

## Scrum Detalhado - SITI-Natal com Django, Python e BI

*Duração do Projeto: 27 de setembro a 21 de novembro (8 semanas, 8 sprints)*

*Objetivo Geral: Desenvolver o sistema web SITI-Natal, usando Python/Django, técnicas de webcrawling, e ferramentas de análise de dados, permitindo cruzamento de informações com o sistema tributário Directa para identificar divergências e subsidiar ações fiscais.*

### **Equipe:**

1. **Engenheiro e Cientista de Dados:** Modelagem de dados, processamento e análise com BI.
2. **Desenvolvedor Backend:** Desenvolvimento com Django, API e banco de dados.
3. **Desenvolvedor Frontend:** Interface web com visualização interativa.
4. **Especialista em Webcrawling:** Coleta automatizada de dados dos portais imobiliários.

### **Sprint 1 (27 Set - 3 Out)**

- **Objetivo:** Planejamento inicial e configuração de ambientes de desenvolvimento.
  - **Backend (Django):** Configurar o projeto Django, estrutura básica do modelo de dados, conexão com banco de dados.
  - **Frontend:** Começar o design do sistema (prototipagem de interfaces com ferramentas como Figma ou Adobe XD).
  - **Webcrawling:** Identificar portais imobiliários e coletar requisitos para o scraping (definir os campos como área, valor, localização etc.).
  - **Cientista de Dados:** Planejamento das etapas de tratamento e análise de dados, selecionar bibliotecas de BI (como Pandas, Matplotlib).
  - **Todos:** Reunião inicial de alinhamento com stakeholders, definição de metas e criação do backlog completo.

### **Sprint 2 (4 Out - 10 Out)**

- **Objetivo:** Implementar a primeira versão de coleta de dados.
  - **Webcrawling:** Desenvolver scripts de web scraping com bibliotecas como Scrapy ou BeautifulSoup, coletando dados preliminares.

- **Backend (Django):** Estruturar os modelos no banco de dados (PostgreSQL ou MySQL), incluindo tabelas para imóveis e coleta de dados.
- **Frontend:** Implementar as primeiras páginas, com layout básico e navegação inicial.
- **Cientista de Dados:** Definir como os dados coletados serão armazenados e processados no banco de dados, mapeando as relações.
- **Todos:** Revisar o backlog e ajustar conforme o progresso inicial.

### Sprint 3 (11 Out - 17 Out)

- **Objetivo:** Primeira integração completa do webcrawling com o backend.
  - **Webcrawling:** Melhorar os scripts de coleta de dados, automatizar a coleta diária e enviar dados diretamente para o banco via Django ORM.
  - **Backend (Django):** Criar APIs para inserção, consulta e atualização de dados imobiliários. Implementar sistema de autenticação (se necessário).
  - **Frontend:** Desenvolver o dashboard inicial com visualização básica dos dados coletados.
  - **Cientista de Dados:** Iniciar o tratamento dos dados, removendo duplicatas e verificando consistência. Criar rotinas de análise preliminar.
  - **Todos:** Reunião para avaliar os primeiros dados coletados e ajustá-los com base no feedback da SEMUT.

### Sprint 4 (18 Out - 24 Out)

- **Objetivo:** Processamento e enriquecimento de dados, e melhoria no front-end.
  - **Webcrawling:** Garantir que a coleta de dados cubra todos os portais imobiliários relevantes e que as informações estejam consistentes.
  - **Backend (Django):** Melhorar a performance da inserção de dados, ajustar queries e criar endpoints mais complexos (filtros por localização, preço).
  - **Frontend:** Adicionar filtros e visualizações dinâmicas no dashboard (filtrar por localização, valor, número de quartos etc.).
  - **Cientista de Dados:** Começar a integração dos dados com o sistema Directa, definir métricas para cruzamento de dados.
  - **Todos:** Avaliar a coleta e processamento de dados, ajustar scripts e pipeline conforme a necessidade.

## Sprint 5 (25 Out - 31 Out)

- **Objetivo:** Integração com ferramentas de BI e sistema Directa.
  - **Webcrawling:** Continuar aprimorando a coleta automática e tratar qualquer inconsistência.
  - **Backend (Django):** Estabelecer integração com APIs ou bancos de dados do sistema Directa, se aplicável.
  - **Frontend:** Melhorar a interface gráfica, adicionar gráficos e visualizações interativas (utilizando frameworks como Chart.js ou D3.js).
  - **Cientista de Dados:** Desenvolver dashboards interativos com ferramentas como Power BI ou Tableau (se integrados) ou desenvolver soluções customizadas com Python.
  - **Todos:** Revisar o progresso da integração com o Directa e ajustar os fluxos de trabalho.

## Sprint 6 (1 Nov - 7 Nov)

- **Objetivo:** Testes extensivos e otimizações.
  - **Todos:** Foco em testes de carga, performance, e usabilidade. Revisar todos os módulos para garantir eficiência.
  - **Webcrawling:** Testar o webcrawler em ambientes reais, simulando picos de uso e diferentes condições de rede.
  - **Backend:** Revisar a segurança, otimizar queries e APIs para garantir a resposta em tempo real.
  - **Frontend:** Testar a usabilidade do dashboard e corrigir problemas de layout e performance.
  - **Cientista de Dados:** Revisar os algoritmos de processamento e análise de dados, garantir precisão nas divergências detectadas.

## Sprint 7 (8 Nov - 14 Nov)

- **Objetivo:** Refino final e preparação para o deployment.
  - **Webcrawling:** Garantir a estabilidade e escalabilidade do sistema de coleta.
  - **Backend:** Realizar últimos ajustes de integração e performance.
  - **Frontend:** Implementar feedbacks finais de design e usabilidade.
  - **Cientista de Dados:** Validar os relatórios finais gerados pelo sistema, garantir a clareza e precisão das divergências detectadas.
  - **Todos:** Testes finais de pré-deployment.

## Sprint 8 (15 Nov - 21 Nov)

- **Objetivo:** Deployment, documentação e entrega.
  - **Deployment:** Subir o sistema para o ambiente de produção.
  - **Documentação:** Produzir manuais de uso e documentação técnica para suporte contínuo.
  - **Treinamento:** Realizar treinamentos com os analistas da SEMUT e da Procuradoria para uso das ferramentas de BI e dashboards.
  - **Entrega Final:** Sprint de revisão e aceitação da entrega.

## Reuniões do Scrum

- **Daily Standups:** Check-ins rápidos diários para verificar o progresso.
- **Sprint Planning:** Definir os objetivos e tarefas detalhadas no início de cada sprint.
- **Sprint Review:** Apresentar o trabalho concluído ao final de cada sprint.
- **Sprint Retrospective:** Discutir melhorias para os próximos ciclos.

## Tecnologias Utilizadas

- **Python/Django:** Backend e APIs.
- **Scrapy/BeautifulSoup:** Webcrawling.
- **PostgreSQL/MySQL:** Banco de dados.
- **Pandas, Matplotlib, Power BI:** Ferramentas de análise e visualização de dados.
- **HTML/CSS/JavaScript (Chart.js/D3.js):** Frontend para visualização.

Esse cronograma detalhado garante que todas as partes críticas do sistema serão desenvolvidas e integradas, com foco em entregas incrementais e feedback constante da equipe da SEMUT e Procuradoria.