

PROCEDIMENTO OPERACIONAL
PADRÃO – POP PARA DOWNLOAD
DOS DADOS DO E-SUS NOTIFICA
UTILIZANDO O *LOGSTASH*

Utilizando o *Logstash* para realização do download dos dados do e-SUS Notifica

Utilizando o Logstash para realização do download dos dados do e-SUS Notifica

e-SUS Notifica

O e-SUS Notifica é um sistema de informação desenvolvido pela Secretária de Vigilância em Saúde (SVS), com o apoio do Departamento de Informática do SUS (DATASUS), para registro das notificações dos casos suspeitos e confirmados da doença causada pelo novo Coronavírus, Covid-19. No e-SUS Notifica são registrados todos os casos de síndrome gripal leve à moderados que se enquadram na definição de caso, para acompanhamento pelas secretarias estaduais e municipais, bem como pelo Ministério da Saúde por meio da Secretaria de Vigilância em Saúde (SVS).

De acordo com o Regulamento Sanitário Internacional - RSI, por estarmos em uma Pandemia e no Brasil, como nação signatária seguimos os protocolos do RSI, a epidemia foi classificada pelo Ministério da Saúde como Emergência de Saúde Pública Nacional (ESPN) após avaliação da gravidade da situação da saúde no país. Dado a declaração pela OMS de Emergência de Saúde Pública Internacional (ESPIN), e a declaração pelo Ministério da Saúde de ESPN, os casos suspeitos ou confirmados de Covid-19 devem ser notificados ao ministério da saúde através do sistema e-SUS Notifica em até 24 horas para acompanhamento e consolidação dos dados nacional e a resposta oportuna em termos de gestão em saúde.

Os dados do e-SUS são analisados por gestores estaduais e municipais diariamente para a divulgação dos números oficiais que indicam a quantidade de novos casos surgiram no país, bem como para o acompanhamento da situação epidemiológica do Covid-19.

INTERFACE DE PROGRAMAÇÃO DE APLICAÇÃO – API

Uma Interface de Programação de Aplicação, ou do inglês *Application Programming Interface* – API é um conjunto específico funcionalidades estabelecidas pelos desenvolvedores de uma aplicação, onde é possível a utilização de funções apenas por meio de programação. Desta forma, é possível se utilizar de características e recursos mais específicos de um sistema ou software, recursos esses muitas vezes menos acessíveis a seus utilizadores em geral.

As API's são formas de se disponibilizar consultas a dados ou acesso às informações de um sistema, podendo em alguns casos a inclusão de dados, ou até mesmo permitir que desenvolvedores criem funcionalidades através da API. Deste modo, a API torna-se um mecanismo muito popular para a disseminação de dados de sistemas de informações. Através da API é possível, dentre outras coisas, acessar os dados de um sistema de maneira automatizada, sem a necessidade de navegar dentro do sistema ou utilizar suas funcionalidades internas.

A Secretaria de Vigilância em Saúde – SVS juntamente como Departamento de Informática do SUS - Datasus, disponibilizou aos estados e municípios com mais de 300 mil habitantes o acesso à API para consulta e download dos dados, do sistema e-SUS Notifica, onde são registrados os casos de Covid-19 no Brasil, essa API visa dar acesso aos estados e municípios aos dados que são de sua gestão para que os mesmos possam analisar a situação epidemiológica auxiliando-os na tomada de decisão e iniciativas relacionadas aos casos de Covid-19.

LOGSTASH

O Logstash é uma ferramenta open source (gratuita e de código aberto) de tratamento e processamento de dados do lado servidor, capaz de realizar a importação, filtragem e o tratamento de dados das mais diversas fontes, padronizando-os e exportando-os para as mais diversas ferramentas de armazenamento e de gerenciamento de dados, a mais comumente utilizada é o *Elasticsearch*, servindo como fonte de integração de dados.

O Logstash faz parte do pacote ELK Stack, pacote este que contempla ainda os seguintes produtos: *Elasticsearch*, que é um produto de busca e análise, o *Kibana* é uma ferramenta de visualização de dados através de gráficos e diagramas e o *Logstash* é um pipeline de processamento de dados do lado servidor, desenvolvido para incorporar dados das mais diversas fontes de dados simultaneamente enviando-os mais comumente ao *Elasticsearch*.

DOWNLOAD DOS DADOS DO E-SUS NOTIFICA

INSTALAÇÃO E CONFIGURAÇÃO DO JAVA

Para a realização do download dos dados do e-SUS Notifica através do *Logstash* é necessário realizar o download da versão 7.13 da ferramenta que está disponível no seguinte link: <https://www.elastic.co/downloads/past-releases/logstash-7-13-0>, é estritamente recomendada o uso desta versão do Logstash. Ao finalizar o download teremos um arquivo no formato .zip que dever ser descompactado para a utilização, por se tratar de um programa stand-alone não necessita de instalação do programa, e o mesmo pode ser utilizado assim que o conteúdo do arquivo for extraído numa pasta de fácil acesso em seu computador.

O único pré requisito para a utilização do *Logstash* é a instalação do JAVA, e recomendamos a instalação do JDK, recomendamos a instalação da versão 16.0.2 do Java JDK, disponível no seguinte link: <https://www.oracle.com/java/technologies/javase/16-0-2-relnotes.html>. Após a instalação do JAVA JDK, é necessário realizar o procedimento de configuração da variável de ambiente. A variável de ambiente é a configuração que mapeia a instalação do JAVA para que o mesmo funcione adequadamente quando

requisitado pelos mais diversos programas que o utilizam, tais como o Logstash, o *Elasticsearch* o *Kibana* entre outros.

Para a configuração da variável de ambiente no Windows podemos realizar utilizando o prompt de comando pressionando a tecla Windows, digitar “cmd” sem aspas e pressionar enter, será apresentada a tela de prompt de comando.

Na tela do prompt de comando podemos executar o seguinte comando: “setx -m JAVA_HOME "C:\Program Files\Java\jdk-xx.x.x"” substituindo o “xx.x.x” pela versão instalada em seu computador.

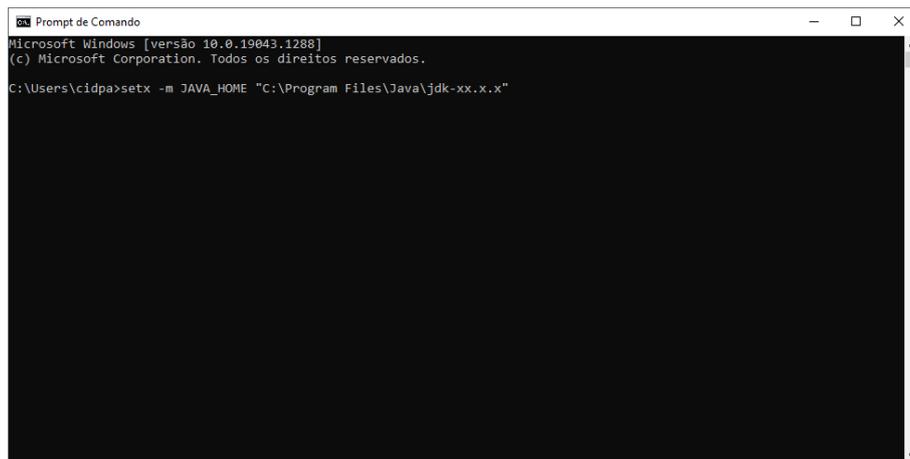


Figura 1 - Prompt de comando do Windows

Outra forma de realizar a configuração da variável de ambiente JAVA_HOME, pressionar a tecla Windows ou clicar na barra de pesquisa do Windows e digitar “variáveis” sem aspas, será aberta uma janela de propriedades do sistema conforme a figura 2.

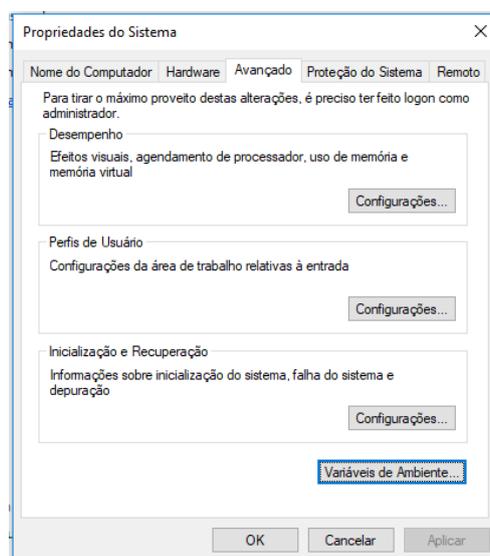


Figura 2 - Propriedades do Sistema

Ao clicar na opção “Variáveis de Ambiente” será apresentada uma tela com as variáveis configuradas no sistema, conforme apresentado na figura a seguir.

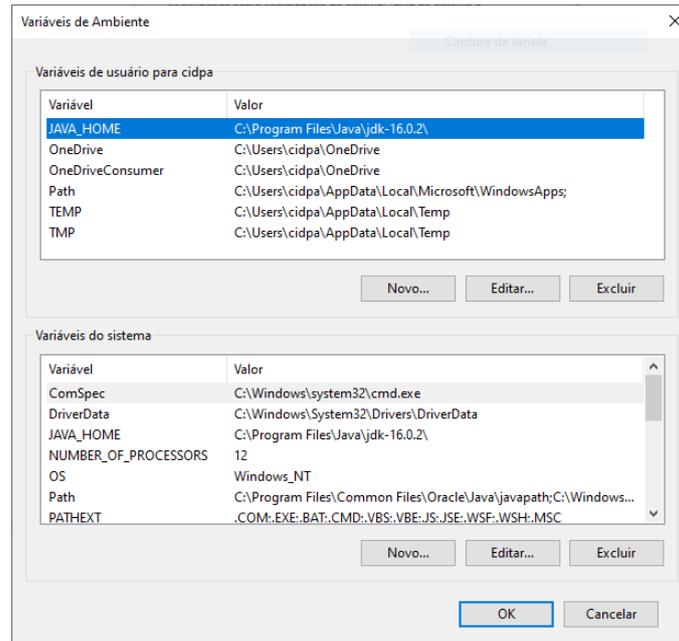


Figura 3 - Variáveis de Ambiente

Em variáveis de ambiente, na parte inferior da tela clique na opção “Novo...” para inserir uma nova variável de ambiente no sistema Windows., será apresentada uma nova janela denominada “Nova Variável de Sistema”.

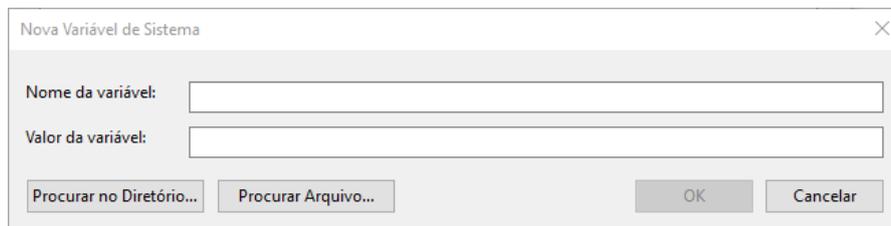


Figura 4 - Nova Variável de Sistema

No primeiro campo “Nome da variável” preencher com “JAVA_HOME” e no segundo campo, preencher com o caminho para a instalação do JAVA

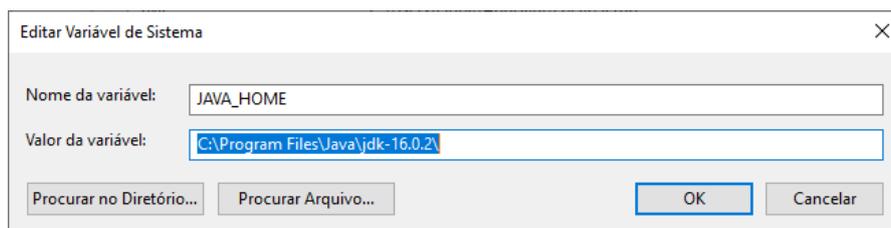


Figura 5 - Editando Variável de Sistema

Ao finalizar basta clicar no botão “OK” e reiniciar o computador para que as alterações sejam aplicadas.

UTILIZANDO O LOGSTASH

Após a configuração da variável de ambiente é necessário a realização do download do Logstash e a criação do script para a o download e tratamento dos dados do e-SUS. Com a configuração realizada da instalação do JAVA e a configuração. A versão do Logstash recomendada é a 7.13.0 disponível para download no seguinte link: <https://www.elastic.co/downloads/past-releases/logstash-7-13-0>. Clicando na opção “Windows” se inicia o download de um arquivo com o nome “logstash-7.13.0-windows-x86_64.zip”.

Após o download do arquivo, o próximo passo é descompactar o arquivo em uma pasta de fácil acesso, recomendável utilizar na pasta “Download” ou “Meus Documentos” ou uma pasta de sua preferência.

Esta versão do Logstash não requer demais configurações específicas, após descompactar o mesmo está pronto para o uso, sendo necessário apenas a criação do arquivo utilize o “Winrar” (<https://www.winrar.com/start.html?&L=9>) ou outro programa de sua preferência para descompactar o arquivo no formato zip obtido através do download. Com o Winrar instalado, basta clicar com botão direito do mouse, e clicar em “Extrair aqui”.

Após descompactar o arquivo, será criado um diretório denominado “logstash-7.13.0”, é recomendável renomear a pasta para “logstash”, após o final do processo teremos um diretório denominado “logstash” com a seguinte estrutura de pastas e arquivos.

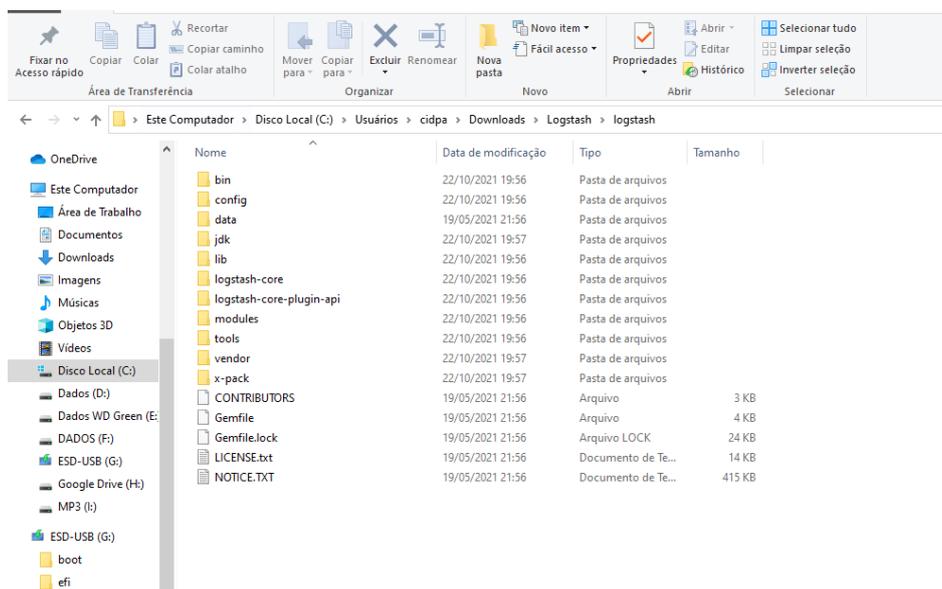


Figura 6 - Estrutura de arquivos do Logstash

No próximo passo vamos criar um novo diretório denominado “conf” para salvar o nosso script que realiza o download dos dados do e-esus.

Ao final teremos a seguinte estrutura de pastas e arquivos no diretório “logstash”.

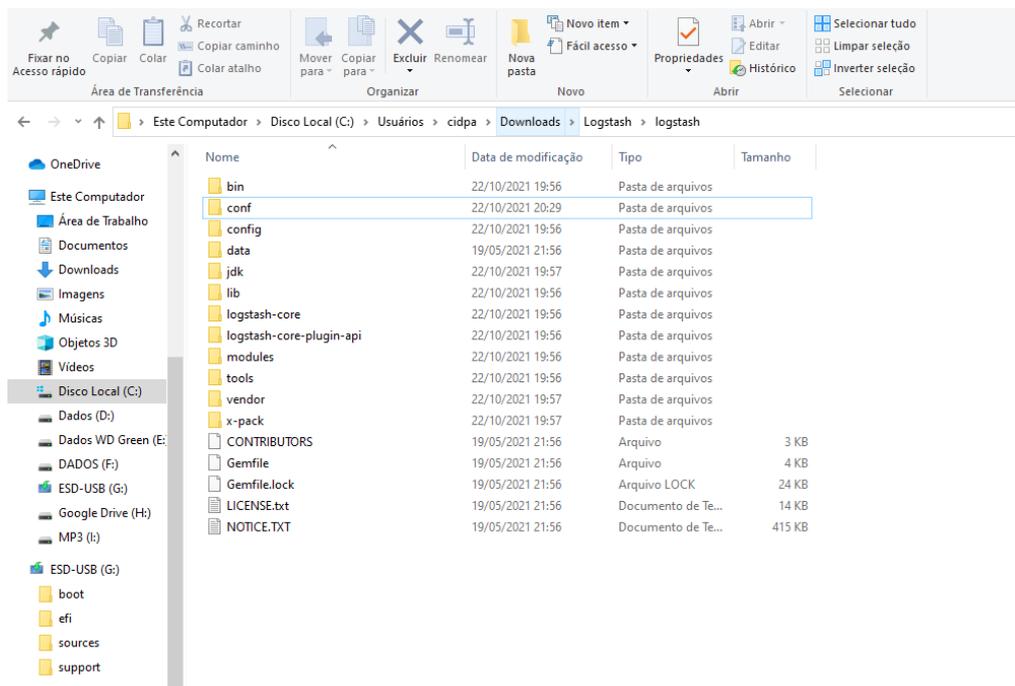


Figura 7 - Estrutura final do Logstash

Utilizando o “Sublime text” <https://www.sublimetext.com/>, ou um editor de texto de sua preferência, crie um novo arquivo e salve o mesmo na pasta “conf”, criada no passo anterior, com o seguinte nome e extensão: “download-esus.conf”. A extensão do arquivo deve ser “.conf” para que o script funcione adequadamente.

O script é dividido em dois blocos “input” e “output”. O bloco “input” faz a conexão com a fonte de dados sendo ela um banco de dados, uma API ou até mesmo um arquivo nos mais diversos formatos. O bloco “output” faz a exportação dos dados para as mais variadas formas de armazenamento de dados tais como: Elasticsearch, arquivos nos formatos .csv xls xlsx, e os mais diversos Sistemas de Gerenciamento de Banco de Dados SGBD.

No trecho input vamos inserir as informações de login, tais como usuário e senha, e ainda o “index” e a “query” para realizar a consulta nos dados, sendo possível uma busca por todos os dados, por período ou mesmo por um atributo específico. O trecho de código do bloco input apresentado faz a busca de todos os dados do estado do Acre notificados no e-SUS Notifica que foram publicados no OpenDatasus.

```
input {
  elasticsearch {
    user => "user"
    password => "senha"
    hosts => ["https://elasticsearch-saps.saude.gov.br"]
    index => "desc-notificacoes-esusve-ac"
    size => 10000
```

```

scroll => "5m"
codec => "json"
docinfo => true
query => '{"query": {"match_all": {}}}'
}
}

```

O segundo trecho do código apresentando é o trecho do output, e no script vamos exportar os dados para um arquivo “.csv”, entre as opções disponíveis podemos personalizar o separador do arquivo csv utilizando a opção “col_sep”, neste caso vamos utilizar o ponto e virgula “;”, indicamos também o caminho onde deverá ser salvo o arquivo, utilizando a opção “path” e as colunas que desejamos que sejam exportadas para o arquivo csv utilizando a opção “fields”.

```

output {
  csv {
    csv_options => {"col_sep" => ";"}
    path => "E:/Esus/Opendatasus/esusve-ac.csv"
    fields => ["dataNotificacao", "dataInicioSintomas", "dataTeste", "estrangeiro",
"profissionalSaude", "cbo", "sexo", "estado", "municipio", "sintomas", "outrosSintomas",
"condicoes", "estadoTeste", "tipoTeste", "resultadoTeste", "estadoNotificacao",
"municipioNotificacao", "dataRegistro", "dataAtualizacao", "source_id", "idade", "classificacaoFinal",
"evolucaoCaso", "dataEncerramento", "testeSorologico", "dataTesteSorologico",
"tipoTesteSorologico", "resultadoTesteSorologicoIgA", "resultadoTesteSorologicoIgG",
"resultadoTesteSorologicoIgM", "resultadoTesteSorologicoTotais", "profissionalSeguranca",
"racaCor", "@timestamp", "estadoIBGE", "estadoNotificacaoIBGE", "municipioIBGE",
"municipioNotificacaoIBGE"]

  }
  stdout { codec => rubydebug }
}

```

Na exportação para o formato “.csv” apresentada neste exemplo, apesar de listarmos os campos que devem ser exportados para o arquivo, o *Logstash* não cria a primeira linha com o nome das colunas, sendo recomendado neste caso, deixar um arquivo com o mesmo local e nome apresentado na opção “path” do trecho output do código, com a primeira linha contendo os campos e com o mesmo separador definido no script, deste modo o script identifica o arquivo no local indicado e insere os dados a partir da segunda linha.

Para realizar a execução do script e realizar o download dos dados devemos utilizar o prompt de comando ou o Power Shell do Windows 10, é recomendável a utilização deste pelas facilidades que o mesmo oferece.

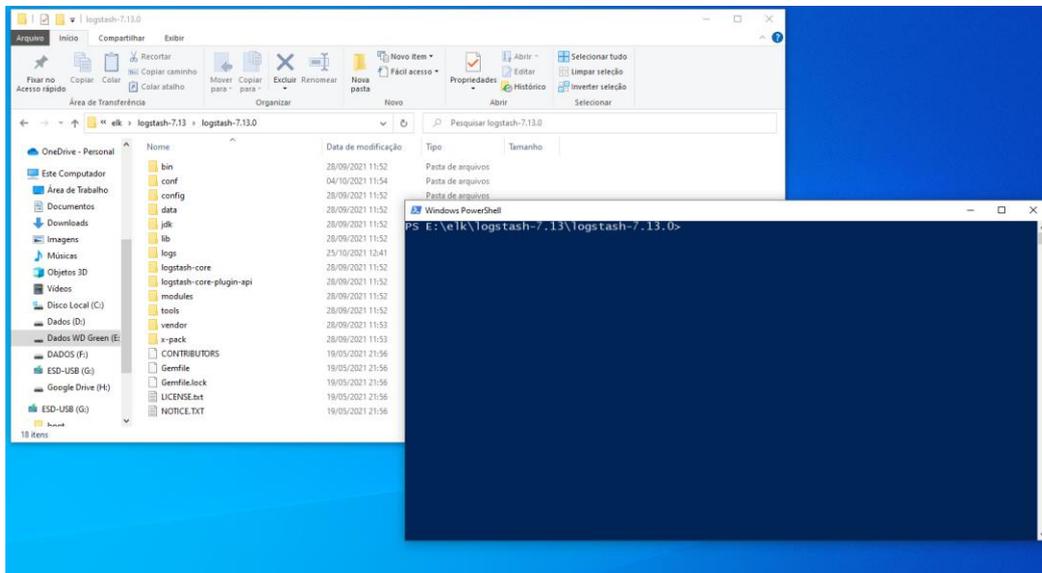


Figura 8 - Utilizando o Power Shell pra rodar o Script

Para a execução do script vá até o diretório raiz onde se encontram os arquivos baixados do *Logstash*, o comando para a execução do script tem a seguinte sintaxe, local do arquivo de execução do Logstash, o mesmo por padrão fica na pasta bin/ e para a versão do Windows é o "logstash.bat", o segundo parâmetro para a execução é o "-f" onde indicamos ao logstash que o mesmo deverá executar um script em um arquivo, e o terceiro parâmetro é o arquivo do tipo "conf" que apresentamos anteriormente denominado "download-esus.conf", por padrão recomendamos que o mesmo seja salvo na pasta "conf/". Deste modo a sintaxe completa para a execução do script seria a seguinte: `.\bin\logstash.bat -f .\conf\download-esus.conf`.

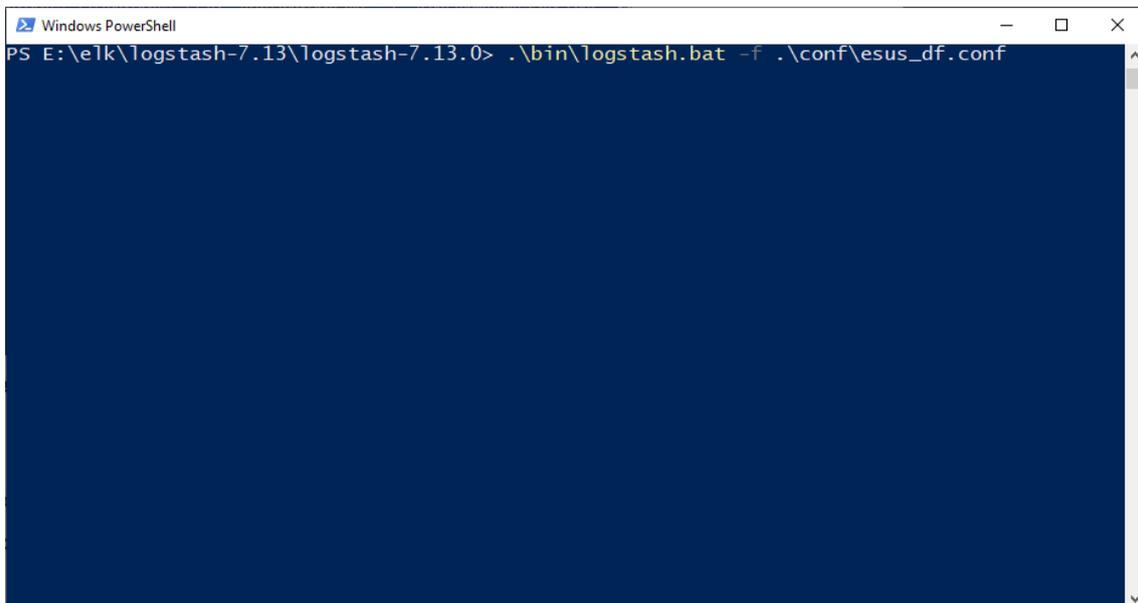


Figura 9 - Comando para a execução do script de download dos dados

Caso não seja encontrado nenhum erro e o script seja executado corretamente será apresentada uma tela demonstrando o processo de download dos dados conforme apresentado na figura a seguir.

```
Windows PowerShell
"dataInicioSintomas" => "2020-05-06T03:00:00.000Z",
"outrasCondicoes" => nil,
"dataNotificacao" => "2020-05-06T20:26:15.774Z",
"localRealizacaoTestagem" => nil,
"outroBuscaAtivaAssintomatico" => nil,
"codigoEstrategiaCovid" => nil,
"idade" => 50,
"@timestamp" => 2021-08-26T19:39:36.666Z,
"id" => "S4lvPTRC4U",
"condicoes" => nil,
"codigoLocalRealizacaoTestagem" => nil,
"municiplioNotificacaoIBGE" => "5300108",
"estadoTeste" => "Concluido",
"racaCor" => "Parda",
"tipoTeste" => "TESTE RÁPIDO - ANTICORPO",
"@version" => "1",
"estrategiaCovid" => nil,
"estadoNotificacao" => "Distrito Federal",
"municiplioIBGE" => "5300108",
"municiplioNotificacao" => "Brasilia",
"dataEncerramento" => "2020-05-09T03:00:00.000Z",
"outroTriagemPopulacaoEspecific" => nil,
"dataNascimento" => "1971-03-14T03:00:00.000Z",
"paisOrigem" => nil,
"estadoNotificacaoIBGE" => "53",
"resultadoTeste" => "Negativo",
"registroAtual" => true
}
```

Caso o download esteja sendo executado corretamente será apresentada a tela com os campos e o seu respectivo conteúdo, a depender das configurações do computador que está executando o script pode ser que não seja possível ler as informações na tela, pois as mesma são atualizadas rapidamente a cada registro baixado, bastando aguardar a finalização e a exportação dos dados para o arquivo csv.

SCRIPT NO FORMATO TXT

```
input {
  elasticsearch {
    user => "usuario"
    password => "senha"
    hosts => ["https://elasticsearch-saps.saude.gov.br"]
    index => "desc-notificacoes-esusve-df"
    size => 10000
    scroll => "5m"
    codec => "json"
    docinfo => true

    query => '{"query": {"match_all": {}}}'
  }
}

output {
  csv {
    csv_options => {"col_sep" => ";"}
    path => "E:/Esus/Opendatasus/2021/esusve-df.csv"
    fields => ["dataNotificacao", "dataInicioSintomas", "dataTeste", "estrangeiro",
"profissionalSaude", "cbo", "sexo", "estado", "municipio", "sintomas", "outrosSintomas",
"condicoes", "estadoTeste", "tipoTeste", "resultadoTeste", "estadoNotificacao",
"municipioNotificacao", "dataRegistro", "dataAtualizacao", "source_id", "idade",
"classificacaoFinal", "evolucaoCaso", "dataEncerramento", "testeSorologico",
"dataTesteSorologico", "tipoTesteSorologico", "resultadoTesteSorologicoIgA",
"resultadoTesteSorologicoIgG", "resultadoTesteSorologicoIgM", "resultadoTesteSorologicoTotais",
"profissionalSeguranca", "racaCor", "@timestamp", "estadoIBGE", "estadoNotificacaoIBGE",
"municipioIBGE", "municipioNotificacaoIBGE"]
  }
  stdout { codec => rubydebug }
}
```