Machine Learning Workshop I LABS

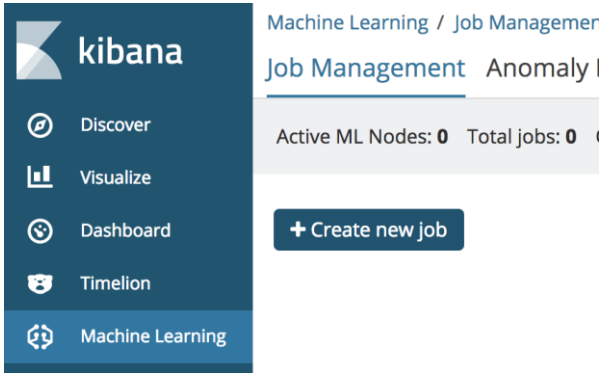
Elastic Stack | ML | Labs

Shawn Hooton | Solution Architect

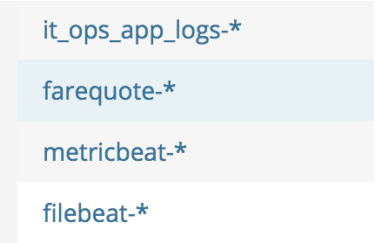
Lab 1: The Simplest job

Steps to Complete

1) In Machine Learning,
Create new job



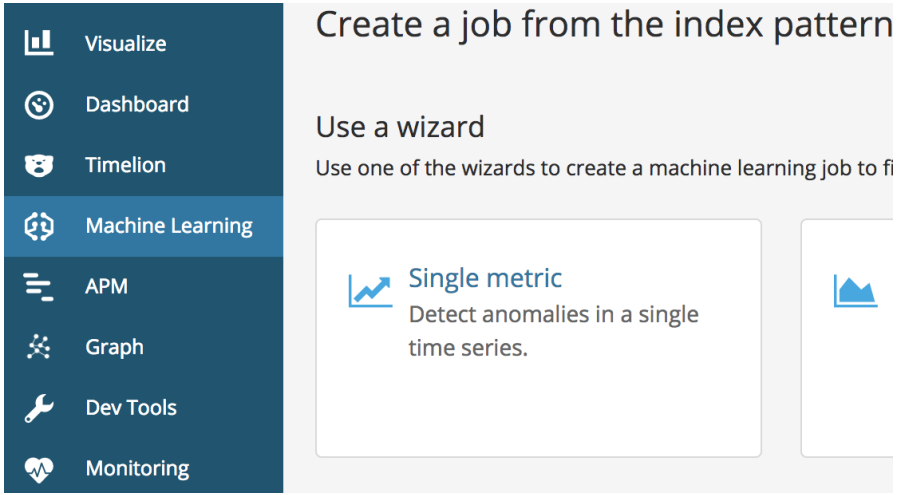
2) Choose the “farequote-*”
Index pattern



elastic

Steps to Complete

3) pick the *Single Metric* wizard



elastic

Steps to Complete

4) Aggregation: count

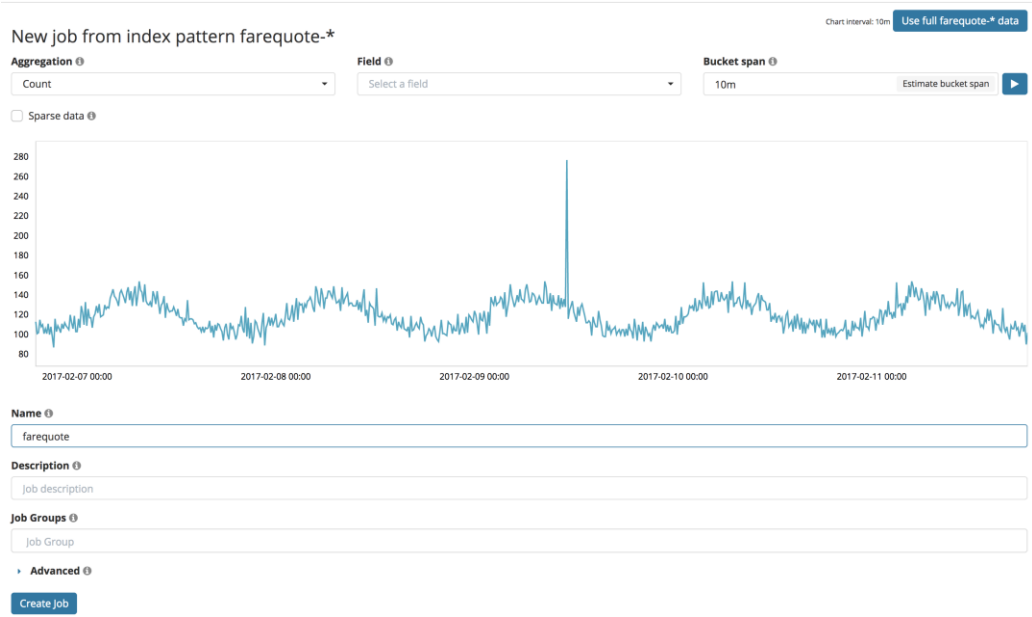
5) Field: *<leave blank>*

6) Bucket span: 10m

7) Click the *“use full farequote data”*

8) Name: *“farequote”*

9) click “Create Job”

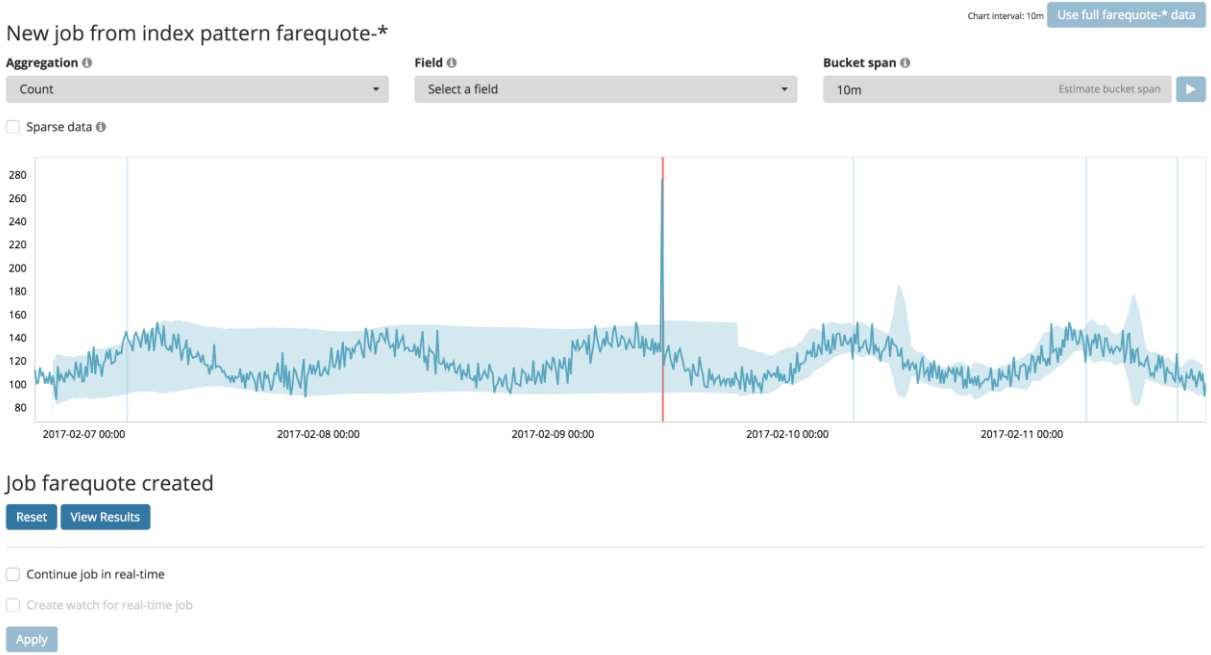


elastic

Steps to Complete

10) See animated learning

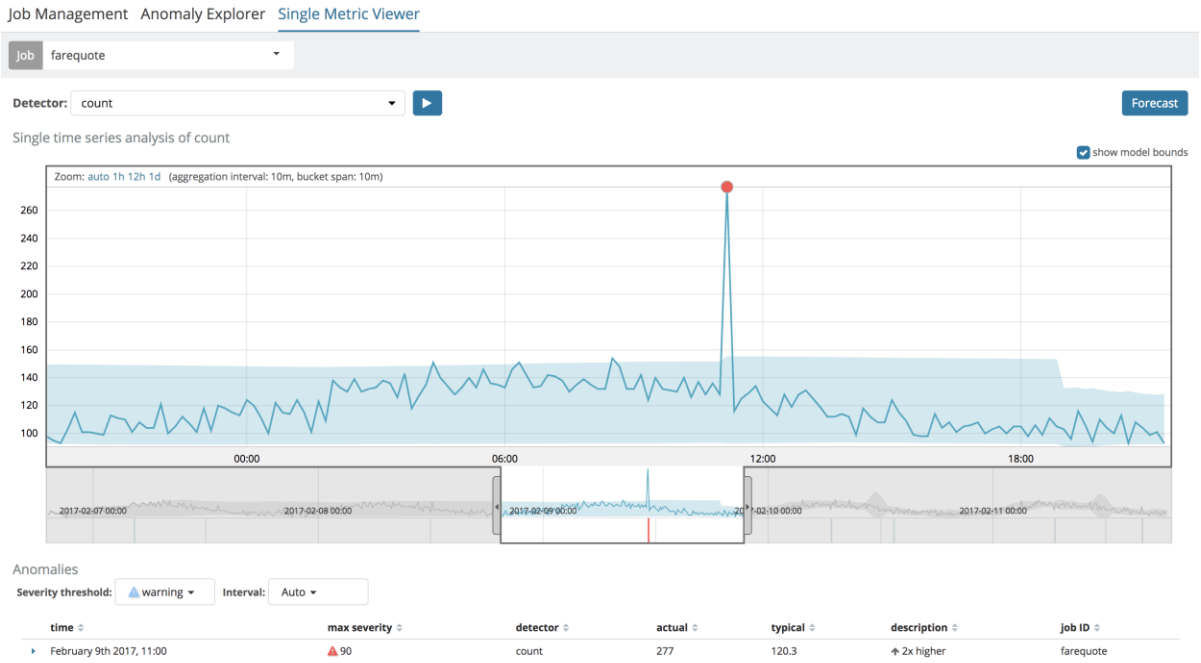
11) click *“View Results”*



elastic

Steps to Complete

12) Zoom in on anomaly



elastic

Lab 2: Advanced Jobs

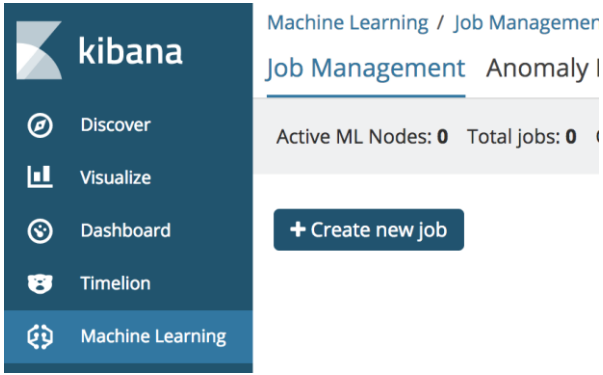
Steps to Complete

- Index: “farequote”
- ML job type: *“advanced”*
  - job name: “farequote_response”
  - bucket_span: *10m*
  - Detector:
    - function: *max*
    - field_name: *responsetime*
    - partition_field_name: *airline*
  - influencer: *airline*
- Run the job over the entire data set (data is not real-time)

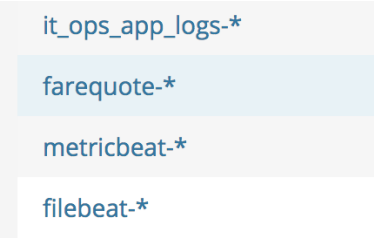
elastic

Steps to Complete

1) In Machine Learning,
Create new job



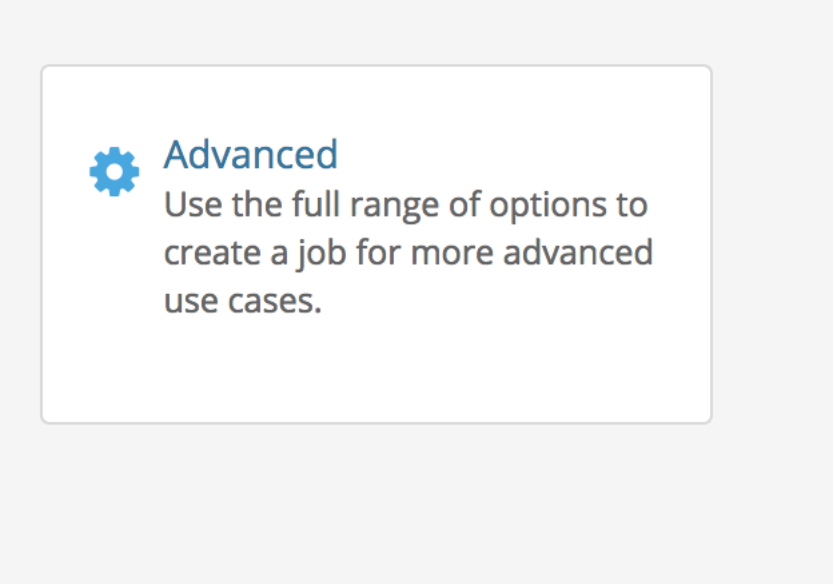
2) Choose the “farequote-*”
Index pattern



elastic

Steps to Complete

3) Choose Advanced Job Wizard

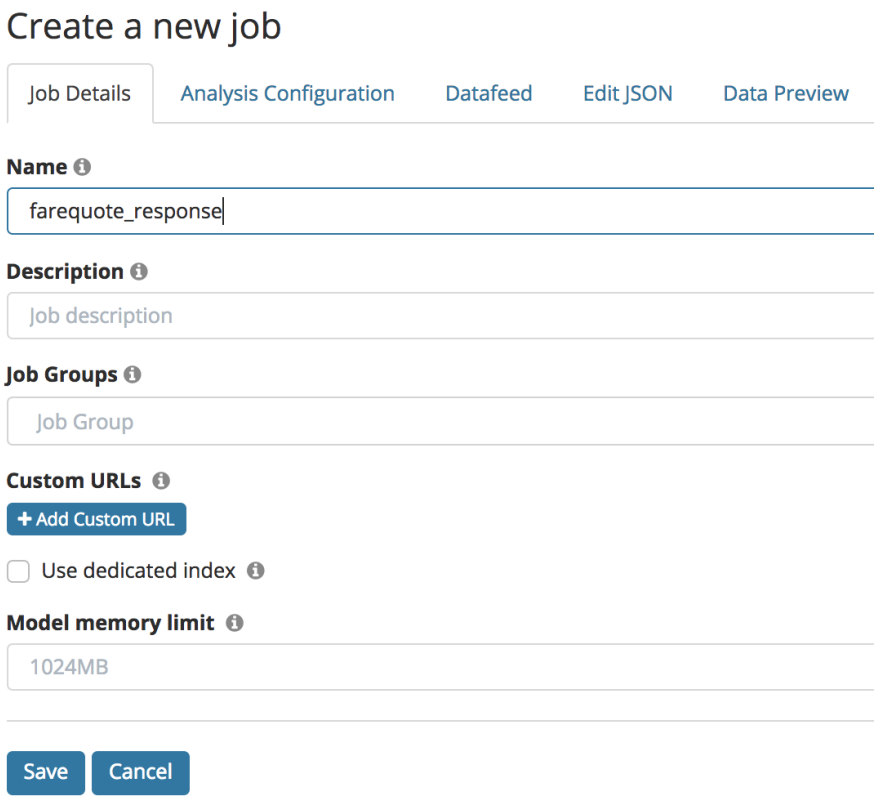


elastic

Steps to Complete

5) Name job

"farequote_response"



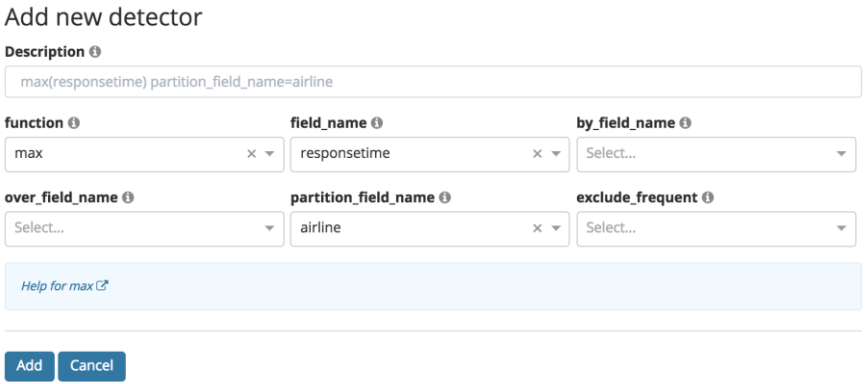
elastic

Steps to Complete

6) Set bucket_span= *10m*

7) Add a Detector:

function: *max*
field_name: *responsetime*
partition_field_name: *airline*



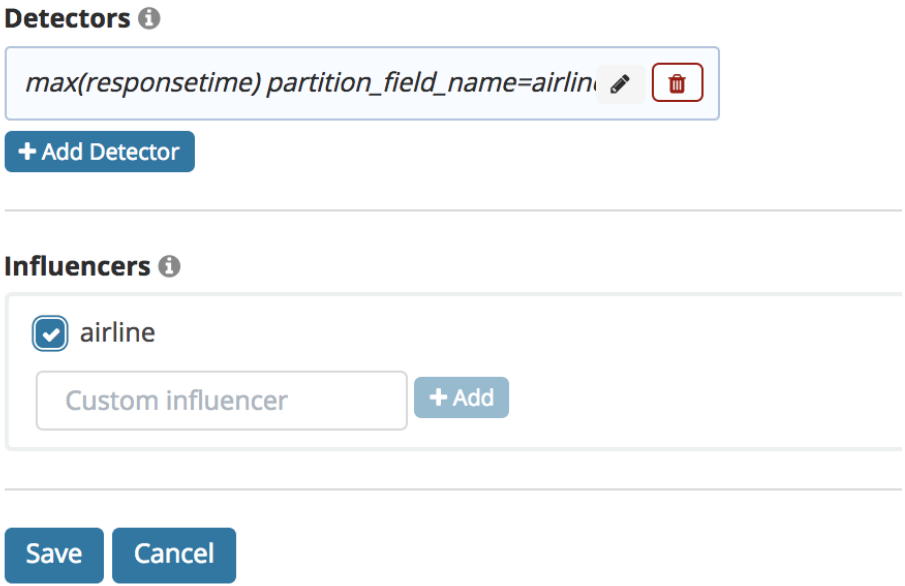
elastic

Steps to Complete

7) Select *"airline"* as Influencer

8) Save job

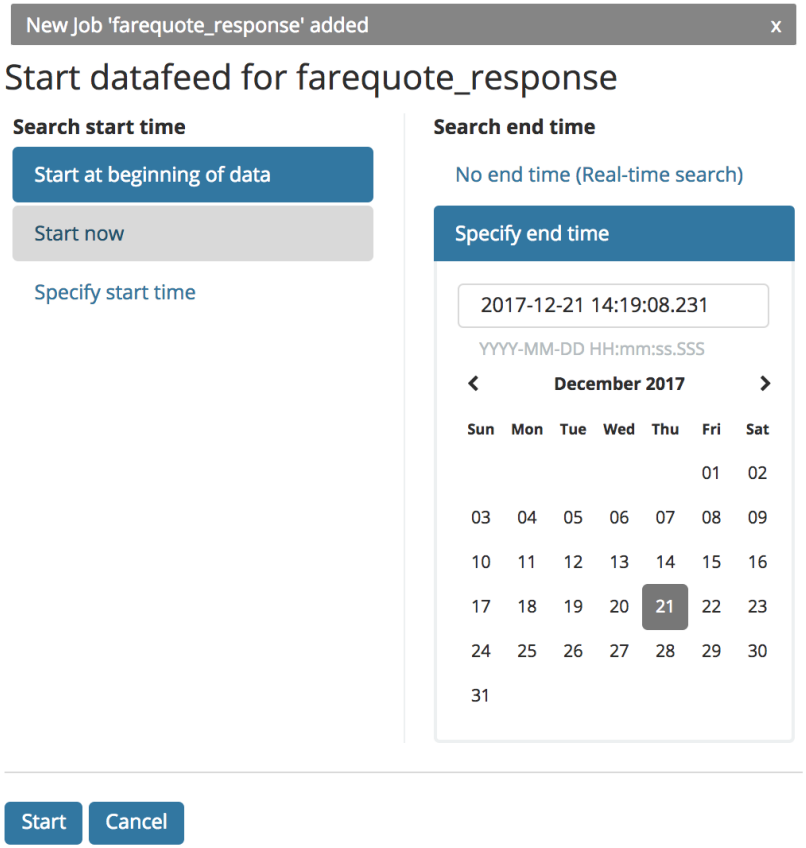
9) Start datafeed



elastic

Steps to Complete

8) Start at beginning of data
leave “now” as end time



elastic

Steps to Complete

9) wait until all 86274 events are processed

Job Management Anomaly Explorer Single Metric Viewer

Active ML Nodes: **1** Total jobs: **2** Open jobs: **0** Closed jobs: **1** Active datafeeds: **0**

New Job 'farequote_response' added

+ Create new job

Job filter

Job ID	Description	Processed records	Memory status	Job state	Datafeed state	Latest timestamp	Actions
farequote		719	ok	closed	stopped	2017-02-11 18:49:56	
farequote_response		86,274	ok	closing	stopped	2017-02-11 18:59:54	

Page Size 10

elastic

Elastic Stack Overview

Anomaly Detection Concepts

Concepts & Configuration

Interactive Logs Analysis

Exploring Results with Explorer View

Comparison of Job Type

Other kind of Analysis

Advanced Topics

elastic

Explorer View of a Job

Job Management | Anomaly Explorer | Single Metric Viewer

Active ML Nodes: **1** Total jobs: **2** Open jobs: **0** Closed jobs: **1** Active datafeeds: **0**

New Job 'farequote_response' added

+ Create new job

Job fil

Job ID	Description	Processed records	Memory status	Job state	Datafeed state	Latest timestamp	Actions
farequote		719	ok	closed	stopped	2017-02-11 18:49:56	
farequote_response		86,274	ok	closing	stopped	2017-02-11 18:59:54	

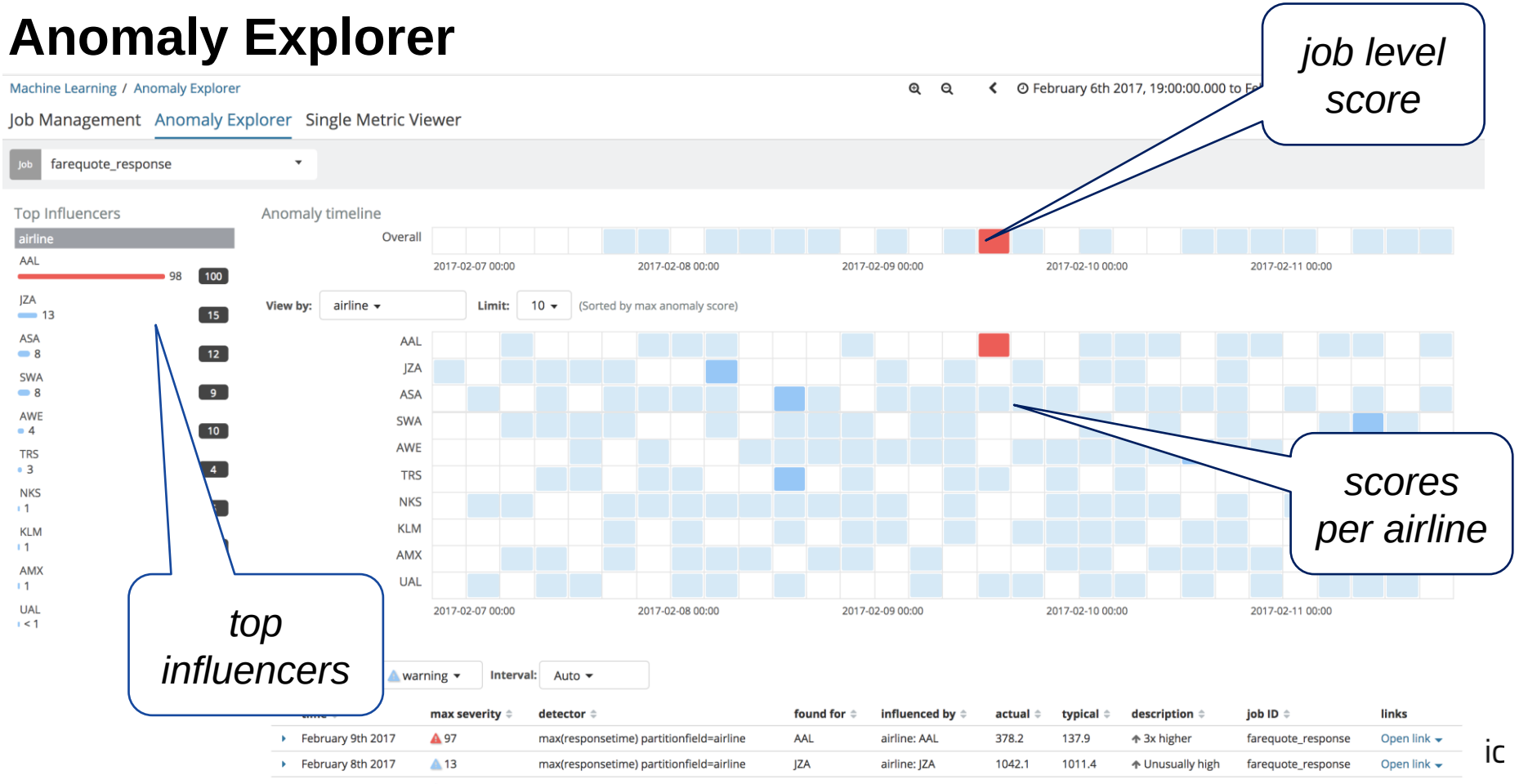
Page Size 10

Or here

click here

elastic

Anomaly Explorer



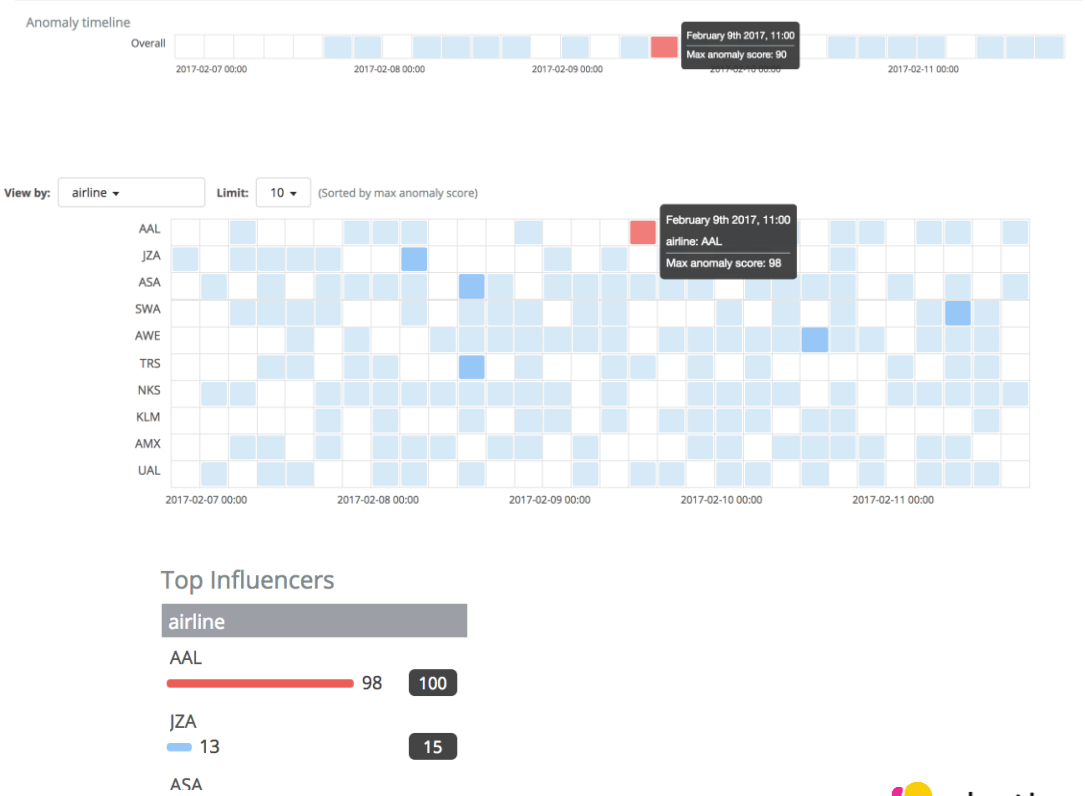
Concept: What is an Influencer?

- An Influencer is a field, selected at configuration time, that would be a logical **entity ”to blame”** if an anomaly were to exist
- Doesn’t have to be a field in the actual detector, but **fields used to split the data are often good candidates**
- Will get its own score based upon how influential that entity is on the anomaly

elastic

Scoring

- Overall Job score is 90
  - How unusual is that bucket, given all airlines?
- Detector score is 98
  - How unusual is the response time of airline=AAL?
- airline=AAL is the top influencer in this time range
  - 98 is the max anomaly score
  - 100 is the sum of anomaly scores in this time range



elastic

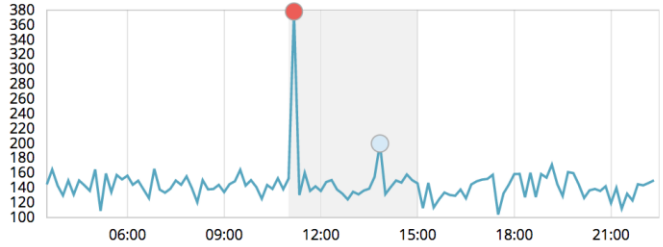
Anomaly Details

view of response time for AAL

Anomalies

Severity threshold: warning **Interval:** Auto **Show Charts**

max(responsetime) partitionfield=airline - airline AAL View



time	max severity	detector	found for	influence	actual	typical	description	job ID	links
February 9th 2017, 11:00	97	max(responsetime) partitionfield=airline	AAL	AAL	378.2	137.9	3x higher	farequote_res	

Description:
critical anomaly in max(responsetime) partitionfield=airline found for airline AAL

Details on highest severity anomaly:

airline:	AAL
time:	February 9th 2017, 11:10:00 to February 9th 2017, 11:20:00
function:	max
fieldName:	responsetime
actual:	378.2
typical:	137.9
job ID:	farequote_response
probability:	3.95474803120103e-22

Influenced by:

airline	AAL

actual vs. "typical"

raw probability

elastic

Lab 3: Multi-Metric Jobs

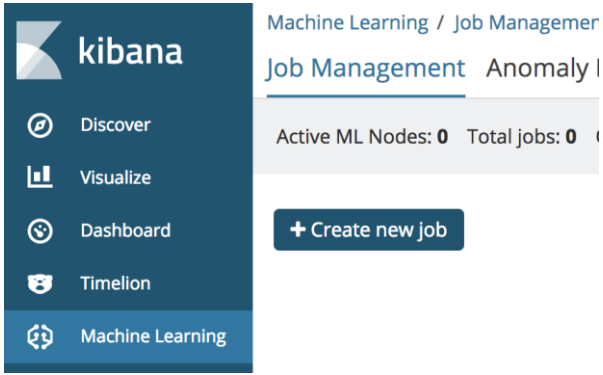
Steps to Complete

- Again, use index: *“farequote”:*
- Job type: *“multi-metric”*
  - Re-create the *“max(responsetime) per airline”* job
- Also add *“count per airline”* in the same job

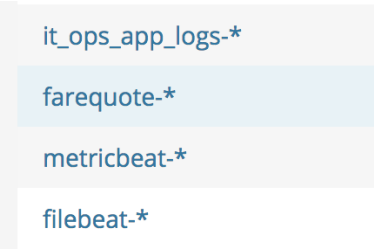
elastic

Steps to Complete

1) In Machine Learning,
Create new job



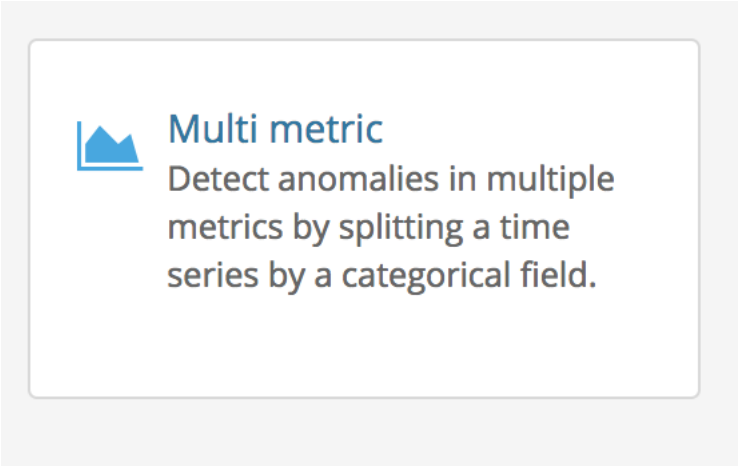
2) Choose the “farequote-*”
Index pattern



elastic

Steps to Complete

3) pick Multi-Metric
Job Wizard



elastic

Steps to Complete

4) Choose
   - *event rate, count*
   - *responsetime, max*

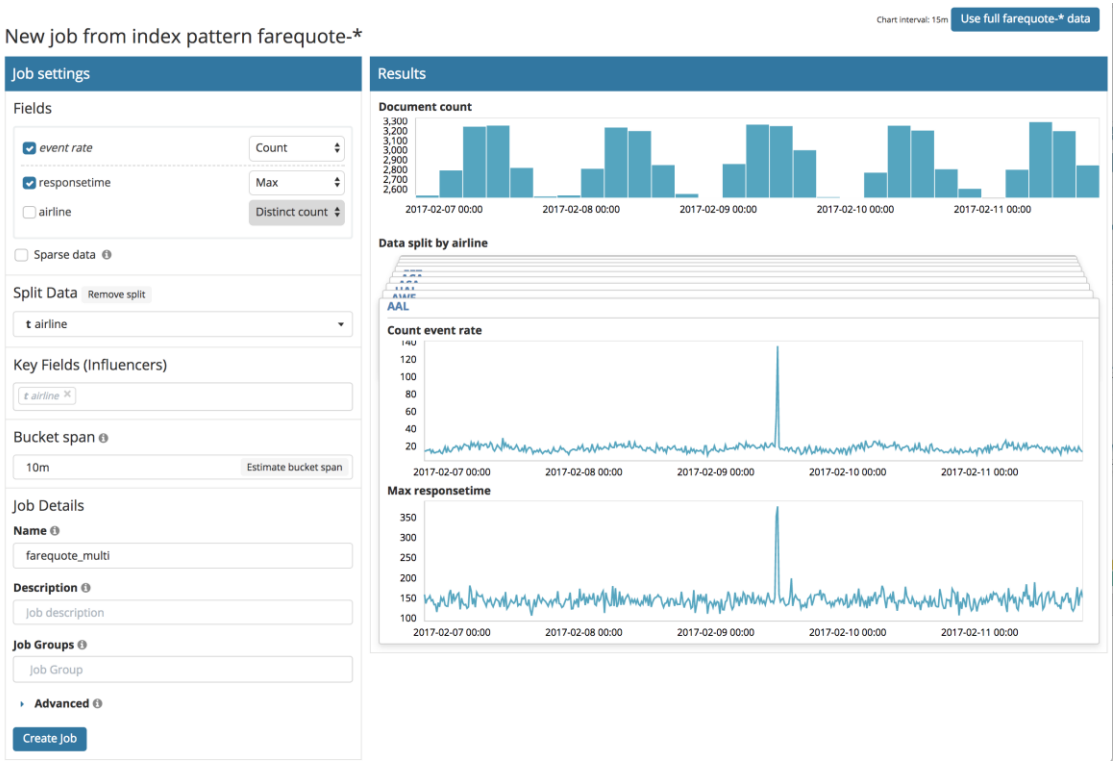
5) Bucket span: *10m*

6) Split Data: *airline*

7) Click *"use full farequote data"*

8) Name: *"farequote_multi"*

9) Click *"Create Job"*

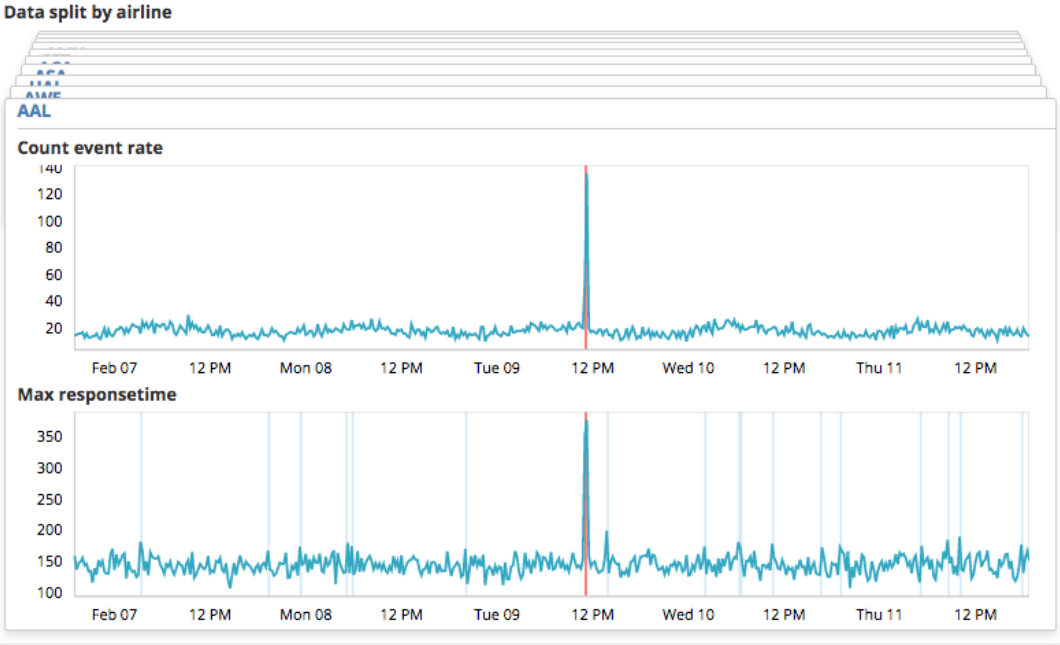


elastic

Steps to Complete

10) See animated learning

11) Click “View Results”

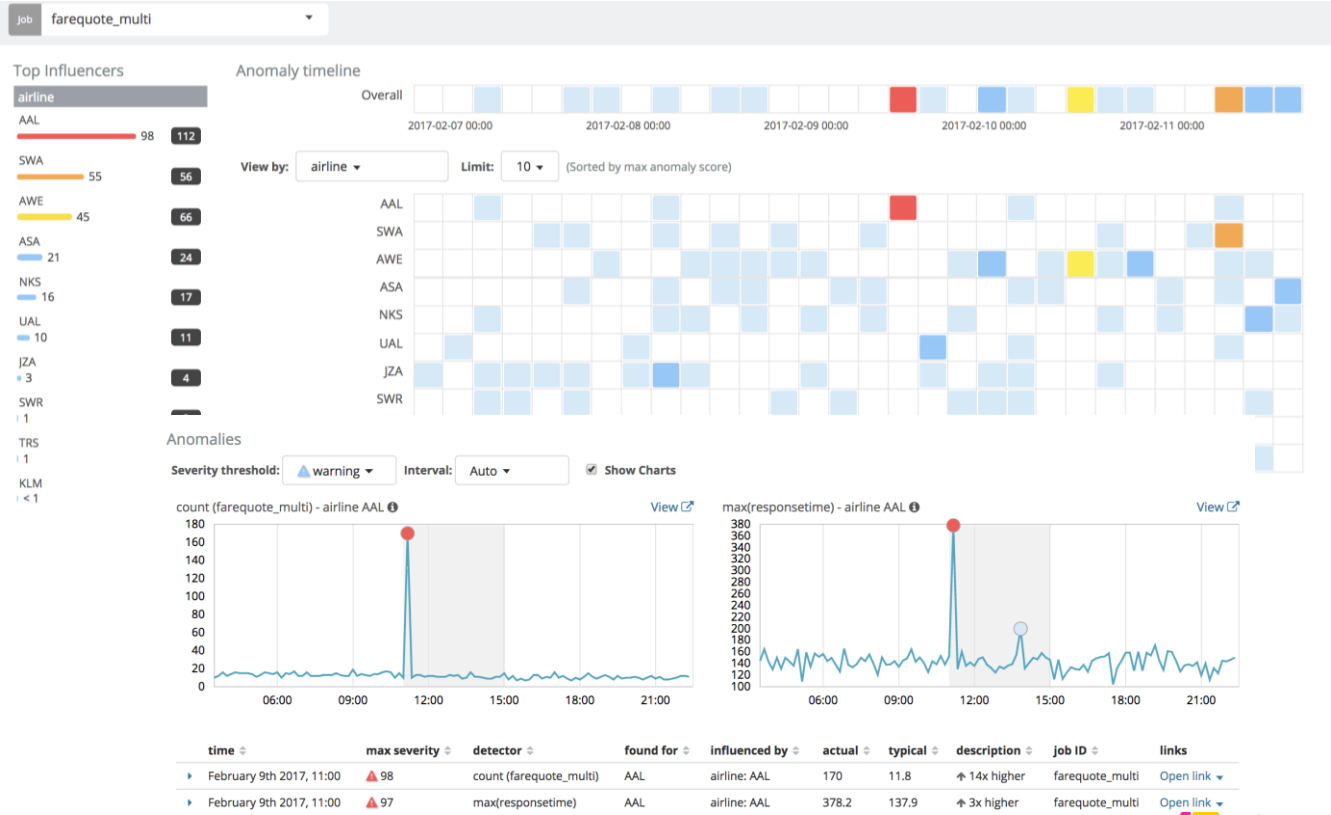


elastic

Steps to Complete

12) Result:

anomalies for AAL in both count and response time



elastic

Lab 4: Multi-Job Analysis

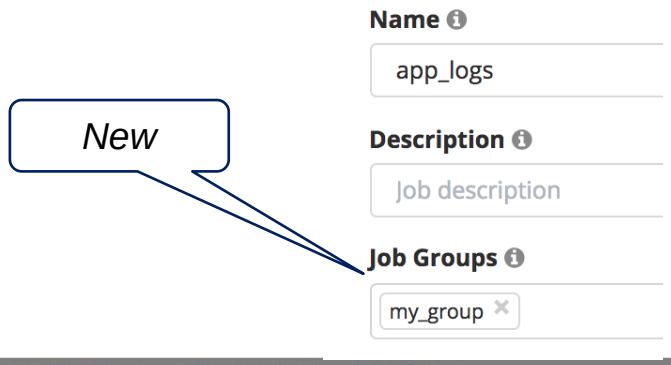
Steps to Complete

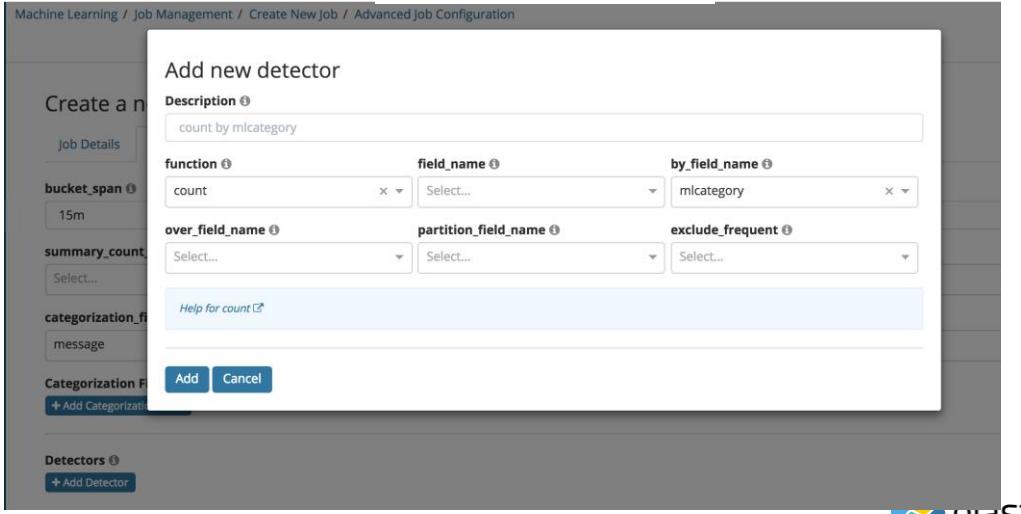
- Create and Advanced job for:
  - “it_ops_app_logs-*”
  - Create a “count by mlcategory” job for the log events
    - use “message” as the categorization_field_name
- Create Multi-Metric jobs for the following indices:
  - it_ops_sql-*
  - it_ops_network-*
- Create Single-Metric job for the following :
  - it_ops_kpi-*
- Group all jobs into the same Job Group name
- View all jobs overlaid in the Explorer View

elastic

Steps to Complete

- For index:it_ops_app_logs
  - create an advanced job
  - put job in a group

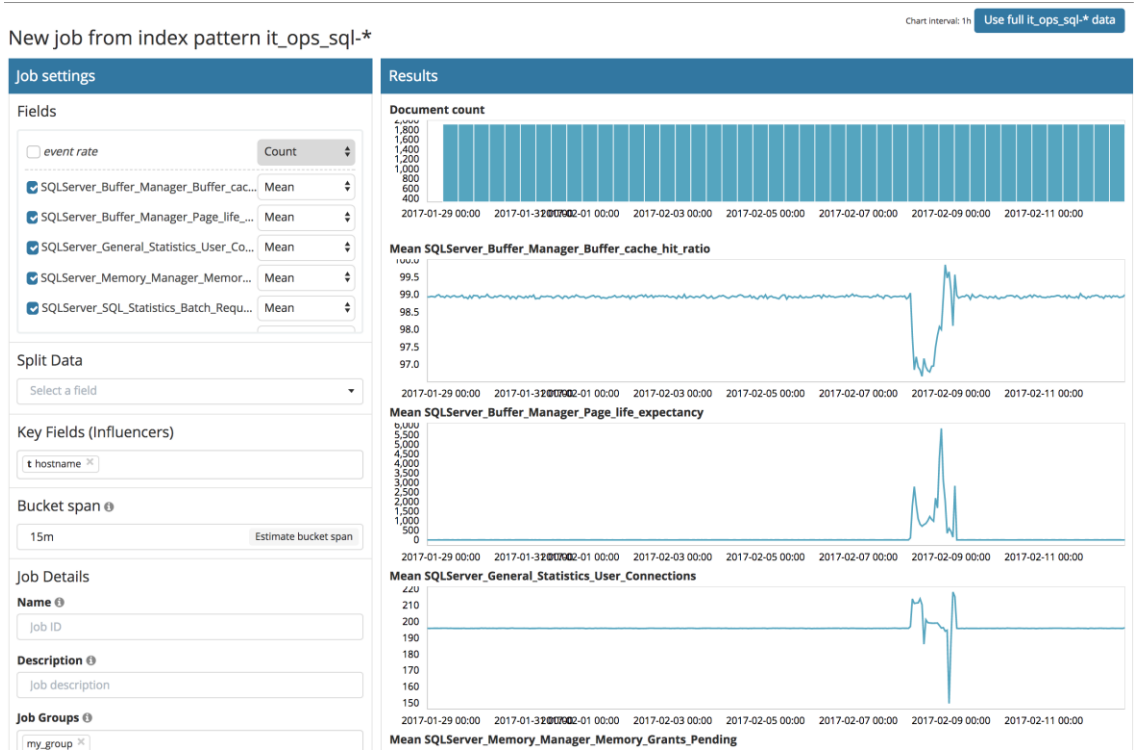




  - make sure you choose “message” for categorization_field_name
  - detector is: count with by_field_name of “mlcategory”

Steps to Complete

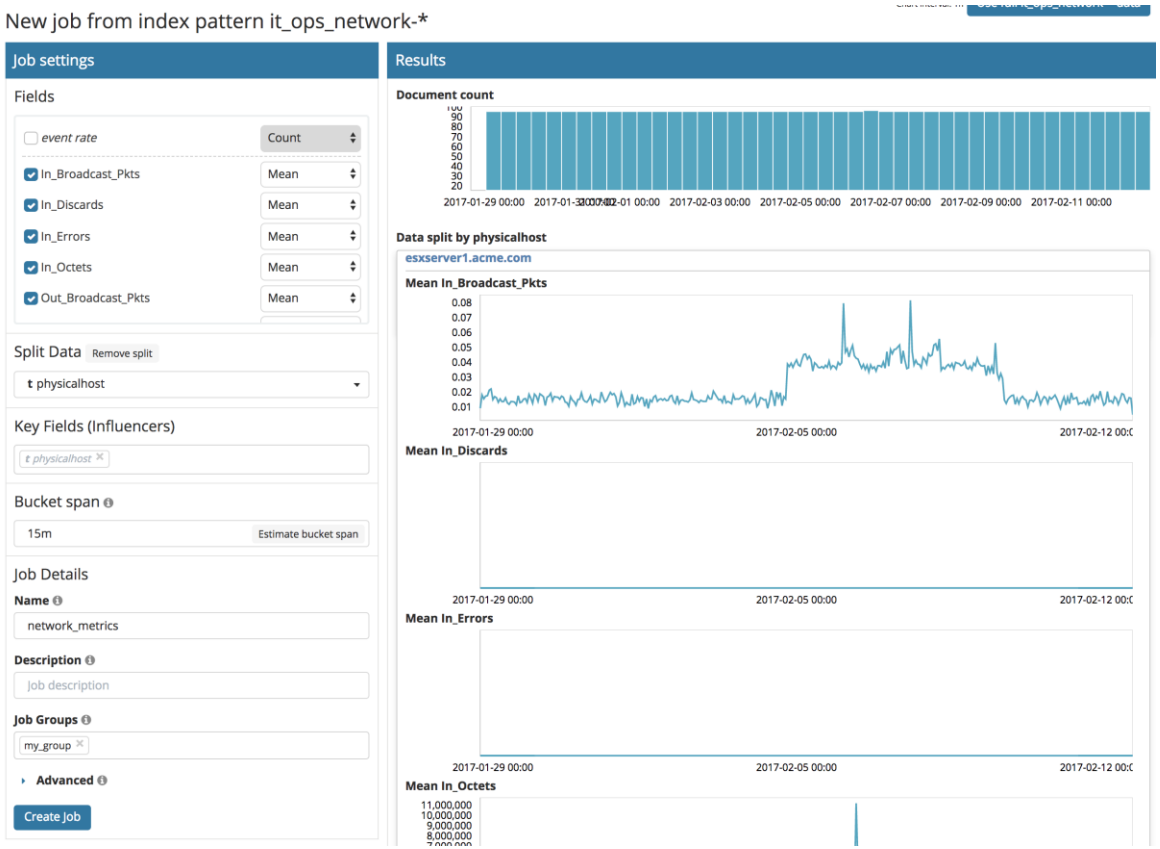
- For index:it_ops_sql-*
  - create multi-metric job
  - “mean” for all SQL metrics
- Also put job in same Job Group as the app logs job



elastic

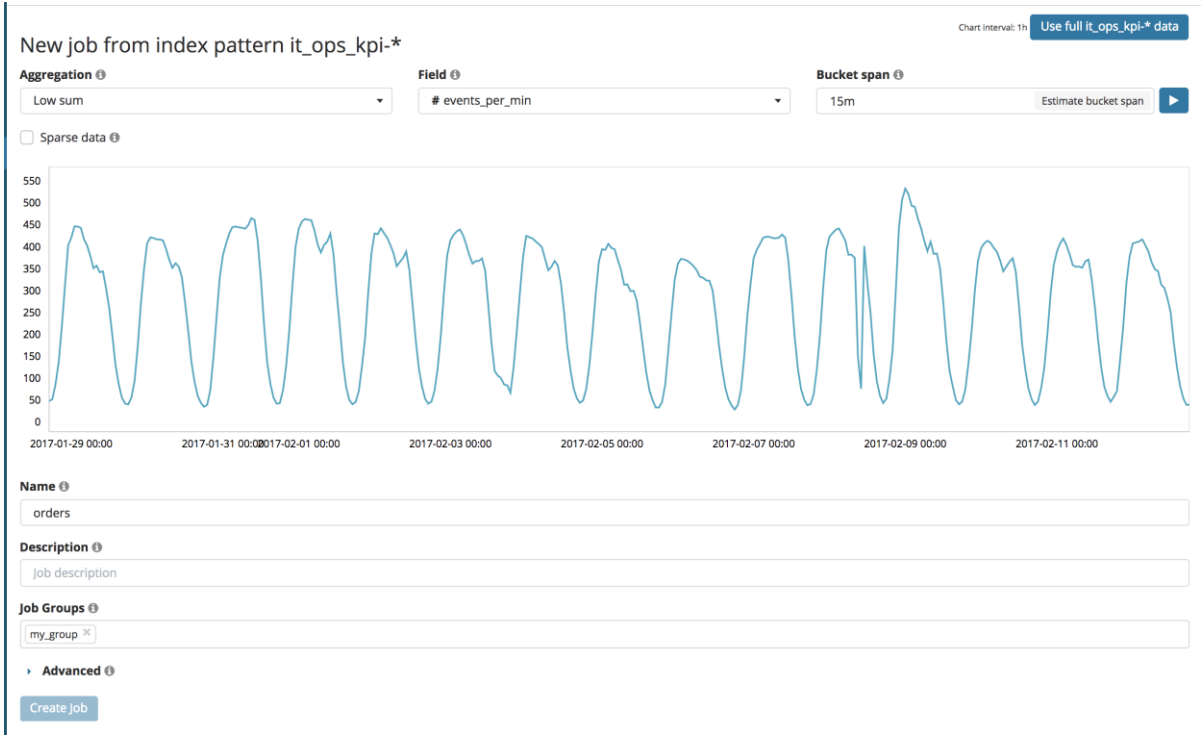
Steps to Complete

- For index:it_ops_network-*
  - create multi-metric job
  - “mean” for all metrics
- Also put job in same Job Group as the app logs job



Steps to Complete

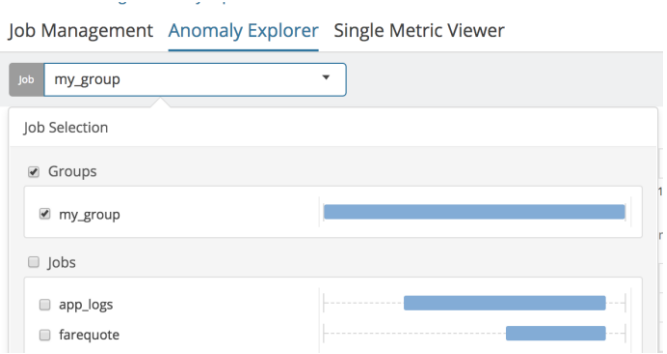
- For index:it_ops_kpi-*
  - create single-metric job
  - “low_sum” for field “events_per_min”
- Also put job in same Job Group as the app logs job



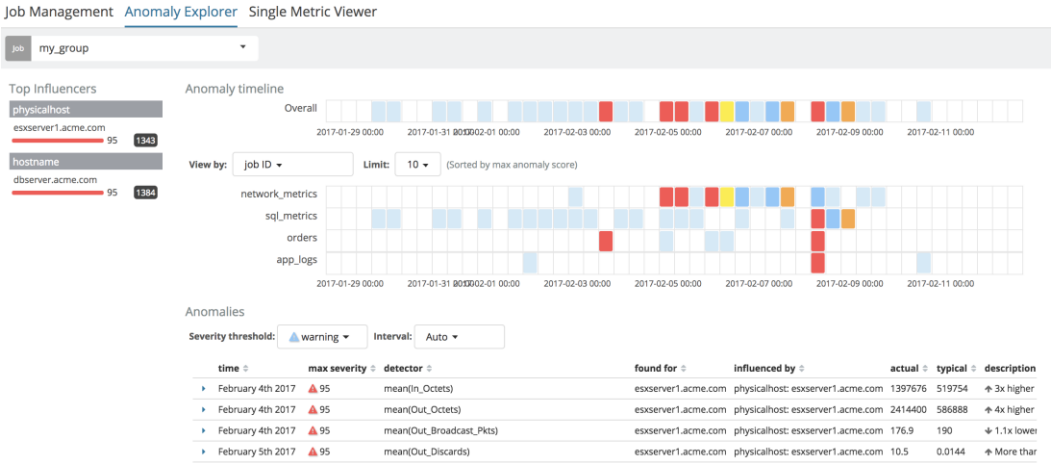
elastic

Steps to Complete

- View all jobs in Group



- See correlated anomalies



elastic

elastic

THANK YOU