For some reason my pivot table has "Year" in the column even though I didn't specify that in the column or index arguments and it's adding the years together instead of the number of medals. Does anyone know how to fix this? My pivot table looks like below

                                    Year

                Medal           Bronze   Gold     Silver

Country    Sport

AFG         Taekwondo     4020     0           0

Hi Joe,

Thanks so much! I have a problem on your problem and on problem 3.

For problem 3, for some reason my pivot table has "Year" in the column even though I didn't specify that in the column or index arguments and it's adding the years together instead of the number of medals.

Year

Medal Bronze Gold Silver

Country Sport

AFG Taekwondo 4020 0 0

This is the code I have for it: olympics.pivot\_table(index = ["Country", "Sport"], columns = "Medal", fill\_value = 0, aggfunc = np.sum)

I can't find anything online about why a variable would show a variable that wasn't specified in the index and columns. I've tried searching for things like "column showing up in pivot table that shouldn't be pandas" but am getting no good results. Is there a resource or something you can point me to? This never happened with any of the pivot tables we did in class.

The other part I'm having trouble with in question 3 is part 5.2. On piazza people said it should be a histogram, not a bar graph, but when you do histogram it does the frequency of the numbers and not the country and the number of medals they got. I've played with doing stuff like x = olympics\_pivot2.index, but am having no luck

# https://plot.ly/python/histograms/

# https://plot.ly/python/histograms/

trace0 = go.Histogram(y = olympics\_pivot2["Medal"]["Summer"], name = "Summer")

trace1 = go.Histogram(y = olympics\_pivot2["Medal"]["Winter"], name = "Winter")

data = [trace0, trace1]

layout = go.Layout(barmode = "stack")

fig = go.Figure(data = data, layout = layout)

plotly.offline.iplot(fig, filename = "stacked histogram")

Lastly, on your problem for add\_percentiles(df) I get a Type Error saying: ('cannot operate on a series without a rhs of a series/ndarray of type datetime64[ns] or a timedelta', 'occurred at index Close-Apple')

My code is : data\_frame["Apple percentile"] = data\_frame.apply(lambda x: percents(x[0], data\_frame["Date"], data\_frame[["Close-Apple", "Close-Tesla", "Close-Netflix"]]))

I haven't done the Tesla or netflix percentile because I was trying to just get one of them to work to start.