Spark Simulation

A simulation, composing a *Data Source ~> Spark Flow ~> Data Sink pipeline*.

- Source ~ Cassandra, Hdfs, Kafka, Text
- Flow ~ Transformations and Action
- Sink ~ Cassandra, Hdfs, Kafka, Text

Technologies

- Scala
- ScalaFX
- Spark
- Kafka
- Cassandra
- Cassandra-Spark Connector

Rationale

- Spark engineers work with Scala.
- Spark simulations can be built with ScalaFX, a JavaFX wrapper.
- JavaFX includes Charts, 2D, 3D and Animations.
- JavaFX apps and controls are decorated by CSS.
- Third party JavaFX charts and components exist.
- JavaFX includes a WebView component that allows for the rendering of Javascript content. For instance, D3 could be used to build an exotic chart.

Visual Spark

Executes and visualizes a *Data Source ~> Spark Flow ~> Data Sink* pipeline as:
- Source: Kafka Topic Messages
- Flow: Time Series Chart
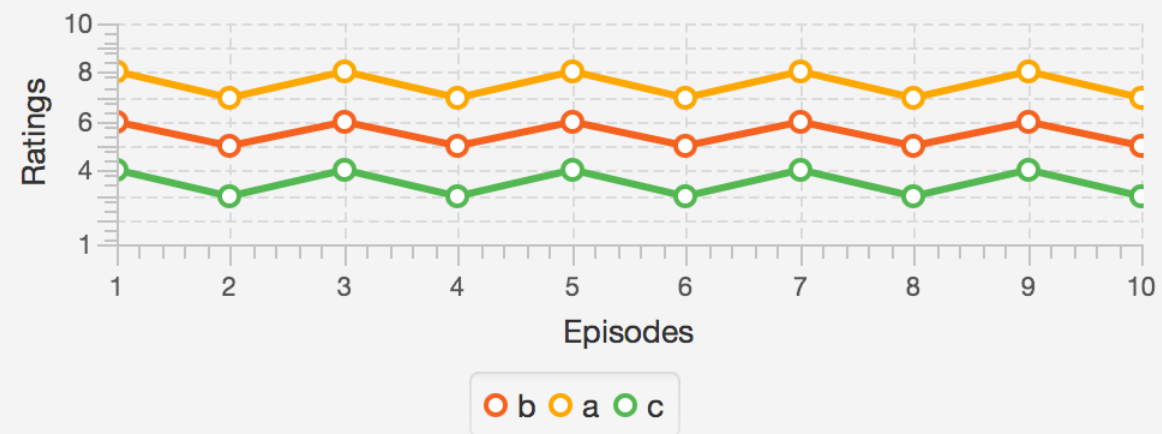- Sink: Aggregate Pie Chart

via a ScalaFX application.

Visual Spark

Play

30 messages processed in 22 seconds.

Source

Program	Season	Episode	Rating	
a	1	1	8	
a	1	2	7	
a	1	3	8	
a	1	4	7	
a	1	5	8	
a	1	6	7	
a	1	7	8	
a	1	8	7	

Flow

Episode Ratings

Ratings

10
8
6
4
2

1 2 3 4 5 6 7 8 9 10

Episodes

○ b ○ a ○ c

Sink

Program Ratings

b(33%)   c(21%)

a(45%)

● a(45%)   ● b(33%)   ● c(21%)