

CH07_Drug_Overdose_Python_Supplement

September 3, 2026

1 Chapter 7 — Drug overdose data (Python supplement)

Condensed notebook for the **Python / Plotly** portion of *Chapter 7: Drug Overdose Data*.

Target graphics:

- Overdoses by gender, bar/line combo in Express and graph objects (**Figures 7.41–7.42**)
- Selected drug trends over time (**Figure 7.43**)
- Dual-axis total rate and annual percent change (**Figure 7.44**)
- Overdose rates by race/ethnicity (**Figures 7.45–7.50**)

Data: Drug Overdose Data by Gender.xlsx and Drug Overdose Data by Demographic.xlsx in Data For Condensed Notebooks.

Dependencies:

- pandas
- plotly
- openpyxl

1.1 Imports and display options

```
[1]: # Import packages.
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go

[2]: # Set output options.
import plotly.io as pio
pio.renderers.default = "pdf+jupyterlab+notebook"

[3]: # Set the maximum number of DataFrame rows to display.
pd.options.display.max_rows = 8
pd.options.display.max_columns = 6
```

1.2 Loading the data

```
[4]: df_gender = pd.read_excel('../Data For Condensed Notebooks/Drug Overdose Data_
    ↳by Gender.xlsx')
```

```
[5]: df_gender
```

```
[5]:
```

	Category	1999	2000	...	\
0	Total Overdose Deaths	16849	17415	...	
1	Female	5591	5852	...	
2	Male	11258	11563	...	
3	Any Opioid1	8050	8407	...	
..	...	...	...	...	
99	Male	61	46	...	
100	Antidepressants WITHOUT Synthetic Opioids othe...	1627	1675	...	
101	Female	865	907	...	
102	Male	762	768	...	

	2021	2022	2015-2022 Fold Change
0	106699	107941	2.059786
1	32398	32127	1.652029
2	74301	75814	2.300391
3	80411	81806	2.472153
..	...	...	...
99	1374	1461	4.234783
100	3138	3106	0.760157
101	1996	1931	0.789452
102	1142	1175	0.716463

[103 rows x 26 columns]

```
[6]: df_demographic = pd.read_excel('../Data For Condensed Notebooks/Drug Overdose_
↳Data by Demographic.xlsx')
df_demographic
```

```
[6]:
```

	Category	1999	2000	...	2021	\
0	Total Overdose Deaths	6.1	6.2	...	32.4	
1	Female	3.9	4.1	...	19.6	
2	Male	8.2	8.3	...	45.1	
3	White (Non-Hispanic)	6.2	6.6	...	36.8	
..	...	...	...	...	...	
80	Asian* (Non-Hispanic)	NaN	NaN	...	1.5	
81	Native Hawaiian or Other Pacific Islander* (...)	NaN	NaN	...	11.8	
82	Hispanic	0.2	0.2	...	6.4	
83	American Indian or Alaska Native (Non-Hispanic)	NaN	NaN	...	27.4	

	2022	Fold Change 2015 to 2022
0	32.6	2.000000
1	19.4	1.644068
2	45.6	2.192308
3	35.6	1.687204
..	...	...

```

80    1.7                NaN
81   13.7                NaN
82    6.9             4.928571
83   32.7             6.055556

```

```
[84 rows x 26 columns]
```

1.3 Combined bar/line chart — overdoses by gender

Melt year columns, then combine a line chart (male/female) with bars (total). **Figure 7.41** (Express); **Figure 7.42** (graph objects).

```
[7]: year_cols = list(range(1999, 2023))
```

Slice male/female rows and the total row; melt each to long form.

```
[8]: df_mf = df_gender.loc[1:2, :]
df_mf
```

```
[8]:
```

	Category	1999	2000	...	2021	2022	2015-2022 Fold Change
1	Female	5591	5852	...	32398	32127	1.652029
2	Male	11258	11563	...	74301	75814	2.300391

```
[2 rows x 26 columns]
```

```
[9]: df_total = df_gender.loc[0:0, :]
df_total
```

```
[9]:
```

	Category	1999	2000	...	2021	2022	\
0	Total Overdose Deaths	16849	17415	...	106699	107941	
	2015-2022 Fold Change						
0							2.059786

```
[1 rows x 26 columns]
```

```
[10]: df_total_melted = df_total.melt(id_vars='Category', value_vars=year_cols,
    ↪var_name='Year', value_name='Overdoses')
df_total_melted
```

```
[10]:
```

	Category	Year	Overdoses
0	Total Overdose Deaths	1999	16849
1	Total Overdose Deaths	2000	17415
2	Total Overdose Deaths	2001	19394
3	Total Overdose Deaths	2002	23518
..	...	...	...
20	Total Overdose Deaths	2019	70630
21	Total Overdose Deaths	2020	91799

```

22    Total Overdose Deaths  2021    106699
23    Total Overdose Deaths  2022    107941

```

[24 rows x 3 columns]

```

[11]: df_mf_melted = df_mf.melt(id_vars='Category', value_vars=year_cols,
    ↪var_name='Year', value_name='Overdoses')
df_mf_melted

```

```

[11]:   Category  Year  Overdoses
0     Female  1999      5591
1        Male  1999     11258
2     Female  2000      5852
3        Male  2000     11563
..      ...  ...      ...
44    Female  2021     32398
45        Male  2021     74301
46    Female  2022     32127
47        Male  2022     75814

```

[48 rows x 3 columns]

Figure 7.41 — Express: lines plus add_bar for totals.

```

[12]: # Create the line chart for the male and female genders.
fig = px.line(df_mf_melted, x='Year', y='Overdoses', color='Category')

# Add the bar chart for all genders.
fig.add_bar(x=df_total_melted['Year'], y=df_total_melted['Overdoses'],
    ↪name='Total')
fig.update_layout(title='Overdoses by Gender', width=1000)
fig

```

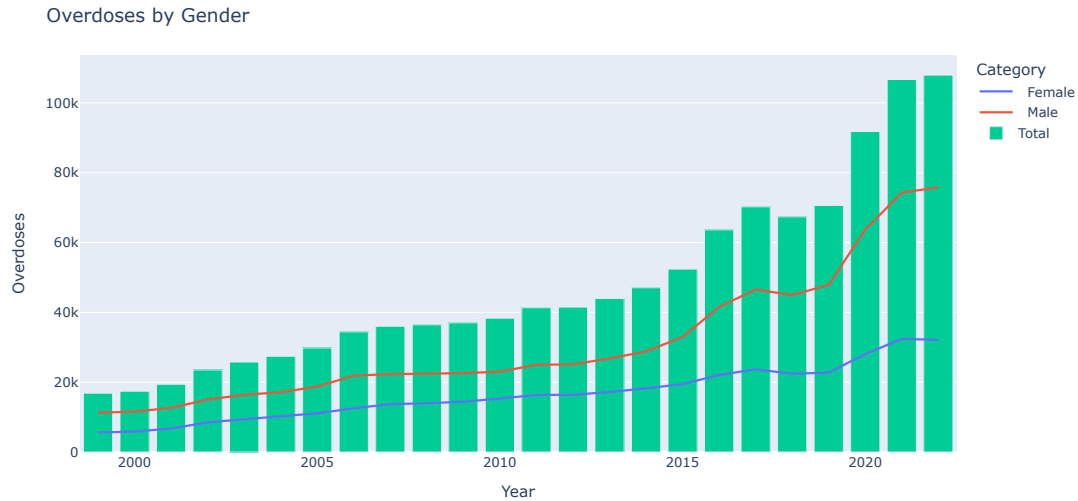


Figure 7.42 — Graph objects: melt each gender row separately, then combine traces.

```
[13]: df_female = df_gender.loc[1:1]
df_male = df_gender.loc[2:2]

df_female_melted = df_female.melt(value_vars=year_cols, var_name='Year',
    ↪value_name='Overdoses')
df_male_melted = df_male.melt(value_vars=year_cols, var_name='Year',
    ↪value_name='Overdoses')
```

```
[14]: # Create an empty figure.
fig = go.Figure()

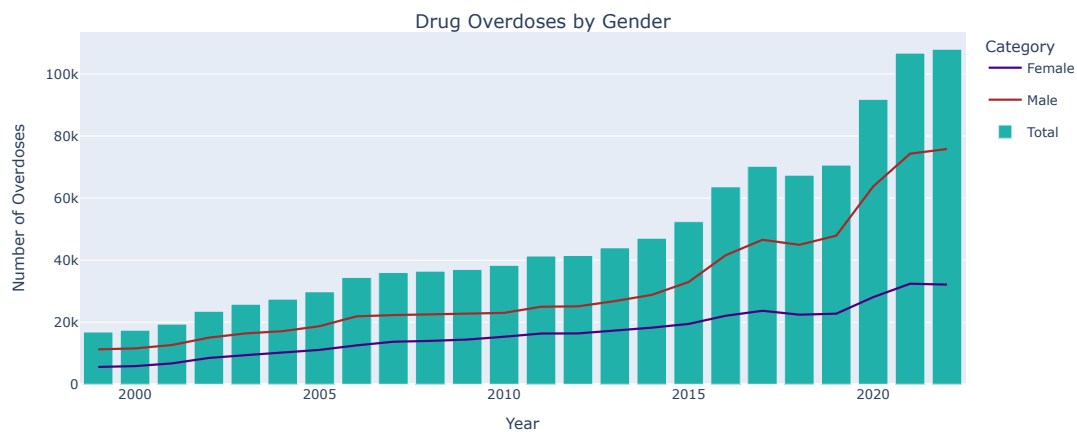
female_trace = go.Scatter(x=df_female_melted['Year'],
    y=df_female_melted['Overdoses'],
    mode='lines',
    name='Female',
    legendgroup='Female',
    marker_color='indigo')
male_trace = go.Scatter(x=df_male_melted['Year'],
    y=df_male_melted['Overdoses'],
    mode='lines',
    name='Male',
    legendgroup='Male',
    marker_color='brown')
total_trace = go.Bar(x=df_total_melted['Year'],
    y=df_total_melted['Overdoses'],
    name='Total',
    legendgroup='Total',
```

```

        marker_color='lightseagreen')

fig.add_traces([female_trace, male_trace, total_trace])
fig.update_xaxes(title='Year')
fig.update_yaxes(title='Number of Overdoses')
fig.update_layout(title={'text': 'Drug Overdoses by Gender',
                          'y': 0.83, 'x': 0.5,
                          'xanchor': 'center', 'yanchor': 'top'},
                  legend_title='Category', width=1000)
fig

```



1.4 Line chart — selected drugs over time

Select rows by index, shorten labels, melt, and plot. **Figure 7.43.**

```

[15]: df_select = df_gender.loc[[3, 7, 58, 43, 73, 19, 88]]
      df_select

```

```

[15]:
      Category  1999  2000  ...  2021  \
3      Any Opioid1  8050  8407  ...  80411
7      Prescription Opioids2  3442  3785  ...  16706
58  Psychostimulants With Abuse Potential (primar...  547  578  ...  32537
43      Cocaine5  3822  3544  ...  24486
73      Benzodiazepines7  1135  1298  ...  12499
19      Heroin4  1960  1842  ...  9173
88      Antidepressants8  1749  1798  ...  5859

      2022  2015-2022 Fold Change
3  81806          2.472153

```

7	14716	0.963026
58	34022	5.952064
43	27569	4.063827
73	10964	1.247185
19	5871	0.451998
88	5863	1.197998

[7 rows x 26 columns]

```
[16]: df_select.loc[58, 'Category'] = 'Psychostimulants'
df_select
```

```
[16]:
```

	Category	1999	2000	...	2021	2022	\
3	Any Opioid1	8050	8407	...	80411	81806	
7	Prescription Opioids2	3442	3785	...	16706	14716	
58	Psychostimulants	547	578	...	32537	34022	
43	Cocaine5	3822	3544	...	24486	27569	
73	Benzodiazepines7	1135	1298	...	12499	10964	
19	Heroin4	1960	1842	...	9173	5871	
88	Antidepressants8	1749	1798	...	5859	5863	

	2015-2022 Fold Change
3	2.472153
7	0.963026
58	5.952064
43	4.063827
73	1.247185
19	0.451998
88	1.197998

[7 rows x 26 columns]

```
[17]: df_select['Category'] = df_select['Category'].replace(r'\d+', '', regex=True)
df_select
```

```
[17]:
```

	Category	1999	2000	...	2021	2022	\
3	Any Opioid	8050	8407	...	80411	81806	
7	Prescription Opioids	3442	3785	...	16706	14716	
58	Psychostimulants	547	578	...	32537	34022	
43	Cocaine	3822	3544	...	24486	27569	
73	Benzodiazepines	1135	1298	...	12499	10964	
19	Heroin	1960	1842	...	9173	5871	
88	Antidepressants	1749	1798	...	5859	5863	

	2015-2022 Fold Change
3	2.472153
7	0.963026

```

58          5.952064
43          4.063827
73          1.247185
19          0.451998
88          1.197998

```

[7 rows x 26 columns]

```

[18]: df_select_melted = df_select.melt(value_vars=year_cols,
                                     id_vars='Category',
                                     var_name='Year',
                                     value_name='Overdoses')

df_select_melted

```

```

[18]:
      Category  Year  Overdoses
0      Any Opioid  1999      8050
1  Prescription Opioids  1999      3442
2    Psychostimulants  1999       547
3         Cocaine  1999      3822
..          ...   ...      ...
164        Cocaine  2022     27569
165  Benzodiazepines  2022     10964
166         Heroin  2022       5871
167  Antidepressants  2022       5863

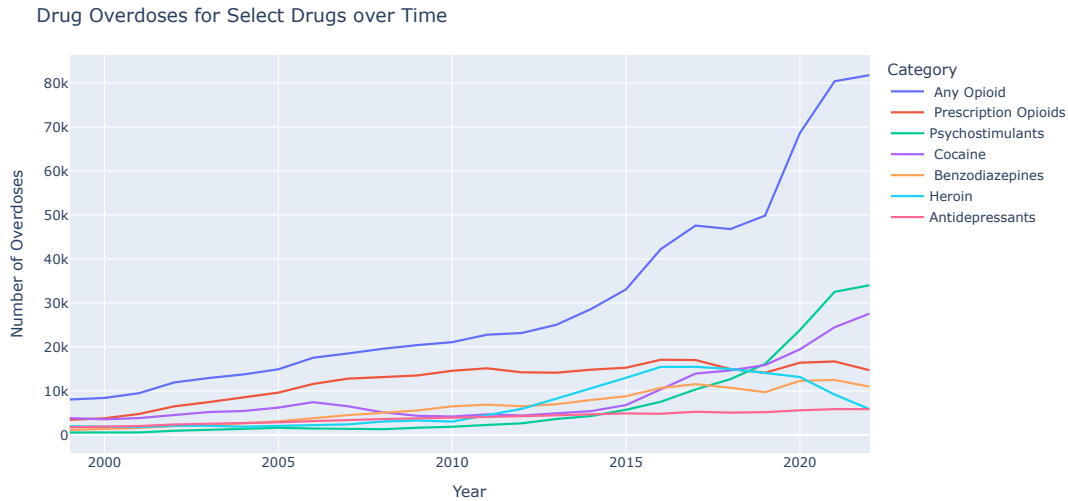
```

[168 rows x 3 columns]

```

[19]: fig = px.line(df_select_melted,
                  x='Year',
                  y='Overdoses',
                  color='Category',
                  width=1000)
fig.update_yaxes(title='Number of Overdoses')
fig.update_layout(title='Drug Overdoses for Select Drugs over Time')
fig

```



1.5 Total overdose rate and annual percent change

Melt the total-rate row, compute year-over-year percent change with `shift`, then dual y-axes. **Figure 7.44.**

```
[20]: df_tot_rate = df_demographic.loc[0:0, :]
df_tot_rate_melted = df_tot_rate.melt(id_vars='Category', value_vars=year_cols,
    ↪value_name='Overdose Rate', var_name='Year')
df_tot_rate_melted
```

```
[20]:
```

	Category	Year	Overdose Rate
0	Total Overdose Deaths	1999	6.1
1	Total Overdose Deaths	2000	6.2
2	Total Overdose Deaths	2001	6.8
3	Total Overdose Deaths	2002	8.2
..	...	...	...
20	Total Overdose Deaths	2019	21.6
21	Total Overdose Deaths	2020	28.3
22	Total Overdose Deaths	2021	32.4
23	Total Overdose Deaths	2022	32.6

[24 rows x 3 columns]

```
[21]: current_year = df_tot_rate_melted['Overdose Rate']
previous_year = df_tot_rate_melted['Overdose Rate'].shift(1)
df_tot_rate_melted['Percent Change'] = (current_year - previous_year) /
    ↪previous_year
df_tot_rate_melted
```

```
[21]:
```

	Category	Year	Overdose Rate	Percent Change
0	Total Overdose Deaths	1999	6.1	NaN
1	Total Overdose Deaths	2000	6.2	0.016393
2	Total Overdose Deaths	2001	6.8	0.096774
3	Total Overdose Deaths	2002	8.2	0.205882
..	...	...	...	...
20	Total Overdose Deaths	2019	21.6	0.043478
21	Total Overdose Deaths	2020	28.3	0.310185
22	Total Overdose Deaths	2021	32.4	0.144876
23	Total Overdose Deaths	2022	32.6	0.006173

[24 rows x 4 columns]

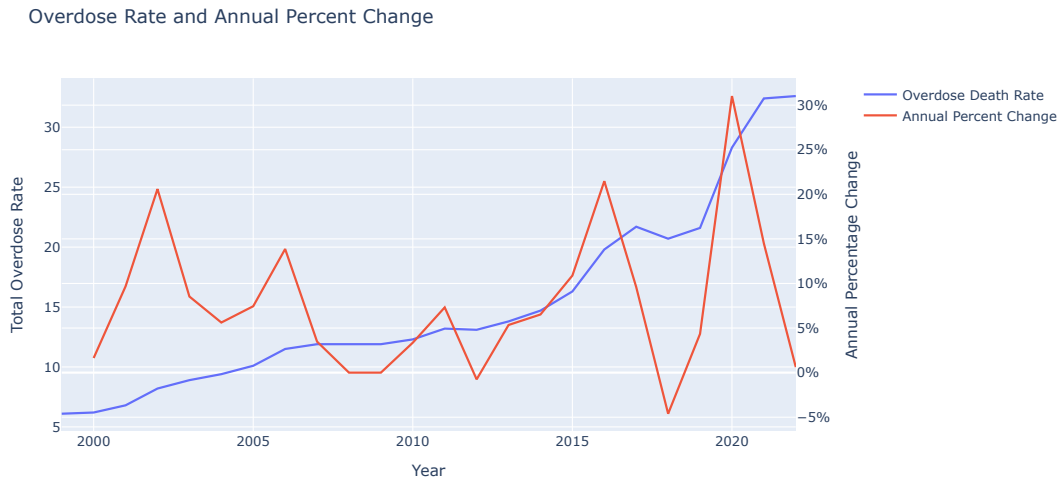
```
[22]: from plotly.subplots import make_subplots

fig = make_subplots(specs=[[{"secondary_y": True}]])

fig.add_trace(
    go.Scatter(x=df_tot_rate_melted['Year'],
               y=df_tot_rate_melted['Overdose Rate'],
               mode='lines',
               name='Overdose Death Rate'),
    secondary_y=False,
)

fig.add_trace(
    go.Scatter(x=df_tot_rate_melted['Year'],
               y=df_tot_rate_melted['Percent Change'],
               mode='lines',
               name='Annual Percent Change'),
    secondary_y=True,
)

fig.update_layout(title_text='Overdose Rate and Annual Percent Change',
                  width=1000)
fig.update_xaxes(title_text='Year')
fig.update_yaxes(title_text='Total Overdose Rate', secondary_y=False)
fig.update_yaxes(title_text='Annual Percentage Change', tickformat='.0%',
                  secondary_y=True)
fig
```



1.6 Overdose rates by race/ethnicity

Fold-change bar chart (**Figures 7.45–7.46**), then 2022 rates with a custom multiindex (**Figures 7.47–7.50**).

```
[23]: df_dem_select = df_demographic.loc[[0, 3, 6, 15, 18]]
      df_dem_select
```

```
[23]:
```

	Category	1999	2000	...	2021	\
0	Total Overdose Deaths	6.1	6.2	...	32.4	
3	White (Non-Hispanic)	6.2	6.6	...	36.8	
6	Black (Non-Hispanic)	7.5	7.3	...	44.2	
15	Hispanic	5.4	4.6	...	21.1	
18	American Indian or Alaska Native (Non-Hispanic)	6.0	5.5	...	56.5	

	2022	Fold Change 2015 to 2022
0	32.6	2.000000
3	35.6	1.687204
6	47.5	3.893443
15	22.7	2.948052
18	65.2	3.075472

[5 rows x 26 columns]

Figure 7.45 — First attempt (long category labels).

```
[24]: fig = px.bar(df_dem_select, x='Category',
                  y='Fold Change 2015 to 2022', color='Category', width=1000)
      fig
```

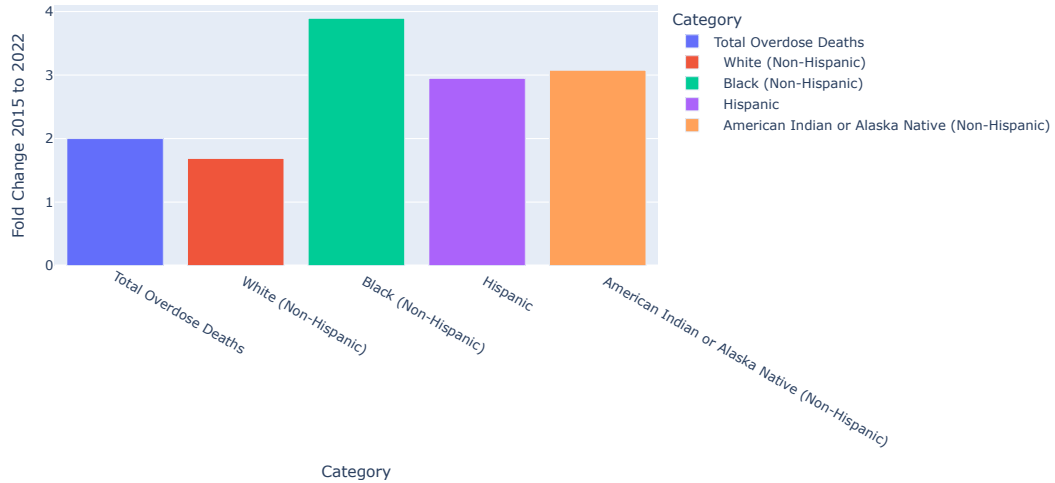
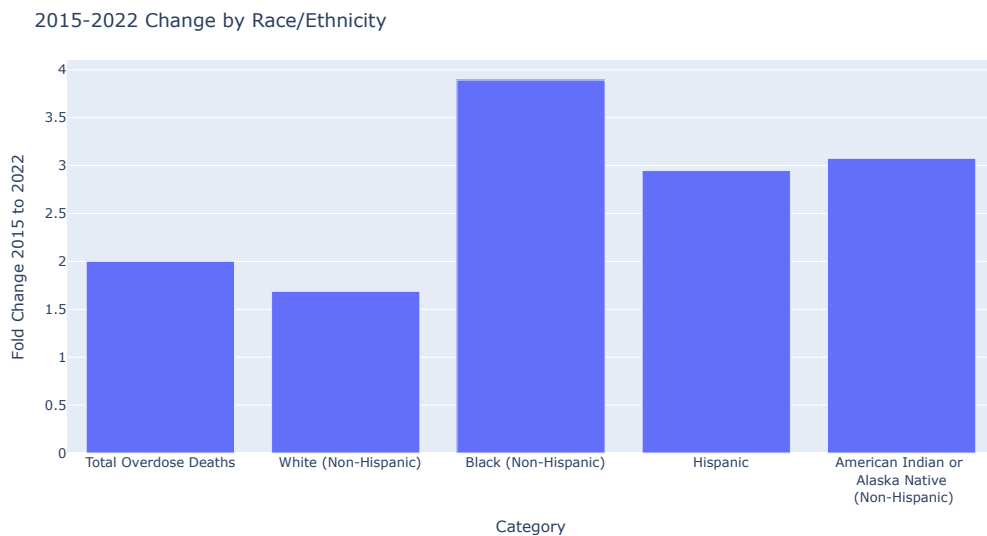


Figure 7.46 — Shorten labels; disable tick rotation and legend.

```
[25]: df_dem_select.loc[18, 'Category'] = 'American Indian or <br> Alaska Native <br>
      ↪(Non-Hispanic)'

fig = px.bar(df_dem_select,
             x='Category',
             y='Fold Change 2015 to 2022',
             width=1000)
fig.update_xaxes(tickangle=0)
fig.update_layout(title='2015-2022 Change by Race/Ethnicity')
fig
```



Build a race × gender multiindex for 2022 overdose rates.

```
[26]: df_dem_tot = df_demographic.head(21)
df_dem_tot
```

```
[26]:
```

		Category	1999	2000	...	\
0		Total Overdose Deaths	6.100	6.200	...	
1		Female	3.900	4.100	...	
2		Male	8.200	8.300	...	
3		White (Non-Hispanic)	6.200	6.600	...	
..		...	...	...	...	
17		Male	8.600	7.100	...	
18	American Indian or Alaska Native (Non-Hispanic)		6.000	5.500	...	
19		Female	5.200	4.300	...	
20		Male	6.704	6.661	...	

	2021	2022	Fold Change 2015 to 2022
0	32.4	32.6	2.000000
1	19.6	19.4	1.644068
2	45.1	45.6	2.192308
3	36.8	35.6	1.687204
..	...	...	...
17	32.4	35.2	3.229358
18	56.5	65.2	3.075472
19	44.1	47.1	2.803571
20	69.3	83.4	3.233185

[21 rows x 26 columns]

```
[27]: dem_cat = ['All', 'White',
                'Black', 'Asian',
                'Native Hawaiiin or <br> Other Pacific Islander',
                'Hispanic',
                'American Indian or <br> Alaska Native']

sub_cat = ['All', 'Female', 'Male']
idx = pd.MultiIndex.from_product([dem_cat, sub_cat], names=('Race', 'Gender'))
df_dem_tot = df_dem_tot.set_index(idx)
df_dem_tot.reset_index(inplace=True)
df_dem_tot
```

```
[27]:
```

	Race	Gender	\
0	All	All	
1	All	Female	
2	All	Male	

3		White	All
..		...	...
17		Hispanic	Male
18	American Indian or 	Alaska Native	All
19	American Indian or 	Alaska Native	Female
20	American Indian or 	Alaska Native	Male

		Category	...	2021	2022	\
0	Total Overdose Deaths	...	...	32.4	32.6	
1		Female	...	19.6	19.4	
2		Male	...	45.1	45.6	
3	White (Non-Hispanic)	...	...	36.8	35.6	
..		...	...	...	...	
17		Male	...	32.4	35.2	
18	American Indian or Alaska Native (Non-Hispanic)	...	...	56.5	65.2	
19		Female	...	44.1	47.1	
20		Male	...	69.3	83.4	

	Fold Change 2015 to 2022
0	2.000000
1	1.644068
2	2.192308
3	1.687204
..	...
17	3.229358
18	3.075472
19	2.803571
20	3.233185

[21 rows x 28 columns]

Figure 7.47 — Facet by race; color by gender.

```
[28]: fig = px.bar(df_dem_tot,
                 x='Gender',
                 y=2022,
                 facet_col='Race',
                 color='Gender',
                 labels={'Race': '', 'Gender': ''},
                 width=1000,
                 title='Overdose Rates by Race/Ethnicity (2022)')
fig.for_each_annotation(lambda a: a.update(text=a.text.split('=')[-1]))
# Make space for an annotation.
fig.update_layout(margin=dict(l=20, r=20, t=60, b=220))

# Add the annotation.
fig.add_annotation(dict(font=dict(color='black', size=13),
```

```

x=0,
y=0,
xref="paper",
yref="paper",
xanchor='left',
yanchor='top',
yshift=-40,
showarrow=False,
text="Note: Categories other than Hispanic and All do
↳not include Hispanics.",
textangle=0))

fig.update(layout_showlegend=False)
fig

```

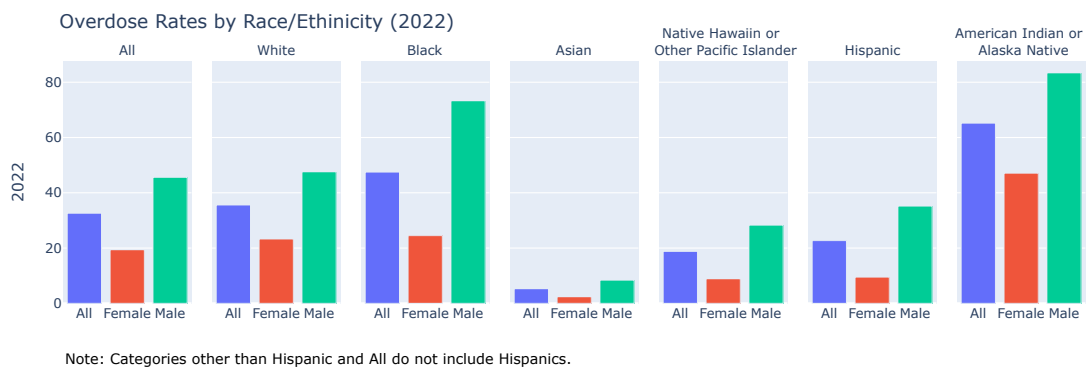


Figure 7.48 — Grouped bars: race on x-axis, color by gender.

```

[29]: fig = px.bar(df_dem_tot,
                x='Race',
                color='Gender',
                y=2022,
                barmode='group',
                width=1000,
                title='Overdose Rates by Race/Ethnicity (2022)')
# Make space for an annotation.
fig.update_layout(margin=dict(l=20, r=20, t=60, b=220))

# Add the annotation.

```

```
fig.add_annotation(dict(font=dict(color='black', size=13),
                        x=0,
                        y=0,
                        xref="paper",
                        yref="paper",
                        xanchor='left',
                        yanchor='top',
                        yshift=-40,
                        showarrow=False,
                        text="Note: Categories other than Hispanic and All do_
↳not include Hispanics.",
                        textangle=0))

fig
```

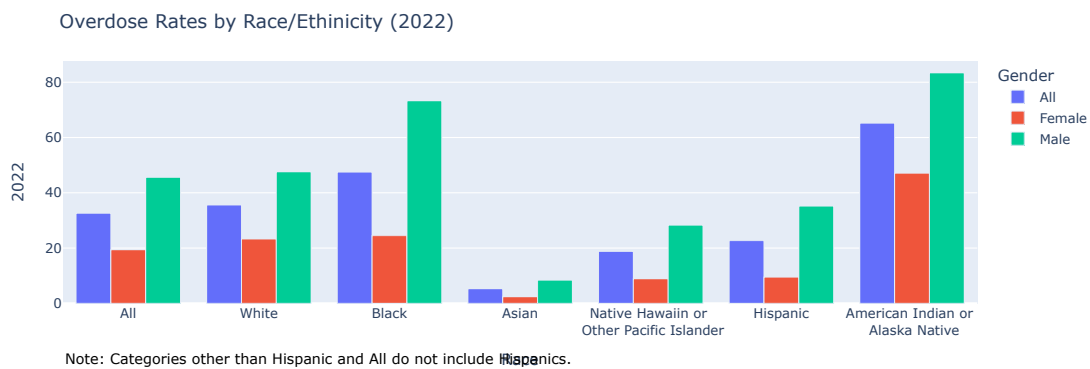


Figure 7.49 — Grouped bars: gender on x-axis, color by race.

```
[30]: fig = px.bar(df_dem_tot,
                  x='Gender',
                  color='Race',
                  y=2022,
                  barmode='group',
                  width=1000,
                  title='Overdose Rates by Race/Ethnicity (2022)')
# Make space for an annotation.
fig.update_layout(margin=dict(l=20, r=20, t=60, b=220))

# Add the annotation.
```

```
fig.add_annotation(dict(font=dict(color='black', size=13),
                        x=0,
                        y=0,
                        xref="paper",
                        yref="paper",
                        xanchor='left',
                        yanchor='top',
                        yshift=-40,
                        showarrow=False,
                        text="Note: Categories other than Hispanic and All do not include Hispanics.",
                        textangle=0))

fig
```

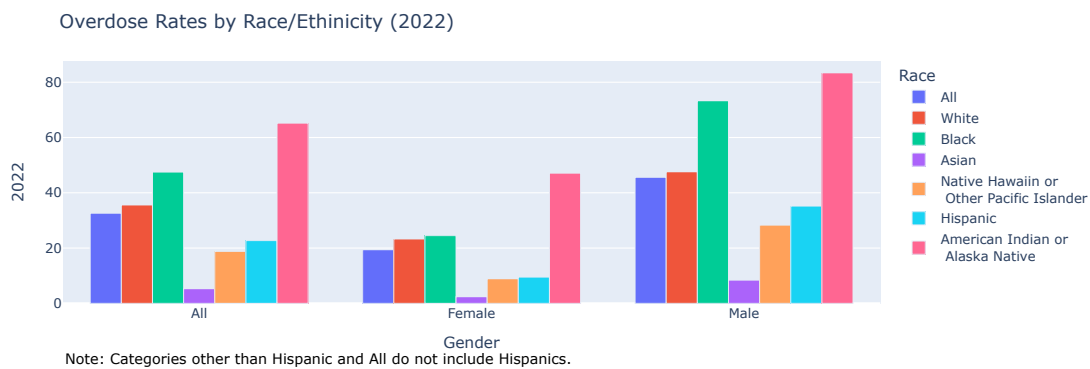


Figure 7.50 — Graph objects with multi-level x-axis ($x=[\text{Race}, \text{Gender}]$).

```
[31]: x = [df_dem_tot['Race'], df_dem_tot['Gender']]
color_map = {'All': 'brown', 'Female': 'indigo', 'Male': 'steelblue'}
colors = [color_map[g] for g in df_dem_tot['Gender']]

fig = go.Figure()
fig.add_bar(x=x, y=df_dem_tot[2022], marker_color=colors)
fig.update_layout(barmode='group', title='Overdose Rates by Race/Ethnicity',
                  width=1150)
# Make space for an annotation.
fig.update_layout(margin=dict(l=20, r=20, t=60, b=220))
```

```
# Add the annotation.
fig.add_annotation(dict(font=dict(color='black', size=13),
                        x=0,
                        y=0,
                        xref="paper",
                        yref="paper",
                        xanchor='left',
                        yanchor='top',
                        yshift=-40,
                        showarrow=False,
                        text="Note: Categories other than Hispanic and All do
↳not include Hispanics.",
                        textangle=0))

fig
```

