

参考 https://blog.csdn.net/fwi_ntu/article/details/87863788

```
import csv
from elasticsearch import Elasticsearch
from datetime import timedelta, datetime

yesterday = datetime.today() + timedelta(-1)
index_name = yesterday.strftime('monitor-%Y.%m.%d')
starttime=datetime.now()

es = Elasticsearch(hosts="http://192.168.101.203:9200/", http_auth=('big_data_query','big_data_query'))
query_json = {"query":{"match_all":{}}}
query = es.search(index=index_name,body=query_json,scroll='5m',size=1000)

results = query['hits']['hits'] # es查询出的结果第一页
total = query['hits']['total'] # es查询出的结果总量
scroll_id = query['_scroll_id'] # 游标用于输出es查询出的所有结果

for i in range(0, int(total/1000)+1):
    # scroll参数必须指定否则会报错
    query_scroll = es.scroll(scroll_id=scroll_id,scroll='5m')['hits']['hits']
    results += query_scroll

with open('/tmp/esdata/'+index_name+'.csv','w',newline='',encoding='utf-8') as flow:
    csv_writer = csv.writer(flow)
    for res in results:
        sponsor= "" if not res['_source'].get("sponsor") else res['_source']['sponsor']
        traceLogId= "" if not res['_source'].get("traceLogId") else res['_source']['traceLogId']
        offset= "" if not res['_source'].get("offset") else res['_source']['offset']
        level= "" if not res['_source'].get("level") else res['_source']['level']
        txnSubType= "" if not res['_source'].get("txnSubType") else res['_source']['txnSubType']
        input_type= "" if not res['_source'].get("input_type") else res['_source']['input_type']
        requestIp= "" if not res['_source'].get("requestIp") else res['_source']['requestIp']
        txnType= "" if not res['_source'].get("txnType") else res['_source']['txnType']
        source= "" if not res['_source'].get("source") else res['_source']['source']
        type= "" if not res['_source'].get("type") else res['_source']['type']
        resultDesc= "" if not res['_source'].get("resultDesc") else res['_source']['resultDesc']
        result= "" if not res['_source'].get("result") else res['_source']['result']
        timeLength= "" if not res['_source'].get("timeLength") else res['_source']['timeLength']
        _timestamp= "" if not res['_source'].get("@timestamp") else res['_source']['@timestamp']
        csv_writer.writerow([res['_id']+'#+sponsor'+#+traceLogId+'#+str(offset)+'#+
level+'#+txnSubType+'#+input_type+'#+requestIp+'#+txnType+'#+source+'#+type+'#+resultDesc+'#+result+'#+
str(timeLength)+'#+_timestamp])

endtime=datetime.now()
print('done!' + str((endtime-starttime).seconds))
```