

## Estudo de Caso: Inclusão Digital na cidade de Juititiba

### Identificando os Envolvidos

Vamos imaginar que estamos em uma sala de aula, onde o tema do dia é a inclusão digital e o impacto social. O nosso projeto é como uma ponte que conecta duas margens: de um lado, temos a Azza, uma das maiores empresas de telecomunicações da região de SP, e do outro, a comunidade da cidade Juititiba.

A AZZA é como um grande barco que navega pelas águas das comunicações, oferecendo serviços de telefonia, internet, TV streaming e monitoramento de câmeras. No entanto, esse barco quer ir além de simplesmente vender serviços. A Azza deseja ser um agente de transformação social, ajudando a comunidade a ter acesso a serviços digitais e promovendo a inclusão digital.

A Comunidade Local: A cidade de Juititiba é como um pequeno mundo, onde vivem cerca de 31.444 pessoas. A maioria dos moradores enfrenta desafios financeiros, com uma grande parte vivendo com menos de um salário mínimo por mês. Imagine uma sala de aula onde muitos alunos não conseguiram terminar o ensino médio; isso é o que se vê aqui, onde a escolaridade predominante é o ensino fundamental incompleto.

### Participantes do Projeto

Neste projeto, esperamos que cerca de 150 pessoas se envolvam diretamente nas atividades. Entre elas, teremos jovens que buscam se qualificar profissionalmente, adultos que querem aprender novas habilidades digitais e famílias que desejam melhorar o acesso à internet e à informação. É como uma grande festa onde todos são convidados a aprender e crescer juntos!

### A Situação-Problema

No entanto, nem tudo é fácil. A comunidade da cidade de Juititiba enfrenta uma verdadeira tempestade quando se trata de acesso a serviços digitais. Imagine que a internet é como uma estrada que leva a novas oportunidades. Infelizmente, essa estrada está cheia de buracos e sem sinalização, dificultando a jornada dos moradores. A falta de infraestrutura e de iniciativas para inclusão digital resulta em grandes desafios, limitando as oportunidades educacionais e econômicas.

Sem acesso à internet e às ferramentas digitais, os moradores têm dificuldade em encontrar empregos, continuar seus estudos ou até mesmo realizar tarefas do dia a dia, como pagar contas. É como tentar navegar em um mar revolto sem um mapa!

### Motivação Acadêmica e Demanda Sociocomunitária

Nosso projeto surge como uma resposta a essa demanda por inclusão digital. Queremos capacitar a população local em habilidades tecnológicas básicas e facilitar o acesso à

internet. É uma oportunidade de aplicar o que aprendemos em sala de aula sobre tecnologia da informação, comunicação e desenvolvimento comunitário. Aqui, a educação técnica se torna uma ferramenta poderosa para promover a equidade social.

### Objetivos a Serem Alcançados

Agora, vamos falar sobre o que queremos alcançar:

1. Acesso à Internet e monitoramento de câmeras: Queremos garantir que 80% das residências da cidade tenham acesso à internet de alta qualidade até o final do semestre e que a cidade tenha o monitoramento para segurança 24hrs.
2. Capacitação: Nossa meta é capacitar 300 moradores em habilidades digitais básicas até o fim do projeto. Imagine essas pessoas se tornando navegadores habilidosos no vasto oceano da internet!
3. Centro Comunitário Digital: Queremos facilitar a criação de um centro comunitário digital, um espaço onde a população poderá acessar recursos tecnológicos e receber orientação, tudo isso até o término do semestre.

Esses objetivos serão monitorados e avaliados em parceria com a Azza e representantes da comunidade local, garantindo que os resultados atendam às necessidades e proporcionem um impacto positivo duradouro na vida dos moradores.

### Conclusão

Este projeto é uma oportunidade incrível para todos nós. Imagine um futuro onde a comunidade tenha as ferramentas necessárias para navegar com confiança no mundo digital. Juntos, podemos transformar essa visão em realidade!

### Estudo de Caso: Capacitação em Big Data

#### Objetivo do Projeto:

Neste projeto, nosso objetivo é capacitar a comunidade do Bairro XXX a utilizar tecnologias de Big Data, como Hadoop e Spark, para analisar dados socioeconômicos. Isso não só promove a inclusão digital, mas também impulsiona o desenvolvimento local. Imagine que estamos fornecendo ferramentas para que os moradores possam desvendar histórias escondidas nos números da sua própria comunidade!

#### Ação 1: Obtenção e Preparação dos Dados

##### O que vamos fazer?

Vamos coletar dados socioeconômicos da comunidade e prepará-los para análise. Pense nisso como reunir peças de um quebra-cabeça que, juntas, revelarão uma imagem clara da realidade do bairro.

Quando?

Nas primeiras duas semanas.

Como?

- Coleta de Dados: Vamos buscar informações em fontes confiáveis, como o IBGE e registros municipais.
- Armazenamento: Usaremos o Apache Hadoop, que é como um grande armário onde armazenaremos todos esses dados de forma organizada.

Requisitos:

Precisamos de um cluster Hadoop, acesso a bases de dados e uma equipe de TI para nos ajudar.

Onde?

No setor de T.I da empresa.

#### Ação 2: Processamento de Dados com Hadoop

O que faremos?

Vamos processar e transformar os dados brutos coletados. Imagine que estamos limpando e preparando ingredientes antes de cozinhar um prato delicioso.

Quando?

Semanas 3 a 4.

Como?

- Criação de Jobs: Vamos criar "jobs" no Hadoop, que são como receitas que nos ajudam a mapear e reduzir grandes volumes de dados.
- Limpeza: Precisamos garantir que nossos dados estejam prontos para análise, eliminando qualquer informação desnecessária.

Requisitos:

Um cluster Hadoop configurado e scripts de MapReduce.

Onde?

Na infraestrutura de TI da Universidade.

#### Ação 3: Análise de Dados com Apache Spark

O que vamos fazer?

Analisaremos os dados processados usando o Apache Spark. Pense nisso como usar uma lupa para observar detalhes que antes não eram visíveis.

Quando?

Semanas 5 a 6.

Como?

- Implementação de Scripts: Vamos executar scripts que nos ajudarão a agregar, filtrar e agrupar os dados.

- Geração de Insights: Utilizaremos o Spark para descobrir informações valiosas sobre a comunidade.

Requisitos:

Um cluster Spark e acesso aos dados processados.

Onde?

No setor de T.I da empresa.

Ação 4: Criação de Visualizações e Relatórios

O que faremos?

Desenvolveremos dashboards e relatórios que traduzem os insights obtidos em informações visuais. Imagine um mapa que mostra não apenas onde você está, mas também as melhores rotas para chegar ao seu destino.

Quando?

Semanas 7 a 8.

Como?

- Ferramentas de Visualização: Usaremos ferramentas como Apache Zeppelin ou Tableau para criar gráficos interativos.

- Relatórios Detalhados: Vamos gerar documentos que ajudem na tomada de decisões.

Requisitos:

Acesso aos resultados da análise e ferramentas de visualização.

Onde?

na sala de reuniões da empresa..

Ação 5: Treinamento e Apresentação dos Resultados

O que faremos?

Apresentaremos os resultados à comunidade e realizaremos treinamentos. Imagine que estamos abrindo as portas de uma nova sala de aula, onde todos podem aprender e se beneficiar.

Quando?

Semanas 9 a 10.

Como?

- Workshops: Realizaremos encontros presenciais e online para ensinar a comunidade a utilizar as ferramentas desenvolvidas.

- Discussão dos Insights: Vamos compartilhar os aprendizados e discutir como eles podem impactar a comunidade.

Requisitos:

Material didático e um espaço para os workshops.

Onde?

No centro comunitário da cidade.

#### Envolvimento da Comunidade

##### Planejamento:

Começaremos com uma reunião inicial com os líderes comunitários para entender suas expectativas e necessidades.

##### Desenvolvimento:

A comunidade participa de testes e validações, oferecendo feedback ao longo do processo.

##### Avaliação:

Coletamos opiniões e sugestões após cada fase do projeto, usando questionários e discussões em grupo.

#### Avaliação dos Resultados

##### Como vamos medir o sucesso?

- Comparação de Indicadores: Vamos analisar os dados socioeconômicos antes e depois do projeto.
- Questionários de Satisfação: Aplicamos questionários para entender a satisfação da comunidade.
- Observação Direta: Verificaremos como as ferramentas e insights gerados estão sendo utilizados para melhorar as condições locais.

#### Resultados Esperados

- Acesso Melhorado: A comunidade terá acesso a dados e informações críticas.
- Capacitação Digital: Aumentaremos a capacitação digital e o uso de tecnologias avançadas.
- Impacto Positivo: As decisões e o desenvolvimento local serão impactados positivamente com base nas análises realizadas.

Com este projeto, não estamos apenas utilizando tecnologia; estamos construindo um futuro mais informado e participativo para o Bairro XXX. Juntos, podemos transformar dados em ações que beneficiam a todos!

#### Relato da Experiência Individual no Projeto: Um Estudo de Caso

##### 1. Contextualização

Vamos imaginar que a comunidade da cidade é como um grande quebra-cabeça, onde cada peça representa um aspecto da vida dos moradores: saúde, educação, emprego, entre outros. No meu projeto, trabalhei para juntar essas peças usando tecnologias de Big Data, como Hadoop e Spark. O objetivo era claro: ajudar a comunidade a entender melhor seus dados socioeconômicos e promover a inclusão digital.

Minha função foi como a de um mecânico que, ao invés de consertar carros, ajustava ferramentas tecnológicas para coletar, processar e analisar dados. Ao aplicar o que aprendi na sala de aula, pude ver como a teoria se transforma em prática, ajudando a comunidade a tomar decisões mais informadas.

## 2. Metodologia

Para que esse projeto funcionasse, combinamos métodos técnicos com a participação ativa da comunidade. Aqui está um resumo do que fizemos:

- Coleta de Dados: Usamos fontes públicas, como o IBGE, que são como uma biblioteca cheia de informações. Armazenamos esses dados no Hadoop, que é como um grande armário que guarda tudo de forma organizada.
- Processamento e Análise: Criamos scripts de MapReduce, que são como receitas de bolo, para processar os dados. Usamos o Apache Spark para análises mais complexas. Imagine que o Spark é um superchef que consegue preparar vários pratos ao mesmo tempo! Utilizamos ferramentas como PySpark para fazer as análises em Python e o Apache Zeppelin para visualizar os dados de forma amigável.
- Participação Comunitária: Realizamos workshops, que foram como rodas de conversa, onde os moradores puderam dar suas opiniões e sugerir melhorias. Também aplicamos questionários para entender como eles estavam se sentindo em relação ao projeto.
- Implementação Técnica: Configuramos um cluster Hadoop e um cluster Spark, que funciona como um time de supercomputadores trabalhando juntos. Escrevemos o código em Python usando bibliotecas como PySpark e Pandas, que são ferramentas essenciais para quem trabalha com dados.

## 3. Resultados e Discussão

Os resultados foram muito animadores! O projeto teve um impacto positivo na comunidade da cidade. Aqui estão alguns pontos que destacam o que conseguimos:

- Impacto na Comunidade: Conseguimos capacitar 300 moradores em habilidades básicas de Big Data. Isso é como dar a eles uma nova chave para abrir portas que antes estavam fechadas, permitindo que eles interpretem e utilizem dados para tomar decisões.
- Facilidades: A configuração de ambientes de Big Data na nuvem facilitou muito nosso trabalho. A participação ativa da comunidade foi um grande trunfo, como um time jogando junto em um campeonato.
- Dificuldades: No início, alguns moradores tiveram dificuldades para se adaptar às novas tecnologias. Isso nos levou a oferecer sessões adicionais de treinamento, como se estivéssemos dando aulas extras para garantir que todos estivessem na mesma página.

- Descobertas e Aprendizagens: Aprendi que a simplicidade e a clareza na comunicação são essenciais, especialmente quando falamos com pessoas de diferentes níveis de conhecimento. O projeto também me ensinou a importância de adaptar soluções tecnológicas ao contexto local, como um alfaiate que ajusta uma roupa para que fique perfeita em alguém.

### Evidências do Processo

1. Código Hadoop, Spark e pandas: O arquivo `projeto\_juquitiba.py` contém o código que usamos para processar e analisar os dados. É como o manual de instruções que nos guiou durante todo o processo.
2. Fotografias das Reuniões e Workshops: Os arquivos anexados são registros visuais das nossas interações com a comunidade, mostrando a energia e o engajamento dos participantes.
3. E-mails trocados entre as Partes Envolvidas: O PDF `emails\_troca.pdf` reúne as principais comunicações entre nossa equipe e os representantes da comunidade, documentando a colaboração contínua.
4. Link para o Repositório Git: O link leva ao repositório Git onde todo o código do projeto está armazenado. É como uma biblioteca virtual onde todos podem acessar o que foi feito.

### Conclusão

Este projeto não só documenta a execução e o impacto que tivemos, mas também mostra como a teoria aprendida em sala de aula foi essencial para alcançarmos nossos objetivos. As evidências coletadas destacam a colaboração contínua com a comunidade e a eficácia das tecnologias de Big Data que utilizamos. Espero que essa experiência sirva como inspiração para futuras iniciativas, tanto para educadores quanto para alunos, mostrando que a tecnologia pode realmente fazer a diferença na vida das pessoas.

```
from pyspark.sql import SparkSession
from pyspark.sql.functions import col, avg
import pandas as pd
import pyarrow as pa
import pyarrow.hdfs as hdfs

# Criando uma sessão Spark
spark = SparkSession.builder \
    .appName("Análise Socioeconômica Avançada") \
    .getOrCreate()

# Carregando os dados socioeconômicos do Hadoop (HDFS)
```

```
dados = spark.read.csv("hdfs:///AZZA/T.I/dados_socioeconomicos.csv",
header=True, inferSchema=True)

# Exibindo o esquema dos dados
dados.printSchema()

# Filtrando dados por faixa etária
dados_filtrados = dados.filter((col("idade") >= 18) & (col("idade") <=
65))

# Calculando a média de renda por nível educacional
media_renda =
dados_filtrados.groupBy("nivel_educacional").agg(avg("renda").alias("me
dia_renda"))

# Salvando o resultado intermediário no HDFS
media_renda.write.csv("hdfs:///AZZA/T.I/resultados/media_renda_por_educ
acao.csv", header=True)

# Encerrando a sessão Spark
spark.stop()

# Conectando ao HDFS e carregando o resultado intermediário com pandas
fs = hdfs.HadoopFileSystem("hdfs://seu-nodo-hadoop:8020")
with fs.open("./AZZA/T.I/media_renda_por_educacao.csv/part-00000",
"rb") as f:
    # Carregar o arquivo CSV no pandas
    media_renda_df = pd.read_csv(f)

# Exibindo o DataFrame do pandas
print("DataFrame do pandas:")
print(media_renda_df.head())

# Análise detalhada com pandas
# Exemplo: Encontrando o nível educacional com a maior média de renda
max_renda = media_renda_df.loc[media_renda_df['media_renda'].idxmax()]

print("\nNível educacional com a maior média de renda:")
print(max_renda)

# Salvando o resultado final em um arquivo CSV local
media_renda_df.to_csv("./AZZA/T.I/media_renda_por_educacao_local.csv",
index=False)
```



```
print("\nResultado final salvo em  
./AZZA/T.I/media_renda_por_educacao_local.csv")
```

## Explicação do Script

### **Criação da Sessão Spark:**

- Inicializa uma sessão Spark para processamento distribuído.

### **2. Carregamento dos Dados:**

- Lê um arquivo CSV do HDFS em um DataFrame do PySpark.

### **3. Filtragem e Agregação:**

- Filtra os dados para idades entre 18 e 65 anos.
- Calcula a média de renda por nível educacional.

### **4. Salvamento Intermediário:**

- Salva o resultado da agregação como um arquivo CSV no HDFS.

### **5. Encerramento da Sessão Spark:**

- Finaliza a sessão Spark.

### **6. Leitura com pandas:**

- Usa `pyarrow` para ler o arquivo CSV do HDFS para um DataFrame do pandas.

### **7. Análise Detalhada:**

- Exibe o DataFrame e encontra o nível educacional com a maior média de renda.

### **8. Salvamento Final:**

- Salva o DataFrame final em um arquivo CSV local.

Esse processo combina PySpark para processamento em larga escala e pandas para análises detalhadas.