Initial EDA

Before running any unsupervised algorithm, we should look at basic descriptive stats that answer questions like:

* what are the most commonly taken elective courses across the entire population?

|  |  |
| --- | --- |
| Capital Markets & Investm | FINCB8306 |
| 6166 | 6062 |

* How does that change when aggregated by [insert any other field e.g. gender, race, future industry etc.].
* Who are the most popular instructors?
* Does this change over time? Etc.
* Maybe we want to do a heatmap with all elective courses to find which courses are most highly correlated initially?