

**INSTITUTO FEDERAL**  
**SÃO PAULO**

# Impacto das especificações técnicas de dispositivos móveis Android no seu preço final: uma análise estatística e de ciência de dados em um estudo de caso na Índia

## Tecnologia em Análise e Desenvolvimento de Sistemas

Aluno: Murillo Teixeira Adão

Professor: Dr. Glauber da Rocha Balthazar



INSTITUTO FEDERAL  
SÃO PAULO

# INTRODUÇÃO

- Avanço rápido de hardware e software.
- Substituição precoce de dispositivos.
- Interrupção de updates acelera a troca de aparelhos.
- Vida útil de 2 a 3 anos.
- Influência das especificações no preço final e consequentemente substituição.





INSTITUTO FEDERAL  
SÃO PAULO

# JUSTIFICATIVA

- Ciclo de vida curto dos smartphones aumenta os resíduos eletrônicos.
- Ausência de updates expõe consumidores a ameaças de segurança.
- Troca precoce de dispositivos é motivada pela falta de suporte.
- Relação entre especificações técnicas e preço final ainda não explorada.





# HIPÓTESE

- **Definição da hipótese:** Determinadas especificações técnicas de smartphones influenciam no seu preço final com base na percepção dos consumidores
- **H0:** Algumas especificações técnicas influenciam o preço final dos smartphones.
- **H1:** Todas as especificações influenciam significativamente o preço final.





INSTITUTO FEDERAL  
SÃO PAULO

# OBJETIVOS

- Identificar as variáveis mais influentes no preço final de smartphones usando estatística e machine learning.
- Tratar dados, realizar análise descritiva, aplicar Análise Discriminante Canônica e utilizar modelos de machine learning para predição de preços.



# FUNDAMENTAÇÃO TEÓRICA

- **Python:** Aplicado na ciência de dados com bibliotecas como Pandas, NumPy e Matplotlib.
- **Estatística Descritiva:** Organiza e resume dados por meio de medidas como média e desvio padrão.
- **Análise Discriminante Canônica:** Técnica multivariada de análise de dados, visa separar grupos de dados e identifica variáveis mais relevantes (discriminantes).



# METODOLOGIA

- **Levantamento Bibliográfico:** Busca de artigos relacionados a aparelhos Android, estatística em geral e estatística descritiva.
- **Dataset:** Obtenção de base de dados que correspondesse a necessidade da análise (Open source) - Copyright [2024] [Yamin Hossain].
- **Tratamento de Dados:** Limpeza, remoção de outliers e normalização para garantir integridade.
- **Ferramentas Utilizadas:** Python e suas bibliotecas (Pandas, Matplotlib), software estatístico IBM® SPSS® Statistics (versão 25).
- **Predição do Preço Final:** Predizer o preço final do aparelho utilizando aprendizagem supervisionada com Python.





# METODOLOGIA

- **Análise Estatística Descritiva:** Cálculos envolvendo MTC, MD e MA.
- **Análise Discriminante Canônica:** Utilizada para identificar as variáveis que mais influenciaram o preço final dos smartphones.

$$Z_n = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Equação 1 - Modelo da Análise Discriminante Canônica

$Z_n$ : variável dependente

$\alpha$ : intercepto

$X_i$ : variáveis explicativas

$\beta_i$ : coeficientes discriminantes de cada variável explicativa





# RESULTADOS

- **Memory (GB):** Média de 4,37 GB, Desvio padrão de 1,89 GB, Variância de 3,99 GB, Amplitude de 8,33 GB, Mediana de 4,00 GB, IQR de 2,08 GB, Curtose de 0,53, Assimetria de 0,53.
- **Storage (GB):** Média de 76,41 GB, Desvio padrão de 54,87 GB, Variância de 3.441,42 GB, Amplitude de 286,33 GB, Mediana de 64 GB, IQR de 54,67 GB, Curtose de 3,91, Assimetria de 1,43.
- **Selling Price Dollar:** Média de 260,10 USD, Desvio padrão de 149,37 USD, Variância de 22.799,49 USD, Amplitude de 870,90 USD, Mediana de 228,56 USD, IQR de 160,15 USD, Curtose de 1,57, Assimetria de 1,57.

# RESULTADOS

- Variáveis para ADC: Color, MemoryGB, StorageGB e SellingPriceDollar
- Aplicando a equação no dataset para a análise geral as variáveis independentes analisadas foram: Colors, MemoryGB, StorageGB, e SellingPriceDollar. Já a variável dependente foi Brands.
- A análise revelou duas funções discriminantes significativas, sendo que a primeira explicou 60% da escolha da marca em função da quantidade de memória (MemoryGB), e a segunda explicou 40% em função do preço final de venda (SellingPriceDollar).
- Análise individual pra cada marca.



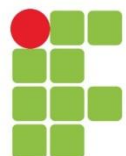
# RESULTADOS

Fabricante	Valores (US\$)		
	Baixo	Médio	Alto
Asus	54-93	94-188	189-761
Google Pixel	145.02-483.37	543.79-858.01	918.43-1111.78
LG	60-120	121-241	241-726
Lenovo	42-90	91-130	130-363
Motorola	65-120	121-181	182-1087
Nokia	56-120	121-157	158-680
Oppo	60-167	168-229	230-737
Poco	96-145	146-204	205-375
Samsung	57-170	171-289	290-1909
Xiaomi	78-138	139-228	229-664
Realme	78-132	133-217	218-580
Vivo	96-193	194-277	278-966

Tabela: Categorização da variável dependente

Fabricante	Determinação da função			
	Lambda de Wilks	Correlação Canônica	Variáveis discriminantes	
			Primeira	Segunda
Asus	0.437	0.740	Memory (0.694)	Storage (0.228)
Google Pixel	0.488	0.671	Memory (0.521)	Rating (0.409)
LG	0.695	0.540	Memory (0.359)	Rating (0.275)
Lenovo	0.746	0.549	Memory (0.971)	Storage (0.261)
Motorola	0.622	0.615	Rating (0.625)	Storage (0.259)
Nokia	0.692	0.741	Storage (0.510)	Memory (0.492)
Oppo	0.483	0.717	Memory (0.947)	Storage (0.083)
Poco	0.367	0.776	Memory (0.629)	Rating (0.255)
Samsung	0.966	0.673	Memory (0.683)	Rating (0.226)
Xiaomi	0.289	0.842	Memory (0.858)	Storage (0.114)
Realme	0.464	0.729	Memory (0.611)	Storage (0.356)
Vivo	0.819	0.356	Inconclusivo	

Tabela: Discriminação individual por marca



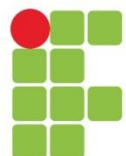
# RESULTADOS

Fabricante (Brands)	DT			KNN			LR		
	AC	PC	PE	AC	PC	PE	AC	PC	PE
Asus	0.94	94.12%	5.88%	0.79	79.41%	20.59%	0.88	88.24%	11.76%
Google Pixel	0.77	77.78%	22.22%	0.77	77.78%	22.22%	0.88	88.89%	11.11%
LG	0.58	58.62%	41.38%	0.51	51.72%	48.28%	0.48	48.28%	51.72%
Lenovo	0.5	50%	50%	0.61	61.11%	38.89%	0.5	50.00%	50.00%
Motorola	0.51	51.61%	48.39%	0.58	58.06%	41.94%	0.41	41.94%	58.06%
Nokia	0.57	57.14%	42.86%	0.51	51.43%	48.57%	0.54	54.29%	45.71%
OPPO	0.71	71.05%	28.95%	0.65	65.79%	34.21%	0.61	61.84%	38.16%
POCO	0.59	59.09%	40.91%	0.5	50.00%	50.00%	0.5	50.00%	50.00%
SAMSUNG	0.71	71.81%	28.19%	0.68	68.09%	31.91%	0.58	58.51%	41.49%
Xiaomi	0.75	75.86%	24.14%	0.68	68.97%	31.03%	0.75	75.86%	24.14%
Realme	0.72	72.53%	27.47%	0.70	70.33%	29.67%	0.79	79.12%	20.88%
Vivo	0.45	45.71%	54.29%	0.34	34.29%	65.71%	0.4	40.00%	60.00%

**DT:** DecisionTree; **KNN:** KNeighborsClassifier; **LR:** LogisticRegression; **AC:** Acuracia; **PC:** Porcentagem de acerto; **PE:** Porcentagem de erro;

Tabela: Predição de dados





# RESULTADOS

Fabricante (Brands)	RF			SVM		
	AC	PC	PE	AC	PC	PE
<i>Asus</i>	0.70	70.59%	29.41%	0.73	73.53%	26.47%
<i>Google Pixel</i>	0.88	88.89%	11.11%	0.77	77.78%	22.22%
<i>LG</i>	0.55	55.17%	44.83%	0.48	48.28%	51.72%
<i>Lenovo</i>	0.5	50%	50%	0.55	55.56%	44.44%
<i>Motorola</i>	0.51	51.61%	48.39%	0.54	54.84%	45.16%
<i>Nokia</i>	0.54	54.29%	45.71%	0.65	65.71%	34.29%
<i>OPPO</i>	0.68	68.42%	31.58%	0.51	51.32%	48.68%
<i>POCO</i>	0.59	59.09%	40.91%	0.54	54.55%	45.45%
<i>SAMSUNG</i>	0.65	65.96%	34.04%	0.59	59.04%	40.96%
<i>Xiaomi</i>	0.79	79.31%	20.69%	0.74	74.14%	25.86%
<i>Realme</i>	0.75	75.82%	24.18%	0.63	63.74%	36.26%
<i>Vivo</i>	0.51	51.43%	48.57%	0.37	37.14%	62.86%

**RF:** *RandomForest*; **SVM:** *SupportVectorMachine*; **AC:** Acurácia; **PC:** Porcentagem de acerto; **PE:** Porcentagem de erro;

Tabela: Predição de dados



# CONCLUSÕES

- A análise estatística revelou a alta diversidade do mercado de smartphones.
- A ADC identificou memória e armazenamento como as especificações mais influentes no preço dos smartphones, destacando padrões claros entre grupos de dispositivos.
- Modelos de aprendizagem de máquina reforçaram as análises, com destaque para RandomForest, que combinou estabilidade e precisão, e SVM, eficaz em capturar padrões complexos.
- $H_0$  estava correta.
- Contribuição para entendimento da dinâmica de preços no mercado de smartphones Android.
- Publicação na revista Journal of Retailing and Consumer Services

<https://www.sciencedirect.com/journal/journal-of-retailing-and-consumer-services>





# BIBLIOGRAFIA PRINCIPAL

- PM3. Python para análise de dados: como fazer e principais bibliotecas. Disponível em: <https://www.cursospm3.com.br/blog/python-para-analise-de-dados/>. Acesso em: 25 set. 2024.
- Guedes, D., Vicente, R., Siqueira, C. Estatística Descritiva. IME-USP, 2021. Disponível em: [https://www.ime.usp.br/~rvicente/Guedes\\_etal\\_Estatistica\\_Descritiva.pdf](https://www.ime.usp.br/~rvicente/Guedes_etal_Estatistica_Descritiva.pdf). Acesso em: 23 set. 2024.
- SAS Institute. Canonical Discriminant Analysis. Disponível em: <https://www.sfu.ca/sasdoc/sashtml/insight/chap40/sect7.htm>. Acesso em: 23 set. 2024.
- MOORE, David S.. A estatística básica e sua prática. 9. ed. Rio de Janeiro: Ltc, 2023. 626 p.
- BRUCE, Peter. Estatística prática para cientistas de dados: 50 conceitos essenciais. Rio de Janeiro: Alta Books, 2019. 299 p.
- Fávero LP, Belfiore SP, Silva FL, Chan BL (2009) Análise de dados: modelagem multivariada para tomada de decisões. Campus-Elsevier, Rio de Janeiro.
- ESCOVEDO, Tatiana. Introdução a data science: algoritmos de machine learning e métodos de análise. São Paulo: Casa do Código, 2022. 272 p.

