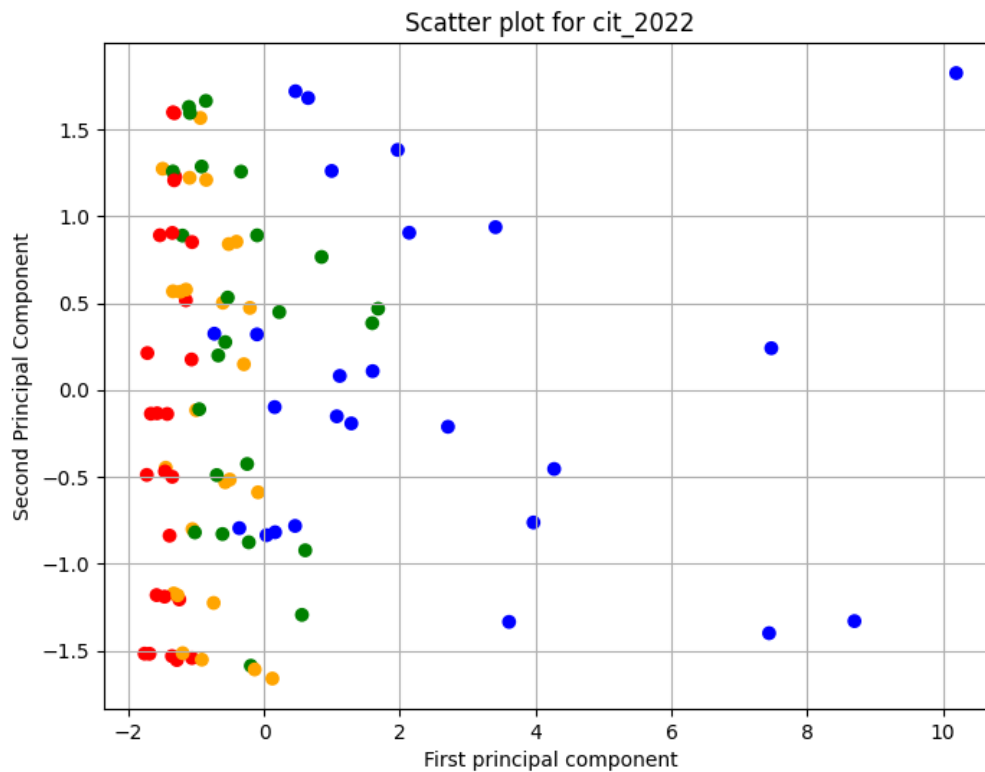


CIS662 HW2 Report

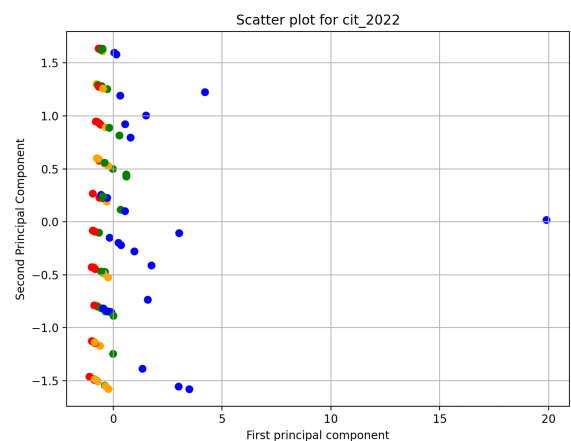
Gina Roh

* Note that I removed one data point (outlier) to enhance the clarity of the image representation. The original scatter plots, including the outlier, are also included in a smaller size in this report.

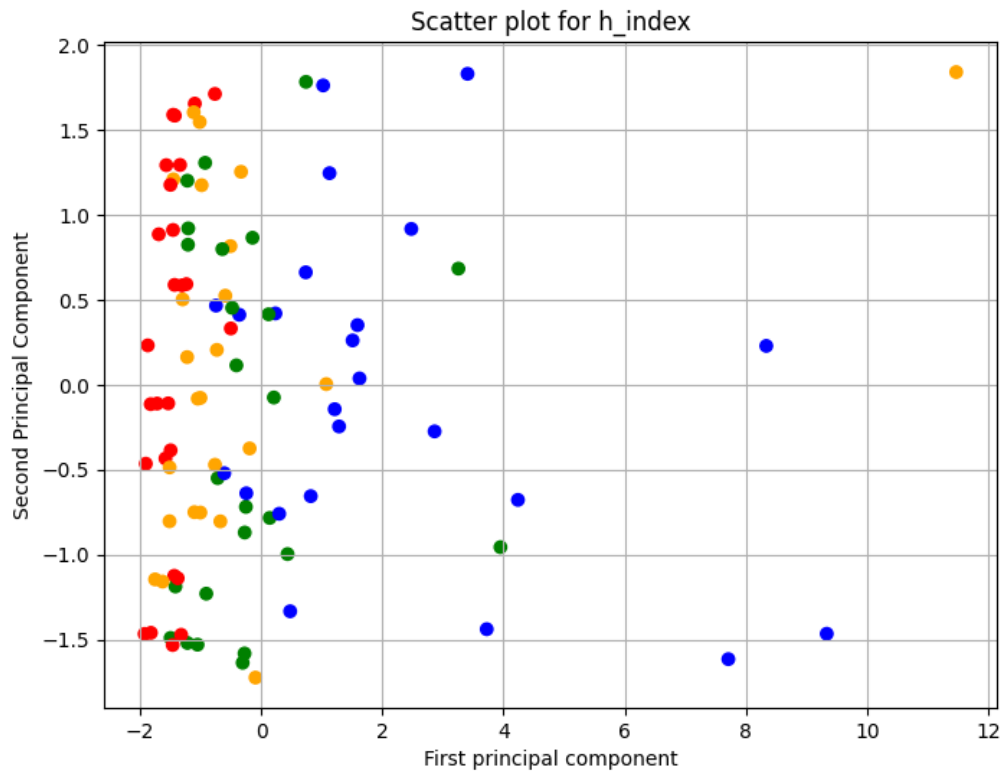
Problem 1: Plot quartile of 2022 citation numbers using the university rank and 2017-2021 citation numbers.



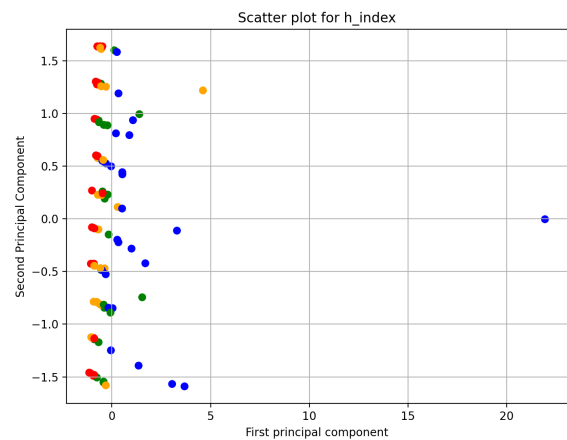
The first principal component is effective in predicting the 2022 citation numbers, whereas the second principal component doesn't appear to be useful for 2022 citation numbers prediction. The values are scattered and mixed when associated with the second principal component.



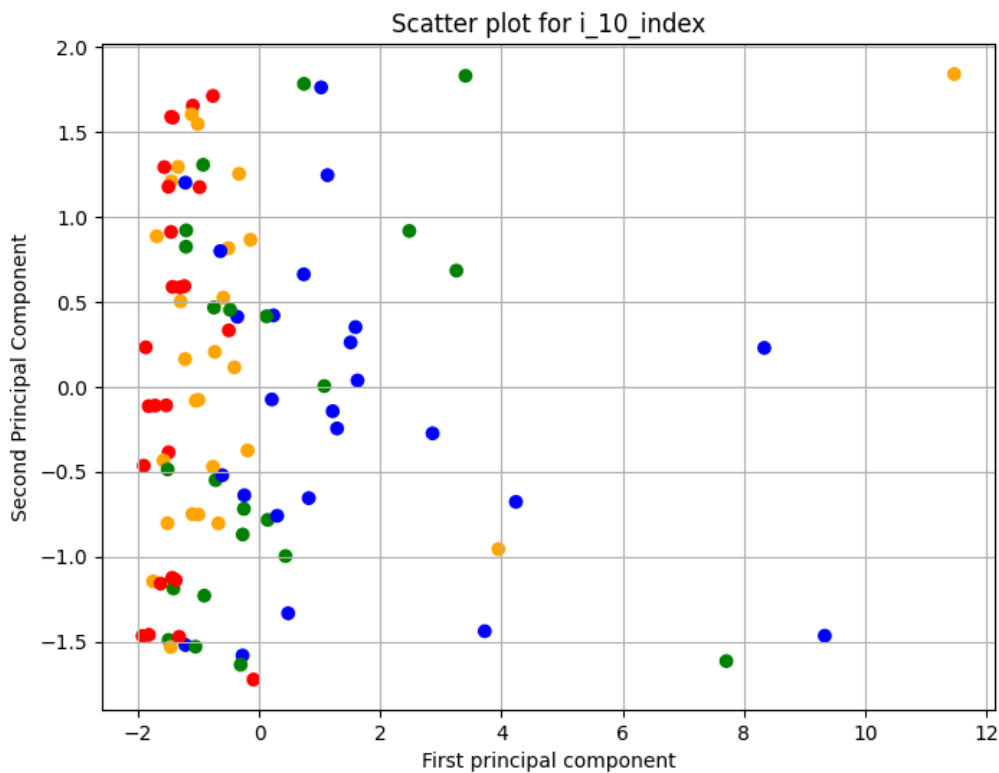
Problem2: Plot quartile of the h-index using the university rank and all (2017-2022) citation numbers.



The first principal component is effective in predicting the h-index, whereas the second principal component doesn't appear to be useful for h-index prediction. The values are scattered and mixed when associated with the second principal component.



Problem3: Plot quartile of the i10-index using the university rank and all (2017-2022) citation numbers.



The first principal component is quite effective in predicting the i10-index, whereas the second principal component doesn't appear to be useful for i10-index prediction. The values are scattered and mixed when associated with the second principal component.

