

UNIVERSIDADE DE SÃO PAULO
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

JOÃO FONTANEZI
RENAN BEZERRA GUIMARÃES
GUSTAVO HENRIQUES VIEIRA
GABRIEL MONTEIRO DE SOUZA

ANÁLISE DE GASTOS COM EXPLORAÇÃO ESPACIAL
ACH0021 - TRATAMENTO E ANÁLISE DE DADOS / INFORMAÇÕES (2023)
Docente: Karla Roberta P. S. Lima

SÃO PAULO
2023



EACH

Escola de Artes, Ciências e Humanidades
da Universidade de São Paulo

SUMÁRIO

1 DEFINIÇÃO DO PROBLEMA.....	3
2 PLANEJAMENTO.....	4
3 DADOS	5
4 POPULAÇÃO E AMOSTRAS.....	7
5 DISTRIBUIÇÕES DE FREQUÊNCIAS.....	9
6 REPRESENTAÇÃO GRÁFICA.....	12
7 MEDIDAS DE POSIÇÃO.....	14
REFERÊNCIAS BIBLIOGRÁFICAS.....	16



1 DEFINIÇÃO DO PROBLEMA

Após o fim da Segunda Guerra Mundial, no fim dos anos 50, a humanidade entrou em uma nova era de evolução tecnológica, sendo definida como a “Corrida Espacial”. Tendo como protagonistas os Estados Unidos e a URSS, com início formal em 1957, o período da Corrida Espacial foi um episódio de disputa de poder, aonde aquele que conseguisse atravessar a fronteira espacial teria ao seu lado não somente a admiração e atenção global, mas também o domínio científico-militar. Logicamente, ambas as potências não mediram esforços para alcançarem seus objetivos, criando agências como a NASA, coordenando programas espaciais, formando e captando profissionais e construindo satélites e foguetes.

Tais operações precisaram de investimentos massivos, tendo em vista a grandeza de tais conquistas, além das dificuldades dos processos em si. Com gastos nas casas dos milhões, e pesquisas que perduram até a atualidade, é lógico ponderar sobre sua importância, já que, o planeta e a sociedade como um todo possuem diversos problemas que seriam reduzidos significativamente com essa proporção de investimentos. Nesta análise de dados, busca-se compreender a proporção dos gastos com exploração espacial na história e na atualidade, seus resultados, sua viabilidade financeira, sua significância para a humanidade e para o mundo, com base na argumentação pautada sobre os dados apresentados.

2 PLANEJAMENTO

O planejamento desta análise foi pautado nos seguintes tópicos:

- Definir as motivações para a exploração espacial.
- Compreender os processos de exploração e pesquisa.
- Coletar, por meio de banco eletrônico, dados brutos.
- Analisar os dados coletados, contemplando a magnitude dos investimentos na área, bem como seus resultados, taxas de sucesso e retorno em forma de conhecimento científico, além das contribuições para a humanidade.
- Argumentar, com base nos resultados da análise, se os processos de exploração são, ou não, viáveis financeiramente e cientificamente, e se são significativos para o avanço da sociedade como um todo.

3 DADOS

- Coleta**

Para a coleta dos dados, o grupo optou pela pesquisa bibliográfica, analisando a base de dados brutos hospedada no portal “[Kaggle](#)”, denominada como “[Space Mission Launches: All the space missions since 1947](#)”, além do website “[Next Spaceflight](#)”, especializado em registro de dados sobre lançamentos de foguetes. O portal foi escolhido por sua densidade de dados, apresentando bases densas e complexas, que proporcionam uma análise mais precisa e auxiliam na compreensão dos resultados, permitindo por consequência uma melhor argumentação para os problemas definidos.

A base de dados está no seguinte formato:

	X	Organisation	Location	Date	Detail	Rocket_Status	Price	Mission_Status
1	0	SpaceX	LC-39A, Kennedy Space Center, Florida, USA	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5 Starlink V1 L9 & BlackSky	StatusActive	50.00	Success
2	1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Center, ...	Thu Aug 06, 2020 04:01 UTC	Long March 2D Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success
3	2	SpaceX	Pad A, Boca Chica, Texas, USA	Tue Aug 04, 2020 23:57 UTC	Starship Prototype 150 Meter Hop	StatusActive	N/A	Success
4	3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:25 UTC	Proton-M/Briz-M Ekspress-80 & Ekspress-103	StatusActive	65.00	Success
5	4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541 Perseverance	StatusActive	145.00	Success
6	5	CASC	LC-9, Taiyuan Satellite Launch Center, China	Sat Jul 25, 2020 03:13 UTC	Long March 4B Ziyuan-3 03, Apocalypse-10 & NJU...	StatusActive	64.68	Success
7	6	Roscosmos	Site 31/6, Baikonur Cosmodrome, Kazakhstan	Thu Jul 23, 2020 14:26 UTC	Soyuz 2.1a Progress MS-15	StatusActive	48.50	Success
8	7	CASC	LC-101, Wenchang Satellite Launch Center, China	Thu Jul 23, 2020 04:41 UTC	Long March 5 Tianwen-1	StatusActive	N/A	Success
9	8	SpaceX	SLC-40, Cape Canaveral AFS, Florida, USA	Mon Jul 20, 2020 21:30 UTC	Falcon 9 Block 5 ANASIS-II	StatusActive	50.00	Success
10	9	JAXA	LA-Y1, Tanegashima Space Center, Japan	Sun Jul 19, 2020 21:58 UTC	H-IIA 202 Hope Mars Mission	StatusActive	90.00	Success
11	10	Northrop	LP-08, Wallops Flight Facility, Virginia, USA	Wed Jul 15, 2020 13:46 UTC	Minotaur IV NROL-129	StatusActive	46.00	Success
12	11	ExPace	Site 95, Jiuquan Satellite Launch Center, China	Fri Jul 10, 2020 04:17 UTC	Kuaizhou 11 Jilin-1 02E, CentiSpace-1 S2	StatusActive	28.30	Failure
13	12	CASC	LC-3, Xichang Satellite Launch Center, China	Thu Jul 09, 2020 12:11 UTC	Long March 3B/E Apstar-6D	StatusActive	29.15	Success
14	13	IAI	Pad 1, Palmachim Airbase, Israel	Mon Jul 06, 2020 01:00 UTC	Shavit-2 Ofek-16	StatusActive	N/A	Success
15	14	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Center, ...	Sat Jul 04, 2020 23:44 UTC	Long March 2D Shiyang-6 02	StatusActive	29.75	Success
16	15	Rocket Lab	Rocket Lab LC-1A, M7h1a Peninsula, New Zealand	Sat Jul 04, 2020 21:19 UTC	Electron/Curie Pics Or It Didn't Happen	StatusActive	7.50	Failure
17	16	CASC	LC-9, Taiyuan Satellite Launch Center, China	Fri Jul 03, 2020 03:10 UTC	Long March 4B Gaofen Duomo & BF-02	StatusActive	64.68	Success
18	17	SpaceX	SLC-40, Cape Canaveral AFS, Florida, USA	Tue Jun 30, 2020 20:10 UTC	Falcon 9 Block 5 GPS III SV03	StatusActive	50.00	Success
19	18	CASC	LC-2, Xichang Satellite Launch Center, China	Tue Jun 23, 2020 01:43 UTC	Long March 3B/E Beidou-3 G3	StatusActive	29.15	Success
20	19	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Center, ...	Wed Jun 17, 2020 07:19 UTC	Long March 2D Gaofen-9 03, Pixing III A & HEAD-5	StatusActive	29.75	Success
21	20	SpaceX	SLC-40, Cape Canaveral AFS, Florida, USA	Sat Jun 13, 2020 09:21 UTC	Falcon 9 Block 5 Starlink V1 L8 & SkySat 16 to 18	StatusActive	50.00	Success
22	21	Rocket Lab	Rocket Lab LC-1A, M7h1a Peninsula, New Zealand	Sat Jun 13, 2020 05:12 UTC	Electron/Curie Don't stop me now!	StatusActive	7.50	Success

- Organização e apresentação dos dados**

Em questão de organização e apresentação dos dados, foram selecionados para o estudo do atributo estatístico e posterior tratamento estatísticos, as orientações de análise de dados do livro (TAVARES, MARCELO, 2014), o software de planilhas eletrônicas, “Microsoft Excel”,

e a linguagem de programação “R”, orientada à manipulação, análise e visualização de dados, por meio do software “RStudio”. Serão apresentados dados tanto por tabelas, quanto por gráficos.

4 POPULAÇÃO E AMOSTRAS

- **População**

A população selecionada para a análise consiste nos países que já realizaram lançamentos espaciais, tendo em vista que possuem em comum altos valores direcionados para investimento, agências espaciais dedicadas, e lançamentos de equipamentos para pesquisa e exploração extraterrestre.

- **Amostras**

Dentro dos dados referentes à exploração espacial, serão analisados como amostras as tendências de gastos com as operações em cada organização e a comparação com a média global, utilizando as variáveis disponibilizadas na base de dados.

- **Variáveis**

- Variáveis Qualitativas Nominais: Nome da organização, localização, dia, data e hora de lançamento, nome do equipamento, status do foguete, e status da missão.
- Variáveis Qualitativas Ordinais: Posição das organizações ao nível de gastos, posição dos foguetes por custo e posição das organizações em taxa de sucesso.
- Variáveis Quantitativas Discretas: Número de organizações envolvidas em pesquisas espaciais, número de lançamentos de foguetes por organização, número de missões bem sucedidas e fracassadas e número de missões ativas e desativadas.

- Variáveis Quantitativas Contínuas: Gastos com lançamentos por organização.

- **Representatividade**

A representatividade selecionada para a população e amostras na análise foi a amostragem probabilística, já que todos os elementos da população e suas amostras têm a mesma possibilidade de serem analisados, tendo em vista que estão todos limitados por uma mesma base de dados, de características comuns, além de ser uma população finita e acessível.



5 DISTRIBUIÇÕES DE FREQUÊNCIAS

Para distribuição de frequências, foi determinado uma análise no total investido pelas organizações pelas tentativas de lançamento, valores dados em milhões de dólares

Classes	Freq. Absoluta	Freq. Relativa	Freq. Acumulada	Freq. Acumulada (relativa)
[0,12)	28	0.03	28	0.03
[12,24)	20	0.021	48	0.051
[24,36)	169	0.178	217	0.229
[36,48)	131	0.138	348	0.367
[48,60)	103	0.109	451	0.475
[60,72)	112	0.118	563	0.593
[72,84)	7	0.007	570	0.601
[84,96)	39	0.041	609	0.642
[96,108)	0	0	609	0.642
[108,120)	52	0.055	661	0.697
[120,132)	18	0.019	679	0.715
[132,144)	12	0.013	691	0.728
[144,156)	18	0.019	709	0.747
[158,168)	15	0.016	724	0.763
[168,180)	0	0	724	0.763
[180,192)	3	0.003	727	0.766
[192,204)	75	0.079	802	0.845
[204,216)	0	0	802	0.845
[216,228)	0	0	802	0.845
[228,240)	0	0	802	0.845
[240,252)	0	0	802	0.845
[252,264)	0	0	802	0.845
[264,276)	0	0	802	0.845



[276,288)	0	0	802	0.845
[288,300)	0	0	802	0.845
[300,312)	0	0	802	0.845
[312,324)	0	0	802	0.845
[324,336)	0	0	802	0.845
[336,348)	0	0	802	0.845
[348,360)	11	0.012	813	0.857
[360,372)	0	0	813	0.857
[372,384)	0	0	813	0.857
[384,396)	0	0	813	0.857
[396,408)	0	0	813	0.857
[408,420)	0	0	813	0.857
[420,432)	0	0	813	0.857
[432,444)	0	0	813	0.857
[444,456)	136	0.143	949	1
Total	949	1.001		

Cálculos:

$$A = \max - \min = 450 - 5.3 = 444.7$$

Como $n > 100$, então:

$$k = 5 \times \log(949) = 34.27704$$

$$c = \frac{A}{k - 1} \approx 12$$

$$\text{Limite inf. 1ª classe} = 5.3 - \frac{c}{2} \approx 0$$

$$fr_i = \frac{fa_i}{\sum_{i=1}^n fa_i} = \frac{fa_i}{949}$$

Também foi analisado a taxa de sucesso das tentativas de lançamento:



EACH

Escola de Artes, Ciências e Humanidades
da Universidade de São Paulo

Status da missão	Freq. Absoluta	Freq. relativa
Sucesso	3879	0.9
Falha	339	0.08
Falha parcial	102	0.02
Falha no pré lançamento	4	0.001
Total	4324	1.001



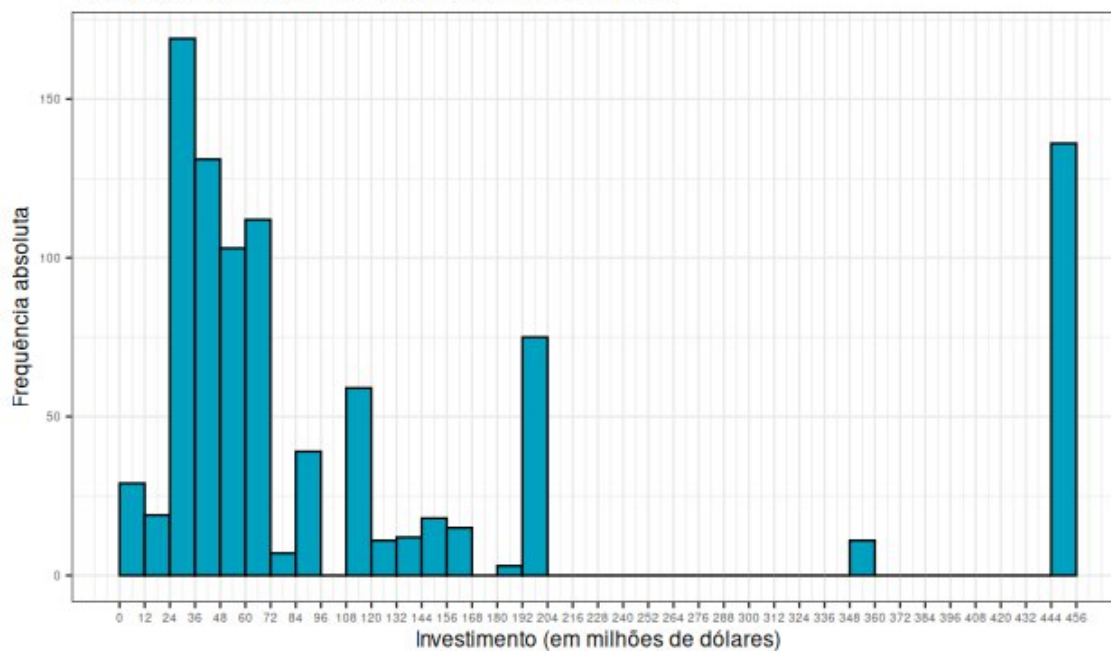
EACH

Escola de Artes, Ciências e Humanidades
da Universidade de São Paulo

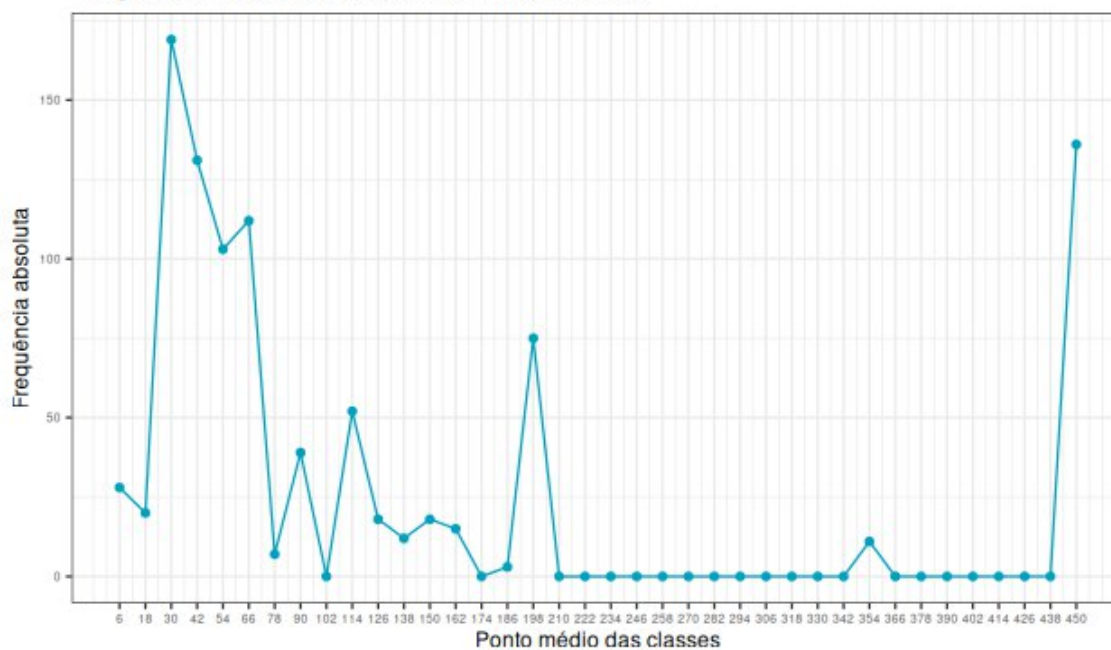
6 REPRESENTAÇÃO GRÁFICA

Os dois gráficos abaixo, mostram uma análise da frequência pelas classes de investimentos totais das organizações

Histograma da frequência absoluta de investimentos



Polígono da frequência absoluta de investimentos

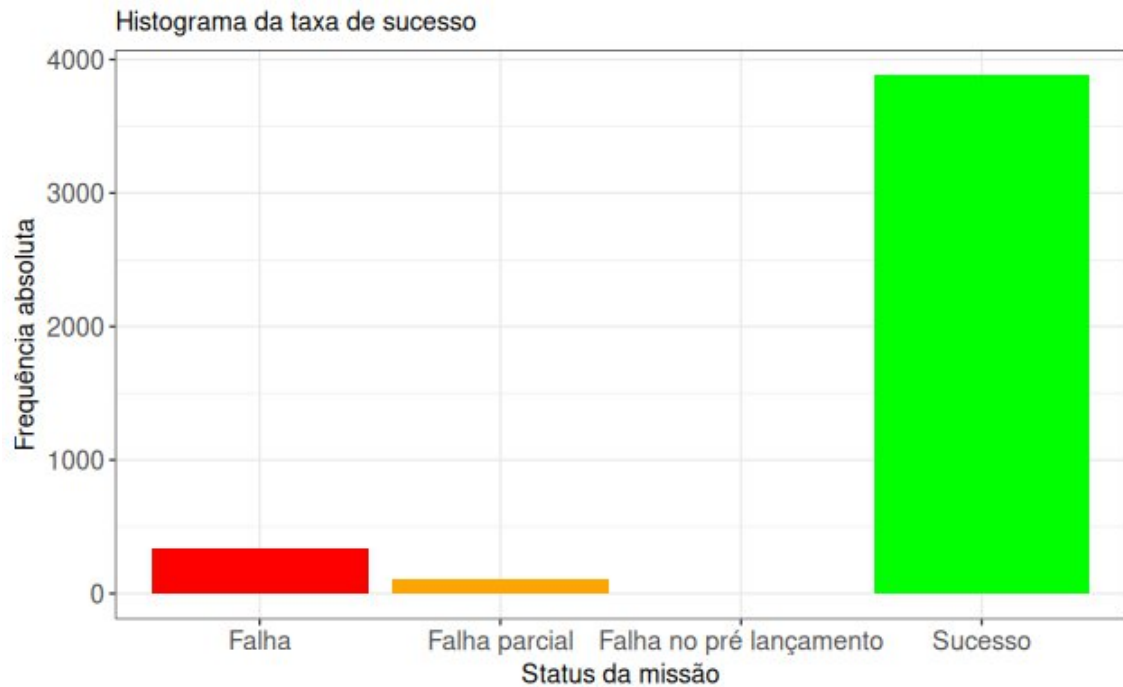




EACH

Escola de Artes, Ciências e Humanidades
da Universidade de São Paulo

O gráfico abaixo, mostra uma análise na frequência absoluta pelo status das missões



7 MEDIDAS DE POSIÇÃO

- **Média**

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Utilizando o cálculo da média aritmética simples, verificou-se o seguinte resultado:

MÉDIA DE GASTOS COM LANÇAMENTOS (EM MILHÕES DE DÓLARES): **129,7952**

- **Moda**

$$Mo = L_i + \frac{d_1}{d_1 + d_2} \times c$$

Utilizando o cálculo da moda, verificou-se o seguinte resultado:

MODA DE GASTOS COM LANÇAMENTOS (EM MILHÕES DE DÓLARES): **450**

Segundo o resultado do cálculo, verifica-se um conjunto de dados UNIMODAL.



EACH

Escola de Artes, Ciências e Humanidades
da Universidade de São Paulo

- **Mediana**

$$E \text{ (elemento central)} = (n+1) / 2$$

Utilizando o cálculo da mediana, verificou-se o seguinte resultado:

MEDIANA DE GASTOS COM LANÇAMENTOS (EM MILHÕES DE
DÓLARES): **62**

REFERÊNCIAS BIBLIOGRÁFICAS

TAVARES, Marcelo. **Estatística aplicada à administração**. 3. ed. rev. ampl. Florianópolis: Departamento de Ciências da Administração/UFSC; [Brasília]: CAPES: UAB, 2014.