

# TRABALHO DE CONCLUSÃO DE CURSO - TURMA 2022

## ORIENTAÇÕES

---

Prezados alunos,

Nesse documento estão apresentadas as informações necessárias para o desenvolvimento do seu Trabalho de Conclusão de Curso. O TCC nos cursos de especialização da PUC Minas Virtual é um trabalho interdisciplinar. Nosso propósito é consolidar os conhecimentos aprendidos no curso, dando ao aluno mais uma oportunidade de colocá-los em prática em um contexto de trabalho.

**Nesse TCC você terá duas opções.** Você poderá escolher entre duas opções: (1) escrever um trabalho teórico (com a entrega de um artigo de no máximo 20 páginas) ou (2) desenvolver um projeto de Engenharia de Dados utilizando os conceitos, técnicas e ferramentas estudadas durante o curso.

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, mas que seja um problema REAL e RELEVANTE para sua organização ou para a sociedade. 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 “**O que deve ser entregue**”? que fica ao final deste documento.

O trabalho deve ser feito individualmente.

Em nosso modelo de especialização, não há a orientação acadêmica/metodológica através de um professor orientador. Com base no formato escolhido, artigo ou projeto, você irá elaborar o texto do seu trabalho. Em caso de dúvidas sobre as regras ou procedimentos do TCC, o aluno deve enviar uma mensagem no Fórum de Discussões da disciplina. Desta forma, espera-se que este espaço sirva de orientação para todos os alunos, uma vez que todos terão acesso às perguntas realizadas pelos colegas e os esclarecimentos sobre elas.

## OPÇÃO 1: ARTIGO

---

Caso você faça a opção por escrever um trabalho teórico, você deverá submeter seu artigo no formato da Revista Abakós, do Instituto de Ciências Exatas e Informática. Os templates para WORD ou LATEX estão disponíveis [aqui](#). Seu artigo deve ter no máximo 20 páginas e tem tema aberto. O trabalho teórico deve ter qualidade científica e apresentar uma visão diferenciada sobre algum tema, lançando novas luzes a respeito. Isso vai exigir grande profundidade e poder de questionamento. Você pode falar sobre novas tecnologias, aplicações, soluções de problemas em Engenharia de Dados; também pode fazer um estudo comparativo entre métodos, técnicas ou padrões, de caráter teórico e/ou prático, adotando como referência um benchmark. Deve ser fortemente apoiado em referências.

## OPÇÃO 2: PROJETO PRÁTICO

---

A seguir estão as etapas que você deve seguir caso tenha optado por fazer um projeto prático. Nesse caso você deverá entregar um relatório técnico detalhando todas as etapas do projeto.

### 1. Definição do Problema

No seu TCC, primeiramente você deverá escolher um **PROBLEMA REAL** de seu interesse em qualquer área: transporte, economia, consumo, educação, saúde, etc. Você tem liberdade para escolher o tema que achar mais interessante, mas é importante que esse seja um problema **RELEVANTE** para sua organização ou para a sociedade. É muito importante que nessa primeira etapa você defina muito bem a pergunta que o seu trabalho vai resolver usando as técnicas e ferramentas da Engenharia de Dados.

### 2. Modelagem conceitual e definição das Tecnologias/Ferramentas/Arquitetura

A partir da clara definição do seu problema, você deve agora buscar as fontes de dados para o seu projeto, fazer a modelagem de dados, identificar as tecnologias/ferramentas que serão utilizadas em cada etapa (preparação e ingestão, orquestração, processamento, e visualização).

Como sugestão para obtenção de *datasets*, você pode buscar em sites como o Kaggle, o famoso repositório de dados da Universidade da Califórnia em Irvine ([UCI](https://www.uci.edu/)), os dados abertos do governo federal ([dados.gov.br](https://dados.gov.br)) ou de governos de outros países (EUA – [data.gov](https://data.gov), Reino Unido – [data.gov.uk](https://data.gov.uk)). O artigo “[Best Public Datasets for Machine Learning and Data Science](#)” dá excelentes dicas de repositórios de dados. Sugerimos o uso de diferentes formatos para que você possa colocar em prática tudo o que aprendeu: formato estruturado, não-estruturado e semi-estruturado).

*Observação: caso você queira utilizar dados da empresa onde você trabalha, sugerimos que você remova ou camufle qualquer dado que identifique pessoas ou organizações.*

### 3. Ingestão de Dados

Após a obtenção dos seus dados, é momento de elaborar rotinas para a ingestão de dados (preparação, ETL, workflows, carga e armazenamento, streaming). A sua 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.

### 4. Orquestração de Dados

Nessa etapa você vai elaborar rotinas para a orquestração de dados. Deixe bem claro em seu documento quais foram as estratégias utilizadas.

## 5. Visualização dos dados

Nessa última etapa você deve elaborar rotinas para a visualização dos dados através de painéis de indicadores, dashboards, etc.

## ENTREGA DO TRABALHO

---

Para realizar a entrega do trabalho, você deve postar o seu TCC conforme o template disponibilizado, em formato PDF. Os links para o vídeo no Youtube e para o repositório contendo dados e demais arquivos (scripts, etc) devem estar contidos nesse documento.

Após avaliação das entregas postadas no Canvas, os professores da banca irão indicar os alunos aptos para a apresentação final do TCC.

## O QUE DEVE SER ENTREGUE?

---

- Artigo ou relatório técnico conforme template disponibilizado (em formato PDF). Lembre-se de detalhar o máximo possível. Informe as ferramentas utilizadas, onde e quando coletou os dados, etc...
- **Link para vídeo no Youtube.** Esse vídeo terá tempo máximo de 5 minutos e deverá apresentar de forma sucinta o seu projeto, desde a definição do problema, a obtenção e tratamento dos dados, até a apresentação dos resultados.

## APRESENTAÇÃO DO TRABALHO PARA A BANCA

---

- Você deve preparar um conjunto de slides no PowerPoint (ou algum software equivalente) para apresentar seu trabalho para a banca de professores. Geralmente o tempo de apresentação é de 15 a 20 minutos.

## DÚVIDAS?

---

Nosso objetivo foi disponibilizar todos os materiais necessários para a execução do trabalho. Entretanto, entendemos que dúvidas podem surgir. Neste caso, mande uma mensagem para a gente no fórum de discussão. Bom trabalho!