

DATA THON'S 2025

DIÁRIO DE BORDO

Patrocínio:



Realização:



Boas vindas!

- Neste Diário de Bordo, você vai encontrar os **conteúdos explicativos** e as **ferramentas** para utilizar em cada etapa do DatathONS
- Aqui você também vai registrar todas as atividades realizadas pelo grupo durante o processo.
- Documentar os aprendizados em uma jornada de inovação é crucial para o desenvolvimento de uma solução!

Vamos lá?

 [CLIQUE AQUI](#) PARA ACESSAR
OS MATERIAIS COMPLETOS DO
DATATHONS 2025

Boas práticas



Vocês são
protagonistas!
É importante ter um
líder e ele tomar frente
nas atividades

Dica: Alguém pode
compartilhar
a tela para todos
estarem na mesma
“página”



Trabalhem em equipe!
Todos devem participar e
contribuir para o
desenvolvimento do
projeto

Desafio e Resultados Esperados

Desafio: **Como podemos aplicar inteligência para transformar dados abertos em valor para o setor elétrico?**

Contexto

- Grande volume e complexidade dos dados do Setor Elétrico Brasileiro
- Busca por novas abordagens para torná-los **acessíveis, compreensíveis e úteis** para diferentes empresas e perfis de usuários
- Oportunidade de usar o Portal de Dados Abertos do ONS (<https://dados.ons.org.br>), que reúne informações essenciais para o entendimento da operação do Sistema Interligado Nacional (SIN)

Objetivo

- Gerar valor real a partir dos dados abertos
- Permitir que qualquer pessoa consiga interpretar, comparar e visualizar informações estratégicas com facilidade
- Usar IA em soluções para transformar a experiência de uso do portal, tornando-a mais intuitiva, interativa e produtiva

Resultado Esperado

- Mais transparência, ao aproximar a sociedade dos dados públicos
- Melhores decisões, ao permitir análises rápidas e embasadas
- Maior usabilidade e valor agregado do Portal de Dados Abertos
- Experimentação tecnológica, ao explorar o potencial de inteligência nos dados do setor elétrico

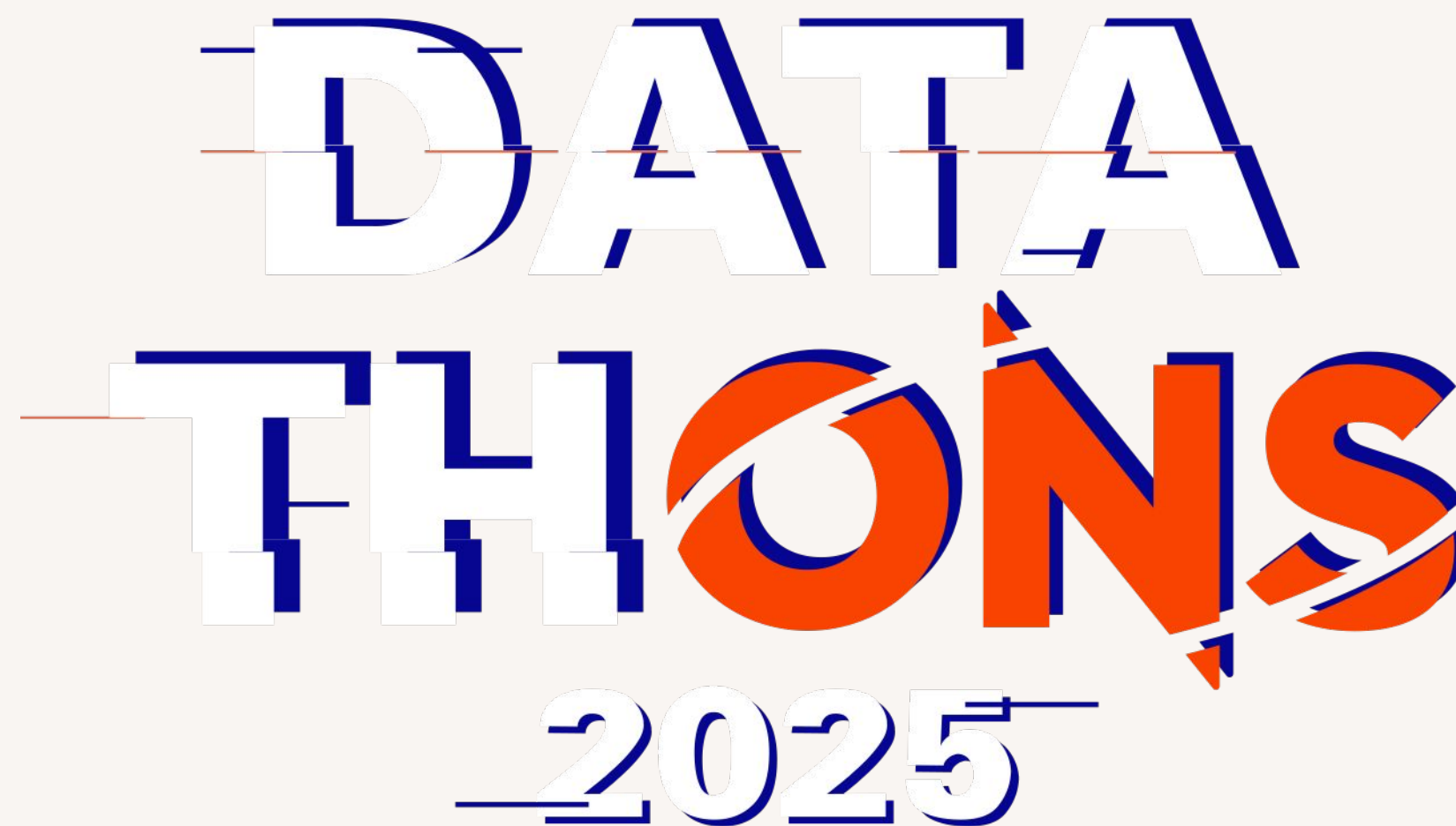
Apresentações de Pitch (30/09)



Critérios de Avaliação

Alinhamento ao Desafio	Inovação	Solução	Impacto	Pitch
o quanto a solução apresentada é capaz de solucionar o desafio proposto?	o quanto a solução se diferencia positivamente de outras que já existem no mercado?	o quanto o produto/processo e a tecnologia são viáveis e factíveis de implementação?	o quanto a solução trará benefícios e impactará o ONS positivamente?	o quanto o pitch foi bem apresentado, didático e com as informações necessárias?

Identidade Visual



#090A59

#F94300

#002246

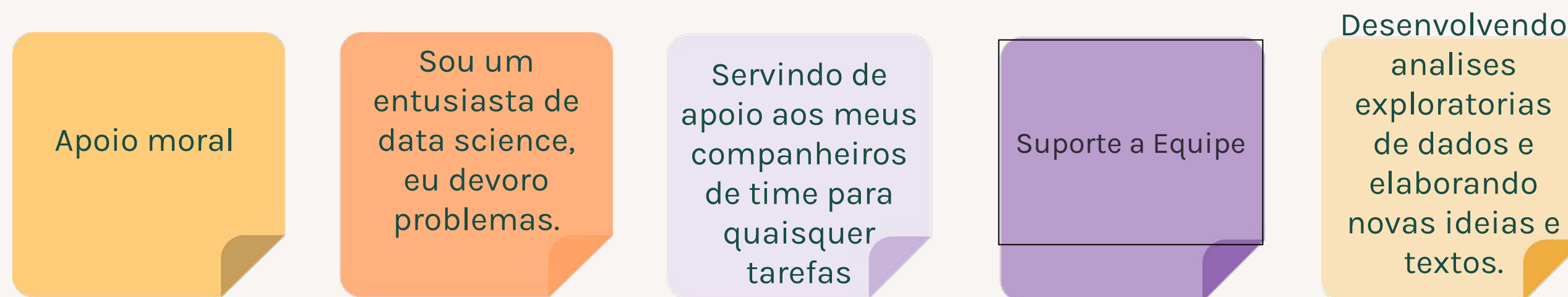
#002A75

**Fonte:
Montserrat**

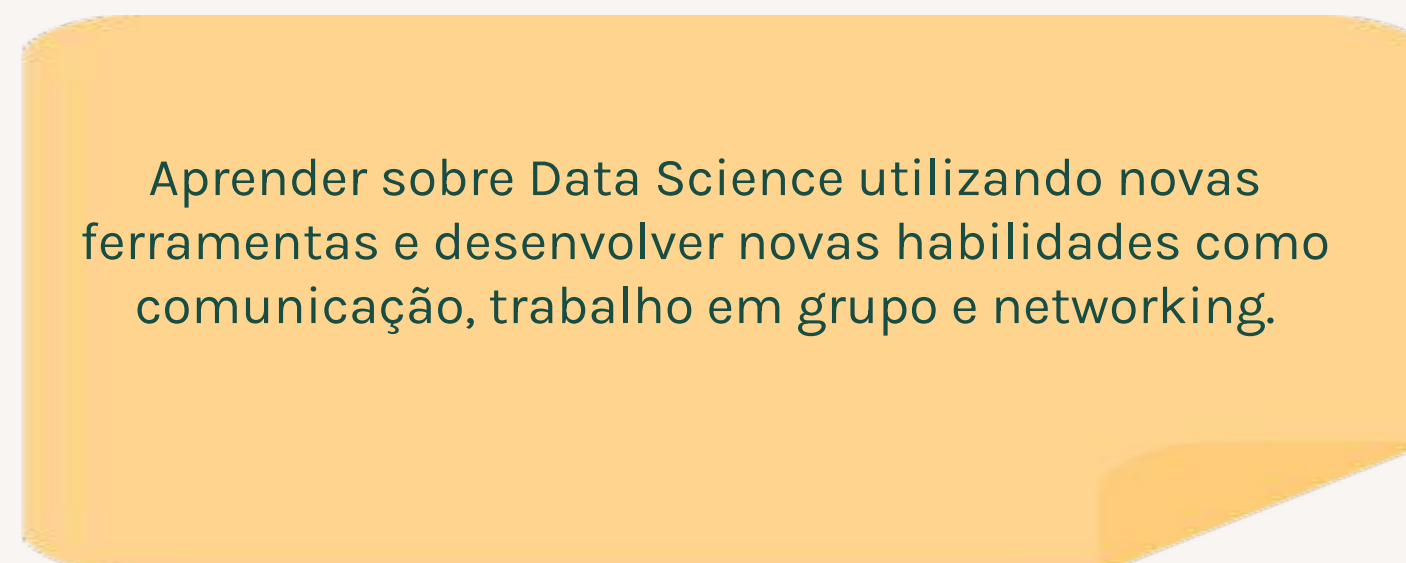
Atividades DatathONS

Grupo 10

Como cada um pode ajudar



Expectativa do Grupo



Dúvidas

Tópico 1

Pode ser algo mais geral ou vocês recomendam que a gente selecione só alguns datasets?

Preencha aqui

Preencha aqui

Tópico 2

Preencha aqui

Preencha aqui

Preencha aqui

Tópico 3

Preencha aqui

Preencha aqui

Preencha aqui

Insights

Front-end

Back-end

Dash

Dash bootstrap

(html, css e
ReactJs)

Python

Plotly, pandas

API do Gemini e
API do ONS

insights

Pensar em
pré-processam
ento da tabela
de dados, para
agilizar o
processo

Investigar a
capacidade de
API do Gemini
(complementar
com código
externo)

Preencha aqui

Levantamento de Problemas

Ponto: converter dados filtrados pro excel

Qual é o problema?

Dificuldade em interagir, analisar e entender os dados.

Por que isso é um problema?

Por que, qualquer indivíduo que queira consultar algum dado, precisa ir no site, procurar, baixar e aí sim começar a entender e explorar.

De quem é o problema?

Entusiastas do setor elétrico que buscam informações mais precisas e embasadas.

Qual é o problema?

A dificuldade para jornalistas, pesquisadores e público no geral de entender e interpretar os dados, que são muito técnicos.

Por que isso é um problema?

Isso gera um vácuo de informação que é frequentemente preenchido por desinformação, suposições e narrativas simplistas.

De quem é o problema?

Jornalistas, pesquisadores, estudantes e público no geral.

Qual é o problema?

Há uma grande quantidade de dados gerados todos os dias, entretanto, não há como interpretar eles na mesma velocidade que são gerados.

Por que isso é um problema?

A lentidão na interpretação dos dados cria uma defasagem perigosa entre o que está acontecendo no sistema elétrico e a capacidade de reação dos agentes.

De quem é o problema?

Operadores do sistema, empresas de geração e transmissão de energia, analistas de mercado e grandes consumidores industriais.

Data Product Canvas

Problema

O público geral e a imprensa têm dificuldade em entender a saúde do sistema elétrico brasileiro, que é vulnerável a secas.

Não há um indicador simples para se o nível dos reservatórios está crítico ou se estamos dependendo excessivamente de usinas termelétricas (mais caras e poluentes), o que impacta diretamente a conta de luz e o risco de racionamento.

Dados

Dados hidráulicos por Reservatório - Base horária

Dados do Balanço de Energia do DESSEM

Hipótese

Se uma IA conseguir interpretar dados de níveis de reservatórios e da matriz de geração de energia, respondendo a perguntas simples em português, então jornalistas poderão criar matérias mais embasadas e o público entenderá melhor os motivos trás de uma bandeira tarifária vermelha ou do aumento do custo de energia.

Solução

Desenvolver uma Assistente de IA que a partir da base de dados da ONS gera as informações solicitadas pelo usuário.

Passo 1.

Problema

- É o ponto de partida. Se o problema não estiver claro, todo o trabalho do time pode se perder. Aqui a equipe define qual problema dentro do desafio vai atacar e por que ele é relevante.

Perguntas que devem ser respondidas:

- Qual é o problema que queremos resolver?

O público geral e a imprensa têm dificuldade em entender a saúde do sistema elétrico brasileiro,

- Por que isso é importante para o ONS e/ou sociedade?

As variações no sistema elétrico brasileiro geram diversas mudanças no valor tarifário, o que não é bem comunicado para a sociedade, o que gera desentendimento.

- Quem é impactado por esse problema?

Toda a sociedade, principalmente o público leigo relacionado ao assunto (a maioria das pessoas, que não lida com eletricidade no dia a dia).

- O que acontece hoje se esse problema não for resolvido?

Eventualmente, de um mês para o outro, a mudança na conta da eletricidade da casa do cidadão comum é intensificada e sem explicação válida.

Passo 2.

Dados

- Quais datasets estão disponíveis para atacar esse problema?

“Dados Hidráulicos por Reservatório - Base horária” e o “Geração de energia verificada por tipo de usina”

- Que variáveis parecem mais úteis?

Dados Hidráulicos por Reservatório:

- niv_armaz_percentual: Mostra o nível do reservatório (o quão cheio está).
- nom_subsistema: Será utilizado para filtrar e comparar por região.
- dta_instante: Será utilizado para analisarmos um período específico

Dados do Balanço de Energia do DESSEM - Geral:

- val_geracao_hidraulica: indicador principal da energia vinda dos reservatórios;
- val_geracao_termica: ajuda a identificar a energia mais cara e poluente que é acionada quando a fonte hídrica não é suficiente.

- Como vamos combinar/explorar os dados disponíveis?

Primeiro, uma IA (Gemini) escolhe o dataset correto para a sua pergunta; em seguida, outra IA (Gemini) analisa os dados coletados para te dar uma resposta em texto.

Passo 3.

Hipótese

O que acreditamos que os dados podem mostrar?

Os dados mostrarão um "raio-X" simplificado e quase em tempo real da saúde do Sistema Interligado Nacional. Eles revelarão a correlação direta entre o nível dos reservatórios e a matriz de geração de energia, permitindo a criação de insights claros para os leitores.

Que relação ou padrão esperamos encontrar?

O padrão central que espera-se encontrar é a **correlação inversa** entre o nível dos reservatórios e a geração de energia termelétrica. Ou seja, espera-se provar com dados que, à medida que o nível de armazenamento hídrico diminui, a geração por usinas térmicas (mais caras e poluentes) aumenta para compensar.

Como essa hipótese Ajuda a resolver o problema?

A hipótese, de que podemos extrair e simplificar automaticamente essa correlação, resolve o problema de duas maneiras principais:

- **Combate a Dificuldade de Interpretação:** Em vez de exigir que um jornalista analise múltiplas planilhas, A IA fará a análise e entrega a conclusão pronta.
- **Reduzir a Defasagem:** A análise será feita sob demanda e em segundos, facilitando a distribuição de informação do setor elétrico..

O que consideramos um resultado de sucesso?

- Um usuário não-técnico conseguir fazer uma pergunta complexa e receber uma resposta clara, baseada em dados e contextualizada de forma rápida.
- O "tempo para obter o insight" (time-to-insight) cai de horas ou dias (análise manual) para segundos, aumentando drasticamente a transparência e combatendo a desinformação sobre o setor elétrico

Passo 4.

Solução

- Aqui entra a ideia que o time vai desenvolver. Não precisa ser 100% funcional, mas deve mostrar como a hipótese gera uma solução aplicável.

Perguntas que devem ser respondidas:

- Qual é a ideia de solução/protótipo que estamos propondo?

Nossa ideia é atacar o problema a partir de um site que nos permite enviar uma solicitação humanizada para obter informações como uma explicação dos datasets e sua visualização clara e direta.

- Como ela usa os dados para atacar o problema?

Os dados são processados utilizando algumas tecnologias como pandas e plotly para facilitar a normalização e a visualização dos dados, além da API do gemini para realizar solicitações de forma intuitiva.

- Como a solução poderia ser aplicada no dia a dia do ONS?

O site para a visualização poderia ser uma aba dedicada dentro do próprio ONS.

- Qual é a saída esperada (dashboard, modelo preditivo, automação, visualização, alerta...)?

A saída esperada é uma resposta no modelo Gemini, que contém uma interação humanizada baseada na pergunta do usuário juntamente a um dashboard com os dados com a visualização facilitada.

Ferramentas:

- **n8n:**
 - https://github.com/n8n-io/n8n?utm_source=chatgpt.com
 - <https://www.youtube.com/watch?v=4cQWJViybAQ>
- **Open WebUI:**
 - <https://docs.openwebui.com/>
- **Notebook LM:**
 - <https://notebooklm.google/>
- **Amazon Q**
 - <https://aws.amazon.com/pt/q/?refid=c7e94da8-23ce-4250-b0c9-fca76d75aa90>
- **Hugging Face:**
 - <https://huggingface.co/>
 - <https://www.youtube.com/watch?v=QEaBAZQCtwE>
- **Modelo Chronos:**
 - <https://huggingface.co/collections/amazon/chronos-models-and-datasets-65f1791d630a8d57cb718444>
- **Gemini CLI**
 - <https://developers.google.com/gemini-code-assist/docs/gemini-cli?hl=pt-br>

DATA THON'S 2025

Obrigado!

Patrocínio:



Realização:

