作業1: 資料進得來,分析出得去,信用金融大數據

1. 請依”clean required”一欄對CRX.csv進行資料之匯入與清理及轉換(transfer to)。(43%)

|  |  |  |  |
| --- | --- | --- | --- |
| Attribute | Defined | Transfer to | Clean required |
| Gender  a/b | Female/male | 0/1 | Judge yourself (5%) |
| Age | age |  |  |
| Debt | Outstanding debt (feature has been scaled) |  |  |
| Married | Married, u=Single/Divorced/etc, y=Married | 0/1 (5%) |  |
| BankCustomer  g/p | BankCustomer  Bank customer, 0=does not have a bank account, 1=has a bank account | 0/1 | Judge yourself (5%) |
| Industry  c, d, cc, i, j, k, m, r, q, w, x, e, aa, f |  | 1. CommunicationServices 2. ConsumerDiscretionary 3. ConsumerStaples 4. Education 5. Energy 6. Financials 7. Healthcare 8. Industrials 9. InformationTechnology 10. Materials 11. Real Estate 12. Research 13. Transport   Utilities | Judge yourself (10%) |
| Ethnicity  v, h, bb, j, n, z, dd, ff, o | Ethnicity | |  | | --- | | Asian | | Black | | Latino | | Other | | White | | 1. 請移除缺漏值. 2. 請移除焮數過少的值. 3. Judge yourself (10%) |
| YearsEmployed | Years employed |  |  |
| Prior default  t, f. | Prior default, 0=no prior defaults, 1=prior default | 0/1 | Judge yourself (2%) |
| Employed  t, f. | Employed, 0=not employed, 1=employed | 0/1 | Judge yourself (2%) |
| Credit score | Credit score (this feature has been scaled) |  |  |
| Drivers license,  t, f. | 0=no license, 1=has license | 0/1 | Judge yourself (2%) |
| Citizenship  g/p/s | Citizenship,  either ByBirth, ByOtherMeans or Temporary | ByBirth, ByOtherMeans or Temporary | Judge yourself (5%) |
| ZipCode |  |  | Remove |
| Income | Income (this feature has been scaled) |  |  |
| Approved, +/- | 0=not approved, 1=approved | 0/1 | Judge yourself (2%) |

請貼出您的process. (12%)

2. 太棒了,現在我們終於可以開始進行一些敍述統計了,

2-1 (10%)請問您最後留下幾筆資料? 還有缺漏值嗎? 那些欄位?

2-3 (10%) 您建議怎麼處理這些缺漏值,請說明並進行必要之處置

3. 有圖有真相,請畫出2張圖,並加以說明。(25%)



