

PUBLICAÇÃO DE MAPAS AGRÍCOLAS INTERATIVOS NA WEB

Mário Balan, Pedro Alves Quilici Coutinho, Jaudete Daltio, Márcia Dompieri

Empresa Brasileira de Pesquisa Agropecuária – Embrapa
{mario.balan, pedro.coutinho}@colaborador.embrapa.br
{jaudete.daltio, marcia.dompieri}@embrapa.br

RESUMO

A análise integrada de dados espaciais contribui na caracterização e compreensão dos territórios. Contudo, nem todo dado espacial disponível possui seu componente geográfico explícito, o que dificulta consideravelmente a análise do espaço geográfico. Devido ao volume de dados, mapas dinâmicos oferecem maior possibilidades de exploração e permitem uma melhor interpretação dos dados. O objetivo desse artigo é apresentar uma solução para a elaboração e publicação de painéis agrícolas interativos na web. Os painéis podem conter, além de tabelas, mapas e gráficos interativos, manipuláveis por meio de controles embutidos em páginas web. A solução processa os dados das pesquisas agropecuárias, associa as feições cartográficas, aplica processamentos e cálculos estatísticos para construir um banco de dados que serve de base para a elaboração dos painéis. As páginas resultantes, que incorporam todos os elementos dinâmicos, podem ser publicadas na web ou em gerenciadores de conteúdo.

Palavras-chave – Dados Geográficos, Pesquisas Agrícolas Anuais, Mapas Dinâmicos, Tableau.

ABSTRACT

Integrated spatial data analysis contributes to the characterization of territories. However, not all spatial data has its geographical component explicit, which makes it very hard to produce geographic analysis (topological) and create maps. Due data volume, dynamic maps offer more possibilities of exploration and allow a better data interpretation. The goal of this paper is to present a solution for the preparation and publication of agricultural interactive sheets/dashboards on the web. These sheets/dashboards may contain tables, graphics and maps, filtered through web page controls. The solution processes agricultural surveys data, join cartographic features, applying correction, processing and statistical routines to build a database – the data source of the sheets/dashboards. The resulting pages, which incorporate all the dynamic elements, can be published on the web or in content management systems.

Key words – Geographic data, Annual Agricultural Survey, Dynamic maps, Tableau.

1. INTRODUÇÃO E MOTIVAÇÃO

Dados espaciais ganharam maior popularidade nas últimas décadas, e hoje constituem uma parcela significativa dos dados disponíveis na Web. Esses dados, que representam

a distribuição espacial de diversos fenômenos, possuem um papel importante na compreensão de um território. Análises de caracterização territorial comumente demandam a correlação de diversas fontes de dados e utilizam a componente espacial como unidade integradora [1]. Caracterizações dessa natureza são adotadas tanto para o planejamento de atividades políticas como administrativas, visando o desenvolvimento social, ambiental e econômico de uma sociedade.

É comum, contudo, a disponibilização de dados espaciais em formatos “não geográficos”, ou seja, formatos nos quais o componente espacial não é explícito. Nesses casos, é necessário uma etapa adicional para viabilizar análises geográficas, como operações topológicas de adjacência, pertinência, interseção, cruzamento ou proximidade. O Instituto Brasileiro de Geografia e Estatística (IBGE) disponibiliza grandes volumes de dados espaciais desta forma [4]. As pesquisas anuais e dados censitários gerados pela instituição são organizados e disponibilizados essencialmente em formato tabular. Adicionalmente, o IBGE disponibiliza dados cartográficos com as malhas territoriais de todo o Brasil. Com o intuito de promover a integração de ambos dados, o IBGE instituiu um identificador único para cada território – o *geocódigo*. Embora não se trate de um processo complexo, a transformação dos dados espaciais tabulares do IBGE em dados geográficos requer essa etapa adicional de integração, de acordo com as variáveis e os territórios de interesse.

Além disso, o volume de dados publicado pelo IBGE é considerável, tanto em diversidade como em extensão espacial e temporal. Algumas pesquisas estão disponíveis desde 1974, por exemplo. E, apenas em termos da produção agropecuária, o IBGE divulga várias pesquisas com dados de aproximadamente 200 cadeias diferentes, entre lavouras, pecuária, extrativismo vegetal e silvicultura. A riqueza de dados, tanto em relação ao volume quanto à quantidade de variáveis, permite a geração de diversos mapas para uma mesma cadeia. A exploração dinâmica destes mapas trás inúmeras vantagens na compreensão e interpretação dos dados [2] [3].

O objetivo deste trabalho é apresentar uma solução para a elaboração e publicação de painéis agrícolas interativos na web. Trata-se de uma solução capaz de, a partir de dados tabulares, gerar tabelas, mapas e gráficos dinâmicos, compondo uma ou mais páginas web. A principal contribuição da solução é viabilizar a publicação deste conteúdo de forma interativa na web, além de proporcionar consistência e agilidade neste processo. O intuito é facilitar a navegação e interpretação dos dados agropecuários, de forma que diferentes usuários possam realizar análises específicas de acordo com suas necessidades.

2. ASPECTOS TÉCNICOS ENVOLVIDOS

2.1. Dados Cartográficos

Os dados cartográficos utilizados provêm das malhas digitais oficiais do IBGE. Essas malhas são constituídas por arquivos vetoriais com a divisão das unidades da federação compatíveis com a escala original de trabalho 1:250.000. Os mapas dinâmicos elaborados pela solução podem agregar diferentes escalas territoriais – municípios, microrregiões, mesorregiões ou estados, por exemplo. Para generalizar a denominação, utilizamos o termo *território* para designar a entidade espacial qualificada. Isto presume que, em alguns casos, os dados originais qualitativos sejam agregados.

2.2. Dados Qualitativos

Os dados qualitativos utilizados oriundam das pesquisas permanentes promovidas pelo IBGE, promovidas anualmente e que abrangem todo o território nacional. Esses dados são disponibilizados pelo Sistema IBGE de Recuperação Automática - SIDRA¹. A unidade de investigação é o município ou a microrregião. Os dados provêm das seguintes pesquisas: (i) *Produção Agrícola Municipal (PAM)*: dados sobre a quantidade produzida, área plantada e colhida e valor de produção, organizados em dois grandes grupos: lavouras temporárias e lavouras permanentes; (ii) *Pesquisa Pecuária Municipal (PPM)*: dados sobre os efetivos das espécies animais criadas e a quantidade produzida de subprodutos de pecuária, organizados em três principais grupos: aquicultura, efetivo de rebanho e produtos de pecuária; e (iii) *Produção de Extração Vegetal e da Silvicultura (PEVS)*: dados sobre a quantidade e o valor das produções obtidas mediante o processo de exploração dos recursos florestais, organizados em dois grandes grupos: extração vegetal e silvicultura.

2.3. Tableau

O Tableau é um software para produção de painéis interativos para visualização de dados em duas dimensões. Esses painéis são construídos a partir da combinação de múltiplas fontes de dados, desde de planilhas, passando por Sistemas Gerenciadores de Banco de Dados Relacionais (SGBDs) até dados espaciais em formato vetorial, como arquivos shapefile.

Conceitualmente, o Tableau categoriza os atributos em dois grupos: *dimensões* e *medições*. Dimensões correspondem a valores qualitativos utilizados para categorização e segmentação, que podem ter ou não algum critério de ordenação (por exemplo, meses do ano). Medições correspondem a valores numéricos e quantitativos, passíveis de agregação.

3. SOLUÇÃO PARA A ELABORAÇÃO DE MAPAS AGRÍCOLAS INTERATIVOS

3.1. Visão Geral

A Figura 1 apresenta uma visão geral das etapas para a elaboração dos mapas agrícolas interativos em páginas

web. A solução recebe três entradas, indicadas à esquerda da figura como E1, E2 e E3, possui cinco etapas de processamento, indicados na parte central da figura com a respectiva numeração e fornece como saída uma página web (ou um conjunto de páginas) contendo mapas e painéis interativos, indicado à direita na figura. A execução de cada etapa é realizada manualmente em diferentes plataformas, sendo possível diferentes níveis de customização de acordo com os requisitos necessários.

A entrada **E1** corresponde aos dados qualitativos obtidos das pesquisas anuais do IBGE (descritos na seção 2.2), contendo, em cada registro, os códigos de cada território (geocódigos) e os dados relacionados ao quantitativo estimado de cada cadeia considerada (quantidade produzida para lavouras ou efetivo de rebanho para pecuária, por exemplo). É importante ressaltar a obrigatoriedade do geocódigo, identificador espacial unívoco oficial do IBGE para determinação de um território. A entrada **E2** corresponde aos arquivos vetoriais contendo as malhas digitais oficiais do IBGE (descritos na seção 2.1). A escala territorial de **E2** precisa estar compatível com a escala dos dados qualitativos de **E1** (ambos em nível municipal, por exemplo).

A entrada **E3** corresponde aos requisitos desejáveis para os mapas e painéis interativos, assim como os elementos complementares que irão compor sua apresentação em uma ou mais páginas web. Esses requisitos irão influenciar, por exemplo, nos procedimentos de transformação a serem aplicados aos dados, nos processamentos e cálculos estatísticos adotados ou na ênfase desajada na construção dos elementos de simbologia dos mapas. Além disso, podem determinar os elementos de controle e navegação da página web, influenciando indiretamente a granularidade espacial dos painéis ou o nível de especificidade/generalidade das entidades durante o processo de modelagem de dados.

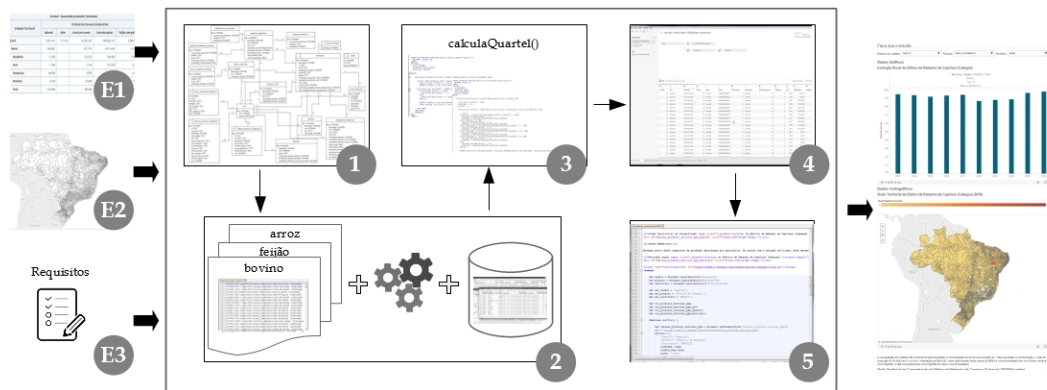
Os próximos parágrafos detalham o funcionamento de cada uma das etapas.

Etapa 1: Modelagem de Dados

Esta etapa considera as entradas **E1**, **E2** e **E3** e consiste na modelagem de um banco de dados objeto-relacional normalizado que irá comportar todos os dados necessários. Consideramos o banco de dados como objeto-relacional no sentido de que partimos de uma abordagem relacional, porém acrescentamos entidades espaciais na modelagem, visando comportar tanto os dados qualitativos quanto os cartográficos. A modelagem considera a granularidade territorial desejada (município ou microrregião, por exemplo) e os níveis de agregação pertinentes (estados, regiões ou recortes territoriais, como o semiárido ou o MATOPIBA). Além disso, considera as variáveis desejadas na interação com os painéis e mapas dinâmicos. Finalizada a modelagem, o banco é criado utilizando-se o SGBD PostgreSQL com a extensão espacial PostGIS. O PostGIS adiciona ao PostgreSQL a capacidade de armazenamento e recuperação de entidades geográficas, além de implementar diversas funcionalidades para análises espaciais e topológicas que estendem o SQL, a linguagem de consulta padrão.

¹sidra.ibge.gov.br

Figura 1: Visão Geral



Etapa 2: ETL

Esta etapa também considera todas as entradas e a modelagem da etapa anterior e realiza os procedimentos de ETL (*Extract - Transform - Load*). O passo de extração segmenta **E1** por cadeia agropecuária – ou seja, cria-se um arquivo dedicado para cada cadeia considerada adequada ao modelo de dados definido. Cada arquivo contém dados de produção, da pesquisa a que pertencem (PAM, PPM, por exemplo), da categoria da cadeia (lavoura ou pecuária, por exemplo) e o ano em que ela foi coletada. Apenas são listados os territórios que possuem dados de produção.

O passo de transformação aplica as correções necessárias tanto em **E1** como **E2**. No caso de **E1**, trata-se, por exemplo, as alterações em unidades de medida indicadas pelo IBGE ou de consenso da comunidade agrícola visando viabilizar a análise integrada de múltiplos anos. Para a produção de frutas, por exemplo, houve uma mudança em 2001: anteriormente os dados retratavam o número de frutos produzidos e, após esse ano, passou a retratar o peso total da produção. Para o café, por outro lado, o IBGE apenas indica que a partir de 2002 a quantidade publicada é em relação ao grão e, antes dessa data, em coco. Transformações também podem ser aplicadas nos dados cartográficos visando correções topológicas ou remoção de ilhas cuja produção é praticamente nula.

O passo de carga utiliza os arquivos fragmentados e corrigidos para povoar o banco de dados modelado na etapa 1. A carga é realizada por meio de instruções SQL de inclusão massiva (COPY). Os passos descritos nessa etapa são executados via rotinas implementadas em Java.

Etapa 3: Processamento e Cálculos Adicionais

Esta etapa atua diretamente no banco de dados e realiza processamentos ou cálculos estatísticos, de acordo com os requisitos de **E3**. Em dados de lavouras, por exemplo, atributos calculados como produtividade ou área perdida (diferença entre área plantada e área colhida) são comumente pertinentes. Outra análise normalmente aplicada é o cálculo da concentração de produção de cada cadeia [5]. O resultado dessa análise rotula cada território em quatro categorias (Q4, Q3, Q2 e Q1) de acordo com seu *ranking* de importância. Esses resultados serão posteriormente utilizados na definição da simbologia dos painéis de mapas interativos temáticos ou apresentados como atributos dos territórios definidos nesses painéis. Os cálculos são realizados diretamente no banco de

dados por meio de *stored procedures* na linguagem *pgSQL* ².

Etapa 4: Elaboração dos Painéis Interativos

Esta etapa consiste na elaboração dos painéis interativos no Tableau de acordo com os requisitos **E3**. Os painéis podem conter gráficos, tabelas ou mapas. A etapa consiste no estabelecimento de uma conexão com o banco de dados criado, na seleção de todas as relações pertinentes para os painéis e na definição dos critérios de junção entre essas relações. Após essa definição das variáveis a serem apresentadas, define-se também no painel as simbologias quantitativas e qualitativas associadas aos mapas. É possível, também, a criação via Tableau de novos campos calculados quando necessário. Esses campos podem ser relevantes na apresentação de sumarizações por agregação territorial ou cálculo de percentuais relativos. Após a elaboração, os painéis são publicados em uma plataforma de hospedagem ³ para posterior inserção em páginas web.

Etapa 5: Elaboração das Páginas Web

Esta etapa realiza a implementação de uma ou mais páginas navegáveis e interativas contendo, de forma integrada, conteúdos textuais e painéis dinâmicos. Todas as informações apresentadas são provenientes do banco de dados gerado nas etapas anteriores. Os requisitos **E3** são determinantes para a definição dos seletores e as opções de navegabilidade, assim como a integração entre os painéis interativos. A página web construída nessa etapa é especificada em HTML, CSS e utiliza JavaScript para a implementação das atualizações de conteúdo. Os painéis interativos publicados são acessados e filtrados também via API JavaScript ⁴. A publicação dessa página – em containers web ou gerenciadores de conteúdo (por exemplo, o Liferay) – materializa o resultado gerado pela solução.

4. RESULTADOS

A solução vem sendo incrementalmente desenvolvida e já foi aplicada em diferentes ocasiões, mostradas na Figura 2 itens (a) e (b). Ambos resultados foram publicados no gerenciador de conteúdo Liferay ⁵, juntamente com as páginas corporativas.

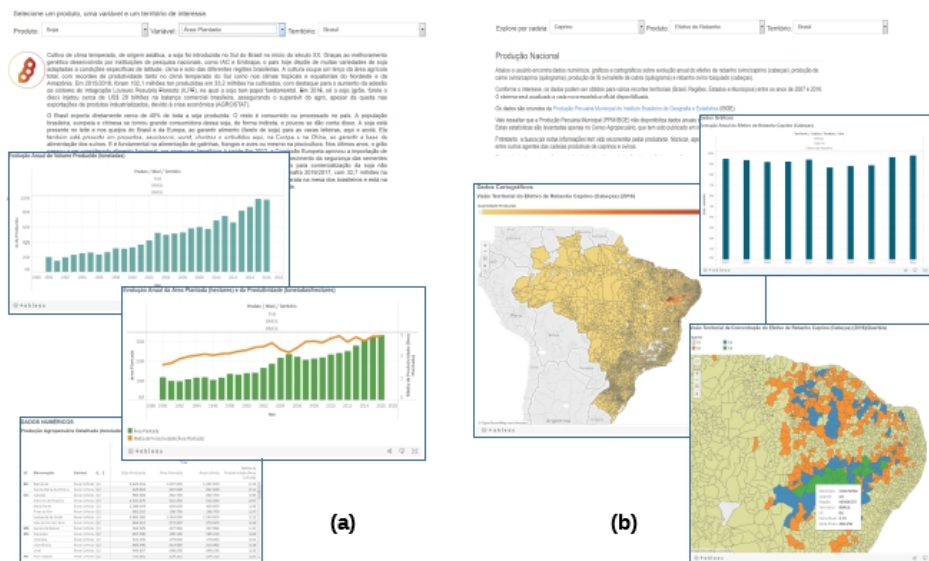
² www.postgresql.org/docs/9.6/static/plpgsql.html

³ public.tableau.com

⁴ https://onlinehelp.tableau.com/samples/en-us/js_api/tutorial.html

⁵ www.liferay.com

Figura 2: Resultados gerados pela ferramenta



A Figura 2(a) ilustra sua utilização no Sistema de Inteligência Territorial Estratégica da Macrologística Agropecuária Brasileira ⁶. Esse sistema apresenta as 10 maiores cadeias de produtos agropecuários em termos de volume exportado, utilizando dados do PAM, PPM e PEVS, no intervalo temporal de 1990 e 2016. O nível territorial adotado foi microrregião e considerou-se as agregações espaciais: Brasil, Regiões, Unidades Federativas e Bacias Logísticas (um recorte específico do projeto). Tinha-se como requisito uma interface genérica (e, por consequência, painéis únicos) para apresentar o volume de produção de todas as cadeias e a seleção de variáveis deveria ser de acordo com a cadeia selecionada. Para as lavouras, são apresentadas área plantada, área colhida e produtividade. Esses resultados estão disponíveis na página *Produção Agropecuária* do portal.

A Figura 2(b) apresenta um resultado análogo no Centro de Inteligência e Mercado de Caprinos e Ovinos ⁷. Essa página agrega dados sobre pecuária e produtos de pecuária relativos às culturas de caprinos e ovinos, no intervalo temporal de 2007 a 2016. Dois mapas interativos compõem a página: um deles com simbologia de quebras naturais, que mostra uma visão geral do Brasil; outro com as quatro classes referentes aos quartéis de concentração de produção [5], mostrado com ênfase no semiárido, área concentradora da produção de caprinos no Brasil.

5. CONCLUSÕES

Este trabalho apresentou uma solução para a publicação de dados agropecuários na web por meio de painéis dinâmicos, compostos de tabelas, mapas e gráficos. Os casos de utilização da solução tiveram resultados positivos e boa aceitação pelo público geral – aceitação avaliada pela popularidade/número de acessos mensurados para os portais onde as páginas foram hospedadas.

Apesar de utilizar como base os dados tabulares oriundos

das pesquisas anuais do IBGE, não há restrições para a utilização da solução em outros contextos. Diferentes requisitos podem ser considerados como entrada da solução e influenciar desde a modelagem do banco de dados à elaboração dos painéis. A única premissa assumida pela solução é a existência de valores quantitativos associados a um território passível de espacialização via dados cartográficos oficiais.

Dentre os próximos passos para aprimoramento da solução destacamos a automatização do processo de obtenção de dados do IBGE. O SIDRA possui uma API ⁸ que permite a extração de dados automática por agentes de software, o que poderia agilizar o processo. Outra extensão prevista é a utilização de dados espaciais publicados como geosserviços (WMS/WFS) para a elaboração dos mapas interativos no Tableau, dado o grande volume de dados geográficos institucionais disponíveis nesse formato ⁹.

6. REFERÊNCIAS

- [1] MIRANDA, E. E. de; MAGALHÃES, L. A.; CARVALHO, C. A. de. *Um Sistema de Inteligência Territorial Estratégica Para o MATOPIBA*. 2014.
- [2] SILVA, W. C. et al. Pastagem.org: um portal de dados e informações georeferenciadas para uma pecuária mais produtiva e ambientalmente sustentável. In: *Anais do Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*. [S.l.: s.n.], 2015. p. 2521–2528.
- [3] LOURENÇO, R. B. et al. Atlas: A visualization and analysis framework for geospatial datasets. In: *Anais Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*. [S.l.: s.n.], 2017. p. 7301–7308.
- [4] IBGE, I. B. de Geografia e E. *Pesquisas Agropecuárias. Série Relatórios Metodológicos, volume 6*. 2002.
- [5] GARAGORRY F. L.; CHAIB FILHO, H. *Elementos de agrodinâmica (Manuscrito)*. 2008.

⁶www.embrapa.br/macrologistica

⁷www.embrapa.br/cim-inteligencia-e-mercado-de-caprinos-e-ovinos

⁸api.sidra.ibge.gov.br

⁹www.inde.gov.br