ACADGILD - SESSION7(EXPLORING APACHE PIG)

To-do

Problem Statement:

What are the movie titles that the user has rated?

How many times a movie has been rated by the user?

In question 2 above, what is the average rating given for a movie?

Down loaded all the related files of the case study and and places them on Hadoop as well as local.

Executed Task1_movienotrated on local

```
acadgild@localhost:~/PIG
File Edit View Search Terminal Help
[acadgild@localhost PIG]$ pig -brief -x local Task1_movienotrated.pig
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/
/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/li
LoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple bindings for an explanati
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/11/26 04:42:51 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/11/26 04:42:51 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49
Logging error messages to: /home/acadgild/PIG/pig_1543187571589.log
user.name is deprecated. Instead, use mapreduce.job.user.name
Default bootup file /home/acadgild/.pigbootup not found
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
File Edit View Search Terminal Help
(Bridget Everett: Gynecological Wonder (2015))
(Jay Pharoah: Can I Be Me? (2015))
(Natasha Leggero: Live at Bimbo's (2015))
(Jeff Dunham: Unhinged in Hollywood (2015))
(Steve Rannazzisi: Breaking Dad (2015))
(Brian Regan: Live From Radio City Music Hall (2015))
(Jimmy Dore: Sentenced To Live (2015))
(Paul F. Tompkins: Crying and Driving (2015))
(Sarah Colonna: I Can't Feel My Legs (2015))
(Miles Jupp: Is The Chap You're Thinking Of (2015))
(John Bishop Supersonic Live (2015))
(Kevin Bridges Live: A Whole Different Story (2015))
(Marc Maron: More Later (2015))
(Elliott Morgan: Premature (2015))
(Mike Epps: Don't Take It Personal (2015))
(Jerrod Carmichael: Love at the Store (2014))
(Keith Lowell Jensen: Atheist Christmas (2014))
(Kurt Metzger: White Precious (2014))
(Tracy Morgan: Bona Fide (2014))
(Joe Rogan: Rocky Mountain High (2014))
(Robert Newman's History of Oil (2006))
(Robert Schimmel: Unprotected (1999))
(Bill Maher: Be More Cynical (2000))
(Wyatt Cenac: Comedy Person (2011))
("Nick Swardson: Seriously)
(Eugene Mirman: An Evening of Comedy in a Fake Underground Laboratory (2012))
(Moshe Kasher: Live in Oakland (2012))
(Chris Hardwick: Mandroid (2012))
(Poodle Springs (1998))
(Aurélie Laflamme: Les pieds sur terre (2015))
("Dr. Syn)
(Nukie (1988))
(Decanted (2016))
(Columbus (2017))
(Queerama (2017))
(title)
Pig script completed in 2 minutes, 51 seconds and 205 milliseconds (171205 ms)
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost PIG]$ |
acadoild
                                                             | acadoild@localbost:~/
                                PIG
```

Executed the same on MapReduce and got the same output.

```
[acadgild@localhost PIG]$ pig -brief Taskl movienotrated map.pig
18/11/26 04:52:31 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/11/26 04:52:31 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
18/11/26 04:52:31 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49
Logging error messages to: /home/acadgild/PIG/pig 1543188151750.log
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/comm
/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1
LoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Unable to load native-hadoop library for your platform... using builtin-java classes where app
Default bootup file /home/acadgild/.pigbootup not found
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
fs.default.name is deprecated. Instead, use fs.defaultFS
Connecting to hadoop file system at: hdfs://localhost:8020
```

```
File Edit View Search Terminal Help

(Bridget Everett: Gynecological Wonder (2015))
(Jay Pharoah: Can I Be Me? (2015))
(Natasha Leggero: Live at Bimbo's (2015))
(Seteve Rannazzisi: Breaking Dad (2015))
(Steve Rannazzisi: Breaking Dad (2015))
(Steve Rannazzisi: Breaking Dad (2015))
(Srian Regan: Live From Radio City Music Hall (2015))
(Jimmy Dore: Sentenced To Live (2015))
(Paul F. Tompkins: Crying and Driving (2015))
(Sarah Colonna: I Can't Feel My Legs (2015))
(Miles Jupp: Is The Chap You're Thinking Of (2015))
(John Bishop Supersonic Live (2015))
(Kevin Bridges Live: A Whole Different Story (2015))
(Marc Maron: More Later (2015))
(Elliott Morgan: Premature (2015))
(Elliott Morgan: Premature (2015))
(Mike Epps: Don't Take It Personal (2015))
(Jerrod Carmichael: Love at the Store (2014))
(Keith Lowell Jensen: Atheist Christmas (2014))
(Kurt Metzger: White Precious (2014))
(Tracy Morgan: Bona Fide (2014))
(Joe Rogan: Rocky Mountain High (2014))
(Robert Newman's History of Oil (2006))
(Robert Newman's History of Oil (2006))
(Robert Newman's History of Oil (2006))
(Myatt Cenae: Comedy Person (2011))
("Wick Swardson: Seriously)
(Eugene Mirman: An Evening of Comedy in a Fake Underground Laboratory (2012))
(Moshe Kasher: Live in Oakland (2012))
(Poodle Springs (1998))
(Aurélie Laflamme: Les pieds sur terre (2015))
("Dr. Syn)
(Nukie (1988))
(Decanted (2016))
(Columbus (2017))
(Culumbus (2017))
(Culumbus (2017))
(Culumbus (2017))
(Title)
Pig script completed in 5 minutes, 2 seconds and 605 milliseconds (302605 ms)
You have new mail in /var/spol/mail/acadgild
lacadgild@localhost PIGIS
```

Task1_movierated ran on local

```
rile buil view search lettilliai neip
You have new mail in /var/spool/mail/acadgild [acadgild@localhost PIG]$ pig -brief -x local Task1_movierated.pig SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/
/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lil
LoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation of type [org.slf4j.impl.Log4jLoggerFactory]
18/11/26 05:04:43 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/11/26 05:04:43 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49 Logging error messages to: /home/acadgild/PIG/pig_1543188883425.log
user.name is deprecated. Instead, use mapreduce.job.user.name
Default bootup file /home/acadgild/.pigbootup not found
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
fs.default.name is deprecated. Instead, use fs.defaultFS
Connecting to hadoop file system at: file:///
Pig Script ID for the session: PIG-Taskl movierated.pig-89f2a69a-a2c1-49a4-9e
ATS is disabled since yarn.timeline-service.enabled set to false
fs.default.name is deprecated. Instead, use fs.defaultFS fs.default.name is deprecated. Instead, use fs.defaultFS
Unable to load native-hadoop library for your platform... using builtin-java pig features used in the script: COGROUP,FILTER
fs.default.name is deprecated. Instead, use fs.defaultFS
```

Task1_movierated_map ran on map reduce

```
File Edit View Search Terminal Help

[acadgild@localhost PIG]$ pig -brief Task1_movierated_map.pig

18/11/26 05:11:11 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

18/11/26 05:11:11 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

18/11/26 05:11:11 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

Apache Pig version 0.16.0 (17146530) compiled Jun 01 2016, 23:10:49

Logging error messages to: /home/acadgild/PIG/pig 1543189271261.log

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j
/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/or
LoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

Inable to load native-badoon library for your platform. Using builtin-java classes where applicable
```

```
File Edit View Search Terminal Help

TechnoCalyps (2006))

TachnoCalyps (2006))

Starquest II (1996))

Rivers of Sand (1974))

Altar of Fire (1976))

Šilenė smutná princezna (1968))

The Wonders of Aladdin (1961))

Phobos. Fear Kills (2010))

The Final Storm (2010)

In a Heartbeat (2017))

"Blood)

Jungle Woman (1944))

To Be Fat Like Me (2007))

Cadet Kelly (2002))

The Scheming Gambler's Paradise (1905))

The Man with the Rubber Head (1901))

The Living Playing Cards (1905))

The Hilarious Posters (1906))

The Devilish Tenant (1909))

The Untameable Whiskers (1904))

The Imperceptable Transmutations (1904))

Pooh's Heffalump Halloween Movie (2005))

The One Ann Band (1900))

The Fat and Lean Wrestling Match (1900))

Deep Hearts (1981))

Mom (2017))

The Morning After (2015))

St. Michael Had a Rooster (1972))

House of Horrors (1946))

Shadow of the Blair Witch (2000))

The Burkittsville 7 (2000))

Caged Heat 3000 (1995))

Robin Hood (1991))

Subduel

Century of Birthing (2011)

Betrayal (2003))

Satan Triumphant (1917))

ig script completed in 4 minutes, 51 seconds and 782 milliseconds (291782 ms)

ou have new mail in /var/spool/mail/acadgild

acadgild@localhost PIG|$
```

Task2_countmovierated (How many times a movie was rated) ran on local

```
File Edit View Search Terminal Help

[acadgild@localhost PIG]$ pig -brief -x local Task2_countmovierated.pig

5LF4J: Class path contains multiple SLF4J bindings.

5LF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/h/org/slf4j/impl/StaticLoggerBinder.class]

5LF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j
LoggerBinder.class]

5LF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

5LF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]

18/11/26 05:13:57 INFO pig.ExecTypeProvider: Trying ExecType: LOCAL

18/11/26 05:13:57 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49

Logging error messages to: /home/acadgild/PIG/pig_1543189437506.log

user.name is deprecated. Instead, use mapreduce.job.user.name

Default bootup file /home/acadgild/.pigbootup not found

mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
```

Task2_countmovierated_map(How many times a movie was rated) ran on map reduce

```
File Edit View Search Terminal Help

[acadgild@localhost PIG]$ pig -brief Task2_countmovierated_map.pig

18/11/26 05:20:34 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

18/11/26 05:20:34 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

18/11/26 05:20:34 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49

Logging error messages to: /home/acadgild/PIG/pig_1543189834339.log

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7
LoggerBinder.class]

SLF4J: See http://www.slf4i.org/codes.html#multiple bindings for an explanation.
```

```
("Crimson Rivers 2: Angels of the Apocalypse (Rivières pourpres II - Les anges de l'apocalypse,240)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou kodomo) (Journey to Agartha) (2011),102)
(Dragon Ball Z: Bio-Broly (boragon bóru Z 11: Súpă senshi gekihal Katsun ow aver da) (1994),33)
(The Garden of Sinners - Chapter 7: — not nothing heart. (Murder Speculation Part B) (2009),132,24746)
(The Garden of Sinners - Chapter 7: — not nothing heart. (Murder Speculation Part B) (2009),132,77)
(Dragon Ball: Sleeping Princess in Devil's Castle (Doragon bóru: Majinjó einem compagnie) (1973),7)
(Dragon Ball: Sleeping Princess in Devil's Castle (Doragon bóru: Majinjó einem Supernacht) (2008),13)
(Ellen ten Damme: As I Was Wondering Where This Mixed-up Little Life of Mine Was Leading To (2007),1)
(Films to Keep You Awake: The Christmas Tale (Peliculas para no dormir: Cuento de navidad) (2005),10)
(Images of the World and the Inscription of War (Bilder der Welt und Inschrift des Krieges) (1989),4)
(Neon Genesis Evangelion: Death & Rebirth (Shin seski Evangelion Gekijó-ban). Shito shinsei) (1997),43)
(Dragon Ball Z: The Return of Cooler (Doragon bóru Z 6: Gekitotsul Hyakuoku pawa no senshi) (1992),44)
(Bicycle Thieves (a.k.a. The Bicycle Thielf) (a.k.a. The Bicycle Thieves) (Ladri di biciclette) (1948),3953)
(Dragon Ball Z: Bojack Unbound (Doragon bóru Z 9: Ginga girigiri!! Butchigiri no sugoi yatsu) (1993),35)
(Amazon Jack 2: The Movie Star (a.k.a. Hugo the Movie Star) (Jungledyret 2 den store filmhelt) (1996),3)
(History of Postwar Japan as Told by a Bar Hostess (Nippon Sengoshi - Madamu onboro no Seikatsu) (1990),3)
(Hysex Life... or How I Got Into an Argument (Comment je me suis disputé. (ma vica exuelle)) (1996),16)
(All the Boys Are Called Patrick (Charlotte et Veronique ou tous les garcons s'appellent Patrick) (1996),16)
(All the Boys Are Called Patrick (Charlotte et Veronique ou tous les garcons s'appellent Patrick) (1996),16)
(All the Boys Are Called Patrick) (a.k.a. Hugo the Movie Granting Fe. - Wir Kinner
```

```
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost PIG]$ pig -brief -x local Task3_averagemovierating.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.
LoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an exp
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/11/26 05:37:42 INFO pig.ExecTypeProvider: Trying ExecType: LOCAL
18/11/26 05:37:42 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecTy
Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49
```

Output:

```
"Crimson Rivers 2: Angels of the Apocalypse (Rivières pourpres II - Les anges de l'apocalypse, 2.8583333333333333)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou Kodomo) (Journey to Agartha) (2011), 3.2941176470580234)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou Kodomo) (Journey to Agartha) (2011), 3.2941176470580234)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou Kodomo) (Journey to Agartha) (2011), 3.2941176470580234)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou Kodomo) (Journey to Agartha) (2011), 3.2941176470580234)
(Children Who Chase Lost Voices from Deep Below (Hoshi o ou Kodomo) (Journey to Agartha) (2001), 3.2941176470580234)
(Children Who Chase Lost Voices from Children (Hordren Speculation Part B) (2009), 3.1333333333333333)
(Children Who Chase Lost Voices from Children (Hordren Speculation Part B) (2009), 3.1333333333333333333)
(Children Who Chase Lost Voices from Children (Hordren Speculation Part B) (2009), 3.133333333333333333333)
(Children Who Chase Lost Voices from Children (Hordren Children Children Children (Hordren Children Children Children Children (Hordren Children Children Children Children (Hordren Children Children Children Children Children Children (Hordren Children Children Children Children Children Children (Hordren Children Children Children Children Children Children Children Children (Hordren Children C
```

Task3_averagemovierating_map(Average rating for movie) ran on map reduce

```
File Edit View Search Terminal Help

[acadgild@localhost PIG]$ pig -brief Task3_averagemovierating_map.pig

18/11/26 05:44:29 INFO pig.ExecTypeProvider: Trying ExecType: LOCAL

18/11/26 05:44:29 INFO pig.ExecTypeProvider: Trying ExecType: MAPREDUCE

18/11/26 05:44:29 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49

Logging error messages to: /home/acadgild/PIG/pig 1543191269414.log

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4
/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/o
LoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.slf4j.impl.Log4jlcoggerFactory]

Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

Default bootup file /home/acadgild/.pigbootup not found

mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

fs.default.name is deprecated. Instead, use mapreduce.jobtracker.address

fs.default.name is deprecated. Instead, use fs.defaultFS

Connecting to badoon file system at: hdfs://localhost:8020
```

File Edit View Search Terminal Help

("Crimson Rivers 2: Angels of the Apocalypse (Rivières pourpres II - Les anges de l'apocalypse, 2.858333333333333)

(Children Who Chase Lost Voices from Deep Below (Hoshi o ou kodomo) (Journey to Agartha) (2011), 3.2941176470888234)

(Dragon Ball Z: Bio-Broly (Doragon böru Z 11: Spå senshi gekhal Katsu no wa ore da) (1994), 2.7878787878787881)

(Harry Potter and the Sorcerer's Stone (a.k.a. Harry Potter and the Philosopher's Stone) (2001), 3.484603572294512)

(The Garden of Sinners - Chapter 7:not nothing heart. (Murder Speculation Part B) (2009), 3.1333333333333)

(Now Where Did the Seventh Company Get to? (Mais où est donc passée la 7ême compagnie) (1973), 3.0)

(Dragon Ball: Sleeping Princess in Devil's Castle (Doragon böru: Majinjö no nemuri hime) (1987), 2.5)

(Henry Kissinger: Secrets of a Superpower (Henry Kissinger - Geheimmisse einer Supermacht) (2008), 3.0)

(Ellen ten Damme: As I Was Wondering Where This Mixed-up Little Life of Mine Was Leading To (2007), 2.0)

(Films to Keep You Awake: The Christmas Tale (Peliculas para no dormir: Cuento de navidad) (2005), 2.8)

(Images of the World and the Inscription of War (Bilder der Welt und Inschrift des Krieges) (1989), 3.5)

(Neon Genesis Evangelion: Death & Rebirth (Shin seiki Evangelion Gekijō-ban: Shito shinsei) (1997), 3.64284415960265)

(Dragon Ball Z: The Return of Cooler (Doragon böru Z 6: Gekitotsul Hyakuoku pawā no senshi) (1992), 3.00

(Oblivion Island: Maruka and the Magic Mirror (Hottarake no shima - Haruka to maho no kagami) (2009), 2.6)

(Bicycle Thieves) (a.k.a. The Bicycle Thiefy) (a.k.a. The Bicycle Thieves) (Ladri di biciclette) (1940), 3.987857323551733)

(Dragon Ball Z: Bojack Unbound (Doragon böru Z 9: Ginag girigiri!! Butchigiri no sugoi yatsu) (1993), 2.942857142857143)

(Amazon Jack 2: The Movie Star (a.k.a. Hugo the Movie Star) (Jungledyret 2 - den store filmhelt) (1996), 3.66666666666666)

(Mistory of Postwar Japan as Told by a Bar Hostess (Nipon Sengoshi - Madamu onboro no Seikatsu) (1996) (1987),3.3333333333333333333333333333333)
(Revolutionary Girl Utena: Adolescence of Utena (a.k.a. Revolutionary Girl Utena the Movie) (Shoujo kakumei Utena: Adolescenc mokushiroku) (1999),3.4193548387096775)
(Dragon Ball Z the Movie: The World's Strongest (a.k.a. Dragon Ball Z: The Strongest Guy in The World) (Doragon bôru Z: Kono do de ichiban tsuyoi yatsu) (1990),2.7960976744186047)
Pig script completed in 11 minutes, 47 seconds and 50 milliseconds (707050 ms)
fou have new mail in /var/spool/mail/acadgild
[acadgild@localhost PIG]\$

[BB acadgild

[BB acadgild] PIG acadgild ☐ acadgild@localhost;~/... → Task3_averagemovier...

File Edit View Search Terminal Help