elastic

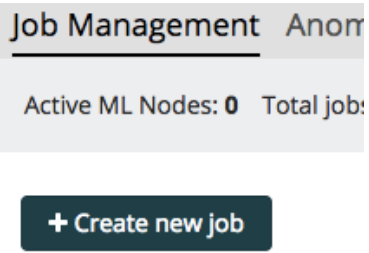
X-Pack
Machine Learning Labs Guide

Lab 1: The Simplest Job

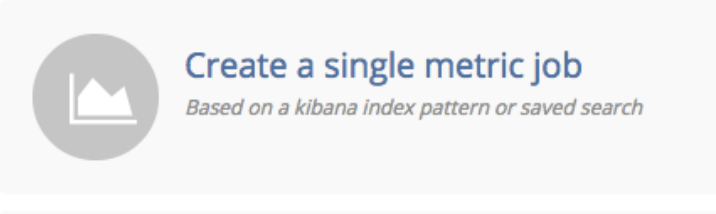
elastic

Steps to Complete

1) In Machine Learning, Create new job



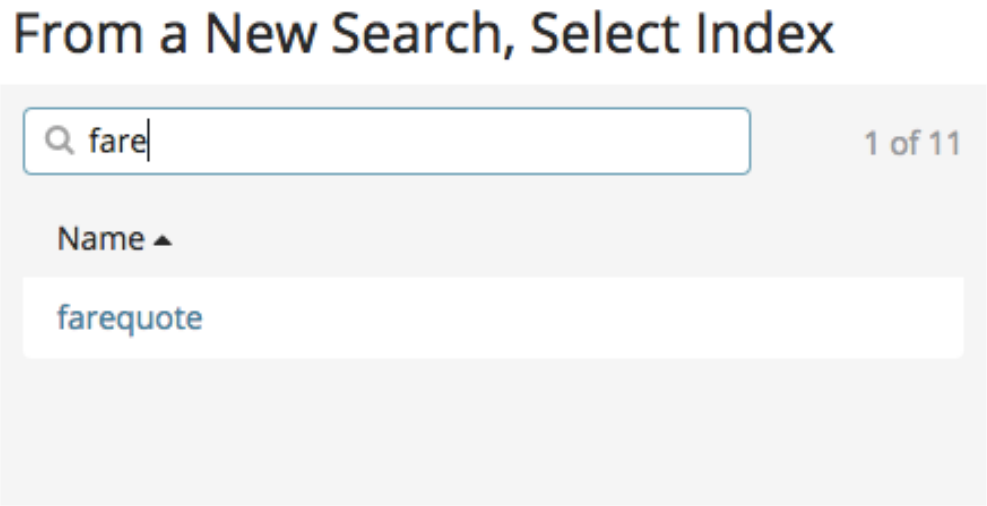
2) Choose single metric



elastic

Steps to Complete

3) pick farequote index



elastic

Steps to Complete

4) choose the “Count” aggregation

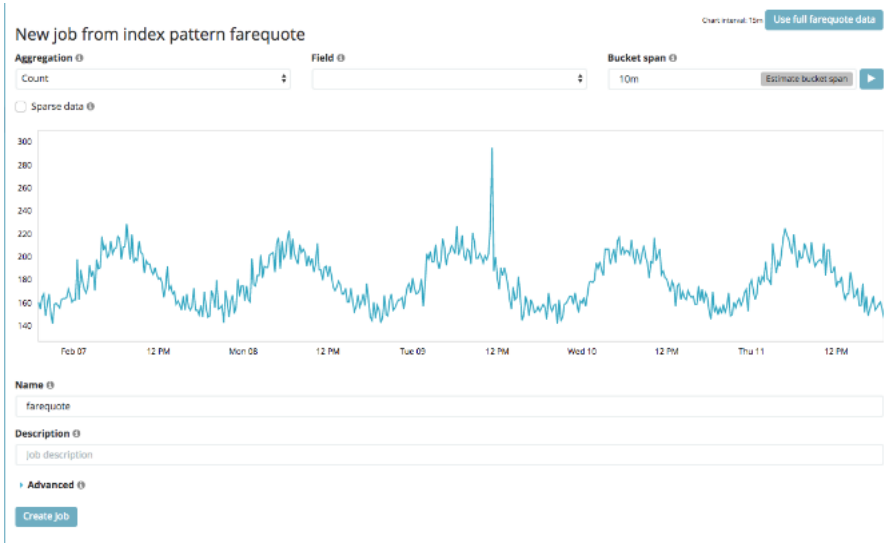
5) select 10m for bucket span

6) leave “field” blank (we don’t count fields, we’re counting documents)

7) click the “use full farequote data” button

8) name job “farequote”

9) click “Create Job”

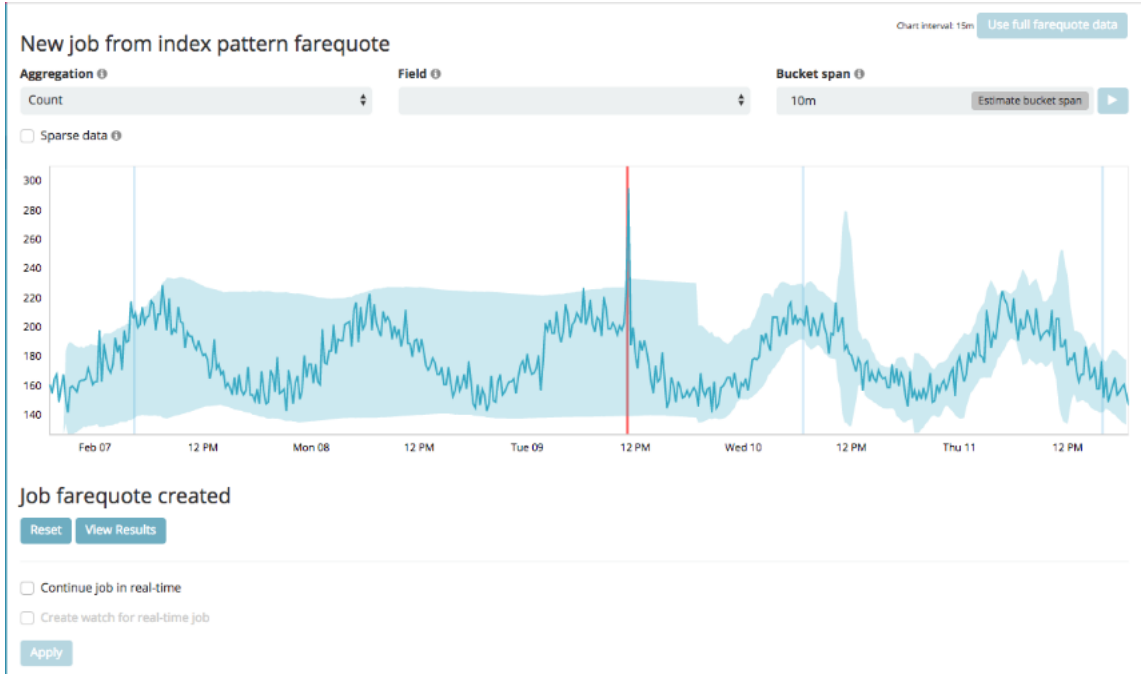


elastic

Steps to Complete

10) See animated learning

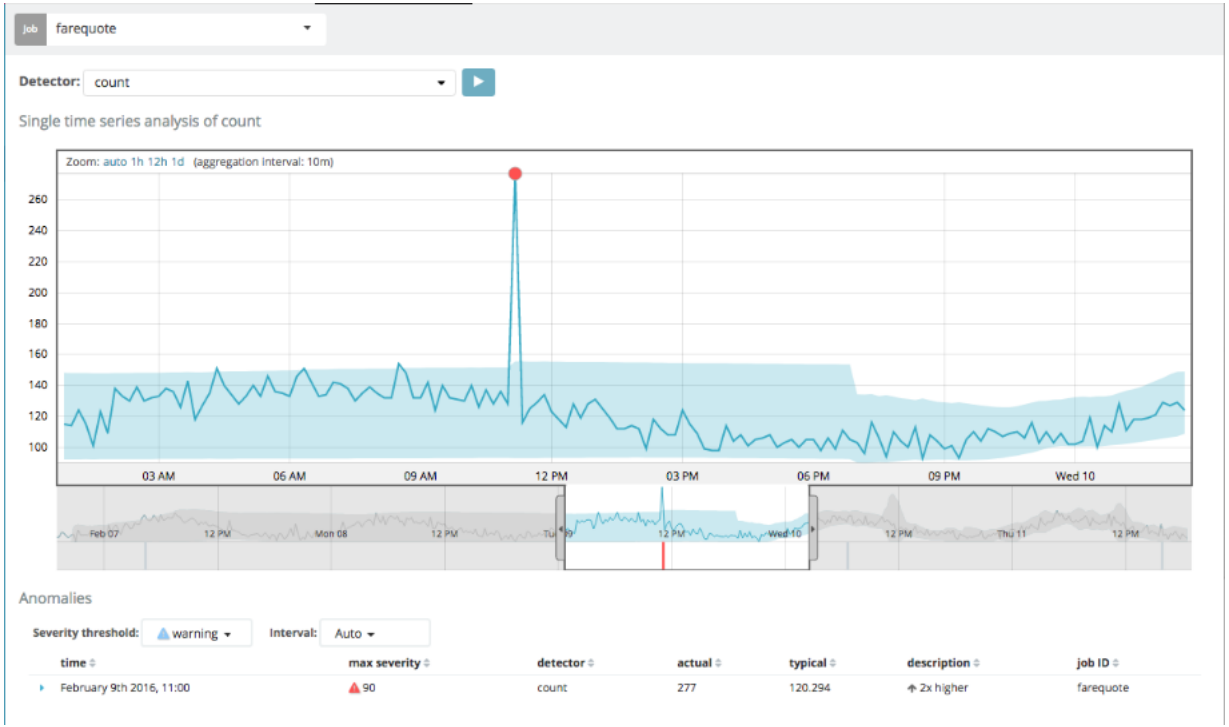
11) click “View Results”



elastic

Steps to Complete

12) Zoom in on anomaly



elastic

Lab 2: Advanced Jobs

elastic

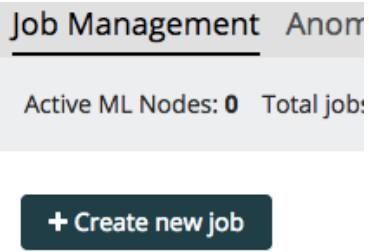
Steps to Complete

- Using the “farequote” data set:
- Get familiar with how the raw data looks in Kibana Discover
- Load the ML home screen and create a new “advanced” job to analyze unusually slow response times per airline
  - choose a bucket_span of 10m
  - pick an appropriate way to split data (hint: partition using airline)
  - select airline as influencer
- Run the job over the entire data set (data is not real-time)

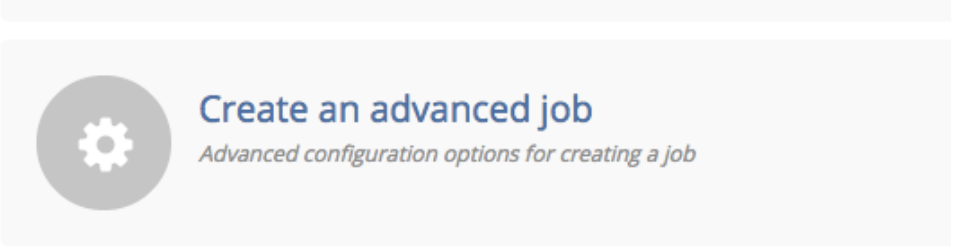
elastic

Steps to Complete

1) Create new job



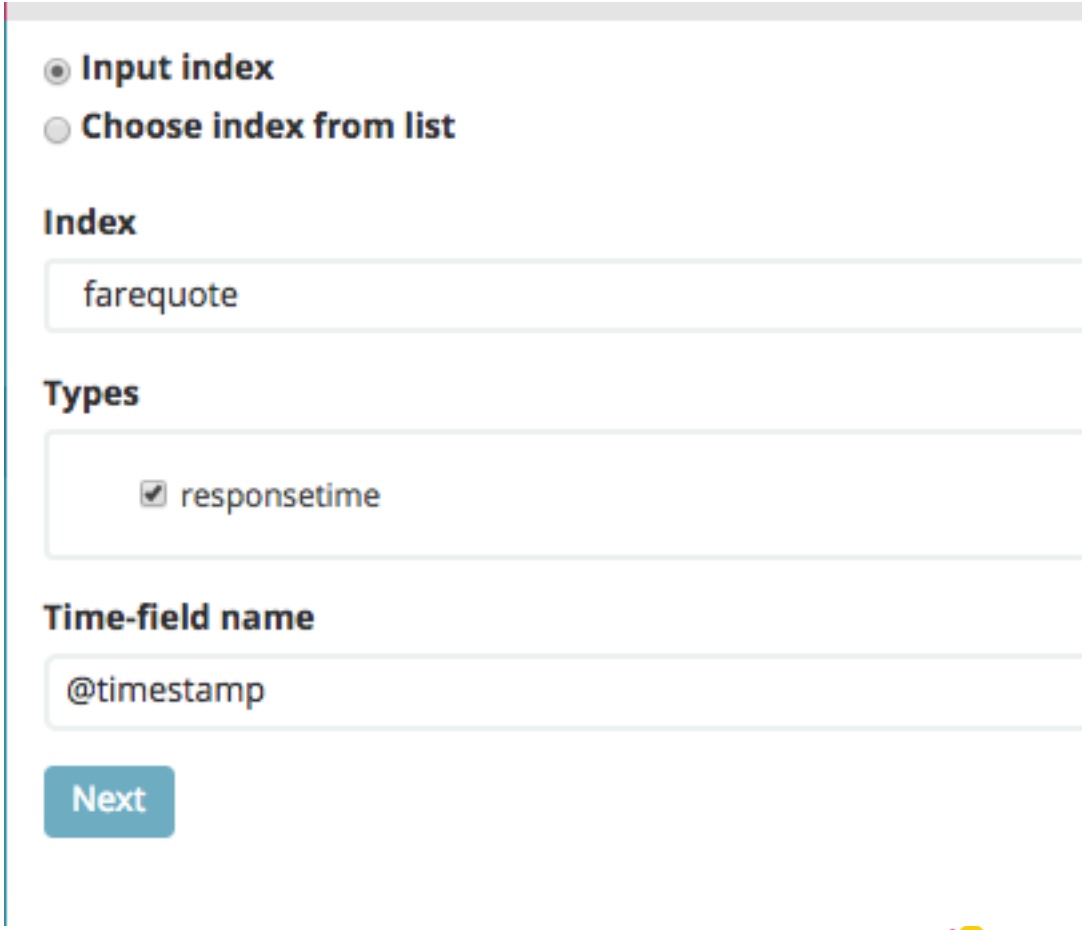
2) Choose advanced



elastic

Steps to Complete

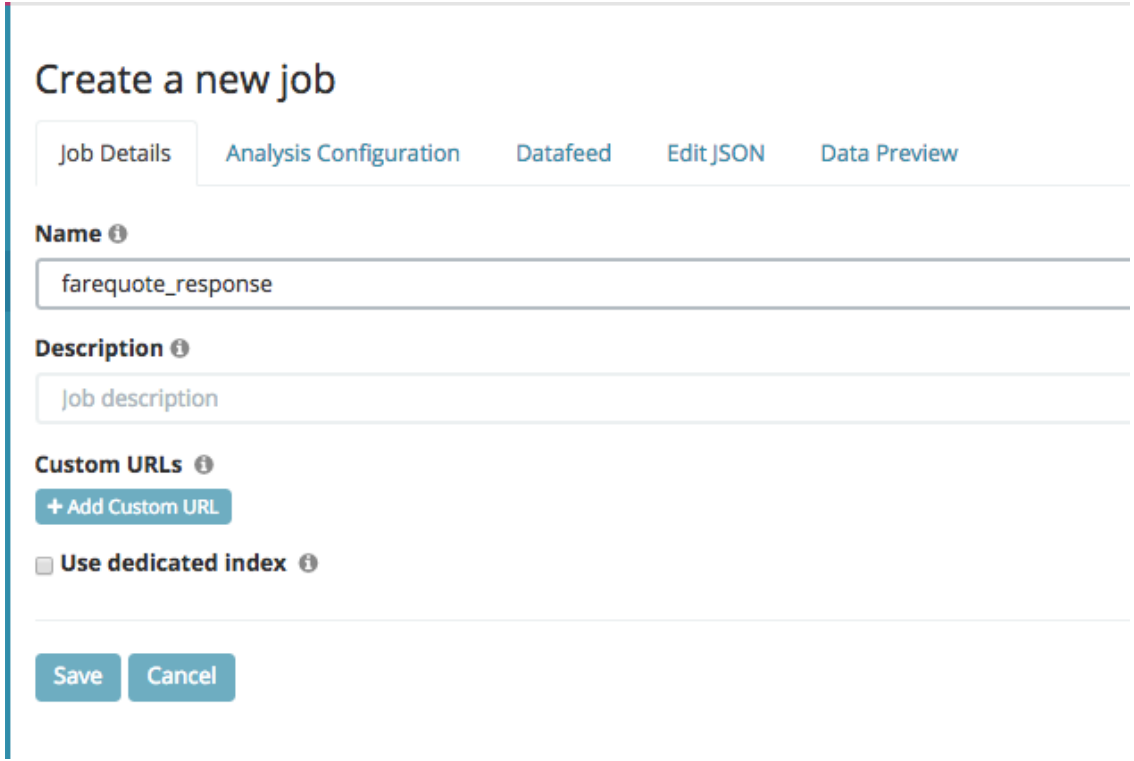
3) Pick farequote index



elastic

Steps to Complete

5) Name job
“farequote_response”



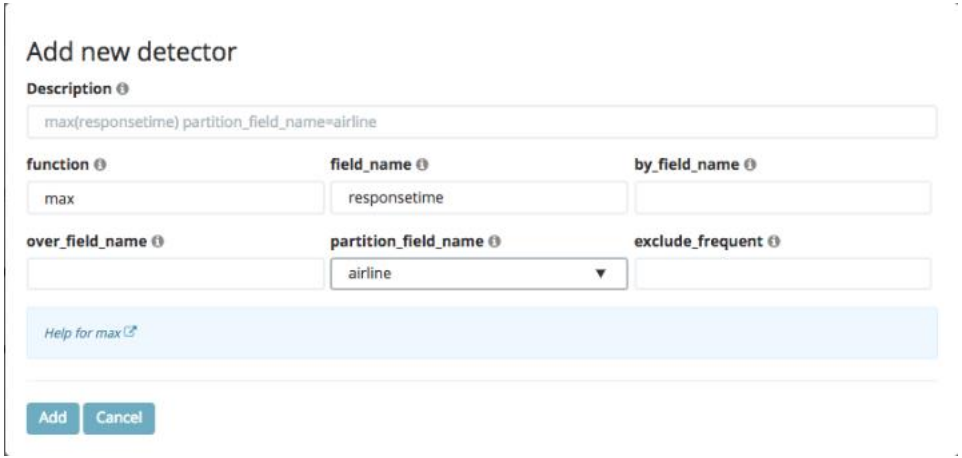
elastic

Steps to Complete

6) set bucket_span=10m

7) Add a Detector:

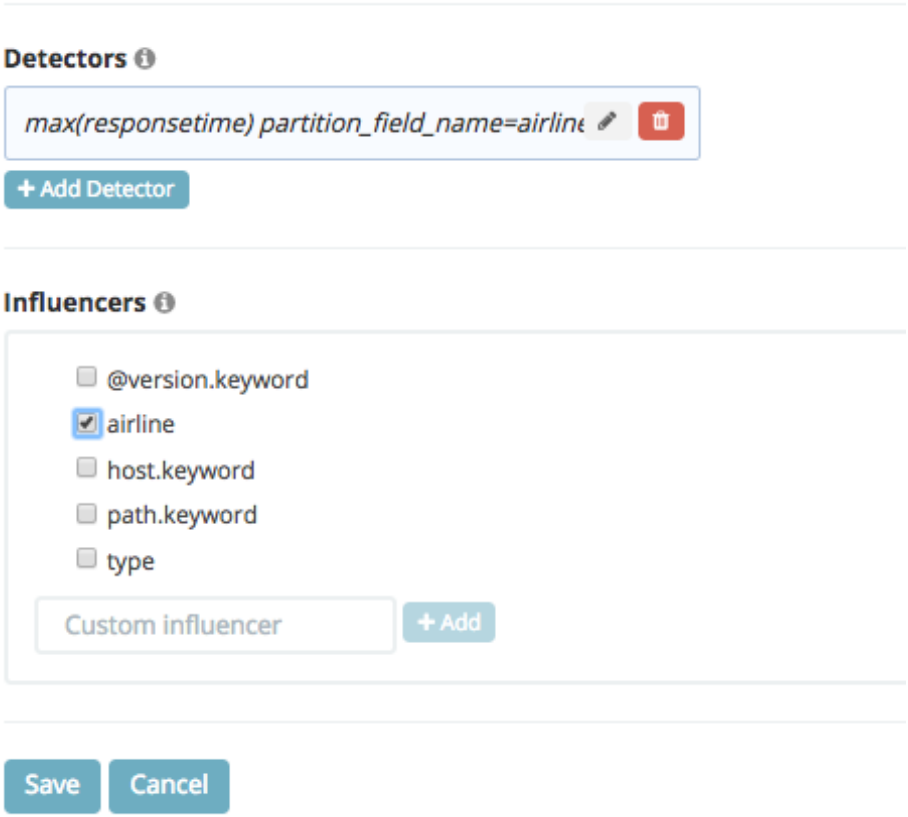
function: max
field_name: responsetime
partition_field_name: airline



elastic

Steps to Complete

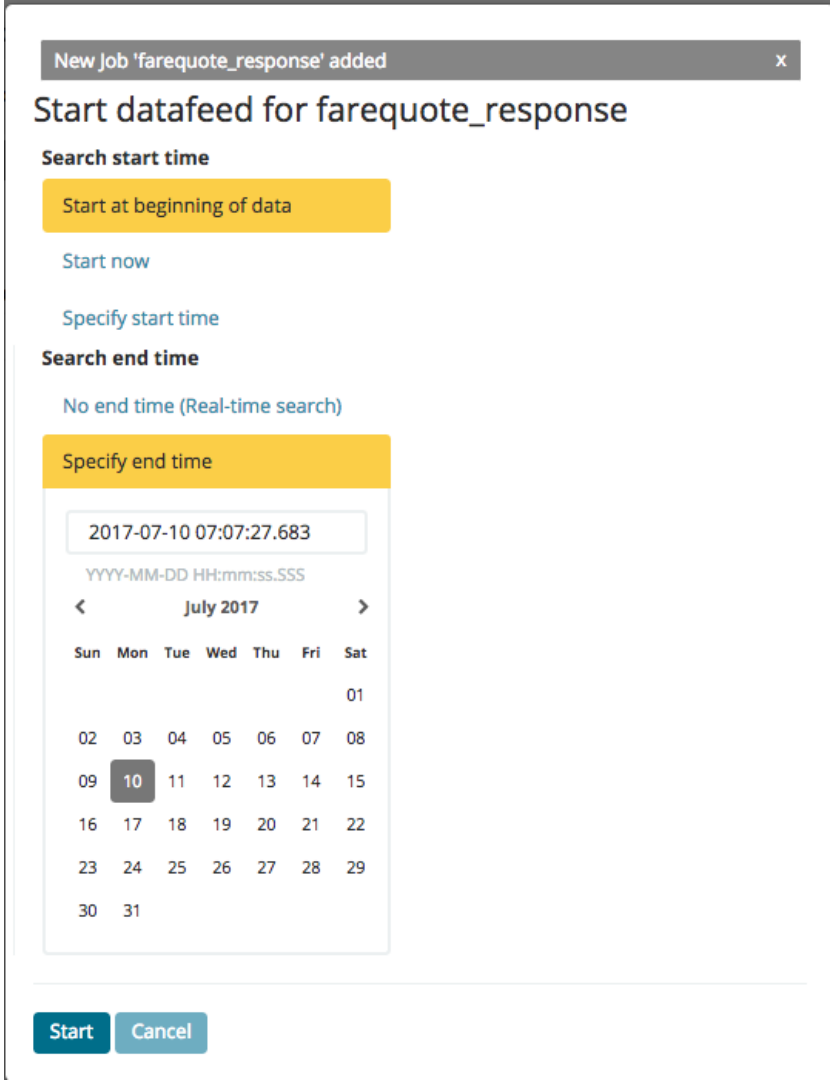
7) select Influencer,
save job
start datafeed



elastic

Steps to Complete

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



elastic

Steps to Complete

9) wait until all 86275 events are processed

+ Create new job

Job filter

Job ID	Description	Processed records	Memory status	Job state	Datafeed state	Latest timestamp	Actions
farequote		720	ok	closed	stopped	2016-02-11 18:59:54	
farequote_response		86,274	ok	closed	stopped	2016-02-11 18:59:54	

Page Size 10

elastic

Lab 3: Multi-Metric Jobs

elastic

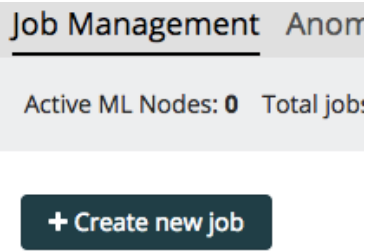
Steps to Complete

- Again, using the “farequote” data set:
- Re-create the “max(responsetime) per airline” job using a “multi-metric” job
- Also add “count per airline” in the same job

elastic

Steps to Complete

1) Create new job

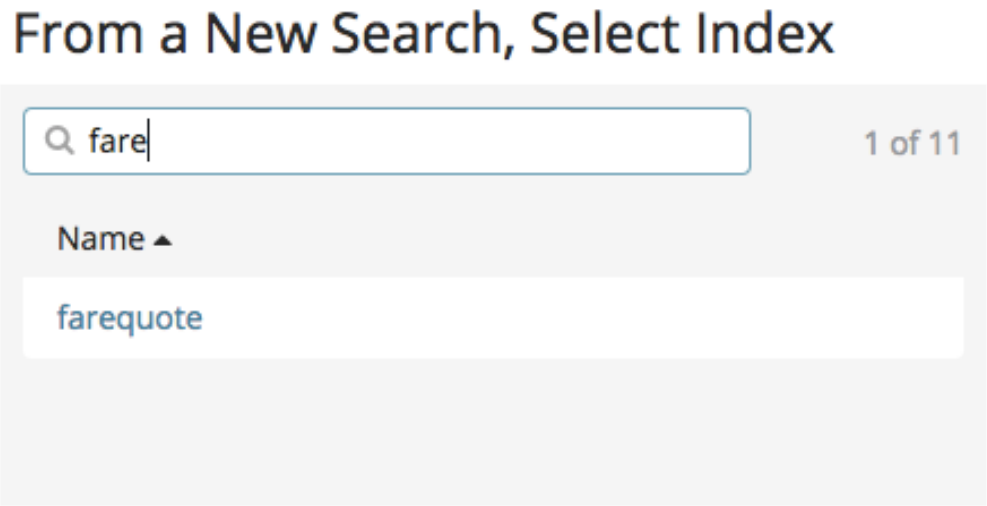


2) Choose multi metric

elastic

Steps to Complete

3) pick farequote index



elastic

Steps to Complete

4) choose
- event rate, count
- responsetime, max

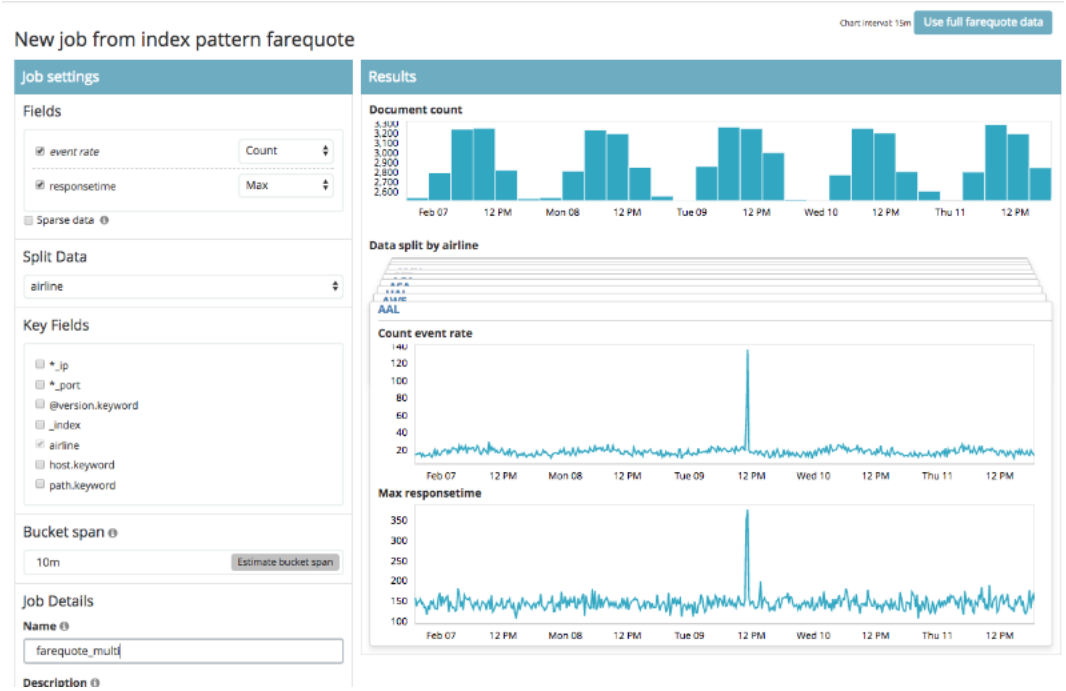
5) select 10m for bucket span

6) Split Data by airline (influencer for airline is chosen for you)

7) click “use full farequote data”

8) name job “farequote_multi”

9) click “Create Job”

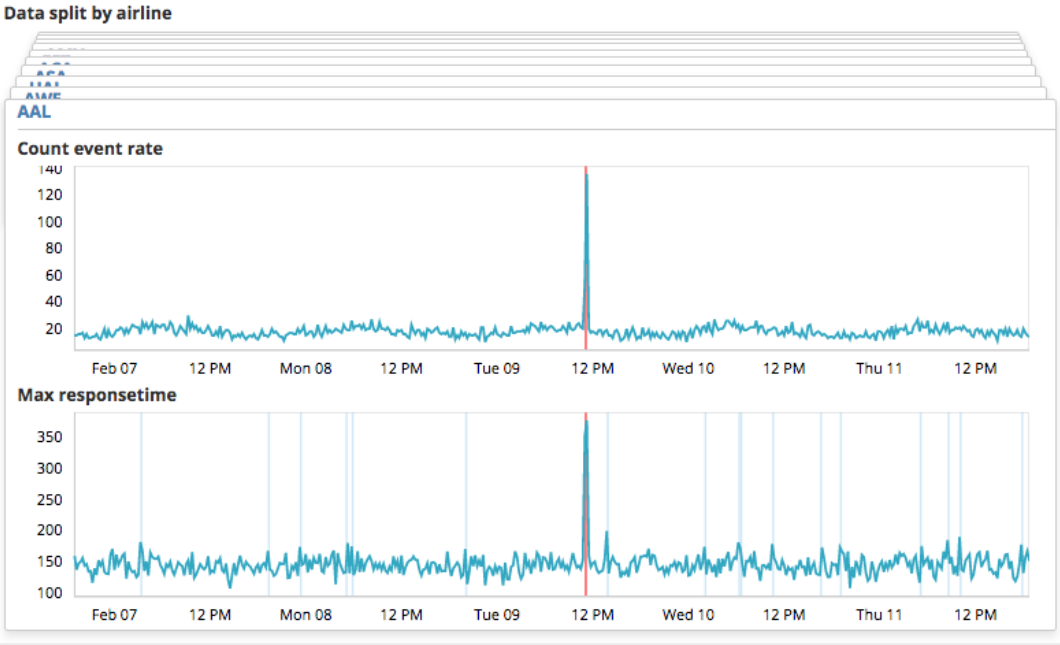


elastic

Steps to Complete

10) See animated learning

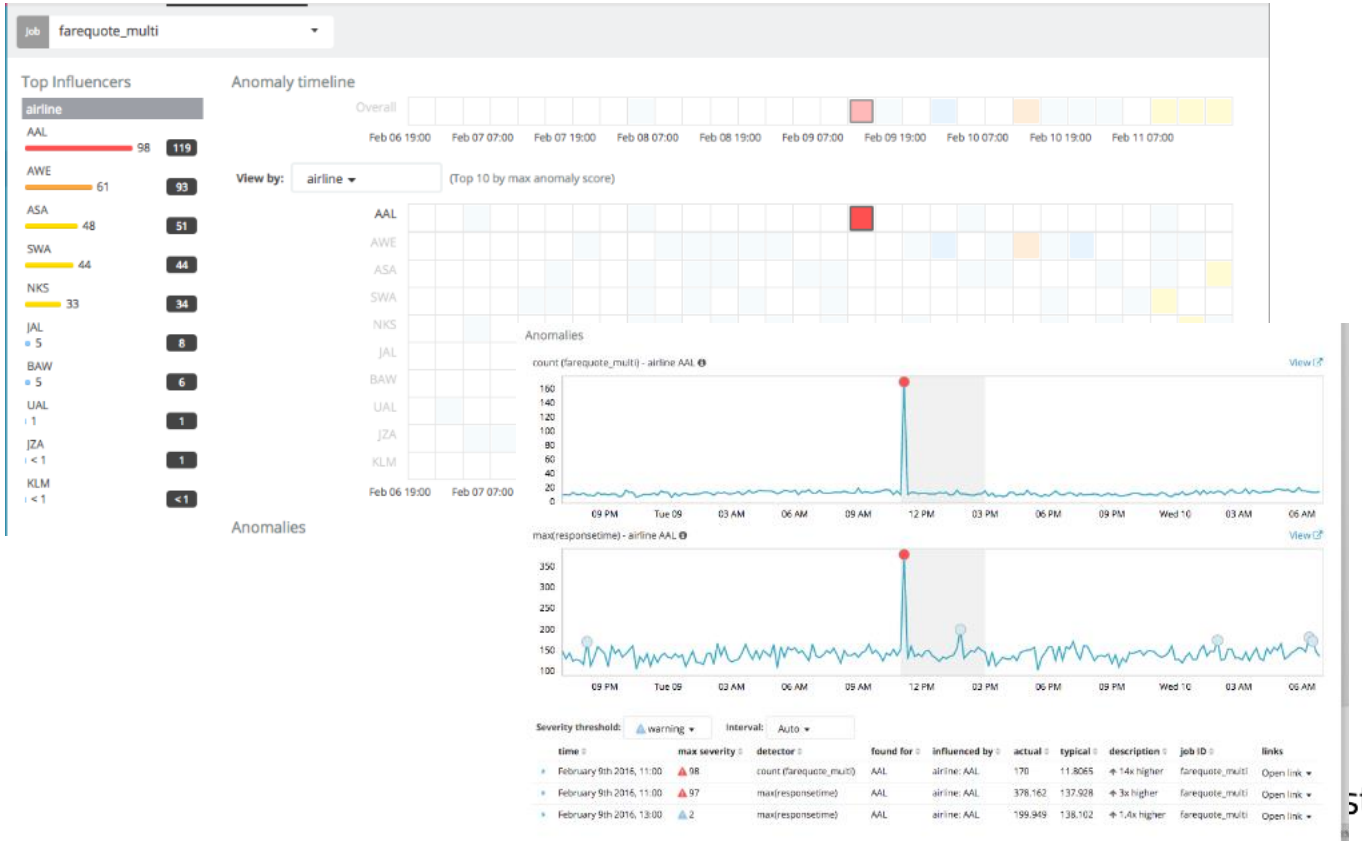
11) click “View Results”



Steps to Complete

12) Result:

anomalies for AAL
in both count
and response time



stic

Lab 4: Multi-Job Analysis

elastic

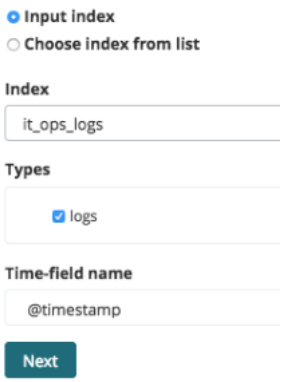
Steps to Complete

- Using the “it_ops_logs” data set:
  - Create a “count by mlcategory” job for the log events
    - use “message” as the categorization_field_name
- Using the “it_ops_metrics” data set:
  - Create a “mean(metricvalue) by metricname” job for the metrics
- View both jobs overlaid in the Explorer View

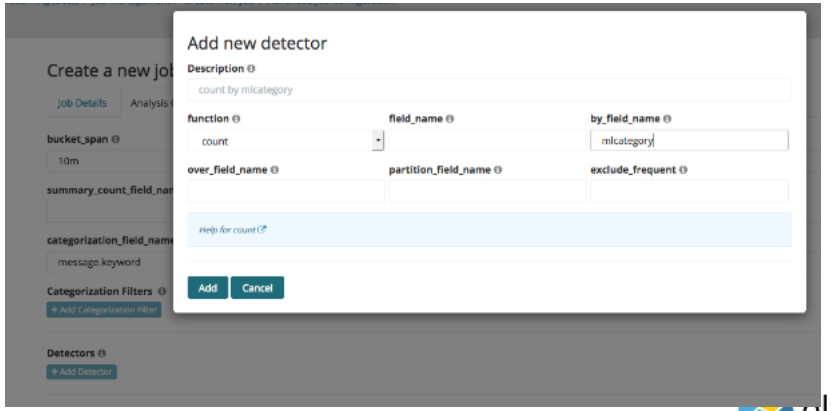
elastic

Steps to Complete

- Answer
  - For index:it_ops_logs
    - create an advanced job



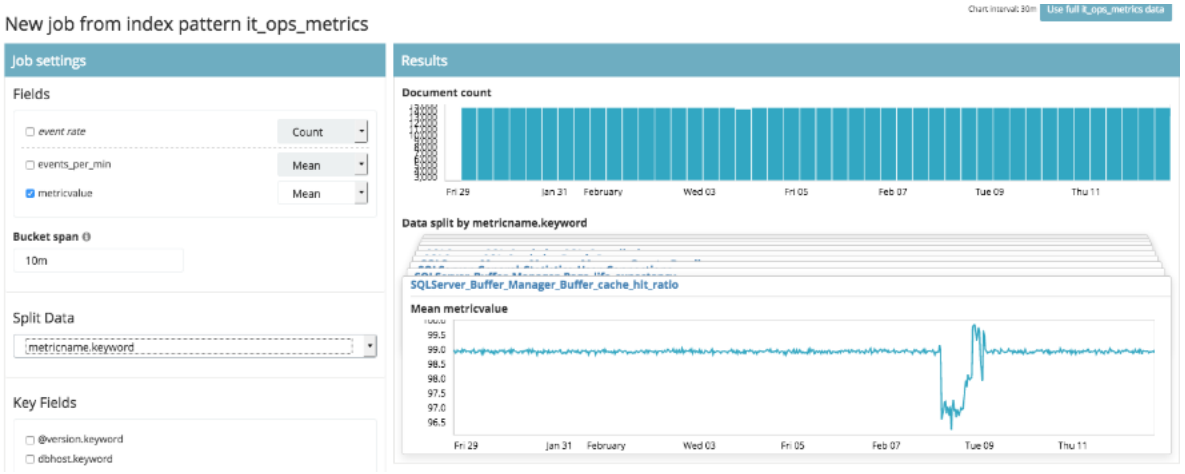
    - make sure you choose “message” for categorization_field_name
    - detector is: count with by_field_name of “mlcategory”



elastic

Steps to Complete

- Answer
  - For index:it_ops_metrics
    - create multi-metric job
    - mean of metricvalue split on metricname.keyword



Steps to Complete

- Your goal is to get this View: