

Relatório do trabalho da disciplina de Integração e Sistemas de
Informação

Covid19 - Knime

Francisco Ferreira - 14061

João Oliveira- 16979

Licenciatura Engenharia Sistemas Informáticos

Novembro de 2020

Índice

INTRODUÇÃO	1
INTRODUÇÃO AO PROBLEMA	3
IMPORTAR OS DADOS	4
TRATAMENTO DE DADOS	6
GUARDAR DADOS FILTRADOS	9
HISTÓRICO CONSOANTE PAÍS SELECIONADO	11

Lista de Figuras

Figura 1 Dashboard Tratamento de dados	6
Figura 2 GET Request API summary	6
Figura 3 Ungroup	7
Figura 4 JsonPath Countries	7
Figura 5 Row Filter procurar Portugal	8
Figura 6 Dashboard guardar ficheiros em XML e JSON	9
Figura 7 Dashboard GET total sobre o país	11
Figura 8 String manipulation	12
Figura 9 GET Request com o output do string manipulation	13
Figura 10 Bar Chart com os dados recebido desde o primeiro dia	14
Figura 11 Dashboard Base Dados	15
Figura 12 Dados que foram inseridos na base dados	15
Figura 13 Boletim covid-19 no email	16

Lista de Fragmentos

Fragmento 1 Get do summary em JSON	4
Fragmento 2 de JSON guardado	9
Fragmento 3 Estrutura de XML Guardado	10

Introdução

Estamos a viver numa nova realidade, está a ser um ano atípico devido ao covid 19. Por ser algo recente e com muitos dados, e como o tema deste trabalho era à nossa escolha, decidimos utilizar a covid19 como tema. Utilizando assim as bases de dados e API'S que estão ao nosso alcance decidimos fazer o tratamento e filtragem de dados para aquilo que é o nosso objetivo.

Introdução ao problema

A Johns Hopkins University Center for Systems Science and Engineering tem uma base de dados que é pública, onde estão todos os casos desde o primeiro dia em que cada país teve um caso de covid. Ao todo esta base de dados tem 190 países e todo o histórico de covid nesse mesmo país. Como são muitos dados e não fazia sentido estar a exportar toda a base de dados e todos os csv, vamos utilizar uma api que se chama covid19api.

O Nosso objetivo é então filtrar toda a informação para só trazer informação de um país, neste caso Portugal, para depois ser guardado em XML e JSON.

No fim de toda a filtragem será enviado um email com todos os casos, como por exemplo os novos confirmados, as novas mortes, os novos recuperados, o total de confirmados, o total de mortes e por fim o total de recuperados.

Importar os dados

Nesta primeira fase importamos os dados da API com um Get deste link

<https://api.covid19api.com/summary>, que nos traz o resumo dos casos novos e do total por país atualizado diariamente.

```
{
  "Message": "",
  "Global": {
    "NewConfirmed": 665991,
    "TotalConfirmed": 52119182,
    "NewDeaths": 12240,
    "TotalDeaths": 1284482,
    "NewRecovered": 382236,
    "TotalRecovered": 33919628
  },
  "Countries": [{
    "Country": "Afghanistan",
    "CountryCode": "AF",
    "Slug": "afghanistan",
    "NewConfirmed": 146,
    "TotalConfirmed": 42609,
    "NewDeaths": 4,
    "TotalDeaths": 1581,
    "NewRecovered": 13,
    "TotalRecovered": 34967,
    "Date": "2020-11-12T17:27:30Z",
    "Premium": {}
  }],
}
```

Fragmento 1 Get do summary em JSON

Figura 1 JsonPath CountriesFragmento 2 Get do summary em JSON

Podemos concluir que com este GET que retornou o código 200 OK,

- 1) Global, em todo o mundo diariamente.
 - a) Casos novos confirmados (*NewConfirmed*)
 - b) Casos totais confirmados (*TotalConfirmed*)
 - c) Novas mortes (*NewDeaths*)
 - d) Total de mortes (*TotalDeaths*)
 - e) Novos recuperados (*NewRecovered*)

f) Total recuperados (*TotalRecovered*)

2) Casos por país diariamente.

- a) País (*Country*)
- b) Código de país (*CountryCode*)
- c) Novas mortes (*NewDeaths*)
- d) Total de mortes (*TotalDeaths*)
- e) Novos recuperados (*NewRecovered*)
- f) Total recuperados (*TotalRecovered*)
- g) Data (*Date*)

Depois de compreendermos os dados que recebemos esta na hora de os filtrar

Tratamento de dados

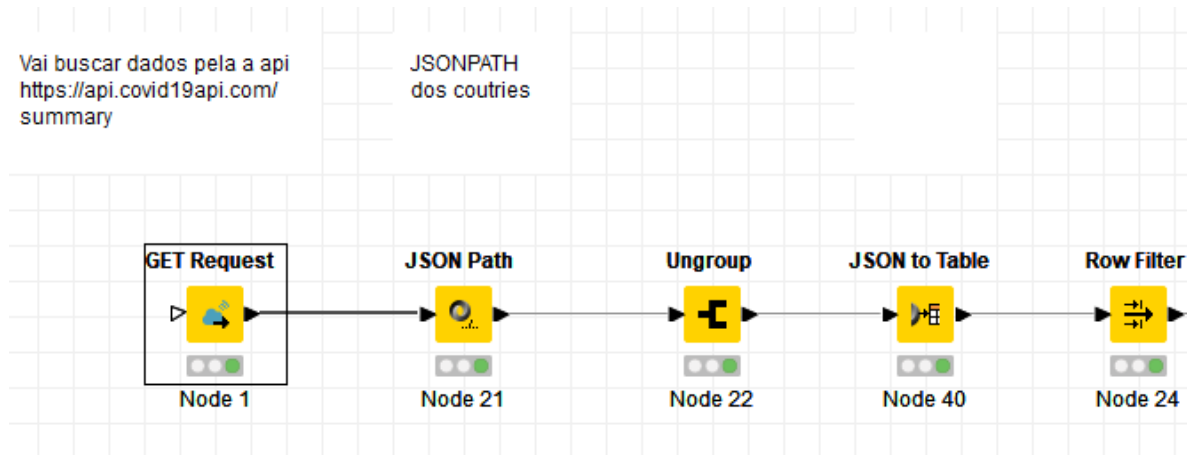


Figura 2 Dashboard Tratamento de dados

Utilizamos nó GET Request este só pode ser usado para emitir solicitações HTTP GET. As solicitações GET são usadas para pedir dados de um serviço web sem enviar outros dados além dos parâmetros de solicitados.

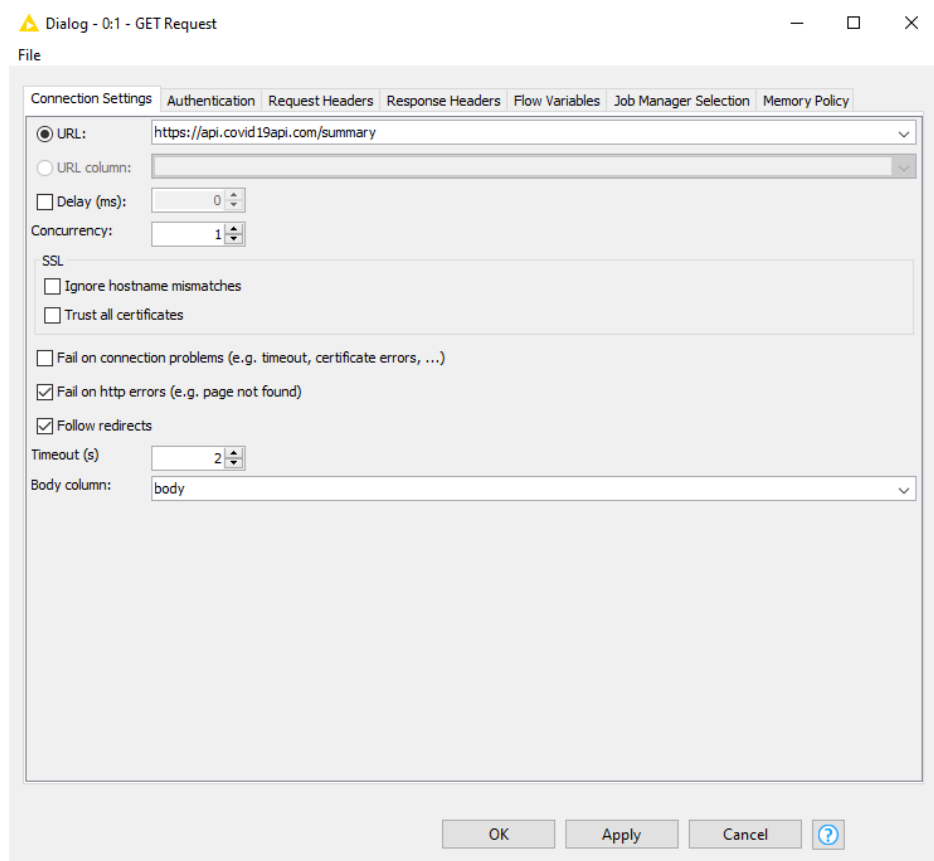


Figura 3 GET Request API summary

No Knime utilizamos o JSON Path para filtrar só os países, isto é, removemos a informação do estado atual globalmente sobre o covid19. Depois utilizamos o Ungroup que é para meter cada país da lista recebida numa linha.

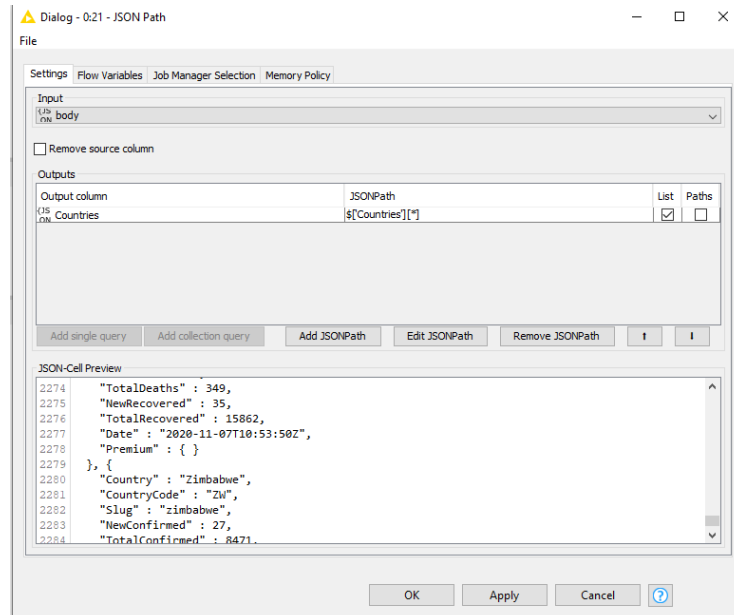


Figura 5 JsonPath Countries

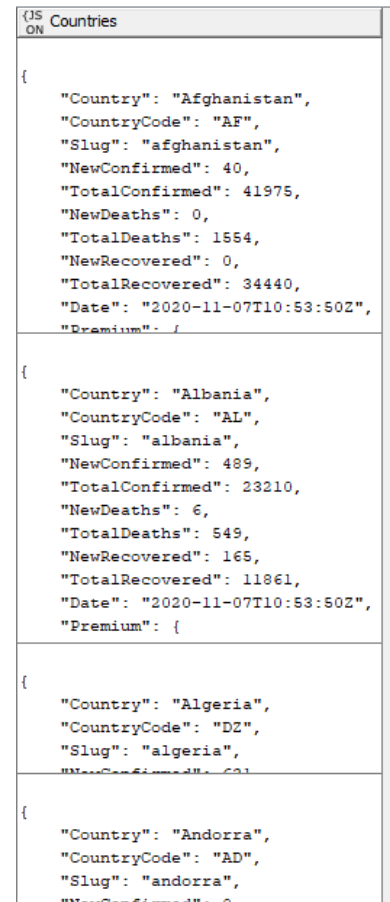


Figura 4 Ungroup

Depois de tudo filtrado iremos colocar o JSON para uma tabela o JSON to Table converte uma coluna JSON em várias colunas, por meio das quais a lista de colunas é extraída heurísticamente da estrutura JSON. Ele pode extrair os elementos folha primitivos (como strings e números), omitindo o caminho da árvore JSON ou a estrutura JSON completa. O último, no entanto, pode gerar uma saída confusa, pois os tipos das colunas são novamente JSON ou coleções de JSON.

Agora vamos a parte de filtrar para o país que queremos para isso utilizamos o Row Filter neste caso vamos seleccionar o Country e procurar Portugal, ira nos devolver a linha de Portugal e todos os dados contidos nele.

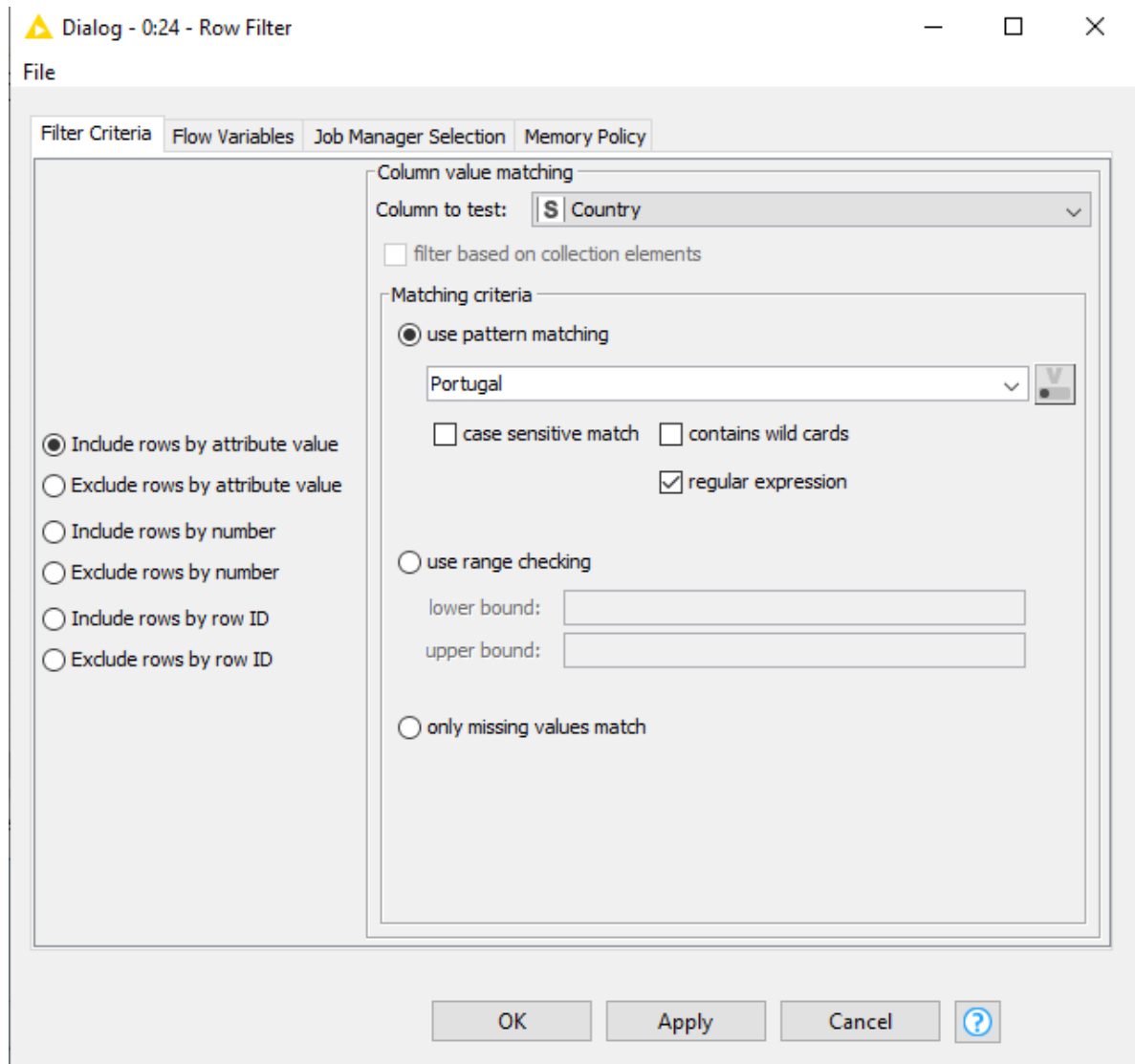


Figura 6 Row Filter procurar Portugal

Guardar dados Filtrados

Depois de já termos um país neste caso Portugal, vamos escrever para dois tipos de ficheiros JSON e XML.

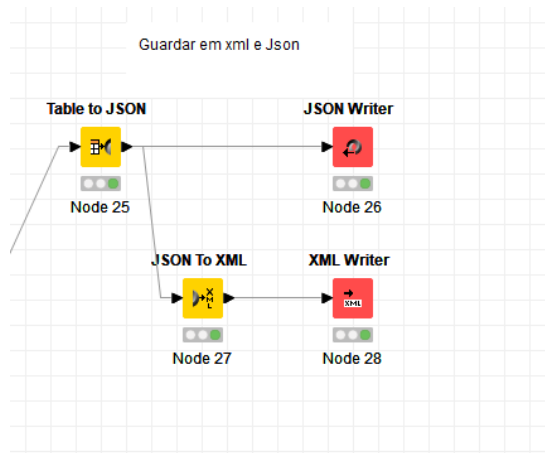


Figura 7 Dashboard guardar ficheiros em XML e JSON

```
[ {
  "Status" : 200,
  "Content type" : "application/json;
charset=UTF-8",
  "Country" : "Portugal",
  "CountryCode" : "PT",
  "Slug" : "portugal",
  "NewConfirmed" : 5550,
  "TotalConfirmed" : 166900,
  "NewDeaths" : 52,
  "TotalDeaths" : 2792,
  "NewRecovered" : 2301,
  "TotalRecovered" : 93754,
  "Date" : "2020-11-07T10:53:50Z"
} ]
```

Fragmento 3 de JSON guardado

```
<?xml version="1.0" encoding="UTF-8"?>

<root xmlns="portugal">
  <item>
    <Status>200</Status>
    <Contenttype ns:originalKey="Content type" xmlns:ns="http://www.knime
.org/json2xml/originalKey/">application/json; charset=UTF-8</Contenttype>
    <Country>Portugal</Country>
    <CountryCode>PT</CountryCode>
    <Slug>portugal</Slug>
    <NewConfirmed>5550</NewConfirmed>
    <TotalConfirmed>166900</TotalConfirmed>
    <NewDeaths>52</NewDeaths>
    <TotalDeaths>2792</TotalDeaths>
    <NewRecovered>2301</NewRecovered>
    <TotalRecovered>93754</TotalRecovered>
    <Date>2020-11-07T10:53:50Z</Date>
  </item>
</root>
```

Fragmento 4 Estrutura de XML Guardado

Histórico consoante país selecionado

Já com o país selecionado vamos fazer uma string manipulation, vamos acrescentar o país a este link `"https://api.covid19api.com/total/country/"` que é para irmos buscar todos os dados desde o primeiro dia que houve o primeiro registo de covid no país.

Dentro do string manipulation utilizamos esta expressão `string("https://api.covid19api.com/total/country/" + $Country$` e criamos uma coluna nova com o nome `GetAPIContry` nesta coluna ira ser guardado o link do get já com o país, exemplo:

`"https://api.covid19api.com/total/country/portugal"`

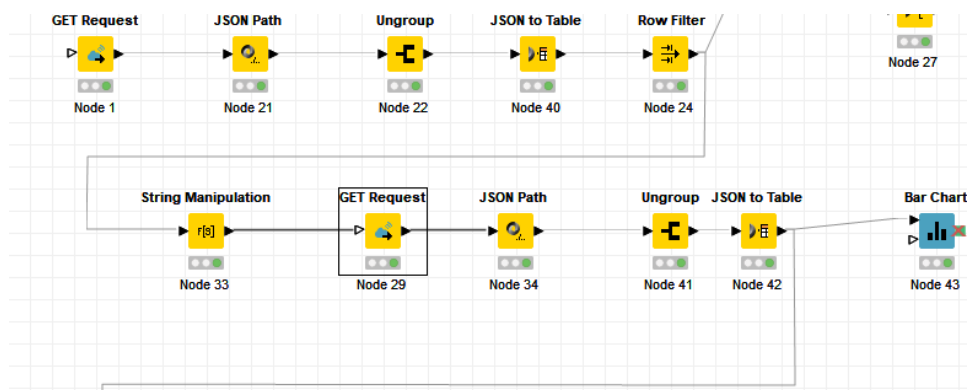


Figura 8 Dashboard GET total sobre o país

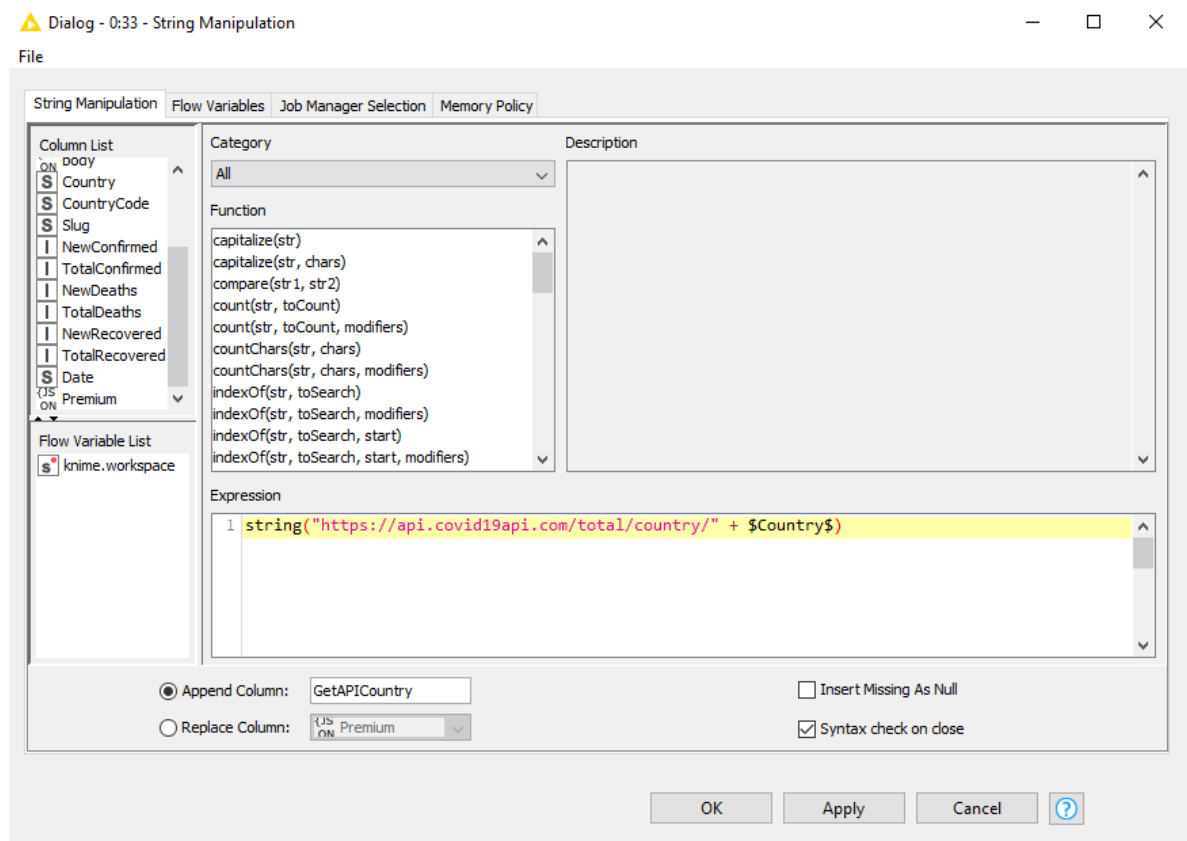


Figura 9 String manipulation

Depois de o link feito com a string manipulation utilizamos novamente o GET Request, mas desta vez com o URL Column com o GetAPICountry como mostra a seguinte figura.

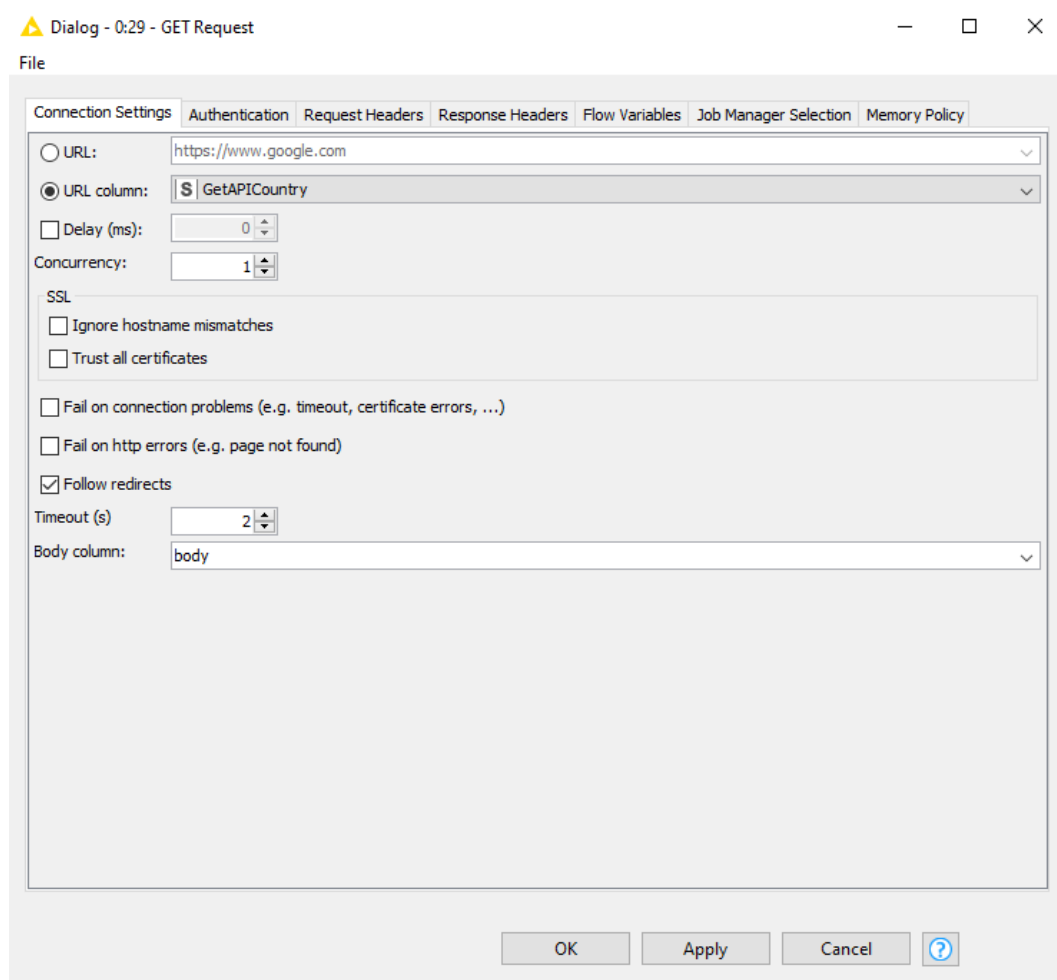


Figura 10 GET Request com o output do string manipulation

Depois repetimos o mesmo processo que nos passos anteriores. Um grande poder que o knime tem é poder criar gráficos consoante os dados que temos. Neste caso utilizei um BarChart com as mortes recovered e Active desde o primeiro dia.

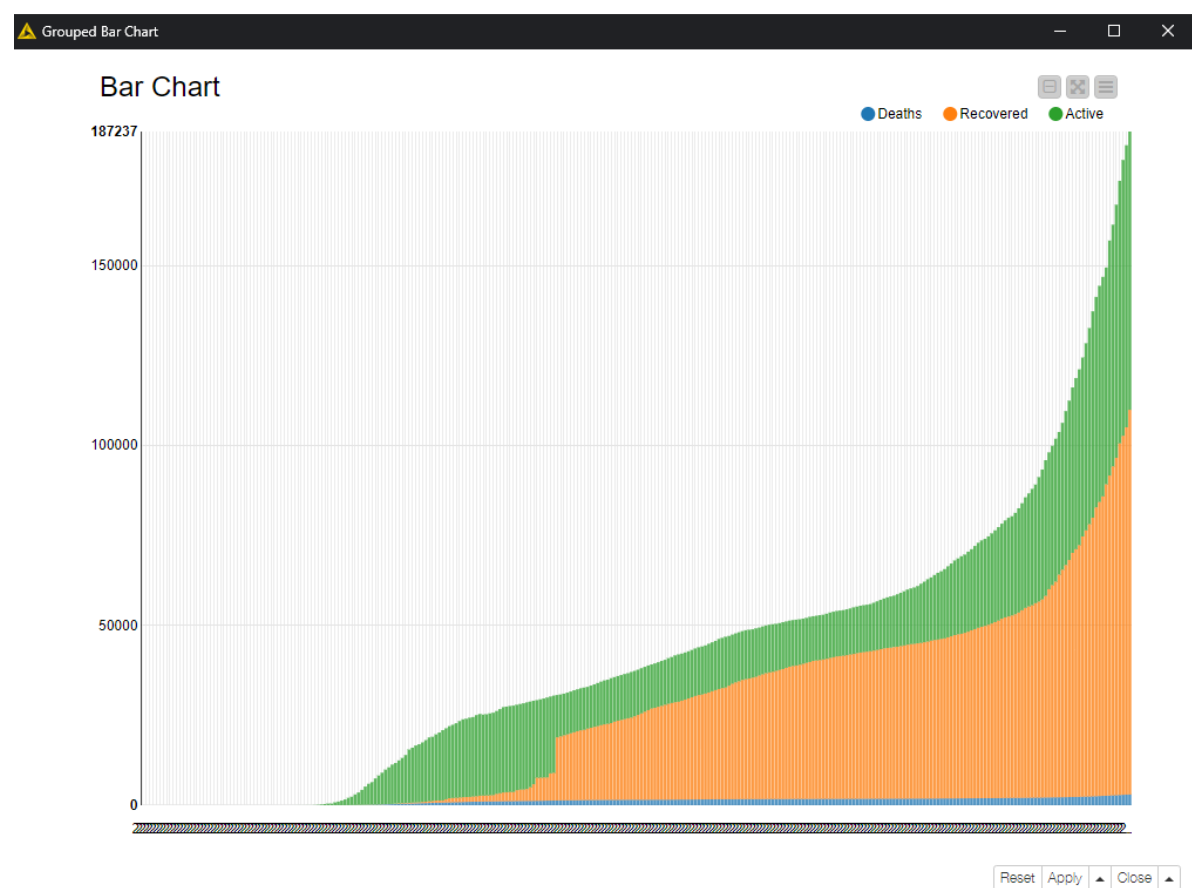


Figura 11 Bar Chart com os dados recebido desde o primeiro dia

A verde temos os casos ativos, a laranja os casos recuperados e a azul as mortes em Portugal.

Base de dados

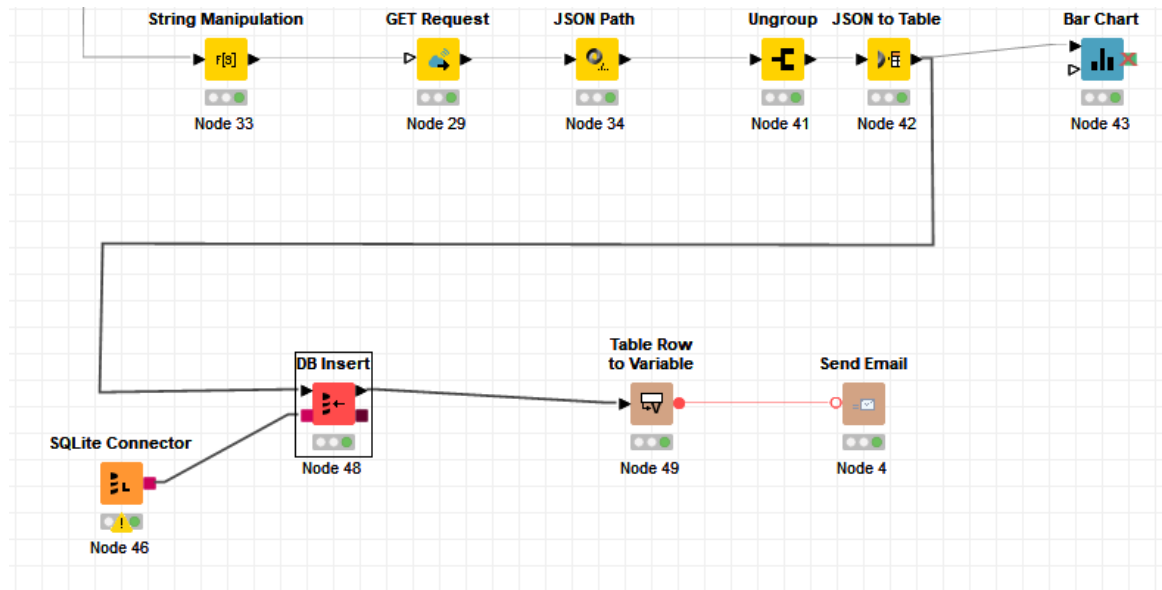


Figura 12 Dashboard Base Dados

Depois de ter em nossa mão estes dados todos filtrados decidimos fazer a nossa própria base de dados, criamos uma tabela que se chama país com o Nome do país, data, casos Ativos, casos recuperados, mortes e o Total confirmado.

Depois de criada conectamos a base dados no knime e fazemos o DB Insert para inserirmos os dados na tabela.

	Country	Date (#1)	Active	Recovered	Deaths	TotalConfirmed
	Filter	Filter	Filter	Filter	Filter	Filter
61	Portugal	2020-03-22T00:00:00Z	1581	5	14	166900
62	Portugal	2020-03-23T00:00:00Z	2032	5	23	166900
63	Portugal	2020-03-24T00:00:00Z	2307	22	33	166900
64	Portugal	2020-03-25T00:00:00Z	2930	22	43	166900
65	Portugal	2020-03-26T00:00:00Z	3441	43	60	166900
66	Portugal	2020-03-27T00:00:00Z	4149	43	76	166900
67	Portugal	2020-03-28T00:00:00Z	5027	43	100	166900
68	Portugal	2020-03-29T00:00:00Z	5800	43	119	166900
69	Portugal	2020-03-30T00:00:00Z	6225	43	140	166900

Figura 13 Dados que foram inseridos na base dados

Email

Depois da nossa base dados já conter dados filtrados com só o queremos esta na altura de enviamos um email com os Novos Confirmados, Novas mortes, Novos Recuperados, Total Confirmados, Total Mortes, Total Recuperados e terá este aspeto.

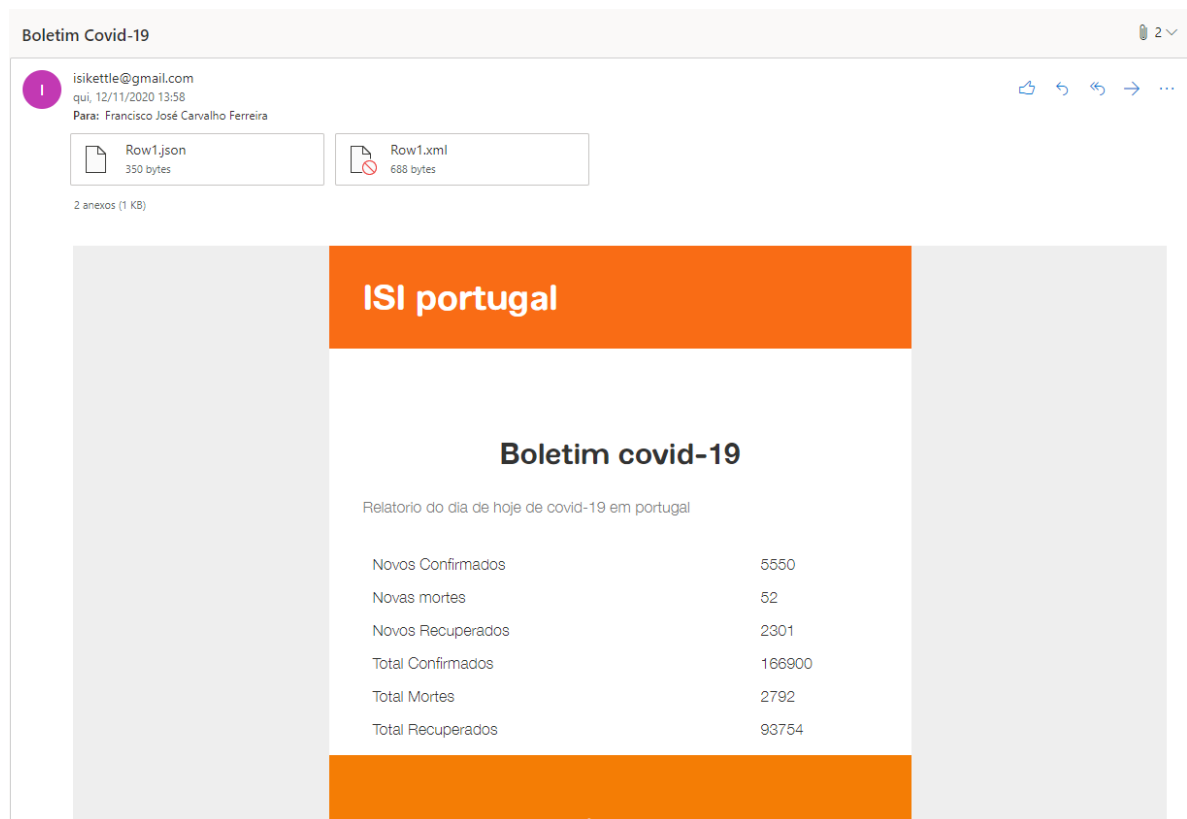


Figura 14 Boletim covid-19 no email

Conclusão

Com este trabalho conseguimos aprofundar a matéria dada na nesta unidade curricular, achamos também que o objetivo foi alcançado. Como fizemos o trabalho em grupo isso obrigou-nos a usar dois programas o que é muito bom pois existem comandos e maneiras diferentes de fazer as coisas, mesmo sendo o mesmo tema.

Referências

Forum Knime. (s.d.). Obtido de Forum Knime: <https://forum.knime.com/>

Hub Knime. (s.d.). Obtido de Hub Knime: <https://hub.knime.com/>

Knime. (s.d.). Obtido de Knime: <https://www.knime.com/>

NodePit. (s.d.). Obtido de NodePit: nodepit.com

Stripo. (s.d.). Obtido de Stripo Email: <https://stripo.email/>