```
--Below code uses 3 input tables, which can be dowloaded here:
--https://hikognition.com/target-variable-risk-models/

--1. A Vintage table is created
--it combines all loans with all possible reporting dates and loan status on those dates

use a2
if exists (select * from INFORMATION_SCHEMA.TABLES s
where s.TABLE_CATALOG = 'X2' and s.TABLE_NAME = 'vintage')
drop table x2.dbo.vintage

select b.*
,DPD = case when b.ReportingDate < dr.MinDate then -1
                    when dr.MinDate is Null and w.LoanID is Null then -2
                    when dr.MinDate is Null and w.LoanID is Not Null then 8888
                    --when b.ReportingDate > dr.MaxDate then dr.DPD
                    when b.ReportingDate > dr.MaxDate and w.LoanID is Not Null then
dr.DPD
                    when b.ReportingDate > dr.MaxDate and w.LoanID is Null then -3
                    else dd.DPD end
,Balance = case when b.ReportingDate < dr.MinDate then Null
                           when b.ReportingDate > dr.MaxDate and w.LoanID is Not Null
then dr.Balance
                           when b.ReportingDate > dr.MaxDate and w.LoanID is Null then 0
                           else dd.Balance end
,WO = case when b.LoanID is Null then 0 else 1 end
into x2.dbo.vintage
       from (select c.LoanID, c.CustID
       ,CustVintage = year(c.CustStart) *100 + month(c.CustStart)
       ,LoanVintage = year(c.LoanDate)*100 + month(c.LoanDate)
       ,LoanType, c.LoanAmount as LoanAmount--, RodzajUmowy = case when c.RodzajUmowy =
'Nowa' then c.RodzajUmowy else 'Kolejna' end
       , MOB = left(dates.ReportingDate,4)*12 + right(dates.ReportingDate,2) -
year(c.LoanDate)*12 - month(c.LoanDate)
       ,CMOB = left(dates.ReportingDate,4)*12 + right(dates.ReportingDate,2) -
year(c.CustStart) *12 - month(c.CustStart)
       ,dates.ReportingDate
              from x2.dbo.Loans c
              ,(select dd.ReportingDate
                     from x2.dbo.LoansPastDue dd group by ReportingDate) dates
                     where 0 <= left(dates.ReportingDate,4)*12 +
right(dates.ReportingDate,2) -year(c.LoanDate)*12 - month(c.LoanDate)
       ) b
       left join x2.dbo.LoansPastDue dd on b.LoanID = dd.LoanID and b.ReportingDate =
dd.ReportingDate
       left join (select a.*, b.DPD, b.Balance
              from (select dd.LoanID, min(dd.ReportingDate) MinDate, max(dd.ReportingDate)
MaxDate
                     from x2.dbo.LoansPastDue dd group by dd.LoanID) a
              join (select * from x2.dbo.LoansPastDue dd) b on a.LoanID = b.LoanID and
a.MaxDate = b.ReportingDate) dr
              on b.LoanID = dr.LoanID
       left join (select distinct LoanID from x2.dbo.wo) w on b.LoanID = w.LoanID


--2. VintagePlus table is created
--It has additional important vatiables at customer level
```

```
if exists (select * from INFORMATION_SCHEMA.TABLES s
where s.TABLE_CATALOG = 'X2' and s.TABLE_NAME = 'VintagePlus')
drop table x2.dbo.VintagePlus

select v.*, va.DPDcust, va.BadExposureCust, va.TotalBorrowed, va.LoansSoFar
into x2.dbo.VintagePlus
from x2.dbo.vintage v
       left join
       (select v.CustID, v.ReportingDate
       ,max(v.DPD) as DPDcust, round(sum(v.Balance),2) as BadExposureCust
       ,round(sum(v.LoanAmount),2) as TotalBorrowed, count(*) LoansSoFar
       from x2.dbo.vintage v
       group by v.CustID, v.ReportingDate) va
       on v.CustID = va.CustID and v.ReportingDate = va.ReportingDate

--3 Beh vars table is created
--it contains behavioral variables

use x2
if exists (select * from INFORMATION_SCHEMA.TABLES s
where s.TABLE_CATALOG = 'X2' and s.TABLE_NAME = 'BehVars')
drop table x2.dbo.BehVars

SELECT
c.LoanID, c.CustID
,max(c_past.LoanDate) LastLoanDate
,DATEDIFF(day,max(c_past.LoanDate),c.LoanDate) DaysFromLastL
,DATEDIFF(day,min(c_past.LoanDate),c.LoanDate) DaysFromFirstL
,floor(DATEDIFF(day,min(c_past.LoanDate),c.LoanDate)/360) YearsAsCust
,max(c_past.LoanAmount) MaxLoanAmt, count(distinct c_past.LoanID) ClosedLoans
,sum(c_past.UmPrzed) ClosedLoansA
,max(d.DPD) MaxDPDever            ,max(d12.DPD) MaxDPD12m,          max(d6.DPD) MaxDPD6m
,max(d.Balance) MaxBalPD    ,max(d12.Balance) MaxBalPD12m     ,max(d6.Balance) MaxBalPD6m

,MonthsSincePD = DATEDIFF(month,max(cast(cast(d.ReportingDate as varchar)+'28' as
date)),c.LoanDate)
,MonthsSince30 = DATEDIFF(month,max(cast(cast(dd30.ReportingDate as varchar)+'28' as
date)),c.LoanDate)
,MonthsSince90 = DATEDIFF(month,max(cast(cast(dd90.ReportingDate as varchar)+'28' as
date)),c.LoanDate)
into x2.dbo.BehVars
FROM x2.dbo.Loans c

       left join (--All earlier Loans of this customer
       SELECT c.LoanID, c.LoanType, c.CustID, c.LoanDate
       ,c.LoanAmount
       ,1 as UmPrzed
       FROM x2.dbo.Loans c
       ) c_past
       on c.CustID = c_past.CustID and c.LoanDate > c_past.LoanDate

       left join x2.dbo.LoansPastDue d --All loans past due of this customer on earlier
reporting dates
       on c_past.LoanID = d.LoanID and c.LoanDate > cast(cast(d.ReportingDate as
varchar)+'28' as date)

       left join x2.dbo.LoansPastDue d6 --All loans past due of this customer last 6
months
```

```
	on c_past.LoanID = d6.LoanID and c.LoanDate > cast(cast(d6.ReportingDate as
varchar)+'28' as date)
	and c.LoanDate <  dateadd(month,6,cast(cast(d6.ReportingDate as varchar)+'28' as
date))

	left join x2.dbo.LoansPastDue d12 --All loans past due of this customer last 12
months
	on c_past.LoanID = d12.LoanID and c.LoanDate > cast(cast(d12.ReportingDate as
varchar)+'28' as date)
	and c.LoanDate <  dateadd(month,12,cast(cast(d12.ReportingDate as varchar)+'28' as
date))

	left join (--All 30+ loans of this customer on earlier reporting dates
	select * from x2.dbo.LoansPastDue dd where dd.DPD > 30) dd30
	on c_past.LoanID = dd30.LoanID and c.LoanDate > cast(cast(dd30.ReportingDate as
varchar)+'28' as date)

	left join (--All 90+ loans of this customer on earlier reporting dates
	select * from x2.dbo.LoansPastDue dd where dd.DPD > 90) dd90
	on c_past.LoanID = dd90.LoanID and c.LoanDate > cast(cast(dd90.ReportingDate as
varchar)+'28' as date)

group by c.LoanID, c.LoanID, c.CustID, c.LoanDate


--4 Finally, LoansWithPredictors table is created
--it is a dataset ready to use for credit scoring models

use x2
if exists (select * from INFORMATION_SCHEMA.TABLES s
where s.TABLE_CATALOG = 'X2' and s.TABLE_NAME = 'LoansWithPredictors')
drop table x2.dbo.LoansWithPredictors

select l.*
,IncPrec = case		when round(l.Income,0) != l.Income then 1
				when round(l.Income,-1) != l.Income then 2
				when round(l.Income,-2) != l.Income then 3
				when round(l.Income,-3) != l.Income then 4 else 5 end
,b.ClosedLoans, b.DaysFromLastL, DaysFromFirstL, YearsAsCust
,isnull(MaxLoanAmt,0) MaxLoanAmt, isnull(last_l.LoanAmount,0) as LastLoanAmt
,MaxDPDever, MaxDPD12m, MaxDPD6m
,MaxBalPD, MaxBalPD12m, MaxBalPD6m, MonthsSincePD, MonthsSince30,   MonthsSince90
,f.BadFlag
into x2.dbo.LoansWithPredictors
from x2.dbo.Loans l
left join x2.dbo.BehVars b on l.LoanID = b.LoanID
left join (select v.LoanID
	,BadFlag = max(case when v.DPDcust > 90 and v.BadExposureCust > 0.5* v.LoanAmount
	then 1 else 0 end)
	from x2.dbo.VintagePlus v
	where v.MOB = 12
	group by v.LoanID) f
	on l.LoanID = f.LoanID
left join x2.dbo.Loans last_l on b.LastLoanDate = last_l.LoanDate and last_l.CustID =
b.CustID
where f.BadFlag is not Null
```