

6.8 - Elaborando a página de análise por gênero

00:00:00:00 - 00:00:01:00

Olá, pessoal. Bom, na última aula, nós fizemos a página resumo do nosso dashboard. Colocamos algumas informações importantes e que já apresentavam, de certa forma, o que a gente gostaria de aprofundar nesse dash, que, no caso, é a análise de diversidade por cima dos dados da pesquisa do Data Hackers com os profissionais da área de dados. Na aula de hoje, nós vamos criar uma outra página, que vai ser uma página do tipo detalhada, que tem como objetivo aprofundar em alguma área específica da análise. Lá no nosso rascunho, a gente já fez um rascunho do que a gente está pensando em colocar nessa página. A gente pode voltar aqui para ele e ver quais são os gráficos que vamos colocar. Então aqui vamos ver as páginas, a segunda página... Agora sim, com o rascunho aberto podemos ir para o Looker e aí a gente vai revisando o que a gente vai colocar nessa página detalhada. Então vamos lá. A primeira coisa é criar uma nova página. Nós fazemos isso clicando em adicionar página.

00:00:01:00 - 00:00:02:01

Aí já adicionou uma segunda página. Bom, se a gente clicar aqui no cantinho, a gente vai ver que está tudo sem título. A gente pode aproveitar esse momento para colocar os títulos. A gente vem aqui em Página e Gerenciar Páginas. Nada aconteceu porque a gente ocultou aqui aquela barra lateral. Vamos mostrar os painéis aqui. Agora sim apareceu as páginas do nosso relatório. Aí está aqui que a primeira página está sem título e a segunda também. A gente clica aqui na primeira, que é o nosso resumo, vamos clicar nesses três pontinhos aqui e ir em renomear. Aí vamos colocar aqui tipo "página resumo" mesmo, porque a gente é muito criativo e tal. E a segunda página já vamos aproveitar também e colocar o nome dela, tipo "análise de gênero", né, que a gente combinou lá no rascunho. Então beleza, a gente já nomeou.

00:00:02:01 - 00:00:03:04

Agora, voltando para essa página que a gente criou, ela está em branco. Então, vamos começar criando um gráfico de gênero e escolaridade. Assim, podemos apresentar se, por exemplo, pessoas do gênero masculino têm nível de escolaridade maior que as do gênero feminino. Nós vimos em aulas anteriores que tem vários tipos de gráficos, com seus prós e contras. Nesse caso aqui, como queremos uma comparação entre os gêneros para cada categoria de escolaridade, nós podemos fazer um gráfico de barras empilhado na horizontal. Isso facilita a gente a facilitar a proporção de gênero por cada nível. Outra possibilidade seria a gente usar gráfico de pizza, mas nesse caso precisaríamos de um gráfico para cada escolaridade, ou um gráfico para cada gênero, o que dificultaria a comparação porque a gente teria que ficar comparando entre dois gráficos, e a gente quer concentrar tudo num gráfico só. Então, a gente vai lá em adicionar gráfico e depois clica em adicionar gráficos de barra empilhados na horizontal. Vamos lá. Adicionar gráfico. Aqui tem os gráficos empilhados. A gente não quer na vertical, a gente quer na horizontal.

00:00:03:04 - 00:00:04:04

Então, a gente clica aqui na horizontal. E... A gente coloca em algum lugar, não vamos pensar em questões estéticas ainda. Eu vou só dar um zoom aqui na página, colocar um pouquinho mais próxima para a gente conseguir ver o gráfico. Ok. Bom, depois que a gente colocou aqui o gráfico, a gente vai trabalhar naquelas questões de configuração, que são as dimensões. Para dimensão mesmo, aqui a primeira, sem ser a detalhada, a gente pode colocar, não sei, nível de ensino. É mais para a questão de que a gente não sabe exatamente como vai ficar. Aí se não ficar legal, a gente inverte, tá? A gente já fez isso em outra aula. Mas vou colocar aqui, por exemplo, nível de ensino. E aí na dimensão detalhada eu vou colocar gênero. Vamos ver se vai ficar bom. Bom, eu acho que já ficou bacana, ficou legal.

00:00:04:04 - 00:00:05:01

Só que os rótulos, vamos dar uma olhada se os rótulos estão corretos também, porque os rótulos precisam estar certinhos. A métrica, vamos ver se está certo, vamos colocar o ID aqui, porque a gente quer a quantidade de pessoas, ou seja, os IDs únicos. A gente vai aqui, vamos fazer a contagem. Beleza, certinho. Então, o que a gente tem aqui? No eixo Y a gente tem aqui as categorias: graduação, pós-graduação, estudante, etc e tal. E aqui no eixo X a gente tem as quantidades. O roxo representa o masculino e o azulzinho aqui representa o feminino. E tem uma camadinha aqui de uma cor indeterminada, deixa eu ver se a produção me fala, é um azul meio com roxo, que é a parte de prefiro não informar. E aí se a gente passar aqui por cima o mouse, a gente consegue ver as quantidades e tudo mais.

00:00:05:01 - 00:00:06:07

Ok. Mas gente, olhando assim, não está deixando claro pra gente qual que é a proporção mesmo de gênero dentro de cada categoria de escolaridade. Por exemplo, em doutorado, que temos uma menor quantidade de pessoas no geral, fica difícil saber se o percentual de pessoas do gênero feminino é maior do que o percentual de pessoas do gênero feminino de mestrado. Para melhorar essa visualização, podemos usar o gráfico empilhado com porcentual. Aí, para poder trocar, a gente clica aqui no gráfico, aí a gente clica aqui em gráfico, nessa setinha para baixo. Do lado vai ter: gráfico de barras 100% empilhadas, esse 100% indicando que vai ser porcentual. Vamos clicar aqui nele e vai atualizar. Ok, já atualizou para porcentagem. Então, quando a gente passa o mouse aqui, a gente vê o número, a quantidade, mas entre parênteses a gente consegue ver a porcentagem também. Já fica um pouco melhor. Mas para melhorar ainda mais, a gente pode colocar esse número de porcentagem aqui no gráfico.

00:00:06:07 - 00:00:07:05

Para fazer isso, a gente pode ir na barra de estilo, de configurações, estilo e a gente desce aqui um pouco e marca essa opção de mostrar rótulos de dados. E aí já vai aparecer os numerozinhos. Já ficou bacana. Não é o percentual, mas a gente consegue ver, comparar a quantidade de mulheres na graduação com mulheres na pós-graduação. A gente vê que uma é 365, a outra a quantidade é 370. Bom, gente, uma coisa também é que não está dando para a gente poder ver essa lateral aqui completa, né?

Se a gente passar o mouse aqui por dentro e quando chegar nessa partezinha aqui, que é bem o início do nosso gráfico, o mouse vai mudar, vai ficar uma setinha para um lado e uma para o outro, e a gente consegue arrastar até aparecer os nomes completos, beleza? E uma coisa interessante que a gente pode ver aqui é que, apesar do percentual de mulheres não ser maior, assim, grande, a gente consegue ver que, por exemplo, em pós-graduação e mestrado, a gente tem uma grande quantidade de mulheres.

00:00:07:05 - 00:00:08:30

Elas meio que continuam buscando conhecimento depois da graduação. A gente tem uma quantidade alta, não reduz depois da graduação. Bom, contudo, a gente também pode concluir que muitas pessoas na área de dados, de forma geral, têm alguma formação. Não podemos afirmar que a formação é na área de computação, mas podemos afirmar que alguma formação tem, né? Bom, antes de a gente ir para o próximo gráfico, eu só vou colocar o título, né? Ficou faltando aqui. Vamos lá, clicar no gráfico e em estilo. Clicar em mostrar título. E esse gráfico é um gráfico... Eita. Gráfico de gênero e escolaridade. Perfeito. Depois a gente pensa na estética. Vou arredar aqui um pouquinho para o lado. E, bom, vamos fazer um gráfico agora de gênero por senioridade, para a gente conseguir ver se as pessoas do gênero masculino estão em cargos mais altos do que as pessoas do gênero feminino. Vamos fazer outro gráfico de barras, mas dessa vez vertical e sem ser empilhado, para a gente aprender se outras visualizações ficariam boas para esses tipos de dados. Bom, vamos lá. Adicionar gráfico. E gráfico de barras, sem ser empilhado.

00:00:08:30 - 00:00:10:07

Vamos colocar ele aqui e vamos nessa aba de configuração trabalhar com as nossas dimensões. Aqui em dimensão mesmo, a gente pode colocar aquela coluna de novo nível que a gente já criou numa aula de Featuring Engineering, não sei se vocês lembram. Então, vamos lá, novo nível. Na dimensão detalhada, o que a gente está comparando? Gênero. Então, vamos colocar aqui: gênero. Beleza. Na métrica, vamos colocar o ID e vamos fazer a contagem. Perfeito, vamos aumentar aqui um pouquinho. Bom, vamos aqui na barra de estilo para poder colocar também um título, né? Vamos colocar aqui: mostrar título e gráfico gênero por senioridade. Beleza. Observando agora os rótulos do eixo X, a gente consegue ver que nas últimas barrinhas não tem rótulo, né? Se a gente passar aqui o cursor nessa barra, a gente vê que esses dados são dados nulos. Lembram da aula em que falamos sobre valores faltantes e a importância de aprender a tratar esses valores? Quando temos muitos nulos, eles podem influenciar de forma indevida nas nossas análises. Nós vimos que podemos deletar os nulos ou substituir por outros valores, mas em alguns casos o melhor é não tratar, por falta de informação para substituir. Então, para tirar esses dados nulos do nosso gráfico, nós vamos criar um filtro. E vamos criar um filtro para excluir todos os nulos.

00:00:10:07 - 00:00:11:10

Vamos aqui na aba de configuração, adicionar um filtro, criar um filtro. Vamos chamar de excluir nulos. Vamos colocar excluir. A gente vai selecionar qual que é a coluna que a gente está querendo excluir esses nulos, né? Então: novo nível. Inclusive, vou colocar no nome excluir nulos nível, para a gente poder saber que é excluir os nulos dessa coluna. E aí a gente coloca aqui essa opção de nulo. Então, vai excluir da coluna novo nível todos os nulos. Vamos salvar.

Já vai aplicar, atualizando aqui e sumiu aí as nossas barrinhas de nulo. Olha, com esse gráfico a gente consegue fazer uma boa comparação entre gênero e nível de senioridade. Vocês acham que se fizéssemos gráfico do tipo empilhado ficaria melhor? A gente poderia empilhar para, de novo, ficar melhor ver as proporções como fizemos anteriormente. Mas nesse caso, se o nosso objetivo fosse apenas visualizar a quantidade de pessoas por gênero em cada área ou verificar.

00:00:11:10 - 00:00:12:08

Por exemplo, qual a senioridade tem mais mulheres, esse gráfico aqui já seria o suficiente. Tudo depende da pergunta que a gente se propõe a responder. Novamente, podemos ver que tem muitas pessoas do gênero masculino em todos os níveis de senioridade, mas temos uma diferença absurda se comparamos o item de pessoa gestora. Sabemos que temos mais pessoas do gênero masculino do que feminino na pesquisa. Então, é normal esperar que a diferença continue aqui. Porém, notem de pessoa gestora, a barra roxa parece ser quatro ou cinco vezes maior que a barra azul, que é de gênero feminino. Enquanto em júnior, a barra roxa é duas ou três vezes maior. Isso significa que a diferença aumentou de júnior para pessoa gestora. Então, podemos até dizer que ir para um nível de líder é mais difícil para uma pessoa do gênero feminino. Então, ok. Vamos para um próximo gráfico, fazer um gráfico de média salarial por gênero, para saber se tem uma maior quantidade de pessoas do gênero masculino que ganham mais.

00:00:12:08 - 00:00:13:41

Vamos fazer um outro gráfico de barras. Então, vamos aqui... Deixa eu colocar aqui um pouquinho mais para baixo agora, já posicionar mais ou menos onde eu vou colocar o meu gráfico. Adicionar gráfico, vamos colocar um gráfico de barras aqui. Beleza. Aqui em dimensão agora vai ser salário e em dimensão detalhada vai ser gênero. Bom, é isso que a gente tinha feito nos outros gráficos, né? Mas veja bem, a gente quer a média de salário. Então, o salário é uma métrica, é a média do salário, não a coluna de salário em si. Então, cancela. No lugar de salário aqui na dimensão, a gente coloca gênero e tira o gênero lá da dimensão detalhada. E agora o salário vai para a métrica. Então aqui a gente coloca salário. E aqui nesse cantinho, nessa canetinha, a gente vai colocar aqui, ó: médio. Agora sim. Opa. Aqui. Bom, aí ficou assim o nosso gráfico. Todas as barras da mesma cor, né? Para a gente poder colocar cada barrinha ali de uma cor diferente, a gente pode colocar gênero também na dimensão detalhada. Vamos colocar lá. Gênero. Top, cada barrinha com uma cor, e aí representando a média salarial de cada gênero.

00:00:13:41 - 00:00:15:13

Beleza? Vamos colocar um título no nosso gráfico aqui ainda em estilo: mostrar título, vamos colocar gráfico, média salarial por gênero. E aí, quais conclusões podemos tirar desse gráfico? Que pessoas do gênero masculino tendem a ganhar mais? Vemos umas barrinhas aqui, indicadoras das médias, e a gente pode ver que a barrinha de masculino está acima do feminino. Inclusive, aproveitando essa interatividade, a gente pode passar aqui na barrinha feminina e ver exatamente a média salarial. E, se comparada com a masculina, dá uma diferença de quase dois mil reais. Então, realmente, pelos nossos dados aqui, as pessoas do gênero masculino tendem a ganhar a mais. Para finalizar, a gente pode colocar um gráfico de gênero e etnia. Mas dessa vez usando um gráfico de barras horizontal e com os valores em porcentagem.

Novamente, pensando naquela questão de proporção. A gente vai lá em cima: adicionar um gráfico. Aí a gente quer empilhados e do tipo 100%, que é porcentagem para a gente conseguir fazer uma proporção. Vou aumentar um pouquinho, ok. Aí na dimensão aqui... Na primeira, que é a dimensão, a gente coloca a cor, raça e etnia. Na segunda a gente coloca gênero. Vamos ver como que vai ficar. Ok.

00:00:15:13 - 00:00:16:45

Na métrica aqui, a gente coloca ID. Aqui na métrica também a gente coloca contagem. E ok. Só para poder colocar os numerozinhos também, a gente vai em estilo e aqui nessa opção de mostrar rótulos dos dados. Show, galera. Vamos aproveitar que a gente está aqui em estilo já para colocar o título. Gráfico gênero etnia. Acho que foi, né? Beleza. Bom, a gente consegue concluir que o gênero masculino é a maioria independente de qual etnia, né? A gente pode até pensar que é porque a maioria das pessoas que responderam o formulário assinalaram que são do gênero masculino e tal, mas a gente já sabe o contexto. E podemos levantar um questionamento interessante. Então, na área de TI, tem mais pessoas do gênero masculino. Bom, mas também existe uma coisa muito interessante, que nesse gráfico empilhado, no percentual, a gente consegue ver a concentração, a porcentagem dentro de cada categoria. Por exemplo, dentro da categoria de pessoas brancas, a gente tem 75% masculino e 23% feminino. Mas o que eu quero apontar é que essa porcentagem é grande dentro da categoria parda, da categoria preta, por exemplo, a gente tem 27% de um público feminino, e dentro da categoria amarela, que a gente tem 29% feminino.

00:00:16:45 - 00:00:18:13

E se a gente pensar em juntar preta e parda, fica até uma porcentagem possivelmente maior. Bom, agora eu vou organizar esteticamente nossa página de análise de gênero naquele mesmo esquema lá de pegar um retângulo, colocar o título, colocar os gráficos bonitinhos no fundo, e eu vou deixar essa parte estética para vocês organizarem aí. E é bom até que compartilhem entre vocês como que está ficando, algumas ideias, quem conseguir fazer alguma coisa diferente para ficar legal. Beleza? Bom, gente, a minha página de análise de gênero ficou assim, com os gráficos e tal. Para poder ficar esse roxo, eu só coloquei o retângulo que eu já tinha dado aquela dica, né, no fundo. E, bom, olhando aqui agora, uma coisa que ficou faltando é um gráfico ou algo para a gente saber sobre a idade dessas pessoas, né? Mas eu guardei o melhor para o final. Aqui no Looker tem uma ferramenta legal que possibilita realizar filtros. A gente aqui vai nos três pontinhos e vai adicionar um controle. A gente pode clicar aqui nessa primeira opção, que é a lista suspensa. Eu vou colocar aqui aleatoriamente, depois a gente reorganiza na página. Ok. Quando colocamos ela aqui na nossa página, vai aparecer um retângulo, a gente vai aqui na parte de configuração com esse retângulo selecionado e na parte de campo de controle, a gente coloca qual que é a coluna que a gente quer trabalhar e filtrar, né?

00:00:18:13 - 00:00:19:46

No nosso caso, como eu dei essa ideia aí da idade, a gente pode colocar a faixa de idade aqui. E na parte métrica a gente pode colocar o ID e o tipo de contagem, né? Métrica, já está o ID aqui mesmo, e é do tipo contagem. Ok. E eu vou deixar para mexer no estilo depois, porque aí é só ir na barra de estilo e vocês podem também fazer essa brincadeira aí de mexer no estilo sozinhos.

E bom, esse filtro, deixa eu só ocultar aqui e ver se vai dar um zoom. Ótimo, perfeito. Ô gente, eu vou colocar aqui para o outro lado. Vocês podem colocar aí no lugar que vocês quiserem. Aí quando a gente clica nessa caixinha, ela estava assim, quando eu clico nela, aparecem ali todas as faixas de idades possíveis. E aqui nessa página a gente tem um gráfico de forma geral, de todos os dados, mas se eu quiser analisar uma faixa de idade específica, por exemplo, de 22 a 24 anos, eu posso abrir aqui e aqui entre 22 e 24, eu posso clicar nesse somente, e vai atualizar todos os dados somente para essa faixa de idade. Então aqui, a gente vai ter todos os dados para analisar uma faixa específica. Olha que doideira aqui, se a gente puder, por exemplo, ver essa parte de senioridade. A gente vê que tem uma diferença muito grande no nível júnior.

00:00:19:46 - 00:00:21:03

Às vezes, quando está em uma idade mais velha, sei lá, a partir de 40 anos, não tenha essa diferença. Enfim, a gente consegue fazer algumas coisas mais legais aqui, algumas comparações bem interessantes. Então, gente, a minha página de análise de gênero ficou assim. Mas como falei na aula anterior, vocês têm liberdade criativa para explorar o Looker e se quiser colocar mais gráficos ou outros tipos de recursos que o Looker oferece, fiquem à vontade. É só lembrar aí que dashboards servem para contar uma história sobre os dados, tentar mostrar de forma visual para as pessoas o que está acontecendo com os dados. E agora vou deixar um desafio aqui para vocês. Nós já vimos como acrescentar gráficos e tudo mais. A próxima página do nosso dashboard tem como título Análise Detalhada de Etnia. Eu gostaria que vocês já se adiantassem e fizessem essa página. Terão quatro gráficos no total: Etnia e Salário, Etnia e Região do Brasil, Etnia e Nível de Senioridade e Etnia e Escolaridade. A estética vai seguir a mesma dessa página que a gente fez agora. Qualquer dúvida, vocês podem perguntar para a tutora do curso, tá? Então na próxima aula, eu já venho aqui mostrando como ficou a minha página, quais os tipos de gráfico eu escolhi e a gente finaliza esse módulo. Combinades? E foi isso nessa aula, pessoal. Até mais.