1. Create a vintage curve of 30+ DPD (Days Past Due) from the following data. We have provided you with 2 files here.

Payment\_history: This file contains 4 fields

* Loan.ID - Unique identifier of loan contract
* DPD- The number of days past the EMI due date when the payment was done
* DPAmount– The amount of payment due
* Mth – The EMI payment month

contract\_file : This file contains

* CL.Contract..CL.Contract.ID – Unique identifier of loan contract , same as loan ID
* Status – The current Status of loan. The status as of 30th Aug 2018. Active – Good standing means the cmr is current and has paid all his dues as of 30th Aug. 2018. Active – Bad standing mean the cmr has not paid any/some/all of his dues as of 30th Aug. 2018. Closed -obligations met mean the cmr has prepaid his loan and his last.payment date will be the date of his loan closure
* Contract.Date – The date on which the loan was booked
* Oldest.Unpaid..Due.Date – In case of accounts with payment due, it reflects the oldest EMI due date for which the payment is to be received. In case of good accounts, the date reflects the coming due date. (Note: the data is pulled as of 30th Aug 2018). This field does not have significance in case of closed accounts
* Next.Due.Date – The next EMI due date. Since the data is pulled as of 30th Aug. 2018, the next due date for all the accounts is October 1st , except for closed accounts.
* Last.Payment.Date – The last payment date of emi from the customer

1. In the following data set, create a model to predict the active good standing customer. The target variable can be derived from the field “ pymt\_standing”. The description for other variables is given in sheet 1

