ABSTRACT

At the 1979 SUGI meeting, Bill Gjertsen presented a series of SAS macros for United States map reports. We have found this to be a very useful piece of software for research done by our admissions office. We have added a feature of our own in which the map is shaded to allow quicker, more visual interpretation. The intensity of shading can correspond to the actual numbers printed, to per capita ratios, or to ranks.

INTRODUCTION

The student population of Washington University tends to be a non-local one, with some representatives from every one of the fifty states. The broad geographical diversity of applicants can be a challenge because interests and attitudes tend to vary from one region to another. For some years our admissions office has recognized this fact and has cross tabulated information by regions of the United States. In connection with this effort, we have found the United States map macro [1] presented at last year's SUGI meeting to be very helpful because it gives us a way of organizing data in more detailed form than regional, yet keeping it digestible. We have, in fact, added to the impact of the map report by allowing five shades of gray scale that can be directly or indirectly related to the numbers printed on the map. Outlines of the states are kept and printed on top of the shading. In case of missing data, the state is left blank except for the name.

METHODOLOGY

The first step is to label all points of the printer grid (52 rows by 132 columns) with state codes. To accomplish this, we divided the grid into two regions, west and east, and keyed alphanumeric into card image records for each region separately. For example, in the western region, A represents Alaska, B represents Washington, C represents Idaho, and so on. These codes can be seen in Figures 1 and 2. We then create two SAS data sets, named EASTCODE and WESTCODE, containing numeric codes corresponding to the order of the state postal abbreviations. Two user-defined formats, $EASTUS, and $WESTUS, allow us to convert the alphanumeric into the numeric codes, ranging from 1 to 52 (to allow for the fifty states, the District of Columbia, and the region surrounding the country. See the upper portion of Figure 3 for details.

The next step is to create corresponding SAS data sets containing the state outlines and the postal codes. These are obtained by reading the outlines of the states shown in Figures 4 and 5 into the SAS data sets WESTBORD and EASTBORD. Following these operations, the four data sets are combined into a single one called MAPINFO.

At this stage, we are ready to incorporate numeric information into the process. In the example shown in Figure 3, all we have done is read data from cards, then classify it into five groups. In a practical application, this data would probably come from the FREQ, MEANS, or SUMMARY procedures followed by the RANK procedure, but we have left out such details so the reader can concentrate on the process of making the map.

In the final step, we use array notation together with the OVERPRINT feature of the PUT statement to convert the values 1 through 5 associated with the states to five levels of gray scale. The process starts with the values, which are stored in the variables R1 through R52. The index variable for the array R defined on R1-R52 is STATEKEY, and its values are varied as we sweep across each line of the map. The value of R corresponding to STATEKEY is stored in the variable called RANK, which in its turn serves as index variable for the arrays C1, C2, and C3 which contain the gray scale characters. The gray scale characters are then placed into an array called SHADE, and from there are dumped into the PRINT file. At the same time, in a rather inelegant fashion, we examine each line number to see if the particular line should contain any numeric data. If so, we overprint that as well.
The results can be seen in Figure 6. The particular data used was a ranking of Washington University on the idealness of distance from home. The map shows, to a large degree, that the greater the distance, the less the prospective students consider it ideal (contrary to one theory that claims there is such a thing as being "too close to home"). The north and middle eastern states evidently show much more sensitivity to distance than do the wide open western states. The response from Mississippi is quite surprising. The only state from which we had no information is North Dakota, which is blank except for the ND.

Figure 1. Alphanumeric codes for the western region.
CONCLUSIONS

We have found this program to be a valuable tool, with a relatively moderate cost of about one dollar per map. Considerable effort was spent keying the information in Figures 1 and 2, but we will be happy to save others that effort by making available card decks of Figures 1 through 5 to anyone requesting. We would appreciate, however, if those requesting a copy would please send a mailer suitable for return mail containing approximately 300 cards plus return postage.

REFERENCE


Figure 2. Alphanumeric codes for the eastern region.
OPTIONS NODATE MISSING='.';
PROC FORMAT;
VALUE $WESTUS A=1 B=48 C=14 D=21 £=30 F=24 G=38 H=S1 I=42 J=5 K=34 L=45 M=6 N=28 O=13 P=21 Q=39 R=45 S=33 T=44 U=37 V=3 W=19 X=12 Y=52;
DATA WESTCODE; LENGTH ALPHA1-ALPHA75 $ CODE1-CODE75 8;
ARRAY ALPHA(I) ALPHA1-ALPHA75; ARRAY CODE(I) CODE1-CODE75;
INPUT ALPHA1-ALPHA15 (75*$1.);
DO 1=1 TO 75; CODE = PUT(ALPHA,$WESTUS.); END;
KEEP CODE1-CODE75;
CARDS;
**** PLACE DATA FROM FIGURE 1 HERE ****
DATA EASTCODE; LENGTH ALPHA76-ALPHA132 $ CODE76-CODE132 8;
ARRAY ALPHA(I) ALPHA76-ALPHA132; ARRAY CODE(I) CODE76-CODE132;
INPUT (ALPHA76-ALPHA132) (51*$1.);
DO 1=1 TO 51; CODE = PUT(ALPHA,$EASTUS.); END;
KEEP CODE76-CODE132;
CARDS;
**** PLACE DATA FROM FIGURE 2 HERE ****
DATA WESTBORD; INPUT WESTLINE $CHAR75.; CARDS;
**** PLACE DATA FROM FIGURE 4 HERE ****
DATA EASTBORD; INPUT EASTLINE $CHAR57.; CARDS;
**** PLACE DATA FROM FIGURE 5 HERE ****
DATA MAPINFO; MERGE WESTCODE EASTCODE WESTBORD EASTBORD;
DATA RATINGS; INPUT T1-T51;
DATA RATINGS; SET RATINGS;
ARRAY T(I) T1-T51; ARRAY R(I) R1-R51;
DO 1=1 TO 51; R = 1 + FLOOR(T-1)/10; IF R=. THEN R=6; END; R52;
DROP I;
DATA NULL; MERGE MAPINFO RATINGS; BY RETAIL C1-C12 X C13 + C14 X C15 1 C16 1 C17 1 C18 1 C19 1 C20 1 C21 1 C22 1 C23 1 C24 1 C25 1 C26 1 C27 1 C28 1 C29 1 C30 1 C31 1 C32 1 C33 1 C34 1 C35 1 C36 1 C37 1 C38 1 C39 1 C40 1 C41 1 C42 1 C43 1 C44 1 C45 1 C46 1 C47 1 C48 1 C49 1 C50 1 C51 1 C52 1 C53 1 C54 1 C55 1 C56 1 C57 1 C58 1 C59 1 C60 1 C61 1 C62 1 C63 1 C64 1 C65 1 C66 1 C67 1 C68 1 C69 1 C70 1 C71 1 C72 1 C73 1 C74 1 C75 1 C76 1 C77 1 C78 1 C79 1 C80 1 C81 1 C82 1 C83 1 C84 1 C85 1 C86 1 C87 1 C88 1 C89 1 C90 1 C91 1 C92 1 C93 1 C94 1 C95 1 C96 1 C97 1 C98 1 C99 1 C100 1 C101 1 C102 1 C103 1 C104 1 C105 1 C106 1 C107 1 C108 1 C109 1 C110 1 C111 1 C112 1 C113 1 C114 1 C115 1 C116 1 C117 1 C118 1 C119 1 C120 1 C121 1 C122 1 C123 1 C124 1 C125 1 C126 1 C127 1 C128 1 C129 1 C130 1 C131 1 C132 1 C133 1 C134 1 C135 1 C136 1 C137 1 C138 1 C139 1 C140 1 C141 1 C142 1 C143 1 C144 1 C145 1 C146 1 C147 1 C148 1 C149 1 C150 1 C151 1 C152 1 C153 1 C154 1 C155 1 C156 1 C157 1 C158 1 C159 1 C160 1 C161 1 C162 1 C163 1 C164 1 C165 1 C166 1 C167 1 C168 1 C169 1 C170 1 C171 1 C172 1 C173 1 C174 1 C175 1 C176 1 C177 1 C178 1 C179 1 C180 1 C181 1 C182 1 C183 1 C184 1 C185 1 C186 1 C187 1 C188 1 C189 1 C190 1 C191 1 C192 1 C193 1 C194 1 C195 1 C196 1 C197 1 C198 1 C199 1 C200 1 C201 1 C202 1 C203 1 C204 1 C205 1 C206 1 C207 1 C208 1 C209 1 C210 1 C211 1 C212 1 C213 1 C214 1 C215 1 C216 1 C217 1 C218 1 C219 1 C220 1 C221 1 C222 1 C223 1 C224 1 C225 1 C226 1 C227 1 C228 1 C229 1 C230 1 C231 1 C232 1 C233 1 C234 1 C235 1 C236 1 C237 1 C238 1 C239 1 C240 1 C241 1 C242 1 C243 1 C244 1 C245 1 C246 1 C247 1 C248 1 C249 1 C250 1 C251 1 C252 1 C253 1 C254 1 C255 1 C256 1 C257 1 C258 1 C259 1 C260 1 C261 1 C262 1 C263 1 C264 1 C265 1 C266 1 C267 1 C268 1 C269 1 C270 1 C271 1 C272 1 C273 1 C274 1 C275 1 C276 1 C277 1 C278 1 C279 1 C280 1 C281 1 C282 1 C283 1 C284 1 C285 1 C286 1 C287 1 C288 1 C289 1 C290 1 C291 1 C292 1 C293 1 C294 1 C295 1 C296 1 C297 1 C298 1 C299 1 C300 1 C301 1 C302 1 C303 1 C304 1 C305 1 C306 1 C307 1 C308 1 C309 1 C310 1 C311 1 C312 1 C313 1 C314 1 C315 1 C316 1 C317 1 C318 1 C319 1 C320 1 C321 1 C322 1 C323 1 C324 1 C325 1 C326 1 C327 1 C328...
Figure 4. State borders and labels for the western region.
Figure 5. State borders and labels for the eastern region.
Figure 6. A shaded map report.