

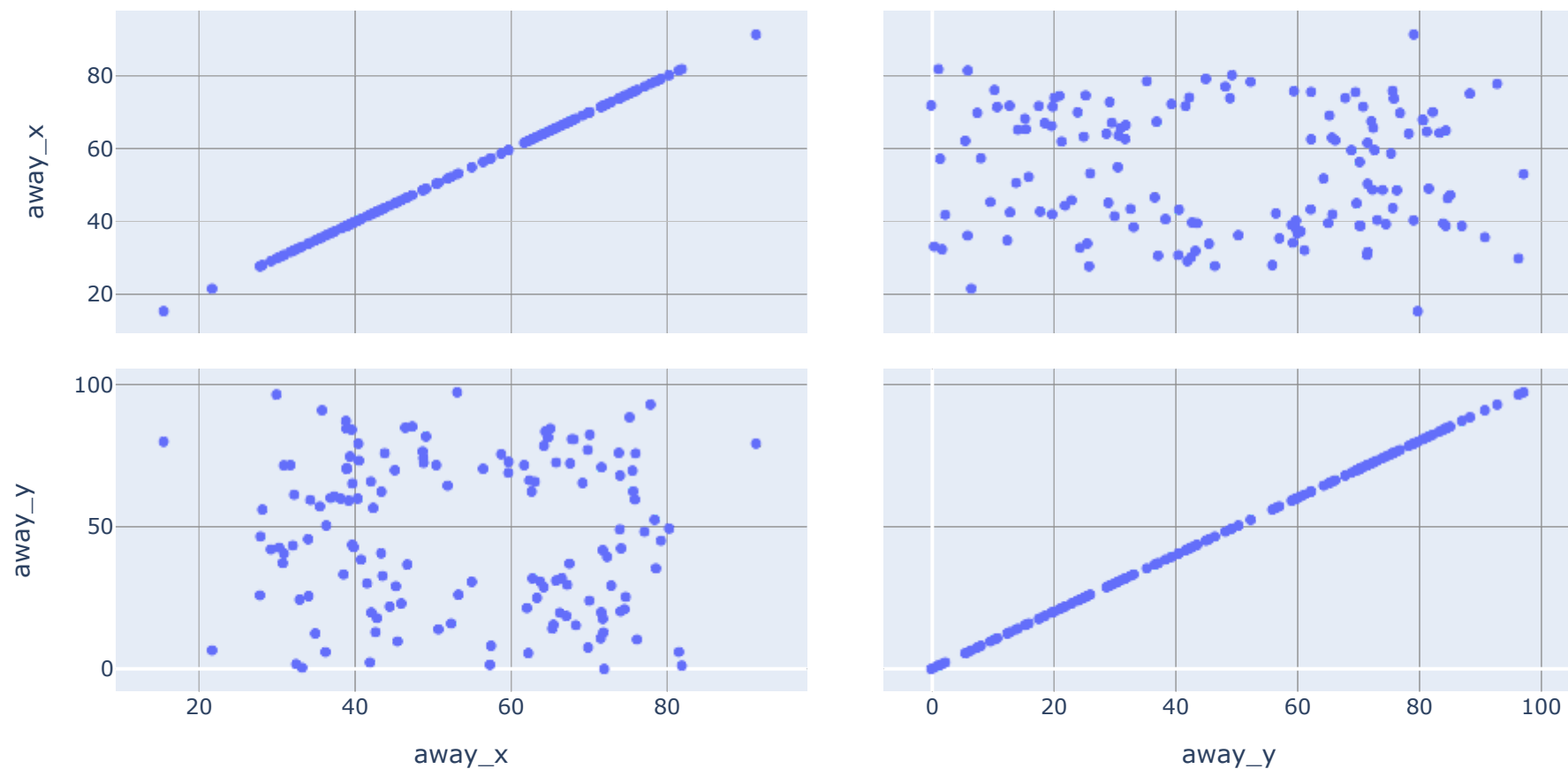
In [1]: # Problem 5

```
import pandas as pd
import plotly.express as px
from numpy import cov
data = pd.read_csv('/Users/nivethida/Downloads/same(1).csv')

print("Corelation coefficient for away_x and away_y: ", cov(data['away_x'], data['away_y']))
columns = data[['away_x', 'away_y']]
fig = px.scatter_matrix(columns)
fig.show()

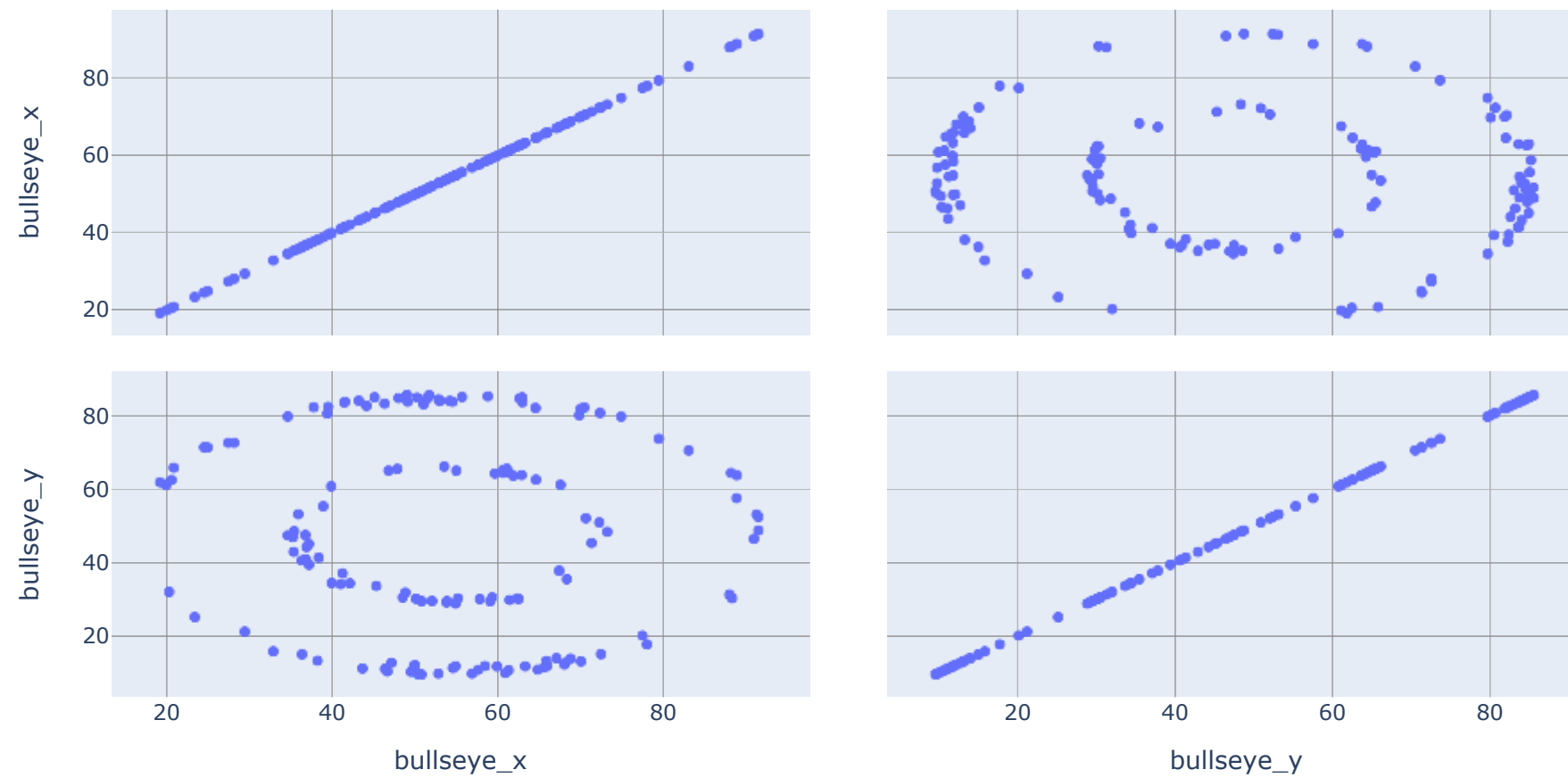
# The from the below results we sould see that corelation coefficient is almost same for the variables
# But the scatter plot is different which explains the Anscombe's Quartet.
```

Corelation coefficient for away\_x and away\_y: [[281.22702899 281.22702899]  
[281.22702899 281.22702899]]



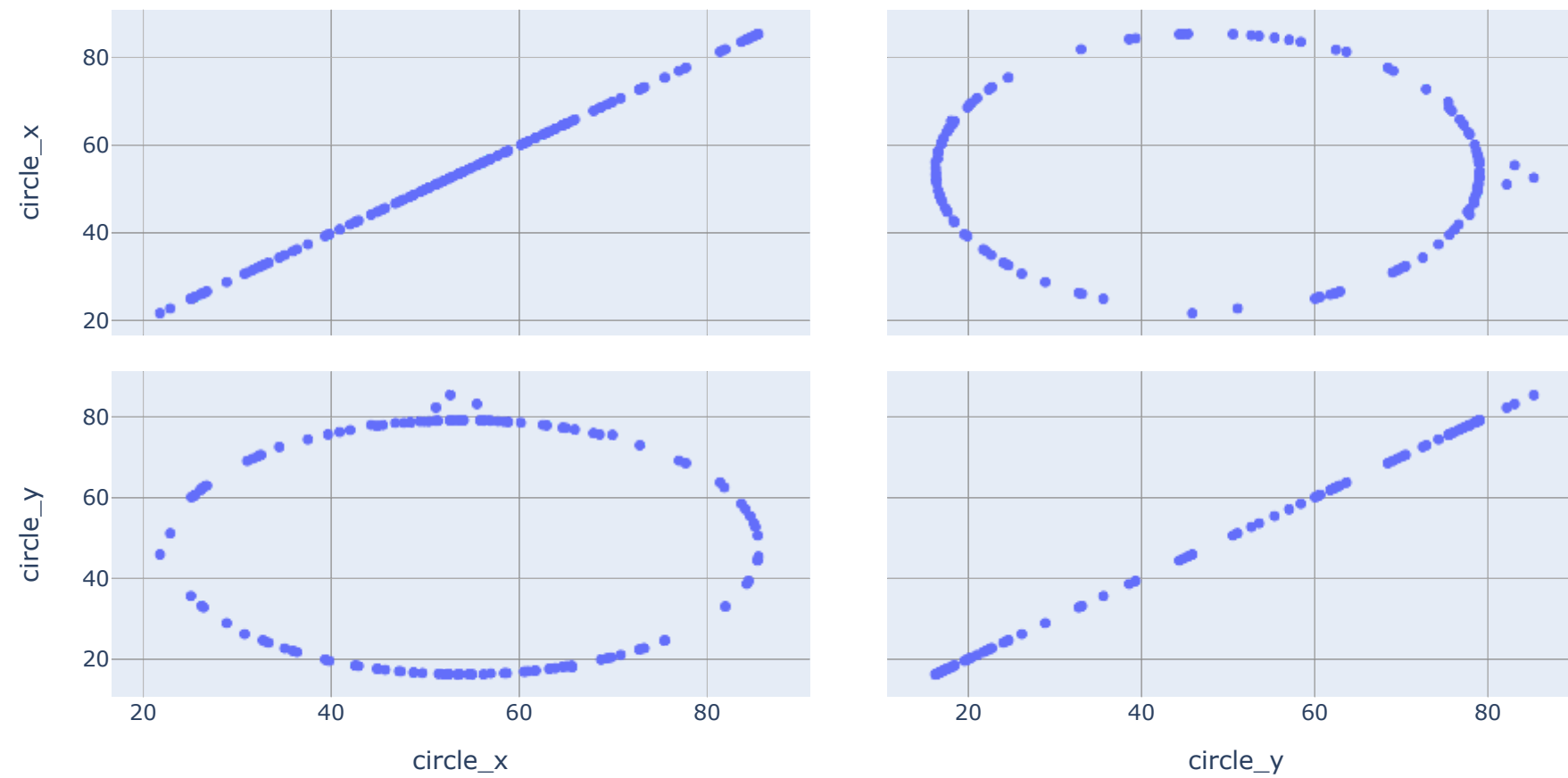
```
In [2]: print("Corelation coefficient for bullseye_x and bullseye_y: ", cov(data['bullseye_x'], data['bullseye_y']))
columns1 = data[['bullseye_x', 'bullseye_y']]
fig = px.scatter_matrix(columns1)
fig.show()
```

Corelation coefficient for bullseye\_x and bullseye\_y:  $\begin{bmatrix} 281.20739319 & -30.97990171 \\ -30.97990171 & 725.53337231 \end{bmatrix}$



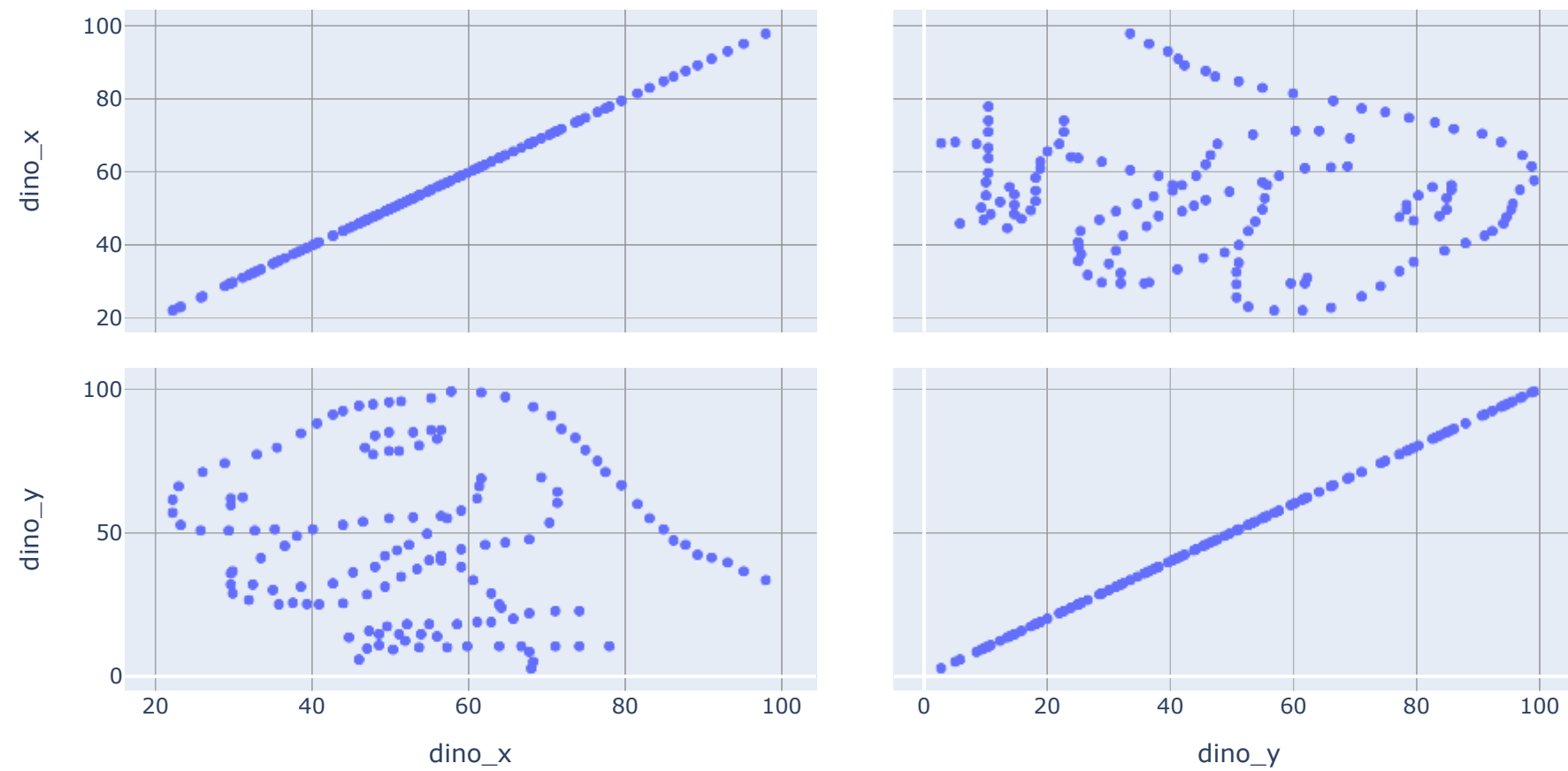
```
In [3]: print("Corelation coefficient for circle_x and circle_y: ", cov(data['circle_x'], data['circle_y']))
columns1 = data[['circle_x', 'circle_y']]
fig = px.scatter_matrix(columns1)
fig.show()
```

```
Corelation coefficient for circle_x and circle_y: [[280.89802436 -30.84661989]
[-30.84661989 725.22684369]]
```



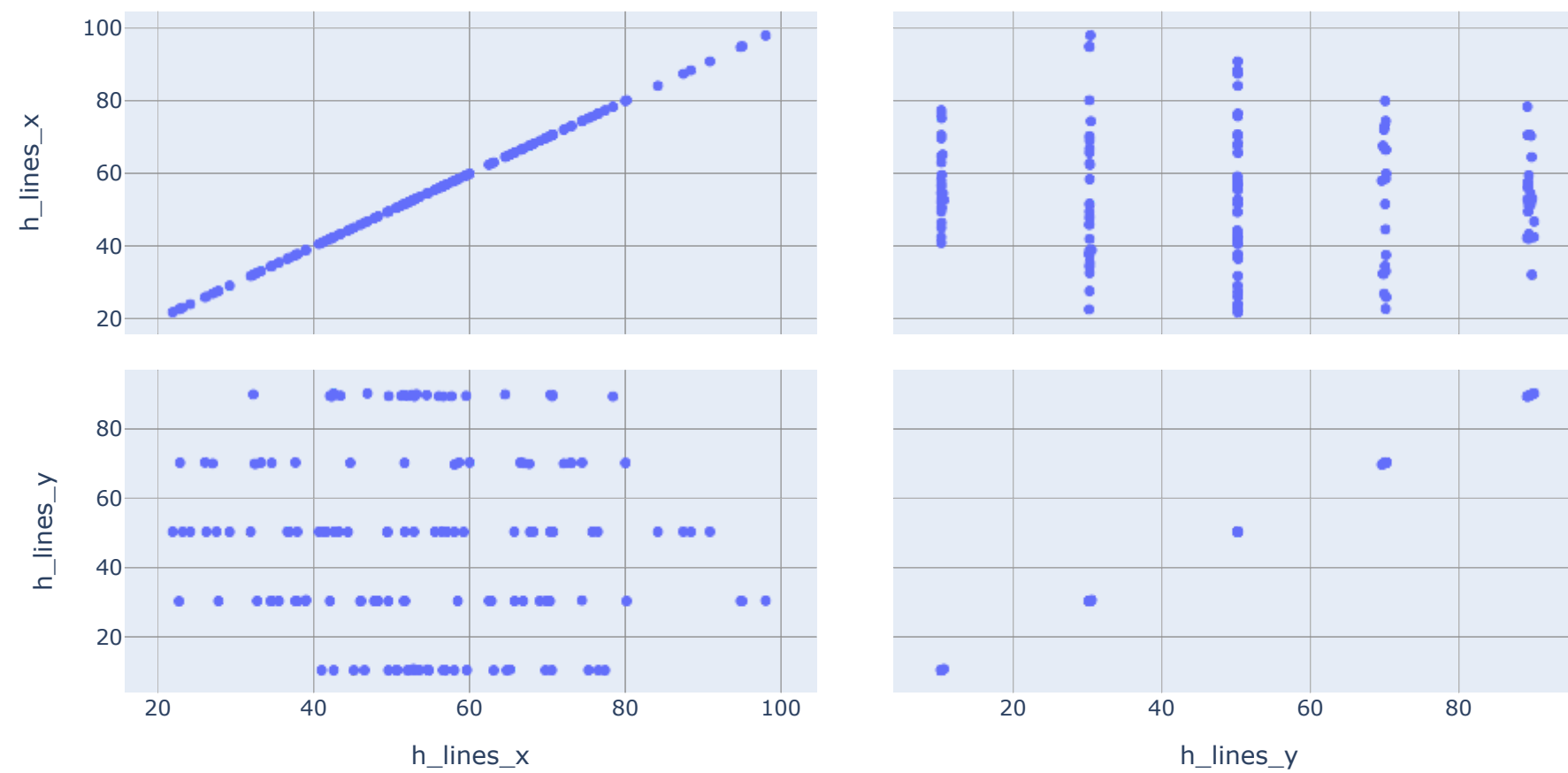
```
In [4]: print("Corelation coefficient for dino_x and dino_y: ", cov(data['dino_x'], data['dino_y']))
columns1 = data[['dino_x', 'dino_y']]
fig = px.scatter_matrix(columns1)
fig.show()
```

```
Corelation coefficient for dino_x and dino_y: [[281.06998759 -29.11393267]
[-29.11393267 725.515961  ]]
```



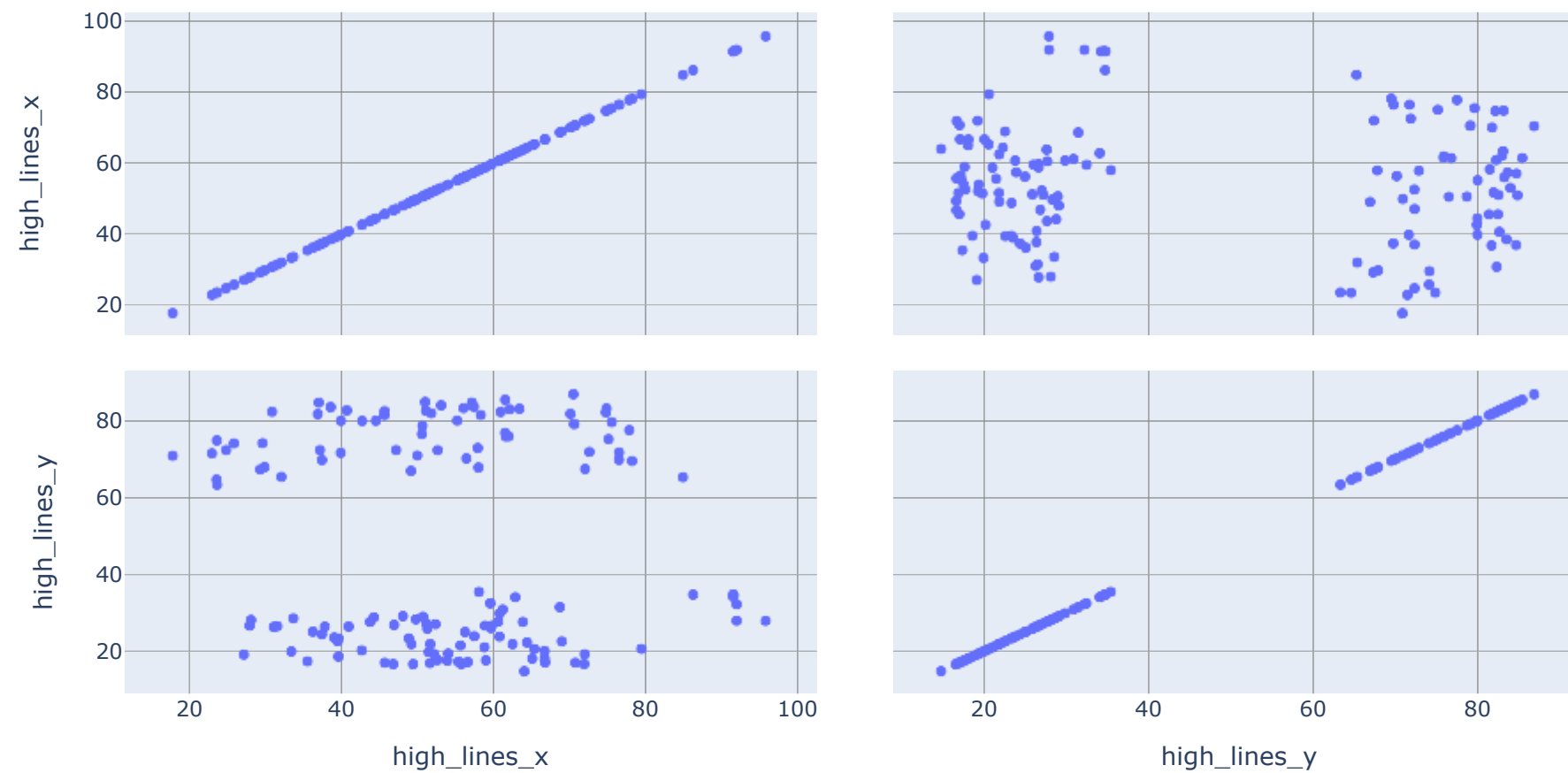
```
In [33]: print("Corelation coefficient for h_lines_x and h_lines_y: ", cov(data['h_lines_x'], data['h_lines_y']))
columns1 = data[['h_lines_x', 'h_lines_y']]
fig = px.scatter_matrix(columns1)
fig.show()
```

Corelation coefficient for h\_lines\_x and h\_lines\_y:  $\begin{bmatrix} 281.09533252 & -27.8748158 \\ -27.8748158 & 725.75693077 \end{bmatrix}$



```
In [34]: print("Corelation coefficient for high_lines_x and high_lines_y: ", cov(data['high_lines_x'], data['high_lines_y']))  
columns1 = data[['high_lines_x', 'high_lines_y']]  
fig = px.scatter_matrix(columns1)  
fig.show()
```

Corelation coefficient for high\_lines\_x and high\_lines\_y:  $\begin{bmatrix} 281.12236356 & -30.94301196 \\ -30.94301196 & 725.76349016 \end{bmatrix}$



In [ ]: