

The background of the slide is a photograph of a modern building's interior courtyard. The structure is characterized by a series of thick, dark red, angled steel beams that create a complex, geometric framework. The walls and ceiling are composed of large glass panels, which reflect the surrounding environment and the red structural elements. The floor is made of light-colored, rectangular tiles. In the distance, through the glass, a reception area is visible, featuring a dark counter with the Santander logo and some seating. The overall lighting is warm, with a strong red hue dominating the scene.

Santander Data Challenge



O que é um Data Challenge?

Sobre o desafio

O desafio Santander Data Challenge faz uma série de provocações em relação a como podemos ajudar os pequenos empreendedores e a sociedade em geral a superarem essa crise que está sendo gerada pela pandemia. Sabidamente, as empresas no Brasil tem perdido clientes, encontrado dificuldades para se reinventar ou montar estratégias para superar a crise, e os pequenos negócios ou comerciantes autônomos e especial encontram ainda mais dificuldade para entender como reagir a esse cenário. No Data Challenge você terá a oportunidade de utilizar todo o seu conhecimento de dados para analisar em profundidade o que está acontecendo com a saúde, a economia, e os negócios em geral para propor uma solução inovadora com potencial de ajudar nossa sociedade. O resultado do seu trabalho pode ser um produto, serviço, estudo, modelo de machine learning, mas independente da abordagem do seu time, é importante que seja um produto de dados ou construído com sólido respaldo no uso dos dados.

A photograph of a modern building's entrance at dusk or night. The building features a prominent red-tinted glass facade with a grid-like structure. The entrance is a large glass atrium with a red carpet. The text "Quais dados posso utilizar?" is overlaid in white. In the background, a sign with the Santander logo is visible.

Quais dados posso utilizar?

DADOS PÚBLICOS



<https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>

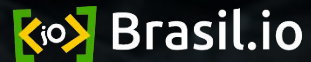


Datasets do sus: <https://opendatasus.saude.gov.br/dataset>

Dados sobre covid do gov: <https://covid.saude.gov.br/>

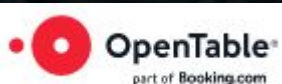
Diversos datasets de São Paulo (epicentro): <http://dados.prefeitura.sp.gov.br/sv/dataset>

Dados de contágio e isolamento social de São Paulo (clique no download do csv): <https://www.seade.gov.br/coronavirus/>



Boletins informativos e casos do coronavírus por município por dia: https://brasil.io/dataset/covid19/caso_full/

Google Trends (dados que indicam principais buscas no google)



Base de dados de reservas de restaurantes nos períodos antes e depois do covid (não tem Brasil, mas pode ser possível pra achar uma relação com outros países com economia parecida)

<https://www.opentable.com/state-of-industry>

Dados do IBGE sobre população: <https://cidades.ibge.gov.br/brasil/sp/panorama>



Outros exemplos de bases públicas:

<https://transparencia.registrocivil.org.br/especial-covid>

<http://www.sptrans.com.br/desenvolvedores/api-do-olho-vivo-guia-de-referencia/documentacao-api/>

https://www.prefeitura.sp.gov.br/cidade/secretarias/transportes/institucional/sptrans/aceso_a_informacao/index.php?p=152415

https://www.tomtom.com/en_gb/traffic-index/recife-traffic

<http://www.fenabreve.org.br/portal/conteudo/emplacamentos>

http://www.antt.gov.br/passageiros/Dados_Operacionais

http://www.antt.gov.br/passageiros/Estatisticas_e_Estudos_de_Servicos_de_Transporte_de_Passageiros.html

<https://www.semanticscholar.org/paper/To-mask-or-not-to-mask%3A-Modeling-the-potential-for-Eikenberry-Mancuso/13802ecd53a008564b81736a9b932b191918bbdb>

https://ai2-semanticscholar-cord-19.s3-us-west-2.amazonaws.com/antiviral_with_properties.sdf.gz

<https://www.semanticscholar.org/cord19/download>

<https://scisight.apps.allenai.org/>

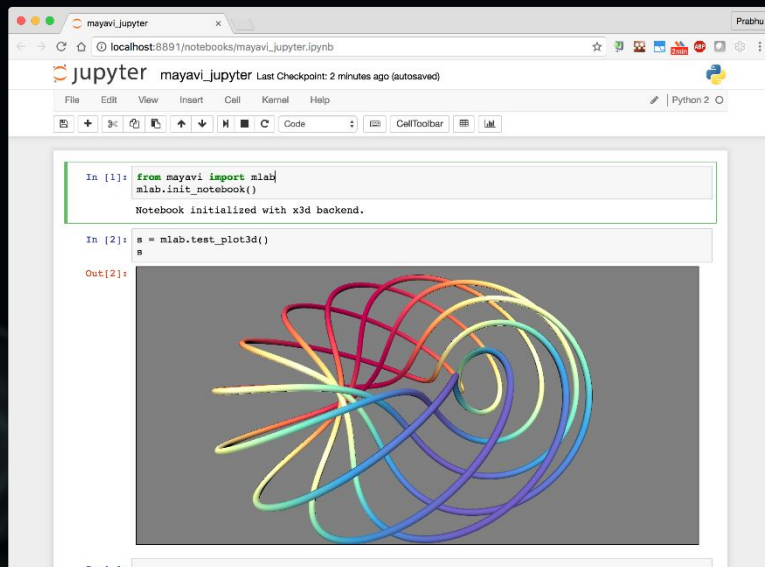
Todos os datasets que utilizar devem ser públicos e acessíveis aos demais participantes, para evitar que times tenham informações e vantagens privilegiadas

A photograph of a modern building's entrance at dusk or dawn. The building features a prominent red-tinted glass facade with a grid-like structure. The entrance is a large glass atrium with a red carpet. The text "Entrega técnica?" is overlaid in white. In the background, a sign for "Santander" is visible.

Entrega técnica?

Entrega

A entrega do Hackathon pressupõe uma entrega mais técnica, que deve ser realizada na forma de notebook jupyter (.ipynb)



O time deve apresentar um notebook bem estruturado com toda a parte de dados desenvolvida pela equipe contendo os seguinte tópicos separados em células:

- Análise exploratória (gráficos e estatísticas válidas)
- Data preparation / ETL (transformações pertinentes)
- Modelagem / Estudos (Machine Learning ou análises)
- Validação dos resultados (teste, avaliação do modelo, testes de hipótese, etc.)

O notebook gerado deve ser salvo em um projeto único do GIT de um dos participantes e o link anexo à apresentação.