



**PROPOSTA TÉCNICA E COMERCIAL**

UNIMED SANTOS

# **Estruturação de Dados – UNIMED SANTOS**



**Unimed**   
Santos

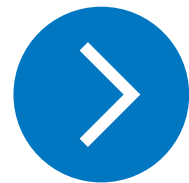
# OBJETIVO DESTA PROPOSTA

- Trabalhar a jornada de Dados da Unimed Santos, internalizando o tratamento, controle, gestão e disponibilização dos dados.
- Atuaremos como parceiro estratégico nas concepção dos Projetos, bem como na execução técnica do trabalho de entrega das mudanças para garantir a visão futura na Jornada de Dados.
- E, para darmos o primeiro passo, a entrega já mapeada desta jornada será a automação e disponibilização da visão comercial de consumo/utilização/dados do contrato do cliente de forma online, eliminando a necessidade de criação manual, por meio do PowerBI.

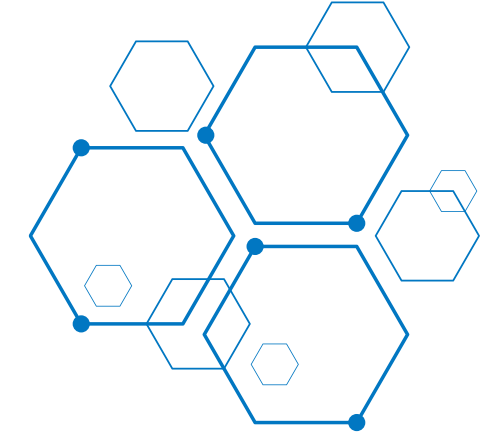


# ESCOPO DA SOLUÇÃO **inicial**





# FLUXO DETALHADO DO PROCESSO



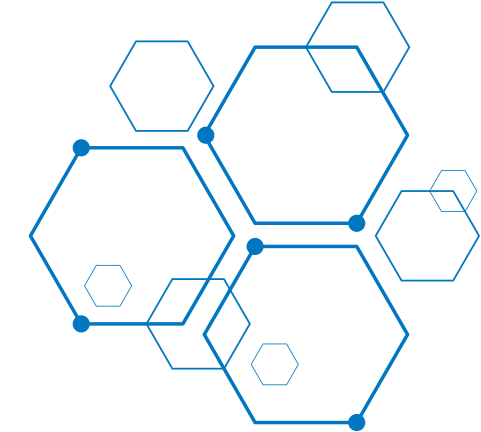
## 1. Extração e Carregamento de Dados (Input Inicial):

- Os dados serão lidos diretamente de uma staging area conectada ao banco de dados principal da UNIMED, proporcionando maior integração e eliminação do processo manual.





# FLUXO DETALHADO DO PROCESSO



## 2. Armazenamento e Processamento no DataLake (ETL na Cloud)

- **Transformação Inicial:**

- Os dados serão carregados do DataLake na estrutura de Cloud.
- Serão feitas a padronização dos dados, incluindo tratamentos como:
  1. Remoção de inconsistências e duplicidades.
  2. Validação de formatos e valores.
- Conversão de dados brutos para um formato adequado para análises e relatórios.

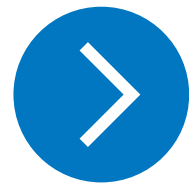
- **Cálculos e Agregações:**

- Serão realizados os cálculos e agregações necessários, como somatórios, médias, índices, e cruzamentos de informações específicas para atender aos requisitos de cada relatório.

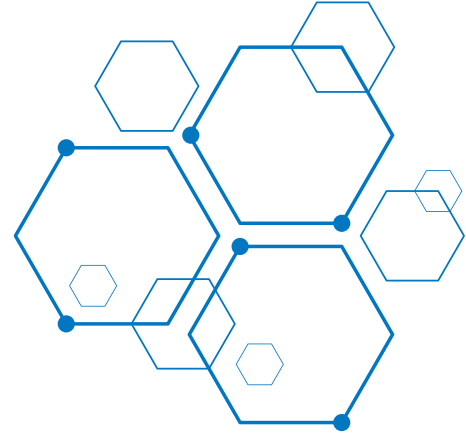
- **Futuro**

- A estrutura será replicada no ambiente do cliente, garantindo domínio completo dos dados pela UNIMED.



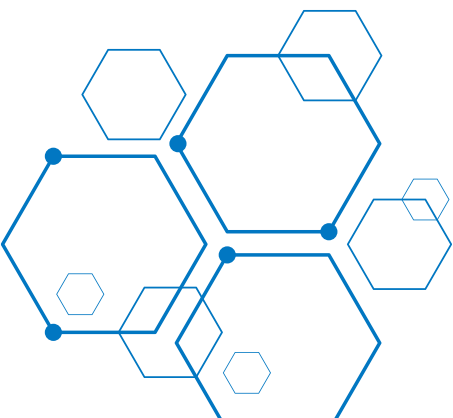


# FLUXO DETALHADO DO PROCESSO



## 3. Armazenamento no DataWarehouse

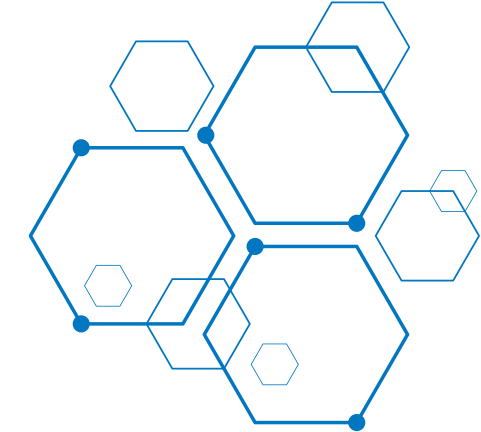
- **Estruturação dos Dados:**
  - Os dados transformados no DataLake serão consolidados em um DataWarehouse.
  - Cada conjunto de informações será organizado por cliente e empresa, assegurando segmentação clara e de fácil acesso.
- **Periodicidade:**
  - Os dados no DataWarehouse serão atualizados diariamente, desde que o processo de extração inicial ocorra no mesmo período.







# FLUXO DETALHADO DO PROCESSO



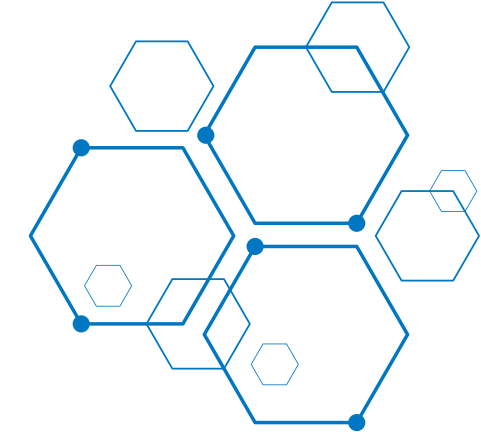
## 4. Consumo e Visualização no Power BI

- **Configuração de Relatórios:**
  - No Power BI, serão criados relatórios interativos que:
    1. Apresentem os dados processados de forma visualmente clara e intuitiva.
    2. Possuam um filtro por empresa, permitindo à diretoria selecionar as informações específicas de cada cliente.
  - O design dos relatórios será feito para destacar métricas-chave, tendências e valores relevantes para decisões estratégicas.
- **Distribuição:**
  - Os relatórios estarão acessíveis aos membros da diretoria, mediante aquisição de licenças específicas do Power BI para visualização e desenvolvimento.
- **Atualização Automática:**
  - Os relatórios no Power BI serão atualizados automaticamente a cada novo ciclo de dados carregados (D-1).





# BENEFÍCIOS DO PROJETO



## 1. Automação do Processo:

- Eliminação da necessidade de solicitações manuais e geração sob demanda.

## 2. Redução de Erros Humanos:

- Dados processados automaticamente, minimizando inconsistências e falhas manuais.

## 3. Velocidade na Geração de Relatórios:

- Relatórios disponibilizados com maior agilidade, reduzindo prazos de entrega.

## 4. Dados Atualizados Diariamente:

- Informações confiáveis e atualizadas D-1, assegurando suporte a decisões estratégicas com base em dados recentes.





# Premissas & Requisitos

## Premissas Técnicas

### 1. Extração e Carregamento de Dados:

- Os dados serão extraídos manualmente e carregados no SharePoint.
- Caso a automação avance para o ambiente de produção, a integração com uma staging area deve ser priorizada.

### 2. Licenciamento do Power BI:\*

- É necessário adquirir licenças do Power BI para o time de desenvolvimento e para cada usuário da diretoria que acessará os relatórios.

### 3. Infraestrutura:\*

- O DataLake e DataWarehouse estarão em uma estrutura de Cloud, que deverá ser fornecida pelo cliente para a replicação no ambiente interno.

### 4. Frequência de Atualização:

- O processo de extração e carregamento de dados no SharePoint deve ser diário para garantir relatórios atualizados.

\* Estes itens estarão, para este momento, sendo gerido pela ModalGR, conforme pagina 22.

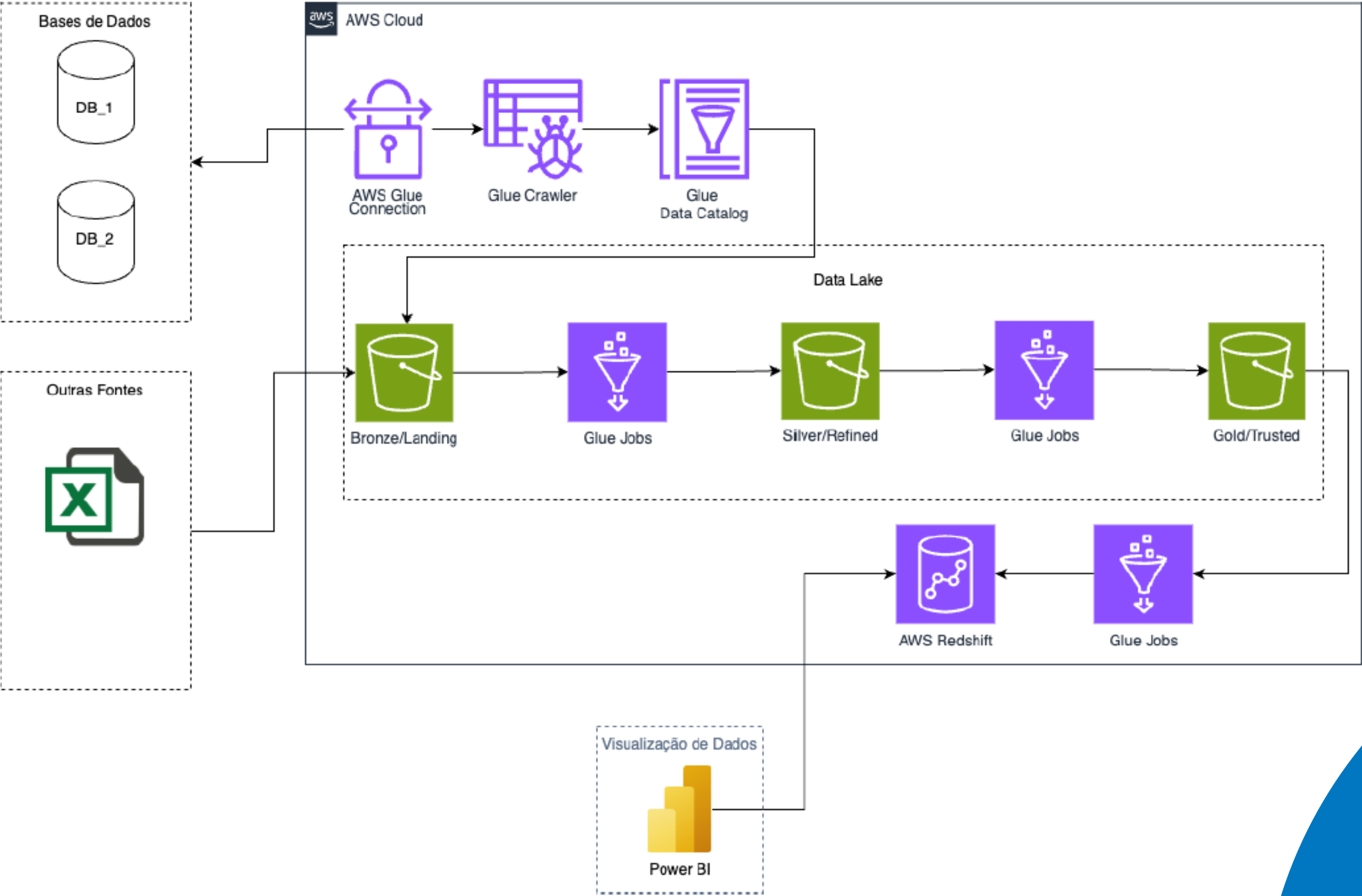


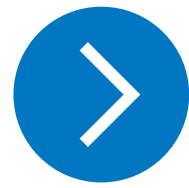
# ARQUITETURA DA SOLUÇÃO



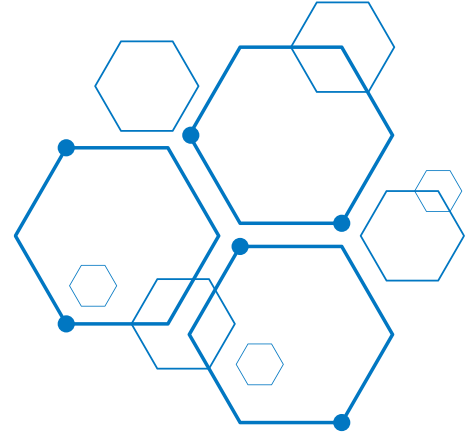


# DIAGRAMA DA SOLUÇÃO





# DIAGRAMA DA SOLUÇÃO



## 1. Fontes de Dados

### 1. Bases de Dados (DB\_1 e DB\_2):

- Representam os sistemas de origem contendo informações relevantes, como sistemas transacionais ou bancos de dados utilizados pela **UNIMED**.
- Os dados são extraídos desses bancos através de uma conexão com o **AWS Glue Connection**.

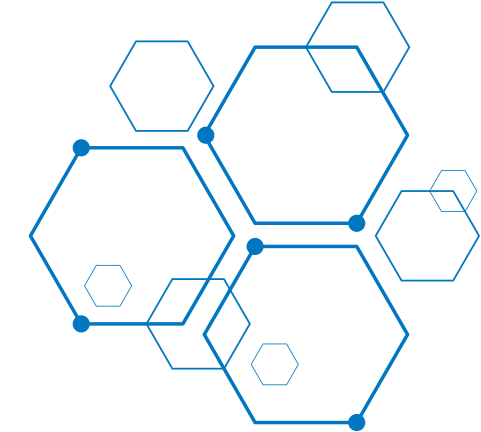
### 2. Outras Fontes (Excel):

- Refere-se a dados que não estão nas bases de dados principais, mas são complementares e fornecidos em planilhas Excel.
- Esses dados são carregados diretamente para o DataLake.





# DIAGRAMA DA SOLUÇÃO



## 2. AWS Glue Connection e Glue Crawler

### 1. AWS Glue Connection:

- Estabelece a conexão segura com os bancos de dados de origem (DB\_1 e DB\_2), permitindo a extração de dados.
- Os dados extraídos são estruturados para serem processados pelo Glue.

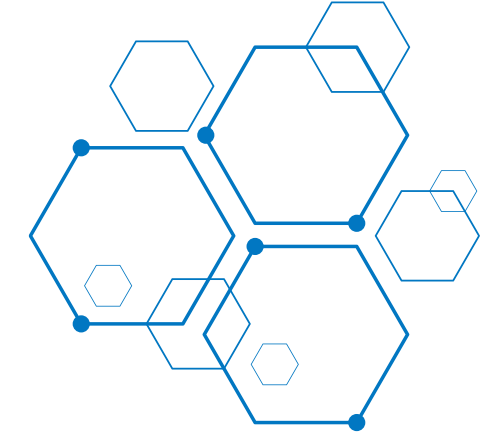
### 2. Glue Crawler:

- Responsável por varrer os dados nas fontes configuradas e gerar automaticamente o **Glue Data Catalog**.
- Esse catálogo cria uma tabela de metadados que facilita o consumo dos dados no DataLake..





# DIAGRAMA DA SOLUÇÃO



## 3. DataLake (Bronze, Silver, Gold)

O DataLake é dividido em três camadas para organização e processamento incremental dos dados:

### 1. Bronze/Landing:

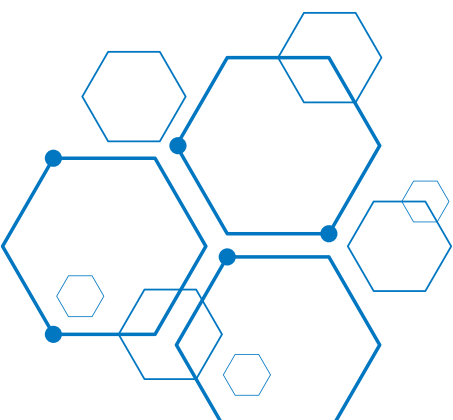
- É a camada de entrada, onde os dados são armazenados exatamente como foram extraídos, sem transformações.
- Permite a retenção de um histórico bruto dos dados, essencial para auditorias ou reproprocessamentos.

### 2. Silver/Refined:

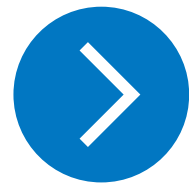
- Nesta camada, os dados passam por tratamentos e transformações básicas.
- Inclui processos como limpeza, padronização e estruturação para análises mais simples.

### 3. Gold/Trusted:

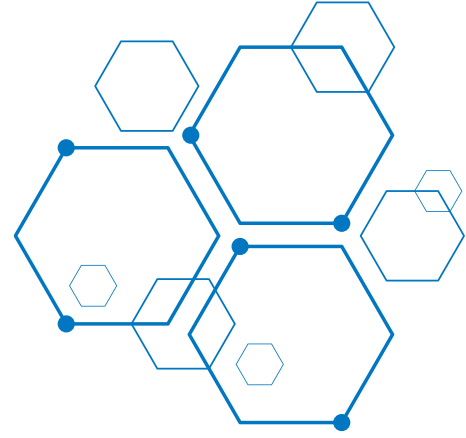
- Representa a camada final de dados confiáveis, prontos para consumo por aplicações analíticas.
- Aqui, os dados já foram agregados, calculados e estão no formato ideal para geração de relatórios.







# DIAGRAMA DA SOLUÇÃO



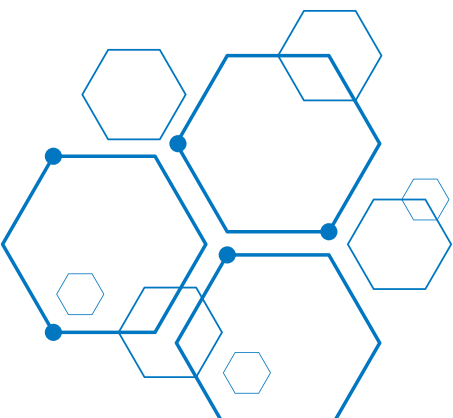
## 4. Glue Jobs

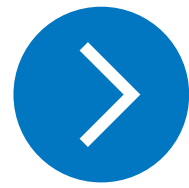
São os processos de transformação de dados que ocorrem entre as camadas do DataLake:

- **Bronze → Silver:** Aplicação de tratamentos básicos, como remoção de duplicidades e ajustes de formatação.
- **Silver → Gold:** Execução de cálculos complexos, agregações e enriquecimento dos dados.

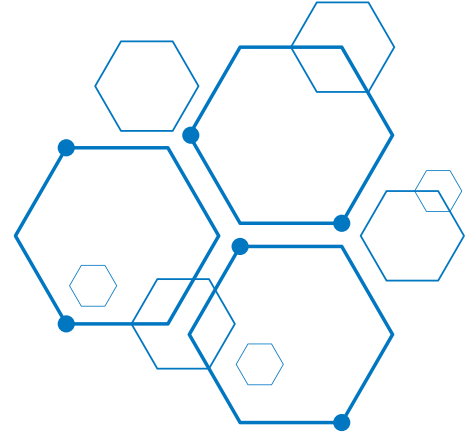
### Integração com AWS Redshift:

- Após as transformações, os Glue Jobs carregam os dados da camada Gold no **AWS Redshift**, que é o DataWarehouse.





# DIAGRAMA DA SOLUÇÃO



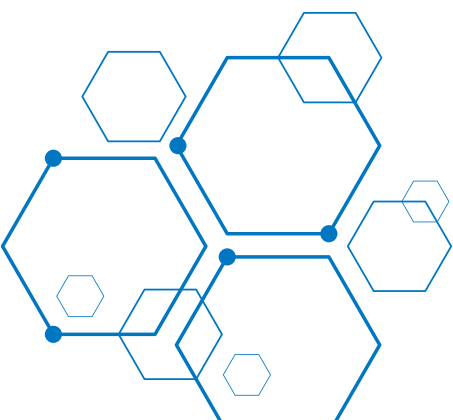
## 5. AWS Redshift

### 1. DataWarehouse:

- Centraliza os dados transformados e refinados para consultas analíticas rápidas.
- Permite a execução de consultas SQL complexas para alimentar relatórios no Power BI.

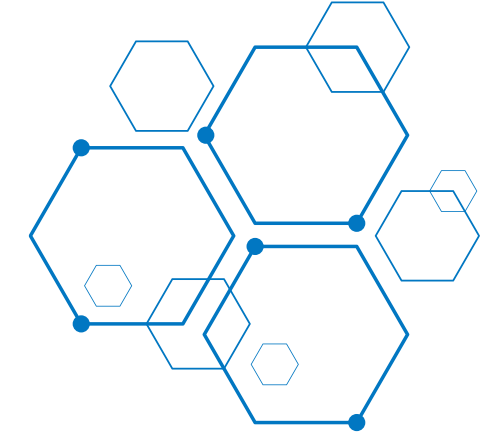
### 2. Interação com Glue Jobs:

- Os dados no Redshift podem ser atualizados continuamente a partir dos Glue Jobs para manter a consistência com a camada Gold.





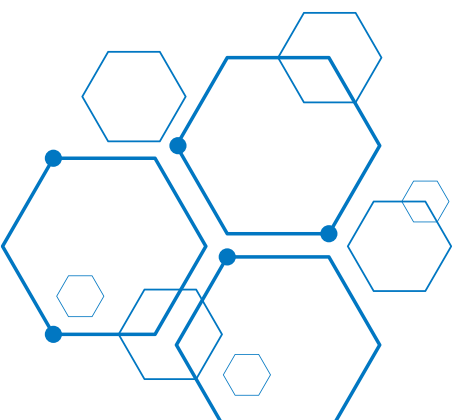
# DIAGRAMA DA SOLUÇÃO

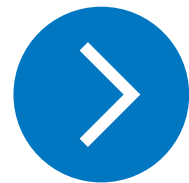


## 6. Power BI

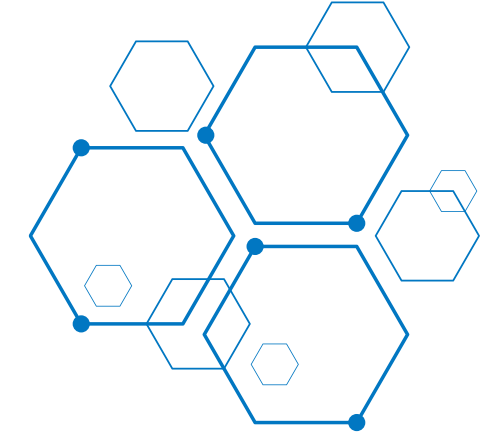
### 1. Visualização de Dados:

- Conecta-se ao **AWS Redshift** para consumir os dados já transformados e disponibilizá-los para a diretoria em forma de relatórios interativos.
- Os relatórios incluem filtros por empresa e cliente, possibilitando análises personalizadas.





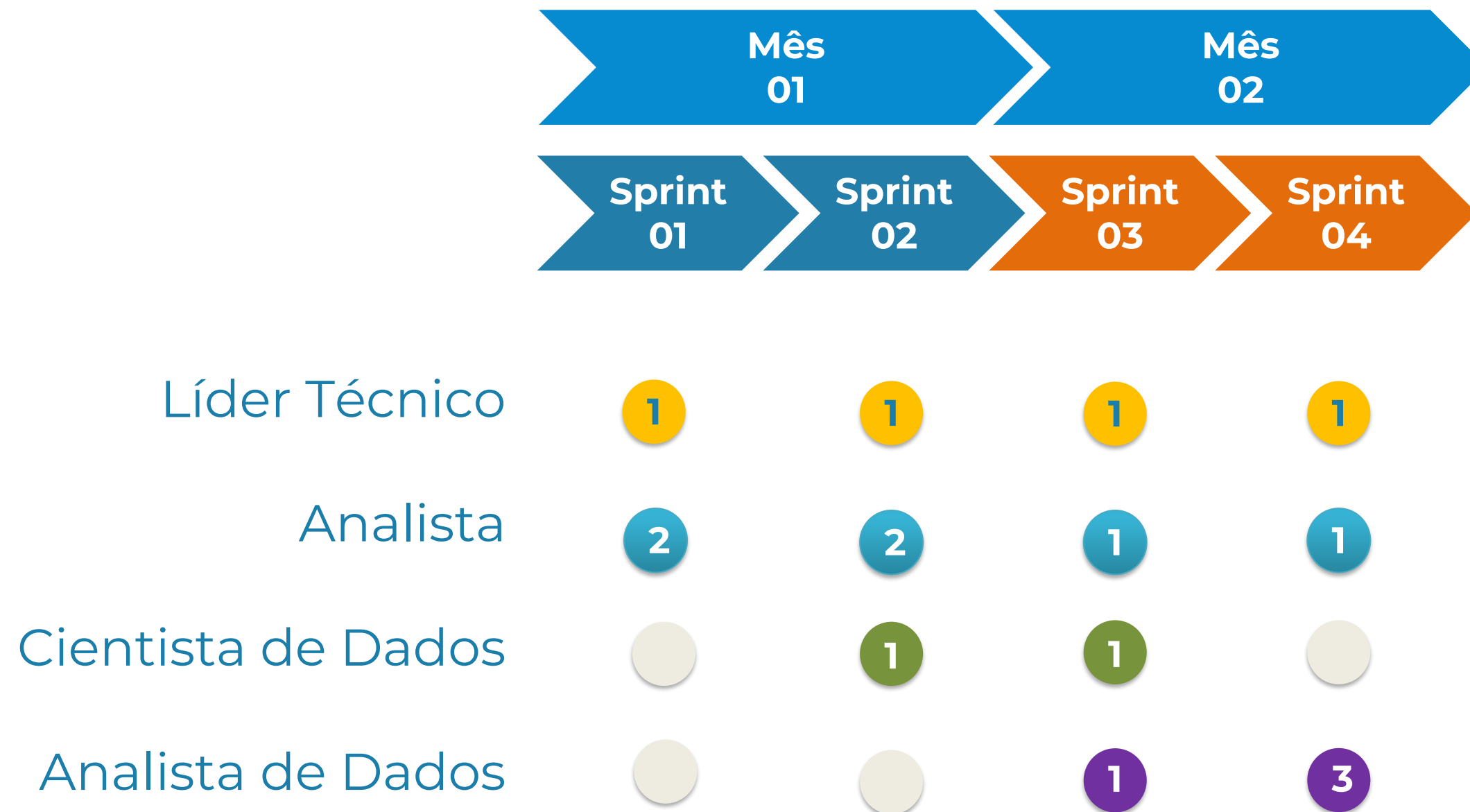
# RESUMO DO FLUXO



1. Dados são extraídos de bancos de dados (DB\_1 e DB\_2) e planilhas Excel.
2. **AWS Glue Connection** conecta-se às fontes, e o **Glue Crawler** organiza os metadados no **Data Catalog**.
3. Dados são armazenados e processados no **DataLake** em três etapas (Bronze, Silver, Gold).
4. Dados refinados (Gold) são carregados no **AWS Redshift** para consultas analíticas.
5. **Power BI** consome os dados do Redshift para gerar relatórios interativos.



# Estruturação de uma Squad

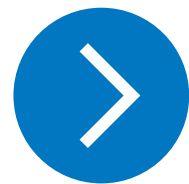




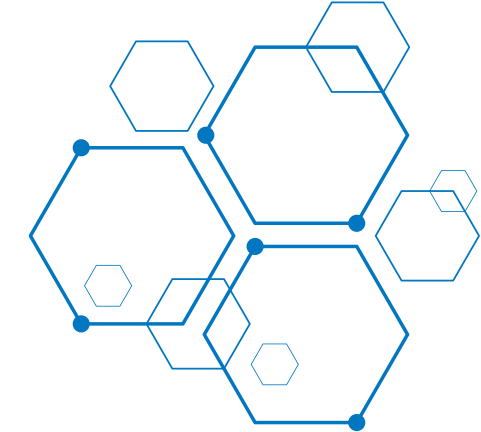
# ESTIMATIVA de TEMPO e ESFORÇO







# ESTIMATIVA DE ESFORÇO



## Esforço da 1ª entrega

Esforço Estimado <sup>1</sup>	400 Horas	Até 2 meses para entrega
Esforço LIMITE <sup>2</sup>	740 Horas	

- **Esforço Estimado<sup>1</sup>** – Não considera riscos, sendo o mais provável, mediante nosso trabalho.
- **Esforço LIMITE<sup>2</sup>** – O esforço maior possível que a Unimed Santos poderá ser cobrada pelo trabalho, após alinhamento da reunião em 21/jan/25.
- **PRAZO DE ENTREGA:** O trabalho será entregue em até 2 meses após seu início, sendo no Kick-Off informada a data mais realista de sua entrega.

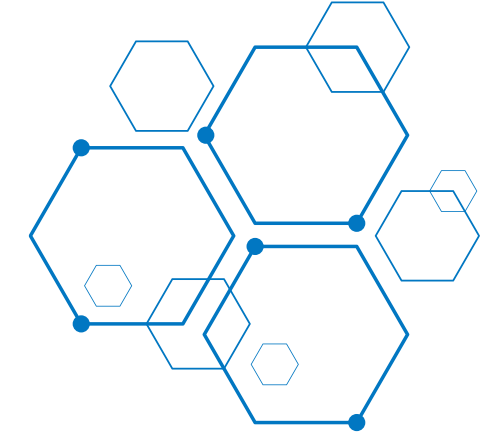


# Custos Adicionais **(Ambiente Cloud/Licenças)**





# Ambiente e Licenças



- Para atender ao modelo inicial, de avaliação e validação financeira, foi colocado na Reunião que a ModalGR poderia assumir a Gestão (em caráter inicial e temporário), a gestão de licenças do PowerBI e ambiente Cloud.
- Com o objetivo de apoiar a Unimed Santos nesta jornada, o valor cobrado será apenas repassado do fornecedor para pagamento pela Unimed Santos, via Nota de Débito. Caso seja por Nota Fiscal, acrescentaremos os custos dos Impostos.
  - Licenças Power BI:  
Custo Licenças: R\$ 59,90/usuario/mes  
Fonte: <https://www.microsoft.com/pt-br/power-platform/products/power-bi/pricing>
  - Cloud Azure: Ambiente/Transação/Repositorio  
Custo: R\$ 1.385,00/mês (Estimado)  
Fonte: <https://azure.microsoft.com/pt-br/pricing/calculator/>
- A Unimed Santos poderá, a qualquer momento, solicitar e executar a transferência do ambiente. A ModalGR apoiará em todo o processo para garantir a qualidade esperada, dentro da presente proposta.



# Impulsione seu negócio com **tecnologia de resultado!**



Escaneie para falar  
com o nosso time



MODALGR®



+55 (13) 4101.0010



comercial@modalgr.io



## MATRIZ BRASIL

R. Visc. de Rio Branco,  
02 - 6º Andar  
Centro, Santos - SP



## PORTUGAL

Rua Fernanda Seno, 6  
7005-485 / Évora

