

Documentação do projeto ‘Execução  
Orçamentária e Financeira do Governo  
Federal’ (“Projeto Caixa”)

*R6 Estatística e Treinamentos LTDA*

*2019-12-29*



# Contents

<b>1</b>	<b>Introdução</b>	<b>5</b>
1.1	Questões investigadas . . . . .	5
1.2	As bases de dados . . . . .	6
1.3	Análise descritiva . . . . .	7
1.4	Análise preditiva . . . . .	7
1.5	Oficina de repasses e implementação dos produtos . . . . .	7
<b>2</b>	<b>Validação e Compreensão dos Dados</b>	<b>9</b>
2.1	Introdução . . . . .	9
2.2	Questões elementares . . . . .	9
2.3	Bases de dados levantadas para endereçar as questões . . . . .	10
2.4	Base 1: Movimentações diárias do Limite de Saque . . . . .	11
2.5	Base 2: Pagamentos diários . . . . .	11
2.6	Base 3: Movimentações diárias em obrigações a pagar . . . . .	12
2.7	Tabelas Derivadas . . . . .	13
<b>3</b>	<b>Quadro analítico histórico</b>	<b>15</b>
3.1	Disponibilidades Líquidas . . . . .	15
3.2	Classificações Orçamentárias e Financeiras . . . . .	15
3.3	Disponibilidades Líquidas por Vinculação . . . . .	15
3.4	Comportamento do Caixa por Movimentação . . . . .	16
<b>4</b>	<b>Quadro analítico preditivo</b>	<b>17</b>
	<b>Referências</b>	<b>19</b>



# Chapter 1

## Introdução

O objetivo deste projeto é analisar o comportamento do caixa e das obrigações financeiras dos órgãos federais, com a finalidade de fornecer informações para a gestão da programação financeira por parte do Tesouro Nacional, além de identificar oportunidades de melhorias nesse processo, e possivelmente fundamentar a criação de indicadores para avaliação da gestão financeira das unidades do Governo Federal.

Elencamos aqui alguns aspectos importantes do projeto.

### 1.1 Questões investigadas

De início, analisamos o perfil das despesas e receitas das unidades do Governo Federal, começando com os órgãos do Ministério da Justiça (que já possui um sistema de acompanhamento de despesas bem estruturado).

A análise procurou compatibilizar as informações orçamentárias com as informações financeiras. Chamamos de *classificadores orçamentários*: Função, Subfunção, Programa, Ação, Grupo de Despesa, Modalidade de Aplicação, Elemento de Despesa, Indicador de Resultado EOF, Indicador de Exceção Decreto. Como *classificadores financeiros*, nos referimos, essencialmente, à Vinculação de Pagamento.

A *Fonte de Recurso* é um classificador comum a esses dois contextos, orçamentário e financeiro.

Com esse escopo em mente, em parceria com a equipe do GT-CEAD, as seguintes questões foram abordadas neste projeto:

- a. Qual o comportamento do caixa e das obrigações a pagar (e da disponibilidade líquida) no período analisado<sup>1</sup>?
- b. Existem casos em que unidades de um mesmo órgão permanecem com disponibilidade líquida negativa, enquanto outras unidades desse mesmo órgão encontram-se com disponibilidade positiva? E se considerar a fonte de recursos, há casos em que o órgão passa por períodos com disponibilidade negativa em uma fonte enquanto há recursos disponíveis em outra fonte? E se considerar as duas situações conjuntamente<sup>2</sup>?
- c. Como as classificações orçamentárias se relacionam com as classificações financeiras? Especificamente, é possível identificar certos tipos de despesas que são sempre (ou frequentemente) pagas com recursos de determinadas vinculações?
- d. Caso seja possível a identificação mencionada em (c), as questões (a) e (b) seriam revisitadas para estimar a disponibilidade líquida para cada vinculação, considerando as classificações orçamentárias das obrigações. Nesse cenário, existem unidades com saldo total suficiente para cobrir todas as suas obrigações, porém com insuficiência em algumas vinculações?
- e. Qual o comportamento do caixa das unidades em termos de movimentações?
- f. Qual o intervalo entre duas operações<sup>3</sup> de grande porte?

Para responder essas perguntas, utilizamos ferramentas descritivas e de modelagem preditiva, descritas nos Capítulos 3 e 4, respectivamente.

## 1.2 As bases de dados

Para responder as questões levantadas no item anterior, os dados do SIAFI foram divididos em três bases de dados:

- 1. base de movimentações diárias do limite de saque;
- 2. base de pagamentos diários;
- 3. base de movimentações diárias nas obrigações a pagar.

---

<sup>1</sup>Por órgão, por unidade e por fonte de recursos.

<sup>2</sup>Ou seja, uma unidade de um mesmo órgão fica com disponibilidade negativa em uma fonte, enquanto outra unidade desse mesmo órgão possui disponibilidade positiva nessa mesma fonte.

<sup>3</sup>Uma despesa alta seguida de um recebimento de recursos também alto.

Com as Tabelas 1 e 3, calculamos a *disponibilidade líquida diária* a partir da subtração entre os saldos diários do caixa (Tabela 1) e as obrigações a pagar (Tabela 3), para cada unidade gestora ou órgão, e para cada fonte de recursos.

Com as Tabelas 1 e 2, relacionamos no contexto dos pagamentos, as informações orçamentárias (Tabela 2) com os vínculos de pagamento (Tabela 1).

Finalmente com a Tabela 1, analisamos as movimentações para tipo de documento, obtendo um histórico das movimentações de cada órgão e cada Unidade Gestora.

Na primeira etapa do projeto, essas bases foram estudadas e validadas. Bases auxiliares também foram construídas para facilitar as análises subsequentes. A documentação dessa etapa se encontra no Capítulo 2.

## 1.3 Análise descritiva

Dada a riqueza de informações e granularidade das bases de dados, todos os resultados descritivos foram construídos em uma aplicação online, que permite a manipulação das visualizações a partir de filtros e seletores.

O aplicativo pode ser acessado a partir do seguinte link: [https://rseis.shinyapps.io/explorador\\_disponibilidades\\_liquidas\\_v2/](https://rseis.shinyapps.io/explorador_disponibilidades_liquidas_v2/)

O código-fonte e manutenção do aplicativo foram repassados à equipe do GT-CEAD ao fim do projeto.

Um resumo dos principais resultados se encontra no Capítulo 3.

## 1.4 Análise preditiva

Os modelos utilizados neste projeto consideraram os métodos em estado da arte dentro do contexto de modelagem preditiva, como florestas aleatórias, o algoritmo XGBoost e redes neurais.

A descrição dos modelos ajustados e seus resultados se encontram no Capítulo 4.

## 1.5 Oficina de repasses e implementação dos produtos

A última etapa do projeto consistiu de uma oficina de repasses, realizada presencialmente no GT-CEAD, no Tesouro Nacional, em Brasília. Nessa oficina, foi discutida a teoria por trás dos métodos aplicados, tal como apresentada neste

relatório. Também foram apresentados e explicados os scripts em linguagem de programação R utilizados para implementar os modelos.



## Chapter 2

# Validação e Compreensão dos Dados

### 2.1 Introdução

A base de dados do SIAFI concentra as informações geradas pelo processo de execução orçamentária e programação financeira do Governo Federal. Este projeto tem como objetivo gerar conhecimento sobre o comportamento do caixa e das obrigações financeiras dos órgãos federais para fins de gestão da programação financeira por parte do Tesouro Nacional, além de identificar oportunidades de melhorias nesse processo e, possivelmente, fundamentar a criação de indicadores para avaliação da gestão financeira das unidades do Governo Federal.

Este é o primeiro dos quatro relatórios que irão compor o projeto e descreve o processo de conhecimento, ajuste e preparo das bases de dados para análises subsequentes.

### 2.2 Questões elementares

Para nortear as análises das próximas etapas, nessa primeira fase de conhecimento do problema e da base, foram levantadas as seguintes questões:

- a. Qual o comportamento do caixa e das obrigações a pagar (e da disponibilidade líquida) no período analisado<sup>1</sup>?
- b. Existem casos em que unidades de um mesmo órgão permanecem com disponibilidade líquida negativa, enquanto outras unidades desse mesmo

---

<sup>1</sup>Por órgão, por unidade e por fonte de recursos.

órgão encontram-se com disponibilidade positiva? E se considerar a fonte de recursos, há casos em que o órgão passa por períodos com disponibilidade negativa em uma fonte enquanto há recursos disponíveis em outra fonte? E se considerar as duas situações conjuntamente<sup>2</sup>?

- c. Como as classificações orçamentárias se relacionam com as classificações financeiras? Especificamente, é possível identificar certos tipos de despesas que são sempre (ou frequentemente) pagas com recursos de determinadas vinculações?
- d. Caso seja possível a identificação mencionada em (c), as questões (a) e (b) seriam revisitadas para estimar a disponibilidade líquida para cada vinculação, considerando as classificações orçamentárias das obrigações. Nesse cenário, existem unidades com saldo total suficiente para cobrir todas as suas obrigações, porém com insuficiência em algumas vinculações?
- e. Qual o comportamento do caixa das unidades em termos de movimentações?
- f. Qual o intervalo entre duas operações<sup>3</sup> de grande porte?

## 2.3 Bases de dados levantadas para endereçar as questões

O perfil das despesas do **Ministério da Justiça** será utilizado como ponto de partida das análises descritivas e exploratórias. A escolha deste órgão foi em virtude do seu excelente sistema de acompanhamento das despesas.

Ao todo, três bases foram extraídas:

- **lim\_saque** com as movimentações de limites de saque;
- **obrigacoes** com as movimentações de obrigações a pagar; e
- **pagamentos** com as movimentações de pagamentos;

A partir destas, outras três tabelas derivadas foram construídas. As descrições detalhadas estão na seção seguinte.

- **disponibilidades\_liquidas\_diarias** com as informações diárias de **saldo disponível** e **obrigações a pagar** de cada UG para cada fonte de recursos;

---

<sup>2</sup>Ou seja, uma unidade de um mesmo órgão fica com disponibilidade negativa em uma fonte, enquanto outra unidade desse mesmo órgão possui disponibilidade positiva nessa mesma fonte.

<sup>3</sup>Uma despesa alta seguida de um recebimento de recursos também alto.

- `vinculacao_de_pagamentos` com informações diárias pareadas de **paga-mentos**, **saldo disponível** e **vinculações de pagamentos** de cada documento;
- `lim_saque_por_tipo_de_documento` com as informações diárias de **saldo disponível** por tipo de documento (e.g. NS, OB, PF, etc.).

Verificou-se que as extrações **encontram-se prontas para análise** e com replicações em formatos `.rds` para serem lidas pelo software R. Abaixo estão listados os seus respectivos campos.

## 2.4 Base 1: Movimentações diárias do Limite de Saque

**Filtro:** item de informação “LIMITES DE SAQUE”.

**Campos:**

- Órgão Máximo
- Órgão
- UG
- **Vinculação de Pagamento**
- Fonte de Recursos Detalhada
- Fonte de Recursos
- Documento Lançamento
- Movimento / Valor Financeiro

## 2.5 Base 2: Pagamentos diários

**Filtro:** item de informação “PAGAMENTOS TOTAIS”.

**Campos:**

- Órgão Máximo
- Órgão
- UG
- Fonte de Recursos Detalhada
- Fonte de Recursos
- **Função**
- **Subfunção**
- **Programa**
- **Ação**
- **Grupo de Despesa**

- **Modalidade de Aplicação**
- **Elemento de Despesa**
- **Indicador de Resultado EOF** (indica se a despesa é primária ou financeira, entre outras coisas)
- **Indicador de Exceção Decreto**
- **Ano do Empenho**
- **Empenho**
- **Órgão Máximo da UO**
- Documento Lançamento
- Movimento / Valor Financeiro

## 2.6 Base 3: Movimentações diárias em obrigações a pagar

**Filtro:** item de informação “VALORES LIQUIDADOS A PAGAR (EXERCÍCIO + RP)”.

**Campos:**

- Órgão Máximo
- Órgão
- UG
- Fonte Detalhada
- Fonte (posições 3 e 4 da fonte detalhada – exemplo: se a fonte detalhada é: 0100123456, a fonte será 00)
- **Função**
- **Subfunção**
- **Programa**
- **Ação**
- **Grupo de Despesa**
- **Modalidade de Aplicação**
- **Elemento de Despesa**
- **Indicador de Resultado EOF** (indica se a despesa é primária ou financeira, entre outras coisas)
- **Indicador de Exceção Decreto**
- **Ano do Empenho**
- **Empenho**
- **Órgão Máximo da UO**
- Movimento / Valor Financeiro

## 2.7 Tabelas Derivadas

### 2.7.1 Disponibilidades líquidas Diárias

- Informações diárias de **saldo disponível** e **obrigações a pagar** de cada UG para cada fonte de recursos.
- Útil para as questões **a**, **c** e **e**.
- Cruzamento entre as bases `lim_saque` e `obrigacoes` pelas chaves `NO_DIA_COMPLETO`, `NO_FONTE_RECURSO` e `NO_UG`.
- Observações: 352.037
- Campos: 11

* NO_DIA_COMPLETO	`<date>	2017-08-22, 2017-08-23, 2017-08-24, 2017-08...
* NO_UG	`<chr>	"ACADEMIA NACIONAL DA POLICIA RODOV. FEDERAL..."
* NO_ORGAO	`<chr>	"DEPARTAMENTO DE POLICIA RODOVIARIA FEDERAL/..."
* NO_FONTE_RECURSO	`<chr>	"RECEITAS DE CONCURSOS DE PROGNOSTICOS", "RE..."
* saldo_diario	`<dbl>	-2779.98, -5654.96, -7356.26, -5601.60, -560...
* obrigacoes_a_pagar	`<dbl>	0.000000e+00, 0.000000e+00, 0.000000e+00, 0.0...
* disponibilidade_liquida	`<dbl>	-2779.98, -5654.96, -7356.26, -5601.60, -560...
* ano	`<dbl>	2017, 2017, 2017, 2017, 2017, 2017, 2017, 20...
* mes	`<dbl>	8, 8, 8, 8, 8, 8, 8, 8, 8, 9, 9, 9, 9, ...
* dia	`<int>	22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 1, 2...
* padded	`<lgl>	TRUE, TRUE, TRUE, TRUE, FALSE, FALSE, TRUE, ...

### 2.7.2 Vinculação de Pagamentos

- Informações diárias pareadas de **pagamentos**, **saldo disponível** e **vinculações de pagamentos** de cada documento.
- Útil para as questões **b**, **c** e **e**.
- Cruzamento entre as bases `lim_saque` e `pagamentos` pelas chaves `NO_DIA_COMPLETO` e `ID_DOCUMENTO`.
- Observações: 1.005.187
- Campos: 7

```
* NO_DIA_COMPLETO      ^<chr> "01/02/2017", "01/02/2017", "01/02/2017", "0..."
* ID_DOCUMENTO         ^<chr> "1940031920820170B800016", "1940031920820170..."
* pagamento           ^<dbl> 20.79, 724.21, 8000.00, 76.95, -15188.99, -9..."
* NO_VINCULACAO_PAGAMENTO ^<chr> "CUSTEIO/INVESTIMENTO - RESUL.PRIM = 2", "CU..."
* vinculacoes_distintas ^<int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
* saldo_diario         ^<dbl> -20.79, -724.21, -8000.00, -76.95, 15188.99,...
* pagamento_por_saldo  ^<dbl> -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, ...
```

### 2.7.3 Limimtes de Saque Por Tipo De Documento

- Informações diárias de **saldo disponível** por tipo de documento. Os tipos de documentos são: DF, DR, GF, GP, GR, NL, NS, OB, PF e RA.
- Útil para a questão **d**.
- Derivada da base `lim_saque`.
- Observações: 254.667
- Campos: 7

```
* tipo_de_documento  `<chr> "DF", "DF", "DF", "DF", "DF", "DF", "DF", "DF", "...`
* NO_DIA_COMPLETO    `<chr> "01/02/2017", "01/02/2017", "01/02/2017", "01/02/..."
* NO_UG              `<chr> "COORD. REG. NOROESTE DO MATO GROSSO/MT", "COORD..."
* NO_ORGAO           `<chr> "FUNDACAO NACIONAL DO INDIO", "DEPARTAMENTO DE PO..."
* NO_ITEM_INFORMACAO `<chr> "LIMITES DE SAQUE (OFSS, DIVIDA, BACEN E PREV)", "...`
* NO_FONTE_RECURSO   `<chr> "RECURSOS ORDINARIOS", "RECURSOS ORDINARIOS", "TX..."
* saldo_diario       `<dbl> -35.47, -2089.65, -3644.19, -12.01, -14.94, -31.3..."
```

## Chapter 3

# Quadro analítico histórico

### 3.1 Disponibilidades Líquidas

Para responder as questões (a) e (b) sobre disponibilidades líquidas das UGs, elaboramos:

- Gráficos das séries temporais das disponibilidades líquidas de cada UG e Fonte de Recurso.
- Gráficos das séries temporais das disponibilidades líquidas consolidadas por UG.
- Indicadores de disponibilidades líquidas negativas para as UGs:
  - Proporção do tempo do período analisado com disponibilidade negativa.
  - Duração da disponibilidade negativa atual.

### 3.2 Classificações Orçamentárias e Financeiras

(c) Despesas e Vinculações

### 3.3 Disponibilidades Líquidas por Vinculação

(d) revisitamos...

### 3.4 Comportamento do Caixa por Movimentação

- Gráficos das séries temporais dos saldos e obrigações de cada UG e Fonte de Recurso.
- Gráficos das séries temporais dos saldos e obrigações consolidados por UG.
- Gráfico dos tempos entre duas movimentações de grande porte por UG e por Órgão.



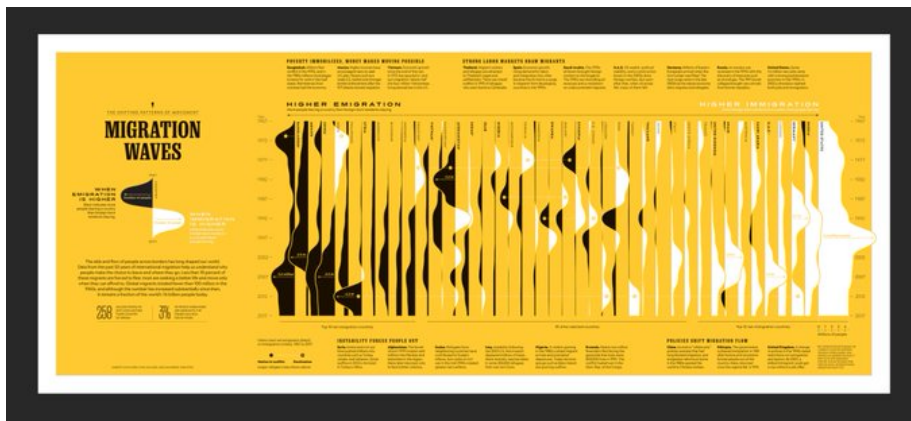
## Chapter 4

# Quadro analítico preditivo



# Referências

- NatGep migration waves



<https://twitter.com/aLucasLopez/status/1153646875427385344?s=20>

- Diagrama de bolhas em D3 do Jim Vallandingham.
- <https://www.nature.com/articles/s41598-017-13448-3>
- <https://github.com/gm-spacagna/deep-ttf/>