# Checklist – PROJETO DE ANÁLISE DE DADOS

Nome do Projeto: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Criar uma pasta no computador ou na nuvem;

Criar um ambiente virtual (Virtualenv), caso seja, um projeto que envolva Python

Criar um arquivo requeriments.txt, caso seja, um projeto que envolva Python

Criar um repositório GitHub e inicializa-lo

Criação do arquivo LICENSE (MIT 2.0)

Criação do arquivo README.md

Criação da pasta VENV (Com o arquivo GITIGNORE)

Criar um arquivo dos tópicos para resolver o problema de negócio (Python: Usar pseudocódigos)

Criar um arquivo com as definições do planejamento do projeto (Python: planejamento.txt)

Pré-requisitos:

Quais serão as fontes de dados para encontrar as respostas?

Quais serão as ferramentas utilizadas para encontrar as respostas?

Como será o Processo? (Escolher quais serão os passos para encontrar as respostas: Como serão feitas as próximas etapas)

Qual será o Produto Final? (E-mail, Planilha, Gráfico / Mapas, Dashboards, etc.)

Criar as subpastas que serão utilizadas no projeto:

Pasta Dados (DATASETS): Externos, Provisórios, Processados e Brutos

Pasta Relatórios (REPORTS): Gráficos, Imagens, Pasta Páginas em HTML e Pasta Estático

Pasta Referência ou Documentos (REFERENCES ou DOCS)

Pasta Notebooks (SRC), caso seja, um projeto que envolva Python

Criação do arquivo init.py

Refatore: Criar um utilitario.py (utility.py)

Pasta Diagnóstico (DIAGNOSTICS)

Notabook\_01\_fgc\_DataCollect (Etapa de coleta de dados)

Notabook\_02\_fgc\_DataCleaning (Etapa de limpeza de dados)

Notebook\_03\_fgc\_DataProcessing (Etapa de processamento de dados)

Pasta Recursos (FEATURES)

Criação do arquivo build\_features.py

Notebook\_04\_fgc\_DataPreparationModeling (Etapa de preparação dos dados - modelagem)

Pasta Modelos (MODELS)

Criação do arquivo predict\_model.py

Criação do arquivo train\_model.py

Pasta Resultados (RESULTS ou LOGS)

Notebook\_05\_fgc\_ResultsReadAndAnalysis (Etapa de análise e leitura de resultados)

Pasta Visualização (VISUALIZATION)

Notebook\_06\_fgc\_DataVisualization (Etapa de visualização de dados)

Pasta Testes (TESTS): Criação do arquivo teste.py

Pasta Configuração (CONF) (Arquivos de configuração do projeto)

Arquivo credenciais.py (Armazenar as senhas/chaves de acessos)

Arquivo setup.py (Torna o projeto PIP instalável com "pip install -e".)

Arquivo makefiles.txt (Contém instruções sobre como criar e vincular (ou compilar) um conjunto de arquivos de código-fonte)

OBS: Criar um projeto comentado - Foque em explicar "porque" tomou as decisões. É mais importante explicar seu raciocício do que ter um modelo com performance absurdamente boa.