

CAPTURING LOGIC OF DERIVED MEASURES

① Top 100 Tax Defaulters

[year-hist-paid]: max(year) where amount is non-zero.

[pending-tax-current-year]: avg. tax paid historically $\times$ no. of years where payment is pending.

Example:

08-09	10	10-11	10
09-10	20	11-12	20

default from (12-13) to (16-17) = 5 years.

avg. hist payment = $(10 + 20 + 10 + 20) / 4 = 15$

pending tax = $15 \times 5 = 75$.

② Top 100 Mismatch in Tax paid in 2016-17

[avg-tax-before-16-17]: avg. of tax payments per year using historical info.

[amt-tax-16-17]: tax paid in 2016-17.

[amount-mismatch]: (avg-tax-before-16-17) - (amt-tax-16-17).

Example:

Year	Amt
08-09	100
09-10	~~110~~
10-11	100
11-12	~~110~~
12-13	100
13-14	~~110~~
14-15	100
15-16	~~110~~
16-17	65

$$\begin{aligned} \text{avg} &= (100 + 110 + 100 + 110 + \dots + 110) / 8 \\ &= 115. \end{aligned}$$

$$\text{tax-16-17} = 65.$$

$$\text{amount-mismatch} = 115 - 65 = 50.$$

③ Top 100 duplicate application Number

[mismatch-year]: assessment year where there is duplication of App. No.

[mismatch-app-no]: App. No. for which there is a (mismatch) b/w NEW & OLD sources [~~in tax paid~~] (duplicate records)

[total-mismatch-amt]: sum total of all amounts [tax paid] where the given (mismatch-app-no) occurs.

Example

Year	App. No.	08-09	09-10	...	16-17
08-09	1002	10	100		500
08-09	1002	10	100		500
08-09	1002	10	100		500
16-17	8006	20	200		1000
16-17	8006	20	200		1000

mismatch-year	mismatch-app-no	total-mismatch-amt
08-09	1002	30.
16-17	8006	2000.

④ Top 100 Properties with High Variance
in historical tax payment

[variance-cov]: Ratio of std. Dev (σ) by
 Mean (μ) for a given property
 () based on hist. tax paid.

$$\boxed{COV = \frac{\sigma}{\mu}} \rightarrow \text{filter applied here.} \\ \underline{COV > 50\%}$$

[potential]: Deviation/fluctuation in hist
 tax payment [Amt].

$$\text{potential} = COV \times \text{avg. tax paid.} \\ = \frac{\sigma}{\mu} \times \mu = \sigma$$

Example.

Year	Tax
11-12	20
12-13	150
13-14	180
14-15	40
15-16	170
16-17	50

$$COV = \text{std. Dev} [20, 150, 180, 40, 170, 50]$$

$$\text{Avg} [20, 150, 180, 40, 170, 50]$$

$$= \frac{66}{102} = 65\%$$

$$\text{potential} = 65\% \times 102 (\text{avg}) \\ = \underline{\underline{66}}$$

⑤ Top 10 streets with high variance
in historical tax payments

[street-potential]: overall potential opportunity
for additional tax collection

$$= \left\{ 80\% \times \left[\begin{array}{l} \text{Max tax} \\ \text{paid by} \\ \text{any property} \\ \text{on a street} \end{array} \right] \times \left[\begin{array}{l} \text{no. of} \\ \text{properties} \\ \text{on the} \\ \text{street} \end{array} \right] \right\} - \text{tax paid by street}$$

Example

Street no	# pcts	PID	owner	total tax 05-17
M001	5	001	A	100
M001	5	002	B	80
N001	5	003	C	70
N001	5	004	D	70
N001	5	005	E	65

Max tax paid by any property = 100.

no. of properties = 5

tax paid by street = $100 + 80 + 70 + 70 + 65 = 385$

$$\begin{aligned} \text{street-potential} &= (80\% \times 100 \times 5) - 385 \\ [M1001] &= 400 - 385 = \underline{\underline{15}} \end{aligned}$$

(b) Top 100 mismatch b/w NEW & OLD
Source [in Amt/tax paid]

[mismatch-year]: Year where for a given SAS ID & App. No. there is a mismatch in Tax Paid Info. b/w NEW & OLD sources

[app-no-mismatch-amt]: App. No. where there is a mismatch in Amt/tax paid [within an assessment year - across NEW & OLD info. sources]

[new-amount]: Amt according to NEW PDFs

[old-amount]: Amt according to old MySQL table(s)

[abs-potential]: Absolute difference b/w (new-amount) & (old-amount)

Example

Source	SAS ID	Year	App No	Amt/Tax
New	1001	2008-09	100	100
Old	1001	2008-09	100	50
New	2001	2016-17	252	300
Old	2001	2016-17	252	4000

mismatch year	app-no-mismatch-amt	new amount	old amount	abs potential
2008-09	100	100	50	+ 50
2016-17	252	300	4000	+ 3500

⑦ Top 100 tax payers across years

[min-rank]: highest rank [1 > 10]
obtained on the tax payer list
based on tax paid [~~with~~for that year]

[max-rank]: lowest rank [10 < 1]
obtained on the tax payer list
based on tax paid [for that year]

[potential]: potential for additional tax
collection

$$= \left[80\% \times \text{max tax ever paid} \times \text{no. of years active} \right]$$

— [total tax paid to date].

Example:

Year	Amnt	Rank
10-11	1000	6
11-12	300	56
12-13	500	21
13-14	800	12
14-15	50	121
15-16	350	45
16-17	1005	6

Note: Rank based on all
properties within the
ward, for that year

$$\text{min-rank} = 6$$

$$\text{max-rank} = 121$$

$$\text{potential} = \left[80\% \times 1005 \times 7 \right]$$

$$= \left[1000 + 300 + \dots + 1005 \right]$$

$$= \textcircled{1623}$$

⑧ Top 100 PIDs with duplicate information

[no-of-records]: No. of entries with same PID (for a given PID)

[total-sas-tax]: sum total of tax paid by a given SASID, across all years.

[total-pid-tax]: sum total of tax paid by all SASIDs with a given PID.

Example

PID	SAS ID	08-09	09-10	10-11	11-12	12-13	13-14	14-15
P01	1001	10	10	10	10	10	10	10
P01	1002	20	20	20	20	20	20	20
P01	1010	30	30	30	30	30	30	30
P02	2031	100	100	100	100	100	100	100
P02	2089	200	200	200	200	200	200	200

PID	SAS	total sas tax	total pid tax
P01	1001	70	420
P01	1002	140	420
P01	1010	210	420
P02	2031	700	2100
P02	2089	1400	2100

$$420 = 70 + 140 + 210$$

$$2100 = 700 + 1400$$