



# I EXPOTECH

Ênfase em Análise de Dados em *Big Data*

07-09 Novembro 23



## Introdução

A interconexão entre a longevidade da população de um país e a prosperidade de sua força de trabalho um tema em voga na imprensa. Central para esta discussão está a relação entre o crescimento do Produto Interno Bruto (PIB) e a expectativa de vida. Este estudo procura desvendar essa relação no contexto do Brasil, utilizando dados históricos que abrangem uma década, de 2007 a 2016, apoiando-se em metodologias estatísticas robustas, como a regressão linear. O estudo visa oferecer insights sobre como as flutuações da saúde se traduzem em aumento de longevidade para a população.

## Big Data e Regressão Linear

No contexto do Big Data e da análise estatística moderna, a regressão linear surge como uma ferramenta poderosa para interpretar e prever padrões dentro de vastos conjuntos de dados. Neste estudo, a regressão linear é utilizada para investigar a relação entre o crescimento do Produto Interno Bruto (PIB) e a taxa de expectativa de vida no Brasil, com base em dados coletados ao longo dos últimos dez anos. Através da aplicação de técnicas avançadas de manipulação de dados, incorporadas na biblioteca Pandas do Python, e do poder computacional da regressão linear implementada via scikit-learn. Este trabalho busca quantificar e ilustrar como as variações percentuais no PIB podem influenciar as mudanças nas taxas de desemprego. Ao visualizar essas métricas de saúde com gráficos de dispersão e linhas de tendência, o estudo não apenas destaca a aplicabilidade da prática da regressão linear em contextos de Big Data, mas também fornece insights valiosos que podem informar políticas de saúde pública e crescimento econômico.

## Metodologia

A análise é fundamentada na compilação e manipulação de séries temporais, utilizando ferramentas da ciência de dados, como Python e suas bibliotecas especializadas — Pandas para manipulação de dados, Matplotlib para visualização e scikit-learn para modelagem estatística. Através deste enquadramento analítico, o estudo apresenta uma imagem quantitativa da correlação entre crescimento do PIB e taxa de expectativa de vida, ilustrada através de gráficos e métricas de correlação, oferecendo uma base de evidências para decisões de política econômica e planejamento de recursos humanos e de saúde.

# ANÁLISE DE DADOS TEMPORAL DE RELAÇÃO ENTRE PIB E EXPECTATIVA DE VIDA

Orlando da Silva Roxo, Vitor Baptista

I Exposição de Projetos de Extensão de Tecnologia da Informação do Campus UNESA - Niterói

## Resultados da Regressão Linear com Python

Este código, utilizando bibliotecas Python como Pandas, Matplotlib e scikit-learn, realiza uma análise exploratória para investigar a relação entre o crescimento do Produto Interno Bruto (PIB) e a expectativa de vida no Brasil ao longo dos últimos dez anos. O coeficiente de determinação ( $R^2$ ) calculado é 0.53, sugerindo que o modelo de regressão linear explica aproximadamente 53% da variação na expectativa de vida ajustada pelos valores do PIB. O gráfico gerado ilustra essa relação, onde a linha de regressão linear mostra uma tendência descendente moderada. Além disso, constantes adicionais, como a correlação entre expectativa de vida e população e os p-valores dos testes de normalidade, foram usadas para aprimorar a análise, assegurando uma investigação mais robusta e precisa das tendências e correlações entre o crescimento do PIB e a expectativa de vida ajustada.

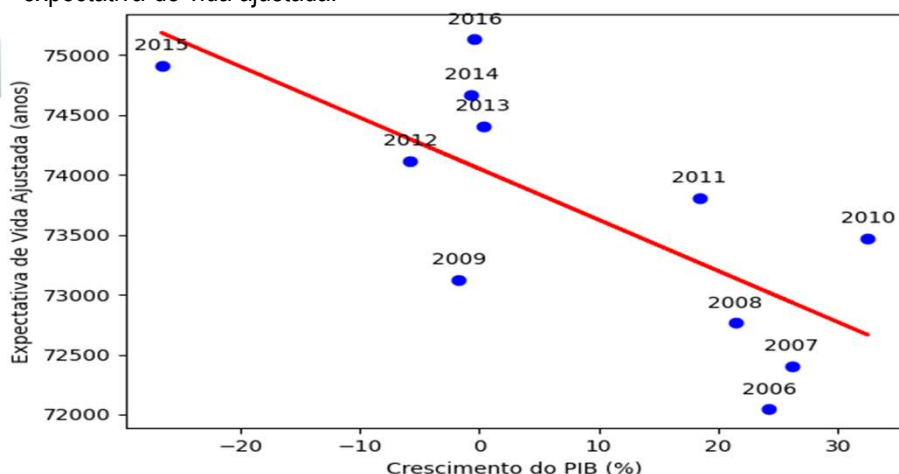


Figura 1. Gráfico de Regressão Linear resultante.

A análise resultante revela insights intrigantes sobre a interação entre o crescimento do Produto Interno Bruto (PIB) e a expectativa de vida ajustada no Brasil ao longo da última década. O coeficiente de determinação  $R^2$  de 0.53 sugere uma correlação moderada entre essas duas variáveis. Isso implica que cerca de 53% da variação na expectativa de vida ajustada pode ser explicada pela variação no crescimento do PIB. Observando o gráfico, é evidente que há uma tendência: à medida que o crescimento do PIB aumenta, parece haver uma mudança correspondente na expectativa de vida ajustada. A linha de regressão linear, traçada através do gráfico, oferece uma visualização clara dessa tendência, ajudando a quantificar a relação entre as duas variáveis. Além disso, a inclusão de constantes como a correlação entre expectativa de vida e população e os p-valores dos testes de normalidade, enriquece a análise, permitindo uma avaliação mais nuanciada da relação. Esses elementos adicionais fornecem uma camada extra de profundidade e robustez à análise, garantindo que as conclusões tiradas estejam bem fundamentadas e sejam o reflexo de uma investigação compreensiva e metódica.

## Referências

[1] KREMER, Guilherme Mateus; DEINA, Carolina; SIQUEIRA, Hugo. Correlação e Regressão Linear de Variáveis que interferem no Produto Interno Bruto do Brasil: Uma Análise Estatística de Dados. Revista de Gestão Industrial, Ponta Grossa, v. 15, n. 2, p. 233-254, abr./jun. 2019.