

```
In [1]: import pandas as pd
import sqlite3 as sql
import plotly.express as px
import os
```

```
In [2]: #Connecting to database with sqlite3
database = "college.db"
connection = sql.connect(database)
```

```
In [3]: pd.read_sql('select * from college;', connection)
```

Out[3]:

	index	Rank	Name	City	State	Public/Private	Undergraduate Population	Student Population	Net Price	Average Grant Aid	Total Annual Cost	Alumni Salary	Acceptance Rate	SAT Lower	SAT Upper	ACT Lower	ACT Upper	Website	
	0	0	1.0	Harvard University	Cambridge	MA	Private	13844.0	31120.0	14327.0	49870.0	69600.0	146800.0	5.0	1460.0	1590.0	32.0	35.0	www.harvard.edu
	1	1	2.0	Stanford University	Stanford	CA	Private	8402.0	17534.0	13261.0	50134.0	69109.0	145200.0	5.0	1390.0	1540.0	32.0	35.0	www.stanford.edu
	2	2	3.0	Yale University	New Haven	CT	Private	6483.0	12974.0	18627.0	50897.0	71290.0	138300.0	7.0	1460.0	1580.0	32.0	35.0	www.yale.edu
	3	3	4.0	Massachusetts Institute of Technology	Cambridge	MA	Private	4680.0	11466.0	20771.0	43248.0	67430.0	155200.0	7.0	1490.0	1570.0	33.0	35.0	web.mit.edu
	4	4	5.0	Princeton University	Princeton	NJ	Private	5659.0	8273.0	9327.0	48088.0	66150.0	139400.0	6.0	1430.0	1570.0	31.0	35.0	www.princeton.edu

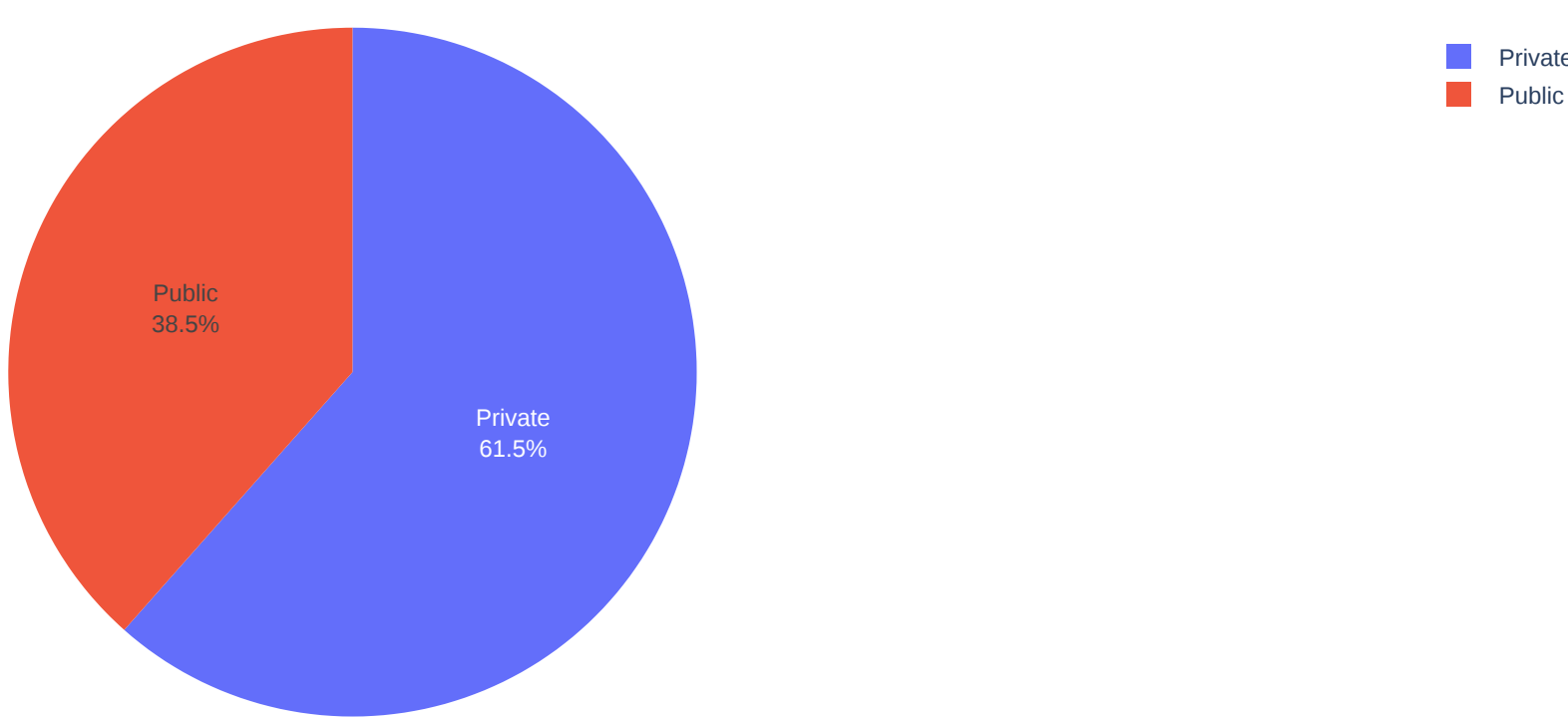
	645	645	646.0	New Mexico State University	Las Cruces	NM	Public	13379.0	14432.0	8625.0	9582.0	34720.0	96700.0	64.0	910.0	1160.0	18.0	23.0	www.nmsu.edu
	646	646	647.0	Indiana State University	Terre Haute	IN	Public	13626.0	13763.0	13012.0	9297.0	32938.0	85600.0	85.0	900.0	1110.0	17.0	23.0	www.indstate.edu
	647	647	648.0	Emory & Henry College	Emory	VA	Private	1094.0	1226.0	19340.0	27155.0	48100.0	70700.0	72.0	988.0	1170.0	19.0	25.0	www.ehc.edu
	648	648	649.0	Wells College	Aurora	NY	Private	488.0	516.0	22828.0	30207.0	55180.0	NaN	80.0	NaN	NaN	NaN	NaN	www.wells.edu
	649	649	650.0	Catawba College	Salisbury	NC	Private	1336.0	1463.0	19807.0	23662.0	44921.0	78100.0	42.0	930.0	1100.0	18.0	23.0	www.catawba.edu

650 rows x 19 columns

650 rows x 18 columns

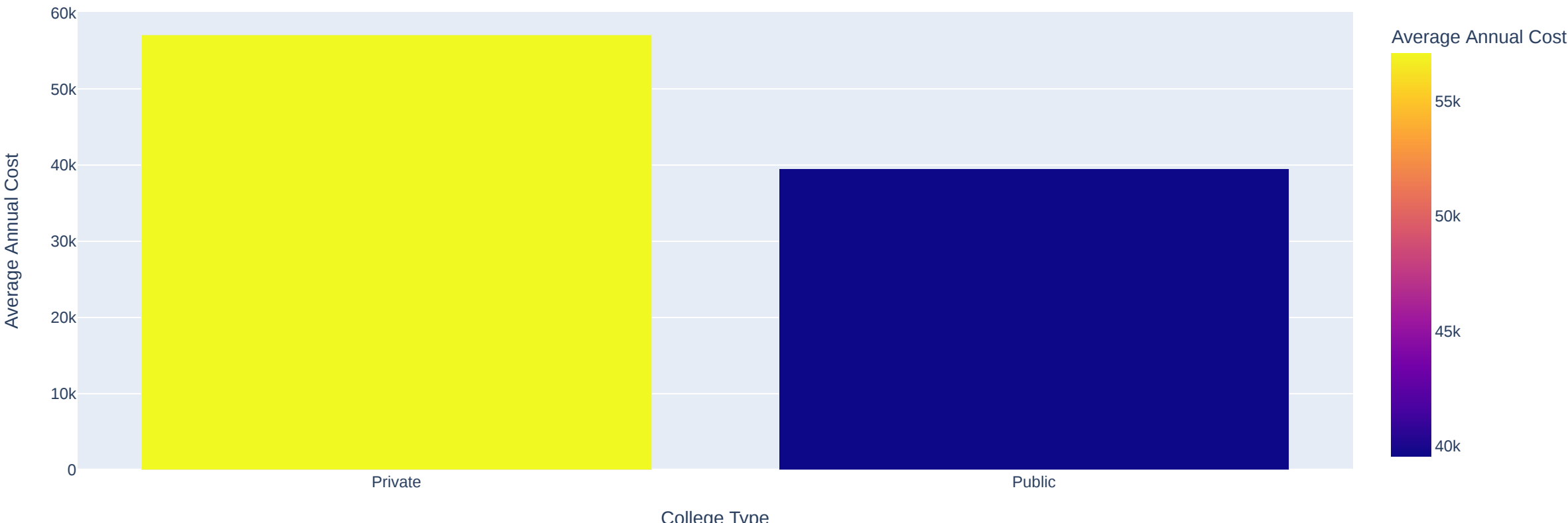
```
In [4]: pr_pu = pd.read_sql('select "Public/Private" as "College Type", count(*) as "Colleges in list" from college group by "College Type";', connection)
fig_1 = px.pie(pr_pu, values="Colleges in list", names="College Type", title="Private Colleges vs Public Colleges")
fig_1.update_traces(textposition="inside", textinfo="label+percent")
fig_1.show()
```

Private Colleges vs Public Colleges



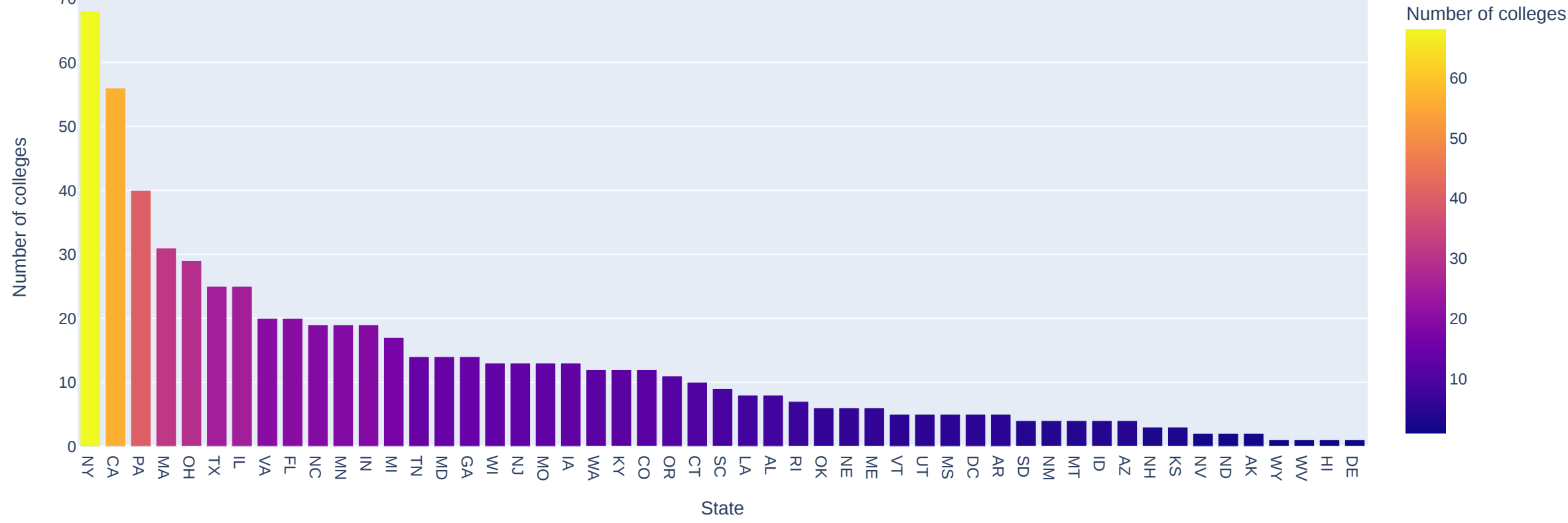
```
In [5]: # Average cost of studying in Public vs Private
cost = pd.read_sql('select "Public/Private" as "College Type", avg("Total Annual Cost") as "Average Annual Cost" from college group by "College Type";', connection)
px.bar(cost, x="College Type", y="Average Annual Cost", title="Average Cost of study Publivs vs Private", color="Average Annual Cost")
```

Average Cost of study Publivs vs Private



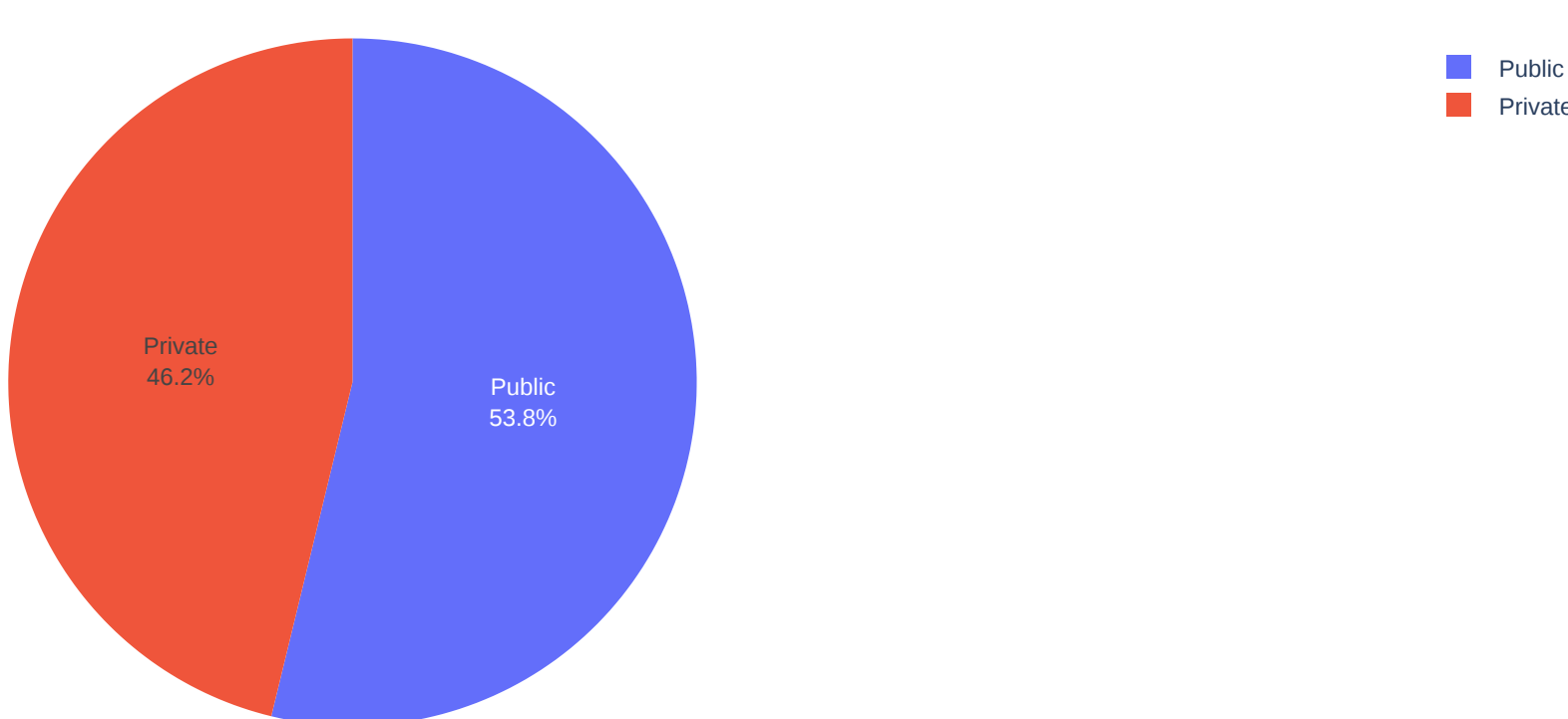
```
In [6]: # Which States have maximum number of colleges
state = pd.read_sql('select State, count(*) as "Number of colleges" from college group by "State" order by "Number of colleges" desc;', connection)
px.bar(state, x="State", y="Number of colleges", title="Number of colleges for each state", color="Number of colleges")
```

Number of colleges for each state



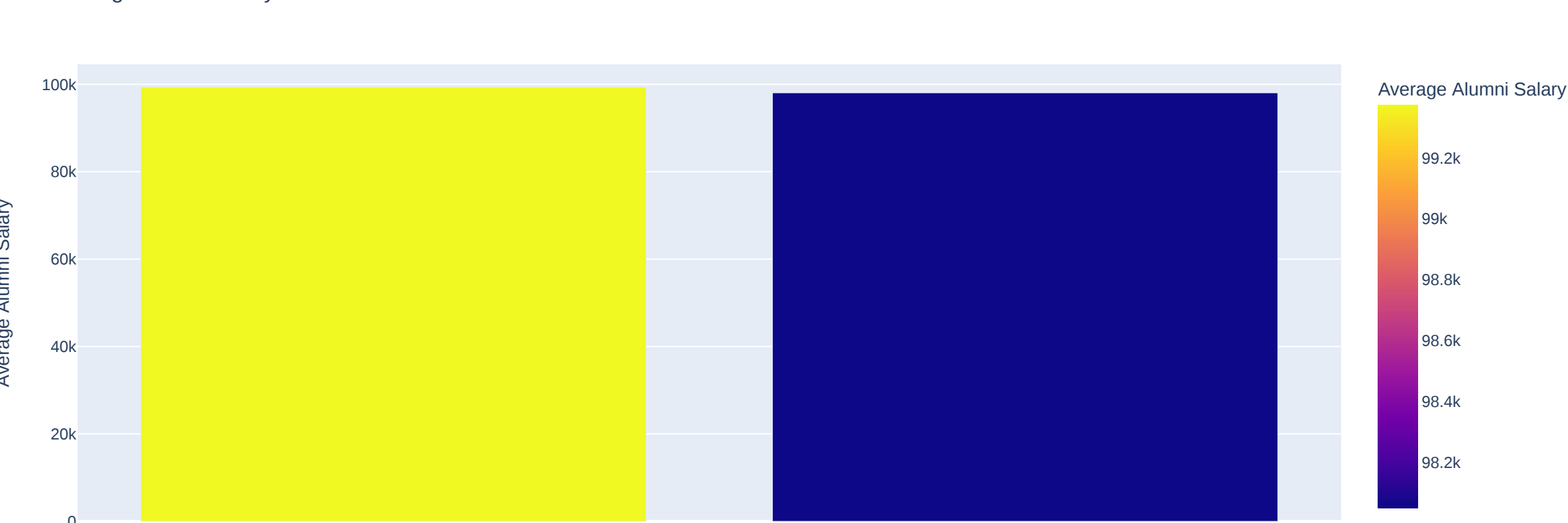
```
In [7]: # Average acceptance rate Public Vs Private
acc = pd.read_sql('select "Public/Private" as "College type", avg("Acceptance Rate") as "Average acceptance rate" from college group by "College Type";', connection)
fig_2 = px.pie(acc, values="Average acceptance rate", names="College type", title="Average acceptance rate Public Vs Private")
fig_2.update_traces(textposition="inside", textinfo="label+percent")
fig_2.show()
```

Average acceptance rate Public Vs Private



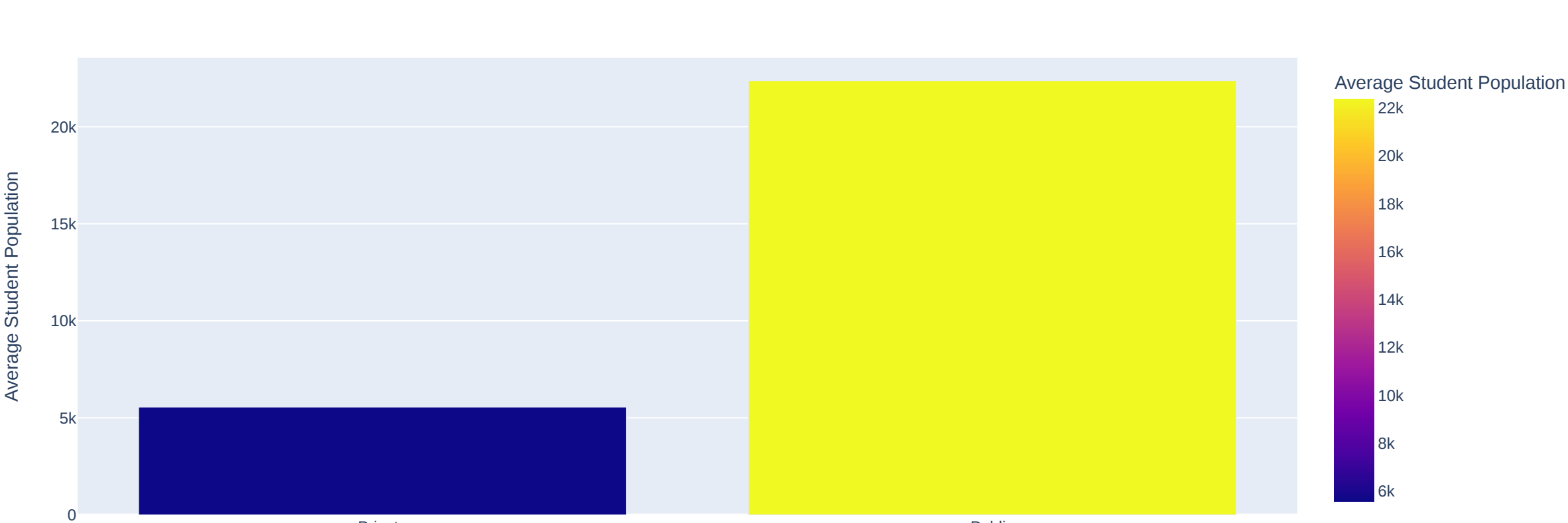
```
In [8]: # Average alumni salary Public Vs Private
alu_sal = pd.read_sql('select "Public/Private" as "College type", avg("Alumni Salary") as "Average Alumni Salary" from college group by "College type";', connection)
px.bar(alu_sal, y="Average Alumni Salary", x="College type", title="Average alumni salary Public Vs Private", color="Average Alumni Salary")
```

Average alumni salary Public Vs Private



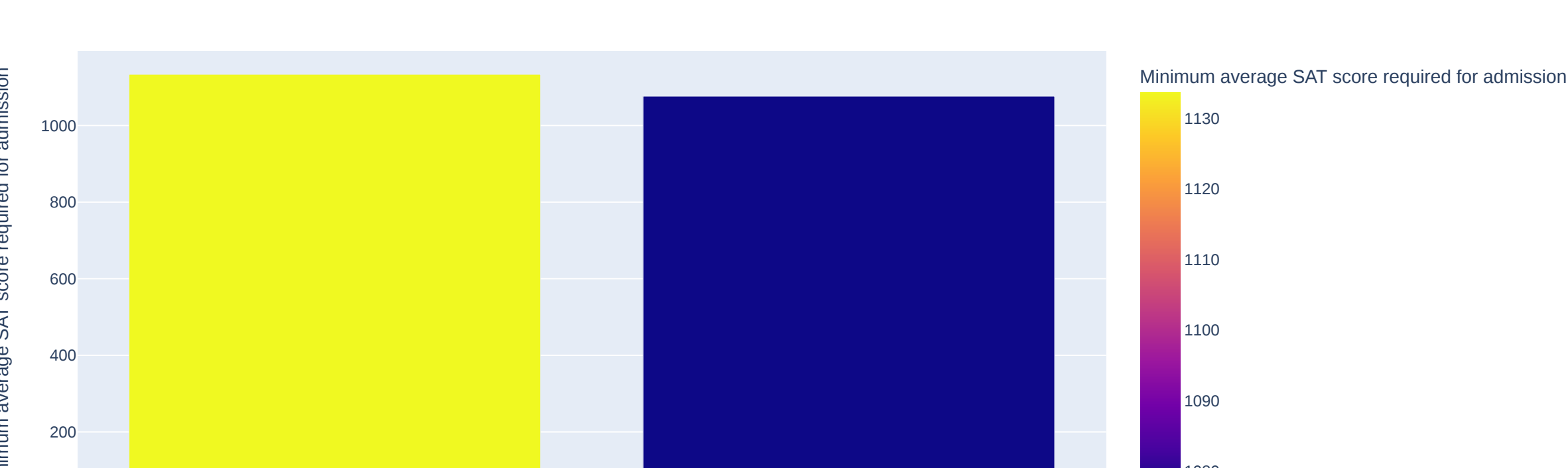
```
In [9]: # Student population Public Vs Private
stu_pop = pd.read_sql('select "Public/Private" as "College type", avg("Student Population") as "Average Student Population" from college group by "College type";', connection)
px.bar(stu_pop, x="College type", y="Average Student Population", title="Average Student Population Public Vs Private", color="Average Student Population")
```

Average Student Population Public Vs Private



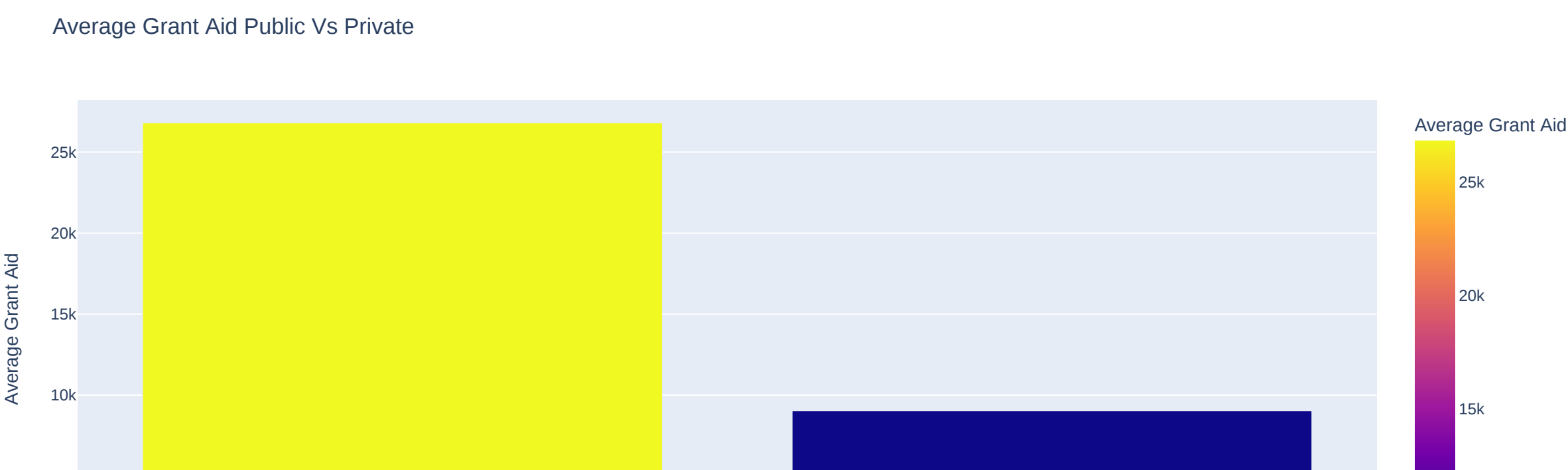
```
In [10]: # Minimum average SAT score for admission Public Vs Private
sat = pd.read_sql('select "Public/Private" as "College type", avg("SAT Lower") as "Minimum average SAT score required for admission" from college group by "College type";', connection)
px.bar(sat, x="College type", y="Minimum average SAT score required for admission", color="Minimum average SAT score required for admission", title="Minimum average SAT score required for admission Public Vs Private")
```

Minimum average SAT score required for admission Public Vs Private



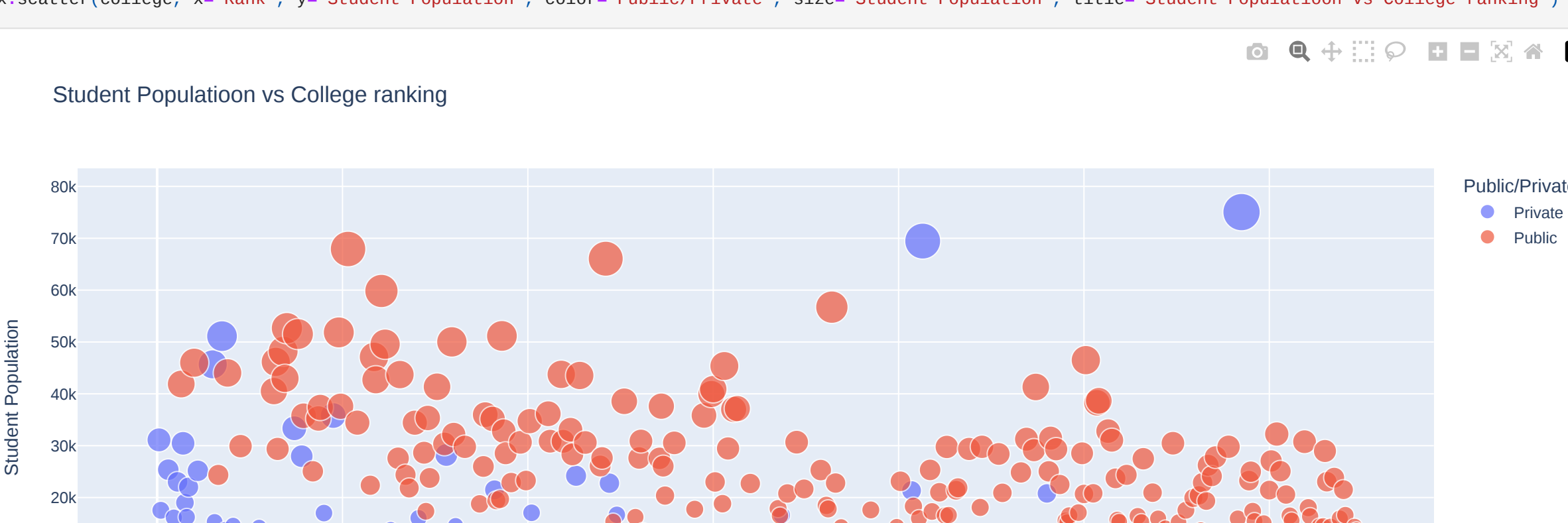
```
In [11]: # Average aid grant Public vs Private
aid = pd.read_sql('select "Public/Private" as "College type", avg("Average Grant Aid") as "Average Grant Aid" from college group by "College type";', connection)
px.bar(aid, x="College type", y="Average Grant Aid", color="Average Grant Aid", title="Average Grant Aid Public Vs Private")
```

Average Grant Aid Public Vs Private



```
In [12]: # Student vs College rank (Concentration of students)
college = pd.read_sql('select * from college;', connection)
px.scatter(college, x="Rank", y="Student Population", color="Public/Private", size="Student Population", title="Student Population vs College ranking")
```

Student Population vs College ranking



```
In [13]: # Closing the connection
connection.close()
```

```
In [ ]:
```