Effect of Personal Visits on Response Rates to Mail Surveys

By Cecil C. Smith

Achieving a high rate of response to a mail survey is important from several viewpoints. Aside from any statistical considerations, it is obviously a waste of time and money to address envelopes and to mail questionnaires to people who do not return them. The statistical aspects of the problem are even more serious. A low return usually means that the reported data are not representative of the universe the investigator is trying to sample. The various geographic areas may not be covered in their proper proportions and the few people who do respond from each locality may differ considerably from the average in that locality with respect to the item being estimated. Such differences tend to be predominantly in one direction. They lead to serious biases for which satisfactory adjustments cannot always be made. Even when the bias is kept under control by interviewing samples of nonrespondents, investigators find that it pays to have a high response to the mailed inquiry. The higher the response, the smaller the sample of interviewed nonrespondents needs to be to attain the desired level of statistical precision. Rates of response can be stimulated by several different devices. This paper considers one of these—the effects of personal visits to the individuals on a mailing list.

STATISTICIANS who conduct mail surveys know that individuals who receive mail questionnaires are more likely to fill them out and return them when they have had some previous personal contacts with representatives of the agency that sends the questionnaires. There has been some question as to the length of time the stimulating effect of a visit to a potential respondent will be maintained when he is asked thereafter to return questionnaires regularly at periodic intervals. This is the situation with the general crop reporters who are asked to report crop conditions and miscellaneous other agricultural data once a month. After they are recruited and show some interest in reporting, they are requested to continue reporting as long as their interest in the work justifies their being kept on the mailing list.

Many reporters report regularly over long periods. Others report only occasionally unless some action is taken from time to time to stimulate their interest. With the limited resources available in many State Statisticians' offices it is difficult to maintain close personal contacts with a large number of reporters in a State. It is therefore well to know how much improvement from such contacts can be expected in the reporters' performance and how long the effects of a single visit will hold up.

In May 1949 the State Statistician of Idaho started a test on the 74 general crop reporters who were carried on the mailing list for Ada County, Idaho. By taking every other name on the list half of the number was selected for personal visits. The visitor was able to call on 13 reporters a day. These visits made in May 1949, were the only ones made.

When the selected reporters were visited most of them appeared to be interested in discussing crop reporting and related activities. One reporter remarked that he would be more inclined to report regularly if he were personally acquainted with the people to whom he was reporting.

A number of the reporters were obviously not much interested in crop reporting. Judging from their general attitudes and the nature of their farming operations they would not be likely to become good regular reporters. Uncovering such dead wood is itself a useful byproduct of personal visits to the people on a mailing list. Ordinarily such individuals would be culled from the list, but to make this a fair test they were allowed to remain.

Before any reporters were visited, 28 percent of those on the half of the list selected for visiting returned their completed May crop-report schedules. Of those on the other half of the list, 26 percent returned completed schedules. In other words, before the test was started the two groups of reporters were about equally good. But immediately after the May visits the percentage return jumped to 77 percent for the visited group, whereas the return for the unvisited group was only 29 percent.

This is a rather spectacular demonstration of the effectiveness of personal visits, particularly when it is remembered that some visited reporters were hopeless cases who would ordinarily have been dropped from the mailing list.

But it is more important to look at the degree to which this stimulating effect holds up. The chart shows the percentage return for each group of reporters for 21 months after the visits were made. It is easy to see that after almost 2 years the percentage return for the visited group averages in the neighborhood of 50 percent while the percentage return for the unvisited group averages in the neighborhood of 25 percent, or about the same as at the start of the test. The trend shown for the visited group indicates that this difference would remain about the same for quite a while longer.

Normal list mortality prevented the test from being continued any longer. Although 28 of the visited reporters were still active, the unvisited group had dwindled to 19. This indicates another benefit from personal contacts. In addition to stimulating returns, visits appear to lower the mortality rate on lists.

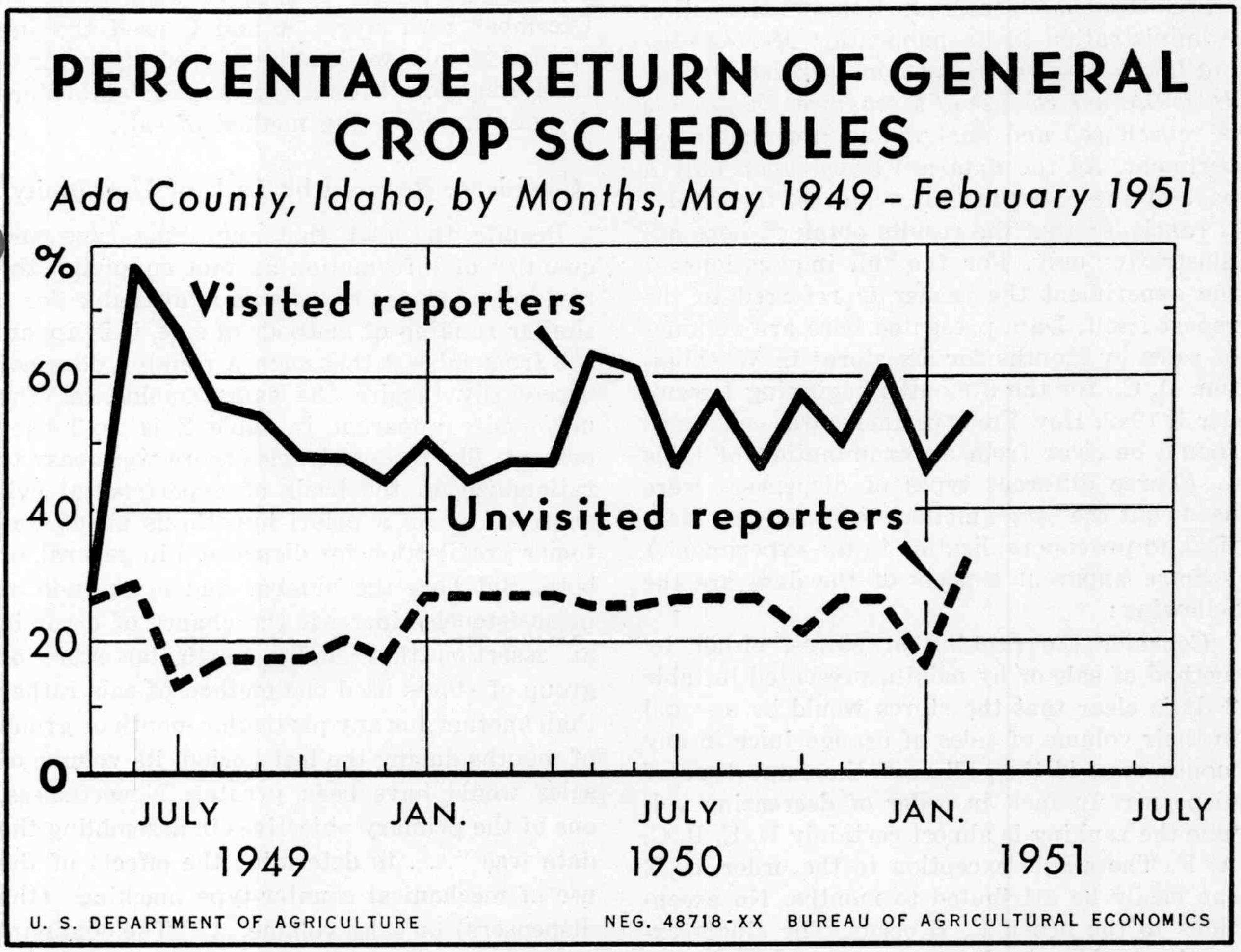

Figure 1

An Experiment in Marketing

By Glenn L. Burrows

The procedure of rating alternative retailing practices by totaling sales over a number of test stores and over a number of sales reporting periods implies certain tacit assumptions that are not always valid. In this article, several important requirements for an analysis of variance of retail-store data are discussed and shown to be not unusually restrictive. Variance analysis of sales data obtained from a latin square experiment is illustrated. The results are similar to, but more meaningful than, those obtained from analyses lacking clearly stated assumptions.

THIS PAPER deals with some of the uses and limitations of the latin square design with special reference to its use in the study of retail merchandising.

Presentation of Data

By way of illustration, part of the data reported by the Marketing and Facilities Research Branch of the Production and Marketing Administration in its publication *Merchandising Reconstituted Frozen Concentrated Orange Juice through the use of Mechanical Dispensers* is reproduced and analyzed as a complete experiment. As the data are based upon only a part of the experimental evidence, the reader is cautioned that the results obtained here are illustrative only. For the full implications of the experiment the reader is referred to the report itself. Data presented here are volumes of sales by months for six stores in Washington, D. C., for the 6 months beginning December 1, 1949. How the experiment was conducted should be clear from an examination of table 1. (Three different types of dispensers were used; but the "jug" methods of sale were identical, to provide replication in the experiment.)

Some apparent aspects of the data are the following:

Consider the ranking of stores, either by method of sale or by month, presented in table 2. It is clear that the stores would be unequal in their volume of sales of orange juice in any month even if they all used the same type of dispenser. In fact, in order of decreasing volume the ranking is almost certainly D, E, B, C, A, F. The single exception to the order D, E can easily be attributed to months. No exceptions to the order E, B occur. The single exception to the order B, C can easily be attributed to months. One of the two exceptions to the order C, A among stores ranked by method of sale (see column for dispenser 3) is easily attributable to months; of the three exceptions when ranked by months, that for February is inconsequential and that for April is attributable to the method of dispensing. The two exceptions to F occurring in last place are both attributable to the method of dispensing. In December both stores A and C used the jug method of sale, while store F used dispenser 3. In March, store F used dispenser 1, while store A again used the jug method of sale.

Confidence Reduced by Lack of Uniformity

Despite the fact that the same type and quantity of information as that supplying the ranking of stores by volume is available for a similar ranking of methods of sale, it is apparent from table 3 that such a ranking does not necessarily inspire the same confidence. The uniformity apparent in table 2 is no longer present. The inconsistencies there were easy to rationalize on the basis of experimental evidence or on an a priori hypothesis about customer predilection for dispensers in general, or both. But here the number and magnitude of inconsistencies increase the chance of error in an assertion that, had a particular store or group of stores used one method of sale rather than another for any particular month or group of months during the test period, its volume of sales would have been greater. Nevertheless, one of the primary objectives in assembling the data was ". . . to determine the effects of the use of mechanical counter-type machines (the dispensers) on sales volume. . . ." The dominant influence of store volume was anticipated; the

TABLE 1.—*Sales of reconstituted frozen concentrated orange juice from jug and mechanical dispensers in 6 test stores, Washington, D. C. by months, December 1, 1949, to May 31, 1950*

Method of sale (Juice container)	December		January		February		March	
	Sales	Store identification	Sales	Store identification	Sales	Store identification	Sales	Store identification
	Gallons		*Gallons*		*Gallons*		*Gallons*	
Jug 1	46.0	A	128.3	D	64.0	B	95.0	E
Jug 2	76.0	B	43.5	F	86.0	E	47.5	A
Jug 3	35.8	C	58.5	A	28.0	F	83.0	B
Dispenser 1	155.2	D	66.9	C	66.7	A	66.6	F
Dispenser 2	123.0	E	79.0	B	106.5	D	75.5	C
Dispenser 3	64.0	F	120.4	E	63.5	C	167.8	D
Total	500.0		496.6		414.7		535.4	

Method of sale (Juice container)	April		May		Total		
						By stores	
	Sales	Store identification	Sales	Store identification	Sales by containers	Store identification	Sales
	Gallons		*Gallons*		*Gallons*		*Gallons*
Jug 1	65.0	C	38.6	F	436.9	A	356.7
Jug 2	185.8	D	71.5	C	510.3	B	513.0
Jug 3	108.0	E	171.0	D	484.3	C	378.2
Dispenser 1	104.5	B	134.5	E	594.4	D	914.6
Dispenser 2	43.5	F	61.5	A	489.0	E	666.9
Dispenser 3	76.5	A	106.5	B	598.7	F	284.2
Total	583.3		583.6		3,113.6		

design incorporates special means of removing such influence and of assessing the reliability of the comparisons among the monthly totals. For example, in the first three rows of table 1 every month is linked to every other month by at least one identical store employing the same method of sale.

What is it that guides one to the choice of low- and high-volume months or of low- and high-volume methods of sale? Do the data conform to a preconceived hypothesis, or do they suggest hypotheses about low- and high-volume months and low- and high-volume methods of sale? Certainly, in the absence of any hypothesis, they are useless. Furthermore, the process of totaling—leading, as it does, to averaging—by months, by method of sale, or by stores, is not only confusing but also misleading unless some rather well-defined hypotheses are kept in mind, particularly in view of the numerous inconsistencies. When do totals (averages) and differences among them have real meaning?

Nature of the Analysis

It is clear that certain differences among stores have already been interpreted as real, and the only question to which this author seeks the answer is: Are not similar conclusions about months and methods of sale also possible? From the way in which this experiment was conducted, and in view of the stated objectives it is fairly obvious that some investigators will answer yes. Certainly some investigators will answer no, and the author conjectures that among them are those who will find little fault with his conclusion concerning store volumes or with its justification. Nevertheless, despite the number of inconsistencies and perhaps the lesser magnitudes among differences between sales for different methods of sale and between different months, the same methods are available for comparisons as were available for comparisons between store volumes of sale.

Total sales for the three dispensers were 1,682.1 gallons as compared with a total of 1,431.5 gallons for the jugs; also dispenser 2, which is the only dispenser for which total sales did not exceed total sales of each of the three jugs, was used by the three largest-volume stores (as established earlier) in the three poorest months (ranked on the basis of monthly total sales). As an excuse for the low volume of sales for dispenser 2, however, such an argument is defenseless unless it is established that any increase in sales with the dispenser over those with jugs is larger for large-volume stores than for small-volume stores. This question could be asked regarding all the dispensers. An affirmative answer would invalidate the ordinary analysis of variance for this experiment. Rather, it would suggest the desirability of investigating constant ratios. For this reason, among others, readers are warned not to be misguided in the use of the latin square design by overenthusiastic advocates, who do not always state clearly the assumptions necessary to the proper interpretation of its analysis.

These assumptions are not necessarily formidable. In fact, they are no more restrictive than those required, although not so frequently stated, in less formalized analyses. An attempt is made to state simply the important assumptions and to illustrate how, through formalized analysis, they lead to useful conclusions.

It is assumed that the data are free from gross reporting errors. From the standpoint of the marketing researcher two requirements, called here constancy and equal variability, are particularly important, because they are frequently not met and because they affect the practical decisions to be made as a result of the analysis.

Constancy in this latin square means that volume of sales observed for a particular store in a given month depends upon the effect of the particular month, the same for every observation occurring in that month regardless of store or method of sale; upon the effect of the par-

TABLE 2.—*Stores ranked by volume of sales*

Volume of sales	Method of sale						Month					
	Jug			Dispenser								
	1	2	3	1	2	3	Dec.	Jan.	Feb.	Mar.	Apr.	May
Rank												
1	D	D	D	D	E	D	D	D	D	D	D	D
2	E	E	E	E	D	E	E	E	E	E	E	E
3	C	B	B	B	B	B	B	B	A	B	B	B
4	B	C	A	C	C	A	F	C	B	C	A	C
5	A	A	C	A	A	C	A	A	C	F	C	A
6	F	F	F	F	F	F	C	F	F	A	F	F

TABLE 3.—*Methods of sale ranked by volume of sales*

Volume of sales	Store						Month					
	A	B	C	D	E	F	Dec.	Jan.	Feb.	Mar.	Apr.	May
Rank												
1	D 3	D 3	D 2	J 2	D 1	D 1	D 1	J 1	D 2	D 3	J 2	J 3
2	D 1	D 1	J 2	D 3	D 2	D 3	D 2	D 3	J 2	J 1	J 3	D 1
3	D 2	J 3	D 1	J 3	D 3	[1] D 2, J 2	J 2	D 2	D 1	J 3	D 1	D 3
4	J 3	D 2	J 1	D 1	J 3	D 2, J 2	D 3	D 1	J 1	D 2	D 3	J 2
5	J 2	J 2	D 3	J 1	J 1	J 1	J 1	J 3	D 3	D 1	J 1	D 2
6	J 1	J 1	J 3	D 2	J 2	J 3	J 3	J 2	J 3	J 2	D 2	J 1

[1] Dispenser 2 and Jug 2 sold equal amounts in store F.

ticular method of sale, the same for any month and in any store in which the method is used; and upon the effect of the particular store.

The store effect has already been demonstrated to exist independently of the method of sale employed or the month in which it is observed. The combination of these separate effects constitutes all accountable reasons for the differences among the observations. It is possible to conceive of other effects that contribute to differences among the observed results; but formal analysis of the latin square provides for testing for the existence and estimation of the magnitude of only such effects as are postulated above.

The assumption of equal variability is necessary only if tests for the existence of effects are to be made in order to decide objectively such questions as the following: Is dispenser 3 really superior to dispenser 1? Is dispenser 2 really no better than the jug method of sale? Even in the absence of the assumption of constancy, certain comparisons of sale volumes, such as those for stores C and D in April with their respective values in May, measure only the effect of months. Surely the variability among such differences is some indication of the precision of measurement of the month effect. For store C the difference is 6.5 gallons and for store D it is –14.8 gallons.

Is it not necessary for the absolute difference (that is, disregarding sign) between volumes for two different methods of sale used by any store in these 2 months to exceed 14.8 gallons before a real effect can be established between methods of sale? It would certainly seem so if the differences for every store are measured with the same precision. This is precisely the reason why the assumption of equal variability is made. The measure of error inherent in such differences is a guide to the proper significance to be attached to differences observed among the effects of months. Similarly, such differences as those between volume of sales for stores A and B in December and in March should measure the effect of stores.

These two assumptions can be written succinctly as follows:

$$y_{ijk} = f(M, s_i, m_j, d_k, e_{ijk})$$

where: y_{ijk} = volume of sales (or a transformation thereof) observed for the ith store (i = A,B, . . ., F) in the jth month (j = Dec., Jan., Feb. . . . May) using the kth method of dispensing juice (k = jug 1, jug 2, jug 3, Dispenser 1, . . . Dispenser 3)

f indicates a relationship with (dependency upon or function of) the quantities enclosed in its parentheses.

M = mean (average) sales for all stores over the test.

s_i = adjustment to M for the ith store because of its size.

m_j = adjustment to M for all stores in the jth month.

d_k = adjustment to M for all stores using the kth type of dispenser.

e_{ijk} = discrepance between observed sales (or its transform) and the combined postulated effects.

Note that the s_i is the same for store i in any month and when using any dispenser; similar statements hold for m_j and d_k. Also, it has been assumed that the variability among differences between two e's in the same row (that is, having same i = same store) is neither more nor less than that between two e's in the same column (that is, having same j = same month) and the same for that between e's with the same k (that is, for same dispenser). Now if sales (or some simple transformation thereof) can be shown to follow a simple additive relation (that is, $z_{ijk} = M + s_i + m_j + d_k + e_{ijk}$), where z_{ijk} is either y_{ijk} or some simple transform thereof, and the assumptions are correct, differences among totals for months, for stores, and for dispensers are efficient estimates of differences among the m_j's, s_i's, and d_k's, respectively. This is quite clear, for it is obvious that in differencing two such totals all effects subtract out except the particular one in question and some terms measuring discrepance from hypothesis; these are compared with differences that measure solely discrepance from hypothesis.

It is only under such assumptions that a latin square arrangement – and many other useful designs – eliminates the effects of nontest variables (store size, seasonal variation) from the comparison of the effects of the use of different dispensers upon volume of sales. The assumption of equal effects on stores of all volumes for different months and different dis-

TABLE 4.—*Expected sales of reconstituted frozen concentrated orange juice from jug and mechanical dispensers in 6 test stores, Washington, D. C., by months, December 1, 1949, to May 31, 1950*

Method of sale (Juice container)	December		January		February		March	
	Sales	Store identification	Sales	Store identification	Sales	Store identification	Sales	Store identification
	Gallons		*Gallons*		*Gallons*		*Gallons*	
Jug 1	42.6	A	135.0	D	54.5	B	100.2	E
Jug 2	80.9	B	42.2	F	92.4	E	60.8	A
Jug 3	54.1	C	50.0	A	24.2	F	82.5	B
Dispenser 1	161.9	D	71.9	C	54.7	A	62.7	F
Dispenser 2	103.0	E	76.8	B	130.1	D	60.8	C
Dispenser 3	57.5	F	120.7	E	59.0	C	168.5	D
Total	500.0		496.6		414.9		535.5	

Method of sale (Juice container)	April		May		Total		
	Sales	Store identification	Sales	Store identification	Sales by containers	By stores: Store identification	By stores: Sales
	Gallons		*Gallons*		*Gallons*		*Gallons*
Jug 1	60.1	C	44.5	F	436.9	A	356.8
Jug 2	161.7	D	72.4	C	510.4	B	513.1
Jug 3	116.1	E	157.5	D	484.4	C	378.3
Dispenser 1	108.8	B	134.5	E	594.5	D	914.7
Dispenser 2	53.1	F	65.2	A	489.0	E	666.9
Dispenser 3	83.5	A	109.6	B	598.8	F	284.2
Total	583.3		583.7		3,114.0		

pensers may not be appropriate. Perhaps the assumption of a constant percentage increase for all stores is more suitable. Thus, a change from the jug method of sale to a dispenser might be expected to effect a greater absolute increase in volume of sales for store D than for store F, even though the percentage increases are equal. A simple transformation to logarithms of the original observations renders the data suitable to the previous assumptions if constant rates can be assumed for stores, months, and methods of sale.

Adequacy of the Analysis

It may be instructive to examine the adequacy of the proposed hypotheses to represent the observed data. For such purposes, the sales to be expected on the basis of the hypotheses are presented in table 4, using the volumes of sales without transformation. Except for rounding errors, the totals for months, for stores, and for methods of sale are identical with those of table 1. The sensitivity of the hypotheses is apparent in reflecting the differences among the observed data. Table 5 shows the difference between the reported and expected sales.

To determine whether large errors are associated with large volumes of sales a number of measures might be thought to be useful. A simple correlation of the discrepancies in table 5 with the reported sales in table 1 would be expected to be positive, under the hypothesis of constant percentage increases for stores, months, and methods of sale. Its value is .24. But, in applying the usual test of the correlation coefficient to decide whether this value differs significantly from zero, one difficulty arises from the fact that the discrepancies in table 5 are not statistically independent. Thus the test—a one-tailed test is appropriate—with n = N-2 = 36-2 = 34 would indicate correlations significantly greater than zero too frequently; in this instance a nonsignificant result is indicated. Readers who are familiar with the concept of degrees of freedom might propose to test this correlation coefficient at n = 20 degrees of freedom—the reason for the 20 degrees of freedom appears later. But even though this would yield fewer significant correlations, such an approximate test is still far from adequate.

A simple correlation of the discrepancies in table 5 with the expected sales in table 4 is known to be zero by virtue of the hypothesis. It is therefore of no use as a test measure of degree of association. A simple correlation of the discrepancies in table 5 with a measure of store volume of sales will provide no criterion for rejecting the hypothesis of equal store effects, for it too must be zero, no matter what measures are used for store size.

Another proposal might be a simple correlation of the absolute values of the discrepancies in table 5 with either reported sales, as in the first measure considered, or with some measure of store size, as in the third measure considered. The latter was computed, using inverted ranks as an indication of store size, and found to be .0015. The difficulty of interpreting the significance of this measure would usually be as great or greater than that encountered with the first measure discussed. But it would hardly be necessary to investigate the significance of so small a correlation coefficient.

One simple test would be appropriate and adequate in one special instance, although it is an over-all test of the adequacy of the combined hypotheses. The chi-square test would require that one size of serving glass be used in all stores with all methods of sale during all test periods. Except for a constant multiplier, depending upon the size of glass used, it involves squaring the discrepances in table 5, dividing by the corresponding entry in table 4, and summing over-all entries.

Unfortunately, this test would not generally be applicable, nor is it in the present study. Here the same choice of two sizes of glasses was offered in every store, and the price for each choice was the same for all stores. If such a test were to indicate that the data departed significantly from the hypothesis, it would doubtless be desirable, and at first glance possible, to separate chi square into portions attributable to various possible causes as, for example, the contribution to the measure of departure made by large and small stores as opposed to those of average volume. In the present state of statistical knowledge this is too much to require, for the same lack of independence among the discrepances prevents a separation of chi square into such meaningful portions. It is possible to decompose it, but only into components whose practical significance is difficult to interpret.

From the foregoing discussion it is seen that a number of frequently used techniques must be interpreted with caution. They are mentioned

TABLE 5.—*Differences between reported and expected sales of reconstituted frozen concentrated orange juice from jug and mechanical dispensers in six test stores, Washington, D. C., by months, December 1, 1949, to May 31, 1950*

Method of sale (Juice container)	December		January		February		March		April		May	
	Sales	Store identification	Sales	Store identification	Sales	Store identification	Sales	Store identification	Sales	Store identification	Sales	Store identification
	Gallons		*Gallons*		*Gallons*		*Gallons*		*Gallons*		*Gallons*	
Jug 1	3.4	A	− 6.7	D	9.5	B	− 5.2	E	4.9	C	− 5.9	F
Jug 2	− 4.9	B	1.3	F	− 6.4	E	−13.3	A	24.1	D	− 9.9	C
Jug 3	−18.3	C	8.5	A	3.8	F	0.5	B	− 8.1	E	13.5	D
Dispenser 1	− 6.7	D	− 5.0	C	12.0	A	3.9	F	− 4.3	B	0.0	E
Dispenser 2	20.0	E	2.2	B	−23.6	D	14.7	C	− 9.6	F	− 3.7	A
Dispenser 3	6.5	F	− 0.3	E	4.5	C	− 0.7	D	− 7.0	A	− 3.1	B

here only to illustrate the dangers in applying inefficient, and even inappropriate and uninterpretable, techniques.

How to Perform the Analysis

Regardless of the size or sizes of glasses employed in dispensing the juice—as long as the same sizes are used with all dispensers—an analysis of variance is possible for deciding the question of whether some methods of sale really sell more juice than do others. Such analyses under the two different sets of hypotheses, together with some comments on their interpretation, follow.

TABLE 6.—*An analysis of variance for stores, months, and methods of sale, by constant-additive and by constant-percentage effects*

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio to error of mean square
	Additive			
Stores	5	46,617.79	9,323.56	55.2
Months	5	3,386.31	677.26	4.0
Methods of sale	5	3,492.90	698.58	4.1
Error	20	3,380.68	169.03	
	Percentage			
Stores	5	1.1124	0.2225	42.9
Months	5	.0792	.0158	3.1
Methods of sale	5	.1271	.0254	4.9
Error	20	.1037	.0052	

In the analysis of variance in table 6, the total variability exhibited by the data in table 1, that is, the sum of squares of the differences between each entry and the mean of all entries, has been split up into four portions—stores about the mean, months about the mean, methods of sale about the mean, and a portion designated as error. This error sum of squares is nothing more than the sum of the squares of the discrepancies in table 5. Except for rounding errors, totals for stores, months, and methods of sale in table 5 are all zero.

Comparable to the chi-square test discussed above, one might suggest the ratio of the error sum of squares to total sum of squares as a measure of adequacy of the hypothesis. For the additive hypothesis this ratio is .0594 and for the percentage hypothesis .0739. Thus the hypothesis of constant additive effects accounts for about 94 percent of the total variability among the data in table 1, and the hypothesis of constant percentage effects accounts for about 93 percent of the variability among the logarithms of the data in the same table.

Incidentally, one often finds in marketing literature objections to the use of correlation and regression techniques, yet the same authors claim advantages for the analysis of variance; the test of the measure of adequacy just described, however, is precisely the test of the multiple correlation of sales with store, month, and dispenser effects. The use of the latin square and the analysis of variance is only a different formal presentation of the results of a regression or correlation analysis; the tests available with one method are available with the other and are equivalent.

The ratios of the mean squares for the separate effects to that for error are the test criteria for establishing the significance of those effects. The large ratio for stores as compared to that for dispensers merely reflects what was observed earlier by the method of ranking. Despite the predominance of variability among stores, opportunity is afforded here for some definite conclusions with respect to dispensers. In fact, the F-test indicates highly significant differences among the effects of dispensers on sales under either set of assumptions. Numerous exact or approximate tests are available for testing for real differences between specific stores or groups of stores and between different methods or grouped methods of sale.

Table 7 shows the ranked mean sales for the various methods of sale. Two different methods of averaging produce two different rankings, a fact that should caution against any hasty adoption of either ranking to the exclusion of the other, as neither method has any unusual claim to superiority.

An approximate method by Tukey for comparing means was unsuccessful in either case in establishing any superior or inferior methods or groups of methods; a method due to Duncan indicated that both dispenser 3 and dispenser 1 were superior to jug 1, jug 3, and dispenser 2.

TABLE 7.—*Method of sales of orange juice ranked by arithmetic mean and by geometric mean monthly sales*

Arithmetic mean		Geometric mean	
Method of sale	Mean monthly sales	Method of sale	Mean monthly sales
	Gallons		*Gallons*
Dispenser 3	99.8	Dispenser 3	93.5
Dispenser 1	99.1	Dispenser 1	93.0
Jug 2	85.1	Dispenser 2	77.0
Dispenser 2	81.5	Jug 2	75.1
Jug 3	80.7	Jug 1	66.9
Jug 1	72.8	Jug 3	66.9

The mean sales for jug 2 did not differ significantly from any of the other means. But the exact student test for comparing the mean for dispensers with the mean of jug sales indicates in both cases that, on the average, dispensers have a highly significant superiority over the jug method of sale. An approximate test suggested by the data indicate further that dispensers 1 and 3 are superior to dispenser 2. As might be suspected from the data, no test indicated definitely a superiority of dispenser 3 over dispenser 1.

Conclusion

It has been shown that two different sets of assumptions lead to approximately the same conclusions; the only valid measures of adequacy of the hypotheses were about the same for both sets. Many researchers suppose that some statistical test should exist to aid in the selection of one or the other set. But in both instances, the conclusions are derived from the assumptions, and they clearly cannot serve in addition as a basis for determining which set is the more acceptable.

Book Reviews

Land for Tomorrow. By L. DUDLEY STAMP. American Geographical Society, New York. Indiana University Press, Bloomington, and 230 pages. 1951. $4.00.

THIS IS ANOTHER BOOK about the world food problem (with a few comments on minerals and a plug for mapping present land-use) by a professor of geography at the London School of Economics. It is based on a series of lectures given by the author at Indiana University in 1950.

Progress toward the solution of this problem is such a stern and urgent challenge to Western Civilization that any book that may encourage people to study it is welcome. This one will be useful to those who have not yet studied the problem, provided they follow it up with specific discussions of the factors involved and of constructive suggestions about how to develop the potentialities of land and people.

Dr Stamp's intentions are obviously in the public interest as viewed by an English professor who has travelled widely (and well) and who wants to be helpful. The book has many good paragraphs about ideas that need frequent restatement and repetition. Yet these are diluted with others that are too general or too platitudinous to be rewarding.

The point emphasized on the dust jacket is surprising, as its author suggests, ". . . that the most important undeveloped lands are not in the Tropics and uninhabited latitudes, as we might expect; the hope for land development lies in the middle latitudes, in the United States, the Soviet Union, Canada, Australia, and the Argentine . . ." Indeed surprising! Perhaps enough to make the book sell. Fortunately, however, this is not what the author says. He writes, ". . . that there are greater immediate prospects of increasing agriculture output in the middle latitudes than there is of securing immediate help in the world food