

CURSO SOBRE MANUSEIO DE BASES DE DADOS DO GOVERNO FEDERAL NO SOFTWARE “R”

Desafios na análise dados

Professores:

Marcelo Prudente

Rafael Giacomini

Cadastrros e registros administrativos

- Principais desafios do cruzamento de dados de registros administrativos:
 - **Qualidade dos dados**
 - **Ausência de informação sobre a base de dados**
 - **Ausência de informação de campos na base de dados**
 - **Tratamento de dados – registros duplicados**
 - **Características especiais**
 - **Escolha das chaves**
 - **Escolha do método**
 - **Análise do resultado**

Principais desafios do cruzamento de dados de registros administrativos

➤ Qualidade dos dados

Exemplo: Registros do SDPA com problemas, cruzamento com dados do CPF (CPF como chave primária)*

nomerequerente_SDPA	nom_contribuinte_CPF	nomemaerequerente_SDPA	nom_mae_CPF	dt_nasc_sdpa	dt_nasc_cpf
ERISVALDO FRAZAO COSTA	ELISVALDO FRAZAO COSTA	MARIA POMPEU DOS SANTOS	MARIA CARLIZA DE OLIVEIRA	19810906	19800906
JACINILDE BELO NOGUEIRA	JACILEIDE BELO FERREIRA	LAURA GOMES DE SENA	LAURA GOMES DE SENA	19750806	19750906
ALUIZIO SOBRINHO DE OLIVEIRA	ALUISIO SOBRINHO DE OLIVEIRA	ROZALIA FERREIRA MAFRA	ROSALIA FERREIRA MAFRA	19630331	19630313
MARTILIANE DO SOCORRO DOS SANTOS BACE	MARTIANE DO SOCORRO DOS SANTOS BACELAR	MARIA DA GLORIA FRAZAO COSTA	MARIA DA GLORIA FRAZAO COSTA	19701110	19721210
ROSILENE POMEU DOS SANTOS	LENALDO DIAS DOS SANTOS	NULL	ANTONIA PAVAO	19780314	19841112
ZACARIAS FURTADO MARQUES	ADEMIR MACHADO COSTA	VALDIRENE DE JESUS CORDEIRA PESSOA	CATARINA PAIM DE SOUZA	19830730	19861028
BRAZ MARINHO MAFRA	BRAS MARINHO MAFRA	MARIA IRACI SOBRINHO DE OLIVEIRA	MARIA IRACI SOBRINHO DE OLIVEIRA	19901026	19900402
GUSTAVO FELIPE SARGES RAMOS	CRISTIANA DA SILVA DIAS	MARIA DOLORES DOS SANTOS FURTADO	MARIA MACHADO COSTA	19900712	19560303
EDIVALDO GOMES DA SILVA	EDVALDO GOMES DA SILVA	GREGORIA BELO NOGUEIRA	GREGORIA PEREIRA BELO	19570604	19570704
EDICLEIDE PESSOA FERREIRA	MARIA DA SILVA	MARIA DAS GRACAS SARGES RAMOS	LUZIA DE JESUS DA SILVA DIAS	19770414	19650223
LAERCIO GONCALVES BENASSULI	LAELCIO GONCALVES BENASSULI	MARIA ONEIDE PEREIRA DOS SANTOS	MARIA ONEIDE PEREIRA DOS SANTOS	19580513	19570513

Informações incorretas:

- Nem sempre o documento informado pertence a própria pessoa;
- Grafias diversas dos nomes, (erro de digitação);
- Dados pessoais divergentes;
- Alterações de nomes não refletidas em todas as bases de dados; etc.

Provável alteração de nome por casamento, não atualizada em um dos cadastros

*dados alterados para preservar o sigilo das informações

Principais desafios do cruzamento de dados de registros administrativos

➤ Ausência de informação sobre a base de dados

Em geral, o acesso as bases de dados não vem acompanhado de documentos como layout, dicionário de dados, descrição das variáveis, etc.

OB1201607 - Bloco de notas

Arquivo Editar Formatar Exibir Ajuda

0000C300008000000307020160628000000000JOAZ ZOANTAS SIROS	J CANA DS SAS DOZ AANTOS	192912282016062638351447587
000C1100246000000202920160513000000000MAR23QIMMOES DA SIZAM CARDOSO	D JINA SEODES DAM ALVA	194407302016040900199661561
0000C200080000000148720160629000000000DJAI OAO PEREIRA PA OS	EDE TRGB000PEREIRA0PASSOS	192211142016062650305557572
0000C900128000000125620160614000000000NLS5002OLIVIMIL S OOS	J210TE OE JEIRA029ATOS	196112242016051961939994500
0000C900129000000125720160615000000000ARG0000 MANDS DE GRITO	SURDA JUSTANA K DJESUS	193212102016060776852610578
0000C900130000000125820160617000000000RAIN OO VEMPA DO B'BCUMENTO	GI AKSTIA SJACHAVA JESUS	197603282016060495936610500
000C12001390000004586201606070000000000MARIIP ODRIGOLS DA SOUZDA JSTA	NAT23A ROZUOGUE JTA SILVA	193310052016051788335593515
000C12001400000004587201606080000000000ANA O JESU2560UZA	JHE JE DE JESUS	197404272016053001888151552
000C12001410000004588201606081283991974ZITA DERNAL257 DA SIHA	VIA SWIA SIMNANA DE SOUSA	192601222016052673446874534
000C12001420000004589201606080946434131JOVIA ANO JTA DA DA VA	REIA RODRIGUES JESCRUZ	191703162016040702943553590
000C12001430000004590201606090000000000JURJO AMARQUNO DIERSOZA	HE JESTA PAIEIBRIAE SOUZA	194007122016060888782336504
000C12001440000004591201606100000000000JOSIARH VES JEDIM	M3 A JAFOS CAIO	195006122016060262359916572
000C12001450000004592201606100000000000FILIJJO PEBA SA JUDA O	ROJELINAE TRIA DE JESUS	194011292016060507398298587
000C12001460000004593201606100000000000MAR3ROGNEP DA IZ VA	MEIRA SACINA DA JZCVA	195012122016052712674661895
000C12001470000004594201606130000000000DEONS E ALMPEBA JZJTA JTOA REIRA	C LUZANA CORREIA HE ALMEIDA	193601222016060435698985553
000C12001480000004595201606130000000000LEOCAND NEVES DB000UJA	SA A MAOA DE JDA S	192711122016060927856356634
000C12001490000004596201606130000000000LUCIVA FERRA SI CARREJO	BALFERRIOA TRIMESUE	199706172016061000000000000
000C120015000000045972016062104975807565INVAC SOUT BRANDAO	MARIA RIBOSRO BRARIO	192611102016061606306608591
000C12001510000004598201606220000000000FABLEOCARDOCERDA HA VA SANTOS	ARTE ESUSA DA RIVA	198607172016061102502189500
000C12001520000004599201606300000000000ASTRIZILDO JIVIER ROS	BERNARDIA MARIA DA JESUS	194210182016061602788633530
0000C6001650000001922201606010000000000LINERARA BAVZOSA C MEIA	7 OITA JJA BARBRIA	193603112016052787730278553
0000C6001660000001923201606080000000000ROSIEVE) ANTOSTIO XOTIER	ALVPOLDIM CELES112 DA LUZ	191508302016050501887817573
0000C60016700000019242016060800000000000MARIIPDAS DA ES ENO SILVE OLIVEIRA	ORRA HODMINA SAN JNA	192704082016052574944878591
0000C6001680000001925201606150000000000ZELO CANB2A DQJISANYOS	CAYJIDJ ENDES DE ARJUS	195303292016051944623577520
0000C6001700000001927201606290000000000JOAA M JO JPMARJ	TE BOM CEJULINILVAIRES	193006122016062214545357504
0000C5002300000004054201606270000000000MIL MOHA DA DAZITUA	AMIRNANRWJ RA SA MATA	196811202016062109526661893
0000C5002310000004055201606270000000000OSVATA SOTROI DE JS	CRECITA PAO LUZOUZA JESUS	195605152016062011871016134
0000C5002310000004056201606280000000000ANA BRANDAO RAMOS	JORAM M AIA FERN	192008202016010467432972591
0000C5002320000004057201606280000000000DEOC MAIANDMAJSSIDO. QZ OB ELUZA	DOMINGAROSA SZ JESUS	191612202016052501258567504
0000C5002320000004058201606290000000000MARVA PRAN AVCA NEVAOVAB LCAEIRA	ANA FRIJES CA DAS NEVES	195002022016061803183893550
0000C9000070000003245201606200000000000PAU M GERGILU SANTARJIAS	JO 19RC MARA CS DIAS	196810292016061200000000000
0000C9000080000003246201606200000000000JOSELVAULO AS S JOZ	JOVA TSJ ARTIVA A CRA DTOS	194711212016061320703686534
0000C9000090000003247201606290000000000LAUVIAINO AEIR000RA2LVA	LUCLMEIA DA 23LVA JESEIRA	194012202016060600000000000
0000C90001000000032482016062900000000000ALM CALVE PERIRCA	MARI320A SES 23JANA	194505182016062527935957568
0000C90001100000032492016062900000000000OLIMTHA RC GUOZ DAVJILVA	VINA JNADA AKOZ DAM ALVA	192810032016062518199488549
0000C9000110000003250201606300000000000FRATJSCO DE SILA INOT	MAS INA 30 A O CLUZITO	193210132016062646296565534

Ln 1, Col 1

Principais desafios do cruzamento de dados de registros administrativos

➤ Ausência de informação sobre a base de dados

Importante:

✓ Contato com responsáveis pela base de dados

✓ Conhecer forma de inserção das informações na base de dados

Exemplos:

O formulário GFIP da Caixa Previdência Social é dividido em várias seções para coleta de dados pessoais, profissionais e financeiros. No topo, há campos para identificação básica como nome, CPF e data de nascimento. Abaixo, há uma seção para dados de emprego e salário, com campos para data de início, data de fim, valor do salário e tipo de contrato. A parte inferior do formulário contém uma tabela para o registro de contribuições mensais, com colunas para o mês, o valor da contribuição e o valor do salário. No rodapé, há campos para a data, a assinatura e a autenticação do documento.

A tela de uma aplicação web de Imposto Pago/Retido apresenta uma interface com uma barra de navegação superior e uma lista de itens no lado esquerdo. O conteúdo principal é uma tabela com colunas para o tipo de imposto, o valor pago e o valor retido. A tabela contém várias linhas de dados, com alguns campos destacados em verde. No rodapé, há uma barra de status e uma barra de ferramentas.

A Certidão de Nascimento da República Federativa do Brasil é um documento oficial que registra o nascimento de uma criança. O formulário contém campos para o nome da criança, a data e o local de nascimento, o sexo, a filiação dos pais e a assinatura do responsável. No topo, há o brasão de armas do Brasil e o nome do registro. No rodapé, há uma seção para observações e averbações.

A Declaração de Nascimento Vivo do Ministério da Saúde é um formulário que registra o nascimento de uma criança. O formulário contém campos para o nome da criança, a data e o local de nascimento, o sexo, a filiação dos pais e a assinatura do responsável. No topo, há o brasão de armas do Brasil e o nome do registro. No rodapé, há uma seção para observações e averbações.

Principais desafios do cruzamento de dados de registros administrativos

➤ Ausência de informação sobre a base de dados

República Federativa do Brasil
Ministério da Saúde
2ª VIA - CARTÓRIO

Declaração de Nascido Vivo 30-52287570-1

I Cartório

1 Cartório Código 2 Registro 3 Data

4 Município 5 UF

II Local da Ocorrência

6 Local da Ocorrência 7 Estabelecimento Código

1 Hospital 2 Outros Estab. Saúde 3 Domicílio 4 Outros 9 Ignorado

8 Endereço da ocorrência, se fora do estab. ou da resid. da mãe (Rua, praça, avenida, etc) Número Complemento 9 CEP

10 Bairro/Distrito Código 11 Município de ocorrência Código 12 UF

III Mãe

13 Nome da Mãe 14 Cartão SUS

15 Idade (anos) 16 Estado Civil 17 Escolaridade (Em anos de estudo concluídos) 18 Ocupação habitual e ramo de atividade 19 Núm. de filhos tidos em gestações anteriores (obs.: utilizar 99 se ignorados)

1 Solteira 2 Casada 3 Viúva 4 Separada judicialmente/divorciada 5 Ignorado 1 Nenhuma 2 De 1 a 3 3 De 4 a 7 4 De 8 a 11 5 12 e mais 9 Ignorado

20 Residência da mãe Logradouro Número Complemento 21 CEP

22 Bairro/Distrito Código 23 Município Código 24 UF

IV Gestação e Parto

25 Duração da gestação (em semanas) 26 Tipo de gravidez 27 Tipo de parto 28 Número de consultas de pré-natal

1 Menos de 22 2 De 22 a 27 3 De 28 a 31 4 De 32 a 36 5 De 37 a 41 6 42 e mais 9 Ignorado 1 Única 2 Dupla 3 Tripla e mais 9 Ignorado 1 Vaginal 2 Cesáreo 9 Ignorado 1 Nenhuma 2 De 1 a 3 3 De 4 a 6 4 7 e mais 9 Ignorado

V Recém Nascido

29 Nascimento Data Hora 30 Sexo M - Masculino F - Feminino I - Ignorado 31 Índice de Apgar 1º minuto 5º minuto

32 Raça/cor 33 Peso ao nascer em gramas

1 Branca 2 Preta 3 Amarela 4 Parda 5 Indígena

34 Detectada alguma malformação congênita e/ou anomalia cromossômica? 35 Código

1 Sim 2 Não 9 Ignorado Qual 7

36 Polegar direito da mãe 37 Pé direito da criança

Principais desafios do cruzamento de dados de registros administrativos

➤ Ausência de informação de campos na base de dados

Exemplo: Registros de pensionistas do SIAPE, cruzamento com dados do CPF e BPC (CPF como chave primária)*

nome_cpf	nome_pens_siape	nome_bpc	nom_mae_cpf	nom_mae_siape	nom_mae_bpc	dt_nasc_cpf	dt_nasc_siape	dt_nasc_bpc
RUTE AUGUSTA DA SILVA	RUTH AUGUSTA DA SILVA	RUTE AUGUSTA DA SILVA	LEDIR RODRIGUES COURA	NULL	LEDIR RODRIGUES COURA	13.09.1946	13.09.1949	13.09.1946
OLIVIA SENA MOURA	OLIVIA SENA MOURA	OLIVIA SENA MOURA	GONCALA RODRIGUES SENA	NULL	GONCALA RODRIGUES SENA	15.04.1946	05.07.1949	15.04.1946
ALEXANDRE RICARDO DA SILVA	ALEXANDRE RICARDO DA SILVA	ALEXANDRE RICARDO DA SILVA	ANTONIETA BANI DA SILVA	NULL	ANTONIETA BANI DA SILVA	20.04.1947	17.04.1947	20.04.1947
MARTA ROCHA	MARTA ROCHA	MARTA ROCHA	MARIA IVONIZA ROCHA	NULL	MARIA IVONIZA ROCHA	10.11.1935	10.11.1924	10.11.1934
DANIEL JOMAR FIGUEIREDO	DANIEL JOMAR FIGUEREDO	DANIEL JOMAR FIGUEIREDO	ENEDINA JOMAR FIGUEIREDO	NULL	ENEDINA JOMAR FIGUEIREDO	21.06.1952	21.06.1937	21.06.1937
TERESINHA DO NASCIMENTO	TEREZINHA DO NASCIMENTO	TEREZINHA DO NASCIMENTO	JURACI FERREIRA MEDEIROS	N INFORMADO	JURACI FERREIRA MEDEIROS	20.08.1973	28.08.1973	20.08.1973
MARIA SILVA DE JESUS	MARIA SILVA DE JESUS	MARIA SILVA DE JESUS	JACY SILVA DE JESUS	NULL	JACY SILVA DE JESUS	01.01.1948	01.08.1948	01.01.1948
MARIZA DE SOUZA GOMES	MARIZA DE SOUZA GOMES	MARIZA DE SOUSA GOMES	IRACEMA MOTA GOMES	NULL	IRACEMA MOTA GOMES	24.08.1943	21.10.1954	24.08.1943
SONIA MARIA	SONIA MARIA	SONIA MARIA	SEVERINA AMELIA DA SILVA	NULL	SEVERINA AMELIA DA SILVA	08.07.1948	03.06.1925	08.07.1948
LUCIA DOMINGUES CARDOSO	LUCIA DOMINGUES CARDOSO	LUCIA DOMINGUES CARDOSO	JULIA DE AZEVEDO	LUCIA DOMINGUES CARDOSO	JULIA DE AZEVEDO	23.04.1979	02.04.1979	23.04.1979
MARIA BATISTA NASCIMENTO	MARIA BATISTA NASCIMENTO	MARIA BATISTA CAVALCANTI	LUZIA BATISTA CAVALCANTI	NULL	LUZIA BATISTA CAVALCANTI	08.06.1985	08.06.1986	08.06.1985
MARIA DA CONCEICAO SOUSA	MARIA DA CONCEICAO SOUSA	MARIA DA CONCEICAO SOUSA	ISABEL RESENDE BRAGA	NULL	IZABEL REZENDE BRAGA	13.04.1948	03.05.1948	13.04.1948
ANA LUIZA DE BARROS	ANA LUIZA DE BARROS	ANA LUIZA DE BARROS	MARIA DE LOURDES BARROS	NULL	MARIA DE LOURDES BARROS	20.10.1942	22.10.1943	20.10.1942
ELIANE DE AZEVEDO SOUZA	ELIANE DE AZEVEDO SOUZA	ELIANE DE AZEVEDO SOUZA	MARIA CAROLINA DE JESUS	NULL	MARIA CAROLINA DE JESUS	15.04.1949	15.04.1929	15.04.1929
LILIAN COURA DA SILVA	LILIAN COURA DA SILVA	LILIAN COURA DA SILVA	ANTONIA GUIOMAR DA SILVA	NULL	ANTONIA GUIOMAR DA SILVA	25.04.1974	03.05.1974	25.04.1974
ANTONIA BARBOSA GOMES	ANTONIA BARBOSA GOMES	ANTONIA BARBOSA GOMES	MARIA DE LOROTA GOMES	NULL	MARIA DE LORETA GOMES	22.09.1981	22.07.1981	22.09.1981
BETINA NOVAIS DOS SANTOS	BETINA NOVAIS DOS SANTOS	BETINA NOVAES DOS SANTOS	EULINA OLIVEIRA NOVAIS	BETINA NOVAES DOS SANTOS	EULINA OLIVEIRA NOVAIS	03.12.1976	03.12.1975	03.12.1976
MARIA SOARES FERREIRA	MARIA SOARES FERREIRA	MARIA SOARES FERREIRA	LUCINDA SOARES FERREIRA	NULL	LUCINDA SORES FERREIRA	07.09.1943	22.03.1974	07.09.1943

A ausência de informação dificulta a identificação da pessoa, gera dúvidas quanto a qualidade da base e pode atrapalhar a conclusão dos resultados.

*dados alterados para preservar o sigilo das informações

Principais desafios do cruzamento de dados de registros administrativos

➤ Registros duplicados

- ✓ Antes de cruzar os dados de uma tabela com outra, os dados duplicados devem ser tratados.

Exemplo: Registros duplicados do BPC, cruzamento com dados do CPF (CPF como chave primária)*

CPF (fictício)	nom_Maciça	nom_CPF	nom_mae_Maciça	nom_mae CPF	dt_nasc_Maciça	dt_nasc_CPF
14704280114	JORGE HENRIQUE FREITAS DA SILVA	JORGE HENRIQUE FREITAS DA SILVA	MARIA DO SOCORRO DA SILVA	MARIA RODRIGUES DA SILVA	21/10/1994	21/10/1994
14704280114	JORGE HENRIQUE FREITAS DA SILVA	JORGE HENRIQUE FREITAS DA SILVA	MARIA RODRIGUES DA SILVA	MARIA RODRIGUES DA SILVA	21/10/1994	21/10/1994
1590510777	RAFAELA BERNARDES PEREIRA	RAFAELA BERNARDES PEREIRA	ROSILANE BERNARDES CABRAL	ROSILANE BERNARDES CABRAL	03/05/1948	03/05/1948
1590510777	RAFAELA BERNARDES PEREIRA	RAFAELA BERNARDES PEREIRA	ROSILANE BERNARDES CABRAL	ROSILANE BERNARDES CABRAL	03/05/1948	03/05/1948
4469464605	NATALIA SOARES A SILVA	NATALIA SOARES DA SILVA	AMARILDA ROSANGELA S DA SILVA	AMARILDA ROSANGELA SOARES DA SILVA	14/12/1945	14/12/1945
4469464605	NATALIA SOARES DA SILVA	NATALIA SOARES DA SILVA	AMARILDA ROSANGELA S DA SILVA	AMARILDA ROSANGELA SOARES DA SILVA	14/12/1945	14/12/1945
14891035416	JOSE CARLOS DE OLIVEIRA	JOSE CARLOS DE OLIVEIRA	ANTONIA GARCEZ DE OLIVEIRA	ANNA MARIA DE OLIVEIRA	29/07/1998	29/07/1998
14891035416	JOSE CARLOS DE OLIVEIRA	JOSE CARLOS DE OLIVEIRA	FRANCISCA CORREIA DA SILVA	ANNA MARIA DE OLIVEIRA	29/07/1998	29/07/1998
176626130	IVANILDA NEVES DA SILVA	IVANILDA NEVES DA SILVA	NOEMIA NEVES DOS SANTOS	NOEMIA NEVES DOS SANTOS	03/07/1999	03/07/1999
176626130	MARIA DA CONCEICAO NEVES DA SILVA	IVANILDA NEVES DA SILVA	IVANILDA NEVES DA SILVA	NOEMIA NEVES DOS SANTOS	03/07/1999	03/07/1999
1928672781	JUNIOR SANTOS CERQUEIRA	JUNIOR DOS SANTOS CERQUEIRA	LUZIA RODRIGUES DE JESUS	LUZIA RODRIGUES DE JESUS	20/09/2001	20/09/2001
1928672781	JUNIOR SANTOS CERQUEIRA	JUNIOR DOS SANTOS CERQUEIRA	LUZIA RODRIGUES DE JESUS	LUZIA RODRIGUES DE JESUS	20/09/2001	20/09/2001
2361365601	EUNICE CUNHA DE SOUZA	EUNICE CUNHA DE SOUZA	ALMERINDA CARIUS DA CUNHA	ALMERINDA CARIUS DA CUNHA	03/10/1942	03/10/1942
2361365601	EUNICE CUNHA SOUZA	EUNICE CUNHA DE SOUZA	ALMERINDA CARMO DA CUNHA	ALMERINDA CARIUS DA CUNHA	03/10/1942	03/10/1942
16391800366	MARIA NILZA DE ANDRADE	MARIA NILZA DE ANDRADE	ELIZABETE DOS SANTOS PEREIRA	ELIZABETE DOS SANTOS PEREIRA	18/11/1961	18/11/1961
16391800366	MARIA NILZA DE ANDRADE	MARIA NILZA DE ANDRADE	ELIZABETE DOS SANTOS PEREIRA	ELIZABETE DOS SANTOS PEREIRA	18/11/1961	18/11/1961
753899461	LUIZA GOMES DE OLIVEIRA	LUIZA GOMES BARBOSA	ADRIANA DE JESUS FREITAS	ADRIANA DE JESUS FREITAS	20/01/1941	20/01/1941
753899461	LUIZA GOMES BARBOSA	LUIZA GOMES BARBOSA	ADRIANA DE JESUS FREITAS	ADRIANA DE JESUS FREITAS	08/09/1974	20/01/1941

Ex.: Filha cadastrada com CPF da mãe

*dados alterados para preservar o sigilo das informações

Principais desafios do cruzamento de dados de registros administrativos

➤ Características especiais

Exemplo de utilização da base do CPF como referência (supondo ser mais confiável)

- ✓ CPF não localizados na base da RFB
- ✓ Casos com informação de óbito na base da RFB
- ✓ Casos em que o CPF se encontra cancelado, suspenso ou anulado.

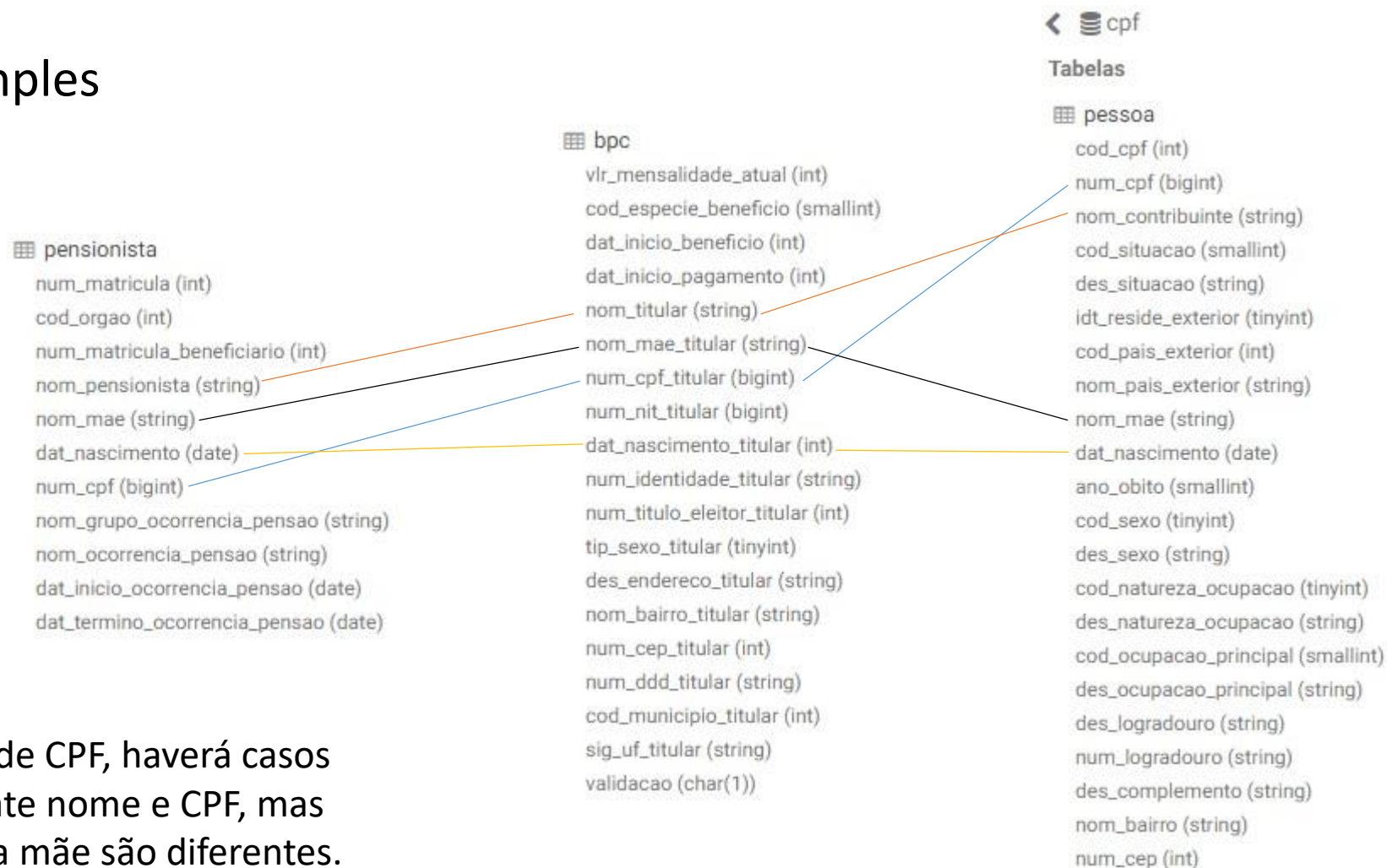
Situação do CPF na base da Receita Federal

- REGULAR: não há nenhuma pendência no cadastro do contribuinte.
- PENDENTE DE REGULARIZAÇÃO: o contribuinte deixou de entregar alguma Declaração do Imposto Renda da Pessoa Física (DIRPF) a que estava obrigado em pelo menos um dos últimos cinco anos.
- SUSPENSA: o cadastro do contribuinte está incorreto ou incompleto.
- CANCELADA: o CPF foi cancelado, em virtude de decisão administrativa ou judicial ou por falecimento do contribuinte
- NULA: foi constatada fraude na inscrição e o CPF foi anulado.

Principais desafios do cruzamento de dados de registros administrativos

➤ Escolha das chaves

- Não é uma tarefa simples



Ex.: Mesmo que bata informação de CPF, haverá casos em que o nome é diferente, ou bate nome e CPF, mas a data de nascimento e o nome da mãe são diferentes.

Principais desafios do cruzamento de dados de registros administrativos

➤ Escolha das chaves

- É preciso conhecer as bases de dados
 - Exemplo: vários benefícios para a mesma pessoa -> vários registros para a mesma pessoa
- É preciso saber como estão preenchidas as informações
 - Exemplo: determinado campo só é preenchido para determinadas pessoas, benefícios, etc.
- Cuidado para não criar elos errados
- Regras de **Localização e Unicidade**
 - Como localizar uma pessoa de uma base de dados em outra base de dados e ter certeza de que é a mesma pessoa?
 - A eficiência da Regra de Localização e Unicidade está diretamente ligada à qualidade dos dados inseridos na base pelos usuários.
- **Avaliação de Unicidade:** rotina sistêmica, composta por um conjunto de critérios e regras, que possibilita a comparação de dados cadastrais da pessoa e a verificação de equivalência entre os registros avaliados, com a finalidade de confirmar se pertencem ou não a um mesmo titular.

Principais desafios do cruzamento de dados de registros administrativos

➤ Escolha das chaves

- Regras de **Localização** e **Unicidade** - Matriz de Avaliação – Exemplo do Cadastro NIS/CAIXA

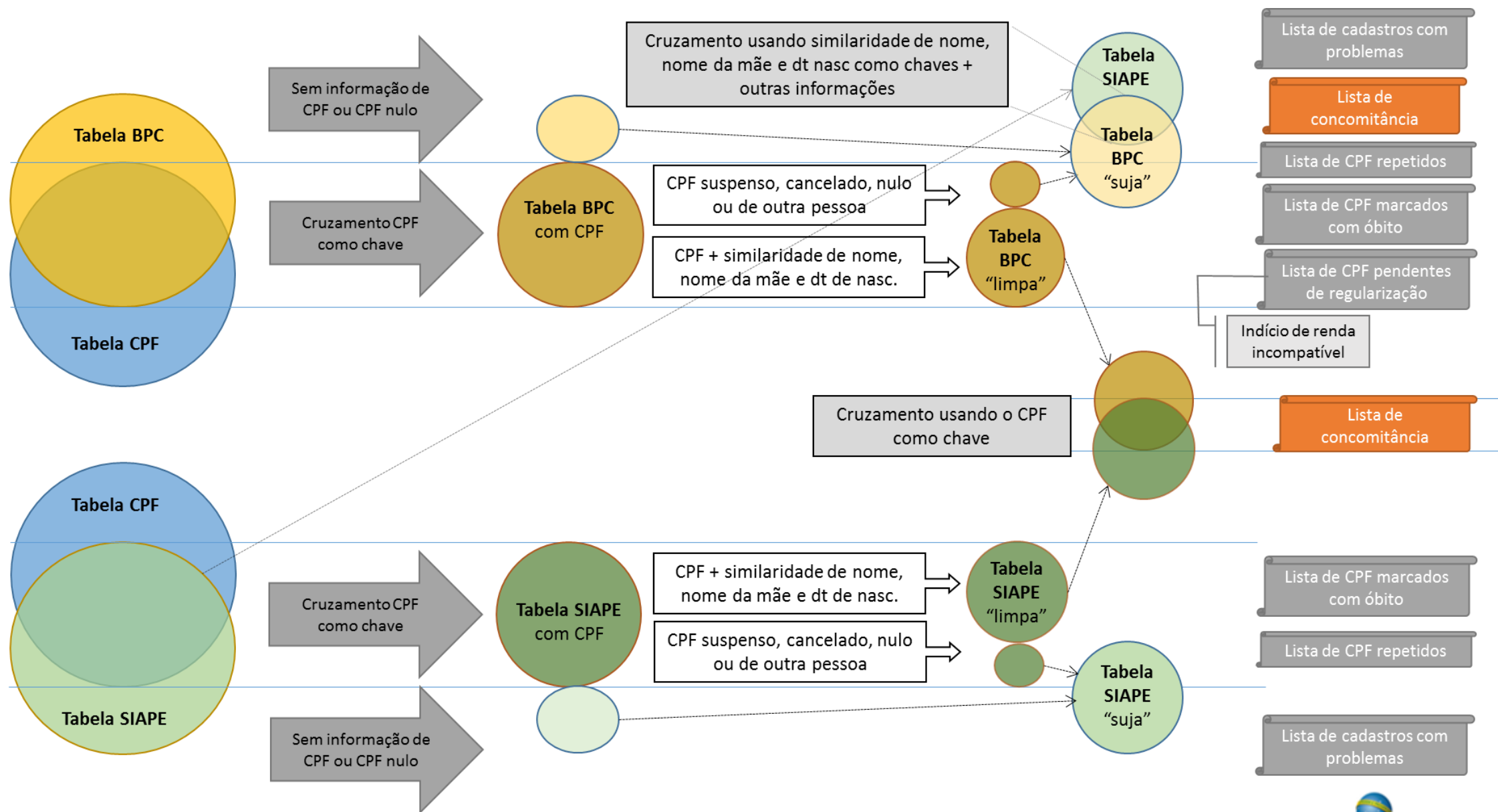
Árvore de Decisão		Matriz de Decisão				Matriz de Comparação			
NIS Informado		Atributos Prioritários				Atributos de Complementares			
Nome da Pessoa				UF	S			UF	S
1	Fonético Completo	1	CPF	-	-	1	Nome do Pai	-	-
2	Fonético Abreviado	2	Título de Eleitor	-	-	2	Município Nascimento	-	-
3	Fonético Primeiro e Último	3	Carteira de Trabalho	S		3	UF de Nascimento	-	-
4	Fonético Primeiro Nome	4	Identidade	S	S	4	Certidão Casamento Certidão de Óbito Certidão do Índio	S	-
		5	Certidão de Nascimento	S		5	Passaporte	-	-
		6	Data de Nascimento		-	6	Portaria Naturalização	-	-
		7	Nome da Mãe	-	-				

Unicidade Cadastral

Coincidência de:

- 1 forma fonética do nome na Árvore de Decisão
+
- 2 ou + atributos da Matriz de Decisão
+
- utilização dos atributos complementares, quando necessário.

Esquema de cruzamento de dados: SIAPE X CPF X BPC



Resultados dos cruzamentos

Resultado dos cruzamentos		Tabela BPC	
		4.364.500	
		BPC validado	BPC não validado
		3.971.373	393.127
Tabela SIAPE_Servidor 1.190.770	SIAPE_Servidor validado 1.190.765	196	14
	SIAPE_Servidor não validado 5		
Tabela SIAPE_Pensionista 611.033	SIAPE_Pensionista validado 503.945	5.520	23
	SIAPE_Pensionista não validado 107.088		

Notas:

Validado - Com CPF regular ou pendente de regularização na base da RFB e com informações compatíveis com o cadastrado no CPF (nome, nome da mãe e data de nascimento);

Não validado - Sem CPF ou CPF nulo, cancelado ou suspenso na base da RFB ou com informações incompatíveis com o cadastrado no CPF.

Resultados dos cruzamentos

- Análise tem que levar em consideração as regras dos programas:
 - Possibilidade de acumulação de benefícios indenizatórios
 - Possibilidade de trabalho de estágio
- São indícios - resultado precisa ser apurado individualmente

Situação no SIAPE	Categoria no BPC		Total
	Pessoa com deficiência	Idoso	
ANIST.PRIVADO L10559	3	20	23
ANIST.PUBLICO L10559	1	5	6
APOSENTADO	1	47	48
ATIVO PERMANENTE	47	2	49
CEDIDO SUS/LEI 8270		1	1
CELETISTA/EMPREGADO	8	1	9
CLT ANS -DEC 6657/08	1	1	2
CONT.PROF.SUBSTITUTO	5		5
CONTRATO TEMPORARIO	2		2
ESTAGIARIO	51	2	53
NOMEADO CARGO COMIS.	2		2
REFORMA CBM / PM		4	4
SEM VINCULO	4	2	6
PENSIONISTA	1.640	3.903	5.543
Total Geral	1.765	3.988	5.753