**Credit One - Exploratory Data Analysis**

Informal Report

(AKA Lessons Learned)

Samantha Goodman

**Potential Business Value found during EDA**

* Customers who were granted lower credit limits account for most of the defaults.

Men who were granted less than 80,000NT in credit defaulted at a rate of almost 34%.

For this reason, maybe raising the lowest credit offered to over 90,000NT to eliminate many of the problem customers.

* Because so many people with lower credit limits are defaulting, it seems like there is a problem with the decision making process whether or not to grant them a credit line.
* Once a customer is two months behind in bill payments, they are 50% likely to default. Every month they are behind contributes to the chance that they will default on the loan.

**Main Lessons learned from doing EDA**

* It’s easy to reuse code from previous datasets if comments are clear and variables are clearly labeled.
* Adding text to a plot can make additional details like a horizontal line representing the average default rate more clear.
* Sometimes when you isolate a feature, it doesn’t lead to any great insights. I found many times I would dig into the data, only to find that the smaller grouped data mimicked the larger pattern, and thus wasn’t a valuable segment to examine.

**Recommendations for Guido**

* At this point I haven’t built a model or even know what kind of model I will choose, but it is apparent that many people are granted credit lines who are later in default, particularly those granted lower value credit lines.
* It appears that customers who do not pay their bills are likely to default on the loan. This trend starts at 2 months behind and continues the longer they do not pay.