

GOIÁS MARKET

A Goiás Market é uma empresa com uma gama de produtos diversificada que atua no ramo de varejo.

A empresa se deu conta que suas operações vêm aumentando significativamente ao longo dos anos e percebeu a importância da transformação digital.

Buscando um melhor posicionamento no mercado, a Goiás Market iniciou um projeto de cultura *data-driven*, para processar seus dados e entender melhor o seu passado e ser mais assertiva no futuro.

Assim, o projeto se inicia com a implantação do setor de *Business Intelligence*, visando os seus dados históricos para diagnosticar suas operações e, futuramente, a aplicação de *Data Science* para modelar seus negócios para o futuro, produzindo prognósticos e previsões de suas operações.

O objetivo geral da Goiás Market é utilizar seus dados gerados de forma mais adequada e segura e tomar decisões baseada em dados.

O início do projeto será a modelagem de um *Data Warehouse*, que comportará os dados históricos da empresa para análise dos gestores da empresa e também servirá como fonte de dados para criação de *Pipelines* para ciência de dados, aplicando técnicas de *Data Mining*, *Machine Learning* e processamento de *Big Data* para compreender melhor o comportamento do seu público.

A Goiás Market visa a transformação digital como vantagem competitiva.

OBJETIVOS DE NEGÓCIOS

- ✓ Analisar dados históricos e extrair informação de valor relevante para a gestão de negócios.
- ✓ Definir clientes e prever comportamento de compras, visando maior volume de vendas.
- ✓ Realizar previsões de vendas e lucro.

OBJETIVOS TÉCNICOS

- **Modelagem de um *Data Warehouse*:**
 - Modelagem da *Stage Area*;
 - Modelagem do *Data Warehouse*;
 - Criação do Cubo Analítico;
 - Automação de processos e BI Self-Service;
- **Modelagem de Pipeline de dados para *Data Science*:**
 - Pipelines com linguagem R e Python, acessando bancos em SQL;
 - Análise Exploratória de Dados;
 - Aplicação de *Data Mining* para regras de associação e sistema de recomendação;
 - Aplicação de *Machine Learning* para previsão de vendas e lucro;
 - Aplicação de *Machine Learning* para classificação de vendas por regiões



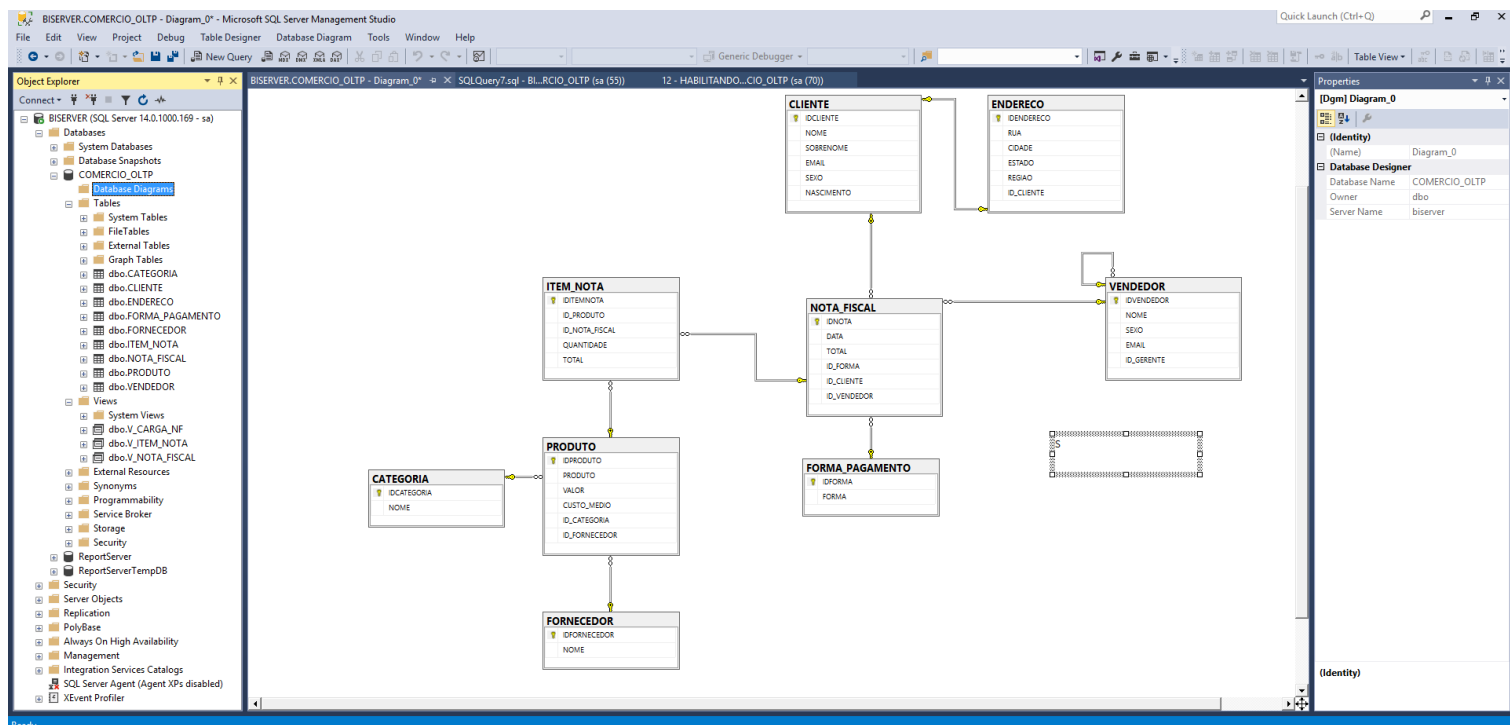
- Modelagem do banco de dados transacional da empresa (ambiente OLTP);
- Modelagem lógica e física da *Staging Area* e ETL de carga;
- Modelagem lógica e física do *Data Warehouse* e ETL de carga;
- Análise de dados através do Microsoft Power BI;
- Automação do processo e *BI Self-Service*.

AMBIENTE OLTP

O Ambiente OLTP, que é o banco de dados transacional da empresa Goiás Market, é constituído por algumas tabelas que auxiliam o negócio através de registro de suas atividades de vendas de produtos e cadastro dos clientes.

O ambiente OLTP foi construído sob algumas regras, ou *constraints*, para que haja integridade relacional durante sua carga.

Com a baixa redundância no ambiente OLTP, as tabelas apresentam dispersão de informações ali escritas, como segue abaixo:



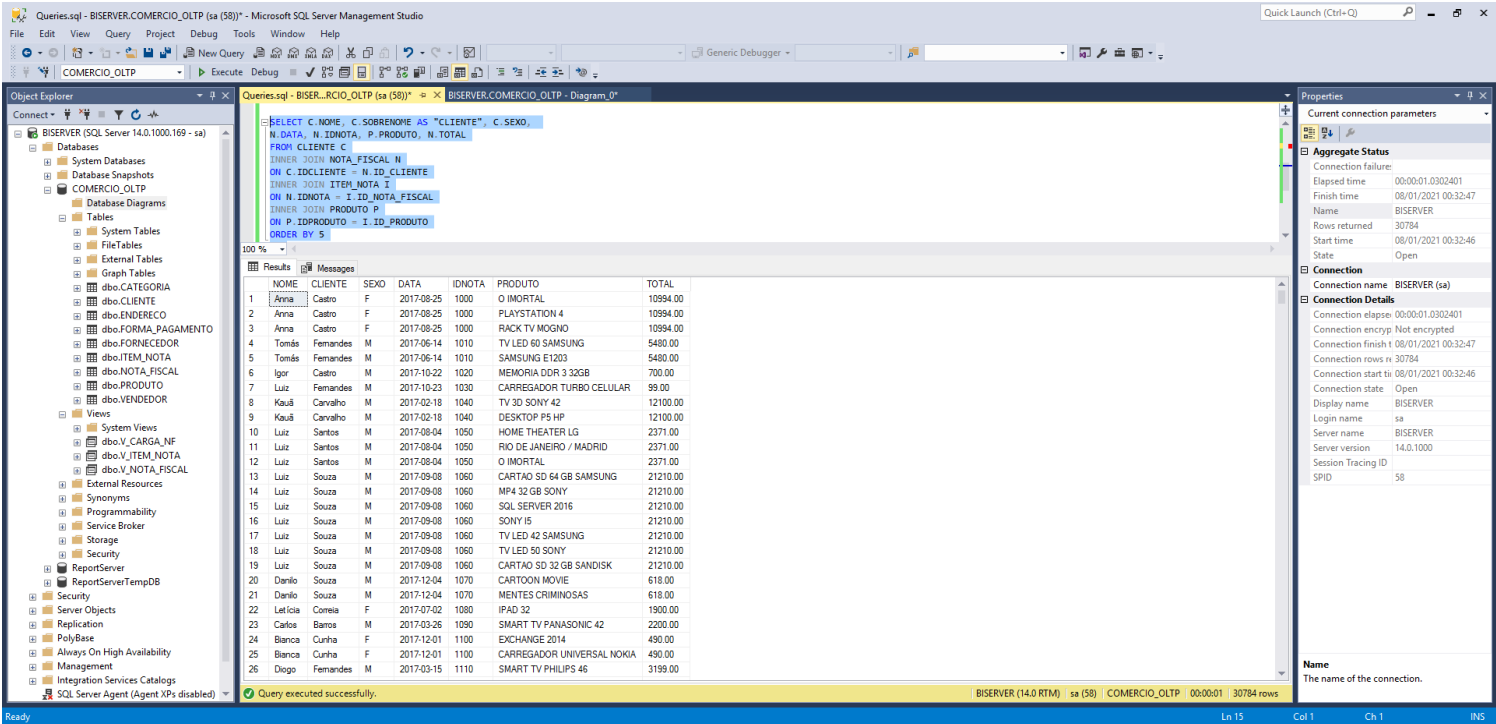
O desenvolvimento do projeto visa a construção de um *Data Warehouse* para retirar a carga do ambiente OLTP e disponibilizar um fácil acesso aos dados, com transformações feitas segundo requisitos do negócio e com maior granularidade para cruzamento de dados.

Para uma rápida consulta no banco de dados OLTP, algumas *queries* são executadas no banco, a fim de produzir um relatório das vendas totais realizadas.

A *Query* pode ser definida a seguir:

```
SELECT C.NOME, C.SOBRENOME AS CLIENTE, C.SEXO AS
SEXO_CLIENTE, P.PRODUTO AS PRODUTO, CA.NOME AS CATEGORIA,
N.IDNOTA AS NOTA_FISCAL, N.DATA AS DATA_NOTA, QUANTIDADE,
I.TOTAL AS TOTAL_ITEM, N.TOTAL AS TOTAL_NOTA, FP.FORMA AS
FORMA_PAGAMENTO
FROM CLIENTE C
INNER JOIN NOTA_FISCAL N
ON C.IDCLIENTE = N.ID_CLIENTE
INNER JOIN FORMA_PAGAMENTO FP
ON IDFORMA = ID_FORMA
INNER JOIN ITEM_NOTA I
ON N.IDNOTA = I.ID_NOTA_FISCAL
INNER JOIN PRODUTO P
ON P.IDPRODUTO = I.ID_PRODUTO
INNER JOIN CATEGORIA CA
ON IDCATEGORIA = ID_CATEGORIA
ORDER BY 5
GO
```

QUERY OLTP



Outras queries que se pode extrair algumas informações rapidamente do banco de dados:

FATURAMENTO POR VENDEDOR

Queries.sql - BISER_RCIO_OLTP (sa (61))

```
SELECT V.NOME, SUM(NF.TOTAL) AS TOTAL_VENDAS_VENDEDOR FROM VENDEDOR V
INNER JOIN NOTA_FISCAL NF
ON V.IDVENDEDOR = NF.ID_VENDEDOR
GROUP BY V.NOME
ORDER BY 2 DESC
GO
```

	NOME	TOTAL_VENDAS_VENDEDOR
1	PILAR SANCHES	3051969.00
2	HERMES MACEDO	3006790.00
3	TERESA CRISTINA	2977228.00
4	RENE CHIARI	2948115.00
5	DEBORA ALMEIDA	2939122.00
6	ANTONIO JESUS	2910430.00
7	SILVIO BAHIANA	2891517.00
8	RODRIGO MARQUES	2859663.00
9	FABIO AGUSTO	2850705.00
10	SERGIO AGUDO	2842360.00
11	DAVID BORGES	2835116.00
12	CLARA MAFRA	2824465.00
13	MARCELO WITEMBERG	2804635.00
14	WALLACE MOREIRA	2786713.00
15	JOAO NUNES	2764184.00
16	CARLOS SENRA	2763675.00
17	LILIAN SALINAS	2727369.00
18	RODRIGO BIO	2694228.00
19	CELA TAVARES	2688515.00
20	EDUARDO SANTOS	2633845.00
21	RENATA MONTEIRO	2607341.00
22	PEDRO NUNES	2558538.00

Query executed successfully. BISERVER (14.0 RTM) sa (61) COMERCIO_OLTP 00:00:00 24 rows

FATIAMENTO POR PRODUTO

Queries.sql - BISER_RCIO_OLTP (sa (61))

```
SELECT F.NOME AS FORNECEDOR, P.PRODUTO, C.NOME AS CATEGORIA, P.CUSTO_MEDIO, P.VALOR FROM PRODUTO P
INNER JOIN CATEGORIA C
ON P.ID_CATEGORIA = C.IDCATEGORIA
INNER JOIN FORNECEDOR F
ON P.ID_FORNECEDOR = F.IDFORNECEDOR
ORDER BY 1
GO
```

	FORNECEDOR	PRODUTO	CATEGORIA	CUSTO_MEDIO	VALOR
1	ACER	DESKTOP I3	INFORMATICA	996.00	1800.00
2	ACER	DESKTOP I5	INFORMATICA	1020.00	2100.00
3	ACER	DESKTOP I7	INFORMATICA	1800.00	2300.00
4	ACER	NOTEBOOK ACER I5	INFORMATICA	800.00	1500.00
5	ACER	NOTEBOOK ACER I7	INFORMATICA	1700.00	2100.00
6	ACER	ADAPTADOR WIRELESS	INFORMATICA	60.00	89.00
7	ACER	HUB SWITCH 6 PORTAS	INFORMATICA	60.00	79.00
8	AMERICA LINES EXPRESS	SAO PAULO / MIAMI	PASSAGENS	2156.00	2300.00
9	AMERICA LINES EXPRESS	RIO DE JANEIRO / MIAMI	PASSAGENS	1763.00	1900.00
10	AMERICA LINES EXPRESS	RIO DE JANEIRO / NEW YORK	PASSAGENS	2354.00	2600.00
11	APPLE	IPOD NANO 16	INFORMATICA	110.00	300.00
12	APPLE	IPOD NANO 32	INFORMATICA	320.00	450.00
13	APPLE	IPAD MINI 16	INFORMATICA	790.00	1200.00
14	APPLE	IPAD 32	INFORMATICA	1300.00	1900.00
15	APPLE	IPAD 64	INFORMATICA	1700.00	2500.00
16	APPLE	IPHONE 6	CELULARES	1100.00	1900.00
17	APPLE	IPHONE 7	CELULARES	1900.00	2500.00
18	ARNO	JUCER ARNO	ELETRODOMESTICOS	367.00	450.00
19	ARNO	CENTRIFUGA ARNO	ELETRODOMESTICOS	2030.00	2800.00
20	BILLZAR GAMES	CAVALIEIRO DAS TREVAS	GAMES	50.00	110.00
21	BILLZAR GAMES	CARTOON MOVIE	GAMES	34.00	99.00
22	BILLZAR GAMES	CAVALEROS DE PRATA	GAMES	89.00	145.00

Query executed successfully. BISERVER (14.0 RTM) sa (61) COMERCIO_OLTP 00:00:00 232 rows

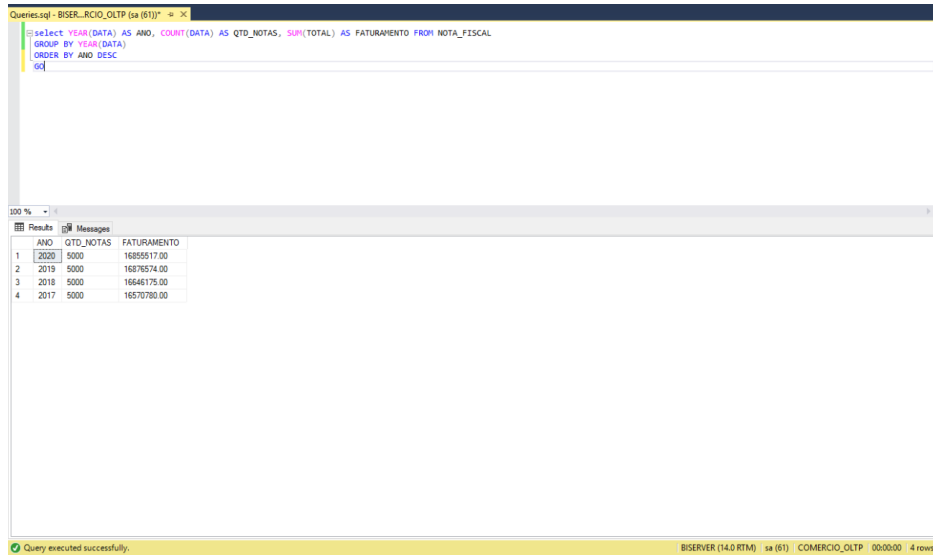
FATURAMENTO PRO REGIÃO

Queries.sql - BISER_RCIO_OLTP (sa (61))

```
SELECT E.REGIAO, COUNT(DISTINCT E.ID_CLIENTE) AS QTD_CLIENTES, SUM(NF.TOTAL) AS FATURAMENTO_REGIAO FROM ENDERECO E
INNER JOIN NOTA_FISCAL NF
ON E.ID_CLIENTE = NF.ID_CLIENTE
GROUP BY E.REGIAO
ORDER BY 3 DESC
GO
```

	REGIAO	QTD_CLIENTES	FATURAMENTO_REGIAO
1	SUDESTE	524	34305715.00
2	NORDESTE	175	11939891.00
3	SUL	136	8791921.00
4	CENTRO OESTE	111	8191901.00
5	NORTE	54	3719618.00

Query executed successfully. BISERVER (14.0 RTM) sa (61) COMERCIO_OLTP 00:00:01 5 rows



```
Query: select YEAR(DATA) AS ANO, COUNT(DATA) AS QTD_NOTAS, SUM(TOTAL) AS FATURAMENTO FROM NOTA_FISCAL
GROUP BY YEAR(DATA)
ORDER BY ANO DESC
```

	ANO	QTD_NOTAS	FATURAMENTO
1	2020	5000	16855517.00
2	2019	5000	16876574.00
3	2018	5000	16646175.00
4	2017	5000	16570780.00

Query executed successfully. BSERVER (14.0 RTM) sa (81) COMERCIO_OLTP 00:00:00 4 rows

CONCLUSÃO AMBIENTE OLTP

O ambiente transacional é a fonte de dados principal da empresa Goiás Market. É no OLTP onde todas informações são lançadas mediante um software de interface de vendas e registros.

Esse ambiente deve continuar livre e rápido para processos de escritas e leituras (como consulta de estoque por exemplo) nas rotinas diárias da empresa. Isso torna o *Data Warehouse* uma solução bastante útil para gerenciar os negócios de forma rápida e simples, sem congestionar o sistema da companhia.

Nesta etapa, obtemos uma visão geral do banco de dados e das consultas à serem feitas. Também, começamos a compreender os dados e definir o escopo do projeto, o qual será apresentado na fase de modelagem do ambiente de *Stage*.