

Chapter 9 – Processing Documents in ingest Pipelines

Example Request - create_pipeline_range_ratio

PUT http://localhost:9200/_ingest/pipeline/range_ratio

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "description": "calculate the range of high and low and divide by the close
price",
  "processors": [
    {
      "script": {
        "source": "ctx.range_ratio = (ctx.high-
ctx.low)/ctx.close"
      }
    }
  ]
}
```

```
curl --location --request PUT "http://localhost:9200/_ingest/pipeline/range_ratio" --
header "Content-Type: application/json" --header "accept: application/json" --data
${'{"description\":"calculate the range of high and low and divide by the close
price", \"processors\":[ { \"script\": { \"source\": \"ctx.range_ratio = (ctx.high-
ctx.low)/ctx.close\" } } ] }'}
```

Example Request - delete_pipeline_range_ratio

DEL http://localhost:9200/_ingest/pipeline/range_ratio

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "description": "calculate the range of high and low and divide by the close
price",
  "processors": [
    {
      "set": {
        "field": "range_ratio",
```

```

        "value": "({{high}}-{{low}})/{{close}}"
    }
}
]
}

```

```

curl --location --request DELETE "http://localhost:9200/_ingest/pipeline/range_ratio"
--header "Content-Type: application/json" --header "accept: application/json" --data
$'{\ "description\":"calculate the range of high and low and divide by the close
price", \ "processors\":[ { \ "set\": { \ "field\": \ "range_ratio", \ "value\":
\ "({{high}}-{{low}})/{{close}}\ " } } ] }'

```

Example Request - delete_pipeline_wildcard_expression

DEL http://localhost:9200/_ingest/pipeline/range_ratio*

HEADERS

```

Content-Type
application/json
accept
application/json

```

BODY

```

{
  "description":"calculate the range of high and low and divide by the close
price",
  "processors":[
    {
      "set": {
        "field": "range_ratio",
        "value": "({{high}}-{{low}})/{{close}}"
      }
    }
  ]
}

```

```

curl --location --request DELETE "http://localhost:9200/_ingest/pipeline/range_ratio*"
--header "Content-Type: application/json" --header "accept: application/json" --data
$'{\ "description\":"calculate the range of high and low and divide by the close
price", \ "processors\":[ { \ "set\": { \ "field\": \ "range_ratio", \ "value\":
\ "({{high}}-{{low}})/{{close}}\ " } } ] }'

```

Example Request - simulate_pipeline_range_ratio

POST http://localhost:9200/_ingest/pipeline/range_ratio/_simulate

HEADERS

```

Content-Type
application/json
accept
application/json

```

BODY

```
{
  "docs" :[
    { "_source": {
      "date": "2018-12-26",
      "open": 24.8706,
      "high": 25.6553,
      "low": 24.7117,
      "close": 25.6553,
      "volume": 422158,
      "change": 0.84425,
      "changePercent": 3.403,
      "label": "Dec 26, 18",
      "changeOverTime": 0,
      "symbol": "ACWF"
    }
  ]
}
```

```
curl --location --request POST
"http://localhost:9200/_ingest/pipeline/range_ratio/_simulate" --header "Content-Type:
application/json" --header "accept: application/json" --data
'${"docs" :[ {"_source": {"date": "2018-12-26", "open": 24.8706, "high":
25.6553, "low": 24.7117, "close": 25.6553, "volume": 422158, "change":
0.84425, "changePercent": 3.403, "label": "Dec 26, 18", "changeOverTime": 0,
"symbol": "ACWF"}} ] }'
```

Example Request - simulate_pipeline_ad_hoc_range_ratio

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "pipeline": {
    "description": "calculate the range of high and low and divide by the
close price",
    "processors": [{
      "script": {
        "source": "ctx.range_ratio = (ctx.high-
ctx.low)/ctx.close"
      }
    ]
  },
  "docs" :[
    { "_source": { "date": "2018-12-26", "open": 24.8706, "high": 25.6553,
"low": 24.7117, "close": 25.6553, "volume": 422158, "unadjustedVolume": 422158,
"change": 0.84425, "changePercent": 3.403, "vwap": 25.2926, "label": "Dec 26, 18",
"changeOverTime": 0, "symbol": "ACWF"}
  ]
}
```

```
}
```

```
curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --header "Content-Type: application/json" --header "accept: application/json" --data '${\ "pipeline": { \ "description": "\ calculate the range of high and low and divide by the close price", \ "processors": [{ \ "script": { \ "source": \ "ctx.range_ratio = (ctx.high-ctx.low)/ctx.close" } } ] }, \ "docs" : [ { \ "_source": { \ "date": \ "2018-12-26", \ "open": 24.8706, \ "high": 25.6553, \ "low": 24.7117, \ "close": 25.6553, \ "volume": 422158, \ "change": 0.84425, \ "changePercent": 3.403, \ "label": \ "Dec 26, 18", \ "changeOverTime": 0, \ "symbol": \ "ACWF" } } ] }'
```

Example Request - index_with_pipeline_range_ratio

POST http://localhost:9200/cf_etf_pipeline/_doc/1?pipeline=range_ratio

HEADERS

Content-Type
application/json
accept
application/json

PARAMS

pipeline
range_ratio

BODY

```
{
  "date": "2018-12-26",
  "open": 24.8706,
  "high": 25.6553,
  "low": 24.7117,
  "close": 25.6553,
  "volume": 422158,
  "change": 0.84425,
  "changePercent": 3.403,
  "label": "Dec 26, 18",
  "changeOverTime": 0,
  "symbol": "ACWF"
}
```

```
curl --location --request POST "http://localhost:9200/cf_etf_pipeline/_doc/1?pipeline=range_ratio" --header "Content-Type: application/json" --header "accept: application/json" --data '${\ "date": \ "2018-12-26", \ "open": 24.8706, \ "high": 25.6553, \ "low": 24.7117, \ "close": 25.6553, \ "volume": 422158, \ "change": 0.84425, \ "changePercent": 3.403, \ "label": \ "Dec 26, 18", \ "changeOverTime": 0, \ "symbol": \ "ACWF" }'
```

Example Request - index_with_pipeline_range_ratio_doc

DEL http://localhost:9200/cf_etf_pipeline/_doc/1

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "date": "2018-12-26",
  "open": 24.8706,
  "high": 25.6553,
  "low": 24.7117,
  "close": 25.6553,
  "volume": 422158,
  "change": 0.84425,
  "changePercent": 3.403,
  "label": "Dec 26, 18",
  "changeOverTime": 0,
  "symbol": "ACWF"
}
```

```
curl --location --request DELETE "http://localhost:9200/cf_etf_pipeline/_doc/1" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{\"date\": \"2018-12-26\", \"open\": 24.8706, \"high\": 25.6553, \"low\": 24.7117,
\"close\": 25.6553, \"volume\": 422158, \"change\": 0.84425, \"changePercent\":
3.403, \"label\": \"Dec 26, 18\", \"changeOverTime\": 0, \"symbol\": \"ACWF\"}'
```

Example Request - get_document_from_index_with_pipeline_range_ratio

GET http://localhost:9200/cf_etf_pipeline/_doc/1

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
}
```

```
curl --location --request GET "http://localhost:9200/cf_etf_pipeline/_doc/1"
```

Example Request - simulate_pipeline_ad_hoc_kv_processor

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "pipeline": {
    "description": "kv processor demo",
    "processors": [
      { "kv": {
        "field": "input_field",
        "field_split": ",",
        "value_split": ":",
        "target_field": "output_field",
        "trim_key": " ",
        "trim_value": " "
      } },
      { "remove": { "field": ["input_field"] } }
    ],
    "docs": [
      { "_source": { "input_field": " a:b , c:d " } }
    ]
  }
}
```

```
curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{\"pipeline\": { \"description\": \"kv processor demo\", \"processors\": [ {\"kv\":
{ \"field\": \"input_field\", \"field_split\": \"\\\", \"value_split\": \"\\\":\",
\\\"target_field\": \"output_field\", \"trim_key\": \" \", \"trim_value\": \" \" }},
{ \"remove\": { \"field\": [\"input_field\"] } } ] }, \"docs\": [ {\"_source\":
{ \"input_field\": \" a:b , c:d \" } } ] }'
```

Example Request - simulate_pipeline_grok_processor

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "pipeline": {
    "description": "grok processor demo",
    "processors": [
      { "grok": {
        "field": "message",
        "patterns": [ "\\[%{TIMESTAMP_ISO8601:timestamp}\\]\\[%{LOGLEVEL:log-
level}%{SPACE}\\]\\[%{JAVACLASS:class}\\]\\[%{SPACE}\\]\\[%{DATA:node}\\]\\[%{SPACE}\\]\\[%{DATA:i
ndex_doc}\\]\\[%{GREEDYDATA:action}\\]",
        "ignore_missing": true
      } },
      { "remove": { "field": ["message"] } }
    ],
    "docs": [
```

```

        {"_source": {"message": "[2019-06-20T00:04:58,856][INFO ][o.e.c.m.MetadataDeleteIndexService] [WTW.local] [cf_etf/RiYRsKyxTGStNcHczbUxcQ] deleting index"}}
    ]
}

```

```

curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{"pipeline": { \ "description\":"grok processor demo",
\ "processors\":[ { \ "grok\": { \ "field\": \ "message\","patterns\":
[ \ "\ \ \ \ [%{TIMESTAMP_ISO8601:timestamp} \ \ \ \ ] \ \ \ \ [%{LOGLEVEL:log-
level}%{SPACE} \ \ \ \ ] \ \ \ \ [%{JAVACLASS:class} \ \ \ \ ] \ \ \ \ [%{SPACE} \ \ \ \ [%{DATA:node} \ \ \ \ ] \ \ \ \ [%{SPACE} \
\ \ \ \ [%{DATA:index_doc} \ \ \ \ ] \ \ \ \ [%{GREEDYDATA:action} \ "], \ "ignore_missing\": true }},
{ \ "remove\": { \ "field\": [ \ "message \ " ] } }, \ "docs\": [ { \ "_source\":
{ \ "message\": \ "[2019-06-20T00:04:58,856][INFO ][o.e.c.m.MetadataDeleteIndexService]
[WTW.local] [cf_etf/RiYRsKyxTGStNcHczbUxcQ] deleting index \ " } } ] }'

```

Example Request - simulate_pipeline_dissect_processor

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```

{
  "pipeline": {
    "description": "dissect processor demo",
    "processors": [
      {
        "dissect": {
          "field": "message",
          "pattern":
            "[%{timestamp}][%{log-level}][%{class}] [%{node}]"
        },
        "[%{index_doc}][%{action}]",
        "ignore_missing": true
      },
      {
        "remove": { "field": ["message"] }
      }
    ],
    "docs": [
      {
        "_source": {
          "message": "[2019-06-20T00:04:58,856][INFO ][o.e.c.m.MetadataDeleteIndexService] [WTW.local] [cf_etf/RiYRsKyxTGStNcHczbUxcQ] deleting index"
        }
      }
    ]
  }
}

```

```

curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{"pipeline": { \ "description\":"dissect processor demo",
\ "processors\":[ { \ "dissect\": { \ "field\": \ "message\","pattern\":
\ "[%{timestamp}][%{log-level}][%{class}] [%{node}] [%{index_doc}][%{action} \ "],

```

```
"ignore_missing": true }}, {"remove": {"field": ["message"]} } ] },
"docs": [ {"_source": {"message": "[2019-06-
20T00:04:58,856][INFO ][o.e.c.m.MetadataDeleteIndexService] [WTW.local]
[cf_etf/RiYRsKyxTGStNcHczbUxcQ] deleting index\"}} ] }'
```

Example Request - simulate_pipeline_simple_conditional_execution

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

```
Content-Type
application/json
accept
application/json
```

BODY

```
{
  "pipeline": {
    "description": "remove the rating field is the rating is equal to a
single space string",
    "processors": [{
      "remove": {
        "field": "rating",
        "if": "ctx.rating == ' '"
      }
    }]
  },
  "docs" : [
    { "_source": { "symbol": "AGT", "fund_name": "iShares MSCI Argentina and
Global Exposure ETF", "rating": " ", "morningstar_category": "Miscellaneous
Region", "category": "International", "family": "iShares", "market_cap": "mid", "description":
"The investment seeks to track the investment results of the MSCI All Argentina
25\\50 Index (the \\\"underlying index\\\" ).\\n The fund generally will invest at
least 90% of its assets in the component securities of the underlying index and in
investments that have economic characteristics that are substantially identical to the
component securities of the underlying index. The index is designed to measure the
broad Argentina equity universe, while including a minimum number of constituents, as
defined by MSCI, Inc. (the \\\"index provider\\\" or \\\"MSCI\\\" ). The fund is non-
diversified.", "exchange": "Cboe Global Markets EDGX"} }
  ]
}
```

```
curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --header "Content-Type: application/json" --header "accept: application/json" --data '{"pipeline": { "description": "remove the rating field if the rating is equal to a single space string", "processors": [ { "remove": { "field": "rating", "if": "ctx.rating == ' '" } } ], "docs": [ { "_source": { "symbol": "AGT", "fund_name": "iShares MSCI Argentina and Global Exposure ETF", "rating": " ", "morningstar_category": "Miscellaneous Region", "category": "International", "family": "iShares", "market_cap": "mid", "description": "The investment seeks to track the investment results of the MSCI All Argentina 25/50 Index (the \"underlying index\"). The fund generally will invest at least 90% of its assets in the component securities of the underlying index and in investments that have economic characteristics that are substantially identical to the component securities of the underlying index. The index is designed
```



```
to measure the broad Argentina equity universe, while including a minimum number of constituents, as defined by MSCI, Inc. (the \\\\\"index provider\\\\\\" or \\\\\"MSCI\\\\\\""). The fund is non-diversified.\\",\\"exchange\\":\\"Cboe Global Markets EDGX\\"}} ] }'
```

Example Request - get_document_empty_rating_from_cf_etf

POST http://localhost:9200/cf_etf/_search

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "query": {
    "match": {
      "symbol": "AGT"
    }
  }
}
```

```
curl --location --request POST "http://localhost:9200/cf_etf/_search" --header "Content-Type: application/json" --header "accept: application/json" --data '${"query": { "match": { "symbol": "AGT" } } }'
```

Example Request - simulate_pipeline_conditional_execution_null_safe_operator

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "pipeline": {
    "description": "calculate the range of high and low and divide by the close price",
    "processors": [
      {
        "set": {
          "field": "new_id",
          "value": "{{id}}",
          "if": "ctx.parent?.child != null"
        }
      }
    ],
    "docs": [
      { "_source": { "id": 1, "parent": { "child": "value" } } },
      { "_source": { "id": 2, "parent": { } } },
    ]
  }
}
```

```
        {"_source": {"id":3}}
    ]
}
```

```
curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --header "Content-Type: application/json" --header "accept: application/json" --data '$'{\"pipeline\": { \"description\": \"calculate the range of high and low and divide by the close price\", \"processors\": [ {\"set\": { \"field\": \"new_id\", \"value\": \"{id}\" }, \"if\": \"ctx.parent?.child != null\" } } ], \"docs\": [ {\"_source\": {\"id\":1, \"parent\": {\"child\": \"value\"}}, {\"_source\": {\"id\":2, \"parent\": {\"_source\": {\"id\":3}} } ] }'
```

Example Request - define_for_market_cap_null_processor

PUT http://localhost:9200/_ingest/pipeline/for_market_cap_null_processor

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "description": "for_market_cap_null_processor",
  "processors": [
    { "set": {
      "field": "market_cap_type",
      "value": "no_market_cap ETF"
    }
  ]
}
```

```
curl --location --request PUT
"http://localhost:9200/_ingest/pipeline/for_market_cap_null_processor" --header
"Content-Type: application/json" --header "accept: application/json" --data
'$'{\"description\": \"for_market_cap_null_processor\",
\"processors\": [ {\"set\": { \"field\": \"market_cap_type\", \"value\":
\"no_market_cap ETF\" } } ] }'
```

Example Request - define_for_market_cap_processor

PUT http://localhost:9200/_ingest/pipeline/for_market_cap_processor

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
```

```

    "description": "for_market_cap_processor",
    "processors": [
      { "set": {
        "field": "market_cap_type",
        "value": "market_cap ETF"
      } }
    ]
  }
}

```

```

curl --location --request PUT
"http://localhost:9200/_ingest/pipeline/for_market_cap_processor" --header "Content-
Type: application/json" --header "accept: application/json" --data
$'{\"description\": \"for_market_cap_processor\",
\\\"processors\\\": [ { \\\"set\\\": { \\\"field\\\": \\\"market_cap_type\\\", \\\"value\\\":
\\\"market_cap ETF\\\" } } ] }'

```

Example Request - symbo_bil_market_cap_null_etf

GET http://localhost:9200/cf_etf/_search?q=symbol:BIL

HEADERS

Content-Type
application/json
accept
application/json

PARAMS

q
symbol:BIL

```

curl --location --request GET "http://localhost:9200/cf_etf/_search?q=symbol:BIL" --
header "Content-Type: application/json" --header "accept: application/json"

```

Example Request - simulate_pipeline_conditional_execution_switch_case

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```

{
  "pipeline": {
    "description": "calculate the range of high and low and divide by the
close price",
    "processors": [
      { "pipeline": {
        "name": "for_market_cap_null_processor",
        "if": "ctx.market_cap == null"
      } }
    ]
  }
}

```

```

        }},
        {"pipeline":{
            "name": "for_market_cap_processor",
            "if": "ctx.market_cap != null"
        }}
    ]
},
"docs" :[
    {"_source": {"id":1, "market_cap": null}},
    {"_source": {"id":2, "market_cap": "large"}}
]
}

```

```

curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{\"pipeline\": { \"description\": \"calculate the range of high and low and divide by
the close price\", \"processors\": [ {\"pipeline\": { \"name\":
\\\"for_market_cap_null_processor\\\", \\\"if\\\": \\\"ctx.market_cap == null\\\" }},
{\\\"pipeline\": { \"name\": \\\"for_market_cap_processor\\\", \\\"if\\\": \\\"ctx.market_cap !=
null\\\" } } ] }, \\\"docs\\\" : [ {\\\"_source\\\": {\\\"id\\\":1, \\\"market_cap\\\": null}},
{\\\"_source\\\": {\\\"id\\\":2, \\\"market_cap\\\": \\\"large\\\"}} ] }'

```

Example Request - simulate_pipeline_hit_exception_error

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```

{
  "pipeline": {
    "description": "test of hitting the exception error",
    "processors": [{
      "remove": {
        "field": "rating"
      }
    }]
  },
  "docs" :[
    {"_source": {"id":3}}
  ]
}

```

```

curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --
header "Content-Type: application/json" --header "accept: application/json" --data
$'{\"pipeline\": { \"description\": \" test of hitting the exception error\",

```

```
\\"processors\\": [{ \\"remove\\": { \\"field\\": \\"rating\\"}}] }, \\"docs\\" : [ {\\\"_source\\": {\\\"id\\":3}} ] }'
```

Example Request - simulate_pipeline_on_failure_hit_exception_error

POST http://localhost:9200/_ingest/pipeline/_simulate

HEADERS

Content-Type
application/json
accept
application/json

BODY

```
{
  "pipeline": {
    "description": "test parameter on_failure when error occurs",
    "processors": [{
      "remove": {
        "field": "rating",
        "on_failure": [
          { "set": { "field": "error", "value": -1 } },
          { "set": { "field": "message", "value": "remove processor failure: { { _ingest.on_failure_message } }" } }
        ]
      }
    ]
  },
  "docs" : [
    { "_source": { "id": 3 } }
  ]
}
```

```
curl --location --request POST "http://localhost:9200/_ingest/pipeline/_simulate" --header "Content-Type: application/json" --header "accept: application/json" --data '${\\"pipeline\\": { \\"description\\":\\"test parameter on_failure when error occurs\\", \\"processors\\": [{ \\"remove\\": { \\"field\\": \\"rating\\", \\"on_failure\\": [ {\\\"set\\": { \\"field\\": \\"error\\", \\"value\\": -1}}, {\\\"set\\": { \\"field\\": \\"message\\", \\"value\\":\\"remove processor failure: { { _ingest.on_failure_message } }\\\"}} ] } ] }, \\"docs\\" : [ {\\\"_source\\": {\\\"id\\":3}} ] } }'
```