

INFORMAÇÕES SOBRE A AVALIAÇÃO

LAB04	Laboratório 04 - 20 pontos
--------------	-----------------------------------

INFORMAÇÕES DOCENTE						
CURSO: ENGENHARIA DE SOFTWARE	DISCIPLINA: LABORATÓRIO DE EXPERIMENTAÇÃO DE SOFTWARE	TURNO	MANHÃ	TARDE	NOITE	PERÍODO/SALA: 6º
					x	
PROFESSOR (A): João Paulo Carneiro Aramuni						

Visualização de dados utilizando uma ferramenta de bi

Introdução

Business Intelligence (BI) refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que apoiam tomadas de decisão baseadas em dados. Dessa forma, BI é dado por um conjunto de soluções que buscam converter um grande volume de dados em informações relevantes para o negócio, facilitando a tomada de decisões estratégicas em relação às atividades de uma empresa. Os dashboards são artefatos atualizados automaticamente em frequências definidas, carregando os dados provenientes dos sistemas em visualizações gráficas e possibilitando que os seus usuários tenham informações úteis e atualizadas em tempo hábil. Por meio desse artefato, o usuário pode detectar falhas e identificar possibilidades de melhorias, aumentando e melhorando a produtividade na tomada de decisões para o negócio. No contexto de experimentação em engenharia de software, as técnicas de BI e os dashboards produzidos podem ser úteis para melhor explorar os dados coletados e responder às questões de pesquisa levantadas de maneira clara e objetiva.

Assim, o objetivo deste laboratório é utilizar uma ferramenta de BI (Microsoft Power BI, Tableau ou Google Data Studio) para realizar uma melhor apresentação de resultados de trabalhos de experimentação, explorando os dados de TIS 6.

Especificação do Dashboard

1. Caracterização do Dataset

A primeira parte do dashboard produzido deve caracterizar o dataset utilizado no trabalho. Ou seja, apresentar, através de visualizações adequadas, as características mais relevantes dos objetos de estudo utilizados na pesquisa. Por exemplo, se o dataset é composto por repositórios do GitHub, apresentar dados tais quais os coletados no



PUC Minas

Laboratório 01. Outros elementos que podem constituir o dataset (lista não exaustiva): usuários do GitHub, sites da web, posts no StackOverflow, etc.

Importante: se o estudo utiliza alguma característica dos dados para particionar o dataset em análises comparativas, é essencial que a caracterização do dataset apresente as características do dataset completo e também dos subgrupos formados. Por exemplo, em estudos que comparam linguagens de programação diferentes, é necessário apresentar as características de todos os repositórios coletados, assim como dos grupos formados por cada linguagem. Em estudos de avaliam períodos de tempo diferentes, deve-se apresentar os dados de todo o período analisado, bem como dos intervalos que serão comparados.

2. Questões de Pesquisa

Para cada pergunta de pesquisa proposta no GQM do trabalho, deve-se incluir uma ou mais visualizações que a responda. É essencial que todas as métricas propostas sejam apresentadas de modo que as perguntas sejam respondidas de maneira completa. Além disso, o dashboard deve ser auto-explicativo. Ou seja, apresentar a pergunta proposta e, logo em seguida, as visualizações que a respondem. Portanto, é essencial que o dashboard conte a história expressa pelos dados analisados.

Atenção: É importante escolher de maneira adequada as medidas de tendência central que serão utilizadas para sumarizar os dados. Além disso, as labels dos eixos de cada visualização também devem ser claras e expressar a medida escolhida.

Apresentações em Aula

A fim de acompanhar o desenvolvimento do dashboard e discutir melhorias, **as entregas parciais serão apresentadas em aula**. Assim, todos os grupos deverão apresentar o incremento produzido durante o horário de aula, com o resultado do que é esperado para a Sprint. Os grupos que não apresentarem sofrerão penalidade na nota da Sprint. Além disso, grupos que não estiverem completos durante a apresentação também serão penalizados (conforme definição abaixo).

Relatório Final

Além do dashboard produzido, cada grupo deverá atualizar o artigo final que está produzindo na disciplina de TIS 6 para incluir os gráficos criados. Recomenda-se que os gráficos de caracterização do dataset sejam incluídos na Seção 3 (Metodologia), enquanto os gráficos das RQs sejam inseridos na Seção 4 (Resultados). Cada gráfico deve ser citado e explicado no texto. **A entrega final consistirá do dashboard (salvo em pdf) e do o artigo de TIS6 atualizado.**

Trabalho Alternativo

Para os alunos que não estão matriculados em TIS 6, um dashboard deve ser preenchido utilizando alguma base de dados pública (dados do governo, dados abertos, ou outras bases disponíveis gratuitamente). A entrega final deverá contemplar o dashboard e um relatório contendo (i) introdução; (ii) metodologia/descrição da base; (iii) resultados; (iv) discussão.

Processo de Desenvolvimento

1. Sprints e Entregas

Lab04S01: Caracterização do dataset + Apresentação em aula (5 pontos)

Lab04S02: Visualizações para as RQs 1 e 2 + Apresentação em aula (5 pontos)

Lab04S03: Dashboard final (Caracterização + RQs) + Artigo de TIS 6 atualizado (10 pontos)

Grupo 1 - Terça: Prazo final: 02/06 | Apresentação: 03/06

Grupo 2 - Sexta: Prazo final: 29/05 | Apresentação: 30/05

Valor total: 20 pontos | Desconto de 50% da nota da sprint para grupos que não apresentarem em aula | É obrigatória a participação de todos os membros do grupo na apresentação em aula (sob pena do desconto de 50% da nota da sprint para todo o grupo).

2. Material de Apoio

- Tableau: <https://help.tableau.com/current/guides/get-started-tutorial/pt-br/get-started-tutorial-home.htm>
- Power BI: <https://docs.microsoft.com/en-us/power-bi/fundamentals/service-get-started>
- Google Data Studio: <https://datastudio.google.com/>