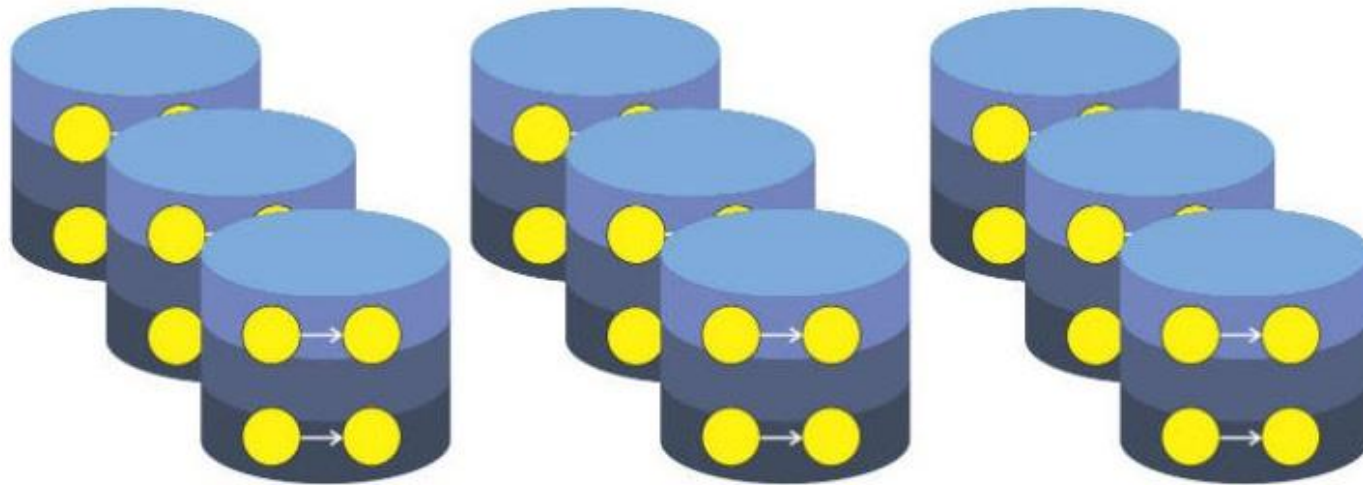
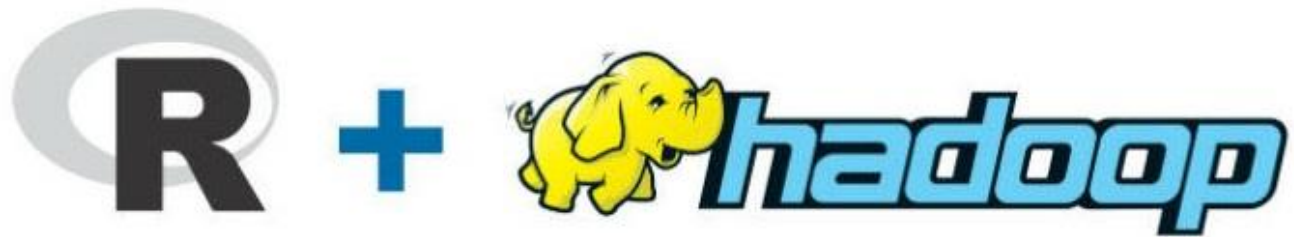


R Studio dan Data Lake

Overview



Koneksi R Studio ke Hive di Data Lake (1)

Langkah 1: Menambahkan Package Pada R Console:

```
install.packages("rJava")  
install.packages("RJDBC", dep=TRUE)
```

Setelah terinstal akan muncul User Library seperti gambar berikut:

User Library			
<input type="checkbox"/>	DBI	R Database Interface	0.7
<input type="checkbox"/>	rJava	Low-Level R to Java Interface	0.9-9
<input type="checkbox"/>	RJDBC	Provides access to databases through the JDBC interface	0.2-5

Load library tersebut pada session R

```
library("DBI")  
library("rJava")  
library("RJDBC")
```

Koneksi R Studio ke Hive di Data Lake (2)

Langkah 2: Menyalin Hadoop Jars yang Diperlukan dan Mendefinisikan Lokasinya agar dapat Digunakan oleh R

mkdir R-java (membuat direktori untuk menampung hadoop jars pada server R Studio)

scp solusi247@hdpdl01:/home/solusi247/R-java/* R-java (menyalin hadoop jars ke direktori tersebut)

Skrip untuk mendefinisikan lokasi hadoop jars agar dapat digunakan oleh R

```
cp = c("/home/solusi247/R-java/hive-jdbc-1.2.1.jar"  
, "/home/solusi247/R-java/hadoop-common-2.7.1_yava_2.3.0.0.jar"  
, "/home/solusi247/R-java/libthrift-0.9.3.jar"  
, "/home/solusi247/R-java/hive-service.jar"  
, "/home/solusi247/R-java/httpclient-4.2.5.jar"  
, "/home/solusi247/R-java/httpcore-4.2.5.jar"  
, "/home/solusi247/R-java/hive-jdbc-1.2.1-standalone.jar")  
  
.jinit(classpath=cp)
```

Koneksi R Studio ke Hive di Data Lake (3)

Langkah 3: Menjalankan Skrip Koneksi ke Hive Menggunakan IP Master Node

```
drv = JDBC("org.apache.hive.jdbc.HiveDriver"  
  , "/home/solusi247/R-java/hive-jdbc-1.2.1.jar")  
conn = dbConnect(drv  
  , "jdbc:hive2://10.0.26.151:10000/exim_nscs"  
  , "hive", "")
```

Hasil Eksplorasi R Studio dan Data Lake (1)

1. Menampilkan Daftar Database yang Ada

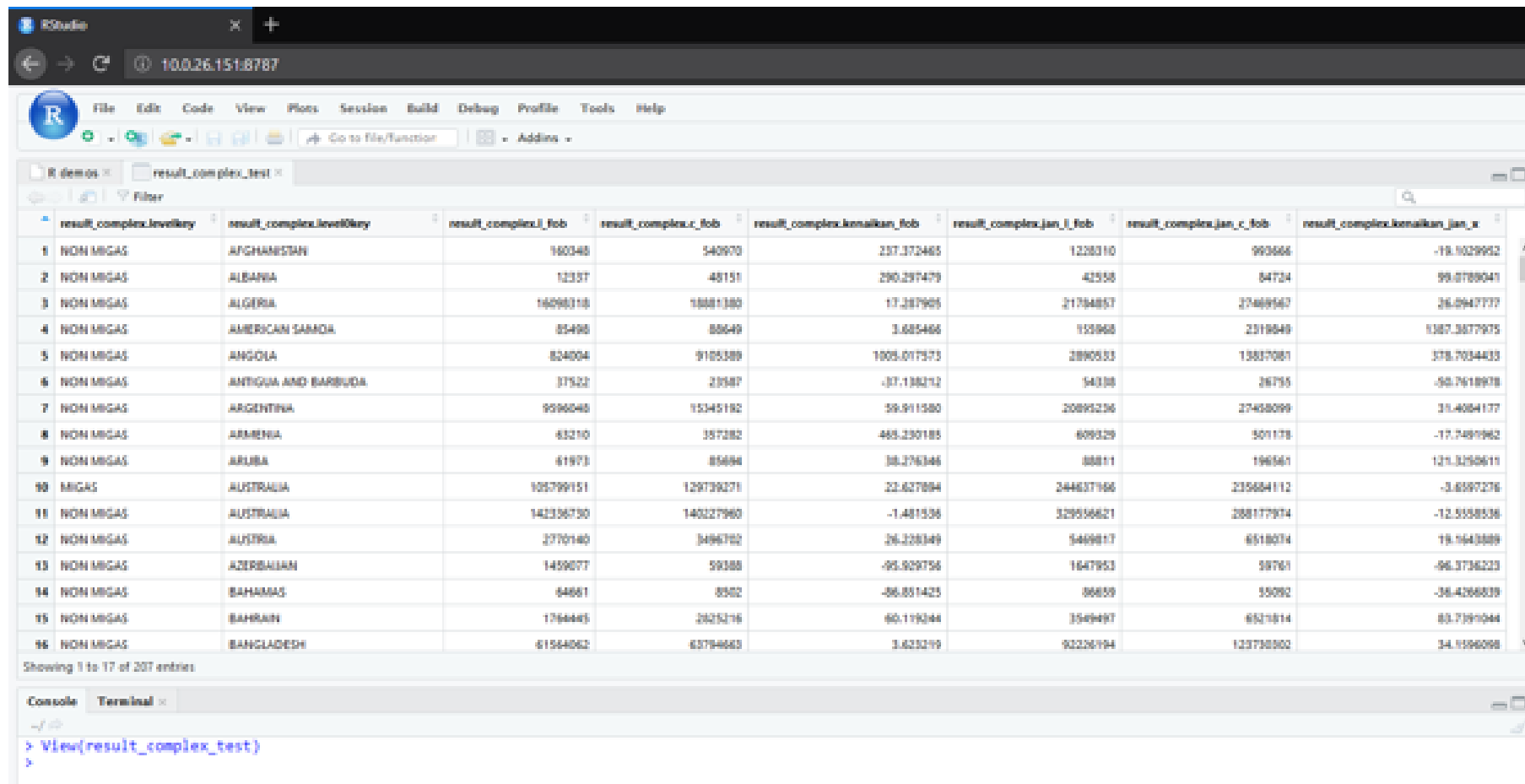
```
> .jclassLoader()$setDebug(0L)
> dbGetQuery(conn, "show databases")
database_name
1      crawling
2      dbtest
3      default
4      exim_nsc
5      sp2010
6      training
> |
```

2. Menampilkan Daftar Tabel yang Ada

```
> dbGetQuery(conn, "show tables")
  tab_name
1 commodity
2 country
3 country_orc
4 data_source
5 data_source_orc
6 fact5juta_nol
7 fact_complex
8 fact_complex_orc
9 fact_complex_orc2
10 fact_exim
11 fact_exim_orc
12 fact_heavy
13 fact_heavy_orc
14 fact_new
15 fact_new_orc
16 fact_pqt
17 hs
18 period
19 result_complex
20 result_heavy
21 rpt_fact01
22 rpt_fact02
23 rpt_fact03
24 sitc
25 view10juta_exim
> |
```

Hasil Eksplorasi R Studio dan Data Lake (2)

3. Melihat Isi Sebuah Tabel



The screenshot shows the R Studio interface with a data table titled 'result_complex_test'. The table has 17 columns and 17 rows of data. The columns are: result_complex.levelkey, result_complex.levelkey, result_complex.fob, result_complex.c_fob, result_complex.kenaikan_fob, result_complex.jan_fob, result_complex.jan_c_fob, and result_complex.kenaikan_jan_x. The rows represent different countries, starting with 'AFGHANISTAN' and ending with 'BANGLADESH'. The data is displayed in a grid format with alternating light and dark gray rows. The R Studio interface includes a menu bar at the top with options like File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with various icons. The console at the bottom shows the command 'View(result_complex_test)' being executed.

	result_complex.levelkey	result_complex.levelkey	result_complex.fob	result_complex.c_fob	result_complex.kenaikan_fob	result_complex.jan_fob	result_complex.jan_c_fob	result_complex.kenaikan_jan_x
1	NON MIGAS	AFGHANISTAN	160348	540970	237.372485	1228310	993666	-19.1029932
2	NON MIGAS	ALBANIA	12337	48151	290.297479	42558	84734	99.0789041
3	NON MIGAS	ALGERIA	16098318	10881380	17.287905	21784857	27489567	26.2947737
4	NON MIGAS	AMERICAN SAMOA	85498	88649	3.685466	153968	2318649	1387.3877975
5	NON MIGAS	ANGOLA	824004	9105389	1005.017373	2890533	13837081	378.7034433
6	NON MIGAS	ANTIGUA AND BARBUDA	37523	23587	-37.188212	54338	26755	-50.7618978
7	NON MIGAS	ARGENTINA	9396048	15345182	58.911580	20895236	27468099	31.4084137
8	NON MIGAS	ARMENIA	63210	357282	443.230185	609329	501178	-17.7481962
9	NON MIGAS	ARUBA	61873	85684	38.276346	88811	196561	121.3258611
10	MIGAS	AUSTRALIA	105799151	129739271	22.627884	244637166	235684112	-3.6597276
11	NON MIGAS	AUSTRALIA	142356730	140227960	-1.481536	329556621	288177974	-12.5558536
12	NON MIGAS	AUSTRIA	2770140	3496762	26.228349	5469817	6518074	19.5643889
13	NON MIGAS	AZERBAIJAN	1459077	59388	-95.929756	1647953	58761	-96.3736223
14	NON MIGAS	BAHAMAS	84881	8502	-86.851425	86859	55092	-38.4286839
15	NON MIGAS	BAHRAIN	1764445	2625216	60.119344	3549497	6521814	83.7391044
16	NON MIGAS	BANGLADESH	61564062	63784683	3.623219	92236194	123738302	34.1596098

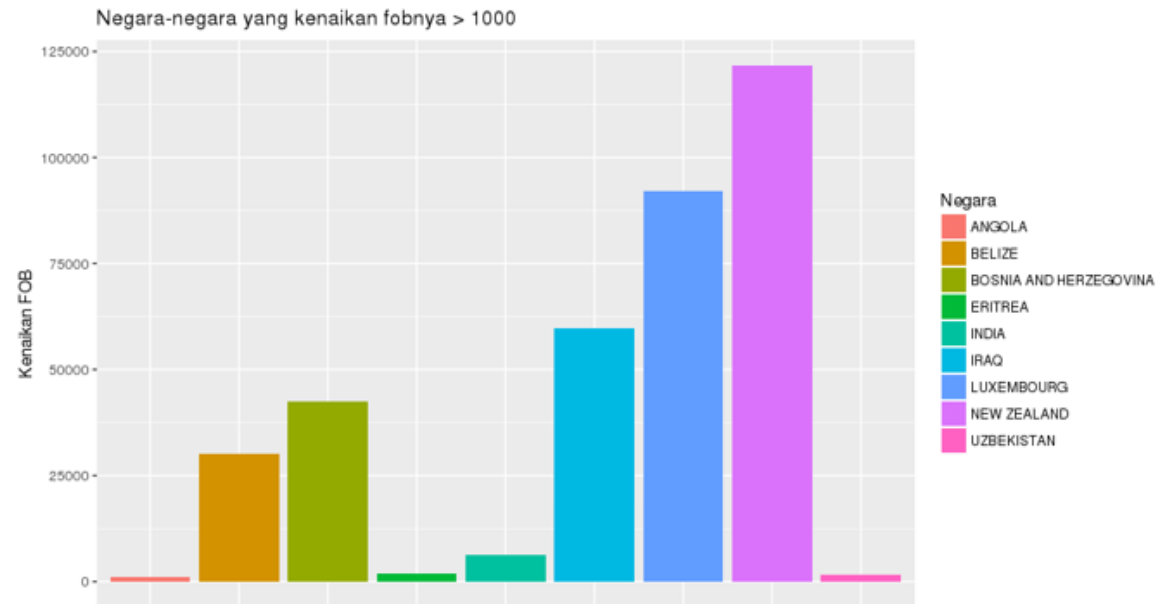
Hasil Eksplorasi R Studio dan Data Lake (3)

4. Membuat Bar Chart

Sebelumnya diperlukan Instalasi package "ggplot2" untuk membuat berbagai macam chart.

```
test <- dbGetQuery(conn, "select * from result_complex where result_complex.kenaikan_fob > 1000")  
(Skrip query untuk memilih negara dengan kenaikan fob lebih dari 1000)
```

```
ggplot(test, aes(x=result_complex.level0key, y=result_complex.kenaikan_fob,  
fill=result_complex.level0key)) + geom_bar(stat = "identity") + labs(x="Negara", y="Kenaikan FOB",  
title="Negara-negara yang kenaikan fobnya > 1000", fill="Negara")+theme(axis.title.x=element_blank(),  
axis.text.x=element_blank(), axis.ticks.x=element_blank())  
(Skrip untuk membuat bar chart seperti gambar berikut)
```



Eksplorasi Lanjutan

- ▶ Beberapa hal lain yg dapat dilakukan menggunakan R Studio namun masih memerlukan eksplorasi lebih lanjut antara lain:
- ▶ Analisis statistik, machine learning, web app yang terhubung ke R session (shiny), transformasi data, dan lainnya.
- ▶ Informasi lebih lanjut dapat dilihat pada link berikut:
- ▶ <https://www.rstudio.com/resources/cheatsheets/>

Shiny Web App

- ▶ <http://shiny.rstudio.com/>
- ▶ Shiny is an R package that makes it easy to build interactive web apps straight from R.
- ▶ You can host standalone apps on a webpage or embed them in R Markdown documents or build dashboards.
- ▶ You can also extend your Shiny apps with CSS themes, htmlwidgets, and JavaScript actions.

