

IMPORTANDO BIBLIOTECAS

```
In [22]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import ttest_ind
```

IMPORTANDO BANCO DE DADOS

```
In [15]: control_group = pd.read_csv('control_group.csv', sep=';')
test_group = pd.read_csv('test_group.csv', sep=';')
control_group = control_group.dropna()
test_group = test_group.dropna()
```

```
In [16]: control_group
```

Out[16]:

	Campaign Name	Date	Spend [USD]	# of Impressions	Reach	# of Website Clicks	# of Searches	# of View Content	# of Add to Cart	# of Purchase
0	Control Campaign	1.08.2019	2280	82702.0	56930.0	7016.0	2290.0	2159.0	1819.0	618.0
1	Control Campaign	2.08.2019	1757	121040.0	102513.0	8110.0	2033.0	1841.0	1219.0	511.0
2	Control Campaign	3.08.2019	2343	131711.0	110862.0	6508.0	1737.0	1549.0	1134.0	372.0
3	Control Campaign	4.08.2019	1940	72878.0	61235.0	3065.0	1042.0	982.0	1183.0	340.0
5	Control Campaign	6.08.2019	3083	109076.0	87998.0	4028.0	1709.0	1249.0	784.0	764.0
6	Control Campaign	7.08.2019	2544	142123.0	127852.0	2640.0	1388.0	1106.0	1166.0	499.0
7	Control Campaign	8.08.2019	1900	90939.0	65217.0	7260.0	3047.0	2746.0	930.0	462.0
8	Control Campaign	9.08.2019	2813	121332.0	94896.0	6198.0	2487.0	2179.0	645.0	501.0
9	Control Campaign	10.08.2019	2149	117624.0	91257.0	2277.0	2475.0	1984.0	1629.0	734.0
10	Control Campaign	11.08.2019	2490	115247.0	95843.0	8137.0	2941.0	2486.0	1887.0	475.0
11	Control Campaign	12.08.2019	2319	116639.0	100189.0	2993.0	1397.0	1147.0	1439.0	794.0
12	Control Campaign	13.08.2019	2697	82847.0	68214.0	6554.0	2390.0	1975.0	1794.0	766.0
13	Control Campaign	14.08.2019	1875	145248.0	118632.0	4521.0	1209.0	1149.0	1339.0	788.0
14	Control Campaign	15.08.2019	2774	132845.0	102479.0	4896.0	1179.0	1005.0	1641.0	366.0
15	Control Campaign	16.08.2019	2024	71274.0	42859.0	5224.0	2427.0	2158.0	1613.0	438.0
16	Control Campaign	17.08.2019	2177	119612.0	106518.0	6628.0	1756.0	1642.0	878.0	222.0
17	Control Campaign	18.08.2019	1876	108452.0	96518.0	7253.0	2447.0	2115.0	1695.0	243.0
18	Control Campaign	19.08.2019	2596	107890.0	81268.0	3706.0	2483.0	2098.0	908.0	542.0
19	Control Campaign	20.08.2019	2675	113430.0	78625.0	2578.0	1001.0	848.0	1709.0	299.0
20	Control Campaign	21.08.2019	1803	74654.0	59873.0	5691.0	2711.0	2496.0	1460.0	800.0
21	Control Campaign	22.08.2019	2939	105705.0	86218.0	6843.0	3102.0	2988.0	819.0	387.0
22	Control Campaign	23.08.2019	2496	129880.0	109413.0	4410.0	2896.0	2496.0	1913.0	766.0
23	Control Campaign	24.08.2019	1892	72515.0	51987.0	4085.0	1274.0	1149.0	1146.0	585.0
24	Control Campaign	25.08.2019	1962	117006.0	100398.0	4234.0	2423.0	2096.0	883.0	386.0
25	Control Campaign	26.08.2019	2233	124897.0	98432.0	5435.0	2847.0	2421.0	1448.0	251.0
26	Control Campaign	27.08.2019	2061	104678.0	91579.0	4941.0	3549.0	3249.0	980.0	605.0
27	Control Campaign	28.08.2019	2421	141654.0	125874.0	6287.0	1672.0	1589.0	1711.0	643.0
28	Control Campaign	29.08.2019	2375	92029.0	74192.0	8127.0	4891.0	4219.0	1486.0	334.0
29	Control Campaign	30.08.2019	2324	111306.0	88632.0	4658.0	1615.0	1249.0	442.0	670.0

```
In [17]: test_group
```

Out[17]:

	Campaign Name	Date	Spend [USD]	# of Impressions	Reach	# of Website Clicks	# of Searches	# of View Content	# of Add to Cart	# of Purchase
0	Test Campaign	1.08.2019	3008	39550	35820	3038	1946	1069	894	255
1	Test Campaign	2.08.2019	2542	100719	91236	4657	2359	1548	879	677
2	Test Campaign	3.08.2019	2365	70263	45198	7885	2572	2367	1268	578
3	Test Campaign	4.08.2019	2710	78451	25937	4216	2216	1437	566	340
4	Test Campaign	5.08.2019	2297	114295	95138	5863	2106	858	956	768
5	Test Campaign	6.08.2019	2458	42684	31489	7488	1854	1073	882	488
6	Test Campaign	7.08.2019	2838	53986	42148	4221	2733	2182	1301	890
7	Test Campaign	8.08.2019	2916	33669	20149	7184	2867	2194	1240	431
8	Test Campaign	9.08.2019	2652	45511	31598	8259	2899	2761	1200	845
9	Test Campaign	10.08.2019	2790	95054	79632	8125	2312	1804	424	275
10	Test Campaign	11.08.2019	2420	83633	71286	3750	2893	2617	1075	668
11	Test Campaign	12.08.2019	2831	124591	10598	8264	2081	1992	1382	709
12	Test Campaign	13.08.2019	1972	65827	49531	7568	2213	2058	1391	812
13	Test Campaign	14.08.2019	2537	56304	25982	3993	1979	1059	779	340
14	Test Campaign	15.08.2019	2516	94338	76219	4993	2537	1609	1090	398
15	Test Campaign	16.08.2019	3076	106584	81389	6800	2661	2594	1059	487
16	Test Campaign	17.08.2019	1968	95843	54389	7910	1995	1576	383	238
17	Test Campaign	18.08.2019	1979	53632	43241	6909	2824	2522	461	257
18	Test Campaign	19.08.2019	2626	22521	10698	7617	2924	2801	788	512
19	Test Campaign	20.08.2019	2712	39470	31893	6050	2061	1894	1047	730
20	Test Campaign	21.08.2019	3112	133771	109834	5471	1995	1868	278	245
21	Test Campaign	22.08.2019	2899	34752	27932	4431	1983	1131	367	276
22	Test Campaign	23.08.2019	2407	60286	49329	5077	2592	2004	632	473
23	Test Campaign	24.08.2019	2078	36650	30489	7156	2687	2427	327	269
24	Test Campaign	25.08.2019	2928	120576	105978	3596	2937	2551	1228	651
25	Test Campaign	26.08.2019	2311	80841	61589	3820	2037	1046	346	284
26	Test Campaign	27.08.2019	2915	111469	92159	6435	2976	2552	992	771
27	Test Campaign	28.08.2019	2247	54627	41267	8144	2432	1281	1009	721
28	Test Campaign	29.08.2019	2805	67444	43219	7651	1920	1240	1168	677
29	Test Campaign	30.08.2019	1977	120203	89380	4399	2978	1625	1034	572

FILTRANDO OS CONJUNTOS DE DADOS DE NÚMEROS DE IMPRESSÕES

```
In [18]: control_impressions = control_group.iloc[:,3].values
        test_impressions = test_group.iloc[:,3].values

In [19]: control_impressions

Out[19]: array([ 82702., 121040., 131711.,  72878., 109076., 142123.,  90939.,
        121332., 117624., 115247., 116639.,  82847., 145248., 132845.,
        71274., 119612., 108452., 107890., 113430.,  74654., 105705.,
        129880.,  72515., 117006., 124897., 104678., 141654.,  92029.,
        111306.])

In [20]: test_impressions

Out[20]: array([ 39550, 100719,  70263,  78451, 114295,  42684,  53986,  33669,
        45511,  95054,  83633, 124591,  65827,  56304,  94338, 106584,
        95843,  53632,  22521,  39470, 133771,  34752,  60286,  36650,
        120576,  80841, 111469,  54627,  67444, 120203])
```

TESTE T E VALOR P

```
In [21]: t, p = ttest_ind(control_impressions, test_impressions)

        print(f"t = {t:.2f}")
        print(f"p = {p:.2f}")

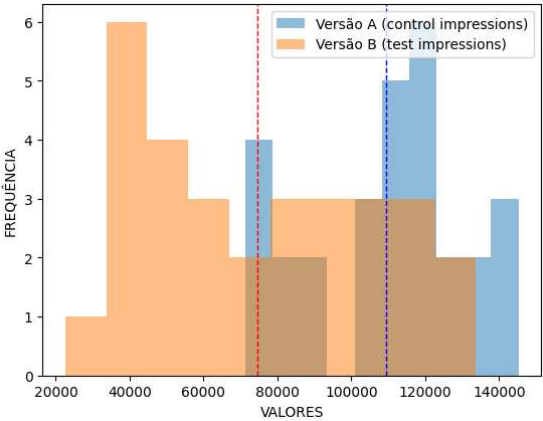
t = 4.88
p = 0.00
```

Por meio dos valores t e p, observa-se que a média da amostra A (Dados de controle) é superior que a média da amostra B (Dados de teste), tendo e vista o valor positivo de t. Ademais, o indicador valor-p de 0.00 expressa que há uma significância estatística entre as diferenças observadas nos dois grupos de dados.

VIZUALIZAÇÃO GRÁFICA

```
In [25]: media_A = np.mean(control_impressions)
        media_B = np.mean(test_impressions)

        plt.hist(control_impressions, alpha=0.5, label='Versão A (control impressions)')
        plt.hist(test_impressions, alpha=0.5, label='Versão B (test impressions)')
        plt.axvline(media_A, color='b', linestyle='dashed', linewidth=1)
        plt.axvline(media_B, color='r', linestyle='dashed', linewidth=1)
        plt.xlabel('VALORES')
        plt.ylabel('FREQUÊNCIA')
        plt.legend(loc='upper right')
        plt.show()
```



No gráfico acima, pode-se observar o histograma de frequência dos dois conjuntos de dados estudados, bem como suas médias representadas pelas linhas tracejadas verticais.