

ANÁLISE PNAD CONTÍNUA IBGE

Relatório Técnico - Processo de ETL e Análise

Data de geração: 04/10/2025 19:34

Vaga: Analista de Dados Sênior

1. RESUMO EXECUTIVO

Este relatório documenta o processo completo de análise dos dados da PNAD Contínua do IBGE, desde a extração via API até a criação de dashboards interativos no Power BI.

Objetivo: Análise da taxa de desocupação brasileira através dos microdados oficiais

Período: 2012-2024 (162 trimestres)

Tecnologias: Python, Pandas, SQLite, Power BI

Entregáveis: Pipeline de ETL automatizado e dashboard com insights estratégicos

2. METODOLOGIA

2.1. EXTRAÇÃO (E)

Dados obtidos através da API SIDRA do IBGE:

```
# API endpoint
url = "https://apisidra.ibge.gov.br/values/t/6381/n1/all/v/4099/p/all"

# Parâmetros:
# t/6381 - PNAD Contínua Trimestral
# v/4099 - Taxa de desocupação
# p/all - Todos os períodos disponíveis
```

2.2. TRANSFORMAÇÃO (T)

RELATÓRIO TÉCNICO - ANÁLISE PNAD CONTÍNUA IBGE

Processos aplicados nos dados:

```
# Script de transformação
- Limpeza de dados faltantes e inconsistentes
- Conversão de tipos (string para numérico)
- Criação de métricas derivadas (variação percentual)
- Normalização de nomes de colunas
- Cálculo de médias móveis (4 períodos)
- Classificação por categorias (Baixa/Moderada/Alta)
```

2.3. CARREGAMENTO (L)

Destinos dos dados processados:

```
# Pipeline de carregamento
1. Banco SQLite: ibge_analise.db (armazenamento local)
2. Conexão direta Power BI via ODBC
3. Dataset otimizado para análise em tempo real
```

3. ARQUITETURA DA SOLUÇÃO

Fluxo completo de dados implementado:

ARQUITETURA DO PROJETO:

```
API IBGE (REST)
  |
Python Scripts (ETL)
  |
SQLite Database
  |
Power BI (conexão direta ODBC)
  |
Power BI Dashboard
  |
Insights e Relatórios
```

3.1. Scripts Python Desenvolvidos

pnad_etl.py - Pipeline principal de extração e transformação
verificar_dados.py - Validação e controle de qualidade
preparar_dados_powerbi.py - Otimização para visualização
powerbi_final.py - Geração do dataset final
gerar_relatorio_pdf.py - Este relatório

3.2. Estrutura do Banco de Dados

```
-- Tabelas no SQLite
pnad_historico (dados brutos da API)
powerbi_otimizado (dados tratados para análise)
dados_ibge_api (metadados e controles)
```

4. ANÁLISE E INSIGHTS

4.1. Principais Indicadores Calculados

- Taxa de desocupação trimestral
- Variação percentual período a período
- Média histórica (2012-2024)
- Classificação por nível de desocupação
- Tendência (Alta/Baixa/Estável)
- Média móvel (suavização de tendências)

4.2. Insights Identificados

1. TENDÊNCIA DE LONGO PRAZO: Análise da evolução da taxa de desocupação ao longo de 12 anos
2. SAZONALIDADE: Identificação de padrões trimestrais recorrentes
3. IMPACTO DE EVENTOS: Análise do efeito de crises econômicas e pandêmicas
4. COMPARATIVO HISTÓRICO: Posicionamento atual em relação à média do período

5. VISUALIZAÇÃO NO POWER BI

5.1. Conexão Power BI com SQLite

CONEXÃO DIRETA VIA ODBC:

- Configuração de conexão ODBC com SQLite
- Importação direta das tabelas do banco
- Atualização em tempo real dos dados
- Maior performance e integridade dos dados

BENEFÍCIOS DA CONEXÃO DIRETA:

- Elimina intermediários (CSV)
- Dados sempre atualizados
- Melhor performance
- Maior confiabilidade

5.2. Estrutura do Dashboard

PÁGINA 1 - VISÃO GERAL

- Gráfico de linha evolutivo da taxa de desocupação
- KPIs principais (taxa atual, média histórica, variação)
- Indicadores de status e tendência

PÁGINA 2 - ANÁLISE DETALHADA

- Média móvel para suavização de tendências
- Análise por categorias de desocupação
- Tabela detalhada com todos os períodos

PÁGINA 3 - RECOMENDAÇÕES

- Insights acionáveis para gestores públicos
- Estratégias para empresas privadas

5.2. Medidas DAX Implementadas

```
// Taxa Atual (último período)
Taxa Atual =
    CALCULATE(
        MAX([taxa_desocupação]),
        LASTDATE([data_referencial])
    )

// Tendência
Tendência =
    VAR TaxaAtual = [Taxa Atual]
    VAR TaxaAnterior = [Taxa Período Anterior]
    RETURN
    IF(TaxaAtual > TaxaAnterior, "Alta",
        IF(TaxaAtual < TaxaAnterior, "Baixa", "Estável"))
```

6. RESULTADOS E IMPACTOS

6.1. Para Gestores Públicos

- Identificação de períodos críticos para intervenção
- Base data-driven para políticas de geração de emprego
- Monitoramento da efetividade de programas sociais
- Análise de impacto de medidas econômicas

6.2. Para Empresas Privadas

- Análise de mercado para decisões de expansão
- Planejamento estratégico de contratações
- Estudo de cenários econômicos para investimentos
- Benchmarking do setor

7. TECNOLOGIAS UTILIZADAS

+-----+-----+-----+			
Camada	Tecnologia	Finalidade	
+-----+-----+-----+			
Extração	Python Requests	Consumo da API IBGE	
Processamento	Pandas, NumPy	Transformação de dados	
Armazenamento	SQLite	Banco de dados local	
Visualização	Power BI	Dashboard interativo	
Documentação	FPDF (Python)	Relatório técnico	
+-----+-----+-----+			

8. CONSIDERAÇÕES FINAIS

8.1. Lições Aprendidas

- Importância da validação de dados na fonte para qualidade
- Eficiência do pipeline Python para processos ETL complexos
- Valor da documentação completa do processo analítico
- Flexibilidade do SQLite para prototipagem e análise

8.2. Próximos Passos Recomendados

- Inclusão de dados regionais (análise por UF)
- Cruzamento com outras bases do IBGE (renda, educação)
- Automação da atualização trimestral dos dados
- Desenvolvimento de modelos preditivos simples
- Criação de alertas para tendências significativas

ANEXOS

Anexo A - Estrutura de Pastas do Projeto

```
IBGE/  
|-- pnad_etl.py (script principal)  
|-- verificar_dados.py (validação)  
|-- preparar_dados_powerbi.py (otimização)  
|-- powerbi_final.py (dataset final)  
|-- gerar_relatorio_pdf.py (este relatório)  
|-- ibge_analise.db (banco SQLite)  
|-- pnad_powerbi_pronto.csv (dados Power BI)  
|-- instrucoes_powerbi.txt (guia de uso)
```

Anexo B - Metadados dos Dados

Fonte: PNAD Contínua - IBGE

Indicador principal: Taxa de desocupação (%)

Período coberto: 2012-2024

Frequência: Trimestral

Total de registros: 162 períodos

Última atualização: Dados mais recentes disponíveis

Documento gerado automaticamente via Python - Demonstrando habilidades técnicas em ETL e análise de dados