



RIQUEZA E BEM-ESTAR: O QUE REALMENTE PODE PREDIZER A FELICIDADE?

SANTIAGO, Giulia¹, MOURA, Gabriel², MOURA, Rafael³,
SOUZA, Deivid⁴, PINEDA, Aruane⁵

^{1, 2, 3, 4, 5} Faculdade SENAI-SP Suiço Brasileira, São Paulo, Brasil

Resumo: A relação entre o desenvolvimento econômico e o bem-estar subjetivo populacional é alvo de intensos debates nas ciências sociais e econômicas. Este trabalho analisa o quanto fatores econômicos e sociais influenciam o nível de felicidade global, mensurado pelo *Ladder Score*. Utilizando a base de dados do *World Happiness Report 2024* compreendendo 143 países, aplicou-se um pipeline estatístico e de ciência de dados envolvendo análise de correlação de *Pearson* e Regressão Linear, validada por divisão de treino/teste (80/20) e validação cruzada (*5-fold Cross-Validation*). Os resultados isolados indicam que o suporte social ($R^2 = 0,66$) e o Log do PIB per capita ($R^2 = 0,59$) são os preditores individuais mais fortes. Contudo, no modelo multivariado (que explicou 82% da variação global, com $R^2 = 0,819$), o efeito direto do PIB reduziu-se drasticamente, tornando-se o segundo menor coeficiente devido a efeitos de multicolinearidade. Conclui-se que a riqueza econômica não compra a felicidade de forma direta, mas atua como vetor estrutural indispensável para o financiamento de redes de suporte social, saúde pública e liberdades individuais, que efetivamente consolidam o bem-estar coletivo.

Palavras-chave: Regressão Linear; World Happiness Report; Modelagem Estatística; Análise Preditiva; Ciência de Dados.

INTRODUÇÃO

A relação entre riqueza material e felicidade é uma das questões mais persistentes nas ciências sociais e econômicas. Desde os trabalhos pioneiros de Richard Easterlin na década de 1970, que constataram que o crescimento da renda nacional não significava necessariamente um aumento proporcional da satisfação subjetiva com a vida: o chamado "paradoxo de Easterlin", pesquisadores de diferentes áreas têm buscado entender quais fatores efetivamente determinam o bem-estar das populações. Essa discussão ganhou ainda mais relevância com a consolidação de índices internacionais de felicidade, como o *World Happiness Report* (WHR), que desde 2012 mensura anualmente o bem-estar subjetivo de mais de 140 países por meio do *Ladder Score*, uma escala baseada na "escada de Cantril" em que o respondente avalia sua própria vida entre 0 (pior vida possível) e 10 (melhor vida possível).

O próprio WHR decompõe esse índice em seis variáveis explicativas Log do PIB per capita, suporte social, expectativa de vida saudável, liberdade para fazer escolhas de vida, generosidade e percepção de corrupção, o que abre espaço para investigações quantitativas sobre o peso relativo de cada fator. Estudos anteriores sugerem que a renda per capita está grandemente correlacionada com a felicidade em comparações entre países, mas também apontam que

essa relação não é linear nem isolada: fatores como capital social, saúde e liberdades civis frequentemente emergem como preditores tão ou mais relevantes do que a riqueza econômica, sobretudo quando se controla estatisticamente o efeito conjunto das variáveis. Esse é um ponto metodologicamente sensível, pois indicadores econômicos e sociais tendem a ser altamente correlacionados, onde países com PIB elevado tendem também a ter melhores sistemas de saúde e redes de suporte social mais robustas, gerando assim a multicolinearidade que pode distorcer a interpretação de modelos que analisam essas variáveis de forma isolada.

É nesse contexto que se insere este trabalho, que tem como problema de pesquisa investigar em que medida fatores econômicos e sociais explicam o nível de felicidade subjetiva entre nações, utilizando a base de dados do *World Happiness Report 2024*, compreendendo 143 países. Mais especificamente, busca-se compreender se o poder explicativo do PIB per capita sobre a felicidade varia quando analisado de forma isolada e quando inserido em um modelo multivariado que considera simultaneamente os demais determinantes sociais do bem-estar.

Assim, este trabalho tem como objetivos: (i) identificar, por meio de correlação de *Pearson*, os preditores individuais mais fortemente associados ao *Ladder Score*; (ii) construir e validar um modelo de



regressão linear múltipla, onde aplicou-se a divisão treino/teste (80/20) e validação cruzada de 5 *folds*, sendo capaz de estimar o peso conjunto dessas variáveis sobre a felicidade; e (iii) discutir, à luz da multicolinearidade identificada entre os preditores, em que medida a riqueza econômica atua como causa direta ou como condição estrutural indireta para a construção do bem-estar coletivo.

MATERIAL E MÉTODOS

Essa pesquisa é um estudo quantitativo e exploratório, baseado em modelagem estatística e técnicas de Ciência de Dados aplicadas a indicadores socioeconômicos globais. O fluxo metodológico foi estruturado em quatro etapas principais: tratamento dos dados, análise estatística descritiva e de correlação, modelagem preditiva por regressão multivariada e validação cruzada. O conjunto de dados utilizado foi obtido a partir do *World Happiness Report 2024*, com informações de 143 países. A variável dependente (*target*) definida para o estudo foi o *Ladder Score* (índice de felicidade), enquanto as variáveis independentes (preditores) englobaram seis dimensões: (i) Log do PIB per capita; (ii) Suporte Social; (iii) Expectativa de Vida Saudável ao Nascer; (iv) Liberdade para Fazer Escolhas de Vida; (v) Generosidade; e (vi) Percepção de Corrupção. O *pipeline* de processamento e modelagem foi inteiramente desenvolvido na linguagem *Python*, utilizando as bibliotecas *Pandas* e *NumPy* para manipulação de dados, *Matplotlib* e *Seaborn* para visualização gráfica, e *Scikit-Learn* para aplicação dos modelos estatísticos e de *machine learning*.

Na etapa de preparação dos dados, realizou-se a limpeza do *dataset* com a remoção de duplicidade e o tratamento de registros nulos. A análise inicial consistiu na aplicação da Estatística Descritiva para sumarização das variáveis e no cálculo do Coeficiente de Correlação de *Pearson*, avaliando a força e a direção da associação linear individual entre cada preditor e o *Ladder Score*. Com o intuito de evitar o sobreajuste (*overfitting*) e garantir a capacidade de generalização das estimativas, a base de dados foi dividida aleatoriamente utilizando a função *train_test_split*, na proporção de 80% para o de treinamento e 20% para o teste. Adicionalmente, para mitigar vieses decorrentes da divisão amostral e obter uma métrica de desempenho estatisticamente robusta, foi aplicada a Validação Cruzada *K-Fold* com 5 *folds* ($k=5$). O modelo foi avaliado com base no Coeficiente de Determinação (R^2), no Erro Quadrático Médio (MSE) e no Erro Absoluto Médio (MAE). Por fim, foi realizado o diagnóstico de multicolinearidade para validar as suposições do

modelo multivariado e interpretar adequadamente os pesos de cada variável.

RESULTADOS E DISCUSSÃO

A etapa inicial de Estatística Descritiva caracterizou o comportamento do *Ladder Score* entre os 143 países da amostra (Figura 1). A distribuição mostrou-se moderadamente assimétrica à esquerda, com média de 5,53 e mediana de 5,79, concentrando a maior parte dos países na faixa intermediária entre 5 e 7 pontos. O ranking dos extremos destaca a magnitude da desigualdade global de bem-estar: o grupo dos cinco países mais felizes (Finlândia, Dinamarca, Islândia, Suécia e Israel) apresenta índices entre 7,34 e 7,74, enquanto o grupo dos cinco menos felizes (Congo-Kinshasa, Serra Leoa, Lesoto, Líbano e Afeganistão) registra valores entre 1,72 e 3,29 – uma diferença de até seis pontos na escada de Cantril entre os polos da distribuição. Esse panorama descritivo fornece o pano de fundo empírico para as análises de correlação e regressão discutidas a seguir.

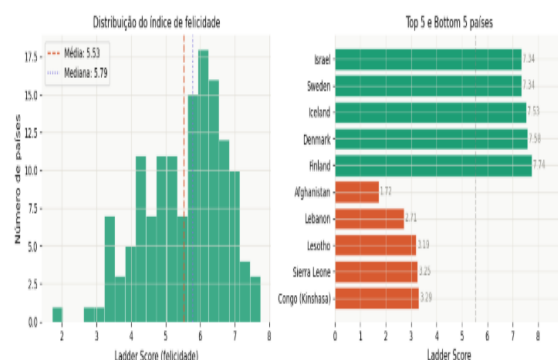


Figura 1. Distribuição do índice de felicidade e ranking dos cinco países mais e menos felizes.

A análise inicial de correlação linear simples, revelou que o Suporte Social apresenta a associação linear mais forte com o índice de felicidade dos países ($r=0,814$; $R^2 = 0,662$), superando levemente o Log GDP per capita ($r = 0,769$; $R^2 = 0,591$). A Expectativa de Vida Saudável posiciona-se logo em seguida, praticamente empatada com o fator econômico ($r = 0,760$; $R^2 = 0,577$). Todas as três variáveis demonstraram significância estatística ($p < 0,001$). Contudo, embora a Liberdade de Escolha ($r = 0,644$) e a Baixa Percepção de Corrupção ($r = 0,452$) também exibam associações positivas significativas, a variável Generosidade não demonstrou relação linear estatisticamente significativa com a felicidade nesta análise isolada ($p = 0,126$, classificada como não significativa).

A Figura 2 ilustra graficamente essa relação entre riqueza econômica e bem-estar subjetivo: cada ponto representa um país posicionado conforme seu Log do PIB per capita e seu *Ladder Score*, com a reta de



regressão simples evidenciando a tendência linear positiva ($r = 0,769$; $R^2 = 0,591$). O gráfico de dispersão deixa evidente que, embora a relação seja consistente em escala global, ela está longe de ser determinística: países como Costa Rica e México alcançam níveis de felicidade superiores ao esperado para sua renda, enquanto Hong Kong e Singapura, apesar do PIB elevado, posicionam-se abaixo da reta de tendência. Esse padrão de dispersão reforça, já nesta etapa exploratória, a ideia central deste trabalho de que a riqueza econômica favorece a felicidade em média, mas não a garante isoladamente, sendo moderada por outros fatores estruturais que serão explorados nas seções seguintes.

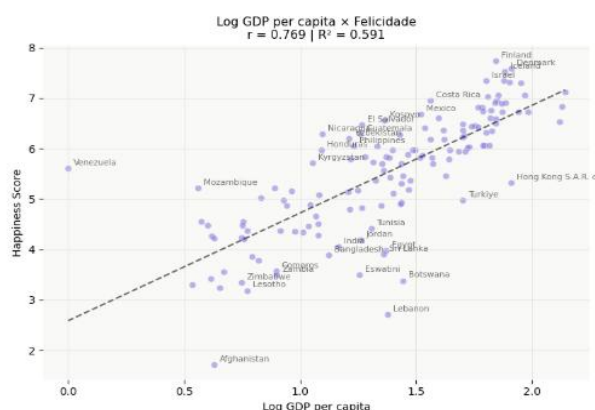


Figura 2. Relação entre Log do PIB per capita e o índice de felicidade.

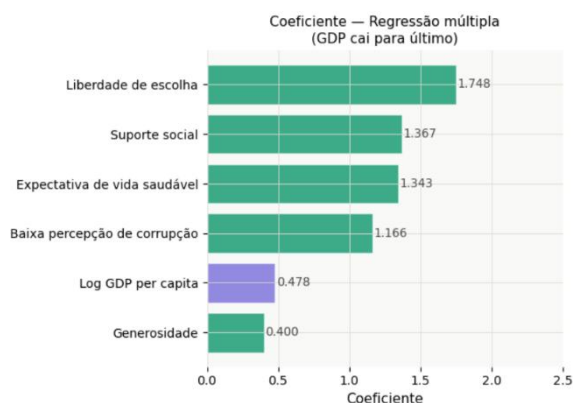


Figura 3. Correlação de cada fator com o índice de felicidade (regressão simples).

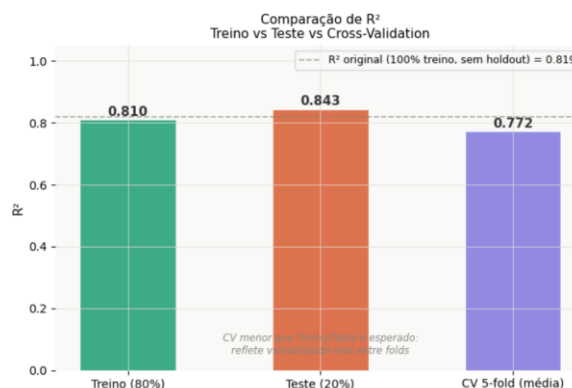


Figura 4. Comparação de Resultados Treino X Teste X Cross Validation.

Para complementar a análise bivariada, construiu-se a matriz de correlação de *Pearson* entre todas as variáveis do estudo (Figura 5), permitindo visualizar simultaneamente a relação do *Ladder Score* com cada preditor e as inter-relações entre os próprios preditores. Observa-se que, além das fortes associações já discutidas entre o índice de felicidade e o Suporte Social ($r = 0,81$), o Log do PIB per capita ($r = 0,77$) e a Saúde ($r = 0,76$), os preditores econômicos e sociais também se correlacionam fortemente entre si, com destaque para o par PIB–Saúde ($r = 0,83$) e PIB–Suporte Social ($r = 0,73$). Esse padrão é o indicio visual da multicolinearidade mencionada na seção de Material e Métodos: países mais ricos tendem sistematicamente a apresentar também melhor saúde e maior coesão social, de modo que essas três variáveis carregam informação estatística sobreposta. Em contrapartida, a Generosidade revela correlações fracas e até negativas com as demais variáveis (incluindo $r = -0,06$ com o PIB), confirmando seu comportamento estatisticamente atípico já identificado na análise individual.

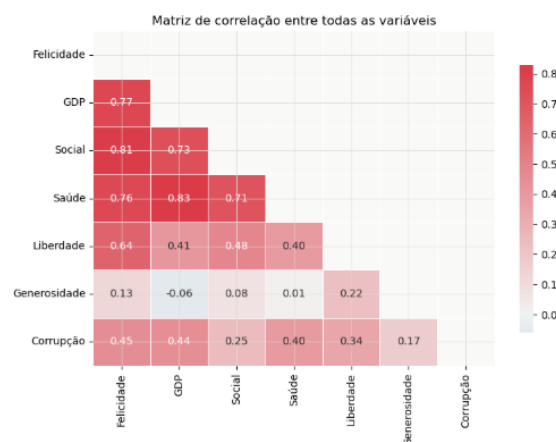


Figura 5. Matriz de correlação entre todas as variáveis do estudo.

O diagnóstico formal de multicolinearidade, conduzido por meio do Fator de Inflação da Variância (VIF), confirma quantitativamente o padrão observado na matriz de correlação (Figura 5). O Log do PIB per capita apresentou o maior VIF do modelo (46,33), seguido pela Expectativa de Vida Saudável (38,53) e pelo Suporte Social (33,14): valores muito acima do limiar de referência ($VIF \geq 10$) costumeiramente adotado na literatura para indicar colinearidade severa entre preditores. Os elevados valores de VIF sugerem que os coeficientes individuais devem ser interpretados com cautela, uma vez que pequenas variações na amostra podem produzir mudanças significativas nas estimativas dos parâmetros. A Liberdade de Escolha apresentou multicolinearidade moderada (18,27), enquanto Generosidade (5,09) e Percepção de Corrupção (3,37) permaneceram dentro de níveis aceitáveis. Esse resultado fornece a explicação estatística para o fenômeno relatado na regressão múltipla: como PIB, Saúde e Suporte Social compartilham grande parte de sua variância, o modelo distribui o poder explicativo entre essas variáveis correlacionadas, fazendo com que o coeficiente direto do PIB per capita seja artificialmente reduzido, mesmo sendo ele, isoladamente, um dos preditores mais fortes da felicidade. Em outras palavras, o efeito do PIB não desaparece: ele é parcialmente absorvido pelas variáveis de saúde e suporte social que historicamente caminham junto com o desenvolvimento econômico.

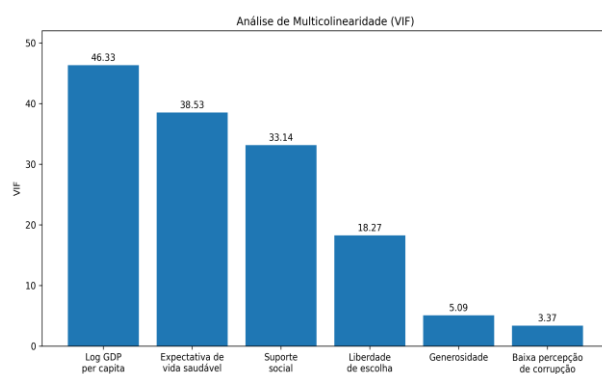


Figura 6. Análise de multicolinearidade entre os preditores (Fator de Inflação da Variância – VIF).

CONCLUSÃO

Este estudo investigou os determinantes do bem-estar subjetivo global por meio de modelagem estatística multivariada, buscando responder se o desenvolvimento econômico é o fator isolado e definitivo para a felicidade de uma nação. Os resultados obtidos rejeitam a premissa de que a renda seja a única variável determinante do bem-estar, revelando uma dinâmica muito mais complexa e interconectada entre fatores materiais, sociais e

institucionais. A análise bivariada evidenciou que o Suporte Social ($R^2 = 0,66$) atua como o preditor isolado mais forte da felicidade, superando o Log GDP per capita ($R^2 = 0,59$), seguido pela Expectativa de Vida Saudável. Essa hierarquia transforma-se significativamente na Regressão Linear Múltipla: ao controlar simultaneamente todas as variáveis, o impacto direto do PIB reduz-se substancialmente, apresentando uma contribuição marginal menor quando comparado aos demais preditores no modelo ajustado. Países com elevado desempenho econômico tendem a apresentar, por consequência, sistemas de saúde pública robustos e maior coesão social, diluindo o efeito puramente monetário na equação da felicidade.

Do ponto de vista metodológico, o modelo multivariado final demonstrou alto poder explicativo e excelente consistência, sendo capaz de explicar 82% da variância global do índice de felicidade ($R^2 = 0,8194$). A aplicação da validação cruzada 5-fold com embaralhamento corrigiu os vieses de ordenação prévia do *dataset* e atestou a estabilidade preditiva do modelo ($R^2 \sim 0,77$). Os desempenhos semelhantes entre treino, teste e validação cruzada sugerem boa capacidade de generalização do modelo. Dessa forma, a resposta à pergunta central deste projeto consolida-se em uma perspectiva estrutural: Os resultados indicam que a relação entre riqueza e felicidade é mediada por fatores sociais e institucionais. Portanto, conclui-se que há, sim, um impacto significativo e estatisticamente relevante da variável econômica, desde que interpretada não como um fim, mas como a base sobre a qual se está a liberdade de escolha, a saúde e o suporte social de uma sociedade.

AGRADECIMENTOS

O grupo agradece à Faculdade de Tecnologia SENAI - Campus Santo Amaro, pela oportunidade de aprendizado e por fornecer as ferramentas e o ambiente acadêmicos necessários para o desenvolvimento das competências em Inteligência e Análise de Dados.

REFERÊNCIAS

- AJAYPAL SinghLO. World Happiness Report 2024. Kaggle, 2024. Disponível em: <https://www.kaggle.com/datasets/ajaypalsinghlo/world-happiness-report-2024>. Acesso em: 30 abr. 2026.
- DIENER, Ed; SELIGMAN, Martin E. P. Beyond money: toward an economy of well-being. Psychological Science in the Public Interest, v. 5,



p. 1-31, 2004. DOI: 10.1111/j.0963-7214.2004.00501001.x.

HAIR JR., Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E.; TATHAM, Ronald L. Análise multivariada de dados. Bookman, Porto Alegre, 2009.

HELLIWELL, John F.; PUTNAM, Robert D. The social context of well-being. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, v. 359, p. 1435-1446, 2004. DOI: 10.1098/rstb.2004.1522.

JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. An Introduction to Statistical Learning: with Applications in R. Springer, New York, 2013.