



Especialização Pós-graduação em Big Data
Visualização de Dados
Professor Mestre Rodrigo Sant'Ana

PBL – Problem Based Learning

Aprendizado baseado em problema

Alunos

Wagner Correia,

Márcio Fossa,

Eric Ruiz

Julho/2020

MOTIVAÇÃO

Proposta do trabalho apresentado

No momento em que estamos passando por uma pandemia, a informação correta, adequada, de fácil compreensão e acessível, é um diferencial que auxilia consideravelmente para que cada pessoa saiba como está a situação em sua cidade, bem como seu bairro.

Atualmente, a divulgação dos boletins epidemiológicos oficiais do município de Itajaí-SC, é feita de forma online, porém textual e descritiva (ex.: <https://www.itajai.sc.gov.br/noticia/25324>), o que pode dificultar sua interpretação e compreensão de forma clara.

Assim, propomos a publicação dos boletins através da elaboração de painéis online, com gráficos, mapas e tabelas de dados, a fim de facilitar o acompanhamento e a compreensão da evolução da doença no município por parte da população em geral e todos os demais interessados.

Público alvo

População em geral, principalmente a itajaiense, mas não se limitando a ela, considerando levar ao conhecimento da maior parte da população possível, com acesso online, via internet, seja por meio de computador, laptops, tablets ou smartphones.

TECNOLOGIAS

Business Intelligence (BI), ou Inteligência de Negócios, vem sendo cunhada a décadas e desde a década de 90 sempre muito alinhada à tecnologia de informação e comunicação. Se caracteriza por um conjunto de técnicas aplicadas à gestão e administração. Fomenta a premissa de se tomarem decisões com base em fatos, números, métricas e estatísticas. As informações se constroem primariamente na coleta e armazenamento de dados, seguido de uma etapa de análise e validação, por sua vez, são aplicadas técnicas de ciência de dados e da administração, possibilitando assim, a construção das informações e conhecimento na forma, principalmente, de métricas que nortearão tomadas de decisão a que se espera, desse modo, serem mais assertivas.

Com relação a escolha das tecnologias empregadas, dado ao caráter e interesse público da informação, bem como o possível interesse de utilização dos órgãos públicos, uma das premissas foi o emprego somente de ferramentas e tecnologias livres e de código aberto, desse modo, sem ônus de licenças e afins para todo e qualquer interessado fazer sua própria implementação, bem como trabalhos derivados.

R é uma linguagem e ambiente para computação e gráficos estatísticos. É um projeto GNU (software livre) e é semelhante à linguagem e ambiente S, desenvolvido nos Laboratórios Bell (anteriormente AT&T, agora Lucent Technologies) por John Chambers e colegas.

O ambiente R é altamente extensível por meio de bibliotecas, está disponível como software livre sob os termos da Free Software Foundation's GNU General Public License, pode ser executado em diversas plataformas como Linux, Windows e MacOS, está disponível em <https://www.r-project.org/>.

Shiny, é uma biblioteca R que auxilia na criação de aplicações web interativas a partir do R, aceita a utilização de recursos web como HTML, CSS e Javascript para aprimorar as aplicações criadas, ou até mesmo interagir com aplicações independentes, maiores informações podem ser encontradas em <https://shiny.rstudio.com/>.

Plotly é uma biblioteca para a geração de gráficos interativos, que se integra com R e Shiny, sendo também responsivo, o que facilita a apresentação dos gráficos em dispositivos móveis. Disponível em <https://plotly.com/r/>.

Shiny Server, é um servidor opensource, para execução das aplicações desenvolvidas com Shiny, permite que você possa publicar e tornar on-line essas aplicações, está disponível em <https://shiny.rstudio.com/deploy/>.

Técnicas de apresentação de dados escolhidas

Com base nas métricas encontradas no boletim epidemiológico oficial em sua forma original textual, mantendo em mente também um conceito de dashboard, ou seja, um painel de monitoramento com essas métricas dispostas. Foi proposto um dashboard com as 4 métricas mais relevantes para o acompanhamento da evolução da COVID-19 no município de Itajaí numeradas a seguir em ordem de destaque:

1. Casos Ativos;
2. Total de Confirmados;
3. Total de Curados;
4. Total de Óbitos.

Desse modo, classificamos essas métricas como sendo os mais importantes KPIs, Key Performance Indicators, Indicadores-chave de Desempenho nesse cenário.

Assim sendo, foram elaborados blocos retangulares em destaque no topo da página inicial, com alto contraste, com uma classificação de cores, bem como ícones indicativos. Esses recursos gráficos foram cuidadosamente escolhidos, pois consideramos também um público com baixa visão e também um público com daltonismo. Desse modo, possibilitando o acesso e fácil navegação para essas pessoas.

Dado que Casos Ativos, que corresponde ao Total de Casos Confirmados, menos o Total de Óbitos e menos o Total de Curados, não se trata de uma métrica acumulativa, ou seja, esse dado não se soma ao longo do tempo, pois a cada dia novos casos surgem, bem como pessoas são curadas e, também, infelizmente, algumas perdem suas vidas face à COVID-19, caracteriza-se assim uma métrica não-cumulativa ao longo do tempo. Desse modo, para apresentar a evolução cronológica dos casos ativos, foi utilizado um gráfico de barras verticais, com seu eixo x na linha do tempo, e seu eixo y apresentando os Casos Ativos em cada dia. Seguindo a identidade visual do KPI Casos Ativos, foram implementados o mesmo ícone e cor na identidade visual do gráfico.

Por se tratar de uma métrica acumulativa, diferentemente do gráfico de Casos Ativos, no gráfico de Casos Confirmados ao longo do tempo, foi utilizado um gráfico de linha para representar visualmente esse acúmulo. Sempre seguindo a premissa da identidade visual como elemento de apoio à navegação nos dados, foi utilizado o mesmo ícone e cor do KPI Total de Casos Confirmados na produção do gráfico de linha. Do mesmo modo, ancorando assim a informação total à sua visualização distribuída ao longo do tempo.

Também foi dedicada uma página específica para a visualização dos dados de Casos Confirmados, onde é possível verificar sua distribuição por gênero, faixa etária e bairro. Cada uma dessas visualizações foi disposta cada uma em sua aba específica dentro dessa página.

Na aba gênero, foi utilizado um gráfico de rosca para apresentar o comparativo de Casos Confirmados distribuídos por gênero. Por haver apenas dois elementos de comparação e tendo em vista que o número de elementos é imutável, não foi escolhido o gráfico de barras. Para diferenciar cada porção do gráfico, foram utilizadas diferentes nuances da cor de referência implementada no KPI Casos Confirmados, mantendo assim a diferenciação de cada área do gráfico e sua legenda, mas também respeitando a identidade visual do KPI a que se refere.

Nas abas faixa etária e bairro, mais uma vez utilizamos gráfico de barras verticais, por se tratar de dados categóricos. Também seguindo o esquema de cores conforme já descrito, bem como ícone de referência. Em ambos os casos, a categoria se apresenta no eixo x e a quantidade no eixo y.

Já na aba Mapa dos Bairros, fizemos uso do georreferenciamento do município de Itajaí e seus bairros para quantificar os casos confirmados por bairro, porém numa visão de mapa de densidade, onde quanto mais casos, mais avermelhado fica o bairro. Desse modo dando a dimensão geoespacial do mesmo dado para o usuário, visão essa, que no gráfico de barras verticais por bairro é impossível de se observar. Vale lembrar que, até onde sabemos, a COVID-19 se dissemina principalmente no contato social humano, ou seja, cada ser humano infectado é um vetor de transmissão e essa transmissão só ocorre com a aproximação ou mesmo contato físico. Portanto, saber e conhecer a distribuição dos Casos Confirmados geoespacializados é de suma importância.

Com relação aos Casos de Óbitos, semelhantemente aos Casos Confirmados, uma página específica foi dedicada somente a esse tema, onde se pode observar os dados em distribuições por gênero, faixa-etária e uma linha do tempo. Semelhantemente aos gráficos de distribuições por gênero e por faixa etária dos Casos Confirmados, as distribuições de Casos de Óbitos foram feitas seguindo os mesmos conceitos, porém é claro, respeitando seu próprio ícone e cor de referência.

Ainda no tema dos Casos de Óbitos, considerando sua característica acumulativa, foi feito um gráfico de linha para apresentar a evolução dos óbitos numa linha do tempo. Para tal, apresentamos no eixo x as datas dia a dia e, no eixo y, as quantidades acumuladamente. Seguindo também as cores e o ícone de referência.

Foi criada também uma página inteiramente dedicada a explicar cada métrica. As informações foram retiradas do próprio boletim textual oficial do município de Itajaí.