



# Formação Engenheiro de Dados

Contagem

## D. Gerenciar diariamente o número de contratos de risco

- ❖ Contratos eletrônicos
- ❖ Os de risco possuem esta descrição

```
{
  "id" : ObjectId("5d3b13edd6b1f54a1234108a"),
  "idcliente" : "2",
  "idcontrato" : "22082",
  "datacadastro" : "15/01/2019",
  "contrato" : "CONTRATO DE LOCAÇÃO DE AUTOMÓVEL DE PRAZO DETERMINADO IDENTIFICAÇÃO DAS PARTES CONTRATANTES LOCADORA: RoboCar Tipo de Contrato: de risco
to e acertado o presente Contrato de Locação de Automóvel de Prazo Determinado, que se regerá pelas cláusulas seguintes e pelas condições descritas no presente
ação1 do automóvel de propriedade da LOCADORA. DO USO Cláusula 2ª. O automóvel, objeto deste contrato, será utilizado exclusivamente pelo LOCATÁRIO, não sendo
o da multa prevista na Cláusula 7ª.DA DEVOLUÇÃO Cláusula 3ª. O LOCATÁRIO deverá devolver o automóvel à LOCADORA nas mesmas condições em que estava quando o rec
ízos causados.Cláusula 5ª. Se o LOCATÁRIO não restituir o automóvel na data estipulada, deverá pagar, enquanto detiver em seu poder, o aluguel que a LOCADORA a
niente de caso fortuito3.DA RESCISÃO Cláusula 6ª. É assegurado às partes a rescisão do presente contrato a qualquer momento, desde que haja comunicação à outra
quer das cláusulas por parte dos contratantes ensejará a rescisão deste instrumento e o devido pagamento de multa, pela parte inadimplente no valor de R$ 5000,
s vias de igual teor, juntamente com 2 (duas) testemunhas."
```

# pyspark

◆ Contar apenas contratos de Risco

```
pyspark
pesquisa = sc.textFile("file:///home/cloudera/Downloads/contratos.json")
contagem = pesquisa.flatMap(lambda palavra: palavra.split(" ")).filter(lambda palavra: (palavra == "risco"))
contagem = contagem.map(lambda pal: (pal, 1))
contagem = contagem.reduceByKey(lambda a, b: a + b)
contagem.saveAsTextFile("risco")
quit()
hdfs dfs -ls /user/cloudera/risco/
hdfs dfs -cat /user/cloudera/risco/part-00000
```