@MelAkwule

GitHub makwule08

WomSA

Big Data Search - The ELK Trifecta

Demo: Searching Your Spend

Melanie Akwule

Agenda

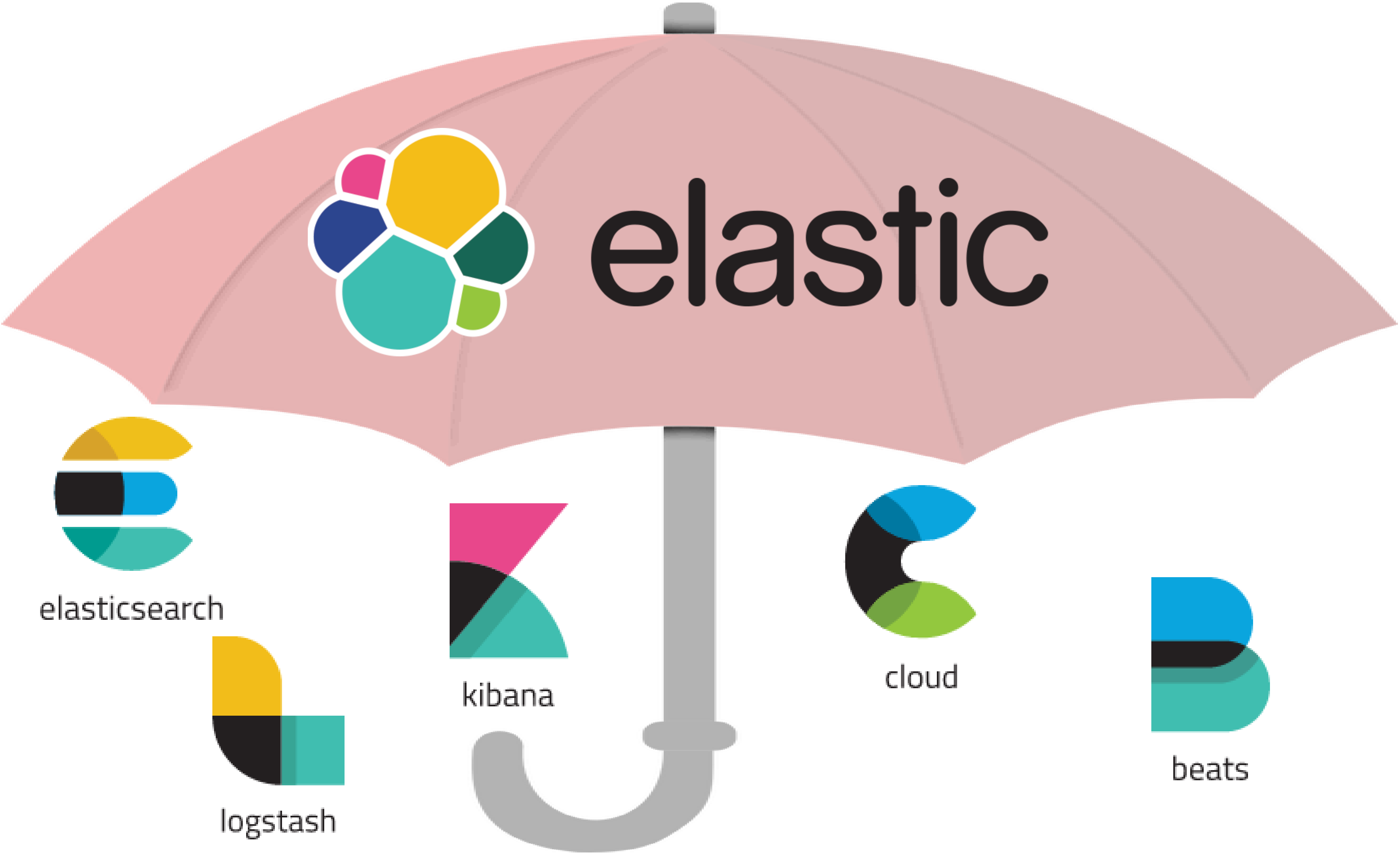
Elastic Overview:
Elastic and its
open source tools

Demo Summary:
Searching your spend

Getting Started:
The ELK trifecta

Elastic Overview

elastic and its open source tools



Elastic Overview

The ELK Trifecta

kibana

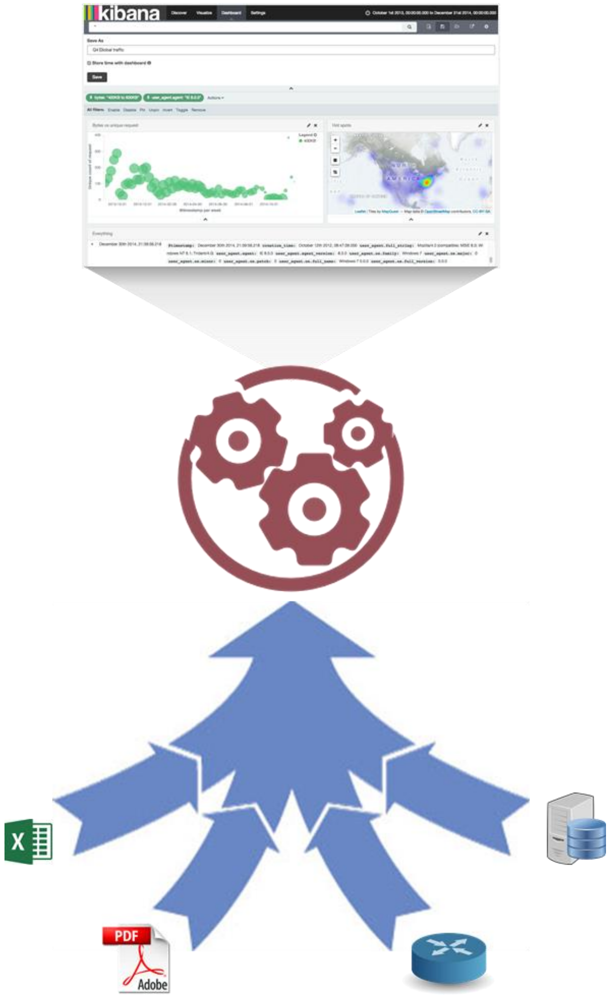
Visualization platform

elasticsearch

Search and analytics engine

logstash

Data collection, enrichment, and transportation pipeline



Demo Summary

searching your spend

Today's Demo

- Builds out a dashboard that highlights insightful metrics
- Walks you through how to build a visualization

kibana

Visualization platform

- Creates an elasticsearch index called "womsa" to analyze each record
- Maintains out of the box configuration

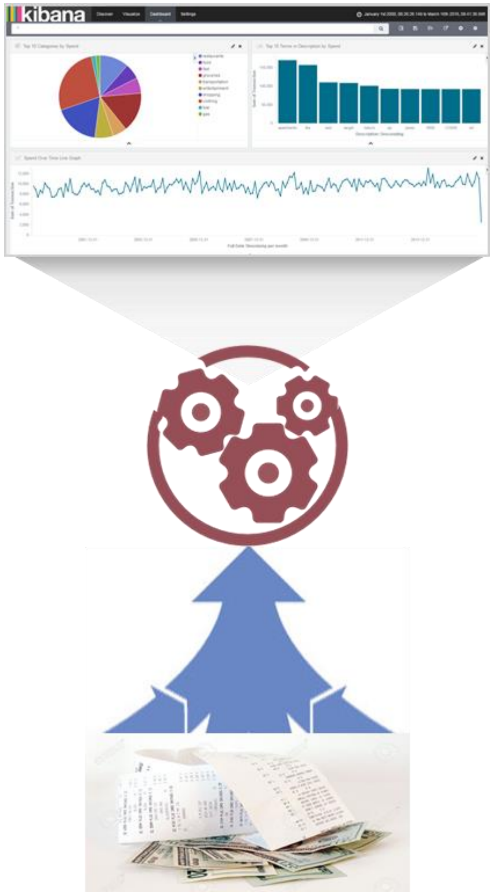
elasticsearch

Search and analytics engine

- Ingests a ~28K record file of mock financial data
- Key Parameters: Full Date Timestamp, Transaction, Description, Account Name, & Category

logstash

Data collection, enrichment, and transportation pipeline



Getting Started

trying it at home

What you'll find on GitHub…

@MelAkwule

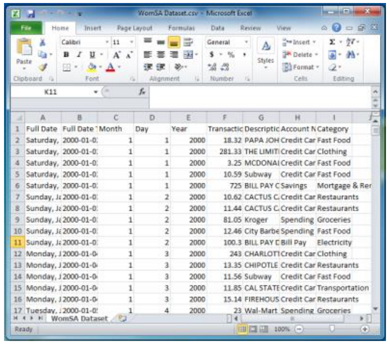
GitHub makwule08

Big Data Search - The ELK Trifecta

Demo: Searching Your Spend

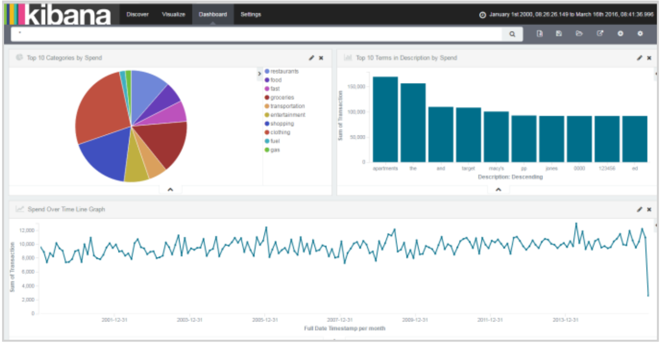
Melanie Akwule – Product Leader, Data & Analytics

This Presentation

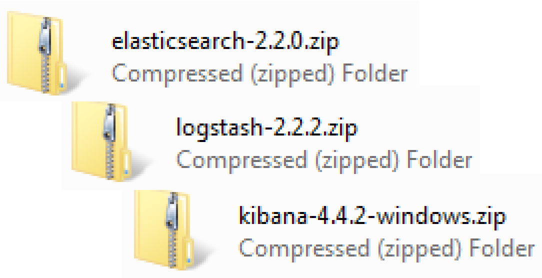


	A	B	C	D	E	F	G	H	I
1	Full Date	Full Date T	Month	Day	Year	Transactic	Descriptic	Account N	Category
2	Saturday,	2000-01-0:	1	1	2000	18.32	PAPA JOH	Credit Car	Fast Food
3	Saturday,	2000-01-0:	1	1	2000	281.33	THE LIMITI	Credit Car	Clothing
4	Saturday,	2000-01-0:	1	1	2000	3.25	MCDONAI	Credit Car	Fast Food
5	Saturday,	2000-01-0:	1	1	2000	10.59	Subway	Credit Car	Fast Food
6	Saturday,	2000-01-0:	1	1	2000	725	BILL PAY C	Savings	Mortgage & Rer
7	Sunday, Ja	2000-01-0:	1	2	2000	10.62	CACTUS C.	Credit Car	Restaurants
8	Sunday, Ja	2000-01-0:	1	2	2000	11.44	CACTUS C.	Credit Car	Restaurants
9	Sunday, Ja	2000-01-0:	1	2	2000	81.05	Kroger	Spending	Groceries
10	Sunday, Ja	2000-01-0:	1	2	2000	12.46	City Barbe	Spending	Fast Food
11	Sunday, Ja	2000-01-0:	1	2	2000	100.3	BILL PAY C	Bill Pay	Electricity
12	Monday, J	2000-01-0	1	3	2000	243	CHARLOTT	Credit Car	Clothing
13	Monday, J	2000-01-0	1	3	2000	13.35	CHIPOTLE	Credit Car	Restaurants
14	Monday, J	2000-01-0	1	3	2000	11.56	Subway	Credit Car	Fast Food
15	Monday, J	2000-01-0	1	3	2000	11.85	CAL STATE	Credit Car	Transportation
16	Monday, J	2000-01-0	1	3	2000	15.14	FIREHOUS	Credit Car	Restaurants
17	Tuesday, J	2000-01-0:	1	4	2000	23	Wal-Mart	Spending	Groceries

Dataset used to build the Dashboard



The dashboard from Kibana



elasticsearch-2.2.0.zip
Compressed (zipped) Folder

logstash-2.2.2.zip
Compressed (zipped) Folder

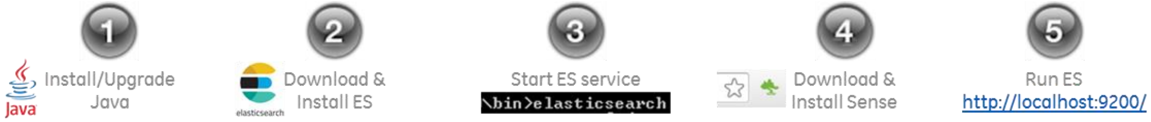
kibana-4.4.2-windows.zip
Compressed (zipped) Folder

Installation Files Used

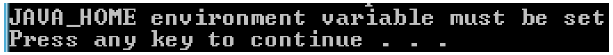
womsa.conf
CONF File
664 bytes

Logstash configuration file

Getting Started
elasticsearch



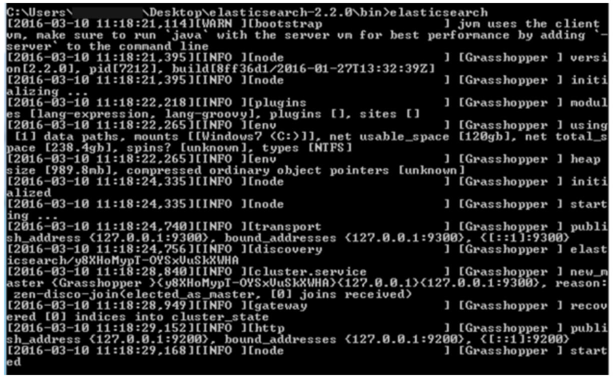
1
- Java Install
- Setting JAVA_HOME Variable Documentation



```
JAVA_HOME environment variable must be set
Press any key to continue . . .
```

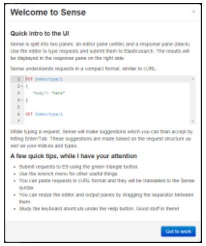
2
- Elasticsearch 2.2.0 Download
- Getting Started Documentation

3
- Start ES: \elasticsearch-2.2.0\bin>elasticsearch
- Helpful CMD Documentation



```
C:\Users\[redacted]\Desktop\elasticsearch-2.2.0\bin>elasticsearch
[2016-03-10 11:18:21,114][WARN ][bootstrap                ] jvm uses the client
vm, make sure to run `java` with the server vm for best performance by adding `-
server` to the command line
[2016-03-10 11:18:21,395][INFO ][node                     ] [Grasshopper ] versi
on[2.2.0], pid[7212], build[8ff36d1/2016-01-27T13:32:39Z]
[2016-03-10 11:18:21,395][INFO ][node                     ] [Grasshopper ] initi
alizing ...
[2016-03-10 11:18:22,218][INFO ][plugins                  ] [Grasshopper ] modul
es [lang-expression, lang-groovy], plugins [], sites []
[2016-03-10 11:18:22,265][INFO ][env                      ] [Grasshopper ] using
 [1] data paths, mounts [[Windows7 (C:)]], net usable_space [120gb], net total_s
pace [238.4gb], spins? [unknown], types [NTFS]
[2016-03-10 11:18:22,265][INFO ][env                      ] [Grasshopper ] heap
size [989.8mb], compressed ordinary object pointers [unknown]
[2016-03-10 11:18:24,335][INFO ][node                     ] [Grasshopper ] initi
alized
[2016-03-10 11:18:24,335][INFO ][node                     ] [Grasshopper ] start
ing ...
[2016-03-10 11:18:24,740][INFO ][transport                ] [Grasshopper ] publi
sh_address {127.0.0.1:9300}, bound_addresses {127.0.0.1:9300}, {[::1]:9300}
[2016-03-10 11:18:24,756][INFO ][discovery                ] [Grasshopper ] elast
icsearch/y8XHoMypT-OYSxVuSkXWHA
[2016-03-10 11:18:28,840][INFO ][cluster.service          ] [Grasshopper ] new_m
aster {Grasshopper }{y8XHoMypT-OYSxVuSkXWHA}{127.0.0.1}{127.0.0.1:9300}, reason:
 zen-disco-join(elected_as_master, [0] joins received)
[2016-03-10 11:18:28,949][INFO ][gateway                  ] [Grasshopper ] recov
ered [0] indices into cluster_state
[2016-03-10 11:18:29,152][INFO ][http                     ] [Grasshopper ] publi
sh_address {127.0.0.1:9200}, bound_addresses {127.0.0.1:9200}, {[::1]:9200}
[2016-03-10 11:18:29,168][INFO ][node                     ] [Grasshopper ] start
ed
```

4
- Sense Chrome Plugin Download
- Query DSL Documentation



Welcome to Sense

Quick intro to the UI

Sense is split into two panes: an editor pane (white) and a response pane (black). Use the editor to type requests and submit them to Elasticsearch. The results will be displayed in the response pane on the right side.

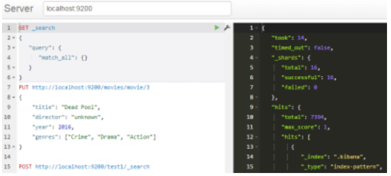
Sense understands requests in a compact format, similar to cURL:

```
PUT index/type/1
{
   "body": "here"
}

GET index/type/1
```

While typing a request, Sense will make suggestions which you can than accept by hitting Enter/Tab. These suggestions are made based on the request structure as well as your indices and types.

A few quick tips, while I have your attention

- Submit requests to ES using the green triangle button.
- Use the wrench menu for other useful things.
- You can paste requests in cURL format and they will be translated to the Sense syntax.
- You can resize the editor and output panes by dragging the separator between them.
- Study the keyboard shortcuts under the Help button. Good stuff in there!

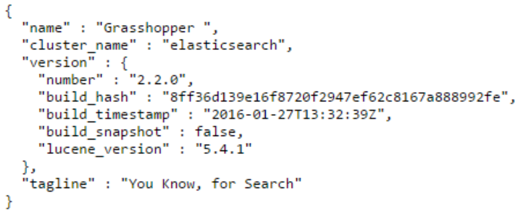


Server: localhost:9200

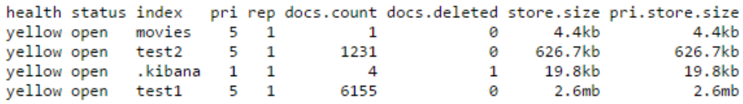
```
GET _search
{
   "query": {
      "match_all": {}
   }
}
PUT http://localhost:9200/movies/movie/3
{
    "title": "Dead Pool",
    "director": "unknown",
    "year": 2016,
    "genres": ["Crime", "Drama", "Action"]
}

POST http://localhost:9200/test1/_search
```

```
{
   "took": 14,
   "timed_out": false,
   "_shards": {
      "total": 16,
      "successful": 16,
      "failed": 0
   },
   "hits": {
      "total": 7394,
      "max_score": 1,
      "hits": [
         {
            "_index": ".kibana",
            "_type": "index-pattern",
```

5
- Run ES: http://localhost:9200/
- List of ES Indicies: http://localhost:9200/_cat/indices?v
- Looking inside an index: http://localhost:9200/test2/_search?pretty



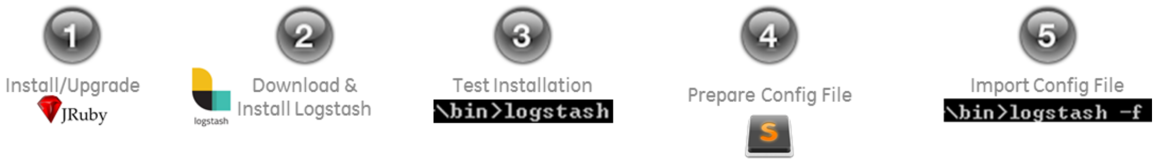
```
{
  "name" : "Grasshopper ",
  "cluster_name" : "elasticsearch",
  "version" : {
    "number" : "2.2.0",
    "build_hash" : "8ff36d139e16f8720f2947ef62c8167a888992fe",
    "build_timestamp" : "2016-01-27T13:32:39Z",
    "build_snapshot" : false,
    "lucene_version" : "5.4.1"
  },
  "tagline" : "You Know, for Search"
}
```



```
health status index   pri rep docs.count docs.deleted store.size pri.store.size
yellow open   movies    5   1          1            0      4.4kb          4.4kb
yellow open   test2     5   1       1231            0    626.7kb        626.7kb
yellow open   .kibana   1   1          4            1     19.8kb         19.8kb
yellow open   test1     5   1       6155            0      2.6mb          2.6mb
```

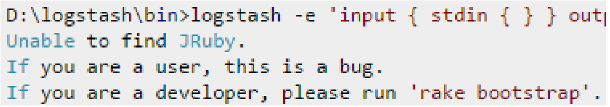
Getting Started

logstash


1 Install/Upgrade JRuby
2 Download & Install Logstash
3 Test Installation `\bin>logstash`
4 Prepare Config File
5 Import Config File `\bin>logstash -f`

1

- JRuby Install

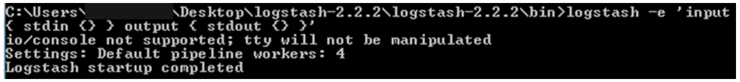

```
D:\logstash\bin>logstash -e 'input { stdin { } } out
Unable to find JRuby.
If you are a user, this is a bug.
If you are a developer, please run 'rake bootstrap'.
```

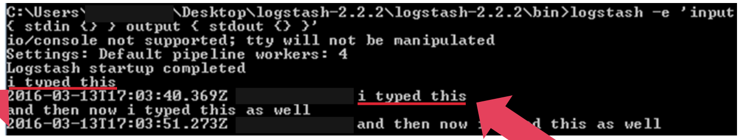
2

- Logstash 2.2.2 Download
- Getting Started Documentation

3

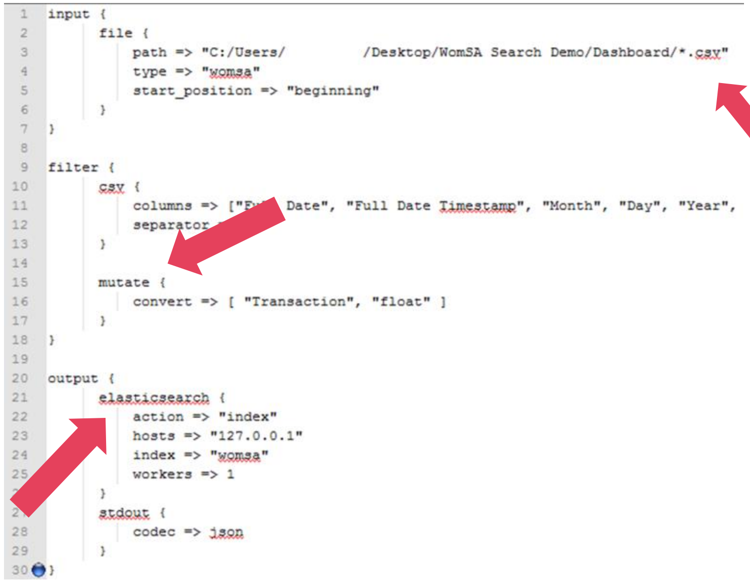
- Enter Standard I/O Command: logstash -e 'input { stdin { } } output { stdout {} }'
- Should be able to type and have it echoed back


```
C:\Users\[redacted]\Desktop\logstash-2.2.2\logstash-2.2.2\bin>logstash -e 'input
{ stdin {} } output { stdout {} }'
io/console not supported; tty will not be manipulated
Settings: Default pipeline workers: 4
Logstash startup completed
```


```
C:\Users\[redacted]\Desktop\logstash-2.2.2\logstash-2.2.2\bin>logstash -e 'input
{ stdin {} } output { stdout {} }'
io/console not supported; tty will not be manipulated
Settings: Default pipeline workers: 4
Logstash startup completed
i typed this
2016-03-13T17:03:40.369Z [redacted] i typed this
and then now i typed this as well
2016-03-13T17:03:51.273Z [redacted] and then now i     d this as well
```

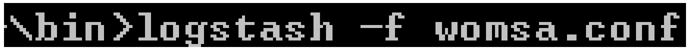
4

- Preparing a Configuration File Documentation
- Configuring for CSV Documentation
- List of Input Plugins
- List of Output Plugins


```
input {
        file {
                path => "C:/Users/          /Desktop/WomSA Search Demo/Dashboard/*.csv"
                type => "womsa"
                start_position => "beginning"
        }
}

filter {
        csv {
                columns => ["F    Date", "Full Date Timestamp", "Month", "Day", "Year",
                separator
        }

        mutate {
                convert => [ "Transaction", "float" ]
        }
}

output {
        elasticsearch {
                action => "index"
                hosts => "127.0.0.1"
                index => "womsa"
                workers => 1
        }
        stdout {
                codec => json
        }
}
```

5

- Import File: logstash -f <filename>.conf
- Troubleshooting Field Format Documentation
- Troubleshooting Data Ingestion Documentation


```
\bin>logstash -f womsa.conf
```

Getting Started
kibana

1 Download & Install Kibana

2 Start ES service
`\bin>elasticsearch`

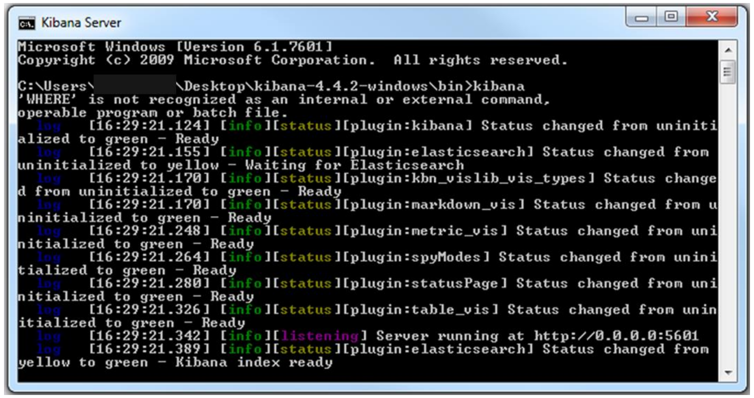
3 Start kibana service
`\bin>kibana`

4 Configure Kibana
http://localhost:5601/

1
- Kibana 4.4.2 Download
- Getting Started Documentation

2
- Start ES: \elasticsearch-2.2.0\bin>elasticsearch

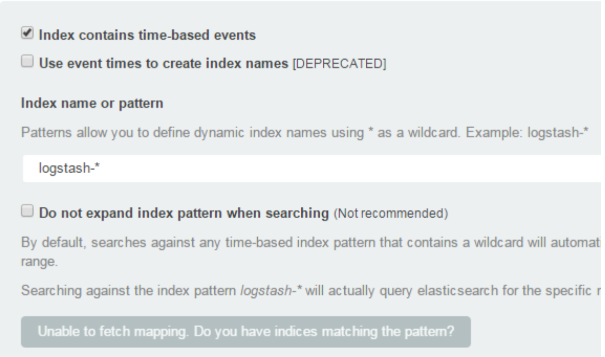
3
- Start Kibana: \kibana-4.4.2\bin>kibana



Kibana Server

```
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation.  All rights reserved.

C:\Users\[redacted]\Desktop\kibana-4.4.2-windows\bin>kibana
'WHERE' is not recognized as an internal or external command,
operable program or batch file.
  log   [16:29:21.124] [info][status][plugin:kibana] Status changed from uniniti
alized to green - Ready
  log   [16:29:21.155] [info][status][plugin:elasticsearch] Status changed from
uninitialized to yellow - Waiting for Elasticsearch
  log   [16:29:21.170] [info][status][plugin:kbn_vislib_vis_types] Status change
d from uninitialized to green - Ready
  log   [16:29:21.170] [info][status][plugin:markdown_vis] Status changed from u
ninitialized to green - Ready
  log   [16:29:21.248] [info][status][plugin:metric_vis] Status changed from uni
nitialized to green - Ready
  log   [16:29:21.264] [info][status][plugin:spyModes] Status changed from unini
tialized to green - Ready
  log   [16:29:21.280] [info][status][plugin:statusPage] Status changed from uni
nitialized to green - Ready
  log   [16:29:21.326] [info][status][plugin:table_vis] Status changed from unin
itialized to green - Ready
  log   [16:29:21.342] [info][listening] Server running at http://0.0.0.0:5601
  log   [16:29:21.389] [info][status][plugin:elasticsearch] Status changed from
yellow to green - Kibana index ready
```

4
- Change the name of your index to match ES
- Map time field – "yyyy-mm-ddThh:mm:ss"



☑ **Index contains time-based events**

□ **Use event times to create index names** [DEPRECATED]

Index name or pattern

Patterns allow you to define dynamic index names using * as a wildcard. Example: logstash-*

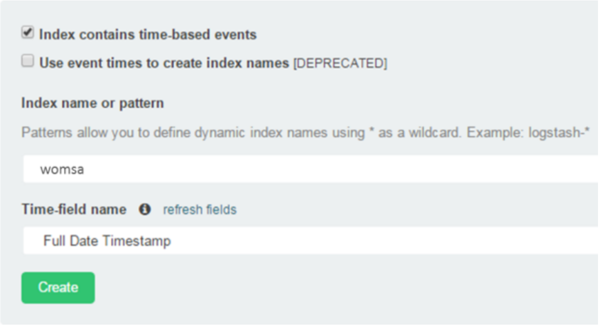
logstash-*

□ **Do not expand index pattern when searching** (Not recommended)

By default, searches against any time-based index pattern that contains a wildcard will automat
range.

Searching against the index pattern *logstash-** will actually query elasticsearch for the specific r

Unable to fetch mapping. Do you have indices matching the pattern?



☑ **Index contains time-based events**

□ **Use event times to create index names** [DEPRECATED]

Index name or pattern

Patterns allow you to define dynamic index names using * as a wildcard. Example: logstash-*

womsa

Time-field name refresh fields

Full Date Timestamp

Let’s check it out..