

ANALISE PNAD CONTINUA IBGE

Relatorio Tecnico - Processo de ETL e Analise

Data de geracao: 04/10/2025 00:08

Vaga: Analista de Dados Senior

1. RESUMO EXECUTIVO

Este relatorio documenta o processo completo de analise dos dados da PNAD Continua do IBGE, desde a extracao via API ate a criacao de dashboards interativos no Power BI.

Objetivo: Analise da taxa de desocupacao brasileira atraves dos microdados oficiais

Periodo: 2012-2024 (162 trimestres)

Tecnologias: Python, Pandas, SQLite, Power BI

Entregaveis: Pipeline de ETL automatizado e dashboard com insights estrategicos

2. METODOLOGIA

2.1. EXTRACAO (E)

Dados obtidos atraves da API SIDRA do IBGE:

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

# Parametros:
# t/6381 - PNAD Continua Trimestral
# v/4099 - Taxa de desocupacao
# p/all - Todos os periodos disponiveis
```

2.2. TRANSFORMACAO (T)

RELATORIO TECNICO - ANALISE PNAD CONTINUA IBGE

Processos aplicados nos dados:

```
# Script de transformacao
- Limpeza de dados faltantes e inconsistentes
- Conversao de tipos (string para numerico)
- Criacao de metricas derivadas (variacao percentual)
- Normalizacao de nomes de colunas
- Calculo de medias moveis (4 periodos)
- Classificacao 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. Arquivo CSV: pnad_powerbi_pronto.csv (Power BI)
3. Dataset otimizado para analise
```

3. ARQUITETURA DA SOLUCAO

Fluxo completo de dados implementado:

ARQUITETURA DO PROJETO:

```
API IBGE (REST)
  |
Python Scripts (ETL)
  |
SQLite Database
  |
CSV Otimizado
  |
Power BI Dashboard
  |
Insights e Relatorios
```

3.1. Scripts Python Desenvolvidos

pnad_etl.py - Pipeline principal de extracao e transformacao
verificar_dados.py - Validacao e controle de qualidade
preparar_dados_powerbi.py - Otimizacao para visualizacao
powerbi_final.py - Geracao do dataset final
gerar_relatorio_pdf.py - Este relatorio

3.2. Estrutura do Banco de Dados

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

4. ANALISE E INSIGHTS

4.1. Principais Indicadores Calculados

- Taxa de desocupacao trimestral
- Variacao percentual periodo a periodo
- Media historica (2012-2024)
- Classificacao por nivel de desocupacao
- Tendencia (Alta/Baixa/Estavel)
- Media movel (suavizacao de tendencias)

4.2. Insights Identificados

1. TENDENCIA DE LONGO PRAZO: Analise da evolucao da taxa de desocupacao ao longo de 12 anos
2. SAZONALIDADE: Identificacao de padroes trimestrais recorrentes
3. IMPACTO DE EVENTOS: Analise do efeito de crises economicas e pandemias
4. COMPARATIVO HISTORICO: Posicionamento atual em relacao a media do periodo

5. VISUALIZACAO NO POWER BI

5.1. Estrutura do Dashboard

PAGINA 1 - VISAO GERAL

- Grafico de linha evolutivo da taxa de desocupacao
- KPIs principais (taxa atual, media historica, variacao)
- Indicadores de status e tendencia

PAGINA 2 - ANALISE DETALHADA

- Media movel para suavizacao de tendencias
- Analise por categorias de desocupacao
- Tabela detalhada com todos os periodos

PAGINA 3 - RECOMENDACOES

- Insights acionaveis para gestores publicos
- Estrategias para empresas privadas

5.2. Medidas DAX Implementadas

```
// Taxa Atual (ultimo periodo)
Taxa Atual =
    CALCULATE(
        MAX([taxa_desocupacao]),
        LASTDATE([data_referencia])
    )

// Tendencia
```

RELATORIO TECNICO - ANALISE PNAD CONTINUA IBGE

```
Tendencia =  
    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 Publicos

- Identificacao de periodos criticos para intervencao
- Base data-driven para politicas de geracao de emprego
- Monitoramento da efetividade de programas sociais
- Analise de impacto de medidas economicas

6.2. Para Empresas Privadas

- Analise de mercado para decisoes de expansao
- Planejamento estrategico de contratacoes
- Estudo de cenarios economicos para investimentos
- Benchmarking do setor

7. TECNOLOGIAS UTILIZADAS

Camada	Tecnologia	Finalidade
Extracao	Python Requests	Consumo da API IBGE
Processamento	Pandas, NumPy	Transformacao de dados
Armazenamento	SQLite	Banco de dados local
Visualizacao	Power BI	Dashboard interativo
Documentacao	FPDF (Python)	Relatorio tecnico

8. CONSIDERACOES FINAIS

8.1. Licoes Aprendidas

- Importancia da validacao de dados na fonte para qualidade
- Eficiencia do pipeline Python para processos ETL complexos
- Valor da documentacao completa do processo analitico
- Flexibilidade do SQLite para prototipagem e analise

8.2. Proximos Passos Recomendados

- Inclusao de dados regionais (analise por UF)
- Cruzamento com outras bases do IBGE (renda, educacao)
- Automacao da atualizacao trimestral dos dados
- Desenvolvimento de modelos preditivos simples
- Criacao de alertas para tendencias significativas

ANEXOS

Anexo A - Estrutura de Pastas do Projeto

```
IBGE/  
|-- pnad_etl.py (script principal)  
|-- verificar_dados.py (validacao)  
|-- preparar_dados_powerbi.py (otimizacao)  
|-- powerbi_final.py (dataset final)  
|-- gerar_relatorio_pdf.py (este relatorio)  
|-- 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 Continua - IBGE

Indicador principal: Taxa de desocupacao (%)

Periodo coberto: 2012-2024

Frequencia: Trimestral

Total de registros: 162 periodos

Ultima atualizacao: Dados mais recentes disponiveis

Documento gerado automaticamente via Python - Demonstrando habilidades tecnicas em ETL e analise de dados