

Mapeamento Habitacional de Famílias e Atividades Econômicas nos Bairros e Favelas da Cidade do Rio de Janeiro

**Ellen A. de Souza, Gabriel T. Corrêa, Jefferson Maxwell,
Kevin S. de Andrade, Riquelme Gomes**

¹Instituto da Computação – Universidade Federal do Rio de Janeiro (UFRJ)

Resumo. *Este artigo descreve o mapeamento de famílias e suas atividades econômicas de áreas habitacionais por bairros do Rio de Janeiro contemplando favelas através de um banco de dados relacional integrado a uma aplicação Web.*

Abstract. *This article describes the mapping of families and their economic activities of housing areas by neighborhoods in Rio de Janeiro encompassing communities / “slums”(favelas) through a relational database integrated to a Web application.*

1. Contextualização

Nosso projeto teve como objetivo fazer a criação e o gerenciamento de um sistema de banco de dados relacional que contém dados reais a fim de serem usados para exibição e consultas de informações referentes à temática escolhida. No desenvolvimento do trabalho, escolhemos dois temas de categoria de grande relevância para a coletividade, os quais foram: Habitação-Urbanismo e Economia. E com base na escolha, fizemos a coleta de duas fontes de dados oficiais do governo que são bases do banco de dados.

A primeira fonte de dados a ser utilizada é o CECAD (Consulta, Seleção e Extração de Informações do CadÚnico) que é uma ferramenta permite conhecer as características socioeconômicas das famílias e pessoas incluídas no Cadastro Único bem como saber quais famílias são beneficiárias do Programa Bolsa Família. Fizemos uso das seguintes filtrasgens:

- (i) Famílias em situação de extrema pobreza;
- (ii) Faixa da renda familiar per capita;
- (iii) Quantidade de famílias beneficiárias do Programa Bolsa Família.

A segunda fonte de dados a ser utilizada é o Portal DataRio que integra um novo modelo de planejamento, integração, gestão e disseminação da informação da Prefeitura do Rio de Janeiro. Dentre as opções de conjuntos de dados desta segunda fonte, as utilizadas foram:

- Limites de favela e Urbanização
- Número de empregados por atividade econômica segundo as Áreas de Planejamento (AP), Regiões Administrativas (RA) e Bairros no Município do Rio de Janeiro em 2005-2020

A ideia do nosso projeto era trabalhar o conjunto de dados coletados e fazer o mapeamento de famílias e com seus atributos e Número de empregados por atividade econômica segundo a Área de Planejamento (AP), Bairro e suas relações, com a finalidade de realizar exibições gráficas e apresentar consultas contendo tais dados.

2. Projeto de Banco de Dados

2.1 Modelo Conceitual

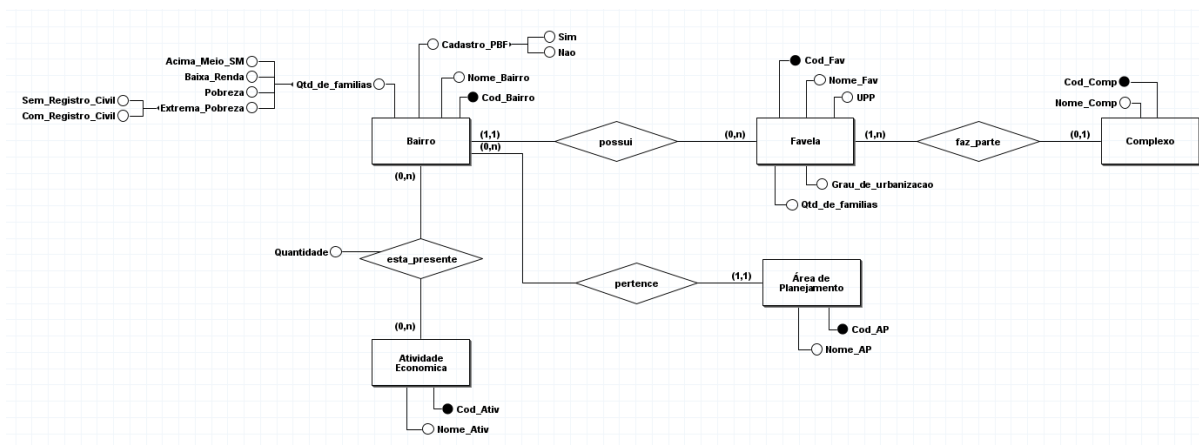


Figura 1. Modelo Conceitual

Para definir os relacionamentos entre os elementos de nosso projeto de forma mais clara e concisa, nosso grupo desenvolveu, em conjunto e com o uso da ferramenta “brModelo”, o Modelo Entidade Relacionamento. Através desse modelo, conseguimos especificar entidades, que concentram informações do banco na forma de atributos, e os relacionamentos entre elas, definindo cardinalidades, atributos, entre outros. Baseado nessas informações, nosso Modelo ER é composto por cinco entidades e quatro relacionamentos. São eles:

- **BAIRRO:** Entidade central do modelo, reunindo as informações dos datasets do CECAD na forma de múltiplos atributos, simples ou compostos. Devido à estrutura do projeto e da organização dos dados obtidos, essa entidade participa de três relacionamentos e da grande maioria das consultas realizadas mais para frente do projeto.
- **ATIVIDADE ECONÔMICA:** Entidade formada através do dataset homônimo do DataRio, representando os diferentes tipos de empregos existentes nos bairros. Seus dados são úteis para analisar o nível de vida dos moradores dos bairros, relacionando com faixas de renda e famílias PBF. Participa em apenas um relacionamento, com a entidade BAIRRO.

- FAVELA: Outra entidade chave no modelo, foi formada através do dataset homônimo do DataRio. Apresenta informações quantitativas e de localização, que são úteis para analisar tanto o nível de vida da população de um bairro quanto a distribuição de pobreza no município do Rio de Janeiro. Participa de dois relacionamentos: um com a entidade BAIRRO e outro com a entidade COMPLEXO.
- COMPLEXO: Entidade formada através do dataset de Limites de Favelas e Urbanização, representa uma divisão da área do município focada nas favelas ao invés dos bairros. Essa entidade, em conjunto com a entidade ÁREA DE PLANEJAMENTO, tem como objetivo providenciar formas diferentes de contemplar as informações das outras entidades por diferentes divisões geográficas. Participa de um relacionamento com a entidade FAVELA.
- ÁREA DE PLANEJAMENTO: Entidade formada através do dataset de atividades econômicas por bairro e AP, representando uma divisão da área do município mais abrangente do que os bairros. Essa entidade, em conjunto com a entidade COMPLEXO, tem como objetivo providenciar formas diferentes de contemplar as informações das outras entidades por diferentes divisões geográficas. Participa de um relacionamento com a entidade BAIRRO.
- ESTA_PRESENTE: Relacionamento entre BAIRRO e ATIVIDADE ECONÔMICA, contém o atributo “quantidade”, o qual indica a quantidade de empregos da atividade econômica naquele bairro. Trata-se de um relacionamento de cardinalidade muitos para muitos, uma vez que um bairro pode conter diversas atividades econômicas e uma mesma atividade econômica pode estar presente em vários bairros.
- POSSUI: Relacionamento entre BAIRRO e FAVELA. Apresenta cardinalidade um para muitos, já que um bairro pode possuir diversas favelas, mas cada favela pertence à necessariamente um único bairro.
- FAZ PARTE: Relacionamento entre FAVELA e COMPLEXO. Apresenta cardinalidade um para muitos, já que um complexo pode possuir diversas favelas, mas cada favela pode fazer parte de no máximo um complexo.
- PERTENCE: Relacionamento entre BAIRRO e ÁREA DE PLANEJAMENTO. Apresenta cardinalidade um para muitos, já que uma área de planejamento pode possuir diversos bairros mas cada bairro é pertencente a somente uma área de planejamento.

2.2 Modelo Lógico

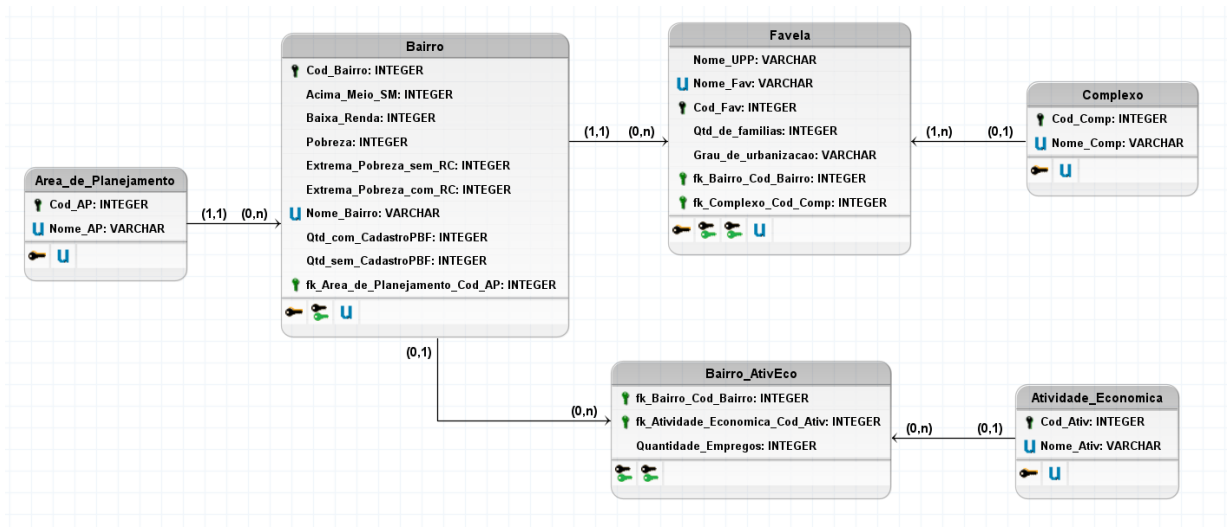


Figura 2. Modelo Lógico

O Modelo Lógico do nosso banco de dados foi formado através da conversão automática do Modelo ER disponível na aplicação brModelo. Durante esse processo fomos capazes de decidir sobre a formação de tabelas por relacionamento, de renomear entidades e atributos e de definir características de atributos como UNIQUE, PRIMARY KEY, FOREIGN KEY, etc. Não foram realizadas muitas mudanças no modelo pela conversão. Cada entidade se tornou uma tabela com suas respectivas chaves primárias e estrangeiras, os atributos compostos se tornaram atributos simples e apenas o relacionamento entre BAIRRO e ATIVIDADE ECONÔMICA se tornou uma tabela e passou a ser nomeado como “Bairro_AtivEco”, por possuir cardinalidade de muitos para muitos.

2.3 Modelo Físico

O Modelo Físico do nosso banco de dados também foi criado com o uso da aplicação BrModelo, porém foi modificado posteriormente. A mudança mais significativa na definição do banco de dados foi na tabela Favela, em que, por padrão, a ferramenta criou a chave estrangeira para Complexo com a cláusula ON DELETE RESTRICT, mas que, no nosso caso, precisávamos que fosse ON DELETE SET NULL. Isso porque na fonte do DataRio havia dois nomes de Complexos que não faziam sentido em nossa aplicação, um chamado "Isolada", o que representaria as favelas que não pertencem a nenhum complexo, isto é, pertencem ao complexo "Isolada" e outro chamado "NaN", causado por strings vazias presentes na fonte dos dados. Ambos casos passaram a ser NULL após a alteração do esquema e a remoção desses 2 registros da tabela Complexo.

Link: [Modelo Físico](#).

3. Instanciação do banco de dados

Uma das tarefas cumpridas foi a de popular o banco de dados e para isso fizemos a utilização de algumas ferramentas web de modo a facilitar a obtenção dos dados. Uma das maiores dificuldades em fazer a instanciação do banco foi o tratamento de dados contidos nos repositórios fonte. Ambos os sites de fontes de dados colhidos tinham arquivos com grandes inconsistências de dados, muitos com nomes de bairros quase repetidos porém com grandes erros gramaticais ou com variações da escrita correta do bairro, por exemplo diferentes dados de ANCHIETA como: ANCHIEA, ANCHIETA, ANCHIIETA, etc, além de informações referentes a outros campos da tupla conterem dados incoerentes, interferindo na obtenção direta dos dados para o nosso repositório. Os dados retirados do CECAD tem formato XLS e os advindos do Portal Data.rio tem formato XLS e CSV.

Dessa forma, tivemos que fazer a limpeza dos dados e prepará-los antes da sua utilização, que foi feita uma parte através da ferramenta PANDAS, uma biblioteca Python para manipulação e análise de dados e também pelo método força bruta de modo que o tratamento foi feito, selecionando os dados à mão sem utilizar script algum. A ferramenta principal que utilizamos foi o google planilhas (google sheets) para fazer a armazenagem dos dados utilizados como base das consultas, a escolha dessa ferramenta foi usada de forma geral pela familiaridade com o programa.

Depois do tratamento dos dados tivemos que identificar seus relacionamentos e assim montar as tabelas e preenchê-las conforme as especificidades dos temas. Para fazer isso utilizamos scripts automáticos em python, em que, para cada arquivo csv, isto é, para cada aba da nossa planilha, preenchamos uma tabela do nosso banco de dados.

4. Aplicação web

A aplicação foi desenvolvida objetivando a apresentação gráfica e simplificada do resultado de consultas feitas sobre a base de dados criada. As tecnologias utilizadas foram: db4free, o qual foi usado para subir o banco de dados de forma remota; React.js, uma biblioteca front-end; Next.js, um framework front-end, mas que também traz suporte para funcionalidades de back-end; e Vercel, que é um serviço web para hospedagem de aplicações web na nuvem.

A ideia principal da aplicação é permitir que o usuário possa visualizar informações gerais e específicas sobre bairros e favelas. Dessa forma, organizamos nosso site em duas abas principais, uma sobre bairros e outra sobre favelas. Em cada uma podemos ver gráficos, tabelas ou uma organização em grupos sobre as consultas ao banco de dados. Além disso, podemos pesquisar por um bairro ou uma favela em específico, caso queiramos consultar dados mais singulares.

Link do site: [Favelas RJ](#)

5. Principais consultas

Para a construção da aplicação web foram criadas dezenas de consultas, mas destacamos a seguir as principais, as quais visam contemplar os requisitos obrigatórios do trabalho.

1. Top 3 bairros com mais favelas

Consulta:

```
SELECT Quant_Favelas, Nome_Bairro
FROM Bairro NATURAL JOIN (
    SELECT COUNT(Cod_Fav) Quant_Favelas, Cod_Bairro
    FROM Favela
    GROUP BY Cod_Bairro
    ORDER BY Quant_Favelas DESC
    LIMIT 3
) AS Bairros_Com_Mais_Favelas
```

Resultado na aplicação: [Favelas RJ - Bairro - Aba Faixa de Renda](#)

2. Quantidade de famílias em cada faixa de renda por área de planejamento

Consulta:

```
SELECT
    Nome_AP,
    SUM(Acima_Meio_SM) as Acima_Meio_SM,
    SUM(Baixa_Renda) as Baixa_Renda,
    SUM(Pobreza) as Pobreza,
    SUM(Extrema_Pobreza_sem_RC + Extrema_Pobreza_com_RC) as
Extrema_Pobreza
FROM Bairro NATURAL JOIN Area_de_Planejamento
GROUP BY Cod_AP
```

Resultado na aplicação: [Favelas RJ - Bairro - Aba Faixa de Renda](#)

3. Atividades econômicas presentes em cada cada bairro, incluindo aqueles que não possuem atividades econômicas

Consulta:

```
SELECT Nome_Bairro, Nome_Ativ, Quantidade_Empregos
FROM Bairro_AtivEco NATURAL JOIN Atividade_Economica
    NATURAL RIGHT JOIN Bairro
ORDER BY Quantidade_Empregos DESC;
```

Resultado na aplicação: [Favelas RJ - Bairro - Aba Atividades Econômicas](#)

4. Quantidade de favelas por área de planejamento

Consulta:

```
SELECT Nome_AP, COUNT(cod_fav) as Qtd_Favelas
FROM Bairro NATURAL JOIN Area_de_Planejamento
      NATURAL JOIN Favela
GROUP BY Cod_AP
```

Resultado na aplicação: [Favelas RJ - Favela - Aba Panorama Geral](#)

5. Complexos e suas respectivas favelas e favelas isoladas

Consulta:

```
SELECT Cod_Fav, Nome_Fav, Cod_Comp, Nome_Comp
FROM Favela NATURAL LEFT JOIN Complexo
```

Resultado na aplicação: [Favelas RJ - Favela - Aba Complexo](#)

6. Atividades de cada integrante

1. Ellen Almeida

Responsável por elaborar o modelo conceitual do banco de dados, fazer o tratamento dos dados do CECAD e DataRio, elaborar consultas SQL, desenvolver a aplicação web e escrever o relatório final.

2. Gabriel Trindade

Responsável por elaborar o modelo conceitual, lógico e físico do banco de dados, fazer o tratamento dos dados DataRio, elaborar consultas SQL e escrever o relatório final.

3. Jefferson Maxwell

Responsável por elaborar o modelo conceitual e lógico do banco de dados, fazer o tratamento dos dados DataRio, popular o banco de dados e escrever o relatório final.

4. Kevin Sena

Responsável por elaborar o modelo conceitual, lógico e físico do banco de dados, fazer o tratamento dos dados do CECAD e DataRio, popular o banco de dados, elaborar consultas SQL, desenvolver a aplicação web e escrever o relatório final.

5. Riquelme Gomes

Responsável por elaborar o modelo conceitual, lógico e físico do banco de dados, elaborar consultas SQL, preparar o material de apresentação do grupo e escrever o relatório final.

7. Conclusão

O projeto da criação de um banco de dados relacional foi interessante em todos seus aspectos, desde a sua idealização até sua elaboração. A parte mais satisfatória foi colocar em prática o conteúdo ensinado na disciplina desde o início do período, compreendendo as diversas modelagens envolvidas no planejamento do banco de dados e os conhecimentos e cuidados com a sintaxe da estrutura SQL e MySQL. No entanto, o que mais estimulou e desafiou o grupo foram as dificuldades enfrentadas ao longo do desenvolvimento do projeto, como abordado no dia da apresentação.

Um grande desafio no desenvolvimento do banco de dados foi o estado dos dados do CECAD, cuja disponibilização dos repositórios de dados se encontrava de maneira bastante desarranjada. Essa experiência nos mostrou de forma prática a importância dos cuidados na coleta de dados e o desafio do tratamento dos mesmos, constituindo uma tarefa detalhada e demorada porém necessária a fim de estabelecer o banco de dados de maneira consistente.

Outro desafio foi a implementação da aplicação WEB do banco de dados do projeto, que demandou grande esforço dos membros da equipe cuja tarefa foi atribuída. A necessidade de aprender e mesclar diferentes ferramentas, organizar os dados para visualização em tabela e em gráfico, disponibilizar o banco de dados em um servidor remoto através da ferramenta db4free, testar e resolver bugs representa um exemplo prático das tarefas existentes nos processos de implementação e manutenção de uma aplicação integrada aos bancos de dados. Foi uma importante experiência para uma carreira nessa área.

Concluindo, o trabalho em geral serviu de grande aprendizado para aprofundar o nosso conhecimento teórico e prático em sistemas de banco de dados e ampliar a nossa visão em uma área extremamente relevante da computação.

Referências

Lista cras/creas/localidade.CECAD,2020. Disponível em: https://cecad.cidadania.gov.br/cras_local.php Acesso em: 29 nov. 2022.

Limites e favela e urbanização. Data.rio,2020. Disponível em: <https://www.data.rio/datasets/PCRJ::limites-de-favelas-e-urbaniza%C3%A7%C3%A3o/explore?location=-22.893227%2C-43.418274%2C11.73&showTable=true> Acesso em: 29 nov. 2022.

Número de empregados por atividade econômica segundo as Áreas de Planejamento (AP), Regiões Administrativas (RA) e Bairros no Município do Rio de Janeiro em 2005-2020 - Data.rio,2018. Disponível em: <https://www.data.rio/documents/n%C3%BAmero-de-empregados-por-atividade-econ%C3%B4mica-segundo-as-%C3%A1reas-de-planejamento-ap-regi%C3%B5es-administrativas-ra-e-bairros-no-munic%C3%ADpio-do-rio-de-janeiro-em-2005-2020/about> Acesso em: 29 nov. 2022.

DB4Free. Disponível em: db4free.net - Banco de dados MySQL de graça. Acesso em: 30 nov. 2022.

React.js. Disponível em: [React – Uma biblioteca JavaScript para criar interfaces de usuário \(reactjs.org\)](https://reactjs.org). Acesso em: 30 nov. 2022

Next.js. Disponível em: [Next.js by Vercel - The React Framework \(nextjs.org\)](https://nextjs.org). Acesso em: 30 nov. 2022.

Primereact. Disponível em: [PrimeReact - React UI Component Library \(primefaces.org\)](https://primefaces.org). Acesso em: 30 nov. 2022.