

Incorporação dos log de mudanças

O que engenheiros, autores e analistas de dados têm em comum? A mudança.

Os engenheiros usam Ordens de Alteração de Engenharia (ECO, na sigla em inglês) para acompanhar os detalhes de projeto do novo produto e as mudanças propostas aos produtos atuais. Os autores usam históricos de revisão do documento para acompanhar as mudanças nas edições e fluxo do documento. Os analistas de dados, por sua vez, usam log de mudanças para acompanhar a limpeza e a transformação dos dados.

O CONTROLE AUTOMATIZADO DE VERSÃO ORIENTA VOCÊ EM GRANDE PARTE DO CAMINHO

A maioria dos aplicativos de software contam com um tipo de monitoramento de histórico integrado. No Planilhas Google, por exemplo, é possível verificar o histórico de versão da planilha toda ou de uma única célula, bem como retornar a uma versão anterior. No Microsoft Excel, você pode usar um recurso chamado Controlar alterações. Já no BigQuery, você tem acesso às mudanças feitas pelo histórico.

Veja como é na prática:

Planilhas Google	1. Clique com o botão direito na célula e selecione "Exibir histórico de edição". 2. Clique nas setas para a esquerda (<) ou direita (>) para ir para frente ou para trás no histórico, conforme necessário.
Microsoft Excel	1. Se o recurso de Controlar Alterações estiver habilitado na planilha: clique em "Revisão". 2. Em Controlar alterações, clique na opção "Aceitar/Rejeitar alterações" para aceitar ou rejeitar as mudanças feitas.
BigQuery	Abra uma versão anterior (sem voltar para ela) e compare-a à versão atual para ver o que mudou.

OS LOGS DE MUDANÇAS VÃO A FUNDO COM VOCÊ

O log de mudanças pode aproveitar seu histórico automatizado de versão, dando a você um registro mais detalhado do seu trabalho. É nele que os analistas de dados registram todas as alterações feitas nos dados. Veja de outro ângulo. Os históricos de versão registram o que foi feito na alteração de um dado do projeto, mas não indicam o porquê. Os logs de mudanças são mais do que úteis para nos ajudar a entender os motivos por trás das alterações. Eles não têm um formato definido; se você quiser, pode até mesmo inserir suas informações em um documento em branco. Se, no entanto, estiver usando um log de mudanças compartilhado, o melhor é definir com outros analistas de dados o formato de todas as entradas no registro.

Normalmente, o log de mudanças registra o seguinte tipo de informação:

- Dados, arquivo, fórmula, consulta ou outro componente modificado
- Descrição do que foi alterado
- Data da alteração
- Responsável pela mudança
- Responsável pela aprovação da mudança
- Número de versão
- Justificativa para a alteração

Digamos que você alterou uma fórmula de uma planilha, pois viu em outro relatório e quis que seus dados fossem coerentes e consistentes. Se, em outro momento, você descobrisse que o relatório usava na verdade a fórmula errada, o histórico automatizado de versão poderia ajudar você a desfazer a alteração. Mas, se também registrou a justificativa para a alteração em um log de mudanças, você poderia informar os criadores do relatório quanto à fórmula incorreta. Se a mudança foi feita há um tempo atrás, pode ser que não se lembre a quem se dirigir. Felizmente, seu log de mudanças poderia dar essa informação para você! Ao dar sequência, você garantiria a integridade dos dados fora do projeto e demonstraria sua integridade pessoal como alguém em que pode ser confiado com dados. Esse é o potencial de um log de mudanças!

Por fim, o log de mudanças é importante no caso de várias alterações terem sido feitas em uma planilha ou consulta. Imagine a situação: um analista fez quatro mudanças e ele quer reverter a mudança nº 2. Ao invés de clicar três vezes no recurso para desfazer a mudança nº 2 (e acabar perdendo as mudanças nº 3 e 4), o analista para desfazer apenas a mudança nº 2 e manter todas as demais. Nosso exemplo foi de apenas quatro alterações, mas pense na importância desse log de mudanças se houvessem centenas de alterações para controlar.

O QUE TAMBÉM ACONTECE NA PRÁTICA



Provavelmente, o analista júnior só precisa saber o acima com uma exceção. Se o analista está fazendo alterações a uma consulta SQL que já existe e é compartilhada na empresa, muito provavelmente ela usa o que chamamos de sistema de controle de versão, como uma consulta que extrai a receita diária para criar um painel para a gerência sênior.

Veja como um sistema de controle de versão afeta uma alteração em uma consulta:

1. Uma empresa conta com versões oficiais de consultas importantes em seu sistema de controle de versão.
2. O analista verifica se a versão mais recente da consulta é aquela que será modificada, o que chamamos de sincronização.
3. O analista faz uma alteração na consulta.
4. O analista pode pedir que outra pessoa revise a alteração. A isso damos o nome de revisão de código, que pode ser realizado informal ou formalmente. Uma revisão informal pode ser tão simples quanto pedir a um analista sênior para dar uma olhada na alteração.
5. Assim que o revisor dá sua aprovação, o analista envia a versão atualizada da consulta a um repositório no sistema de controle de versão da empresa, o que chamamos de alocação de código. Uma prática recomendada é documentar exatamente qual mudança foi feita e por que na área de comentários. Voltando ao nosso exemplo de uma consulta que extrai a receita diária, vejamos um possível comentário: Receita atualizada de modo a incluir a receita oriunda do novo produto, Calypso.
6. Após o envio da alteração, todos da empresa poderão acessar e usar a nova consulta ao sincronizar com as consultas mais recentes armazenadas no sistema de controle de versão.
7. Se a consulta apresenta um problema ou as necessidades da empresa sofrem uma mudança, o analista pode desfazer a alteração da consulta pelo sistema de controle de versão. O analista pode conferir a lista cronológica de todas as alterações feitas na consulta e o autor de cada mudança. Após encontrar sua própria alteração, o analista pode reverter para a versão anterior.
8. A consulta volta ao que era antes do analista fazer a alteração, e todos da empresa conseguem ter acesso a essa mesma consulta também.