

Licenciatura em Sistemas Multimédia

**SIG**

*Consultas Cuidados Saude Primários*

*Bruno Faria*

Índice

[1. Sistema Informação para a Gestão 2](#_Toc75127920)

[1.1 ETL 3](#_Toc75127921)

[1.2 Analysis Services 9](#_Toc75127922)

[1.3 Data Mining 14](#_Toc75127923)

[2. Power BI 15](#_Toc75127924)

[3. Referência Bibliográficas Electronicas 17](#_Toc75127925)

# 1. Sistema Informação para a Gestão

Este trabalho tem como fonte de análise dados do serviço nacional de saude que visa avaliar registos da ARS (Administração Regional de Saude) o nº de consultas programadas e não programadas nos cuidados saude primários, o total de consultas nos cuidados saude primários e as consultas em que a gripe é o sintoma nestes mesmos cuidados de saude primários.

Para a análise foi usado um ficheiro CSV baixado no sitio do Serviço nacional de saude, que foi devidamente formatado e transformado por via do power query do excel.

Nos proximos itens será explicado de uma forma genérica todos os passos para transformar converter e carregar os dados. Depois da manipulação destes, os mesmos são usados para análises de indicadores e tendencias utilizando modelos de mineração de dados tais como tree viwer, neural network e clustering. Por fim são mostrados em gráficos assente na plataforma PowerBI para uma apresentação sobre forma de dashboard.

## 1.1 ETL

O Crescimento exponencial dos dados colocou uma grande pressão sobre a arquitetura dos Data warehouse. Nos dias de hoje as organizações lidam com grandes volumes e diferentes tipos de dados, redes sociais, comportamento do cliente e big data.



Figura 1 – Conexões a BD (Extração) e à fonte de dados (Importação)

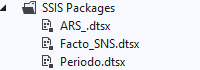


Figura 2 – packages SQL Server integration Service

No mundo empresarial a integração dos dados extrair, transformar, carregar (ETL) ou extrair, carregar, transformar (ELT).

ETL é 1 dos métodos mais populares de carregamento de dados de várias fontes e armazenamento num data warehouse que pode ser consultado por todos os utilizadores de uma organização.

No ETL, os dados são extraídos de fontes distintas, como sistemas ERP e CRM, transformados (cálculos são aplicados, dados brutos são alterados para o formato / tipo necessário, etc.) e, em seguida, carregados para o Data warehouse, também chamado de base de dados de destino.

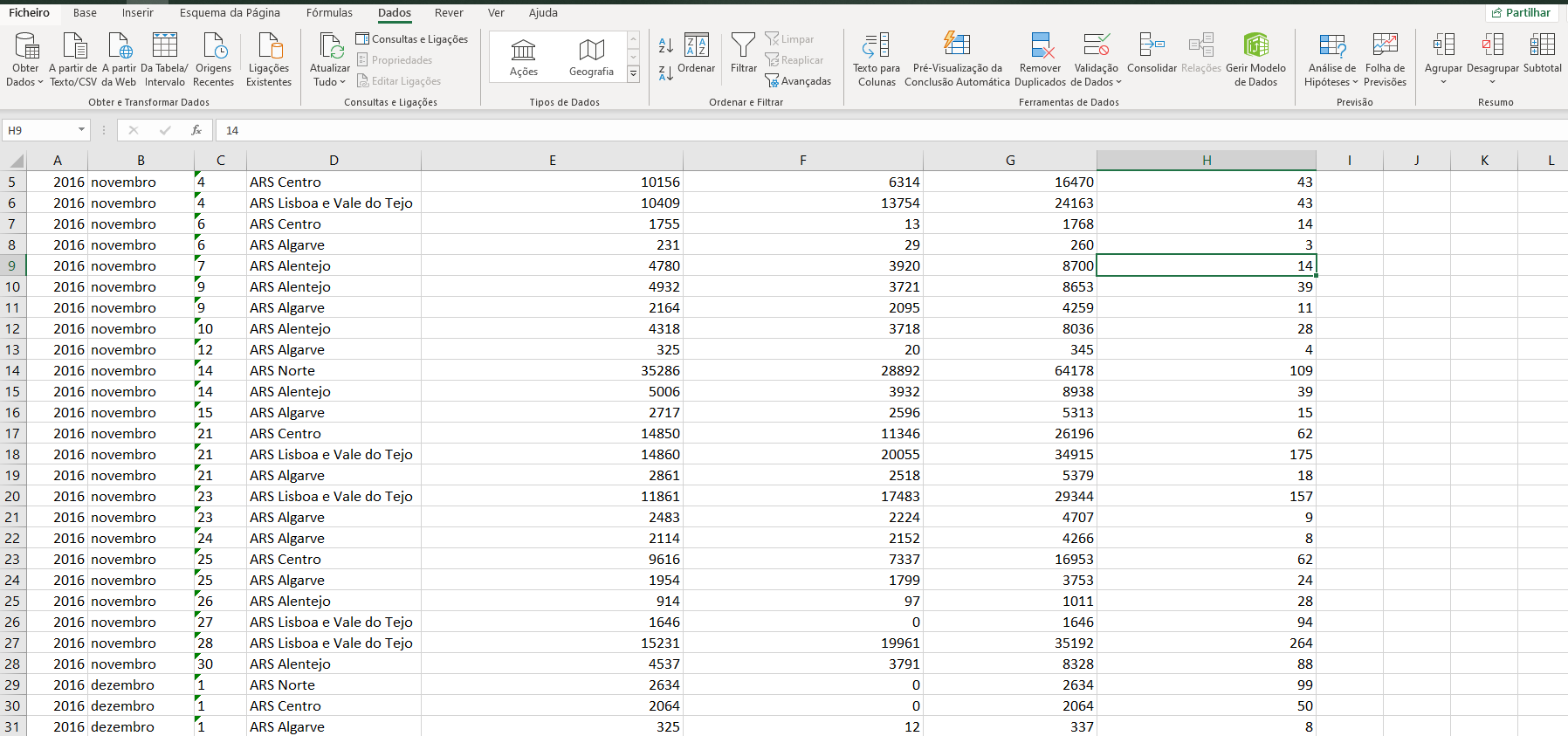


Figura 3 – Ficheiro CSV imporatdo no power query do excel - **origem dos nossos dados.**

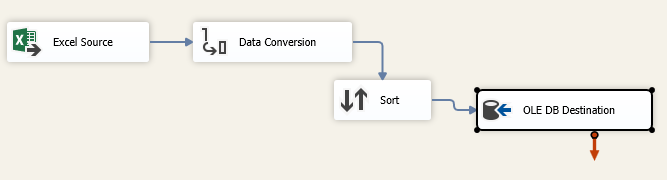
Exemplo: 

Figura 4 – Fluxo de dados de uma tabela Dimensão

Neste exemplo os dados da tabela extraida tem como origem um ficheiro CSV previamente importado formatado e trabalhado no Power-Query do Excel. Aqui está representada uma tabela dimensão, os dados são extraidos, renomeados e convertidos para string, ordenados e depois exportados para o DatawareHouse.

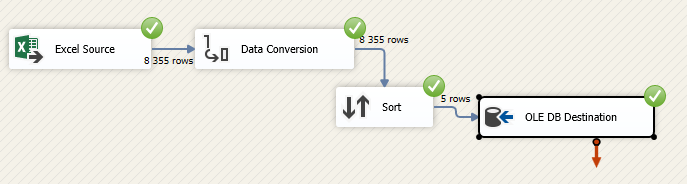
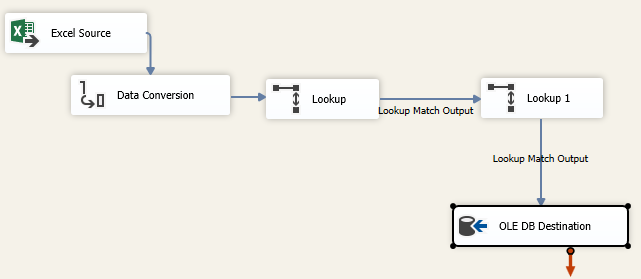


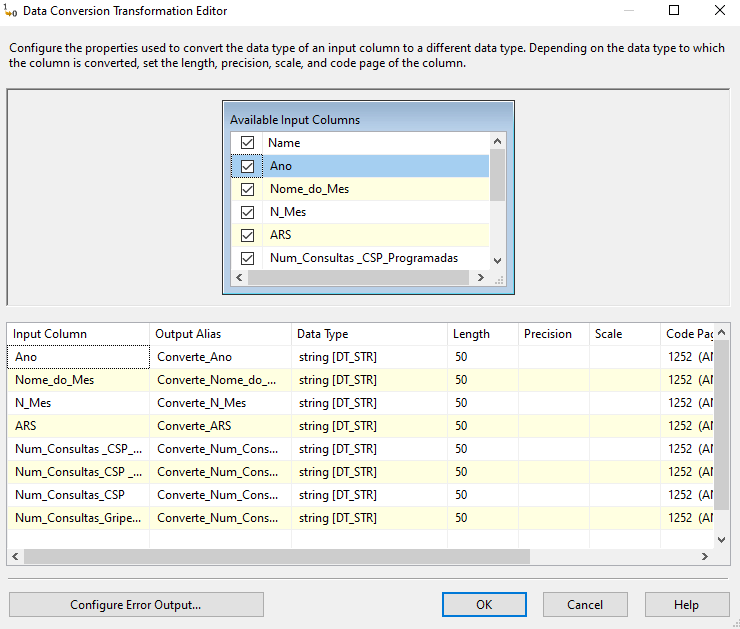
Figura 5 – Fluxo de dados de uma tabela Dimensão- Exportação

Quando falamos de tabelas Facto vs Tabelas dimensão convém desmistificar as suas diferenças:

* A Tabela Fato, como o próprio nome já diz, armazena os fatos históricos de um acontecimento, que pode ser uma venda, um evento, ou qualquer outra situação em que seja possível registar uma ação(quantidades). Nesse sentido, as informações contidas nela repetem-se e normalmente são colocadas periodicamente.
* Já a Tabela Dimensão, armazena as informações de um único item, ou seja, as informações referentes a cada produto e por tanto não podem se repetir. Quando partimos para uma utilização mais analítica das ferramentas e usamos validação de dados por exemplo, essa base exclusiva é chamada de Dimensão. Então, estas tabelas podem ser caso do nosso exemplo, as informações de um único remédio, ou seja, seu fabricante, laboratório não irão se repetir.

Figura 6 – Fluxo de dados de uma tabela Facto

A tabela Facto armazena grande quantidade de dados históricos, em função do tempo, obtidos a partir da intersecção de todas as tabelas dimensões num modelo de dados em estrela. A dimensão tempo (uma das tabelas auxiliares) é sempre integrante da chave primária e é na tabela Facto onde armazenamos os indicadores de desempenho por exemplo de um negócio.

Figura 7 – Conversão de Dados

As ferramentas ETL ajudam a integrar dados para atender aos requisitos de data warehouses tradicionais que são alimentados por cubos de dados de processamento analítico online (OLAP) e / ou tecnologias de sistema de gestão de base de dados relacional (RDBMS). Ferramentas OLAP e consultas de linguagem de consulta estruturada (SQL) dependem da padronização de dimensões entre conjuntos de dados para fornecer resultados agregados. Isso significa que os dados devem passar por uma série de transformações, como:

* Conversão de valores numéricos
* Edição de strings de texto
* Correspondência de linhas e colunas
* Encontrar e substituir itens de dados
* Mudança dos nomes das colunas
* Recombinação de colunas de diferentes tabelas e base de dados
* Pré-cálculo de agregados intermediários

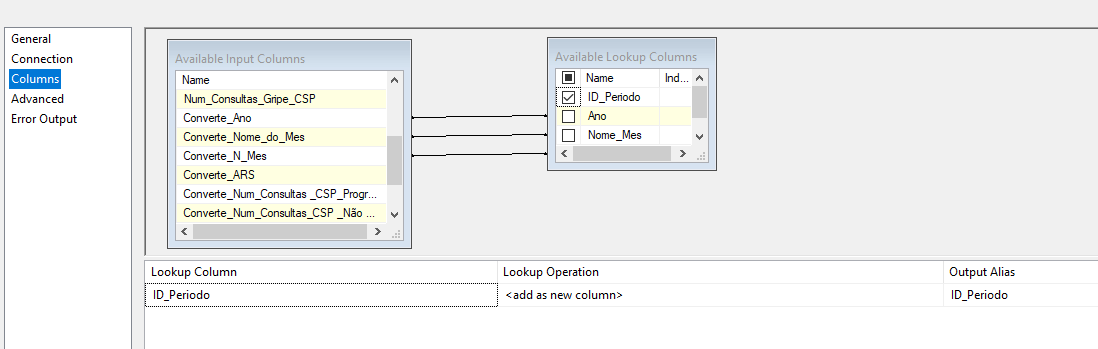


Figura 8 – Transformação permite o desempenho de junções simples entre os dados de entrada e de referência



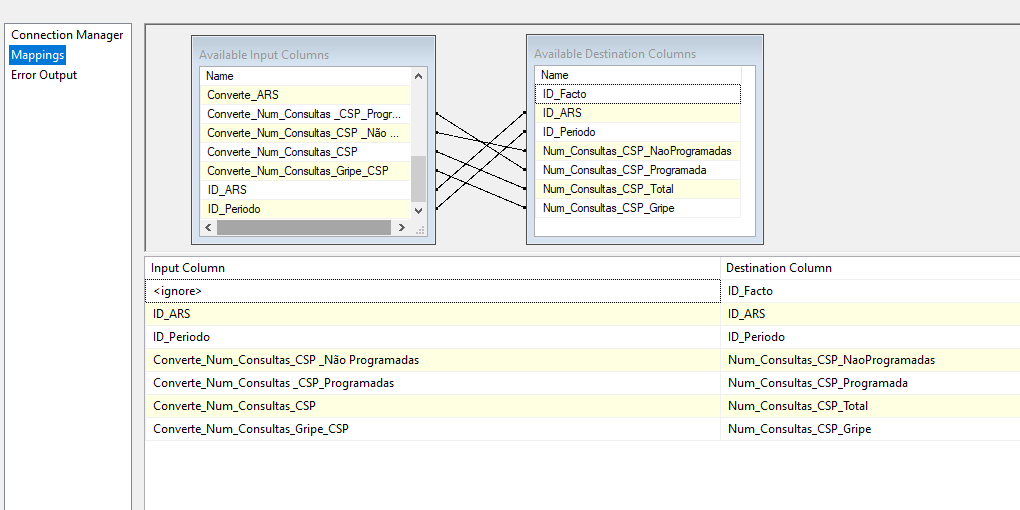


Figura 9 – Exportação dados via OLE DB provider, dados inseridos no DataWareHouse

Para data warehouses tradicionais, essas transformações são realizadas antes de carregar os dados no sistema de destino, normalmente um data warehouse relacional. Este é o processo seguido no ETL.

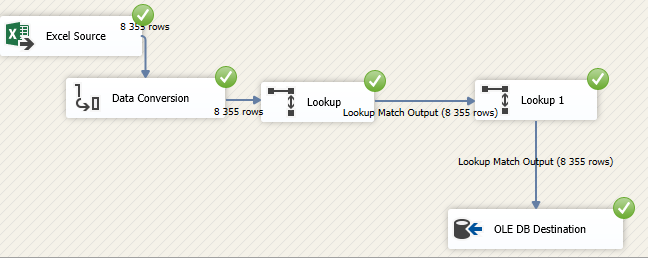


Figura 10 – Fluxo de dados de uma tabela Facto - EXportação

## 1.2 Analysis Services

O Analysis Services é um mecanismo de dados analíticos (VertiPaq) usado no suporte à decisão e análise de negócios. Ele fornece recursos de modelo de dados semânticos de nível empresarial para business intelligence (BI), análise de dados e aplicativos de relatório, como Power BI, Excel, Reporting Services e outras ferramentas de visualização de dados.

Instalado como uma instância de servidor local, o SQL Server Analysis Services oferece suporte a modelos tabulares em todos os níveis de compatibilidade (dependendo da versão), modelos multidimensionais, mineração de dados e Power Pivot para SharePoint.

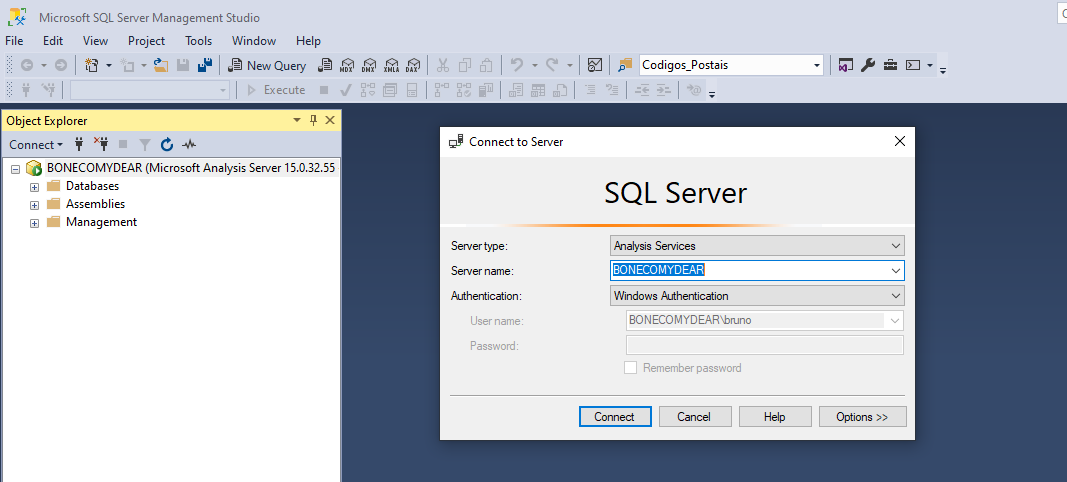


Figura 11 – instância do SQL Server Analysis Services

Um fluxo de trabalho de implementação típico inclui a instalação de uma instância do SQL Server Analysis Services, a criação de um modelo de dados tabular ou multidimensional, a implantação do modelo como uma base de dados numa instância de servidor, o processamento da base de dados para carregá-lo com os dados e a atribuição de permissões para permitir o acesso aos dados. Quando estiver pronto para uso, o modelo de dados pode ser consultado por qualquer aplicativo cliente que suporte o Analysis Services como uma fonte de dados.

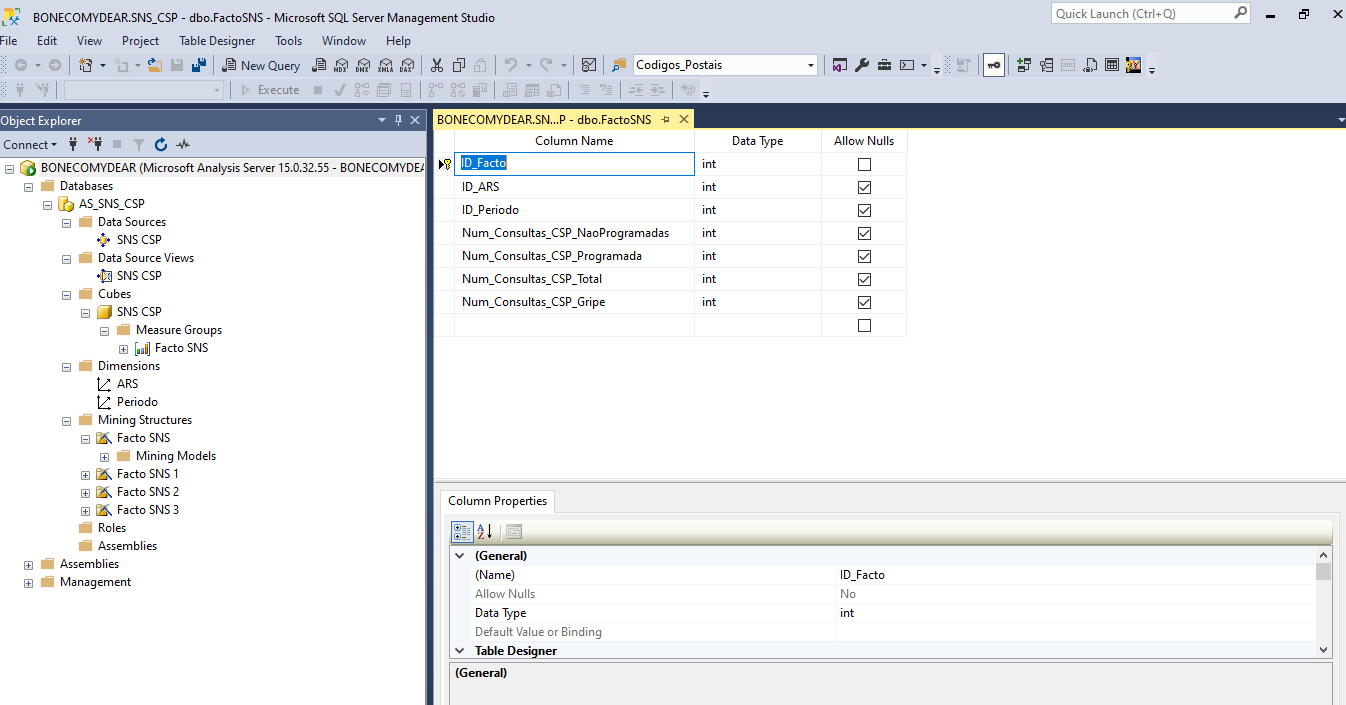


Figura 12 – instância do SQL Server Analysis Services

Para criar um modelo, usamos o Visual Studio com a extensão de projetos do Analysis Services, escolhendo um modelo de projeto Tabular ou Multidimensional. O modelo de projeto contém pastas para todos os objetos necessários num modelo.

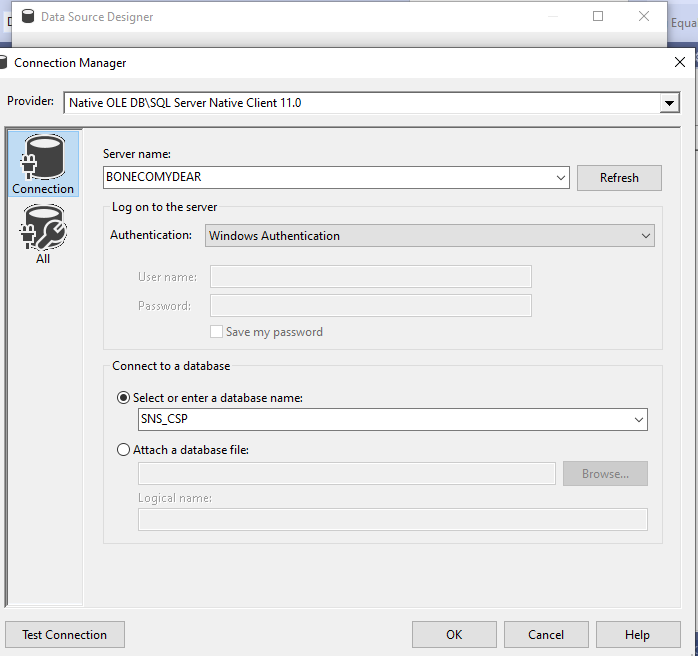


Figura 13 – Data Source Designer – connection manager

Os modelos são preenchidos com dados de sistemas de dados externos, geralmente data warehouses hospedados num mecanismo de base de dados relacional SQL Server. Os modelos especificam objetos de consulta, como cubos, mas também especificam dimensões que podem ser usadas em vários cubos, cálculos

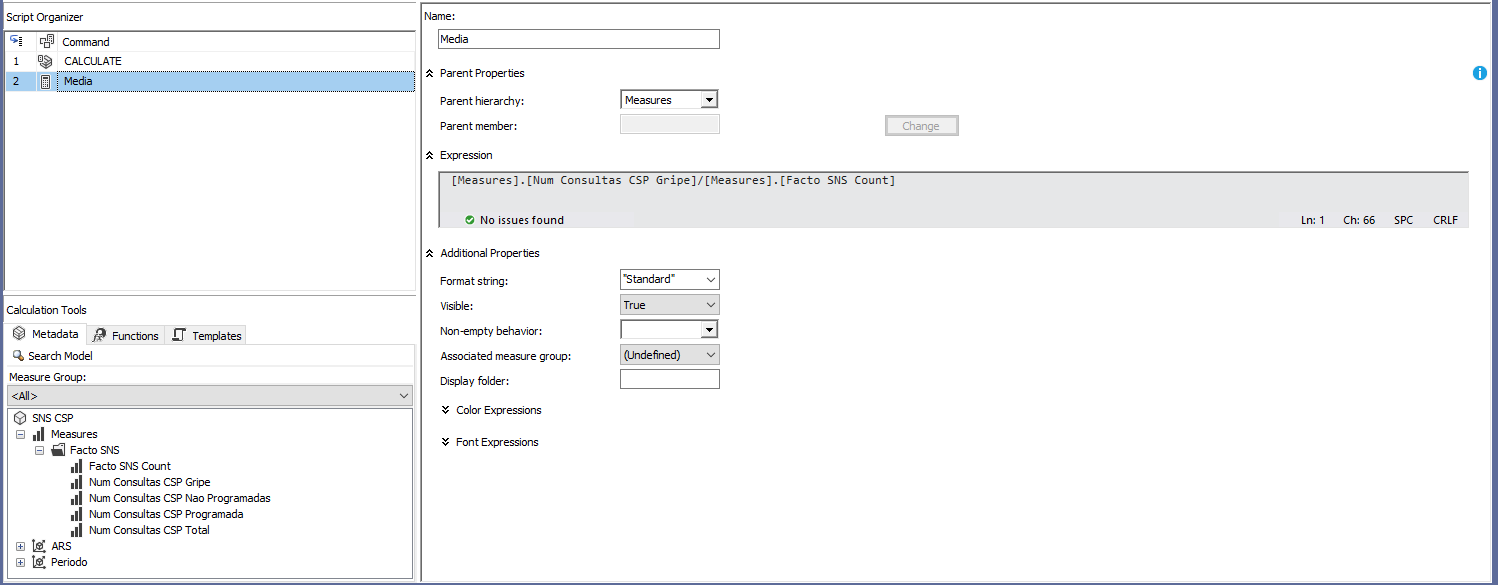
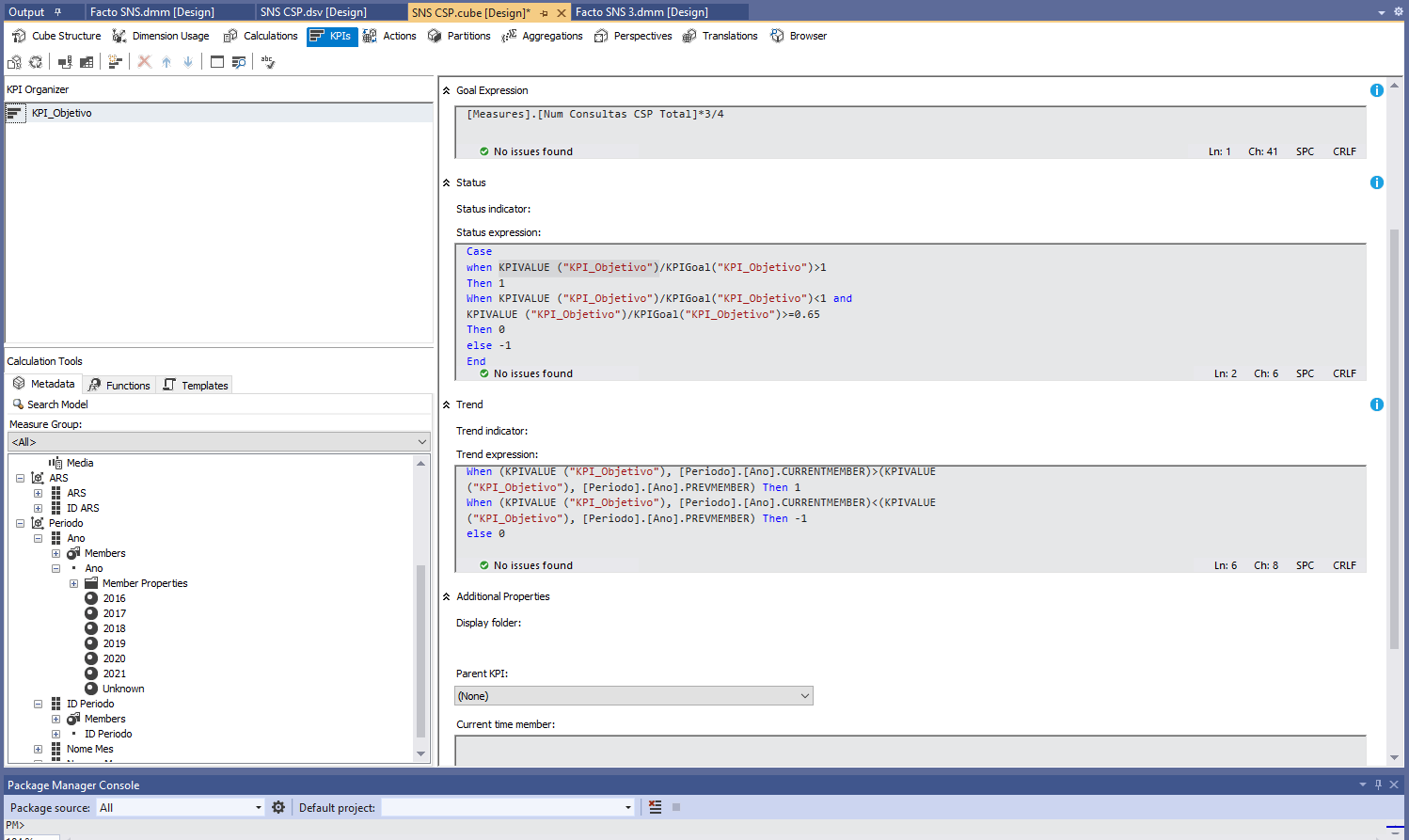


Figura 14 – Calculos – Média Consultas Gripe

e KPIs



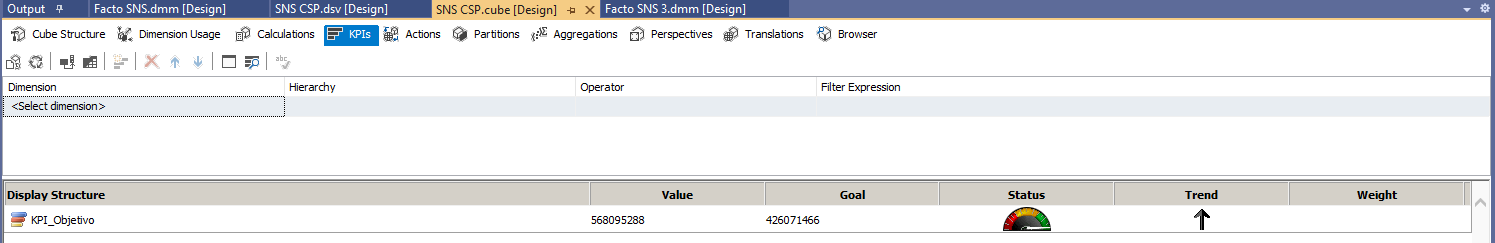


Figura 15 – KPI para objetivo de redução de consultas em ¾

que encapsulam a lógica de negócios e interações, como navegação e comportamentos de drill-through.

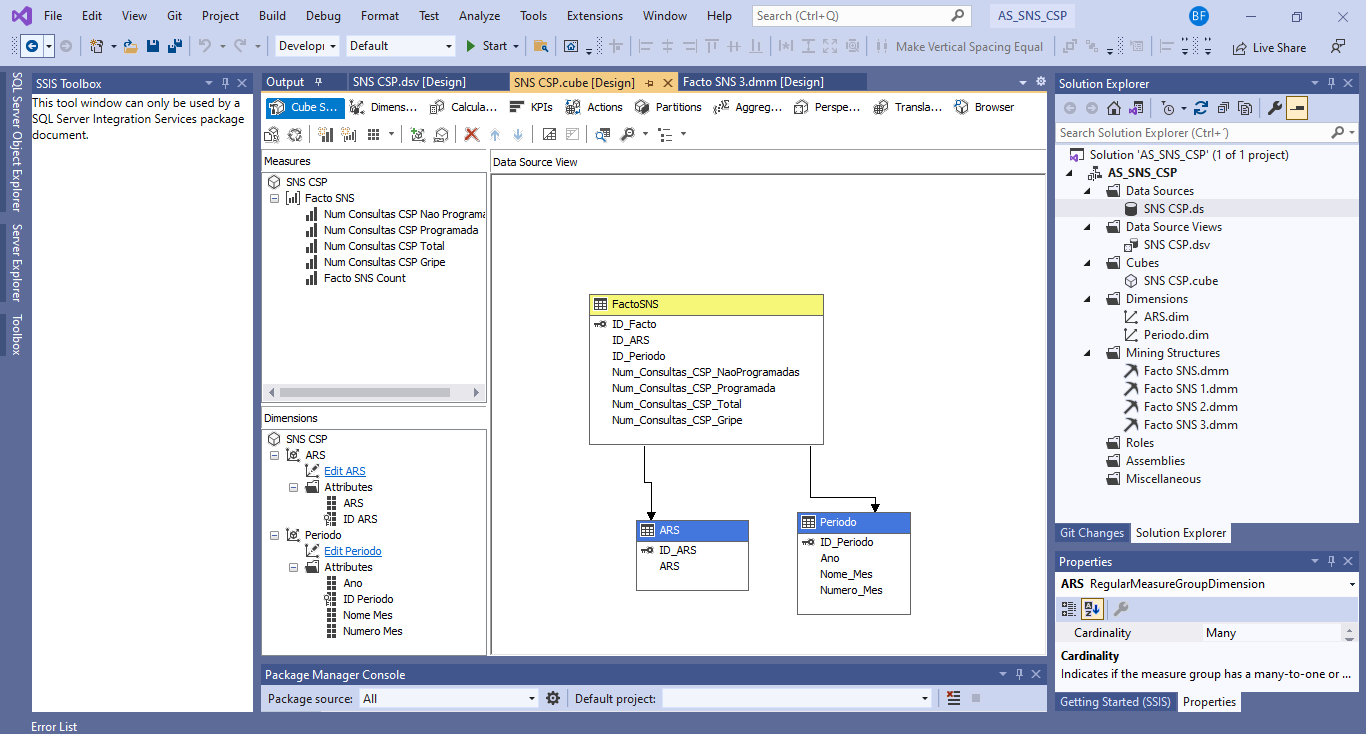
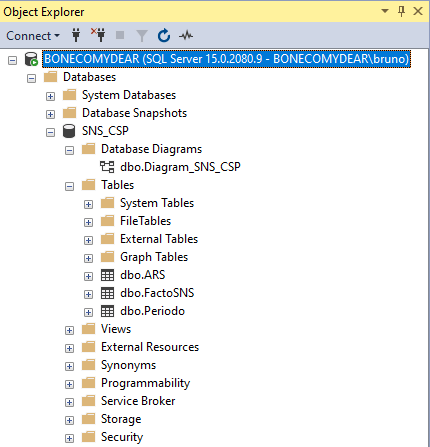


Figura 16 – Cubo – Data Source View

Para usar um modelo, ele é implementado numa instância de servidor que executa base de dados num modo de servidor específico, disponibilizando os dados para utilizadores autorizados que se ligam por meio do Excel ou outros aplicativos.

 Figura 17 – Instância SQL

1.3 Data Mining

Definição: em palavras simples, a mineração de dados é definida como um processo usado para extrair dados utilizáveis ​​de um conjunto maior de quaisquer dados brutos. Implica analisar padrões de dados em grandes lotes de dados usando um ou mais softwares. A mineração de dados tem aplicações em vários campos, como ciência e pesquisa. Aplicando a mineração de dados, as empresas podem aprender mais sobre as tendencias dos seus clientes e desenvolver estratégias mais eficazes relacionadas a várias funções de negócios e, por sua vez, aproveitar os recursos de uma maneira mais otimizada e perspicaz. Isso ajuda as empresas a estar mais perto de seus objetivos e a tomar melhores decisões. A mineração de dados envolve coleta e armazenamento eficazes de dados, bem como processamento de computador. Para segmentar os dados e avaliar a probabilidade de eventos futuros, a mineração de dados usa algoritmos matemáticos sofisticados.

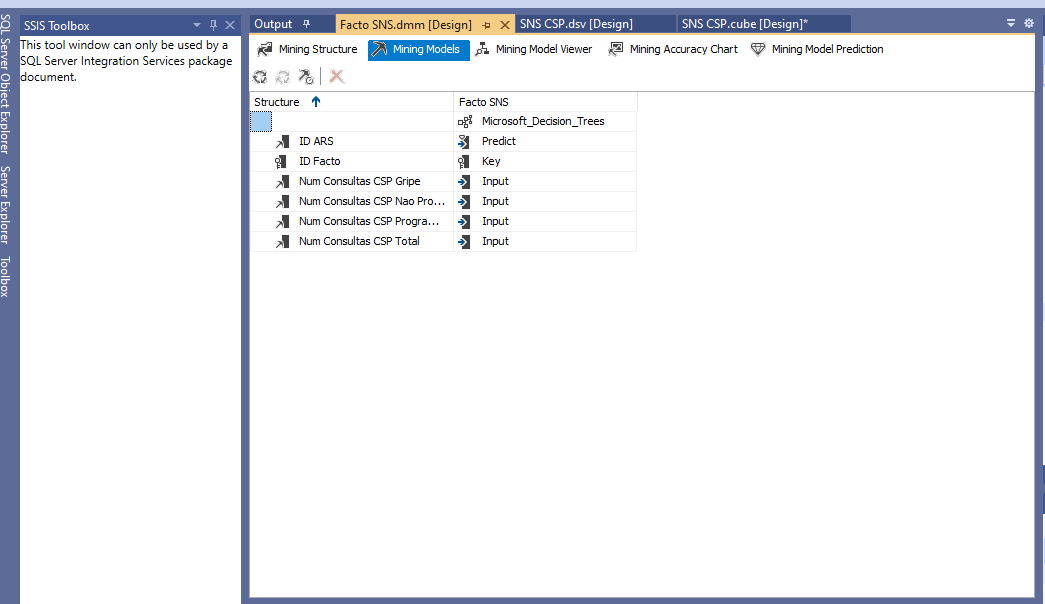


Figura 18 – Modelo mineração – Predict (Previsão) – Inputs (Filtro Entradas) a aplicar

A mineração de dados envolve explorar e analisar grandes blocos de informações para obter padrões e tendências significativos. Ele pode ser usado de várias maneiras, como marketing de base de dados, gestão de risco de crédito, detecção de fraude, filtragem de e-mail de spam ou mesmo para discernir o sentimento ou opinião dos utilizadores.

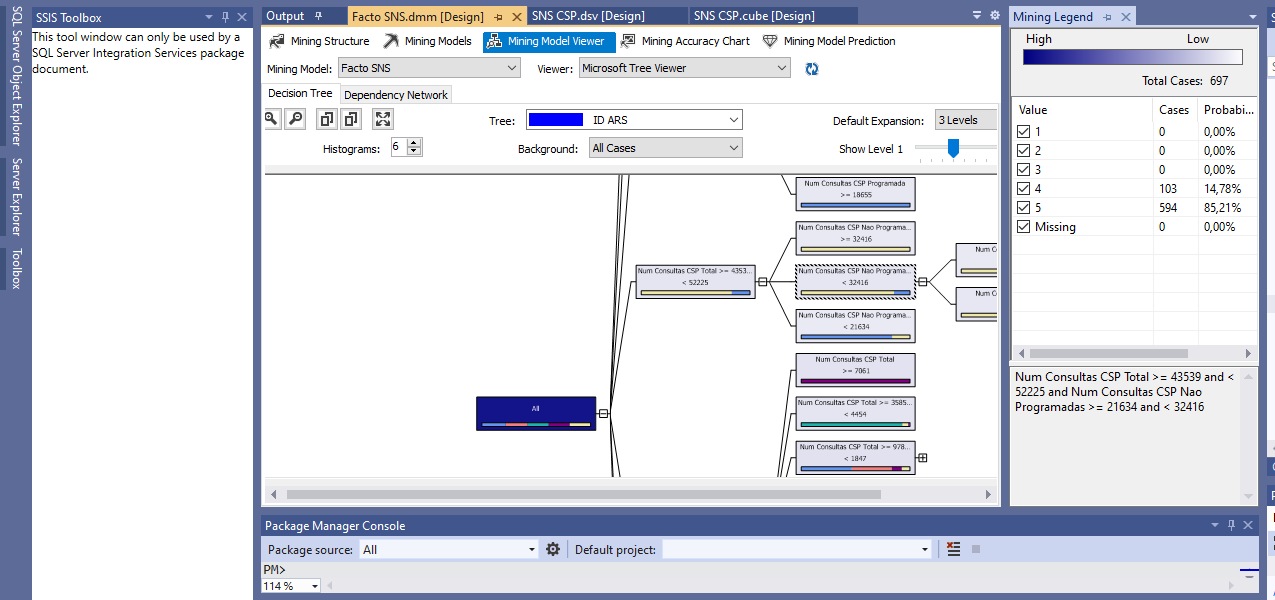


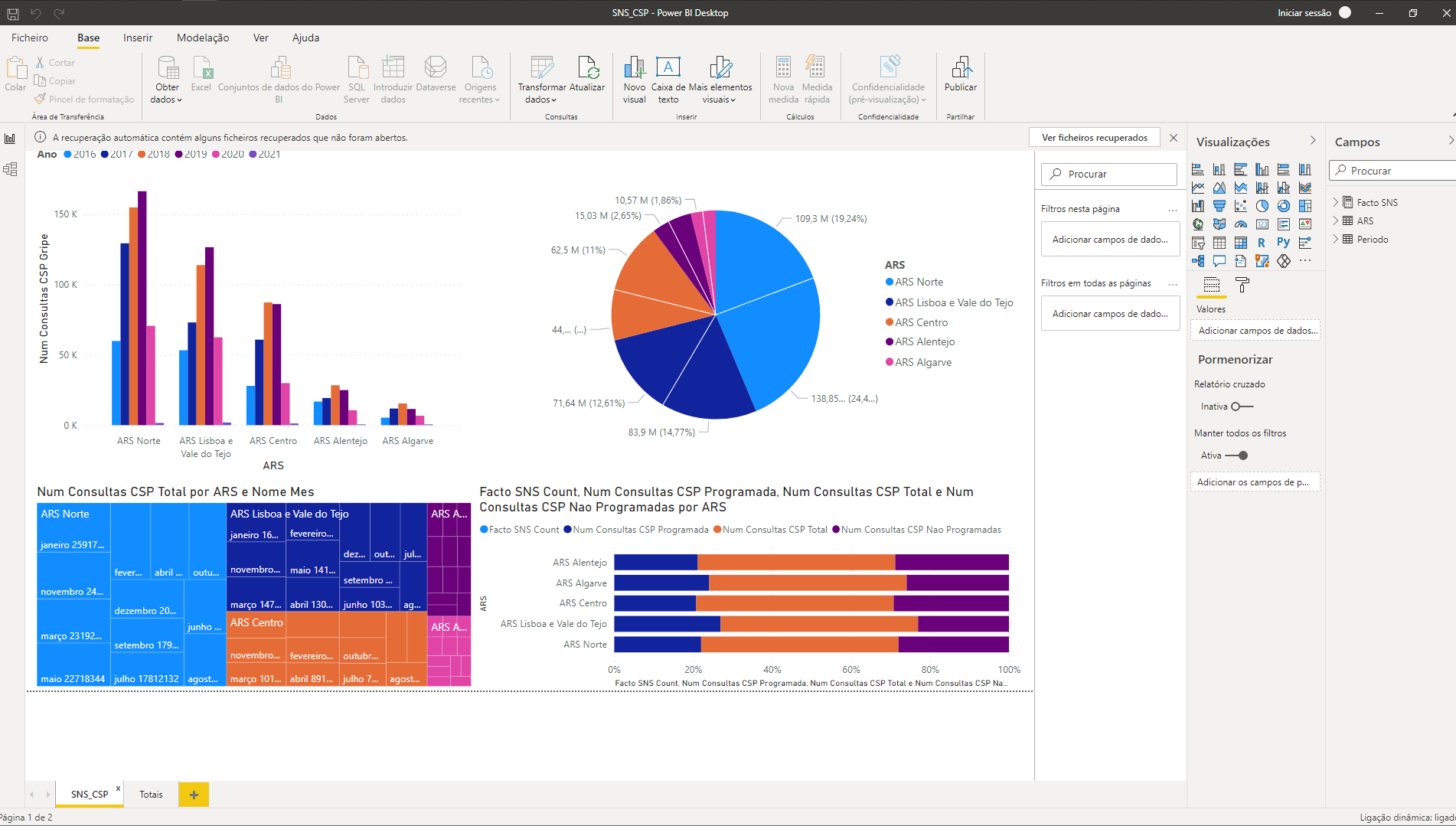
Figura 19 – Microsoft Tree Viwer

O processo de mineração de dados divide-se em cinco etapas. Primeiro, as organizações obtem-se dados e são carregados nos seus data warehouses. Em seguida, eles armazenam e é feita a gestão dos mesmos, seja em servidores internos ou na nuvem.

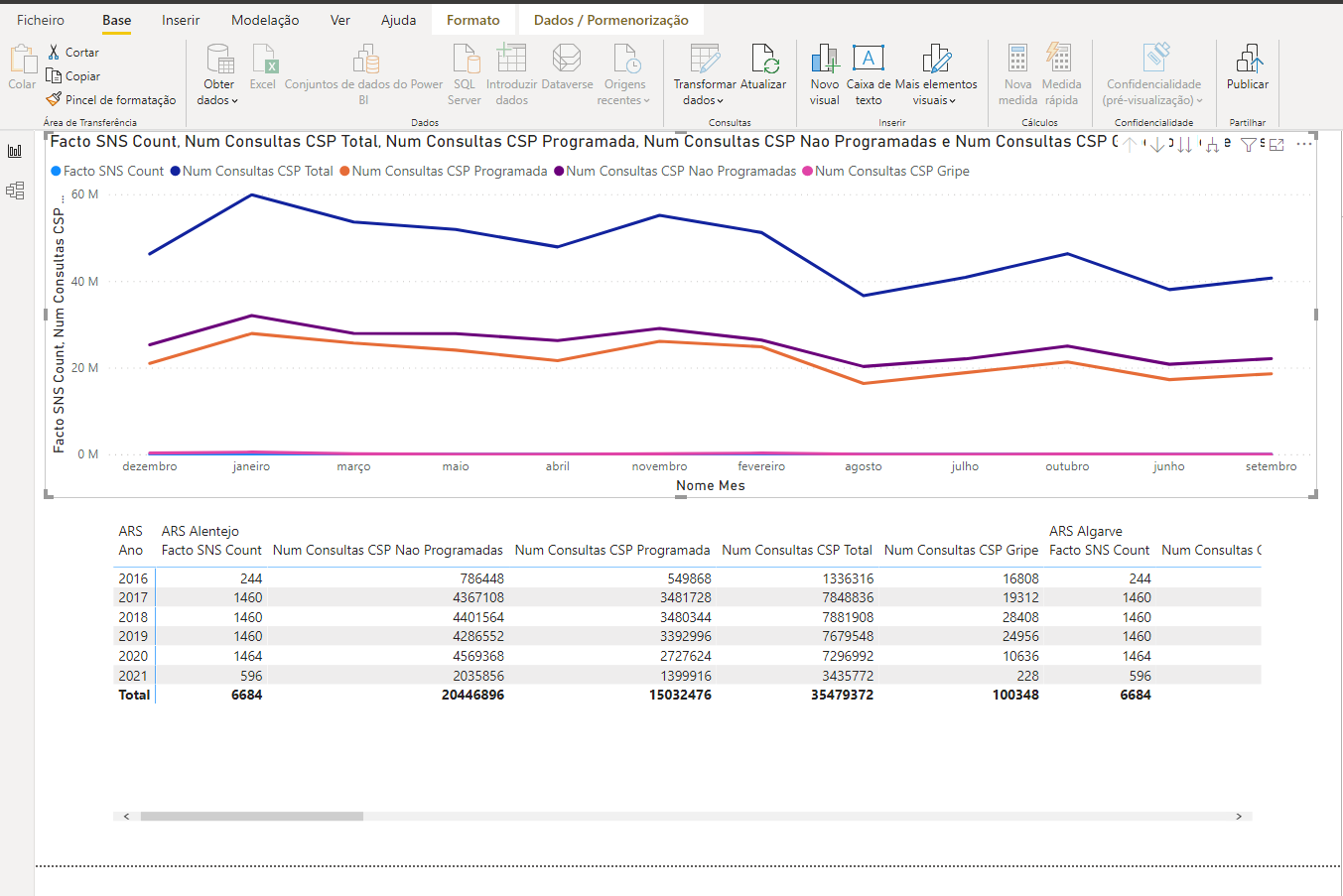
Analistas de negócios, equipas de gestão e profissionais de tecnologia da informação consultam os dados e determinam comoestes são organizados. De seguida, o software aplicativo classifica os dados com base nos resultados do utilizador e, finalmente, o utilizador final apresenta os dados num formato fácil de compartilhar, como um gráfico ou tabela.

## 2. Power BI

O Power BI é uma coleção de serviços de software, aplicativos e conectores que funcionam juntos para transformar suas fontes de dados não relacionadas em percepções coerentes, visualmente imersivas e interativas. Seus dados podem ser um ficheiro do Excel ou uma coleção de data warehouses híbridos baseados em nuvem e no local. O Power BI permite ligação fácil às fontes de dados.



“Power BI”, diz a Microsoft, “é uma solução de análise de negócios que permite que você visualize seus dados e compartilhe percepções em sua organização ou incorpore-as em seu aplicativo ou site”. A solução inclui vários produtos e serviços, e o valor e a versatilidade vêm do aproveitamento dos elementos individuais e do aproveitamento de como eles funcionam juntos.



## 3. Referência Bibliográficas Electronicas