

UNIGOIÁS – Análise e Desenvolvimento de Sistemas

Trabalho Prático N2 – Estatística com Python

Importações e Exportações de Derivados de Petróleo (2000–2025)

Alunos:

- Thomaz Arthur Correia de Oliveira
- Christopher Fernandes de Barros
- Tomaz Alcides Cavalheiro
- Fabricio de Freitas Oliveira

Goiânia – 2025

Etapa 1 – Limpeza e Organização dos Dados

- Foram removidas duplicatas, registros com valores nulos e volumes iguais a zero.
- Volume importado/exportado convertido de string para float (substituindo vírgula por ponto).
- Total de registros válidos após limpeza: 9.252
- Período disponível: 2000 a 2025

Etapa 2 – Quartis, Percentis e Outliers

- Q1: 96,37 m³ | Mediana (Q2): 7.663,93 m³ | Q3: 133.938,16 m³
- Amplitude Interquartílica (IQR): 133.841,79 m³
- Outliers (Q3 + 1,5xIQR): 991 registros
- Meses com mais registros no quartil superior: Abril, Agosto, Maio, Junho
- Indícios de sazonalidade observados entre Maio e Agosto

Etapa 3 – Distribuição Normal e Z-score

- Z-score calculado para todos os volumes
- 962 valores estão acima de 1 desvio-padrão da média
- Probabilidade de $Z > 1,96 \approx 2,5\%$
- Distribuição assimétrica positiva: muitos valores extremos altos

Etapa 4 – Intervalo de Confiança (IC 95%)

- Amostra aleatória com 300 registros
- Volume: $140.641,74 \pm 31.108,64$ m³ (contém a média populacional)
- Dispêndio: $61.403.455,19 \pm 17.726.191,13$ USD (contém a média populacional)

Etapa 5 – Testes de Hipóteses

Hipótese 1:

- H0: Volume médio de exportação da Gasolina A = Óleo Diesel
- H1: Volume médio é diferente
- Resultado: $p < 0,05 \rightarrow$ Rejeita H0 \rightarrow Diferença estatisticamente significativa

Hipótese 2:

- H0: Proporção de dispêndio > US\$ 500 mil = 0,5
- H1: Proporção diferente de 0,5
- Proporção observada: 67,12%
- Resultado: $p < 0,05 \rightarrow$ Rejeita H0 \rightarrow Proporção estatisticamente diferente de 50%

Conclusão Geral

A análise estatística dos dados da ANP evidenciou a existência de valores extremos relevantes, sazonalidade em alguns meses, e diferenças estatisticamente significativas entre produtos. O projeto permitiu aplicar conceitos de limpeza, quartis, distribuição normal, inferência estatística e testes de hipóteses com dados reais.