splunk> .conf2017

Fun with Analytics

Marcello Lino | SVP, Security Analytics Engineering

James Sullivan | VP, Security Analytics Engineering

Sep/2017| Washington, DC

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Agenda

Fun with Analytics

- Intro – About Us
- What is this “fun” all about?
- Hardware
- The Solution
- Results

Who Are We?

Marcello Lino

- Senior Security Engineer
  - 25+ years of IT experience
  - Background in database, development, *NIX
  - Using Splunk for 3 years
  - Splunk Certified Architect
- Hobbies
  - Play guitar (mostly metal \m/)

James Sullivan

- Senior Security Engineer
  - 15+ years of IT experience
  - Background in*NIX, Python, Security
  - Using Splunk for 3 years
  - Splunk Certified Architect
- Hobbies
  - Hiking

What Is This Fun All About?

Science Project

The idea came from a 6th grade science project.

- Objective was to grow plants on different soil types and analyze the results
  - Soil matter→ Clay, Sand, Dirty and Silt.
  - And more and more questions started to come up …
    - How do we measure the amount of water required?
    - Do we need a green house to ensure a constant, stable environment?
    - How do we know if the environment is healthy or not?
      - Sun light?
      - Temperature?
      - Moisture?
      - Humidity?

Let’s Make It Fun

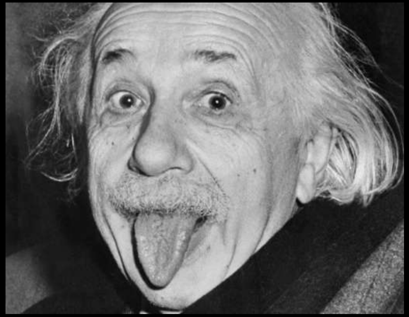
Science Project

So we thought … Let’s collect all the required data automatically!

- Having this data collected allows near real-time analysis on:
  - Illumination (lux)
  - Soil moisture
  - Current green house temperature and humidity
- Data is streamed to Splunk for:
  - Analytics
  - Visualizations

“A **Theory** Can Be Proved By **Experiment**; But No Path Leads From Experiment To The Birth Of A Theory.”

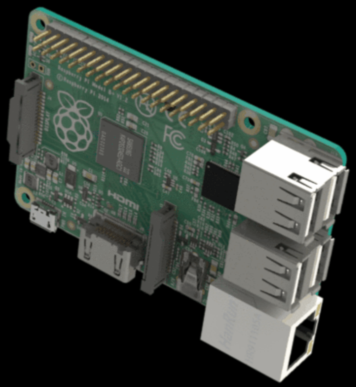
Albert Einstein

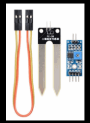
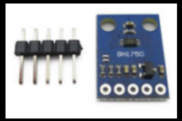


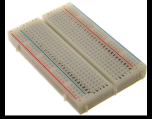
Equipment Used

Science Project

- Raspberry Pi 3
- Sensors, chips, etc:
  - Light intensity sensor BH1750
  - MCP3008 Microchip 8 Channel 10 bit
  - Breadboard MB102 & jumper cables
  - Temperature and Humidity sensor AM2302
  - Soil Moisture Sensor And Automatic Watering System (AWS was not implemented)
  - Traffic light LEDs



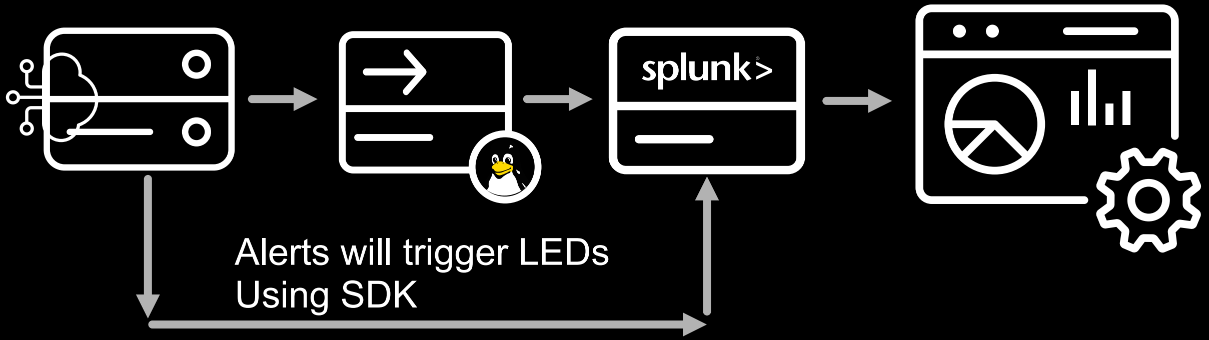


Software Used

Science Project

- Python scripts created for data collection and alerts
- Splunk Universal Forwarder
- Splunk Enterprise (free version!)

Process Flow

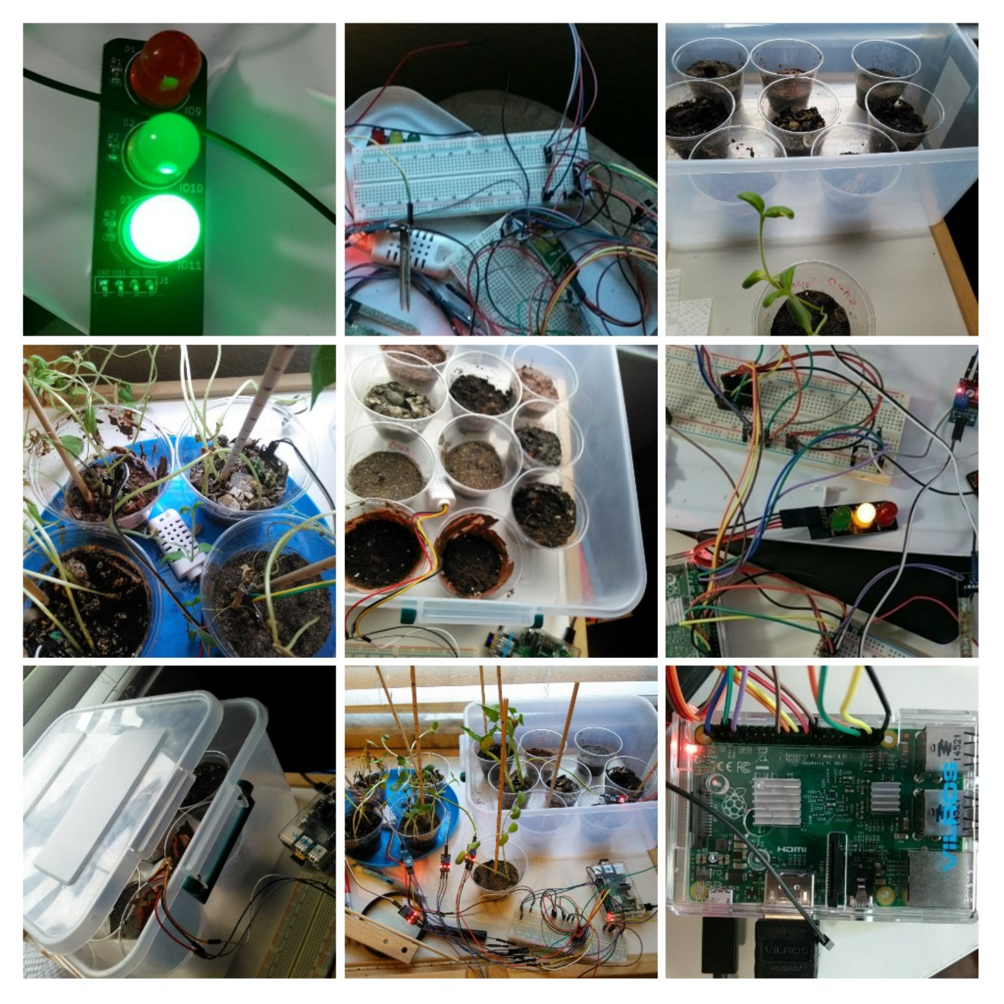


Alerts will trigger LEDs
Using SDK

And When We Put Everything Together

Science Project

- First… Isabella received A+ as final grade (applause…)
- LEDs light up whether the plants are in optimal (Green) or bad conditions (Red).
  - Need water
  - Has too much water
  - Temperature
  - Too humid could indicate plants cannot breath
- Splunk visualizations allows real time analytics



The Sourcetypes

Science Project

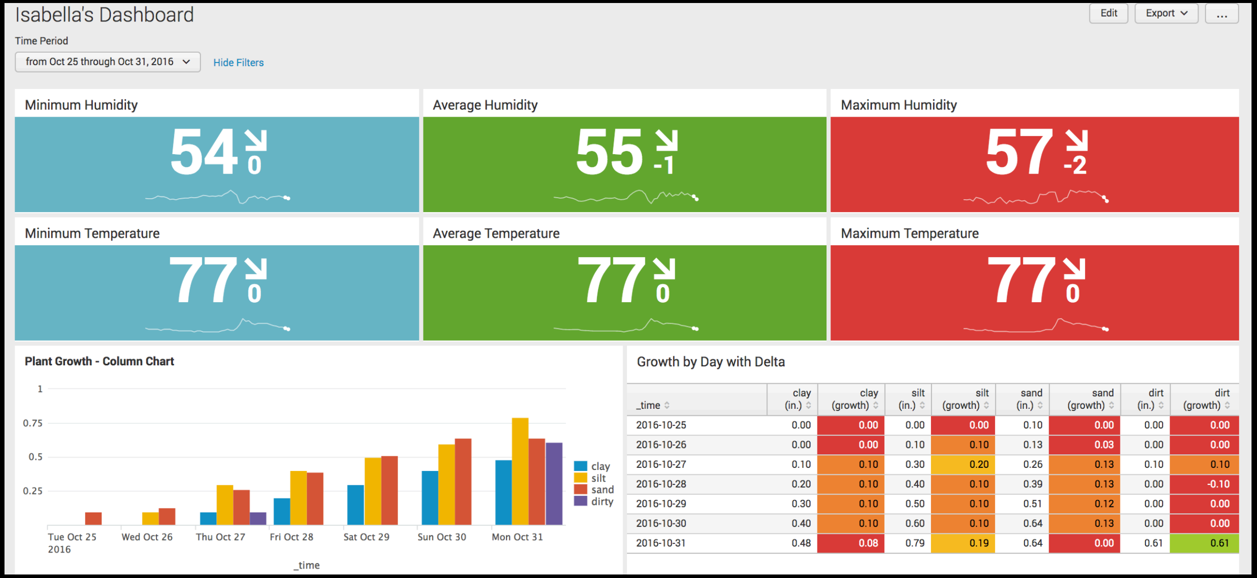
- Greenhouse – Temperature and Humidity
  - Near real-time collection using sensors and Splunk UF
- Growth – Daily plant measurements (in inches)
  - Isabella measured daily and fed results into Splunk via dashboard form input
- Soil – Moisture for each of the soil types
  - Near real-time collection using sensors and Splunk UF

- Data (output) was written w/ normalized timestamps, line breaks and key=value (or JSON) pairs to make indexing and field extraction automatic.

Isabella’s Dashboard

Science Project

- A series of dashboards and reports were built based on Isabella’s requirements
  - Temperature and Humidity: Show me the min, avg and max by day
  - How much did the plant grow for each soil type by day
  - Moisture levels by soil matter. Let’s make sure they are at the right level.



Isabella's Dashboard

Time Period

from Oct 25 through Oct 31, 2016

Minimum Humidity 54 ↘ 0

Average Humidity 55 ↘ -1

Maximum Humidity 57 ↘ -2

Minimum Temperature 77 ↘ 0

Average Temperature 77 ↘ 0

Maximum Temperature 77 ↘ 0

Plant Growth - Column Chart

Growth by Day with Delta

_time	clay (in.)	clay (growth)	silt (in.)	silt (growth)	sand (in.)	sand (growth)	dirt (in.)	dirt (growth)
2016-10-25	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00
2016-10-26	0.00	0.00	0.10	0.10	0.13	0.03	0.00	0.00
2016-10-27	0.10	0.10	0.30	0.20	0.26	0.13	0.10	0.10
2016-10-28	0.20	0.10	0.40	0.10	0.39	0.13	0.00	-0.10
2016-10-29	0.30	0.10	0.50	0.10	0.51	0.12	0.00	0.00
2016-10-30	0.40	0.10	0.60	0.10	0.64	0.13	0.00	0.00
2016-10-31	0.48	0.08	0.79	0.19	0.64	0.00	0.61	0.61

Alerting w/ the Python SDK

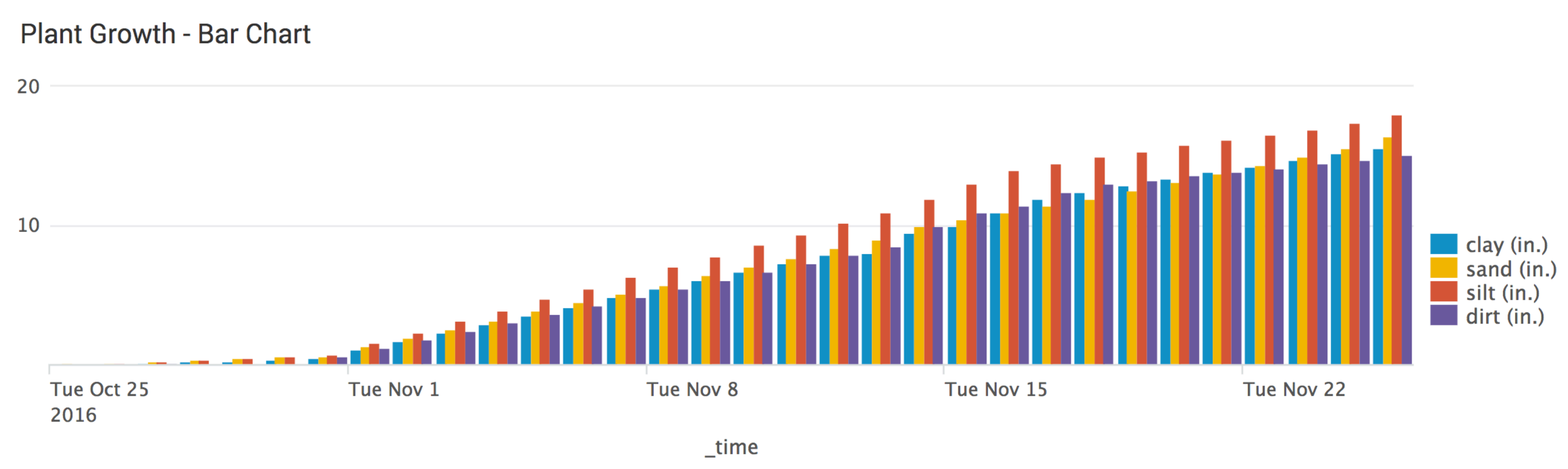
Science Project

- The Splunk SDK for Python was installed on the Raspberry Pi device.
- Every 30 seconds, a script would:
  - Search moisture levels and trigger LED lights.
    - >1000 (Red) = Too Dry!
    - Between 800 and 1000 (Yellow)
    - Between 600 and 800 (Green)
    - Between 100 and 500 (Yellow)
    - <100 (Red) = Too Wet!
- The Python SDK package includes sample scripts (eg. search.py) that helped us get up and running quickly.

Measure Growth by Day

Science Project

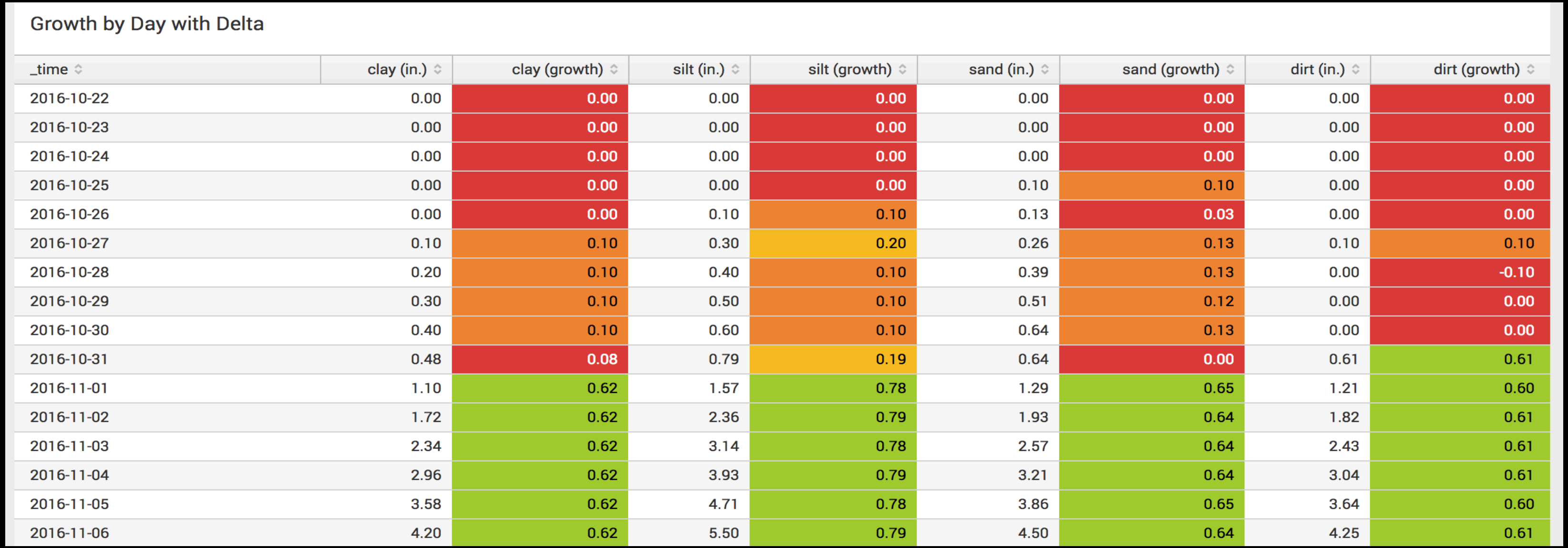
- Use the **timechart** command to visually measure growth by day.
  - **Sample search**: … | timechart max(clay) as “clay (in.)”



Measure Growth by Day

Science Project

- Use the **delta** command to compute the difference in growth by day.
  - Powerful splunk command that computes the difference, **in search order**, between the field value for the event and the field value for the previous event.
  - **Sample search**: … | delta sand as “sand (growth)”



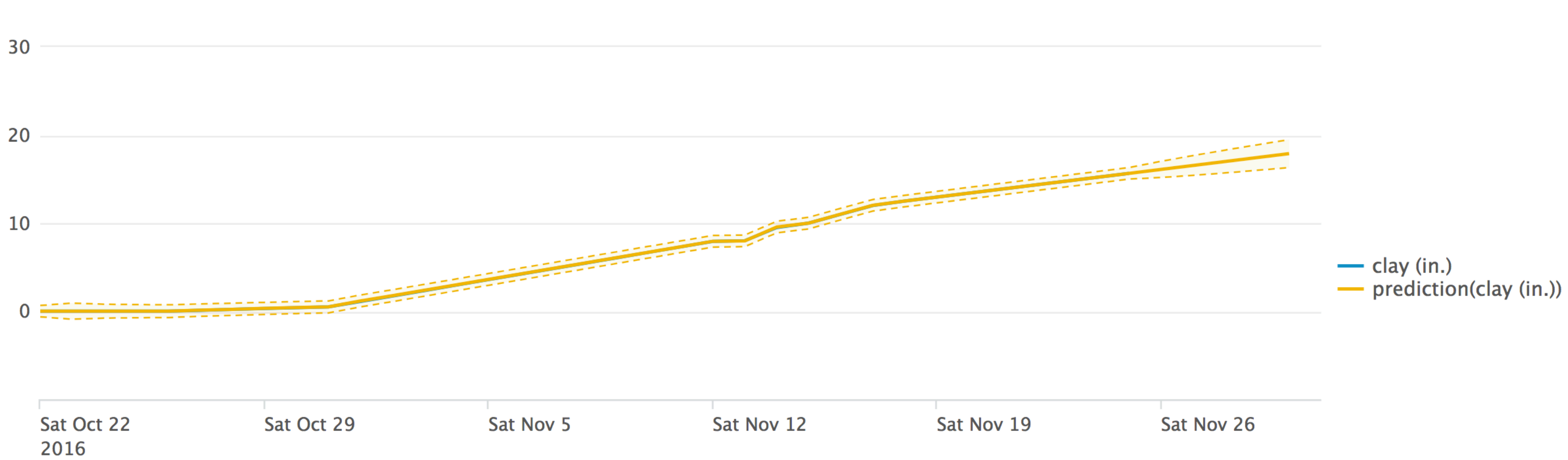
Growth by Day with Delta

_time	clay (in.)	clay (growth)	silt (in.)	silt (growth)	sand (in.)	sand (growth)	dirt (in.)	dirt (growth)
2016-10-22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2016-10-23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2016-10-24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2016-10-25	0.00	0.00	0.00	0.00	0.10	0.10	0.00	0.00
2016-10-26	0.00	0.00	0.10	0.10	0.13	0.03	0.00	0.00
2016-10-27	0.10	0.10	0.30	0.20	0.26	0.13	0.10	0.10
2016-10-28	0.20	0.10	0.40	0.10	0.39	0.13	0.00	-0.10
2016-10-29	0.30	0.10	0.50	0.10	0.51	0.12	0.00	0.00
2016-10-30	0.40	0.10	0.60	0.10	0.64	0.13	0.00	0.00
2016-10-31	0.48	0.08	0.79	0.19	0.64	0.00	0.61	0.61
2016-11-01	1.10	0.62	1.57	0.78	1.29	0.65	1.21	0.60
2016-11-02	1.72	0.62	2.36	0.79	1.93	0.64	1.82	0.61
2016-11-03	2.34	0.62	3.14	0.78	2.57	0.64	2.43	0.61
2016-11-04	2.96	0.62	3.93	0.79	3.21	0.64	3.04	0.61
2016-11-05	3.58	0.62	4.71	0.78	3.86	0.65	3.64	0.60
2016-11-06	4.20	0.62	5.50	0.79	4.50	0.64	4.25	0.61

Predict Future Growth by Day

Science Project

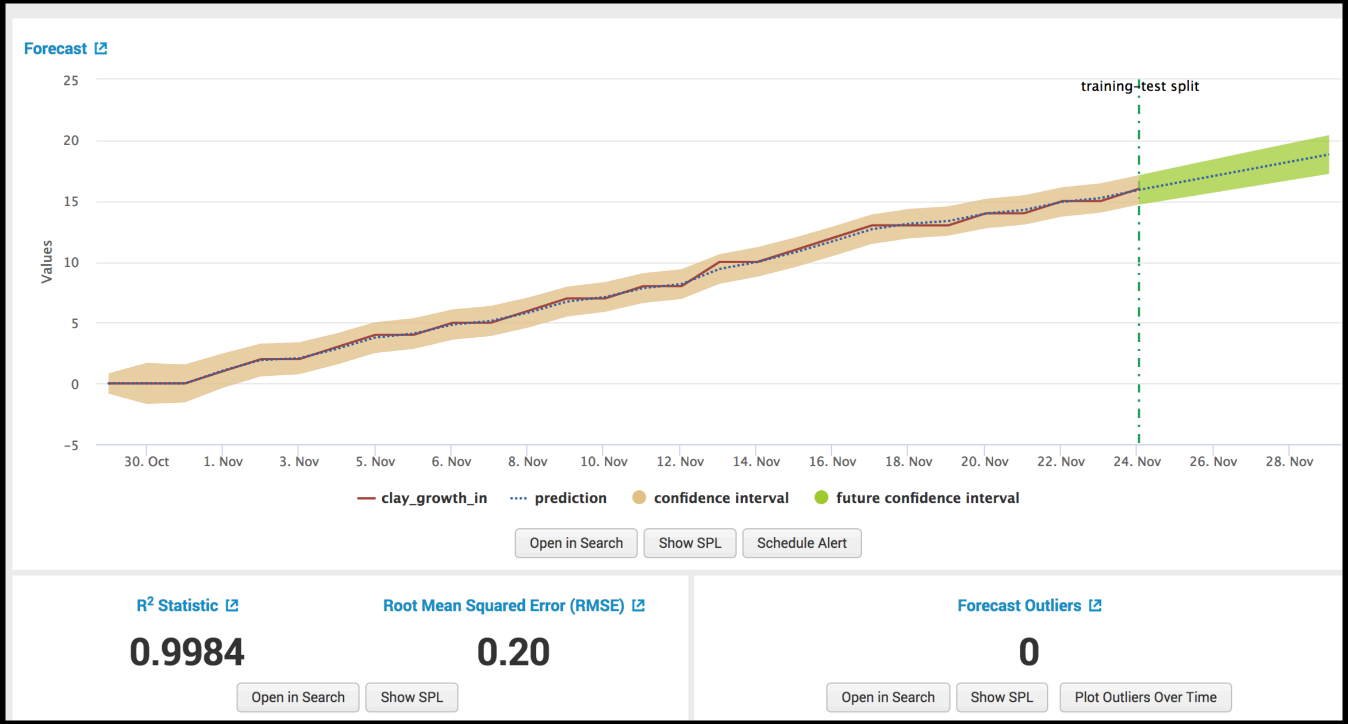
- Use the **predict** search command to predict future growth.
  - **Sample search**: … | timechart max(clay) as “clay (in.)” **| predict “clay (in.)”**



Predict Using The Splunk MLT

Science Project

- Use the Forecast assistant in the Splunk Machine Learning Toolkit
  - Prettier visualization!



R^2 Statistic: 0.9984

Root Mean Squared Error (RMSE): 0.20

Forecast Outliers: 0

Thank You

Don't forget to rate this session in the .conf2017 mobile app

splunk> .conf2017