

# House of Excellence

**Consultores Responsáveis:**

Estatiano 1

Estatiano 2

...

Estatiano n

**Requerente:**

ESTAT

Brasília, 13 de outubro de 2024.



## Sumário

	Página
1 Não sei se todas as questões de fontes e cores estão corretas, gostaria de saber se tem algo que estou fazendo errado e que o pacote da estat não cobre e preciso fazer manualmente. . . . .	3
2 Introdução . . . . .	4
3 Referencial Teórico . . . . .	5
4 Análises . . . . .	6
4.1 Top 5 países com maior número de mulheres medalhistas . . . .	6
4.2 2. Valor IMC por esporte, estes sendo, ginástica, futebol, judô, atletismo e badminton (20/10) . . . . .	6
4.3 3. Top 3 medalhistas gerais por quantidade de cada tipo de medalha (27/10) . . . . .	7
5 Conclusões . . . . .	8

- 1 Não sei se todas as questões de fontes e cores estão corretas, gostaria de saber se tem algo que estou fazendo errado e que o pacote da estat não cobre e preciso fazer manualmente.**

## 2 Introdução

O seguinte projeto tem como objetivo trazer uma análise descritiva e exploratória de dados fornecidos pela “House of Excellence”, visando otimizar o desempenho de seus atletas.

O banco de dados conta com 9 variáveis e 67.474 observações, sendo as mais importantes (tirando as informações básicas de cada atleta, como nome, peso e altura); “Sport”: Esporte praticado pelo atleta, “Event”: Nome do evento esportivo que o atleta participou, “Medal”: medalha adquirida pelo atleta na competição.

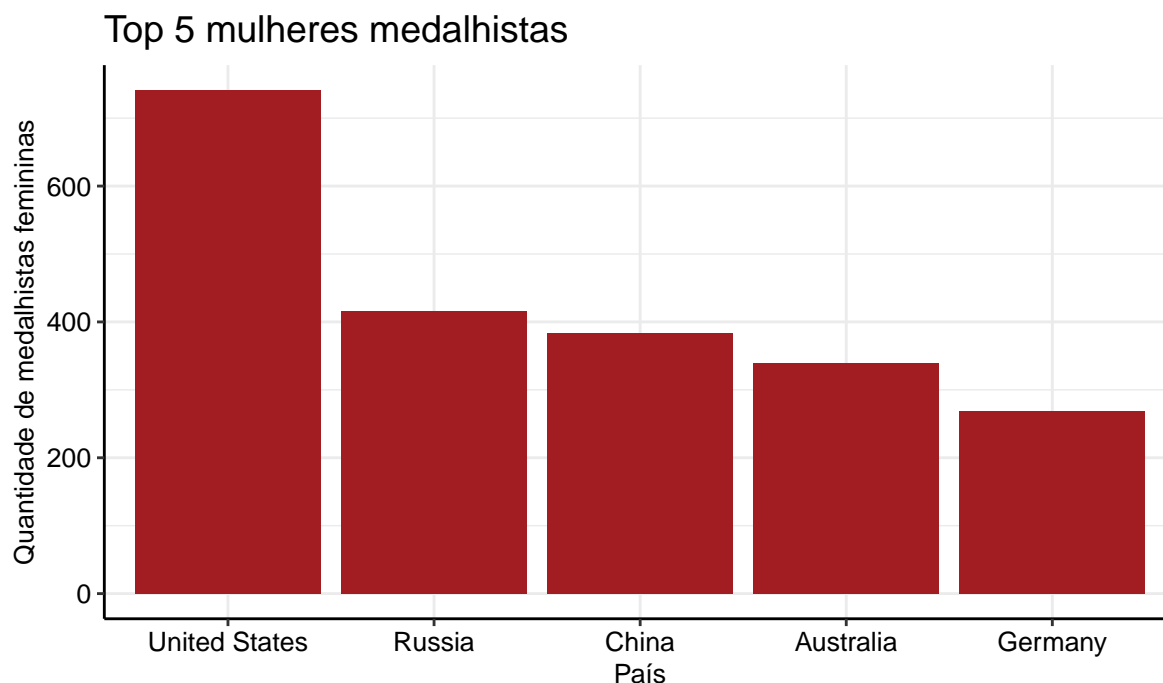
Todas as análises foram feitas utilizando o software RStudio

### 3 Referencial Teórico

?

## 4 Análises

### 4.1 Top 5 países com maior número de mulheres medalhistas



É possível observar, pela (adicionar o endereço da figura) que os Estados Unidos estão na frente com uma vantagem esmagadora em comparação à equipe mais próxima (Rússia), tendo quase o dobro de medalhas conquistadas pelas mulheres de sua equipe. Em contraste com a primeira equipe colocada, as outras 4 equipes que estão no topo, tem número de medalhas muito próximas entre si.

### 4.2 2. Valor IMC por esporte, estes sendo, ginástica, futebol, judô, atletismo e badminton (20/10)

Quantidade de “na’s” nas variáveis de altura e peso respectivamente

```
[1] 6
```

```
[1] 13
```

Como a quantidade de “na’s” é muito pequena, os “na’s” serão considerados um evento aleatório e as linhas serão removidas.

```
{r}
```

```
tabela1 <- imc_calculado %>% group_by(Sport) %>% summarise(media = round(mean(imc), 2), variancia = round(var(imc), 2), mediana = round(median(imc), 2))
```

```
tabela1 %>% ggplot(aes(x = "Sport", y = "media")) + geom_col() + theme_estat()
{r}
```

```
imc_calculado %>% ggplot(aes(x = reorder(Sport, imc, FUN = median, decreasing  
= T), y = imc, fill = Sport[1])) + geom_boxplot() + theme_estat() + labs(y = "IMC", x =  
"Esporte", fill = "Esporte")
```

### 4.3 3. Top 3 medalhistas gerais por quantidade de cada tipo de medalha (27/10)

```
{r} df <- df %>% mutate(Medal_value = case_when( Medal == "Gold" ~ 3, Medal ==  
"Silver" ~ 2, Medal == "Bronze" ~ 1 ))  
{r}  
top_medalhistas <- df %>% group_by(Nomes) %>% summarise(n = n(), ) %>% ar-  
range(desc(n)) %>% head()
```

```
{r}
```

```
top_nomes <- top_medalhistas %>%  
  filter(n == 4) %>%  
  select(Nomes) %>%  
  as_vector()
```

```
terceiro <- df %>%  
  filter(Nomes %in% top_nomes) %>%  
  group_by(Nomes) %>%  
  summarise(Medal_value = sum(Medal_value)) %>%  
  arrange(desc(Medal_value)) %>%  
  head(1) %>%  
  select(Nomes)
```

## 5 Conclusões

?