

dr Karol Przanowski

kprzan@sgh.waw.pl

Opis projektu zaliczeniowego do zajęć

nr przedmiotu: 223110-1234

„Podstawowe i zaawansowane programowanie oraz statystyka w SAS”

Ogólny opis zadań

Celem projektu jest zapoznanie studenta z budową w pełni automatycznych procesów raportowych wykorzystując zarówno metody przetwarzania danych jak i analizy statystyczne. Projekt oparty jest na podstawowych sposobach raportowania portfela kredytów detalicznych przydatnych w zarządzaniu ryzykiem kredytowym.

Teoretyczne podstawy

Analizę vintage wykonuje się by badać stabilność ryzyka kredytowego w czasie i by znaleźć asymptotyczną jego maksymalną wartość. Statystykę vintage liczy się zarówno ilościową jak i kwotową. Jest ona zależna od następujących parametrów:

m – liczba miesięcy po uruchomieniu kredytu

due – minimalna liczba opóźnionych rat

m_prod – miesiąc w którym wszystkie kredyty uruchomiono.

Vintage ilościowy (m_prod, m, due) = (liczba wszystkich kredytów uruchomionych w miesiącu m_prod dla których m -miesięcy po uruchomieniu liczba opóźnionych rat $\geq due$) / (liczba wszystkich kredytów uruchomionych w miesiącu m_prod).

Statystykę kwotową uzyskujemy wążąc licznik i mianownik kwotami branych kredytów.

Wzór liczenia vintage3 w kodzie SAS

```
libname wyj '..\pd2\projekt\dane\';
data vin;
set wyj.Transactions;
seniority=intck('month',input(fin_period,yyymm6.),input(period,yyymm6.));
vin3=(due_installments>=3);
output;
if status in ('B','C') and period<='200812' then do;
    n_steps=intck('month',input(period,yyymm6.),input('200812',yyymm6.));
    do i=1 to n_steps;
        period=put(intnx('month',input(period,yyymm6.),1,'end'),yyymm6.);
```

```

seniority=intck('month',input(fin_period,yymmnn6.),input(period,yymmnn6.));
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;

```

Lista zadań

Ze zbioru production należy przeanalizować wszystkie zmienne opisujące rachunek (aid) i klienta (cid).

Każdą ze zmiennych należy podzielić na kategorie, maksymalnie 3. Dla przykładu zmienną act_age na kategorie: młodzi (<18), starsi (<35) i najstarsi (>=35). Dla każdej skategoryzowanej zmiennej należy wyznaczyć miernik współzależności ze zmienną zerojedynkową vintage3 po 12 miesiącach (w kodzie powyżej zmienna vin3 dla seniority=12).

Należy posortować wszystkie zmienne według tej statystyki i dla pierwszych 5 najlepszych zmiennych z grup app_ , act_ , ags_ , agr_ , wykonać wykresy w różnych formatach: HTML – interaktywne raporty z wygodnymi linkami i obrazkami png (ods graphics) , PDF – raport wykonany przez jakąś technikę (być może wykorzystując LaTeX) dającą ładne i wysokiej rozdzielczości raporty, Excel – raporty z gotowymi ładnymi elementami graficznymi.

Należy policzyć vintage dla due= 1, 2 i 3.

Linki reportów HTML powinny wskazywać także Excele i PDFy.

Finalny zestaw raportów powinien składać się z kilku warstw:

- warstwa zbiorcza, gdzie przedstawione są raporty całego portfela i dwóch produktów: css i ins.
- warstwa identyfikacji zmiennych i ich porządku
- warstwa każdej zmiennej.

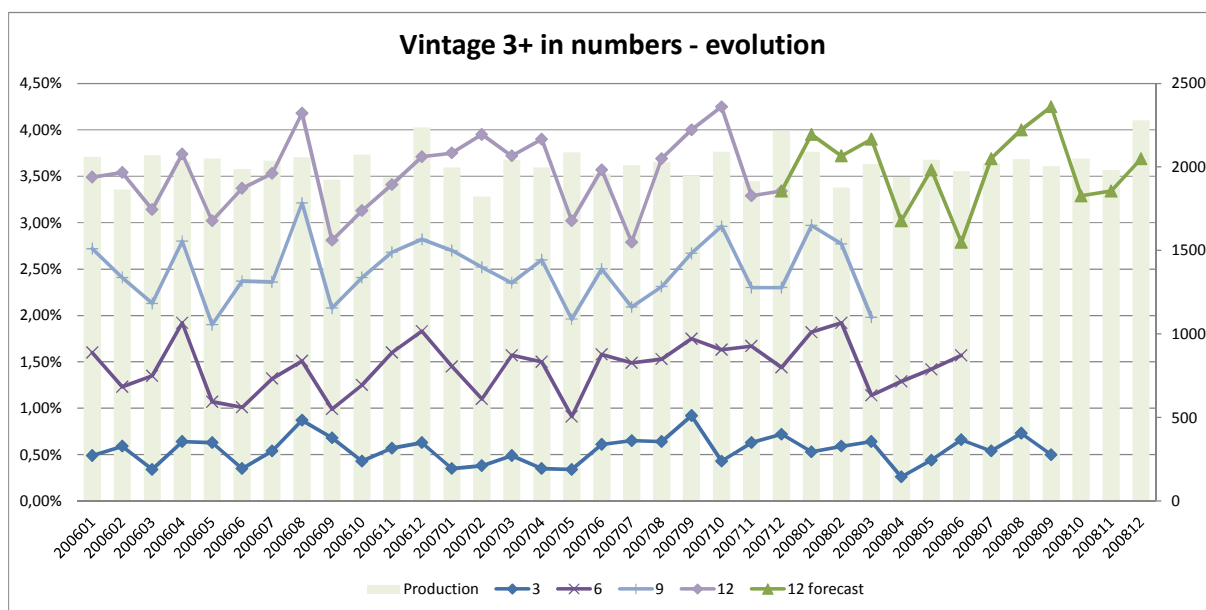
Kreatywność

Istotą projektu jest nie tylko napisanie właściwych kodów, ale także umiejętne rozwiązanie problemu dużej liczby raportów i wielu obliczeń. Dodatkowo należy wpasć na pomysł jak policzyć prognozę vintage3 po 12 miesiącach zarówno ilościową jak i kwotową. Wykonawcy projektu powinni postawić

się w roli zarządzającego ryzykiem kredytowym i umieć na podstawie wykonanych raportów jednoznacznie stwierdzić, czy ryzyko portfela jest stabilne, czy profil klientów jest stabilny, a jeśli nie to co się zmienia. Można też zdecydować się na inną granulację, zamiast po miesiącach to po kwartałach lub półroczach. Finalne zestawy raportów muszą być czytelne i w powinny być tak dobrane do danych, by można było w szybki sposób zauważyć najważniejsze zmiany.

Projekt może wykonywać kilka osób odpowiednio dzieląc się zadaniami, byleby każdy student miał mniej więcej tyle samo pracy do wykonania.

Przykładowe raporty vintage mogą wyglądać następująco (patrz excel vintage3_example.xlsx załączony w projekcie):



Prognoza jest tu wstawiona dla przykładu, bynajmniej nie jest poprawnie policzona.

