Stef. cards, '69 num. - Gaffney

Page -

Here is a real challenge!

Orig. and 4

cc: Paul Tsao

P. I. Prentice ("To show part of the enormous effort required to develop usable valid data.")

Complete Programming Instructions for Staffing analysis of Indl. ~~Plant Survey~~ and the (3 signif. figures in results is plenty.) Survey, Milw. City.

I. Variables

II. Operations on single variables

III. Pairings

IV. Method of pairing.

I. The Variables

	Col. Nos.	Unit	Range
A. Plant Area (PA).	8-10	000 sf	1-999
~~Short: 000 sq. ft.~~ "0" indicates tenant status			
B. Years in Business (YB)	21, 22	Years	1-99
~~(For tenants, use this as age of bldg.)~~			
C. Years at location (YL)	23, 24	Years	1-99
~~(For owners, use this as age of bldg.)~~			
D. Depreciated Plant Area (DPA)	Derived	~~000 sf~~ 1/10 acre	1-999

$$DPA \equiv PA \times (.95)^{age} * \div 4.356$$

Where Age = [YL + 60] for tenants ‡

Age = YL for owners. [Except if col. 20 is 2, age = YL + 30] [If YL blank, use YB]

Tenant status shown by "0" in cols. 8-10, 11-13, and 14-19

---

‡ Rationale: a tenant in an old bldg. probably moved there, so plant age > his YL. An owner without a 2 in col. 20 probably built the plant, so plant age = YL. An owner with a 2 in col. 20 bought the plant. Bought plants are probably newer than tenant occ. plants.

	Cols	Unit	Range
E. Land Assmt. (LAS)	11-13	$(000)	1-999
F. Improvemt. Assmt. (IAS)	14-19	$(000)	1-999,999
G. Total Assmt. (TAS)	Derived	$(000)	1-999,999

$$TAS = LAS + IAS$$

H. Land Area (La)	25-28	1/10 acre	.1-999.9
I. Census Tract (CT)	30-32	#	1-221

J. Assessment ratio. (AR)

A set of 18 numbers, (appendix F) (separate listing). ~~Divide~~

Every census tract falls in one of 18 groups with a separate assessment ratio.

~~Divide (G), total assessment, by appropriate (J), assessment ratio's, to get:~~

K. Equalized Total Assessment (ETas)	Derived (see J)	$(000)	1-999,999

Divide G, (Tas) by appropriate J, (AR), so: $ETAS = \frac{TAS}{AR}$

L. Expansion Area (Ea)	33,34	1/10 acre	0-9.9
M. 2dry. Exp. Area (E2a)	35,36	acre	0-99

N. Parking Places (PgP) — 37-40 — # — 0 - 9,999

O. Parking area (Pg. a) — Derived — $\frac{1}{10}$ acre — 0 - 9,999

$Pg.a = PgP \times \frac{1}{10}$ ≠

~~P. Employees — 46-50 — # — 5 - 99,999~~

P. Total Land area (TLa) — Derived — $\frac{1}{10}$ acre — .1 - 9,999.

(P) = (H) + (L) + 10(M) + (O)

or $TLa = La + Ea + 10 \cdot [E\,2\,a] + PgA$

Q. Employees (EMP) — 46-50 — # — 5 - 99,999

R. % Use autos (%ua) — 56,57 — % — 0 - 99

S. No. of autos (A$\frac{3}{8}$u) — Derived — # — 5 - 99,999

Multiply ~~Q x R = AU~~ EMP x (%ua) = au

(Q) x (R) = (S)

T. Need for Pg a (NPgA) — Derived — $\frac{1}{10}$ acre — .5 - 9,999

Multiply (S) $\times \frac{1}{10}$ = NPgA in $\frac{1}{10}$ acres

---

≠ Rationale: One Pg. P. requires 400 sf ≅ $\frac{1}{100}$ acre.

So Pg a in $\frac{1}{10}$ acres = $\frac{1}{10}$ PgP

U. Sewer (S) — 62, 63 — "1" means outside sewered area — 0 - 1

V. Truck Traffic (TT) — 64-66 — ## — 0 - 999

~~W. Rail Traffic (RT)~~ — ~~67~~ — ~~"1" means yes 2-4 means no~~ — ~~1 - 4~~

W. Preferred Location (PLO) — 85-87 — Census Tract — 1 - 221

X. Sewered Land Area (SLA) — Derived — $\frac{1}{10}$ acre — .1 - 999.9

Count all LA [(H), Cols. 25-28)] where ($U_{(Cols. 62,63)}$) shows sewer.

Y. Unsewered Land Area (USLA) — Derived — $\frac{1}{10}$ acre — .1 - 999.9

Count all LA where $U_{(Cols. 62,63)}$ shows ~~sewer~~ outside sewered area.

Z. Land Area Adjusted for sewer (LAAS) — $\frac{1}{10}$ acre — .1 - 999.9

Weight LA by $\frac{1}{2}$ if unsewered (i.e. if Cols. 62, 63 show "1.")

	Code	Unit	Range
AA. Land area adjusted for value (LAaV)	Derived	$(000)	1 - 99,999

For each Census Tract, Land Value per sq. ft. is given on separate listing, supplemental sheet (Appendix II)
Convert to value per 1/10 acre unit, so: mult. x $4,356.
Key the values with CT nos. (I)
Multiply $LA_A^{(H)}$ of each firm x value per 1/10 acre for its CT

BB. Total Vicinal Land Area $(TVLA)$	Derived	1/10 acre	.1 - 9,999.9

$TVLA = LA + EA + P_g A$
or, $(BB) = (H) + (L) + (O)$
as check
$TVLA = TLA - 10 \times E2A$
or, $(BB) = (P) - 10(M)$

CC. Total Vicinal Land area Adjusted for Value (TVLAaV)	Derived	$(000)	1 - 99,999.

Adjust area for value as in AA.
i.e. 1. Convert CT values per sf from App. II to values per 1/10 acre of $1,000 units by multiplying by $4,356.
2. Key these values with CT nos. (I)
3. Multiply TVLA (BB) of each firm by corresponding value per 1/10 acre for its CT.

II. Operations on single variables

A. Lorenz Concentration Data

1. Rank firms by size of variable (x)
2. Break into deciles
3. Report max. value each decile
4. " Σx for each decile ~~(also grand total, mean, median, σ, and σ ÷ mean.)~~
5. " each decile Σx as % of grand total Σx
6. Break top decile into tenths, report on each percentile as above:
   a. Max value
   b. Σx by percentile
   c. ~~Dec~~ Percentile Σx ÷ grand total Σx
      (Use same grand total as in #5; not decile total)
7. Report Lorenz Conc. Ratio

See Supplemental ~~Sheet 5A~~ app. III for definitions.

B. Other misc. data needed: (grand total, mean, median, st. dev., and st. dev. ÷ mean.)

C. ~~Bz~~ Variables wanted for ~~Lorenz decile analysis as in (A)~~ above analysis.

1. A, Plant Area, PA.
2. D, Depree. Plant Area, DPA
3. G, Total Assessment, TAS
4. H, Land Area, LA
5. K, Equalized Total Assessment, ETA
6. P, Total Land Area, TLA
7. Q, Employees, EMP
8. T, Need for Parking Area, NPgA
9. O, Parking Area, PgA

~~15. Land~~
14. CC, TVLaav
13. M, Ldvg Exp. Area, ELa
12. L, Expansion Area, Ea.

10. V, Truck Traffic, TT

11. Y, Unmet Land Area, USLA

D. ~~R~~. Repeat all above for tenants only. (Tenant ~~station~~ shown by "0" in Cols. 8-10, etc.)

E. ~~D~~. Other operations on single variables

1. Summarize ETa (K) by Census Tract, CT, (I).

i.e., report sum of ETa in each CT.

2. Preferred Location, PL (W)

Rank CT in order of ETa

Report sum of ETa preferred location in each CT

Rank CT ~~by~~ in order of preferred location

~~3. Ratio of Ki / above 2 i.t, above, by CT~~

~~(This involves 2 variables)~~

3. See F, 4, below

F. Operations on single ratios

1. $\frac{ETa}{TVLaav}$ by CT

Rank CT in order of $\frac{ETa}{TVLaAV}$

2. $\frac{DPa}{TVLaav}$ by CT

Rank CT in order of $\frac{DPa}{TVLaav}$

3. $\frac{EMP}{TVLa}$ by CT

Rank CT in order of $\frac{EMP}{TVLa}$

4. $\frac{\Sigma(\text{Preferred Location of ETa})}{\Sigma(\text{ETa in location})}$ by CT (From E, above)

Rank CT in order of ratio.

III. Pairings (of additive variables) Rank by sum of two variables.

A. ~~K, PCa~~ D, DPa (Common unit: $\frac{1}{10}$ acre)

~~D, DPa~~ H, La

~~Repeat~~

B. D, DPa (C.U.: $\frac{1}{10}$ acre)

P, TLa

~~C. K, ETa~~ ~~(C.U.: $(000)~~

~~cc, TVLaaV~~

~~D. TLa~~ ~~(C.U.: $\frac{1}{10}$ acre)~~

~~La~~

~~C~~ C. K, ETa (C.U.: $(000))

cc, TVLaaV

D. $P_g a$ (C.U.: ~~000 s.f.~~) $\frac{1}{10}$ acre)

$NP_g a$ ≠

E. EMP ~~÷ 5~~ ★

TLa

(C.U.: ~~000 s.f.~~ $\frac{1}{10}$ acre)

III ~~E~~ F. Repeat all pairings above for tenants only. (A and B are optional) Tenant status shown by "0", Cols. 8-10 etc.

≠ Rationale: autos are capital, like bldgs., so we are adding capl. & land. Bldgs. are msrd. in floor area; autos in NP_g area.

★ Rationale: same as above, but with labor converted to a land quota of ~~1000~~ $\frac{1}{50}$ acre per man.

IV. Pairings of non-additive variables* Rank by variable listed first.‡

A. P, TLa (C.U.: $\frac{1}{10}$ acre)
D, DPa

B. K, ETa (C.U.: $(000)
H, La × 8.8

C. K, ETa (C.U.: $(000))
P, TLa × 8.8

D. K, ETa (C.U.: $(000))
CC, TVLaav

E. P, TLa (C.U.: $\frac{1}{10}$ acre)
H, La

F. O, PgA (C.U.: $\frac{1}{10}$ acre)
T, NPgA

---

‡ Some pairs are also additive, but I want them ~~listed~~ here ranked by first variable.

		C.U.
G. BB, TVLa cc, TVLaaV	None	$\frac{1}{10}$ acre $(000)
H. P, TLa Q, EMP	None	$\frac{1}{10}$ acre ##
I. cc, TVLaaV Q, EMP	None	$(000) ##
J. K, ETa Q, EMP	None	$(000) ##
K. D, DPa Q, EMP	None	$\frac{1}{10}$ acre ##
L. K, ETa D, DPa	None	$(000) $\frac{1}{10}$ acre
M. E, LaS cc, TVLaaV		$(000)
N. Q, EMP cc, TVLaaV	None	## $(000)
O. P, TLa V, TT	None	$\frac{1}{10}$ acre ##

		C.U.
P. cc, TVLaaV V, TT	None	$\frac{\$(000)}{\#}$
Q.K+cc, ETa + TVLaaV P TLa × 8.8		$\$(000)$
R.D+P #DPa + TLa K ETa	None	$\frac{\frac{1}{10}\text{ acre}}{\$(000)}$
S. cc, TVLaaV D, DPa	None	$\frac{\$(000)}{\frac{1}{10}\text{ acre}}$
T. cc, TVLaaV K, ETa		$\$(000)$

IV. Pairings with ranking by one variable, ~~with~~ to compare two other variables.

Rank by variable listed first

A. ETa
EMP
TLa

B. ETa
TT
TLa

C. TVLaaV
EMP
TLa

D TVLaaV
TT
TLa

VI. Pairings ranked by ratio of variables.
Rank by ratio shown first

A. $\frac{TLa}{DPa}$

Report ~~holdings~~ amounts, absolute & %, by decile, of TLa, DPa, TVLaaV, ETa,

B. $\frac{TVLaaV}{ETa}$

Report ~~holdings~~ amts, absolute and %, by decile, of TVLaaV, ETa, TLa, DPa

C. $\frac{EMP}{TLa}$

Report amts, absolute & %, by decile, of EMP, TLa, PA(A), DPa,
TVLaaV, TT, ETa, NPgA

D. $\frac{PgA}{NPgA}$

Report amts, abs. & %, by decile, of PgA, NPgA.

VII. Method of Pairing

This a variation of the LCR technique.
Let x be the variable used for ranking; y the other.
Report as ff:

Decile	x: Amt	x: 90 LV‡	y: Amt	y: 90 LV	x ÷ y
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
Grand Total					
LCR					

(III) If x & y are additive, (Section II), rank by x + y.
(IV) When x & y are in a common unit, also report $\frac{x}{x+y}$ and $\frac{y}{x+y}$
(V) Rank by x, but report on y and z, as in III and IV.

---

‡ LV = Last Value in decile

Whenever more than 35% of an amt. falls in top decile, break it into percentiles for separate report, repeating same pattern.

(VI) Ranks by $\frac{x}{y}$

Report on x, y, z, and others specified.

Programmer. 3 Signif. Figures in Results — Clerk

Before any runs, check cards for some known quality problems.

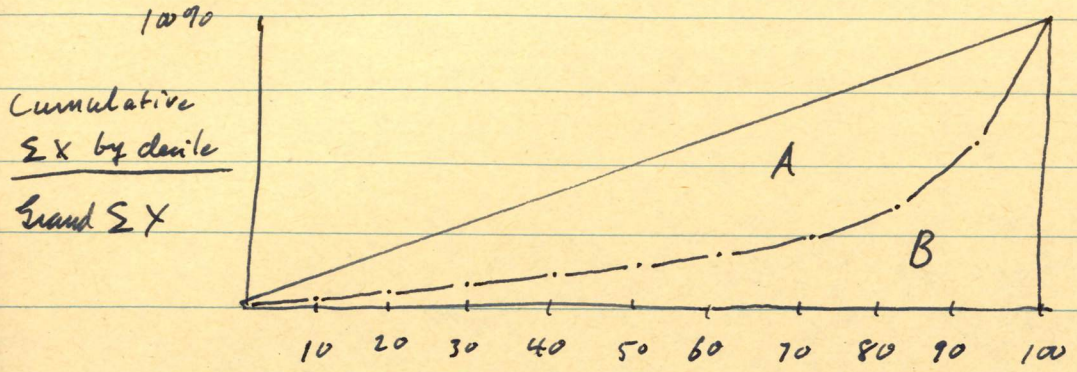
1. Card # 234 is wild, needs correcting & repunching
Proper data: Census Tract 2; Improvement Assessment $12,000; Employees, 10; Land Area, $\frac{1}{10}$.

2. Check for repeaters
529
321
261
243

There may probably be med, since data differed – it was just plant nos. That repeated. But check present deck to see if already corrected. ~~Then~~ If not, ask N. Stefaniak.

3. Plant 613 was off.
Had PA of 430,000 sf
Should be 43,000 sf or $\frac{10}{10}$ acre.
(Ruby Products, 800 E. Locust)

Appendix III: Lorenz Concentration Ratio



Variable (x) ranked by size of x, grouped by deciles

$$LCR = \frac{A}{A+B} \quad \text{by area}$$

$$A + B = \tfrac{1}{2}$$

$$\therefore \quad LCR = 2A$$

$$A = \tfrac{1}{2} - B \quad ; \quad \therefore \quad LCR = 1 - 2B$$

Let $a_1, a_2 \cdots a_{10}$ be altitudes; b be base, always .10.

$$B = \tfrac{1}{2} b a_1 + \tfrac{1}{2} b(a_1 + a_2) + \tfrac{1}{2} b(a_2 + a_3) + \cdots + \tfrac{1}{2} b(a_9 + 1)$$

$$= \tfrac{1}{2} b\left[2a_1 + 2a_2 + \cdots + 1\right] = b\left(a_1 + a_2 + \cdots + a_9 + \tfrac{1}{2}\right)$$

But $b = \frac{1}{10}$

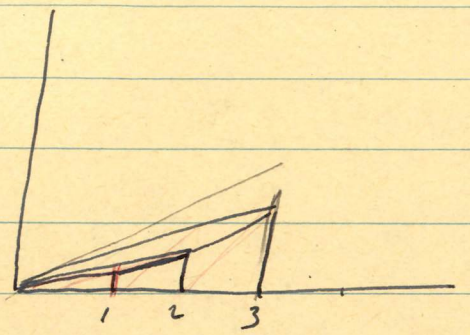
So

$$B = \tfrac{1}{10}\left(a_1 + a_2 + \cdots + a_9\right) + \tfrac{1}{20}$$

$$LCR = 1 - 2B = 1 - \tfrac{1}{5}\left(a_1 + a_2 + \cdots + a_9\right) - \tfrac{1}{10}$$

$$= .9 - \tfrac{1}{5}\left(a_1 + a_2 + \cdots + a_9\right)$$

Alternative approach ~~method~~ to find LCR. prefered by Dave Wilson



He used areas of succeeding triangles

Method 3, possibly easiest to program.

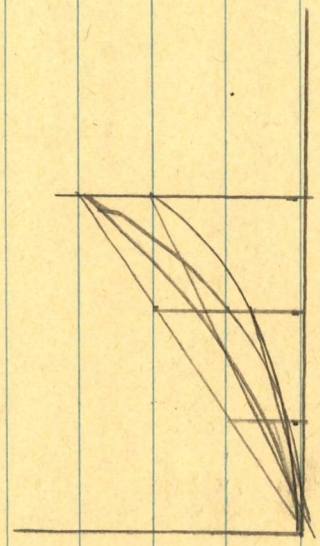
$B = \left(\sum_0^{\ell} \text{of Cum.} \sum_0^{n} x\right)$, where ℓ is limit no.

$A + B = \frac{1}{2}(\Sigma x) \times \ell$

$$A = (A+B) - B = \tfrac{1}{2}\ell \cdot \sum_0^{\ell} x - \left[\sum_0^{\ell} \text{of Cum.} \sum_0^{n} x\right]$$

$$LCR = \frac{A}{A+B} = \frac{\tfrac{1}{2}\ell \cdot \sum_0^{\ell} x - \left[\sum_0^{\ell} \text{of Cum} \sum_0^{n} x\right]}{\tfrac{1}{2}\ell \cdot \sum_0^{\ell} x}$$

$$= 1 - \frac{\sum_0^{\ell} \text{of Cum} \sum_0^{n} x}{\tfrac{1}{2}\ell \cdot \sum_0^{\ell} x}$$



Card 234

Census Tract No. 2
Block No. 11
Non-resid. area ~~786~~ 786

Plant no. 234

Annual mfgr. ~~$5 000~~ $12,000

Floor area 12,000 ft.

No. empls. 10

Land area $\frac{1}{10}$

529
321
267
243.

Plant nos. of cards that "repeated" - same plant no. w. 2 sets of data

~~4-911~~

~~Call Nash — 8-5-96~~

Correct
Land amnt = $1,000
Imps. 51,000

The error of having in one plant at 430,000 that should be 43,000 may effect some data. Checks if want any re-runs.

43,000 s.f. is correct.
430,000 s.f. is wrong

5-155119600001 RUS

‡‡JOB

‡‡XEQ ALPHA 02402
EXECUTION

Pl. Area. Land Area = .2 acre

```
12345678901234567890123456789012345678901234567890123456789012345678901234567890
        56   16132699430 70     57326 1   2          32   1567992153       11   7486
        56   14      1                82613                                      7487
```

END OF JOB

[There was a corrected card. Must have been misplaced]

-.PddddddddddPE-PP*

Non-res. area - 748
Tract No. 56
Plant No. 613

74 | 42 | 6 | 102 | 8 | 110

Ruby Products, c. 800 E. Locust

Dave W.
326

Indl. tenancy complicates picture.

Some big owner-occ. firms also hold older plants, leave them out.

E.g. K&T's old plant. Also have tents. in main bldgs.

→ Also, how does Nvb. treat multi-plant firms?

Some also hold other land: e.g. Greendale syndicate; Schlitz Tdtn.; Cudahy family holdings; etc. And of course holdings in other jurisdns., and other classes of property.

Stef. survey doesn't show who owns tent. plants.

But it shows area & value.

a. Get total area — what % of city indy. is tnt. occ?

" " value — " " " " " " No can do.

b. Get asst. value psf of total land & imp. of land. N.B.

Probably maybe it will be below o-occ, despite centrality and smaller size, bes. older bldgs. But probably above, bes. smaller parcels.

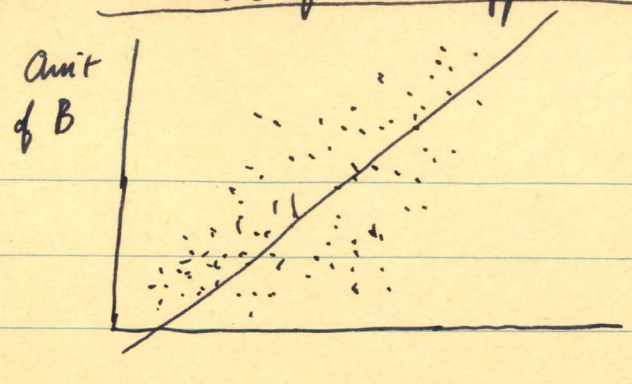
Get asst. value psf of floor. This will be below o-occ, bes. older floor, higher bldgs. N.B.

Thus show tents. do not leave much land. Land is held by owners, a reserve for spec. g. ~~So it is~~ So indl. demand for land is inflated by spec.

c. Get empl./acre — tntd. land is labor intensive.

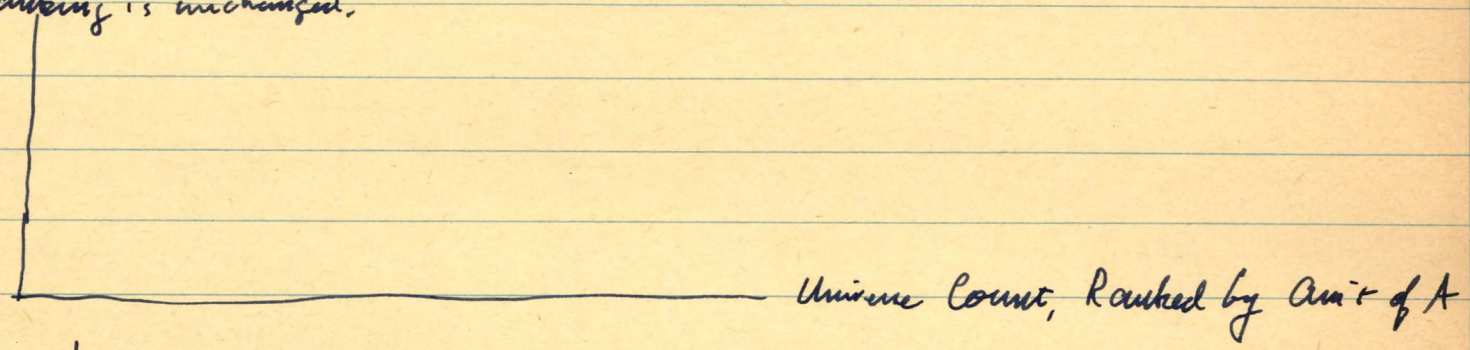
d. Net tax shift. ~~Int~~ Indl. landlords may bear more, bes. old bldgs. But they will not go up as much as reserve-holders.

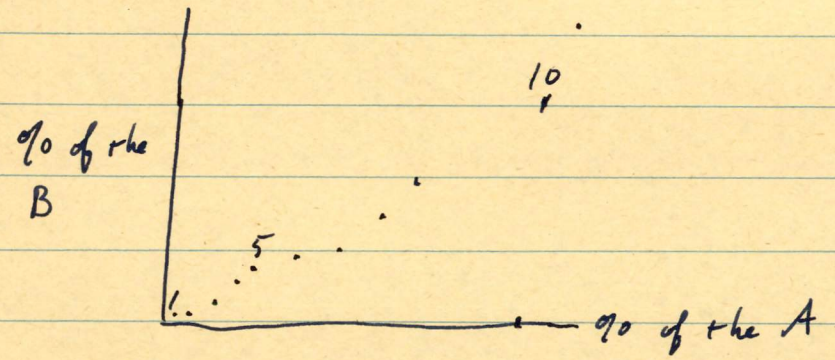
Rationale of decile approach.

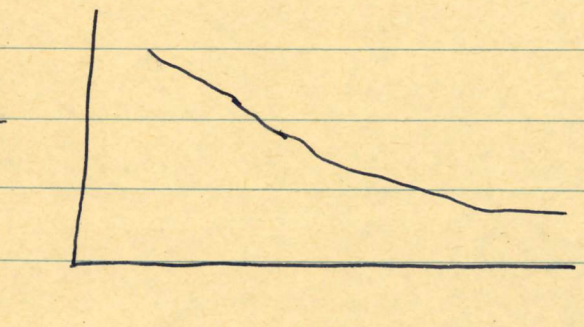


A regr. line is artificially smoothed & straightened

Instead, try Lorenz type approach by deciles
Use deciles instead of intervals of A, to equalize no. of cases used to find each point
Ranking is unchanged.

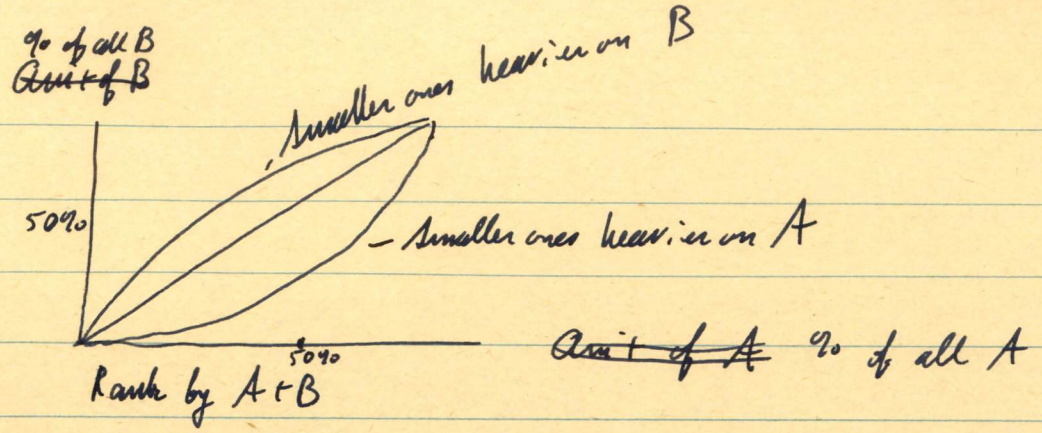






This is good approach if I get reasonably consistent curve. But if points bounce around, will have to go back to LCR, which avoids bouncing by cumulating. Or regression. Or hand-fit a curve. Or change to quintiles or larger.

I don't see how can fail to show a trend, if it is there.



This requires more explaining. If I run into trouble in ratio approach, be prepared to try it.

I. Operations on single items. (1964 Procedures)

A. Plant area (000 ft.²)

1. Mean, md., total, range, σ, $\frac{\sigma}{\bar{x}}$, measure of skewness, $\frac{\sigma^2}{\bar{x}}$,
% of area by deciles; LCR

(6) 21-23

2. all of above for
~~a. all plants~~
b. Owner-occupied
~~c. Tenants~~ (do this thrice, using 3 indications of tenant status as checks on one another:) i) Under Plant Area
ii) Under Occupancy
iii) " Land Area

3. Depreciated plant area (over)

~~B. Assessment ($000)~~

~~Same as (A, 1, and A. 2).~~

(6) 24-26 27-32

a. Land assmt
b. Improvement assmt
c. Land + Improvement assmt.

C. Land Area (.1 acre = 4356 ft.² Convert to 000 ft.², by multiplying by 4.356)

(6) 38-41

Same as (A, 1 and ~~A. 2~~)

(6) 38-41 43-44 45-46 47-50

D. Σ land areas = "Land Area" $\times 4.356$ + "Acres-Expansion" $\times 4.356$ + "Acres available for expansion at other sites" $\times 4.356$ + ["No. of employee off-street parking spaces"] $\times$ ~~400 ft²~~ .4 (000) ft.²

Same as (A, 1 and ~~A. 2~~)

E. Employees

(6) 53-57

Same as (A, 1 and ~~A. 2~~)

Looks like I goofed here in '64; bcs. acres for expansion is in whole acres, while others are $\frac{1}{10}$ acre. ∴ my '64 runs are off.

A, 3. Depreciated plant area

36-37 "Years at location" ≡ n

21-23 D.P.a. ≡ Plant area x .95^{n} to 2 signif. figures.

Same as 1, ~~for owner-occupied only~~

(6) F. Truck Traffic

71-73 Same as (A, 1 and ~~4, 5~~)

II. Ratios – attached pp. 2a & 2b

III. Regressions & correlations. (Every one to be run ~~both~~ ways: ~~1 & 2~~. a on b; ~~and~~ b on a; ~~3 & 4~~ a on $\frac{a}{b}$; b on $\frac{a}{b}$) (and 5, using method of Black's suggestion, least squares direct route to regn. line.)

A. Plant area & ~~improvements assmt~~.

.. .. land area

~~B.~~

.. .. years at location

.. .. Σ land areas

" " acres for expansion at other sites

" " Employees

" " % use autos

" " sewer

" " Truck traffic

" " Water

" " Rail siding (1 vs 2, 3, 4)

" " Fncl. rating

" " Composite credit rating

B. ~~Improvements Assmt.~~ & land area

land assmt.

employees

fncl. rating

composite credit rating

II. Ratios

A. $\frac{\text{Plant area}}{\text{Land area}}$ ≡ plant density, p.d. [Convert to 000 ft.²]

1. % of plants & % of plant area in ff. groups:

pd > 5
5 > pd > 4
4 > pd > 3
3 > pd > 2
2 > pd > 1
1 > p.d > ½
½ > pd

2. Range, mean (weighted, as a check), σ, $\frac{\sigma}{\bar{X}}$, md

3. LCR (The plant ft.² w. the least land are the "poorest", etc.)

B. $\frac{\text{Improvements Assmt}}{\text{Plant} \mid \text{Area}}$ ≡ ~~ff~~ and. value. per ft.² of plant

1. Arrange by deciles. % of plants, % of imprmts. assmt., & % of plant area in each decile. Computer shd. ~~record~~ print out boundary values of deciles, too.

2. Range, mean (weighted, as a check), σ, $\frac{\sigma}{\bar{X}}$, md.

3. LCR

C. $\frac{\text{Improvements Assmt}}{\text{Depreciated Plant Area}}$

Same as B; owner-occupied only

~~D. Plant Area / Land Area~~ (Cf. LCR of land & imprmts.)

D. $\frac{\text{Plant Area}}{\Sigma \text{ Land Areas}}$ ≡ plant - gross - density, PgD.

Same as A, 1, 2, 3

E. $\frac{\text{Employees}}{\text{Plant Area}}$ ≡ labor ~~density~~ floor density, L-FD

1. % of employees & % of plant area by deciles
2. Range, Me (weighted, as a check), σ, $\frac{\sigma}{\bar{x}}$, Md.
3. LCR

F. $\frac{\text{Imprmts. Asmnt}}{\text{Land Area}}$ (Steps more.)

Same as B, 1 & 2 & 3

G. $\frac{\text{Depreed. Plant Area}}{\text{Land Area}}$

Same as B, 1, & 2 & 3

C. Improvements Assmt / Plant Area and years at location

and rail siding (1 vs. 2, 3, & 4.)

D. Land area and land assmt
yrs. at location
employees
% use autos
truck traffic
water
rail siding (1 vs. 2, 3, & 4.)
fuel. rating
comp. credit rating

E. Items at location and employees

% use custos

truck traffic

water

rail siding

fuel. rating

comp. credit rating

7. Σ land areas and employees

fuel. rating

comp. credit rating