Learnings from EDA:

* There are more females than males represented in the data
* University and Graduate level education make up a majority of those who have received loans
* We are seeing a 22% default rate next month, is this considered bad for the organization?
* Divorced people are heavily underrepresented in the data, do they ever receive loans? Are their conditions where they are granted loans?
* In the heat map, I am seeing very few (if any) strong correlations between fields
* Men tend to have a slightly more density of older ages than women I this data set

Main Lesson: The data is heavily limited and suffers from selection bias meaning we know who gets loans but we don’t know who is not getting loans.

Guido: We need to understand how it was decided who received loans and what a concerning default rate is.