Produto CEPAL #01: - Product 1. SAS and Python R programming and description of the first results.

Avaliação da Cobertura dos Dados do BNDES Transparência na Receita Federal

Patrick Alves

**Resumo**:

Apresentamos a organização em formato SAS dos dados abertos da Receita Federal (RF) e do Banco Nacional de Desenvolvimento Econômico e Social (BNDES). Essas informações são microdados de empresas disponibilizados, respectivamente, pelo portal de dados abertos[[1]](#footnote-1) do Governo Federal e pelo portal da Transparência[[2]](#footnote-2) do BNDES. A organização e documentação desses microdados é uma etapa importante em avaliações de financiamentos do BNDES, performance econômica de empresas e probabilidade de inadimplência.

Neste documento apresentamos em cada uma das tabelas existentes, os campos disponíveis e os relacionamentos existente entre as tabelas. Será detalhado também o processo de recuperação dos CNPJs que estão censurados na base de dados de operações indiretas do BNDES. Por fim, apresentados algumas estatísticas descritas advindas do cruzamento de informações entre o BNDES e Receita Federal.

**Introdução**

A necessidade de maior *accountability* tem estimulado as iniciativas de bases de dados abertos contendo microdados de interesse público. Entre essas iniciativas, podemos citar portal o Cadastro Nacional de Pessoas Jurídicas da Receita Federal e o Portal Transparência do BNDES, contendo microdados de empresas. A unidade observacional dos dados disponibilizados é a operação de crédito, contendo o CNPJ censurado das empresas e as condições da operação de crédito (taxa de juros, prazo, amortização, valor do contrato). A unidade observacional das informações da Receita Federal é a matriz empresa, a empresa final e o sócio da empresa, dependendo da tabela relacional. Nosso na concatenação desses dados é desenvolver uma proxy para o risco de crédito dessas operações, relacionando esta medida de risco com as operações do BNDES.

A Receita Federal disponibiliza 03 tabelas da microdados: Empresa, Filiais e Sócios (Tabela 1). A tabela de empresas (RF\_EMPRESA) possui aproximadamente 40 milhões registros correspondendo a matriz das empresas, os dados de filiais (RF\_FILIAIS) possuem aproximadamente 42 milhões de observações correspondendo as plantas locais, e a tabela de sócios (RF\_SOCIOS) possuem aproximadamente 18 milhões de observações, correspondendo a estrutura societárias das empresas. As informações de empresas (RF\_EMPRESA), filiais (RF\_FILIAIS) e sócios (RF\_SOCIOS) possuem como chave comum o CNPJ raiz da empresa (08 primeiros dígitos do CNPJ).

A tabela de operações indiretas (BNDES\_INDIR\_AUTO) do BNDES possui em torno de 269 mil observações, correspondendo a todas as operações de crédito indiretas no período de 2017 a 2022 (Tabela 1). A tabela de operações não-automáticas (BNDES\_NAOAUTO) possuem aproximadamente 20 mil observações, onde o CNPJ não se encontra censurado. A tabela de operações da administração pública (BNDES\_ADM\_PUB), tais como prefeituras, e autarquias, possuindo aproximadamente 4 mil observações. As informações do BNDES possuem como chave comum o CNPJ raiz da empresa (08 primeiros dígitos do CNPJ), o qual pode ser utilizado também para concatenar informações da Receita Federal.

Tabela 1 – Tabelas de Microdados utilizados

|  |  |  |  |
| --- | --- | --- | --- |
| **Origem** | **Tabela** | **Nº de OBS** | **CNPJ** |
| Receita Federal (RF) | RF\_EMPRESA | 40.453.740 | Completo |
|  | RF\_FILIAIS | 42.780.915 | Completo |
|  | RF\_SOCIOS | 18.172.350 | Completo |
| BNDES | BNDES\_INDIR\_AUTO | 269.143 | Censurado |
|  | BNDES\_NAOAUTO | 20.366 | Completo |
|  | BNDES\_ADM\_PUB | 4.632 | Não tem CNPJ |

1. **Metodologia de Recuperação dos CNPJs censurados**

Os dados sobre as operações de financiamento indireto do BNDES possuem os seus CNPJs[[3]](#footnote-3) parcialmente censurados (Tabela 2). Por este motivo foi necessário aplicar um tratamento para recuperação dos CNPJs completos. Esse tratamento utiliza as razões sociais das empresas e 5 dígitos não censurados existente nos dados do BNDES (ver exemplo na Tabela 2).

A razão social vir escrita de formas diferentes em bancos de dados diferentes, dependendo de abreviações, acentuação ortográficas e uso de caracteres especiais. Assim, o primeiro passo será retirar todas as acentuações gráficas, espaços, caracteres especiais e converter todos os caracteres para o formato minúsculo. Após esse tratamento, também escolhemos utilizar somente os 15 primeiros dígitos da razão social das firmas.

A Tabela 2 apresenta alguns exemplos do formato dos CNPJ censurados nas operações indiretas. As colunas “Razão Social Tratada” e “CNPJ 5 Dígitos” demostram os passos iniciais de tratamento dos campos visando recuperação o CNPJ. As mesmas colunas “Razão Social Tratada” e “CNPJ 5 Dígitos” são criadas na base de firmas da Receita Federal. Em seguida as duas tabelas são concatenadas através desses dois campos. As linhas de intercessão entre as tabelas possuem também o CNPJ completo, existente nos dados da Receita Federal.

Tabela 2 – Exemplo do Padrão de Censura do CNPJ

|  |  |  |  |
| --- | --- | --- | --- |
| Razão Social | CPF/CNPJ | Razão Social Tratada | CNPJ 5 Dígitos |
| ROISTER ALIMENTOS E BEBIDAS LTDA | \*\*.\*24.881/0001-\*\* | ROISTERALIMENTO | 24881 |
| TELES HIDRAULICA LTDA | \*\*.\*75.059/0001-\*\* | TELESHIDRAULICA | 75059 |
| WOLFSTORE INDUSTRIA TEXTIL LTDA | \*\*.\*19.663/0001-\*\* | WOLFSTOREINDUST | 19663 |
| ST.MODAS INDUSTRIA E COMERCIO S/A. | \*\*.\*84.693/0001-\*\* | STMODASINDUSTRI | 84693 |
| M.NOBRE PRODUTO DE LIMPEZA LTDA | \*\*.\*56.185/0001-\*\* | MNOBREPRODUTODE | 56185 |
| W.P.X. LOCACOES S.A. | \*\*.\*12.519/0001-\*\* | WPXLOCACOESSA | 12519 |

A Tabela 3 apresenta o resultado da recuperação de CNPJs. Vemos que a quantidade de CNPJs recuperados aumenta ao longo do tempo. Em média foi possível recuperar 78,8% da quantidade de CNPJs, correspondendo a 74% do valor das operações e 73,6% dos valores desembolsados.

Tabela 3 – Tabelas de Microdados utilizados

|  |  |  |  |
| --- | --- | --- | --- |
|  | Nº CNPJs | Valor Contrato (R$) | Valor Desembolso (R$) |
| CNPJ recuperados | 269.143 | 121.418.642.690 | 112.846.872.723 |
| % | 78,80% | 73,92% | 73,58% |

Tabela 4 – Tabelas de Microdados utilizados

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Número de CNPJ** | | **Valor Contratado (R$ Milhão)** | | **%** | |
| ANO | **Recuperado** | **Total** | **Recuperado** | **Total** | **Recuperado** | **Total** |
| 2017 | 22.178 | 31.420 | 15.039 | 21.930 | 70,6% | 68,6% |
| 2018 | 18.079 | 24.034 | 12.270 | 17.316 | 75,2% | 70,9% |
| 2019 | 14.351 | 17.685 | 8.198 | 11.013 | 81,1% | 74,4% |
| 2020 | 30.069 | 35.168 | 12.788 | 16.554 | 85,5% | 77,2% |
| 2021 | 19.355 | 22.338 | 14.162 | 18.117 | 86,6% | 78,2% |
| 2022 | 24.569 | 29.960 | 26.229 | 34.937 | 82,0% | 75,1% |

Figura 1 – Taxa de Recuperação de CNPJ ao longo do tempo

1. **Campos da Receita Federal**

A Tabela 5 apresenta os campos existentes na tabela de empresas (RF\_EMPRESA). Os campos NUMCNPJ\_RAIZ e RAZAO\_SOCIAL são utilizados no processo de recuperação dos CNPJ censurados do BNDES. O capital social das empresas (VR\_CPT\_SOC), a natureza jurídica (COD\_NATJUR) e o porte das empresas (PORTE\_EMPRESA) constituem importantes variáveis de controle para inclusão em modelos de aprendizado de máquina.

Tabela 5 – Tabela RF\_EMPRESA

|  |  |
| --- | --- |
| **Campo** | **Descrição** |
| NUMCNPJ\_RAIZ | CNPJ Raiz (8 DÍGITOS DO CNPJ) |
| RAZAO\_SOCIAL | NOME EMPRESARIAL DA PJ |
| COD\_NATJUR | Código da NATUREZA JURÍDICA (aba COD\_NATJUR) |
| QUAL\_RESPO | QUALIFICAÇÃO DO RESPONSÁVEL (aba COD\_QUALSOCIOS) |
| VR\_CPT\_SOC | CAPITAL SOCIAL DA EMPRESA (R$) |
| PORTE\_EMPRESA | 00 – NÃO INFORMADO;  01 - MICROEMPRESA;  03 - PEQUENO PORTE;  05 - DEMAIS |
| ENTE\_RESPO | ENTE FEDERATIVO RESPONSÁVEL (Prefeituras, Autarquias etc.) |
| CPF\_ENTE\_RESPO | CPF DO ENTE FEDERATIVO RESPONSÁVEL |

A Tabela 6 apresenta os campos da tabela de filiais (RF\_FILIAIS). Os alguns domínios existentes nos campos SIT\_CAD e COD\_MOTIVO podem caracterizar empresa com possibilidade de inadimplência. A data da empresa (DT\_INICIO), setor de atividade econômica (CNAE\_FISCAL), unidade da federação (UF) e município (MUNICÍPIO) podem ser aproveitadas como variáveis de controle.

Tabela 6 – Tabela RF\_FILIAIS

|  |  |
| --- | --- |
| **Campo** | **Descrição** |
| NUMCNPJ | CNPJ (OITO DÍGITOS DO CADASTRO NACIONAL DA PESSOA JURÍDICA). |
| SIT\_CAD | SITUAÇÃO CADASTRAL:  01 – NULA;  02 – ATIVA;  03 – SUSPENSA;  04 – INAPTA;  08 – BAIXADA |
| DATA\_SITUACAO | DATA DO EVENTO DA SITUAÇÃO CADASTRAL |
| COD\_MOTIVO | CÓDIGO DO MOTIVO DA SITUAÇÃO CADASTRAL (ABA COD\_MOTIVO) |
| DT\_INICIO | DATA DE INÍCIO DA ATIVIDADE |
| CNAE\_FISCAL | PRINCIPAL CÓDIGO DA ATIVIDADE ECONÔMICA |
| UF | SIGLA DA UNIDADE DA FEDERAÇÃO DO ESTABELECIMENTO |
| CEP | CEP DO LOGRADOURO DO ESTABELECIMENTO |
| MUNICÍPIO | MUNICÍPIO DO ESTABELECIMENTO |
| NUMCNPJFILIAL | CNPJ COMPLETO (14 dígitos) da filial |

Tabela 7 – Tabela RF\_SOCIOS

|  |  |
| --- | --- |
| **Campo** | **Descrição** |
| NUMCNPJ\_RAIZ | CNPJ (OITO DÍGITOS DO CADASTRO NACIONAL DA PESSOA JURÍDICA) |
| COD\_SOC | CÓDIGO DO IDENTIFICADOR DE SÓCIO (1=PJ, 2=PF, 3=ESTRANGEIRO) |
| NOME\_SOC | NOME DO SÓCIO (NOME PARA PF OU RAZÃO SOCIAL PARA CASO PJ) |
| NUMCPFCNPJ\_SOC | CNPJ/CPF DO SÓCIO CPF OU CNPJ DO SÓCIO (EXCETO ESTRANGEIRO) |
| QUAL\_SOC | QUALIFICAÇÃO DO SÓCIO CÓDIGO DA QUALIFICAÇÃO DO SÓCIO |
| DATA\_SOC | DATA DE ENTRADA SOCIEDADE DATA DE ENTRADA NA SOCIEDADE |
| CPF\_REPR\_LEGAL | CPF DO REPRESENTANTE LEGAL (SOMENTE PARA SÓCIO PJ) |
| NOME\_REPR\_LEGAL | NOME DO REPRESENTANTE LEGAL (SOMENTE PARA SÓCIO PJ) |
| QUAL\_LEGAL | CÓDIGO DA QUALIFICAÇÃO DO REPRESENTANTE LEGAL |
| FAIXA\_ETARIA | FAIXA ETARIA DO SOCIO (Baseada na data de nascimento) |
| - 0 para não se aplica |
| - 1 Entre 0 a 12 anos; |
| - 2 Entre 13 a 20 anos; |
| - 3 Entre 21 a 30 anos; |
| - 4 Entre 31 a 40 anos; |
| - 5 Entre 41 a 50 anos; |
| - 6 Entre 51 a 60 anos; |
| - 7 Entre 61 a 70 anos; |
| - 8 Entre 71 a 80 anos; |
| - 9 Maiores de 80 anos. |

1. **Campos do BNDES**

A Tabela 8 apresenta os campos existente na tabela de operações envolvendo a administração pública (BNDES\_ADM\_PUB).

Tabela 8 – Operações da Administração Publica

|  |  |
| --- | --- |
| **Coluna** | **Descrição** |
| DATA\_NIVEL\_ATUAL | Data da operação de crédito |
| ENTE\_PUBLICO | Nome do ente público |
| MODALIDADE\_OPERACAO | Modalidade da Operação de Crédito  Ex: BNDES FINEM DIRETO, BNDES FINAME |
| MUNICIPIO | Nome do Município |
| NIVEL\_ATUAL | Ex: APROVADA; C/CONSULTA; CONTRATADA; EM ANÁLISE |
| OBJETIVO\_PROJETO | Descrição textual do objetivo do projeto |
| PROGRAMA | Ex: PROJ INV ST PUBLICO, FINEM GENERICO, PMATS,  SEG. PUBLICA, DESENV INTEG MUNICIP, BK AQUISICAO |
| SALDO\_LIBERAR | Valor do projeto ainda a ser liberado |
| SITUACAO\_OPERACAO | ATIVA ou ATIVA-UTILIZADA |
| UF | Unidade da Federação do Ente Público |
| VALOR\_DESEMBOLSADO | VALOR já DESEMBOLSADO (R$) |
| VALOR\_OPERACAO | VALOR da Operação (R$) |

A Tabela 9 apresenta os campos existente na tabela de operações não-automáticas do BNDES (BNDES\_NAO\_AUTOM).

Tabela 9 – Operações Não-Automáticas BNDES

| **Coluna** | **Descrição** |
| --- | --- |
| CNPJ | CNPJ (não censurado) da empresa |
| CNPJ\_Instituicao | CNPJ (não censurado) da instituição intermediadora |
| Cliente | Razão Social da Empresa |
| Codigo\_Municipio | Código do Município |
| Custo\_Financeiro | Custo Financeiro do Projeto:  SELIC, TLP, TAXA FIXA, TJLP, US$ / CESTA, 75% da SELIC, FAN, OUTROS |
| Data\_Contrato | Data da contratação da operação |
| Descricao\_Projeto | Descrição textual do objetivo do projeto |
| Fonte\_Recurso | Ex: RECURSOS LIVRES PRÓPRIOS, RECURSOS LIVRES FAT, RECURSOS VINCULADOS FAT DEPÓSITOS ESPECIAIS, RECURSOS LIVRES TESOURO, RECURSOS VINCULADOS FUNDO CLIMA |
| Forma\_Apoio | DIRETA |
| Inovacao | Sim ou Não |
| Instituicao\_Credenciada | Nome da Instituição Financeira Credenciada |
| Instrumento\_financeiro | Instrumento Financeiro |
| Juros | Taxa de Juros (%) |
| Modalidade\_Apoio | REEMBOLSÁVEL, Não-REEMBOLSÁVEL, etc. |
| Municipio | Município da empresa |
| Natureza\_Cliente | Público ou Privado |
| Numero\_Contrato | Número do Contrato |
| Porte\_Cliente | Porte do Cliente |
| Prazo\_Amortizacao | Prazo de Amortização da Operação |
| Prazo\_Carencia | Prazo de Carência da Operação |
| Produto | BNDES FINEM  BNDES NÃO REEMBOLSÁVEL  BNDES FINAME  BNDES DEBENTURES SIMPLES  BNDES LIMITE DE CRÉDITO  BNDES DEBENTURES SUSTENTAVEIS E DE INFRAESTRUTURA  BNDES MICROCRÉDITO  BNDES GARANTIA  OPERAÇÃO FINANCEIRA  BNDES DIRETO 10  BNDES CRÉDITO DIRETO MÉDIAS EMPRESAS  BNDES EMPRÉSTIMO PONTE  BNDES PROJECT FINANCE  BNDES EXIM PÓS-EMBARQUE  OUTROS |
| Setor\_BNDES | ENERGIA ELÉTRICA  COMÉRCIO E SERVIÇOS  AGROPECUÁRIA  TELECOMUNICAÇÕES  MECÂNICA  TRANSPORTE FERROVIÁRIO  METALURGIA E PRODUTOS  ATV. AUX. TRANSPORTES  QUÍMICA E PETROQUÍMICA  OUTROS TRANSPORTES  CONSTRUÇÃO  SERV. UTILIDADE PÚBLICA  ALIMENTO E BEBIDA  TRANSPORTE RODOVIÁRIO  TÊXTIL E VESTUÁRIO  OUTRAS  CELULOSE E PAPEL  MATERIAL DE TRANSPORTE  EXTRATIVA |
| Setor\_CNAE | Setor de Atividade Econômica CNAE |
| Situacao\_Contrato | ATIVO ou LIQUIDADO |
| Tipo\_Excepcionalidade | CONDIÇÕES FINANCEIRAS E OPERACIONAIS, CONDIÇÕES DE CRÉDITO, COBRANÇA DE COMISSÕES, GARANTIAS, ou NENHUMA |
| Tipo\_Garantia | REAL, PESSOAL, NÃO SE APLICA, DE NATUREZA ESPECÍFICA OU MISTA, COMPROMISSO DE DESEMPENHO FINANCEIRO, 'COVENANTS', DEBÊNTURE QUIROGRAFÁRIA, SEM REGISTRO DE GARANTIA, DEFINIDA PELO AGENTE FINANCEIRO |
| UF | Unidade da Federação da Empresa |
| Valor\_Contrato | Valor do Contrato (R$) |
| Valor\_Desembolso | Valor do Desembolso (R$) |

A Tabela 10 apresenta os campos da tabela de sócios (RF\_ SOCIOS). Os alguns domínios existentes nos campos SIT\_CAD e COD\_MOTIVO podem caracterizar empresa com possibilidade de inadimplência. A data da empresa (DT\_INICIO), setor de atividade econômica (CNAE\_FISCAL), unidade da federação (UF) e município (MUNICÍPIO) podem ser aproveitadas como variáveis de controle.

Tabela 10 – Operações Indiretas BNDES

| **Cliente** | **Razão Social da Empresa** |
| --- | --- |
| CPF/CNPJ | CNPJ censurado |
| UF | Unidade da Federação da Empresa |
| Município | Município |
| Município - código | Código do Município (Código do IBGE) |
| Data da contratação | Data da contratação da operação |
| Valor da Operação em R$ | Valor da Operação em R$ |
| Valor Desembolsado R$ (\*) | Valor Desembolsado R$ (\*) |
| Fonte de recurso (desembolsos) | RECURSOS LIVRES - PRÓPRIOS  RECURSOS LIVRES - FAT  RECURSOS VINCULADOS - FAT DEPÓSITOS ESPECIAIS  RECURSOS LIVRES - TESOURO  RECURSOS VINCULADOS - FUNDO CLIMA |
| Custo financeiro | Ex: SELIC, TLP, TAXA FIXA, TJLP, US$ / CESTA, 75% da SELIC, FAN, OUTROS |
| Juros | Juros anuais (%) |
| Prazo - Carência (meses) | Prazo de Carência em meses |
| Prazo - Amortização (meses) | Prazo de Amortização em meses |
| Modalidade de Apoio | REEMBOLSÁVEL |
| Forma de Apoio | INDIRETA |
| Produto | BNDES AUTOMÁTICO  BNDES FINAME  BNDES FINAME AGRÍCOLA  BNDES FINAME LEASING  BNDES CRÉDITO RURAL |
| Instrumento Financeiro | Instrumento Financeiro |
| Inovação | Sim / Não |
| Área Operacional | AREA DE OPERACOES E CANAIS DIGITAIS |
| Setor CNAE | Setor CNAE |
| Setor BNDES | Setor BNDES |
| Porte do cliente | MÉDIA  GRANDE  PEQUENA  MICRO |
| Natureza do cliente | PRIVADA  ADMINISTRAÇÃO PÚBLICA DIRETA - GOVERNO MUNICIPAL  PÚBLICA INDIRETA |
| Instituição Financeira Credenciada | Instituição Financeira Credenciada |
| CNPJ do Agente Financeiro | CNPJ do Agente Financeiro |
| Situação da Operação | Situação da Operação |

1. Cruzamento do BNDES com a RECEITA FEDERAL

Considerando as operações indireta do BNDES, foram financiadas aproximadamente 107 mil empresas no período 2017 a 2022, correspondendo a um capital social somado de 354 bilhões de reais.

Tabela 11 – Porte da Firma segundo Classificação da Receita Federal

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Firmas Financiadas pelo BNDES | | | | Demais Firmas | | | |
| Porte da Firma | Total  (Mil) | % | Cap. Social  (Milhões) | % | Total  (Mil) | % | Cap. Social  (Milhões) | % |
| Não informado | 0 | 0% | 0 | 0% | 57 | 0% | 125 | 0% |
| Microempresa | 70 | 66% | 58.626 | 17% | 29.198 | 72% | 23.123.268 | 62% |
| Pequeno Porte | 21 | 20% | 5.349 | 2% | 1.054 | 3% | 2.247.097 | 6% |
| Demais | 16 | 15% | 289.993 | 82% | 10.038 | 25% | 12.121.567 | 32% |
| total | 107 | 100% | 353.967 | 100% | 40.347 | 100% | 37.492.057 | 100% |

No período de 2017 a 2022 o percentual somado de micro e pequenas empresas (85,3%) é maior que o percentual de micro e pequenas empresas na população (75,0%), indicando uma orientação para apoio de empresas de menor porte.

Figura 2 – Situação Cadastral da Receita Federal

Entre as empresas financiadas pelo BNDES, a quantidade de empresas sob situação de risco (3 – SUSPENSA, 4 – INAPTA, 08 – BAIXADA) é muito baixa, comparado com as demais empresas. Para as firmas financiadas pelo BNDES, o percentual de empresas suspensas, inaptas ou baixadas é de 9,2%, enquanto para as demais firmas o mesmo percentual é de 60,2%.

Vamos considerar os status da receita federal: 3 – SUSPENSA, 4 – INAPTA, 08 – BAIXADA, como uma proxy de inadimplência das empresas. Assim será possível mensurar o risco de crédito das firmas financiadas pelo BNDES.

Tabela 12 – Descrição de Motivo da Situação Cadastral

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **DESCRIÇÃO MOTIVO** | Firmas Financiadas pelo BNDES | | Demais  Firmas | |
| TOTAL | % | TOTAL | % |
| Baixa Indeferida RFB, Aguardando Análise Sefaz e Indeferida SEFIN | 60.260 | 56,29% | 8.424.789 | 20,88% |
| Baixa Indeferida RFB, Deferida Sefaz E Aguardando Análise RFB | 27.553 | 25,74% | 25.254.694 | 62,59% |
| Registro Cancelado | 15.595 | 14,57% | 1.193.927 | 2,96% |
| Sem Motivo | 2.211 | 2,07% | 1.584 | 0,00% |
| Anulação Por Vícios | 777 | 0,73% | 1.478.584 | 3,66% |
| Encerramento Da Falência | 498 | 0,47% | 278.637 | 0,69% |
| Localização Desconhecida | 90 | 0,08% | 8.067 | 0,02% |
| Anulação Online De Ofício | 32 | 0,03% | 6.641 | 0,02% |
| DeferidopPelo Convenente, Aguardando Análise RFB | 28 | 0,03% | 534.606 | 1,33% |
| Transferência Filial Condição Matriz | 7 | 0,01% | 97.530 | 0,24% |
| Omissa Contumaz | 2 | 0,00% | 1.984 | 0,00% |
| Demais Situações | 6 | 0 | 3.065.638 | 0 |
| **Total** | 107.065 |  | 43.412.319 |  |

1. **Metodologia de Estimação de Risco de Crédito**

A essência da atividade bancária consiste na captação de depósitos através poupanças e outras fontes, e na concessão de empréstimos, o que ajuda na alocação de fundos de maneira mais eficiente na economia. A intermediação bancária garante às instituições financeiras um uma margem de lucro, também denominada spread bancário e comumente definida como receita de juros menos despesas de juros. A adequada mensuração do risco de crédito é uma das atividades fundamentais das instituições financeiras, a qual permite mensurar os juros bancários de forma justa, e alocar eficientemente os recursos financeiro no mercado.

As instituições financeiras consistentemente perseguem metodologias mais eficientes para mensuração do risco de crédito. A busca por mensurações de risco de crédito mais precisas faz uso intenso da exploração de novas formas de dados e de novas metodologias de estimação da probabilidade de inadimplência. Nossa proposta da estimação do risco de crédito caminha em ambos os sentidos. Primeiramente explorando novas fontes de dados, e também pela utilização do aprendizado de máquina.

Tabela 13 – Variáveis Explicativas (*Features*) e Dependentes (Target) do Modelo de Aprendizado de Máquina para Risco de Crédito

|  |  |  |
| --- | --- | --- |
|  | Variáveis Explicativas  BNDES | Variáveis Explicativas  Receita Federal |
| Situações Cadastrais (SIT\_CAD):  3 – SUSPENSA  4 – INAPTA  8 – BAIXADA  Independente da data da situação | * Porte Receita * Idade da Firma * Idade da Situação Cadastral * Natureza Jurídica * Número de Sócios * Idade média dos sócios * Estrangeiros Sócio (Dummy) * Qualificação do Sócio Responsável | * Taxa de Juros * Valor Contratado * Valor Desembolsado * Prazo de Amortização * Prazo de Carência * Modalidade de Apoio * Produto * Inovação (Dummy) * Porte BNDES (Peq, Média, Grande) * UF * CNAE |
| Situações Cadastrais (SIT\_CAD):  3 – SUSPENSA  4 – INAPTA  8 – BAIXADA  Onde data da situação é posterior à data de financiamento, mas não é posterior à data do fim do contrato |

1. **Conclusão**:

Este documento apresentou a metodologia de tratamento dos dados públicos disponibilizados no site BNDES Transparência, visando tornar estes microdados operacionais para estimação do risco de crédito das operações do BNDES. Neste sentido, detalhamos as informações dos financiamentos do BNDES, e as informações obtidas da Receita Federal. Utilizando ambas as fontes de dados, descrevemos a metodologia de recuperação dos CNPJs censurados das operações indiretas do BNDES.

1. **Descrição dos Scripts Entregues**:

Tabela 14 – Tabela de artefatos entregues como produto.

|  |  |
| --- | --- |
| Script | Descrição |
| extraindodados.html | Script de Download dos dados da Receita Federal |
| importacao.html | Script de Unificação e Organização dos dados da Receita Federal |
| criacao bndes.sas | Importação e Organização dos dados do BNDES |
| Organizacao BNDES.sas | Metodologia de Recuperação dos CNPJ Censurados |
| DICIONARIO\_RECEITA.xlsx | Dicionário Receita Federal |

1. https://dadosabertos.bcb.gov.br/ [↑](#footnote-ref-1)
2. <http://www.bndes.gov.br/wps/portal/site/home/transparencia/centraldedownloads> [↑](#footnote-ref-2)
3. Cadastro Nacional de Pessoa Jurídica da Receita Federal. [↑](#footnote-ref-3)