



Capítulo 5/6

Projeto do Curso

# Parte 1. Elaboração de Requisitos Técnicos

Você trabalha como analista de anúncios em vídeo na agência de publicidade Sterling & Draper. Você dedica muito tempo analisando vídeos de tendências no YouTube para determinar qual conteúdo merece atenção de marketing.

Cada vídeo tem uma categoria específica (entretenimento, música, notícias e política etc.), região e data de tendência.

Um vídeo pode estar na seção de tendências por vários dias seguidos.

Toda semana, os novos funcionários Melanie e Ashok fazem as mesmas perguntas:

- Quais categorias de vídeo foram tendências na semana passada?
- Como foram distribuídos entre as várias regiões?
- Quais categorias foram especialmente populares nos Estados Unidos?

Em sua sexta semana de trabalho, você decide que está na hora de automatizar o processo. Você vai fazer um dashboard para Melanie e Ashok.

Pergunta

Quais etapas precisam ser tomadas para projetar e construir o dashboard?



Converse com os administradores de banco de dados e descubra quais dados eles realmente coletam. Em seguida, trabalhe com os engenheiros de dados para projetar tabelas agregadas e a estrutura do pipeline e espere que eles desenvolvam esse pipeline. Em seguida, construa o

dashboard. Finalmente, anuncie solenemente que Melanie e Ashok agora podem estudar os dados por conta própria.

- ☐ Você precisa desenvolver o código do dashboard imediatamente. Ele deve aceitar dados brutos diretamente. Nenhuma tabela agregada ou pipelines. Vamos poupar o tempo dos engenheiros de dados. Quando o dashboard estiver pronto, avise Melanie e Ashok.
- ☐ Você precisa conversar com Melanie e Ashok e descobrir onde obter os dados e em quais bancos de dados armazenar as informações agregadas. Você cuidará do design dos elementos do dashboard. Não há necessidade de discutir nada com os administradores de banco de dados e engenheiros de dados, eles estão fora de contato com problemas analíticos.

Você precisa discutir o conteúdo do dashboard, seu layout e os dados que precisam ser exibidos com Melanie e Ashok. Em seguida, converse com os administradores de banco de dados e engenheiros de dados e descubra onde e como os dados necessários são coletados e como podem ser transformados. Não se esqueça de perguntar onde armazenar as tabelas agregadas. Por fim, desenvolva o pipeline e o dashboard.

Isso mesmo! Primeiro, discutimos o rascunho do dashboard e os requisitos técnicos com os gerentes. Então vamos aos administradores. O código vem bem no final.


✓ **Você conseguiu!**

Depois de conversar com os gerentes e administradores de banco de dados, você elaborou breves requisitos técnicos:

- Objetivo dos negócios: analisar o histórico de vídeos de tendências no YouTube
- Com que frequência o dashboard será usado: pelo menos uma vez por dia

- Usuário do dashboard de destino: gerentes de planejamento de anúncios em vídeo
- Conteúdo de dados do dashboard:
  - Vídeos de tendências do passado, divididos por dia e categoria
  - Vídeos de tendências, divididos por países
  - Uma tabela de correspondência entre categorias e países
- Parâmetros segundo os quais os dados devem ser agrupados:
  - Data e hora da tendência
  - Categoria de vídeo
  - País
- Os dados:
  - Histórico de tendências – valores absolutos com divisão por dia (dois gráficos: números absolutos e proporção percentual)
  - Eventos, discriminados por países – valores relativos (% de eventos)
  - A correspondência entre as categorias e os países – valores absolutos (uma tabela)
- Importância: todos os gráficos são igualmente importantes
- Fontes de dados para o dashboard: os engenheiros de dados prometeram criar uma tabela agregada chamada `trending_by_time`. Segue sua estrutura:
  - `record_id` – chave primária
  - `region` – país / região geográfica
  - `trending_date` – data e hora
  - `category_title` – a categoria de vídeo
  - `videos_count` – o número de vídeos na seção de tendências
- A tabela fica armazenada no banco de dados do `youtube`, criado especialmente para suas necessidades
- Intervalo de atualização de dados: uma vez a cada 24 horas, à meia-noite UTC
- Gráficos, controles do dashboard e sua disposição:

Título e descrição do dashboard	Filtro de data e hora	Filtro de país
Gráfico "Histórico de tendências" - vídeos de tendências divididos por tempo e categoria de tendências (valores absolutos, gráfico de áreas empilhadas)	Gráfico "Vídeos em alta por país" - vídeos em alta divididos por país (valores relativos, gráfico de pizza)	
Gráfico "Histórico de tendências" - vídeos de tendências divididos por tempo e categoria de tendências (% do total, gráfico de áreas empilhadas)	Tabela "Vídeos em alta por país e categoria". As colunas devem representar os países, as linhas devem indicar as categorias de tendências. As células da tabela devem conter números absolutos de vídeos em alta. As células da tabela devem ser destacadas em relação aos seus valores (tabela de realce).	

 Avaliar