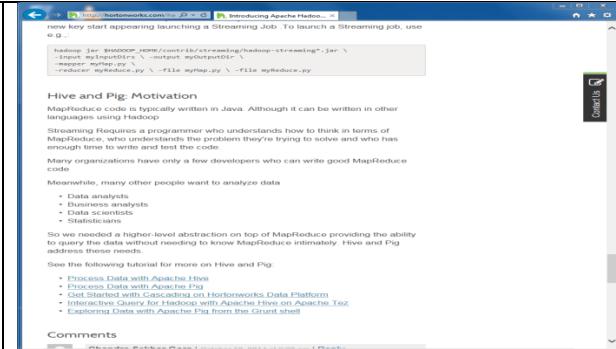
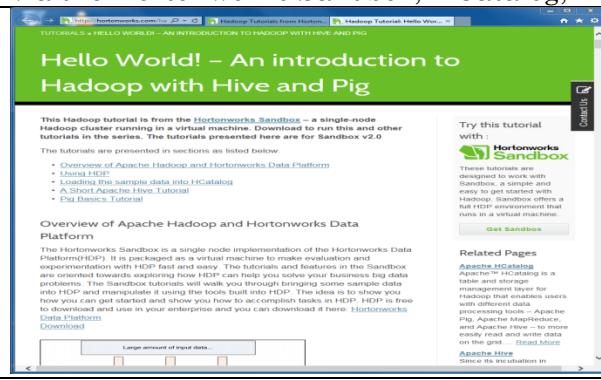
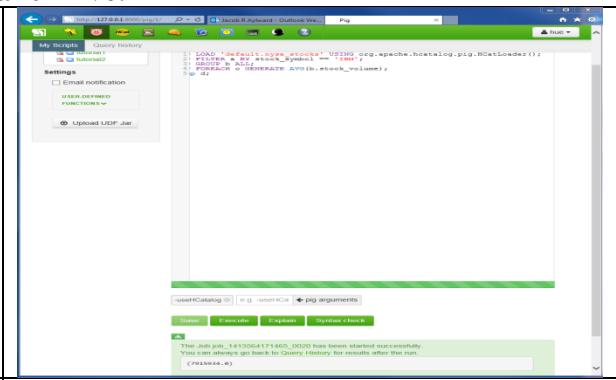
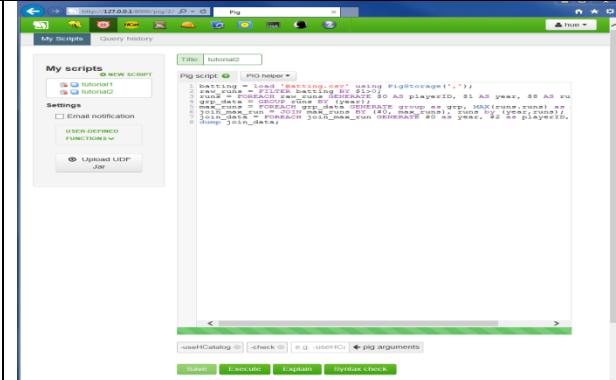


OCTOBER 27, 2014

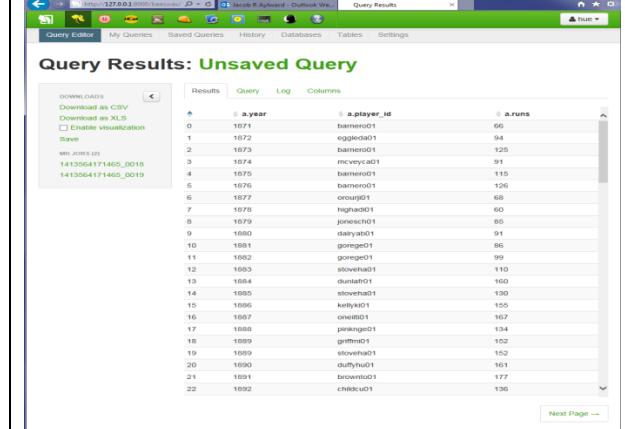
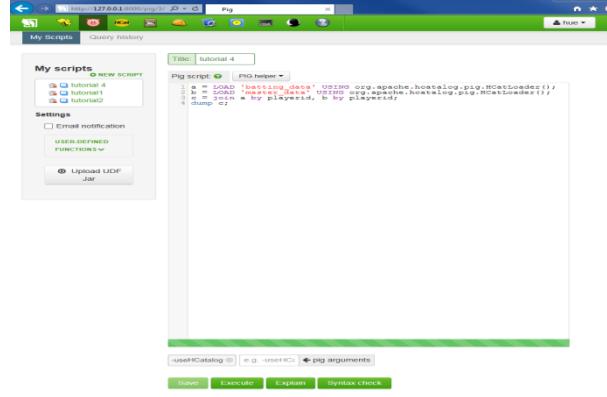
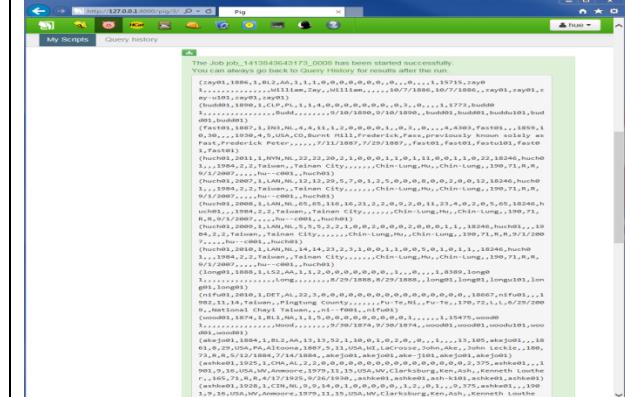
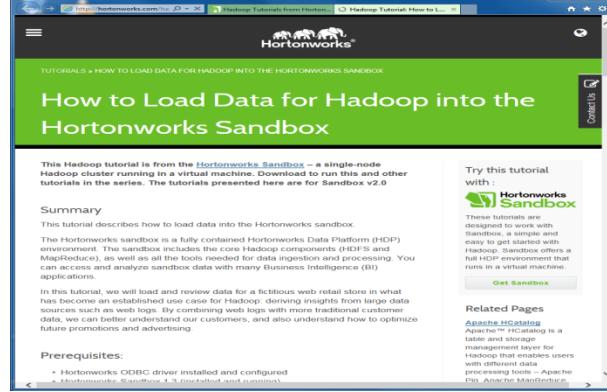
# JACOB AYLWARD

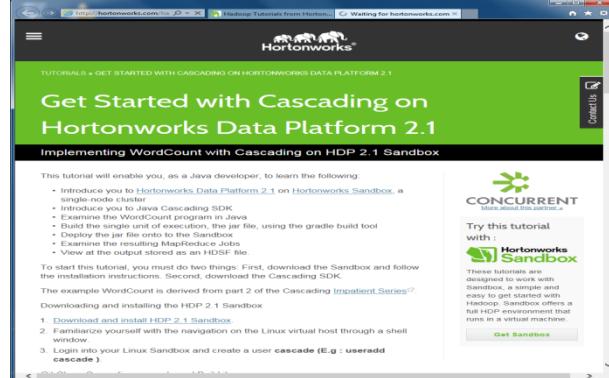
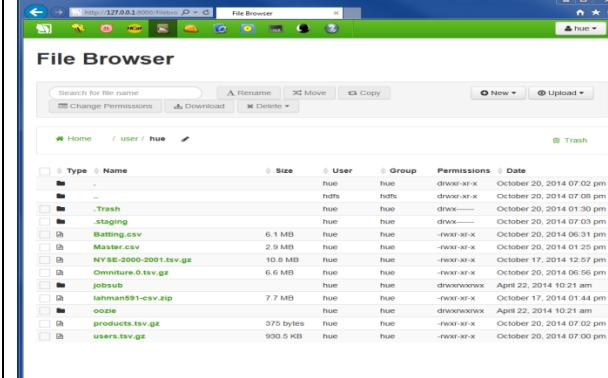
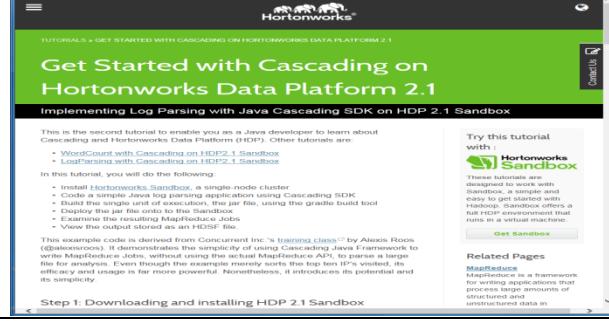
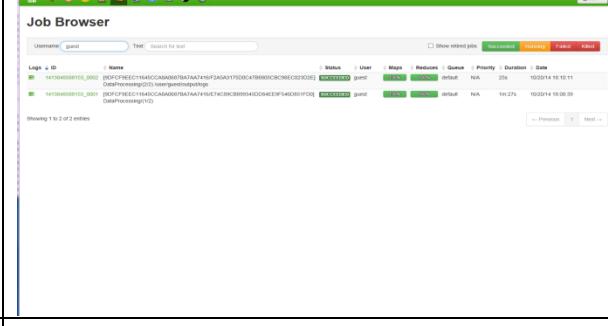
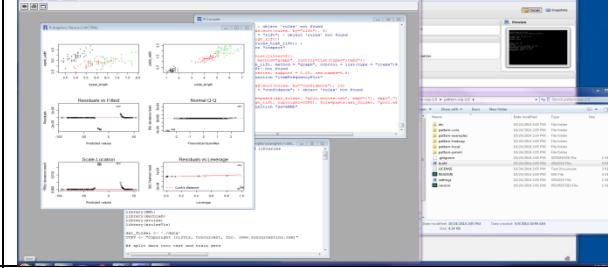
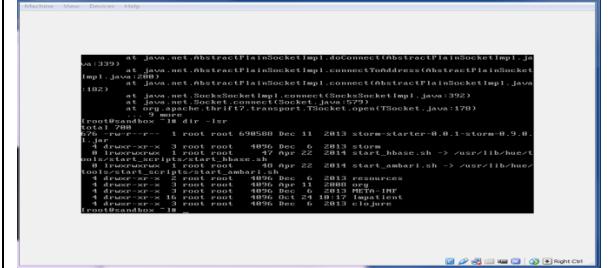
## HORTONWORKS TUTORIAL

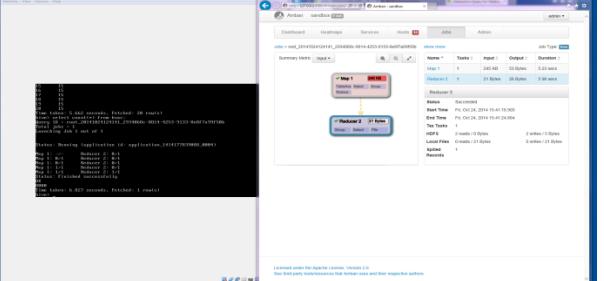
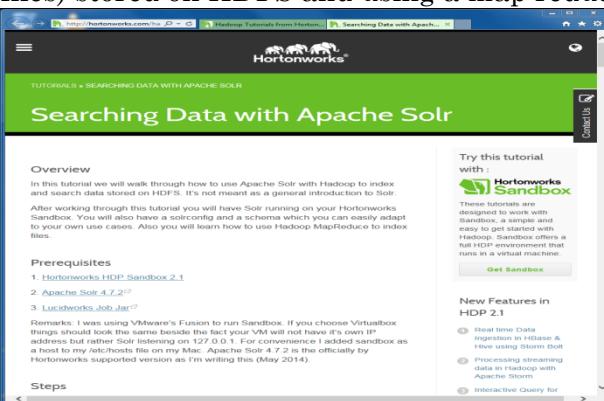
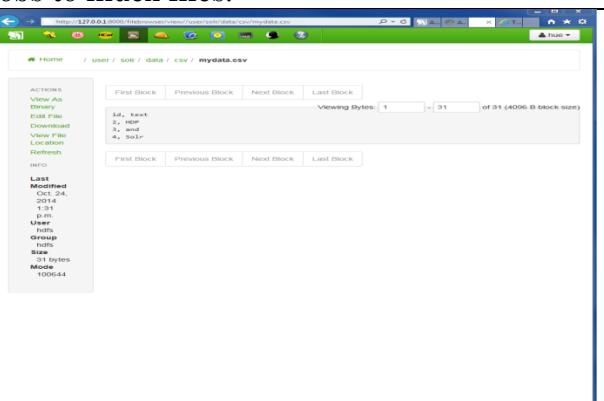
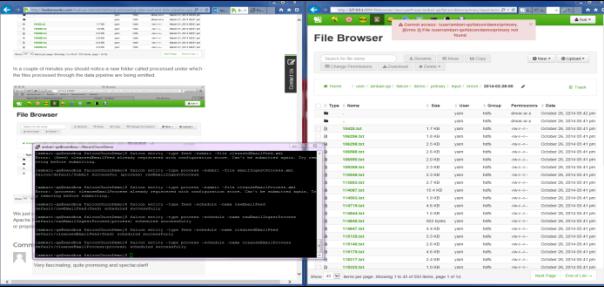
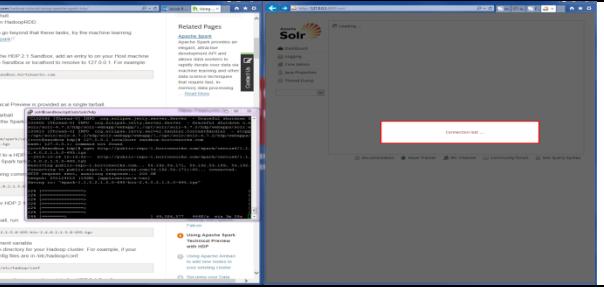
JACOB R. AYLWARD  
FROSTBURG STATE UNIVERSITY  
101 Braddock Rd, Frostburg, MD 21532

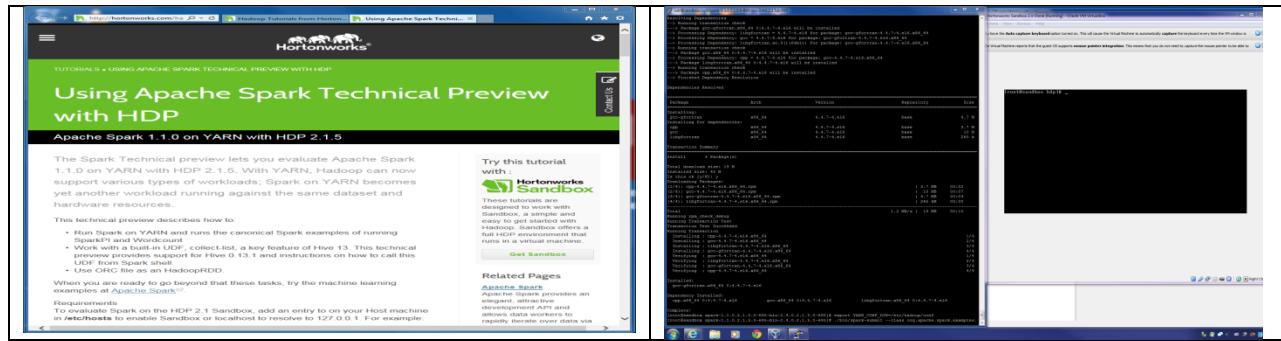
<b>Hello World Tutorials:</b>	
<b>Tutorial 1</b>	Introducing Apache Hadoop to Developers
<p>Apache Hadoop is a community driven open-source project governed by the Apache Software Foundation. It was originally implemented at Yahoo...</p>	
	
<b>Tutorial 2</b>	Hello World! – An introduction to Hadoop with Hive and Pig
<p>This Hadoop tutorial provides a short introduction into working with big data in Hadoop via the Hortonworks Sandbox, HCatalog, Pig and Hive.</p>	
	
<b>Tutorial 3</b>	How to Process Data with Apache Hive
<p>This Hadoop tutorial shows how to process Data with Hive using a set of Baseball statistics on American players from 1871-2011.</p>	
	
<b>Tutorial 4</b>	How to Process Data with Apache Pig
<p>This Hadoop tutorial shows how to process Data with Apache Pig using a set of Baseball statistics on American Players from 1871-2011.</p>	

## Jacob Aylward

 <p><b>Tutorial 5</b></p> <p>In this tutorial, you will learn how to load a data file into HDFS; Learn about 'FILTER', 'FOREACH' with examples; storing values into HDFS and Grunt shell's file commands.</p>	 <p><b>Exploring Data with Apache Pig from the Grunt shell</b></p>
 <p><b>Tutorial 6</b></p> <p>In this tutorial, we will load and review data for a fictitious web retail store in what has become an established use cause for Hadoop: deriving insight from large data sources such as web logs.</p>	 <p><b>How to Load Data for Hadoop into the Hortonworks Sandbox</b></p>
 <p><b>Tutorial 7</b></p> <p>How to get started with Cascading and Hortonworks Data Platform using the Word Count Example.</p>	 <p><b>Get Started with Cascading on Hortonworks Data Platform 2.1</b></p>

	
<h2>Tutorial 8</h2>	<h2>Get Started with Cascading on Hortonworks Data Platform 2.1</h2>
<p>This is the second tutorial to enable you as a Java developer to learn about Cascading and Hortonworks Data Platform (HDP). Other tutorials are:....</p>	
	
<h2>Tutorial 9</h2>	<h2>Cascading Pattern</h2>
<p>Learn how to use Cascading Pattern to quickly migrate Predictive Models (PMML) from SAS, R, MicroStrategy onto Hadoop and deploy them at scale.</p>	
	
<h2>Tutorial 10</h2>	<h2>Processing streaming data in Hadoop with Apache Storm</h2>
<p>How to use Apache Storm to process real-time streaming data in Hadoop with Hortonworks Data Platform.</p>	
	
<h2>Tutorial 11</h2>	<h2>Interactive Query for Hadoop with Apache</h2>

		<b>Hive on Apache Tez</b>
How to use Apache Tez and Apache Hive for Interactive Query with Hadoop and Hortonworks Data Platform 2.1.		
		
<b>Tutorial 12</b>	Searching Data with Apache Solr	
In this tutorial we will walk through how to run Solr in Hadoop with the index (solr data files) stored on HDFS and using a map reduce jobs to index files.		
		
<b>Tutorial 13</b>	Define and Process Data Pipelines in Hadoop with Apache Falcon	
Use Apache Falcon to define an end-to-end data pipeline and policy for Hadoop and Hortonworks Data Platform 2.1.		
		
<b>Tutorial 14</b>	Using Apache Spark Technical Preview with HDP	
This guide describes how to run Spark on YARN. It also provides the canonical examples of running SparkPL and Wordcount with Spark shell.		



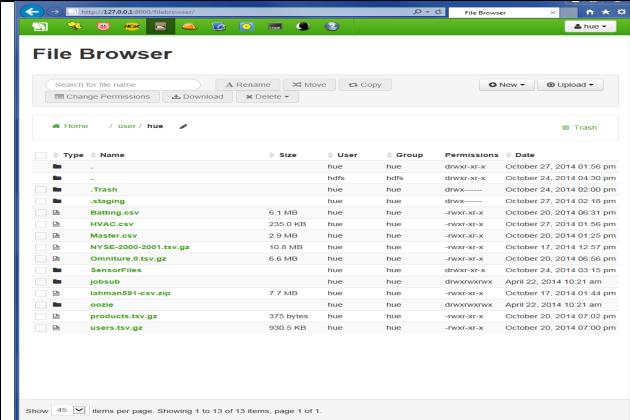
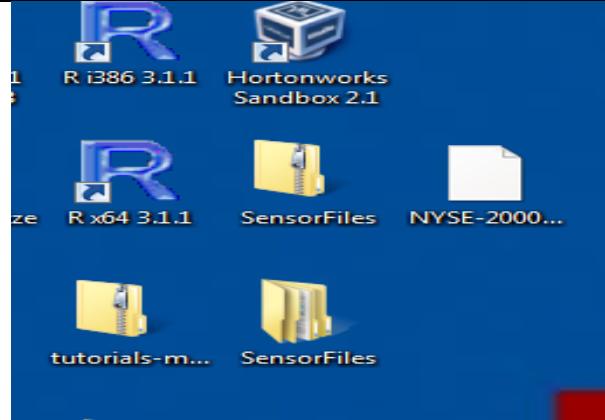
## Real World Examples:

### Example 7

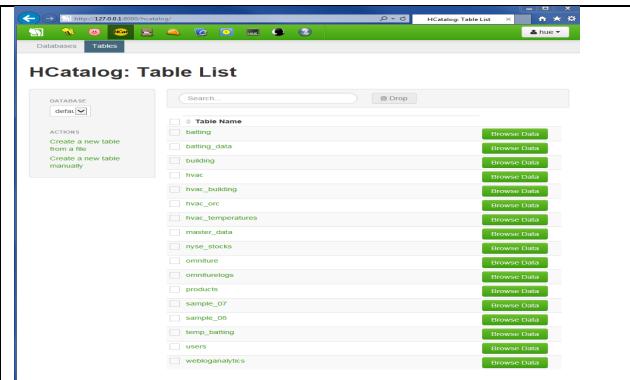
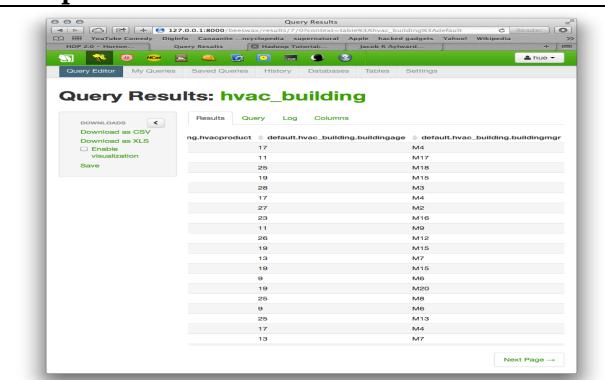
### How to Analyze Machine and Sensor Data

The tutorial describes how to refine data from heating, ventilation, and air conditioning (HVAC) systems using the Hortonworks Data Platform, and how to analyze the refined sensor data to maintain optimal building temperatures.

#### Step 1: Download and Extract the Sensor Data Files



#### Step 2: Load the Sensor Data into the Hortonworks Sandbox



#### Step 3: Run Two Hive Scripts to Refine the Sensor Data

**Query Results: hvac\_temperatures**

default.hvac_temperatures.date	default.hvac_temperatures.time	default.hvac_temperatures
6/1/13	0:00:01	66
6/2/13	2:00:01	70
6/4/13	3:00:01	67
6/5/13	4:00:01	66
6/6/13	5:00:01	67
6/7/13	6:00:01	70
6/8/13	7:00:01	70
6/9/13	8:00:01	66
6/10/13	9:00:01	65
6/11/13	10:00:01	67
6/12/13	11:00:01	69
6/13/13	12:00:01	66
6/14/13	13:00:01	66
6/15/13	14:00:01	67
6/16/13	15:00:01	65
6/17/13	16:00:01	67
6/18/13	17:00:01	66
6/19/13	18:00:01	69
6/20/13	19:00:01	67
6/21/13	20:00:01	69
6/22/13	21:00:01	66
6/23/13	22:00:01	67

**Query Results: hvac\_building**

default.hvac_building.date	default.hvac_building.time	default.hvac_building.targettemp
6/1/13	0:00:01	66
6/2/13	2:00:01	70
6/4/13	3:00:01	67
6/5/13	4:00:01	68
6/6/13	5:00:01	67
6/7/13	6:00:01	70
6/8/13	7:00:01	70
6/9/13	8:00:01	66
6/10/13	9:00:01	65
6/11/13	10:00:01	67
6/12/13	11:00:01	69
6/13/13	12:00:01	69
6/14/13	13:00:01	65
6/15/13	14:00:01	67
6/16/13	15:00:01	65
6/17/13	16:00:01	67
6/18/13	17:00:01	66
6/19/13	18:00:01	69
6/20/13	19:00:01	67
6/21/13	20:00:01	69
6/22/13	21:00:01	66
6/23/13	22:00:01	67

#### Step 4: Access the Refined Sentiment Data with Microsoft Excel

query\_result [Last saved by user] - Excel

Book2 - Microsoft Power Map for Excel

Tour 1

Power Map needs data to work with.

Select a table or a range in Excel and click on "Add Selected Data to Power Map" under Insert > Map.

Map

Insert

Send a Smile 😊

Outlook

11:28p

5:26p

3:59p

3:10p

2:22p

11:05p

10:05p

7:46 PM  
10/29/2014

#### Step 5: Visualize the Sensor Data Using Excel Power View

