

CH04_COVID_Python_Supplement

September 3, 2026

1 Chapter 4 — COVID vaccination data (Python supplement)

Condensed notebook for the **Python / Plotly** portion of *Chapter 4: COVID Data*.

Target graphics:

- Absolute death risk by age and vaccination status (**Figures 4.26–4.27**)
- Relative risk (**Figure 4.28**)
- Combined death rate by vaccine status only (**Figure 4.29**)

Data: Covid.csv in Data For Condensed Notebooks.

Dependencies:

- pandas
- plotly

1.1 Imports and display options

```
[1]: import pandas as pd
import plotly.express as px
```

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

```
[3]: pd.options.mode.copy_on_write = True
```

```
/var/folders/_x/nbw2t83x75d0lr3jn4pg_cx00000gn/T/ipykernel_25991/2126689829.py:1
: Pandas4Warning: The 'mode.copy_on_write' option is deprecated. Copy-on-Write
can no longer be disabled (it is always enabled with pandas >= 3.0), and setting
the option has no impact. This option will be removed in pandas 4.0.
pd.options.mode.copy_on_write = True
```

1.2 Loading the data

```
[4]: data_link = '../Data For Condensed Notebooks/Covid.csv'
df = pd.read_csv(data_link)
```

1.3 Grouping and reshaping

groupby → value_counts → unstack outcomes; compute rates with crosstab.

```
[5]: grouped = df.groupby(['age_group', 'vaccine_status', 'outcome'])
grouped = pd.DataFrame(df.groupby(['age_group', 'vaccine_status', 'outcome']).
    ↪value_counts())
grouped = grouped.unstack(level=2)
grouped.columns = grouped.columns.get_level_values(1)
```

```
[6]: rates = pd.crosstab(index=df['age_group'], df['vaccine_status'],
    columns=df['outcome'],
    normalize='index')
rates.columns = ['death_rate', 'survival_rate']
risk = rates.unstack(level=1)
death_risk = risk['death_rate']
death_risk['relative_risk'] = death_risk['unvaccinated'] /
    ↪death_risk['vaccinated']
death_risk
```

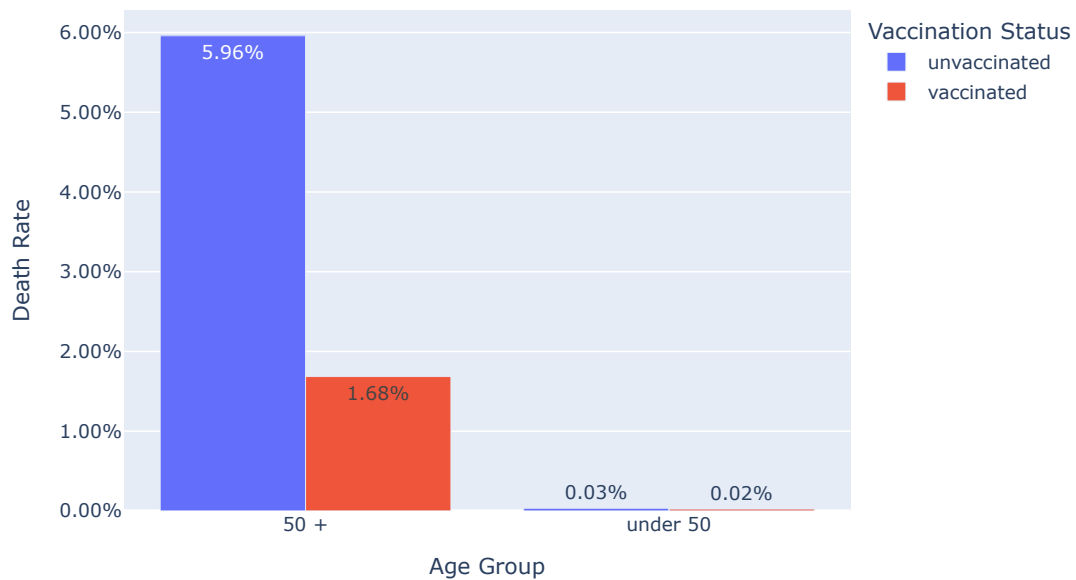
```
[6]: vaccine_status  unvaccinated  vaccinated  relative_risk
age_group
50 +                0.059593      0.016845      3.537623
under 50            0.000325      0.000234      1.390626
```

1.4 Absolute risk bar charts

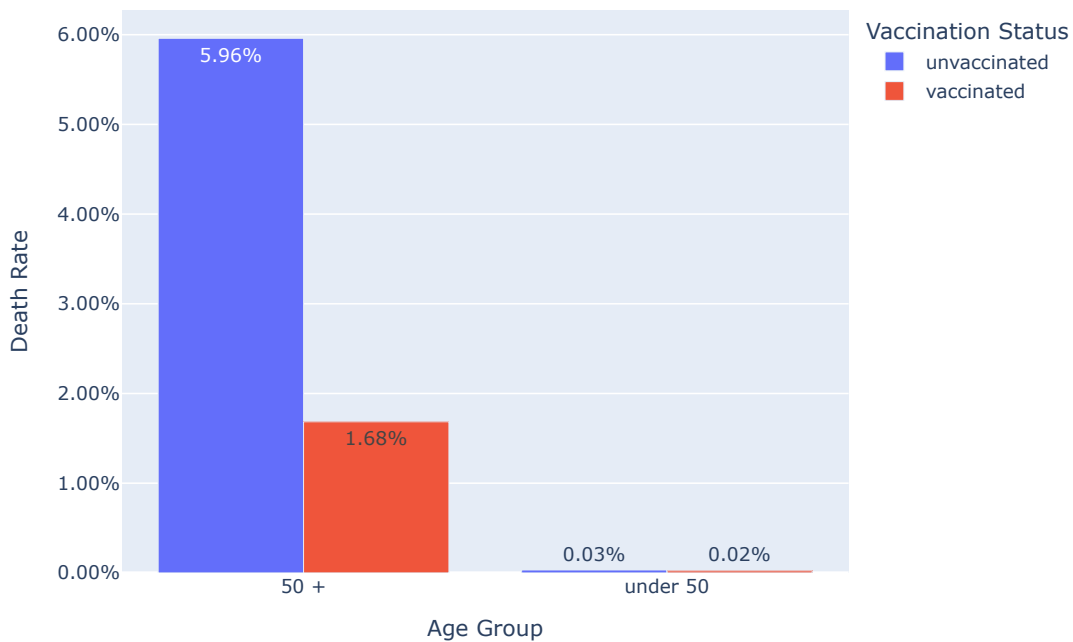
Wide form, then long form after melt. **Figure 4.26** (wide); **Figure 4.27** (long).

```
[7]: abs_risk = death_risk[['unvaccinated', 'vaccinated']]
fig = px.bar(abs_risk,
    text=['unvaccinated', 'vaccinated'],
    labels={'value': 'Death Rate',
    'age_group': 'Age Group',
    'vaccine_status': 'Vaccination Status'},
    barmode='group',
    text_auto='.2%',
    title='Death Rate by Age Group and Vaccination Status')
fig.update_yaxes(tickformat='.2%')
fig
```

Death Rate by Age Group and Vaccination Status



```
[8]: abs_risk_long = pd.DataFrame(abs_risk.stack(level=-1))
abs_risk_long.reset_index(inplace=True)
abs_risk_long.columns = ['age_group', 'vaccination_status', 'percent_dead']
fig = px.bar(abs_risk_long, x='age_group',
             y='percent_dead',
             color='vaccination_status',
             text='percent_dead',
             barmode='group',
             labels={'percent_dead': 'Death Rate',
                    'age_group': 'Age Group',
                    'vaccination_status': 'Vaccination Status'},
             text_auto='.2%')
fig.update_yaxes(tickformat='.2%')
fig
```



1.5 Relative risk bar chart

Reference line at 1 with annotation boxes. **Figure 4.28.**

```
[9]: relative_risk = death_risk[['relative_risk']]
```

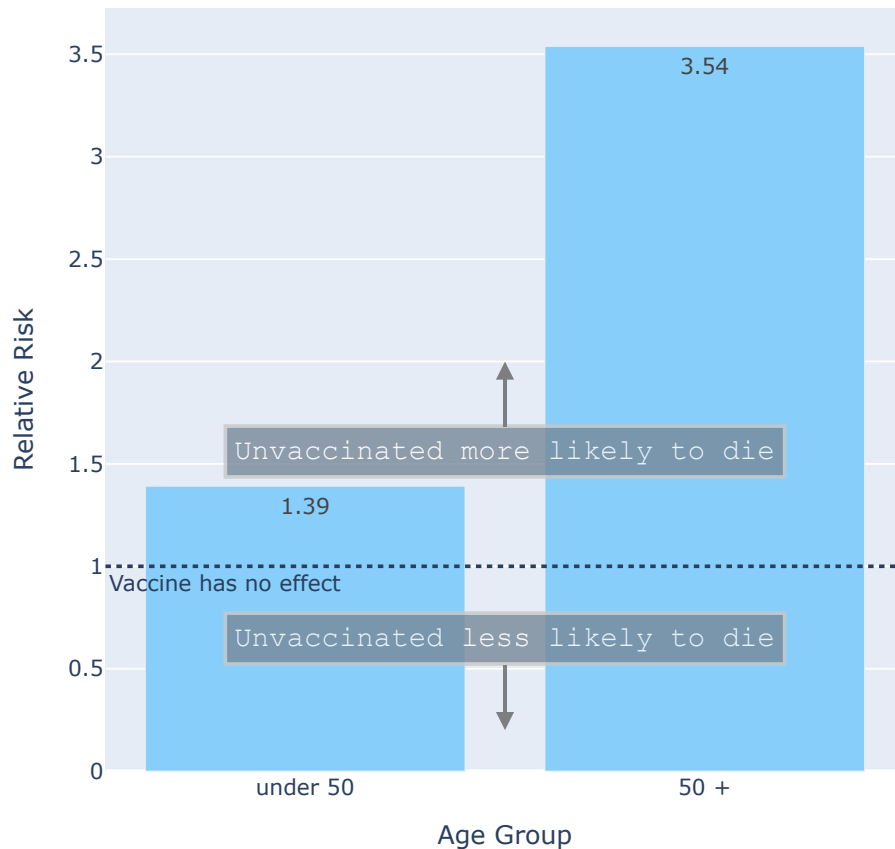
```
[10]: fig = px.bar(relative_risk,
                  x=relative_risk.index,
                  y='relative_risk',
                  text='relative_risk',
                  text_auto='.3',
                  labels={'age_group': 'Age Group',
                          'relative_risk': 'Relative Risk'},
                  title='Relative Risk for Unvaccinated People',
                  height=600,
                  width=600)
fig.update_traces(marker_color='lightskyblue')
fig.add_hline(y=1, line_dash='dot',
              annotation_text='Vaccine has no effect',
              annotation_position='bottom left')
fig.add_annotation(x=0.5, y=2,
                  text='Unvaccinated more likely to die',
```

```

        showarrow=True,
        arrowhead=2,
        arrowsize=1,
        arrowwidth=2,
        arrowcolor='#636363',
        ax=0,
        ay=50,
        font=dict(family='Courier New, monospace', size=16,
↪color='#ffffff'),
        bordercolor='#c7c7c7',
        borderwidth=2,
        borderpad=4,
        bgcolor='lightslategray',
        opacity=0.8)
fig.add_annotation(x=0.5, y=0.2,
        text='Unvaccinated less likely to die',
        showarrow=True,
        arrowhead=2,
        arrowsize=1,
        arrowwidth=2,
        arrowcolor='#636363',
        ax=0,
        ay=-50,
        font=dict(family='Courier New, monospace', size=16,
↪color='#ffffff'),
        bordercolor='#c7c7c7',
        borderwidth=2,
        borderpad=4,
        bgcolor='lightslategray',
        opacity=0.8)
fig.update_xaxes(categoryorder='array', categoryarray=['under 50', '50 +'])
fig

```

Relative Risk for Unvaccinated People



1.6 Death rate by vaccine status only

Aggregate across age groups. Contrasts with the stratified charts above. **Figure 4.29.**

```
[11]: combined_rates = pd.crosstab(index=df['vaccine_status'], columns=df['outcome'],
    ↪normalize='index')
combined_rates.columns = ['death_rate', 'survival_rate']
fig = px.bar(combined_rates[['death_rate']],
             labels={'value': 'Death Rate',
                    'vaccine_status': 'Vaccine Status'},
             text_auto='.2%',
             title='Death Rate by Vaccine Status',
             height=600,
```

```
width=400)  
fig.update_layout(showlegend=False)  
fig.update_yaxes(tickformat='.2%')  
fig
```

Death Rate by Vaccine Status

