

Especificação do Projeto Final da Disciplina

Descrição Geral do Projeto

Inicialmente, a turma deverá se reunir em grupos de no mínimo 3 e no máximo 4 alunos. Um recurso de “escolha de grupo” foi disponibilizado no Moodle para facilitar esta tarefa. Depois, cada grupo deverá escolher um tema e trabalhar na solução do problema proposto. Um recurso de “escolha” foi disponibilizado no Moodle para esta tarefa. É obrigatório que **um aluno por grupo preencha a escolha do tema** do grupo.

O objetivo principal deste projeto é proporcionar aos alunos a oportunidade de projetar e implementar uma solução para análise visual de dados de diálise do DATASUS. Cada grupo deverá escolher entre um dos seguintes temas, sendo obrigatório ter dois grupos por tema:

1. **Perfil e Internação de Pacientes em Diálise em Porto Alegre**

Questões iniciais: Qual a origem dos pacientes de diálise (hemodiálise e diálise peritoneal) que internam em POA/RS? Quais são os principais motivos de hospitalização? O perfil dos pacientes varia entre os diferentes hospitais? Existe alguma mudança nesse comportamento ao longo dos anos?

Objetivo: Desenvolver um dashboard interativo que permita analisar dados de um hospital e comparar dados de diferentes hospitais, identificando tendências e padrões de internação dos pacientes em diálise ao longo do tempo.

2. **Tempo de Espera para Acesso Vascular na Diálise Crônica em Porto Alegre**

Questões iniciais: Após o início da diálise crônica, qual é o tempo médio até a confecção da Fístula Arterial Venosa (FAV) nos hospitais de POA/RS? Existem perfis de pacientes que apresentam maior demora para a confecção da FAV? Houve mudanças na demora ao longo dos anos, especialmente em períodos críticos como a pandemia de COVID-19?

Objetivo: Criar um dashboard que permita analisar a série temporal de dados sobre a confecção da FAV, visando identificar padrões e fatores que influenciam a demora no acesso vascular.

3. **Frequência e Distribuição de Partos em Pacientes em Diálise em Porto Alegre**

Questões iniciais: Quantas pacientes em diálise realizaram partos durante o tratamento de substituição renal (hemodiálise, diálise peritoneal e transplante) em POA/RS? Quais hospitais de Porto Alegre registraram mais partos dessas pacientes? Houve variação significativa no número de partos ao longo dos anos, tanto no total quanto por hospital?

Objetivo: Desenvolver um dashboard que permita visualizar a evolução temporal e a distribuição dos partos entre pacientes em diálise, destacando possíveis tendências e variações entre hospitais.

O detalhamento de cada tema, bem como o processo de coleta de dados, será apresentado pelos proponentes na aula do dia 10/04, mas os entregáveis de cada parte do projeto são apresentados a seguir.

Parte I

Descrição

A primeira parte do projeto, simplificada, consiste: na coleta, análise e preparação dos dados; na definição das perguntas de negócio; na especificação do ambiente de desenvolvimento; e no projeto da solução.

Considerando o processo CRISP-DM, o *pipeline* de visualização de dados e o processo de *visual analytics*, nesta primeira parte do desenvolvimento do projeto devem ser seguidas as seguintes etapas:

1. Entendimento do negócio e do tema proposto;
2. Coleta dos dados;
3. Entendimento dos dados através de uma análise exploratória visual;
4. Preparação e pré-processamento de dados, que pode também incluir o enriquecimento dos dados se necessário;
5. Estudo do ambiente de desenvolvimento e projeto das representações visuais que serão implementadas.

Espera-se que sejam utilizadas as ferramentas apresentadas nas aulas (por exemplo, ydata-profiling e Tableau Prep) para auxiliar a analisar as características e a qualidade dos dados, e a necessidade de fazer algum pré-processamento ou enriquecimento dos dados.

A partir do tema proposto e da coleta e análise dos dados, devem ser definidas as perguntas de negócio, ou seja, o que se quer descobrir a partir das visualizações propostas. Algumas perguntas já foram apresentadas, mas cada grupo deve propor ao menos uma questão adicional. Não há um número máximo de perguntas, cada grupo deve elaborar a quantidade de perguntas que achar mais adequada.

Também deve ser escolhido e estudado o ambiente de desenvolvimento do *dashboard* que será criado na parte II do projeto. É obrigatório o uso de um ambiente gratuito, como, por exemplo, o Streamlit (<https://streamlit.io/>) ou o Dash (<https://dash.plotly.com/>).

A última etapa da parte I do projeto consiste na escolha, justificativa e descrição das visualizações interativas que serão implementadas na parte II do projeto. Cada grupo deve escolher diferentes formas de visualização (*heatmaps*, gráficos de linha, *small multiples*, etc.), **justificar** as escolhas e descrever como os dados serão utilizados/mapeados, identificando qual pergunta de negócio responde e porque a técnica de visualização escolhida é adequada para analisar os dados.

Os entregáveis da parte I do projeto são:

- **Relatório com as seguintes informações:** (1) Nomes dos componentes do grupo; (2) descrição do conjunto de dados, dos *insights* obtidos com a análise exploratória visual (incluindo gráficos) e do enriquecimento dos dados (se realizado); (3) apresentação das perguntas de negócio; (4) ambiente de desenvolvimento escolhido; (5) descrição das visualizações e interações que serão implementadas com uma **justificativa** para escolha de cada uma e os dados que serão utilizados. Não há um formato pré-definido para o relatório, mas ele deve ser um documento escrito (não uma apresentação de PowerPoint) que deve conter todas as informações especificadas apresentadas de uma forma objetiva. Cada item descrito acima deve ser colocado como um título. Recomenda-se o uso de folha tamanho A4 com orientação retrato, uma coluna, fonte tamanho 12 e espaçamento simples de parágrafo. Títulos devem estar com uma fonte maior e em negrito.
- **Pitch:** o grupo deve preparar e apresentar um *pitch* de 7 a 9 minutos. O objetivo é mostrar a análise dos dados, apresentar as perguntas de negócio e as vantagens da solução proposta para o problema apresentado, justificando as escolhas feitas.

Data

Apresentação do *Pitch* e entrega do relatório no Moodle: 15/05/2025 (durante o horário de aula).

Critérios de avaliação

Serão considerados para avaliação da parte I os seguintes critérios:

- Relatório: inclusão de todo conteúdo solicitado; relevância e profundidade da análise dos dados; relevância, originalidade, eficácia e **justificativa** das visualizações interativas propostas; capacidade de síntese e escrita.
- *Pitch*: qualidade da apresentação (clara e fluente); respeito ao tempo estabelecido; capacidade de síntese para apresentação da solução proposta e justificativa das escolhas.

Parte II

Descrição

A segunda parte do projeto consiste na implementação, teste e validação da solução proposta na parte I.

Considerando o processo CRISP-DM, o *pipeline* de visualização de dados e o processo de *visual analytics*, inicialmente deve ser feito o mapeamento dos dados para as representações visuais especificadas na parte I. A implementação das visualizações e interações projetadas deve ser feita usando o ambiente de desenvolvimento escolhido. Estas visualizações interativas devem ser organizadas, obrigatoriamente, em um ou mais *dashboards* e permitir a exploração dos dados de maneira efetiva e intuitiva, possibilitando responder as perguntas de negócio especificadas. É importante explorar diferentes formas de visualização, como, por exemplo, *heatmaps*, gráficos de linha e *small multiples*, além de outros gráficos que possam apresentar informações complementares, tais como informações estatísticas. Também devem ser consideradas as boas práticas que serão vistas em aula para o desenvolvimento das visualizações e do(s) *dashboard(s)*, incluindo os *design patterns* para *dashboards* (<https://dashboarddesignpatterns.github.io/>).

Depois, a solução implementada deve ser testada e avaliada para verificar se responde a todas as perguntas de negócio definidas na parte I do projeto. Nesta etapa pode ser necessário fazer alguns ajustes nas visualizações interativas propostas, pois é importante que elas possibilitem extrair as informações necessárias, permitindo solucionar os problemas propostos.

Os entregáveis da parte II do projeto são:

- **Link ou vídeo com uma demonstração da solução proposta:** se a solução estiver disponível *online*, deve ser incluído o *link* no relatório. Caso contrário, deve ser preparado um vídeo de no **máximo 2 minutos** para demonstrar as funcionalidades da solução proposta e como elas permitem responder às perguntas de negócio e obter *insights*. O vídeo deve ser colocado em algum drive compartilhado ou no YouTube e o *link* para ele deve ser colocado no relatório.
- **Relatório com as seguintes informações:** (1) Nomes dos componentes do grupo; (2) descrição das visualizações e interações implementadas, demonstrando a facilidade de uso das mesmas e os padrões utilizados para o desenvolvimento do(s) *dashboard(s)*; (3) análise dos dados utilizando as visualizações implementadas, incluindo as respostas para as perguntas de negócio e os *insights* obtidos com o uso da solução implementada; (4) discussão dos benefícios e limitações da solução proposta e apresentação de possíveis melhorias como trabalho futuro; (5) *link* para a solução

implementada ou vídeo com a demonstração. Não há um formato pré-definido para o relatório, mas ele deve ser um documento escrito (não uma apresentação de PowerPoint) e todas as informações especificadas devem ser apresentadas de uma forma objetiva. Recomenda-se o uso de folha tamanho A4 com orientação retrato, uma coluna, fonte tamanho 12 e espaçamento simples de parágrafo. Títulos devem estar com uma fonte maior e em negrito.

- **Pitch:** o grupo deve preparar e apresentar um *pitch* de 7 a 9 minutos. O objetivo é mostrar o funcionamento, as vantagens e a inovação da solução implementada, bem como os resultados obtidos (respostas das perguntas de negócio e *insights* identificados). As limitações da solução também devem ser abordadas.

Data

Apresentação do *Pitch* e entrega do relatório no Moodle: 24/06/2025 (durante o horário de aula).

Critérios de avaliação

Serão considerados para avaliação da parte II os seguintes critérios:

- Solução implementada: qualidade visual seguindo as boas práticas estudadas, variedade de representações visuais interativas, funcionalidades desenvolvidas e *insights* obtidos.
- Relatório: inclusão de todo conteúdo solicitado; relevância, originalidade e eficácia das visualizações interativas propostas; relevância e profundidade da análise dos resultados obtidos; criatividade e qualidade da solução proposta; capacidade de síntese e escrita.
- *Pitch*: qualidade da apresentação (clara e fluente); respeito ao tempo estabelecido; capacidade de síntese para demonstração da solução proposta e dos *insights* obtidos.