

Universidade Presbiteriana Mackenzie

Datatrends Insights

Storytelling

2023

Integrantes

LEONARDO DOS REIS OLHER - TIA: 22510249;

THAINÁ VIEIRA DOS SANTOS - TIA: 22500081;


VINICIUS CAUMO SEGATTO - TIA: 22506861;

NICOLAS PINOTTI - TIA 22514112;

VINÍCIUS VIEIRA DA CUNHA OLIVEIRA - TIA: 22505865.



Sumário

- 1) Introdução e Problemática
 - 2) O storytelling
 - 3) Dados Utilizados e Abordagem Analítica
 - 4) Equipes-Chave
 - 5) Análise Exploratória de Dados
 - 6) Aprendizado de Máquina: Métodos Aplicados
 - 7) Entrega do Relatório Inicial e Aprimoramento da Narrativa
 - 8) Possível modelo de negócios
 - 9) Conclusão
- 
- A decorative graphic in the bottom right corner consisting of numerous thin, overlapping, wavy lines in a light red or pink color, creating a sense of motion and depth.

1) Introdução e Problemática

Introdução

- **DataTrend Insights** atua em consultoria em remuneração e desenvolvimento de soluções de software para gestão de remuneração em Ciência de Dados.
- Oferece análises de mercado, faixas salariais e recomendações específicas para clientes.

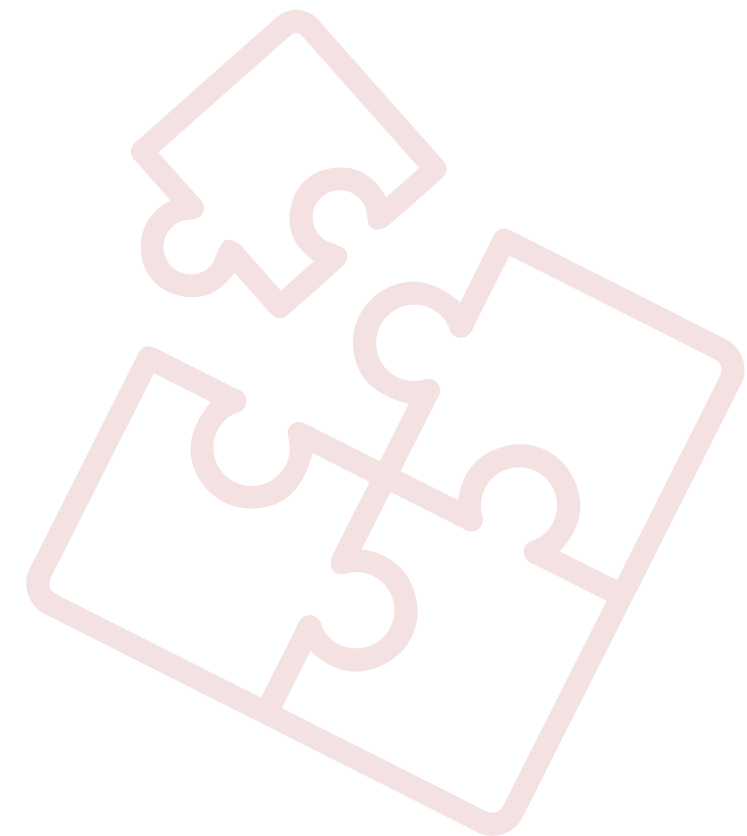
Problemática

- Desafio de retenção de Cientistas de Dados Júnior.
- Perda de talentos para concorrentes devido a discrepâncias salariais e benefícios.
- Necessidade de identificar faixa salarial competitiva em 2023 e desenvolver estratégia eficaz de retenção.

2) O storytelling

Aplicação da narrativa de **storytelling** para comunicar de forma envolvente o processo de análise e ação.

A Ciência de Dados ganha vida através do **storytelling**, uma poderosa ferramenta que transforma números em narrativas envolventes.



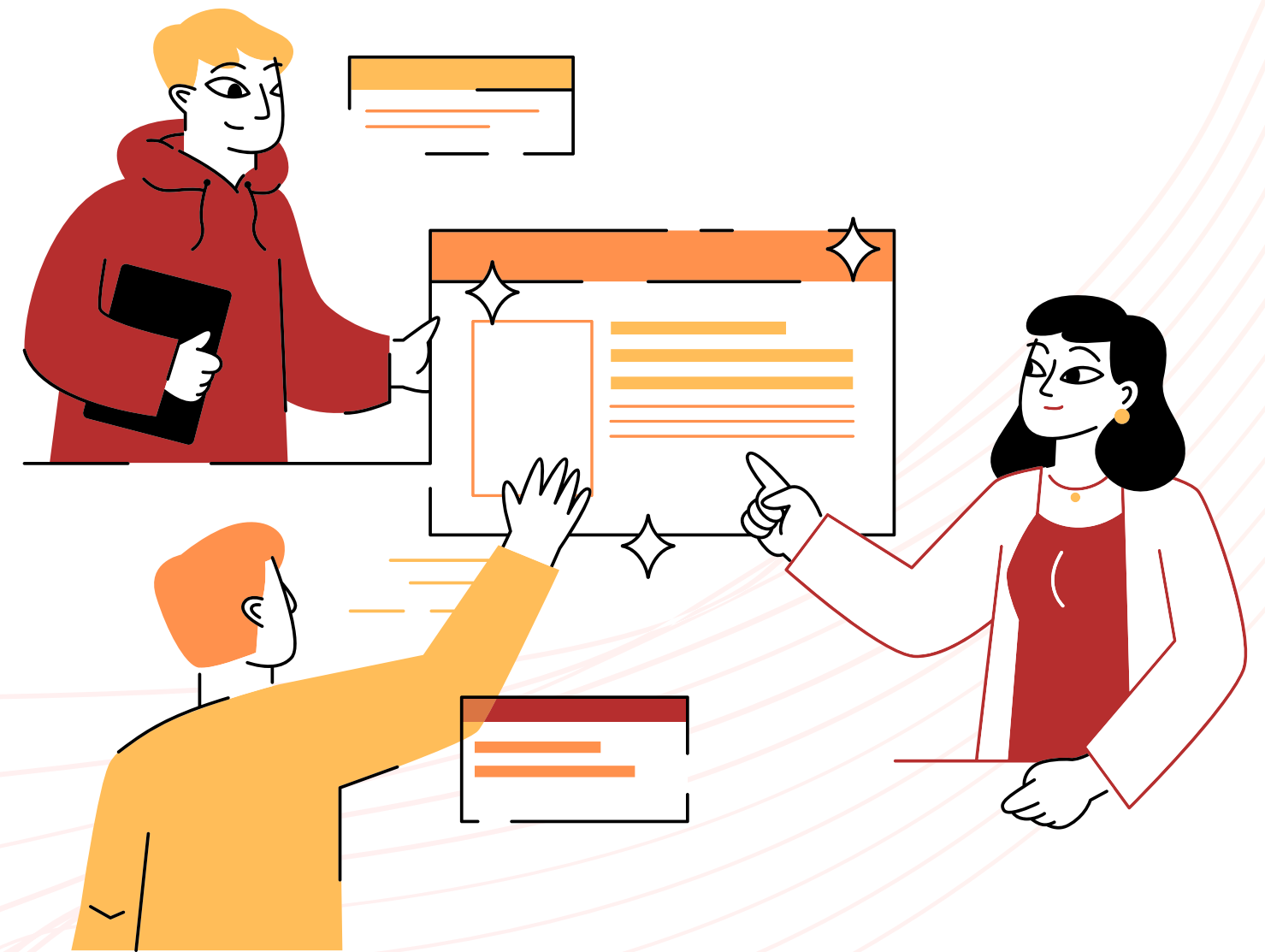
3) Dados Utilizados e Abordagem Analítica

- Conjunto de dados abrangente como espinha dorsal das análises de tendências salariais em Ciência de Dados em 2023.
- Componentes essenciais incluem Ano de Trabalho, Nível de Experiência, Tipo de Emprego, Cargo, Salário, Moeda do Salário, Salário em USD, Localização da Empresa e Tamanho da Empresa.
- Salário não é apenas uma variável numérica, mas uma variável contínua.
- Consideração de variáveis textuais é crucial para explorar relacionamentos e tendências significativas.

4) Equipes-Chave

Equipe de Análise de Dados (EAD):

- Cientistas de Dados, Analistas e Programadores.
- Responsáveis pela coleta, limpeza, preparação e análise aprofundada do conjunto de dados.
- Fornecem os insights fundamentais que fundamentam toda a narrativa.



4) Equipes-Chave

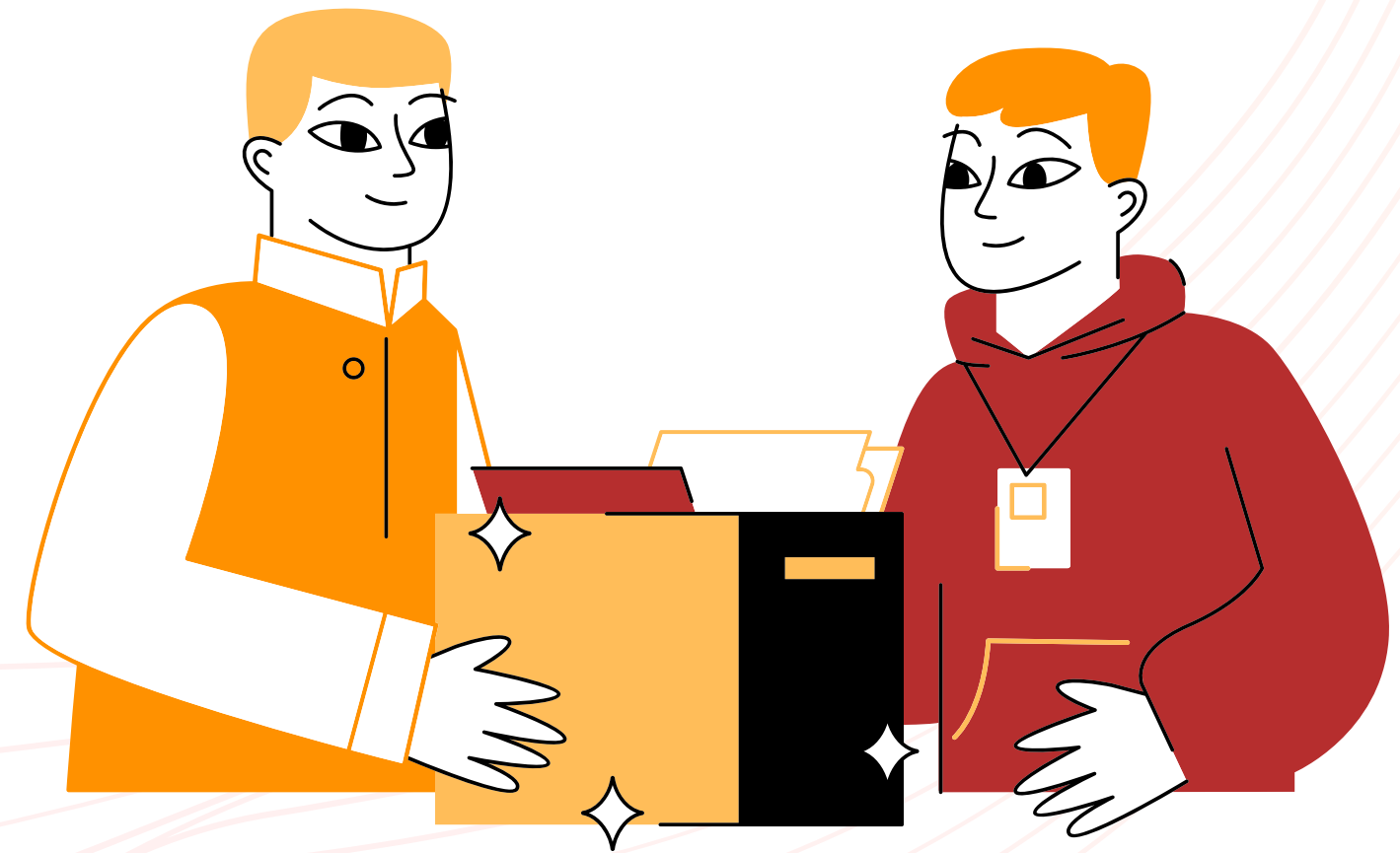
Equipe de Comunicação de Dados (ECD):

- Especialistas em Storytelling, Designers e Analistas de BI.
- Transformam insights em uma narrativa envolvente.
- Utilizam visualizações, gráficos e explicações claras para comunicar resultados aos stakeholders.



5) Análise Exploratória de Dados

- A AED é o pilar essencial para o sucesso do projeto.
- Fornece a base para decisões e estratégias informadas.



5) Análise Exploratória de Dados

Grupo 1: Carregamento e Informações Iniciais

- Garante entendimento da estrutura e qualidade dos dados.

Grupo 2: Limpeza e Transformação de Dados:

- Prepara conjunto de dados para análises avançadas.

Grupo 3: Transformações de Variáveis Categóricas:

- Facilita interpretação e análises mais profundas.

Grupo 4: Visualização dos Dados:

- Identifica padrões e tendências essenciais.
- Direciona decisões estratégicas.

5) Análise Exploratória de Dados

Benefícios para as Equipes:

- **Equipe de Análise de Dados (EAD):**
 - Insights valiosos para análises futuras.
- **Equipe de Comunicação de Dados (ECD):**
 - Base sólida para storytelling envolvente.
- **Desenvolvimento:**
 - Guiado pelo design de ferramentas analíticas.
- **Recursos Humanos:**
 - Informações cruciais para estratégias de retenção.
- **Liderança:**
 - Decisões estratégicas fundamentadas em dados confiáveis.

Conclusão

- A AED não só revela insights essenciais, mas também **capacita equipes a conduzir análises avançadas** e implementar estratégias impactantes de retenção de talentos.

6) Aprendizado de Máquina: Métodos Aplicados

- Aplicamos técnicas de Machine Learning para previsões e insights sobre salários em empregos de Data Science.

Abordagens Utilizadas:

1. **k-Nearest Neighbors (k-NN):**

- Técnica supervisionada para entender proximidade entre salários de diferentes empregos de Data Science.
- Ajuda a identificar padrões de similaridade entre cargos.

2. **Random Forest:**

- Modelo de aprendizado de máquina menos propenso a overfitting.
- Combinação de várias árvores de decisão para resultados mais robustos.
- Enfoque em reduzir viés e variância para previsões precisas.

6) Aprendizado de Máquina: Métodos Aplicados

Detalhes sobre Random Forest

- Utilização de métricas como impureza de Gini, ganho de informação e erro quadrático médio (MSE) para avaliar qualidade das divisões.
- Hiperparâmetros cruciais: tamanho dos nós, quantidade de árvores aleatórias, e tamanho da amostra.

Processo de Treinamento

1. Pré-processamento de Dados para AM:

- Normalização e tratamento de variáveis categóricas.

2. Divisão do Conjunto de Dados:

- Separação entre dados de treinamento e teste.

3. Treinamento do Modelo:

- Aplicação do algoritmo Random Forest.
- Ajuste de hiperparâmetros para otimização do desempenho.

4. Avaliação do Modelo:

- Métricas como precisão, recall, e F1-score para avaliar performance.

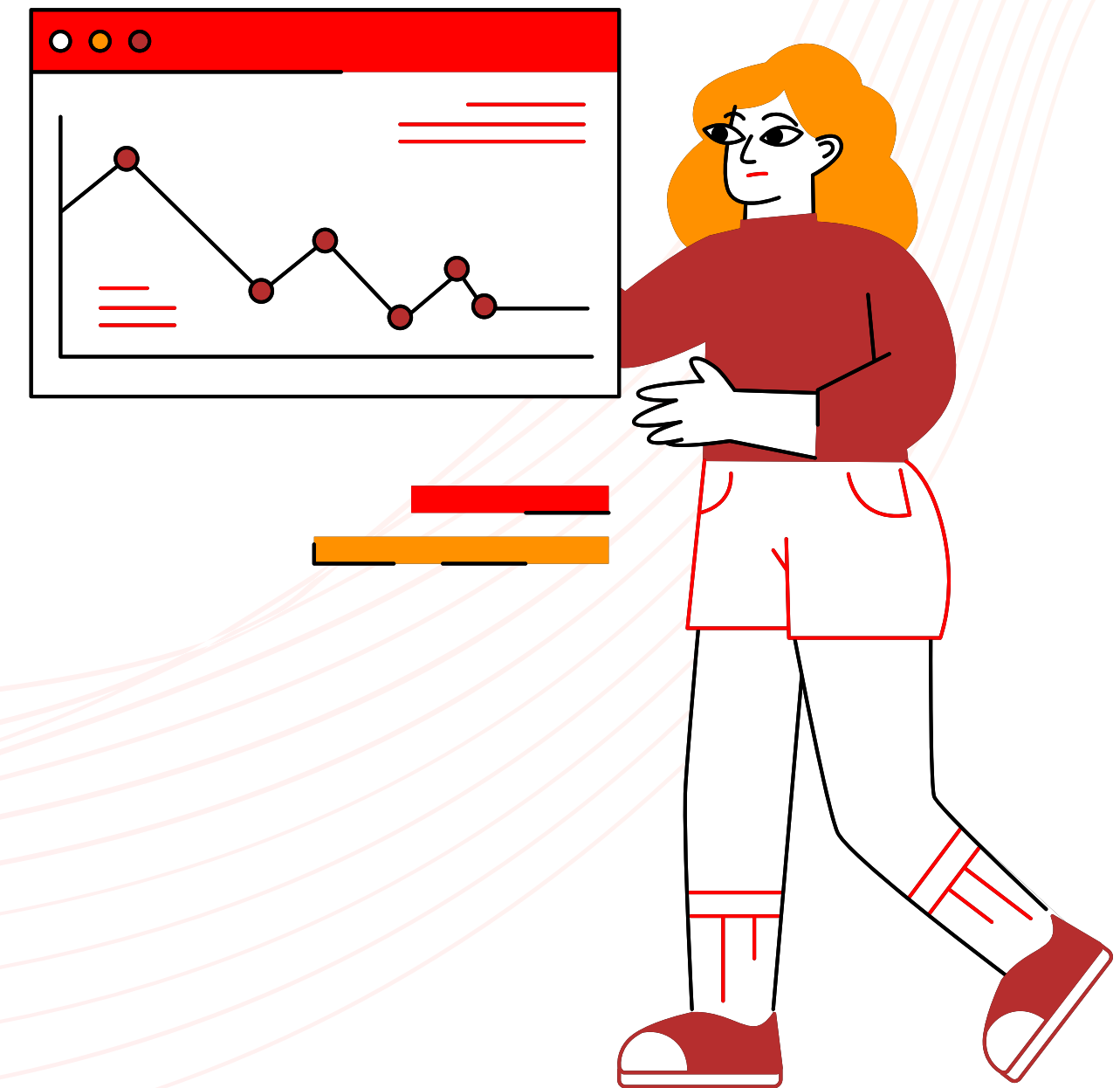
6) Aprendizado de Máquina: Métodos Aplicados

Benefícios para as Equipes

- **Equipe de Análise de Dados (EAD):**
 - Utilização avançada de técnicas preditivas.
- **Equipe de Desenvolvimento:**
 - Integração de resultados em ferramentas analíticas.

Conclusão

- A combinação de k-NN e Random Forest proporciona uma visão abrangente dos padrões salariais em Data Science, capacitando as equipes com previsões precisas e insights relevantes.



7) Entrega do Relatório Inicial e Aprimoramento da Narrativa

Principais Destaques

1. Insights Relevantes:

- Destacamos os insights mais significativos obtidos durante a análise.
- Incluímos visualizações impactantes para facilitar a compreensão.

2. Previsões do Machine Learning:

- Apresentamos as previsões geradas pelo k-Nearest Neighbors e Random Forest.
- Destacamos a confiabilidade e a utilidade das previsões para estratégias futuras.

3. Estratégias de Retenção de Talentos:

- Propomos estratégias específicas com base nos dados e previsões.
- Consideramos não apenas os aspectos salariais, mas também outros fatores identificados durante a análise.

7) Entrega do Relatório Inicial e Aprimoramento da Narrativa

Feedback Contínuo

- Valorizamos a contribuição das equipes e partes interessadas.
- Solicitamos feedback para aprimorar ainda mais a narrativa e as estratégias apresentadas.

Próximos Passos

- Iniciar discussões sobre a implementação das estratégias propostas.
- Preparar versões futuras do relatório incorporando novos dados e análises.

Benefícios para as Equipes

- **Equipe de Análise de Dados (EAD):**
 - Reconhecimento do impacto das análises.
- **Equipe de Comunicação de Dados (ECD):**
 - Oportunidade de ajustar a narrativa para maior eficácia.
- **Liderança e Recursos Humanos:**
 - Base sólida para decisões estratégicas e implementação de medidas de retenção.

7) Entrega do Relatório Inicial e Aprimoramento da Narrativa

Feedback Contínuo

- Valorizamos a contribuição das equipes e partes interessadas.
- Solicitamos feedback para aprimorar ainda mais a narrativa e as estratégias apresentadas.

Próximos Passos

- Iniciar discussões sobre a implementação das estratégias propostas.
- Preparar versões futuras do relatório incorporando novos dados e análises.

Conclusão

- A entrega do relatório inicial marca uma etapa crucial, e a busca constante por aprimoramento reflete nosso compromisso com a excelência na análise de dados e na comunicação eficaz.

8) Possível Modelo de Negócios

1. Consultoria em Remuneração:

- Serviços personalizados baseados em tendências salariais e estratégias de retenção.

1. Soluções de Software Especializadas:

- Desenvolvimento de ferramentas para análise, modelagem e previsão de tendências salariais em Ciência de Dados.

Benefícios para Clientes:

- Tomada de decisões estratégicas informadas.
- Abordagem personalizada para retenção de talentos.
- Acesso a ferramentas avançadas para gestão de remuneração.

8) Possível Modelo de Negócios

Conclusão

DataTrend Insights não apenas revela insights, mas oferece soluções estratégicas e tecnológicas, destacando-se como um parceiro indispensável na compreensão e otimização das dinâmicas salariais em Ciência de Dados.



9) Conclusão

A DataTrend Insights, ao enfrentar desafios de retenção de talentos, não apenas revelou as complexidades das tendências salariais em Ciência de Dados, mas também ofereceu soluções estratégicas e tecnológicas.

A análise detalhada transformou dados em narrativas significativas, culminando na oferta de consultoria personalizada e ferramentas avançadas. Nossa história é de **inovação, compreensão profunda e compromisso contínuo em moldar o panorama salarial** e impulsionar o sucesso de nossos clientes. A jornada continua, prometendo novos capítulos de descobertas e impacto.

