

Regulamento do TCC - Business Intelligence e Analytics

1 - Introdução

Nesse documento, apresentamos as informações necessárias para o desenvolvimento do Trabalho de Conclusão de Curso (TCC) para o curso de Pós Graduação Lato Sensu em **Business Intelligence e Analytics**. O TCC se constitui de um trabalho interdisciplinar realizado individualmente em que o aluno coloca em prática tudo o que foi visto nas outras disciplinas. O propósito deste trabalho é consolidar os conhecimentos aprendidos durante o curso, dando ao aluno a oportunidade de verificar sua competência em atender demandas reais por meio do desenvolvimento de uma solução completa.

O contexto e problema a ser abordado no TCC deve ser escolhido pelo aluno. Dessa forma, espera-se que os conhecimentos possam ser aplicados em um projeto alinhado com os interesses do aluno. Nesse ponto, é importante ressaltar que, mesmo tendo a possibilidade de escolha do tema, o aluno deverá observar cuidadosamente e cumprir o conjunto de requisitos e as restrições técnicas que fazem parte deste trabalho, descritos no item **Escopo do Trabalho**.

Os itens que devem ser produzidos para o fechamento da disciplina são descritos no item **Entrega** que fica ao final deste documento.

Em caso de dúvidas sobre as regras ou procedimentos do TCC, o aluno deve utilizar o [fórum da disciplina \(https://pucminas.instructure.com/courses/1754/discussion_topics/152768\)](https://pucminas.instructure.com/courses/1754/discussion_topics/152768). Desta forma, espera-se que este espaço sirva de orientação para todos os alunos uma vez que podem ver as perguntas realizadas pelos demais alunos e os esclarecimentos sobre elas.

2 - Escopo do trabalho

Para o contexto escolhido pelo aluno, deverá ser apresentado um objetivo a partir da base de dados selecionada para o projeto. Disponibilizamos uma [lista de base de dados \(https://pucminas.instructure.com/courses/1754/files/1367675/download?wrap=1\)](https://pucminas.instructure.com/courses/1754/files/1367675/download?wrap=1) ↓ https://pucminas.instructure.com/courses/1754/files/1367675/download?download_frd=1) que podem ser consultada pelos alunos como ideias de possíveis trabalhos.

IMPORTANTE: Caso o aluno utilize uma base de dados proprietária de alguma empresa, ressaltamos a necessidade de autorização específica para a utilização dos dados, bem como sugerimos que os dados sejam alterados para preservar a privacidade das pessoas que possam estar representadas nos dados.

O aluno deverá desenvolver uma solução de análise de dados composta pelos seguintes itens:

- **Tratamento e cargas de dados:** conjunto de itens que inclui a documentação da estrutura das fontes de dados utilizadas pelo projeto, bem como todos os artefatos utilizados para o tratamento e carga dos dados (ETL) no ambiente utilizado pelas ferramentas de análise.
- **Estrutura de dados dimensional:** modelo de dados dimensional ou estrutura resultante da etapa inicial de tratamento e cargas de dados que defina um conjunto de fatos e dimensões.
- **Dashboard, relatórios e visualizações:** ambiente integrado do usuário que apresente as análises descritivas, preditivas ou prescritivas criadas com a respectiva documentação de cada uma das saídas do projeto que são utilizadas pelos usuários alvo da solução.

Para a avaliação de assimilação dos conhecimentos passados nas disciplinas do curso, foi formulado um conjunto obrigatório de requisitos e restrições técnicas que a solução a ser desenvolvida deve apresentar. Assim sendo, embora seja possível escolher o tema, objeto do trabalho, é necessário cumprir os itens que são apresentados a seguir.

2.1 - Requisitos Funcionais

Esta seção relaciona um conjunto genérico de funcionalidades que a sua aplicação deve apresentar. Cabe ao aluno escolher em quais funcionalidades irá incorporar os requisitos apresentados na relação a seguir. São eles:

ID	Requisito Funcional	Descrição
RF-01	Integração de tratamento e carga de dados	A solução deve apresentar um mecanismo automatizado (processos ETL) para o processamento das fontes de dados utilizadas de tal forma que, caso os dados sejam atualizados e a estrutura seja mantida, não seja necessário o desenvolvimento de procedimento adicional para visualizar as alterações dos dados
RF-02	Painel / Dashboard	O painel/dashboard deve apresentar pelo menos três abas integradas e relacionadas entre si, conforme estrutura descrita a seguir: <ul style="list-style-type: none"> • Dados estratégicos - apresenta indicadores (pelo menos 10) para o problema escolhido agrupados via soma, médias, etc; • Dados táticos - apresenta o acompanhamento histórico das informações por meio de gráficos (pelo menos 8) de pizza, linha, barras, dispersão, etc para representação de dados mensalmente; • Dados operacionais - apresenta tabelas (pelo menos 3) com o detalhamento de dados.
RF-03	Tipos diferenciados de visualizações de dados	O ambiente de análise dos dados deve fornecer pelo menos 10 visualizações de tipos diferentes. Deve prover cenários com combinações diferenciadas de variáveis: uma variável numérica, uma variável categórica, duas variáveis numéricas, combinações de variáveis numéricas e categóricas.
RF-04	Técnica de	Deve ser aplicada algumas das técnicas de aprendizado de máquina

	aprendizado de máquina	na solução proposta (regras de associação, clusterização, classificação etc.) .
RF-05	Filtros de dados	A solução deve apresentar filtros baseados nas dimensões a serem aplicados nas diversas abas do painel/dashboard de forma integrada refletindo nas diversas visualizações associadas.

2.2 - Requisitos não funcionais

Além das funcionalidades apresentadas acima, o sistema deverá apresentar as seguintes características:

ID	Requisito Não Funcional	Descrição
RNF-01	Desempenho da solução	A interface da solução deve apresentar os resultados esperados em cada uma das telas em no máximo 10 segundos, sendo recomendado que os resultados sejam exibidos em tempo menor (ideal: 3 segundos). Desta forma, espera-se que as bases de dados utilizadas em consultas sejam consolidadas para permitir o rápido acesso às informações disponíveis.
RNF-02	Documentação das análises	Todas as métricas devem ser documentadas no próprio painel, informando o cálculo que foi realizado para apresentação do cenário.

2.3 - Restrições Técnicas

As restrições técnicas limitam o que poderá ser empregado no desenvolvimento da solução. São elas:

ID	Restrição	Descrição
RT-01	Estrutura da base de dados	A base de dados escolhida para o projeto deve apresentar uma estrutura que permita extrair no mínimo tabelas de fato com 10 dimensões e dados de 24 meses de histórico.

3 - Entrega

Neste trabalho, a entrega se divide em três partes: (1) **relatório técnico**, (2) **artefatos da aplicação**, e (3) **vídeo demonstrativo**. Os detalhes de cada um destas partes são apresentados a seguir. Na tarefa associada à entrega do trabalho, o aluno deve submeter apenas a documentação do projeto que indicará a localização das outras partes em servidores na Internet.

3.1 - Relatório Técnico

Os itens relacionados com o detalhamento do problema a ser resolvido, projeto e modelagem da solução, arquitetura da solução e testes devem ser descritos e formatados em um único documento no formato PDF, o que constitui a documentação do Projeto. No ambiente virtual de aprendizagem você tem acesso ao [modelo de relatório técnico](https://pucminas.instructure.com/courses/1754/files/1377024/download?wrap=1)

(<https://pucminas.instructure.com/courses/1754/files/1377024/download?wrap=1>)_ ↓

(https://pucminas.instructure.com/courses/1754/files/1377024/download?download_frd=1) , sugerido para esta parte do trabalho. Neste documento devem constar os seguintes itens:

1. **Introdução** - visão geral sobre o problema a ser tratado e a justificativa para a escolha da base do projeto;
2. **Modelo de Dados** - modelo dimensional com a devida descrição das tabelas de fato e dimensão definidas para o projeto;
3. **Processo ETL** - descrição dos processos de transformação e carga de dados com um print dos fluxos criados;
4. **Camada de Apresentação** - descrição das métricas definidas (conceitos e regras de cálculo), das dimensões utilizadas e dos gráficos escolhidos para a composição do dashboard;
5. **Registros de homologação**: testes da solução desenvolvida mostrando que o dado apresentado no dashboard é o mesmo dos sistemas fonte;
6. **Conclusões**: achados mais relevantes nos dados feitos a partir do uso do dashboard e as experiências adquiridas no processo de desenvolvimento.

3.2 - Artefatos da Aplicação

Os códigos fonte/artefatos criados para o projeto devem ser disponibilizado para acesso por parte dos avaliadores. Estes arquivos podem ser disponibilizados em um servidor de armazenamento na Internet como GitHub, GitLab, BitBucket ou ainda repositórios mais genéricos como o DropBox, Google Drive, OneDrive, entre outros. O aluno deve garantir que o acesso ao link fornecido ficará ativo por todo o período de avaliação do TCC.

Importante: O link para o código fonte deve ser incluído no arquivo da Documentação do Projeto.

3.3 - Vídeo demonstrativo

Um vídeo com a apresentação sucinta do painel e seus recursos principais, limitado a 5 minutos. Sugere-se que o aluno apresente um ciclo completo de consulta aos dados pelas diversas visualizações do dashboard. O software OBS Studio é uma boa sugestão de software para gravação de screencast.

Importante: O link para o vídeo deve ser incluído no arquivo da Documentação do Projeto.

4 - Planejamento do TCC

O [cronograma do TCC \(https://pucminas.instructure.com/courses/1754/pages/cronograma-tcc\)](https://pucminas.instructure.com/courses/1754/pages/cronograma-tcc) apresenta todas as datas possíveis para a participação dos alunos do curso nesta disciplina. No processo de desenvolvimento deste trabalho, você deve escolher uma das quatro turmas, cujas datas são apresentadas no cronograma.

Uma vez escolhida a turma, o trabalho é dividido conforme o plano de etapas que é descrito a seguir. As etapas são:

1. **Desenvolvimento do trabalho** - etapa em que o aluno define o problema a ser tratado e desenvolve a solução.
2. **Entrega** - o aluno submete o trabalho finalizado na tarefa aberta no ambiente virtual de aprendizagem.
3. **Avaliação Preliminar** - Os avaliadores verificam se o aluno está apto ou não para participar da etapa de banca.
4. **Agendamento de Bancas** - O aluno é notificado do resultado da avaliação preliminar e é feito um trabalho de marcação das bancas.
5. **Banca Avaliadora** - O aluno apresenta seu trabalho para uma banca examinadora que fará as críticas e recomendações ao trabalho, aplicando uma nota final para o aluno na disciplina de TCC. Os avaliadores repassam a nota para a secretaria que repassará para o aluno.