

Regulamento do Projeto Integrado

1. Introdução

Este documento apresenta as informações necessárias sobre o desenvolvimento do Projeto Integrado para o curso de Pós-graduação *Lato Sensu* em **Analytics e Business Intelligence**.

O Projeto Integrado se constitui de um trabalho interdisciplinar realizado individualmente em que o aluno coloca em prática tudo o que foi visto nas disciplinas do curso. 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 de mercado por meio do desenvolvimento de uma solução completa. O problema a ser abordado no Projeto Integrado é escolhido pelo aluno e, dessa forma, espera-se que os conhecimentos possam ser aplicados em um projeto alinhado com os interesses do aluno.

A disciplina é organizada por **Professores Coordenadores** do Projeto Integrado. Dúvidas administrativas, dúvidas sobre o ambiente ou dúvidas sobre os procedimentos do trabalho devem ser encaminhadas à coordenação por meio do **Fórum da Coordenação**.

Durante todo o trabalho, o aluno será acompanhado por um **Professor Mentor**. Todas as dúvidas sobre a escolha do tema ou da base de dados a ser utilizada, ou ainda, dúvidas sobre os critérios de aceitação apresentados no regulamento ou sobre o processo de entrega e avaliação devem ser encaminhadas por meio do **Fórum de Mentoria**. Esse acompanhamento é feito de forma coletiva com o grupo alocado ao Professor Mentor. Desta forma, espera-se que este espaço sirva de troca de informações para todos os alunos uma vez que todos podem ver as perguntas realizadas pelos demais alunos e os esclarecimentos sobre elas.

A estrutura do trabalho e os itens que devem ser produzidos são descritos no item **2. Escopo do Trabalho**. As etapas de entrega do trabalho e suas regras são descritas no item **3. Entrega**. Os detalhes do processo de avaliação, bem como os critérios de aceitação das entregas, estão descritos no item **4. Avaliação**.

2. Escopo do Trabalho

O Projeto Integrado envolve o desenvolvimento de uma solução de Analytics e Business Intelligence para um cenário organizacional à escolha do aluno. Para esse cenário, o aluno deverá identificar os objetivos que motivam o desenvolvimento desse projeto e se atentar para a disponibilidade das fontes de dados que deem condições para a realização de um trabalho como esse.

Com o objetivo de fornecer ideias para os projetos, disponibilizamos uma relação de bases de dados públicas. Caso o aluno utilize uma base de dados proprietária de alguma empresa,

ressaltamos a necessidade de autorização específica para a utilização dos dados. Sugerimos, ainda, que os dados sejam alterados para preservar a privacidade das pessoas que possam estar representadas nos dados.

O trabalho é composto pela sequência de 3 módulos que se complementam de forma evolutiva (A, B e C) e o conteúdo de cada um dos módulos é apresentado a seguir.

2.1 - Módulo A - Discovery e Projeto da Solução

Nesse módulo, será realizado o mapeamento do espaço problema e dos objetivos associados a este contexto na perspectiva dos principais interessados (stakeholders). O aluno deve detalhar este contexto e, em seguida, deve criar os modelos de dados e implementar os processos de integração, tratamento e carga dos dados que vão suportar a solução.

Os itens constantes do módulo A são:

- **Contexto do projeto** - apresentação da visão geral do contexto organizacional abordado, identificando a motivação para realização do projeto, os objetivos estratégicos associados, os indicadores para o acompanhamento destes objetivos, os principais interessados (stakeholders), a relação das fontes de dados a serem utilizadas, os modelos de dados a serem gerados pelo projeto, as integrações de dados entre as diversas fontes e outros itens relevantes para o trabalho. Estas informações podem ser detalhadas textualmente ou via algum modelo de Canvas voltado projetos de dados;
- **Modelos de dados**
 - **Fontes de dados** - diagramas dos modelos relacionais de bancos de dados ou estruturas dos arquivos utilizados pelo projeto como fontes;
 - **Base Dimensional** - diagramas dos modelos dimensionais com a devida descrição das tabelas de fato e dimensões definidas para o projeto;
- **Processos de Integração, Tratamento e Carga de Dados**
 - **Ingestão de Dados e Processos ETL** - descrição dos processos de preparação, tratamento, transformação e carga de dados com imagens dos fluxos criados;
 - **Códigos Fonte** ([Link para repositório externo](#)) - Os códigos fonte ou os artefatos criados para o projeto devem ser disponibilizados para acesso por parte dos avaliadores. Estes arquivos podem ser disponibilizados em um servidor externo de armazenamento na Internet como GitHub, GitLab, BitBucket ou ainda repositórios mais genéricos como o DropBox, Google Drive, OneDrive, entre outros.

2.2 - Módulo B - Painel de Controle (Dashboard)

Nesse módulo, será desenvolvido um dashboard com diversas análises distribuídas em três perspectivas distintas, porém interligadas. A primeira perspectiva tem foco na gestão em nível estratégico que acompanha as informações de maneira resumida, mas utilizando indicadores de desempenho que atestam a saúde da organização como um todo. Na segunda perspectiva, o foco é na gestão em nível tático que avalia a evolução de contextos distintos da organização

identificando possíveis melhorias dos processos. Na terceira perspectiva, o foco está na gestão operacional de processos específicos e que entram num nível mais granular das informações.

Os itens constantes do módulo B são:

- **Análises de Dados** - Detalhamento dos painéis criados com as perspectivas estratégica, tática e operacional, as análises que os compõem, as métricas definidas (conceitos e regras de cálculo), as dimensões utilizadas e dos gráficos escolhidos para a composição do dashboard;
- **Dashboard** - painéis de controle criados por meio de uma ferramenta de Business Intelligence e Analytics com as análises nas perspectivas:
 - Painel Estratégico - indicadores diversos, KPIs e principais visualizações dados consolidados para o contexto analisado;
 - Painel Tático - acompanhamento histórico e segmentado das informações apresentadas no Painel Estratégico;
 - Painel Operacional - detalhamento dos dados apresentados nos painéis anteriores.
- **Registros de homologação**: testes da solução desenvolvida mostrando que o dado apresentado no dashboard é o mesmo dos sistemas fonte;
- **Códigos Fonte** ([Link para repositório externo](#)) - Os códigos fonte/artefatos criados para o projeto devem ser disponibilizados 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.
- **Vídeo de Apresentação - Módulo B** ([Link para repositório externo](#)) - 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*.

2.3 - Módulo C - Análises Avançadas e Proposta de Intervenção

Nesse módulo vamos tratar de duas entregas muito relevantes para os gestores que demandam soluções de análise de dados. Inicialmente, vamos abordar as análises avançadas a partir da base de dados estruturada nos módulos anteriores identificando novos conhecimentos no contexto do projeto. Posteriormente, vamos fazer uma análise de tudo o que foi obtido com o projeto para indicar as possíveis intervenções que podem alavancar diferenciais competitivos para a organização.

Os itens constantes do módulo C são:

- **Análises avançadas**
 - **Padrões e/ou Tendências**: análises de dados criadas a partir da aplicação de técnicas de aprendizado de máquina. As análises podem ser independentes ou incorporadas no dashboard criado anteriormente.

• Conclusões

- **Análise crítica:** texto detalhando os principais achados a partir das análises criadas e as experiências adquiridas no processo de desenvolvimento.
- **Proposta de Intervenção:** texto descrevendo possíveis ações a serem tomadas por gestores do contexto analisado no intuito de melhorar o desempenho da organização.
- **Lições aprendidas:** texto descrevendo os aprendizados do percurso e que podem ser empregados em outros projetos.
- **Códigos Fonte** ([Link para repositório externo](#)) - Os códigos fonte/artefatos criados para o projeto devem ser disponibilizados 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.
- **Vídeo de Apresentação - Módulo C** ([Link para repositório externo](#)) - Um vídeo com a apresentação sucinta das análises avançadas e com a análise crítica dos principais achados no projeto e a proposta de intervenção.

3. Entrega

Este item trata do processo de entrega dos resultados do projeto. Conforme apresentado anteriormente, o trabalho é dividido em três módulos distintos. As regras e o cronograma de entrega devem ser observados rigorosamente pelos alunos.

3.1. Regras das Entregas

Os professores mentores e, em especial, os alunos da disciplina devem observar as seguintes regras sobre a entrega dos artefatos:

1) Arquivo do Relatório Técnico a ser entregue no AVA (Canvas)

Todas as entregas aceitam apenas um arquivo no formato PDF que se constitui do Relatório Técnico do projeto. Neste documento, são descritos ou relacionados todos os demais itens que fazem parte do trabalho. Para itens externos (ambientes hospedados na Internet, vídeos de apresentação, códigos fonte, entre outros), deve ser apresentado no referido documento os links de acesso.

Importante: O aluno deve certificar que o professor mentor terá acesso ao item externo com as informações fornecidas no referido documento. Tais itens adicionais devem permanecer disponíveis durante o intervalo de tempo definido no **Cronograma da Disciplina** para que possam ser avaliados pelos envolvidos no processo.

2) Participação nos módulos/entregas

O aluno deve entregar todos os módulos para ser aprovado na disciplina. Uma vez que um artefato depende do seguinte, a não entrega de qualquer um dos módulos se constitui na

reprovação do aluno, ainda que no somatório das notas, o aluno consiga obter o mínimo estabelecido pela PUC Minas.

3) Locais e datas de entrega

O ambiente disponibiliza tarefas específicas para cada um dos módulos com datas distintas com o objetivo de receber as entregas dos alunos. Os resultados do módulo A somente poderão ser entregues nas tarefas relacionadas ao módulo A. O aluno pode escolher as datas de entrega de cada módulo conforme o progresso do seu curso. O aluno pode fazer a entrega os três módulos em uma única data, ou em datas subsequentes, obedecendo a ordem, A, B e C. Independentemente da sua escolha, cada módulo deve ser entregue separadamente na tarefa específica.

3.2. Cronograma da Disciplina

Como temos alunos entrando no curso em momentos distintos do ano, o cronograma apresenta diversas opções de datas para a entrega de cada módulo. De acordo com o planejamento de cada aluno, é possível escolher a data mais conveniente para realizar a entrega dos módulos.

O cronograma da disciplina é apresentado na Figura a seguir.

4. Avaliação

Nesta seção, são abordados o processo de avaliação do Projeto Integrado e os critérios de aceitação para o trabalho como um todo. É fundamental que o aluno observe estes critérios para planejamento do seu projeto.

4.1 - Processo de Avaliação e Pontuação dos Módulos

Os pontos da disciplina são divididos da seguinte forma:

- Módulo A - 30 pontos,
- Módulo B - 30 pontos,
- Módulo C - 40 pontos.

A solução desenvolvida pelo aluno nesta disciplina é avaliada pelo Professor Mentor. Somente após a data de entrega e mediante a entrega definitiva do módulo, o professor mentor fará a avaliação e lançará a nota e as considerações para que o aluno possa verificar o que está conforme ao regulamento e o que precisa de melhorias. Não são realizadas revisões intermediárias antes da entrega.

O Professor Mentor pode, na sua avaliação, rejeitar a entrega do aluno, caso a entrega não cumpra pelo menos 50% do que foi solicitado para cada módulo. **O aluno poderá ter sua entrega rejeitada pelo professor apenas uma vez. Na segunda entrega rejeitada, o aluno estará reprovado na disciplina.** Nestes casos, o aluno pode contratar um processo de tutoria a parte.

Nesse processo não haverá bancas de apresentação do trabalho e a nota final será a soma das notas obtidas em cada uma das três etapas.

Em situações excepcionais, poderá haver a participação da Coordenação do Projeto Integrado para dirimir sobre dúvidas relacionadas ao processo de avaliação. Nestes casos, a Coordenação do Projeto Integrado poderá substituir o Professor Mentor no processo de avaliação, quando solicitado pelo mesmo ou por motivos de força maior.

4.2 - Critérios de Aceitação

Para orientar a todos os envolvidos no Projeto Integrado quanto ao que precisa ser entregue e forma de avaliação do projeto a ser desenvolvido, foi definido um conjunto de critérios de aceitação.

Estes critérios trazem características funcionais e não funcionais que devem ser observadas pelo aluno na hora de definir e desenvolver seu projeto. Os critérios de aceitação são apresentados na tabela a seguir.

ID	Critério	Descrição
CA-01	Integração, tratamento e carga de dados (ETL)	A solução deve apresentar um mecanismo automatizado 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 ou manutenção dos processos existentes para visualizar as alterações dos dados
CA-02	Painel de Controle (Dashboard)	O dashboard deve apresentar pelo menos três abas integradas e relacionadas entre si , com perspectivas estratégica, tática e operacional.

CA-03	Tipos de visualizações de dados	<p>A solução desenvolvida deve fornecer visualização dos dados e informações contemplando no mínimo os seguintes itens:</p> <ul style="list-style-type: none">• pelo menos 15 visualizações de tipos diferentes (tabelas com dados detalhados, mapas, gráficos diversos - linhas, barras, dispersão, gauges, etc);• cenários com combinações diferenciadas: variável numérica, uma variável categórica, duas variáveis numéricas, combinações de variáveis numéricas e categóricas.
CA-04	Disponibilização de filtros de dados	<p>A solução deve apresentar filtros baseados nas dimensões a serem aplicados nas diversas abas do dashboard de forma integrada refletindo nas diversas visualizações associadas.</p>
CA-05	Aplicação de técnica de aprendizado de máquina	<p>No módulo de análises avançadas, deve ser aplicada alguma técnica de aprendizado de máquina na solução proposta (regras de associação, clusterização, classificação, etc).</p>
CA-06	Desempenho da solução	<p>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).</p> <p>Desta forma, espera-se que as bases de dados utilizadas em consultas sejam consolidadas para permitir o rápido acesso às informações disponíveis.</p>
CA-07	Documentação das análises	<p>Todas as métricas apresentadas devem ser documentadas no próprio painel, informando o cálculo que foi realizado para apresentação do cenário. A documentação pode vir no formato de labels, balões de dicas exibidos ao passar com o mouse ou uma aba específica para detalhamento e orientações sobre as métricas.</p>
CA-08	Estrutura e volume da base de dados	<p>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.</p>

