# Análise de dados do Departamento de recursos humanos

## Etapa 1 - Exemplos de problemas de negócio envolvendo o departamento de RH

Alguns problemas ou necessidades de negócios envolvendo o departamento de RH foram listados abaixo. Alguns serão respondidos na conclusão do projeto, após a análise, dependendo da disponibilidade do conjunto de dados “fornecidos” pelo cliente. Outros poderão ser utilizados como um ponto de partida para serem discutidos com o departamento de RH futuramente ou quando outros dados forem fornecidos:

**Alta rotatividade:** O departamento de RH enfrenta uma taxa de rotatividade significativa, resultando em uma necessidade constante de contratação de novos funcionários e impactando a estabilidade da força de trabalho. Esse problema pode levar ao aumento dos custos de recrutamento, a dinâmica da equipe, o moral dos funcionários, diminuição da produtividade e possíveis lacunas de conhecimento dentro da empresa, além de incorrer em custos associados à recontratação e treinamento de novos funcionários.

**Práticas de remuneração desiguais:** O departamento de RH está enfrentando disparidades nos salários dos funcionários, onde certos grupos podem estar recebendo remuneração injusta em comparação com seus pares. Esse problema pode afetar negativamente a satisfação e o engajamento dos funcionários e potencialmente levar a complicações legais relacionadas à equidade salarial e à discriminação.

**Desigualdade geográfica:** O departamento de RH enfrenta um desequilíbrio nas contratações, com aumento significativo em uma cidade e redução simultânea em outra cidade.

**Tempo de empresa:** O departamento de RH está enfrentando desafios relacionados ao tempo de serviço dos funcionários, incluindo preocupações sobre retenção de funcionários, planejamento de sucessão e transferência de conhecimento.

## Etapa 2 - Definição da Baseline

Foram formuladas perguntas relacionadas aos problemas ou necessidades de negócios envolvendo o departamento de RH, as quais serão respondidas, caso existam dados suficientes para responde-las, ao final da análise.

**Alta rotatividade**

* Quais são as principais razões para a alta taxa de rotatividade e qual são os seus índices?
* Como a empresa pode melhorar a retenção de funcionários e reduzir a rotatividade?
* Quais são os fatores subjacentes que contribuem para a insatisfação dos funcionários e subsequentes rescisões?
* Quais iniciativas podem ser implementadas para melhorar a retenção de funcionários e aumentar a satisfação no trabalho?

**Práticas de remuneração desiguais**

* Existem disparidades salariais significativas entre os diferentes grupos de funcionários? Em caso afirmativo, quais são os fatores que contribuem para essas disparidades?
* Como a empresa pode garantir práticas de remuneração justas e equitativas para todos os funcionários?

**Desigualdade geográfica:**

* Quais fatores contribuíram para o aumento das contratações em uma cidade?
* Existem necessidades comerciais específicas ou condições de mercado que impulsionam o aumento de contratações nessa cidade?
* Que desafios ou obstáculos estão impedindo os esforços de contratação na cidade que está diminuindo?
* Houve uma mudança no foco estratégico da empresa ou nos requisitos operacionais que favorecem uma cidade em detrimento de outra?
* Existem fatores externos, como condições econômicas ou tendências da indústria, que influenciam as disparidades de contratação entre as cidades?
* Que medidas o departamento de RH pode tomar para lidar com o desequilíbrio na contratação e garantir uma distribuição mais equilibrada de talentos nas cidades?
* Existem riscos potenciais associados à concentração de funcionários em uma cidade, como aumento da rotatividade ou dependência excessiva de um único local?
* Que estratégias podem ser implementadas para atrair e reter os melhores talentos na cidade que está enfrentando uma queda nas contratações?
* Existem oportunidades de mobilidade interna ou acordos de trabalho remoto que podem ajudar a mitigar o impacto dos desequilíbrios de contratação entre as cidades?

**Tempo de empresa**

* Qual é o tempo médio de serviço dos funcionários da empresa?
* Existem departamentos ou funções específicas que apresentam maior rotatividade ou períodos médios de permanência mais curtos?
* Que fatores contribuem para que os funcionários deixem a empresa antes de atingirem cargos mais longos?
* Como a empresa pode melhorar a retenção de funcionários e incentivar compromissos de longo prazo?
* Que estratégias podem ser implementadas para atrair e reter os melhores talentos para cargos críticos?
* Existem oportunidades de desenvolvimento de carreira para incentivar os funcionários a permanecer e crescer dentro da empresa?
* Como o departamento de RH pode identificar e abordar quaisquer preocupações ou desafios que possam contribuir para uma permanência mais curta dos funcionários?
* Como a empresa pode garantir uma transferência de conhecimento suave e um processo de planejamento de sucessão para mitigar o impacto da rotatividade de funcionários?
* Existem quaisquer incentivos ou benefícios que possam ser introduzidos para recompensar e reconhecer os funcionários por seus serviços contínuos à empresa?

## Etapa 3 - Planejamento

Essa etapa tem como objetivo estabelecer a abordagem do projeto, incluindo a decomposição dos problemas, a entrega do produto final, as etapas do processo e as ferramentas a serem utilizadas. Para alcançar essa definição, é necessário realizar reuniões diárias e/ou especificas, em conjunto com os demais membros da equipe de análise de dados, por esse motivo, será adotado um processo de planejamento simplificado para este exemplo de estudo, descrito no arquivo “planejamento.txt.”

## Etapa 4 - Coleta de Dados Brutos de Fontes Confiáveis

Os dados utilizados no projeto são dados primários e foram coletados diretamente do cliente e no formato xlsx. Encontram-se na pasta \_datasets.

## Etapa 5 e 6 - Limpeza e Processamento de Dados e Análise Exploratória de Dados

**Etapas realizadas no Power Query**

### Tabela Calendário

A tabela DIMCALENDARIO foi criada realizando as etapas abaixo:

1. No Power BI, entre no Power Query, através do Transformar dados.
2. Adicione uma nova fonte > Consulta nula > Renomeie a consulta para DIMCALENDÁRIO.
3. Adicione a fórmula na Barra de fórmula para fazer uma lista de datas:

= List.Dates(#date (2000, 1, 1), Duration.Days(#date (2018, 12, 31) - #date (2000, 1, 1)) +1, #duration (1,0,0,0))

1. Na aba transformar, clique em converter para a tabela (Criando uma tabela a partir de uma lista de valores) > Abrirá uma janela > Clique OK
2. Formate a coluna como Data e renomeie o nome da coluna para Data
3. Na aba adicionar coluna > Clique na data e acrescente as colunas: Ano, mês, nome do mês, início do mês e dia > Salvar e fechar (Voltar para o Power BI)
4. Relacionar a tabela DIMCALENDÁRIO com a tabela FATOFUNCIONÁRIO

Colunas criadas a partir da coluna da Data:

* Data = CALENDAR (DATE (2000,01,01), DATE (2018,12,31))
* Ano = YEAR (TabelaCalendário [Data]) ou Ano = Date.Year([Data])
* Mês = MONTH (TabelaCalendário [Data]) ou Mês = Date.Month([Data])
* Nome do mês = FORMAT (TabelaCalendário [Data], “mmmm”) ou Nome do mês = Date.MonthName([Data])
* Início do mês = STARTOFMONTH (TabelaCalendário [Data]) ou Início do mês = Date.StartOfMonth([Data])
* Dia = DAY (TabelaCalendário [Data]) ou Dia = Date.Day([Data])
* Mês/Ano = FORMAT (TabelaCalendário [Data], “mmm/yyyy”)

### Tabela Avaliações

A tabela DIMAVALIAÇÃO foi criada a partir da duplicação da tabela FATOFUNCIONÁRIO. Foram excluídas várias colunas, permanecendo apenas as colunas ID RH, Trabalho em Equipe, Liderança, Comunicação, Iniciativa e Organização. As colunas Trabalho em Equipe, Liderança, Comunicação, Iniciativa e Organização foram transformadas em linhas ficando na coluna Competência e seus valores agrupados em uma coluna de Notas.

* ID RH = Contém o número da identificação exclusiva de cada funcionário. ***(Alteração no Tipo de Dado: Número inteiro)***
* Competência = Contém os nomes das competências de cada funcionário, tendo 5 opções: Trabalho em Equipe, Liderança, Comunicação, Iniciativa e Organização ***(Alteração no Tipo de Dado: Texto)***
* Nota = Contém o número da nota individual atribuída para cada competência de cada funcionário ***(Alteração no Tipo de Dado: Número inteiro)***

### Tabela Funcionários

A tabela FATOFUNCIONARIO foi modificada, permanecendo com as seguintes colunas:

* ID RH = Sem modificação nos dados ***(Alteração no Tipo de Dado: Inteiro)***
* Nome Completo = Sem modificação nos dados ***(Alteração no Tipo de Dado: Texto)***
* Estado Civil = Substituições de C para Casado(a) e S para Solteiro(a) ***(Alteração no Tipo de Dado: Texto)***
* Sexo = Substituições de M para Masculino e F para Feminino ***(Alteração no Tipo de Dado: Texto)***
* Data de Nascimento = Sem modificação nos dados ***(Alteração no Tipo de Dado: Data)***
* Endereço = Sem modificação nos dados ***(Alteração no Tipo de Dado: Texto)***
* Data de contratação = Sem modificação nos dados ***(Alteração no Tipo de Dado: Data)***
* Data de demissão = Sem modificação nos dados ***(Alteração no Tipo de Dado: Data)***
* Salário Base = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* Impostos = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* Benefícios = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* VT = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* VR = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* Cargos = Sem modificação nos dados ***(Alteração no Tipo de Dado: Texto)***
* Níveis = Sem modificação nos dados ***(Alteração no Tipo de Dado: texto)***
* Áreas = Sem modificação nos dados ***(Alteração no Tipo de Dado: Texto)***
* Férias acumuladas = Sem modificação nos dados ***(Alteração no Tipo de Dado: Inteiro)***
* Horas Extras = Sem modificação nos dados ***(Alteração no Tipo de Dado: Inteiro)***
* Imagens = Sem modificação nos dados ***(Alteração no Tipo de Dado: Texto)***
* Avaliação do funcionário = Sem modificação nos dados ***(Alteração no Tipo de Dado: Número decimal)***
* Idade = Criação da coluna através da fórmula Date.From(#date (2018, 12, 31)) - [Data de Nascimento], resultando na diferença da Data de Nascimento de cada funcionário e a data 31/12/2018 em dias. Depois foi feito a transformação da quantidade de dias em anos. ***(Alteração no Tipo de Dado: Inteiro)***
* Cidade = Coluna criada utilizando a coluna de exemplo a partir da coluna Endereço. ***(Alteração no Tipo de Dado: Texto)***
* Estado = Coluna criada utilizando a coluna de exemplo a partir da coluna Endereço. ***(Alteração no Tipo de Dado: Texto)***
* Rua = Coluna criada utilizando a coluna de exemplo a partir da coluna endereço. ***(Alteração no Tipo de Dado: Texto)***

As colunas Trabalho em Equipe, Liderança, Comunicação, Iniciativa e Organização foram excluídas da tabela.

### Parâmetro

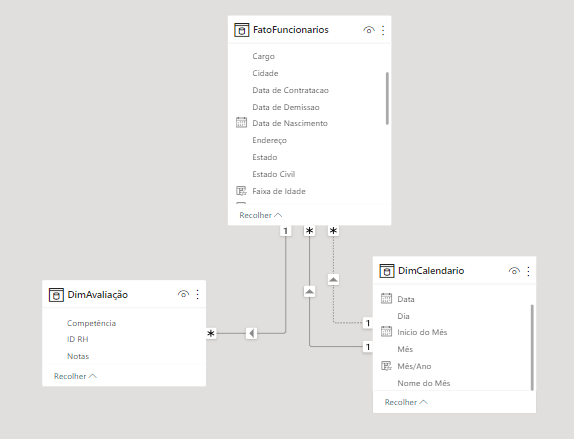
Criação do parâmetro CAMINHO com o caminho das pastas onde se encontra o arquivo em Excel (Dataset.xlxs) e a sua substituição no código Fonte. Assim a conexão do arquivo em Excel com o arquivo em Power BI pode ser atualizada com mais facilidade, caso ocorra alguma modificação de pastas.

## Etapa 7 - Preparação dos Dados (Modelagem dos dados)

**Etapas realizadas no Power BI**

Na exibição do modelo, foi realizado as seguintes relações:

* Tabela DIMCALENDARIO com a tabela FATOFUNCIONARIO de um para muitos, utilizando a coluna DATA como chave primaria
* Tabela DIMAVALIAÇÃO com a tabela FATOFUNCIONARIO de um para muitos, utilizando a coluna ID RH como chave primaria



### Colunas calculadas

1. Coluna Anos de empresa = Se a data de demissão estiver em branco, utilize a fração do ano que representa o número de dias inteiros entre a data de contratação e a data de hoje, caso contrário, utilize a fração do ano que representa o número de dias inteiros entre a data de contratação e a data de demissão.

IF (FatoFuncionarios [Data de Demissao] =BLANK (),

ROUNDDOWN (YEARFRAC (FatoFuncionarios [Data de Contratacao], TODAY ()),0),

ROUNDDOWN (YEARFRAC (FatoFuncionarios [Data de Contratacao], Fato Funcionários [Data de Demissao]),0))

1. Coluna Faixa de Idade = Se o número da coluna Idade for menor que 30, informe “18-30 anos”, se for menor que 40, informe “31-40 anos”, se for menor que 60, informe “41-60 anos”, caso contrário, informe “Acima de 60 anos”.

IF(

    FatoFuncionarios[Idade] <= 30, "18-30 anos",

    IF(

        FatoFuncionarios[Idade] <= 40, "31-40 anos",

        IF(

            FatoFuncionarios[Idade] <= 60, "41-60 anos",

            "Acima de 60 anos"

        )

    )

)

1. Coluna Faixa de tempo de empresa = Se o número da coluna Anos de empresa for menor que 5, informe “0-5 anos”, se for menor que 10, informe “5-10 anos”, se for menor que 15, informe “10-15 anos”, se for menor que 20, informe “15-20 anos”, caso contrário, informe “Acima de 20 anos”.

IF(

    FatoFuncionarios[Anos de empresa] <= 5, "0-5 anos",

    IF(

        FatoFuncionarios[Anos de empresa] <= 10, "5-10 anos",

        IF(

            FatoFuncionarios[Anos de empresa] <= 15, "10-15 anos",

            IF(

                FatoFuncionarios[Anos de empresa] <= 20, "15-20 anos",

                "Acima de 20 anos"

            )

        )

    )

)

### Tabela Medida

A tabela MEDIDA foi criada para armazenar as medidas criadas para serem utilizadas nos Dashboards:

1. Total de contratação = Medida que conta a quantidade de linha que tem a tabela funcionários, pois cada linha representa um registro no RH, um funcionário contratado. = COUNTROWS(FatoFuncionarios)
2. Total de demissões = Medida que conta a quantidade de demissões, usando todas as linhas da tabela (Total de contratação) e filtrando apenas as linhas que não são vazias na coluna Data de Demissão (FatoFuncionarios [Data de Demissao] <> BLANK ()). Além disso, foi necessário especificar uma relação entre a coluna de demissão e a data do calendário para que as demissões fiquem no ano correto. = CALCULATE ([Total de contratação], FatoFuncionarios [Data de Demissao] <> BLANK (), USERELATIONSHIP (DimCalendario [Data], FatoFuncionarios [Data de Demissao]))
3. Total de funcionários ativos = Medida que calcula a quantidade de funcionários ativos fazendo a diferença do total de contratação com o total de demissões = [Total de contratação] - [Total de demissões]
4. % Turnover = Medida que calcula a taxa de rotatividade de funcionários. Somando os funcionários demitidos e admitidos e dividindo por 2. Seu resultado é divido pela quantidade de funcionários ativos. = (([Total de demissões] + [Total de contratação]) /2) / [Total de funcionários ativos]

1. Média salarial = Medida calcula a média do Salário, VR, VT, Benefício e Impostos de todos os funcionários = AVERAGEX (FatoFuncionarios,FatoFuncionarios[Salário Base] + FatoFuncionarios[VR] + FatoFuncionarios[VT] + FatoFuncionarios[Beneficios] + FatoFuncionarios[Impostos])
2. Desvio da Meta = Medida que calcula quanto que desvio da meta de 60% em relação a taxa de turnover. Cidade que tiver a taxa de rotatividade abaixo de 60% atingiu a meta (Fica verde). = 0.6 - [% Turnover]
3. Salário total = Medida calcula o somatório do Salário, VR, VT, Benefício e Impostos de todos os funcionários = SUMX (FatoFuncionarios,FatoFuncionarios[Salário Base] + FatoFuncionarios[VR] + FatoFuncionarios[VT] + FatoFuncionarios[Beneficios] + FatoFuncionarios[Impostos])
4. Nota individual = Medida que calcula o total de todas as notas individuais = SUM('DimAvaliação'[Notas])
5. Nota Geral = Medida que calcula a média das notas de todos os funcionários = CALCULATE(AVERAGE('DimAvaliação'[Notas]), ALL(FatoFuncionarios))
6. Média das notas das competências = Medida que calcula a média da avaliação das notas referente as competências dos funcionários = AVERAGE('DimAvaliação'[Notas])

Funções DAX utilizadas:

* A função **COUNTROWS([<TABLE>])** conta o número de linhas na tabela especificada ou em uma tabela definida por uma expressão.
* A função **CALCULATE(<EXPRESSION>, [<FILTER1>], [<FILTER2>] …)** avalia uma expressão em um contexto de filtro modificado.
* A função **USERELATIONSHIP(<COLUMNNAME1>, <COLUMNNAME2>)** especifica a relação a ser usada em um cálculo específico como aquela que existe entre COLUMNNAME1 e COLUMNNAME2.
* A estrutura condicional **IF (LOGICALTEST, RESULTIFTRUE, [RESULTIFFALSE])** verifica se uma condição é atendida e retornará um valor, se for TRUE, e outro valor, se for FALSE.
* A função **ROUNDDOWN (NÚMERO, NÚMERO DE DÍGITOS)** arredonda um número para baixo, em direção a zero.
* A função **YEARFRAC (STARTDATE, ENDDATE, [BASE])** retorna à fração do ano que representa o número de dias inteiros entre START\_DATE e END\_DATE.
* A função **TODAY ()** retorna a data atual no formato DATETIME.
* A função **AVERAGEX(<TABLE>, <EXPRESSION>)** calcula a média aritmética de um conjunto de expressões avaliadas de uma tabela.
* A função **SUMX(<TABLE>, <EXPRESSION>)** retorna a soma de uma expressão avaliada para cada linha de uma tabela.
* A função **SUM (<COLUMN>)** adiciona todos os números de uma coluna.
* A função **AVERAGE(<COLUMN>)** retorna a média aritmética de todos os números de uma coluna.

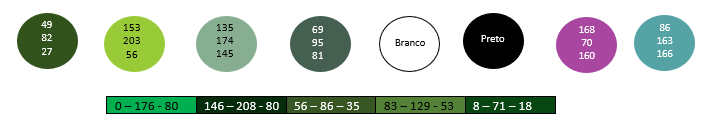
## Etapa 8 - (Pré) Validação dos resultados com os clientes

Essa etapa de pré validação dos resultados com os clientes, para realizar correções e últimos ajustes, não foi feita, pois se trata de um projeto fictício com objetivo de demonstrar as habilidades de análise de dados utilizando Excel e Power BI.

## Etapa 9 - Storytelling dos processos e resultados

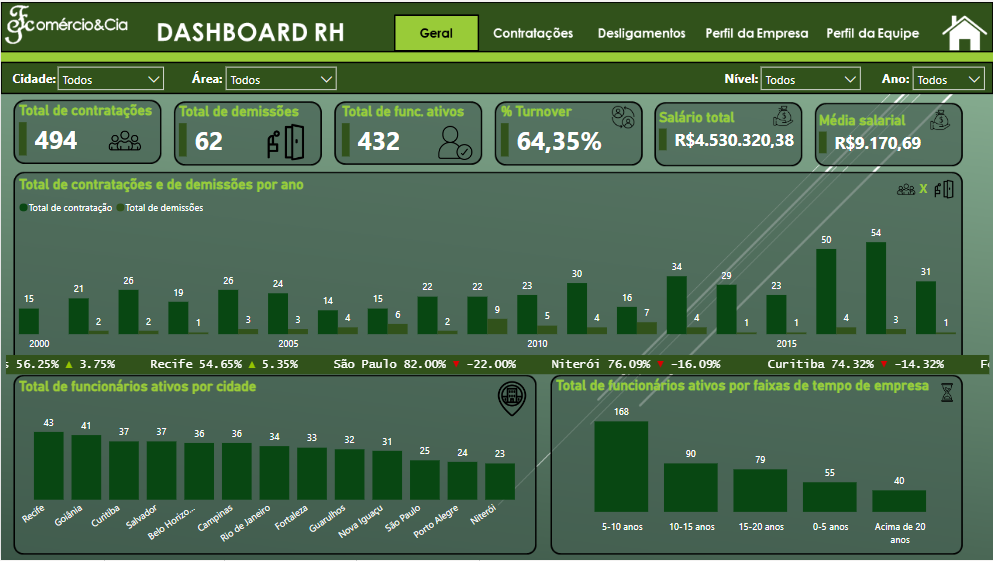
**Etapas realizadas no Power BI**

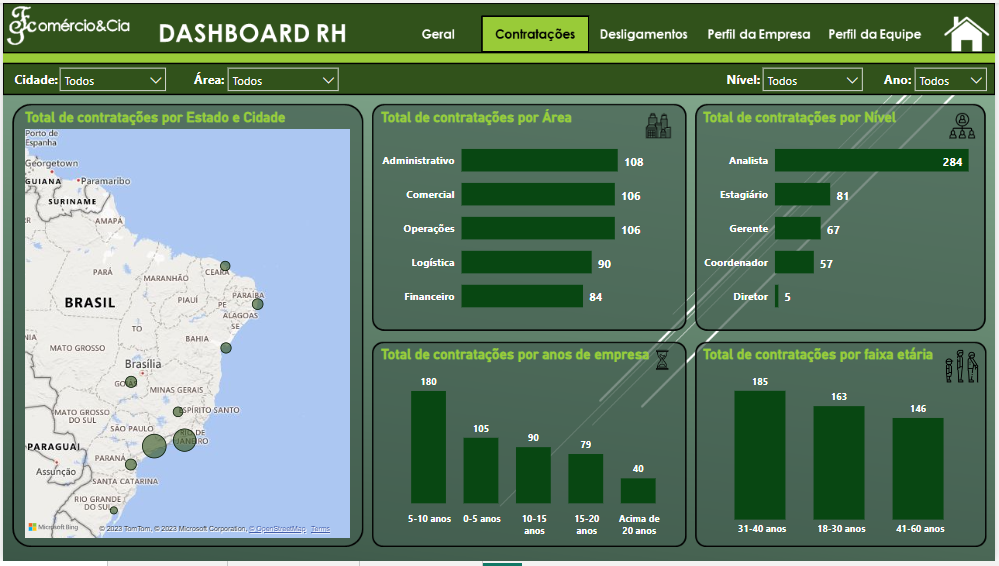
1. Definição das cores a serem utilizadas no Dashboards

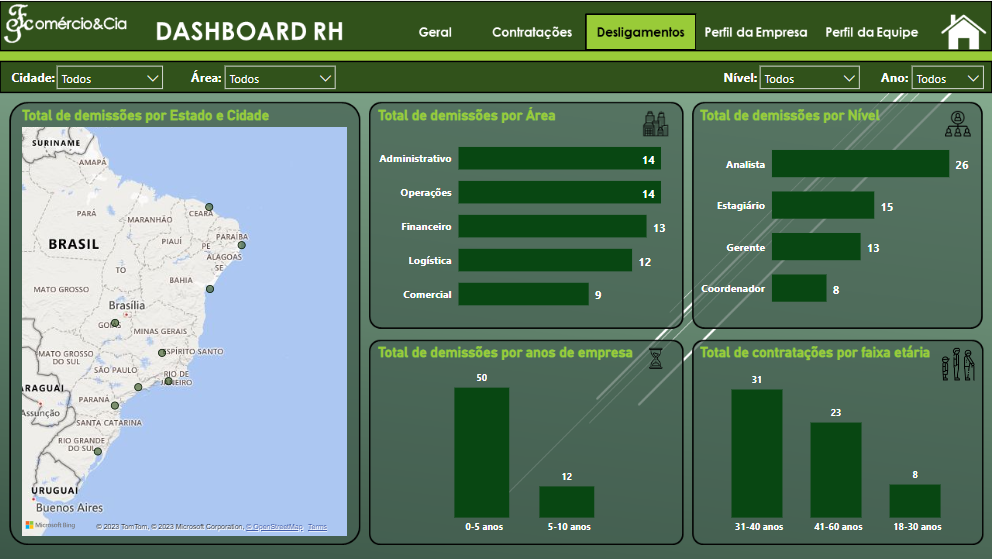


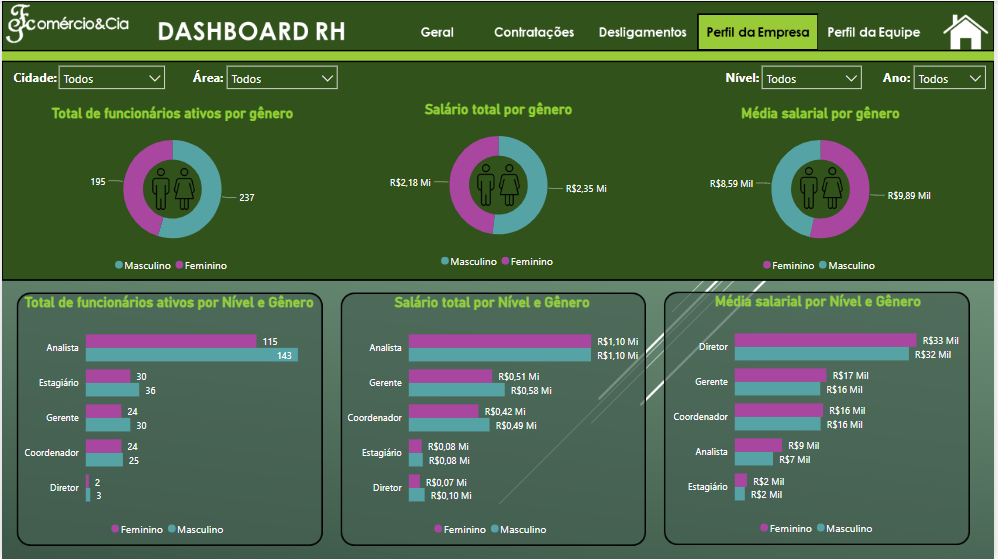
1. Criação das 5 guias: Geral, Contratações, Desligamentos, Perfil da empresa e Perfil da equipe
2. Criação das 5 telas de fundo no PowerPoint
3. Criação da tabela \_Medida
4. Criação das medidas, utilizando a linguagem DAX
   1. % Turnover
   2. Desvio da meta
   3. Média das notas das competências
   4. Média salarial
   5. Nota Geral
   6. Nota Individual
   7. Salário total
   8. Total de contratação
   9. Total de demissões
   10. Total de funcionário ativos
5. Criação de 3 colunas calculadas na tabela FATOFUNCIONÁRIO:
   1. Anos de empresa
   2. Faixa de Idade
   3. Faixas de tempo de empresa
6. Criação das segmentações de dados para as 5 guias: Filtro Cidade, Filtro Área, Filtro Nível e Filtro Ano
7. Criação dos gráficos e cartões de linhas múltiplas
8. Inclusão de ícones no Dashboard.

**DASHBORDS:**











Em conclusão, este projeto demonstra como podemos utilizar o Microsoft Power BI para extrair e analisar dados de um arquivo do Microsoft Excel, que podem ser utilizados ​​para ajudar nas decisões internas para melhorar o ambiente trabalho, oportunidades de crescimento dos colaboradores, um maior nível de satisfação e a retenção de talentos.

**Exemplos de insights após análise:**

**Distribuição geográfica:** Os resultados mostraram que a cidade de Recife possui uma maior concentração de funcionários ativos. O RH pode verificar com a filial de Recife quais são suas particularidades e tentar adaptar e reproduzir para as outras cidades, principalmente para a cidade de Niterói com o menor número de colaboradores.

**Taxa de Turnover:** Está acima da meta estabelecida de 60% indicando que há um número significativo de funcionários deixando a empresa. Isso pode ser um indicativo de problemas relacionados ao ambiente de trabalho, benefícios, oportunidades de crescimento ou outros fatores que influenciam a retenção de talentos. O RH pode investigar as causas desse alto turnover e tomar medidas para melhorar a situação. Além disso, de acordo com o SCROLLER, as regiões com as maiores taxas são as cidades de São Paulo, Curitiba e Niterói, com isso, essa investigação poderia iniciar nessas cidades.

**Tempo de empresa:** A informação sobre o número de funcionários ativos por faixa de tempo de empresa mostra que a maior parte dos colaboradores está entre 5 a 10 anos de experiência na empresa. Esse dado pode ser útil para identificar possíveis problemas de retenção de talentos em períodos específicos da trajetória profissional dos colaboradores. Pode indicar a necessidade de programa de desenvolvimento de carreira para garantir que os funcionários se sintam mais valorizados e tenham oportunidades de crescimento, principalmente nos 5 primeiros anos, período onde ocorre mais desligamentos. Além disso, pode indicar problemas relacionados à integração, treinamento ou satisfação inicial dos colaboradores.

**Contratações por área e hierarquia:** Os resultados mostram que as áreas administrativas e operacionais são as áreas que mais contratam, porém são as áreas também que mais demitem. Investigar os motivos dos desligamentos podem fazer com que as taxas de turnover diminuam consideravelmente, através da implementação de medidas para melhorar a retenção de talentos. Além disso, o mesmo ocorre com o nível hierárquico dos Analistas, categoria com maior número de funcionários e com maior número de desligamento também. Identificar as necessidades desse nível e buscar um maior equilíbrio hierárquico pode ser útil para o planejamento de contratações futuras e um melhor desempenho operacional.

**Diferenças de gênero e salários:** Os resultados mostram um número maior na contratação do gênero masculino em relação ao gênero feminino e consequentemente os maiores salários. Analisar essas diferenças pode exigir uma revisão das políticas e práticas de remuneração e contratação, com intuito de garantir a igualdade de oportunidades e a eliminação de possíveis vieses de gênero.