MVP Engenharia de Dados – Felipe Teixeira Marques

Introdução

Este relatório (MVP) apresenta os resultados de um projeto de Engenharia de Dados que visa analisar a equidade salarial entre homens e mulheres além da checagem do reconhecimento financeiro, através de méritos e promoções para os maiores talentos dentro da Bemobi. O objetivo principal é determinar se as posições nas faixas salarias ou PIR (Position in Rage) entre os gêneros (masculino e feminino) estão sendo feitos de forma justa e identificar quaisquer disparidades salariais que possam existir.

Objetivo

O MVP, visa duas perguntas que são levantadas pela diretoria:

- A distribuição de salários está realizada de forma justa entre gêneros? Em outras palavras, homens e mulheres possuem a mesma distribuição salarial quando comparamos seus salários e suas respectivas medianas?
- Nossos talentos, chamados de Key Talent, foram reconhecidos no último ciclo de mérito?

Para concluir esse objetivo, serão utilizados 3 arquivos em CSV, extraídos de controles da empresa onde temos informações relacionadas à:

- Base de colabores;
- Ciclo de mérito;
- Histórico de méritos e promoções anteriores à 2023;

Com isso será possível realizar os próximos passos da análise, que serão:

1. Avaliar e comparar os salários entre funcionários de diferentes gêneros na organização.
2. Identificar possíveis disparidades salariais que possam indicar falta de reconhecimento dos talentos com base no gênero.
3. Realizar análises estatísticas para determinar se as diferenças salariais são estatisticamente significativas.
4. Apresentar resultados claros e conclusões sobre a equidade salarial, focando especialmente no reconhecimento dos talentos dentro da empresa.
5. Propor recomendações, se necessário, para melhorar a equidade salarial e garantir que os talentos sejam reconhecidos de maneira justa.

Coleta

A coleta de dados foi realizada na empresa Bemobi, e se deu, através de busca por arquivos disponíveis em sistemas internos da empresa.

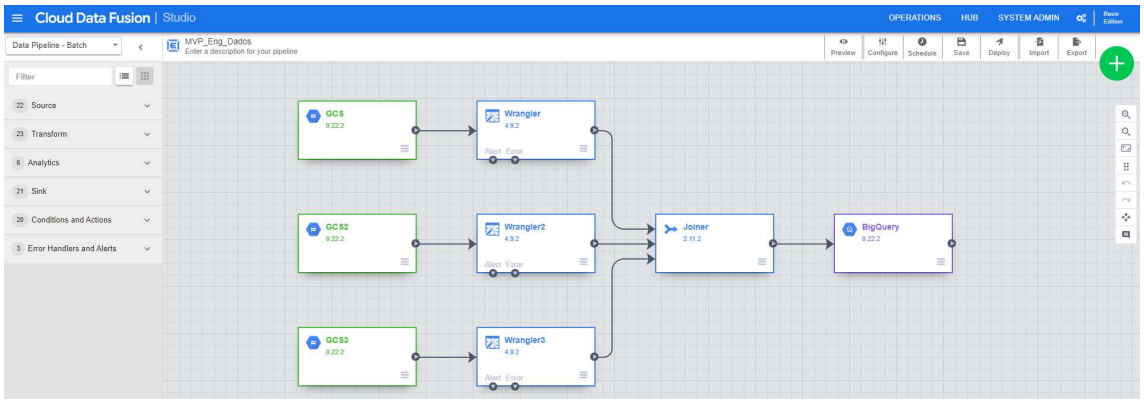
Esses arquivos são gerados em formato XLSX, como existiam dados sensíveis em 2 das 3 tabelas utilizadas foi necessário realizar a remoção, garantindo assim a segurança e privacidade dos dados e por conseguinte dos funcionários da empresa.

Após essa transformação inicial de dados, houve a conversão para o formato CSV, para que só então seguissem para o Google Could, plataforma em nuvem selecionada para o desenvolvimento das análises.

Modelagem

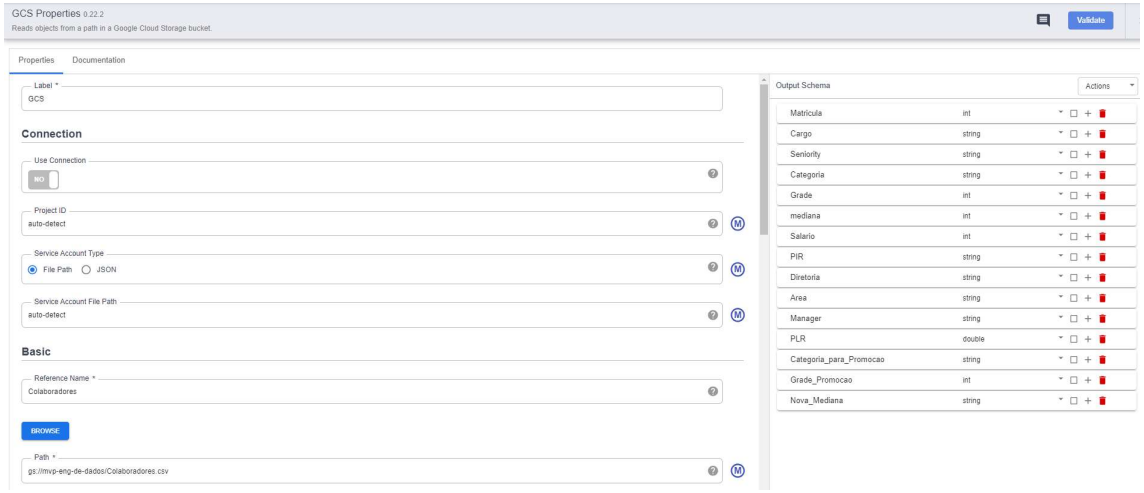
O ETL foi realizado dentro do Google Cloud, inicialmente foi criado um Bucket, denominado de mvp-eng-dados, e dentro dele inseridos os 3 arquivos:

Em seguida os a realização da extração (extract) das informações na etapa 1 (Google Cloud Storage, ou GCS), dos dados da fonte que estavam no bucket.

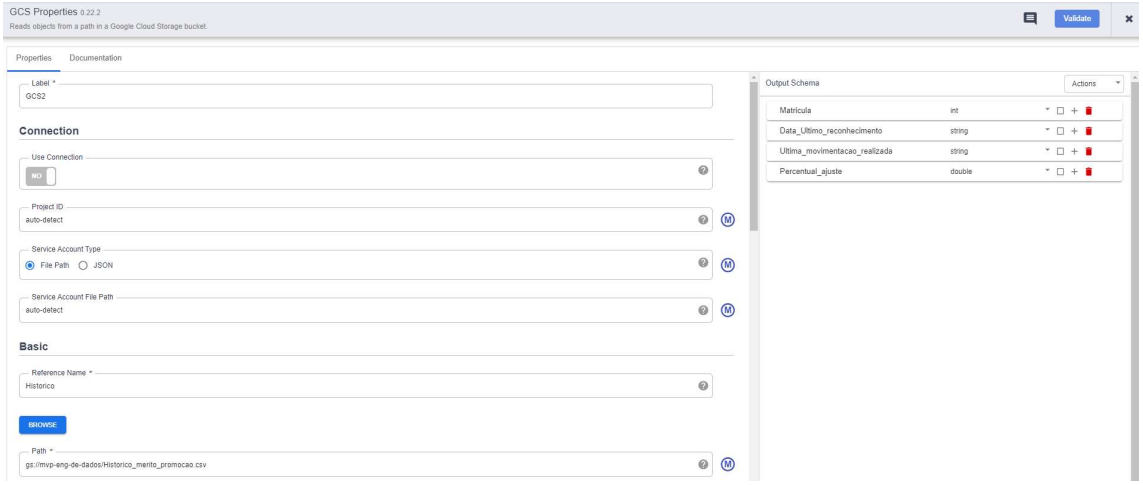


As extrações foram realizadas em 3 GCS distintos:

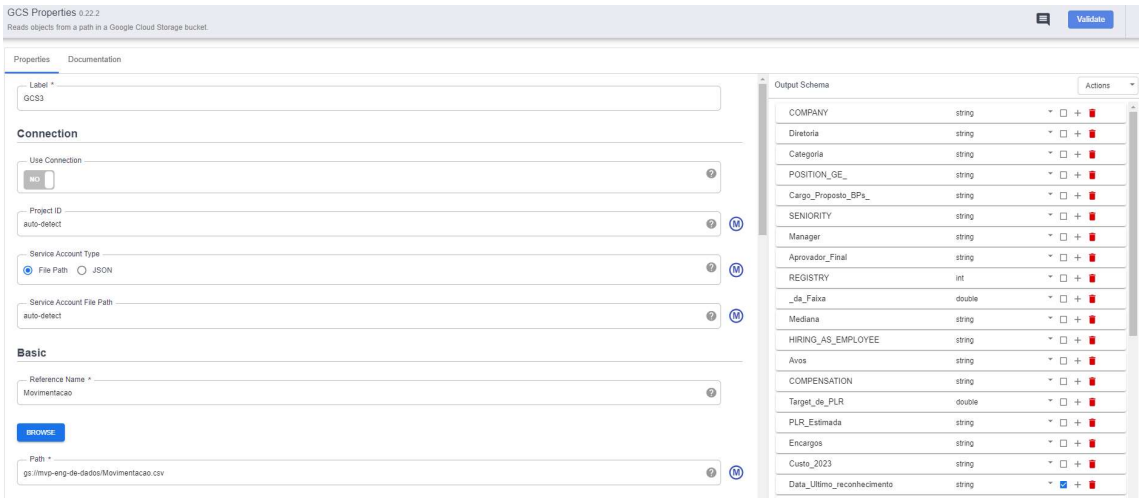
1.



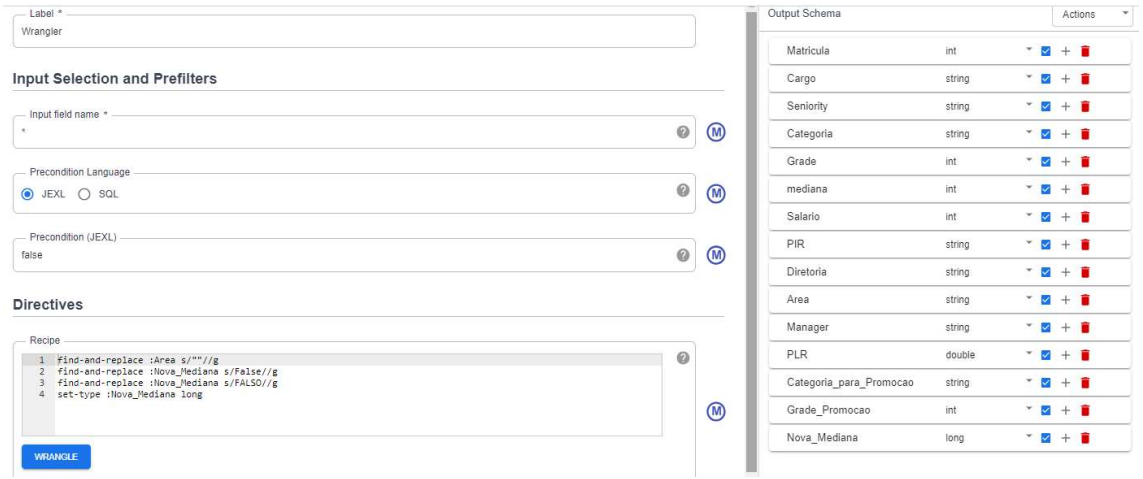
2.



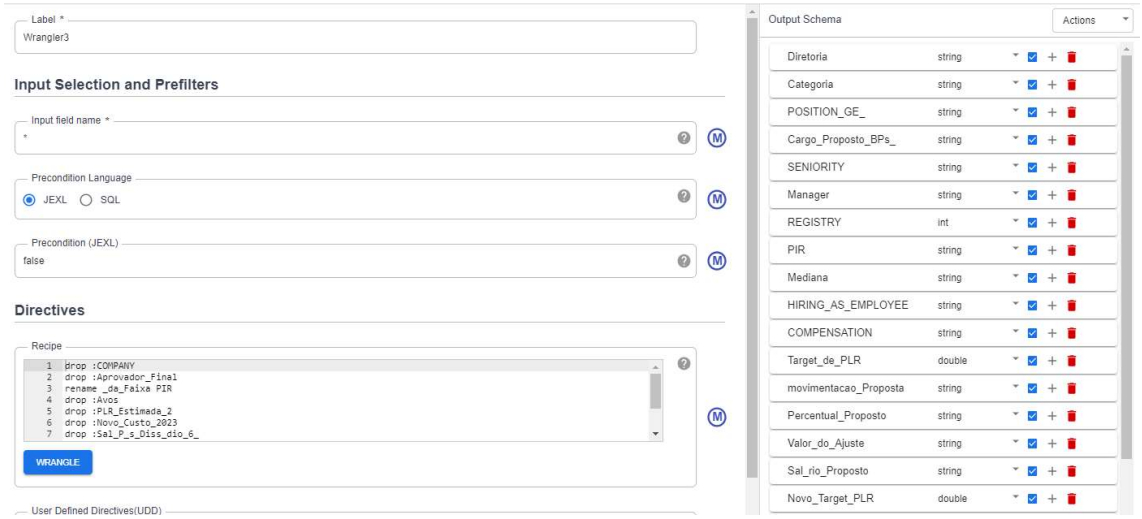
3.



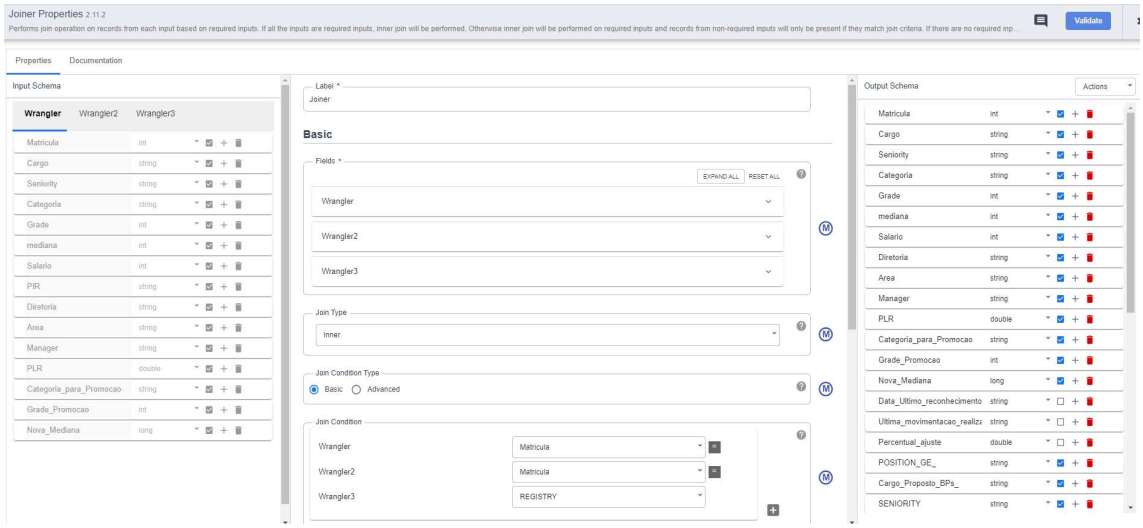
Após essa etapa, o Wrangler é utilizado para realizar as transformações (transform), convertendo Medianas para Int, realizando substituições de “” e “FALSE” e “FALSO” por “vazio”, no primeiro Wrangler.



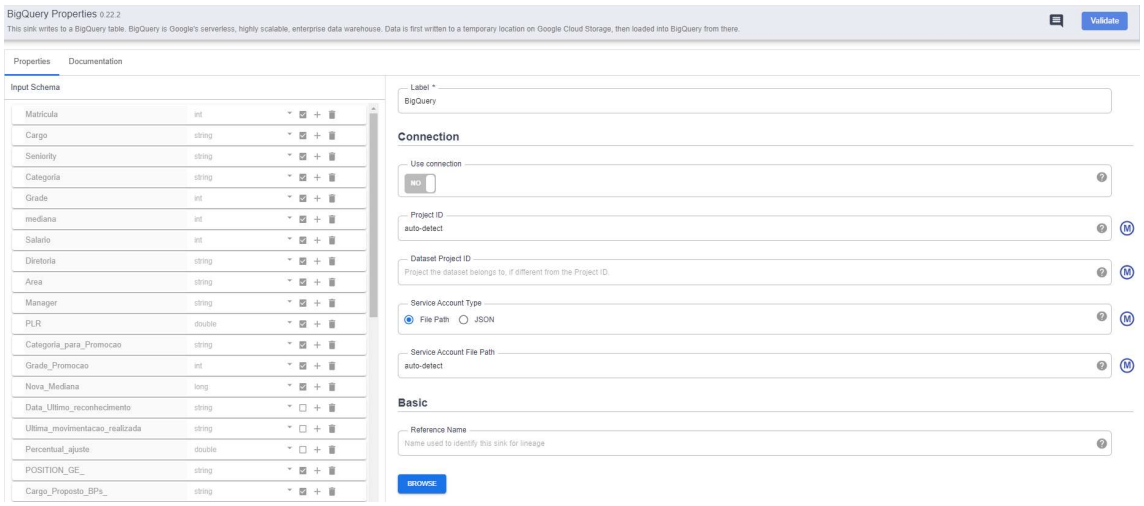
Além disso, foram realizadas diversas transformações no 3 Wrangler, segue a descrição abaixo:



Depois de realizar essas transformações foi utilizado o Joiner Propeties realizando um Inner Join:



E finalmente após essas etapas, o BigQuery é utilizado, para realizar o carregamento de dados (Load).



Na próxima etapa o Deploy do fluxo foi realizado para que os dados possam ser analisados e responder os questionamentos iniciais do MVP.

Análise

Considerando os dados foi possível notar que das 73 promoções ou méritos realizados no último ciclo, 25% (18) ocorreram em pessoas do sexo feminino, o que em primeiro momento pode parecer que essas pessoas estão sendo prejudicadas, entretanto por se tratar de uma empresa de tecnologia, precisamos avaliar o quadro inicial, e nesse temos a seguinte distribuição por sexo:

Gênero	Valor Absoluto	Percentual
Masculino	285	74%
Feminino	99	26%
Total	384	100%

Levando essa tabela em consideração, a perspectiva muda pois se olharmos para cada grupo de forma individual, percebemos que 19% dos homens (pessoas do sexo masculino), receberam alguma promoção ou mérito, e 18% das mulheres (pessoas do sexo feminino), foram promovidas ou contempladas, com isso percebemos que os dois sexos foram considerados de forma semelhante no ciclo de mérito, uma vez que uma diferença de 1% não pode ser considerada grande suficiente para que uma ação seja necessária para gerar uma igualdade de gêneros da empresa.

Em relação ao PIR, notamos que as mulheres possuem uma média de **99%** em suas respectivas faixas, mostrando que de forma geral as mulheres estão sendo pagas de forma justa pelo seu trabalho e consideradas adequadas para suas funções, a média de PIR para homens é de **108%**, o que mostra, que de forma geral estão mais acelerados em suas respectivas faixas e para que ocorra um equilíbrio de gêneros, mostrando que em média homens estão sendo considerados como candidatos a promoções, já que para ser elegível a uma promoção a pessoa precisa estar acima dos 100% de PIR.

Embora a distribuição de promoções e méritos sejam justas uma informação que merece atenção é a distribuição dos talentos, conforme tabela abaixo:

Gênero	Talentos	Percentual
Masculino	37	71%
Feminino	15	28%
Total	52	100%

Ao cruzar os dois quadros notamos que 13% dos homens são considerados talentos e 15% das mulheres são consideradas como talento.

Portanto a empresa não possuí um problema de reconhecimento de mulheres, o que atualmente ainda pode ser considerado um diferencial, especialmente por ser tratar de um mercado majoritariamente masculino.

Autoavaliação

Ao longo do desenvolvimento foram encontradas algumas dificuldades, em geral, relacionadas a plataforma, e senti que não tive assistência através do Discord, por isso, precisei recorrer a colegas que atuam na área e a longas buscas na internet.

Apesar das dificuldades o trabalho foi concluído com sucesso, e para minha surpresa, o resultado obtido foi muito diferente do esperado.

Como citado na introdução foram utilizados aqui dados da empresa em que atuo, a Bemobi. E existe uma percepção de que mulheres não são reconhecidas dentro da empresa, inclusive com propostas de como resolver esse problema sendo levantadas por outras sub áreas dentro do RH.

Ao concluir a análise de dados, pude perceber que na verdade não existe um problema com o reconhecimento de mulheres na empresa, o que ocorre é que majoritariamente temos homens na empresa, tornando o espaço amostral de mulheres muito menor.

Sendo assim, o que podemos pensar é em como captar mais talentos femininos para a empresa, uma vez que não foi notada nenhuma diferença significativa nos percentuais de reconhecimento ou mesmo posicionamento na faixa entre homens e mulheres.