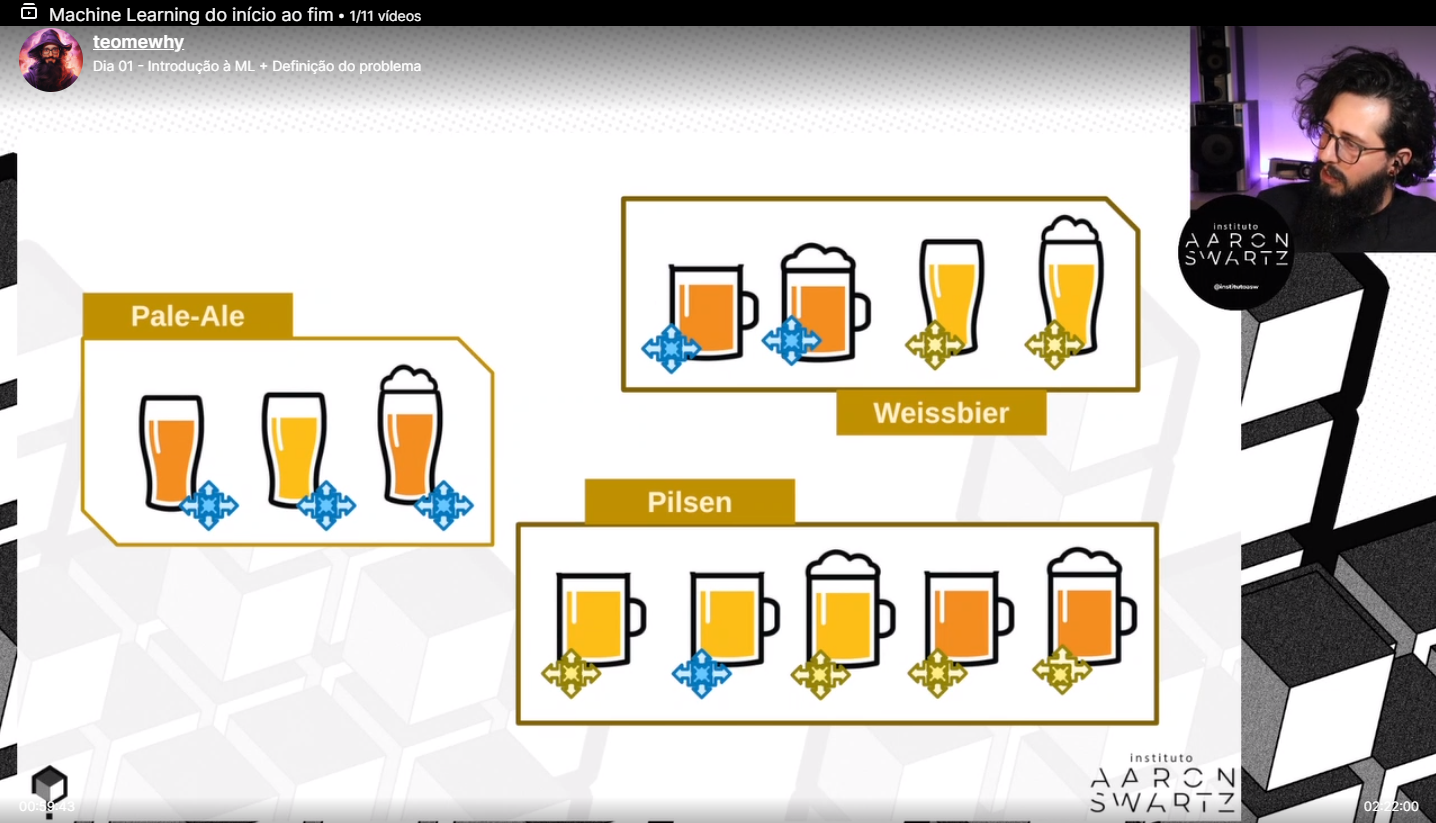
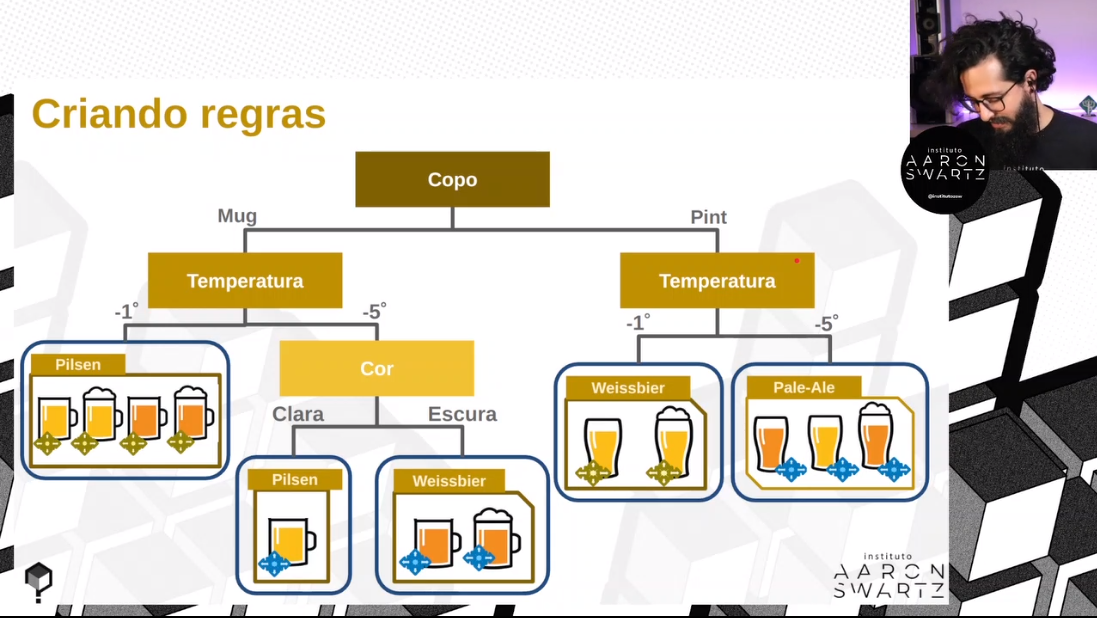
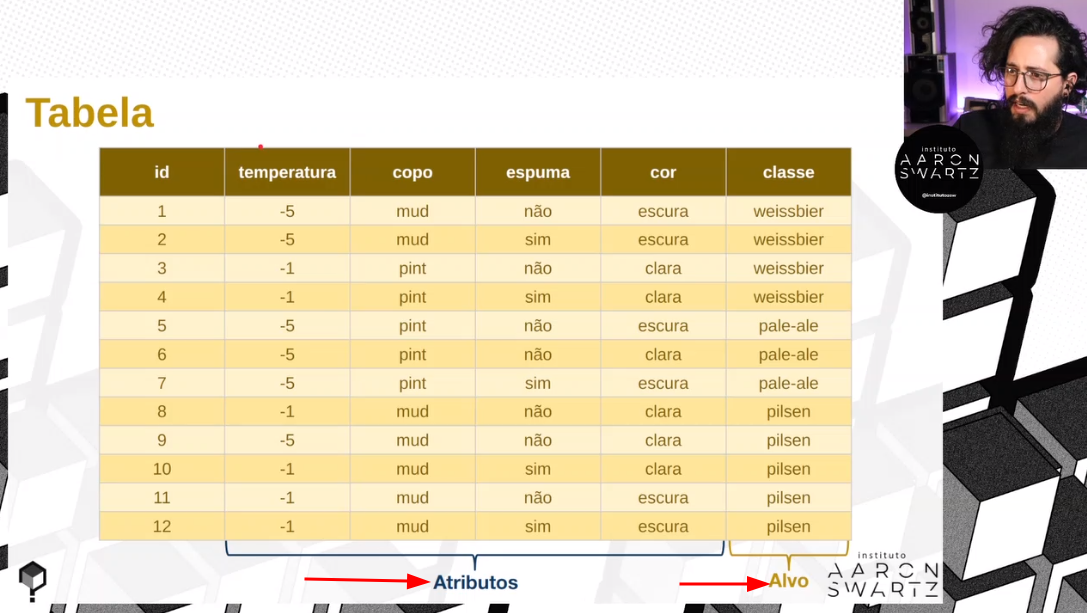
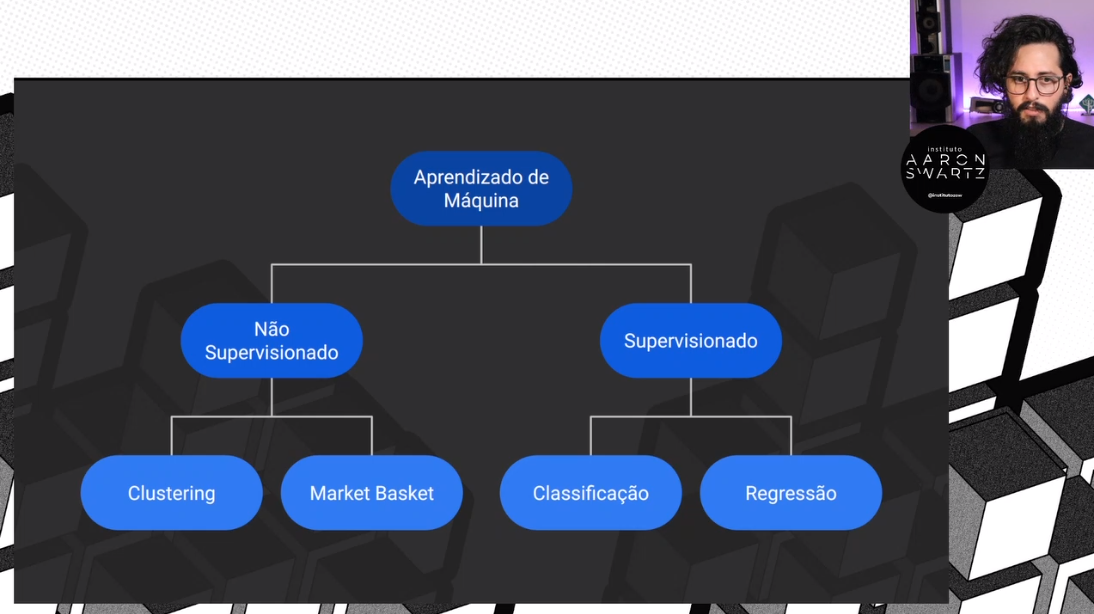
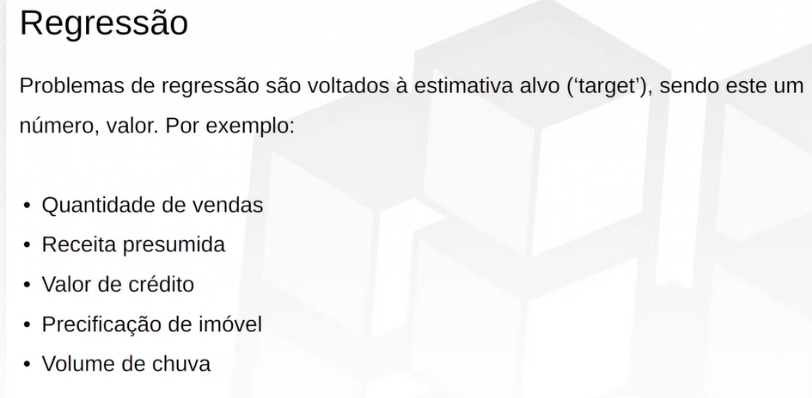
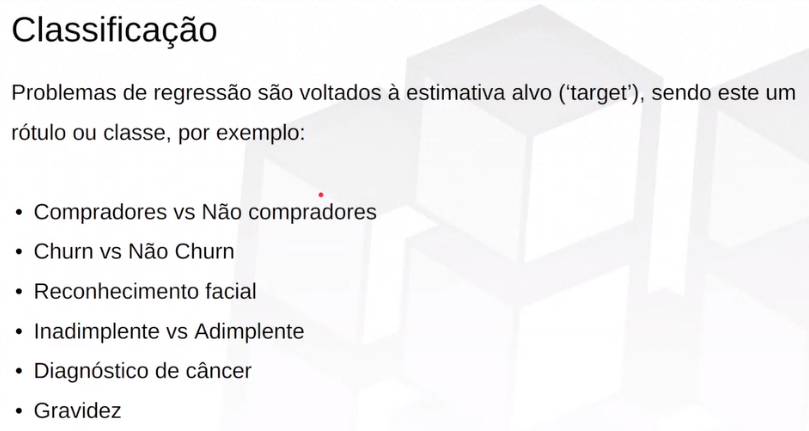
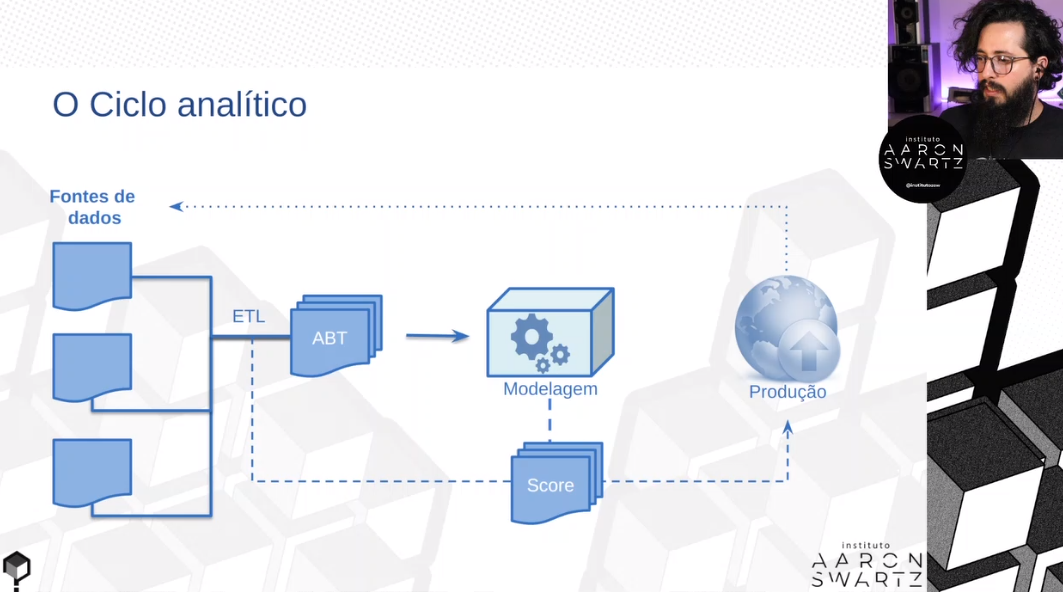
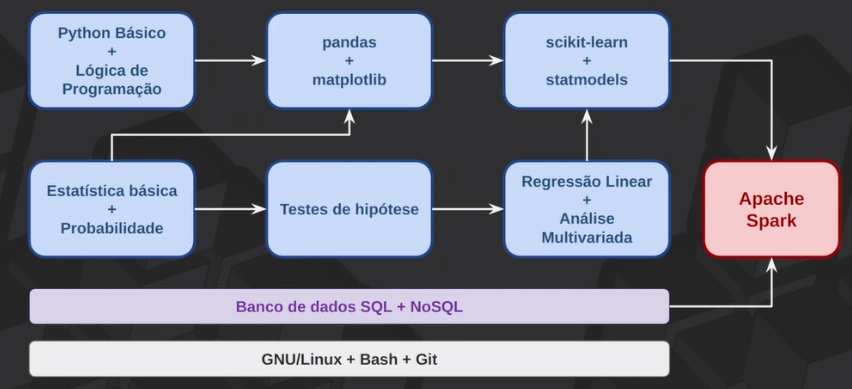
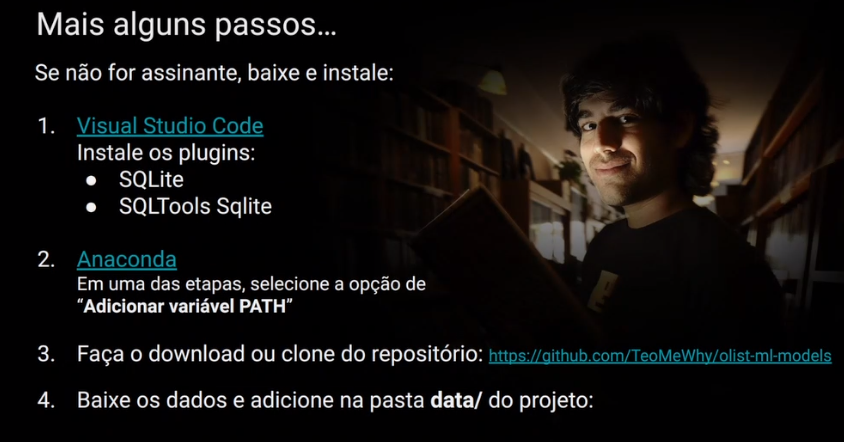
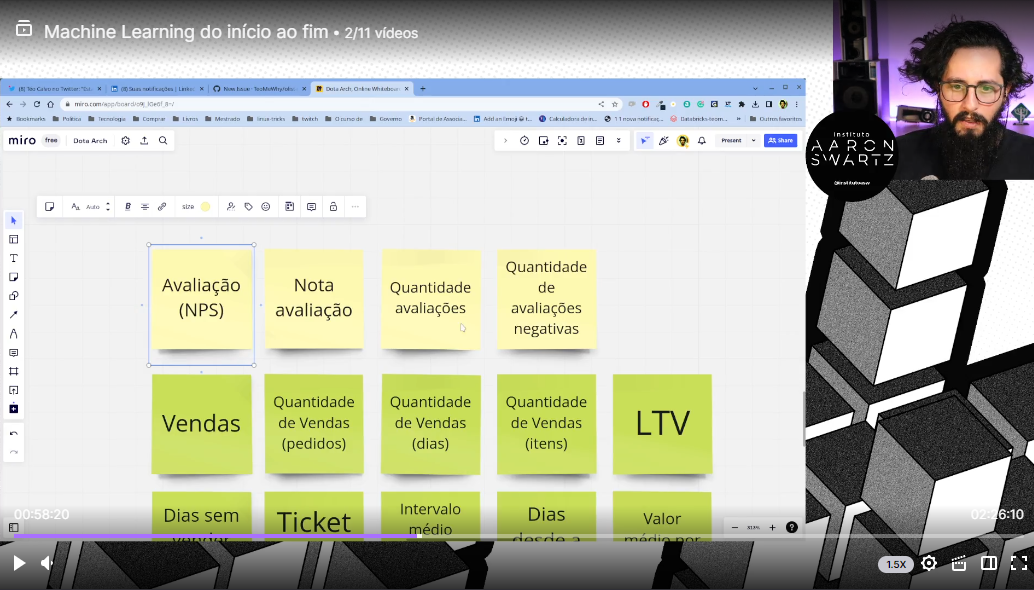
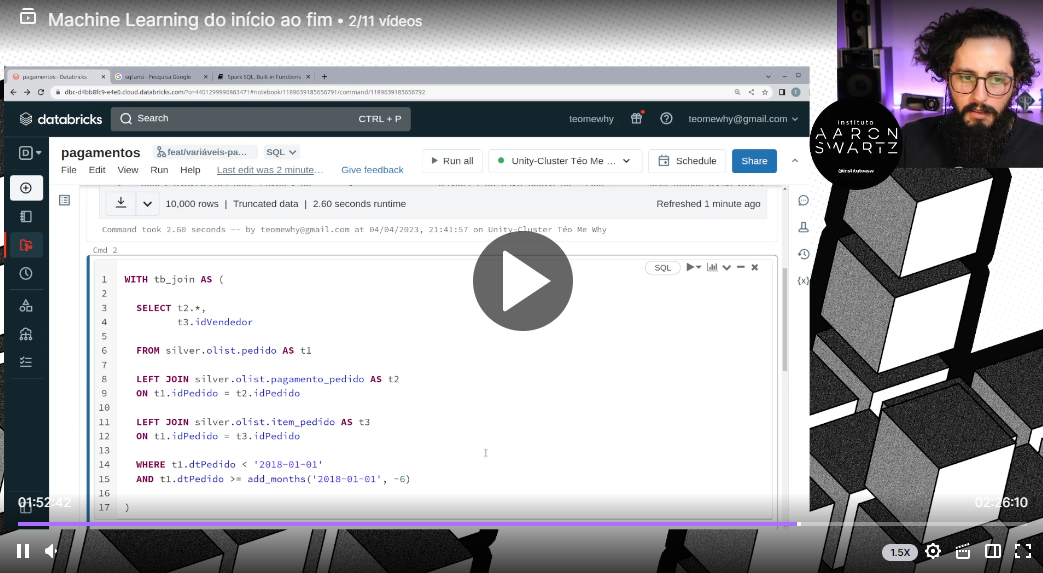
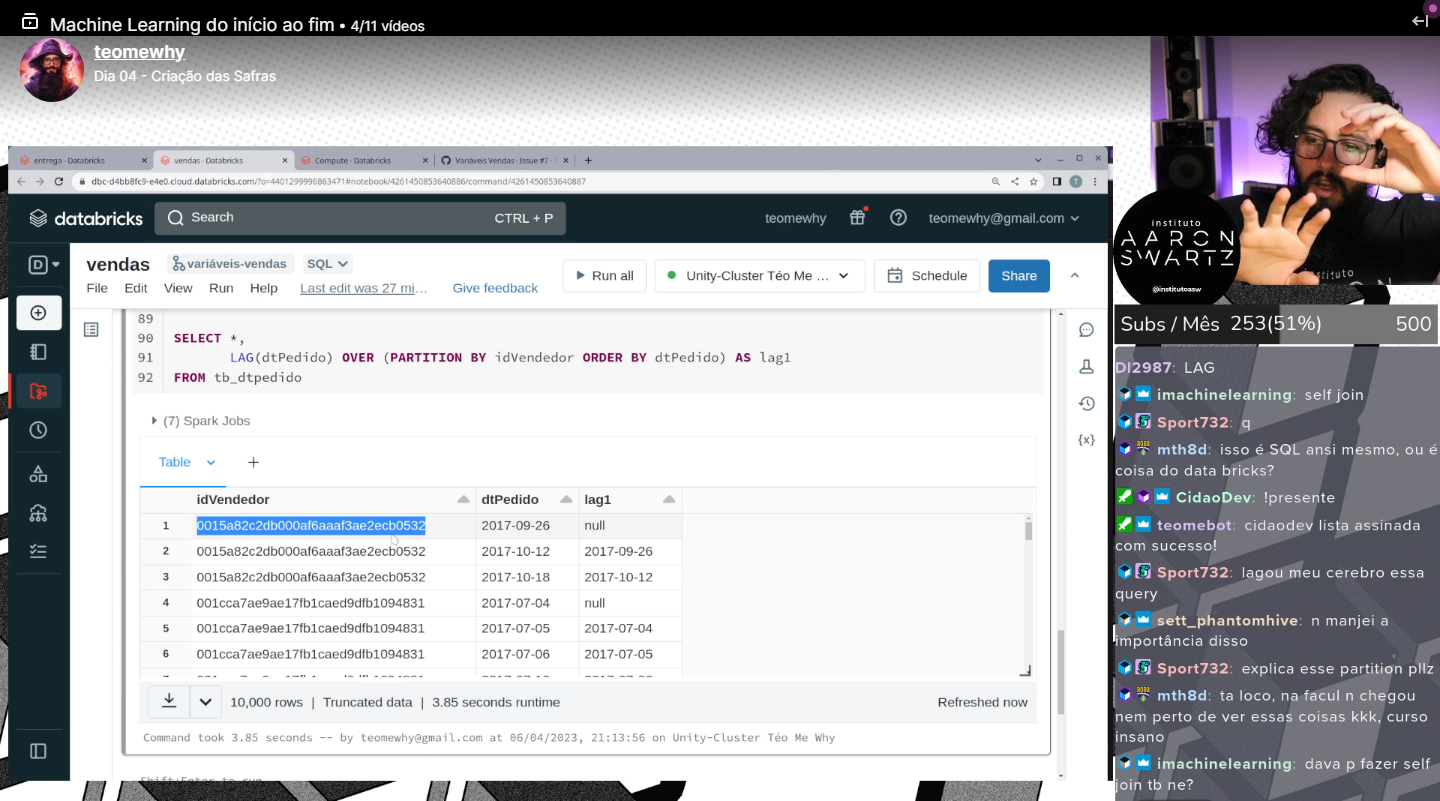
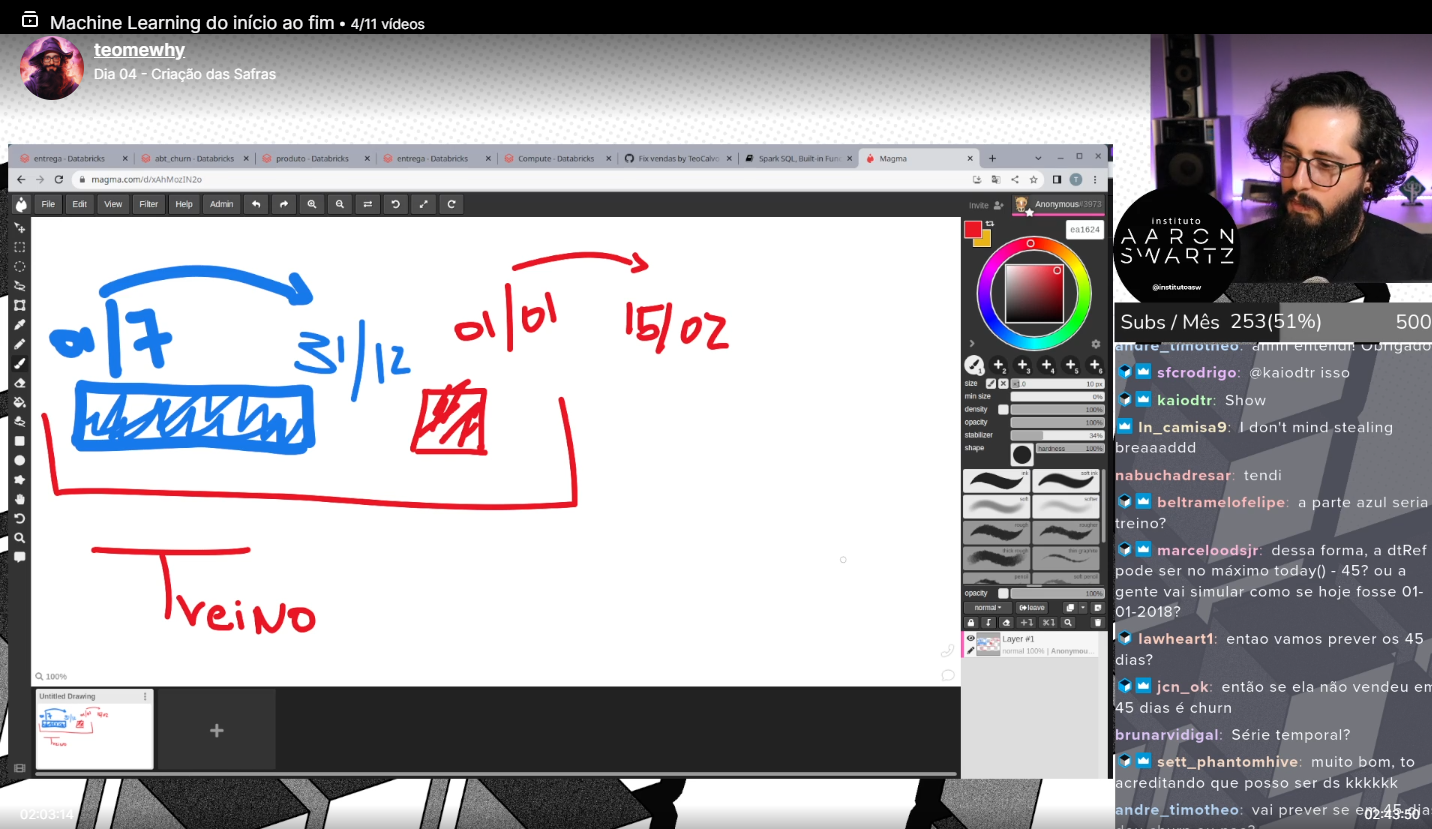
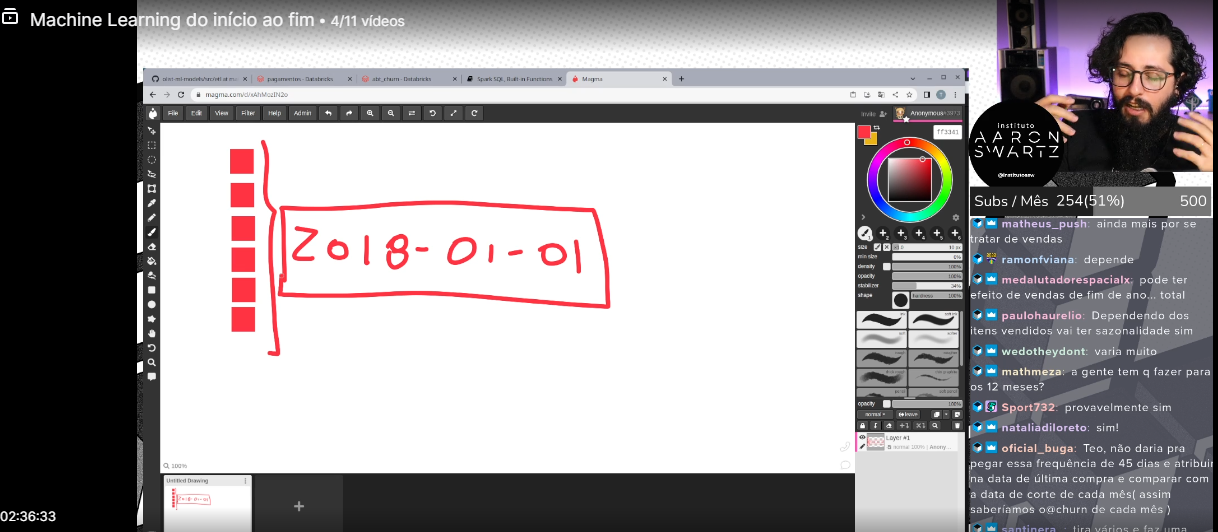
DIA 1  
Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente  
  


Conceito de arvore de decisões. Utilizando estrutura de IF ou CASE das linguagens, classifica determinado dado com base nas suas características quando verdadeiras.  
  
  
  
  
  
  
  
  
  
  
  
  
Trilha de aprendizado:  
  
  
  
  
  
DIA 2  
\*Brainstorm das ideias de quais métricas (variáveis/ feature store) são relevantes para o estudo de ML  
   
  
\*Daqui ele começa a criar as queries em SQL para atender todas as métricas (variáveis) que ele precisa para o estudo de ML.  


DIA 3  
No início do dia 3 o Teo reportou que a query desenvolvida na aula anterior estava incorreta, pois houve uma duplicação de linhas no cruzamento -> Pedido x Item Pedido x Vendedor (se houve mais de um vendedor por pedido, irá duplicadas as linhas de itens tantos quantos forem os vendedores).

DIA 4  
Seguiu no desenvolvimento das queries para formar as variáveis.  
  
Parte importante da aula em que ele utiliza o comando LAG com base no idVendedor, que cria uma nova coluna deslocando uma linha abaixo (em função do vendedor).  
  
  
Todas as queries feitas anteriormente foram na base 01/07 a 31/12.  
Agora ele gerou uma menor base 45 dias pós esse período (01/01 a 15/02) para assim criar a variável resposta. Os primeiros dados serão usados para tentar prever esses 45 dias.  
  
  
Um detalhe importante que ele notou é que existiam vários vendedores “churnados”, pois a data de recência já era maior que 45 dias, portanto, enviesava o modelo e os mesmos foram retirados (parte mais confusa).  
Após isso juntos todas as queries em uma somente.  
  
Aqui o Téo explica que tudo foi feito pra uma “safra” somente e teremos de replicar para demais.  
 

DIA 5  
RESUMO DA ABT: Tabela que contém as características e a variável resposta, que temporalmente é ideal que estejam separadas.  
Criou todos os scripts em python para gerar as tabelas automaticamente no banco de forma automatizada trocando as variáveis de data e afins.  
  
DIA 6  
