REGIONAL SETTLEMENT IN THE VALLEY OF MEXICO

by

Linda Stephen Neff

Department of Anthropology
Northern Arizona University

Exercise 3 "Sampling Strategies," Professor: Dr. Susan Gregg, Flagstaff, Arizona.
November 17, 1994.

Introduction

Quantitative settlement pattern analysis aids in our interpretation of a region. Rather than focusing exclusively on site cores it is necessary to examine the surrounding settlement. In doing so, we obtain a more comprehensive picture of prehistoric lifeways. Does the settlement ebb and flow with changes taking place in the political centers through time? Are sites specializing in tribute commodities for the maintenance of a regional government? Is there a regional administrative center? If not how do groups or communities organize themselves across geographic space? What rules govern placement of sites? Is it politics, the environment, or both? Quantitative settlement pattern analysis can address all these questions. Through descriptive statistics and exploratory data analysis, I will identify patterns in the data and attempt to interpret the new found relationships. Simple statistics, in conjunction with frequency tables, stem-&-leaf diagrams, histograms and box plots will be the primary tools I use to analyze the Formative Period settlement data from the Valley of Mexico (Figure 1).

Purpose of Project

The purpose of this project is threefold. First, I will use the above mentioned techniques of exploratory data analysis as "both analytical and as presentation devices" (Gregg 1994) to understand Formative Period settlement in the Valley of Mexico. Second, I will draw a 20% probabilistic sample from a *universe*, analyze it and then compare it to the results obtained from a 100% settlement analysis. In doing so, I will learn the advantages and disadvantages of sampling. In the conclusion, I will summarize the results and identify issues needing further resolution for future research.

Sampling Design

Probability sampling is a useful tool when attempting to cover a large area in a short amount of time. In this study, I use a combination of two systematic transect samples, crossing the survey area north to south, as well as east to west (Figure 2). In doing so, I obtain a 20% (19.6% rounded up) sample of the area. Areas that cross count twice in the sample. Initially, I labeled each possible transect from left to right in both directions. I selected the first 1 km. transect from a random number table and then roughly spaced the remaining transects by 10. I chose transect versus quadrants for two reasons. First, I believe I will have a better chance of picking up more variability. Second, they will be easier to find on the ground. A stratified random transect sample may have been the better choice as far as being more representative of the population. Yet, due to time constraints, the systematic sample will suffice. The combination of the north-south and east-west transects represents a majority of the environmental and soil depth zones.

Total Survey Area: 837.82 km^2 ≈ 838 km^2
20% Probabilistic Sample: 167.56 km^2
Total of 8 transects:

- Three running west to east - total area covered: 80 km^2
- Five running north to south - total area covered: 84 km^2
- Total actual sample area: 164 km^2
- Total probabilistic sample: 164/838 = .1957 ≈ .20 or 20%

	Transect #1	Transect #2	Transect #3	Transect #4
SW:	2126.4N 480E	2120.2N 490E	2116.03N 502E	2106.65N 512E
SE:	2126.4N 481E	2120.1N 491E	2116.03N 503E	2106.67N 513E
NW:	2128N 480E	2135N 490E	2132.53N 502E	2131.45N 512E
NE:	2129.1N 481E	2135N 491E	2133.28N 503E	2134.74N 513E

	Transect #5	Transect #6	Transect #7	Transect #8
SW:	2106.8N 522E	2133N 488E	2123N 485.22E	2113N 509.24E
SE:	2106.8N 523E	2133N 521.37E	2123N 525E	2113N 530.78E
NW:	2132.1N 522E	2134N 487E	2124N 485.14E	2114N 509.34E
NE:	2132.1N 523E	2134N 521.35E	2124N 525E	2114N 528.69E

Data Analysis

SYSTAT and MS Excel are the primary statistical packages in use for this analysis. I will first stratify the data by time period: Early, Middle, Late and Terminal Formative and pull out the various site types based on the criteria laid out in Table 1. In addition, I plan on stratifying the two categorical variables: environmental zones and soil depths. Exploration of the interval/ratio data and site types through these stratifications will provide valuable information to further inspect the relationships of each period's settlement over time and elaborate on the settlement within a specific time period.

Table 1. Definitions of Site Types

Site Type	**Description**
Primary Regional Center	Several thousands inhabitants, architecturally complex
Secondary Regional Center	Hundreds to perhaps thousands of inhabitants, nucleated, large-scale ceremonial civic architecture
Segregated Elite District	Isolated residential area in a prominent topographical situation
Nucleated Village	Compact village that lacks large-scale architecture: 100-1500 inhabitants
Large Village	Village with 500 or more inhabitants, used to describe nucleated and compact villages
Dispersed Village	Same as nucleated village, but spread over a large area: 100-1500 inhabitants
Hamlet	Community of under 100 people
Isolated Residences	A homestead or farmstead consisting of 5-10 people. Perhaps a single or extended family
Camp	Small, scattered, seasonal/temporary site with no permanent architecture. (Quarries, salt-making sites, maguey-roasting camps, etc.)
Ceremonial Centers	Architecture, but no extensive debris. Sites can be quite small.

To achieve this I will use a variety of techniques. Simple statistics on the following variables, elevation, rainfall, size in hectares and sherd density, provide the base information for further exploratory analysis. Histograms and stem-n-leaf plots of each variable will help identify modalities, outliers and the overall shape of the distribution. And finally, box plots easily convey the comparisons of these variables over time.

Early Formative

No sites in the sample represent an Early Formative occupation.

Middle Formative

Four sites characterize the Middle Formative Period in the 20% sample of the Valley of Mexico (Table 2).

Table 2. Middle Formative Sites

SITE	ENVIRON	ELEV	RAIN	SOILDEP	HA	SHERD	POP	SITETYPE
14	LP-Smooth	2280	690	Deep	4.3	9.3	40	Hamlet
19	Lake-plain	2260	690	Deep	.2	50	10	Camp
25	LP-Rugged	2280	700	Med-Deep	42	38	1600	Secondary
41	LP-Rugged	2275	740	Med	3.7	21.62	80	Hamlet

Three of the sites concentrate on the rich soil of the piedmont with the fourth site located within the lake plain. The elevation in meters does not vary widely, measuring between 2260 meters and 2280 meters. Rainfall generally increases with the elevation, averaging 705 mm. A site hierarchy presents itself based on site sizes, sherd densities and estimated populations. Unfortunately, the structure of the hierarchy is unclear do to the small sample size.

A cursory analysis of the tabular data reveals one site (Site 25) exceeds 1500. This suggests it is a secondary regional center (Table 2). Site 14 and Site 41 both represent hamlets averaging a population estimate of 60. Site 25, located in the lake plain, has 25% more sherds than the secondary regional center, but is only .5% its size. Thus, this site stands out as a specialized lake activity area. The high sherd densities are a result of repeated use whereas the low site size insinuates a restricted resource area.

In summary, the Middle Formative settlement pattern reflects a focus on the rich soils of the Piedmont with one lake plain resource procurement site. A settlement hierarchy is in place with at least one secondary regional center and two hamlets. One hundred hamlets with five secondary centers is a reasonable estimate for the settlement structure of the entire valley. Yet, because the secondary center is an outlier it is probably over represented. Due to the large gap in estimated populations, it is also reasonable to assume we are lacking quite a bit of variability between these two site types.

Late Formative

Sixteen sites portray a sample of the Late Formative Period in the Valley of Mexico (Table 3).

Table 3. Late Formative Sites

SITE	ENVIRON	ELEV	RAIN	SOILDEP	HA	SHERD	POP	SITETYPE
20	Lake_plain	2255	690	Deep	0.2	50.00	10	Camp
106	LP-Smooth	2530	950	Very Shallow	1.3	7.69	10	Isolated
134	Sub-valley	2470	1020	Med	10	20.00	200	Large
145	LP-Rugged	2400	1080	Deep	0.5	20.00	10	Isolated
13	LP-Smooth	2275	690	Deep	60	35.18	2100	Secondary
68	LP-Smooth	2400	800	Fully Eroded	0.3	33.33	10	Isolated
35	LP-Rugged	2275	725	Med	18	44.94	800	Large
51	LP-Smooth	2340	780	Med-Deep	43	39.35	1700	Secondary
80	LP-Smooth	2360	875	Med	7.2	9.72	70	Hamlet
85	LP-Smooth	2380	890	Med	1	10.00	10	Isolated
137	LP-Smooth	2440	1030	Deep	1.9	5.26	10	Isolated
94	LP-Rugged	2450	900	Med	0.6	16.67	10	Isolated
102	UP-Rugged	2510	930	Shallow	1.9	10.53	20	Hamlet
106	LP-Smooth	2530	950	Very Shallow	1.3	7.69	10	Isolated
122	LP-Rugged	2440	990	Fully Eroded	4.5	11.11	50	Hamlet
131	LP-Rugged	2490	1010	Shallow	0.9	11.11	10	Isolated

During this period the Valley of Mexico experiences extensive changes in settlement and demography. First, considerable population growth measures as the highest prehistoric populations for the valley. Estimated population for the survey area triples (Middle Formative sample = 1730 and Late Formative Population = 5030). Second, the Middle Formative hierarchy further differentiates as another secondary center emerges as well as two large villages. Both of these centers are outliers according to their site sizes and estimated populations (Figure 3 and 4). Roughly 56% of the population live in the secondary centers, 20% lives in the large villages, and the remainder live in scattered hamlets and isolated residences. With this increase in population, settlement occurs in all environmental zones and soil depths (Figure 5 and 6). Yet, more of the sites congregate in the piedmont. One site (Site 20) is probably another lake plain resource procurement area. Similar to the Middle Formative lake plain site, it has the highest sherd density (50.00), but an extremely small site area (.2 hectares).

To conclude, the Late Formative experiences a population boom. If we extrapolate the 20% sample to the region, the population was pushing 25,150. The sample clearly picked up a wider range of variability and a more developed settlement hierarchy. The inhabitants occupied all the environmental zones and soil depth zones available concentrating on the piedmont.

Terminal Formative

The Terminal Formative (Table 4) continues its dominant settlement pattern in the Lower Piedmont (Figure 7) witnessing a 5% decrease in population estimates.

Table 4. Terminal Formative Sites

SITE	ENVIRON	ELEV	RAIN	SOILDEP	HA	SHERD	POP	SITETYPE
87	LP-Rugged	2270	890	Shallow	0.9	0.00	0	Ceremonial
58	LP-Rugged	2330	780	Shallow-Med	1.2	8.33	10	Isolated
36	Lake_plain	2245	730	Deep	3.8	13.16	50	Hamlet
9	Lake_plain	2245	680	Deep	20.9	19.14	400	Dispersed
11	LP-Rugged	2280	700	Med	14	20.00	280	Nucleated
28	Lake_plain	2260	700	Deep	0.4	25.00	10	Camp
16	LP-Smooth	2280	690	Deep	2.5	12.00	30	Hamlet
34	LP-Rugged	2280	725	Med	43.4	41.47	1800	Secondary
45	LP-Rugged	2305	760	Shallow-Med	33.8	39.94	1350	Secondary
50	LP-Smooth	2360	775	Med-Deep	8.7	11.49	100	Nucleated
98	LP-Smooth	2365	925	Med-Deep	23.6	10.17	240	Dispersed
92	LP-Rugged	2470	900	Med	0.9	11.11	10	Isolated
100	UP-Rugged	2520	925	Med	1.8	16.67	30	Hamlet
124	LP-Rugged	2450	1000	Shallow-Med	7.5	20.00	150	Nucleated
126	LP-Rugged	2460	1000	Med	2.7	7.41	20	Hamlet
127	LP-Rugged	2500	1000	Shallow	0.9	22.22	20	Hamlet
129	UP-Rugged	2550	1000	Shallow	0.3	33.33	10	Elite
133	Sub-valley	2470	1020	Med	11.2	17.86	200	Nucleated

135	Sub-valley	2470	1020	Deep	4.2	11.90	50	Hamlet

Two extremely large secondary centers (Site 34 and Site 45) house over 66% of the valley's population (Figure 8). Occupation continues to exist in a variety of settings continuing exploitation of the deep to medium - deep soils (Figure 9). While maintaining its original structure, the settlement populations concentrate into numerous large (dispersed and nucleated) villages. A site size stem-&-leaf diagram targeting populations greater than a 100 and less than 1000 (Figure 10) demonstrates how the dispersed villages (Site 98 and Site 9) all fall within the 75th percentile of their distribution. In opposition, the more nucleated villages are lower than the 75th percentile. For the first time a possible elite district and ceremonial center emerge during the Terminal Formative. The elite district (Site 129) presents itself at the tail end of the 75th percentile sherd density distribution (Figure 11) and is the highest elevation value (2550 m). The small size (.9 hectares), peripheral location and lack of artifactual remains suggest Site 87 is a ceremonial center. Similar to the earlier periods, one specialized camp with relatively high artifact densities occupies a lake plain locality.

In summary, the Terminal Formative sees a slight decrease in population. Population increases in the secondary regional centers (56% in the Late versus 66% in the Terminal) and also increases in the large villages (20% in the Late versus 29% in the Terminal). Much less of the population lives in isolated residences or hamlets (4%). Clearly, the political power affects where the frontier settlement chooses to live. Finally, ceremonial centers and elite districts serve both their sacred and restrictive functions.

A Comparison of the 20% Sample to the Original 100% Survey

Reasons for drawing a sample are many. Yet, one goal focuses on obtaining a representative number of sites for the region. In comparing the sampling strategy to the original 100% survey, I chose to analyze the total number of sites per square kilometer as well as look at site representation for each period. In the original 100% survey, I found 153 sites within roughly 838 square kilometers. For the 20% systematic transect sample, I discovered 39 sites out of 167.6 square kilometers, roughly 22% more sites than the original survey (Table 5).

Table 5. Comparison of Total Number of Sites

	Number of Sites Discovered	**Total Kilometers Squared**
20% Sample Survey	39	167.6 km²
Sample Prediction	195	838 km²
100% Survey	153	838 km²

Sample Prediction discovered 22% more sites than original 100% survey.

Clearly if we took more than one systematic random transect sample the number of recorded sites would vary. Yet, 22% will add a lot of cushion to a possible budget. If we take a brief look at the number of sites discovered during the 20% sample survey by environmental zone, and compare the results to the number of sites found for the 100% survey we find some interesting results. Apparently, the systematic transect sample over represented the LP-Rugged by 63%, Lake-Plain by 36% and the Sub-Valley by 27% (Table 6). This overestimation represents 67% of the transect sample. Thus, if I stratified the original 100% sample by environmental type, the results would more accurately portray the total site numbers.

Table 6. A Comparison of Sites Discovered by Environmental Zone

Environmental Zone	**# of sites discovered on 20% Sample Survey**	**# of predicted sites**	**# of sites discovered on 100% survey**	**Total percent difference between the two strategies**

LP-Hill	0	0	3	-100%
LP-Rugged	18	90	33	+63%
Lake-Plain	5	25	16	+36%
LP-Smooth	12	60	63	-5%
UP-Rugged	1	5	16	-69%
Sub-Valley	3	15	11	+27%
Island	0	0	2	-100%
Lake-bed	0	0	8	-100%

A comparison of sites represented by period illustrates a similar phenomenon. Predicted site numbers always over represented the original 100% survey (Table 7).

Table 7. A Comparison of Sites Discovered by Time Period

Time Period	**20% Sample**	**Prediction**	**100% Survey**	**Percent Difference**
Early	0	0	4	-100%
Middle	4	20	17	15%
Late	16	64	60	6%
Terminal	19	95	71	25%

Again, if I stratified my sample by environment my results would be more accurate. To conclude, this sampling strategy over represented the number of sites in the survey area by 22%. A stratified systematic random sample based on environmental zone would prove more accurate. Future analyses could examine the remaining categorical variables, such as soil depths and rainfall to determine if we could better control our sample. Unfortunately, time is the limiting factor in this analysis and I will not address these questions here.

Conclusion

The above comparison illustrated some of the downfalls of sampling but a review of the general settlement patterns through time demonstrates otherwise. The total number of sites for the Early Formative was only four therefore it is not unusual that none of these sites appeared in the sample. The Middle Formative had 17 sites in the 100% sample and four in the sample. These four sites concurred with the original 100% study in that settlement focused on the rich soils of the piedmont with specialized resource procurement sites located on the lake plain. A settlement hierarchy also presented itself. Yet, its structure was unclear due to the small sample size. The sample missed the few lakebed occupations, isolated residences, and elite districts. What is more important, the sample bypassed five large villages. In any case, the transects did identify substantial diversity for such a small sample size suggesting more variability was present than met the eye.

The transect sample did a fine job in portraying the Late and Terminal Periods' diversity and general patterns. The Late Period saw a tremendous increase in population, concentrating its settlement, further differentiating it hierarchy and occupying all environmental zones. The Terminal Period demonstrates a slight decrease in population, settlement increases in secondary centers and large villages, and less frontier occupation. Where the population appeared to quadruple for the 100% sample, it only tripled for the sample. Unfortunately, the transects missed the primary center first appearing in the Late Formative. Moreover, rarities such as ceremonial centers and elite districts were lacking in the Late Formative sample. Yet, both appear in the Terminal Formative. Clearly, a 100% survey sample was the optimal strategy. Yet, the 20% systematic transect sample portrayed the general settlement trends through time, revealing the immense diversity in site types.

Chalco-Xochimilco Survey Area

Figure 1. Location of Study Area

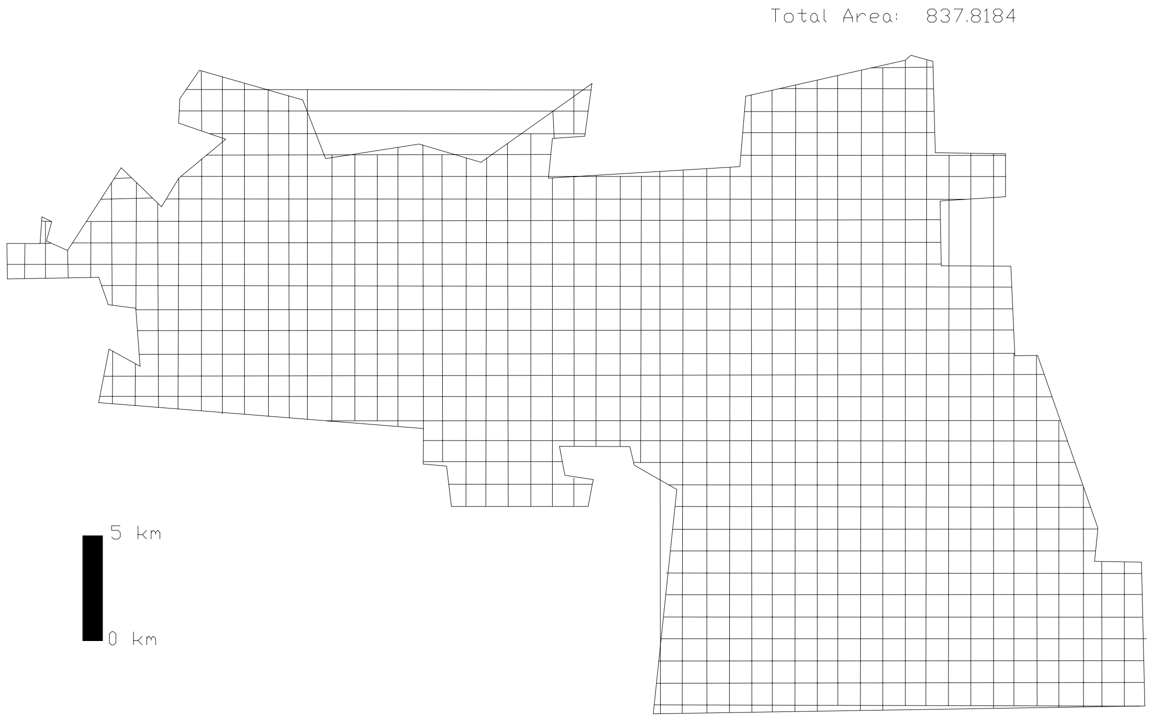

Figure 2. Location of systematic transects within the survey area

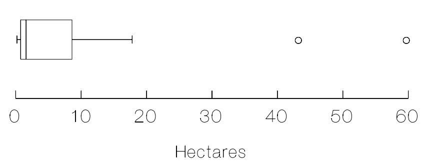

Figure 3. Late Formative Site Size

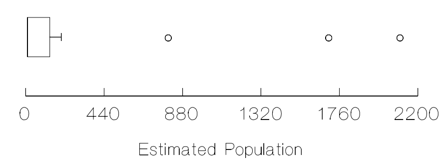

Figure 4. Late Formative Estimated Population

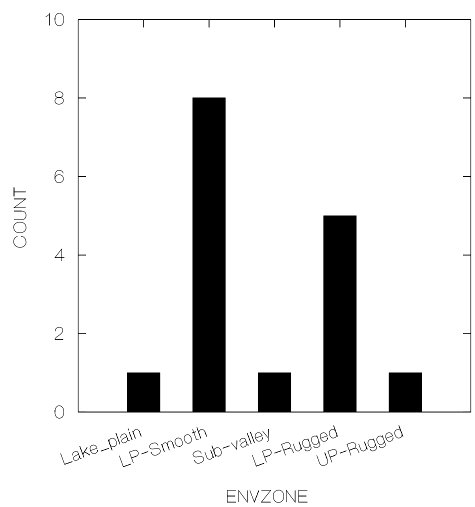

Figure 5. Late Formative Environmental Zones

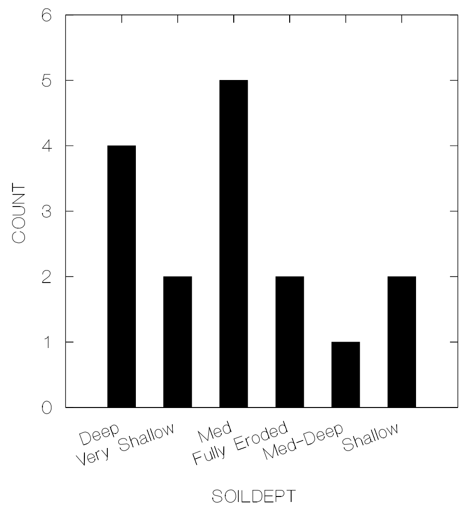

Figure 6. Late Formative Soil Depths

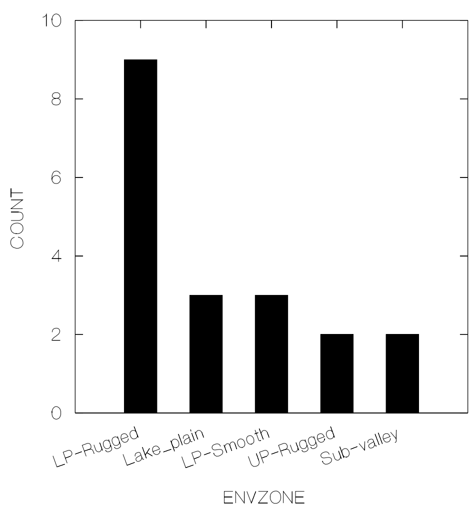

Figure 7. Terminal Formative Environmental Zones

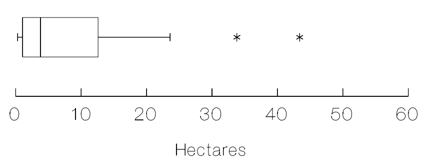

Figure 8. Terminal Formative Site Size

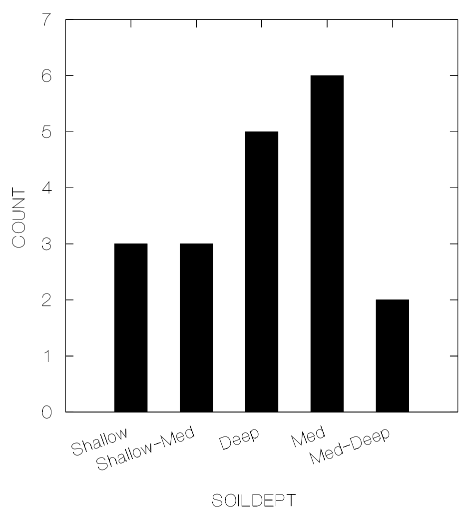

Figure 9. Terminal Formative Soil Depths

0	7
1M	14
1	
2H	03

Figure 10. Terminal Formative Site Size (Population > 100 < 1001)

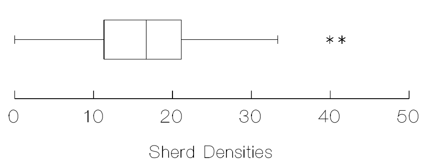

Figure 11. Terminal Formative Sherd Densities