

# **Verificação dos créditos de ICMS – Óleo diesel**

**Módulo: DATA ANALYTICS**

**Alunos: Letícia Álvares Barbalho Barbosa**

**Allan da Costa Andrade**

**Robson Mota Gonçalves**

**Marcelo Dias da Silva**

**Deibe Paiva Lima**

**Curso: Pós-Graduação em Big Data e Machine Learning**

**Turma: III**

**Professor: Marlesson Santana**

# Tema

Identificar indícios de sonegação relativos ao montante dos valores lançados a título de crédito do ICMS (Imposto sobre Circulação de Mercadorias e Prestação de Serviços de transporte e comunicação) na Escrituração Fiscal Digital (EFD) relativo à utilização do óleo diesel consumido em atividades industriais.

# Descrição do problema

Atentado para o montante dos valores lançados a título de crédito de ICMS relativos a aquisição de óleo diesel realizado pelas indústrias goianas, buscamos, com estes procedimentos, reconstruir o valor máximo que pode ser creditado com base na análise dos documentos eletrônicos emitidos relativos à aquisição de óleo diesel efetuados no período de 2014 a 2017.

Através da análise destes documentos eletrônicos (Nota Fiscal Eletrônica - NFE) determinaremos quantidade total de óleo diesel adquirida pelo estabelecimento industrial. Com este dado, poderemos calcular o montante do tributo incidente na operação de aquisição de óleo diesel, que representa o limite máximo de valor objeto de creditamento.

Com base nestes valores reconstituídos de ICMS incidente nas operações de aquisição de óleo diesel, compararemos com os valores lançados a crédito na Escrituração Fiscal Digital, buscando assim inconsistências e desproporções a serem verificados em futuro procedimento de auditoria, lembrando que quanto maior o valor lançado a crédito do imposto, menor o valor apurado e conseqüentemente menor o valor recolhido, causando assim prejuízos ao erário.

Atentando para o fato de que o montante do tributo incidente na operação de aquisição de óleo diesel não vem destacado no documento fiscal, devido ao fato de que a mercadoria (óleo diesel) estar sujeita ao regime de substituição tributária, deveremos calcular o tributo incidente utilizando o Preço Médio Ponderado a consumidor final (PMPF) e a alíquota vigente à época da aquisição, divulgados pelo site da Sefaz ([www.sefaz.go.gov.br](http://www.sefaz.go.gov.br)).

Atentando para questões referentes ao sigilo fiscal e tendo por base os objetivos meramente acadêmicos do presente trabalho, omitidos os dados dos contribuintes e substituímos por dados e valores meramente fictícios.

Em acréscimo aos objetivos acadêmicos do presente trabalho, buscamos demonstrar aos presentes um pouco das atividades desenvolvidas pelos Servidores da Secretaria da Fazenda do Estado de Goiás voltadas à proteção do crédito tributário.

# Storytelling

Buscando atingir o objetivo traçado (detectar inconsistências nos valores lançados a crédito na apuração do valor do imposto relativo às operações de aquisição de óleo diesel gerando prejuízos ao erário), contamos com três bases de dados:



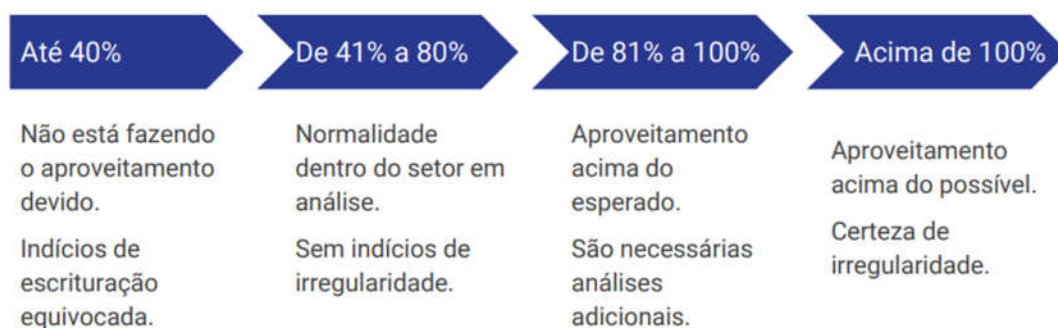
Estas bases de dados referem-se a uma amostra de um determinado setor industrial e foram armazenadas em três arquivos em formato CSV (NFE\_Diesel-71fcd.csv, E111.csv, PMPF.csv). Novamente insistimos em esclarecer que se trata de dados meramente fictícios utilizados apenas para fins acadêmicos.

Partindo dos dados acima descritos, reconstruímos o valor máximo de ICMS relativo às aquisições de óleo diesel (ICMS correspondente às notas de entrada) e comparamos com os dados lançados a crédito pelo contribuinte, gerando assim os percentuais de aproveitamento do crédito lançado em relação ao total calculado em relação aos documentos de entrada (Valor Lançado na Escrituração / Valor Total Calculado Através da análise das Notas Fiscais).

O creditamento de tais valores por parte do contribuinte não ocorre no momento da aquisição da mercadoria e sim no momento da utilização da mercadoria adquirida, onde o contribuinte quantifica o volume aplicado diretamente na produção/industrialização e calcula o valor do respectivo imposto, objeto de creditamento. Os contribuintes industriais só podem aproveitar o crédito referente a aquisições de óleo diesel aplicados diretamente na atividade produtiva ou industrial do estabelecimento (custo de produção). Nas demais atividades (uso e consumo), o creditamento de tais valores só poderá ser utilizada relativo às mercadorias adquiridas a partir de 1º de janeiro de 2020. Porém, muitas vezes, os valores lançados extrapolam do montante real, causando um abatimento indevido do montante do imposto

devido, causando, conseqüentemente, prejuízo ao erário. No setor industrial objeto do presente trabalho, estipulamos que um aproveitamento dentro do intervalo de 41 a 80% do total de óleo diesel adquirido pelo industrial não gera indícios de irregularidades. Um Aproveitamento em percentual menor que 41% gera indícios de lançamento equivocado do valor do crédito, ou seja, o contribuinte pode estar lançando em outra conta ou mesmo creditando direto no registro do próprio documento fiscal de entrada, o que gera a necessidade de maiores averiguações. Um aproveitamento acima de 80% gera indícios de inconsistências no valor lançado, sendo que aproveitamentos acima de 100% apresenta o caso crítico (aproveitamento maior que o limite máximo).

Assim categorizamos os dados colhidos seguindo o seguinte esquema:

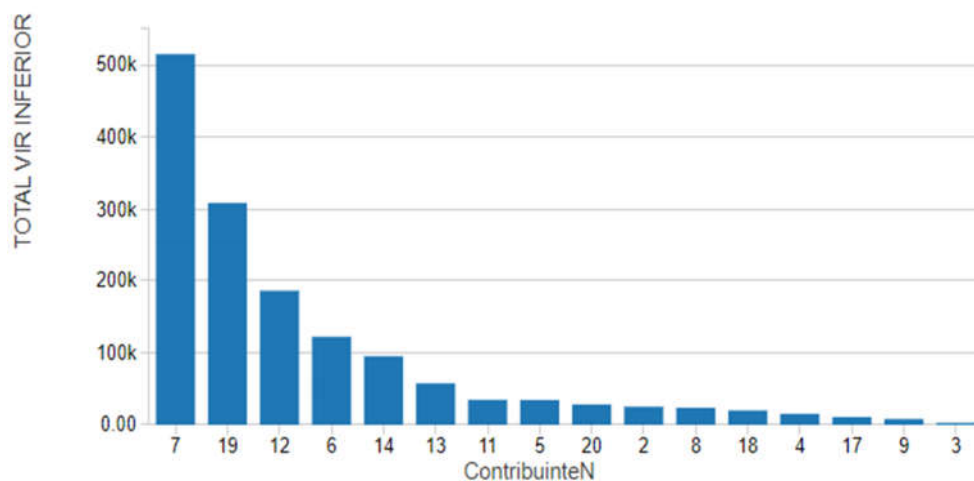


Como a apuração do ICMS é por regra mensal, precisamos analisar as consistências mensalmente, totalizando os resultados de forma a identificar as empresas com as maiores inconsistências no período em análise (2014 a 2017).

Cada situação de inconsistência acima mencionada requer um tratamento específico por parte dos procedimentos de auditoria posteriores. Assim, apresentaremos os resultados gerados individualizados para cada situação:

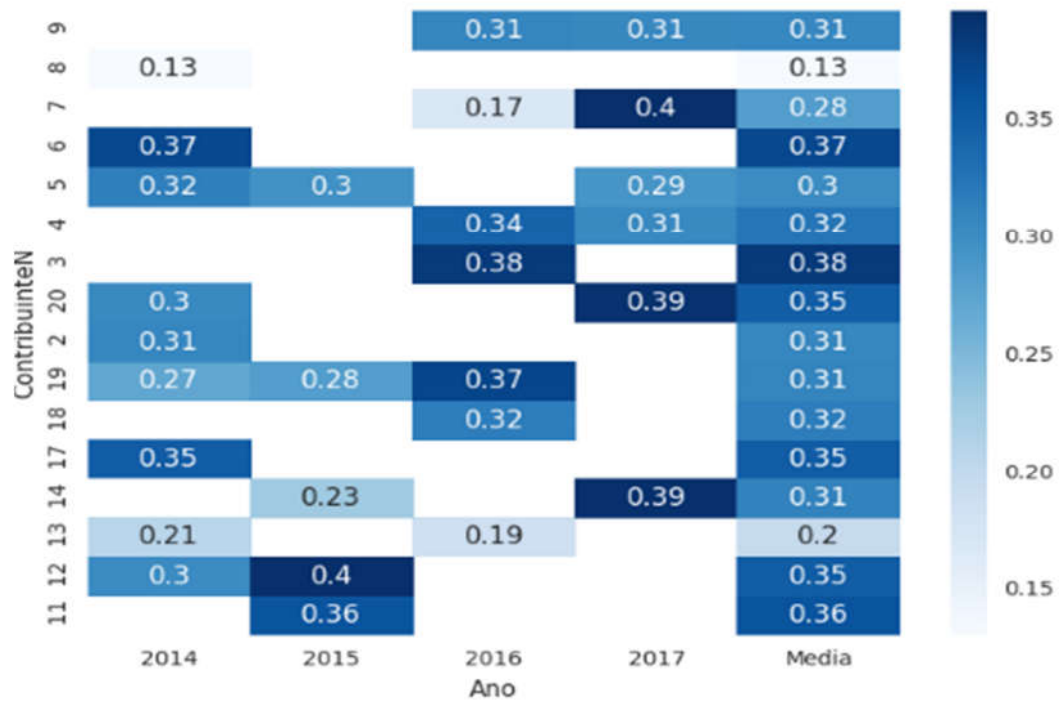
- 1ª Situação de Inconsistência: Aproveitamento de crédito abaixo do esperado (Percentual de aproveitamento até 40%):

**Contribuintes com maior discrepância entre o limite inferior (40%) e o valor creditado (Valores em R\$)**



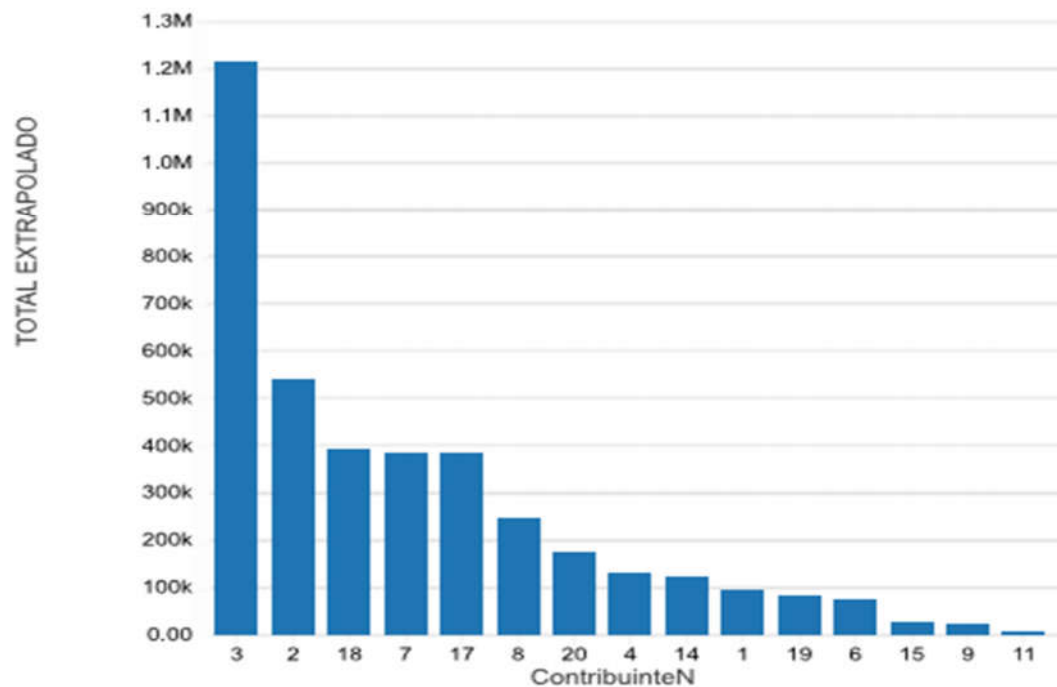
**\*Total VLR Inferior:** diferença entre o limite mínimo convencionado (40%) e o valor creditado pelo contribuinte.

Média anual entre os Percentuais inferiores a 40%



- 2ª Situação de Inconsistência: Aproveitamento de crédito acima do esperado, porém inferior ao montante total apurado nos documentos fiscais (Percentual de aproveitamento no intervalo de 80 a 100%):

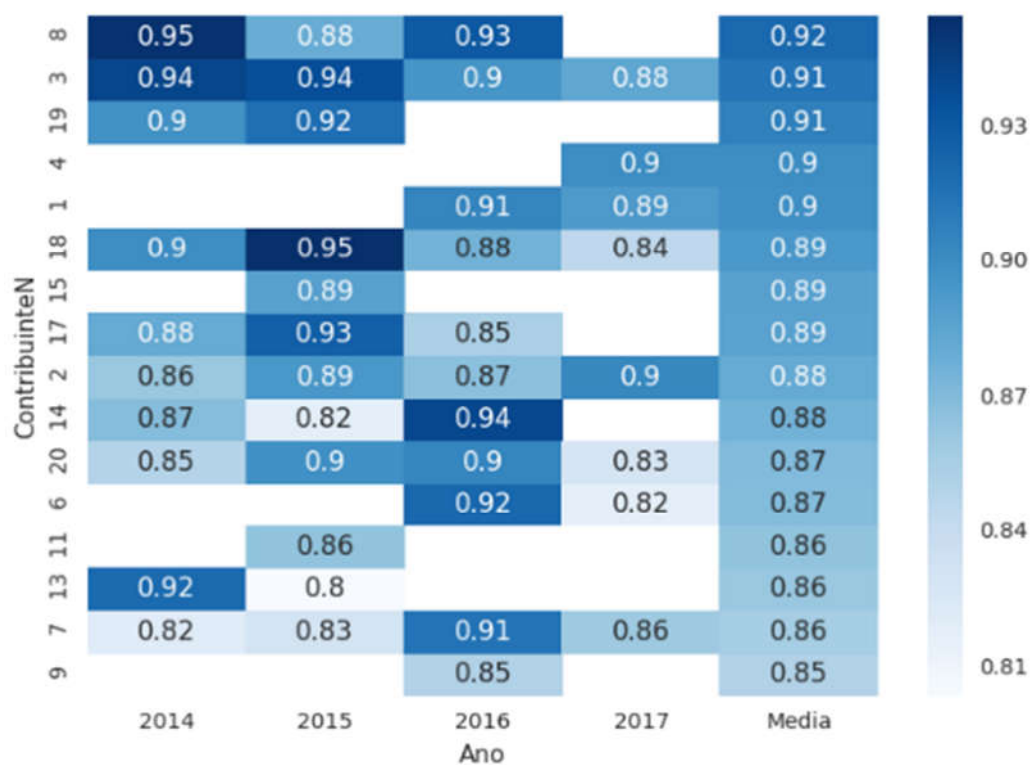
**Contribuintes com maior discrepância entre o valor creditado e o limite superior (80%) - (Valores em R\$)**



**\*Total Extrapolado:** diferença entre o valor creditado pelo contribuinte e limite superior convencionado (80%).

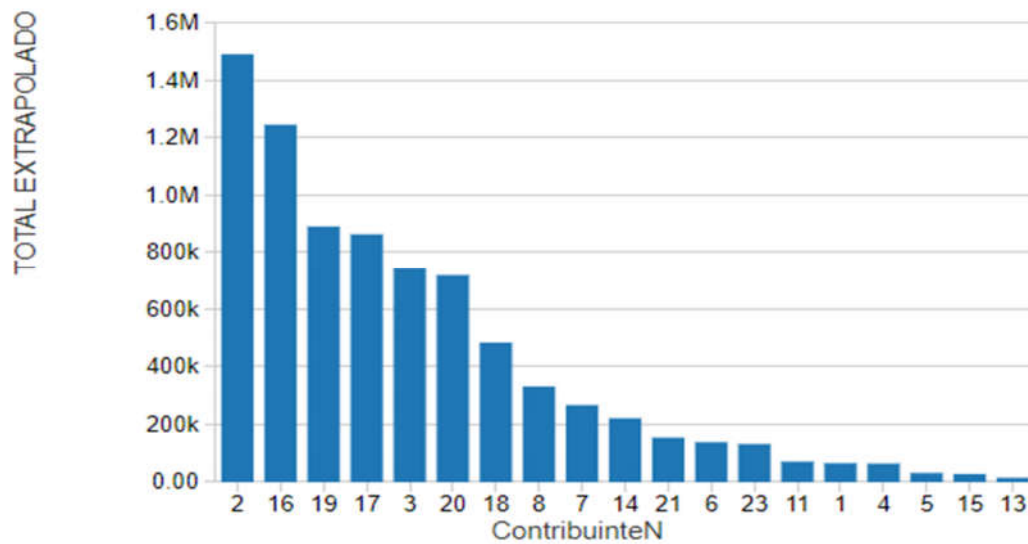


Média anual entre os Percentuais superiores a 80% e inferiores a 100%



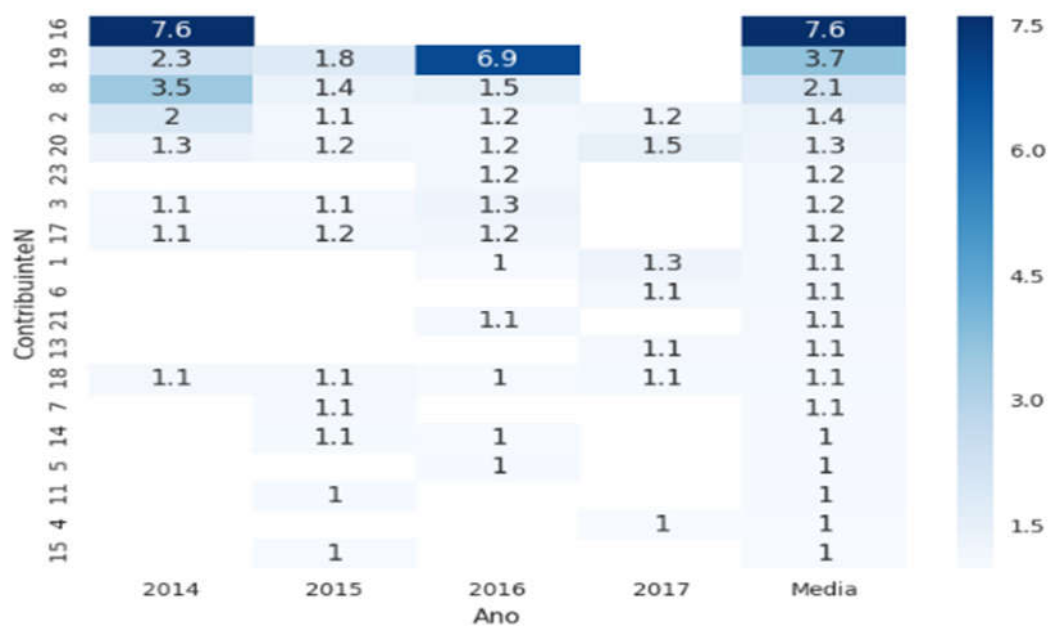
- 3ª Situação de Inconsistência: Aproveitamento de crédito bem acima do esperado (Percentual de aproveitamento acima de 100%):

**Contribuintes com maior discrepância entre o valor creditado e o limite máximo permitido (100% do valor do ICMS referentes às notas fiscais de entrada) - (Valores em R\$)**



**\*Total Extrapolado:** diferença entre o valor creditado pelo contribuinte e limite superior convencionado (80%).

**Média anual entre os Percentuais superiores a 100%**



# Conclusão

Através dos procedimentos e critérios acima mencionados, para cada situação abordada, identificamos os contribuintes que apresentaram maior discrepância entre o valor escriturado do crédito referente à utilização do óleo diesel no estabelecimento industrial e o valor reconstruído comprovado pelas notas fiscais que acobertam as operações de entrada de óleo diesel nos referidos estabelecimentos industriais, constituindo assim situações potencialmente danosas ao Estado, exigindo averiguações adicionais e, se for o caso, a exigência do crédito tributário devido.

# Apêndice

## Scripts gerados:

```
from pyspark.sql.types import DateType, IntegerType, FloatType, StringType,
DoubleType

from pyspark.sql.functions import col, udf, mean, sum

from pyspark.sql.functions import udf, struct

import seaborn as sns

import pandas as pd

import matplotlib.pyplot as plt

dfnfe = spark.read.format("csv") \
    .option("inferSchema", "true") \
    .option("header", "true") \
    .option("sep", ";") \
    .load("/FileStore/tables/NFE_Diesel-71fcd.csv")

dfel11 = spark.read.format("csv") \
    .option("inferSchema", "true") \
    .option("header", "true") \
    .option("sep", ";") \
    .load("/FileStore/tables/E111.csv")

dfpmpf = spark.read.format("csv") \
    .option("inferSchema", "true") \
    .option("header", "true") \
    .option("sep", ";") \
    .load("/FileStore/tables/PMPF.csv")

dfnfe.printSchema()

dfel11.printSchema()

dfpmpf.printSchema()
```

```

root
|-- Contribuinte: string (nullable = true)
|-- Nfe: integer (nullable = true)
|-- Cod CFOP: integer (nullable = true)
|-- Data de Emissao: string (nullable = true)
|-- AnoMes: integer (nullable = true)
|-- Produto: string (nullable = true)
|-- Unidade: string (nullable = true)
|-- Quantidade Comercial: double (nullable = true)

root
|-- Contribuinte: string (nullable = true)
|-- AnoMes: integer (nullable = true)
|-- Cod Ajuste Apuracao ICMS : string (nullable = true)
|-- Valor Ajuste Apuracao ICMS: double (nullable = true)

root
|-- AnoMes: integer (nullable = true)
|-- PMPF: double (nullable = true)
|-- ALIQUOTA: double (nullable = true)

```

```

dfnfe_mensal = dfnfe.groupby(["Contribuinte","AnoMes"]).agg(sum('Quantidade
Comercial')).alias("QUANT DIESEL")).cache()

```

```

dfnfe_mensal = dfnfe_mensal.orderBy(['Contribuinte','AnoMes'])

```

```

dfnfe_mensal.show()

```

```

+-----+-----+-----+
| Contribuinte|AnoMes|      QUANT DIESEL|
+-----+-----+-----+
|Contribuinte 1|201401|      45122.982|
|Contribuinte 1|201402|      135066.11|
|Contribuinte 1|201403|      90229.353|
|Contribuinte 1|201404|      254314.814|
|Contribuinte 1|201405|      584907.351|
|Contribuinte 1|201406| 639776.2760000001|
|Contribuinte 1|201407|      627420.619|
|Contribuinte 1|201408| 639069.5549999999|
|Contribuinte 1|201409| 587193.3740000001|
|Contribuinte 1|201410|      459940.379|
|Contribuinte 1|201411| 194712.64299999998|
|Contribuinte 1|201412| 15739.771999999999|
|Contribuinte 1|201501| 6738.4529999999995|
|Contribuinte 1|201502|      10000.0|
|Contribuinte 1|201503|      30000.0|
|Contribuinte 1|201504|      344998.154|
|Contribuinte 1|201505|      734199.033|

```

```

df_aux=dfel11.join(dfnfe_mensal,["Contribuinte","AnoMes"])

```

```

df_tot=df_aux.join(dfpmpf,["AnoMes"])

```

```

df_tot = df_tot.withColumn("ICMS DIESEL",df_tot['QUANT
DIESEL']*df_tot['PMPF']*df_tot['ALIQUOTA'])

```

```
df_tot = df_tot.withColumn("PERC APROVEITAMENTO",df_tot['Valor Ajuste Apuracao ICMS']/df_tot['ICMS DIESEL'])

df_tot = df_tot.withColumn("ANO", (df_tot['AnoMes']/100).cast(IntegerType()))

df_tot = df_tot.withColumn("ContribuinteN", df_tot['Contribuinte'])[13:5])

df_tot.show()
```

```
0|ContribuinteN|
-----|-----|
|201608| Contribuinte 1|          G0020007|    164413.17|    498471.663| 3.0034|    0.15| 224566.48889813| 0.7321358830047895|2016|
|201609| Contribuinte 1|          G0020007|   146906.71|   361429.475| 3.0034|    0.15| 162827.59278225| 0.9022224519185696|2016|
|201610| Contribuinte 1|          G0020007|   18909.99|   41895.762| 3.0034|    0.15| 18875.81128862| 1.0018107158889018|2016|
|201611| Contribuinte 1|          G0020007|   15140.18|   35695.114| 3.023|    0.15| 16185.9494433| 0.9353902934787132|2016|
|201612| Contribuinte 1|          G0020007|    8764.08|   30000.0| 3.008|    0.15| 13536.0| 0.6474645390070922|2016|
|201701| Contribuinte 1|          G0020007|    8809.71|   45000.0| 3.05285|    0.15| 28606.7375| 0.42751608053137956|2017|
```

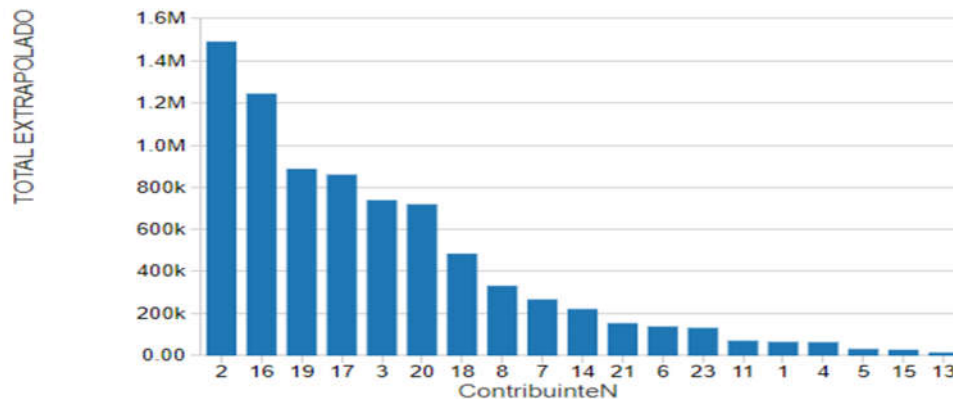
```
df_customer1 = df_tot.filter(df_tot['PERC APROVEITAMENTO'] > 1)

df_customer1 = df_customer1.withColumn("Vlr EXTRAPOLADO", df_customer1['Valor Ajuste Apuracao ICMS']-(df_customer1['ICMS Diesel'])*0.8)

df_customer2 = df_customer1.groupBy("ContribuinteN").agg(sum("Vlr EXTRAPOLADO").alias("TOTAL EXTRAPOLADO"))

df_customer2 = df_customer2.orderBy(['TOTAL EXTRAPOLADO'], ascending=False)

display(df_customer2)
```



```
df_customer3 = df_customer1.groupBy("Ano", "ContribuinteN").agg(sum("Valor Ajuste Apuracao ICMS").alias("Ajustes ICMS"),
sum("ICMS DIESEL").alias("ICMS Diesel"))

df_customer3 = df_customer3.withColumn("Perc Aproveit", df_customer3['Ajustes ICMS']/df_customer3['ICMS Diesel'])

pandas_df = df_customer3.toPandas()

df_graph = pandas_df.pivot("ContribuinteN", "Ano", "Perc Aproveit")

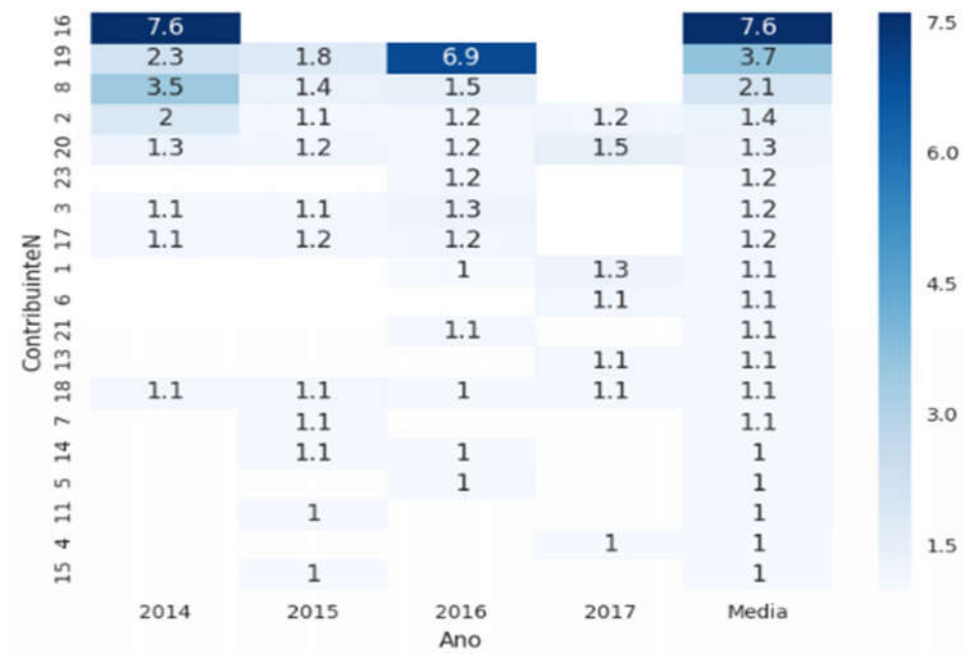
df_graph['Media'] = df_graph.mean(axis=1)

df_graph = df_graph.sort_values(by=['Media'], ascending=False)

df_graph
```

Ano	2014	2015	2016	2017	Media
ContribuinteN					
16	7.629314	NaN	NaN	NaN	7.629314
19	2.259582	1.826272	6.942415	NaN	3.676090
8	3.471192	1.412811	1.519941	NaN	2.134648
2	1.950037	1.139571	1.153033	1.234708	1.369337
20	1.311650	1.197759	1.159890	1.503512	1.293203
23	NaN	NaN	1.181897	NaN	1.181897
3	1.091211	1.092099	1.315907	NaN	1.166406
17	1.133754	1.157325	1.200708	NaN	1.163929
1	NaN	NaN	1.001811	1.296696	1.149254
6	NaN	NaN	NaN	1.133211	1.133211
21	NaN	NaN	1.095858	NaN	1.095858
13	NaN	NaN	NaN	1.091229	1.091229
18	1.100804	1.092623	1.004535	1.061628	1.064898
7	NaN	1.052217	NaN	NaN	1.052217
14	NaN	1.074852	1.007288	NaN	1.041070
5	NaN	NaN	1.039217	NaN	1.039217
11	NaN	1.036296	NaN	NaN	1.036296
4	NaN	NaN	NaN	1.012230	1.012230

```
fig, ax = plt.subplots()
df_graph.sortlevel(level=0, ascending=False, inplace=True)
ax = sns.heatmap(df_graph, cmap="Blues", annot=True)
display(ax.figure)
```

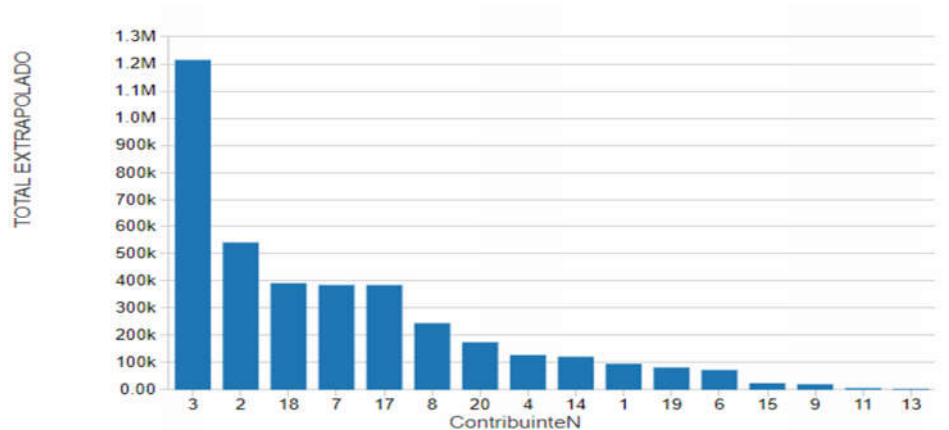




```

df_customer = df_tot.filter(df_tot['PERC APROVEITAMENTO'] > 0.8)
df_customer1 = df_customer.filter(df_tot['PERC APROVEITAMENTO'] <= 1)
df_customer1 = df_customer1.withColumn("VlR EXTRAPOLADO", df_customer1['Valor
Ajuste Apuracao ICMS']-(df_customer1['ICMS Diesel']*0.8))
df_customer2 = df_customer1.groupBy("ContribuinteN").agg(sum("VlR
EXTRAPOLADO").alias("TOTAL EXTRAPOLADO"))
df_customer2 = df_customer2.orderBy(['TOTAL EXTRAPOLADO'], ascending=False)
display(df_customer2)

```



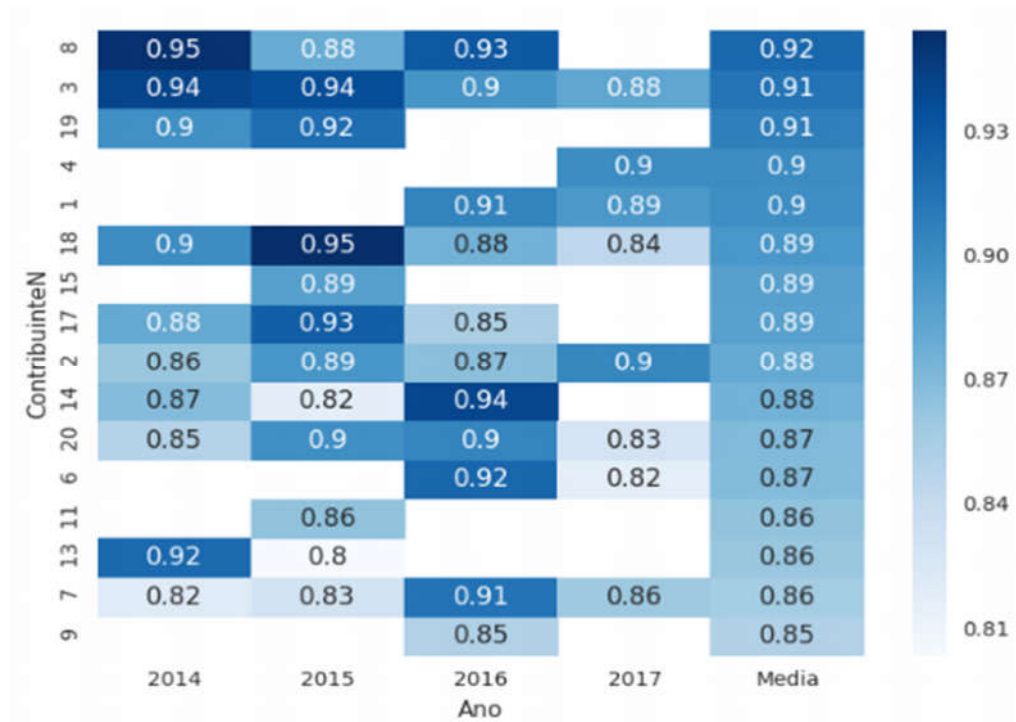
```

df_customer3 = df_customer1.groupBy("Ano", "ContribuinteN").agg(sum("Valor
Ajuste Apuracao ICMS").alias("Ajustes ICMS"),
sum("ICMS DIESEL").alias("ICMS Diesel"))
df_customer3 = df_customer3.withColumn("Perc Aproveit", df_customer3['Ajustes
ICMS']/df_customer3['ICMS Diesel'])
pandas_df = df_customer3.toPandas()
df_graph = pandas_df.pivot("ContribuinteN", "Ano", "Perc Aproveit")
df_graph.sortlevel(level=0, ascending=False, inplace=True)
df_graph['Media'] = df_graph.mean(axis=1)
df_graph = df_graph.sort_values(by=['Media'], ascending=False)
df_graph

```

Ano	2014	2015	2016	2017	Media
ContribuinteN					
8	0.952648	0.880298	0.930461	NaN	0.921135
3	0.942048	0.936736	0.895137	0.883961	0.914471
19	0.897644	0.917753	NaN	NaN	0.907699
4	NaN	NaN	NaN	0.900342	0.900342
1	NaN	NaN	0.905221	0.892198	0.898710
18	0.900072	0.954735	0.876040	0.844413	0.893815
15	NaN	0.887497	NaN	NaN	0.887497
17	0.881461	0.926613	0.853962	NaN	0.887345
2	0.861520	0.894469	0.866843	0.902954	0.881447
14	0.869140	0.817894	0.940646	NaN	0.875893
20	0.849448	0.898605	0.904610	0.829108	0.870443
6	NaN	NaN	0.922243	0.817535	0.869889
11	NaN	0.864607	NaN	NaN	0.864607
13	0.920455	0.803659	NaN	NaN	0.862057
7	0.820923	0.831166	0.914688	0.860048	0.856706
9	NaN	NaN	0.851250	NaN	0.851250

```
fig, ax = plt.subplots()
df_graph.sortlevel(level=0, ascending=False, inplace=True)
ax = sns.heatmap(df_graph, cmap="Blues", annot=True)
display(ax.figure)
```



```

df_customer1 = df_tot.filter(df_tot['PERC APROVEITAMENTO'] < 0.4)

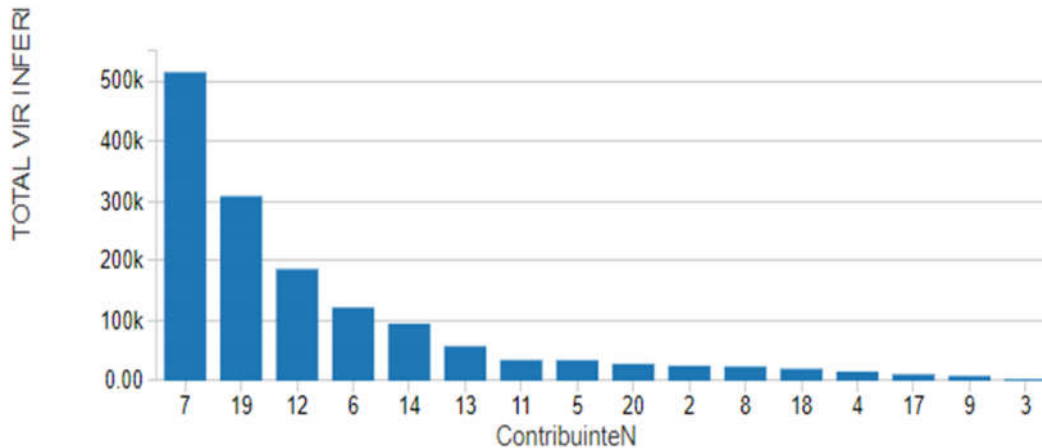
df_customer1 = df_customer1.withColumn("VIR INFERIOR AO MINIMO", (df_customer1['ICMS Diesel']*0.4)-
(df_customer1['Valor Ajuste Apuracao ICMS']))

df_customer2 = df_customer1.groupBy("ContribuinteN").agg(sum("VIR INFERIOR AO MINIMO").alias("TOTAL VIR
INFERIOR"))

df_customer2 = df_customer2.orderBy(['TOTAL VIR INFERIOR'], ascending=False)

display(df_customer2)

```



```

df_customer3 = df_customer1.groupBy("Ano", "ContribuinteN").agg(sum("Valor
Ajuste Apuracao ICMS").alias("Ajustes ICMS"),
sum("ICMS DIESEL").alias("ICMS Diesel"))

df_customer3 = df_customer3.withColumn("Perc Aproveit", df_customer3['Ajustes
ICMS']/df_customer3['ICMS Diesel'])

pandas_df = df_customer3.toPandas()

df_graph = pandas_df.pivot("ContribuinteN", "Ano", "Perc Aproveit")

df_graph.sortlevel(level=0, ascending=False, inplace=True)

df_graph['Media'] = df_graph.mean(axis=1)

df_graph = df_graph.sort_values(by=['Media'], ascending=False)

df_graph

```

Ano	2014	2015	2016	2017	Media
ContribuinteN					
3	NaN	NaN	0.384371	NaN	0.384371
6	0.371024	NaN	NaN	NaN	0.371024
11	NaN	0.359340	NaN	NaN	0.359340
17	0.350779	NaN	NaN	NaN	0.350779
12	0.302691	0.396778	NaN	NaN	0.349735
20	0.304383	NaN	NaN	0.393189	0.348786
4	NaN	NaN	0.342996	0.305233	0.324115
18	NaN	NaN	0.316958	NaN	0.316958
14	NaN	0.228496	NaN	0.394796	0.311646
19	0.272166	0.283217	0.374366	NaN	0.309916
2	0.308027	NaN	NaN	NaN	0.308027
9	NaN	NaN	0.308592	0.306356	0.307474
5	0.315440	0.296381	NaN	0.292545	0.301455
7	NaN	NaN	0.169509	0.396713	0.283111
13	0.206878	NaN	0.186302	NaN	0.196590
8	0.130970	NaN	NaN	NaN	0.130970

```
fig, ax = plt.subplots()
df_graph.sortlevel(level=0, ascending=False, inplace=True)
ax = sns.heatmap(df_graph, cmap="Blues", annot=True)
display(ax.figure)
```

