## 1. 課題

以下のような状況を仮定します;

　あなたはデータサイエンティストとして[Home Credit Group](http://www.homecredit.net/)から業務委託され，事業モデルを提案することになりました。

以下の要件を満たす事業モデルについての資料を提出してください。

### 要件

* 提出はスライド形式で、15スライド以内のPDFとしてください。
* 本講座で学んだ内容（データの可視化・機械学習技術など）を用いてください。
* 機械学習に関して事前知識をもたない人間に対する資料としてふさわしいものとしてください。
* ファイル名は "[GitHub ID].pdf"としてください。例) GitHubIDがHogeFugaならば "hogefuga.pdf"とする。

### データについて

* データは過去にKaggleで開催されたコンペティション「[Home Credit Default Risk](https://www.kaggle.com/c/home-credit-default-risk/rules)」で配布されたものを主として使用してください。（**第二回コンペで配布されたデータセットではなく、Kaggleで配布されているデータセットを使用することに注意してください。**ここで、データセットが複数に分かれていますが、その全てを使用する必要はありません。）
* 誰でも無償でアクセス可能なオープンデータは用いても構いませんが，その際には出典を明記してください。

## 2. ヒント

**- 目的変数について**

題材となっているコンペティションではクレジットリスクを目的変数としていますが，最終課題では自由に目的変数を設定して構いません。

**- 何を求められるのか？**

与えられたデータから収益を上げるための事業を考えます。その際に評価されるのは、事業提案の必然性の論理と事業規模の適切な推定です。

**- 事業を提案するまで**

以下に，事業を提案するまでの一例を挙げます;

1. データ探索 + 可視化
2. 事業のアイデアの模索
3. 考えた事業の必要性，規模等を概算

**- 資料作成の注意点**

要件にあるように，作成した資料は機械学習に関して事前知識をもたない人間に対するものです。そのため専門用語の濫用やソースコードのコピー&ペースト等は著しく評価が低くなる可能性があります。

**- データが多すぎて困った人へ**

application\_{train|test}.csvがメインのcsvです。まずはこれを探索すると良いでしょう。