Visualization 3

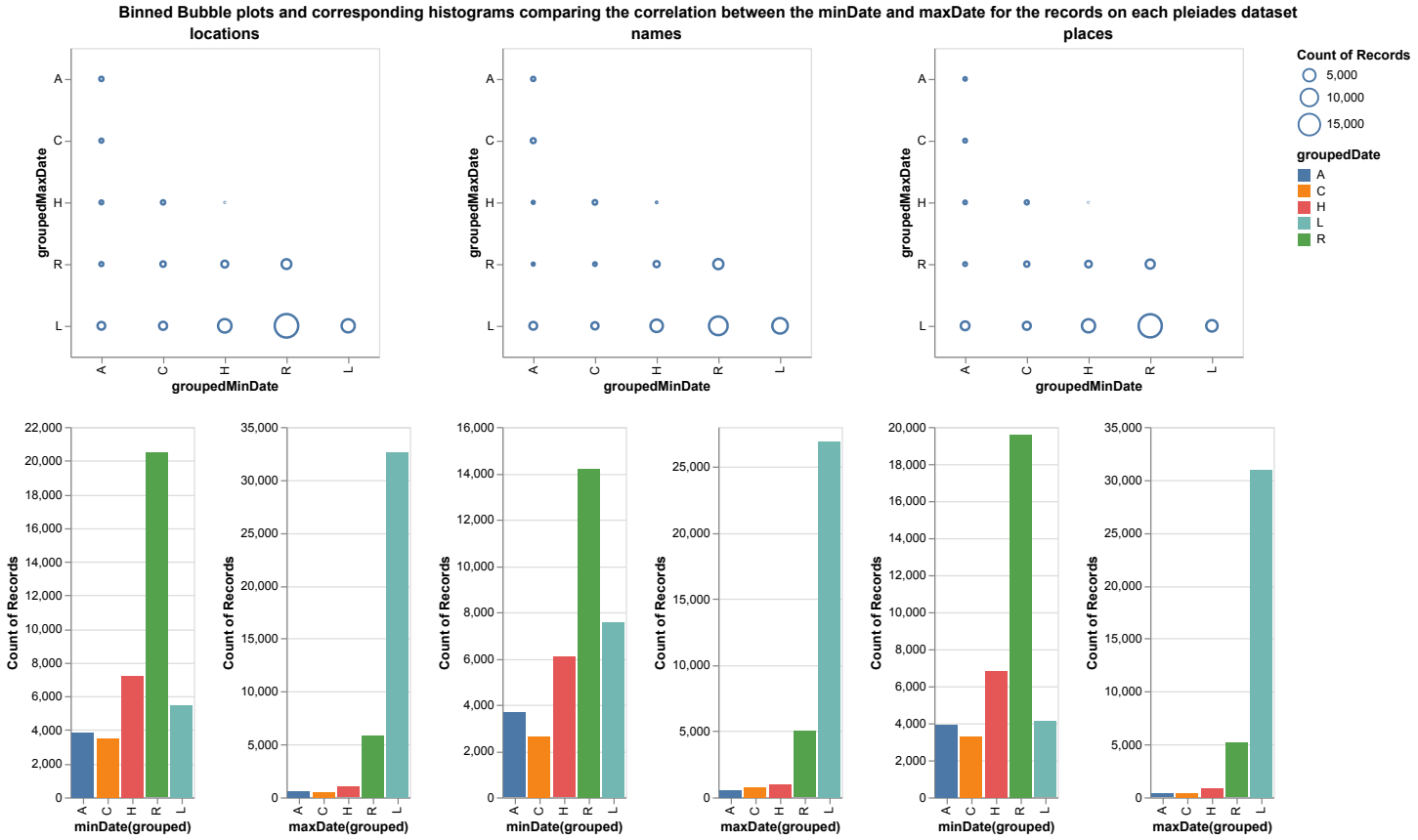
What can we learn from the visualization?

To understand the correlation between the minDate and maxDate for the records across each pleiades dataset

What is the name for the type of visualization(s) used?

Binned bubble plot and histogram

```
import altair as alt

locations = ('https://raw.githubusercontent.com/SwanseaU-TTW/csc337_coursework1/master/
names = ('https://raw.githubusercontent.com/SwanseaU-TTW/csc337_coursework1/master/plei
places = ('https://raw.githubusercontent.com/SwanseaU-TTW/csc337_coursework1/master/ple

def binnedBubblePlot(title, source):
    return (
        alt.Chart(source).transform_calculate(
            groupedMinDate='datum.minDate < -550 ? "A" : datum.minDate < -330 ? "C" : da
        ).transform_calculate(
            groupedMaxDate='datum.maxDate < -550 ? "A" : datum.maxDate < -330 ? "C" : da
        ).mark_point().encode(
            alt.X("groupedMinDate:O", sort=['A', 'C', 'H', 'R', 'L']),
            alt.Y("groupedMaxDate:O", sort=['A', 'C', 'H', 'R', 'L']),
            size='count():N',
            tooltip=['count()']
        ).properties(width=250, height=250, title=f'{title}')
    )

def histo(source, column):
    return (
        alt.Chart(source).transform_calculate(
            groupedDate=f'datum.{column} < -550 ? "A" : datum.{column} < -330 ? "C" : da
        ).mark_bar().encode(
            alt.X("groupedDate:O", sort=['A', 'C', 'H', 'R', 'L'], title=f'{column}(grou
            alt.Y('count()'),
            color='groupedDate:N',
            tooltip=['groupedDate:N', 'count()']
        ).interactive()
    )

def combi(title, source):
    return (
        binnedBubblePlot(title, source) & (histo(source, 'minDate') | histo(source, 'max
    )

alt.hconcat(
    combi('locations', locations), combi('names', names), combi('places', places)
).properties(
    title='Binned Bubble plots and corresponding histograms comparing the correlation be
).configure_title(orient='top', anchor='middle')
```



What are all visual mappings used?

For each pleiades dataset,

binned bubble plot

x position: minDate (grouped and sorted according to the group)

y position: maxDate (grouped and sorted according to the group)

tooltip: count of records

Given that date is minDate or maxDate

histogram

x position: date (grouped and sorted according to the group)

y position: count of records

color: date (grouped and sorted according to the group)

tooltip: count of records

Was there any special data preparation done?

For each dataset, the data has been grouped to match the pleaides README on `timePeriods` such that the records fall into the following bins "*...*'*A*' *(1000-550 BC),* '*C*' *(550-330 BC),* '*H*' *(330-30 BC),* '*R*' *(AD 30-300),* '*L*' *(AD 300-640)*".

What are the limitations of your design?

The size scale could be more carefully chosen for better visibility of all bubbles, especially those on the smaller end of the spectrum. Then, for easier, quicker comparisons between each chart, each bubble could be overlay with the count of records it represent as opposed to relying for tooltips. The charts could also be organised and sized such that for each bubble plot, the corresponding `minDate` histogram is below it and the `maxDate` histogram is 'turned' on its left side and places to the left of the bubble plot and the scales line up.