**Analysing the Yearly Data Split By Clinic**

**Welcome to your workplace...**

There are two maternity wards at the Vienna General Hospital: clinic 1 and clinic 2. Clinic 1 was staffed by all-male doctors and medical students, and clinic 2 was staffed by female midwives.



Challenge 1: The Yearly Data Split by Clinic

Let's turn our attention to the annual data. Use plotly to create line charts of the births and deaths of the two different clinics at the Vienna General Hospital.

* Which clinic is bigger or more busy judging by the number of births?
* Has the hospital had more patients over time?
* What was the highest number of deaths recorded in clinic 1 and clinic 2?

.

.

..

...

..

.

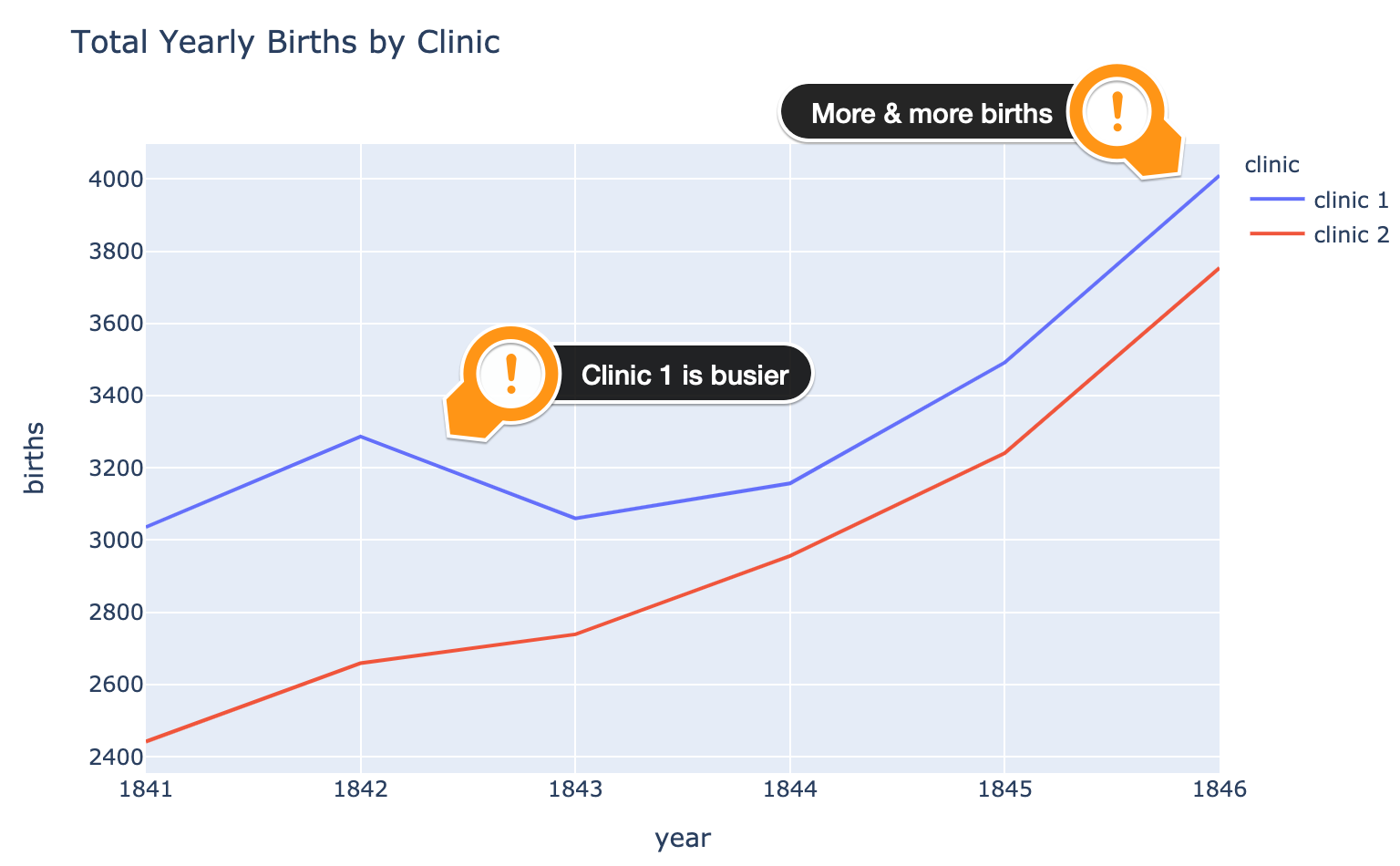
.

**Solution to Challenge 1**

To show two line charts side by side we can use plotly and provide the clinic column as the color.

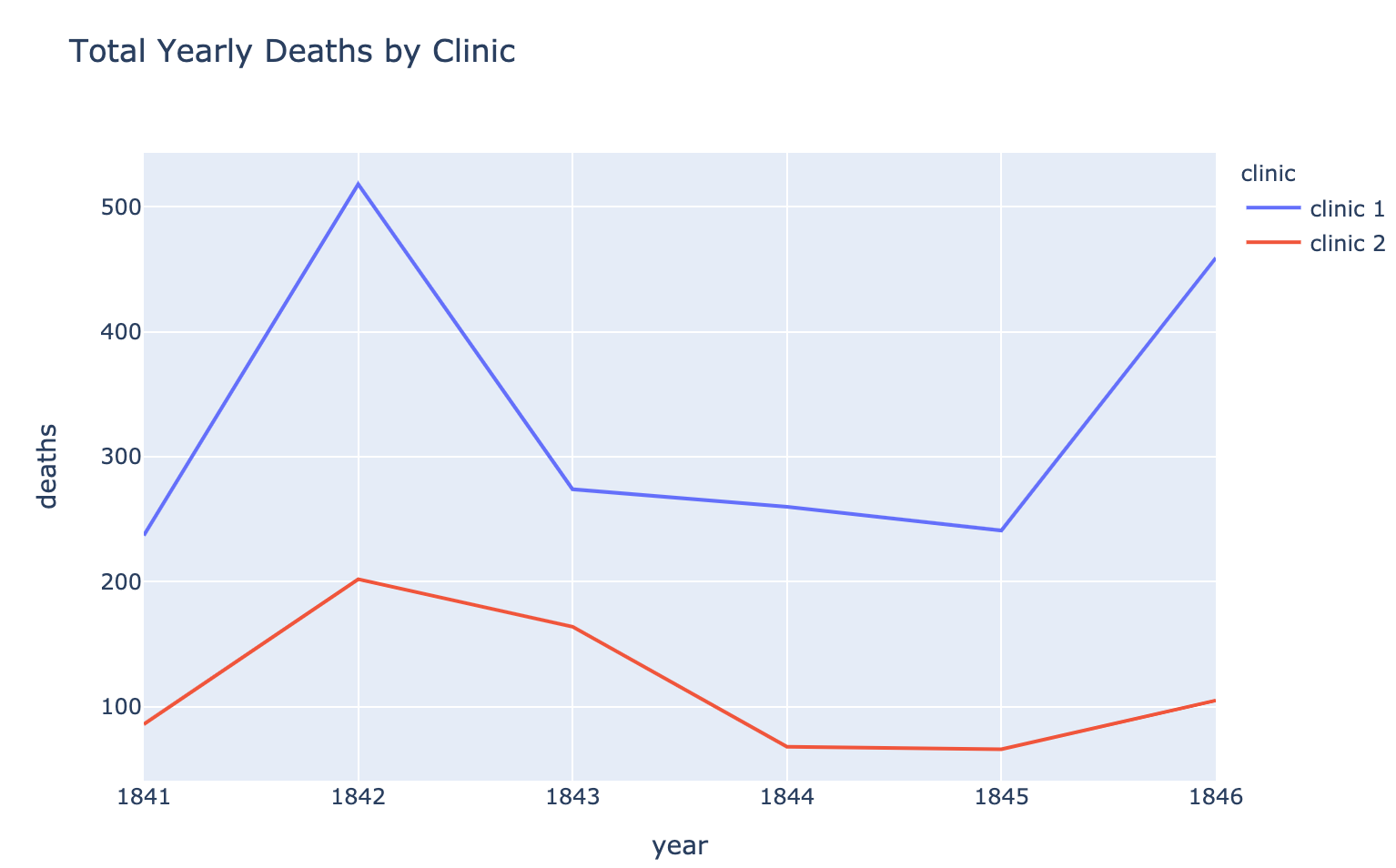
1. line = px.line(df\_yearly,
2. x='year',
3. y='births',
4. color='clinic',
5. title='Total Yearly Births by Clinic')
7. line.show()

We see that more and more women gave birth at the hospital over the years. Clinic 1, which was staffed by male doctors and medical students was also the busier or simply the larger ward. More births took place in clinic 1 than in clinic 2.



We also see that, not only were more people born in clinic 1, more people also died in clinic 1.

1. line = px.line(df\_yearly,
2. x='year',
3. y='deaths',
4. color='clinic',
5. title='Total Yearly Deaths by Clinic')
7. line.show()



To compare apples and apples, we need to look at the proportion of deaths per clinic.

Challenge 2: Calculate the Proportion of Deaths at Each Clinic

Calculate the proportion of maternal deaths per clinic. That way we can compare like with like.

* Work out the percentage of deaths for each row in the df\_yearly DataFrame by adding a column called "pct\_deaths".
* Calculate the average maternal death rate for clinic 1 and clinic 2 (i.e., the total number of deaths per the total number of births).
* Create another plotly line chart to see how the percentage varies year over year with the two different clinics.
* Which clinic has a higher proportion of deaths?
* What is the highest monthly death rate in clinic 1 compared to clinic 2?

.

.

..

...

..

.

.

**Solution to Challenge 2**

We can add a new column that has the percentage of deaths for each row like this:

1. df\_yearly['pct\_deaths'] = df\_yearly.deaths / df\_yearly.births

The average death rate for the entire time period for clinic 1 is:

1. clinic\_1 = df\_yearly[df\_yearly.clinic == 'clinic 1']
2. avg\_c1 = clinic\_1.deaths.sum() / clinic\_1.births.sum() \* 100
3. print(f'Average death rate in clinic 1 is {avg\_c1:.3}%.')

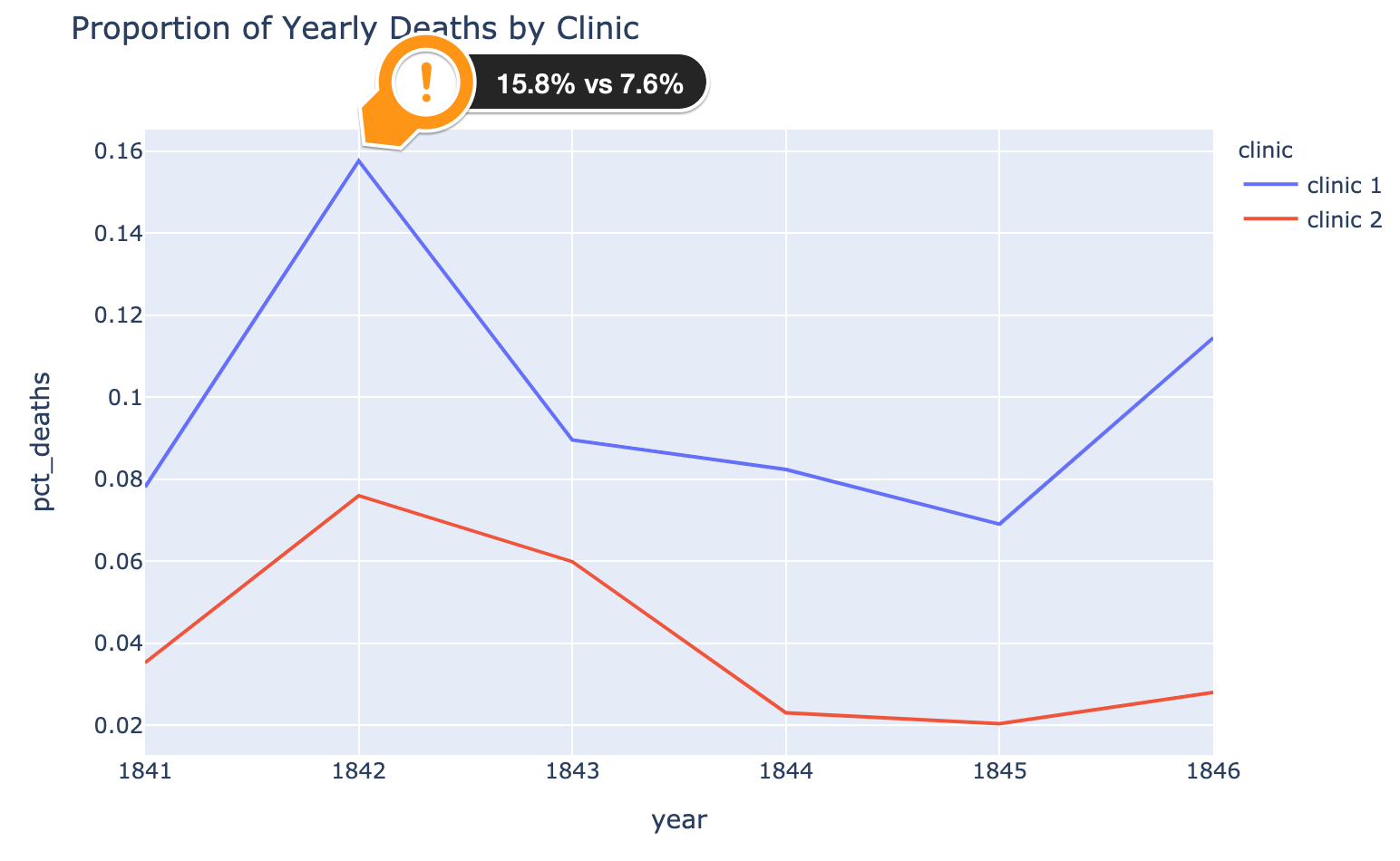
9.92%. In comparison, clinic 2 which was staffed by midwives had a much lower death rate of 3.88% over the course of the entire period. Hmm... 🤔

1. clinic\_2 = df\_yearly[df\_yearly.clinic == 'clinic 2']
2. avg\_c2 = clinic\_2.deaths.sum() / clinic\_2.births.sum() \* 100
3. print(f'Average death rate in clinic 2 is {avg\_c2:.3}%.')

Once again, let's see this on a chart

1. line = px.line(df\_yearly,
2. x='year',
3. y='pct\_deaths',
4. color='clinic',
5. title='Proportion of Yearly Deaths by Clinic')
7. line.show()

1842 was a rough year. About 16% of women died in clinic 1 and about 7.6% of women died in clinic 2.



Still, clinic 2 had a consistently lower death rate than clinic 1! This is what puzzled and frustrated Dr Semmelweis.

**The story continues...**

At first, Dr Semmelweis thought that the position of the women giving birth was the issue. In clinic 2, the midwives' clinic, women gave birth on their sides. In the doctors' clinic, women gave birth on their backs. So, Dr. Semmelweis, had women in the doctors' clinic give birth on their sides. However, this had no effect on the death rate.

Next, Dr Semmelweis noticed that whenever someone on the ward died, a priest would walk through clinic 1, past the women's beds ringing a bell 🔔. Perhaps the priest and the bell ringing terrified the women so much after birth that they developed a fever, got sick and died. Dr Semmelweis had the priest change his route and stop ringing the bell 🔕. Again, this had no effect.

At this point, Dr Semmelweis was so frustrated he went on holiday to Venice. Perhaps a short break would clear his head. When Semmelweis returned from his vacation, he was told that one of his colleagues, a pathologist, had fallen ill and died. His friend had pricked his finger while doing an autopsy on a woman who had died from childbed fever and subsequently got very sick himself and died. 😮

Looking at the pathologist's symptoms, Semmelweis realised the pathologist died from the same thing as the women he had autopsied.  This was his breakthrough: anyone could get sick from childbed fever, not just women giving birth!

This is what led to Semmelweis' new theory. Perhaps there were little pieces or particles of a corpse that the doctors and medical students were getting on their hands while dissecting the cadavers during an autopsy. And when the doctors delivered the babies in clinic 1, these particles would get inside the women giving birth who would then develop the disease and die.