**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**

TRATAMENTO E ANÁLISE DE DADOS E INFORMAÇÕES

SÃO PAULO

2023

**ANÁLISE DE GASTOS COM EXPLORAÇÃO ESPACIAL**

Sumário

[DEFINIÇÃO DO PROBLEMA 3](#_Toc134135302)

[PLANEJAMENTO 4](#_Toc134135303)

[DADOS 5](#_Toc134135304)

[POPULAÇÃO E AMOSTRAS 6](#_Toc134135305)

[DISTRIBUIÇŌES DE FREQUÊNCIAS 7](#_Toc134135306)

[REPRESENTAÇÃO GRÁFICA 8](#_Toc134135307)

[MEDIDAS DE POSIÇÃO 9](#_Toc134135308)

# **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 que alcançassem 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.

# **PLANEJAMENTO**

O planejamento desta análise foi pautado nos seguintes tópicos:

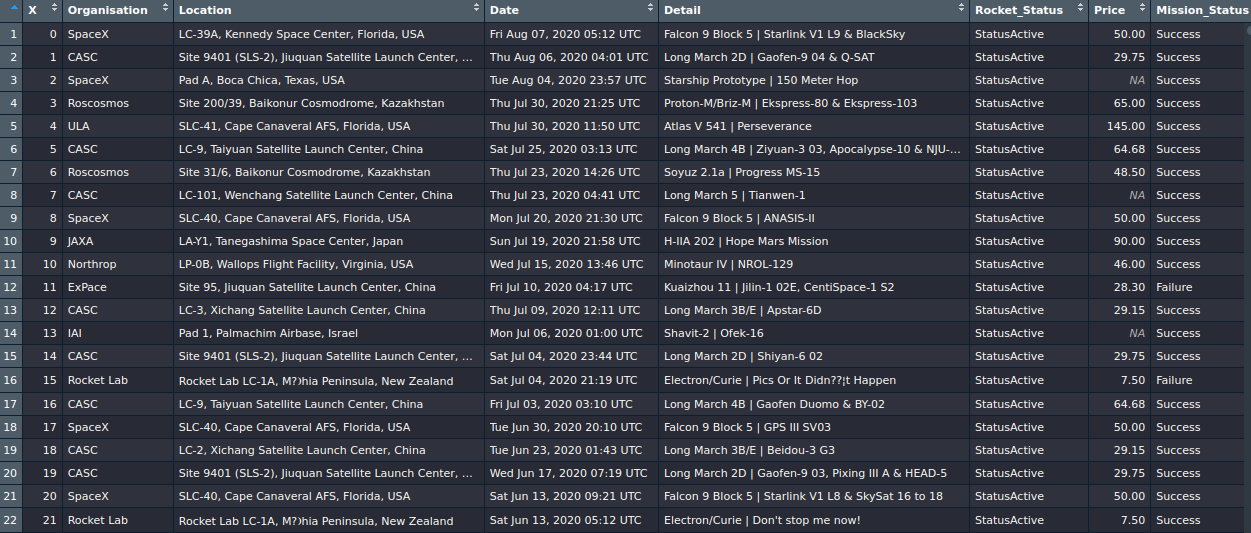
* Definir as motivações para a exploração espacial.
* Compreender os processos de exploração e pesquisa.
* Coletar, por meio de banco eletrônico, dados brutos.
* Analisar os dados coletados, contemplando a magnitude dos investimentos na área, bem como seus resultados, taxas de sucesso e retorno em forma 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.

# **DADOS**

* **Coleta**

Para a coleta dos dados, o grupo optou pela pesquisa bibliográfica, analisando a base de dados brutos hospedada no portal “[Kaggle](https://www.kaggle.com/)”, denominada como “[Space Mission Launches: All the space missions since 1947](https://www.kaggle.com/datasets/salmane/space-missions-launches)”, além do website “[Next Spaceflight](https://nextspaceflight.com/)”, 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:



* **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 da obra “Estatística Aplicada à Administração”, de Marcelo Tavares, 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.

# **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 o número de lançamentos de foguetes por país, a média de lançamentos por ano, as tendências de gastos com as operações em cada nação 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 em nível de gastos, posição dos países em 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, múmero de lançamentos de foguetes por organização, número de lançamento de foguetes por país, número de missões bem sucedidas e fracassadas e número de missões ativas e desativadas.

-Variáveis Quantitativas Contínuas: Gastos com lançamentos por país e 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.

# **DISTRIBUIÇŌES DE FREQUÊNCIAS**

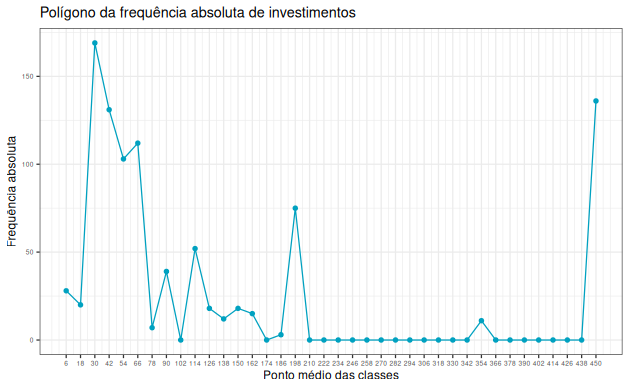
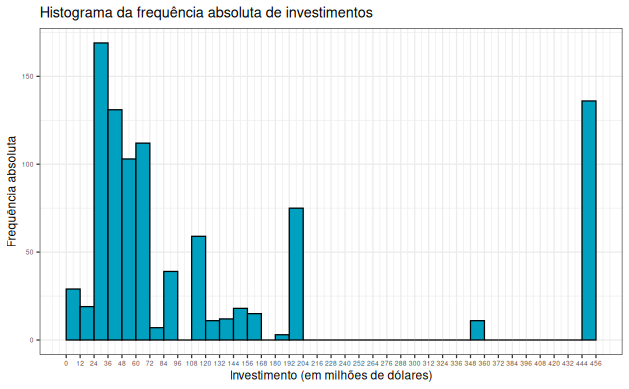
Para distribuição de frequências, foi determinado uma analíse no total investido pelas organizações pelas tentativas de lançamento, valores dados em milhões de dólares

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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:

Como n >100, então

# **REPRESENTAÇÃO GRÁFICA**



# **MEDIDAS DE POSIÇÃO**