# CANADA DEBTOR IQ – INSTRUCTIONS FOR DATA EXTRACTION IN CAMBRIAN

**Using ACTUAL as 201711 - Previous is 201710**

**Changes SIC CODE format – AVG BFRS score rolling average inclusion – Risk Class definition of score cds 0 and bfrs**

# TABLES AT TRADE LEVEL



# CONEPT OF THE DATA SCHEMA FOR CURRENT AND PREVIOUS MONTH

DATA SCHEMA for **ACTUAL TABLE** (current period data) and **PREVIOUS TABLE** (previous period data) have the same structure.

The idea is that on an ongoing basis the **ACTUAL TABLE** will run and the previous ran of that table will become the “PREVIOUS TABLE”. ACTUAL and PREVIOUS could be identified by the ARCHIVE that refers to those periods.

For the first run intermediate tables for previous archive are named dbp\_xxx

# MACRO VARIABLES

*The following values need to be modified according to each Debtor IQ customer and time.*

set hivevar:ARCHIVES\_6='201706','201707','201708','201709','201710', '201711';

set hivevar:ARCHIVES\_3L='201709','201710', '201711';

set hivevar:ARCHIVES\_3H='201706','201707', '201708';

set hivevar:ARCHIVES\_0='201711';

set hivevar:ARCHIVES\_24=201511;

### Step1. Trades selection from commercial file (DB\_1)

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_6 |
| Final Table | No |

create table debtor\_iq\_db.DB\_1

stored as avro as

select archive,

alp.suppliersubjectnumber,

alp.ReferenceNumber ,

alp.subjectnumber,

alp.totalbalanceamount,

alp.currentbalanceamount,

alp.period1balanceamount,

alp.period2balanceamount,

alp.period3balanceamount,

alp.accountsecuredindicator

from **efx\_ca\_prod.comm\_file**

lateral view explode(alp000s) exploded\_table as alp

where archive in ( ${hivevar:ARCHIVES\_6})

and alp.suppressionflag = 'N';

### Step 2. Attribute Calculation for Reporting (DB\_2)

|  |  |
| --- | --- |
| Macro Used | None |
| Final Table | No |

create table debtor\_iq\_db.DB\_2

stored as avro as

select archive ,

suppliersubjectnumber,

subjectnumber,

ReferenceNumber ,

totalbalanceamount,

currentbalanceamount,

period1balanceamount,

period2balanceamount ,

period3balanceamount,

accountsecuredindicator,

cast(totalbalanceamount as int) as TOTAL\_BAL,

cast(currentbalanceamount as int) as CURR\_BAL,

cast(period1balanceamount as int) as BAL\_P1,

cast(period2balanceamount as int) as BAL\_P2,

cast(period3balanceamount as int) as BAL\_P3,

(cast(period1balanceamount as int) + cast(period2balanceamount as int)+

cast(period3balanceamount as int)) as TOTALP1P2P3,

case when cast(period3balanceamount as int)>0 then "90+ Days Past Due"

when cast(period2balanceamount as int)>0 then "60-89 Days Past Due"

when cast(period1balanceamount as int)>0 then "30-59 Days Past Due"

when cast(currentbalanceamount as int)>0 then "Current"

else "Not Applicable" end as DPD,

case when cast(totalbalanceamount as int)>0 then (cast(currentbalanceamount as int)/cast(totalbalanceamount as int))\*100 end as pct\_curr,

case when cast(totalbalanceamount as int)>0 then (cast(period1balanceamount as int)/cast(totalbalanceamount as int))\*100 end as pct\_pd1,

case when cast(totalbalanceamount as int)>0 then (cast(period2balanceamount as int)/cast(totalbalanceamount as int))\*100 end as pct\_pd2,

case when cast(totalbalanceamount as int)>0 then (cast(period3balanceamount as int)/cast(totalbalanceamount as int))\*100 end as pct\_pd3

from debtor\_iq\_db.DB\_1 ;

### Step 3. PI Calculation (DB \_3)

|  |  |
| --- | --- |
| Macro Used | None |
| Final Table | No |

This query creates PI and cap to value 99

create table debtor\_iq\_db.DB\_3

stored as avro as

select \*,

case when (accountsecuredindicator = "N" and TOTAL\_BAL=0) then 0

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)>=99 then 99

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)<=0 then 0

when (accountsecuredindicator = "N") and 0<cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)<99 then

cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int) end as PI

from debtor\_iq\_db.DB\_2 ;

drop table debtor\_iq\_db.DB\_2;

### Step 4. PI aggregation (DB\_4 and DB\_4H)

This query creates the average PI for the most recent 3 months

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_3L |
| Final Table | No |

create table debtor\_iq\_db.DB\_4

stored as avro as

select suppliersubjectnumber,

subjectnumber,

ReferenceNumber ,

avg(pi) as pi\_avg3

from debtor\_iq\_db.DB\_3

where archive in ( ${hivevar:ARCHIVES\_3L})

group by suppliersubjectnumber, subjectnumber,ReferenceNumber;

This query creates the average PI for oldest 3 months

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_3H |
| Final Table | No |

create table debtor\_iq\_db.DB\_4H

stored as avro as

select suppliersubjectnumber,

subjectnumber,

ReferenceNumber ,

avg(pi) as pi\_avg6

from debtor\_iq\_db.DB\_3

where archive in ( ${hivevar:ARCHIVES\_3H})

group by suppliersubjectnumber, subjectnumber,ReferenceNumber;

### Step 5. Province and SIC code Extraction (DB\_SIC) - Changes

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_0 |
| Final Table | No |

Change to adjust to SIC code issue

create table debtor\_iq\_db.DB\_SIC

stored as avro as

select

smp.subjectnumber,

smp.subjectname,

smp.siccode1forthiscompany,

smp.naicscodeforthiscompany,

smp.provinceorstatecode,

smp.subjectaddress,

case when cast(length(smp.siccode1forthiscompany) as int) = 1 then concat ('000', substr(smp.siccode1forthiscompany,1,1))

when cast(length(smp.siccode1forthiscompany) as int) = 6 then concat ('00', substr(smp.siccode1forthiscompany,1,2))

when cast(length(smp.siccode1forthiscompany) as int) = 7 then concat ('0', substr(smp.siccode1forthiscompany,1,3))

else substr(smp.siccode1forthiscompany,1,4) end as substr\_sic4

from efx\_ca\_prod.comm\_file

lateral view explode(smp100s) exploded\_table as smp where comm\_file.archive = ${hivevar:ARCHIVES\_0};

### Step 6. Get Description of SICCODE (DB\_SIC1)

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_0 |
| Final Table | No |

French description and coalesce function included for both is new since there were missing values in the 60 digits description vs 30. ***If they don’t want to change structure then eliminate French.***

create table debtor\_iq\_db.DB\_SIC1

stored as avro

as

select smp.\*,

coalesce(vtp.englishdescription60,vtp.englishdescription30," ") as englishdescription60,

coalesce(vtp.frenchdescription60,vtp.frenchdescription30," ") as frenchdescription60

from debtor\_iq\_db.DB\_SIC as smp

left outer JOIN (select englishdescription30,englishdescription60, siccode4digits,frenchdescription30,frenchdescription60 from efx\_ca\_prod.comm\_vtp318

where archive = ${hivevar:ARCHIVES\_0} ) as vtp

on smp.substr\_sic4 = vtp.siccode4digits;

Drop table debtor\_iq\_db.DB\_SIC;

### Step 7. Merge trade file with SICCODE description (DB\_5)

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_0 |
| Final Table | No |

create table debtor\_iq\_db.DB\_5

stored as avro as

select cfile.\*,

smp.subjectname,

smp.siccode1forthiscompany,

smp.substr\_sic4,

smp.naicscodeforthiscompany,

smp.provinceorstatecode,

smp.subjectaddress,

smp.englishdescription60,

smp.frenchdescription60

from debtor\_iq\_db.DB\_3 as cfile

left outer JOIN (select subjectnumber, subjectname,siccode1forthiscompany,naicscodeforthiscompany,

provinceorstatecode, subjectaddress,substr\_sic4,englishdescription60, frenchdescription60

from debtor\_iq\_db.DB\_SIC1 ) as smp

on cfile.SubjectNumber = smp.subjectnumber

where archive= ${hivevar:ARCHIVES\_0} ;

### Step 8. Extract scores and attributes for current period (DB\_scores1)

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_0 |
| Final Table | No |

create table debtor\_iq\_db.DB\_scores1

stored as avro as

select cfile.suppliersubjectnumber,

cfile.SubjectNumber ,

fullcds\_ot.equifaxsubjectnumber,

fullcds\_ot.ninetydaypi,

fullcds\_ot.ninetydayci,

fullcds\_ot.cdsscorevalue,

fullcds\_ot.marketingbfrsscorevalue as bfrs,

fullcds\_ot.ninetydaypastdueamountperiod1,

fullcds\_ot.ninetydaypastdueamountperiod2,

fullcds\_ot.ninetydaypastdueamountperiod3

from efx\_ca\_prod.comm\_scores

JOIN (select suppliersubjectnumber, SubjectNumber, archive from debtor\_iq\_db.DB\_1

where archive = ${hivevar:ARCHIVES\_0} ) as cfile

on cfile.SubjectNumber = fullcds\_ot.equifaxsubjectnumber

where comm\_scores.archive = ${hivevar:ARCHIVES\_0} ;

### Step 9. CDS aggregation (DB\_scores\_av1 and DB\_scores\_av2)

This query creates the average CDS for the most recent 3 months

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_3L and ARCHIVES\_3H |
| Final Table | No |

Change to adjust rolling average BFRS score

|  |
| --- |
| create table debtor\_iq\_db.DB\_scores\_av1  stored as avro as  select cfile.SubjectNumber,  avg(cast(fullcds\_ot.cdsscorevalue as int)) as scorecds\_avg1,  avg(cast(fullcds\_ot.marketingbfrsscorevalue as int)) as bfrs\_avg1  from efx\_ca\_prod.comm\_scores  JOIN (select suppliersubjectnumber, SubjectNumber, archive  from debtor\_iq\_db.DB\_1 where archive in (${hivevar:ARCHIVES\_3L})) as cfile    on cfile.SubjectNumber = fullcds\_ot.equifaxsubjectnumber  where comm\_scores.archive in (${hivevar:ARCHIVES\_3L})  group by cfile.SubjectNumber ;  create table debtor\_iq\_db.DB\_scores\_av2  stored as avro as  select cfile.SubjectNumber ,  avg(cast(fullcds\_ot.cdsscorevalue as int)) as scorecds\_avg2,  avg(cast(fullcds\_ot.marketingbfrsscorevalue as int)) as bfrs\_avg2  from efx\_ca\_prod.comm\_scores  JOIN (select suppliersubjectnumber, SubjectNumber, archive from debtor\_iq\_db.DB\_1 where archive in (${hivevar:ARCHIVES\_3H})) as cfile  on cfile.SubjectNumber = fullcds\_ot.equifaxsubjectnumber  where comm\_scores.archive in (${hivevar:ARCHIVES\_3H})  group by cfile.SubjectNumber ; |

### Step 10.Create Score Bands (current) merge with average CDS (db\_scores\_3)

|  |  |
| --- | --- |
| Macro Used | No |
| Final Table | No |
|  |  |

Change to adjust rolling average BFRS score

create table debtor\_iq\_db.DB\_scores\_3

stored as avro as

select sc.suppliersubjectnumber,

sc.SubjectNumber,

avg(cast( ninetydaypi as int)) as scoreninetydaypi,

avg(cast( ninetydayci as int)) as scoreninetydayci,

avg(cast( cdsscorevalue as int)) as scorecds,

avg(cast(bfrs as int)) as scoremktbfrs ,

avg(cast( ninetydaypastdueamountperiod1 as int)) as ninetydpdamtP1,

avg(cast( ninetydaypastdueamountperiod2 as int)) as ninetydpdamtP2,

avg(cast( ninetydaypastdueamountperiod3 as int)) as ninetydpdamtP3,

avg(av1.scorecds\_avg1 ) as scorecds\_avg1,

avg(av2.scorecds\_avg2 ) as scorecds\_avg2,

avg(av1.bfrs\_avg1) as bfrs\_avg1,

avg(av2.bfrs\_avg2 ) as bfrs\_avg2

from debtor\_iq\_db.DB\_scores1 as sc

left JOIN (select SubjectNumber, scorecds\_avg1, bfrs\_avg1 from debtor\_iq\_db.db\_scores\_av1) as av1 on sc.SubjectNumber= av1.SubjectNumber

left JOIN (select SubjectNumber, scorecds\_avg2, bfrs\_avg2 from debtor\_iq\_db.db\_scores\_av2) as av2 on sc.SubjectNumber= av2.SubjectNumber

group by sc.subjectnumber, sc.suppliersubjectnumber;

### Step 11. Merge all the tables to create the final one to use in Spotfire

|  |  |
| --- | --- |
| Macro Used | No |
| Final Table | No |

Change to adjust rolling average BFRS score

create table debtor\_iq\_db.DB\_6

stored as avro as

select

act.archive,

act.suppliersubjectnumber,

act.subjectnumber,

act.referencenumber,

act.total\_bal,

act.curr\_bal,

act.bal\_p1,

act.bal\_p2,

act.bal\_p3,

act.totalp1p2p3,

act.dpd,

act.pct\_curr,

act.pct\_pd1,

act.pct\_pd2,

act.pct\_pd3,

act.pi,

act.subjectname,

act.siccode1forthiscompany,

act.substr\_sic4,

act.naicscodeforthiscompany,

act.provinceorstatecode,

act.subjectaddress,

act.englishdescription60,

act.frenchdescription60,

cast(pi.pi\_avg3 as int) as pi\_avg3,

cast (pih.pi\_avg6 as int) as pi\_avg6

from debtor\_iq\_db.DB\_5 as act

LEFT JOIN (select suppliersubjectnumber,subjectnumber,ReferenceNumber, pi\_avg3 from debtor\_iq\_db.DB\_4) as pi

ON act.ReferenceNumber =pi.ReferenceNumber and

act.subjectnumber=pi.subjectnumber and

act.suppliersubjectnumber=pi.suppliersubjectnumber

LEFT JOIN (select suppliersubjectnumber,subjectnumber,ReferenceNumber, pi\_avg6 from debtor\_iq\_db.DB\_4H) as piH

ON act.ReferenceNumber =pih.ReferenceNumber and

act.subjectnumber=pih.subjectnumber and

act.suppliersubjectnumber=pih.suppliersubjectnumber;

### Step 12. Create final table for Visualizations FALTA CORRER

|  |  |
| --- | --- |
| Macro Used | No |
| Final Table | Yes |

create table debtor\_iq\_db.DIQ\_trades

stored as avro

as

select

act.\*,

sc.scoreninetydaypi,

sc.scoreninetydayci,

sc.scorecds,

sc.scoremktbfrs,

sc.ninetydpdamtp1,

sc.ninetydpdamtp2,

sc.ninetydpdamtp3,

scorecds\_avg1,

scorecds\_avg2,

bfrs\_avg1,

bfrs\_avg2,

case when cast(sc.scorecds as int)>509 then "Risk Class 1 (510+)"

when cast(sc.scorecds as int)>451 then "Risk Class 2 (452-509)"

when cast(sc.scorecds as int)>375 then "Risk Class 3 (376-451)"

when cast(sc.scorecds as int)>332 then "Risk Class 4 (333-375)"

when cast(sc.scorecds as int)>100 then "Risk Class 5 (101-332)"

when cast(sc.scorecds as int)=0 then "Risk Class 0 Bankruptcy"

else "No Score" end as scorecds\_bands,

case when cast(sc.scoremktbfrs as int)>1343 then "Risk Class 1 (1344+)"

when cast(sc.scoremktbfrs as int)>1309 then "Risk Class 2 (1310-1343)"

when cast(sc.scoremktbfrs as int)>1264 then "Risk Class 3 (1265-1309)"

when cast(sc.scoremktbfrs as int)>1204 then "Risk Class 4 (1205-1264)"

when cast(sc.scoremktbfrs as int)>1000 then "Risk Class 5 (1001-1204)"

else "No Score" end as bfrs\_bands,

case

when cast(sc.scoremktbfrs as int)>1343 and cast(sc.scorecds as int)>375 then "Low Risk"

when (cast(sc.scoremktbfrs as int)>1343 and cast(sc.scorecds as int)>100 and cast(sc.scorecds as int)<376) or (

cast(sc.scoremktbfrs as int)>1000 and cast(sc.scoremktbfrs as int)<1344 and cast(sc.scorecds as int)>375) then "Medium Risk"

when cast(sc.scoremktbfrs as int)<1344 and cast(sc.scoremktbfrs as int)>0 and cast(sc.scorecds as int)<376 and cast(sc.scorecds as int)>0

then "High Risk" end as Risk\_category,

case

when (act.pi\_avg3-act.pi\_avg6)>=5 then "Deteriorating Trend"

when (act.pi\_avg3-act.pi\_avg6)<=-5 then "Improving Trend"

when (act.pi\_avg3-act.pi\_avg6)>-5 and (act.pi\_avg3-act.pi\_avg6)<5 then "Stable Trend"

else "Trend not Available" end as PI\_trend,

case when cast(substr(substr\_sic4,1,2) as int) in (1,2,7,8,9) then "Agriculture, Forestry and Fishing"

when cast(substr(substr\_sic4,1,2) as int) in (10,12,13,14) then "Mining"

when cast(substr(substr\_sic4,1,2) as int) in (15,16,17) then "Construction"

when cast(substr(substr\_sic4,1,2) as int)>19 and cast(substr(substr\_sic4,1,2) as int)<40 then "Manufacturing"

when cast(substr(substr\_sic4,1,2) as int)>39 and cast(substr(substr\_sic4,1,2) as int)<50 then "Transport, Comms, Electric, Gas, Sanitary Svs"

case when (cast(substr(substr\_sic4,1,2) as int) in (1,2,7,8,9) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('11'))) then "Agriculture, Forestry and Fishing"

when (cast(substr(substr\_sic4,1,2) as int) in (10,12,13,14)or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('21'))) then "Mining"

when (cast(substr(substr\_sic4,1,2) as int) in (15,16,17) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('23'))) then "Construction"

when ((cast(substr(substr\_sic4,1,2) as int)>19 and cast(substr(substr\_sic4,1,2) as int)<40) or

(cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('31', '32', '33'))) then "Manufacturing"

when ((cast(substr(substr\_sic4,1,2) as int)>39 and cast(substr(substr\_sic4,1,2) as int)<50) or

(cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('22','48', '49'))) then "Transportation"

when (cast(substr(substr\_sic4,1,2) as int) in (50,51) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('42'))) then "Wholesale Trade"

when ((cast(substr(substr\_sic4,1,2) as int)>51 and cast(substr(substr\_sic4,1,2) as int)<60) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('44', '45') )) then "Retail Trade"

when (cast(substr(substr\_sic4,1,2) as int) in (60,61,62,63,64,65,66,67) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('52', '53'))) then "Finance, Insurance and Real Estate"

when ((cast(substr(substr\_sic4,1,2) as int) >69 and cast(substr(substr\_sic4,1,2) as int) <90) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('51', '54', '55', '56', '61', '62', '71', '72', '81'))) then "Services"

when (cast(substr(substr\_sic4,1,2) as int) in (91,92,93,94,95,96,97) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('92'))) then "Public Administration"

when cast(substr(substr\_sic4,1,2) as int) in (99) then "Non-classifiable"

else " " end as Sic\_Division

from debtor\_iq\_db.DB\_6 as act

left outer JOIN (select \* from debtor\_iq\_db.DB\_scores\_3) as sc

ON act.subjectnumber=sc.subjectnumber and

act.suppliersubjectnumber=sc.suppliersubjectnumber ;

# DETOKENIZATION

### Step 13. DETOKENIZE TRADES DATA

*ALTER TABLE debtor\_iq\_db.diq\_trades RENAME TO debtor\_iq\_db.diq\_trades\_token;*

*create table debtor\_iq\_db.diq\_trades*

*stored as avro as*

*select*

*archive,*

*suppliersubjectnumber,*

*subjectnumber,*

*referencenumber,*

*total\_bal,*

*curr\_bal,*

*bal\_p1,*

*bal\_p2,*

*bal\_p3,*

*totalp1p2p3,*

*dpd,*

*pct\_curr,*

*pct\_pd1,*

*pct\_pd2,*

*pct\_pd3,*

*pi,*

*default.ptyunprotectstr(subjectname,'TK\_COMMERCIAL\_TEXT') as subjectname,*

*siccode1forthiscompany,*

*substr\_sic4,*

*naicscodeforthiscompany,*

*default.ptyunprotectstr(provinceorstatecode,'TK\_COMMERCIAL\_TEXT') as provinceorstatecode,*

*subjectaddress,*

*englishdescription60,*

*frenchdescription60,*

*pi\_avg3,*

*pi\_avg6,*

*scoreninetydaypi,*

*scoreninetydayci,*

*scorecds,*

*scoremktbfrs,*

*ninetydpdamtp1,*

*ninetydpdamtp2,*

*ninetydpdamtp3,*

*scorecds\_avg1,*

*scorecds\_avg2,*

*bfrs\_avg1,*

*bfrs\_avg2,*

*scorecds\_bands,*

*bfrs\_bands,*

*risk\_category,*

*pi\_trend,*

*sic\_division*

*from debtor\_iq\_db.diq\_trades\_token ;*

### Step 13 (b). DETOKENIZE TRADES DATA FROM PREVIOUS MONTH

**ONLY APPLICABLE FOR THIS FIRST RUN**

*ALTER TABLE debtor\_iq\_db.diq\_tradesprev RENAME TO debtor\_iq\_db.diq\_tradesprev\_token;*

*create table debtor\_iq\_db.diq\_tradesprev*

*stored as avro as*

*select*

*archive,*

*suppliersubjectnumber,*

*subjectnumber,*

*referencenumber,*

*total\_bal,*

*curr\_bal,*

*bal\_p1,*

*bal\_p2,*

*bal\_p3,*

*totalp1p2p3,*

*dpd,*

*pct\_curr,*

*pct\_pd1,*

*pct\_pd2,*

*pct\_pd3,*

*pi,*

*default.ptyunprotectstr(subjectname,'TK\_COMMERCIAL\_TEXT') as subjectname,*

*siccode1forthiscompany,*

*substr\_sic4,*

*naicscodeforthiscompany,*

*default.ptyunprotectstr(provinceorstatecode,'TK\_COMMERCIAL\_TEXT') as provinceorstatecode,*

*subjectaddress,*

*englishdescription60,*

*frenchdescription60,*

*pi\_avg3,*

*pi\_avg6,*

*scoreninetydaypi,*

*scoreninetydayci,*

*scorecds,*

*scoremktbfrs,*

*ninetydpdamtp1,*

*ninetydpdamtp2,*

*ninetydpdamtp3,*

*scorecds\_avg1,*

*scorecds\_avg2,*

*bfrs\_avg1,*

*bfrs\_avg2,*

*scorecds\_bands,*

*bfrs\_bands,*

*risk\_category,*

*pi\_trend,*

*sic\_division*

*from debtor\_iq\_db.diq\_tradesprev\_token ;*

### Step 14 UNION of TRADES current ARCHIVE and PREVIOUS

**ON THE ONGOING BASIS WHAT WILL HAPPEN IS THAT THE PREVIOUS MONTHS’ “CURRENT ARCHIVE” WILL BECOME THE PREVIOUS ARCHIVE OF THIS NEW RUN. SO THE RUN SHOULD BE DONE ONLY FOR CURRENT ARCHIVE and create a UNION with DIQ\_TRADES from previous RUN and keep the most recent archives.**

*ALTER TABLE debtor\_iq\_db.diq\_trades RENAME TO debtor\_iq\_db.diq\_tradesact;*

*create table debtor\_iq\_db.diq\_trades*

*stored as avro as*

*Select \**

*from debtor\_iq\_db.diq\_tradesprev as c*

*UNION ALL SELECT \* from debtor\_iq\_db.diq\_tradesact as b;*

# DIQ\_trades – Contents

|  |  |
| --- | --- |
| **NAME** | **Description** |
| diq\_trades.archive | Archive (KEY) |
| diq\_trades.suppliersubjectnumber | Supplier ID (KEY) |
| diq\_trades.subjectnumber | Subjectnumber |
| diq\_trades.referencenumber | referencenumber |
| diq\_trades.total\_bal | Total Balance |
| diq\_trades.curr\_bal | Balance in Good standing |
| diq\_trades.bal\_p1 | Balance in DQ1 |
| diq\_trades.bal\_p2 | Balance in DQ2 |
| diq\_trades.bal\_p3 | Balance in DQ3 |
| diq\_trades.totalp1p2p3 | Balance Past Due |
| diq\_trades.dpd | Days Past Due |
| diq\_trades.pct\_curr | % Balance in Current Status |
| diq\_trades.pct\_pd1 | % Balance in DQ1 |
| diq\_trades.pct\_pd2 | % Balance in DQ2 |
| diq\_trades.pct\_pd3 | % Balance in DQ3 |
| diq\_trades.pi | Payment Index |
| diq\_trades.subjectname | Subjectname |
| diq\_trades.siccode1forthiscompany | SicCode 1 |
| diq\_trades.substr\_sic4 | Sic Code 4 Digit |
| diq\_trades.naicscodeforthiscompany | NACIS Code |
| diq\_trades.provinceorstatecode | Province or State |
| diq\_trades.subjectaddress | Address (Tokenized) |
| diq\_trades.englishdescription60 | English description of SIC Code |
| diq\_trades.frenchdescription60 | French description of SIC Code |
| diq\_trades.pi\_avg3 | Average PI (last 3 months) |
| diq\_trades.pi\_avg6 | Average PI (months 4,5,6 prior) |
| diq\_trades.scoreninetydaypi | Score NinetyDay PI |
| diq\_trades.scoreninetydayci | Score NinetyDay CI |
| diq\_trades.scorecds | Score CDS |
| diq\_trades.scoremktbfrs | Score BFRS |
| diq\_trades.ninetydpdamtp1 | Amount DPD (DQ 1) with Market |
| diq\_trades.ninetydpdamtp2 | Amount DPD (DQ 2) with Market |
| diq\_trades.ninetydpdamtp3 | Amount DPD (DQ 3) with Market |
| diq\_trades.scorecds\_avg1 | Average CDS Score last 3 months |
| diq\_trades.scorecds\_avg2 | Average CDS Score last 4,5,6 months |
| diq\_trades.bfrs\_avg1 NEW | Average BFRS Score last 3 months |
| diq\_trades.bfrs\_avg2 NEW | Average BFRS Score last 4,5,6 months |
| diq\_trades.scorecds\_bands | CDS Band (per Kens definition) |
| diq\_trades.bfrs\_bands | BFRS Band (per Kens definition) |
| diq\_trades.risk\_category | Risk category (per Ken's definition) |
| diq\_trades.pi\_trend | PI Trend (average last 3 - prior 3 comparison) |
| diq\_trades.sic\_division (NEW LOGIC wit NAICS) | SIC Division according to KEN definition |

# DEBTOR IQ – AGGREGATED DATA – SCHEMAS

Generic RUN as of 201711 as current archive

set hivevar:ARCHIVES\_0n=201711;

set hivevar:ARCHIVES\_24=201511;

## TABLES AGGREGATED

4 levels of aggregations

1. Archive
2. Archive + Suppliersubjectnumber
3. Archive + SIC (4 digit)
4. Archive + Province or State code



### Step1. Balance aggregation, PI calculation at archive level

Keeps current archive + 23 previous archives

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 / ARCHIVES\_0n |
| Final Table | NO |

create table debtor\_iq\_db.DB\_1HIST24

stored as avro as

select archive,

cfile.provinceorstatecode,

alp.suppliersubjectnumber,

alp.ReferenceNumber ,

alp.subjectnumber,

alp.totalbalanceamount,

alp.currentbalanceamount,

alp.period1balanceamount,

alp.period2balanceamount,

alp.period3balanceamount,

alp.accountsecuredindicator ,

cast(alp.totalbalanceamount as int) as TOTAL\_BAL,

cast(alp.currentbalanceamount as int) as CURR\_BAL,

cast(alp.period1balanceamount as int) as BAL\_P1,

cast(alp.period2balanceamount as int) as BAL\_P2,

cast(alp.period3balanceamount as int) as BAL\_P3,

(cast(alp.period1balanceamount as int) + cast(alp.period2balanceamount as int)+cast(alp.period3balanceamount as int)) as TOTALP1P2P3,

case when cast(alp.totalbalanceamount as int)>0 then (cast(alp.currentbalanceamount as int)/cast(alp.totalbalanceamount as int))\*100 end as pct\_curr,

case when cast(alp.totalbalanceamount as int)>0 then (cast(alp.period1balanceamount as int)/cast(alp.totalbalanceamount as int))\*100 end as pct\_pd1,

case when cast(alp.totalbalanceamount as int)>0 then (cast(alp.period2balanceamount as int)/cast(alp.totalbalanceamount as int))\*100 end as pct\_pd2,

case when cast(alp.totalbalanceamount as int)>0 then (cast(alp.period3balanceamount as int)/cast(alp.totalbalanceamount as int))\*100 end as pct\_pd3

from (select \* from efx\_ca\_prod.comm\_file lateral view explode(alp000s) exploded\_table as alp

where cast(archive as int) > ${hivevar:ARCHIVES\_24} and alp.suppressionflag = 'N') as cfile1

left join (select cast(archive as int) as archive1 ,smp.provinceorstatecode, smp.subjectnumber as subjectnumber1

from efx\_ca\_prod.comm\_file lateral view explode(smp100s) exploded\_table as smp

where cast(archive as int)>${hivevar:ARCHIVES\_24}) as cfile

on

cfile.SubjectNumber1 = cfile1.alp.subjectnumber and

cfile.archive1=cfile1.archive

;

(\*) Province is from the subjectnumber at that archive. All subjectnumbers have only 1 province assigned per archive

**create table debtor\_iq\_db.DB\_PI\_SUPPLIER**

stored as avro as

select archive,suppliersubjectnumber,

sum(TOTAL\_BAL) as TOTAL\_BAL,

sum(CURR\_BAL) as CURR\_BAL,

sum(BAL\_P1) as BAL\_P1,

sum(BAL\_P2) as BAL\_P2,

sum(BAL\_P3) as BAL\_P3,

count(\*) as TOT\_TRADES,

sum(case when (BAL\_P3>0)then 1 else 0 end) as NUM\_P3,

sum(case when (BAL\_P2>0 and BAL\_P3<=0) then 1 else 0 end) as NUM\_P2,

sum(case when (BAL\_P1>0 and BAL\_P3<=0 and BAL\_P2<=0) then 1 else 0 end) as NUM\_P1,

sum(case when (CURR\_BAL>0 and (BAL\_P3+BAL\_P2+BAL\_P1)<=0) then 1 else 0 end) as NUM\_CURR,

sum(case when (TOTAL\_BAL>0) then 1 else 0 end) as NUM\_BAL,

sum(BAL\_P1+BAL\_P2+BAL\_P3) as TOTALP1P2P3,

avg(

case when (accountsecuredindicator = "N" and TOTAL\_BAL=0) then 0

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)>=99 then 99

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6)as int)<=0 then 0

when (accountsecuredindicator = "N") and 0<cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6)as int)<99 then

cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)

end) as PI,

sum(case when (accountsecuredindicator = "N" and TOTAL\_BAL>=0) then 1 else 0 end) as PI\_num

from debtor\_iq\_db.DB\_1HIST24

group by suppliersubjectnumber, archive;

(\*) if at some point there is a need for PI without ZEROS then the logic is =PI xPI\_NUM/ NUM\_BAL,

Change to adjust SIC

**create table**

create table debtor\_iq\_db.DB\_PI\_SUPPLIER1

stored as avro as

select sup.\*,

case when cast(length(mmp.newsic) as int) = 1 then concat ('000', substr(mmp.newsic,1,1))

when cast(length(mmp.newsic) as int) = 2 then concat ('00', substr(mmp.newsic,1,2))

when cast(length(mmp.newsic) as int) = 3 then concat ('0', substr(mmp.newsic,1,3))

else substr(mmp.newsic,1,4) end as newsic

from debtor\_Iq\_db.DB\_PI\_SUPPLIER as sup

left outer JOIN (select archive, newsic, membernumber

from efx\_ca\_prod.comm\_mmp000 where archive=${hivevar:ARCHIVES\_24}) as mmp

on sup.suppliersubjectnumber = mmp.membernumber;

### Step 2 Balance aggregation, PI calculation at archive + province level

create table debtor\_iq\_db.DB\_PI\_PROV

stored as avro as

select archive,provinceorstatecode,

sum(TOTAL\_BAL) as TOTAL\_BAL,

sum(CURR\_BAL) as CURR\_BAL,

sum(BAL\_P1) as BAL\_P1,

sum(BAL\_P2) as BAL\_P2,

sum(BAL\_P3) as BAL\_P3,

count(\*) as TOT\_TRADES,

sum(case when (BAL\_P3>0)then 1 else 0 end) as NUM\_P3,

sum(case when (BAL\_P2>0 and BAL\_P3<=0) then 1 else 0 end) as NUM\_P2,

sum(case when (BAL\_P1>0 and BAL\_P3<=0 and BAL\_P2<=0) then 1 else 0 end) as NUM\_P1,

sum(case when (CURR\_BAL>0 and (BAL\_P3+BAL\_P2+BAL\_P1)<=0) then 1 else 0 end) as NUM\_CURR,

sum(case when (TOTAL\_BAL>0) then 1 else 0 end) as NUM\_BAL,

sum(BAL\_P1+BAL\_P2+BAL\_P3) as TOTALP1P2P3,

avg(

case when (accountsecuredindicator = "N" and TOTAL\_BAL=0) then 0

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)>=99 then 99

when (accountsecuredindicator = "N") and cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6)as int)<=0 then 0

when (accountsecuredindicator = "N") and 0<cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6)as int)<99 then

cast(((pct\_curr+(pct\_pd1\*3)+(pct\_pd2\*5)+(pct\_pd3\*7)-100)/6) as int)

end) as PI,

sum(case when (accountsecuredindicator = "N" and TOTAL\_BAL>=0) then 1 else 0 end) as PI\_num

from debtor\_iq\_db.DB\_1HIST24

group by provinceorstatecode, archive;

### Step 3 Keep 24 months of CDS and BFRS score

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 |
| Final Table | NO |

create table debtor\_iq\_db.DB\_scores\_hist

stored as avro as

select

alp.suppliersubjectnumber,

alp.SubjectNumber ,

alp.totalbalanceamount as total\_bal,

sc.archive,

fullcds\_ot.cdsscorevalue,

fullcds\_ot.marketingbfrsscorevalue as bfrs,

alp.suppressionflag,

smp.provinceorstatecode

from (select \* from efx\_ca\_prod.comm\_scores where cast(archive as int) > ${hivevar:ARCHIVES\_24} ) as sc

left join (select \* from efx\_ca\_prod.comm\_file lateral view explode(alp000s) exploded\_table as alp

lateral view explode(smp100s) exploded\_table as smp

where cast(archive as int) > ${hivevar:ARCHIVES\_24} and alp.suppressionflag = 'N') as cfile

on cfile.alp.SubjectNumber = fullcds\_ot.equifaxsubjectnumber and

cfile.alp.SubjectNumber = cfile.smp.SubjectNumber and

cfile.archive = sc.archive

where alp.suppressionflag = 'N' ;

### Step 4. SCORES at archive + province level DB\_SCORE\_PROV

create table debtor\_iq\_db.DB\_SCORE\_PROV

stored as avro as

select

archive,

avg(cast(cdsscorevalue as int)) as avgprov\_cds,

avg(cast(bfrs as int)) as avgprov\_bfrs,

provinceorstatecode

from debtor\_iq\_db.db\_scores\_hist

group by archive, provinceorstatecode;

### Step 5. Create Score bands for each archive for last 24 periods grouped by supplier and archive – mergear con las otras & Compute CDS and BFRS Score for Supplier

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 |
| Final Table | NO |

create table create table debtor\_iq\_db.DB\_AGGCDS1

stored as avro as

select suppliersubjectnumber, archive,

count(1) as avg\_cdsnum,

avg(cast(cdsscorevalue as int)) as avg\_cds,

avg(cast(bfrs as int)) as avg\_bfrs,

sum(case when cast(cdsscorevalue as int)>509 then 1 else 0 end) as NUM\_CDSRiskClass1,

sum(case when cast(cdsscorevalue as int)>451 and cast(cdsscorevalue as int)<=509 then 1 else 0 end) as NUM\_CDSRiskClass2,

sum(case when cast(cdsscorevalue as int)>375 and cast(cdsscorevalue as int)<=451 then 1 else 0 end) as NUM\_CDSRiskClass3,

sum(case when cast(cdsscorevalue as int)>332 and cast(cdsscorevalue as int)<=375 then 1 else 0 end) as NUM\_CDSRiskClass4,

sum(case when cast(cdsscorevalue as int)>100 and cast(cdsscorevalue as int)<=332 then 1 else 0 end) as NUM\_CDSRiskClass5,

sum(case when cast(cdsscorevalue as int)=0 then 1 else 0 end) as NUM\_CDSRiskClass0,

sum(case when cast(cdsscorevalue as int)>509 then total\_bal else 0 end) as BAL\_CDSRiskClass1,

sum(case when cast(cdsscorevalue as int)>451 and cast(cdsscorevalue as int)<=509 then total\_bal else 0 end) as BAL\_CDSRiskClass2,

sum(case when cast(cdsscorevalue as int)>375 and cast(cdsscorevalue as int)<=451 then total\_bal else 0 end) as BAL\_CDSRiskClass3,

sum(case when cast(cdsscorevalue as int)>332 and cast(cdsscorevalue as int)<=375 then total\_bal else 0 end) as BAL\_CDSRiskClass4,

sum(case when cast(cdsscorevalue as int)>100 and cast(cdsscorevalue as int)<=332 then total\_bal else 0 end) as BAL\_CDSRiskClass5,

sum(case when cast(cdsscorevalue as int)=0 then total\_bal else 0 end) as BAL\_CDSRiskClass0,

sum(case when cast(bfrs as int)>1343 then 1 else 0 end) as NUM\_BFRSRiskClass1,

sum(case when cast(bfrs as int)>1309 and cast(bfrs as int)<=1343 then 1 else 0 end) as NUM\_BFRSRiskClass2,

sum(case when cast(bfrs as int)>1264 and cast(bfrs as int)<=1309 then 1 else 0 end) as NUM\_BFRSRiskClass3,

sum(case when cast(bfrs as int)>1204 and cast(bfrs as int)<=1264 then 1 else 0 end) as NUM\_BFRSRiskClass4,

sum(case when cast(bfrs as int)>1000 and cast(bfrs as int)<=1204 then 1 else 0 end) as NUM\_BFRSRiskClass5,

sum(case when cast(bfrs as int)<=1000 then 1 else 0 end) as NUM\_BFRSRiskClass0,

sum(case when cast(bfrs as int)>1343 then total\_bal else 0 end) as BAL\_BFRSRiskClass1,

sum(case when cast(bfrs as int)>1309 and cast(bfrs as int)<=1343 then total\_bal else 0 end) as BAL\_BFRSRiskClass2,

sum(case when cast(bfrs as int)>1264 and cast(bfrs as int)<=1309 then total\_bal else 0 end) as BAL\_BFRSRiskClass3,

sum(case when cast(bfrs as int)>1204 and cast(bfrs as int)<=1264 then total\_bal else 0 end) as BAL\_BFRSRiskClass4,

sum(case when cast(bfrs as int)>1000 and cast(bfrs as int)<=1204 then total\_bal else 0 end) as BAL\_BFRSRiskClass5,

sum(case when cast(bfrs as int)<=1000 then total\_bal else 0 end) as BAL\_BFRSRiskClass0

from debtor\_iq\_db.db\_scores\_hist

group by suppliersubjectnumber, archive ;

### Step 6. Compute CDS and BFRS Score for Market

|  |  |
| --- | --- |
| Macro Used | NO |
| Final Table | NO |

create table debtor\_iq\_db.DB\_AGGCDS2

stored as avro as

select archive,

avg(cast(cdsscorevalue as int)) as avg\_cds\_market,

avg(cast(bfrs as int)) as avg\_bfrs\_market

from debtor\_iq\_db.DB\_scores\_hist

group by archive;

### Step 7. Identify Supplier SICCODE

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 |
| Final Table | NO |

Change to adjust SIC

create table debtor\_iq\_db.db\_PI\_SIC

stored as avro as

select smp.archive ,

smp.subjecttype,

smp.subjectnumber,

case when cast(length(smp.siccode1forthiscompany) as int) = 1 then concat ('000', substr(smp.siccode1forthiscompany,1,1))

when cast(length(smp.siccode1forthiscompany) as int) = 6 then concat ('00', substr(smp.siccode1forthiscompany,1,2))

when cast(length(smp.siccode1forthiscompany) as int) = 7 then concat ('0', substr(smp.siccode1forthiscompany,1,3))

else substr(smp.siccode1forthiscompany,1,4) end as substr\_sic4

from efx\_ca\_prod.comm\_subject\_master as smp

where smp.archive >(${hivevar:ARCHIVES\_24} ) AND smp.archive <=(${hivevar:ARCHIVES\_0n}) and smp.subjecttype="P" ;

### Step 8. Create a table with PI at market and industry level

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 and ARCHIVES\_0 |
| Final Table | NO |

create table debtor\_iq\_db.DB\_PI

stored as avro as

select

cattr.archive,

cattr.subjectnumber,

cast(cattr.alllinespi as int) as PI

from (select cast(archive as int) as archive,sup000.alllinespi, sup000.subjectnumber

from efx\_ca\_prod.comm\_attributes where cast(archive as int) > ( ${hivevar:ARCHIVES\_24}) and cast(archive as int) <=( ${hivevar:ARCHIVES\_0n})) as cattr ;

/\* KEY ARCHIVE – MARKET LEVEL PI \*/

create table debtor\_iq\_db.DB\_PI\_MKT

stored as avro as

select archive,

case when avg(PI) >= 99 then 99

when avg(PI) <= 99 then avg(PI) end as avgmkt\_pi

from debtor\_iq\_db.DB\_PI

group by archive;

/\* KEY ARCHIVE – SIC4 LEVEL PI \*/

**create table debtor\_iq\_db.DB\_PI\_SIC2**

stored as avro as

select

cfile.archive1,

sicfile.substr\_sic4,

case when avg(cattr.PI) >= 99 then 99

when avg(cattr.PI) <= 99 then avg(PI) end as avgsic\_pi,

sum(case when cattr.PI>=0 then 1 else 0 end) as avgsic\_pinum

from (select cast(archive as int) as archive1 ,snum\_fam, smp.subjecttype, smp.subjectnumber as subjectnumber1

FROM efx\_ca\_prod.comm\_file lateral view explode(smp100s) exploded\_table as smp

where cast(archive as int)>${hivevar:ARCHIVES\_24}) as cfile

left JOIN ( select cast(archive as int) as archive,PI,subjectnumber from debtor\_iq\_db.DB\_PI

where cast(archive as int)>${hivevar:ARCHIVES\_24}) as cattr

on cfile.SubjectNumber1 = cattr.subjectnumber and

cfile.archive1=cattr.archive

left join (select substr\_sic4, cast(archive as int) as archive2 ,subjectnumber as subjectnumber2

from debtor\_iq\_db.db\_PI\_sic where cast(archive as int) >${hivevar:ARCHIVES\_24}) as sicfile

on cfile.snum\_fam=sicfile.subjectnumber2

and cfile.archive1=sicfile.archive2

group by sicfile.substr\_sic4, cfile.archive1;

### Step 9. Create a table SCORES at industry level

|  |  |
| --- | --- |
| Macro Used | ARCHIVES\_24 \_0 |
| Final Table | NO |

create table debtor\_iq\_db.DB\_SCORE\_SIC1

stored as avro as

select

cfile.archive1,

sicfile.substr\_sic4,

avg(cast(cattr.cdsscorevalue as int)) as avgsic\_cds,

avg(cast(cattr.bfrs as int)) as avgsic\_bfrs

from (select cast(archive as int) as archive1 ,snum\_fam, smp.subjecttype, smp.subjectnumber as subjectnumber1

FROM efx\_ca\_prod.comm\_file lateral view explode(smp100s) exploded\_table as smp where cast(archive as int)>${hivevar:ARCHIVES\_24} and

cast(archive as int)<=${hivevar:ARCHIVES\_0n}) as cfile

left JOIN ( select cast(archive as int) as archive,cdsscorevalue,bfrs,subjectnumber from debtor\_iq\_db.DB\_scores\_hist

where cast(archive as int)>${hivevar:ARCHIVES\_24}) as cattr

on cfile.SubjectNumber1 = cattr.subjectnumber and

cfile.archive1=cattr.archive

left join (select substr\_sic4, cast(archive as int) as archive2 ,subjectnumber as subjectnumber2

from debtor\_iq\_db.db\_PI\_sic where cast(archive as int) >${hivevar:ARCHIVES\_24}) as sicfile

on cfile.snum\_fam=sicfile.subjectnumber2

and cfile.archive1=sicfile.archive2

group by sicfile.substr\_sic4, cfile.archive1;

### Step 10 FINAL Merge table SCORES and PI at an industry code level

|  |  |
| --- | --- |
| Macro Used | NO |
| Final Table | YES |

create table debtor\_iq\_db.DIQ\_AGGSIC

stored as avro as

select smp.\*,

vtp.englishdescription60,

vtp.frenchdescription60,

score.avgsic\_bfrs,

score.avgsic\_cds

from debtor\_iq\_db.DB\_PI\_SIC2 as smp

left outer JOIN (select englishdescription60, frenchdescription60, siccode4digits , archive from efx\_ca\_prod.comm\_vtp318

where archive = ${hivevar:ARCHIVES\_0} ) as vtp

on smp.substr\_sic4 = vtp.siccode4digits

left outer JOIN (select avgsic\_cds, avgsic\_bfrs, archive1 as archive2, substr\_sic4 as substr\_sic41

from debtor\_iq\_db.DB\_SCORE\_SIC1) as score

on smp.substr\_sic4 = score.substr\_sic41

and smp.archive1=score.archive2

;

### Step 11. FINAL at SUPPLIER + ARCHIVE and MKT

|  |  |
| --- | --- |
| Macro Used | NO |
| Final Table | YES |

create table debtor\_iq\_db.DIQ\_AGGSUPPLIER

stored as avro as

select sup.\*,

cdss.avg\_cdsnum,

cdss.avg\_cds,

cdss.avg\_bfrs,

cdss.num\_cdsriskclass1,

cdss.num\_cdsriskclass2,

cdss.num\_cdsriskclass3,

cdss.num\_cdsriskclass4,

cdss.num\_cdsriskclass5,

cdss.num\_cdsriskclass0,

cdss.bal\_cdsriskclass1,

cdss.bal\_cdsriskclass2,

cdss.bal\_cdsriskclass3,

cdss.bal\_cdsriskclass4,

cdss.bal\_cdsriskclass5,

cdss.bal\_cdsriskclass0,

cdss.num\_bfrsriskclass1,

cdss.num\_bfrsriskclass2,

cdss.num\_bfrsriskclass3,

cdss.num\_bfrsriskclass4,

cdss.num\_bfrsriskclass5,

cdss.num\_bfrsriskclass0,

cdss.bal\_bfrsriskclass1,

cdss.bal\_bfrsriskclass2,

cdss.bal\_bfrsriskclass3,

cdss.bal\_bfrsriskclass4,

cdss.bal\_bfrsriskclass5,

cdss.bal\_bfrsriskclass0,

mktpi.avgmkt\_pi,

mktsc.avg\_cds\_market,

mktsc.avg\_bfrs\_market

from debtor\_Iq\_db.DB\_PI\_SUPPLIER1 as sup

left outer JOIN debtor\_iq\_db.Db\_AGGCDS1 as cdss

on cdss.suppliersubjectnumber=sup.suppliersubjectnumber and

cdss.archive=sup.archive

left outer JOIN debtor\_iq\_db.db\_pi\_mkt as mktpi

on sup.archive = mktpi.archive

left outer JOIN debtor\_iq\_db.db\_aggcds2 as mktsc

on sup.archive = mktsc.archive ;

### Step 12. FINAL at PROVINCE+ ARCHIVE

|  |  |
| --- | --- |
| Macro Used | NO |
| Final Table | YES |

create table debtor\_iq\_db.DIQ\_AGGPROV

stored as avro as

select pip.\*,

sc.avgprov\_cds,

sc.avgprov\_bfrs

from debtor\_iq\_db.DB\_PI\_prov as pip

left outer JOIN debtor\_iq\_db.db\_score\_prov as sc

on pip.provinceorstatecode=sc.provinceorstatecode and

pip.archive=sc.archive;

## DETOKENIZE AGGREGATED PROV TABLE

*ALTER TABLE debtor\_iq\_db.diq\_aggprov RENAME TO debtor\_iq\_db.diq\_aggprov\_token;*

*create table debtor\_iq\_db.diq\_aggprov*

*stored as avro as*

*select*

*archive,*

*default.ptyunprotectstr(provinceorstatecode,'TK\_COMMERCIAL\_TEXT') as provinceorstatecode,*

*total\_bal,*

*curr\_bal,*

*bal\_p1,*

*bal\_p2,*

*bal\_p3,*

*tot\_trades,*

*num\_p3,*

*num\_p2,*

*num\_p1,*

*num\_curr,*

*num\_bal,*

*totalp1p2p3,*

*pi,*

*pi\_num,*

*avgprov\_cds,*

*avgprov\_bfrs*

*from debtor\_iq\_db.diq\_aggprov\_token;*

# CONTENTS

**Aggregation PROVINCE+ ARCHIVE**

|  |  |
| --- | --- |
| diq\_aggprov.archive | Archive (KEY) YYYYMM |
| diq\_aggprov.provinceorstatecode | Province or State |
| diq\_aggprov.total\_bal | Total Balance |
| diq\_aggprov.curr\_bal | Balance in Good standing |
| diq\_aggprov.bal\_p1 | Balance in DQ1 |
| diq\_aggprov.bal\_p2 | Balance in DQ2 |
| diq\_aggprov.bal\_p3 | Balance in DQ3 |
| diq\_aggprov.tot\_trades | Total Trades (Count) |
| diq\_aggprov.num\_p3 | Number of trades with 90+DPD Balance > 0 |
| diq\_aggprov.num\_p2 | Number of trades with 60-90 DPD Balance > 0 |
| diq\_aggprov.num\_p1 | Number of trades with 30-60DPD Balance > 0 |
| diq\_aggprov.num\_curr | Number of trades with Current Balance > 0 (no DPD Balance) |
| diq\_aggprov.num\_bal | Number of trades with Balance > 0 |
| diq\_aggprov.totalp1p2p3 | Total 30+DPD Balance |
| diq\_aggprov.pi | Payment Index |
| diq\_aggprov.pi\_num | Number of trades used to calculate PI |
| diq\_aggprov.avgprov\_cds | Risk Score (CDS) |
| diq\_aggprov.avgprov\_bfrs | BFRS Score |

**Aggregation SUPPLIER + ARCHIVE**

|  |  |
| --- | --- |
| diq\_aggsupplier.archive | Archive (KEY) YYYYMM |
| diq\_aggsupplier.suppliersubjectnumber | Supplier ID (KEY) |
| diq\_aggsupplier.total\_bal | Total Balance |
| diq\_aggsupplier.curr\_bal | Balance in Good standing |
| diq\_aggsupplier.bal\_p1 | Balance in DQ1 |
| diq\_aggsupplier.bal\_p2 | Balance in DQ2 |
| diq\_aggsupplier.bal\_p3 | Balance in DQ3 |
| diq\_aggsupplier.tot\_trades | Total Trades (Count) |
| diq\_aggsupplier.num\_p3 | Number of trades with 90+DPD Balance > 0 |
| diq\_aggsupplier.num\_p2 | Number of trades with 60-90 DPD Balance > 0 |
| diq\_aggsupplier.num\_p1 | Number of trades with 30-60DPD Balance > 0 |
| diq\_aggsupplier.num\_curr | Number of trades with Current Balance > 0 (no DPD Balance) |
| diq\_aggsupplier.num\_bal | Number of trades with Balance > 0 |
| diq\_aggsupplier.totalp1p2p3 | Total 30+DPD Balance |
| diq\_aggsupplier.pi | Payment Index of Portfolio |
| diq\_aggsupplier.pi\_num | Number of trades used to calculate PI |
| diq\_aggsupplier.newsic | Sic Code 4 Digit |
| diq\_aggsupplier.avg\_cdsnum | #trades used to calculate avg. CDS Score, BFRS per supplier and archive |
| diq\_aggsupplier.avg\_cds | Risk Score (CDS) |
| diq\_aggsupplier.avg\_bfrs | BFRS Score |
| diq\_aggsupplier.num\_cdsriskclass1 | Risk Score Class 1 (>509) |
| diq\_aggsupplier.num\_cdsriskclass2 | Risk Score Class 2 (451< X <= 509) |
| diq\_aggsupplier.num\_cdsriskclass3 | Risk Score Class 3 (375< X <=451) |
| diq\_aggsupplier.num\_cdsriskclass4 | Risk Score Class 4 (332< X <= 375) |
| diq\_aggsupplier.num\_cdsriskclass5 | Risk Score Class 5 (100< X <=332 |
| diq\_aggsupplier.num\_cdsriskclass0 | Risk Score Class 0 (X = 0) |
| diq\_aggsupplier.bal\_cdsriskclass1 | Sum of balance with Risk Score Class 1 (>509) |
| diq\_aggsupplier.bal\_cdsriskclass2 | Sum of balance with Risk Score Class 2 (451< X <= 509) |
| diq\_aggsupplier.bal\_cdsriskclass3 | Sum of balance with Risk Score Class 3 (375< X <=451) |
| diq\_aggsupplier.bal\_cdsriskclass4 | Sum of balance with Risk Score Class 4 (332< X <= 375) |
| diq\_aggsupplier.bal\_cdsriskclass5 | Sum of balance with Risk Score Class 5 (100< X <=332 |
| diq\_aggsupplier.bal\_cdsriskclass0 | Sum of balance with Risk Score Class 0 (X = 0) |
| diq\_aggsupplier.num\_bfrsriskclass1 | BFRS Score Class 1 |
| diq\_aggsupplier.num\_bfrsriskclass2 | BFRS Score Class 2 |
| diq\_aggsupplier.num\_bfrsriskclass3 | BFRS Score Class 3 |
| diq\_aggsupplier.num\_bfrsriskclass4 | BFRS Score Class 4 |
| diq\_aggsupplier.num\_bfrsriskclass5 | BFRS Score Class 5 |
| diq\_aggsupplier.num\_bfrsriskclass0 | BFRS Score Class 0 |
| diq\_aggsupplier.bal\_bfrsriskclass1 | Sum of balance with BFRS Score Class 1 |
| diq\_aggsupplier.bal\_bfrsriskclass2 | Sum of balance with BFRS Score Class 2 |
| diq\_aggsupplier.bal\_bfrsriskclass3 | Sum of balance with BFRS Score Class 3 |
| diq\_aggsupplier.bal\_bfrsriskclass4 | Sum of balance with BFRS Score Class 4 |
| diq\_aggsupplier.bal\_bfrsriskclass5 | Sum of balance with BFRS Score Class 5 |
| diq\_aggsupplier.bal\_bfrsriskclass0 | Sum of balance with BFRS Score Class 0 |
| diq\_aggsupplier.avgmkt\_pi | Payment Index |
| diq\_aggsupplier.avg\_cds\_market | Risk Score |
| diq\_aggsupplier.avg\_bfrs\_market | BFRS Score |

**Aggregation SIC code + ARCHIVE**

|  |  |
| --- | --- |
| diq\_aggsic.archive1 | Archive (KEY) YYYYMM |
| diq\_aggsic.substr\_sic4 | SIC code 4 digits (text) |
| diq\_aggsic.avgsic\_pi | Payment Index |
| diq\_aggsic.avgsic\_pinum | Number of trades used to calculate PI |
| diq\_aggsic.englishdescription60 | Sic code description (english) |
| diq\_aggsic.frenchdescription60 | Sic code description (french) |
| diq\_aggsic.avgsic\_bfrs | Risk Score |
| diq\_aggsic.avgsic\_cds | BFRS Score |

# Other types of Categories defined.

## Provinceorstatecode

|  |  |  |  |
| --- | --- | --- | --- |
| provinceorstatecode |  | comm\_file distribution |  |
| AB |  | 14.3% | Alberta (Alb.) |
| BC |  | 14.9% | British Columbia (B.C.) |
| MB |  | 3.5% | Manitoba (Man.) |
| NB |  | 1.8% | New Brunswick (N.B.) |
| NL |  | 1.2% | Newfoundland and Labrador (N.L.) |
| NS |  | 2.6% | Nova Scotia  (N.S.) |
| NT |  | 0.1% | Northwest Territories (N.W.T.) |
| NU |  | 0.0% | Nunavut (Nvt.) |
| ON |  | 33.1% | Ontario (Ont.) |
| PE |  | 0.4% | Prince Edward Island (P.E.I.) |
| QC |  | 23.1% | Quebec  (Que.) |
| SK |  | 4.2% | Saskatchewan (Sask.) |
| YT |  | 0.1% | Yukon (Y.T.) |

Archives 201710 -201709 have 99.3% of referencenumbers falling in those categories.

All rest should be named ‘non-CDN’.

## SIC code Division is addressed in the MRD

Check MRD. Sic code format in the Cambrian databases has no zeros up front so when length was < 8, 0’s were added accordingly. To be able to match appropriatedly with vtp318- Note that although siccode1forthiscompany has 8 digits the last 4 are 0000.

NAICS only 13,000 over a 5.9MM subjectnumbers (1 archive) has a NAICS code and has NO SIC code. Coverage on SIC code is ~49%.

Below is a mapping of 2 digit SIC codes to NAISC codes:

|  |  |  |
| --- | --- | --- |
|  | **Applicable 2 digit codes** | |
| **SIC Divisions** | **SIC** | **NAISC** |
| **Agriculture** | 1-2, 7-9 | 11 |
| **Mining** | 10-14 | 21 |
| **Construction** | 15-17 | 23 |
| **Manufacturing** | 20-39 | 31-33 |
| **Transportation** | 40-49 | 22, 48-49 |
| **Wholesale Trade** | 50-51 | 42 |
| **Retail Trade** | 52-59 | 44-45 |
| **Finance** | 60-67 | 52-53 |
| **Services** | 70-89 | 51, 54-56, 61-62, 71-72, 81 |
| **Public Administration** | 91-97 | 92 |
| **Non-classifiable** | 99 |  |

case when (cast(substr(substr\_sic4,1,2) as int) in (1,2,7,8,9) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('11'))) then "Agriculture, Forestry and Fishing"

when (cast(substr(substr\_sic4,1,2) as int) in (10,12,13,14)or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('21'))) then "Mining"

when (cast(substr(substr\_sic4,1,2) as int) in (15,16,17) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('23'))) then "Construction"

when ((cast(substr(substr\_sic4,1,2) as int)>19 and cast(substr(substr\_sic4,1,2) as int)<40) or

(cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('31', '32', '33'))) then "Manufacturing"

when ((cast(substr(substr\_sic4,1,2) as int)>39 and cast(substr(substr\_sic4,1,2) as int)<50) or

(cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('22','48', '49'))) then "Transportation"

when (cast(substr(substr\_sic4,1,2) as int) in (50,51) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('42'))) then "Wholesale Trade"

when ((cast(substr(substr\_sic4,1,2) as int)>51 and cast(substr(substr\_sic4,1,2) as int)<60) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('44', '45') )) then "Retail Trade"

when (cast(substr(substr\_sic4,1,2) as int) in (60,61,62,63,64,65,67) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('52', '53'))) then "Finance, Insurance and Real Estate"

when ((cast(substr(substr\_sic4,1,2) as int) >69 and cast(substr(substr\_sic4,1,2) as int) <90) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('51', '54', '55', '56', '61', '62', '71', '72', '81'))) then "Services"

when (cast(substr(substr\_sic4,1,2) as int) in (91,92,93,94,95,96,97) or (cast(substr(substr\_sic4,1,2) as int) <1 and substr(naicscodeforthiscompany,1,2) in ('92'))) then "Public Administration"

when cast(substr(substr\_sic4,1,2) as int) in (99) then "Non-classifiable"

else " " end as Sic\_Division

## CDS TREND

1. Logic is based on a rolling 6 month timeframe
2. Calculate the average CDS score for the most recent 3 months in that 6 month window
3. Calculate the average CDS score for the oldest 3 months in that 6 month window
4. Compare the two 3 month average values:
   1. **Improving** trend if the delta between the most recent 3 month average and the older 3 month average is less than or equal to **+**15
   2. **Deteriorating** trend if the delta between the most recent 3 month average and the older 3 month average is greater than or equal to **-**15
   3. **Stable** trend if the delta between the most recent 3 month average and the older 3 month average is greater than -15 but less than +15

## BFRS TREND

The BFRS trend logic should be calculated as follows:

1. Logic is based on a rolling 6 month timeframe

2. Calculate the average BFRS score for the most recent 3 months in that 6 month window

3. Calculate the average BFRS score for the oldest 3 months in that 6 month window

4. Compare the two 3 month average values:

a. Improving trend if the delta between the most recent 3 month average and the older 3 month average is less than or equal to +45

b. Deteriorating trend if the delta between the most recent 3 month average and the older 3 month average is greater than or equal to -45

c. Stable trend if the delta between the most recent 3 month average and the older 3 month average is greater than -15 but less than +45

## PI TREND

case

when (act.pi\_avg3-act.pi\_avg6)>=5 then "Deteriorating Trend"

when (act.pi\_avg3-act.pi\_avg6)<=-5 then "Improving Trend"

when (act.pi\_avg3-act.pi\_avg6)>-5 and (act.pi\_avg3-act.pi\_avg6)<5 then "Stable Trend"

else "Trend not Available" end as PI\_trend,

## CDS class score ranges:

case when (cast(sc.scorecds as int)>509 then "Risk Class 1 (510+)"

when (cast(sc.scorecds as int)>451 then "Risk Class 2 (452-509)"

when (cast(sc.scorecds as int)>375 then "Risk Class 3 (376-451)"

when (cast(sc.scorecds as int)>332 then "Risk Class 4 (333-375)"

when (cast(sc.scorecds as int)>100 then "Risk Class 5 (101-332)"

when (cast(sc.scorecds as int)=0 then "Risk Class 0 Bankruptcy"

else "No Score" end as scorecds\_bands,

## BFRS class score ranges:

BFRS class 1: 1344 – 1722

BFRS class 2: 1310 - 1343

BFRS class 3: 1265 - 1309

BFRS class 4: 1205 - 1264

BFRS class 5: 1001 – 1204

## RISK CLASS DEFINITION

1. **Low Risk** = BFRS class 1 (i.e. BFRS score between 1344 – 1722) **AND** CDS class 1,2 or 3 (i.e. CDS score between 376 – 611)
2. **Medium Risk** (**either** of the scenarios below are considered medium risk):
   1. **(scenario 1)** = BFRS class 1 (i.e. BFRS score between 1344 – 1722) **AND** CDS class 4 or 5 (i.e. CDS score between 101 - 375)  **OR**
   2. **(scenario 2)** = BFRS class 2,3,4 or 5 (i.e. BFRS score between 1001 - 1343) **AND** CDS class 1,2 or 3 (i.e. CDS score between 376 – 611)
3. **High Risk** = BFRS class 2,3,4 or 5 (i.e. BFRS score between 1001 - 1343) **AND** CDS class 4 or 5 (i.e. CDS score between 101 - 375)

**A score of 0 in CDS or BFRS indicates bankruptcy; I think those accounts should be excluded from the low/med/high** risk categories. There are no NULLS

## Collectable Accounts

([Balance in Current Status]>5000) And (Integer([comm\_file\_wo\_7.pi])>30) and (Integer([comm\_file\_wo\_7.scoreninetydaypi])<60 and (Integer([comm\_file\_wo\_7.scoreninetydaypi]) Is Not Null) and (Integer([comm\_file\_wo\_7.pi]) Is Not Null))

([Balance in Current Status]>XXX) And (Integer([comm\_file\_wo\_7.pi])>30) and (Integer([comm\_file\_wo\_7.scoreninetydaypi])<60)

## High Risk Account

([comm\_file\_wo\_7.scorecds]<350) and ([Balance Past Due]>=0) and ([Total Balance]>=1000) – In the attribute I assigned XXX =350  and 1000

(Each Debor IQ Customer to input Values XXX : Score LT XXXX and Total Balance >= XXXX)

## Performing Accounts

(([comm\_file\_wo\_7.scorecds]>=475) and ([Total Balance]<=100000) and Integer([comm\_file\_wo\_7.pi])<=30 and Integer([comm\_file\_wo\_7.pi]) is Not Null)

**BFRS class score ranges:**

Risk class 1: 1344 – 1722

Risk class 2: 1310 - 1343

Risk class 3: 1265 - 1309

Risk class 4: 1205 - 1264

Risk class 5: 1001 – 1204