

20201101_magnetis_analise_conversao

November 3, 2020

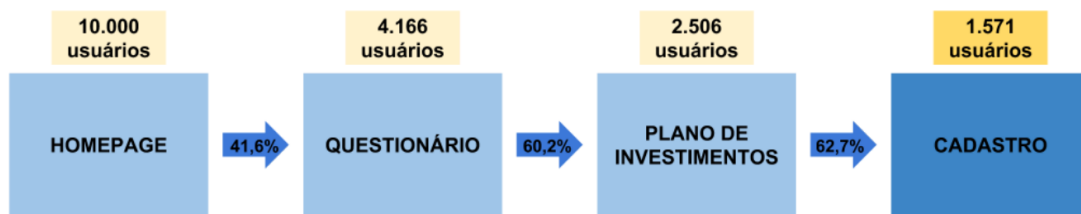
1 Magnetis - Estudo de Conversão

Análise de performance do funil de conversão de novas pessoas investidoras.

Objetivo: Encontrar oportunidades para aumento da taxa de conversão no site.

Etapas: - Conhecer a base de dados para saber quais os atributos disponíveis para estudo e a qualidade dos dados. - Realizar análise exploratória para saber que tipo de perguntas podem ser respondidas. - Formular hipóteses que expliquem o funil de conversão. - Prognosticar ações para aumentar a taxa de cadastro de novas pessoas investidoras.

Perguntas iniciais: - Há alguma diferença significativa entre cada etapa do funil por dispositivo? E por origem? - O valor investido muda significativamente de acordo com a origem ou dispositivo? - Há alguma hora do dia ou dia da semana mais comum para cadastros? - Qual é a faixa de valores com maior número de simulações de investimento? - Qual é o perfil de investidor que mais faz cadastros? E quais os objetivos definidos? E o perfil / objetivo com menor taxa de conversão?



1.1 Preparação do ambiente

Para este estudo, faremos uso das bibliotecas pandas, matplotlib e seaborn.

```
[214]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

1.2 Importação do arquivo de performance do funil

Os nomes das colunas serão alterados para facilitar a seleção de dados: - Id: id - Origem de Tráfego: source - Dispositivo: device - Página: page - Valor Simulado: simulated_investment

```
[147]: conversion_funnel = pd.read_csv('_data/navegacao_usuarios.csv', names=['id', 'source', 'device', 'page', 'simulated_investment'], header=0)
```

1.3 Inspeção dos dados disponíveis para análise

```
[148]: conversion_funnel.head()
```

```
[148]:   id      source  device      page  simulated_investment
0    1  Paid Search  desktop    1 - Homepage             NaN
1    2  Paid Search  desktop    1 - Homepage             NaN
2    3  Paid Search  desktop    4 - Cadastro             2000.0
3    3  Paid Search  desktop    1 - Homepage             2000.0
4    3  Paid Search  desktop  3 - Plano de Investimentos    2000.0
```

Há IDs duplicados. Cada linha representa um estágio do funil. Para fazer análises com dados agregados sem o ruído causado pelas linhas duplicadas, será criado um dataframe sem as duplicações. Este dataframe será usado para estudar as colunas source, device e simulated_investment. O dataframe original será usado para estudar o funil de conversão.

```
[149]: conversion_funnel_unique = conversion_funnel.drop_duplicates(subset=['id']).
      ↪drop('page', axis=1)
```

```
[150]: conversion_funnel_unique.head()
```

```
[150]:   id      source  device  simulated_investment
0    1  Paid Search  desktop             NaN
1    2  Paid Search  desktop             NaN
2    3  Paid Search  desktop             2000.0
6    4  Paid Search  desktop             NaN
7    5  Paid Search  desktop             NaN
```

1.3.1 Análise exploratória do funil de conversão

```
[151]: conversion_funnel['page'].value_counts()
```

```
[151]: 1 - Homepage             10000
      2 - Questionário       4166
      3 - Plano de Investimentos 2506
      4 - Cadastro           1571
      Name: page, dtype: int64
```

```
[152]: conversion_funnel.groupby(['device', 'page']).agg({'id': ['count'],
      ↪'simulated_investment': ['sum', 'mean', 'median']})
```

```
[152]:   device  page      id count  simulated_investment sum      mean \
      count      sum      mean
desktop 1 - Homepage    7379    91612000.0  31384.720795
        2 - Questionário  2919    91612000.0  31384.720795
        3 - Plano de Investimentos 1772    73186000.0  41301.354402
```

	4 - Cadastro	1228	62869000.0	51196.254072
mobile	1 - Homepage	2621	48302000.0	38734.562951
	2 - Questionário	1247	48302000.0	38734.562951
	3 - Plano de Investimentos	734	38260000.0	52125.340599
	4 - Cadastro	343	13161000.0	38370.262391

		median
device	page	
desktop	1 - Homepage	16000.0
	2 - Questionário	16000.0
	3 - Plano de Investimentos	20000.0
	4 - Cadastro	22000.0
mobile	1 - Homepage	9000.0
	2 - Questionário	9000.0
	3 - Plano de Investimentos	10000.0
	4 - Cadastro	20000.0

O valor médio da simulação de investimento aumenta a cada etapa do funil pra quem vem por desktop. No mobile, quem faz o cadastro o faz com valor médio de investimento menor do que quem montou o plano de investimentos, mas a mediana cresce. As medianas são todas distantes das médias ao longo do funil para ambos os tipos de dispositivo, sugerindo a presença de simulações com valores muito distantes da maioria (outliers). A presença de outliers é ainda mais acentuada em dispositivos mobile.

```
[153]: conversion_by_device = pd.DataFrame({
        'desktop': conversion_funnel[conversion_funnel['device'] == 'desktop']
        .value_counts(),
        'mobile': conversion_funnel[conversion_funnel['device'] == 'mobile']
        .value_counts()
    })
conversion_by_device
```

	desktop	mobile
1 - Homepage	7379	2621
2 - Questionário	2919	1247
3 - Plano de Investimentos	1772	734
4 - Cadastro	1228	343

A proporção de pessoas que inicia o questionário no mobile é maior do que no desktop, e após o questionário as proporções caem se comparadas ao desktop.

```
[154]: conversion_funnel.groupby(['source', 'page']).agg({'id': ['count'],
        'simulated_investment': ['sum', 'mean', 'median']})
```

		id	simulated_investment	\
		count	sum	
source	page			

Direct	1 - Homepage	2580	81492000.0
	2 - Questionário	1519	81505000.0
	3 - Plano de Investimentos	1151	74268000.0
	4 - Cadastro	821	48570000.0
Organic Search	1 - Homepage	2547	33967000.0
	2 - Questionário	1229	33319000.0
	3 - Plano de Investimentos	724	20673000.0
	4 - Cadastro	504	16643000.0
Paid Search	1 - Homepage	4445	18973000.0
	2 - Questionário	1176	20029000.0
	3 - Plano de Investimentos	539	13422000.0
	4 - Cadastro	212	9677000.0
Social	1 - Homepage	428	5482000.0
	2 - Questionário	242	5061000.0
	3 - Plano de Investimentos	92	3083000.0
	4 - Cadastro	34	1140000.0

		mean	median
source	page		
Direct	1 - Homepage	58291.845494	22000.0
	2 - Questionário	53657.011192	20000.0
	3 - Plano de Investimentos	64524.761077	22000.0
	4 - Cadastro	59159.561510	50000.0
Organic Search	1 - Homepage	27637.917006	20000.0
	2 - Questionário	27110.659072	19000.0
	3 - Plano de Investimentos	28553.867403	15000.0
	4 - Cadastro	33021.825397	20000.0
Paid Search	1 - Homepage	15081.875994	5000.0
	2 - Questionário	17031.462585	5000.0
	3 - Plano de Investimentos	24901.669759	5000.0
	4 - Cadastro	45646.226415	20000.0
Social	1 - Homepage	19508.896797	12000.0
	2 - Questionário	20913.223140	13000.0
	3 - Plano de Investimentos	33510.869565	23500.0
	4 - Cadastro	33529.411765	20000.0

```
[155]: conversion_by_source = pd.DataFrame({
    'Direct': conversion_funnel[conversion_funnel['source'] == 'Direct']['page'].value_counts(),
    'Organic Search': conversion_funnel[conversion_funnel['source'] == 'Organic Search']['page'].value_counts(),
    'Paid Search': conversion_funnel[conversion_funnel['source'] == 'Paid Search']['page'].value_counts(),
    'Social': conversion_funnel[conversion_funnel['source'] == 'Social']['page'].value_counts()
})
```

```
conversion_by_source
```

```
[155]:
```

	Direct	Organic Search	Paid Search	Social
1 - Homepage	2580	2547	4445	428
2 - Questionário	1519	1229	1176	242
3 - Plano de Investimentos	1151	724	539	92
4 - Cadastro	821	504	212	34

A proporção de pessoas que finaliza o questionário é maior quando a pessoa vem por tráfego direto ou por busca orgânica do que pelas demais origens.

```
[156]: conversion_funnel.groupby(['device', 'source', 'page']).agg({'id': 'count',
    ↳ 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[156]:
```

			id	simulated_investment	\
			count	sum	
desktop	Direct	1 - Homepage	1865	49280000.0	
		2 - Questionário	1109	50047000.0	
		3 - Plano de Investimentos	849	44847000.0	
		4 - Cadastro	682	41668000.0	
	Organic Search	1 - Homepage	1592	22768000.0	
		2 - Questionário	788	22695000.0	
		3 - Plano de Investimentos	491	16236000.0	
		4 - Cadastro	371	13096000.0	
	Paid Search	1 - Homepage	3533	14283000.0	
		2 - Questionário	799	14268000.0	
		3 - Plano de Investimentos	344	9071000.0	
		4 - Cadastro	142	6985000.0	
	Social	1 - Homepage	389	5281000.0	
		2 - Questionário	223	4602000.0	
		3 - Plano de Investimentos	88	3032000.0	
		4 - Cadastro	33	1120000.0	
mobile	Direct	1 - Homepage	715	32212000.0	
		2 - Questionário	410	31458000.0	
		3 - Plano de Investimentos	302	29421000.0	
		4 - Cadastro	139	6902000.0	
	Organic Search	1 - Homepage	955	11199000.0	
		2 - Questionário	441	10624000.0	
		3 - Plano de Investimentos	233	4437000.0	
		4 - Cadastro	133	3547000.0	
	Paid Search	1 - Homepage	912	4690000.0	
		2 - Questionário	377	5761000.0	
		3 - Plano de Investimentos	195	4351000.0	
		4 - Cadastro	70	2692000.0	
	Social	1 - Homepage	39	201000.0	
		2 - Questionário	19	459000.0	

3 - Plano de Investimentos	4	51000.0
4 - Cadastro	1	20000.0

			mean	median
device	source	page		
desktop	Direct	1 - Homepage	48031.189084	22000.0
		2 - Questionário	45128.043282	22000.0
		3 - Plano de Investimentos	52823.321555	22000.0
		4 - Cadastro	61096.774194	50000.0
	Organic Search	1 - Homepage	29189.743590	20000.0
		2 - Questionário	28800.761421	20000.0
		3 - Plano de Investimentos	33067.209776	20000.0
		4 - Cadastro	35299.191375	22000.0
	Paid Search	1 - Homepage	16843.160377	5000.0
		2 - Questionário	17857.321652	5000.0
		3 - Plano de Investimentos	26369.186047	5000.0
		4 - Cadastro	49190.140845	22000.0
	Social	1 - Homepage	19928.301887	12000.0
		2 - Questionário	20636.771300	15000.0
		3 - Plano de Investimentos	34454.545455	25000.0
		4 - Cadastro	33939.393939	20000.0
mobile	Direct	1 - Homepage	86591.397849	22000.0
		2 - Questionário	76726.829268	15000.0
		3 - Plano de Investimentos	97420.529801	12000.0
		4 - Cadastro	49654.676259	25000.0
	Organic Search	1 - Homepage	24942.093541	16000.0
		2 - Questionário	24090.702948	15000.0
		3 - Plano de Investimentos	19042.918455	11000.0
		4 - Cadastro	26669.172932	16000.0
	Paid Search	1 - Homepage	11439.024390	3000.0
		2 - Questionário	15281.167109	3000.0
		3 - Plano de Investimentos	22312.820513	4000.0
		4 - Cadastro	38457.142857	9000.0
	Social	1 - Homepage	12562.500000	12000.0
		2 - Questionário	24157.894737	12000.0
		3 - Plano de Investimentos	12750.000000	11500.0
		4 - Cadastro	20000.000000	20000.0

A média de valores simulados na etapa de cadastro por todas as origens é maior no desktop se comparado ao mobile. A mediana varia ao longo das etapas do funil e sobe para sua máxima na etapa de cadastro para todas as origens e dispositivos.

```
[157]: pd.DataFrame({
        '1 - Homepage': conversion_funnel[conversion_funnel['page'] == '1 -
        ↳Homepage'].groupby('source').agg({'simulated_investment':
        ↳'mean'})['simulated_investment'],
```

```

    '2 - Questionário': conversion_funnel[conversion_funnel['page'] == '2 - Questionário'].groupby('source').agg({'simulated_investment':
    ↳ 'mean'})['simulated_investment'],
    '3 - Plano de Investimentos': conversion_funnel[conversion_funnel['page']
    ↳ == '3 - Plano de Investimentos'].groupby('source').
    ↳ agg({'simulated_investment': 'mean'})['simulated_investment'],
    '4 - Cadastro': conversion_funnel[conversion_funnel['page'] == '4 - Cadastro'].groupby('source').agg({'simulated_investment':
    ↳ 'mean'})['simulated_investment']
})

```

```

[157]:
           1 - Homepage  2 - Questionário  3 - Plano de Investimentos  \
source
Direct                58291.845494        53657.011192        64524.761077
Organic Search        27637.917006        27110.659072        28553.867403
Paid Search           15081.875994        17031.462585        24901.669759
Social                19508.896797        20913.223140        33510.869565

           4 - Cadastro
source
Direct                59159.561510
Organic Search        33021.825397
Paid Search           45646.226415
Social                33529.411765

```

O tráfego que entrou por busca paga e busca orgânica investe mais, em média, do que simulou no questionário.

```

[158]: conversion_funnel.groupby(['page']).agg({'id': ['count']})

```

```

[158]:
           id
           count
page
1 - Homepage      10000
2 - Questionário    4166
3 - Plano de Investimentos  2506
4 - Cadastro       1571

```

```

[159]: conversion_funnel.groupby(['page']).agg({'simulated_investment': ['mean']})

```

```

[159]:
           simulated_investment
           mean
page
1 - Homepage      33584.733557
2 - Questionário    33584.733557
3 - Plano de Investimentos  44471.667997
4 - Cadastro       48395.926162

```

A média geral de simulações de investimento cresce a cada etapa do funil.

```
[160]: '''
        @name set_bucket
        @description classifica cada simulação de investimento em faixas de valores
        @param simulated_investment float
        @return string
        '''
def set_bucket(simulated_investment):
    if np.isnan(simulated_investment):
        return 'no investment'
    elif simulated_investment < 10000:
        return 'below 10.000'
    elif simulated_investment >= 10000 and simulated_investment < 20000:
        return 'between 10.000 and 19.999'
    elif simulated_investment >= 20000 and simulated_investment < 30000:
        return 'between 20.000 and 29.999'
    elif simulated_investment >= 30000 and simulated_investment < 40000:
        return 'between 30.000 and 39.999'
    elif simulated_investment >= 40000 and simulated_investment < 50000:
        return 'between 40.000 and 49.999'
    elif simulated_investment >= 50000 and simulated_investment < 60000:
        return 'between 50.000 and 59.999'
    elif simulated_investment >= 60000 and simulated_investment < 100000:
        return 'between 60.000 and 99.999'
    elif simulated_investment >= 100000 and simulated_investment < 150000:
        return 'between 100.000 and 149.999'
    elif simulated_investment >= 150000 and simulated_investment < 1000000:
        return 'between 150.000 and 999.999'
    else:
        return '1.000.000 or more'
```

```
[161]: conversion_funnel['simulated_investment_range'] =
        ↪ conversion_funnel['simulated_investment'].apply(set_bucket)
```

```
[162]: conversion_funnel.groupby(['page', 'simulated_investment_range']).agg({'id':
        ↪ 'count', 'simulated_investment': ['sum', 'mean']})
```

```
[162]:
```

		id \ count
page	simulated_investment_range	
1 - Homepage	1.000.000 or more	21
	below 10.000	1753
	between 10.000 and 19.999	583
	between 100.000 and 149.999	637
	between 150.000 and 999.999	10
	between 20.000 and 29.999	737

	between 30.000 and 39.999	56
	between 40.000 and 49.999	78
	between 50.000 and 59.999	288
	between 60.000 and 99.999	3
	no investment	5834
2 - Questionário	1.000.000 or more	21
	below 10.000	1753
	between 10.000 and 19.999	583
	between 100.000 and 149.999	637
	between 150.000 and 999.999	10
	between 20.000 and 29.999	737
	between 30.000 and 39.999	56
	between 40.000 and 49.999	78
	between 50.000 and 59.999	288
	between 60.000 and 99.999	3
3 - Plano de Investimentos	1.000.000 or more	21
	below 10.000	842
	between 10.000 and 19.999	385
	between 100.000 and 149.999	550
	between 150.000 and 999.999	10
	between 20.000 and 29.999	464
	between 30.000 and 39.999	10
	between 40.000 and 49.999	11
	between 50.000 and 59.999	210
	between 60.000 and 99.999	3
4 - Cadastro	below 10.000	364
	between 10.000 and 19.999	186
	between 100.000 and 149.999	508
	between 150.000 and 999.999	9
	between 20.000 and 29.999	333
	between 30.000 and 39.999	8
	between 40.000 and 49.999	1
	between 50.000 and 59.999	159
	between 60.000 and 99.999	3

	simulated_investment	\
page	simulated_investment_range	sum
1 - Homepage	1.000.000 or more	21000000.0
	below 10.000	6822000.0
	between 10.000 and 19.999	7703000.0
	between 100.000 and 149.999	67130000.0
	between 150.000 and 999.999	2200000.0
	between 20.000 and 29.999	15679000.0
	between 30.000 and 39.999	1680000.0
	between 40.000 and 49.999	3120000.0
	between 50.000 and 59.999	14400000.0

	between 60.000 and 99.999	180000.0
	no investment	0.0
2 - Questionário	1.000.000 or more	21000000.0
	below 10.000	6822000.0
	between 10.000 and 19.999	7703000.0
	between 100.000 and 149.999	67130000.0
	between 150.000 and 999.999	2200000.0
	between 20.000 and 29.999	15679000.0
	between 30.000 and 39.999	1680000.0
	between 40.000 and 49.999	3120000.0
	between 50.000 and 59.999	14400000.0
	between 60.000 and 99.999	180000.0
3 - Plano de Investimentos	1.000.000 or more	21000000.0
	below 10.000	3899000.0
	between 10.000 and 19.999	4803000.0
	between 100.000 and 149.999	58180000.0
	between 150.000 and 999.999	2200000.0
	between 20.000 and 29.999	9944000.0
	between 30.000 and 39.999	300000.0
	between 40.000 and 49.999	440000.0
	between 50.000 and 59.999	10500000.0
	between 60.000 and 99.999	180000.0
4 - Cadastro	below 10.000	2206000.0
	between 10.000 and 19.999	2538000.0
	between 100.000 and 149.999	53980000.0
	between 150.000 and 999.999	1800000.0
	between 20.000 and 29.999	7096000.0
	between 30.000 and 39.999	240000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	7950000.0
	between 60.000 and 99.999	180000.0

		mean
page	simulated_investment_range	
1 - Homepage	1.000.000 or more	1000000.000000
	below 10.000	3891.614375
	between 10.000 and 19.999	13212.692967
	between 100.000 and 149.999	105384.615385
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21274.084125
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
	no investment	NaN
2 - Questionário	1.000.000 or more	1000000.000000

	below 10.000	3891.614375
	between 10.000 and 19.999	13212.692967
	between 100.000 and 149.999	105384.615385
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21274.084125
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
3 - Plano de Investimentos	1.000.000 or more	1000000.000000
	below 10.000	4630.641330
	between 10.000 and 19.999	12475.324675
	between 100.000 and 149.999	105781.818182
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21431.034483
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
4 - Cadastro	below 10.000	6060.439560
	between 10.000 and 19.999	13645.161290
	between 100.000 and 149.999	106259.842520
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21309.309309
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000

Pessoas que simulam investimentos entre **R\\$ 100.000** e **R\\$ 150.000** e abaixo de **R\\$ 10.000** possuem taxas de conversão maiores no funil se comparadas as demais faixas.

21 pessoas simularam investimento de **R\\$ 1.000.000,00** ou mais. Todas responderam o questionário e passaram pela etapa de definição de plano de investimentos mas nenhuma fez o cadastro.

```
[163]: conversion_funnel[conversion_funnel['device'] == 'desktop'].
↳groupby(['device','page', 'simulated_investment_range']).agg({'id': 'count',
↳'simulated_investment': ['sum', 'mean', 'median']})
```

```
[163]:
```

		id \
		count
device	page	simulated_investment_range
desktop	1 - Homepage	below 10.000 1124
		between 10.000 and 19.999 430
		between 100.000 and 149.999 523
		between 150.000 and 999.999 2
		between 20.000 and 29.999 569
		between 30.000 and 39.999 49

	between 40.000 and 49.999	10
	between 50.000 and 59.999	209
	between 60.000 and 99.999	3
	no investment	4460
2 - Questionário	below 10.000	1124
	between 10.000 and 19.999	430
	between 100.000 and 149.999	523
	between 150.000 and 999.999	2
	between 20.000 and 29.999	569
	between 30.000 and 39.999	49
	between 40.000 and 49.999	10
	between 50.000 and 59.999	209
	between 60.000 and 99.999	3
3 - Plano de Investimentos	below 10.000	478
	between 10.000 and 19.999	257
	between 100.000 and 149.999	468
	between 150.000 and 999.999	2
	between 20.000 and 29.999	390
	between 30.000 and 39.999	3
	between 40.000 and 49.999	10
	between 50.000 and 59.999	161
	between 60.000 and 99.999	3
4 - Cadastro	below 10.000	249
	between 10.000 and 19.999	133
	between 100.000 and 149.999	443
	between 150.000 and 999.999	1
	between 20.000 and 29.999	282
	between 30.000 and 39.999	1
	between 50.000 and 59.999	116
	between 60.000 and 99.999	3
simulated_investment \		
sum		
device page	simulated_investment_range	
desktop 1 - Homepage	below 10.000	
5093000.0		
	between 10.000 and 19.999	
5745000.0		
	between 100.000 and 149.999	
55600000.0		
	between 150.000 and 999.999	
600000.0		
	between 20.000 and 29.999	
12074000.0		
	between 30.000 and 39.999	
1470000.0		
	between 40.000 and 49.999	

400000.0	between 50.000 and 59.999
10450000.0	between 60.000 and 99.999
180000.0	no investment
0.0	
2 - Questionário	below 10.000
5093000.0	between 10.000 and 19.999
5745000.0	between 100.000 and 149.999
55600000.0	between 150.000 and 999.999
600000.0	between 20.000 and 29.999
12074000.0	between 30.000 and 39.999
1470000.0	between 40.000 and 49.999
400000.0	between 50.000 and 59.999
10450000.0	between 60.000 and 99.999
180000.0	
3 - Plano de Investimentos	below 10.000
2575000.0	between 10.000 and 19.999
3143000.0	between 100.000 and 149.999
49860000.0	between 150.000 and 999.999
600000.0	between 20.000 and 29.999
8288000.0	between 30.000 and 39.999
90000.0	between 40.000 and 49.999
400000.0	between 50.000 and 59.999
8050000.0	between 60.000 and 99.999
180000.0	
4 - Cadastro	below 10.000
1511000.0	between 10.000 and 19.999
1807000.0	

47360000.0	between 100.000 and 149.999
200000.0	between 150.000 and 999.999
5981000.0	between 20.000 and 29.999
30000.0	between 30.000 and 39.999
5800000.0	between 50.000 and 59.999
180000.0	between 60.000 and 99.999

		mean
device	page	
desktop	1 - Homepage	
	simulated_investment_range	
	below 10.000	4531.138790
	between 10.000 and 19.999	13360.465116
	between 100.000 and 149.999	106309.751434
	between 150.000 and 999.999	300000.000000
	between 20.000 and 29.999	21219.683656
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
	no investment	NaN
2 - Questionário	below 10.000	4531.138790
	between 10.000 and 19.999	13360.465116
	between 100.000 and 149.999	106309.751434
	between 150.000 and 999.999	300000.000000
	between 20.000 and 29.999	21219.683656
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
3 - Plano de Investimentos	below 10.000	5387.029289
	between 10.000 and 19.999	12229.571984
	between 100.000 and 149.999	106538.461538
	between 150.000 and 999.999	300000.000000
	between 20.000 and 29.999	21251.282051
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
4 - Cadastro	below 10.000	6068.273092
	between 10.000 and 19.999	13586.466165
	between 100.000 and 149.999	106907.449210

	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21209.219858
	between 30.000 and 39.999	30000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
		median
device page	simulated_investment_range	
desktop 1 - Homepage	below 10.000	5000.0
	between 10.000 and 19.999	13000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	300000.0
	between 20.000 and 29.999	20000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
	between 60.000 and 99.999	60000.0
	no investment	NaN
2 - Questionário	below 10.000	5000.0
	between 10.000 and 19.999	13000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	300000.0
	between 20.000 and 29.999	20000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
	between 60.000 and 99.999	60000.0
3 - Plano de Investimentos	below 10.000	5000.0
	between 10.000 and 19.999	11000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	300000.0
	between 20.000 and 29.999	21000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
	between 60.000 and 99.999	60000.0
4 - Cadastro	below 10.000	7000.0
	between 10.000 and 19.999	15000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	200000.0
	between 20.000 and 29.999	21000.0
	between 30.000 and 39.999	30000.0
	between 50.000 and 59.999	50000.0
	between 60.000 and 99.999	60000.0

```
[164]: conversion_funnel[conversion_funnel['device'] == 'mobile'].groupby(['device',
↳ 'page', 'simulated_investment_range']).agg({'id': 'count',
↳ 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[164]:
```

		id \
		count
device	page	
mobile	1 - Homepage	
		simulated_investment_range
		1.000.000 or more
		21
		below 10.000
		629
		between 10.000 and 19.999
		153
		between 100.000 and 149.999
		114
		between 150.000 and 999.999
		8
		between 20.000 and 29.999
		168
		between 30.000 and 39.999
		7
		between 40.000 and 49.999
		68
		between 50.000 and 59.999
		79
		no investment
		1374
	2 - Questionário	
		1.000.000 or more
		21
		below 10.000
		629
		between 10.000 and 19.999
		153
		between 100.000 and 149.999
		114
		between 150.000 and 999.999
		8
		between 20.000 and 29.999
		168
		between 30.000 and 39.999
		7
		between 40.000 and 49.999
		68
		between 50.000 and 59.999
		79
	3 - Plano de Investimentos	
		1.000.000 or more
		21
		below 10.000
		364
		between 10.000 and 19.999
		128
		between 100.000 and 149.999
		82
		between 150.000 and 999.999
		8
		between 20.000 and 29.999
		74
		between 30.000 and 39.999
		7
		between 40.000 and 49.999
		1
		between 50.000 and 59.999
		49
	4 - Cadastro	
		below 10.000
		115
		between 10.000 and 19.999
		53
		between 100.000 and 149.999
		65
		between 150.000 and 999.999
		8
		between 20.000 and 29.999
		51
		between 30.000 and 39.999
		7
		between 40.000 and 49.999
		1
		between 50.000 and 59.999
		43
simulated_investment	\	
sum		
device	page	simulated_investment_range

mobile 1 - Homepage	1.000.000 or more
21000000.0	
	below 10.000
1729000.0	
	between 10.000 and 19.999
1958000.0	
	between 100.000 and 149.999
11530000.0	
	between 150.000 and 999.999
1600000.0	
	between 20.000 and 29.999
3605000.0	
	between 30.000 and 39.999
210000.0	
	between 40.000 and 49.999
2720000.0	
	between 50.000 and 59.999
3950000.0	
	no investment
0.0	
2 - Questionário	1.000.000 or more
21000000.0	
	below 10.000
1729000.0	
	between 10.000 and 19.999
1958000.0	
	between 100.000 and 149.999
11530000.0	
	between 150.000 and 999.999
1600000.0	
	between 20.000 and 29.999
3605000.0	
	between 30.000 and 39.999
210000.0	
	between 40.000 and 49.999
2720000.0	
	between 50.000 and 59.999
3950000.0	
3 - Plano de Investimentos	1.000.000 or more
21000000.0	
	below 10.000
1324000.0	
	between 10.000 and 19.999
1660000.0	
	between 100.000 and 149.999
8320000.0	
	between 150.000 and 999.999

1600000.0	between 20.000 and 29.999
1656000.0	between 30.000 and 39.999
210000.0	between 40.000 and 49.999
40000.0	between 50.000 and 59.999
2450000.0	below 10.000
4 - Cadastro	
695000.0	between 10.000 and 19.999
731000.0	between 100.000 and 149.999
6620000.0	between 150.000 and 999.999
1600000.0	between 20.000 and 29.999
1115000.0	between 30.000 and 39.999
210000.0	between 40.000 and 49.999
40000.0	between 50.000 and 59.999
2150000.0	

	simulated_investment_range	mean
device page	1.000.000 or more	1000000.000000
mobile 1 - Homepage	below 10.000	2748.807631
	between 10.000 and 19.999	12797.385621
	between 100.000 and 149.999	101140.350877
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21458.333333
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	no investment	NaN
2 - Questionário	1.000.000 or more	1000000.000000
	below 10.000	2748.807631
	between 10.000 and 19.999	12797.385621
	between 100.000 and 149.999	101140.350877
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21458.333333
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000

	between 50.000 and 59.999	50000.000000
3 - Plano de Investimentos	1.000.000 or more	1000000.000000
	below 10.000	3637.362637
	between 10.000 and 19.999	12968.750000
	between 100.000 and 149.999	101463.414634
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	22378.378378
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
4 - Cadastro	below 10.000	6043.478261
	between 10.000 and 19.999	13792.452830
	between 100.000 and 149.999	101846.153846
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21862.745098
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
		median
device page	simulated_investment_range	
mobile 1 - Homepage	1.000.000 or more	1000000.0
	below 10.000	2000.0
	between 10.000 and 19.999	12000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	200000.0
	between 20.000 and 29.999	20000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
	no investment	NaN
2 - Questionário	1.000.000 or more	1000000.0
	below 10.000	2000.0
	between 10.000 and 19.999	12000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	200000.0
	between 20.000 and 29.999	20000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
3 - Plano de Investimentos	1.000.000 or more	1000000.0
	below 10.000	3000.0
	between 10.000 and 19.999	12000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	200000.0
	between 20.000 and 29.999	22000.0

4 - Cadastro	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0
	below 10.000	7000.0
	between 10.000 and 19.999	15000.0
	between 100.000 and 149.999	100000.0
	between 150.000 and 999.999	200000.0
	between 20.000 and 29.999	22000.0
	between 30.000 and 39.999	30000.0
	between 40.000 and 49.999	40000.0
	between 50.000 and 59.999	50000.0

```
[165]: conversion_funnel[conversion_funnel['source'] == 'Direct'].groupby(['source',
↳ 'page', 'simulated_investment_range']).agg({'id': 'count',
↳ 'simulated_investment': ['sum', 'mean']})
```

```
[165]:
```

		id \
		count
source	page	
Direct 1 - Homepage	simulated_investment_range	
	1.000.000 or more	21
	below 10.000	348
	between 10.000 and 19.999	181
	between 100.000 and 149.999	378
	between 150.000 and 999.999	10
	between 20.000 and 29.999	265
	between 30.000 and 39.999	44
	between 40.000 and 49.999	68
	between 50.000 and 59.999	80
	between 60.000 and 99.999	3
2 - Questionário	no investment	1182
	1.000.000 or more	20
	below 10.000	439
	between 10.000 and 19.999	195
	between 100.000 and 149.999	381
	between 150.000 and 999.999	10
	between 20.000 and 29.999	284
	between 30.000 and 39.999	41
	between 40.000 and 49.999	62
	between 50.000 and 59.999	84
	between 60.000 and 99.999	3
3 - Plano de Investimentos	1.000.000 or more	20
	below 10.000	231
	between 10.000 and 19.999	170
	between 100.000 and 149.999	356
	between 150.000 and 999.999	10
	between 20.000 and 29.999	272
	between 30.000 and 39.999	7

	between 40.000 and 49.999	3
	between 50.000 and 59.999	79
	between 60.000 and 99.999	3
4 - Cadastro	below 10.000	141
	between 10.000 and 19.999	79
	between 100.000 and 149.999	344
	between 150.000 and 999.999	8
	between 20.000 and 29.999	164
	between 30.000 and 39.999	7
	between 40.000 and 49.999	1
	between 50.000 and 59.999	74
	between 60.000 and 99.999	3
simulated_investment \		
sum		
source page	simulated_investment_range	
Direct 1 - Homepage	1.000.000 or more	
21000000.0		
	below 10.000	
1249000.0		
	between 10.000 and 19.999	
2227000.0		
	between 100.000 and 149.999	
40910000.0		
	between 150.000 and 999.999	
2200000.0		
	between 20.000 and 29.999	
5686000.0		
	between 30.000 and 39.999	
1320000.0		
	between 40.000 and 49.999	
2720000.0		
	between 50.000 and 59.999	
4000000.0		
	between 60.000 and 99.999	
180000.0		
	no investment	
0.0		
2 - Questionário	1.000.000 or more	
20000000.0		
	below 10.000	
1552000.0		
	between 10.000 and 19.999	
2408000.0		
	between 100.000 and 149.999	
41170000.0		
	between 150.000 and 999.999	

2200000.0	between 20.000 and 29.999
6085000.0	between 30.000 and 39.999
1230000.0	between 40.000 and 49.999
2480000.0	between 50.000 and 59.999
4200000.0	between 60.000 and 99.999
180000.0	
3 - Plano de Investimentos	1.000.000 or more
20000000.0	
	below 10.000
1046000.0	between 10.000 and 19.999
2133000.0	between 100.000 and 149.999
38620000.0	between 150.000 and 999.999
2200000.0	between 20.000 and 29.999
5809000.0	between 30.000 and 39.999
210000.0	between 40.000 and 49.999
120000.0	between 50.000 and 59.999
3950000.0	between 60.000 and 99.999
180000.0	
4 - Cadastro	below 10.000
837000.0	between 10.000 and 19.999
1099000.0	between 100.000 and 149.999
37410000.0	between 150.000 and 999.999
1600000.0	between 20.000 and 29.999
3494000.0	between 30.000 and 39.999
210000.0	between 40.000 and 49.999
40000.0	between 50.000 and 59.999
3700000.0	

180000.0

between 60.000 and 99.999

		mean
source page	simulated_investment_range	
Direct 1 - Homepage	1.000.000 or more	1000000.000000
	below 10.000	3589.080460
	between 10.000 and 19.999	12303.867403
	between 100.000 and 149.999	108227.513228
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21456.603774
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
	no investment	NaN
2 - Questionário	1.000.000 or more	1000000.000000
	below 10.000	3535.307517
	between 10.000 and 19.999	12348.717949
	between 100.000 and 149.999	108057.742782
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21426.056338
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
3 - Plano de Investimentos	1.000.000 or more	1000000.000000
	below 10.000	4528.138528
	between 10.000 and 19.999	12547.058824
	between 100.000 and 149.999	108483.146067
	between 150.000 and 999.999	220000.000000
	between 20.000 and 29.999	21356.617647
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000
4 - Cadastro	below 10.000	5936.170213
	between 10.000 and 19.999	13911.392405
	between 100.000 and 149.999	108750.000000
	between 150.000 and 999.999	200000.000000
	between 20.000 and 29.999	21304.878049
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	between 60.000 and 99.999	60000.000000

Simulações de investimento abaixo de R\\$ 10.000,00 são mais comuns no questionário para quem veio por tráfego direto, porém é mais comum que o valor simulado de quem faça o cadastro seja entre R\\$ 100.000,00 e R\\$ 150.000,00.

```
[166]: conversion_funnel[conversion_funnel['source'] == 'Organic Search'].
↳groupby(['source', 'page', 'simulated_investment_range']).agg({'id':_
↳'count', 'simulated_investment': ['sum', 'mean']})
```

```
[166]:
```

source	page	simulated_investment_range	id \ count	
Organic Search	1 - Homepage	below 10.000	272	
		between 10.000 and 19.999	318	
		between 100.000 and 149.999	132	
		between 20.000 and 29.999	350	
		between 50.000 and 59.999	157	
		no investment	1318	
	2 - Questionário	below 10.000	311	
		between 10.000 and 19.999	309	
		between 100.000 and 149.999	135	
		between 20.000 and 29.999	328	
		between 40.000 and 49.999	2	
	3 - Plano de Investimentos	between 50.000 and 59.999	144	
		below 10.000	220	
		between 10.000 and 19.999	181	
		between 100.000 and 149.999	99	
		between 20.000 and 29.999	134	
	4 - Cadastro	between 50.000 and 59.999	90	
		below 10.000	132	
		between 10.000 and 19.999	93	
		between 100.000 and 149.999	83	
		between 150.000 and 999.999	1	
		between 20.000 and 29.999	133	
			between 50.000 and 59.999	62

source	page	simulated_investment_range	simulated_investment \ sum
Organic Search	1 - Homepage	below 10.000	1115000.0
		between 10.000 and 19.999	4418000.0
		between 100.000 and 149.999	13200000.0
		between 20.000 and 29.999	7384000.0
		between 50.000 and 59.999	

7850000.0		no investment
0.0		
	2 - Questionário	below 10.000
1297000.0		
		between 10.000 and 19.999
4279000.0		
		between 100.000 and 149.999
13540000.0		
		between 20.000 and 29.999
6923000.0		
		between 40.000 and 49.999
80000.0		
		between 50.000 and 59.999
7200000.0		
	3 - Plano de Investimentos	below 10.000
1078000.0		
		between 10.000 and 19.999
2240000.0		
		between 100.000 and 149.999
9970000.0		
		between 20.000 and 29.999
2885000.0		
		between 50.000 and 59.999
4500000.0		
	4 - Cadastro	below 10.000
840000.0		
		between 10.000 and 19.999
1247000.0		
		between 100.000 and 149.999
8400000.0		
		between 150.000 and 999.999
200000.0		
		between 20.000 and 29.999
2856000.0		
		between 50.000 and 59.999
3100000.0		
mean		
source	page	simulated_investment_range
Organic Search	1 - Homepage	below 10.000
4099.264706		
		between 10.000 and 19.999
13893.081761		
		between 100.000 and 149.999
100000.000000		

21097.142857		between 20.000 and 29.999
50000.000000		between 50.000 and 59.999
NaN		no investment
	2 - Questionário	
4170.418006		below 10.000
13847.896440		between 10.000 and 19.999
100296.296296		between 100.000 and 149.999
21106.707317		between 20.000 and 29.999
40000.000000		between 40.000 and 49.999
50000.000000		between 50.000 and 59.999
	3 - Plano de Investimentos	
4900.000000		below 10.000
12375.690608		between 10.000 and 19.999
100707.070707		between 100.000 and 149.999
21529.850746		between 20.000 and 29.999
50000.000000		between 50.000 and 59.999
	4 - Cadastro	
6363.636364		below 10.000
13408.602151		between 10.000 and 19.999
101204.819277		between 100.000 and 149.999
200000.000000		between 150.000 and 999.999
21473.684211		between 20.000 and 29.999
50000.000000		between 50.000 and 59.999

Visitas por busca orgânica tendem a simular e a cadastrar simulações de investimento menores do que tráfego direto, com valores de até R\$ 20.000,00.

```
[167]: conversion_funnel[conversion_funnel['source'] == 'Paid Search'].
        ↳groupby(['source', 'page', 'simulated_investment_range']).agg({'id':_
        ↳'count', 'simulated_investment': ['sum', 'mean']})
```

[167]:

			id \	count	
source	page	simulated_investment_range			
Paid Search 1 - Homepage		below 10.000		1037	
		between 10.000 and 19.999		21	
		between 100.000 and 149.999		120	
		between 20.000 and 29.999		61	
		between 50.000 and 59.999		19	
		no investment		3187	
2 - Questionário		1.000.000 or more		1	
		below 10.000		926	
		between 10.000 and 19.999		25	
		between 100.000 and 149.999		113	
		between 20.000 and 29.999		71	
		between 30.000 and 39.999		4	
		between 40.000 and 49.999		7	
		between 50.000 and 59.999		29	
	3 - Plano de Investimentos		1.000.000 or more		1
			below 10.000		387
			between 10.000 and 19.999		21
			between 100.000 and 149.999		88
		between 20.000 and 29.999		19	
		between 40.000 and 49.999		2	
		between 50.000 and 59.999		21	
	4 - Cadastro		below 10.000		91
			between 10.000 and 19.999		14
			between 100.000 and 149.999		78
			between 20.000 and 29.999		12
			between 50.000 and 59.999		17
simulated_investment \					
sum					
source	page	simulated_investment_range			
Paid Search 1 - Homepage		below 10.000			
	4175000.0				
		between 10.000 and 19.999			
	293000.0				
		between 100.000 and 149.999			
	12250000.0				
	between 20.000 and 29.999				
1305000.0					
	between 50.000 and 59.999				
950000.0					
	no investment				
0.0					
	2 - Questionário	1.000.000 or more			
1000000.0					

3750000.0		below 10.000
362000.0		between 10.000 and 19.999
11550000.0		between 100.000 and 149.999
1517000.0		between 20.000 and 29.999
120000.0		between 30.000 and 39.999
280000.0		between 40.000 and 49.999
1450000.0		between 50.000 and 59.999
	3 - Plano de Investimentos	1.000.000 or more
1000000.0		below 10.000
1760000.0		between 10.000 and 19.999
294000.0		between 100.000 and 149.999
8820000.0		between 20.000 and 29.999
418000.0		between 40.000 and 49.999
80000.0		between 50.000 and 59.999
1050000.0		below 10.000
	4 - Cadastro	below 10.000
529000.0		between 10.000 and 19.999
192000.0		between 100.000 and 149.999
7840000.0		between 20.000 and 29.999
266000.0		between 50.000 and 59.999
850000.0		
mean		
source	page	simulated_investment_range
Paid Search 1 - Homepage		below 10.000
4026.036644		between 10.000 and 19.999
13952.380952		between 100.000 and 149.999

102083.333333	between 20.000 and 29.999
21393.442623	between 50.000 and 59.999
50000.000000	no investment
NaN	
2 - Questionário	1.000.000 or more
1000000.000000	below 10.000
4049.676026	between 10.000 and 19.999
14480.000000	between 100.000 and 149.999
102212.389381	between 20.000 and 29.999
21366.197183	between 30.000 and 39.999
30000.000000	between 40.000 and 49.999
40000.000000	between 50.000 and 59.999
50000.000000	1.000.000 or more
3 - Plano de Investimentos	below 10.000
1000000.000000	between 10.000 and 19.999
4547.803618	between 100.000 and 149.999
14000.000000	between 20.000 and 29.999
100227.272727	between 40.000 and 49.999
22000.000000	between 50.000 and 59.999
40000.000000	below 10.000
50000.000000	between 10.000 and 19.999
4 - Cadastro	between 100.000 and 149.999
5813.186813	between 20.000 and 29.999
13714.285714	between 50.000 and 59.999
100512.820513	
22166.666667	
50000.000000	

O comportamento de simulações de visitantes vindos por mídia impulsionada é similar ao de tráfego direto: os valores simulados no topo do funil ficam abaixo de R\\$ 10.000,00 mas ao chegar na etapa de cadastro, valores entre R\\$ 100.000,00 e R\\$ 150.000,00 são o segundo intervalo mais comum.

```
[168]: conversion_funnel[conversion_funnel['source'] == 'Social'].groupby(['source',
↳ 'page', 'simulated_investment_range']).agg({'id': 'count',
↳ 'simulated_investment': ['sum', 'mean']})
```

```
[168]:
```

		id \	
		count	
source	page	simulated_investment_range	
Social	1 - Homepage	below 10.000	96
		between 10.000 and 19.999	63
		between 100.000 and 149.999	7
		between 20.000 and 29.999	61
		between 30.000 and 39.999	12
		between 40.000 and 49.999	10
		between 50.000 and 59.999	32
		no investment	147
	2 - Questionário	below 10.000	77
		between 10.000 and 19.999	54
	3 - Plano de Investimentos	between 100.000 and 149.999	8
		between 20.000 and 29.999	54
		between 30.000 and 39.999	11
		between 40.000 and 49.999	7
		between 50.000 and 59.999	31
		below 10.000	4
	4 - Cadastro	between 10.000 and 19.999	13
		between 100.000 and 149.999	7
		between 20.000 and 29.999	39
		between 30.000 and 39.999	3
		between 40.000 and 49.999	6
		between 50.000 and 59.999	20
simulated_investment \	sum	between 100.000 and 149.999	3
		between 20.000 and 29.999	24
		between 30.000 and 39.999	1
		between 50.000 and 59.999	6
		below 10.000	
		between 10.000 and 19.999	
	source page	simulated_investment_range	
		below 10.000	
		between 10.000 and 19.999	
		between 100.000 and 149.999	

1304000.0	between 20.000 and 29.999
360000.0	between 30.000 and 39.999
400000.0	between 40.000 and 49.999
1600000.0	between 50.000 and 59.999
0.0	no investment
2 - Questionário	below 10.000
223000.0	between 10.000 and 19.999
654000.0	between 100.000 and 149.999
870000.0	between 20.000 and 29.999
1154000.0	between 30.000 and 39.999
330000.0	between 40.000 and 49.999
280000.0	between 50.000 and 59.999
1550000.0	below 10.000
3 - Plano de Investimentos	below 10.000
15000.0	between 10.000 and 19.999
136000.0	between 100.000 and 149.999
770000.0	between 20.000 and 29.999
832000.0	between 30.000 and 39.999
90000.0	between 40.000 and 49.999
240000.0	between 50.000 and 59.999
1000000.0	between 100.000 and 149.999
4 - Cadastro	between 100.000 and 149.999
330000.0	between 20.000 and 29.999
480000.0	between 30.000 and 39.999
30000.0	between 50.000 and 59.999
300000.0	

		mean
source page	simulated_investment_range	
Social 1 - Homepage	below 10.000	2947.916667
	between 10.000 and 19.999	12142.857143
	between 100.000 and 149.999	110000.000000
	between 20.000 and 29.999	21377.049180
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
	no investment	NaN
2 - Questionário	below 10.000	2896.103896
	between 10.000 and 19.999	12111.111111
	between 100.000 and 149.999	108750.000000
	between 20.000 and 29.999	21370.370370
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
3 - Plano de Investimentos	below 10.000	3750.000000
	between 10.000 and 19.999	10461.538462
	between 100.000 and 149.999	110000.000000
	between 20.000 and 29.999	21333.333333
	between 30.000 and 39.999	30000.000000
	between 40.000 and 49.999	40000.000000
	between 50.000 and 59.999	50000.000000
4 - Cadastro	between 100.000 and 149.999	110000.000000
	between 20.000 and 29.999	20000.000000
	between 30.000 and 39.999	30000.000000
	between 50.000 and 59.999	50000.000000

O tráfego que veio por redes sociais possui a menor participação de origens na amostra, mas o valor de investimento simulado de quem chega na etapa de cadastro tem uma tendência mais clara que as demais origens de ficar na faixa entre R\$\\$ 20.000,00 e R\$\\$ 30.000,00.

1.3.2 Análise exploratória de origens, dispositivos e valores simulados

```
[169]: conversion_funnel_unique.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 10000 entries, 0 to 18239
Data columns (total 4 columns):
#   Column                Non-Null Count  Dtype
---  -
0   id                    10000 non-null  int64
1   source                10000 non-null  object
2   device                10000 non-null  object
3   simulated_investment  4166 non-null   float64
dtypes: float64(1), int64(1), object(2)
```


memory usage: 390.6+ KB

A única coluna com dados nulos é a de valor simulado.

```
[170]: conversion_funnel_unique.describe()
```

```
[170]:
```

	id	simulated_investment
count	10000.000000	4166.000000
mean	19648.169800	33584.733557
std	9530.425416	77979.480770
min	1.000000	1000.000000
25%	9987.750000	5000.000000
50%	22654.500000	14000.000000
75%	26065.250000	30000.000000
max	32299.000000	1000000.000000

O menor valor simulado foi de **R\$ 1.000,00** e o maior de **R\$ 1.000.000,00**. A mediana é de **R\$ 14.000,00** e a média é **R\$ 33.584,73**, indicando que os valores de simulação são bem diversos e que há valores extremos (outliers).

```
[171]: conversion_funnel_unique['simulated_investment'].value_counts()
```

```
[171]:
```

100000.0	421
1000.0	416
20000.0	390
5000.0	358
2000.0	301
50000.0	288
22000.0	192
4000.0	170
9000.0	168
10000.0	156
3000.0	118
110000.0	109
6000.0	105
12000.0	100
25000.0	100
7000.0	98
120000.0	87
15000.0	85
40000.0	78
17000.0	63
30000.0	56
21000.0	55
11000.0	47
16000.0	43
14000.0	32
19000.0	23

```

13000.0      21
1000000.0    21
130000.0     20
8000.0       19
18000.0     13
200000.0      9
60000.0      3
400000.0      1
Name: simulated_investment, dtype: int64

```

```

[172]: df_info = pd.DataFrame({
        'percentage': conversion_funnel_unique['simulated_investment'].
        ↪value_counts() / conversion_funnel_unique['simulated_investment'].count()
    }).reset_index()
df_info.rename(columns={'index': 'simulated_investment'}, inplace=True)

```

```

[173]: df_info[:6]['percentage'].sum()

```

```

[173]: 0.5218434949591935

```

Os valores mais simulados são de R\\$ 100.000,00, R\\$ 1.000,00, R\\$ 20.000,00, R\\$ 5.000,00, R\\$ 2.000,00 e R\\$ 50.000,00. Juntos, estes valores representam 52,18% de todas as simulações de investimento.

```

[174]: conversion_funnel_unique['simulated_investment'].value_counts().reset_index().
        ↪sort_values(by=['index'], ascending=False)

```

```

[174]:
   index simulated_investment
27  1000000.0                21
33   400000.0                 1
31  200000.0                 9
28  130000.0                20
16  120000.0                87
11  110000.0               109
0   100000.0               421
32   60000.0                 3
5    50000.0               288
18   40000.0                78
20   30000.0                56
14   25000.0               100
6    22000.0               192
21   21000.0                55
2    20000.0               390
25   19000.0                23
30   18000.0                13
19   17000.0                63
23   16000.0                43

```

17	15000.0	85
24	14000.0	32
26	13000.0	21
13	12000.0	100
22	11000.0	47
9	10000.0	156
8	9000.0	168
29	8000.0	19
15	7000.0	98
12	6000.0	105
3	5000.0	358
7	4000.0	170
10	3000.0	118
4	2000.0	301
1	1000.0	416

21 pessoas simularam investimentos de R\\$ 1.000.000,00.

```
[175]: conversion_funnel_unique[conversion_funnel_unique['simulated_investment'] != 1000000].describe()
```

```
[175]:
```

	id	simulated_investment
count	9979.000000	4145.00000
mean	19646.253633	28688.54041
std	9540.357501	36806.76776
min	1.000000	1000.00000
25%	9982.500000	5000.00000
50%	22665.000000	13000.00000
75%	26070.500000	30000.00000
max	32299.000000	400000.00000

```
[176]: conversion_funnel_unique.describe()
```

```
[176]:
```

	id	simulated_investment
count	10000.000000	4166.000000
mean	19648.169800	33584.733557
std	9530.425416	77979.480770
min	1.000000	1000.000000
25%	9987.750000	5000.000000
50%	22654.500000	14000.000000
75%	26065.250000	30000.000000
max	32299.000000	1000000.000000

Retirando os investimentos de R\\$ 1.000.000,00, o desvio padrão cai pela metade e a média de investimentos cai cerca de R\\$ 5.000,00.

```
[177]: conversion_funnel_unique['device'].value_counts() / conversion_funnel_unique['device'].count()
```

```
[177]: desktop    0.7378
      mobile    0.2622
      Name: device, dtype: float64
```

73,78% das visitas aconteceram em desktop e **26,22%** no mobile.

```
[178]: conversion_funnel_unique['source'].value_counts() /_
      ↪ conversion_funnel_unique['source'].count()
```

```
[178]: Paid Search    0.4367
      Direct         0.2709
      Organic Search  0.2525
      Social         0.0399
      Name: source, dtype: float64
```

43,67% das visitas vieram por mídia impulsionada, **27,09%** vieram por tráfego direto, **25,25%** vieram por busca orgânica e **3,99%** pelas redes sociais.

```
[261]: conversion_funnel_unique.groupby(['source', 'device']).agg({'id': ['count'],_
      ↪ 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[261]:
```

		id	simulated_investment		
		count	sum	mean	median
source	device				
Direct	desktop	1956	51449000.0	46059.982095	22000.0
	mobile	753	32322000.0	78834.146341	18500.0
Organic Search	desktop	1573	23526000.0	30914.586071	20000.0
	mobile	952	11982000.0	26865.470852	19500.0
Paid Search	desktop	3487	12356000.0	15406.483791	5000.0
	mobile	880	3927000.0	10388.888889	3000.0
Social	desktop	362	4181000.0	17567.226891	12000.0
	mobile	37	171000.0	12214.285714	12000.0

```
[180]: conversion_funnel_unique[(conversion_funnel_unique['device'] == 'mobile') &_
      ↪ (conversion_funnel_unique['source'] == 'Direct')]['id'].count() /_
      ↪ conversion_funnel_unique[(conversion_funnel_unique['device'] == 'desktop') &_
      ↪ (conversion_funnel_unique['source'] == 'Direct')]['id'].count()
```

```
[180]: 0.38496932515337423
```

Visitantes mobile que entraram por tráfego direto representam 38,49% do volume de desktop mobile que vieram por esta mesma origem. Quem acessou o site por mobile e veio por tráfego direto fez, em média, simulações maiores do que as demais combinações de fonte de tráfego e dispositivo. A mediana das simulações, por outro lado, é bem mais baixa, indicando que há valores extremos (outliers) nas simulações de mobile por tráfego direto.

```
[181]: conversion_funnel_unique['simulated_investment_range'] =_
      ↪ conversion_funnel_unique['simulated_investment'].apply(set_bucket)
```

```
conversion_funnel_unique
```

```
[181]:      id      source  device  simulated_investment  \
0         1  Paid Search  desktop                NaN
1         2  Paid Search  desktop                NaN
2         3  Paid Search  desktop             2000.0
6         4  Paid Search  desktop                NaN
7         5  Paid Search  desktop                NaN
...      ...      ...      ...      ...
18225    32295      Direct  desktop             120000.0
18229    32296      Direct  desktop             120000.0
18233    32297      Direct  desktop             120000.0
18237    32298      Direct  desktop              50000.0
18239    32299      Direct  desktop             120000.0
```

```
      simulated_investment_range
0                no investment
1                no investment
2             below 10.000
6                no investment
7                no investment
...      ...
18225  between 100.000 and 149.999
18229  between 100.000 and 149.999
18233  between 100.000 and 149.999
18237   between 50.000 and 59.999
18239  between 100.000 and 149.999
```

```
[10000 rows x 5 columns]
```

```
[182]: conversion_funnel_unique.groupby(['device', 'simulated_investment_range']).
      ↪agg({'id': 'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[182]:      id simulated_investment  \
      count                sum
device simulated_investment_range
desktop below 10.000             1124      5093000.0
      between 10.000 and 19.999        430      5745000.0
      between 100.000 and 149.999       522     55500000.0
      between 150.000 and 999.999         2       600000.0
      between 20.000 and 29.999        569     12074000.0
      between 30.000 and 39.999         49      1470000.0
      between 40.000 and 49.999         10       400000.0
      between 50.000 and 59.999        209     10450000.0
      between 60.000 and 99.999          3       180000.0
      no investment          4460           0.0
mobile  1.000.000 or more          21     21000000.0
```

below 10.000	629	1729000.0
between 10.000 and 19.999	153	1958000.0
between 100.000 and 149.999	115	11630000.0
between 150.000 and 999.999	8	1600000.0
between 20.000 and 29.999	168	3605000.0
between 30.000 and 39.999	7	210000.0
between 40.000 and 49.999	68	2720000.0
between 50.000 and 59.999	79	3950000.0
no investment	1374	0.0

		mean	median
device	simulated_investment_range		
desktop	below 10.000	4531.138790	5000.0
	between 10.000 and 19.999	13360.465116	13000.0
	between 100.000 and 149.999	106321.839080	100000.0
	between 150.000 and 999.999	300000.000000	300000.0
	between 20.000 and 29.999	21219.683656	20000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 40.000 and 49.999	40000.000000	40000.0
	between 50.000 and 59.999	50000.000000	50000.0
	between 60.000 and 99.999	60000.000000	60000.0
mobile	no investment	NaN	NaN
	1.000.000 or more	1000000.000000	1000000.0
	below 10.000	2748.807631	2000.0
	between 10.000 and 19.999	12797.385621	12000.0
	between 100.000 and 149.999	101130.434783	100000.0
	between 150.000 and 999.999	200000.000000	200000.0
	between 20.000 and 29.999	21458.333333	20000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 40.000 and 49.999	40000.000000	40000.0
	between 50.000 and 59.999	50000.000000	50000.0
	no investment	NaN	NaN

Simulações de investimento entre R\\$ 40.000,00 e R\\$ 50.000,00 são mais comuns em dispositivos mobile.

```
[183]: conversion_funnel_unique.groupby(['source', 'simulated_investment_range']).
        ↪agg({'id': 'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[183]:
```

		id	simulated_investment	\
		count	sum	
source	simulated_investment_range			
Direct	1.000.000 or more	21	21000000.0	
	below 10.000	414	1543000.0	
	between 10.000 and 19.999	204	2553000.0	
	between 100.000 and 149.999	386	41610000.0	

	between 150.000 and 999.999	9	2000000.0
	between 20.000 and 29.999	282	6045000.0
	between 30.000 and 39.999	44	1320000.0
	between 40.000 and 49.999	68	2720000.0
	between 50.000 and 59.999	96	4800000.0
	between 60.000 and 99.999	3	180000.0
	no investment	1182	0.0
Organic Search	below 10.000	254	984000.0
	between 10.000 and 19.999	300	4162000.0
	between 100.000 and 149.999	152	15300000.0
	between 150.000 and 999.999	1	200000.0
	between 20.000 and 29.999	350	7362000.0
	between 50.000 and 59.999	150	7500000.0
	no investment	1318	0.0
Paid Search	below 10.000	989	4012000.0
	between 10.000 and 19.999	18	243000.0
	between 100.000 and 149.999	96	9890000.0
	between 20.000 and 29.999	60	1288000.0
	between 50.000 and 59.999	17	850000.0
	no investment	3187	0.0
Social	below 10.000	96	283000.0
	between 10.000 and 19.999	61	745000.0
	between 100.000 and 149.999	3	330000.0
	between 20.000 and 29.999	45	984000.0
	between 30.000 and 39.999	12	360000.0
	between 40.000 and 49.999	10	400000.0
	between 50.000 and 59.999	25	1250000.0
	no investment	147	0.0

		mean	median
source	simulated_investment_range		
Direct	1.000.000 or more	1000000.000000	1000000.0
	below 10.000	3727.053140	2000.0
	between 10.000 and 19.999	12514.705882	12000.0
	between 100.000 and 149.999	107797.927461	110000.0
	between 150.000 and 999.999	222222.222222	200000.0
	between 20.000 and 29.999	21436.170213	22000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 40.000 and 49.999	40000.000000	40000.0
	between 50.000 and 59.999	50000.000000	50000.0
	between 60.000 and 99.999	60000.000000	60000.0
	no investment	NaN	NaN
Organic Search	below 10.000	3874.015748	3000.0
	between 10.000 and 19.999	13873.333333	15000.0
	between 100.000 and 149.999	100657.894737	100000.0
	between 150.000 and 999.999	200000.000000	200000.0

Paid Search	between 20.000 and 29.999	21034.285714	20000.0
	between 50.000 and 59.999	50000.000000	50000.0
	no investment	NaN	NaN
	below 10.000	4056.622851	4000.0
	between 10.000 and 19.999	13500.000000	13500.0
	between 100.000 and 149.999	103020.833333	100000.0
	between 20.000 and 29.999	21466.666667	20000.0
	between 50.000 and 59.999	50000.000000	50000.0
Social	no investment	NaN	NaN
	below 10.000	2947.916667	2000.0
	between 10.000 and 19.999	12213.114754	12000.0
	between 100.000 and 149.999	110000.000000	110000.0
	between 20.000 and 29.999	21866.666667	20000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 40.000 and 49.999	40000.000000	40000.0
	between 50.000 and 59.999	50000.000000	50000.0
	no investment	NaN	NaN

Quem entrou por busca orgânica fez mais simulações entre R\\$ 10.000,00 e R\\$ 20.000,00 do que as demais origens, cujos valores de simulação mais comuns foram abaixo de R\\$ 10.000,00.

```
[184]: conversion_funnel_unique[conversion_funnel_unique['source'] == 'Direct'].
      ↳groupby(['source','device', 'simulated_investment_range']).agg({'id':λ
      ↳'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[184]:
```

		id	simulated_investment	\
		count		sum
source	device	simulated_investment_range		
Direct	desktop	below 10.000	280	1144000.0
		between 10.000 and 19.999	133	1677000.0
		between 100.000 and 149.999	348	37720000.0
		between 150.000 and 999.999	2	600000.0
		between 20.000 and 29.999	231	4868000.0
		between 30.000 and 39.999	37	1110000.0
		between 50.000 and 59.999	83	4150000.0
		between 60.000 and 99.999	3	180000.0
		no investment	839	0.0
mobile		1.000.000 or more	21	21000000.0
		below 10.000	134	399000.0
		between 10.000 and 19.999	71	876000.0
		between 100.000 and 149.999	38	3890000.0
		between 150.000 and 999.999	7	1400000.0
		between 20.000 and 29.999	51	1177000.0
		between 30.000 and 39.999	7	210000.0
		between 40.000 and 49.999	68	2720000.0
		between 50.000 and 59.999	13	650000.0
		no investment	343	0.0

		mean	median
source device	simulated_investment_range		
Direct desktop	below 10.000	4085.714286	4000.0
	between 10.000 and 19.999	12609.022556	11000.0
	between 100.000 and 149.999	108390.804598	110000.0
	between 150.000 and 999.999	300000.000000	300000.0
	between 20.000 and 29.999	21073.593074	20000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 50.000 and 59.999	50000.000000	50000.0
	between 60.000 and 99.999	60000.000000	60000.0
mobile	no investment	NaN	NaN
	1.000.000 or more	1000000.000000	1000000.0
	below 10.000	2977.611940	1000.0
	between 10.000 and 19.999	12338.028169	12000.0
	between 100.000 and 149.999	102368.421053	100000.0
	between 150.000 and 999.999	200000.000000	200000.0
	between 20.000 and 29.999	23078.431373	22000.0
	between 30.000 and 39.999	30000.000000	30000.0
	between 40.000 and 49.999	40000.000000	40000.0
	between 50.000 and 59.999	50000.000000	50000.0
	no investment	NaN	NaN

As 21 pessoas que simularam R/\$ 1.000.000,00 o fizeram por mobile vindo por tráfego direto.

```
[185]: conversion_funnel_unique[conversion_funnel_unique['source'] == 'Organic_
↳Search'].groupby(['source','device', 'simulated_investment_range']).
↳agg({'id': 'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[185]:
```

			id simulated_investment \
			count sum
source	device	simulated_investment_range	
Organic Search	desktop	below 10.000	97 630000.0
		between 10.000 and 19.999	234 3275000.0
		between 100.000 and 149.999	97 9770000.0
		between 20.000 and 29.999	236 5001000.0
		between 50.000 and 59.999	97 4850000.0
		no investment	812 0.0
	mobile	below 10.000	157 354000.0
		between 10.000 and 19.999	66 887000.0
		between 100.000 and 149.999	55 5530000.0
		between 150.000 and 999.999	1 200000.0
		between 20.000 and 29.999	114 2361000.0
		between 50.000 and 59.999	53 2650000.0
		no investment	506 0.0

source	device	simulated_investment_range	mean	median
Organic Search	desktop	below 10.000	6494.845361	7000.0
		between 10.000 and 19.999	13995.726496	15000.0
		between 100.000 and 149.999	100721.649485	100000.0
		between 20.000 and 29.999	21190.677966	21000.0
		between 50.000 and 59.999	50000.000000	50000.0
		no investment	NaN	NaN
	mobile	below 10.000	2254.777070	1000.0
		between 10.000 and 19.999	13439.393939	14000.0
		between 100.000 and 149.999	100545.454545	100000.0
		between 150.000 and 999.999	200000.000000	200000.0
		between 20.000 and 29.999	20710.526316	20000.0
		between 50.000 and 59.999	50000.000000	50000.0
		no investment	NaN	NaN

```
[186]: conversion_funnel_unique[conversion_funnel_unique['source'] == 'Paid Search'].
        ↳groupby(['source', 'device', 'simulated_investment_range']).agg({'id': λ
        ↳'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[186]:
```

source	device	simulated_investment_range	id count	simulated_investment sum \
Paid Search	desktop	below 10.000	651	3036000.0
		between 10.000 and 19.999	15	199000.0
		between 100.000 and 149.999	74	7680000.0
		between 20.000 and 29.999	58	1241000.0
		between 50.000 and 59.999	4	200000.0
		no investment	2685	0.0
	mobile	below 10.000	338	976000.0
		between 10.000 and 19.999	3	44000.0
		between 100.000 and 149.999	22	2210000.0
		between 20.000 and 29.999	2	47000.0
		between 50.000 and 59.999	13	650000.0
		no investment	502	0.0

source	device	simulated_investment_range	mean	median
Paid Search	desktop	below 10.000	4663.594470	5000.0
		between 10.000 and 19.999	13266.666667	13000.0
		between 100.000 and 149.999	103783.783784	100000.0
		between 20.000 and 29.999	21396.551724	20000.0
		between 50.000 and 59.999	50000.000000	50000.0
		no investment	NaN	NaN
	mobile	below 10.000	2887.573964	2500.0

between 10.000 and 19.999	14666.666667	17000.0
between 100.000 and 149.999	100454.545455	100000.0
between 20.000 and 29.999	23500.000000	23500.0
between 50.000 and 59.999	50000.000000	50000.0
no investment	NaN	NaN

```
[187]: conversion_funnel_unique[conversion_funnel_unique['source'] == 'Social'].
      ↳groupby(['source', 'device', 'simulated_investment_range']).agg({'id': λ
      ↳'count', 'simulated_investment': ['sum', 'mean', 'median']})
```

```
[187]:
```

		id	simulated_investment	\
		count		sum
source device	simulated_investment_range			
Social desktop	below 10.000	96		283000.0
	between 10.000 and 19.999	48		594000.0
	between 100.000 and 149.999	3		330000.0
	between 20.000 and 29.999	44		964000.0
	between 30.000 and 39.999	12		360000.0
	between 40.000 and 49.999	10		400000.0
	between 50.000 and 59.999	25		1250000.0
	no investment	124		0.0
mobile	between 10.000 and 19.999	13		151000.0
	between 20.000 and 29.999	1		20000.0
	no investment	23		0.0

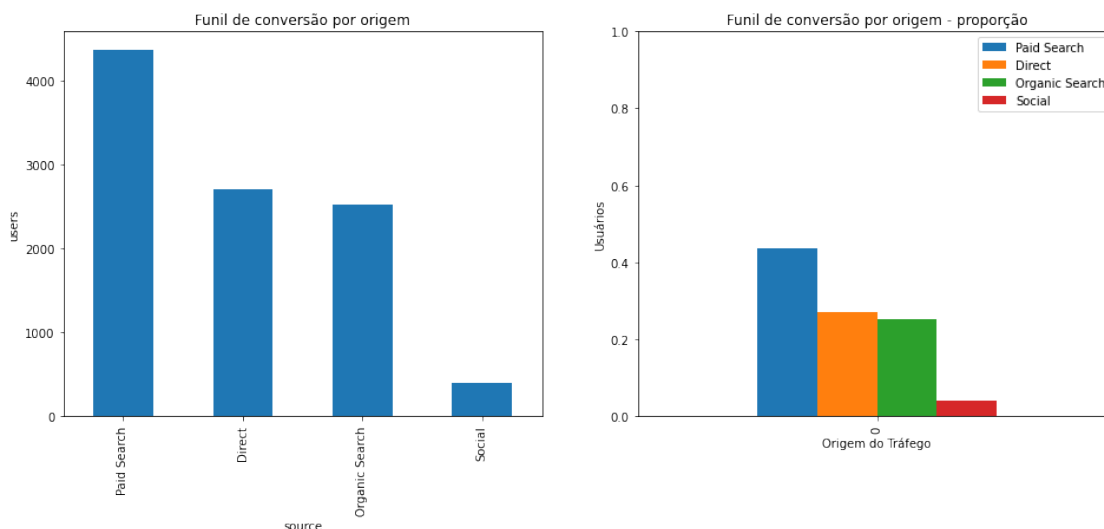
			mean	median
source device	simulated_investment_range			
Social desktop	below 10.000		2947.916667	2000.0
	between 10.000 and 19.999		12375.000000	12000.0
	between 100.000 and 149.999		110000.000000	110000.0
	between 20.000 and 29.999		21909.090909	20000.0
	between 30.000 and 39.999		30000.000000	30000.0
	between 40.000 and 49.999		40000.000000	40000.0
	between 50.000 and 59.999		50000.000000	50000.0
	no investment		NaN	NaN
mobile	between 10.000 and 19.999		11615.384615	12000.0
	between 20.000 and 29.999		20000.000000	20000.0
	no investment		NaN	NaN

A menor porção do tráfego é de social. Dentro deste grupo, praticamente todas as simulações de investimento feitas por dispositivos mobile foram na mesma faixa de valores.

1.4 Visualização dos dados

```
[188]: fig, ax = plt.subplots(ncols=2, figsize=(16, 6))
conversion_funnel_unique['source'].value_counts().plot(ax=ax[0], kind='bar',
→title='Funil de conversão por origem', xlabel='source', ylabel='users')
pd.DataFrame({
    'Paid Search':
→[len(conversion_funnel_unique[conversion_funnel_unique['source'] == 'Paid_
→Search']['source']) / len(conversion_funnel_unique['source'])],
    'Direct': [len(conversion_funnel_unique[conversion_funnel_unique['source']
→== 'Direct']['source']) / len(conversion_funnel_unique['source'])],
    'Organic Search':
→[len(conversion_funnel_unique[conversion_funnel_unique['source'] == 'Organic_
→Search']['source']) / len(conversion_funnel_unique['source'])],
    'Social': [len(conversion_funnel_unique[conversion_funnel_unique['source']
→== 'Social']['source']) / len(conversion_funnel_unique['source'])]
}).plot(ax=ax[1], kind='bar', title='Funil de conversão por origem -
→proporção', xlabel='Origem do Tráfego', ylabel='Usuários', yticks=(0, .2, .
→4, .6, .8, 1))
```

```
[188]: <AxesSubplot:title={'center': 'Funil de conversão por origem - proporção'},
xlabel='Origem do Tráfego', ylabel='Usuários'>
```



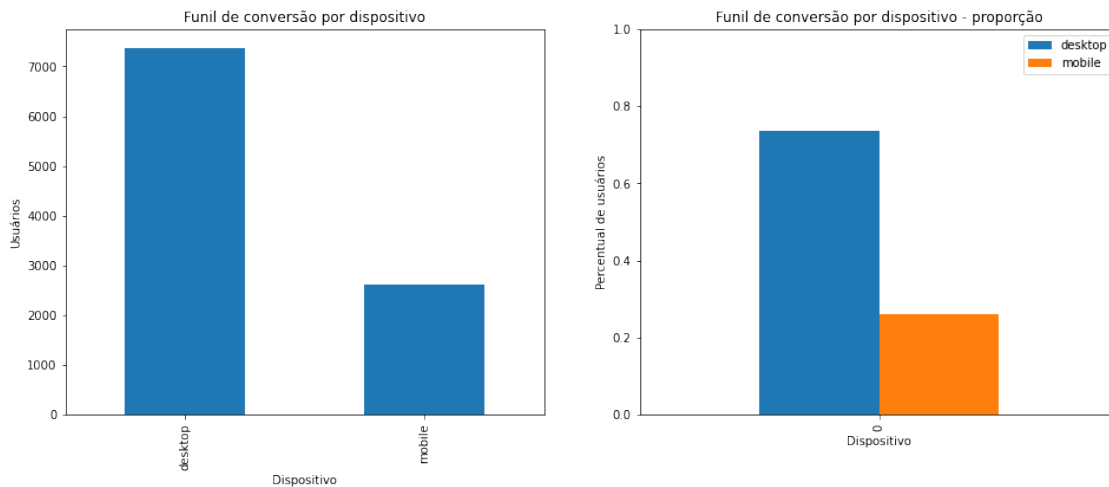
```
[189]: fig, ax = plt.subplots(ncols=2, figsize=(16, 6))
conversion_funnel_unique['device'].value_counts().plot(ax=ax[0], kind='bar',
→title='Funil de conversão por dispositivo', xlabel='Dispositivo',
→ylabel='Usuários')
pd.DataFrame({
```

```

    'desktop': [len(conversion_funnel_unique[conversion_funnel_unique['device']_
↪== 'desktop']['device']) / len(conversion_funnel_unique['device'])],
    'mobile': [len(conversion_funnel_unique[conversion_funnel_unique['device']_
↪== 'mobile']['device']) / len(conversion_funnel_unique['device'])]
}).plot(ax=ax[1], kind='bar', title='Funil de conversão por dispositivo -_
↪proporção', xlabel='Dispositivo', ylabel='Percentual de usuários',_
↪yticks=(0, .2, .4, .6, .8, 1))

```

[189]: <AxesSubplot:title={'center': 'Funil de conversão por dispositivo - proporção'},
 xlabel='Dispositivo', ylabel='Percentual de usuários'>

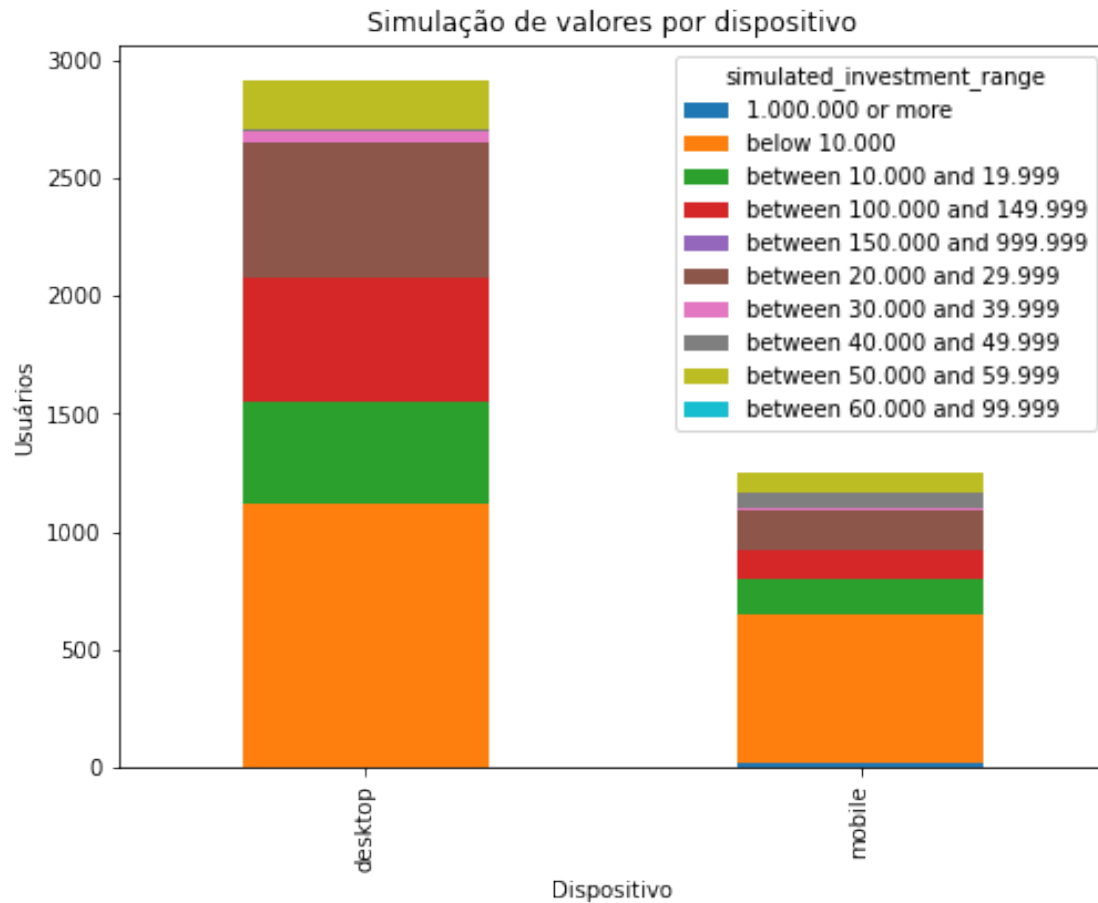


```

[190]: conversion_funnel_unique[conversion_funnel_unique['simulated_investment_range']_
↪!= 'no investment'].groupby(['device', 'simulated_investment_range']).size().
↪unstack().plot(kind='bar', stacked=True, title='Simulação de valores por_
↪dispositivo', xlabel='Dispositivo', ylabel='Usuários', figsize=(8, 6))

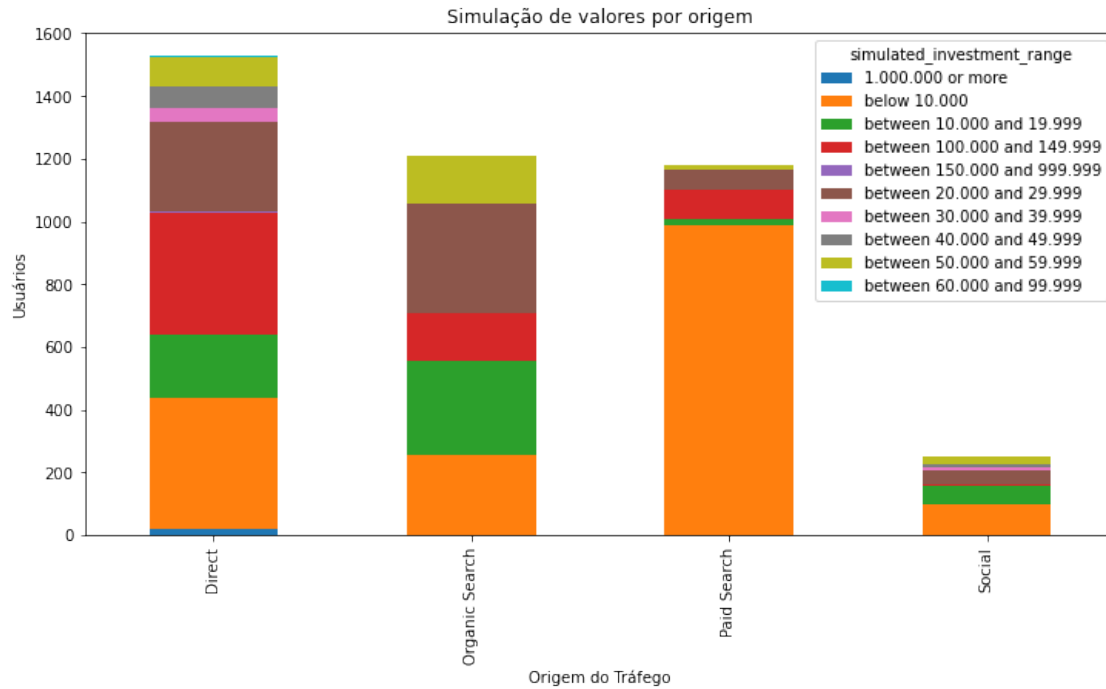
```

[190]: <AxesSubplot:title={'center': 'Simulação de valores por dispositivo'},
 xlabel='Dispositivo', ylabel='Usuários'>



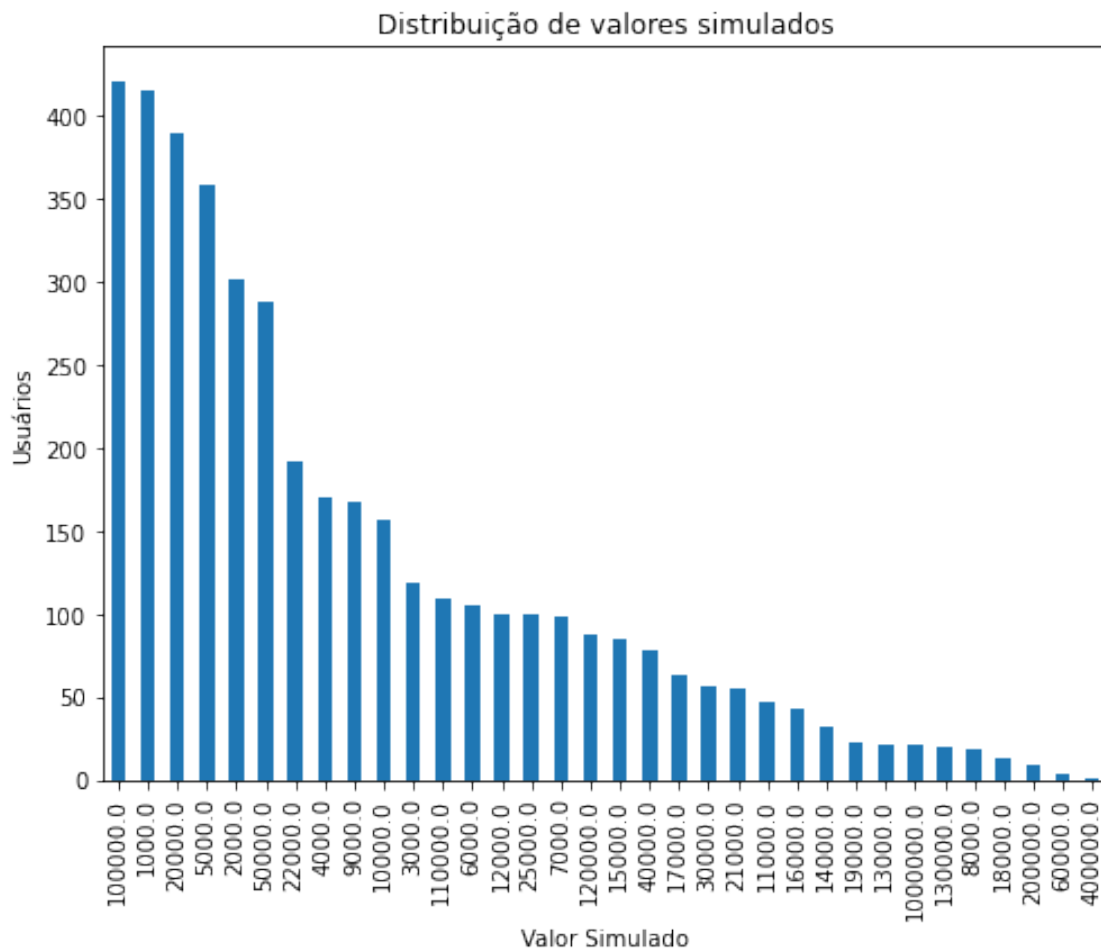
```
[202]: conversion_funnel_unique[conversion_funnel_unique['simulated_investment_range']_
↳ != 'no investment'].groupby(['source', 'simulated_investment_range']).size().
↳ unstack().plot(kind='bar', stacked=True, title='Simulação de valores por_
↳ origem', xlabel='Origem do Tráfego', ylabel='Usuários', figsize=(12, 6))
```

```
[202]: <AxesSubplot:title={'center': 'Simulação de valores por origem'}, xlabel='Origem
do Tráfego', ylabel='Usuários'>
```

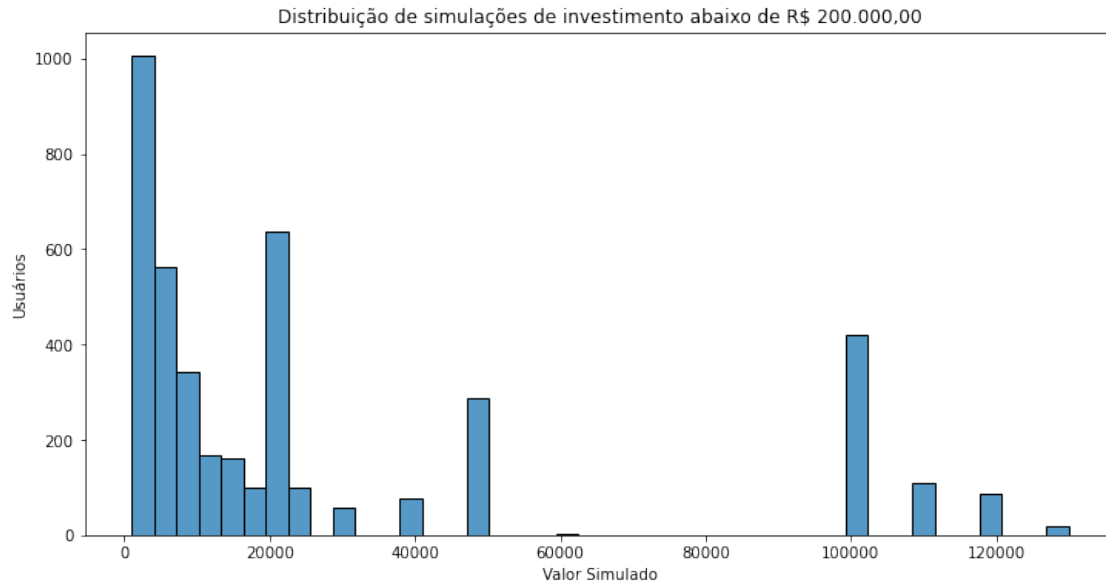


```
[204]: conversion_funnel_unique['simulated_investment'].value_counts().
        ↳plot(kind='bar', title='Distribuição de valores simulados', xlabel='Valor_
        ↳Simulado', ylabel='Usuários', figsize=(8, 6))
```

```
[204]: <AxesSubplot:title={'center': 'Distribuição de valores simulados'}, xlabel='Valor
        Simulado', ylabel='Usuários'>
```



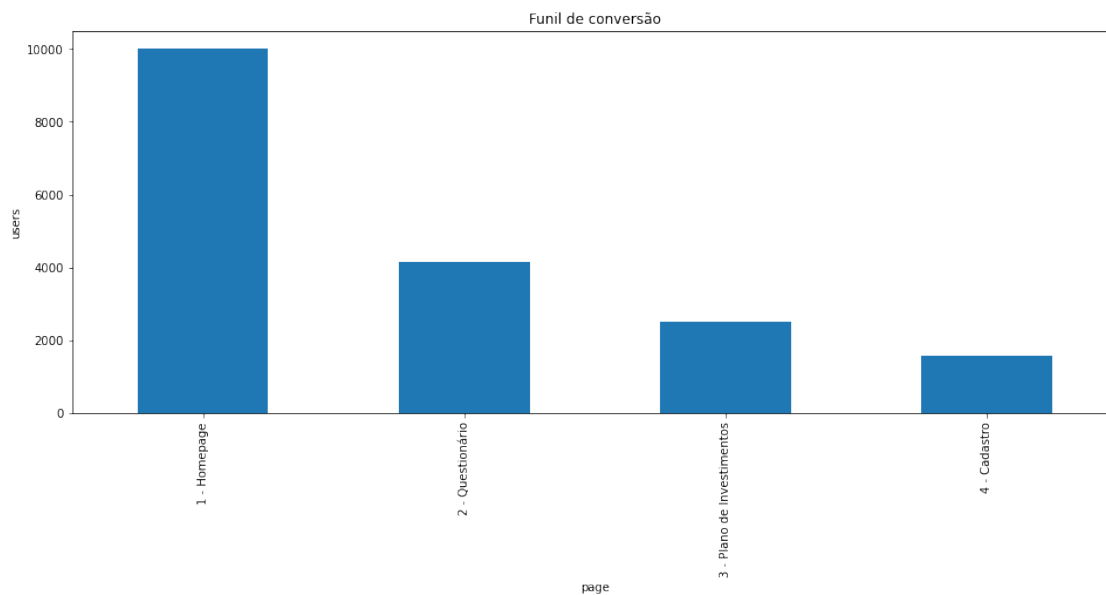
```
[260]: fig, ax = plt.subplots(figsize=(12, 6))
p = sns.histplot(ax=ax,
↳data=conversion_funnel_unique[(conversion_funnel_unique['simulated_investment']
↳notna()) & (conversion_funnel_unique['simulated_investment'] < 200000)],
↳x='simulated_investment').set(title='Distribuição de simulações de
↳investimento abaixo de R\\$ 200.000,00', xlabel='Valor Simulado',
↳ylabel='Usuários')
```

Distribuição de simulações sem a presença de valores extremos. Os valores se concentram entre R\\$ 1.000,00 e R\\$ 25.000,00.

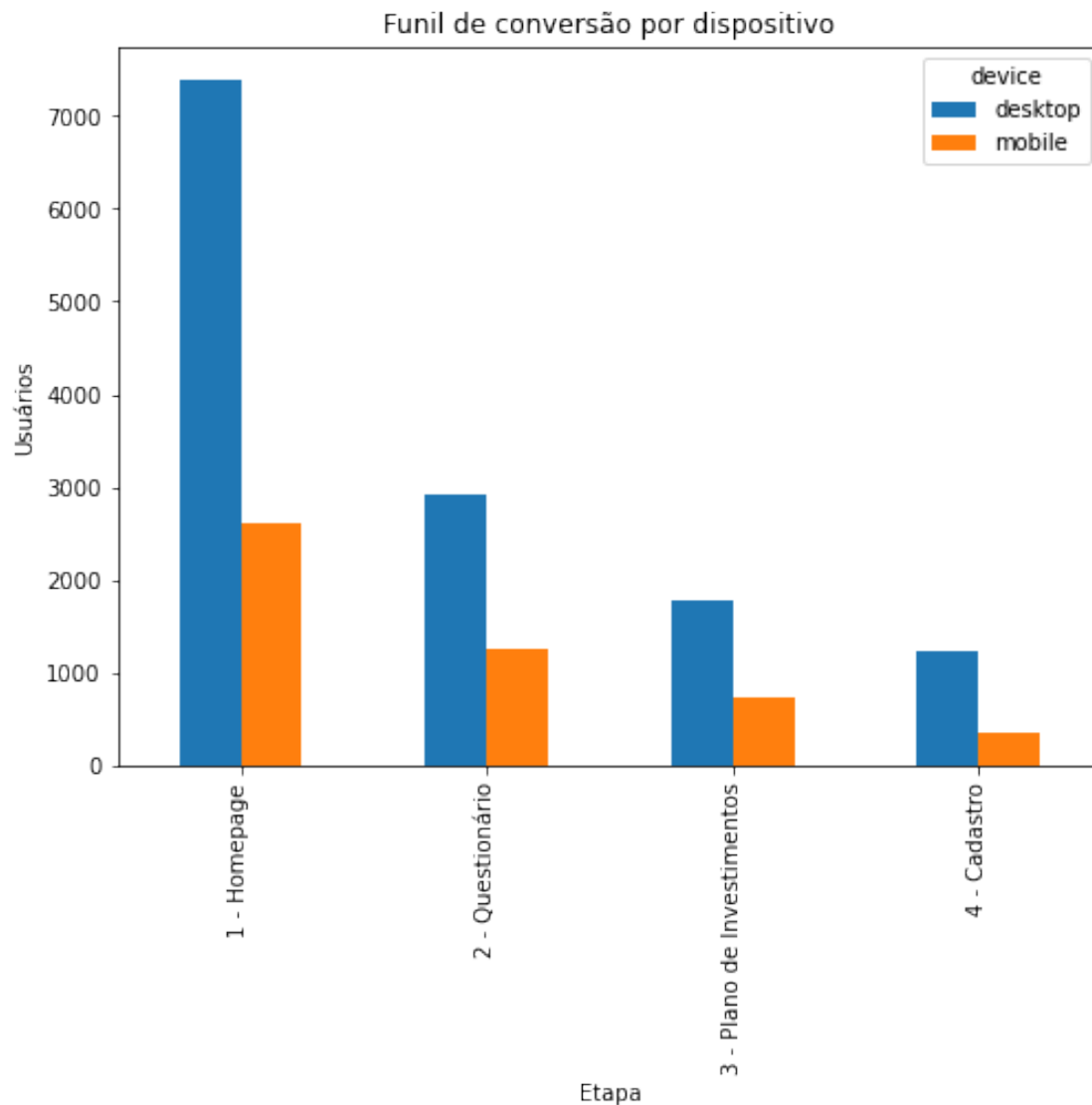
```
[193]: conversion_funnel['page'].value_counts().plot(kind='bar', title='Funil de
↳conversão', xlabel='page', ylabel='users', figsize=(16, 6))
```

```
[193]: <AxesSubplot:title={'center': 'Funil de conversão'}, xlabel='page',
ylabel='users'>
```



```
[194]: conversion_funnel.groupby(['page', 'device']).size().unstack().plot(kind='bar',
↳title='Funil de conversão por dispositivo', xlabel='Etapa',
↳ylabel='Usuários', figsize=(8, 6))
```

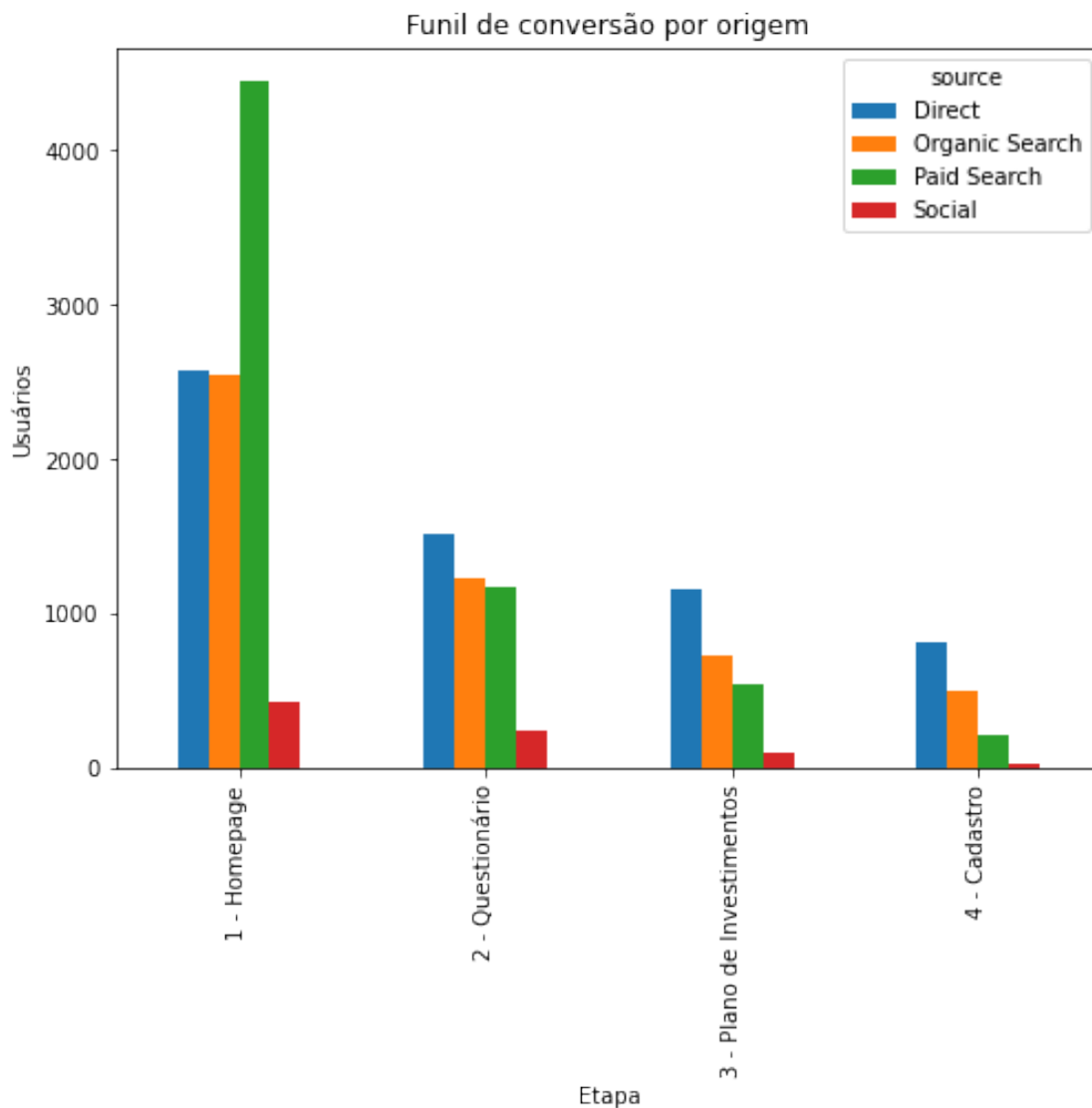
```
[194]: <AxesSubplot:title={'center':'Funil de conversão por dispositivo'},
xlabel='Etapa', ylabel='Usuários'>
```



Visualmente é possível notar que a diferença de cadastros entre dispositivos mobile e desktop é mais acentuada se comparada as etapas de questionário e plano de investimentos.

```
[195]: conversion_funnel.groupby(['page', 'source']).size().unstack().plot(kind='bar',
↳title='Funil de conversão por origem', xlabel='Etapa', ylabel='Usuários',
↳figsize=(8, 6))
```

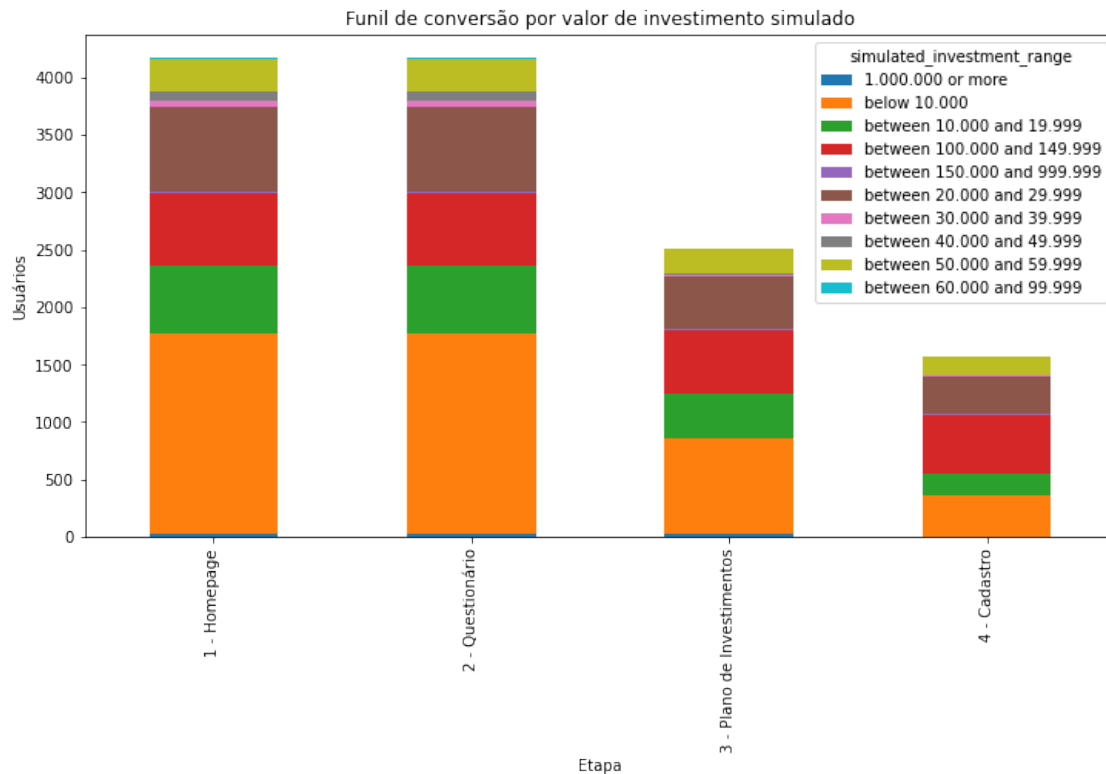
```
[195]: <AxesSubplot:title={'center':'Funil de conversão por origem'}, xlabel='Etapa',
ylabel='Usuários'>
```



A queda de visitas pagas ao longo do funil é mais forte que a queda nas demais origens em todas as etapas do funil.

```
[203]: conversion_funnel[conversion_funnel['simulated_investment_range'] != 'no_
↳investment'].groupby(['page', 'simulated_investment_range']).size().
↳unstack().plot(kind='bar', stacked=True, title='Funil de conversão por valor_
↳de investimento simulado', xlabel='Etapa', ylabel='Usuários', figsize=(12,
↳6))
```

[203]: <AxesSubplot:title={'center': 'Funil de conversão por valor de investimento simulado'}, xlabel='Etapa', ylabel='Usuários'>



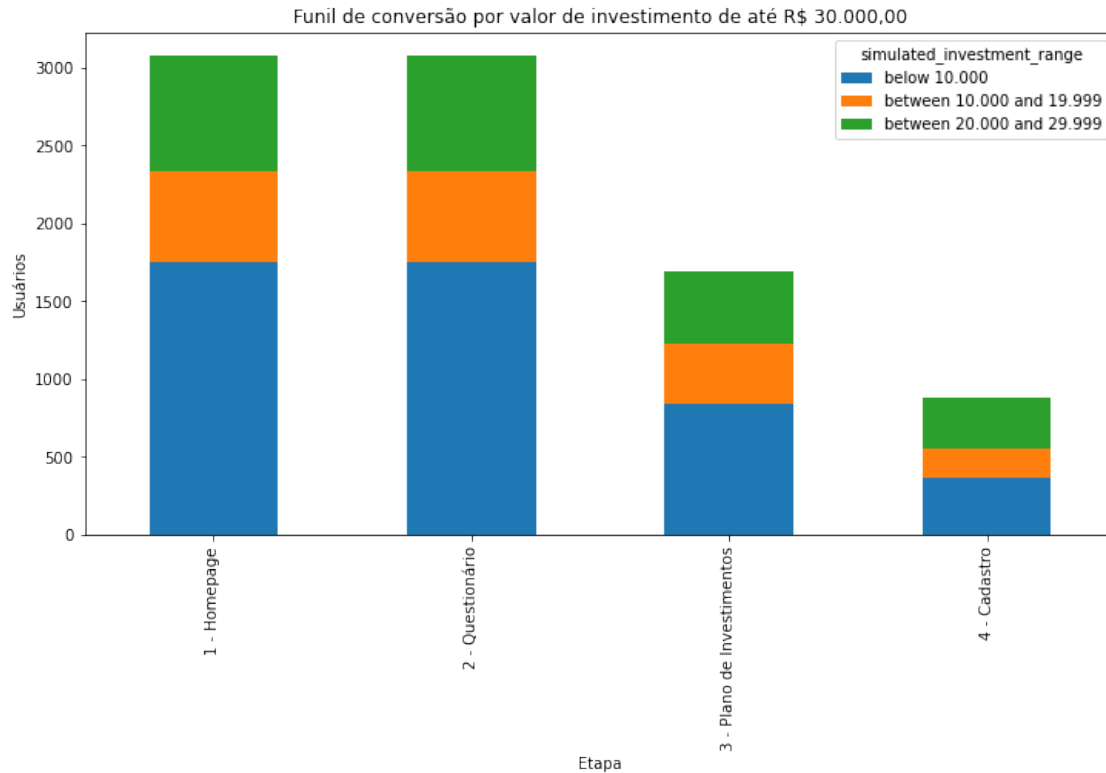
Pessoas que simulam investimentos: - abaixo de R\$ 10.000,00. - entre R\$ 10.000,00 e R\$ 20.000,00. - entre R\$ 20.000,00 e R\$ 30.000,00. - entre R\$ 50.000,00 e R\$ 60.000,00. - entre R\$ 100.000,00 e R\$ 150.000,00.

Compõem praticamente todas as simulações que chegaram até a etapa de cadastro. As maiores proporções de saída ao longo do funil estão nos investimentos simulados de R\$ 30.000,00 ou menos.

[269]:

```
conversion_funnel[(conversion_funnel['simulated_investment'] < 30000) &
↳ (conversion_funnel['simulated_investment'].notna())].groupby(['page',
↳ 'simulated_investment_range']).size().unstack().plot(kind='bar',
↳ stacked=True, title='Funil de conversão por valor de investimento de até
↳ R$ 30.000,00', xlabel='Etapa', ylabel='Usuários', figsize=(12, 6))
```

[269]: <AxesSubplot:title={'center': 'Funil de conversão por valor de investimento de até R\$ 30.000,00'}, xlabel='Etapa', ylabel='Usuários'>



```
[275]: conversion_funnel[(conversion_funnel['simulated_investment'] < 30000) &
↳ (conversion_funnel['simulated_investment'].notna())].groupby(['page']).
↳ agg({'id': ['count']})
```

```
[275]:
```

	id
page	count
1 - Homepage	3073
2 - Questionário	3073
3 - Plano de Investimentos	1691
4 - Cadastro	883

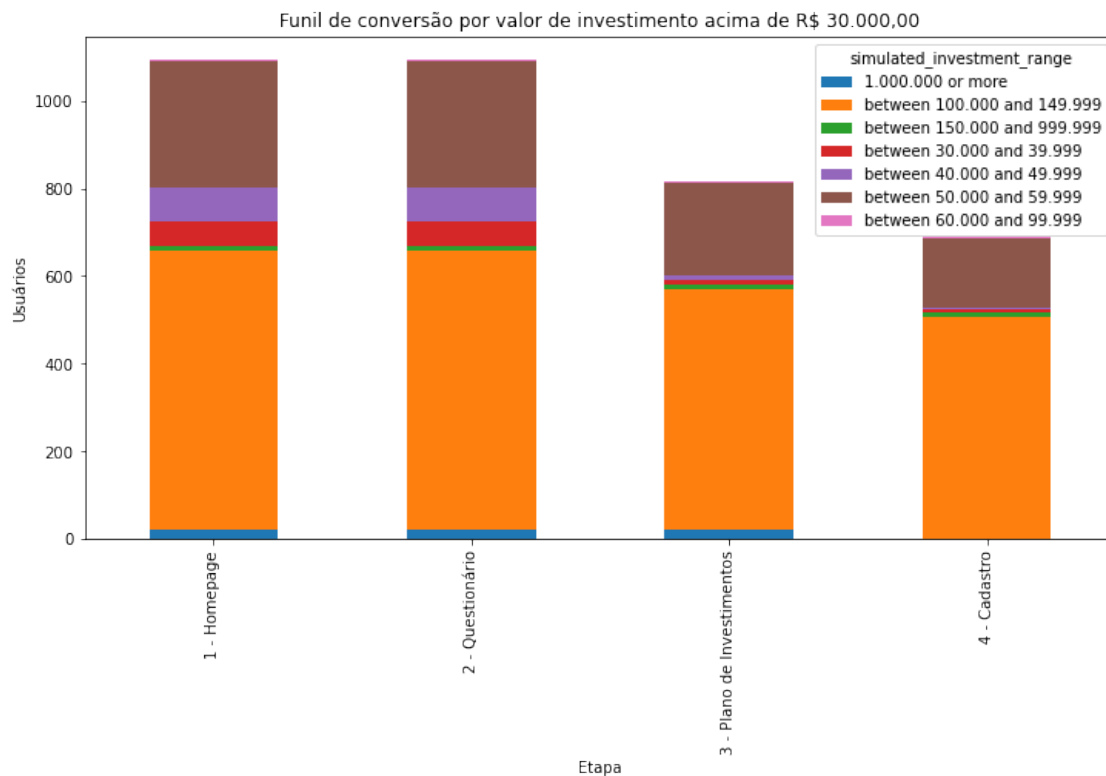
```
[276]: conversion_funnel[(conversion_funnel['simulated_investment'] >= 30000) &
↳ (conversion_funnel['simulated_investment'].notna())].groupby(['page']).
↳ agg({'id': ['count']})
```

```
[276]:
```

	id
page	count
1 - Homepage	1093
2 - Questionário	1093
3 - Plano de Investimentos	815

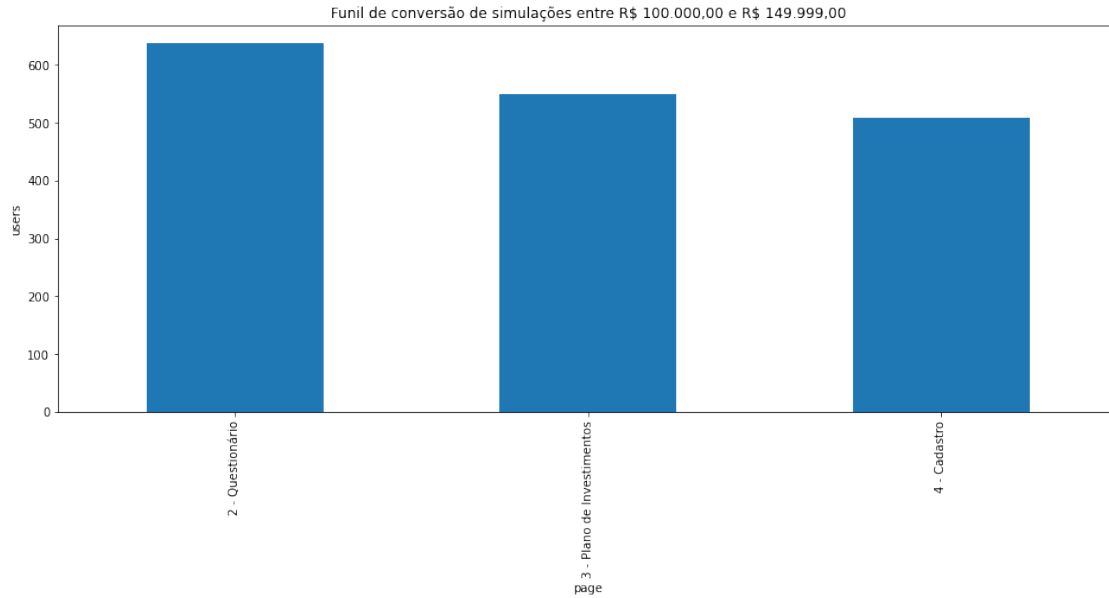
```
[268]: conversion_funnel[(conversion_funnel['simulated_investment'] >= 30000) &
↳ (conversion_funnel['simulated_investment'].notna())].groupby(['page',
↳ 'simulated_investment_range']).size().unstack().plot(kind='bar',
↳ stacked=True, title='Funil de conversão por valor de investimento acima de
↳ R\\$ 30.000,00', xlabel='Etapa', ylabel='Usuários', figsize=(12, 6))
```

```
[268]: <AxesSubplot:title={'center': 'Funil de conversão por valor de investimento acima
de R\\$ 30.000,00'}, xlabel='Etapa', ylabel='Usuários'>
```



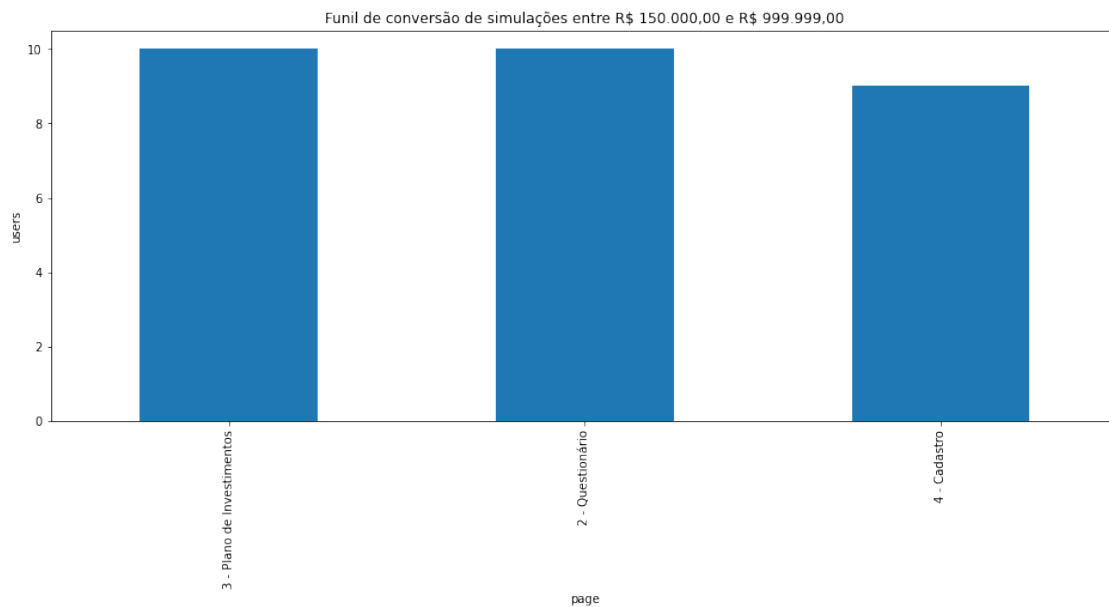
```
[197]: conversion_funnel[(conversion_funnel['simulated_investment_range'] == 'between
↳ 100.000 and 149.999') & (conversion_funnel['page'] != '1 -
↳ Homepage')]['page'].value_counts().plot(kind='bar', title='Funil de
↳ conversão de simulações entre R\\$ 100.000,00 e R\\$ 149.999,00',
↳ xlabel='page', ylabel='users', figsize=(16, 6))
```

```
[197]: <AxesSubplot:title={'center': 'Funil de conversão de simulações entre R\\$
100.000,00 e R\\$ 149.999,00'}, xlabel='page', ylabel='users'>
```



```
[198]: conversion_funnel[(conversion_funnel['simulated_investment_range'] == 'between_
↳150.000 and 999.999') & (conversion_funnel['page'] != '1 -_
↳Homepage')]['page'].value_counts().plot(kind='bar', title='Funil de_
↳conversão de simulações entre R\\$ 150.000,00 e R\\$ 999.999,00',_
↳xlabel='page', ylabel='users', figsize=(16, 6))
```

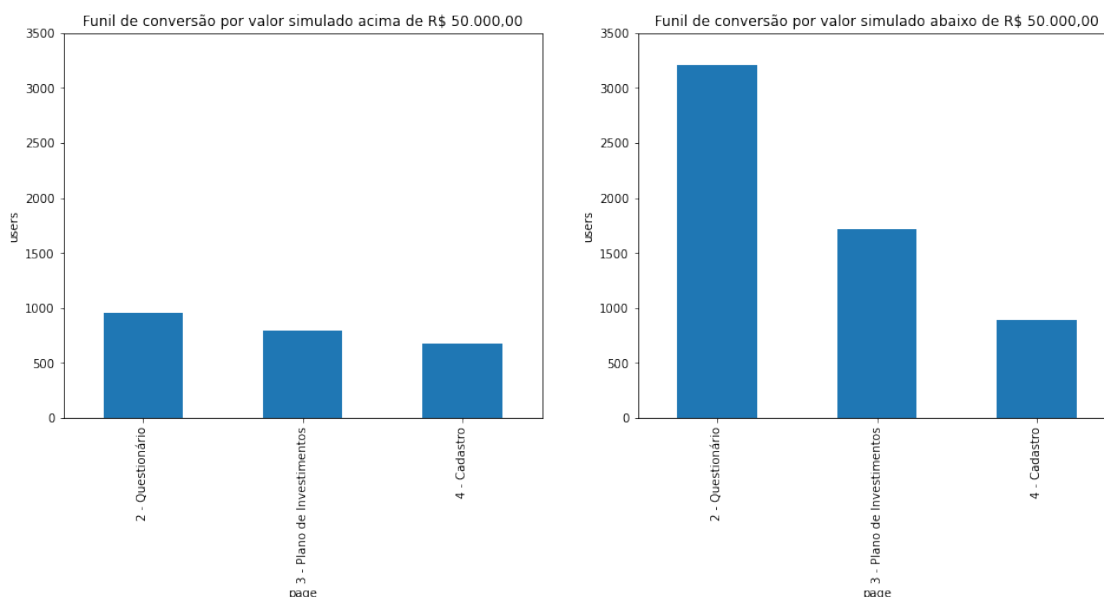
```
[198]: <AxesSubplot:title={'center': 'Funil de conversão de simulações entre R\\$
150.000,00 e R\\$ 999.999,00'}, xlabel='page', ylabel='users'>
```



As taxas de conversão de quem simula investimentos acima de R\$ 150.000,00 são bem maiores que as taxas de todas as pessoas que passaram pelo funil.

```
[199]: fig, ax = plt.subplots(ncols=2, figsize=(16, 6))
conversion_funnel[(conversion_funnel['simulated_investment'] >= 50000) &
↳ (conversion_funnel['page'] != '1 - Homepage')]['page'].value_counts().
↳ plot(ax=ax[0], kind='bar', title='Funil de conversão por valor simulado
↳ acima de R\\$ 50.000,00', xlabel='page', ylabel='users', yticks=(0, 500,
↳ 1000, 1500, 2000, 2500, 3000, 3500))
conversion_funnel[(conversion_funnel['simulated_investment'] < 50000) &
↳ (conversion_funnel['page'] != '1 - Homepage')]['page'].value_counts().
↳ plot(ax=ax[1], kind='bar', title='Funil de conversão por valor simulado
↳ abaixo de R\\$ 50.000,00', xlabel='page', ylabel='users', yticks=(0, 500,
↳ 1000, 1500, 2000, 2500, 3000, 3500))
```

```
[199]: <AxesSubplot:title={'center': 'Funil de conversão por valor simulado abaixo de
R\\$ 50.000,00'}, xlabel='page', ylabel='users'>
```



As taxas de conversão caem substancialmente quando o valor simulado é menor do que R\$ 50.000,00.

1.5 Insights

Perguntas iniciais: - Há alguma diferença significativa entre cada etapa do funil por dispositivo? E por origem? *A composição de pessoas usuárias por dispositivo no funil é de cerca de 70% desktop e 30% mobile. Esta proporção se mantém similar ao longo das etapas do funil e muda para 80% desktop e 20% mobile na última etapa.* - O valor investido muda

significativamente de acordo com a origem ou dispositivo? ***O valor médio da simulação cresce a cada etapa do funil.*** - Há alguma hora do dia ou dia da semana mais comum para cadastros? ***Este dado não existe na base.*** - Qual é a faixa de valores com maior número de simulações de investimento? ***Entre R\\$ 100.000,00 e R\\$ 150.000,00 é a faixa que mais conclui cadastros. Nas demais etapas do funil os valores simulados com maior frequência estão abaixo de R\\$ 10.000,00.*** - Qual é o perfil de investidor que mais faz cadastros? E quais os objetivos definidos? E o perfil / objetivo com menor taxa de conversão? ***Este dado não existe na base.***

Descobertas: - A taxa de conversão em cadastros é maior no desktop. - Mais pessoas que entram por dispositivos móveis iniciam o questionário. - Há mais simulações de valores muito acima da média em dispositivos móveis, mas quando os valores extremos são retirados, os valores da simulação não diferem muito por dispositivo. - A quantidade de entradas no site por mídia paga é a maior mas é uma das que menos chegam até a etapa de cadastro. - As pessoas que chegam até o final do funil estão dispostas a investir valores maiores do que quem fica nas demais etapas. - O tráfego que entra por buscadores e faz o cadastro investe mais, em média, do que simulou no questionário. Em todas as origens a maior mediana é a de quem faz o cadastro. - O menor valor simulado foi de R\\$ 1.000,00 e o maior de R\\$ 1.000.000,00. A mediana é R\\$ 14.000,00 e a média é R\\$ 33.584,73. Há outliers. - 73,78% das visitas vieram de desktop e 26,22% de mobile. - 43,67% de visitas por mídia impulsionada, 27,09% por tráfego direto, 25,25% por busca orgânica e 3,99% por social. - As 21 simulações de investimento de R\\$ 1.000.000,00 aconteceram em dispositivos mobile. - Os seis valores simulados mais comuns representam 52,18% de todas as simulações. - As taxas de conversão caem mais agressivamente quando o valor simulado está abaixo de R\\$ 50.000,00. - A maioria dos visitantes de redes sociais iniciam o questionário mas não o termina. - Visitas por tráfego direto tendem a simular quantias maiores para investimento.

Novas perguntas: - A taxa de conversão entre os passos 1 e 2 é maior no mobile se comparada ao desktop, mas o abandono do questionário é maior no mobile. Há algum problema de usabilidade na experiência mobile? - Quanto maior o valor de simulação mediano, maior a probabilidade de conversão no desktop? - O volume de tráfego que entra por mídia paga no desktop é quase 4x maior que o volume de tráfego de mídia paga mobile, no entanto, o número de finalizações de funil é apenas o dobro do mobile. O tráfego de mídia impulsionada está qualificado? - Simulações abaixo de R\\$ 50.000,00 possuem taxas de conversão menores ao longo do funil. Há algo que não está tornando investimentos de baixo valor atraentes? - A comunicação nas redes sociais está condizente com o conteúdo encontrado no site? - Em quais questões do questionário há maior taxa de abandono?