



PROPOSTA DE UM PROJETO MULTIDISCIPLINAR PARA SER APLICADO COMO UMA INTRODUÇÃO À CIÊNCIA DE DADOS

DOI: 10.37702/2175-957X.COBENGE.2024.5391

Autores: HUGO SAKAI IDAGAWA, LUCAS SANTINO DA SILVA, VICTOR DOS SANTOS RIBEIRO, KAUÃ LAYON LIMA BONFIM, LUIZ EDUARDO BIGOLI E SILVA, SANDRA MÔNICA DO NASCIMENTO MOURA

Resumo: O presente trabalho apresenta uma proposta de projeto para ser utilizado em disciplinas de Ciência de Dados. Esse projeto tem como objetivo a criação de um modelo numérico capaz de prever a resistência mecânica de amostras de compressão fabricadas por impressão 3D em PLA. No desenvolvimento desse projeto, as principais etapas aplicadas em análise de dados são executadas pelos alunos, como a seleção e coleta de dados, eliminação de dados aberrantes ("outliers"), análise de correlação e construção do modelo. Além disso, o projeto incentiva o uso da linguagem de programação Python e as suas bibliotecas de suporte para a criação dos gráficos e a realização dos cálculos de regressão. Como resultado, os procedimentos sugeridos para atingir os objetivos do projeto foram bem sucedidos, onde os outliers foram corretamente identificados e o modelo final foi capaz de prever a resistência mecânica dos corpos de validação com um erro menor do que 10%.

Palavras-chave: Ciência de dados, Construção de modelos numéricos, Compressão de corpos fabricados em impressão 3D, Modelos de resistência mecânica

PROPOSTA DE UM PROJETO MULTIDISCIPLINAR PARA SER APLICADO COMO UMA INTRODUÇÃO À CIÊNCIA DE DADOS

1 INTRODUÇÃO

Na era atual da ciência de dados, professores e cursos superiores são constantemente incentivados a ensinar de maneira integrada a modelagem computacional e a análise de dados. Essa área de modelagem e simulação permite a identificação de propriedades e padrões comuns ao fenômeno estudado e, assim, resultar em novas hipóteses, além da capacidade de gerar previsões para novas situações. Com o intuito de desenvolver as capacidades de pensamento e modelagem computacional nos alunos, a literatura sugere que o uso de um projeto é essencial (ANDERSON et al, 2014; BORNE et al, 2009).

Ao adotar a metodologia de Aprendizado Baseado em Projetos (ou PBL – “Project Based Learning”, em inglês), os alunos têm a oportunidade de participar de situações reais, onde habilidades de solução de problemas, geração de conhecimento e pensamento crítico são desenvolvidas (SHAH RIZODA, 2022). Além disso, essa metodologia proporciona aos alunos uma experiência muito próxima de como um cientista trabalha através da execução de experimentos, coleta e análise de dados, criação e teste de hipóteses e, finalmente, prover evidências de seus resultados (TSIVITANIDOU, 2018).

Essa metodologia incentiva o uso de um projeto multidisciplinar capaz de integrar conteúdos que são normalmente vistos de maneira isolada e, conseqüentemente, apresentar uma melhor compreensão dos fenômenos estudados (SHERNOFF et al, 2017).

Nesse contexto, o trabalho tem como objetivo apresentar um projeto multidisciplinar que pode ser aplicado tanto nos cursos de Engenharia Mecânica, quanto nos cursos de Análise e Desenvolvimento de Sistemas, nas disciplinas correlacionadas à Ciência de Dados.

2 APRESENTAÇÃO DO PROJETO

O projeto proposto consiste em criar um modelo capaz de explicar e prever a resistência mecânica de corpos de prova sujeitos a compressão. Esses corpos de prova são fabricados por impressão 3D em PLA (ácido polilático), com alguns corpos impressos com “defeitos internos” (menor percentual de preenchimento interno).

Como ponto de partida para o projeto, um grande número de corpos de prova pode ser produzido (nesse trabalho foram fabricados 232 corpos) e os alunos, separados em grupos, selecionam aleatoriamente um conjunto de corpos para ensaiar e analisar. Para essa etapa de seleção, os autores recomendam pelo menos um conjunto de 25 corpos por grupo.

Finalizada a etapa de distribuição das amostras entre os grupos, os alunos podem desenvolver as seguintes etapas para concluir o projeto:

- **Seleção e coleta de dados:** decisão sobre quais variáveis devem ser medidas e que podem ser relevantes para explicar a resistência mecânica dos corpos de prova. Nessa etapa, foi instruído que os alunos realizassem a medição das seguintes variáveis: diâmetro, massa, altura, área, densidade e volume.

- **Realização dos ensaios de compressão:** medição da resistência mecânica dos corpos de prova em uma máquina de ensaios universal.
- **Identificação dos corpos de prova com defeitos internos:** etapa em que as variáveis medidas são analisadas para identificar quais corpos de prova apresentam defeitos internos e que devem ser removidos do conjunto de dados.
- **Análise de correlação entre as variáveis:** aplicação de um método para identificar quais variáveis contribuem para a resistência mecânica das amostras. Para essa etapa, aplicou-se a correlação de Pearson.
- **Construção do modelo:** criação de um modelo matemático que relaciona uma variável medida com a resistência dos corpos. Nessa etapa, foi aplicada uma regressão linear para a variável que apresenta a melhor correlação.
- **Validação do modelo:** utilização do modelo criado na etapa anterior para prever a resistência mecânica de uma amostra diferente do conjunto inicial de dados.

Durante a execução do projeto, os alunos podem ser questionados sobre quais variáveis supostamente influenciam na resistência mecânica dos corpos e também realizar uma pesquisa sobre os modelos de compressão apresentados em resistência dos materiais. Para o desenvolvimento do projeto, que inclui a geração dos gráficos e os cálculos de regressão linear, foi utilizada a linguagem de programação Python e as suas várias bibliotecas de suporte. Finalmente, ao término do projeto, os resultados individuais dos grupos podem ser comparados com um resultado onde foi utilizado todo o conjunto de dados, explicitando, assim, a importância do número de amostras para as análises em Ciência de Dados.

3 PROCEDIMENTOS METODOLÓGICOS

Para o projeto proposto, foram inicialmente fabricados, via impressão 3D, 232 corpos de prova de compressão cilíndricos utilizando-se o PLA como material. Para que posteriormente fosse possível realizar uma análise de correlação entre algumas propriedades e a resistência mecânica, os corpos foram fabricados com diâmetro variando de 10,50mm a 14,75mm com passo de 0,25mm e, para a altura, adotou-se uma relação diâmetro/altura de 1:1,5, 1:1,75 ou 1:2, selecionada aleatoriamente para cada corpo. Essas relações foram escolhidas com o objetivo de evitar que os corpos de prova sofressem uma flambagem durante o ensaio de compressão. Todos os corpos foram fabricados utilizando-se uma impressora de extrusão de filamento da marca Ender-3 PRO com parâmetros de impressão apresentados no Quadro 1. A massa total dos 232 corpos de prova foi de 783g e o tempo total de impressão foi de aproximadamente 74h.

Quadro 1 – Parâmetros de impressão dos corpos de prova

Parâmetro	Valor
Altura da camada	0,16mm
Diâmetro do bico extrusor	0,40mm
Nº de linhas de parede	2
Alinhamento da costura	Aleatório
Nº de camadas superiores	3
Nº de camadas inferiores	3
Densidade de preenchimento	80%
Temperatura de impressão	195°C

Fonte: autoria própria.

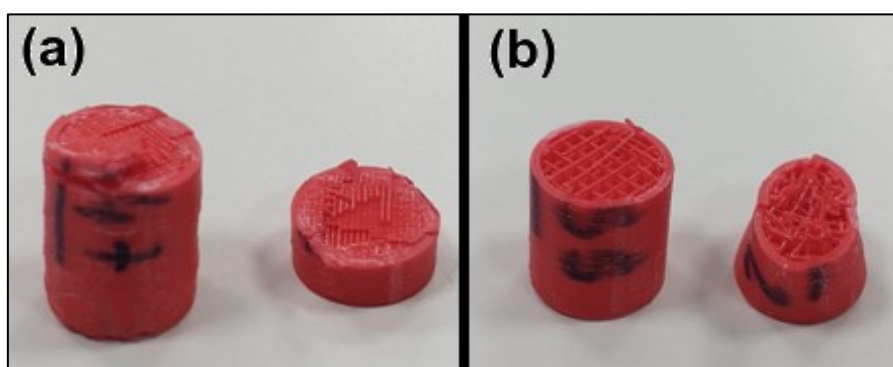
Como um dos objetivos propostos para o projeto é a aplicação de um método de identificação de dados discrepantes (“outliers”), do conjunto de 232 corpos, 13 corpos foram fabricados com defeitos internos, que foram simulados utilizando-se uma densidade de preenchimento igual a 25%. A Figura 1 ilustra uma amostra das peças fabricadas e a Figura 2 apresenta a diferença de preenchimento de uma peça sem defeitos e outra com defeito.

Figura 1 – Amostra dos corpos de prova fabricados.



Fonte: autoria própria.

Figura 2 – Comparação entre um corpo de prova sem defeito interno (a) e outro com defeito interno (b).



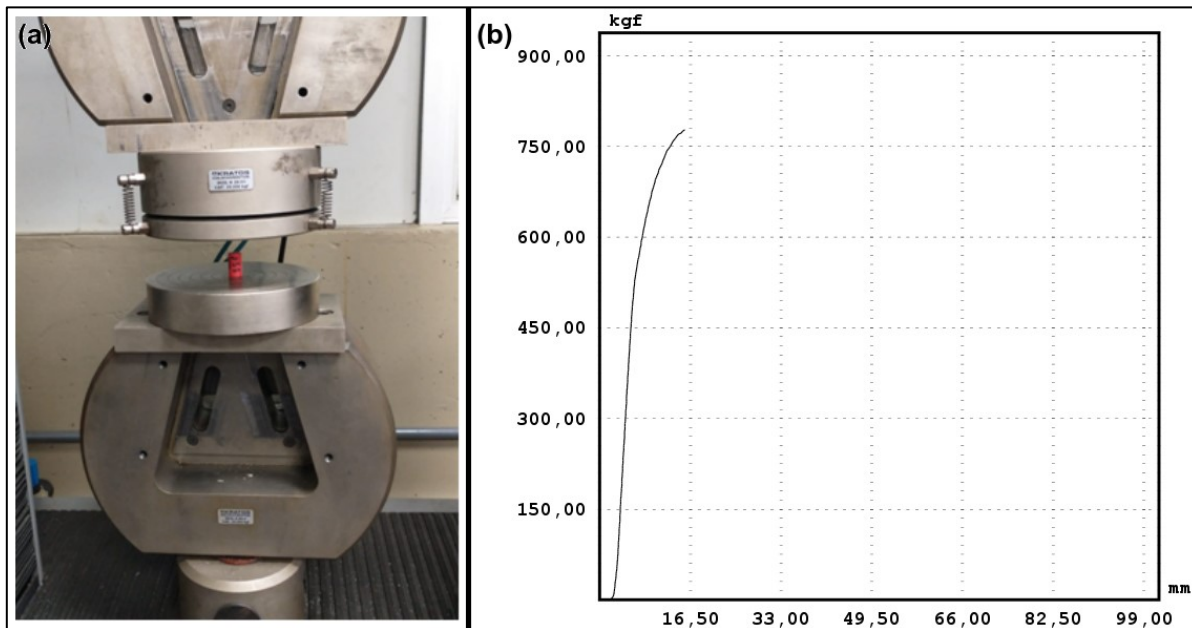
Fonte: autoria própria.

3.1 Coleta de dados e ensaio de compressão

Após a fabricação dos corpos de prova, o próximo passo foi realizar a coleta de dados das variáveis relevantes ao processo. Nesse caso, as variáveis sugeridas para serem avaliadas são: o diâmetro e a altura, os quais foram medidos com um paquímetro, e a massa das peças, que foi medida utilizando-se uma balança com precisão de 0,1g.

Finalizada essa etapa de coleta, procedeu-se ao ensaio de compressão, o qual foi realizado em uma máquina de ensaios universal Kratos equipada com uma célula de carga de 20000kgf (Figura 3a). O ensaio foi realizado com uma velocidade de compressão de 3,0mm/min e o critério adotado para se determinar a resistência mecânica da peça foi o início do escoamento do material, identificado pelo fim da região linear do gráfico (Figura 3b).

Figura 3 – Posicionamento da peça na máquina de ensaios (a); gráfico “Carga X Deslocamento” do ensaio de compressão com o valor de 777kgf adotado como resistência mecânica para a peça (b).



Fonte: autoria própria.

Em seguida, todos os dados coletados foram tabelados e a etapa de análise de dados pode ser iniciada. Além disso, propriedades derivadas das variáveis coletadas podem ser calculadas, como a área da seção dos corpos de prova, o volume e a densidade.

3.2 Análise de dados

A etapa de análise de dados pode ser dividida em três grandes seções: identificação dos “outliers”, análise de correlação e construção de modelo e, finalmente, validação final dos resultados. A primeira seção tem como objetivo a identificação das amostras que foram fabricadas com defeitos internos. Para isso, o método adotado foi a construção dos gráficos de caixa (“boxplots”) das variáveis coletadas e, quando uma peça aparecer como “outlier” em dois ou mais “boxplots”, ela é marcada como uma amostra defeituosa.

Após a identificação das amostras discrepantes, elas podem ser removidas do conjunto de dados inicial e uma análise de correlação pode ser realizada para identificar quais variáveis influenciam na resistência mecânica dos corpos de prova. Nesse projeto, foi utilizado o Coeficiente de Correlação de Pearson (1895) e uma variável foi considerada relevante na previsão da resistência mecânica das peças sempre que esse coeficiente for maior do que 0,70 (AKOGLU, 2018). Em relação a essa análise, os modelos apresentados nas disciplinas de Resistência dos Materiais (BEER, 1995) ditam que a carga suportada por uma estrutura em compressão é proporcional à área da seção transversal, segundo a Equação (1).

$$F = \sigma_{esc} * A \quad (1)$$

Na Equação (1), F corresponde à força que o corpo irá suportar no regime elástico, σ_{esc} corresponde à tensão de escoamento do material e A é a área da seção transversal do corpo. Assim, baseando-se no modelo acima, espera-se que a análise de correlação indique que a área é uma variável relevante na previsão da resistência mecânica dos

corpos ensaiados (coeficiente de Pearson maior do que 0,70), enquanto que as demais variáveis estudadas podem ser descartadas (coeficiente de Pearson menor do que 0,70).

A partir do resultado da correlação, um modelo linear capaz de explicar a resistência mecânica dos corpos de prova pode ser construído aplicando-se o Método dos Mínimos Quadrados (MMQ) (RUGGIERO, 2000).

Finalmente, a última seção do projeto consiste em avaliar a qualidade do modelo obtido na seção anterior comparando-se o resultado experimental de peças que não fazem parte dos dados analisados. Para isso, fabricaram-se dois tipos diferentes de corpos de prova de validação: um cilindro de diâmetro igual a 16mm e altura igual a 24mm e um bloco retangular de base quadrada de 12mm de lado e 18mm de altura. Esses corpos de validação também foram submetidos ao ensaio de compressão nas mesmas condições anteriores e os seus valores de resistência mecânica foram comparados ao valor que o modelo linear previu.

4 RESULTADOS E DISCUSSÕES

A Tabela 1 apresenta uma amostra dos dados coletados os quais foram convertidos para o formato .csv (“comma-separated values”) e toda a análise de dados foi realizada por um script desenvolvido em Python. Nesse script, as bibliotecas “numpy”, “matplotlib”, “pandas” e “seaborn” foram utilizadas para a geração dos gráficos, das correlações e da regressão linear.

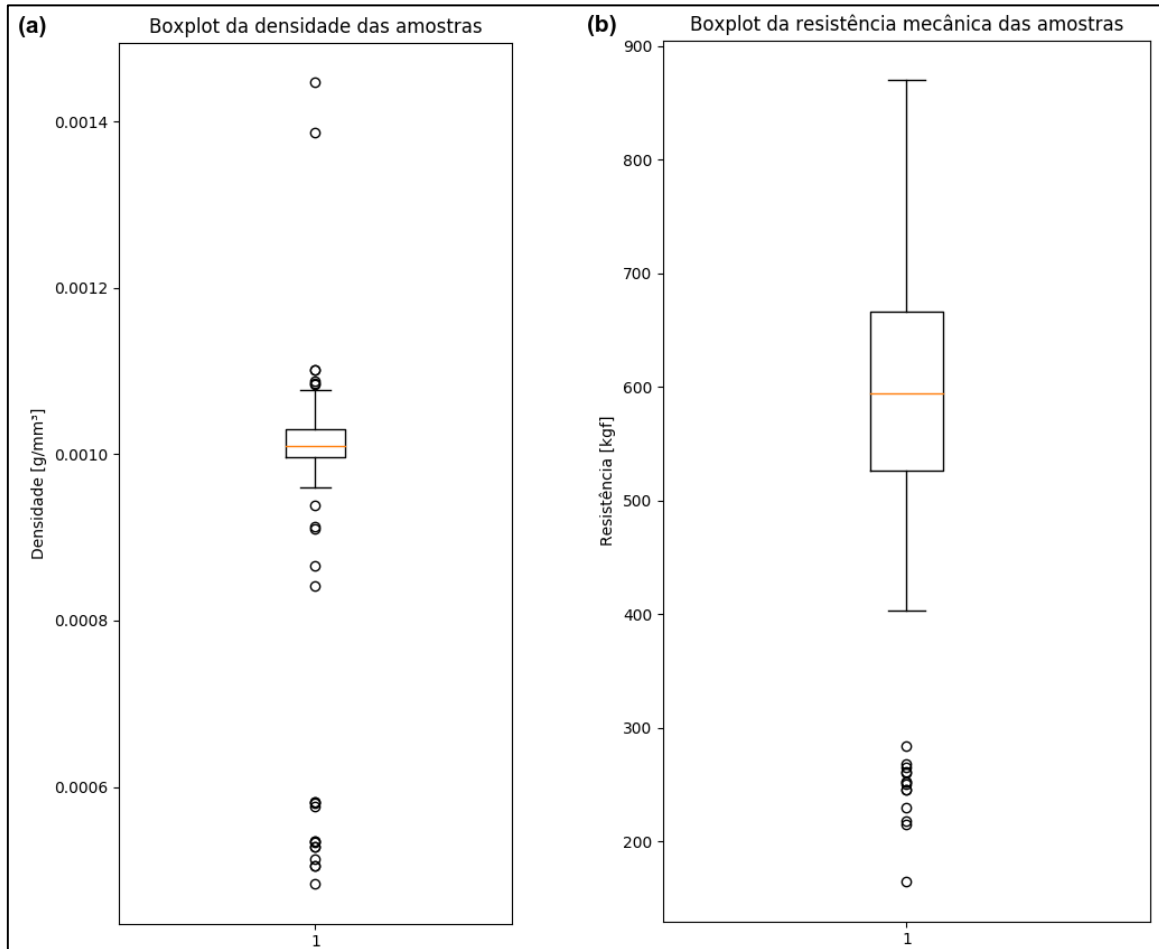
Tabela 1 – Amostra da tabela dos dados coletados para o projeto.

ID	Diâmetro [mm]	Altura [mm]	Massa [g]	Resistência [kgf]
231	14,33	28,46	4,60	670
240	14,30	28,44	4,60	708
114	12,84	25,40	3,30	624
224	14,00	24,42	3,80	754
22	11,59	17,29	1,80	483
47	12,03	25,07	2,40	560
...	...	...	...	...
249	14.82	29.38	5.06	777
117	12.80	25.42	3.33	659

Fonte: autoria própria.

Para a identificação dos “outliers” foram construídos os boxplots da densidade e da resistência mecânica (Figura 4) que identificaram as amostras de número 251 a 263 como pontos discrepantes (13 corpos no total). Esse resultado está de acordo com a fabricação dos corpos e, conforme pode ser observado no diagrama da resistência mecânica (Figura 4b), os “outliers” aparecem todos abaixo do limite inferior boxplot, o que é explicado pelo tipo de defeito que foi induzido nos corpos de prova (preenchimento interno reduzido). Porém, em relação ao diagrama da densidade, existem “outliers” tanto acima quanto abaixo dos limites do boxplot, o que não era esperado, pois o tipo de defeito das peças fabricadas deveria produzir uma densidade mais baixa para o material. Esse comportamento pode ser explicado pela baixa precisão da balança em medir a massa dos corpos de prova, o que poderia ser evitado se corpos de prova com dimensões maiores fossem utilizados para o projeto.

Figura 4 – Boxplots da densidade (a) e da resistência mecânica (b) das 232 amostras fabricadas. Os “outliers” estão representados nos diagramas pelos círculos.

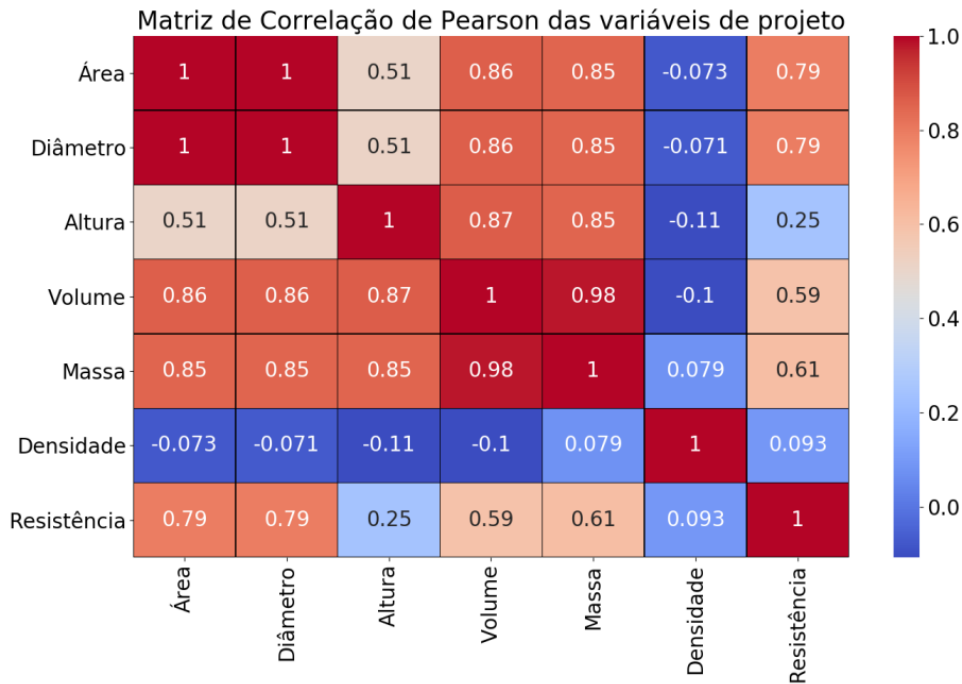


Fonte: autoria própria.

Após a identificação e remoção dos “outliers” do conjunto de dados coletados, procedeu-se para a análise de correlação entre as variáveis. Para essa análise, utilizou-se a função “*corrcoef*” da biblioteca “*numpy*” do Python. A Figura 5 a seguir, apresenta a matriz de correlação dos dados, onde é possível concluir que a resistência mecânica (última linha da matriz) apresenta uma forte correlação com o diâmetro e com a área (ambos com coeficiente de Pearson igual a 0,79). Esse é um resultado esperado, pois está de acordo com os modelos de corpos sujeitos à compressão (Equação 1).

Entre esses resultados, o diâmetro aparece também como uma variável relevante, pois o cálculo da área depende do diâmetro ao quadrado e, apesar dessa relação ser quadrática, não foi possível identificar uma diferença mais relevante na correlação de Pearson porque a diferença entre o menor e o maior diâmetro das peças pode não ter sido o suficiente para influenciar no valor do coeficiente. Assim, uma solução seria a fabricação de corpos com uma maior variação de diâmetros.

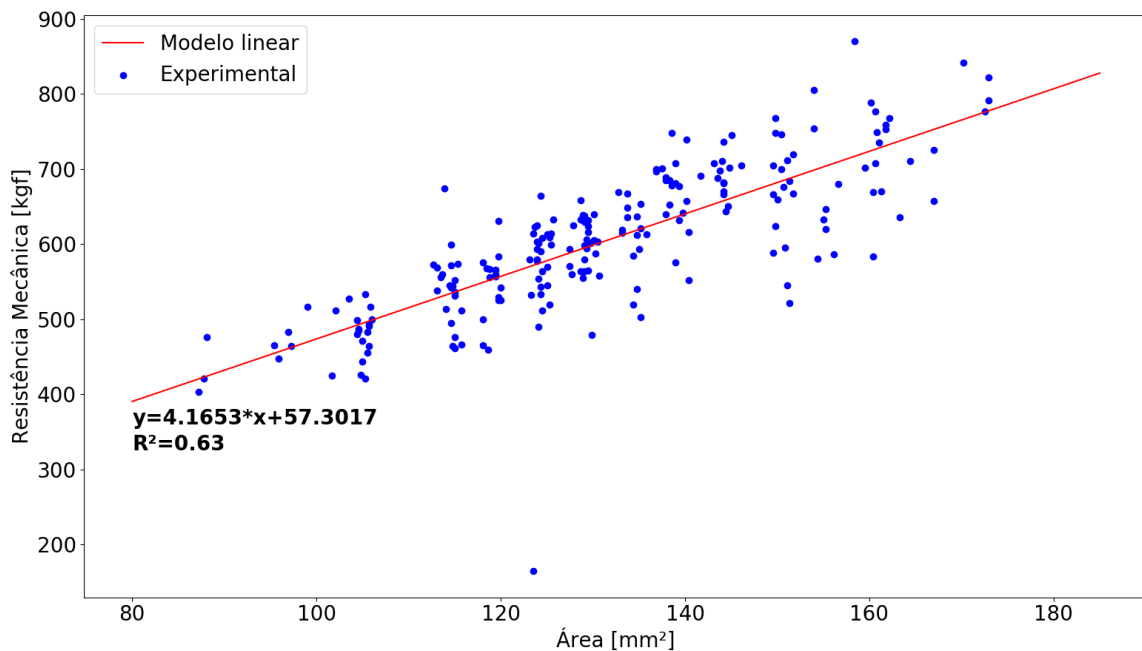
Figura 5 – Matriz de correlação das variáveis analisadas no projeto.



Fonte: autoria própria.

Depois da análise de correlação, foi construído um modelo linear que relaciona a área do corpo de prova com a sua resistência mecânica. Essa função foi obtida utilizando-se a função “polyfit” da biblioteca “numpy” do Python, cujo resultado final está apresentado na Figura 6.

Figura 6 – Comparação entre os dados experimentais e o modelo linear resultante. A equação em destaque na figura corresponde à função obtida com coeficiente de determinação $R^2=0,63$.



Fonte: autoria própria.

Em seguida, com a equação do modelo obtida, compararam-se os resultados dos corpos de validação com os valores previstos pelo modelo (Tabela 2).

Tabela 2 – Comparação entre os resultados experimentais e previstos pelo modelo para os corpos de prova de validação.

Tipo de corpo de prova	Experimental [kgf]	Modelo [kgf]	Erro Relativo [%]
Cilíndrico	909,7	894,8	1,64
Base quadrada	721,3	657,1	8,90

Fonte: autoria própria.

Conforme observado pela Tabela 2, o modelo apresentou um erro relativo de apenas 1,64% para a estimativa de resistência do corpo de prova cilíndrico, enquanto que para o corpo de base quadrada, esse erro foi de 8,90%. Acredita-se que esse erro maior seja devido à geometria do corpo quadrado, onde os cantos podem ter contribuído para uma resistência à compressão mais elevada. Esse efeito talvez possa ser reduzido em corpos de prova cujo lado seja maior do que o fabricado para o projeto.

Uma última etapa de validação foi realizada, onde se comparou os valores de densidade e de resistência à tração dos dados coletados com valores de referência fornecidos por fabricantes (DOUNGKOM, 2017). Esses valores de referência foram multiplicados por 0,80, pois os corpos de prova sem defeito foram fabricados com 80% de preenchimento interno. O resultado dessa comparação pode ser analisado na Tabela 3.

Tabela 3 – Comparação dos valores obtidos de densidade e de resistência à tração com valores de referência de fabricantes.

Propriedade	Valor de referência	Valor de referência ajustado em 80%	Valor coletado nas amostras
Densidade [g/cm³]	1,240	0,992	1,017 ± 0,01
Tensão máxima [MPa]	60,0	48,0	45,2 ± 0,6

Fonte: autoria própria.

Com base nos resultados da Tabela 3, a diferença de menos do que 2,5% entre o valor da densidade de referência e o obtido no projeto é aceitável devido às tolerâncias de medição da balança utilizada. Já em relação à diferença obtida na resistência à tração (menos de 6,5%) pode ser decorrente do processo de impressão 3D, que produz peças com menor resistência mecânica do que as fabricadas via injeção (DOUNGKOM, 2019), principalmente devido à baixa adesão entre as camadas de impressão.

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Nesse artigo, os autores apresentam uma proposta de projeto multidisciplinar que pode ser aplicado aos cursos de Ciência de Dados. O projeto propõe o uso da linguagem de programação Python para realizar toda a análise de dados e a geração dos gráficos, o que se mostrou suficiente para todo o estudo. Além disso, o Python, com suas bibliotecas de suporte, é gratuito, de código aberto e possui uma curva de aprendizado rápida.

Para concluir, os resultados obtidos foram muito satisfatórios, onde a metodologia empregada para identificar as amostras com defeito foi bem sucedida e o modelo final obtido apresentou erros menores do que 10% na previsão da resistência mecânica nos corpos de validação.

Entre as melhorias sugeridas pelos autores seria na fabricação de corpos de prova que abrangessem uma faixa de diâmetros maiores para acentuar o efeito da área na resistência mecânica e, conseqüentemente, gerar um coeficiente de correlação maior

para essa variável. A segunda melhoria seria utilizar um corpo de prova de validação de base quadrada de lado maior, talvez na ordem de 20mm, para reduzir o efeito dos cantos na resistência mecânica.

AGRADECIMENTOS

Os autores agradecem ao SENAI São Paulo pela concessão de auxílio pela participação no Grupo de Pesquisa em Engenharia Elétrica, Eletrônica, Automação e Reindustrialização.

REFERÊNCIAS

AKOGLU, H. User's guide to correlation coefficients. **Turkish Journal of Emergency Medicine**, v.18, n.3, p. 91-93, 2018.

ANDERSON, P.; BOWRING, J.; MCCAULEY, R.; POTHERING, G.; STARR, C. An undergraduate degree in data science: curriculum and a decade of implementation experience. **SIGCSE**, p. 145-150, 2014.

BEER, Ferdinand Pierre; JOHNSTONS, E. Russell Jr. **Resistência dos Materiais**. 3. ed, São Paulo: Pearson Makron Books, 1995.

BORNE, K.; WALLIN, J.; WEIGEL, R. The new computational and data sciences undergraduate program at George Mason University. **Computational Science – ICCS 2009**, v.5545, p. 74-83, 2009.

DOUNGKOM, P.; JIAMIROCH, K. Analysis of temperature distribution for a plastic injection nozzles in Fused deposition model. **FDM, The 2017 Technology Innovation Management and Engineering Science international conference**, p. 245-250, 2017.

DOUNGKOM, P.; JIAMIROCH, K. Analysis of Printing Pattern and Infiltration Percent over the Tensile Properties of PLA Printed Parts by a Fuse Deposition Modelling Printer. **IOP Conference Series: Materials Science and Engineering**, v.501, 2019.

PEARSON, K. Notes on regression and inheritance in the case of two parents. **Proceedings of the Royal Society of London**, v.58, p. 240-242, 1895.

RUGGIERO, Márcia A. Gomes; LOPES, Vera Lúcia da Rocha. **Cálculo numérico: aspectos teóricos e computacionais**. 2. ed, São Paulo: Pearson Makron Books, 2000.

SHAHORIZODA, T. Project-Based Learning as an Effective Teaching Method. **Journal of Ethics and Diversity in International Communication**, v.2, n.2, p. 54-56, 2022.

SHERNOFF, D.J.; SINHA, S.; BRESSLER, D.M.; GINSBURG, L. Assessing teacher education and professional development needs for the implementation of integrated approaches to STEM education. **International Journal of STEM Education**, v.4, n.13, 2017.

TSIVITANIDOU, Olia E. *et al.* **Professional Development for Inquiry-Based Science Teaching and Learning**. 1. ed. Springer Cham. 2018.



COBENGE 2024

52º Congresso Brasileiro de Educação em Engenharia
VII Simpósio Internacional de Educação em Engenharia

Realização



16 a 19/09/2024
Vitória-ES

Organização



PROPOSAL OF A MULTIDISCIPLINARY PROJECT TO BE APPLIED AS AN INTRODUCTION TO DATA SCIENCE

Abstract: *This work proposes a project that can be used in Data Science disciplines. This project aims to build a numerical model capable of predicting the mechanical strength of compression samples manufactured by 3D printing in PLA. During the development of this project, the students apply the main steps used in data analysis, such as data selection and collection, outlier identification, correlation analysis and model construction. Furthermore, the project encourages the use of Python programming language and its support libraries to create graphs and perform regression calculations. As a result, the suggested procedures were successful, where the outliers were correctly identified and the final model was able to predict the mechanical strength of the validation samples with an error of less than 10%.*

Keywords: *data science, numerical model building, 3D printed compression samples, mechanical strength modeling*

Realização:



Organização



