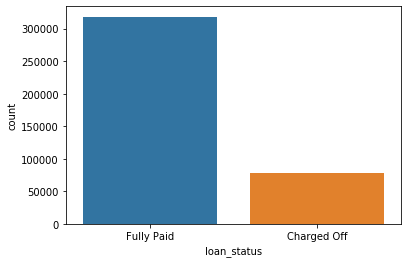
**AI Powered-Loan-Prediction**

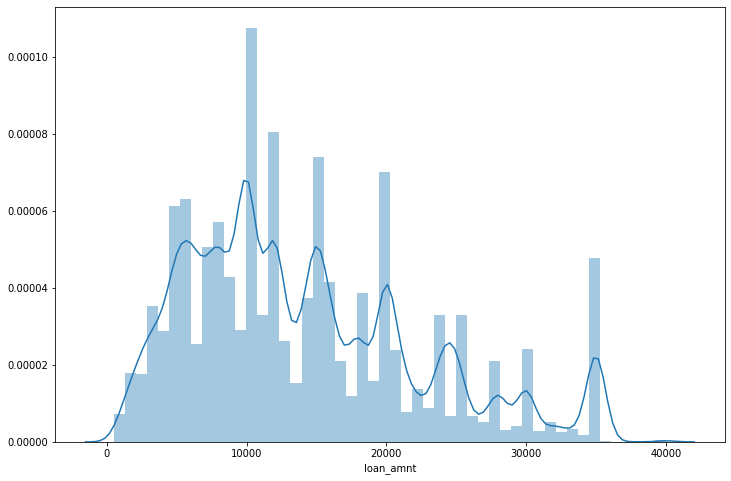
These files contain complete loan data for all loans issued through the 2007-2015, including the current loan status (Current, Late, Fully Paid, etc.) and latest payment information. The file containing loan data through the "present" contains complete loan data for all loans issued through the previous completed calendar quarter. Additional features include credit scores, number of finance inquiries, address including zip codes, and state, and collections among others. The file is a matrix of about 890 thousand observations and 75 variables. A data dictionary is provided in a separate file.

[](https://user-images.githubusercontent.com/57037068/83354391-69ca9880-a369-11ea-810f-51df058d30a2.png)

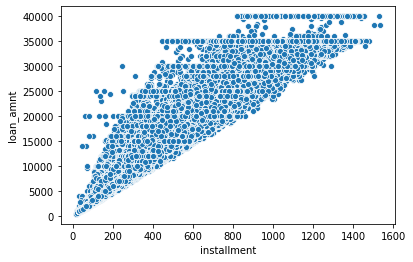
**# Imbalanced data - Classification problem data is not balanced**

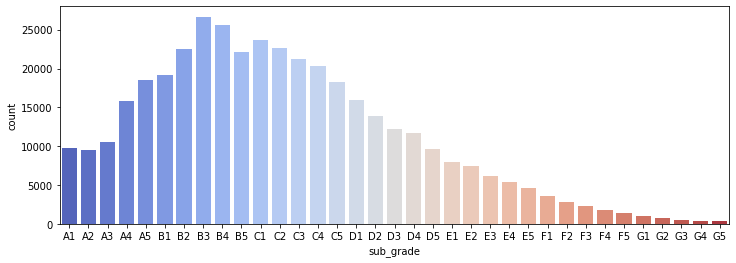
**Means that our accuracy will not be the main metrics that we will be looking for where while model will overfit to the Fully paid status, while underfit for charged off people**

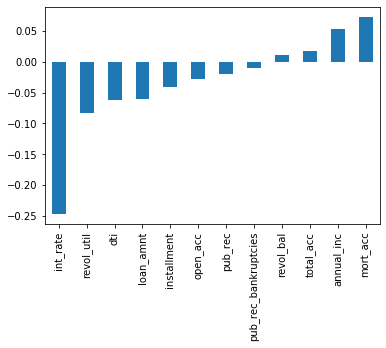
**Precision and recall (F1 score) will be the main representation for the "accuracy" of the model in the end**

[](https://user-images.githubusercontent.com/57037068/83354393-6b945c00-a369-11ea-9910-7ab2309c9008.png)

[](https://user-images.githubusercontent.com/57037068/83354394-6cc58900-a369-11ea-977b-007b86f4df84.png)

[](https://user-images.githubusercontent.com/57037068/83354398-6df6b600-a369-11ea-9831-236f8a326099.png)

[](https://user-images.githubusercontent.com/57037068/83354399-6fc07980-a369-11ea-9e3f-91a62552885e.png)

[](https://user-images.githubusercontent.com/57037068/83354400-70f1a680-a369-11ea-9c14-43a40049c865.png)