In [108]:

```python
import pandas as pd
import seaborn as sns
import plotly.express as px

import matplotlib.pyplot as plt
```

In [109]:

```python
import plotly.io as pio
pio.renderers.default = "plotly_mimetype+notebook"
```

Matplotlib

For this excercise, we have written the following code to load the stock dataset built into plotly express.

In [110]:

```python
stocks = px.data.stocks()
stocks.head()
```

Out[110]:

	date	GOOG	AAPL	AMZN	FB	NFLX	MSFT
0	2018-01-01	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
1	2018-01-08	1.018172	1.011943	1.061881	0.959968	1.053526	1.015988
2	2018-01-15	1.032008	1.019771	1.053240	0.970243	1.049860	1.020524
3	2018-01-22	1.066783	0.980057	1.140676	1.016858	1.307681	1.066561
4	2018-01-29	1.008773	0.917143	1.163374	1.018357	1.273537	1.040708

Question 1:

Select a stock and create a suitable plot for it. Make sure the plot is readable with relevant information, such as date, values.

In [111]:

```python
# YOUR CODE HERE
stocks = px.data.stocks()
fig = px.line(stocks,x="date", y="GOOG", title="GOOGLE stocks").update_layout(yaxis_title="
fig.show()
```


Question 2:

You've already plot data from one stock. It is possible to plot multiples of them to support comparison.
To highlight different lines, customise line styles, markers, colors and include a legend to the plot.

In [176]:

```python
# YOUR CODE HERE
fig, ax = plt.subplots(figsize=(16,6))

for stock in stocks.columns[1:]:
    plt.plot(stocks.date, stocks[stock], label=stock)

for label in ax.get_xaxis().get_ticklabels()[::2]:
    label.set_visible(False)

plt.xticks(rotation=90)
plt.legend()
plt.show()
```


Seaborn

First, load the tips (https://github.com/mwaskom/seaborn-data/blob/master/tips.csv) dataset

In [177]:

```python
tips = sns.load_dataset('tips')
tips.head()
```

Out[177]:

|	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Question 3:

Let's explore this dataset. Pose a question and create a plot that support drawing answers for your question.

Some possible questions:

- Are there differences between male and female when it comes to giving tips?
- What attribute correlate the most with tip?

In [178]:

```python
# YOUR CODE HERE
# What are the differences between smokers and non-smokers when it comes to giving tips?
sns.scatterplot(
    data=tips,
    x='smoker',
    y='tip'
)
plt.show()
```
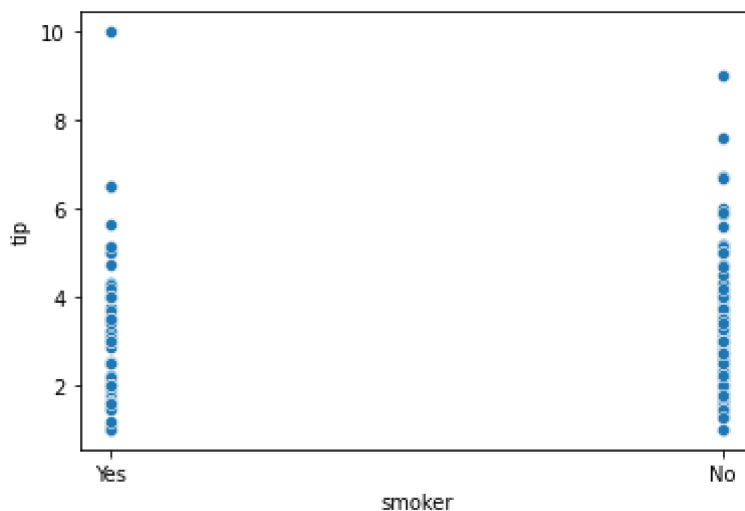
Plotly Express

Question 4:

Redo the above exercises (challenges 2 & 3) with plotly express. Create diagrams which you can interact with.

The stocks dataset

Hints:

- Turn stocks dataframe into a structure that can be picked up easily with plotly express
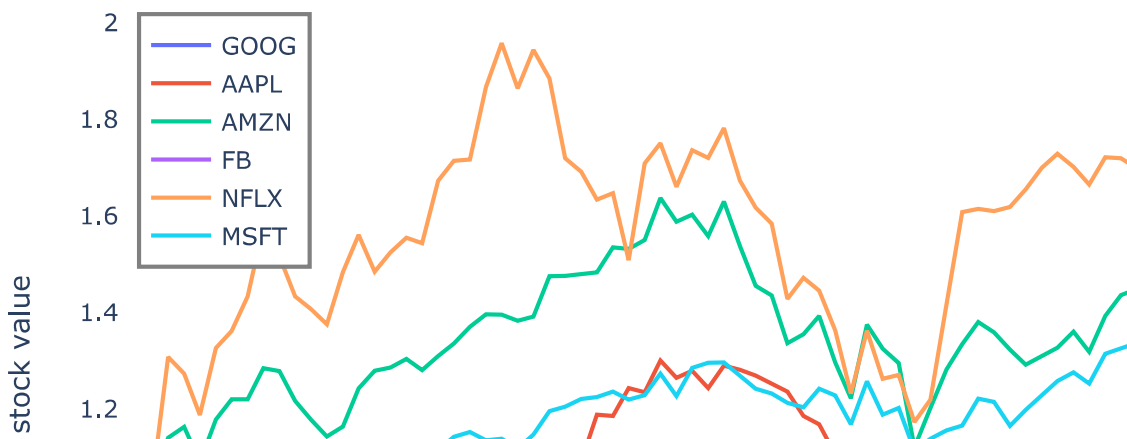
In [179]:

```python
# YOUR CODE HERE
stocks = px.data.stocks()
fig = px.line(stocks,x="date", y=stocks.columns, title="Stocks").update_layout(yaxis_title=

fig.update_layout(
    plot_bgcolor = "white",
    legend=dict(
        title = None,
        bgcolor = "White",
        bordercolor = "Grey",
        borderwidth = 2,
        yanchor="top",
        y=0.995,
        xanchor="left",
        x=0.01
))
fig.update_xaxes(automargin=False)

fig.show()
```
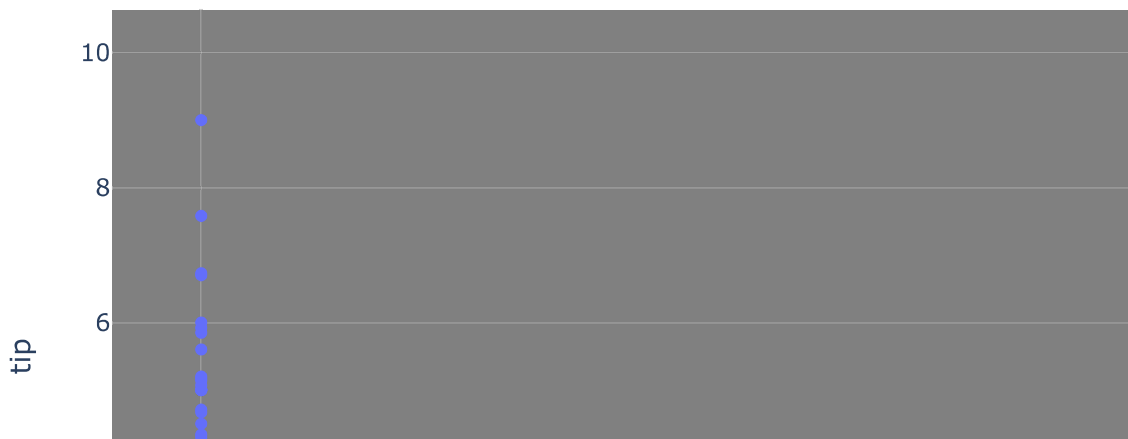
Stocks

The tips dataset

In [194]:

```python
# YOUR CODE HERE
# What are the differences between smokers and non-smokers when it comes to giving tips?
tips = px.data.tips()
fig = px.scatter(
    tips,
    x="smoker",
    y="tip",
    title="Stocks",
    color="smoker"
)

fig.update_layout(
    plot_bgcolor = "Grey",
)

fig.show()
```

Stocks

Question 5:

Recreate the barplot below that shows the population of different continents for the year 2007.

Hints:

- Extract the 2007 year data from the dataframe. You have to process the data accordingly

- use [plotly bar (https://plotly.com/python-api-reference/generated/plotly.express.bar)](https://plotly.com/python-api-reference/generated/plotly.express.bar)
- Add different colors for different continents
- Sort the order of the continent for the visualisation. Use [axis layout setting (https://plotly.com/python/reference/layout/xaxis/)](https://plotly.com/python/reference/layout/xaxis/)
- Add text to each bar that represents the population

In [196]:

```python
#load data
df = px.data.gapminder()
df.head()
```

Out[196]:

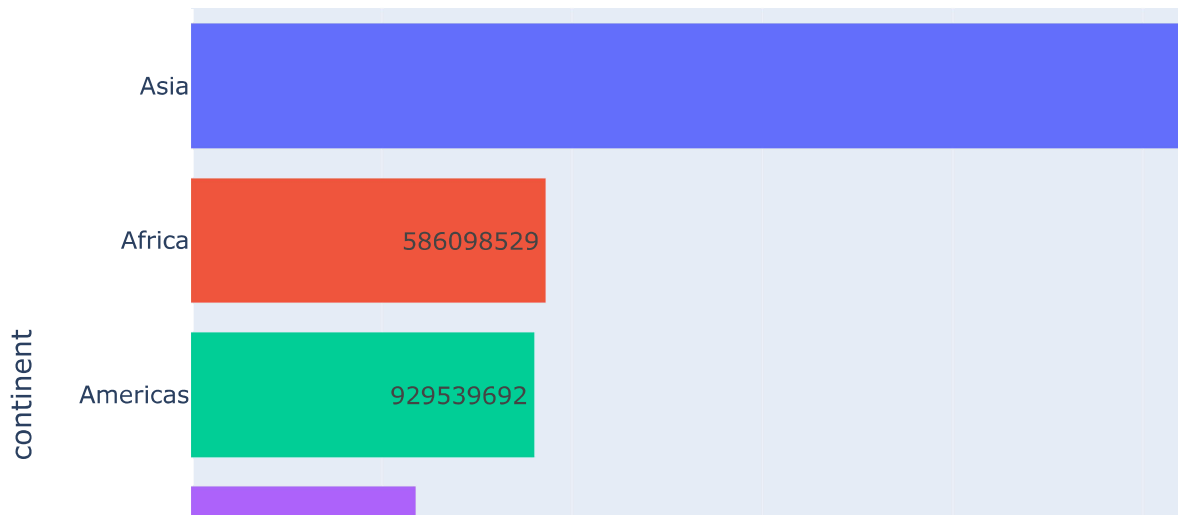|	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
0	Afghanistan	Asia	1952	28.801	8425333	779.445314	AFG	4
1	Afghanistan	Asia	1957	30.332	9240934	820.853030	AFG	4
2	Afghanistan	Asia	1962	31.997	10267083	853.100710	AFG	4
3	Afghanistan	Asia	1967	34.020	11537966	836.197138	AFG	4
4	Afghanistan	Asia	1972	36.088	13079460	739.981106	AFG	4

In [258]:

```python
# YOUR CODE HERE
data_year = df.query("year==2007")
pop = []
continents = data_year.continent.unique()
c = []
for continent in continents:
    pop.append(data_year[data_year.continent == continent]['pop'].sum())
    c.append(continent)

data_grouped = pd.DataFrame([continents,pop] ).transpose()
data_grouped.columns = ['continent', 'pop']
data_grouped = data_grouped.sort_values('pop', ascending = False)

fig = px.bar(
    data_grouped,
    x = 'pop',
    y = 'continent',
    color = 'continent',
    barmode = 'stack',
    text=pop,
    orientation = 'h',
)

fig.show()
```

In [ ]: