International Breweries Analysis

Description

International Breweries Plc manufactures beer and non-alcoholic malt drinks. Its activities include the brewing, packaging, and marketing of beer, alcoholic flavoured and non-flavoured alcoholic beverages, and soft drinks. The company was founded by Lawrence Omole on December 22, 1971 and is headquartered in Lagos, Nigeria.

Project Aim:

In [ ]:
```
from sqlalchemy import create_engine
engine = create_engine('postgresql://postgres:28_March@localhost/international_breweries')
%reload_ext sql
%sql $engine.url
```

Drop Columns if the exist already

In [ ]:
```
%%sql
ALTER TABLE breweries_sales
DROP COLUMN total_cost,
DROP COLUMN total_sale,
DROP COLUMN profit,
DROP COLUMN roi,
DROP COLUMN shortMnthName;
```

```
 * postgresql://postgres:***@localhost/international_breweries
Done.
```
Out[ ]:
```
[]
```

Creating new columns

In [ ]:
```
%%sql
ALTER TABLE breweries_sales
ADD COLUMN total_cost bigint,
ADD COLUMN total_sale bigint,
ADD COLUMN profit bigint,
ADD COLUMN roi NUMERIC(10, 5),
ADD COLUMN shortMnthName character(3);
```

```
 * postgresql://postgres:***@localhost/international_breweries
Done.
```
Out[ ]:
```
[]
```

Populating new columns with values

In [ ]:
```
%%sql
UPDATE breweries_sales
SET total_cost = unit_cost * CAST(quantity_sold AS INTEGER),
    total_sale = unit_price * CAST(quantity_sold AS INTEGER),
    profit = (unit_price * CAST(quantity_sold AS INTEGER)) - (unit_cost * CAST(quantity_sold AS INTEGER)),
    roi = (((unit_price * CAST(quantity_sold AS INTEGER)) - (unit_cost * CAST(quantity_sold AS INTEGER)))/(unit_cost * CAST(quantity_sold A
    shortMnthName = LEFT(month, 3);
```

```
 * postgresql://postgres:***@localhost/international_breweries
1047 rows affected.
```
Out[ ]:
```
[]
```

In [ ]:
```
%%sql

-- Preview

SELECT *
FROM breweries_sales
LIMIT 3
```

```
 * postgresql://postgres:***@localhost/international_breweries
3 rows affected.
```
Out[ ]:

sales_id	region	month	year	sales_rep	brand	unit_cost	unit_price	quantity_sold	country	total_cost	total_sale	profit	roi	shortmnthname
10234	west	February	2019	Howard	trophy	150	200	840	Benin	126000	168000	42000	33.33333	Feb
10236	northwest	April	2019	Gill	castle lite	180	450	832	Ghana	149760	374400	224640	150.00000	Apr
10239	Southeast	July	2019	Morgan	beta malt	80	150	774	Benin	61920	116100	54180	87.50000	Jul

In [ ]:
```
%%sql
SELECT DISTINCT(country) unique_countries
FROM breweries_sales
```

```
 * postgresql://postgres:***@localhost/international_breweries
5 rows affected.
```

Out[ ]:

unique_countries
Nigeria
Benin
Senegal
Ghana
Togo

In [ ]:
```
%%sql
SELECT DISTINCT(region) unique_regions
FROM breweries_sales
```

```
 * postgresql://postgres:***@localhost/international_breweries
6 rows affected.
```

Out[ ]:

unique_regions
southsouth
west
northcentral
northeast
Southeast
northwest

In [ ]:
```
%%sql
SELECT DISTINCT(sales_rep) unique_countries
FROM breweries_sales
```

```
 * postgresql://postgres:***@localhost/international_breweries
11 rows affected.
```

Out[ ]:

unique_countries
Morgan
Jones
Gill
Sorvino
Thompson
Kivell
Jardine
Andrews
Smith
Parent
Howard

In [ ]:
```
%%sql
SELECT DISTINCT(brand) unique_countries
FROM breweries_sales
```

```
 * postgresql://postgres:***@localhost/international_breweries
7 rows affected.
```

Out[ ]:

unique_countries
beta malt
eagle lager
grand malt
hero
budweiser
castle lite
trophy

In [ ]:
```
import pandas as pd
import plotly.express as px
import plotly.offline as po
import plotly.graph_objects as go
import plotly.io as pio
pio.templates.default= 'plotly_dark'
po.init_notebook_mode(connected=True)

# ["plotly", "plotly_white", "plotly_dark", "ggplot2", "seaborn", "simple_white", "none"]
```

In [ ]:
```
query = '''
SELECT
    year,
    SUM(quantity_sold) total_qty,
    SUM(total_cost) total_cost,
    SUM(total_sale) total_sale,
```

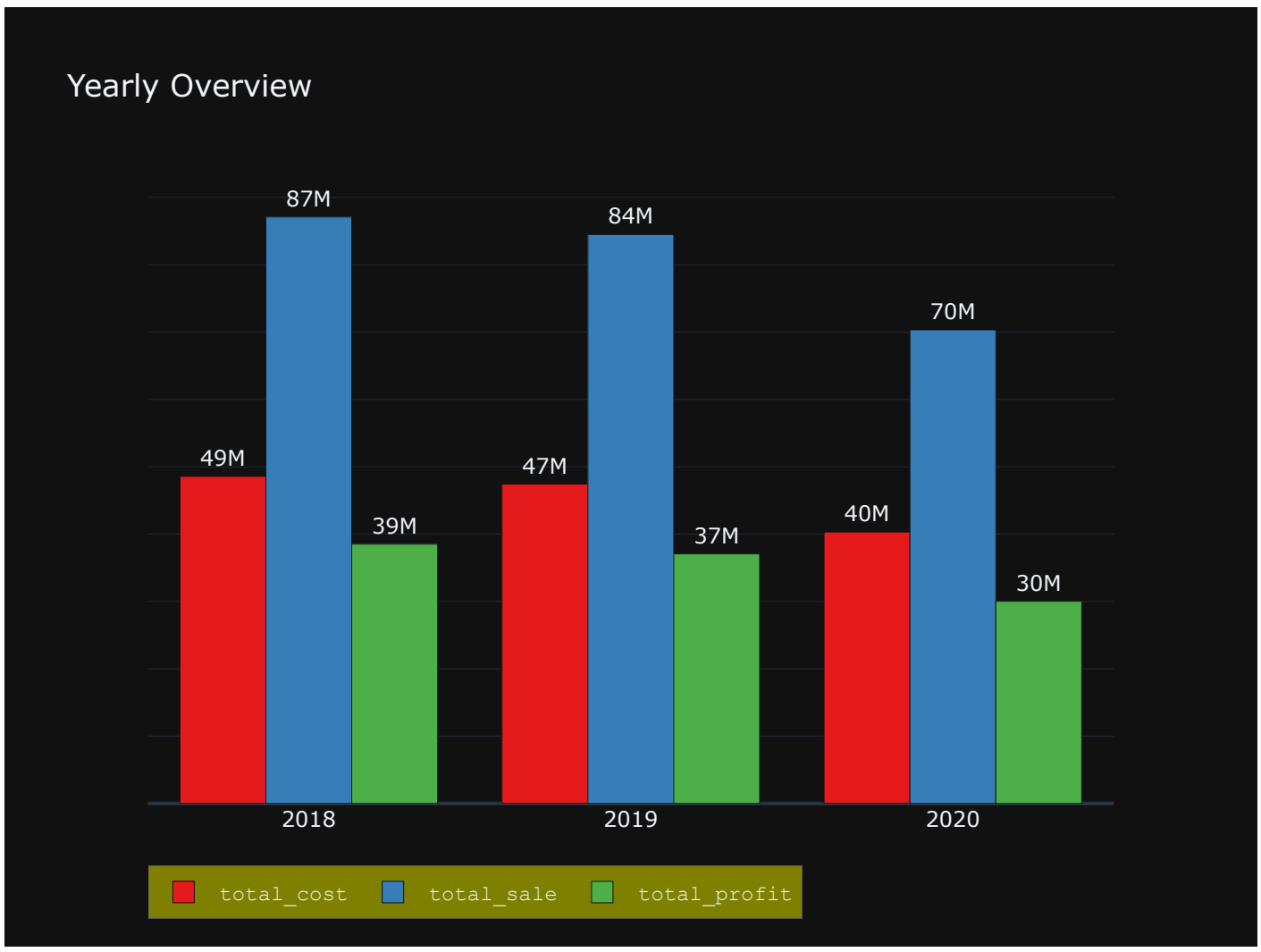
```
    SUM(profit) total_profit,
    AVG(roi) roi

FROM
    breweries_sales
GROUP BY
    year
'''
year_sum = pd.read_sql_query(query, engine)
year_sum
```

Out[ ]:

	year	total_qty	total_cost	total_sale	total_profit	roi
0	2018	317563	48577780.0	87081100.0	38503320.0	75.632404
1	2019	305409	47422350.0	84486200.0	37063850.0	74.662308
2	2020	267750	40305100.0	70325350.0	30020250.0	71.346237

In [ ]:

```
year_sum.year = year_sum.year.astype('str')

fig = px.bar(year_sum, 'year', ['total_cost', 'total_sale', 'total_profit'], width=700,
             barmode='group', text_auto='.2s', labels={'year':'', 'value':''},
             title='Yearly Overview', color_discrete_sequence=px.colors.qualitative.Set1)
fig.update_traces(textposition='outside', cliponaxis=False)
fig.update_yaxes(showticklabels=False)
fig.update_layout(legend = dict(
                        orientation='h',
                        title='',
                        font = dict(
                        family="Courier",
                        size=12,
                        ),
                        bgcolor='olive',
                        bordercolor='blue',
                        borderwidth=.5
                    ))
```



In [ ]:

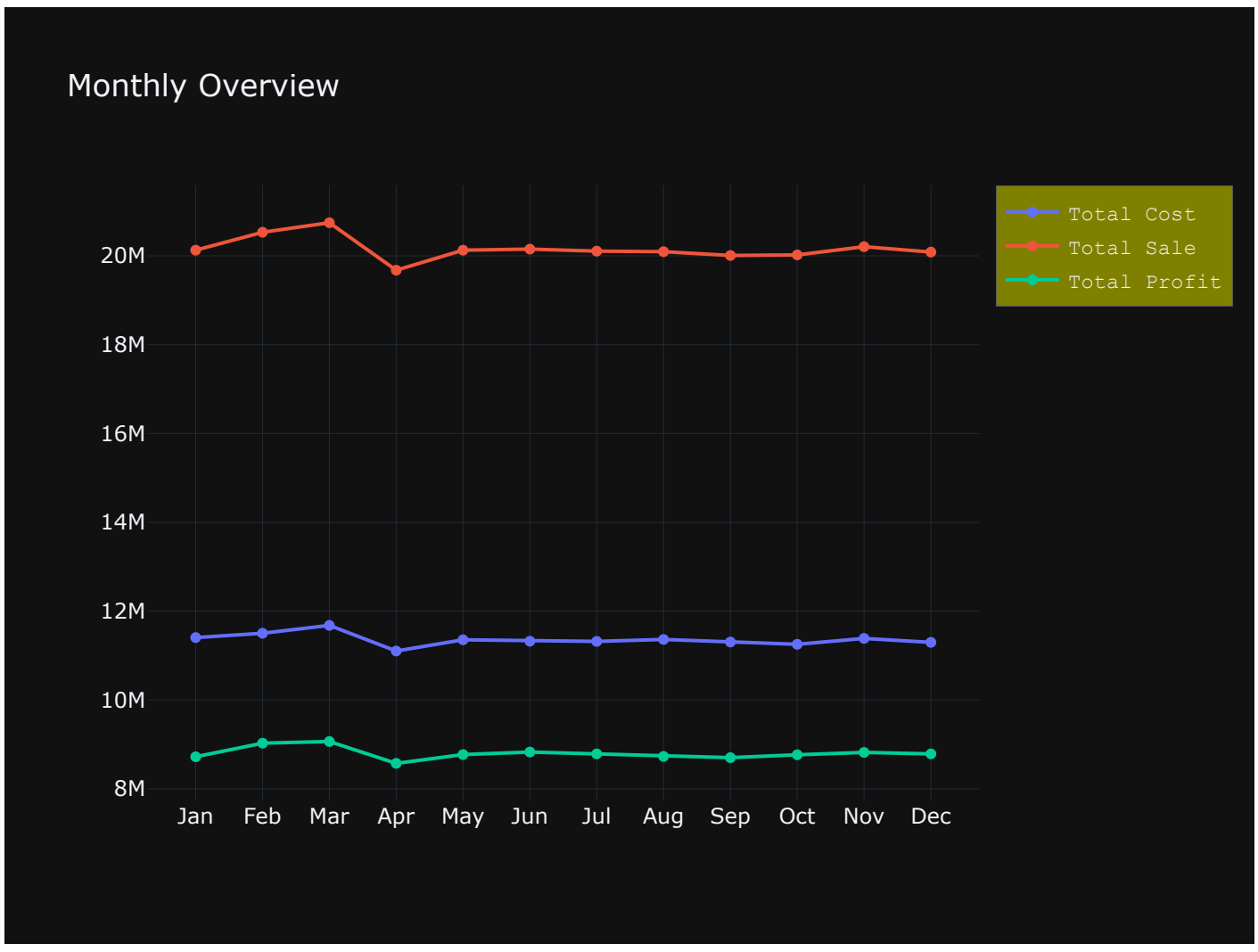
```
query = '''
SELECT
    shortmnthname,
    SUM(quantity_sold) total_qty,
    SUM(total_cost) total_cost,
    SUM(total_sale) total_sale,
    SUM(profit) total_profit,
    AVG(roi) roi

FROM
    breweries_sales
GROUP BY
    shortmnthname
'''
mnth_sum = pd.read_sql_query(query, engine)
mnth_sum
```

Out[ ]:

	shortmnthname	total_qty	total_cost	total_sale	total_profit	roi
0	Aug	74908	11363860.0	20094800.0	8730940.0	73.363195
1	May	74089	11356250.0	20128500.0	8772250.0	73.923821
2	Nov	74120	11386870.0	20207200.0	8820330.0	73.890015
3	Dec	73538	11298510.0	20085800.0	8787290.0	73.923821
4	Feb	74847	11503400.0	20531850.0	9028450.0	74.598930
5	Jun	73637	11324060.0	20152500.0	8828440.0	74.129478
6	Jul	74591	11320340.0	20107350.0	8787010.0	74.306963
7	Oct	73951	11254550.0	20021650.0	8767100.0	74.546427
8	Jan	74373	11406060.0	20129050.0	8722990.0	73.665886
9	Apr	72562	11103620.0	19677450.0	8573830.0	73.890015
10	Mar	76206	11680670.0	20746800.0	9066130.0	74.234068
11	Sep	73900	11307040.0	20009700.0	8702660.0	73.732251

In [ ]:
```
mnth_sum['shortmnthname'] = pd.to_datetime(mnth_sum['shortmnthname'], format="%b").dt.month
mnth_sum.sort_values('shortmnthname', inplace=True)
mnth_order = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun',
              'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']
```

In [ ]:
```
fig = go.Figure()
fig.add_trace(
    go.Scatter(x=mnth_sum.shortmnthname, y=mnth_sum.total_cost,
               name='Total Cost')
),
fig.add_trace(
    go.Scatter(x=mnth_sum.shortmnthname, y=mnth_sum.total_sale,
               name='Total Sale')
),
fig.add_trace(
    go.Scatter(x=mnth_sum.shortmnthname, y=mnth_sum.total_profit,
               name='Total Profit')
)
fig.update_layout(title='Monthly Overview', width=700,
                  xaxis=dict(ticktext=mnth_order,
                             tickvals=[1,2,3,4,5,6,7,8,9,10,11,12]),
                  legend = dict(
                        orientation='v',
                        title='',
                        font = dict(
                        family="Courier",
                        size=12,
                        ),
                        bgcolor='olive',
                        bordercolor='blue',
                        borderwidth=.5
                    )
                  )
```



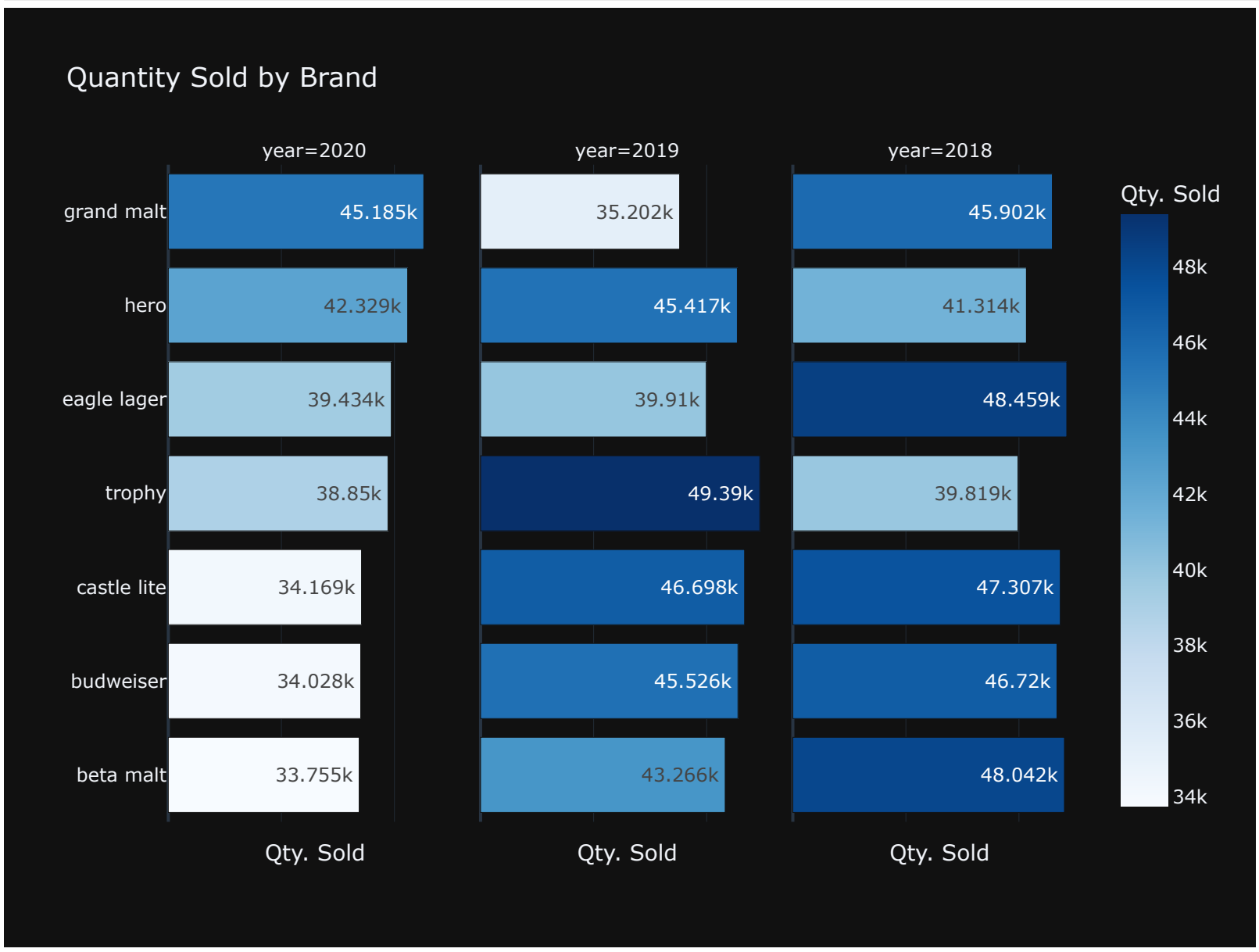
In [ ]:
```
query = '''
SELECT
    brand,
    year,
    SUM(quantity_sold) qty_sold
```

```
FROM
    breweries_sales
GROUP BY
    brand,
    year
'''
brand_qty_sld = pd.read_sql_query(query, engine).sort_values('qty_sold')
brand_qty_sld
```

Out[ ]:

	brand	year	qty_sold
18	beta malt	2020	33755
16	budweiser	2020	34028
7	castle lite	2020	34169
13	grand malt	2019	35202
19	trophy	2020	38850
15	eagle lager	2020	39434
17	trophy	2018	39819
4	eagle lager	2019	39910
11	hero	2018	41314
5	hero	2020	42329
8	beta malt	2019	43266
14	grand malt	2020	45185
3	hero	2019	45417
1	budweiser	2019	45526
0	grand malt	2018	45902
10	castle lite	2019	46698
6	budweiser	2018	46720
2	castle lite	2018	47307
9	beta malt	2018	48042
12	eagle lager	2018	48459
20	trophy	2019	49390

In [ ]:
```
fig = px.bar(brand_qty_sld, 'qty_sold', 'brand', facet_col='year',
             color='qty_sold', title='Quantity Sold by Brand',
             labels={'brand':'', 'qty_sold':'Qty. Sold'}, width=800, height=600,
             text_auto=True, color_continuous_scale='Blues')
fig.update_xaxes(showticklabels=False)
```



In [ ]:
```
query = '''
SELECT
    brand,
    year,
    SUM(total_cost) t_cost,
```

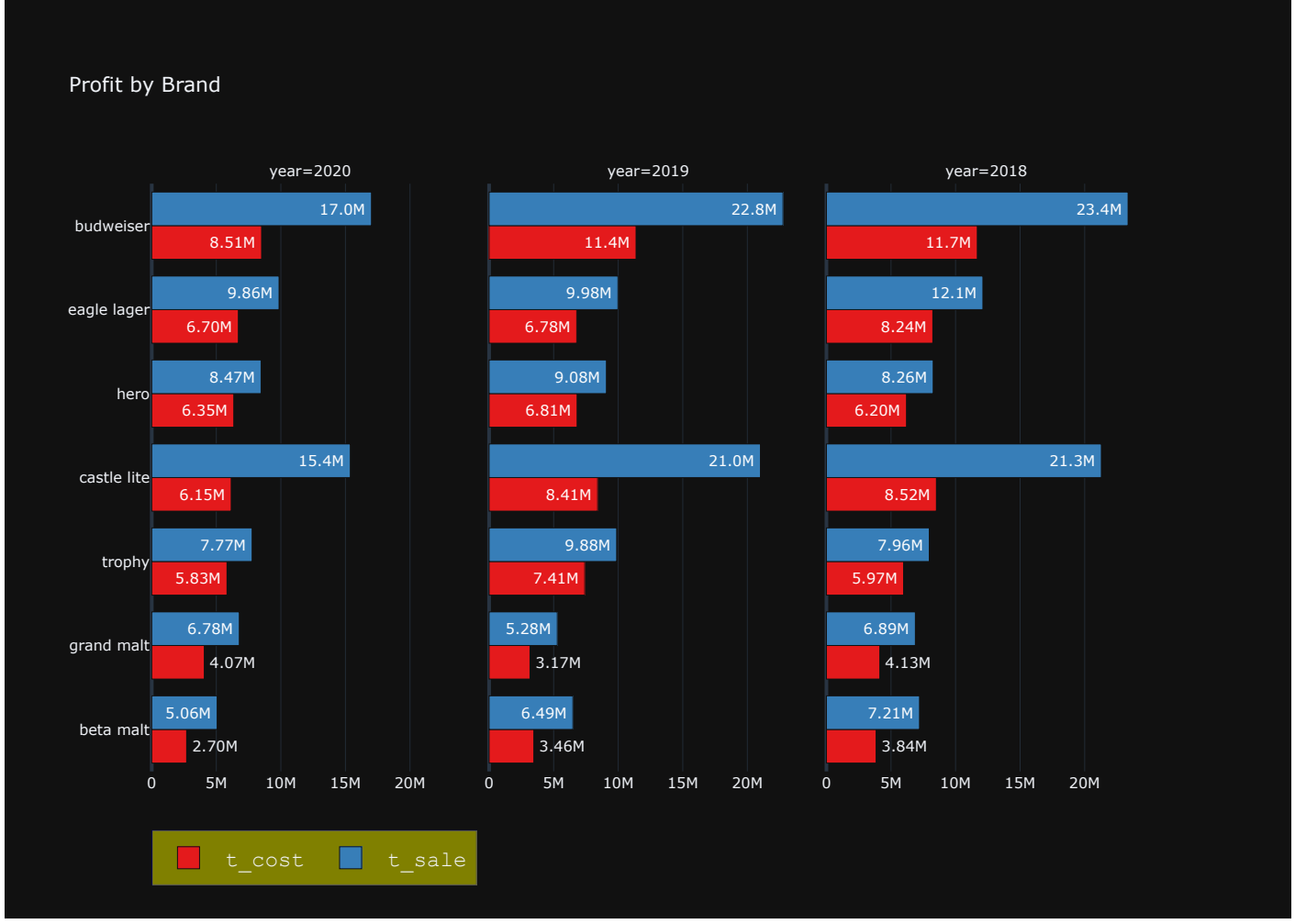
```
    SUM(total_sale) t_sale
FROM
    breweries_sales
GROUP BY
    brand,
    year
'''
brand_cst_sale = pd.read_sql_query(query, engine).sort_values('t_cost')
brand_cst_sale
```

Out[ ]:

	brand	year	t_cost	t_sale
18	beta malt	2020	2700400.0	5063250.0
13	grand malt	2019	3168180.0	5280300.0
8	beta malt	2019	3461280.0	6489900.0
9	beta malt	2018	3843360.0	7206300.0
14	grand malt	2020	4066650.0	6777750.0
0	grand malt	2018	4131180.0	6885300.0
19	trophy	2020	5827500.0	7770000.0
17	trophy	2018	5972850.0	7963800.0
7	castle lite	2020	6150420.0	15376050.0
11	hero	2018	6197100.0	8262800.0
5	hero	2020	6349350.0	8465800.0
15	eagle lager	2020	6703780.0	9858500.0
4	eagle lager	2019	6784700.0	9977500.0
3	hero	2019	6812550.0	9083400.0
20	trophy	2019	7408500.0	9878000.0
12	eagle lager	2018	8238030.0	12114750.0
10	castle lite	2019	8405640.0	21014100.0
16	budweiser	2020	8507000.0	17014000.0
2	castle lite	2018	8515260.0	21288150.0
1	budweiser	2019	11381500.0	22763000.0
6	budweiser	2018	11680000.0	23360000.0

In [ ]:

```
fig = px.bar(brand_cst_sale, ['t_cost', 't_sale'], 'brand', facet_col='year',
             text_auto='.3s', title='Profit by Brand', height=500, width=700,
             labels={'brand':'', 'value':''},barmode='group',
             color_discrete_sequence=px.colors.qualitative.Set1,
             )
fig.update_traces(textposition='auto')
fig.update_layout(font=dict(size=8))
fig.update_layout(legend = dict(
                        orientation='h',
                        title='',
                        font = dict(
                        family="Courier",
                        size=12,
                        ),
                        bgcolor='olive',
                        bordercolor='blue',
                        borderwidth=.5
                    ))
```

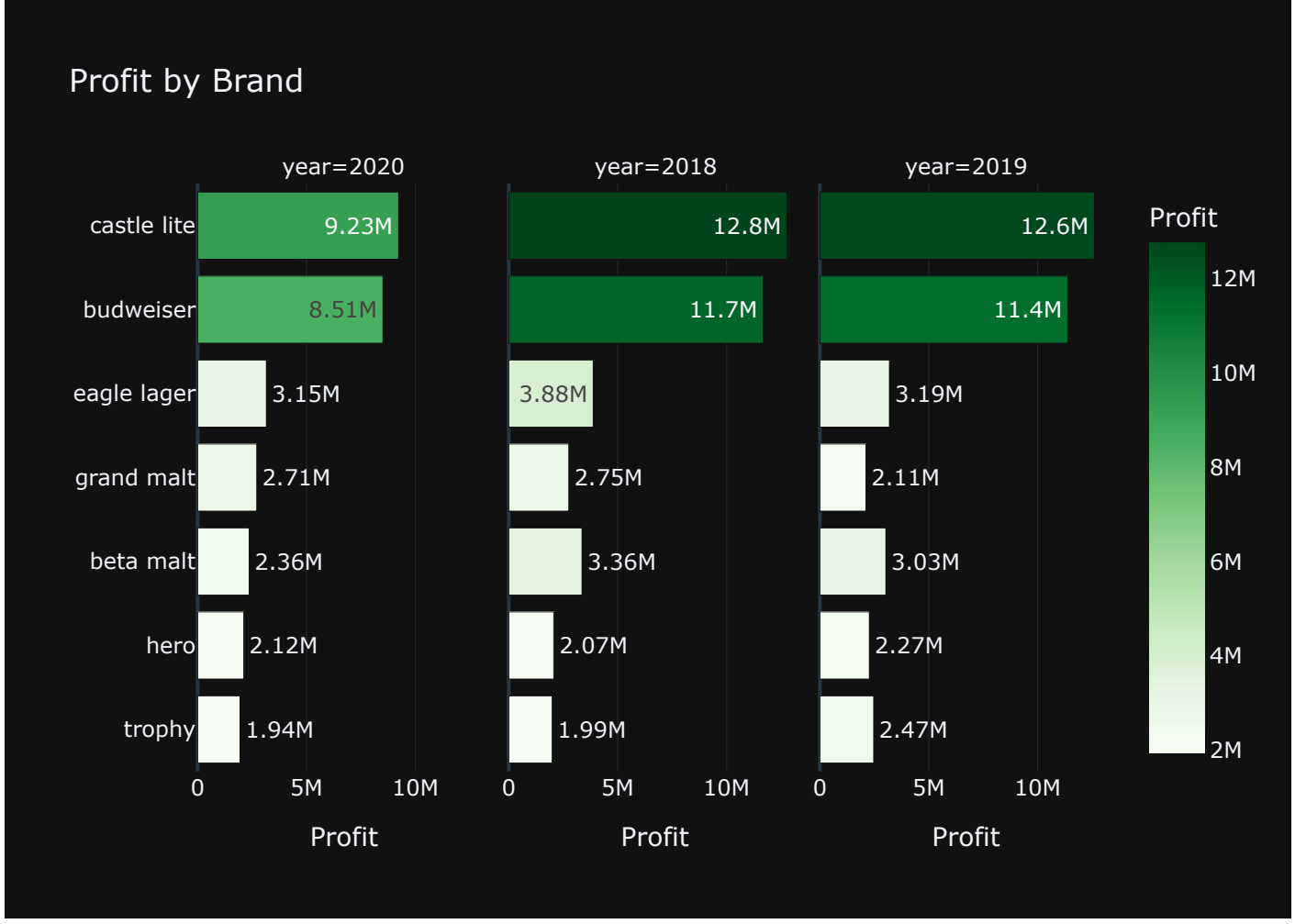


In [ ]:
```
query = '''
SELECT
    brand,
    year,
    SUM(profit) profit
FROM
    breweries_sales
GROUP BY
    brand,
    year
'''
brand_profit = pd.read_sql_query(query, engine).sort_values('profit')
brand_profit
```

Out[ ]:

	brand	year	profit
19	trophy	2020	1942500.0
17	trophy	2018	1990950.0
11	hero	2018	2065700.0
13	grand malt	2019	2112120.0
5	hero	2020	2116450.0
3	hero	2019	2270850.0
18	beta malt	2020	2362850.0
20	trophy	2019	2469500.0
14	grand malt	2020	2711100.0
0	grand malt	2018	2754120.0
8	beta malt	2019	3028620.0
15	eagle lager	2020	3154720.0
4	eagle lager	2019	3192800.0
9	beta malt	2018	3362940.0
12	eagle lager	2018	3876720.0
16	budweiser	2020	8507000.0
7	castle lite	2020	9225630.0
1	budweiser	2019	11381500.0
6	budweiser	2018	11680000.0
10	castle lite	2019	12608460.0
2	castle lite	2018	12772890.0

In [ ]:
```
fig = px.bar(brand_profit, 'profit', 'brand', facet_col='year', text_auto='.3s',
             color='profit', labels={'brand':'', 'profit':'Profit'}, height=500,
             title='Profit by Brand', color_continuous_scale='greens', width=700)
fig
```

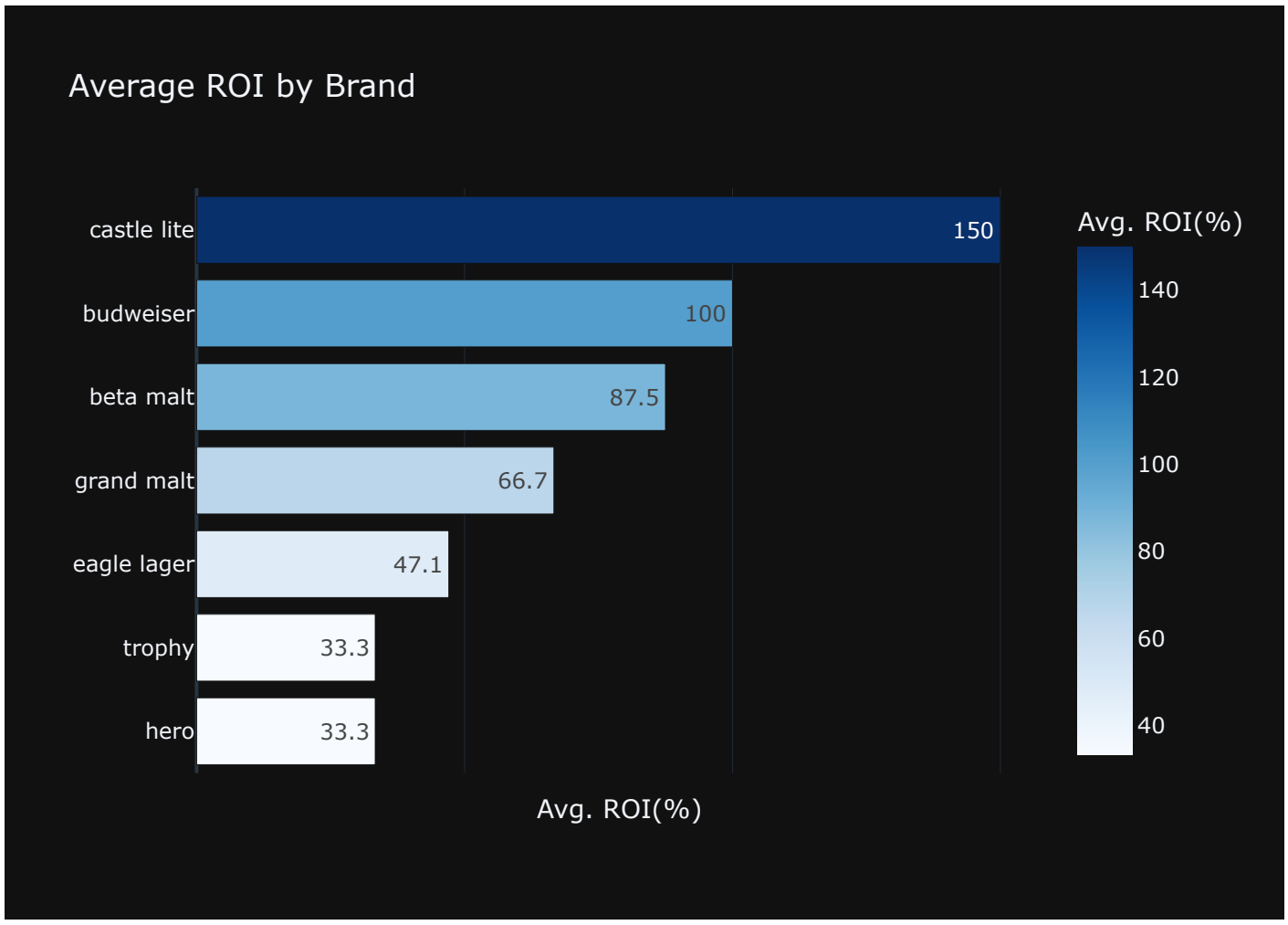


In [ ]:
```
query = '''
SELECT
    brand,
    AVG(roi) avg_roi
FROM
    breweries_sales
GROUP BY
    brand
'''

brand_roi = pd.read_sql_query(query, engine).sort_values('avg_roi')
brand_roi
```

Out[ ]:

	brand	avg_roi
3	hero	33.33333
6	trophy	33.33333
1	eagle lager	47.05882
2	grand malt	66.66667
0	beta malt	87.50000
4	budweiser	100.00000
5	castle lite	150.00000

In [ ]:
```
fig =  px.bar(brand_roi, 'avg_roi', 'brand', text_auto='.3s', color='avg_roi',
              color_continuous_scale='blues', labels={'brand':'','avg_roi': 'Avg. ROI(%)'},
              title='Average ROI by Brand', height=500, width=700)
fig.update_xaxes(showticklabels=False)
```

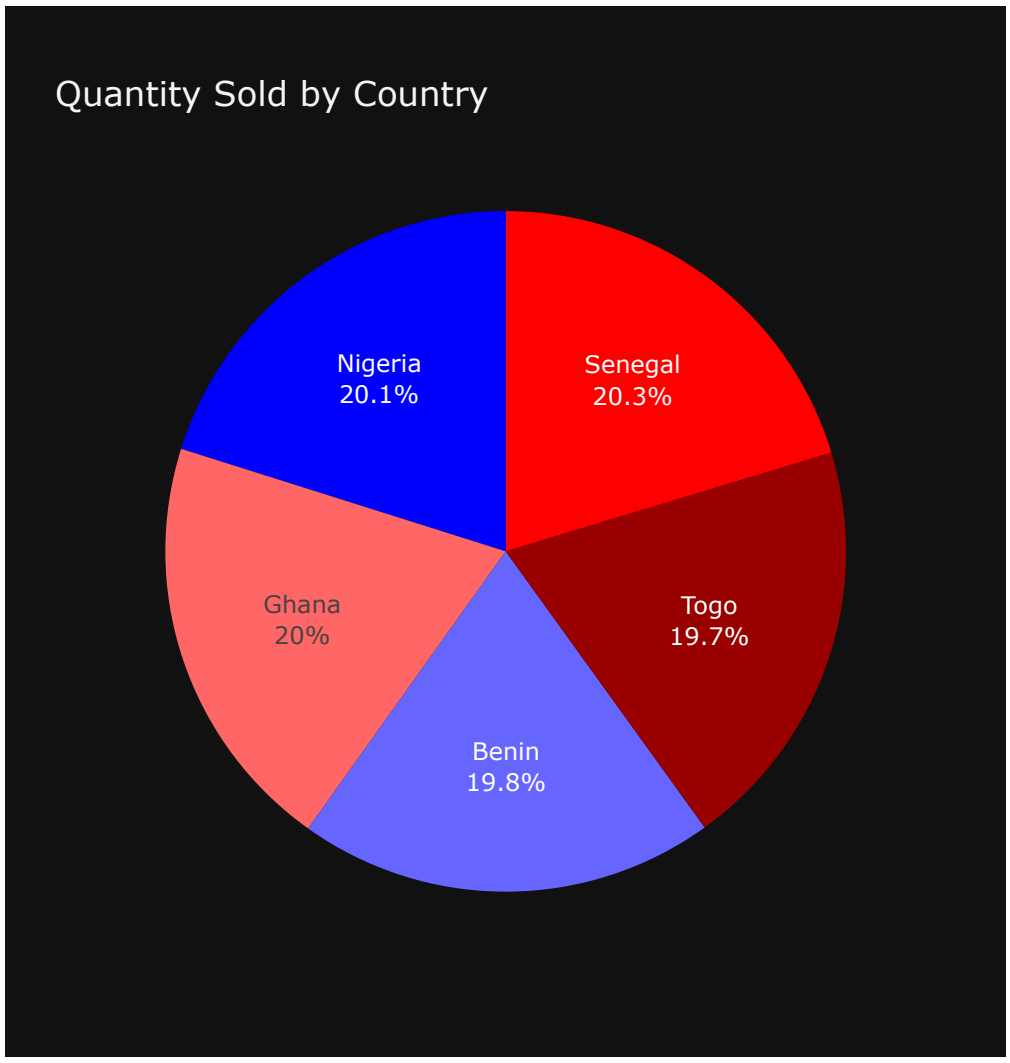


In [ ]:
```
query = '''
SELECT
    country,
    SUM(quantity_sold) total_qty
FROM
    breweries_sales
GROUP BY
    country
ORDER BY
    total_qty DESC
'''
cntry_total_qty = pd.read_sql_query(query, engine)
cntry_total_qty
```

Out[ ]:

	country	total_qty
0	Senegal	180928
1	Nigeria	179307
2	Ghana	178255
3	Benin	176324
4	Togo	175908

In [ ]:
```
fig = px.pie(cntry_total_qty, names='country', values='total_qty', width=500,
       color_discrete_sequence=px.colors.sequential.Bluered_r,
       title='Quantity Sold by Country')
fig.update_traces(textinfo = 'percent+label', showlegend=False)
```

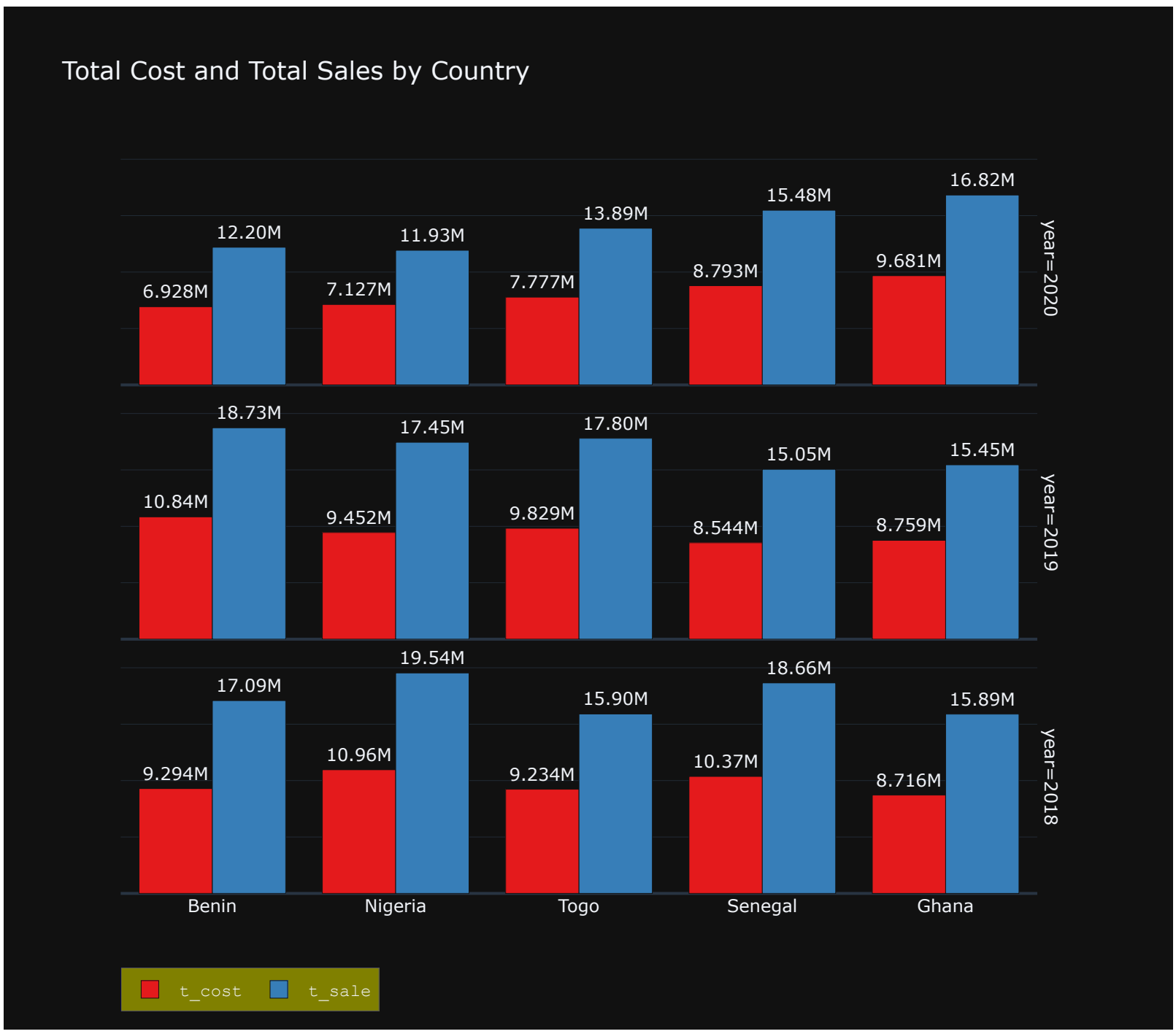


In [ ]:
```
query = '''
SELECT
    country,
    year,
    SUM(total_cost) t_cost,
    SUM(total_sale) t_sale
FROM
    breweries_sales
GROUP BY
    country,
    year
ORDER BY
    t_cost, t_sale
'''
cntry_profit = pd.read_sql_query(query, engine)
cntry_profit
```

Out[ ]:

	country	year	t_cost	t_sale
0	Benin	2020	6927960.0	12201300.0
1	Nigeria	2020	7126630.0	11931950.0
2	Togo	2020	7777190.0	13887150.0
3	Senegal	2019	8544020.0	15052550.0
4	Ghana	2018	8716000.0	15892000.0
5	Ghana	2019	8758770.0	15451750.0
6	Senegal	2020	8792640.0	15480200.0
7	Togo	2018	9234220.0	15903500.0
8	Benin	2018	9294010.0	17089400.0
9	Nigeria	2019	9452460.0	17449800.0
10	Ghana	2020	9680680.0	16824750.0
11	Togo	2019	9829340.0	17801100.0
12	Senegal	2018	10370050.0	18659150.0
13	Benin	2019	10837760.0	18731000.0
14	Nigeria	2018	10963500.0	19537050.0

In [ ]:
```
fig = px.bar(cntry_profit, 'country', ['t_cost', 't_sale'], text_auto='.4s', facet_row='year',
             title='Total Cost and Total Sales by Country', height=700, barmode='group',width=800,
             color_discrete_sequence=px.colors.qualitative.Set1, labels={'value':'', 'country':''})
fig.update_traces(textposition='outside', cliponaxis=False)
fig.update_yaxes(showticklabels=False)
fig.update_layout(legend = dict(
                        orientation='h',
                        title='',
                        font = dict(
                        family="Courier",
                        size=12,
                        ),
                        bgcolor='olive',
                        bordercolor='blue',
                        borderwidth=.5
                    ))
```



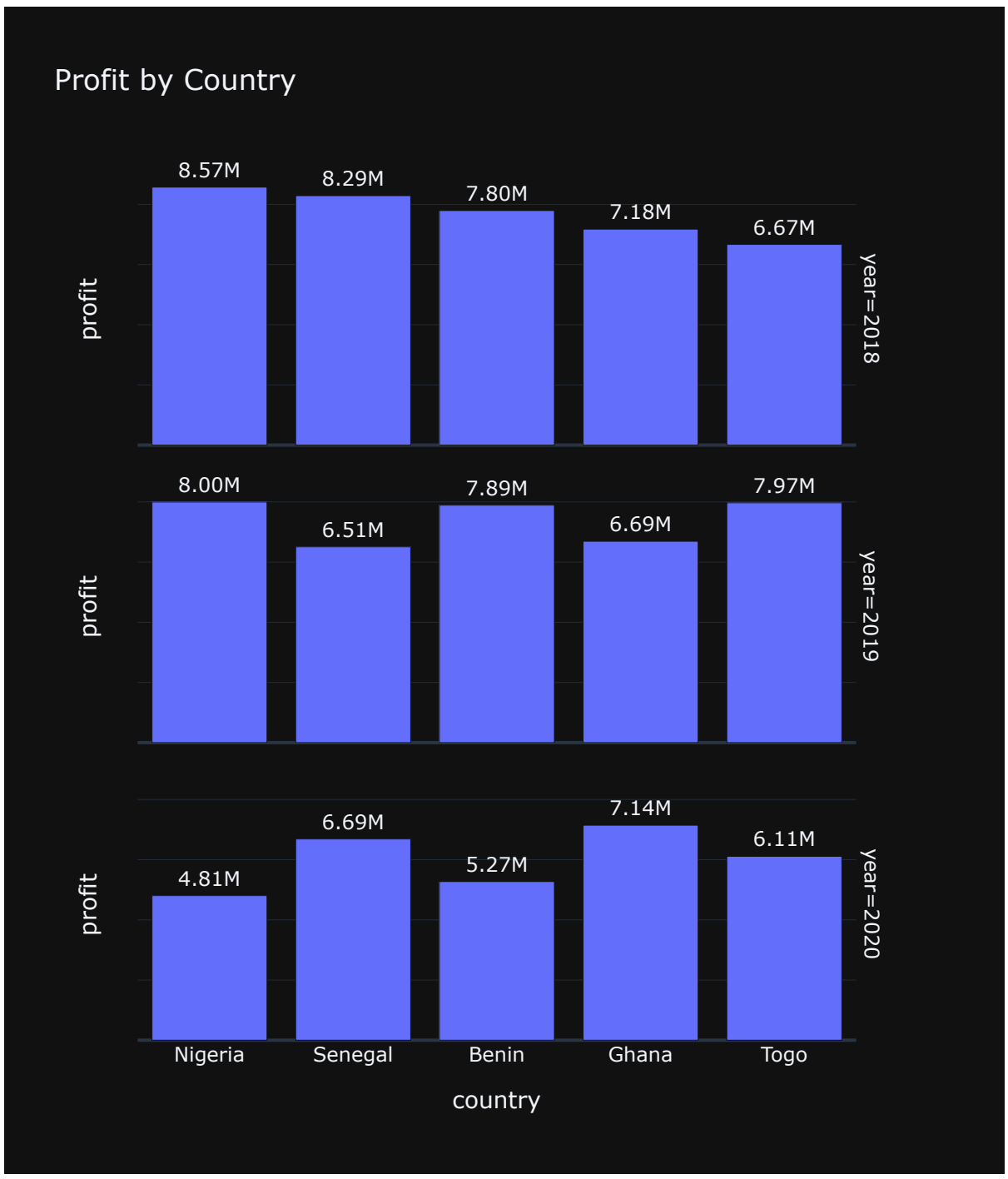
In [ ]:
```
query = '''
SELECT
    country,
    year,
    SUM(profit) profit
FROM
```

```
    breweries_sales
GROUP BY
    country,
    year
ORDER BY
    profit DESC
'''
cntry_profit = pd.read_sql_query(query, engine)
cntry_profit
```

Out[ ]:

	country	year	profit
0	Nigeria	2018	8573550.0
1	Senegal	2018	8289100.0
2	Nigeria	2019	7997340.0
3	Togo	2019	7971760.0
4	Benin	2019	7893240.0
5	Benin	2018	7795390.0
6	Ghana	2018	7176000.0
7	Ghana	2020	7144070.0
8	Ghana	2019	6692980.0
9	Senegal	2020	6687560.0
10	Togo	2018	6669280.0
11	Senegal	2019	6508530.0
12	Togo	2020	6109960.0
13	Benin	2020	5273340.0
14	Nigeria	2020	4805320.0

In [ ]:
```
fig = px.bar(cntry_profit, 'country', 'profit', text_auto='.3s', facet_row='year',
             title='Profit by Country', height=700, width=600)
fig.update_traces(textposition='outside', cliponaxis=False)
fig.update_yaxes(showticklabels=False)
```



In [ ]:
```
%%sql
CREATE TABLE nigeria AS
    SELECT  *
    FROM    breweries_sales
    WHERE   Country = 'Nigeria';

CREATE TABLE senegal AS
    SELECT  *
    FROM    breweries_sales
```

```
    WHERE   Country = 'Senegal';

CREATE TABLE benin AS
    SELECT  *
    FROM    breweries_sales
    WHERE   Country = 'Benin';

CREATE TABLE ghana AS
    SELECT  *
    FROM    breweries_sales
    WHERE   Country = 'Ghana';

CREATE TABLE togo AS
    SELECT  *
    FROM    breweries_sales
    WHERE   Country = 'Togo';
```

```
 * postgresql://postgres:***@localhost/international_breweries
210 rows affected.
209 rows affected.
209 rows affected.
210 rows affected.
209 rows affected.
[]
```

Out[ ]:

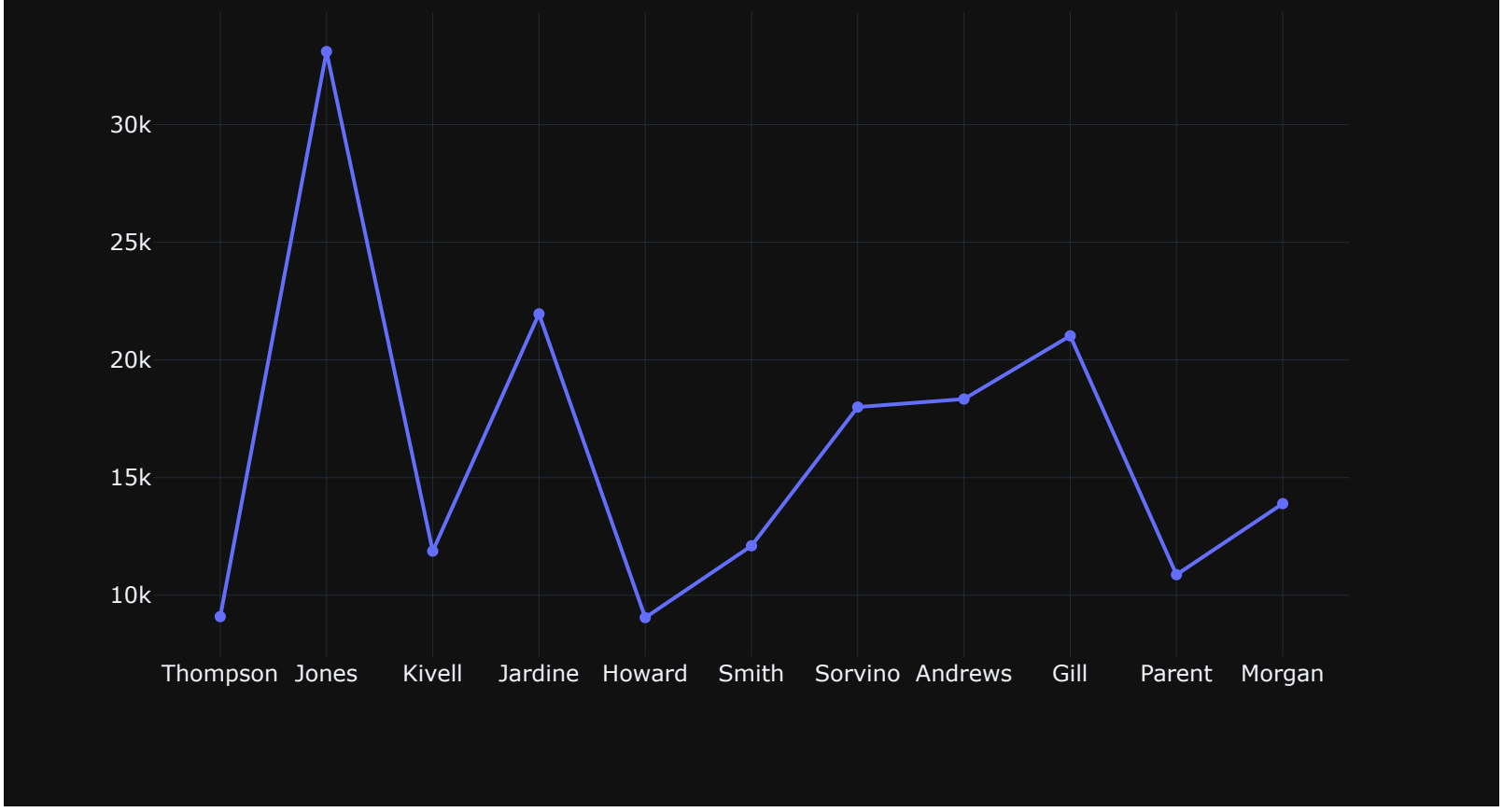
In [ ]:
```
query = '''
SELECT
    sales_rep,
    SUM(quantity_sold) t_qty,
    SUM(total_cost) t_cost,
    SUM(total_sale) t_sale,
    SUM(profit) t_profit,
    ROUND(AVG(roi), 2) avg_roi
FROM
    nigeria
GROUP BY
    sales_rep
'''

ng_sl = pd.read_sql_query(query, engine)
ng_sl
```

Out[ ]:

	sales_rep	t_qty	t_cost	t_sale	t_profit	avg_roi
0	Thompson	9092	1281270.0	2329150.0	1047880.0	77.76
1	Jones	33104	5631170.0	9119600.0	3488430.0	56.88
2	Kivell	11875	1745080.0	2974750.0	1229670.0	68.71
3	Jardine	21958	3447760.0	6081050.0	2633290.0	72.69
4	Howard	9050	1578140.0	2619200.0	1041060.0	60.12
5	Smith	12101	1821450.0	3494650.0	1673200.0	86.48
6	Sorvino	17998	2495590.0	4673100.0	2177510.0	82.85
7	Andrews	18337	2612390.0	4688350.0	2075960.0	75.03
8	Gill	21025	3502700.0	6271950.0	2769250.0	74.88
9	Parent	10874	1485510.0	2891650.0	1406140.0	89.29
10	Morgan	13893	1941530.0	3775350.0	1833820.0	89.03

In [ ]:

In [ ]:
```
fig = go.Figure()

fig.add_trace(
    go.Scatter(x=ng_sl.sales_rep, y=ng_sl.t_qty)
)
fig.update_layout(width=800)
```



In [ ]:
```
%%sql
SELECT
    sales_rep,
    SUM(quantity_sold) t_qty,
    SUM(total_cost) t_cost,
    SUM(total_sale) t_sale,
    SUM(profit) t_profit,
    ROUND(AVG(roi), 2) avg_roi
FROM
    nigeria
GROUP BY
    sales_rep
```

```
 * postgresql://postgres:***@localhost/international_breweries
11 rows affected.
```

Out[ ]:

sales_rep	t_qty	t_cost	t_sale	t_profit	avg_roi
Thompson	9092	1281270	2329150	1047880	77.76
Jones	33104	5631170	9119600	3488430	56.88
Kivell	11875	1745080	2974750	1229670	68.71
Jardine	21958	3447760	6081050	2633290	72.69
Howard	9050	1578140	2619200	1041060	60.12
Smith	12101	1821450	3494650	1673200	86.48
Sorvino	17998	2495590	4673100	2177510	82.85
Andrews	18337	2612390	4688350	2075960	75.03
Gill	21025	3502700	6271950	2769250	74.88
Parent	10874	1485510	2891650	1406140	89.29
Morgan	13893	1941530	3775350	1833820	89.03

In [ ]:
```
%%sql
SELECT * FROM nigeria LIMIT 5
```

```
 * postgresql://postgres:***@localhost/international_breweries
5 rows affected.
```

Out[ ]:

sales_id	region	month	year	sales_rep	brand	unit_cost	unit_price	quantity_sold	country	total_cost	total_sale	profit	roi	shortmnthname
10257	Southeast	January	2019	Morgan	castle lite	180	450	961	Nigeria	172980	432450	259470	150.00000	Jan
10312	west	August	2019	Jones	budweiser	250	500	937	Nigeria	234250	468500	234250	100.00000	Aug
10322	northcentral	June	2019	Jones	hero	150	200	910	Nigeria	136500	182000	45500	33.33333	Jun
10327	northeast	November	2019	Smith	castle lite	180	450	983	Nigeria	176940	442350	265410	150.00000	Nov
10342	west	February	2019	Gill	eagle lager	170	250	922	Nigeria	156740	230500	73760	47.05882	Feb