

**Centro Universitário Estácio de Brasília
Taguatinga-Sul**

Análises de dados com Python aliada ao marketing comercial

**Guilherme Estevão, Gustavo Êrades, Marcos Vinicius, Matheus Oliveira,
Washington Carmo**

Professor: Luan Muruk Ferreira Tonaco

2024

Brasília/Taguatinga/DF

Sumário

1. DIAGNÓSTICO E TEORIZAÇÃO	0
1.1. Identificação das partes interessadas e parceiros	0
1.2. Problemática e/ou problemas identificados	0
1.3. Justificativa	0
1.4. Objetivos/resultados/efeitos a serem alcançados (em relação ao problema identificado e sob a perspectiva dos públicos envolvidos)	0
1.5. Referencial teórico (subsídio teórico para propositura de ações da extensão)	0
2. PLANEJAMENTO E DESENVOLVIMENTO DO PROJETO	0
2.1. Plano de trabalho (usando ferramenta acordada com o docente)	0
2.2. Descrição da forma de envolvimento do público participante na formulação do projeto, seu desenvolvimento e avaliação, bem como as estratégias utilizadas pelo grupo para mobilizá-los.	0
2.3. Grupo de trabalho (descrição da responsabilidade de cada membro)	0
2.4. Metas, critérios ou indicadores de avaliação do projeto	0
2.5. Recursos previstos	0
2.6. Detalhamento técnico do projeto	0
3. ENCERRAMENTO DO PROJETO	0
3.1. Relatório Coletivo (podendo ser oral e escrita ou apenas escrita)	0
3.2. Avaliação de reação da parte interessada	0
3.3. Relato de Experiência Individual	0
3.1. CONTEXTUALIZAÇÃO	0
3.2. METODOLOGIA	0
3.3. RESULTADOS E DISCUSSÃO:	0
3.4. REFLEXÃO APROFUNDADA	0
3.5. CONSIDERAÇÕES FINAIS	0

1. DIAGNÓSTICO E TEORIZAÇÃO

1.1. Identificação das partes interessadas e parceiros

São partes interessadas do projeto de extensão da matéria de Tópicos de Big Data em Python, ministrada pelo professor Luan Muruck Ferreira Tonaco, no Centro Universitário Estácio de Brasília. Dos alunos são:

- Guilherme Estevão: 20 anos, gênero masculino, estudante do ensino superior do curso de Análise e Desenvolvimento de Sistemas
- Gustavo Êrades: 20 anos, gênero masculino, estudante do ensino superior do curso de Análise e Desenvolvimento de Sistemas
- Marcos Vinicius: 20 anos, gênero masculino, estudante do ensino superior do curso de Análise e Desenvolvimento de Sistemas
- Matheus Oliveira: 19 anos, gênero masculino, estudante do ensino superior do curso de Análise e Desenvolvimento de Sistemas
- Washington Carmo: 23 anos, gênero masculino, estudante de Análise e Desenvolvimento de Sistemas

O cliente consiste no Centro Especializado Odontológico de Taguatinga (CEOT Alameda) representado por Dionatas. O cliente foi apresentado à equipe por um dos membros, Guilherme Estevão, na semana do dia 04/03/2024. A fim de firmar o comprometimento entre as partes interessadas bem como suas responsabilidades frente ao projeto, um Termo de Acordo e Cooperação foi elaborado e assinado entre um representante da equipe de trabalho (seu nome) e o representante do cliente (Dionatas). Segue link do drive com o [Termo de Acordo e Cooperação](#).

1.2. Problemática e/ou problemas identificados

O cliente, dono de uma clínica odontológica, informou que pretendia aumentar o alcance e acessibilidade ao seu negócio, portanto dentre suas necessidades/objetivos estavam o marketing digital para o negócio e uso de tráfego pago para a captação de mais clientes e aumento da visibilidade da sua clínica. Com isso em vista a equipe se concentrou em buscar maneiras de aliar a análise de dados com Python com esses objetivos.

A equipe optou por se concentrar na captação de clientes da maneira menos custosa possível, tanto em termos de tempo quanto em termos monetários. Por isso a equipe sugeriu a criação de um relatório que sugerisse expansão do número de clientes através da implementação de diversas estratégias, estratégias essas sugeridas com base na análise de dados coletados por meio de formulários do Google a respeito de satisfação da clientela e sugestões de melhorias no estabelecimento/atendimento e demais pesquisas avulsas referentes a captação de clientes.

Foi proposto ainda pela equipe o desenvolvimento de alguma plataforma de avaliação das práticas já adotadas de marketing na instituição parceira bem como de suas ações para captação de novos clientes, além da proposta do desenvolvimento de alguma plataforma de contato com o cliente, como um web site, que favorecesse a utilização de publicidade paga, melhorasse a identificação do cliente com o estabelecimento, automatizar processos (como o agendamento de consultas) e moderniza o atendimento visando maximizar o impacto e facilitar o alcance do público-alvo.

Todas essas propostas partiram de discussões da equipe na semana do dia 04/03/2024, no entanto nessas discussões também foram identificadas

problemáticas referentes às mesmas propostas que frearam o desenvolvimento do projeto e mostraram a necessidade de uma discussão/planejamento mais detalhados para o pleno desenvolvimento deste projeto de extensão. Dentre essas problemáticas continham o tempo em que a equipe disponha para o projeto, a profundidade de conhecimento necessária para a elaboração de alguns processos, a necessidade de aquisição ou geração de custos adicionais ao cliente e/ou a equipe e, principalmente, a obrigação de gerar um resultado satisfatório, confiável e utilizável para os devidos fins do cliente dentro do prazo acordado (até 14/06/2024).

A equipe detinha de aproximadamente três meses para a discussão, planejamento, desenvolvimento, avaliação e entrega do projeto e, portanto, deveria pensar/desenvolver uma solução para captação de clientes dentro desse prazo. Além disso não parecia viável a equipe, e iria contra a proposta de projeto extensionista do Centro Universitário Estácio, a geração de custos adicionais para o cliente; nenhum membro da equipe detinha experiência com análise ou captação de dados e nem com o uso da linguagem de programação Python para tal fim; levando esses quesitos em consideração os membros da equipe descartaram algumas ideias e optaram, inicialmente, apenas pelo desenvolvimento de formulários de pesquisa de avaliação/satisfação da clientela e análise das práticas de captação de clientes adotadas atualmente e anteriormente pelo cliente afim de fazer uma comparação e dizer se as práticas de marketing anteriormente aplicadas valeram a pena e em que se podia melhorar para melhorar a satisfação dos clientes e garantir a expansão do número da clientela.

1.3. Justificativa

A equipe optou por focar na captação de novos clientes com o uso da análise de dados provindos de pesquisas de formulários e comparações com práticas já adotadas pelo cliente para o mesmo fim usando, como ferramenta para melhor captação de dados e automatização de processos, a linguagem de programação Python. O projeto extensionista proposto corrobora com o conteúdo proposto pela matéria de “Tópicos de Big Data com Python” uma vez que prevê a análise de diversos e diferentes dados provindos de diferentes bases (formulários do Google, avaliações de clientes em redes sociais, cliques no site da empresa em determinado período, aumento/diminuição do tráfego de clientes na clínica etc.) utilizando a linguagem Python para tal fim bem como o uso de outras ferramentas, bibliotecas e conceitos relacionados à Big Data. Além disso, tal projeto também corrobora para a prática real de diversos conceitos e áreas no ramo da Tecnologia da Informação e Comunicação como pensamento computacional com a modularização dos problemas buscando sua simplificação, engenharia de software/UML para um bom planejamento e gerenciamento da aplicação e processos a serem desenvolvidos, boas práticas de programação bem como de seus paradigmas para geração de scripts facilmente legíveis e mais eficazes em termos de velocidade, processamento e resultados, o projeto também inclui diversos requisitos/problemáticas que exigem a utilização de diversas tecnologias e estratégias como o uso de APIs, banco de dados, bibliotecas para melhoria de processos, pesquisa, trabalho em equipe etc. Além de possibilitar o uso de conhecimentos/experiências do ramo da TI adquiridos anteriormente por cada um dos membros do grupo, também ajuda no aprendizado

da linguagem Python, dos conceitos e tecnologias utilizadas na análise de dados, como lidar com prazos, cobranças e clientes. Portanto, a escolha do grupo quanto a problemática se justifica pois possibilita que seus integrantes adquiram conhecimento e experiências reais quanto ao desenvolvimento de sistemas ao mesmo tempo que põem seus conhecimentos e teoria vista em sala de aula em prática.

1.4. Objetivos/resultados/efeitos a serem alcançados (em relação ao problema identificado e sob a perspectiva dos públicos envolvidos)

Dos objetivos esperados com o desenvolvimento do projeto:

- Criação de uma aplicação em linguagem Python que analise com eficácia um conjunto de dados providos de diferentes bases
- Um sistema de uso simples e intuitivo
- Integração desse sistema com um banco de dados que garanta a disponibilidade, integridade, confidencialidade e confiabilidade dos dados
- Geração de um relatório simples e de fácil entendimento que auxilie o cliente na tomada de decisões para futuras ações de marketing e melhorias no seu atendimento, estabelecimento e processos
- Conseguir chegar a um resultado satisfatório tanto para a equipe, com a aquisição de conhecimento e experiência referente a Big Data, como para o cliente, com uma visão clara do que funciona ou não para o atingimento de seus objetivos

1.5. Referencial teórico (subsídio teórico para propositura de ações da extensão)

Como base teórica para o desenvolvimento inicial do projeto bem como das discussões acerca do que poderia ser feito o grupo, por não possuir experiência ou conhecimento aprofundado no que tange análise de dados com python, optou por usar não só as aulas presenciais ministradas pelo professor Luan Muruk Ferreira Tonaco no Centro Universitário Estácio mas também o conteúdo digital disponibilizado pela instituição, videoaulas do YouTube referentes a python e bibliotecas da linguagem para análise de dados e pesquisa por livros (tanto os disponibilizados na biblioteca da instituição quanto aqueles encontrados na internet) como: “Python para Análise de Dados” do autor Wes McKinney, “Python Para Análise de Dados: Tratamento de Dados com Pandas, NumPy e IPython” também do autor Wes McKinney e o livro [“Ciência de Dados do Zero: Primeiras Regras com o Python”](#) de Joel Grus, para que o grupo pudesse conhecer melhor as capacidades do python bem como do funcionamento da análise de Big Data, o Livro Entendendo Algoritmos de Aditya Y. Bhargava importante passo inicial e teórico para entender o que está acontecendo na linguagem de programação python e sua lógica no dia a dia. Sendo assim, em resumo, o referencial teórico que a equipe optou por utilizar em primeiro momento pela equipe foi:

- Aulas presenciais ministradas pelo professor Luan Muruk Ferreira Tonaco
- Slides das aulas ministradas pelo professor Luan Muruk Ferreira Tonaco
- Conteúdo digital disponibilizado pela instituição

- Videoaulas do YouTube (em especial do canal de educação na área de TI Curso em Vídeo)
- Uso de livros da biblioteca da instituição que se referirem a linguagem Python e de suas bibliotecas de desenvolvimento
- Uso de livros encontrados na internet referentes a Análise de Dados com Python

A equipe acreditou que, nesse primeiro momento, esse material e essas fontes seriam suficientes para garantir uma visão geral a cerca de Python, de suas capacidades, bibliotecas úteis em análise de dados e, principalmente, uma visão geral acerca da análise de grandes bases de dados. Tais fontes levaram a equipe a optar, ainda apenas em contexto teórico, em primeiro momento a utilizar formulários do Google e a biblioteca Pandas, já que a quantidade de conteúdo encontrada referente a esses dois assuntos na internet é bem extensa e de fácil acesso e também que biblioteca Pandas do Python estava presente na ementa da matéria de “Tópicos de Big Data em Python” para ser ensinada (e já havia conteúdo digital referente ao tema disponível), além de, em rápida na internet, essa biblioteca aparecer como uma das mais utilizadas na comunidade da linguagem Python. Além disso, com base nas aulas da matéria, a equipe também optou por usar outros frameworks/módulos/bibliotecas do python, como o PySpike, e outras ferramentas para auxílio no desenvolvimento e posterior análise de dados, como o Google Collab e o Excel.

Todo o conteúdo adquirido pelos integrantes do grupo até a primeira semana de teorização do projeto (semana do dia 11/03/2024) no que tange a análise de dados em seus fundamentos, objetivos e aplicações levou o grupo à entender que o projeto cabia dentro do conteúdo de Big Data, isto é, o objetivo do projeto estava de acordo com o que a matéria oferecia. A análise de dados é amplamente utilizada para o marketing pois consegue alinhar (e até fazer previsões sobre) o gosto do cliente com o produto/serviço e estabelecimento que serve de objeto para a análise; como o projeto consistia na captação de mais clientes para aumento de renda e visibilidade do estabelecimento do cliente (clínica odontológica) o uso da análise de dados provindos de formulários de avaliação e de pesquisas de satisfação seriam de grande valia para alinhar as estratégias de marketing com o objetivo do cliente. Já que “Big Data é uma área com intuito de estudar maneiras de tratar, analisar e gerar conhecimentos através de grandes conjuntos de dados” o grupo acreditou que seria plenamente possível gerar um relatório (com gráficos, análises e sugestões) que auxiliasse estratégias de marketing para melhor captação de clientes para a clínica. Todos os conteúdos teóricos acima (em primeiro momento considerando apenas os slides, conteúdo ministrado e videoaulas) confirmavam essa ideia e ajudaram o grupo a pensar em diferentes maneiras de captar os dados, analisar e utilizá-los para o marketing da clínica.

Quanto ao Python, a linguagem é amplamente utilizada no campo da Ciência de Dados, um campo da área de Tecnologia da Informação e Comunicação voltado ao estudo da informação desde a sua captura, transformação, análise e geração de

dados “soltos” em informação bem estruturada e com grande valor para organizações. Além disso, a comunidade da linguagem é ampla e muito ativa (é uma das linguagens mais conhecidas e utilizadas atualmente em todo o mundo), o que ajuda na procura por conteúdo e esclarecimento de dúvidas consequentemente ajudando no desenvolvimento desse projeto de extensão; outro fator a favor do Python é sua simplicidade de sintaxe em relação a outras linguagens, o Python não é fortemente tipado o que, aliado a grande quantidade de conteúdo disponível de forma gratuita e de fácil acesso, ajuda no aprendizado da tecnologia amenizando problemas de falta/pouco conhecimento sobre a linguagem; Um último fator é a quantidade de plataformas disponíveis para facilitação do desenvolvimento de aplicações com Python bem como a quantidade de bibliotecas com as mais diversas aplicações no mundo real. Tais fatores, apoiados por todos os subsídios teóricos listados acima (vale ressaltar que na primeira semana de teorização do projeto de extensão os únicos subsídios teóricos realmente consultados e tomados como base foram as aulas presenciais, slides de conteúdo e videoaulas da internet) levaram a equipe a propositura deste projeto extensionista com a convicção de que seu desenvolvimento seria plenamente possível e estaria dentro do proposto pela matéria além de agregar conhecimento e experiência a todos os integrantes acerca da análise de dados com o uso da linguagem de programação Python.

2. PLANEJAMENTO E DESENVOLVIMENTO DO PROJETO

2.1. Plano de trabalho (usando ferramenta acordada com o docente)

Para este projeto, estabelecemos a meta de entregar uma parte do trabalho por semana, visando uma progressão contínua. Optamos por alinhar os prazos de entrega com os dias de aula, programados para todas as sextas-feiras. Reconhecemos, no entanto, que imprevistos podem surgir, como a ausência de um membro ou eventos específicos, impossibilitando reuniões presenciais. Para contornar essas situações, realizamos reuniões virtuais através de plataformas como Discord ou WhatsApp, programadas uma ou duas vezes por semana, durante as quais esclarecemos dúvidas e atualizamos o progresso de cada membro.

Agora nós estamos usando uma ferramenta de programação muito útil e fácil de manipular que é o Jupyter notebook, através dessa ferramenta nós estamos usando as bibliotecas usadas em aula para podermos criar um gráfico para melhor compreensão do nosso cliente na hora de apresentarmos a ele os dados que coletamos, as formas de gráficos que estamos pensando ainda não foi decidida, mas estamos optando por gráficos pizza ou de barra para que possamos ter uma fácil compreensão dos dados, vamos usar alguns formulários do google para coletar as informações e depois vamos transformar os arquivos do google no formato 'json' para que possamos utilizar ele em nossas bibliotecas, como Eu 'Matheus Oliveira' fiquei designado a fazer as pesquisas necessária para que possamos fazer um projeto o mais profissional possível dentro das nossas capacitações como desenvolvedores e alunos, mesmo assim vamos recorrer a orientação e dicas de nosso supervisor e professor para que possamos avançar com mais eficiência.

Para realizarmos uma entrega de trabalho com informações mais fáceis de ler no dia 03/05/2024, Matheus Oliveira e Gustavo Erades, vão criar alguns gráficos para que possa ser mais simples a compreensão do professor e de nosso cliente caso o mesmo possa comparecer a nossa reunião, Matheus Oliveira ficou encarregado de fazer um gráfico de barra para podermos avaliar e analisar os resultados das perguntas do nosso formulário que criamos, também decidimos criar um QRCode para que possamos compartilhar o nosso formulário com mais eficiência no consultório do cliente, mas ao mesmo tempo que não atrapalhe ou cause incômodos ao mesmo, dessa forma poderemos adquirir feedbacks dos clientes do nosso cliente com mais facilidade, para que possamos criar os gráficos em conjunto com os formulários que criamos decidimos utilizar da ferramenta 'PowerBI', através dessa ferramenta podemos tanto criar os nossos gráficos como também vamos poder fazer mais alguma análise caso o cliente queira ou caso achemos necessário, usando dessas ferramentas podemos criar formas mais intuitivas e fáceis de se compreender para todos para quem vamos apresentar nosso trabalho.

2.2. Descrição da forma de envolvimento do público participante na formulação do projeto, seu desenvolvimento e avaliação, bem como as estratégias utilizadas pelo grupo para mobilizá-los.

Estamos considerando a implementação de uma estratégia de marketing com o intuito de ampliar a base de clientes e o alcance da clínica odontológica do cliente. Planejamos um sistema de avaliação que nos permitirá identificar áreas de melhoria, avaliar o nível de satisfação dos clientes e identificar possíveis lacunas no serviço prestado. Entendemos que a satisfação do cliente desempenha um papel crucial na reputação e no crescimento do negócio, pois clientes satisfeitos tendem a recomendar os serviços da clínica a outros, contribuindo assim para a expansão do negócio.

Dessa maneira, para a mobilização do público, o grupo criou dois formulários para serem preenchidos pela internet. Tais formulários foram disponibilizados em forma de *link* e distribuídos conforme seus devidos objetivos:

Um dos formulários se destina ao público em geral, isto é, a qualquer pessoa que já tenha sido atendido(a) em um clínica odontológica, seu objetivo é coletar dados suficientes para o entendimento dos fatores que levam a maioria do público a buscar um consultório e se manter nele. Esse formulário foi disponibilizado ao Centro Universitário Estácio por intermédio da coordenadora dos cursos de Tecnologia Josyane Lannes sendo enviado no grupo de *whatsapp* da comunidade da instituição, *The Tech Wizards*. Além disso, ele também foi disponibilizado ao ciclo social dos membros do grupo, os quais também incentivaram o compartilhamento massivo do *link* de preenchimento do formulário.

Já o outro formulário possui um objetivo mais específico, ele se refere àqueles que já são pacientes da clínica odontológica (do cliente no qual este trabalho se destina) ou já realizaram algum procedimento no estabelecimento. Para esse formulário, não foi possível uma captação tão veloz de dados quanto do formulário de satisfação geral, porém ele também foi disponibilizado ao público de duas maneiras, por meio de um *QRCode* impresso e disponível no estabelecimento e por meio de um *link* presente no *website* da clínica odontológica.

2.3. Grupo de trabalho (descrição da responsabilidade de cada membro)

O grupo de trabalho foi organizado de forma a permitir a colaboração de todos os membros tanto na parte escrita quanto na parte prática do projeto. A fase inicial do projeto contou com a participação de três membros: Guilherme Estevão, Gustavo Êrades e Matheus Oliveira, responsáveis pelo planejamento inicial. Durante uma reunião na biblioteca da instituição de ensino, ficou estabelecido que a elaboração da primeira parte do trabalho escrito seria dividida entre Gustavo Êrades e Matheus Oliveira. Gustavo ficou encarregado da supervisão e desenvolvimento dos primeiros temas, enquanto Matheus assumiu a redação do "referencial teórico". Por sua vez, o contato com o cliente, desde o planejamento até a conclusão do projeto, foi designado exclusivamente a Guilherme Estevão.

Quanto à redação da segunda parte do roteiro de extensão, decidiu-se que cada membro seria responsável por um tema específico relacionado ao planejamento e desenvolvimento do projeto. Os membros se reuniram via WhatsApp e escolheram os temas de acordo com suas afinidades:

- Guilherme Estevão: redação do tema 2.2 ("Descrição da forma de envolvimento do público participante na formulação do projeto");
- Gustavo Êrades: redação de toda a parte de "Diagnóstico e Teorização", redação dos temas 2.3 ("Grupo de trabalho") e 2.6 ("Detalhamento técnico do projeto") da parte de "Planejamento e Desenvolvimento do projeto", revisão de todo tópico 2 do roteiro de extensão e implementação da solução;
- Marcos Vinicius: redação do tema 2.5 ("Recursos previstos");
- Matheus Oliveira: redação do tema 2.1 ("Plano de trabalho") e supervisão da redação de cada membro referente ao tema 2 ("Planejamento e Desenvolvimento do projeto");
- Washington Carmo: redação do tema 2.4 ("Metas, critérios ou indicadores de avaliação do projeto").

No que diz respeito à parte prática do projeto, foram agendadas reuniões semanais entre os membros do grupo, todas as sextas-feiras, para a entrega e análise do progresso. Como o início do projeto dependia da coleta e análise de dados, foram atribuídas as seguintes responsabilidades para o período de 25/03/2024 a 05/04/2024:

- Guilherme Estevão: formulação de um formulário via Google Forms para uma pesquisa de satisfação, destinado aos pacientes da clínica;
- Gustavo Êrades: formulação de um formulário via Google Forms para uma pesquisa de satisfação sobre clínicas odontológicas, tratamento de pacientes, etc., direcionado ao público em geral;
- Marcos Vinicius: pesquisa sobre o uso da linguagem Python para análise de dados provenientes dos formulários elaborados;
- Matheus Oliveira: pesquisa sobre o uso da linguagem Python para análise de dados provenientes dos formulários elaborados e supervisão da elaboração do tema 2 do roteiro de extensão;
- Washington Carmo: pesquisa sobre o uso da linguagem Python para análise de dados provenientes dos formulários elaborados.

2.4. Metas, critérios ou indicadores de avaliação do projeto

O Projeto Visa utilizar técnicas de Big data em Python para auxiliar o cliente, dono de uma clínica odontológica, a tomar decisões referentes ao marketing para captação de clientes e para isso estamos utilizando formulários e meios de pesquisas para receber e tratar dados de usuários diferentes a fim de obter uma análise do que leva as pessoas a se manterem em determinado consultório, bem como dos fatores que as afastam de determinados consultórios. O objetivo final é a produção de um relatório que resume e detalhe todos os dados coletados e informações geradas, tal relatório será entregue ao cliente e define o maior indicador de sucesso do projeto: a geração de um relatório que possa esclarecer, ou ao menos fornecer uma direção, para o marketing comercial da clínica.

Dentre os critérios de avaliação usados que levaram a equipe a ponderar sobre outras estratégias e pensar ou descartar ideias, estão:

- No tocante a implementação:
 - geração de um código de fácil entendimento e manipulação;
 - código de fácil evolução/alteração;
 - eficiência na hora de executar testes e alterar componentes;
 - rapidez no desenvolvimento.
- No tocante a análise dos dados:
 - geração de dados os mais “precisos” possíveis;
 - possibilitar a análise fácil e rápida;
 - possibilitar a identificação de aspectos negativos e positivos
 - possibilitar a geração de soluções para os aspectos negativos;
 - gerar informações valiosas que pudessem ajudar o cliente na propositura de ações de marketing e captação de clientes.

2.5. Recursos previstos

Dos recursos previstos inicialmente para o desenvolvimento do projeto, estão:

- Uso do Visual Studio Code, Google Collab e Jupyter Notebook para a implementação do projeto;

- Google drive e GitHub para o compartilhamento dos arquivos gerados bem como versionalização do projeto;
- Google Forms e Microsoft Forms para a geração dos formulários (coleta dos dados);
- Microsoft Excel, Google Sheets e Libreoffice Calc para a geração de planilhas com os dados a serem analisados bem como integração com novas respostas dos formulários;
- Bibliotecas do Python como Pandas e Openpyxl para facilitar a codificação e os resultados;
- Power BI e Power Automate para representação visual dos dados e análises.

2.6. Detalhamento técnico do projeto

O desenvolvimento do projeto consistiu em três partes, uma de planejamento, outra de implementação e uma última de análise e entrega dos resultados gerados. Tais etapas deveriam ser realizadas algumas vezes até a entrega final do projeto para que os passos pudessem ser desenvolvidos conforme sua efetividade, isto é, para que pudesse ser analisado se o processo funcionou ou não, se custou muito tempo ou não, se gerou algum resultado relevante ou não, etc. Podendo assim aprimorar o desenvolvimento do projeto até a sua entrega final.

O planejamento do projeto, por ser a primeira etapa e ser essencial para a geração de resultados efetivos deveria ser realizada com cuidado e pensada e repensada quantas vezes fosse necessário, no entanto essa etapa ficou sujeita às rotinas dos membros do grupo o que repercutiu em seu atraso e consequentemente no atraso das demais etapas. Por conta disso o desenvolvimento do projeto ficou comprometido em algumas semanas, mas ainda assim um primeiro ciclo das etapas foi possível ser executado para uma primeira entrega.

Para a coleta dos dados a serem analisados, foram desenvolvidos dois formulários utilizando a plataforma Google Forms, em colaboração com o Chat-GPT na versão 3.5. Um formulário foi direcionado aos clientes do consultório odontológico em questão, enquanto o outro foi disponibilizado ao público em geral. O propósito desses formulários era investigar os motivos que levam o público a escolher determinado consultório e identificar os fatores que os mantêm fiéis a esse estabelecimento. O formulário de satisfação geral (oferecido para o público) foi disponibilizado para a comunidade do Centro Universitário Estácio (The *TechWizards*), para o ciclo social dos integrantes do grupo (amigos e família) e para quem mais tivesse o *link*. Já o formulário específico para o consultório do cliente não foi possível sua impressão e portanto foi disponibilizado (no próprio consultório) como um *QRCode* e também foi colocado no *website* do cliente.

Após a coleta dos dados, apesar da funcionalidade de geração automática de gráficos pelo Google Forms, optou-se por elaborar esses gráficos manualmente utilizando a linguagem de programação Python em conjunto com o módulo "matplotlib". O Google Forms permite o download das respostas em formato .csv, sendo este arquivo posteriormente convertido para o formato .json utilizando o módulo "pandas" do Python. É por meio deste arquivo que os gráficos são gerados e os dados são manipulados. Posteriormente, esses dados devidamente

processados são inseridos em uma planilha criada ou editada utilizando o módulo "openpyxl" também do Python.

A escolha dos formulários do Google para a coleta de dados foi motivada pela sua simplicidade tanto no preenchimento quanto na formatação, captação das respostas e análise em tempo real conforme as respostas são fornecidas. Uma vez que é possível baixar as respostas em formato .csv (Comma Separated Values), os formulários foram disponibilizados ao público por aproximadamente duas semanas e posteriormente os dados foram transferidos para esse formato. Entretanto, o formato .csv se mostrou inadequado para a manipulação e tratamento dos dados devido à sua estrutura desorganizada. Portanto, optou-se por converter o arquivo .csv para o formato .json (JavaScript Object Notation), o qual apresentou uma organização mais adequada e simplificou tanto a leitura quanto a edição e extração de dados para o tratamento necessário.

Os dados extraídos do arquivo .json foram então agregados, analisados, tratados, correlacionados e representados graficamente, com o objetivo de facilitar a análise futura e proporcionar uma visualização clara das informações, ao contrário dos dados dispersos e não relacionados obtidos diretamente dos formulários. Com essa análise completa, o grupo estará preparado para realizar projeções e sugerir mudanças nas estratégias de marketing para o cliente.

Ao baixar as respostas do formulário de satisfação geral como um arquivo .csv, por meio do python, ele foi passado para um arquivo .json para facilitar sua manipulação. Como os formulários do Google salvam data e hora do preenchimento seria necessário retirar esses dados, afinal eles não são necessários para o desenvolvimento do projeto. Além disso os objetos possuíam as perguntas completas dos formulários o que dificultava a leitura e não facilitava na hora de se criar uma planilha - porque dessa maneira os textos ficaram muito grandes nas colunas - deixando tudo muito bagunçado, portanto seria necessário retirar essa parte também. Sendo assim, o arquivo .json foi editado e as perguntas reduzidas e em seguida o arquivo foi salvo novamente (esse processo foi feito inteiramente com Python). Com esse arquivo devidamente editado já era possível começar a gerar gráficos para a análise de dados.

No entanto, esse processo se provou trabalhoso e não muito eficiente, levando o grupo a mudar de estratégia.

A nova estratégia adotada pela equipe foi a utilização do Power BI. Essa ferramenta permite a análise de dados/negócios de maneira simples e interativa, é uma ferramenta desenvolvida pela Microsoft para a geração de relatórios e dashboards que permite a integração com algumas outras plataformas da Microsoft (como o Office 365 por exemplo). A ideia agora consistia em trocar os formulários do Google por uma versão de formulários do office que seria integrado (por meio do Power Automation) à uma planilha Excel e a geração de gráficos e posterior análise seria feita com o uso do Power BI.

Como os formulários já haviam sido disponibilizados e, pelo menos, aberto ao público, já possuía respostas, foi preciso unir as duas ideias. Com o uso do python

as respostas do formulário de satisfação geral - formulário este ainda do Google - foram passadas para uma planilha seguindo o seguinte processo:

- download do arquivo .csv do formulário do Google
- conversão do arquivo para .json
- ajustes nas perguntas (para melhor visualização na planilha) e nos dados retornados (como exclusão dos dados de horário de preenchimento das respostas)
- geração de planilha com os dados do arquivo .json

Para esse processo foram usadas apenas duas bibliotecas: Pandas para leitura do arquivo, conversão e ajustes nos dados e o Openpyxl para a geração da planilha e processo de população dela com os dados do arquivo .json. Esse formulário, agora em formato de planilha (.xlsx), permitiu que fossem gerados gráficos e que pudessem ser feitas análises posteriores tanto com o uso das bibliotecas do Python quanto com o uso do Power BI Desktop. Posteriormente esse formulário foi reescrito, mas agora como um formulário da Microsoft, já permitindo o uso do Power automate e uma maior facilitação na manipulação das próximas respostas já que este estaria integrado a primeira planilha gerada com as primeiras respostas.

O formulário de satisfação geral do Google já estava com o *link* circulando e portanto suas respostas foram vinculadas a uma terceira planilha do Google Sheets, assim caso houvesse mais respostas nele seria possível visualizar esse dado e juntá-lo com os demais já coletados. O outro formulário (específico para para o consultório do cliente) também foi transferido para um formulário da Microsoft e integrado a uma planilha específica. Dessa maneira o processo ficou mais simples e a complexidade/trabalho para geração de gráficos e análises foi reduzida.

Após a conclusão desse processo, a ideia é gerar um documento contendo detalhes da análise realizada, o qual é entregue ao cliente. Esse documento representa o resultado final do projeto e tem como objetivo orientar as decisões de marketing do cliente em relação ao consultório. A escolha de utilizar as bibliotecas Python mencionadas, bem como a forma de utilizá-las de maneira eficaz para o projeto, foi embasada nas aulas ministradas na instituição, no conteúdo digital disponibilizado e em recursos adicionais encontrados na internet, como artigos, vídeos e fóruns relacionados à análise de dados, manipulação de dados em Excel com Python, uso das bibliotecas pandas, matplotlib, openpyxl, big data, Power BI, Power Automate, etc.

No que se refere à persistência dos dados, inicialmente consideramos a utilização do SQLite3 devido à sua natureza como uma biblioteca que oferece um Sistema de Gerenciamento de Banco de Dados (SGBD) de fácil administração, operação e configuração, além de fornecer operações básicas de CRUD (Create, Read, Update, Delete) com o banco de dados. Uma vez que a finalidade era apenas persistir os dados para facilitar sua análise e servir como backup, não julgamos necessário recorrer a um SGBD mais complexo ou robusto. No entanto, como houveram alguns imprevistos durante o desenvolvimento do projeto e o planejamento de como os resultados seriam obtidos precisou ser alterado ao longo do desenvolvimento, essa questão referente a persistência dos dados foi “negligenciada” durante a primeira

etapa do projeto permitindo o grupo em focar na análise dos dados e em formas de sua representação..

Dentre as tecnologias utilizadas para o desenvolvimento do projeto estão:

- Visual Studio Code para codificação do projeto - bem como o Google Collab e Jupyter Notebook que foram utilizados com o mesmo propósito;
- Chat GPT para pesquisas mais rápidas e respostas à dúvidas de maneira mais fácil e rápida;
- Git e GitHub para versionamento do código e gerenciamento dos arquivos do projeto;
- Google drive para gerenciamento dos arquivos referentes a parte escrita do projeto;
- Libreoffice Writer e Google docs para a escrita do roteiro de extensão e demais arquivos de texto gerados no processo;
- Google formulários para a coleta de dados inicial
- Excel e Libreoffice Calc para a organização dos gráficos gerados, geração de tabelas e análise final;
- Power BI para geração de gráficos e maior facilidade na análise dos dados
- Microsoft Forms e Office 365 para construção de formulários e seu gerenciamento;
- Microsoft Power Automate para integração do Excel com os formulários.

Tais ferramentas foram escolhidas ou por já serem de conhecimento dos integrantes do grupo ou por facilidade/necessidade dos mesmos durante o desenvolvimento deste trabalho.

Os códigos usados no projeto foram compartilhados por meio de um repositório público no GitHub:

- [Repositório do código](#)

Inicialmente, foram feitas pesquisas e testes no tocante a implementação do projeto, tais testes servirão para a definição de quais gráficos seriam gerados e como seriam gerados, bem como a definição de como funcionaria a análise deles. Dessa etapa foi gerado uma série de arquivos que seriam usados posteriormente como espécies de templates para consulta de como realizar a implementação do código e gerar resultados efetivos. Como essa etapa foi realizada localmente já estava sendo utilizado o ambiente virtual do python (venv) para evitar qualquer tipo de conflito/danificação na máquina e permitir a instalação de bibliotecas e/ou outras versões no python conforme fosse necessário para o projeto. Após essa etapa a implementação do primeiro ciclo - coleta de dados, implementação, análise de resultados - do projeto foi iniciada. Para a etapa de implementação ficou definido que os alunos:

- Gustavo Êrades
- Matheus Oliveira

Seriam os responsáveis e para a análise e geração de documento com tal análise ficaria de responsabilidade dos alunos:

- Guilherme Estevão
- Washington Carmo
- Marcus Vinícius

3. ENCERRAMENTO DO PROJETO

3.1. Relato Coletivo:

Durante o desenvolvimento do projeto o grupo passou por alguns percalços que prejudicaram a implementação do projeto e consequentemente os resultados obtidos. No entanto, mesmo com dificuldades o projeto não foi parado em momento algum e o grupo buscou por diferentes alternativas a fim de contornar situações problemáticas. O aprendizado adquirido com este projeto foi unânime, todos os membros tiveram a oportunidade de pôr o conhecimento aprendido em sala de aula em prática e, em algum nível, puderam ter uma experiência com a análise de dados. O grupo concordou que a experiência foi interessante e agregou para seu conhecimento acerca da matéria, da linguagem Python e das diversas ferramentas e possibilidades do campo de análise de dados. Além disso, o projeto também serviu como experiência de trabalho em equipe, divisão de tarefas entre os membros de um projeto, lidar com prazos, aprendizado de novas ferramentas, procura por alternativas para contornar problemas, comunicação com clientes e com o público de maneira geral, etc. Enfim, o projeto foi uma ótima oportunidade de aprender e colocar em prática o que foi visto nas aulas da matéria de “Tópicos de Big Data em Python” e também de vivenciar o desenvolvimento de um projeto real em equipe, o trabalho foi fundamental para entendermos todos os desafios e o dia a dia do mercado de trabalho.

3.2. Relato de Experiência Individual

3.2.1. Guilherme Estevão:

Fiquei responsável pelo cliente, entrar em contato com o cliente para ver o que ele mais queria para poder ajudar na clínica odontológica, e o que o cliente mais pediu foi o marketing digital, todos nós nos esforçamos ao máximo para deixar o cliente satisfeito e feliz, e tbm ao mesmo tempo cumprir um bom trabalho, o bom da prática é que você vai aprendendo com seus erros e assim corrigindo o que no futuro poderia ser um problema por já ter passado por uma situação prática fica mais fácil pensar nos erros e ficar mais atento a tudo, em geral esse trabalho ajuda na performance ajuda a saber trabalhar um pouco em equipe, ter um colaborante e compromisso, com uma base um pouco mais sólida não dar tanto um nervosismo mais por pensarem no futuro ter uma equipe, sabendo administrar tudo pode dar certo

3.2.2. Gustavo Êrades:

Fiquei responsável pelo planejamento e desenvolvimento do projeto, incluindo desde a definição do que seria feito até a geração de dados e sua análise. Também fui responsável pela escrita da primeira parte deste documento "Diagnóstico e Teorização" e pela escrita dos tópicos 2.5 e 2.6 e pela supervisão/ajustes nos demais tópicos da parte de "Planejamento e Desenvolvimento do projeto". Minhas atribuições com este trabalho sofreram alterações de acordo com o andamento do projeto e, embora houvessem sido definidas responsabilidades específicas para cada um dos membros em uma reunião online, foi necessário modificar minhas responsabilidades com este trabalho ao longo do tempo para que as entregas constantes e avaliações do projeto não fossem prejudicadas. Apesar disso, gostei de realizar o projeto principalmente porque ele me mostrou uma face do python a qual eu apenas havia ouvido falar, a análise de dados.

O projeto me deu a oportunidade de trabalhar com ferramentas diferentes as quais eu não conhecia e me proporcionou uma visão diferente quanto ao python e suas aplicações, tive a oportunidade de testar e mexer com ferramentas novas para mim como o Power BI, Power Automate e Streamlit, além disso pude ter uma noção, ainda que simples, de como funciona a área de análise de dados e me surpreendi com a quantidade de informações que se é possível obter relacionando dados dispersos e que parecem não ter relação.

Embora tenha gostado de realizar o projeto e tenha aprendido bastante, ele sofreu com alguns percalços que, para mim, prejudicaram seu resultado. O tempo para desenvolvimento do projeto foi grande e o suficiente para reunir dados suficientes para uma análise mais robusta, aprofundada e certa, porém conseguimos poucos dados devido a diversos problemas que nós (o grupo de trabalho) enfrentamos ao longo do desenvolvimento, como problemas com as ferramentas a serem usadas, problemas para a captação de dados nos formulários gerados, falta de tempo e/ou interesse das partes interessadas. Tais fatores acabaram por prejudicar o andamento do projeto e minha experiência com ele também, os resultados obtidos poderiam ter sido mais satisfatórios.

Caso fosse possível continuar o projeto seria necessário o acúmulo de mais dados para que o resultado obtido fosse mais confiável e para isso seria necessário encontrar maneiras mais certas de mobilizar pessoas a fornecerem os dados ou ainda encontrar dados disponíveis na internet que satisfazem os objetivos do projeto. Mesmo assim, fazer parte de um projeto de análise de dados e mexer com ferramentas diferentes foi bem interessante e gostei da experiência.

3.2.3. Matheus Oliveira:

Fiquei responsável por ajudar em algumas partes gráficas, monitorar o que os outros integrantes do grupo escreviam para que o trabalho ficasse com o mínimo de erros possível, tanto em escrita quanto em contexto, também fiquei responsável pelo planejamento do trabalho, onde reunimos para que pudéssemos sempre tentar equilibrar o trabalho o máximo possível, por ficar com a parte de planejamento também fiquei responsável por documentar tudo em nosso documento, também ajudei em algumas ideias do formulário, como também compartilhá-lo para que pudéssemos tentar reunir o maior número de dados possível, assim podemos ter um melhor desempenho para criação de gráficos de dashboard, como resultado desses nossos esforços pudemos aprender mais sobre algumas ferramentas como o Power BI, os formulários do Google, e também nos familiarizamos mais com o excel onde guardamos nossos dados, usamos também de uma ferramenta para criar um QR Code para que os nossos formulários pudessem ser distribuídos com mais facilidade, ao utilizar dessas ferramentas, tivemos um desempenho melhor em grupo e mais facilitado para a elaboração de nosso projeto.

3.2.4. Washington Carmo:

Fiquei responsável por ajudar definir as metas e os critérios do trabalho, e apresentar isto neste roteiro, além de participar de reuniões para definir o andamento, possível resultado e ideias do trabalho, e outras partes do projeto, como coordenação de equipe, comunicação ou gerenciamento de tarefas, entender o que estava acontecendo na parte de I/O pode ter facilitado a comunicação entre você e os membros da equipe que estavam lidando com essa parte do código.

Mesmo não conseguindo ajudar de maneira eficiente na codificação, enfrentei desafios ou problemas durante o projeto. A compreensão dos conceitos por trás das operações de entrada e saída em Python me ajudou a contribuir para a resolução de problemas de forma mais eficaz.

Além disso, o projeto me ajudou a entender como muitos projetos e etc no mercado são trabalhados em grupo, isso me botou no dia a dia nessas empresas e me fez entender melhor como funciona.