



# PROJETO FINAL

Tiago Silva



## Descrição dos Dados Utilizados

Fonte: 'Air\_Traffic\_Passenger\_Statistics.csv'.

Principais variáveis: Contagem de passageiros, companhia aérea, terminal, região geográfica.

Processo de limpeza dos dados: Remoção de duplicados e inconsistências.



## Média de Passageiros por Voo

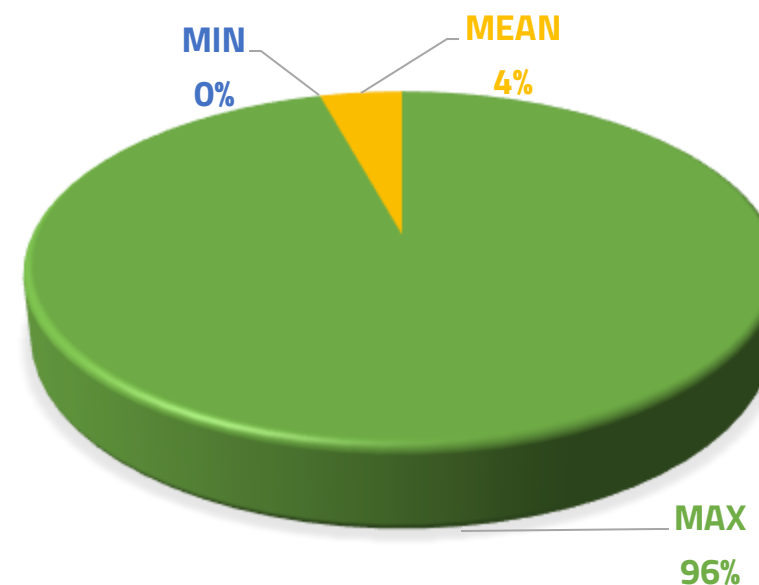
**Valor mínimo:** 1 passageiro.

**Valor máximo:** 659.837 passageiros.

**Média:** 29.240 passageiros por registro.

Há 77 companhias aéreas distintas

## MINIMUM, MAXIMUM AND MEAN PASSENGER PER FLIGHT

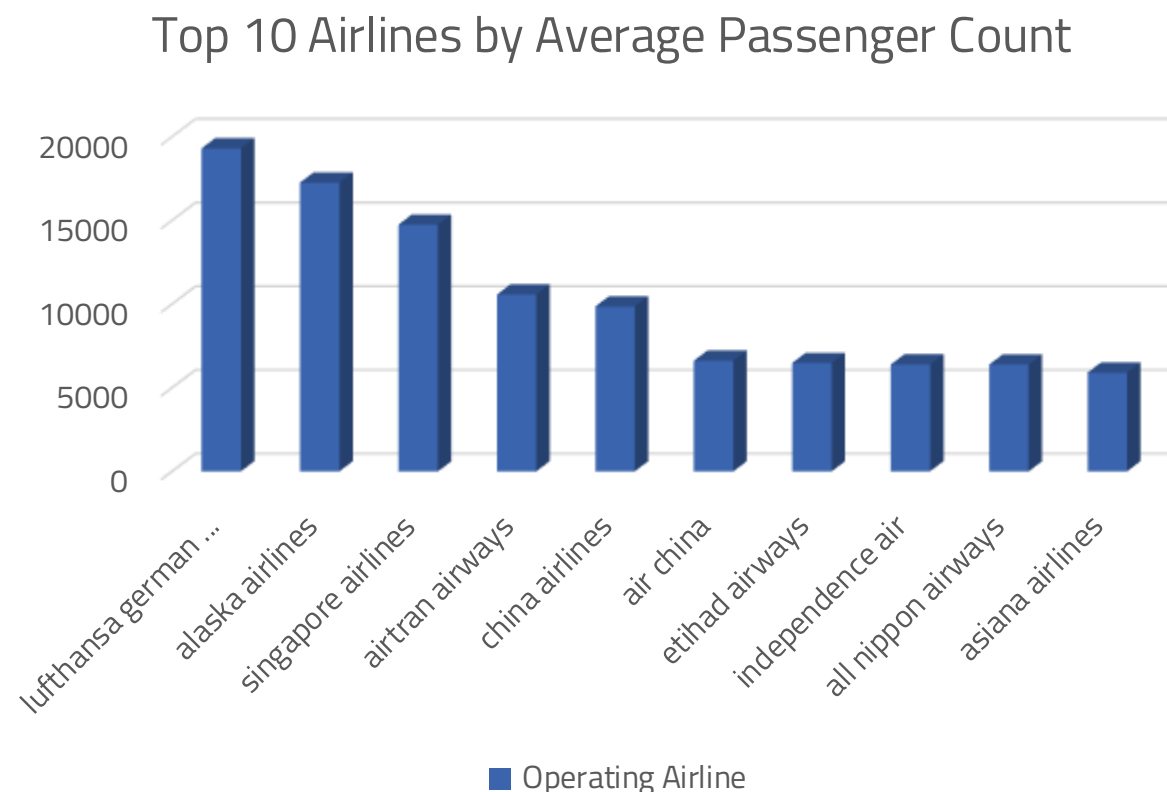




## Companhias Aéreas em Destaque

**Destaque:** Air China (259 registros), Air Berlin (12 registros).

**Comparação de médias:** Alaska Airlines e Singapore Airlines (>14.000 passageiros) vs Belair Airlines (<500 passageiros).





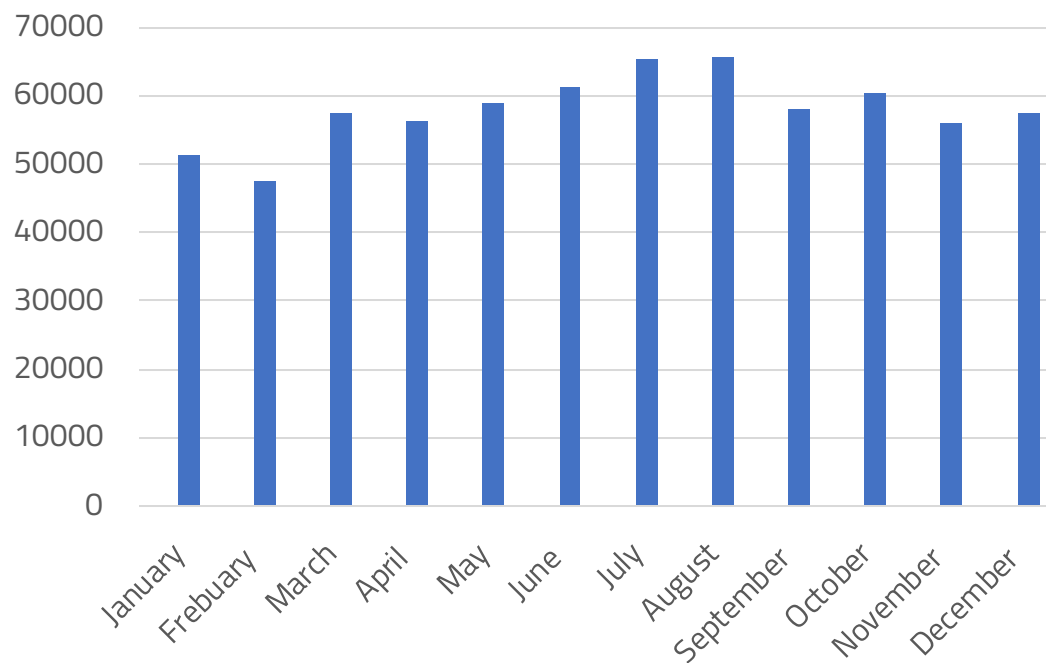
## Média de Passageiros por Ano

Este gráfico ajuda a identificar tendências no aumento ou diminuição do tráfego aéreo, refletindo mudanças na procura, na expansão de rotas, ou até eventos externos que possam ter afetado o transporte aéreo.





## Média de Passageiros por Mês

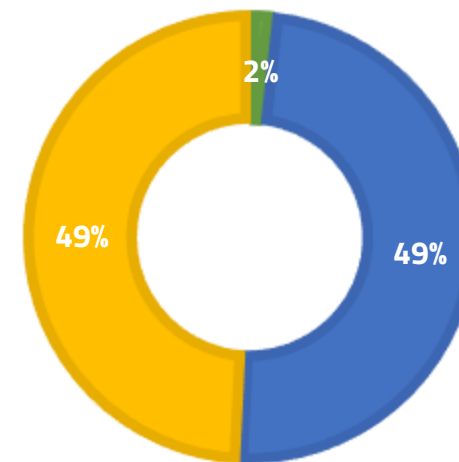


Os meses com maior movimentação podem corresponder a períodos de férias ou eventos especiais, enquanto os meses de menor movimento podem refletir a baixa temporada.



## Média de Passageiros por Tipo de Atividade

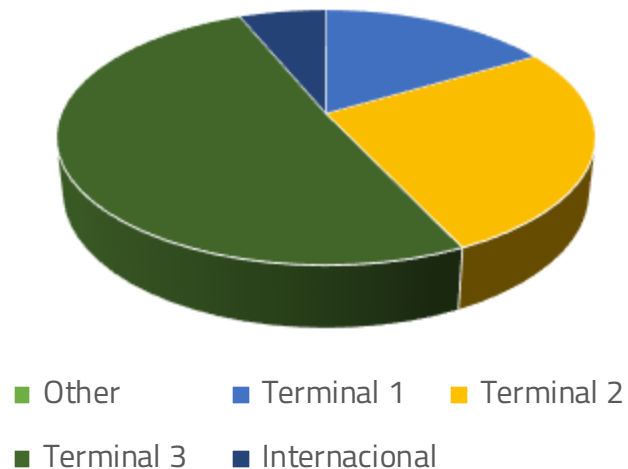
Aqui, observa-se como cada uma dessas operações se distribui no aeroporto, com maior número de passageiros associados aos embarques e desembarques, em relação ao trânsito.



■ Thru / Transit ■ Deplaned ■ Enplaned



## Média de Passageiros por Terminal



Verifica-se que alguns terminais são significativamente mais movimentados do que outros, sugerindo que podem estar associados a operações de maior escala, como voos internacionais ou domésticos de alta procura.





## Principais Correlações Entre Variáveis

Passenger Count vs Adjusted Passenger Count: Correlação quase perfeita (0.999941).

Terminal vs Passenger Count: Correlação moderada (0.433).

GEO Summary vs Passenger Count: Correlação de 0.395.

	Passenger Count	Adjusted Passenger Count	Year	Operating Airline_Indexed	Published Airline_Indexed	GEO Summary_Indexed	GEO Region_Indexed	Type Code_Indexed	Price Category Code_Indexed	Terminal_Indexed	Boarding Area_Indexed	Adjusted Activity Type Code_Indexed	Month_Indexed
Passenger Count	1	0,999941	0,060069	-0,285452	-0,273216	0,395743	-0,299917	-0,071423	0,065047	0,433388	0,344823	-0,071423	-0,010581
Adjusted Passenger Count	0,999941	1	0,059096	-0,286992	-0,274674	0,396856	-0,300785	-0,067804	0,064661	0,434731	0,345585	-0,067804	-0,010639
Year	0,060069	0,059096	1	0,009516	0,038925	-0,066046	0,071299	-0,052364	0,005683	-0,049672	-0,064264	-0,052364	0,080207
Operating Airline_Indexed	-0,285452	-0,286992	0,009516	1	0,867346	-0,069229	0,063805	-0,125313	0,063006	-0,263487	-0,25222	-0,125313	-0,022386
Published Airline_Indexed	-0,273216	-0,274674	0,038925	0,867346	1	-0,212183	0,160969	-0,116691	0,127752	-0,385844	-0,397987	-0,116691	-0,022964
GEO Summary_Indexed	0,395743	0,396856	-0,066046	-0,069229	-0,212183	1	-0,691626	0,02676	0,411498	0,559994	0,583929	0,02676	-0,001066
GEO Region_Indexed	-0,299917	-0,300785	0,071299	0,063805	0,160969	-0,691626	1	0,002738	-0,260386	-0,337081	-0,358086	0,002738	0,003383
Activity Type Code_Indexed	-0,071423	-0,067804	-0,052364	-0,125313	-0,116691	0,02676	0,002738	1	-0,001004	0,080175	0,072956	1	-0,000626
Price Category Code_Indexed	0,065047	0,064661	0,005683	0,063006	0,127752	0,411498	-0,260386	-0,001004	1	0,13971	0,102428	-0,001004	-0,005466
Terminal_Indexed	0,433388	0,434731	-0,049672	-0,263487	-0,385844	0,559994	-0,337081	0,080175	0,13971	1	0,892224	0,080175	0,001095
Boarding Area_Indexed	0,344823	0,345585	-0,064264	-0,25222	-0,397987	0,583929	-0,358086	0,072956	0,102428	0,892224	1	0,072956	0,002206
Adjusted Activity Type Code_Indexed	-0,071423	-0,067804	-0,052364	-0,125313	-0,116691	0,02676	0,002738	1	-0,001004	0,080175	0,072956	1	-0,000626
Month_Indexed	-0,010581	-0,010639	0,080207	-0,022386	-0,022964	-0,001066	0,003383	-0,000626	-0,005466	0,001095	0,002206	-0,000626	1

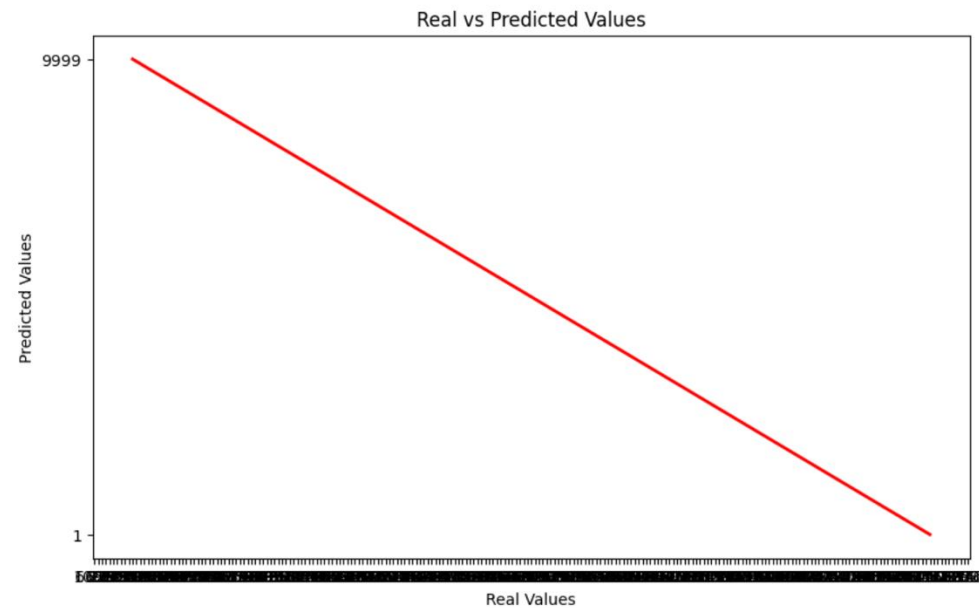


## Resultados da Regressão Linear Múltipla

**R<sup>2</sup> elevado:** Modelo explica grande parte da variação no número de passageiros.

**MSE** alto, mas com boa performance.

**Sugestões de melhorias:** Remoção de colinearidade e normalização.





## Conclusão

A análise revelou que a **infraestrutura** e a **localização geográfica** são fatores chave no volume de passageiros. Companhias como a **Air China** e a **Alaska Airlines** destacaram-se em termos de movimentação. O modelo de **regressão linear** apresentou um bom desempenho, embora melhorias como a normalização dos dados possam aumentar a sua precisão. Estes insights ajudarão a **Tokio School Viagens** a otimizar as suas operações e estratégias de marketing, focando-se nas regiões e nos períodos de maior tráfego.



## Próximos Passos

Deixo aqui próximos passos que podemos fazer:

- **Explorar padrões de crescimento** no tráfego aéreo, identificando tendências ao longo do tempo para antecipar a procura.
- **Analisar hubs de aviação** e o desempenho das companhias aéreas para otimizar as operações e estratégias de marketing.
- **Ajustar o modelo preditivo**, removendo a colinearidade e normalizando os dados, para melhorar a precisão das previsões de passageiros.