



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUÁRIAS



GRUPO:

Vinicius Gottgtroy Bacco

Arthur Teixeira Lima de Santana

Cauã Wylker Alves Santos

José Ramon Santos Nascimento

RELATÓRIO DE TRABALHO FINAL
ANÁLISE EXPLORATÓRIA DE DADOS



1-INTRODUÇÃO

Nessa análise, utilizamos como base um dataset da NBB (Novo Basquete Brasil) com um total de 77 variáveis, que representam as principais estatísticas contabilizadas em um jogo de basquete. Nosso objetivo principal é identificar as variáveis que apresentam maior correlação, seja negativa ou positiva, com a colocação dos times ao final da temporada.

2-FERRAMENTAS UTILIZADAS

O universo de estudo são os jogadores e times que participaram da NBB nas temporadas de 2019-2020, reduzimos as variáveis a apenas 12 que tem mais importância para a análise em questão, são elas:

- 1- Jogador - Nome do Jogador (Variável Qualitativa Nominal)
- 2- Equipe - Nome da Equipe (Variável Qualitativa Nominal)
- 3- Temporada - Ano da temporada (Variável Qualitativa Nominal)
- 4- PTS_TOTAL - Pontos Totais (Variável Quantitativa discreta)
- 5- 3PT_TOTAL - 3 Pontos Totais (Variável Quantitativa discreta)
- 6- 2PT_TOTAL - 2 Pontos Totais (Variável Quantitativa discreta)
- 7- AS_TOTAL - Assistências Totais (Variável Quantitativa discreta)
- 8- RT_TOTAL - Rebotes Totais (Variável Quantitativa discreta)
- 9- EN_TOTAL - Enterradas Totais (Variável Quantitativa discreta)
- 10- TO_TOTAL - Tocos totais (Variável Quantitativa discreta)
- 11- DD_TOTAL - Duplos Duplos Totais, Vezes em que o jogador fez uma mais de 2 dígitos em duas das três seguintes variáveis: pontos, assistências ou rebotes (Variável Quantitativa discreta)
- 12- TD_TOTAL - Triplos Duplos Totais, Vezes em que o jogador fez uma mais de 2 dígitos em pontos, assistências e rebotes (Variável Quantitativa discreta)

Utilizamos bibliotecas do Python para análise exploratória dos dados e para criação de gráficos (Pandas/Matplotlib/Seaborn)



3-ANÁLISE DESCRITIVA DOS DADOS

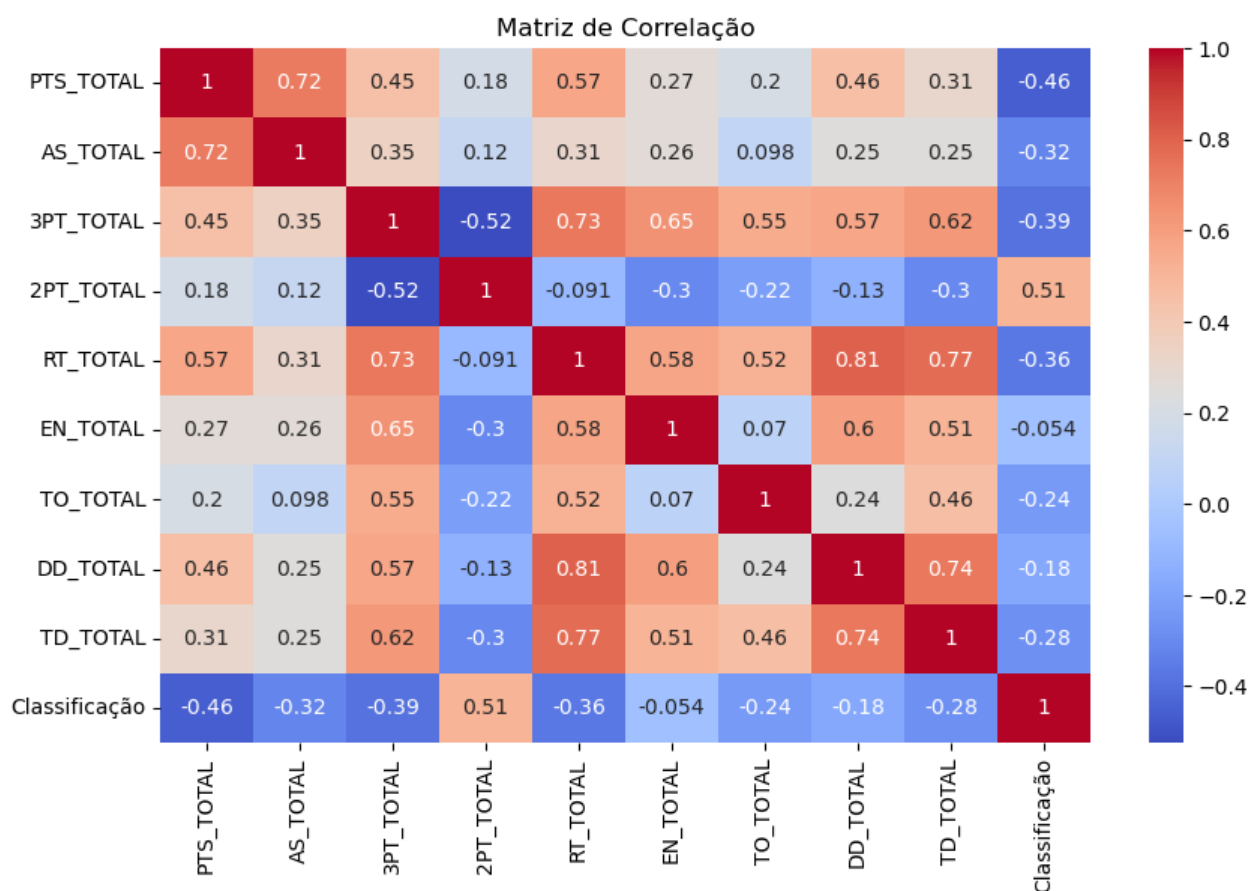
O conjunto de dados analisados é composto de 16 times e 232 jogadores.

Tabela 1 - Tabela de classificação/Variáveis NBB Temporada 2019-20

Equipes	Pontos	Assistências	Cesta de 3 Pontos	Cesta de 2 Pontos	Rebotes	Enterradas	Tocos	Duplos Duplos	Triplos Duplos	Classificação
Flamengo	2088	473	701	881	826	27	45	6	0	1
Sesi Franca	2188	482	686	847	802	40	61	3	0	2
São Paulo	2302	503	841	884	991	67	74	22	5	3
123 Minas	2204	462	684	998	869	51	49	12	0	4
Mogi	2120	432	681	920	824	25	60	7	0	5
Pinheiros	2202	451	644	1030	854	37	44	5	0	6
Corinthians	2141	464	636	1034	832	48	53	3	0	7
Botafogo	2038	447	748	848	791	31	60	3	0	8
Rio Claro	2195	486	632	923	781	56	28	12	0	9
UNIFACISA	2246	496	778	991	906	56	71	14	0	10
Paulistano	2042	450	826	838	873	89	45	10	1	11
Bauru	2084	397	631	1020	827	34	37	10	0	12
BRB/Brasília	1975	421	594	1000	830	33	47	8	0	13
São José	2360	564	672	1159	825	35	48	6	0	14
Fortaleza B.C.	1985	347	653	993	808	35	60	6	0	15
PatoBasquete	1774	430	557	996	751	32	48	3	0	16

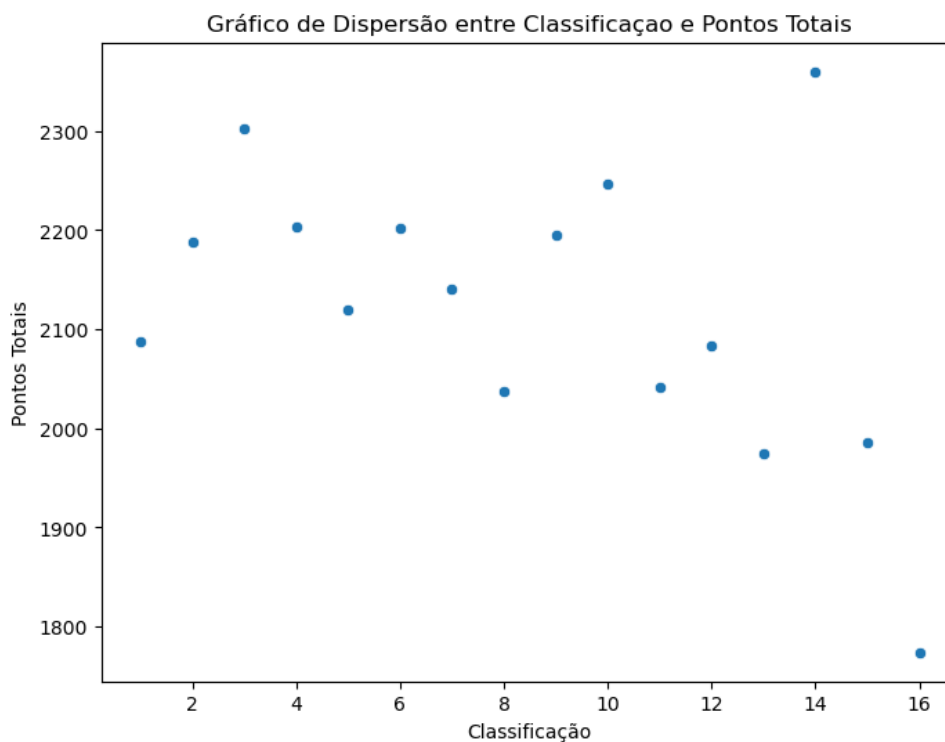
Fonte: NBB-Novo Basquete Brasil (2019-20)

O HeatMap a seguir mostra a correlação entre as variáveis:



Observando essa matriz de correlação, podemos perceber que ela mostra uma correlação negativa entre Pontos Totais e Classificação (-0,46), o que pode não parecer intuitivo. Vamos investigar isso. Outro dado curioso é que os rebotes têm uma forte correlação (0,73) com cestas de três pontos, o que é esperado, já que, no momento do rebote, é comum que o jogador que pegou a bola passe para alguém livre fora da área, buscando a cesta de três.

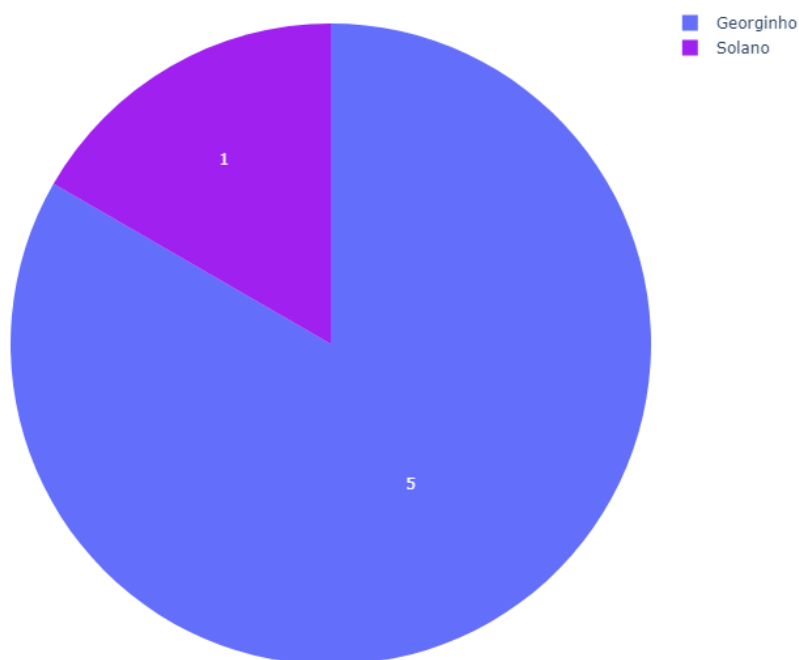
GRÁFICO 1 – CLASSIFICAÇÃO X PONTOS TOTAIS



Como podemos observar pelo gráfico 1 uma alta pontuação não é determinante para uma boa colocação no campeonato o time que fez a maior quantidade de pontos terminou em 14º, adotar uma postura mais defensiva faz sentido na NBB.

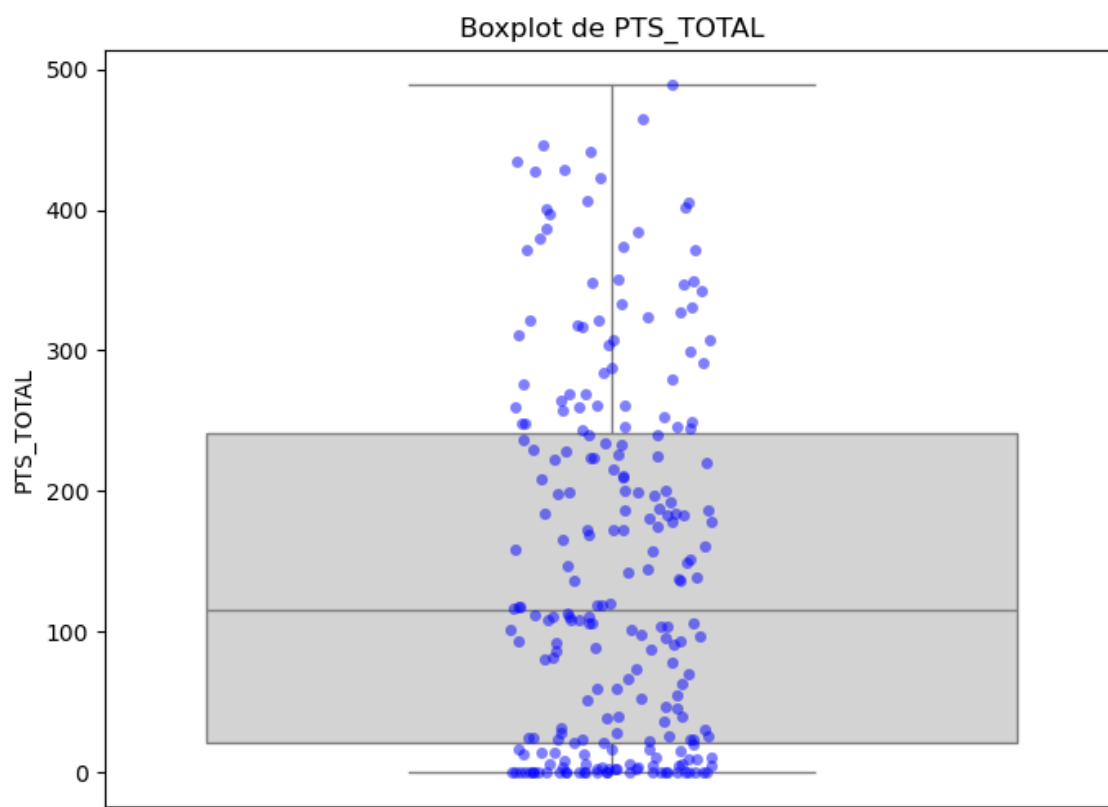
GRÁFICO 2 – DISTRIBUIÇÃO TRIPLOS DUPLOS

Distribuição de Triplos Duplos



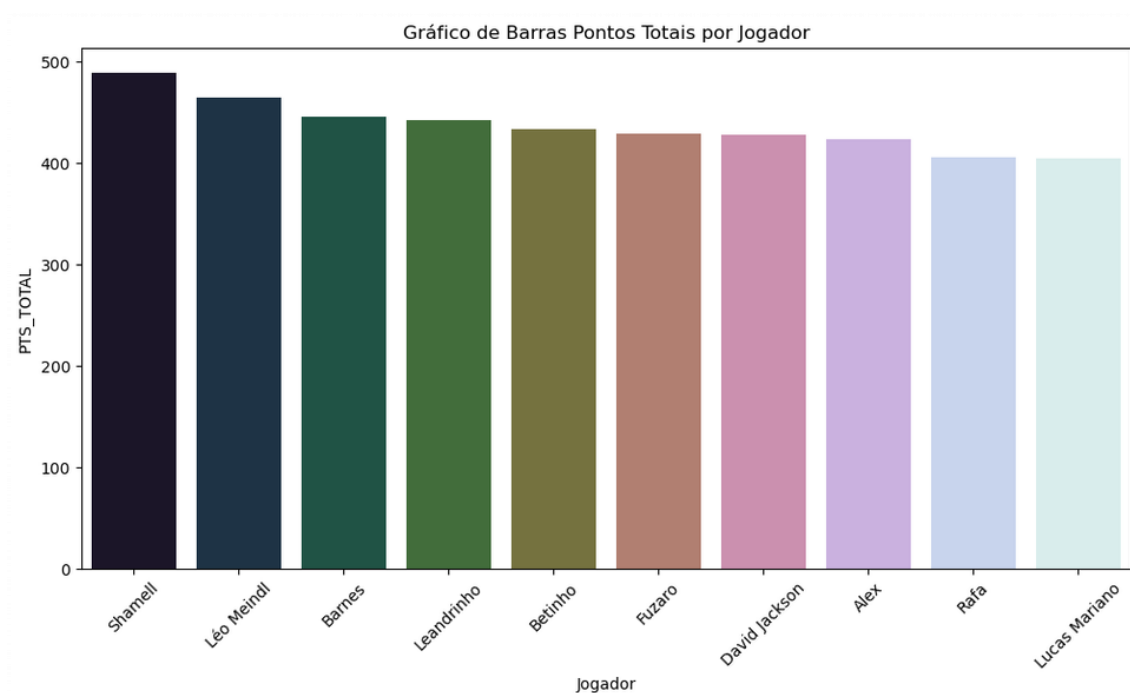
Apenas Dois jogadores fizeram triplos duplos na temporada, de acordo com o gráfico 2, são eles, Solano que joga no Paulistano, e Georginho que garantiu a terceira colocação jogando pelo São Paulo.

GRÁFICO 3 – BOXPLOT PONTOS TOTAIS



Observando o gráfico 3 vemos que metade dos jogadores fizeram 100 ou menos pontos ao longo da temporada.

GRÁFICO 4 – PONTOS TOTAIS POR JOGADOR



O gráfico 4 mostra os 10 jogadores que mais fizeram pontos na temporada, Um dado curioso é que os dois jogadores com maior pontuação são ambos do São Paulo.



4- CONCLUSÃO

De acordo com a análise realizada, podemos concluir que, na liga brasileira de basquete, uma boa defesa é muito mais significativa do que um bom ataque. Times que terminam em boas colocações costumam ganhar as partidas, mas com uma taxa de pontos mais baixa. Manter um ritmo de jogo mais lento, preservar o time fisicamente e focar em cestas de dois pontos parece ser a melhor estratégia a adotar na NBB.