

**FIAP - FACULDADE DE INFORMÁTICA E
ADMINISTRAÇÃO PAULISTA
DATA SCIENCE**

Fábio Pereira de Lima – RM98803
Giovanna Cardoso Satorres – RM99944
Giullia Bianca Rocha Souza – RM552108
Gustavo Semenuk – RM550472
Rafael Luiz Custódio Guimarães – RM98304

**DATA GOVERNANCE & DATA SECURITY
MANAGEMENT**

Dicionário de Dados

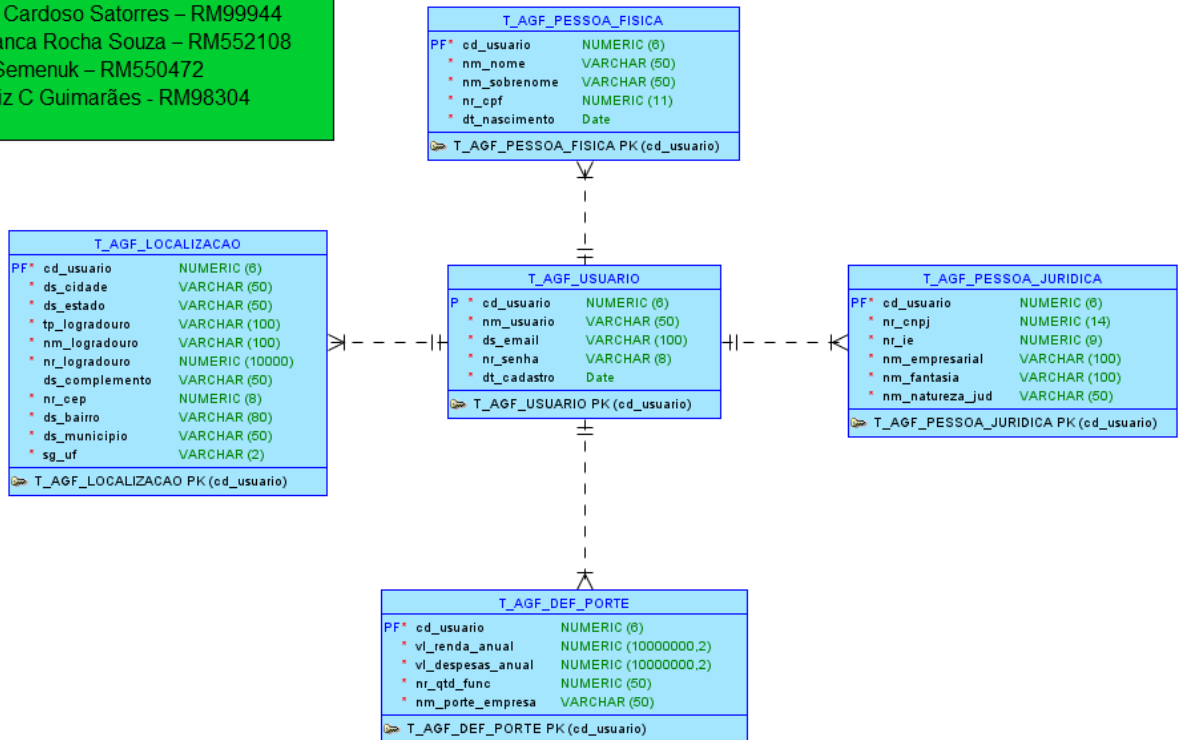
São Paulo
2023

PLANILHA EXCEL

Tabela (ENTIDADE)		T_AGF_USUARIO					
Descrição		Os usuários serão os responsáveis pelo cadastro da empresa e seus respectivos dados					
Política de acesso		DBA pode escrever, ler e editar os dados; os demais colaboradores podem ler					
Volume esperado		Carga inicial de 1200 linhas (registros) e volume diário de até 10 linhas					
Política de backup		Sempre as 2:00 AM e incrementado a cada 1H					
Prazo de retenção		Permanente					
Rotina de limpeza		Não se aplica					
Coluna/campo (atributos)		Tipo de Dados		Tamanho	Constraint	Lista de Interessados	Descrição
cd_usuario		NUMÉRICO - NUMBER		6	PK	DBA	O código de usuario serve para identificar o cliente dentro do Banco de Dados
nm_usuario		TEXTO - VARCHAR2		50	NN		Nome completo do usuário
ds_email		TEXTO - VARCHAR2		100	UN		Cada usuário só pode ter um único e-mail, o e-mail será utilizado para autenticar o acesso do cliente
nr_senha		TEXTO - VARCHAR2		min 8 max 16	NN		A senha serve para o usuário acessar a plataforma
dt_cadastro		DATA - DATE		-x-	NN		Data do momento em que foi realizado o cadastro
Tabela (ENTIDADE)		T_AGF_PESSOA_FISICA					
Descrição		É a pessoa que tem cadastro na empresa e é responsável legal					
Política de acesso		DBA pode escrever, ler e editar os dados; os demais colaboradores podem ler					
Volume esperado		Carga inicial de 1200 linhas (registros) e volume diário de até 10 linhas					
Política de backup		Sempre as 2:00 AM e incrementado a cada 1H					
Prazo de retenção		Permanente					
Rotina de limpeza		Não se aplica					
Coluna/campo (atributos)		Tipo de Dados		Tamanho	Constraint	Lista de Interessados	Descrição
cd_usuario		NUMÉRICO - NUMBER		6	FK	DBA	O código de usuario serve para identificar o cliente dentro do Banco de Dados
nm_nome		TEXTO - VARCHAR2		50	NN		Primeiro nome do usuário
nm_sobrenome		TEXTO - VARCHAR2		50	NN		Sobrenome do usuário
nr_cpf		NUMÉRICO - NUMBER		11	NN		Número do cadastro de pessoa física
dt_nascimento		DATA - DATE		-x-	NN		Data de nascimento do usuário
Tabela (ENTIDADE)		T_AGF_PESSOA_JURIDICA					
Descrição		São os dados da própria empresa					
Política de acesso		DBA pode escrever, ler e editar os dados; os demais colaboradores podem ler					
Volume esperado		Carga inicial de 1200 linhas (registros) e volume diário de até 10 linhas					
Política de backup		Sempre as 2:00 AM e incrementado a cada 1H					
Prazo de retenção		Permanente					
Rotina de limpeza		Não se aplica					
Coluna/campo (atributos)		Tipo de Dados		Tamanho	Constraint	Lista de Interessados	Descrição
cd_usuario		NUMÉRICO - NUMBER		6	FK	DBA	O código de usuario serve para identificar o cliente dentro do Banco de Dados
nr_cnpj		NUMÉRICO - NUMBER		14	UN		Serve para identificar o negócio nos mais diversos tipos de atividades
nr_ie		NUMÉRICO - NUMBER		9	UN		Cadastro formal das empresas que comercializam produtos físicos e que necessitam emitir Nota Fiscal Eletrônica
nm_empresarial		TEXTO - VARCHAR2		100	NN		Serve para identificar o tipo jurídico da empresa
nm_fantasia		TEXTO - VARCHAR2		100	NN		Nome comercial da empresa
nm_natureza_jud		TEXTO - VARCHAR2		50	NN		Serve para identificar atividade comercial que a empresa exerce
Tabela (ENTIDADE)		T_AGF_LOCALIZACAO					
Descrição		A localização se refere ao endereço da fazenda/empresa do cliente					
Política de acesso		DBA pode escrever, ler e editar os dados; os demais colaboradores podem ler					
Volume esperado		Carga inicial de 1200 linhas (registros) e volume diário de até 10 linhas					
Política de backup		Sempre as 2:00 AM e incrementado a cada 1H					
Prazo de retenção		Permanente					
Rotina de limpeza		Não se aplica					
Coluna/campo (atributos)		Tipo de Dados		Tamanho	Constraint	Lista de Interessados	Descrição
cd_usuario		NUMÉRICO - NUMBER		6	FK	DBA	O código de usuario serve para identificar o cliente dentro do Banco de Dados
ds_cidade		TEXTO - VARCHAR2		50	NN		Cidade em que está localizada a fazenda/empresa do cliente
ds_estado		TEXTO - VARCHAR2		50	NN		Estado em que está localizada a fazenda/empresa do cliente
tp_logradouro		TEXTO - VARCHAR2		100	NN		Logradouro da fazenda/empresa do cliente
nm_logradouro		TEXTO - VARCHAR2		100	NN		O Tipo de logradouro para definir se está em uma área de via pública ou dentro de condomínios privados
nr_logradouro		NUMÉRICO - NUMBER		10000	NN		Número para melhor precisão na localização para a fazenda/empresa do cliente
ds_complemento		TEXTO - VARCHAR2		50	NN		Complemento que pode ser usado como ponto de referência para a localização do cliente
nr_cep		NUMÉRICO - NUMBER		[0,9]	NN		O código Postal vai servir para melhor precisão da localização do cliente
ds_bairro		TEXTO - VARCHAR2		50	NN		Bairro da cidade em que está localizada a fazenda/empresa do cliente
ds_municipio		TEXTO - VARCHAR2		50	NN		Município da localização da fazenda/empresa do cliente
sg_uf		TEXTO - VARCHAR2		2	NN		Unidades Federativas de onde está o cliente
Tabela (ENTIDADE)		T_AGF_DEF_PORTE					
Descrição		Define o tamanho da empresa, para pequeno, médio e grande					
Política de acesso		DBA pode escrever, ler e editar os dados; os demais colaboradores podem ler					
Volume esperado		Carga inicial de 1200 linhas (registros) e volume diário de até 10 linhas					
Política de backup		Sempre as 2:00 AM e incrementado a cada 1H					
Prazo de retenção		Permanente					
Rotina de limpeza		Não se aplica					
Coluna/campo (atributos)		Tipo de Dados		Tamanho	Constraint	Lista de Interessados	Descrição
cd_usuario		NUMÉRICO - NUMBER		6	FK	DBA	O código de usuario serve para identificar o cliente dentro do Banco de Dados
vl_renda_anual		NUMÉRICO - NUMBER		1000000	NN		Valor estimado da renda anual do cliente no último ano vigente
vl_despesas_anual		NUMÉRICO - NUMBER		1000000	NN		Valor estimado da despesas(gastos) anual do cliente no último ano vigente
nr_qtd_func		NUMÉRICO - NUMBER		50	NN		Número da quantidade de funcionários da fazenda/empresa do cliente
nr_area_terreno		NUMÉRICO - NUMBER		1000	NN		Área em hectares da fazenda/empresa do cliente
nm_porte_empresa		TEXTO - VARCHAR2		50	CK		Definição de porte da empresa com base no seu faturamento, podendo receber valores como pequeno porte, médio porte, grande porte

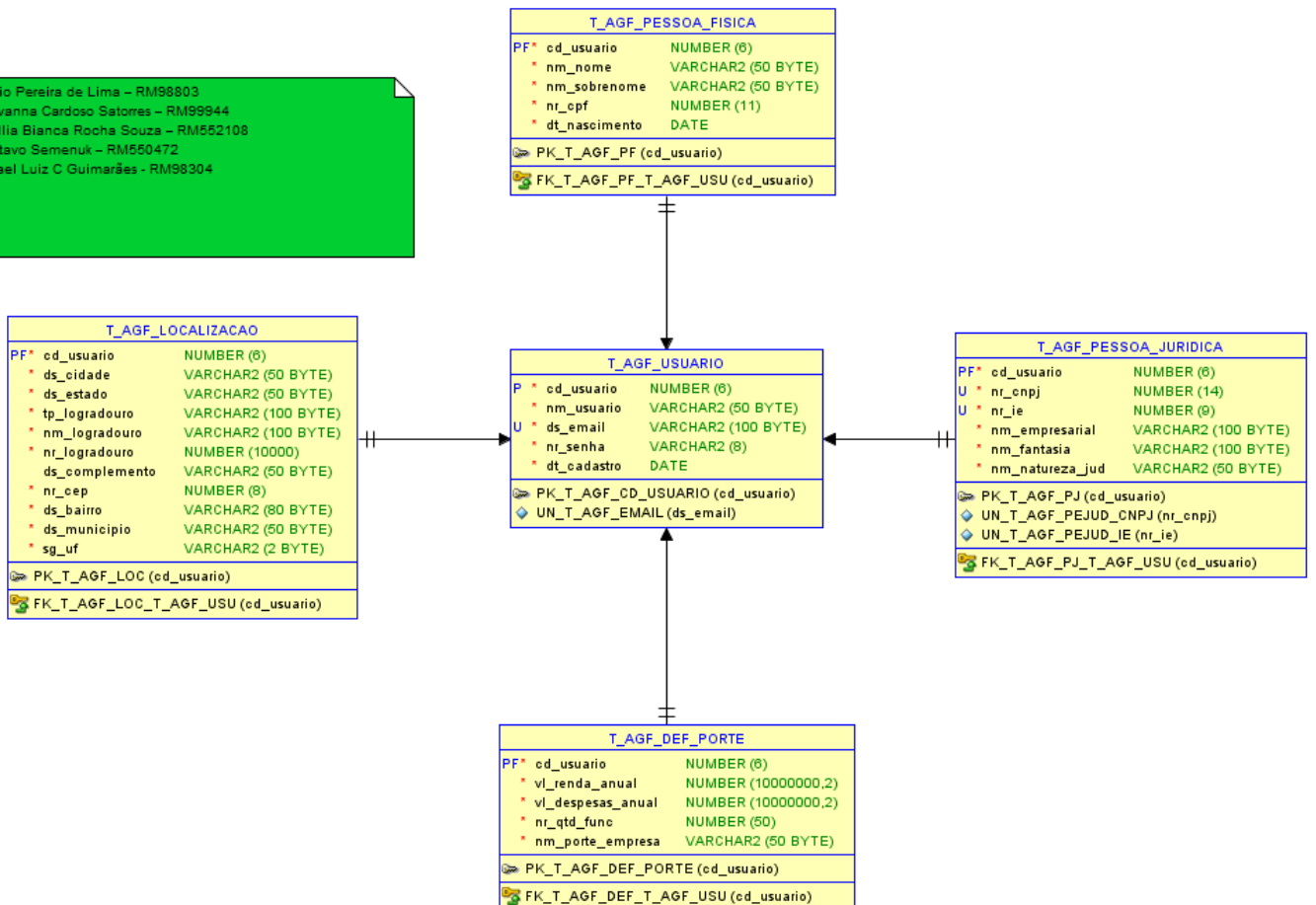
MODELO LÓGICO

Fábio Pereira de Lima – RM98803
Giovanna Cardoso Satorres – RM99944
Giullia Bianca Rocha Souza – RM552108
Gustavo Semenuk – RM550472
Rafael Luiz C Guimarães - RM98304



MODELO RELACIONAL

Fábio Pereira de Lima – RM98803
Giovanna Cardoso Satorres – RM99944
Giullia Bianca Rocha Souza – RM552108
Gustavo Semenuk – RM550472
Rafael Luiz C Guimarães - RM98304



DDL

```
CREATE TABLE t_agf_def_porte (
    cd_usuario          NUMBER(6) NOT NULL,
    vl_renda_anual      NUMBER(10000000, 2) NOT NULL,
    vl_despesas_anual   NUMBER(10000000, 2) NOT NULL,
    nr_qtd_func         NUMBER(50) NOT NULL,
    nm_porte_empresa    VARCHAR2(50 BYTE) NOT NULL
);

ALTER TABLE t_agf_def_porte
    ADD CONSTRAINT ck_t_agf_def_renda_anual CHECK ( vl_renda_anual >=
50000
                                                AND vl_renda_anual <=
250000 );

ALTER TABLE t_agf_def_porte
    ADD CONSTRAINT ck_t_agf_def_dp_anual CHECK ( vl_despesas_anual >=
10000
                                                AND vl_despesas_anual <=
50000 );

ALTER TABLE t_agf_def_porte ADD CONSTRAINT pk_t_agf_def_porte PRIMARY KEY
( cd_usuario );

CREATE TABLE t_agf_localizacao (
    cd_usuario          NUMBER(6) NOT NULL,
    ds_cidade          VARCHAR2(50 BYTE) NOT NULL,
    ds_estado          VARCHAR2(50 BYTE) NOT NULL,
    tp_logradouro      VARCHAR2(100 BYTE) NOT NULL,
    nm_logradouro      VARCHAR2(100 BYTE) NOT NULL,
    nr_logradouro      NUMBER(10000) NOT NULL,
    ds_complemento     VARCHAR2(50 BYTE),
    nr_cep              NUMBER(8) NOT NULL,
    ds_bairro          VARCHAR2(80 BYTE) NOT NULL,
    ds_municipio       VARCHAR2(50 BYTE) NOT NULL,
    sg_uf              VARCHAR2(2 BYTE) NOT NULL
);

ALTER TABLE t_agf_localizacao ADD CONSTRAINT pk_t_agf_loc PRIMARY KEY (
cd_usuario );

CREATE TABLE t_agf_pessoa_fisica (
    cd_usuario          NUMBER(6) NOT NULL,
    nm_nome            VARCHAR2(50 BYTE) NOT NULL,
    nm_sobrenome       VARCHAR2(50 BYTE) NOT NULL,
    nr_cpf             NUMBER(11) NOT NULL,
```

```

        dt_nascimento DATE NOT NULL
    );

ALTER TABLE t_agf_pessoa_fisica ADD CONSTRAINT pk_t_agf_pf PRIMARY KEY (
cd_usuario );

CREATE TABLE t_agf_pessoa_juridica (
    cd_usuario      NUMBER(6) NOT NULL,
    nr_cnpj          NUMBER(14) NOT NULL,
    nr_ie            NUMBER(9) NOT NULL,
    nm_empresarial   VARCHAR2(100 BYTE) NOT NULL,
    nm_fantasia      VARCHAR2(100 BYTE) NOT NULL,
    nm_natureza_jud VARCHAR2(50 BYTE) NOT NULL
);

ALTER TABLE t_agf_pessoa_juridica ADD CONSTRAINT pk_t_agf_pj PRIMARY KEY
( cd_usuario );

ALTER TABLE t_agf_pessoa_juridica ADD CONSTRAINT un_t_agf_pejud_cnpj
UNIQUE ( nr_cnpj );

ALTER TABLE t_agf_pessoa_juridica ADD CONSTRAINT un_t_agf_pejud_ie UNIQUE
( nr_ie );

CREATE TABLE t_agf_usuario (
    cd_usuario  NUMBER(6) NOT NULL,
    nm_usuario  VARCHAR2(50 BYTE) NOT NULL,
    ds_email    VARCHAR2(100 BYTE) NOT NULL,
    nr_senha    VARCHAR2(8) NOT NULL,
    dt_cadastro DATE NOT NULL
);

ALTER TABLE t_agf_usuario ADD CONSTRAINT pk_t_agf_cd_usuario PRIMARY KEY
( cd_usuario );

ALTER TABLE t_agf_usuario ADD CONSTRAINT un_t_agf_email UNIQUE ( ds_email
);

ALTER TABLE t_agf_def_porte
    ADD CONSTRAINT fk_t_agf_def_t_agf_usu FOREIGN KEY ( cd_usuario )
        REFERENCES t_agf_usuario ( cd_usuario );

ALTER TABLE t_agf_localizacao
    ADD CONSTRAINT fk_t_agf_loc_t_agf_usu FOREIGN KEY ( cd_usuario )
        REFERENCES t_agf_usuario ( cd_usuario );

ALTER TABLE t_agf_pessoa_fisica
    ADD CONSTRAINT fk_t_agf_pf_t_agf_usu FOREIGN KEY ( cd_usuario )
        REFERENCES t_agf_usuario ( cd_usuario );

```

```
ALTER TABLE t_agf_pessoa_juridica
  ADD CONSTRAINT fk_t_agf_pj_t_agf_usu FOREIGN KEY ( cd_usuario )
    REFERENCES t_agf_usuario ( cd_usuario );
```

DML

(INSERINDO DADOS NAS TABELAS)

```
INSERT INTO T_AGF_USUARIO (CD_USUARIO, NM_USUARIO, DS_EMAIL, NR_SENHA, DT_CADASTRO)
```

```
VALUES (204056, 'Gorete', 'maria.gorete@gmail.com', 24167854, TO_DATE('20/08/2023','DD/MM/YYYY'));
```

```
INSERT INTO T_AGF_PESSOA_FISICA (CD_USUARIO, NM_NOME, NM_SOBRENOME, NR_CPF, DT_NASCIMENTO)
```

```
VALUES (204056, 'Maria Gorete', 'Bianchi', 96325814770, TO_DATE('10/03/1965','DD/MM/YYYY'));
```

```
INSERT INTO T_AGF_PESSOA_JURIDICA (CD_USUARIO, NR_CNPJ, NR_IE, NM_EMPRESARIAL, NM_FANTASIA, NM_NATUREZA_JUD)
```

```
VALUES (204056, 90324896000190, 123456789, 'Sociedade Empresária Limitada', 'FAZENDA MARIA', 'Agricultura');
```

```
INSERT INTO T_AGF_LOCALIZACAO (CD_USUARIO, DS_CIDADE, DS_ESTADO, TP_LOGRADOURO, NM_LOGRADOURO, NR_LOGRADOURO, DS_COMPLEMENTO, NR_CEP, DS_BAIRRO, DS_MUNICIPIO, SG_UF)
```

```
VALUES (204056, 'São Bernardo', 'São Paulo', 'Avenida', 'Pereira Barreto', '34', 'ponto de ônibus', 09076456, 'Baeta Neves', 'São Bernardo', 'SP' );
```

```
INSERT INTO T_AGF_DEF_PORTE (CD_USUARIO, VL_RENDA_ANUAL, VL_DESPESAS_ANUAL, NR_QTD_FUNC, NM_PORTE_EMPRESA)
```

```
VALUES (204056, 70000, 50000, 34, 'Médio Porte');
```

PERGUNTAS

1. Identifique em seu projeto os Dados Mestre, Dados de Transação e Dados de Referência. Liste as tabelas do seu projeto, classificando-as por esse critério de tipificação.

Dados Mestres:

- > *Tabela 'Usuário'* - PK código do usuário: Identificação única para cada usuário.
- > *Tabela 'Pessoa Física'* - FK código de usuário: Relaciona informações específicas da pessoa física ao usuário.
- > *Tabela 'Pessoa Jurídica'* - FK código de usuário: Semelhante à 'Pessoa Física', relaciona informações da pessoa jurídica ao usuário.

Dados de Transação:

- > *Tabela 'Localização'* - Contém informações que podem mudar ao longo do tempo, como localização e tipo de logradouro.
- > *Tabela 'Definição de Porte da Empresa'* - Valores como renda anual, despesas anuais, etc., podem ser atualizados periodicamente.

Dados de Referência:

- > *Tabela 'Localização'* - Campos relacionados à geografia, como cidade, estado, nome do bairro, etc.
- > *Tabela 'Definição de Porte da Empresa'* - A categoria de porte da empresa com base na renda e despesa anual.

2. Descreva em seguida se você usará Datahub no seu projeto e faça um desenho representando quais dados estarão no HUB e qual o mecanismo de distribuição de dados que você usará. Se não for usar Datahub, justifique o motivo.

Não usamos Datahub

Optamos pelo MySQL em vez de um Data Hub em nosso projeto devido às várias vantagens que impactam diretamente na qualidade e eficiência do nosso projeto. Aqui estão algumas razões pelas quais optamos pelo MySQL e a importância dessa escolha:

- Facilidade de Gerenciamento de Dados: O MySQL fornece uma plataforma robusta e confiável para garantir a precisão e disponibilidade de dados cruciais em nossa análise de crédito e produção.

- Desempenho e Escalabilidade: O MySQL, garante que nosso projeto possa crescer eficientemente, permitindo análises rápidas e resultados atualizados.

- Segurança dos Dados: A robusta segurança do MySQL é crucial para proteger a privacidade e integridade dos dados financeiros e de produção em nosso projeto.

- Flexibilidade de Consulta e Relatórios: A flexibilidade do MySQL na criação de consultas e relatórios complexos é fundamental para a análise de crédito e de produção, permitindo avaliar diversos fatores financeiros e comparar métricas de forma detalhada.

- Custo-Eficiência: O MySQL, como um banco de dados de código aberto, é uma escolha economicamente eficiente para projetos, permitindo controle de custos, crucial ao considerar o crescimento futuro.

Em resumo, escolhemos o MySQL devido à sua capacidade de atender às necessidades presentes e futuras do projeto, oferecendo desempenho, segurança e facilidade de uso. Isso é crítico, especialmente para análises de crédito e produção que exigem precisão e eficácia.