

Lending Environment Simulation Model & Lender Evaluation Tool

Data Exploration

```
In [155]: 1 import glob
          2 import os
          3 import numpy as np
          4 import pandas as pd
          5
          6
          7 pd.set_option('display.max_rows', 50)
          8 pd.set_option('display.max_columns', 500)
          9 pd.set_option('display.width', 1000)
         10
         11 input_path = "../data/original/*.csv"
         12 paths = glob.glob(input_path)
         13 len(paths)
```

Out[155]: 2

```
In [156]: 1 # read all *.csv files from the input path
          2 # drop the headers and replace them with the sequential numbers
          3 df = pd.concat([pd.read_csv(f, names=range(0,38), header=0) for f in paths])
          4 df.shape, df.columns
          5
```

Out[156]: ((29383, 38),
Int64Index([0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37], dtype='int64'))

In [157]:

```
1 # this is a column dictionary
2 columns = {0: "Timestamp",
3            1: "Location ID",
4            2: "Has it become more difficult or easier to find a job in your city?",
5            3: "Is this a good time for people to make a large purchase such as furniture?",
6            4: "Compared to the last 6 months, are you able to spend (more, the same or less) on things?",
7            5: "Will you be able to meet your regular expenses over the next 6 months?",
8            6: "How do you expect your household's income to change over the next 6 months?",
9            7: "How do you expect general economic conditions in your city to change over the next 6 months?",
10           8: "How do you expect general economic conditions in your country to change over the next 6 months?",
11           9: "Gender",
12           10: "Marital status",
13           11: "Age",
14           12: "What's your highest level of education?",
15           13: "Occupation",
16           14: "If you are a student, what level are you currently studying?",
17           15: "Race/Ethnicity",
18           16: "Country",
19           17: "What is the name of the neighborhood where you live?",
20           18: "Over the past 3 months, how many times have you lent someone money?",
21           19: "On average how much do you lend in general?",
22           20: "Who did you lend money to in the past 3 months?",
23           21: "When you lend money, when do you usually expect to get it repaid?",
24           22: "Do you include either interest or a lending fee when you lend?",
25           23: "Do you request guarantees when you lend?",
26           24: "Do you receive your money back in time?",
27           25: "Assuming that you have lent money at least ten times, how often would you expect to be repaid?",
28           26: "What's the most common use of the money you lend?",
29           27: "Have you ever applied for a bank loan?",
30           28: "Are you a tontine / lending club member?",
31           29: "What is the most convenient way to get a loan?",
32           30: "To what extent do you agree with the following sentences [Access to credit is important for economic growth]",
33           31: "To what extent do you agree with the following sentences [Credit is beneficial for the economy]",
34           32: "To what extent do you agree with the following sentences [I would like to have a credit card]",
35           33: "What type of loans are you currently paying off?",
36           34: "Do you have a credit score?",
37           35: "Do you have a credit card?",
38           36: "On average, how much of your total household monthly income do you spend on credit?",
39           37: "If you wanted to take a loan to start a business, how much would you need?",
40 }
```

```
In [158]: 1 df.describe()
```

Out[158]:

	0	1	2	3	4	5	6	7	8	9	10	11	
count	29383	29383	27212	27212	27212	27212	27212	27212	27212	28391	28391	28391	28391
unique	26008	24800	3	3	3	3	3	3	3	4	10	16	16
top	6-24-2017 5:21:25	T19S38b3	Same / Pareil	Maybe / Peut-être	Same / Pareil	Maybe / Peut-être	Stay the same / La même	Stay the same / La même	Stay the same / La même	Male / Homme	Married / Marié	30-34	Bachelor degree / Licence
freq	4	12	13195	16001	14793	15745	14096	14603	14323	17551	12654	9308	7

Verify Quality of the Data that are Included into the Formular

Risk Columns

Recipients

Raw	Description	Categories	Weights
Family	Family	1	-0.05
Friends	Friends	2	0.1
Business colleagues	Business Colleagues	3	0.1
None of the above	Other	4	-0.05

```
In [165]: 1 u=df[20].unique() #20. Who did you lend money to in the past 3 months?
          2 len(u),u
```

```
Out[165]: (214,
array(['Business colleagues', 'Friends', 'Family', "I HAVEN'T LEND MONEY",
      'None', 'no one', 'none', 'no', 'all', 'NONE', 'mtn quick loan',
      'mobile loan', 'and other friends', 'a stranger', 'NO ONE',
      'friends and family', 'Friends and family', 'family and friends',
      'family and colleagues', 'none ', 'workmates', 'Fellow employee',
      'I give money without expecting to be returned.', 'Work Colleagues',
      'No one really.', 'All of the above', nan, 'Nobody', "I haven't ",
      'workers', 'Bank', 'NOBODY', 'None ', 'Cooperative', 'No', 'Nope',
      'No one.', 'Not available', 'non-disclosure.', 'nobody', 'nil',
      'i dont lend money', 'i did not lend any money',
      'Never lent Money', "i didn't lend money to someone", 'never',
      'Family and friends ', 'bothfamily and friends',
      'Friends, Family, and business colleagues', 'Family and friends',
      'FAMILY AND FRIENDS', 'friends, business colleagues and family',
      'both family and friends', 'Neighbors ', 'Brother', 'Sister',
      'society group', 'No One', 'my girl friend.', 'not available',
      'No one', 'work mates', 'Friends and Family', 'nobody',
      'Family and Friends', 'Family & Friends', 'Family And Friebsd',
      'All above', 'All', 'none of the ABOVE', 'did not lend anyone',
      'no body ', 'no lending', 'workmate', 'no body',
      'none of the above', 'nobodi', 'CO-WORKER', 'Did not lend',
      'Anyone who requires financial help', 'Not at all', 'Never',
      'Neighbour ', 'did not lend', 'noone', 'ALICE', 'work mate',
      'No body', 'No one ', 'no 1', 'I have not lend money',
      'I did not lend money', 'havent lend money', 'noi body',
      'I don't loan out money ', 'not', 'not yet', 'i dont',
      'i dont lend', 'non', 'NONE', 'friends and colleagues',
      'friends and colleagues', '0', 'and friends', 'church member',
      'private', 'confidential', 'Not to be disclosed',
      'Never lend to anyone', 'Neither', 'Frinds and Family',
      'Faamily & Friends', 'Fam & Frnds', 'Friends & Family',
      'Friends, Family', 'Nil', 'Micro finance', 'None of them ',
      'Lending Club', 'Loans are a bad idea ', 'E no need ',
      'i have not lent', 'church members', 'Spouse', 'na',
      'i dont lent money', 'Business Colleagues, Friends and Family',
      'Mostly family, friends', 'Campaigners & Business associates',
      "I don't lend money", 'I have never lent any money', 'Both',
      'neighbor', ' None', '-', 'Na', 'n/A', 'o', 'employees',
      'an employee', 'Nobody ', 'I did not lend money ', 'Nobodi ',
      'I lend money to noone ', 'family and friends', 'PARTNERS',
      'I have not lent money over the past 3 months', 'Non', 'student',
      'Husband', 'All concerned', 'Not available.', 'no one!', 'others',
      'ALL ABOVE', 'FRIENDS AND FAMILY', 'BANK', 'no-one', 'kijichi',
      'SACCOS', 'Family /\xa0La famille', 'Friends /\xa0Les amis',
      'Family /\xa0La famille, Friends /\xa0Les amis',
      'Business colleagues / Collegues de travail',
      'Family /\xa0La famille, Friends /\xa0Les amis, Business colleagues / Collegues d
e travail',
      'Friends /\xa0Les amis, Business colleagues / Collegues de travail',
      'Family /\xa0La famille, Business colleagues / Collegues de travail',
      'Family /\xa0La famille, Friends /\xa0Les amis, workmates',
      'school fees',
      'Family /\xa0La famille, Friends /\xa0Les amis, Business colleagues / Collegues d
e travail, MSHWARI',
      'no ', ' non', ' none', "didn't lend", 'no one actually',
      'no one ', 'Family /\xa0La famille, none', ' no one', ' no body',
```

```

'no ', 'nobody', '-----', 'never lend money', ' c', 'zero',
'school mates', 'dash', 'norbodi', 'no ne', '---', 'othee',
'anybogy', 'noo ne', 'no one borrowed ', 'i ddnt borrow money',
'no other', "i don't lend money", 'noboday', '?', 'Les amis',
'La famille', 'Collegues de travail', 'La famille, Les amis',
'Les amis, Collegues de travail',
'La famille, Collegues de travail',
'La famille, Les amis, Collegues de travail', 'personne', 'aucun',
'jamais', 'banque',
'Family /\xa0La famille, Friends /\xa0Les amis, Business colleagues / Collegues d
e travail, Etat',
'ma femme', "J'ai pas emprunté"], dtype=object))

```

Interest

	Raw	Description	Categories	Weights
	Always /Toujours	Always	1	-0.1
	Most of the time/ La moitié du temps	Most of the Time	2	-0.05
	About half the time/ La moitié du temps	About half the time	3	0.05
	Occasionally/ Souvent	Occasionally	4	0.08
	Never / Jamais	Never	5	0.12

```

In [164]: 1 u = df[22].unique() # 22:"Do you include either interest or a lending fee when you lei
          2 len(u), u

```

```

Out[164]: (15, array(['About half the time', 'Occasionally', 'Most of the time', 'Never',
                    'Always', 'Never / Jamais', 'Occasionally / Souvent',
                    'Always /\xa0Toujours',
                    'About half the time /\xa0La moitié du temps',
                    'Most of the time /\xa0La majorité du temps', 'Jamais',
                    'La moitié du temps', 'Souvent', 'La majorité du temps',
                    'Toujours'], dtype=object))

```

Collateral

	Raw	Description	Categories	Weights
	Always /Toujours	Always	1	-0.05
	Most of the time/ La moitié du temps	Most of the Time	2	-0.03
	About half the time/ La moitié du temps	About half the time	3	0.03
	Occasionally/ Souvent	Occasionally	4	0.04
	Never / Jamais	Never	5	0.06

```
In [163]: 1 u = df[23].unique() #23:"Do you request guarantees when you lend?"
          2 len(u),u
```

```
Out[163]: (18, array(['Most of the time', 'Never', 'About half the time', 'Occasionally',
                    'Always', 'Never / Jamais', 'Occasionally / Souvent',
                    'Always /\xa0Toujours',
                    'Most of the time /\xa0La majorité du temps',
                    'About half the time /\xa0La moitié du temps',
                    'About half the time / La moitié du temps', 'Always / Toujours',
                    'Most of the time / La majorité du temps', 'Jamais',
                    'La majorité du temps', 'Souvent', 'La moitié du temps',
                    'Toujours'], dtype=object))
```

Liquidity

Frequency

	Raw	Description	Categories	Weights
	More than 4 times / Plus de 4 fois	More than 4 times	1	0.1
	Between 2-3 times / Entre 2 et 3 fois	2-3 times	2	0.07
	Once / Une fois	Once	3	-0.02
	Never / Jamais	Never	4	-0.05

```
In [166]: 1 u=df[18].unique() #18:"Over the past 3 months, how many times have you lent someone m
          2 len(u),u
```

```
Out[166]: (22, array(['Between 2-3 times', 'Once', 'More than 4 times', 'Never',
                    'Once / Une fois', 'Between 2-3 times /\xa0Entre 2 et 3 fois',
                    'More than 4 times / Plus de 4 fois', 'Never / Jamais',
                    'Plus de 4 fois', 'Entre 2 et 3 fois', 'Aucune', 'Une fois',
                    'Jamais', 'Une fois, Entre 2 et 3 fois',
                    'Entre 2 et 3 fois, Plus de 4 fois', 'Une fois, Plus de 4 fois',
                    'Jamais, Entre 2 et 3 fois', 'Jamais, Une fois, Entre 2 et 3 fois',
                    'Une fois, Entre 2 et 3 fois, Plus de 4 fois',
                    'Jamais, Plus de 4 fois',
                    'Jamais, Une fois, Entre 2 et 3 fois, Plus de 4 fois',
                    'Jamais, Une fois'], dtype=object))
```

Duration

	Raw	Description	Categories	Weights
	Less than a month / Moins d'un mois	Less tan a month	1	-0.01
	At least 1 month but than 3 months / 1 - 2 mois	At least 1 month but less than 3 months	2	0.02
	At least 3 months but less than 6 months / 3 - 5 mois	At least 3 months but less than 6 months	3	0.03
	At least 6 months but less than 12 months / 6 - 11 mois	At least 6 months but less than 12 months	4	0.04
	One year or more / Plus d'un an	One year or more	5	0.05
	???	???	6	-0.02

```
In [167]: 1 u=df[21].unique() #21:"When you lend money, when do you usually expect to get it repa'
2 len(u),u
```

```
Out[167]: (18, array(['At least 3 months but less than 6 months',
'At least 1 month but less than 3 months', 'Less than a month',
'One year or more', 'At least 6 months but less than 12 months',
"Never - I don't expect to get repaid",
'At least 1 month but than 3 months',
'At least 3 months but less than 6 months / 3 - 5 mois',
"Less than a month / Moins d'un mois",
'At least 1 month but than 3 months /\xa01 - 2 mois',
'At least 6 months but less than 12 months / 6 - 11 mois',
"One year or more / Plus d'un an", "Moins d'un mois", '1 - 2 mois',
'3 - 5 mois', '6 - 11 mois', "Plus d'un an", nan], dtype=object))
```

Amount

Raw	Description	Categories	Weights
???	Micro	1	-0.05
???	Small	2	0.03
???	Medium	3	0.05
???	Large	4	0.07

```
In [168]: 1 u=df[19].unique() #19:"On average how much do you lend in general?"
2 len(u),u
```

```
Out[168]: (1056, array(['At least 91 but less than 444', 'Less than 90',
'At least 445 but less than 2200', ..., '150000frs', '0frs',
'5000frs'], dtype=object))
```

Usage

	Raw	Description	Categories	Weights
To pay for one-time or sudden expenses (Wedding, medical emergencies, etc.)		Urgent/Problem	1	-0.1
	???	???	2	0.05
Investment (commerce, merchandise, etc.)		Investment	3	0.15
	???	???	4	-0.05
	I don't know	Do not know	5	0.05

```
In [169]: 1 u=df[26].unique() #26:"What's the most common use of the money you lend?"
          2 len(u),u
```

```
Out[169]: (18,
array(['To invest or cover business expenses (Merchandise, salaries, etc.)',
      'To pay for one-time or sudden expenses (Wedding, medical emergencies, etc.)',
      'To cover regular expenses (Rent, clothing, home appliances, etc.)',
      'To pay off other debts', 'I don't know',
      'Investment (commerce, merchandise, etc.)',
      'Personal or family expenses (Health, education)', nan,
      'Pour payer des dépenses ponctuelles ou soudaines (mariage, urgences médicales, e
tc.)',
      'Pour couvrir les dépenses régulières (loyer, vêtements, appareils ménagers, et
c.)',
      'Payer d'autres dettes',
      'Pour investir ou couvrir les dépenses de votre entreprise (marchandises, salaire
s, etc.)',
      'Je ne sais pas', 'Régler les problèmes subits',
      'Investir dans un projet', 'I don't know / Je ne sais pas',
      'Personal or family expenses (Health, education) /\xa0Régler les problèmes subit
s',
      'Investment (commerce, merchandise, etc.) / Investir dans un projet'],
dtype=object))
```

????

Raw	Description	Categories	Weights
Always /Toujours	Always	1	0.35
Most of the time/ La moitié du temps	Most of the Time	2	0.25
About half the time/ La moitié du temps	About half the time	3	0.15
Occasionally/ Souvent	Occasionally	4	-0.1
Never / Jamais	Never	5	-0.3

```
In [174]: 1 u=df[24].unique() # 24:"Do you receive your money back in time?"
          2 len(u),u
```

```
Out[174]: (18, array(['About half the time', 'Most of the time', 'Occasionally', 'Never',
      'Always', 'Most of the time /\xa0La majorité du temps',
      'Always /\xa0Toujours',
      'About half the time /\xa0La moitié du temps',
      'Occasionally / Souvent', 'Never / Jamais',
      'Most of the time / La majorité du temps', 'Always / Toujours',
      'About half the time / La moitié du temps', 'La moitié du temps',
      'Toujours', 'La majorité du temps', 'Souvent', 'Jamais'],
dtype=object))
```