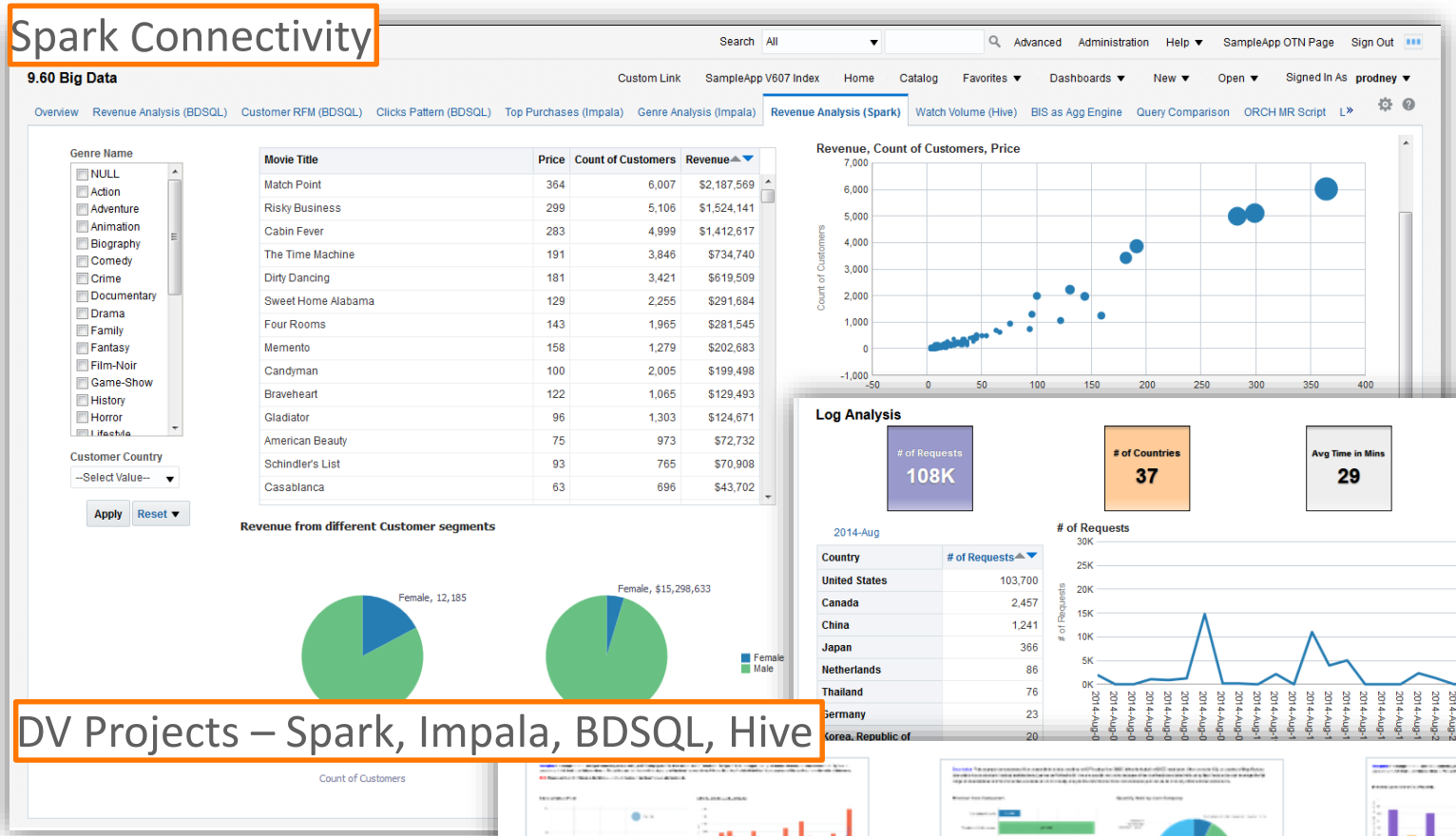
SampleApp BigData Examples

- Contact your local Oracle Sales rep for details on enabling BigData on your SampleAppv607 virtual box image
- See examples below

Big Data Related Samples

Spark Connectivity



CDH 5.7, Spark, BDSQL

cloudera

What's New In CDH 5.7.x

Continue reading:

- What's New in CDH 5.7.0
- What's New in CDH 5.7.1

What's New in CDH 5.7.0

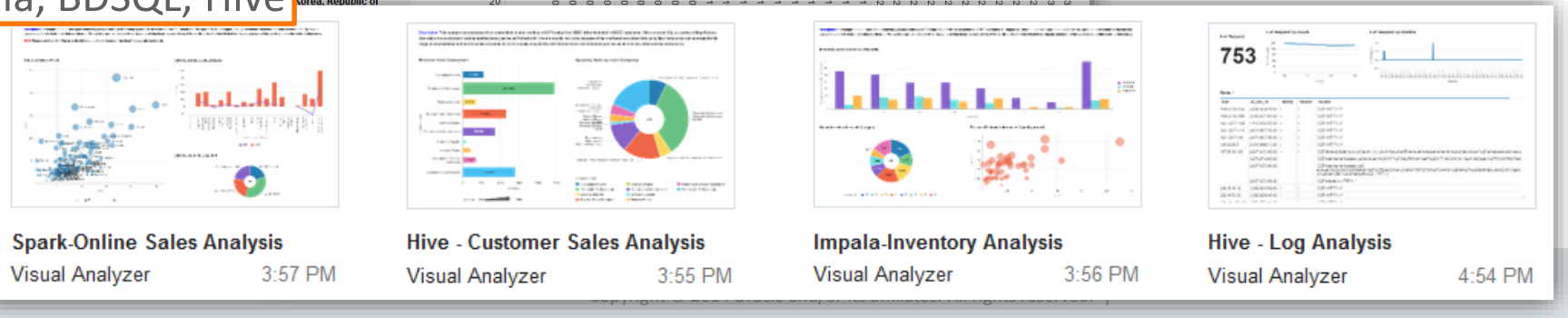
The following sections describe new features introduced in CDH

- Operating System Support
- Apache Hadoop
- Apache HBase
- Apache Hive
- Hue
- Apache Impala (incubating)
- MapReduce
- Apache Oozie
- Cloudera Search
- Apache Spark
- Apache Sentry
- YARN

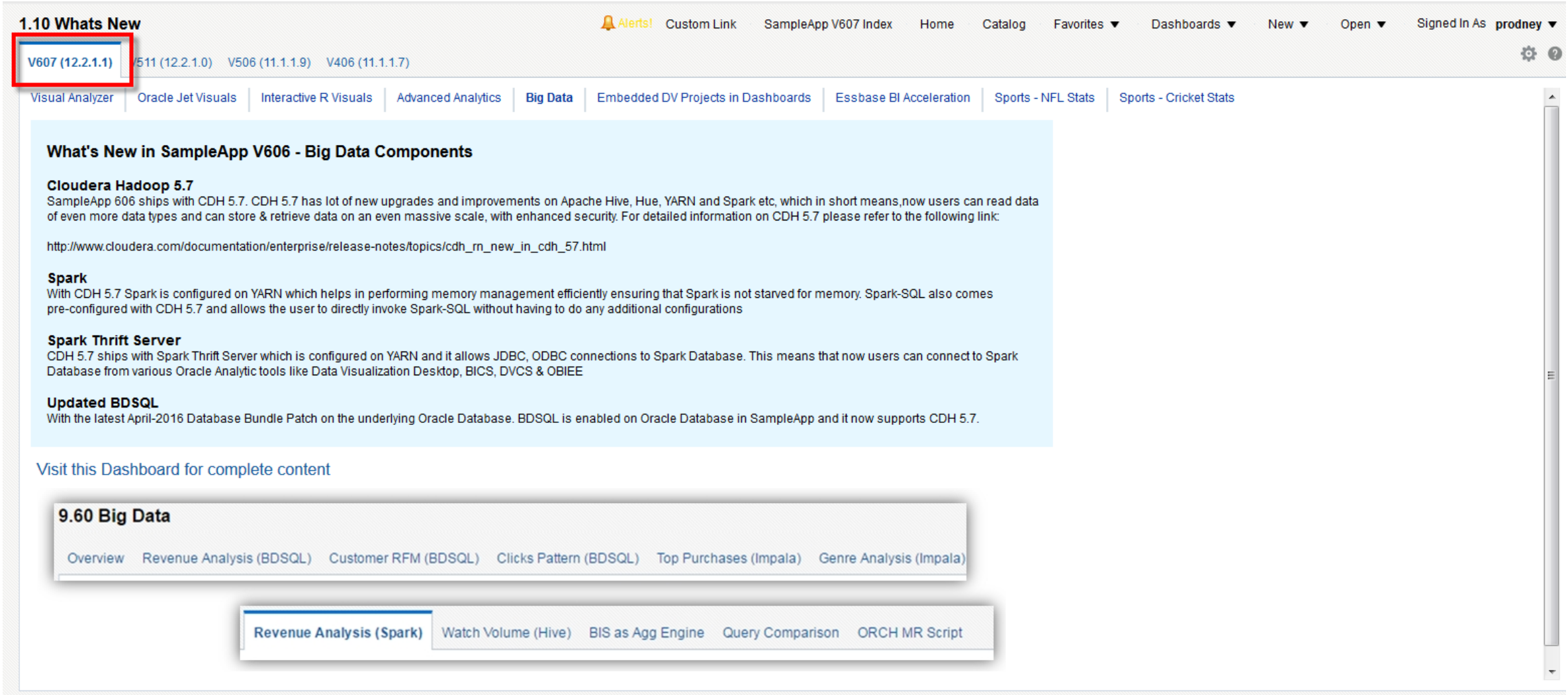
APACHE Spark

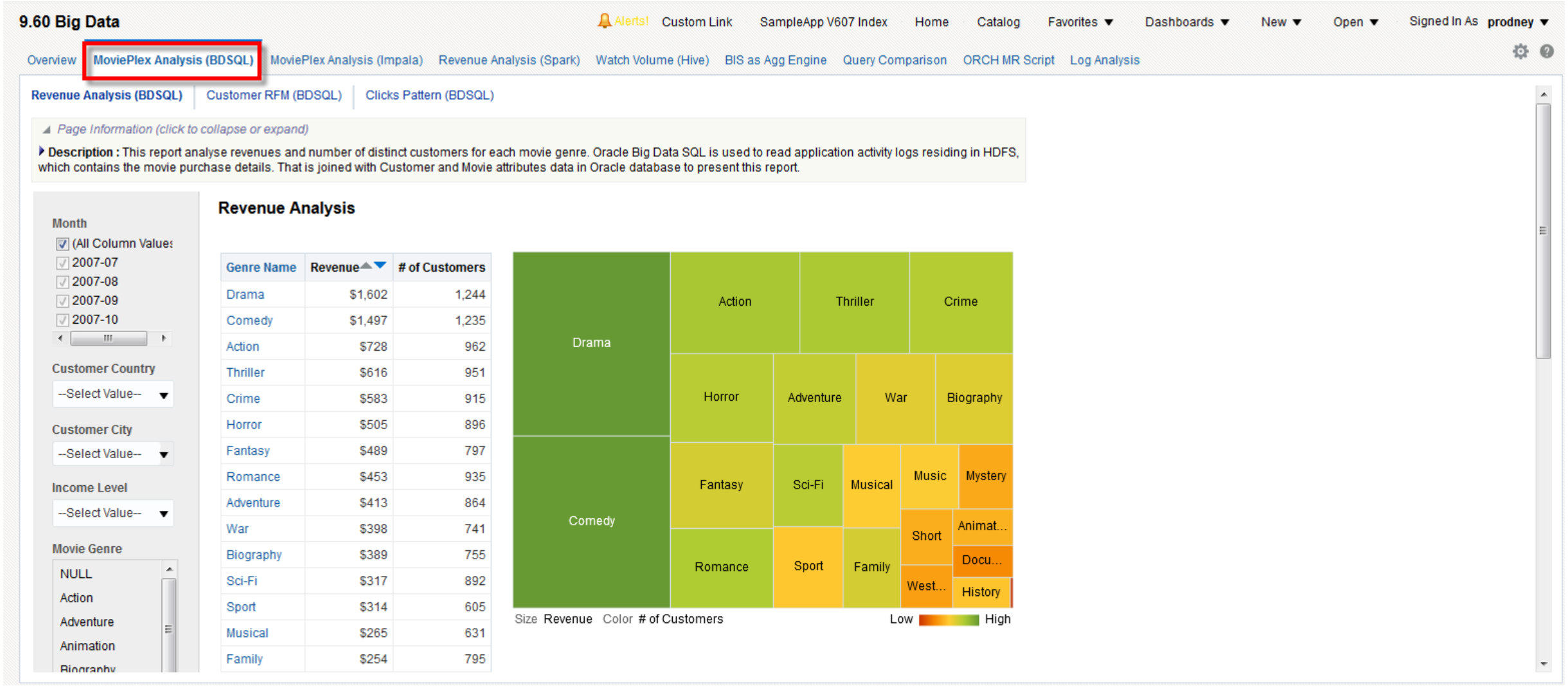
Log Analysis

DV Projects – Spark, Impala, BDSQL, Hive



ORACLE





9.60 Big Data

Alerts! | Custom Link | SampleApp V607 Index | Home | Catalog | Favorites | Dashboards | New | Open | Signed In As **prodney**

Overview | MoviePlex Analysis (BDSQL) | MoviePlex Analysis (Impala) | **Revenue Analysis (Spark)** | Watch Volume (Hive) | BIS as Agg Engine | Query Comparison | ORCH MR Script | Log Analysis

Description : This demonstrates Spark connectivity to data residing in HDFS using Spark ODBC driver included in OBIEE installation.
The Spark SQL dialect supports a range of standard elements, plus some extensions for Big Data use cases related to data loading and data warehousing.
Spark provides a high degree of compatibility with the Hive Query Language (HiveQL). Refer Spark SQL Language Reference documentation from Cloudera for more details.

Genre Name

- NULL
- Action
- Adventure
- Animation
- Biography
- Comedy
- Crime
- Documentary
- Drama
- Family
- Fantasy
- Film-Noir
- Game-Show
- History
- Horror
- Lifestyle

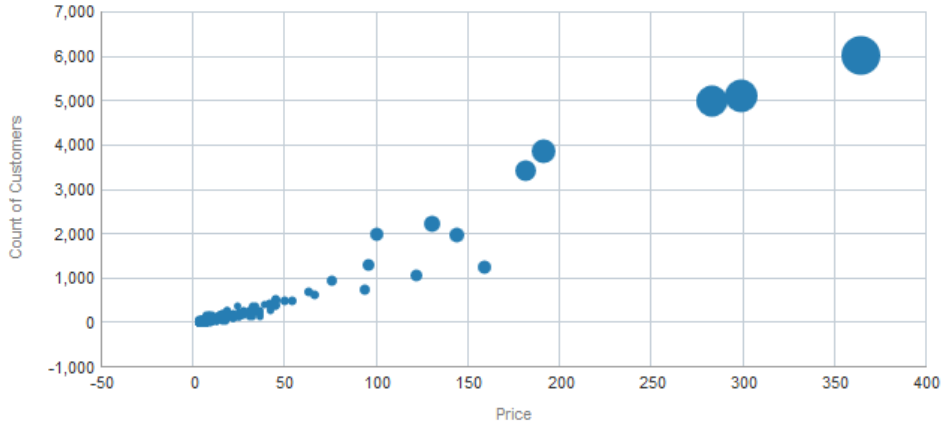
Customer Country

--Select Value--

Apply | Reset

Movie Title	Price	Count of Customers	Revenue
Match Point	364	6,007	$2,187,569
Risky Business	299	5,106	$1,524,141
Cabin Fever	283	4,999	$1,412,617
The Time Machine	191	3,846	$734,740
Dirty Dancing	181	3,421	$619,509
Sweet Home Alabama	129	2,255	$291,684
Four Rooms	143	1,965	$281,545
Memento	158	1,279	$202,683
Candyman	100	2,005	$199,498
Braveheart	122	1,065	$129,493
Gladiator	96	1,303	$124,671
American Beauty	75	973	$72,732
Schindler's List	93	765	$70,908
Casablanca	63	696	$43,702

Revenue, Count of Customers, Price



Revenue from different Customer segments

Revenue

9.60 Big Data

Alerts! Custom Link SampleApp V607 Index Home Catalog Favorites ▼ Dashboards ▼ New ▼ Open ▼ Signed In As prodney ▼

Overview | MoviePlex Analysis (BDSQL) | MoviePlex Analysis (Impala) | Revenue Analysis (Spark) | Watch Volume (Hive) | BIS as Agg Engine | Query Comparison | ORCH MR Script | Log Analysis

Top Purchases (Impala) | Genre Analysis (Impala)

Page Information (click to collapse or expand)

Description : This demonstrates Impala connectivity to data residing in HDFS using Impala ODBC driver included in OBIEE installation (e.g. <BI Home>/bi/common/ODBC/Merant/7.1.4/lib/ARimpala27.so).
The Impala SQL dialect supports a range of standard elements, plus some extensions for Big Data use cases related to data loading and data warehousing.
Impala provides a high degree of compatibility with the Hive Query Language (HiveQL). Refer Impala SQL Language Reference documentation from Cloudera for more details.

Month
- (All Column Values)
- 2007-07
- 2007-08
- 2007-09
- 2007-10

Customer Country
--Select Value--

Income Level
--Select Value--

Genre Name
- NULL
- Action
- Adventure
- Animation
- Biography
- Comedy

Top Customers with Purchases

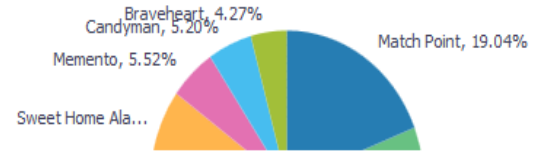
Purchase data from HDFS joined with Customer data in Oracle tables

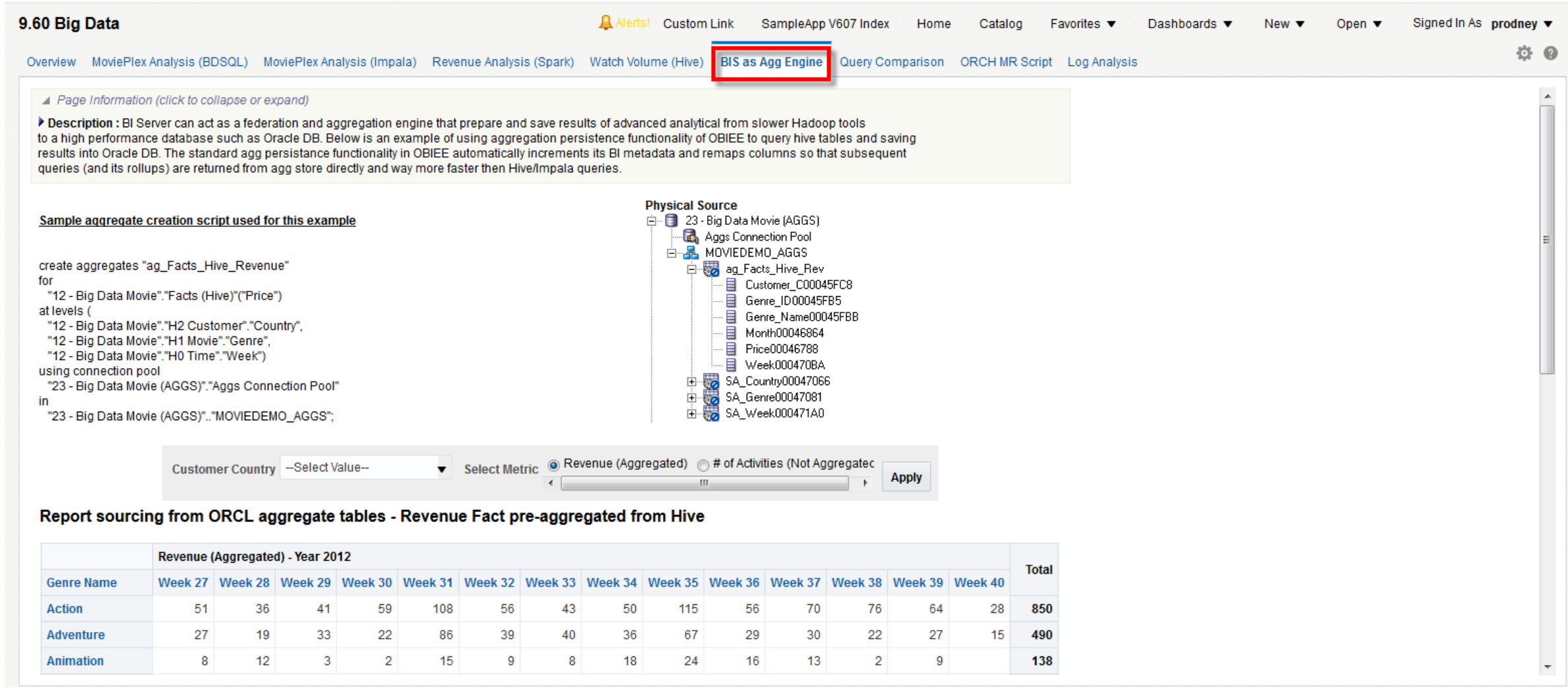
Customer Name	Customer Country	Income Level	Gender	Marital Status	Purchase Cnt
Wilfredo Wilder	United States	B: 30,000 - 49,999	Male	S	24
Carlo Gonzales	United States	B: 30,000 - 49,999	Male	S	23
Maryann Powers	United States	A: Below 30,000	Female	M	21
Phineas Benton	Hungary	C: 50,000 - 69,999	Male	M	21
Jita Nilini	India	A: Below 30,000	Male	S	19
Harold Abe	Japan	C: 50,000 - 69,999	Male	S	18
Heriberto Jong	United Kingdom	B: 30,000 - 49,999	Male	S	18
Jerrold Small	United States	A: Below 30,000	Male	S	18
Rafael Trujillo	United States	E: 90,000 - 109,999	Male	S	18
Tariyel De Vivero	Mexico	B: 30,000 - 49,999	Male	S	18

Top Movies Purchased

Genereted by joining fact data from HDFS with movie attributes in Oracle relational tables

Movie Title	Release Year	No of Purchases	Revenue
Match Point	2005	183	364
Risky Business	1983	150	299
Cabin Fever	2002	142	283
The Time Machine	2002	104	207
Dirty Dancing	1987	91	181
Four Rooms	1995	81	161
Sweet Home Alabama	2002	66	131
Memento	2000	53	158
Candyman	1992	50	99
Braveheart	1995	41	122





	Revenue (Aggregated) - Year 2012														Total
Genre Name	Week 27	Week 28	Week 29	Week 30	Week 31	Week 32	Week 33	Week 34	Week 35	Week 36	Week 37	Week 38	Week 39	Week 40	
Action	51	36	41	59	108	56	43	50	115	56	70	76	64	28	**850**
Adventure	27	19	33	22	86	39	40	36	67	29	30	22	27	15	**490**
Animation	8	12	3	2	15	9	8	18	24	16	13	2	9		**138**