

# Checklist do Projeto de Extensão

## *Tópicos de Big Data em Python*

Professor(a): \_\_\_\_\_

Nome da Equipe: \_\_\_\_\_

Projeto/Tema: \_\_\_\_\_

Parceiro Externo (pessoa/instituição): \_\_\_\_\_

Contato do Parceiro (e-mail/telefone): \_\_\_\_\_

Líder da Equipe: \_\_\_\_\_

Integrantes: \_\_\_\_\_

\_\_\_\_\_

### 1) Diagnóstico e Teorização (2,0 pts)

**Parte interessada definida** (quem é, por que precisa, contexto do problema).

**Problema real descrito** (5–8 linhas): dor, impacto atual, onde ocorre.

**Hipótese de valor:** o que pretendemos melhorar/decidir com dados.

**Levantamento de dados:** fontes disponíveis/possíveis, acesso, formato (CSV, API, planilha), periodicidade, volume.

**Questões éticas/LGPD:** consentimento, anonimização, risco de reidentificação, retenção e segurança.

**KPIs de sucesso** combinados com o parceiro (como mediremos o valor/impacto).

### 2) Planejamento e Desenvolvimento (4,5 pts)

**Escopo claro:** o que será entregue (*script/relatório/dashboard/pipeline simples*) e o que não entra.

**Esboço de arquitetura:** captura/extração → limpeza/tratamento → armazenamento → análise/modelagem → visualização/entrega.

**Tecnologias:** Python (pandas/*PySpark* quando aplicável), notebooks, armazenamento (Drive/S3/DB), libs de visualização. Justificativa breve.

**Plano de dados:** dicionário de dados, faltantes/outliers, amostragem, padronização de tipos e datas.

**Experimentos:** EDA, testes de modelos (se houver), critérios de validação e de comparação.

**Cronograma até o fim do semestre:** marcos e entregas parciais (datas).

**Papéis e responsabilidades:** owner do dado, ETL, analista/modelagem, documentação, relação com parceiro.

**Riscos e mitigação:** acesso aos dados, dependências externas, prazos; plano B.

**Evidências a anexar no Roteiro:** prints do ambiente, links reprodutíveis, amostras anonimizadas, gráficos, versões de código.

**Cadência com o parceiro:** checkpoints (semanal/quinzenal), forma de entrega (apresentação, link, manual).

### 3) Encerramento e Relatos (3,5 pts: 2,5 coletivo + 1,0 individual)

**Produto final utilizável** pelo parceiro (script, dashboard, relatório ou pipeline) + manual rápido de uso.

**Avaliação de impacto:** KPIs antes vs. depois / cenário simulado; limites do estudo e próximos passos.

**Relato coletivo do grupo:** problema, método, arquitetura, resultados, impacto para a comunidade, aprendizados.

**Relato individual (obrigatório):** contribuição pessoal, dificuldades, competências desenvolvidas.

**Entrega do Roteiro de Extensão** (com prazos definidos na disciplina).

### Blocos de Trabalho em Sala (sugestão)

**Roda 1: 20–30 min** — Fechar *Diagnóstico*: problema, dados, KPIs, ética, papéis.

**Roda 2: 30–40 min** — Esboçar *Arquitetura + Cronograma + Riscos*.

**Roda 3: 20–30 min** — Planejar primeiras *análises/experimentos* e evidências para o Roteiro.

*Saída da aula:* documento do grupo atualizado com decisões e próximos passos (**quem faz o quê** até a próxima aula).