|  |  |  |
| --- | --- | --- |
| **Columns** | **Action On the column** | **Question to the Clients** |
| 1. 1:19 and 23:25 | Delete Them, Not necessary in our analysis |  |
| 1. A6 | Calculated the Age(using current year-year given) and thinking of Grouping the age for better analysis |  |
| 1. A9 | Joined it with A9\_2\_TEXT | Are supposed to only consider Cote d’Ivoire or the country doesn’t matter |
| 1. B1 | Drop it since it is empty | Was there a mistake maybe you forgot to send it but even so its not that much needed |
| 1. B3 & B4 | Fill in the missing values with zeros | Question about how to quantify ‘plusiers’ |
| 1. B6 | Added another category: ‘Between 1 and 3 years’ to cater for some values |  |
| 1. B\_6\_TEXT | Calculated well the exact time the biz started and also put some zeros where there is no values |  |
| 1. B7 | Created some categories like:  Production, Processing, Distribution, Aggregation, Transportation, Logistics, Agri-services and dropped the column after |  |
| 1. B8 | Dropped it, not relevant for now |  |
| 1. B9 | Keep it as it is for now but in futire Combined it with the B9\_F9\_TEXT and had more countries: Tonkpi, Tabou, Akoupe, San pedro, Yakasse, Adzope, Alépé, Agnibilekrou, Duékoué, Tiebissou, Abengourou, Guiglo, Zagne, Aboisso, Tchetaly, Bouake, N'Douci, Bouafle, La Mé, Daoukro, Agboville, Toumodi, Yamoussoukro, Grand-Lahou, Boundiali, Dimbokro, Kouibly, Agnébi Tiassa, Katiola, Lakota, Dabou, Didievi, Adieke, Aboisso, Ferke, ZUENOULA, Buyo, Soubre, Tiassale, Assouba, Oume, Adzope, Abidjan, Danane, Duekoue, Daloa, Banahouin | Do they want us to only consider the known cities |
| 1. B11\_1\_TEXT | Text based, taking all the certification companies |  |
| 1. B13\_1,2,3,4,5,6,7\_TEXT | Combine them and filled the missing with zeros |  |
| 1. B14, B14\_2\_TEXT | Combined them |  |
| 1. B15\_1\_TEXT | Text based, will deal with it in the future |  |
| 1. B16 | Text based, Extracted to role from text using nltk. Improve the model efficiency in future |  |
| 1. B17\_1\_TEXT | Extracted the documents use nltk |  |
| 1. C1, C2, C3, C5 | Text based to deal with in the future |  |
| 1. C6 | Crucial but no time to do the text based but in future to go through it |  |
| 1. D2 | Created new columns for Agri - Transport (financing for vehicle)  Insurance  Agri - Equipment (financing for equipment)  Product financing (fertilizers, seeds, pesticides)  No need for financial support  Leasing (equipment, transport/logistics)  Operational loan (working capital)  Agri Cash (liquidities)  Trade finance for export |  |
| 1. D3\_1,2,3,4,5,6,7\_TEXT | Combined them and filled missing values with zeros | So is this a good assumption that the missing values are zeros |
| 1. D4 | Created new columns for all months |  |
| 1. D9 | Filled with zeros where we have missing values | Is it a good assumption that people who didn’t fill have zero loans |
| 1. D11 |  | Many didn’t specify their loans but wrote interest rate. Is there no mistake |
| 1. D13 | Text based, we shall deal with it later on |  |
| 1. D14 | Very Crucial, gave them weight |  |
| 1. D15 | Text based, deal with it later on |  |
| 1. E3 | Created columns for new categories: Pay employees  Pay bills  Other :  Transfer money  Pay business suppliers  Pay producers/ farmers |  |
| 1. E3\_6\_TEXT | Text based, deal with it later |  |
| 1. E4 | Very crucial, Give them weights |  |
| 1. E5 | Text based, deal with it later |  |
| 1. E6\_1\_TEXT | Text based, deal with it later |  |
| 1. F6\_1\_TEXT | Deal with later, text based |  |
| 1. F7\_1\_TEXT | Deal with later |  |
| 1. F8\_1\_TEXT | Deal with later |  |