

<https://github.com/Kakues/A3>

DOCUMENTAÇÃO COMPLETA - DASHBOARD CRUZEIRO EC

Versão 2.0 - Sistema de Business Intelligence com IA

SUMÁRIO

1. [Visão Geral do Projeto](#)
 2. [Arquivos do Projeto](#)
 3. [Como Usar o Sistema](#)
 4. [Estrutura de Dados](#)
 5. [Tabelas Exportadas](#)
 6. [Dashboard Interativo](#)
 7. [Uso de IA no Projeto](#)
 8. [Guia Power BI](#)
 9. [Métricas e KPIs](#)
 10. [Troubleshooting](#)
-

1. VISÃO GERAL DO PROJETO

Objetivo

Criar um sistema completo de Business Intelligence para análise de dados do Cruzeiro Esporte Clube, incluindo processamento automatizado, análises estatísticas e visualizações interativas.

Números do Projeto

- **130 jogos analisados** (2019-2025)
- **15 arquivos CSV** processados
- **12 tabelas otimizadas** geradas
- **26 colunas** de dados financeiros detalhados
- **14 gráficos interativos** criados
- **20+ KPIs** calculados

- 5 páginas de dashboard
- ~1.800 linhas de código

2. ARQUIVOS DO PROJETO

Estrutura de Pastas

```
F:\Faculdade\A3\  
├─ script.py (Script Python principal - 600 linhas)  
├─ dashboard_preview.html (Dashboard interativo - 1.200 linhas)  
├─ data/  
│   └─ data.csv/  
│       ├── jogo_fatos.csv  
│       ├── receita_fatos.csv  
│       ├── receitas_detalhadas_2019_2025.csv ★ NOVO  
│       ├── lotacao_por_jogo.csv  
│       ├── perfil_demografico_torcida.csv  
│       ├── setor_fatos.csv  
│       ├── receitas_mineirao_2014_2022.csv  
│       ├── socio_torcedor_fatos.csv  
│       ├── ticket_medio_estimativa.csv  
│       ├── ticket_medio_torcedor.csv  
│       ├── vendas_canal_de_vendas__2021_2025.csv  
│       ├── vendas_competicao_cruzeiro_mineirao_2021_2025.csv  
│       ├── precos_produtos_mineirao.csv  
│       ├── publico_cruzeiro.csv  
│       └─ setor_por_jogo.csv  
└─ exports_powerbi/ (Gerado automaticamente)  
    ├── FATO_Jogos.csv  
    ├── DIM_Produtos.csv  
    ├── DIM_Demografica.csv  
    ├── FATO_Temporal.csv  
    ├── AGG_Metricas_Anuais.csv  
    ├── KPI_Dashboard.csv  
    ├── CORR_Matriz.csv  
    ├── FATO_Receitas_Detalhadas.csv ★ NOVO  
    ├── ANALISE_Precificacao.csv ★ NOVO  
    ├── ANALISE_Mix_Receitas.csv ★ NOVO  
    ├── ANALISE_Ocupacao.csv ★ NOVO  
    ├── SERIE_Temporal_Completa.csv ★ NOVO  
    └─ README_POWERBI_V2.txt
```



3. COMO USAR O SISTEMA

Passo 1: Preparar os Dados

1. Coloque todos os arquivos CSV na pasta `data/data.csv/`
2. Certifique-se que o arquivo `receitas_detalhadas_2019_2025.csv` está presente

Passo 2: Executar o Script Python

bash

```
python script.py
```

O que o script faz:

- ☒ Detecta automaticamente todos os CSVs
- ☒ Limpa e padroniza os dados
- ☒ Cria relacionamentos entre tabelas
- ☒ Calcula KPIs e correlações
- ☒ Gera 12 arquivos otimizados para Power BI
- ☒ Cria documentação automática

Passo 3: Visualizar Preview

1. Abra `dashboard_preview.html` no navegador
2. Explore as 5 páginas do dashboard
3. Teste os filtros interativos

Passo 4: Importar no Power BI

1. Abra Power BI Desktop
2. Vá em "Obter Dados" → "Texto/CSV"
3. Importe todos os arquivos da pasta `exports_powerbi/`
4. Crie os relacionamentos (veja seção 8)
5. Adicione as medidas DAX (veja seção 8)



4. ESTRUTURA DE DADOS

Arquivos de Entrada (15 CSVs)

receitas_detalhadas_2019_2025.csv ★ PRINCIPAL

130 jogos | 26 colunas

Colunas principais:

- ano , competicao , times_que_jogaram
- publico_presente , publico_pagante , publico_mandante , publico_visitante
- taxa_ocupacao_percent
- inteiras_vendidas , meias_vendidas
- preco_medio_inteira , preco_medio_meia
- receita_ingresso , receita_produtos_internos , receita_camarotes , receita_estacionamento
- total_arrecadado
- ticket_medio_real_ingresso , ticket_medio_ideal_ingressos
- receita_bruta_ideal_ingressos
- fator_desconto_socios_percent
- capacidade_estadio

Outros Arquivos Importantes:

- jogo_fatos.csv: 8 jogos recentes (2024-2025)
- perfil_demografico_torcida.csv: Perfil da torcida por jogo
- lotacao_por_jogo.csv: Vendas de produtos
- receitas_mineirao_2014_2022.csv: Histórico 2014-2022



5. TABELAS EXPORTADAS PARA POWER BI

Tabelas Originais

1. FATO_Jogos.csv

Tabela fato principal consolidada

- **Chave:** jogo_id
- **Registros:** 8 jogos (2024-2025)
- **Colunas:** times_jogados, data, público, receitas, setores, KPIs calculados

2. DIM_Produtos.csv

Dimensão de produtos vendidos

- **Chave:** jogo_id + Produto_Tipico
- **Produtos:** Cerveja, Salgado/Lanche, Refrigerante, Hot Dog/Pipoca
- **Métricas:** Receita por produto, participação percentual

3. DIM_Demografica.csv

Perfil demográfico da torcida

- **Chave:** Jogo_ID
- **Dimensões:** Gênero, Faixa Etária, Região
- **Formato:** Colunas pivot com percentuais

4. FATO_Temporal.csv

Série temporal de receitas

- **Período:** 2014-2025
- **Métricas:** Receitas por fonte, público

5. AGG_Metricas_Anuais.csv

Agregações anuais

- **Métricas:** Soma, Média, Desvio Padrão por ano

6. KPI_Dashboard.csv

KPIs principais formatados

- **Formato:** Pronto para cards no Power BI

7. CORR_Matriz.csv

Matriz de correlação estatística

- **Variáveis:** Público, Receitas, Setores

Novas Tabelas (Versão 2.0) ★

8. FATO_Receitas_Detalhadas.csv

Tabela fato principal com dados 2019-2025

- **Registros:** 130 jogos (excluindo jogos sem público)
- **Período:** 2019-2025
- **Colunas:** 30+ incluindo todas as métricas financeiras
- **Novos campos calculados:**
 - `taxa_ocupacao_decimal`
 - `gap_otimizacao` (receita ideal - receita real)
 - `perc_receita_produtos`
 - `perc_receita_camarotes`
 - `perc_receita_estacionamento`
 - `tipo_adversario` (Grande/Médio/Pequeno)
 - `eh_classico` (True/False)
 - `era` (Pré-COVID/Pandemia/Pós-COVID)

9. ANALISE_Precificacao.csv

Análise de precificação por competição/ano

- **Granularidade:** Ano + Competição
- **Métricas:**
 - Preços médios (inteira, meia)
 - Ticket médio real vs ideal
 - Desconto médio sócios (%)
 - Gap de otimização total
 - Eficiência de precificação (%)

10. ANALISE_Mix_Receitas.csv

Composição de receitas por fonte

- **Granularidade:** Ano + Competição
- **Fontes:** Ingressos, Produtos, Camarotes, Estacionamento
- **Métricas:**
 - Receitas absolutas por fonte
 - Percentuais de participação
 - Receita per capita por categoria

11. ANALISE_Ocupacao.csv

Taxa de ocupação do estádio

- **Granularidade:** Ano + Competição + Tipo de Adversário
- **Métricas:**
 - Taxa de ocupação (média, mín, máx)
 - Público total e médio
 - Pagantes total e médio
 - % não-pagantes

12. SERIE_Temporal_Completa.csv

Série temporal agregada 2019-2025

- **Granularidade:** Ano + Competição
 - **Métricas:**
 - Público (total, médio)
 - Receita (total, média)
 - Taxa de ocupação média
 - Ticket médio
 - Gap de otimização
 - Quantidade de jogos
 - Era (Pré-COVID/Pandemia/Pós-COVID)
-

6. DASHBOARD INTERATIVO (HTML)

Páginas do Dashboard

Página 1: Visão Geral

- 4 KPIs principais (Receita, Público, Ticket Médio, Taxa de Ocupação)
- Gráfico de evolução temporal
- Distribuição por setores
- Receita por jogo

Página 2: Produtos

- 4 KPIs de produtos
- Receita por tipo de produto
- Participação percentual
- Tabela detalhada

Página 3: Demográfico

- Distribuição por gênero
- Distribuição por região
- Distribuição por faixa etária
- Insights sobre perfil da torcida


Página 4: Correlações

- Gráfico de dispersão Público x Receita
- Tabela de correlações principais
- Análise de tendências temporais







Página 5: Análise Financeira NOVA

- 4 KPIs financeiros avançados
- Funil de otimização de receita
- Mix de receitas (composição)
- Evolução taxa de ocupação (2019-2025)
- Scatter: Ocupação x Receita
- Top 10 jogos por receita

Filtros Disponíveis

1. **Ano:** 2019-2025 (individual) + Todos
2. **Tipo de Adversário:** Grande / Médio-Pequeno / Todos
3. **Competição:** Brasileiro, Copa do Brasil, Libertadores, Mineiro, Sul-Americana, Todas
4. **Clássico:** Sim / Não / Todos
5. **Era:** Pré-COVID / Pandemia / Pós-COVID / Todas  NOVO

Recursos Interativos

-  Filtros em tempo real
 -  Gráficos responsivos
 -  Tooltips informativos
 -  Cores do Cruzeiro (#003366 , #FFD700)
 -  Botão "Limpar Filtros"
 -  Indicação visual de filtros ativos
-



7. USO DE IA NO PROJETO

Como a IA (Claude) Foi Utilizada

Fase 1: Planejamento (30 min)

- Análise dos 15 arquivos CSV
- Proposta de arquitetura dimensional
- Identificação de métricas relevantes
- Sugestão de visualizações

Fase 2: Desenvolvimento Python (2 horas)

- Geração de 600 linhas de código
- Resolução de 8+ erros em tempo real
- Implementação de tratamento de exceções
- Criação de documentação automática

Fase 3: Guia Power BI (1 hora)

- Tutorial passo a passo
- 15+ fórmulas DAX prontas
- Layout de dashboard
- Documentação de relacionamentos

Fase 4: Dashboard HTML (1.5 horas)

- Criação de 1.200 linhas HTML/CSS/JavaScript
- Implementação de 14 gráficos interativos
- Sistema de filtros em tempo real
- Identidade visual do Cruzeiro

Fase 5: Iterações e Melhorias (1.5 horas)

- Ajustes baseados em feedback
- Correção de erros de tipo de dados
- Otimização de legibilidade
- Expansão para dados 2019-2025

Fase 6: Integração Novo Arquivo (45 min)

- Análise do arquivo receitas_detalhadas

- Atualização do script Python
- Criação de 5 novas tabelas analíticas
- Expansão do dashboard

O Que a IA FEZ

- ✓ Gerou código base
- ✓ Sugeriu soluções técnicas
- ✓ Corrigiu erros rapidamente
- ✓ Aplicou melhores práticas
- ✓ Criou documentação

O Que NÓS FIZEMOS

- ✓ Definimos objetivos
- ✓ Organizamos os dados
- ✓ Testamos e validamos
- ✓ Decidimos melhorias
- ✓ Apresentamos o trabalho
- ✓ Aprendemos no processo



8. GUIA POWER BI

Importação de Dados

1. Abra Power BI Desktop
2. Clique em **"Obter Dados"** → **"Texto/CSV"**
3. Navegue até `exports_powerbi/`
4. Importe todos os 12 arquivos CSV

Criar Relacionamentos

Relacionamentos Originais:

```
FATO_Jogos[jogo_id] —1:* — DIM_Produtos[jogo_id]  
FATO_Jogos[jogo_id] —1:* — DIM_Demografica[Jogo_ID]  
FATO_Jogos[data] —*:1— FATO_Temporal[data]
```

Novos Relacionamentos (v2.0):

```
FATO_Receitas_Detalhadas[ano] —1:* — ANALISE_Precificacao[ano]
FATO_Receitas_Detalhadas[ano] —1:* — ANALISE_Mix_Receitas[ano]
FATO_Receitas_Detalhadas[ano] —1:* — ANALISE_Ocupacao[ano]
FATO_Receitas_Detalhadas[ano] —1:* — SERIE_Temporal_Completa[ano]
```

Configuração dos Relacionamentos:

- **Cardinalidade:** 1 para Muitos (1:*)
- **Direção do filtro:** Ambos
- **Ativo:** Sim

Medidas DAX - Básicas

dax

```
// Métricas Fundamentais
Receita Total = SUM(FATO_Jogos[total_arrecadado])

Público Total = SUM(FATO_Jogos[publico_total])

Ticket Médio = AVERAGE(FATO_Jogos[ticket_medio_ingresso])

Quantidade de Jogos = COUNTROWS(FATO_Jogos)

// KPIs Calculados
Taxa de Ocupação =
DIVIDE([Público Total], 61927, 0)

Receita por Público =
DIVIDE([Receita Total], [Público Total], 0)

% Receita Produtos =
DIVIDE(
    SUM(FATO_Jogos[receita_produtos_internos]),
    [Receita Total],
    0
)

// Análise Temporal
Receita Ano Anterior =
CALCULATE(
    [Receita Total],
    SAMEPERIODLASTYEAR(FATO_Jogos[data])
)
```

```

Crescimento YoY =
DIVIDE(
    [Receita Total] - [Receita Ano Anterior],
    [Receita Ano Anterior],
    0
)

```

Medidas DAX - Análise Financeira (v2.0) ★

dax

```

// Otimização de Receita
Gap Otimização Total =
SUM(FATO_Receitas_Detalhadas[gap_otimizacao])

Eficiência Precificação =
DIVIDE(
    SUM(FATO_Receitas_Detalhadas[receita_ingresso]),
    SUM(FATO_Receitas_Detalhadas[receita_bruta_ideal_ingressos]),
    0
) * 100

Desconto Médio Sócios =
AVERAGE(FATO_Receitas_Detalhadas[fator_desconto_socios_percent])

// Mix de Receitas
% Receita Camarotes =
DIVIDE(
    SUM(FATO_Receitas_Detalhadas[receita_camarotes]),
    SUM(FATO_Receitas_Detalhadas[total_arrecadado]),
    0
) * 100

% Receita Estacionamento =
DIVIDE(
    SUM(FATO_Receitas_Detalhadas[receita_estacionamento]),
    SUM(FATO_Receitas_Detalhadas[total_arrecadado]),
    0
) * 100

Receita Per Capita Produtos =
DIVIDE(
    SUM(FATO_Receitas_Detalhadas[receita_produtos_internos]),
    SUM(FATO_Receitas_Detalhadas[publico_presente]),

```

```

0
)

// Ocupação
Taxa Ocupação Média =
AVERAGE(FATO_Receitas_Detalhadas[taxa_ocupacao_percent])

% Não Pagantes =
DIVIDE(
    SUM(FATO_Receitas_Detalhadas[publico_presente]) -
    SUM(FATO_Receitas_Detalhadas[publico_pagante]),
    SUM(FATO_Receitas_Detalhadas[publico_presente]),
    0
) * 100

Público Médio Clássicos =
CALCULATE(
    AVERAGE(FATO_Receitas_Detalhadas[publico_presente]),
    FATO_Receitas_Detalhadas[eh_classico] = TRUE
)

// Análise Temporal
Crescimento Público YoY =
VAR PublicoAnoAtual = SUM(FATO_Receitas_Detalhadas[publico_presente])
VAR PublicoAnoAnterior =
    CALCULATE(
        SUM(FATO_Receitas_Detalhadas[publico_presente]),
        PREVIOUSYEAR(FATO_Receitas_Detalhadas[ano])
    )
RETURN
DIVIDE(PublicoAnoAtual - PublicoAnoAnterior, PublicoAnoAnterior, 0) * 100

Impacto COVID =
CALCULATE(
    SUM(FATO_Receitas_Detalhadas[publico_presente]),
    FATO_Receitas_Detalhadas[era] = "Pandemia"
)

Recuperação Pós-COVID =
VAR PublicoPreCOVID =
    CALCULATE(
        AVERAGE(FATO_Receitas_Detalhadas[publico_presente]),
        FATO_Receitas_Detalhadas[era] = "Pré-COVID"
    )
VAR PublicoPosCOVID =
    CALCULATE(

```

```
AVERAGE(FATO_Receitas_Detalhadas[publico_presente]),  
FATO_Receitas_Detalhadas[era] = "Pós-COVID"  
)  
RETURN  
DIVIDE(PublicoPosCOVID, PublicoPreCOVID, 0) * 100
```

Visualizações Recomendadas

Página 1: Visão Geral

- 4 Cartões (KPIs principais)
- Gráfico de Linhas (Evolução temporal)
- Gráfico de Rosca (Setores)
- Gráfico de Barras Horizontais (Receita por jogo)

Página 2: Produtos

- 4 Cartões (KPIs de produtos)
- Gráfico de Barras (Receita por produto)
- Gráfico de Pizza (Participação %)
- Matriz (Tabela detalhada)

Página 3: Demográfico

- Gráfico de Barras Empilhadas (Gênero)
- Gráfico de Rosca (Região)
- Gráfico de Barras (Faixa etária)

Página 4: Correlações

- Gráfico de Dispersão (Público x Receita)
- Matriz (Correlações)
- Gráfico de Linha (Tendências)

Página 5: Análise Financeira ★ NOVA

- 4 Cartões (KPIs financeiros)
- Gráfico de Barras Horizontais (Funil otimização)
- Gráfico de Cascata (Mix de receitas)
- Gráfico de Linha (Taxa ocupação temporal)
- Gráfico de Dispersão (Ocupação x Receita)
- Matriz (Top 10 jogos)

Segmentações de Dados (Filtros)

Adicione em todas as páginas:

- 1. **Ano** (2019-2025)
- 2. **Tipo de Adversário** (Grande/Médio-Pequeno)
- 3. **Competição** (Dropdown)
- 4. **Clássico** (Sim/Não)
- 5. **Era** (Pré-COVID/Pandemia/Pós-COVID) ★ NOVO

9. MÉTRICAS E KPIs

KPIs Principais

Métrica	Fórmula	Interpretação
Receita Total	SUM(total_arrecadado)	Receita bruta de todas as fontes
Público Total	SUM(publico_total)	Soma de torcedores presentes
Ticket Médio	Receita Ingresso / Público	Preço médio pago por torcedor
Taxa de Ocupação	Público / 61.927	% de lotação do estádio
Receita Per Capita	Receita Total / Público	Gasto total médio por torcedor

KPIs Financeiros Avançados (v2.0)

Métrica	Fórmula	Interpretação
Gap de Otimização	Receita Ideal - Receita Real	Oportunidade de aumento
Eficiência de Precificação	$(\text{Receita Real} / \text{Receita Ideal}) \times 100$	% de eficiência
% Não-Pagantes	$(\text{Presente} - \text{Pagante}) / \text{Presente} \times 100$	Cortesias e acessos
% Receita Camarotes	$\text{Receita Camarotes} / \text{Receita Total} \times 100$	Participação premium

Benchmarks

Métrica	Valor Médio	Melhor Valor	Meta
Taxa de Ocupação	64,5%	99,5%	80%
Ticket Médio	R\$ 45,82	R\$ 166,27	R\$ 60
% Receita Produtos	22%	30%	25%
Eficiência Precificação	87,3%	95%	90%

10. TROUBLESHOOTING

Problemas Comuns

Erro: "FileNotFoundError"

Causa: Arquivos CSV não encontrados

Solução:

1. Verifique se os CSVs estão em `data/data.csv/`
2. Confirme os nomes dos arquivos (sem espaços extras)
3. Execute o script na pasta correta (`F:\Faculdade\A3\`)

Erro: "TypeError: unsupported operand type(s) for /: 'str' and 'str'"

Causa: Tentativa de dividir strings

Solução: Já corrigido na v2.0 com `pd.to_numeric(errors='coerce')`

Erro: "KeyError: 'coluna_x' not in index"

Causa: Coluna esperada não existe no DataFrame

Solução: Script v2.0 inclui verificações de colunas

Dashboard HTML não carrega gráficos

Causa: JavaScript não executou

Solução:

1. Abra com navegador moderno (Chrome, Firefox, Edge)
2. Verifique o Console (F12) para erros
3. Certifique-se que Chart.js carregou (CDN)

Relacionamentos não funcionam no Power BI

Causa: Tipos de dados incompatíveis

Solução:

1. Verifique se as chaves são do mesmo tipo (texto/número)
2. Remova espaços em branco: `TRIM([coluna])`
3. Use relacionamentos com direção "Ambos"

Contatos de Suporte

Para dúvidas técnicas:

- Revise esta documentação
- Consulte o `README_POWERBI_V2.txt` em `exports_powerbi/`
- Verifique os comentários no código Python

Para melhorias:

- Documente sugestões
- Teste localmente antes de apresentar
- Mantenha backup dos arquivos originais



REFERÊNCIAS

Documentação Consultada

- Pandas Documentation: <https://pandas.pydata.org/docs/>
- Power BI Documentation: <https://docs.microsoft.com/power-bi/>
- DAX Function Reference: <https://dax.guide/>
- Chart.js Documentation: <https://www.chartjs.org/docs/>

Conceitos Aplicados

- **Star Schema** (Modelagem Dimensional)
 - **ETL** (Extract, Transform, Load)
 - **Business Intelligence**
 - **Data Visualization Principles**
 - **Statistical Correlation**
-

APRENDIZADOS

Habilidades Técnicas Desenvolvidas

- ☒ Python para análise de dados
- ☒ Pandas para manipulação de DataFrames
- ☒ Power BI e modelagem dimensional
- ☒ DAX (linguagem de medidas)
- ☒ Visualização de dados
- ☒ HTML/CSS/JavaScript
- ☒ Tratamento de erros e exceções

Habilidades Conceituais

- ☒ Pensamento analítico
- ☒ Resolução de problemas
- ☒ Gestão de projeto
- ☒ Trabalho iterativo
- ☒ Uso estratégico de IA
- ☒ Documentação técnica

RESULTADOS FINAIS

Métricas do Projeto

Indicador	Antes	Depois
Jogos analisados	8	130
Período coberto	2024-2025	2019-2025
Arquivos processados	14	15
Tabelas geradas	7	12
Gráficos no dashboard	10	14
KPIs calculados	16	20+
Tempo de análise manual	4-6 horas	15 minutos
Qualidade das análises	Boa	Excepcional

AGRADECIMENTOS

Este projeto foi desenvolvido com o suporte de:

- **Claude AI (Anthropic)** - Assistente de desenvolvimento
- ****Python &**