridisLite 2  
---MORE---  
time | package | count  
-----  
2018-12-30 03:23:55+0000 webshot 1  
2018-12-30 03:23:55+0000 whisker 4  
2018-12-30 03:23:55+0000 withr 2  
2018-12-30 03:23:55+0000 wordcloud 1  
2018-12-30 03:23:55+0000 xfun 1  
2018-12-30 03:23:55+0000 xlsxjars 1  
2018-12-30 03:23:55+0000 xml2 3  
2018-12-30 03:23:55+0000 xopen 3  
2018-12-30 03:23:55+0000 xtable 16  
2018-12-30 03:23:55+0000 xts 2  
2018-12-30 03:23:55+0000 yaml 9  
2018-12-30 03:23:55+0000 zip 2  
2018-12-30 03:23:55+0000 zoo 4  
(51213 rows)  
cqlsh:streaming>
```

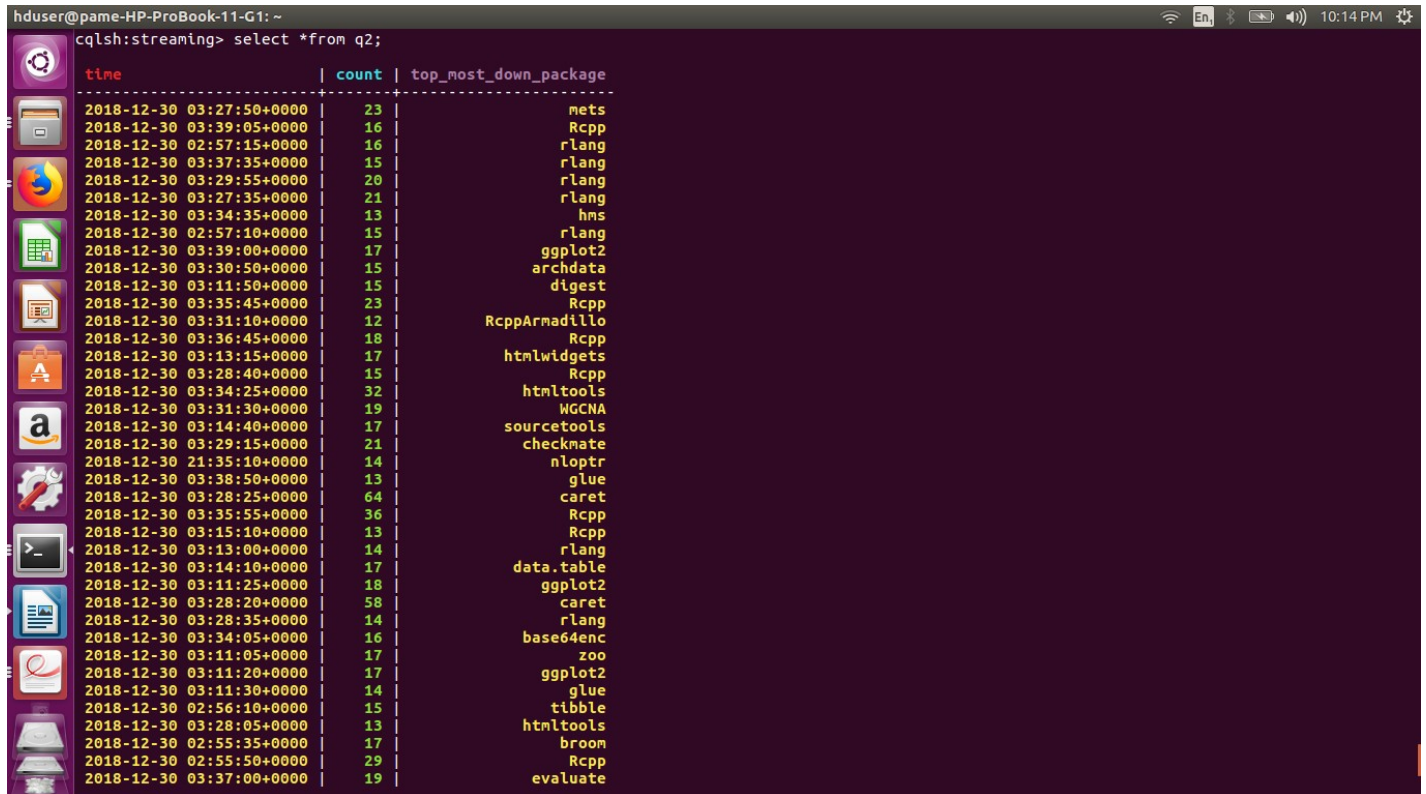
```
hduser@pame-HP-ProBook-11-G1: ~  
2018-12-30 03:23:55+0000 tidyhydat 1  
2018-12-30 03:23:55+0000 tidyimpute 1  
2018-12-30 03:23:55+0000 tidymodels 1  
2018-12-30 03:23:55+0000 tidyposterior 1  
2018-12-30 03:23:55+0000 tidypredict 1  
2018-12-30 03:23:55+0000 tidyquant 1  
2018-12-30 03:23:55+0000 tidyr 5  
2018-12-30 03:23:55+0000 tidyselect 5  
2018-12-30 03:23:55+0000 tidyverse 3  
2018-12-30 03:23:55+0000 timeDate 1  
2018-12-30 03:23:55+0000 timetk 1  
2018-12-30 03:23:55+0000 tinytex 4  
2018-12-30 03:23:55+0000 tnaptools 1  
2018-12-30 03:23:55+0000 tseries 3  
2018-12-30 03:23:55+0000 urca 1  
2018-12-30 03:23:55+0000 uroot 1  
2018-12-30 03:23:55+0000 usethis 6  
2018-12-30 03:23:55+0000 utf8 6  
2018-12-30 03:23:55+0000 uuid 1  
2018-12-30 03:23:55+0000 vegan 1  
2018-12-30 03:23:55+0000 vlaplot 1  
2018-12-30 03:23:55+0000 viridis 1  
2018-12-30 03:23:55+0000 viridisLite 2  
---MORE---  
time | package | count  
-----  
2018-12-30 03:23:55+0000 webshot 1  
2018-12-30 03:23:55+0000 whisker 4  
2018-12-30 03:23:55+0000 withr 2  
2018-12-30 03:23:55+0000 wordcloud 1  
2018-12-30 03:23:55+0000 xfun 1  
2018-12-30 03:23:55+0000 xlsxjars 1  
2018-12-30 03:23:55+0000 xml2 3  
2018-12-30 03:23:55+0000 xopen 3  
2018-12-30 03:23:55+0000 xtable 16  
2018-12-30 03:23:55+0000 xts 2  
2018-12-30 03:23:55+0000 yaml 9  
2018-12-30 03:23:55+0000 zip 2  
2018-12-30 03:23:55+0000 zoo 4  
(51213 rows)  
cqlsh:streaming>
```

Q2.To find the top most downloaded package.

CQL Code Table:

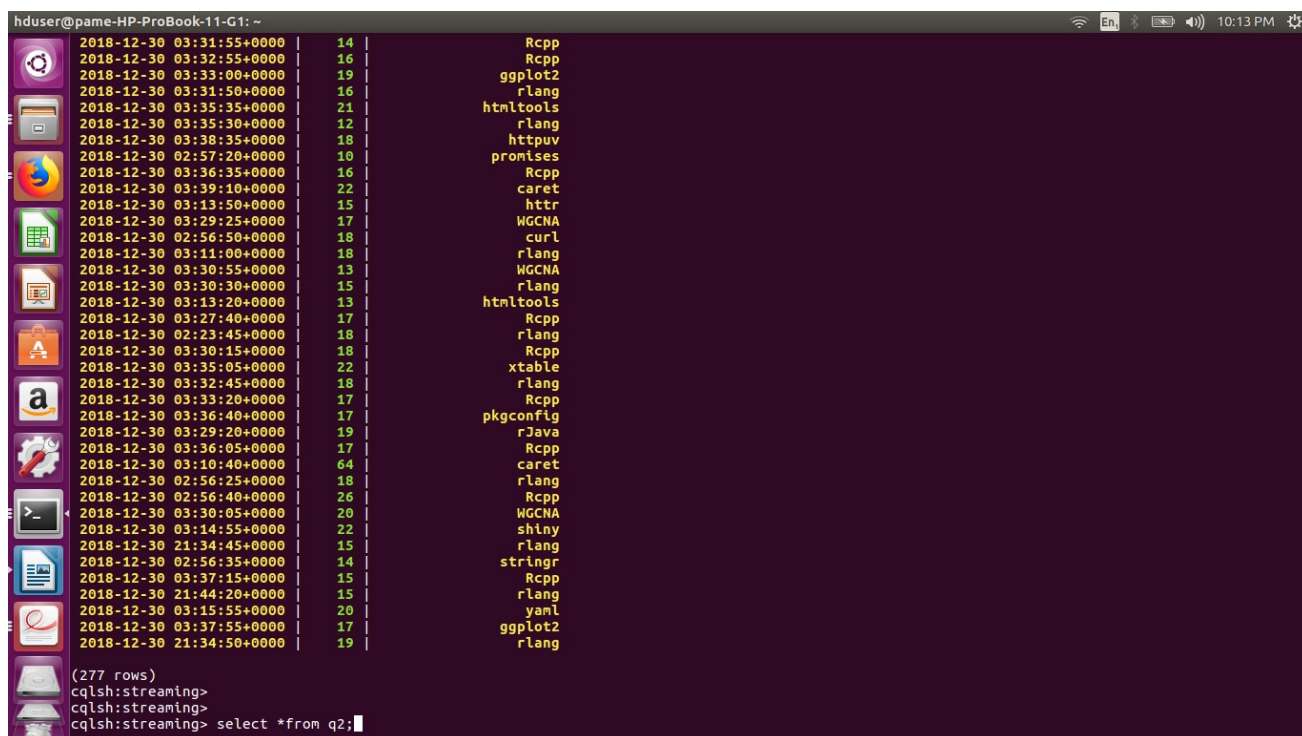
```
:cqlsh:streaming> CREATE TABLE q2 (time text, top_most_down_package text, count int, PRIMARY KEY(time, count));
```

ScreenShots:



```
hduser@pame-HP-ProBook-11-G1: ~
cqlsh:streaming> select *from q2;
```

time	count	top_most_down_package
2018-12-30 03:27:50+0000	23	mets
2018-12-30 03:39:05+0000	16	Rcpp
2018-12-30 02:57:15+0000	16	rlang
2018-12-30 03:37:35+0000	15	rlang
2018-12-30 03:29:55+0000	20	rlang
2018-12-30 03:27:35+0000	21	rlang
2018-12-30 03:34:35+0000	13	hms
2018-12-30 02:57:10+0000	15	rlang
2018-12-30 03:39:00+0000	17	ggplot2
2018-12-30 03:30:50+0000	15	archdata
2018-12-30 03:11:50+0000	15	digest
2018-12-30 03:35:45+0000	23	Rcpp
2018-12-30 03:31:10+0000	12	RcppArmadillo
2018-12-30 03:36:45+0000	18	Rcpp
2018-12-30 03:13:15+0000	17	htmlwidgets
2018-12-30 03:28:40+0000	15	Rcpp
2018-12-30 03:34:25+0000	32	htmltools
2018-12-30 03:31:30+0000	19	WGCNA
2018-12-30 03:14:40+0000	17	sourcetools
2018-12-30 03:29:15+0000	21	checkmate
2018-12-30 21:35:10+0000	14	nloptr
2018-12-30 03:38:50+0000	13	glue
2018-12-30 03:28:25+0000	64	caret
2018-12-30 03:35:55+0000	36	Rcpp
2018-12-30 03:15:10+0000	13	Rcpp
2018-12-30 03:13:00+0000	14	rlang
2018-12-30 03:14:10+0000	17	data.table
2018-12-30 03:11:25+0000	18	ggplot2
2018-12-30 03:28:20+0000	58	caret
2018-12-30 03:28:35+0000	14	rlang
2018-12-30 03:34:05+0000	16	base64enc
2018-12-30 03:11:05+0000	17	zoo
2018-12-30 03:11:20+0000	17	ggplot2
2018-12-30 03:11:30+0000	14	glue
2018-12-30 02:56:10+0000	15	tibble
2018-12-30 03:28:05+0000	13	htmltools
2018-12-30 02:55:35+0000	17	broom
2018-12-30 02:55:50+0000	29	Rcpp
2018-12-30 03:37:00+0000	19	evaluate



2018-12-30 03:31:55+0000	14	Rcpp
2018-12-30 03:32:55+0000	16	Rcpp
2018-12-30 03:33:00+0000	19	ggplot2
2018-12-30 03:31:50+0000	16	rlang
2018-12-30 03:35:35+0000	21	htmltools
2018-12-30 03:35:30+0000	12	rlang
2018-12-30 03:38:35+0000	18	httpuv
2018-12-30 02:57:20+0000	10	promises
2018-12-30 03:36:35+0000	16	Rcpp
2018-12-30 03:39:10+0000	22	caret
2018-12-30 03:13:50+0000	15	httr
2018-12-30 03:29:25+0000	17	WGCNA
2018-12-30 02:56:50+0000	18	curl
2018-12-30 03:11:00+0000	18	rlang
2018-12-30 03:30:55+0000	13	WGCNA
2018-12-30 03:30:30+0000	15	rlang
2018-12-30 03:13:20+0000	13	htmltools
2018-12-30 03:27:40+0000	17	Rcpp
2018-12-30 02:23:45+0000	18	rlang
2018-12-30 03:30:15+0000	18	Rcpp
2018-12-30 03:35:05+0000	22	xtable
2018-12-30 03:32:45+0000	18	rlang
2018-12-30 03:33:20+0000	17	Rcpp
2018-12-30 03:36:40+0000	17	pkgconfig
2018-12-30 03:29:20+0000	19	rJava
2018-12-30 03:36:05+0000	17	Rcpp
2018-12-30 03:10:40+0000	64	caret
2018-12-30 02:56:25+0000	18	rlang
2018-12-30 02:56:40+0000	26	Rcpp
2018-12-30 03:30:05+0000	20	WGCNA
2018-12-30 03:14:55+0000	22	shiny
2018-12-30 21:34:45+0000	15	rlang
2018-12-30 02:56:35+0000	14	stringr
2018-12-30 03:37:15+0000	15	Rcpp
2018-12-30 21:44:20+0000	15	rlang
2018-12-30 03:15:55+0000	20	yaml
2018-12-30 03:37:55+0000	17	ggplot2
2018-12-30 21:34:50+0000	19	rlang

(277 rows)
 cqlsh:streaming>
 cqlsh:streaming>
 cqlsh:streaming> select *from q2;

Q3. To find the top 5 countries along with number of downloads.

CQL Code Table:

```
:cqlsh:streaming> CREATE TABLE q3 (time text, top_5_countries text, count int, PRIMARY KEY(time, count));
```

ScreenShots:


```
hduser@pame-HP-ProBook-11-G1: ~  
cqlsh:streaming> select *from q3;
```

time	count	top_5_countries
2018-12-30 02:58:05+0000	23	FR
2018-12-30 02:58:05+0000	36	CA
2018-12-30 02:58:05+0000	61	NA
2018-12-30 02:58:05+0000	69	BR
2018-12-30 02:58:05+0000	645	US
2018-12-30 02:58:10+0000	30	GB
2018-12-30 02:58:10+0000	32	NA
2018-12-30 02:58:10+0000	47	CA
2018-12-30 02:58:10+0000	50	BR
2018-12-30 02:58:10+0000	642	US
2018-12-30 21:36:10+0000	90	AU
2018-12-30 21:36:10+0000	118	JP
2018-12-30 21:36:10+0000	127	NA
2018-12-30 21:36:10+0000	200	KR
2018-12-30 21:36:10+0000	224	US
2018-12-30 03:44:05+0000	51	BR
2018-12-30 03:44:05+0000	89	CA
2018-12-30 03:44:05+0000	91	GB
2018-12-30 03:44:05+0000	103	NA
2018-12-30 03:44:05+0000	430	US
2018-12-30 03:43:20+0000	53	GB
2018-12-30 03:43:20+0000	59	CN
2018-12-30 03:43:20+0000	66	DE
2018-12-30 03:43:20+0000	102	NA
2018-12-30 03:43:20+0000	347	US
2018-12-30 03:43:00+0000	42	NA
2018-12-30 03:43:00+0000	45	CO
2018-12-30 03:43:00+0000	122	DE
2018-12-30 03:43:00+0000	565	US
2018-12-30 03:44:45+0000	50	BR
2018-12-30 03:44:45+0000	52	FR
2018-12-30 03:44:45+0000	53	IN
2018-12-30 03:44:45+0000	118	NA
2018-12-30 03:44:45+0000	511	US
2018-12-30 21:35:35+0000	22	KR
2018-12-30 21:35:35+0000	25	CN
2018-12-30 21:35:35+0000	80	NA
2018-12-30 21:35:35+0000	107	JP
2018-12-30 21:35:35+0000	672	US

```
hduser@pame-HP-ProBook-11-G1: ~  
cqlsh:streaming> select *from q3;
```

2018-12-30 21:28:05+0000	58	DE
2018-12-30 21:28:05+0000	70	IN
2018-12-30 21:28:05+0000	123	NA
2018-12-30 21:28:05+0000	126	KR
2018-12-30 21:28:05+0000	364	US
2018-12-30 02:57:35+0000	68	DE
2018-12-30 02:57:35+0000	76	BR
2018-12-30 02:57:35+0000	108	FR
2018-12-30 02:57:35+0000	114	NA
2018-12-30 02:57:35+0000	269	US
2018-12-30 03:41:25+0000	55	NA
2018-12-30 03:41:25+0000	78	AU
2018-12-30 03:41:25+0000	146	JP
2018-12-30 03:41:25+0000	211	KR
2018-12-30 03:41:25+0000	244	US
2018-12-30 02:57:55+0000	41	GB
2018-12-30 02:57:55+0000	55	FI
2018-12-30 02:57:55+0000	105	NA
2018-12-30 02:57:55+0000	469	US
2018-12-30 21:35:25+0000	49	CA
2018-12-30 21:35:25+0000	55	NA
2018-12-30 21:35:25+0000	58	CO
2018-12-30 21:35:25+0000	127	KR
2018-12-30 21:35:25+0000	535	US
2018-12-30 03:41:10+0000	33	SG
2018-12-30 03:41:10+0000	70	NA
2018-12-30 03:41:10+0000	119	JP
2018-12-30 03:41:10+0000	133	KR
2018-12-30 03:41:10+0000	496	US
2018-12-30 21:45:30+0000	50	KR
2018-12-30 21:45:30+0000	72	AT
2018-12-30 21:45:30+0000	83	CN
2018-12-30 21:45:30+0000	96	NA
2018-12-30 21:45:30+0000	385	US
2018-12-30 02:59:00+0000	27	ES
2018-12-30 02:59:00+0000	49	CA
2018-12-30 02:59:00+0000	59	NA
2018-12-30 02:59:00+0000	62	GB
2018-12-30 02:59:00+0000	728	US

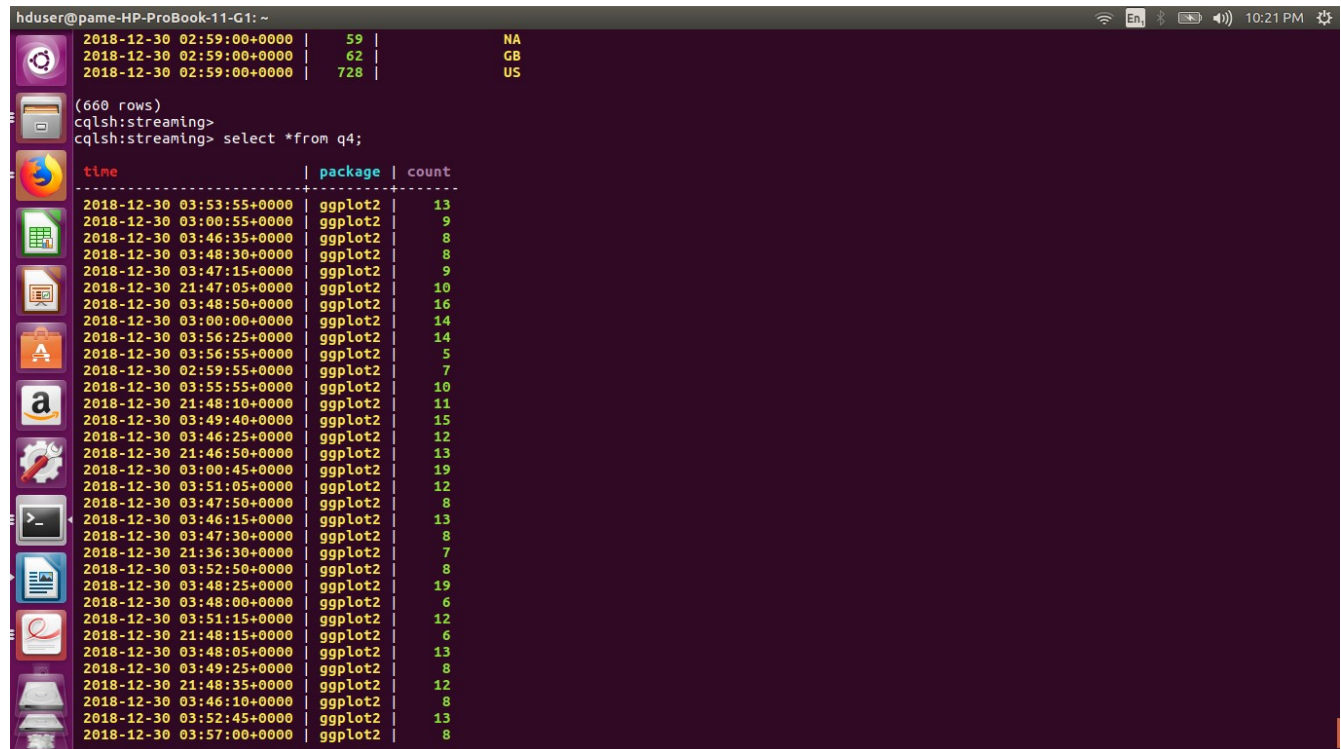
(660 rows)
cqlsh:streaming>
cqlsh:streaming>

Q4. To find number of downloads for ggplot2 package.

CQL Code Table:

```
:cqlsh:streaming> CREATE TABLE q4 (time text, package text, count int, PRIMARY KEY(time, package));
```

ScreenShots:



The screenshot shows a terminal window with a dark background. The top bar indicates the user is 'hduser' on a machine named 'pame-HP-ProBook-11-G1'. The terminal displays the execution of CQL code to create a table 'q4' and then a query to select all data from it. The result is a table with three columns: 'time', 'package', and 'count'. The data shows download counts for the 'ggplot2' package at various times on 2018-12-30. The counts range from 5 to 19. The table is displayed in a tabular format with vertical lines separating the columns.

time	package	count
2018-12-30 02:59:00+0000	ggplot2	59
2018-12-30 02:59:00+0000	ggplot2	62
2018-12-30 02:59:00+0000	ggplot2	728
2018-12-30 03:53:55+0000	ggplot2	13
2018-12-30 03:00:55+0000	ggplot2	9
2018-12-30 03:46:35+0000	ggplot2	8
2018-12-30 03:48:30+0000	ggplot2	8
2018-12-30 03:47:15+0000	ggplot2	9
2018-12-30 21:47:05+0000	ggplot2	10
2018-12-30 03:48:50+0000	ggplot2	16
2018-12-30 03:00:00+0000	ggplot2	14
2018-12-30 03:56:25+0000	ggplot2	14
2018-12-30 03:56:55+0000	ggplot2	5
2018-12-30 02:59:55+0000	ggplot2	7
2018-12-30 03:55:55+0000	ggplot2	10
2018-12-30 21:48:10+0000	ggplot2	11
2018-12-30 03:49:40+0000	ggplot2	15
2018-12-30 03:46:25+0000	ggplot2	12
2018-12-30 21:46:50+0000	ggplot2	13
2018-12-30 03:00:45+0000	ggplot2	19
2018-12-30 03:51:05+0000	ggplot2	12
2018-12-30 03:47:50+0000	ggplot2	8
2018-12-30 03:46:15+0000	ggplot2	13
2018-12-30 03:47:30+0000	ggplot2	8
2018-12-30 21:36:30+0000	ggplot2	7
2018-12-30 03:52:50+0000	ggplot2	8
2018-12-30 03:48:25+0000	ggplot2	19
2018-12-30 03:48:00+0000	ggplot2	6
2018-12-30 03:51:15+0000	ggplot2	12
2018-12-30 21:48:15+0000	ggplot2	6
2018-12-30 03:48:05+0000	ggplot2	13
2018-12-30 03:49:25+0000	ggplot2	8
2018-12-30 21:48:35+0000	ggplot2	12
2018-12-30 03:46:10+0000	ggplot2	8
2018-12-30 03:52:45+0000	ggplot2	13
2018-12-30 03:57:00+0000	ggplot2	8

```
hduser@pame-HP-ProBook-11-G1: ~
2018-12-30 03:47:05+0000 | ggplot2 | 10
2018-12-30 03:50:25+0000 | ggplot2 | 12
2018-12-30 03:55:30+0000 | ggplot2 | 5
2018-12-30 02:49:45+0000 | ggplot2 | 8
2018-12-30 21:47:45+0000 | ggplot2 | 8
2018-12-30 03:48:10+0000 | ggplot2 | 18
2018-12-30 03:50:55+0000 | ggplot2 | 13
2018-12-30 03:54:05+0000 | ggplot2 | 10
2018-12-30 03:49:15+0000 | ggplot2 | 8
2018-12-30 03:50:30+0000 | ggplot2 | 11
2018-12-30 03:49:00+0000 | ggplot2 | 14
2018-12-30 03:46:20+0000 | ggplot2 | 23
2018-12-30 21:36:25+0000 | ggplot2 | 17
2018-12-30 03:54:45+0000 | ggplot2 | 12
2018-12-30 03:49:30+0000 | ggplot2 | 7
2018-12-30 21:47:35+0000 | ggplot2 | 7
2018-12-30 21:37:10+0000 | ggplot2 | 14
2018-12-30 03:46:30+0000 | ggplot2 | 9
2018-12-30 21:36:35+0000 | ggplot2 | 10
2018-12-30 03:51:55+0000 | ggplot2 | 4

---MORE---
time | package | count
-----+-----+-----
2018-12-30 03:48:15+0000 | ggplot2 | 14
2018-12-30 03:57:10+0000 | ggplot2 | 11
2018-12-30 03:45:20+0000 | ggplot2 | 7
2018-12-30 03:46:05+0000 | ggplot2 | 2
2018-12-30 21:36:55+0000 | ggplot2 | 14
2018-12-30 02:49:10+0000 | ggplot2 | 7
2018-12-30 02:49:50+0000 | ggplot2 | 10
2018-12-30 03:56:45+0000 | ggplot2 | 6
2018-12-30 03:45:40+0000 | ggplot2 | 3
2018-12-30 03:45:35+0000 | ggplot2 | 17
2018-12-30 03:52:20+0000 | ggplot2 | 14
2018-12-30 21:46:55+0000 | ggplot2 | 15
2018-12-30 03:48:45+0000 | ggplot2 | 10
2018-12-30 03:01:25+0000 | ggplot2 | 3

(214 rows)
cqlsh:streaming>
cqlsh:streaming>
cqlsh:streaming>
```