Dr. Karol Przanowski

kprzan@sgh.waw.pl

**Project description**

**Course code: 223110-1234**

**Basic and Advanced Programming in SAS with Statistics**

## Project description

Aim of the project is to get familiar with fully automated reporting processes with a use of both methods of data manipulation and statistical analyses. The project is based on basic reporting methods that are useful in credit risk management.

## Theoretical background

Vintage analysis is being used in stability analysis of a credit risk portfolio in time, also in order to find its maximal asymptotical value. Vintage statistics (both quantitative and quota) can be calculated using following parameters:

m – number of months after credit launch

due – minimal number of due instalments

m_prod – month in which all credits were launched

Quantitative vintage (m_prod, m, due)=(the number of all credits launched in m_prod month for which m months after the credit launch a number of instalments due >= due)/ (the number of all credits launched in m_prod month).

Quota vintage is calculated by weighting numerator and denominator with credit amount.

# Vintage3 calculation in SAS

```sas
libname inlib ..\pd2_en\project\data\';
data vin;
set inlib.Transactions;
seniority=intck('month',input(fin_period,yymmn6.),input(period,yymmn6.));
vin3=(due_installments>=3);
output;
if status in ('B','C') and period<='200812' then do;
        n_steps=intck('month',input(period,yymmn6.),input('200812',yymmn6.));
        do i=1 to n_steps;
                period=put(intnx('month',input(period,yymmn6.),1,'end'),yymmn6.);
                seniority=intck('month',input(fin_period,yymmn6.),input(period,yymmn6.));
                output;
        end;
end;
where product='ins';
keep fin_period vin3 seniority;
run;
proc means data=vin noprint nway;
class fin_period seniority;
var vin3;
output out=vintagr(drop=_freq_ _type_) n()=production mean()=vintage3;
format vintage3 nlpct12.2;
run;
proc means data=vin noprint nway;
class fin_period;
var vin3;
output out=production(drop=_freq_ _type_) n()=production;
where seniority=0;
run;
proc transpose data=vintagr out=vintage prefix=months_after_;
by fin_period;
var vintage3;
id seniority;
run;
```

# The list of tasks

Analyse all of the variables out of the *production* set that describe an account (aid) and a client (cid).

Categorize each of the variables into no more than 3 categories. For example, act_age variable could be categorized into such categories: young people (<18), adults (<35) and olds (>=35). For each of the categorized variables calculate the strength of correlation or influence with binary variable Vintage3 after 12 months (in the SAS code above vintage3 for seniority=12).

Order the categorized variables from those with the strongest correlation to whose with the weakest correlation and for 5 most correlated variables out of 5 groups of variables (app_: , act_:, ags:, agr: ) prepare reports in various formats: HTML – interactive reports with convenient link and png graphics (ods graphics), PDF – any technique that allows to produce high resolution graphics (could be LaTeX), Excel – reports with graphical elements.

Calculate vintages for due = 1,2 and 3.

Links to HTML reports should also indicate Excels and PDF files.

The final sets of reports should consist of a few levels of detail:

1. The reports of a whole credit portfolio with distinction of two products: css and ins.
2. Level that identifies variables and its order that depends on the strength of a correlation with vintage3 variable.
3. Each variable level, where vintage reports are prezented.

## Creativity

The aim of the project is not only to write SAS codes but also to solve the problem of a great number of reports and calculations. What is more the project requires to find an adequate method to calculate forecast of both quantitative and quota vintage3 statistics for seniority = 12. Student should stand in credit risk manager's shoes and with a use of the produced reports be able to assess if risk of the portfolio as well as client profile is stable over the time and if not what changes. It is possible to change the aggregation (e.g. prepare the quarterly or semiannually reports). The final sets of reports should be clear and should show key changes in credit risk portfolio without additional analyses.

Project can be prepared by more than one student, but the tasks distribution should be even.

Exemplary vintage reports are shown below (see also excel vintage3_example.xlsx that is included in a project);
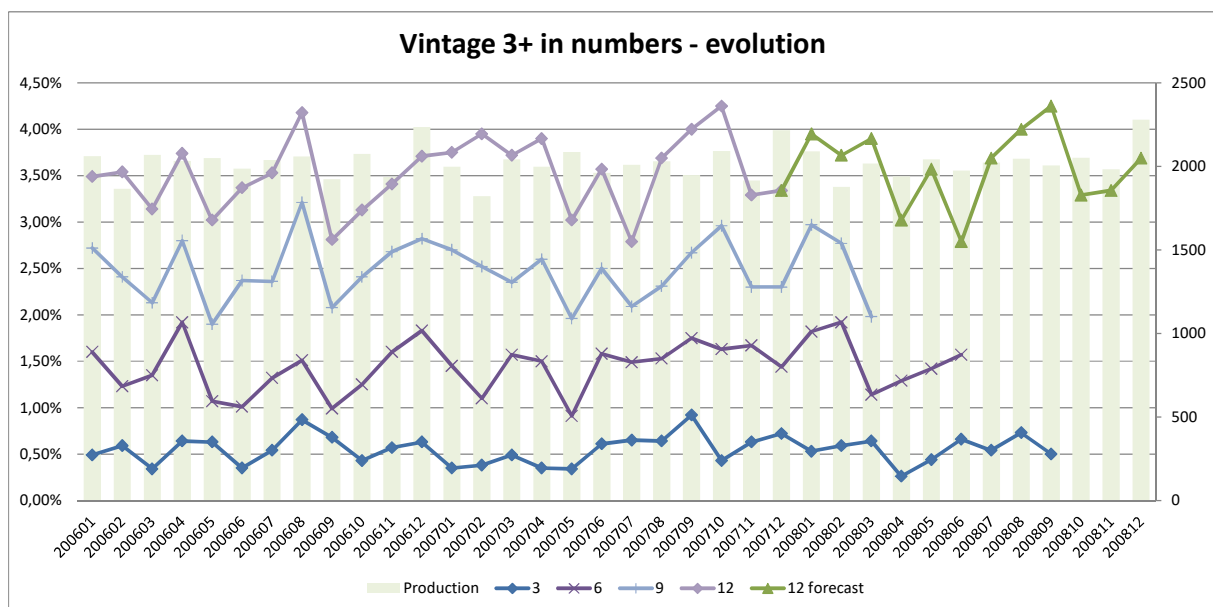


The forecast is showed only as an example and is not correctly calculated.

Vintage 3+ in numbers - asymptotic