POPULATION POTENTIAL
IN
METROPOLITAN AREAS

by

RICHARD WELLBORN ROETHER

B.E., University of Southern California

(1946)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER IN CITY PLANNING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(1949)

Signature of Author ______________________________
Department of City and Regional Planning, February 21, 1949

Certified by ______________________________
Thesis Supervisor

____
Chairman, Department Committee on Graduate Students

Cambridge, Massachusetts
February 21, 1949

Professor Frederick J. Adams
Department of City and Regional Planning
School of Architecture and Planning
Massachusetts Institute of Technology
Cambridge, Massachusetts

Dear Professor Adams:

I herewith submit a thesis entitled Population Potential in Metropolitan Areas. It is a study of the relationship between an idealized mathematical concept and the existing social structure of the Boston Metropolitan Area, directed toward indicating the usability of the concept in analyzing the structure of metropolitan areas, and hence its use as an urban planning tool.

The results obtained in this study are not conclusive as to the practical applicability of the concept, but it is believed that the results do justify further study in this field, which might lead to the development of a truly useful planning technique.

Respectfully,

Richard W. Roether

Acknowledgements

Appreciation is expressed to Professor A. A. Brown of the Department of Mathematics and to Professors Lynch and Rodwin of the Department of City and Regional Planning for their advice, criticism, and suggestions.

Table of Contents

POPULATION POTENTIAL IN METROPOLITAN AREAS

Introduction

This thesis is in the general field of using the qualitative conciseness and quantitative precision of mathematics to describe the mechanics of society, and through the insight gained by this method, to aid in solving the problem of estimating and providing for the needs of society --- the problem of planning.

More particularly, only the variables of number of people and distance as they operate within one type of social unit --- the metropolis --- are considered here.

Specifically, an index of the influence of people at a distance, called population potential, is dealt with by evaluating it and comparing it with selected data for the Boston Metropolitan Area at about the year 1940.

As a first step this study seeks only to show the existence of relationships between population potential and quantities presently used to describe the functioning of society. The subsequent steps of defining such relationships as may be found to exist and showing their specific applications to planning problems are not attempted here.

History

In the field of mathematical analysis of social structure, Auerbach originally indicated an harmonic series distribution of community sizes.[1] A. J. Lotka first applied the rule to U. S. cities, pointed out that it was not an uncommon distribution, and suggested that it shows little or

---

1. Felix Auerbach, Das Gesetz der Bevolkerungskonzentration, Petermanns Mitt., Vol. 59, 1913, pp. 74-76, referred to by Stewart in Geographical Rev., Vol. 37, 1947, p. 462.

nothing about underlying physical relationships.[1] G. K. Zipf assembled considerable and varied data demonstrating the rule in divers fields.[2] E. L. Thorndike points out that the existence of an inverse relationship between community sizes and number of communities is almost indisputable but that the form of such a relationship is still questionable.[3] Zipf, and apparently Lotka and Auerbach, dealt only with the size distribution of communities within large social units such as nations. Zipf however suggests the applicability of such a rule to trading centers within metropolitan areas.[4]

J. Q. Stewart dealt not with the distribution of community sizes but with the effect of size and distance of communities. He first found that over a period of years the proportion of undergraduates and alumni of "national" eastern colleges, such as Princeton, from states east of the Rocky Mountains, varied in proportion to the population of the states each divided by the distance of the state from the college. He found that attendance at the New York World's Fair varied similarly except within the immediate New York area. He studied the circulation of a St. Louis newspaper and found that, outside the central city and within a range not affected by other large cities, a similar relationship held with regard to the number of papers sold in an area.[5]

Zipf found that the length and frequency of appearance of obituaries in the New York Times, and of news items on inside pages of the Chicago

---

1. A. J. Lotka, Science, Vol. 94, 1941, p. 164.
2. G. K. Zipf, National Unity and Disunity.
3. E. L. Thorndike, Science, Vol. 94, 1941, p. 19.
4. Zipf, Amer. Sociological Rev., Vol. 12, 1947, p. 633.
5. J. Q. Stewart, Geographical Rev., Vol. 37, 1947, pp. 471-472.

Tribune, was in proportion to the size of the community where the items originated divided by the distance of the community from the city where the newspaper was published. He found the circulation of the New York Times to vary according to the same rule. Zipf further found that number of bus passengers, number of railroad tickets sold, and number of phone calls and telegrams between pairs of cities was in proportion to the product of the populations of the two cities in each pair divided by the distance between cities.[1]

This evidence led Stewart to formulate a quantity which he called "potential of population", defined at a location, for a group of people at a distance from the location, as the number of people in the group divided by the distance of the group from the location.[2] He then constructed potential maps for the U. S. and for various continents and countries from available data.[3]

Stewart found that in the U. S., for states, rural population density varied directly with potential squared, rural non-farm population density varied directly with potential cubed, average rural non-farm rent varied directly with potential, and farm land value density varied directly with potential squared. He also found that, in rural areas, death rate, suicide rate, and median age tended to vary directly with potential, and birth rate tended to vary inversely with potential.[4] Since the data from which these relationships were derived was averages

---

1. Zipf, Amer. Journ. of Psychology, Vol. 59, 1946, pp. 401-421; and Zipf, Journ. of Psychology, Vol. 22, 1946, pp. 3-8, etc.

2. Stewart, Science, Vol. 93, 1941, p. 89.

3. Stewart, Geographical Rev., Vol. 37, 1947, pp. 475-480.

4. Stewart, Scientific American, Vol. 178, 1948, p. 23.

for states, these relationships indicate nothing about variations with distance from local urban centers.

Stewart went on to derive a quantity called mutual demographic energy, which is created by the proximity of two individuals, groups of people, or communities, and is equal to the product of the two populations divided by the distance between them. He postulated the relationship of demographic energy to income or wealth and interpreted it as a measure of human relations.[1] Stewart apparently is now working on a system of equations to represent the concept of demographic gravitation and its counterbalancing force --- the need for "elbow-room" or "living-space" --- which tends to produce an equilibrium of population distribution.[2]

Theory

The general hypothesis under test here is that people living in a metropolitan area generate some sort of social influence which is spatially distributed over the terrain in direct proportion to number of people and in inverse proportion to distances of people's residences from the location at which this influence is being considered.

The reasoning upon which this hypothesis is based is: that in order to produce influence at a location, an individual must be there in person or be, or have been, in communication, directly or indirectly, with someone at the location; that the probable frequency with which an individual will thus produce influence at a particular location is in inverse proportion to the difficulty of transportation and/or communication to that location; that the difficulty of transportation and

---

1. Ibid.

2. Ibid; and Stewart, Science, Vol. 106, 1947, pp. 179-180.

communication is directly proportional to the straight line distance between the individual's residence and that location; and that all individuals are equally likely to generate such influence in equal amounts.

This reasoning involves some sweeping assumptions.

It is assumed that the social influences considered here are generated only by the concious actions of individuals operating through known means of transportation and communication.

It is assumed that the probable frequency distribution, over a terrain, of the social influences produced by an individual, is related only to difficulty of transportation and communication. This assumes that other environmental conditions necessary for the operation of social influence are distributed at random over the terrain. If this is not the case, a distribution of social influence based on transportation and communication difficulty alone at least serves to isolate the effects of this factor, permitting the study of the effects of remaining factors.[1] It is felt however that all factors tend to be subject to the force of social influence acting through transportation and communication, and that difficulty of transportation and communication itself tends to be changed by social influence.

It should be pointed out that the occurrence of observable social activities resulting from social influences, depends upon the actual initiation of some sort of social activity as well as upon the existence of a suitable environment.[2] However if an area is considered which is

---

1. Stewart, Sociometry, Vol. 5, 1942, p. 71.

2. Kingsley Davis, The Development of the City in Society, First Conference on Long Term Social Trends, Auspices of Social Science Research Council, March 22, 1947, referred to by Stewart in Geographical Rev., Vol. 37, 1947, p. 472; also see Stewart on "susceptibility" in Sociometry, Vol. 5, 1942, pp. 69,71.

large enough to practically eliminate the domination of special local conditions, the occurrence of some sort of social activity is considered to be inexorable, and thus social activity may be considered to be a direct measure of social influence.

The assumption that the frequency (magnitude) of social influence is inversely proportional to the difficulty (effort, cost) of transportation and communication is suggested by the evidence presented by Zipf and Stewart.[1] It seems reasonable that some sort of inverse relationship exists. The particular form of such a relationship being tested here is inverse proportionality.

The assumption that difficulty of transportation and communication is directly proportional to distance is not accurate. "Work-distance" would be a better quantity to use so as to include the effects of barriers and terminal operations. Every trip or message involves some sort of terminal effort ---- getting the car out of the garage, dialing a telephone number, loading a truck ---- which is related only to the number of trips or messages and not to distance. Similarly a portion of the effort expended in overcoming a natural or artificial barrier is not related to distance but simply to the existence of the barrier. However the cost of fuel, or telephone cable, or highway pavement tends to be directly proportional to distance, and thus distance is a dominant factor in many instances, especially as the distances involved become large ---- a case not as marked in a metropolitan area as, for example, in interstate commerce. In any case distance appears to be the only variable related to transportation and communication difficulty which is universal to all types of transportation and communication.

1. See pp. 2-3.

The use of straight line distances, in the assumption of the proportionality of transportation and communication difficulty to distance, further complicates this assumption since transportation and communication lines are generally irregular. However it appears that lines of transportation and communication tend to be developed so as to minimize effort,[1] and in so doing tend to approximate straight lines. Thus a frequency distribution of social influence based upon straight line distances indicates the forces for overcoming barriers to transportation and communication and, to the degree that such barriers have been overcome, represents the existing influences for social activity.

The use of distances measured only from people's residences presupposes that every trip or message which creates social influence is initiated by the individual at his home. It appears that some social influences, such as those which produce wholesale trade and the demand for production goods, are initiated elsewhere. But perhaps even these are indirectly initiated at the home, if the home can be considered to be the place where all human activities are fundamentally supported. In any case it seems evident that many important social functions, such as retail trade, education, recreation, and work, are influenced directly from the home as it is the place from which individuals usually go and come to take part in each of these activities. It is true that two or more functions may be combined in single trips, such as shopping on the way home from work. But in general the most frequent and hence most probable location of the individual would seem to be the home. Hence this is taken as the location from which many, if not all, social

---

1. Zipf, Amer. Sociological Rev., Vol. 12, 1947, pp. 627-650; and Zipf, National Unity and Disunity, pp. 88-178.

influences emanate, either directly or indirectly.

It is assumed that all individuals have equal inert power to produce social influence. It is apparent that individuals actually have varying powers to produce social influence, but if these variations have a random distribution, the distribution of power approximates the distribution of individuals. Age and sex variations tend to be distributed in proportion to population distribution to the extent that family composition tends to be uniform. The tendency for cultural and economic groups to cluster is perhaps a serious threat to the assumption since special customs or income status may limit or magnify the generation of particular social influences. However if a large number of persons are considered, as in a metropolitan area, and if local areas are considered which are large enough to preclude the domination of special groups, then the effect of such variations will be minimized.

If the limits of a metropolitan area, within which population potential is being considered, reach a relatively large area of low population density on all sides, it may be assumed that the potential on the metropolitan area of the population of the remainder of the world has a practically constant value for all locations in the metropolitan area.[1] If all comparisons of data for the metropolitan area are relative rather than absolute, the omission of the influence of the outside world will have no effect upon the results.

By comparing population potential with data on various types of social activity some light may be shed on the validity of the hypothesis and the assumptions which it involves.

---

1. Stewart, Geographical Rev., Vol. 37, 1947, p. 474.

Potential Maps

As a preliminary study, a map,[1] showing the population potential produced in the Town of Lexington by the population within the town, was constructed from a population distribution spot map[2]. This was done by placing, in turn, at each point of intersection of a ½-mile by ½-mile grid superimposed upon the map of the town, the center of an overlay of concentric circles, counting the number of dots within the town lying in each band of the overlay, multiplying the number of dots in each band by the average inverse distance to that band,[3] and summing these products to give a value of potential which was plotted and from which potential contours were interpolated. Due to the spacing of the plotted values the location of the contours approximates, within something less than ½-mile, the correct location of the contours.

The potential distribution thus plotted was compared with a land use map made about 1940 (the date of the potential values), a land use map made in 1948, and a recent field check. It seems that the potential contours produced by this community graphically represent the "neighborhoods" within the town --- in extent by the location of the contours, and in importance by the magnitude of the peak values. It appears that the use of land for common functions, such as shopping, tends to be at

---

1. See p. 22.

2. Mass. State Planning Board, Boston Metropolitan District Spot Map, 1944.

3. If the population density were uniform, the exact equivalent inverse distance would be the inverse of the average distance. Since population density generally tends to decrease as the distance from the center of gravity increases, the use of average inverse distances partially offsets the bias in the approximation which would appear if inverse average distances were used. An exception is made for the central circle of the overlay where the population is assumed to be of uniform density and the inverse average distance is used in order to avoid obtaining an infinite value.

or near peaks of potential.

Adapting Stewart's technique,[1] a population potential map for the Boston Metropolitan Area[2] has been constructed using the 16th U. S. Census of 1940 and a spot map[3] showing the approximate distribution of nearly 2 million persons in 44 cities and towns as derived from this census.

In order to compute potentials, the City of Boston was divided into 14 districts and Cambridge into 2 districts along census tract boundaries, and each of the other cities and towns constituted a single district. The population of each district was then obtained directly from the 1940 Census.

On the map a station was located by eye in each district, which was estimated to be the center of gravity of the population of the district. A scale of inverse distance (from which distances are read off in terms of reciprocal values) was constructed and used to measure the distances from each station to every other station.[4]

For each station the inverse distances were multiplied by the

---

1. Stewart, Science, Vol. 93, 1941, p. 89; and Stewart, Geographical Rev., Vol. 37, 1947, pp. 475-480.

2. See p. 21.

3. Mass. State Planning Board, op. cit.

4. Since the use of inverse distance gives relatively more weight to population at nearer distances than does the use of distance as in finding the center of gravity, the equivalent distance for the purpose of computing the potential of a district will be somewhat less than the distance to the center of gravity of the district. This bias was partially offset by always reading the inverse distances to the nearer mark on the scale rather than interpolating.

corresponding populations and summed[1] to give the potential of the population of all districts except the one in which the station lay. The potential of each district at the station within the district was obtained by applying the concentric overlay method[2] for each district separately at the station within each district.[3]

The 56 values of total potential obtained[4] were then plotted and contours of equal potential were interpolated by eye.

A potential map represents, then, the distribution over a terrain of social influence, as measured in terms of population potential. Thus, for example, at any point lying on the 0.5 contour on the map for the Boston Metropolitan Area there exists an influence equivalent to 0.5 million persons all exactly 1 mile away, or 5 million persons 10 miles away, etc.

The potential map for the Metropolitan Area shows a rather concentric tendency. Local variations in population distribution appear to have little effect, indicating an overall tendency toward regularity in population distribution for the area. An extremely fine-grained map would probably show minor peaks, but the dominant characteristic would

---

1. For some stations a mistake in this computing operation is known to have occurred which increased the computed values of potential. This error is not considered to be serious enough or frequent enough to warrant correction.

2. See p. 9.

3. Discrepancies between the total number of dots counted in each district and the population of the district as found in the census were distributed between concentric bands in proportion to the distribution of the dots counted. The distribution of population in the portion of Walpole which was beyond the edge of the map was estimated from the distribution of population in the portion of the town shown.

4. Potentials were not computed for Stoughton and Cohasset as the centers of these towns were beyond the edges of the map.

still be the peak of the central city.

Correlations

In an attempt to evaluate the significance of population potential as computed for the Boston Metropolitan Area, linear correlations with selected available data have been computed, plotted,[1] and tested for significance.[2]

The data on population density, assessed valuation density, assessed valuation per capita, residential rent for cities and towns, retail sales density, and retail sales per capita are averages for cities and towns and therefore show nothing of local variations of these variables. Total land areas were used in computing the density data and therefore this data is subject to variations due to special local land use conditions. The retail sales data is for 1939, one year previous to the date of the potential values.

This data has been compared directly with potential values computed for stations in each city and town excluding Boston and Cambridge, it being considered that potential variations within cities and towns other than the two largest cities were relatively small enough to permit the

---

1. See pp. 23-34.

2. Linear correlation coefficients have been computed and tested for significance using the formulae shown on p. 35. The significance test is reliable if the data represents random sampling and if the variability of the data is constant over the range of the data. The significance level chosen was 95%; i.e., to be significant the probability must be at least 95% that the true value of the correlation coefficient is not zero. To satisfy this condition the approximate normal standard unit as shown in the table, p. 36, must be greater than 2.0.

   The squared values of the correlation coefficients as shown in the table, p. 36, represent the part of the total variability of the data being compared with potential which can be associated with the variation of potential.

computed values to be used as approximations of average values and hence comparable with other averaged data.

Data on average rents for census tracts or wards and for census blocks for Boston, Cambridge, and 6 other cities was available and used in an attempt to check the degree of correlation of the computed potential values with more restricted, or localized, averages.

The assessed valuation net density data for census tracts was available for Boston at a date 5 years previous to the date of the potential values and therefore is comparable only if the potential distribution and/or assessed valuation densities were relatively static during the intervening 5 years. This density was computed from net land area and therefore is subject to little variation due to special local land use conditions.

It is found that there is a significant correlation of population potential with average population density for cities and towns. It is apparent however that population density is not a direct function of potential and cannot be derived from the potential value[1] for a small area at a specific location.

In view of the lack of correlation with potential of average assessed valuation per capita and average retail sales per capita, it is possible that these quantities tend to be distributed at a constant per capita rate. If this assumption is made, and in view of the correlation of average population density with potential, it follows that assessed valuation density and average retail sales density should correlate with potential, and such is the case. It is noted however that the correlation of average assessed valuation density for cities and towns is

1. Stewart, Geographical Rev., Vol. 37, 1947; p. 474.

stronger than the population density correlation. The weakness of the correlation of average assessed valuation net density for census tracts in Boston (tracts within which stations of computed potential lie) is considered to be due to the time discrepancy of 5 years although special local conditions may have produced non-typical values in some census tracts. Or perhaps since these values are all for locations in a limited central area of the metropolitan area they reflect some non-linear variation which arises in the central area.

The lack of correlation with potential of average residential rents for cities and towns and for census blocks, as well as the weak negative correlation for census tracts, indicates the lack of a simple, direct relationship between potential and residential rent.

The data for traffic flow was derived from maps[1] of the Metropolitan Area showing estimated annual average 24 hour traffic flows on all major routes for the year 1941, one year subsequent to the date of the potential values. The traffic flow data deals only with more or less through traffic and indicates nothing about traffic on local streets, even rather main streets, unless they happen to coincide with major routes.

The total traffic flow crossing each potential contour was measured from the map and correlated with potential, yielding a slightly significant correlation coefficient. When the total traffic flow at each contour is divided by the length of the contour on land, to give average traffic intensity, a strong correlation is found.

Stretches of routes 2 and 9 were selected for study because they

---

1. Mass. Dep't. of Public Works, Traffic Flow Map of the Boston Metropolitan District, 1941.

are direct routes with no major intermediate turnoffs and because they are not combined with other routes. The traffic flow data for these routes has been plotted against potential, showing a tendency toward the proportionality of traffic flow to potential along these routes. From an inspection of all of the traffic flow data, route C1 to the north of Boston was the only other route found to show this same tendency. Most of the other routes appear to become too confused with one another and too indirect to hypothesize that the total traffic flow along any long stretch of any one of them represents a single, simple relationship.

For all of the data plotted only linear correlation has been tested, but the possibility of the data approximating a variation with the square or cube of potential is not discounted in the case of population density, assessed valuation net density, retail sales density, and traffic intensity.

Results

It is apparent that population potential and population density are related since both quantities are derived from the spatial distribution of population. Population density only relates the population within an area to that area, while population potential relates the population outside an area, as well as that within, to the area.

The relative strength of the correlation of average assessed valuation density with potential, in comparison to the correlation of population density, evidences the existence of some relationship between potential and assessed valuation which is not explained by population density alone, even though there is a close correlation (the correlation coefficient is 0.93) between population density and assessed valuation density. This in turn may indicate the existence of a relationship with potential of economic real property values and incomes, to the extent that assessed valuation, economic property value, and income are related.

Since any random through traffic in the Metropolitan Area which was unrelated to the population distribution of the Metropolitan Area would not be expected to show a significant relationship with the population potential distribution within the area, the correlation of total traffic flow with potential and the tendency of traffic flow along direct, separate routes to be proportional to potential indicates the existence of a relationship between potential and traffic flow. The correlation with potential of traffic intensity is considered to evidence a strong direct relationship.[1]

By considering the differences of form and magnitude between the local potential map for Lexington and the potential map for the Metropolitan Area, as well as Stewart's potential map of the U. S.,[2] it is concluded that in using potential for study purposes, the scope of the population from which potential is derived should coincide with the scope of the problem under study --- local potential for local problems, metropolitan potential for metropolitan problems, national potential for national problems; and perhaps car-owner potential for traffic problems,

---

1. For this proportionality of traffic intensity to be reproduced by a random through traffic flow, the total population of the metropolitan area apparently would have to be concentrated at a single point through which all such traffic passed. There is a possibility that traffic intensity and potential vary with respect to space according to different relationships but that these relationships happen to approximately coincide over the range of the data. Traffic intensity might appear to approach infinity at the central potential peak but it is obvious that some sort of saturation must set in at some point near the center due to the physical limitations of space upon the passage of vehicles. And this saturation might well operate in the same way that the limitations of space prevent an infinite potential. Thus traffic intensity and potential may vary according to the same law.

2. Stewart, <u>Geographical Rev.</u>, Vol. 37, 1947, p. 476.

potential weighted by income for land value problems, etc.

Future Research

The evidence presented here is believed to indicate the existence of useful mathematical relationships between population potential and quantities of importance in metropolitan planning, such as land use, land value, and traffic flow. But this evidence is not proof. Therefore in order to better establish the existence of such relationships and in order to study their form in greater detail, greater refinement is needed in both the computation and plotting of potential and in the gathering, selection, and treatment of data with which potential is compared.

The use of averaged data for rather large areas, such as the use of average values for cities and towns in this study, may have hidden important defects in the relationships which did appear, and may have hidden relationships which did not appear. On the other hand, it is evident that the accuracy of statistical predictability decreases as data is localized. Study of the areal limits of accuracy of potential and of each type of data with which it is compared would be necessary before any practical application of such comparisons could be made with confidence.

All of the assumptions and methods used in this study deserve more careful analysis.

Although population density cannot be derived from population potential, potential is a function of population density. The development of an empirical continuous function in plane coordinates to approximately describe population distribution in a metropolitan area would be of great value in itself as well as aid in the predicting and plotting of changes in the potential map.

Further study of the relationship of potential to land value is desirable. Market or economic values of all types of land in all types of locations should be studied, but the availability of this type of data unfortunately is apt to be limited. Where residential rents failed to show a relationship with potential, commercial and industrial rents might prove to be more directly related.

A direct comparison of detailed land use with potential for the metropolitan area is suggested in order to determine whether or not the existence of specific land uses is dependent upon particular potential conditions or levels. It might be hypothesized, for instance, that retail shopping areas tend to occur only at peaks of potential, whether local or central.

The relationship of potential to complete and detailed traffic flow could be usefully studied, but perhaps a more significant relationship is to be found from the comparison of potential with the results of origin and destination surveys. The study of potential in relation to vehicular traffic might well be extended into a study of the relation of potential to traffic in any and all types of transportation and communication.

The ratio of the potential produced by a community upon itself to the potential of the rest of the metropolitan area upon the community might be studied as an index of community self-containment.

The practical use of population potential may depend upon the development of a system of weighting factors for particular subjects to account for local differences due to cultural or economic grouping or natural or artificial barriers to transportation and communication.

The study of the constants involved in relationships which may be found may yield useful concepts. For instance, the slope of the regression line for traffic intensity and potential is fairly closely approximated by a constant value measured in the unit of vehicles per person per day, which suggests this constant as a measure of the rate of use of motor vehicles in the metropolitan area, and may quantitatively represent a general relationship between man and his automobile in a metropolitan society.

Analysis of the meaning of the integral of potential with respect to area, the rate of change of potential with respect to distance, Stewart's concept of demographic energy, and other similar quantities, as they apply to the metropolitan area, should result in a more complete mathematical description of metropolitan social structure than is to be had from a consideration of population potential alone.

In order to permit generalization, future study might well be done with data from urban areas which differ from the Boston area in type and location.

Bibliography

Lotka, Alfred J., The Law of Urban Concentration, Discussion in Science, 94:164, Aug. 15, 1941.

Stewart, John Q., An Inverse Distance Variation for Certain Social Influences, Science, 93:89-90, Jan. 24, 1941.

_____, A Measure of the Influence of Population at a Distance, Sociometry, 5:63-71, Feb. 1942.

_____, Mathematical Rules Concerning the Distribution and Equilibrium of Population, Geographical Review, 37:461-485, July 1947.

_____, Suggested Principles of "Social Physics", Science, 106:179-180, Aug. 29, 1947.

_____, Concerning "Social Physics", Scientific American, 178:20-23, May 1948.

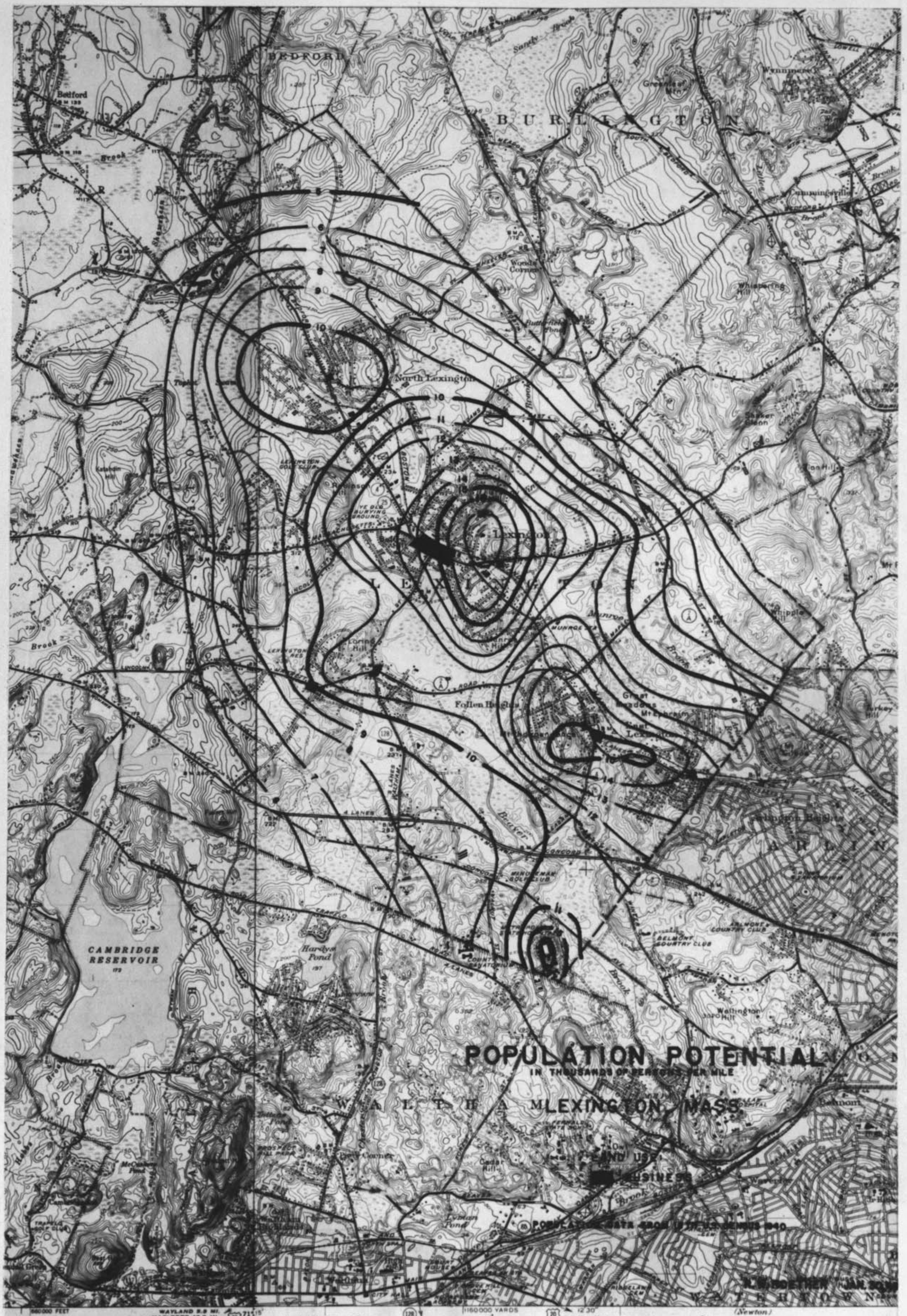
Thorndike, E. L., Review of National Unity and Disunity by G. K. Zipf, Science, 94:19, July 4, 1941.

Zipf, George K., National Unity and Disunity, Bloomington, Ind., The Principia Press, Inc., 1941, 408p.

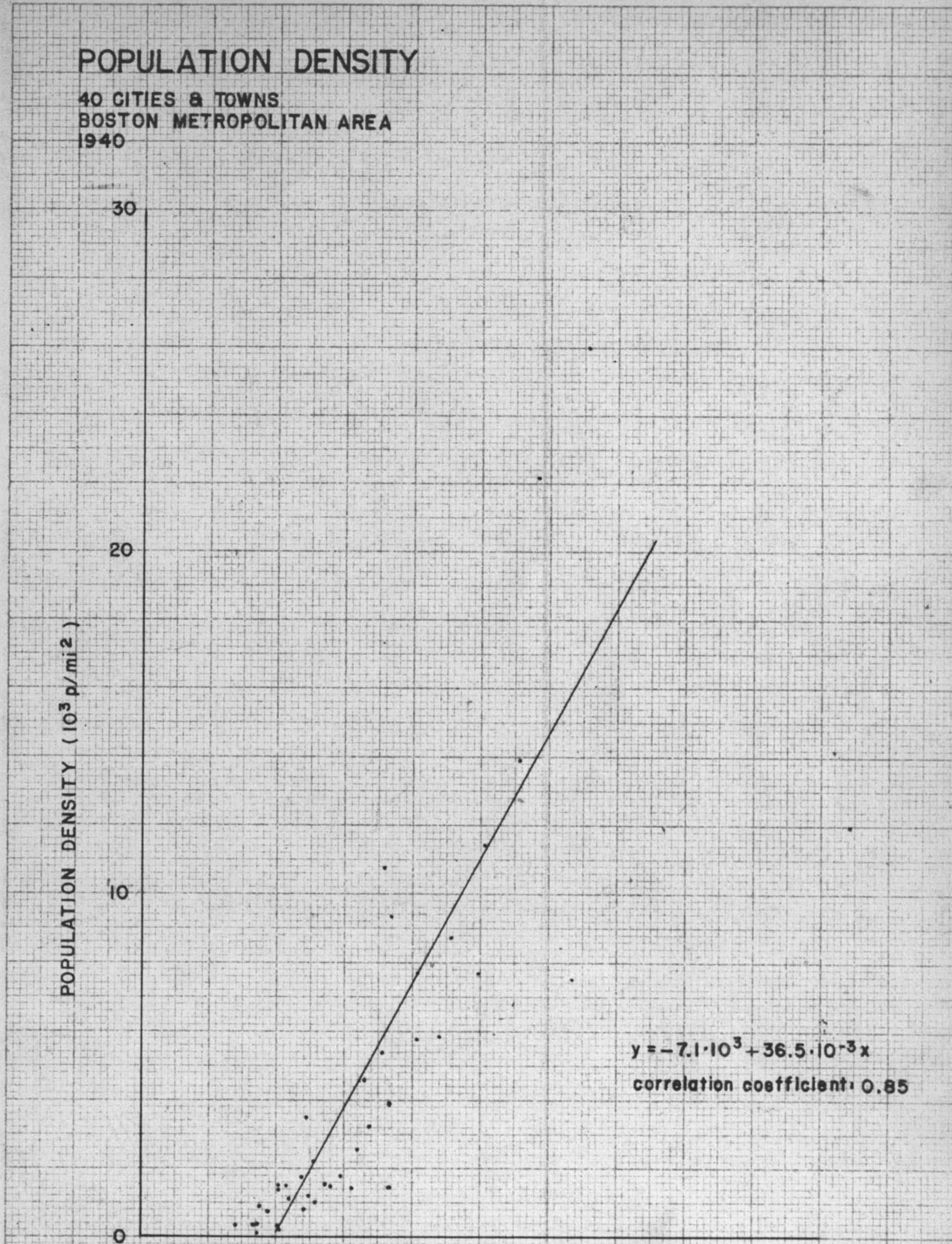
_____, The Unity of Nature, Least-Action, and Natural Social Science, Sociometry, 5:48-62, Feb. 1942.

_____, The Hypothesis of the 'Minimum Equation' as a Unifying Social Principle: With Attempted Synthesis, American Sociological Review, 12:627-650, Dec. 1947.


POPULATION POTENTIAL
IN MILLIONS OF PERSONS PER MILE
BASE MAP SHOWING POPULATION DISTRIBUTION FROM
THE COMMONWEALTH OF MASSACHUSETTS
STATE PLANNING BOARD
THE METROPOLITAN DISTRICT
OF BOSTON
· 1 DOT PER 100 PERSONS
O STATION OF COMPUTED POTENTIAL
DATA FROM 16TH U.S. CENSUS 1940
0.7
0.6
0.5
0.4
0.3
0.2
MARBLEHEAD
NAHANT
NORWELL
HANOVER
R. W. ROETHER JAN. 30, 1949

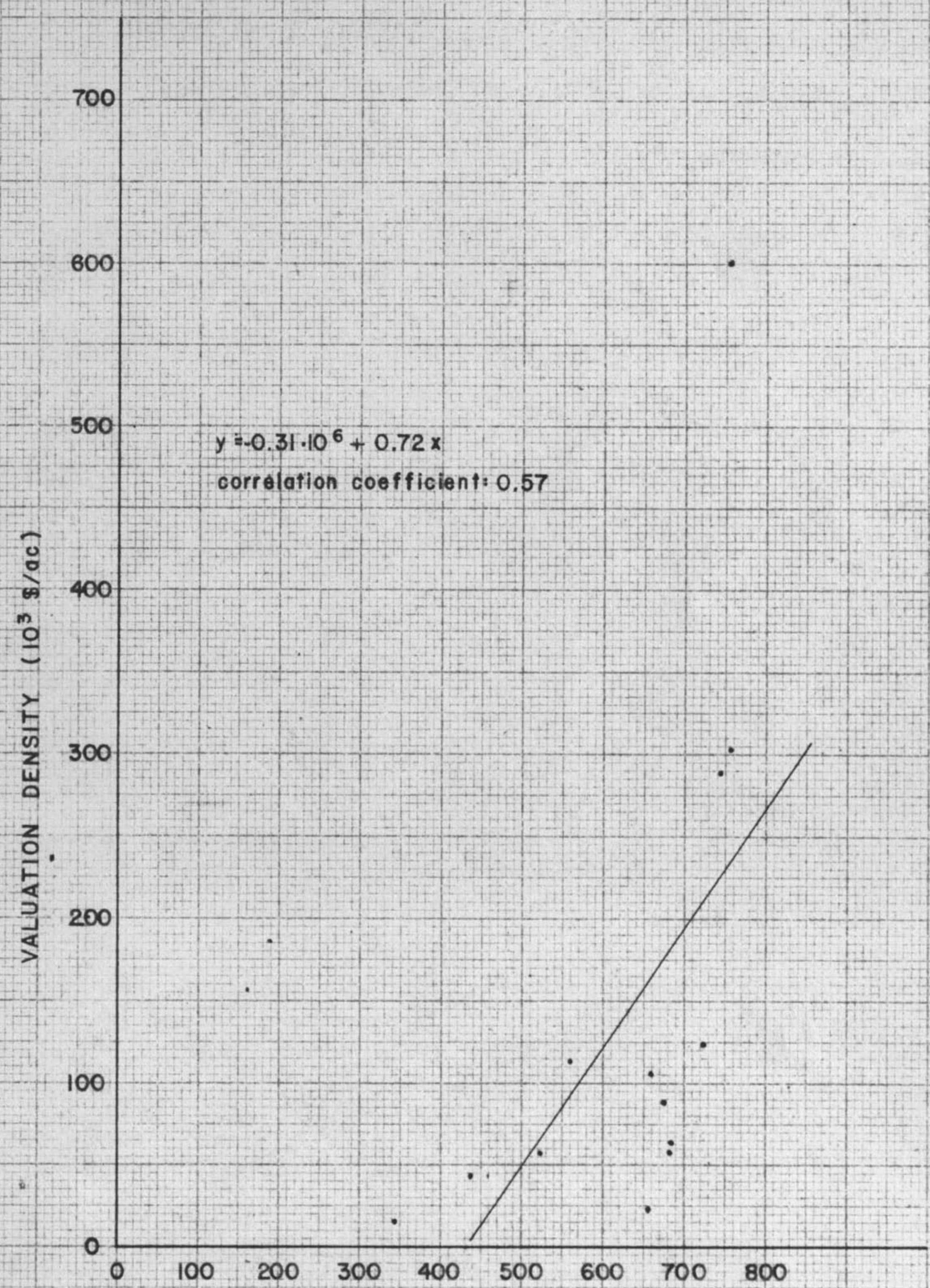

POPULATION POTENTIAL
IN THOUSANDS OF PERSONS PER MILE
LEXINGTON, MASS.
LAND USE
BUSINESS
BEDFORD
B U R L I N G T O N
North Lexington
Follen Heights
CAMBRIDGE RESERVOIR
Mapped by the Geological Survey 1943
Gray tint indicates areas in which only landmark buildings are shown
Polyconic projection. 1927 North American datum
CONCORD MASS.

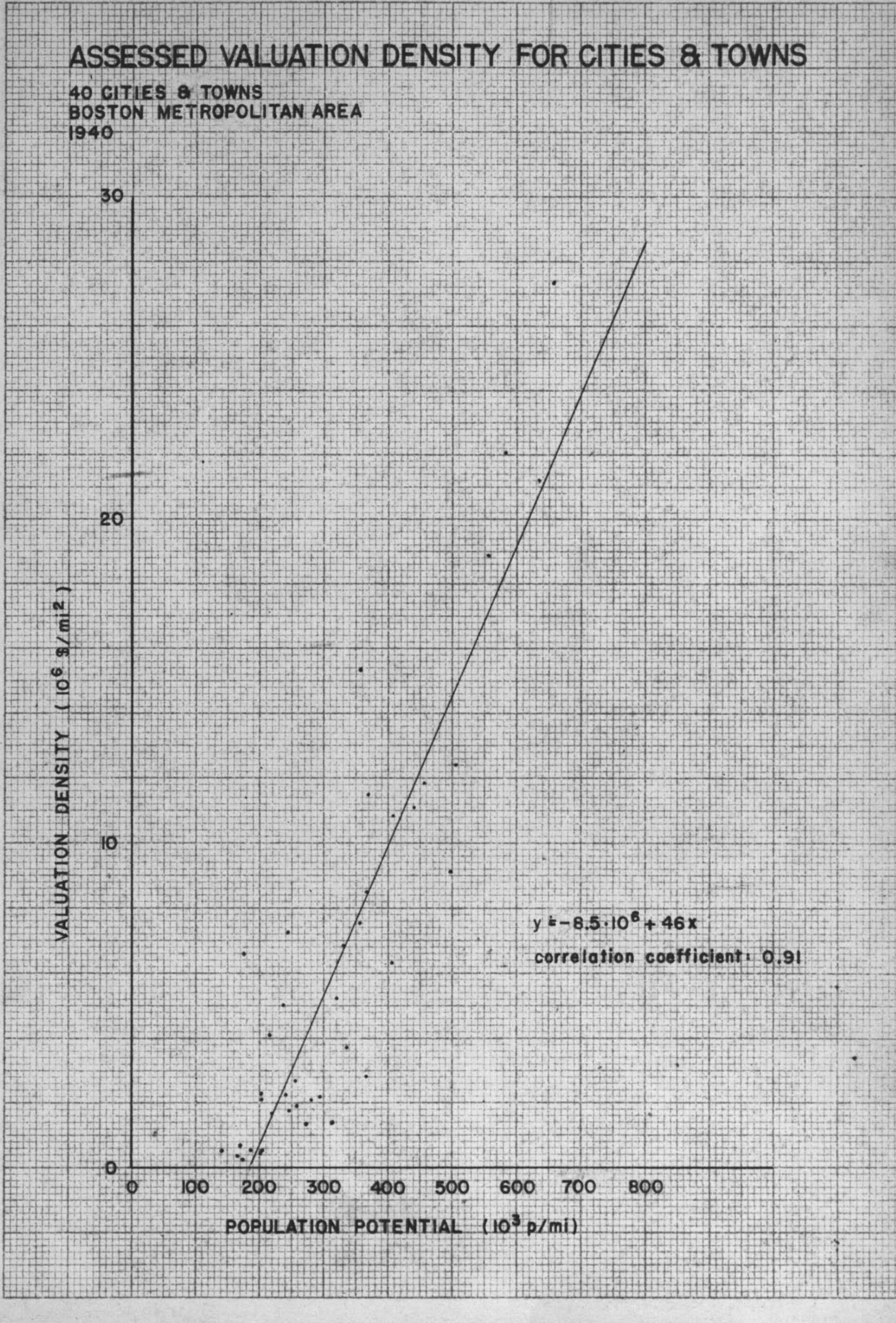
dots clearer

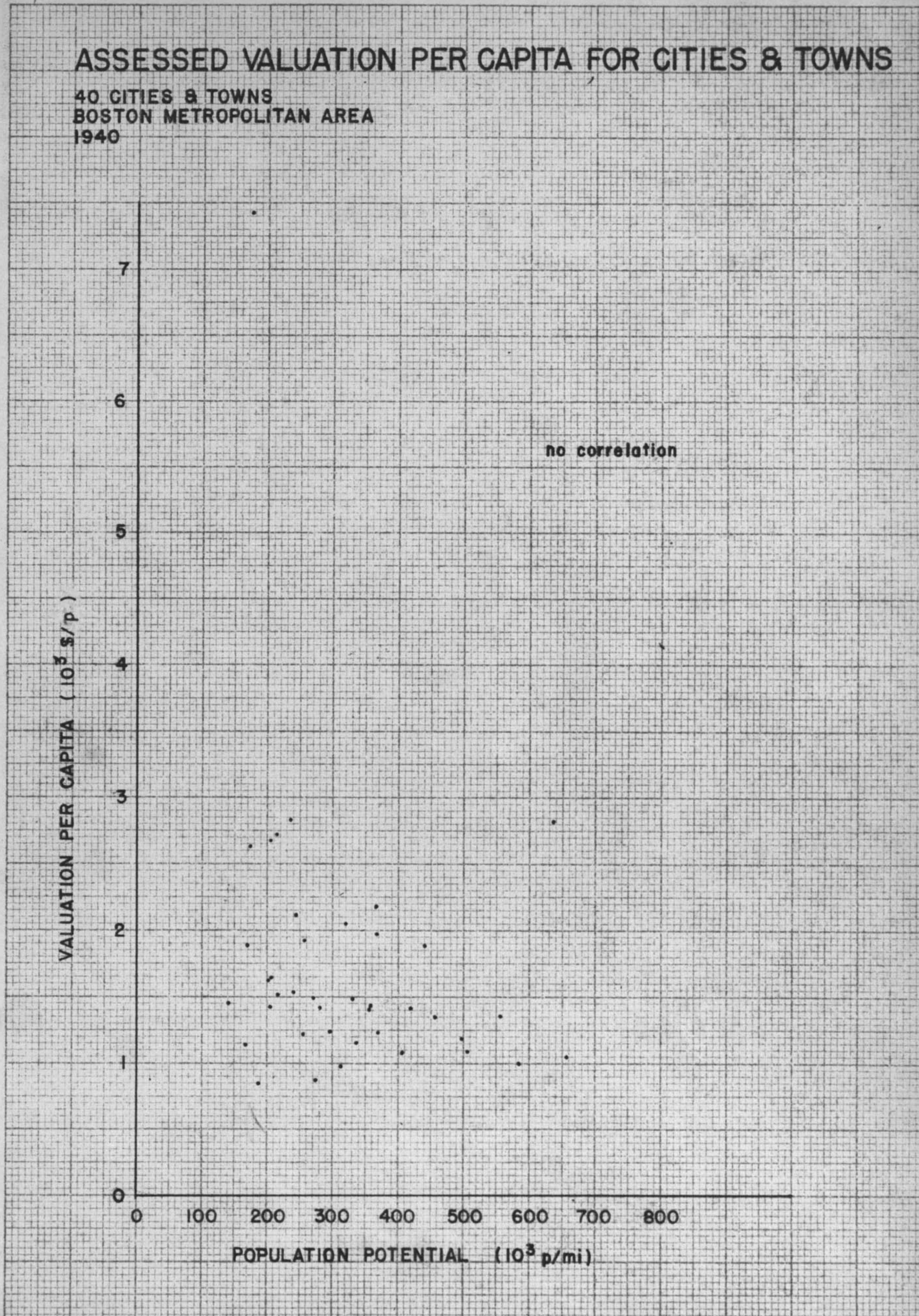


ASSESSED VALUATION NET DENSITY FOR CENSUS TRACTS

13 CENSUS TRACTS
CITY OF BOSTON
1935

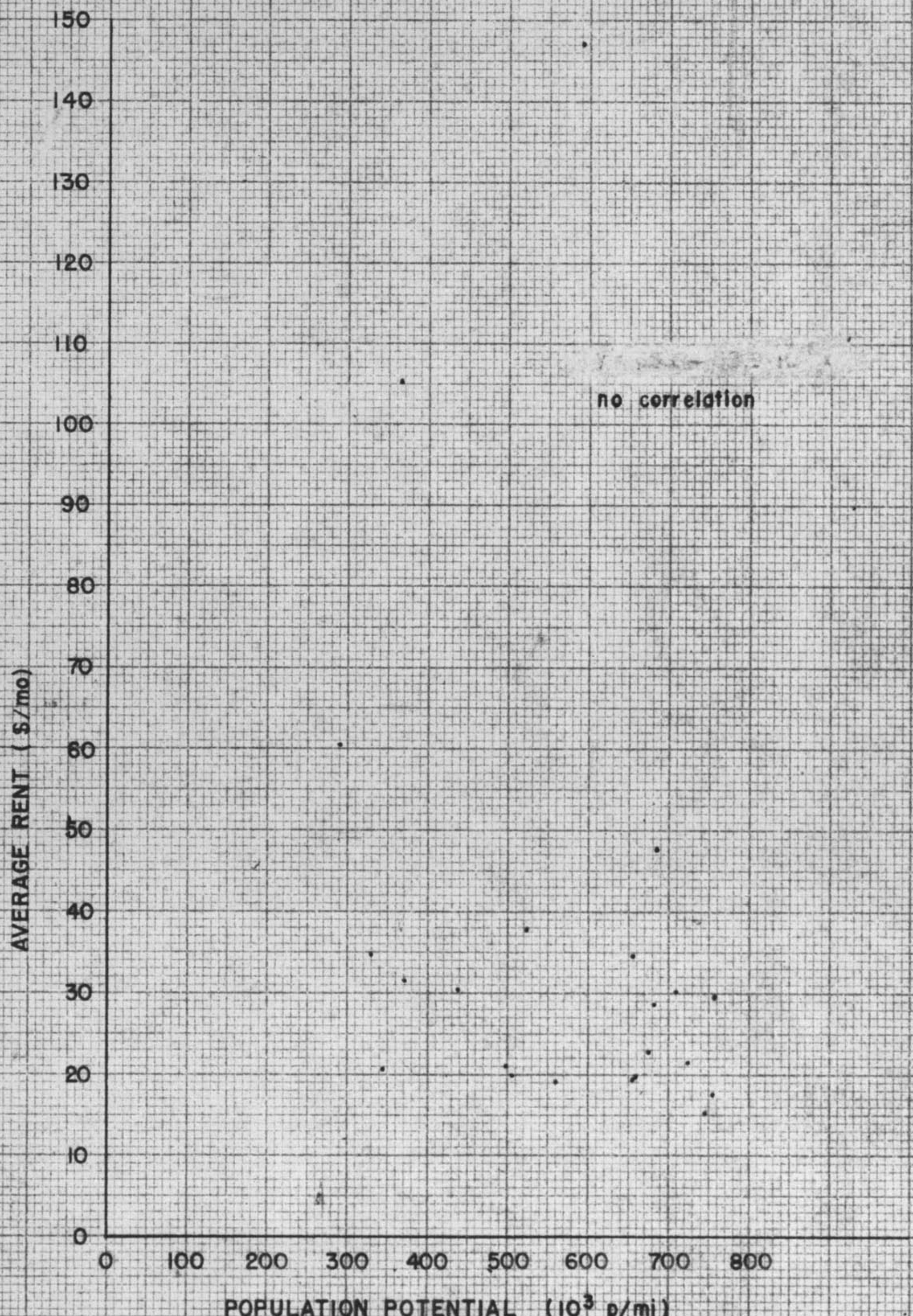


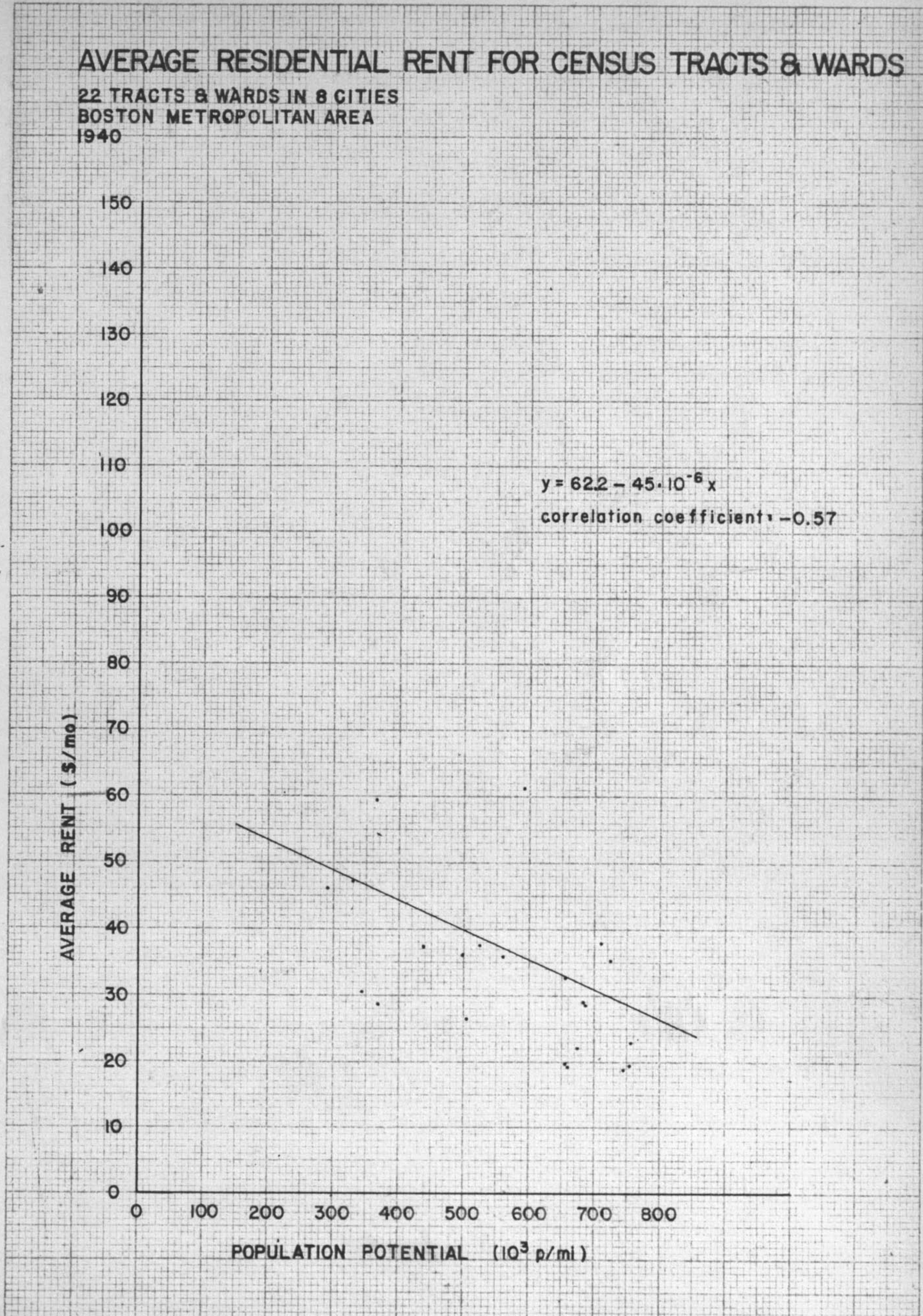

ASSESSED VALUATION DENSITY FOR CITIES & TOWNS
40 CITIES & TOWNS
BOSTON METROPOLITAN AREA
1940
VALUATION DENSITY (10^6 $/mi^2)
0
10
20
30
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10^3 p/mi)
y ≐ −8.5·10^6 + 46x
correlation coefficient: 0.91

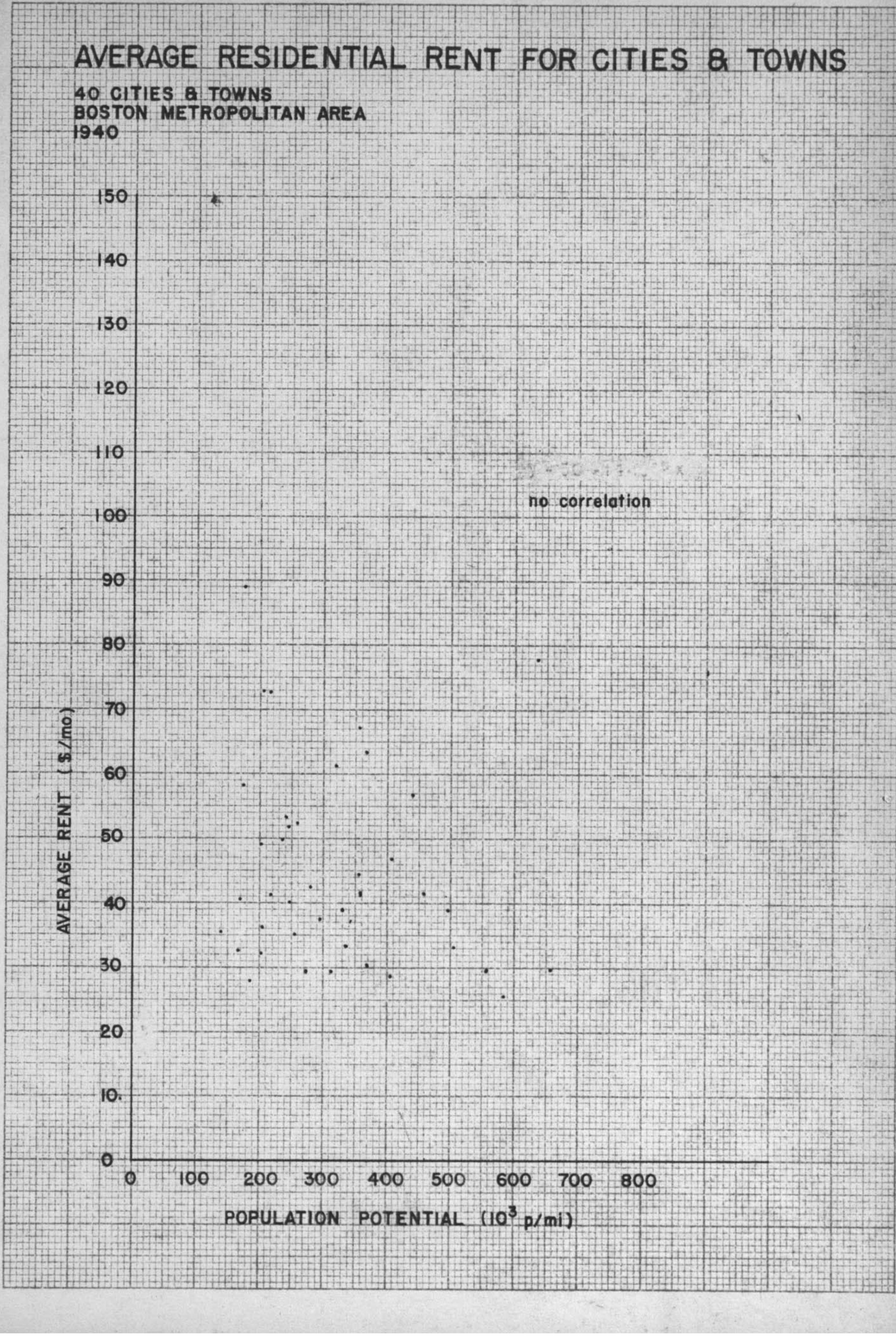

ASSESSED VALUATION PER CAPITA FOR CITIES & TOWNS
40 CITIES & TOWNS
BOSTON METROPOLITAN AREA
1940
no correlation
VALUATION PER CAPITA (10³ $/p)
0
1
2
3
4
5
6
7
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10³ p/mi)

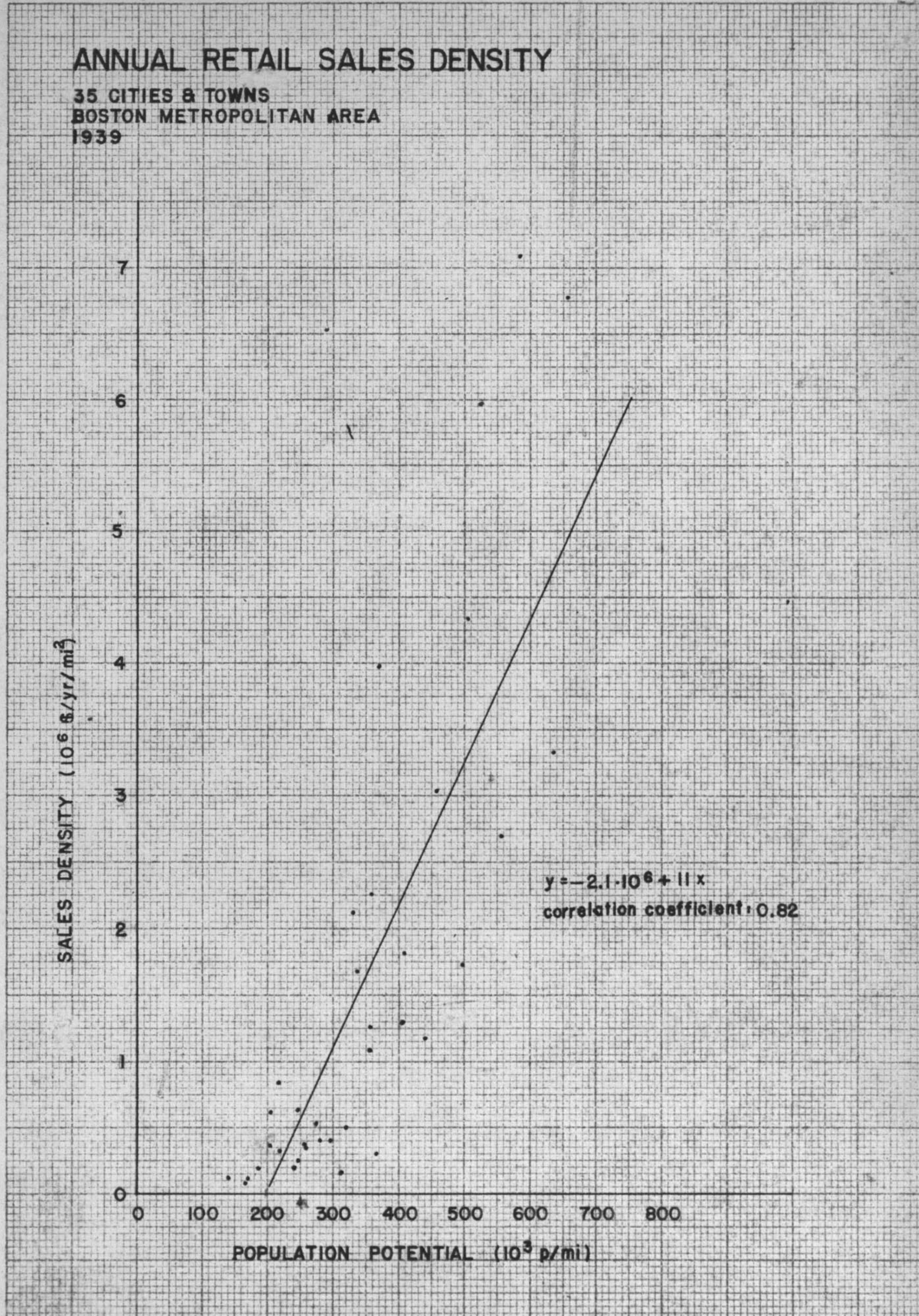
AVERAGE RESIDENTIAL RENT FOR CENSUS BLOCKS

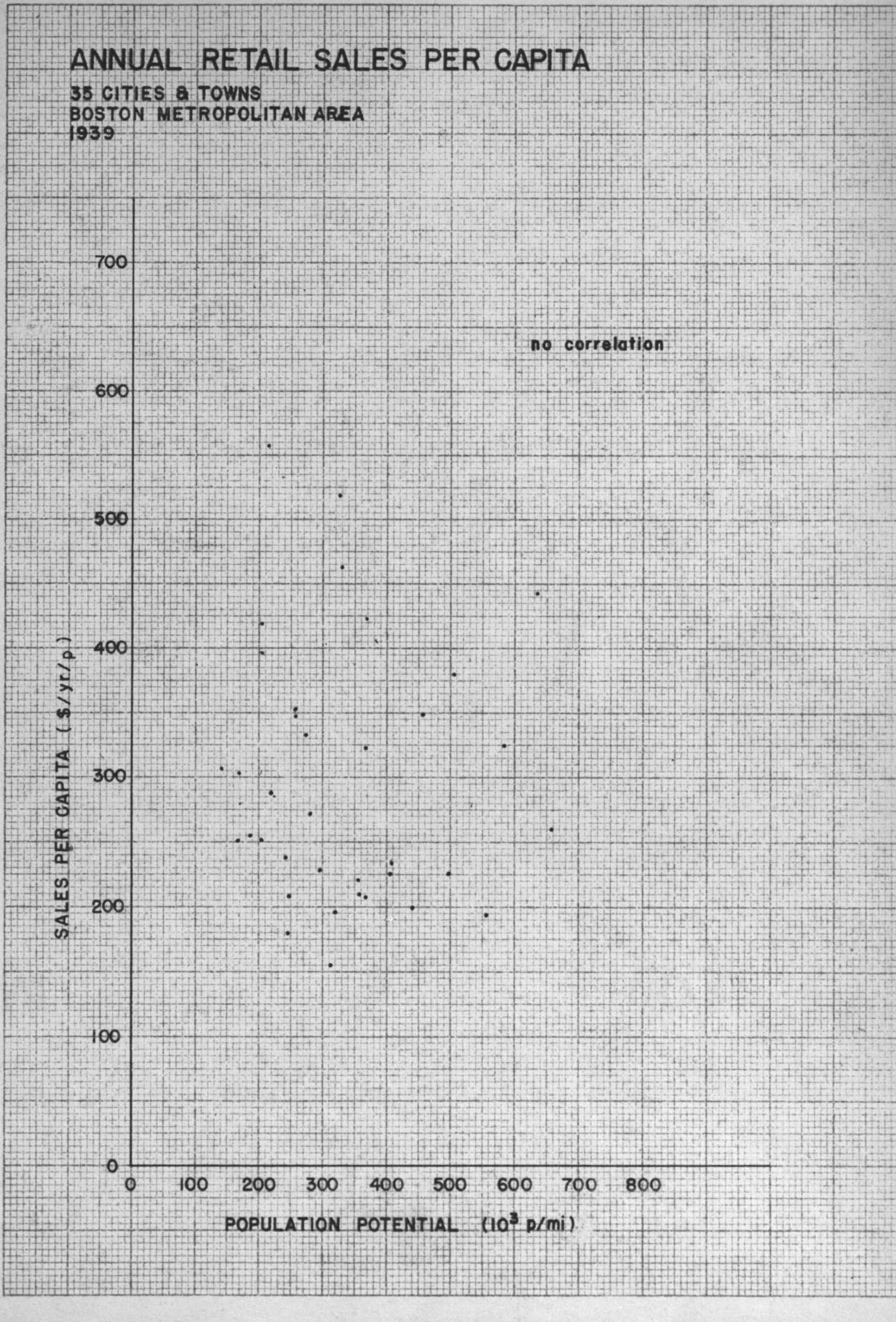
22 BLOCKS IN 8 CITIES
BOSTON METROPOLITAN AREA
1940

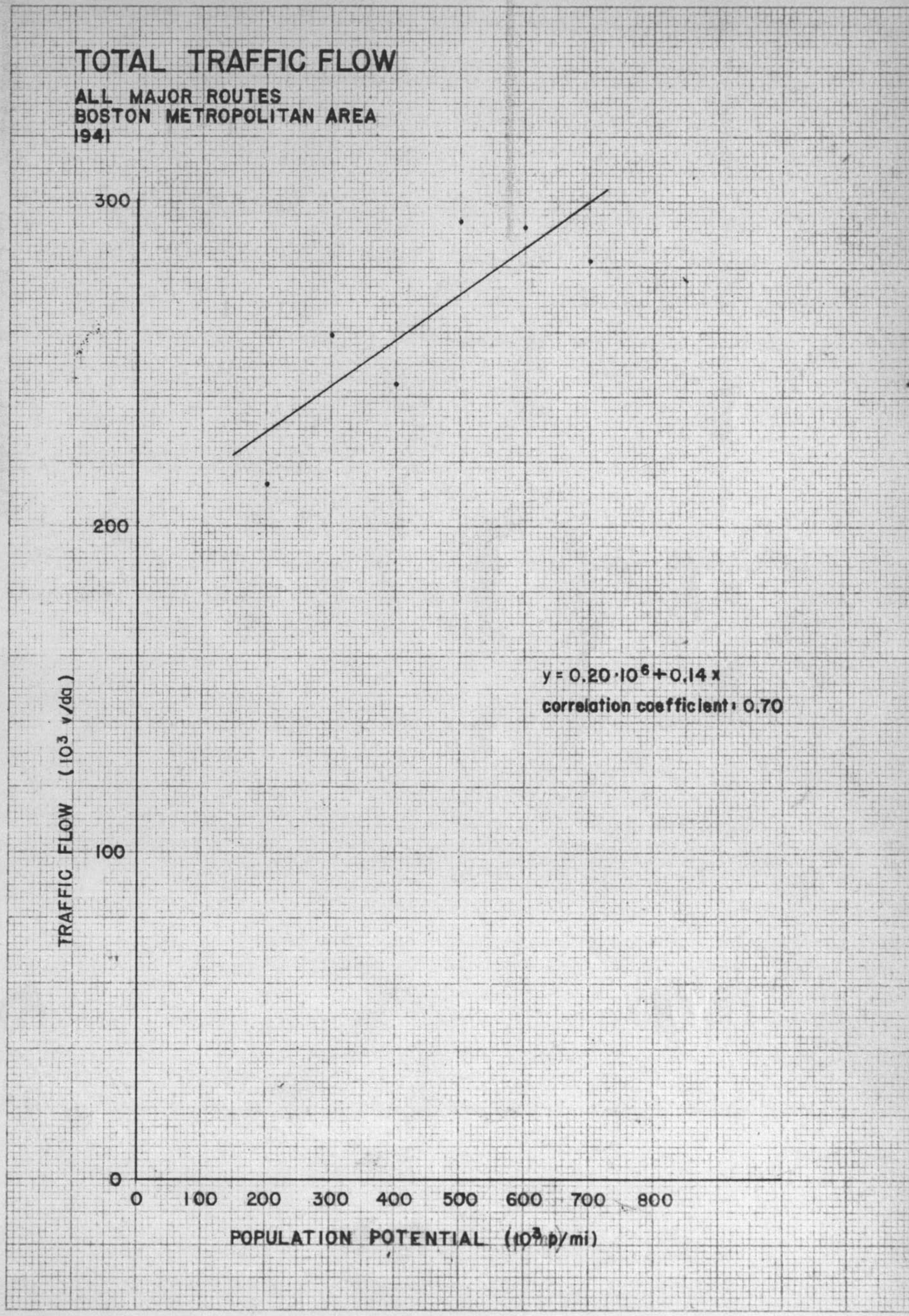


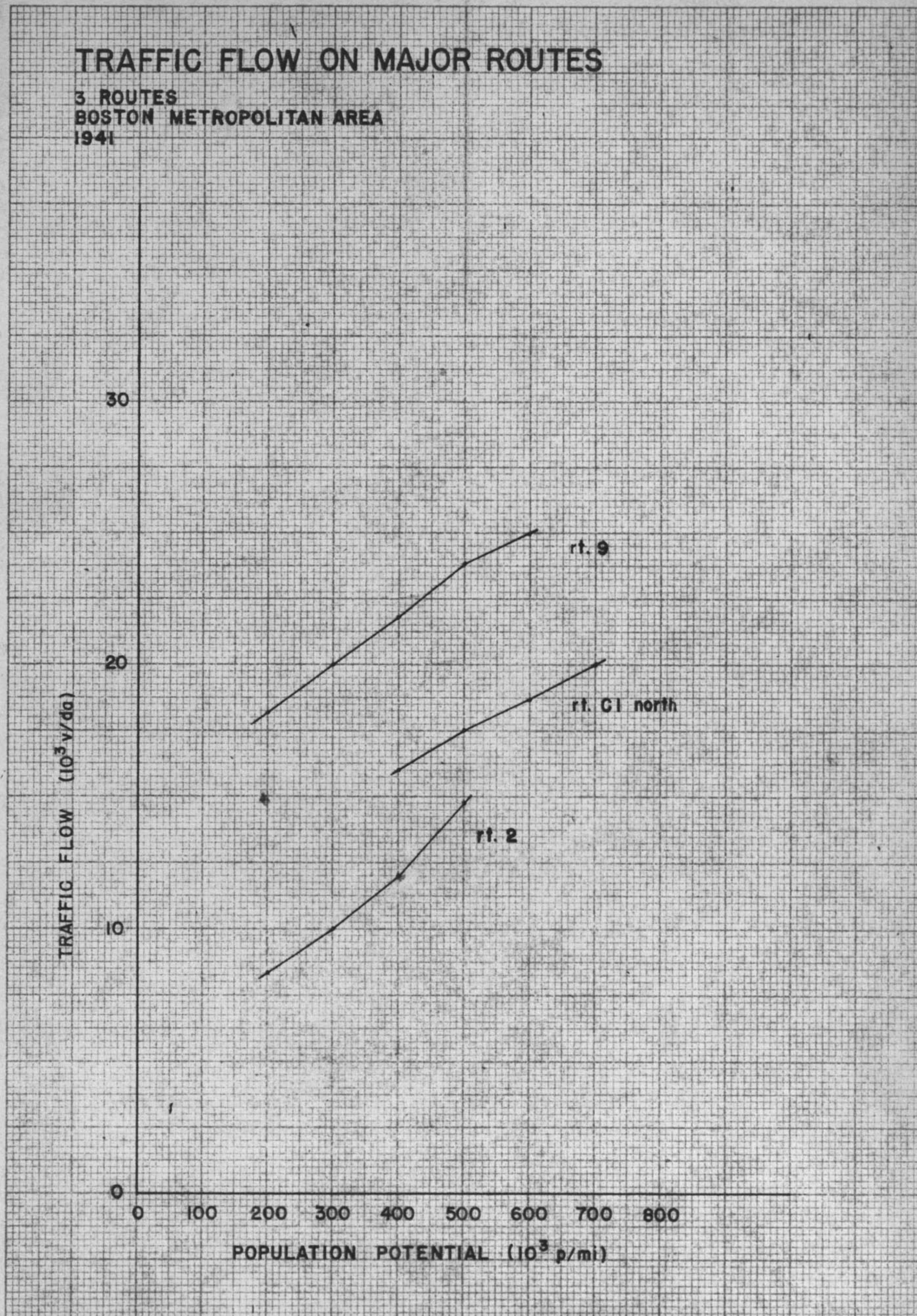

AVERAGE RESIDENTIAL RENT FOR CENSUS TRACTS & WARDS
22 TRACTS & WARDS IN 8 CITIES
BOSTON METROPOLITAN AREA
1940
y = 62.2 − 45·10^-6 x
correlation coefficient: −0.57
AVERAGE RENT ($/mo)
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10^3 p/mi)

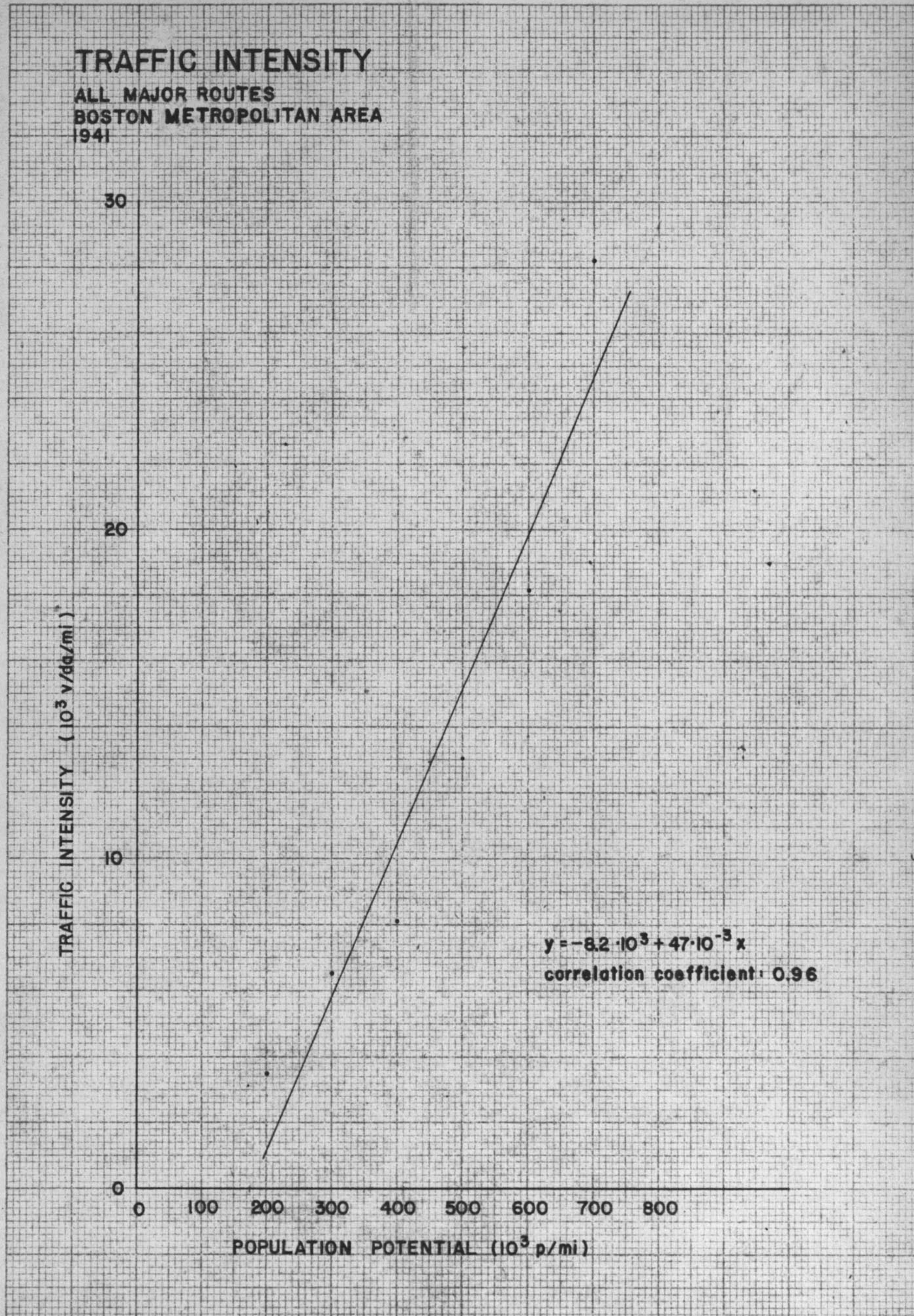

AVERAGE RESIDENTIAL RENT FOR CITIES & TOWNS
40 CITIES & TOWNS
BOSTON METROPOLITAN AREA
1940
no correlation
AVERAGE RENT ($/mo)
POPULATION POTENTIAL (10^3 p/mi)
0
10.
20
30
40
50
60
70
80
90
100
110
120
130
140
150
100
200
300
400
500
600
700
800


ANNUAL RETAIL SALES DENSITY
35 CITIES & TOWNS
BOSTON METROPOLITAN AREA
1939
SALES DENSITY (10^6 $/yr/mi^2)
0
1
2
3
4
5
6
7
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10^3 p/mi)
y = −2.1·10^6 + 11 x
correlation coefficient: 0.82


ANNUAL RETAIL SALES PER CAPITA
35 CITIES & TOWNS
BOSTON METROPOLITAN AREA
1939
no correlation
SALES PER CAPITA ($/yr./p)
700
600
500
400
300
200
100
0
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10^3 p/mi)


TOTAL TRAFFIC FLOW
ALL MAJOR ROUTES
BOSTON METROPOLITAN AREA
1941
300
200
100
0
TRAFFIC FLOW (10³ v/da)
y = 0.20·10⁶ + 0.14 x
correlation coefficient: 0.70
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10³ p/mi)


TRAFFIC FLOW ON MAJOR ROUTES
3 ROUTES
BOSTON METROPOLITAN AREA
1941
TRAFFIC FLOW (10³ v/da)
0
10
20
30
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10³ p/mi)
rt. 9
rt. C1 north
rt. 2


TRAFFIC INTENSITY
ALL MAJOR ROUTES
BOSTON METROPOLITAN AREA
1941
30
20
10
0
TRAFFIC INTENSITY (10^3 v/da/mi)
0
100
200
300
400
500
600
700
800
POPULATION POTENTIAL (10^3 p/mi)
$y = -8.2 \cdot 10^3 + 47 \cdot 10^{-3}x$
correlation coefficient: 0.96

Correlation Formulae

x - abscissa

y - ordinate

n - number of plotted points

Mean x : $\bar{x} = \frac{\Sigma x}{n}$

Mean y : $\bar{y} = \frac{\Sigma y}{n}$

Regression

Slope : $k = \frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma x^2 - n\bar{x}^2}$

y intercept : $y_0 = \bar{y} - k\bar{x}$

Correlation Coefficient

True coefficient : ρ

Sample coefficient : $r = k\sqrt{\frac{\Sigma x^2 - n\bar{x}^2}{\Sigma y^2 - n\bar{y}^2}}$

Significance test

Change of variable : $z = \frac{1}{2} \log_e \frac{1 + r}{1 - r}$

z distribution assumed approx. normal

Mean : m_z

Std. deviation : $\sigma_z = \frac{1}{\sqrt{n - 3}}$

Approx. normal std. unit : $t = \frac{z - m_z}{\sigma_z}$

Hypothesis : $\rho = 0$

Then : $m_z = 0$

And : $t = \frac{z}{\sigma_z}$

Significance level : 5 %

If $|t| < 2.0$, hypothesis accepted (probability of no correlation $>$ 5 %)

If $|t| > 2.0$, hypothesis rejected (probability of no correlation $<$ 5 %)

	1	2	3	4	5	6	7	8	9	10
	POPULATION	POPULATION POTENTIAL LOCAL	METROPOLITAN	TOTAL	POPULATION DENSITY	ASSESSED VALUATION DENSITY	ASSESSED VALUATION PER CAPITA	AVERAGE RESIDENTIAL RENT	ANNUAL RETAIL SALES DENSITY	ANNUAL RETAIL SALES PER CAPITA
SOURCE	1	5	5	5	6	6	6	3	7	7
YEAR	1940	1940	1940	1940	1940	1940	1940	1940	1939	1939
UNIT	10^3p	10^3p/mi	10^3p/mi	10^3p/mi	10^3 p/mi²	10^6 $/mi²	10^3 $/p	$/mo.	10^6 $/yr/mi²	$/yr/p
SWAMPSCOTT	10.8	31	212	243	3.5	7.4	2.11	61.9	0.63	180
NAHANT	1.8	11	227	237	1.8	5.0	2.84	49.9	-	-
LYNN	98.1	176	193	369	9.4	11.5	1.23	30.3	3.97	423
SAUGUS	14.8	32	280	311	1.4	1.4	0.98	29.3	0.15	155
REVERE	34.4	67	338	405	5.8	6.3	1.09	28.8	1.30	225
WINTHROP	16.8	46	311	358	10.7	15.4	1.43	41.7	2.26	210
READING	10.9	30	187	218	1.1	1.7	1.51	41.3	0.32	289
WAKEFIELD	16.2	31	225	255	2.2	2.7	1.22	35.2	0.76	347
STONEHAM	10.8	32	265	296	1.8	2.2	1.24	37.7	0.41	228
MELROSE	25.3	58	298	355	5.4	7.6	1.41	44.5	1.19	221
MALDEN	58.0	117	387	504	11.4	12.4	1.09	33.1	4.34	380
EVERETT	46.8	123	432	556	13.9	18.9	1.36	29.6	2.70	194
CHELSEA	41.3	137	445	582	22.2	22.1	1.00	25.8	7.19	325
WOBURN	19.8	37	234	272	1.5	1.3	0.87	29.3	0.51	332
WINCHESTER	15.1	30	290	320	2.6	5.2	2.05	61.4	0.50	196
LEXINGTON	13.2	20	220	240	0.8	1.2	1.54	53.4	0.19	238
BELMONT	26.9	57	384	440	5.9	11.1	1.89	56.9	1.17	200
ARLINGTON	40.0	66	342	408	7.7	10.9	1.41	46.9	1.81	234
MEDFORD	63.1	91	407	498	7.7	9.1	1.19	39.0	1.74	226
SOMERVILLE	102.2	189	467	656	26.0	27.4	1.05	29.8	6.78	260
WALTHAM	40.0	76	262	337	3.2	3.7	1.16	33.3	1.67	518
WESTON	3.6	5	199	203	0.2	0.6	2.68	73.0	-	-
WELLESLEY	15.1	19	195	214	1.5	4.1	2.72	72.8	0.84	558
NEEDHAM	12.4	30	228	258	1.0	1.9	1.92	52.4	0.35	352
NEWTON	69.9	72	294	366	3.9	8.5	2.18	67.2	1.26	323
WATERTOWN	35.4	69	389	458	8.7	11.8	1.36	41.4	3.04	349
BROOKLINE	49.8	114	521	634	7.5	21.2	2.82	77.9	3.33	443
DOVER	1.4	4	168	172	0.1	0.2	2.63	58.2	-	-
WALPOLE	7.4	13	127	140	0.4	0.5	1.45	35.4	0.11	306
NORWOOD	15.4	36	166	202	1.5	2.1	1.43	36.3	0.61	419
WESTWOOD	3.4	4	197	201	0.3	0.5	1.62	49.1	-	-
DEDHAM	15.5	26	254	280	1.5	2.1	1.42	42.4	0.40	271
CANTON	6.4	15	152	167	0.3	0.4	1.14	32.8	0.08	231
STOUGHTON	8.6	-	147	-	0.5	0.5	0.88	26.2	0.18	330
RANDOLPH	7.6	18	169	187	0.8	0.5	0.84	28.0	0.19	255
MILTON	18.7	21	345	366	1.4	2.8	1.97	63.5	0.30	208
QUINCY	75.8	80	251	330	4.6	6.8	1.49	39.0	2.12	463
BRAINTREE	16.4	20	227	247	1.2	1.8	1.50	40.1	0.25	208
WEYMOUTH	23.9	26	176	202	1.4	2.3	1.63	32.2	0.36	252
HINGHAM	8.0	12	158	170	0.4	0.7	1.88	40.6	0.11	304
COHASSET	3.1	-	132	-	0.3	1.0	3.03	53.5	-	-
HULL	2.2	5	170	175	0.9	6.6	7.42	89.0	-	-

	11	12	13	14	15		16	17	18
	POPULATION	POPULATION POTENTIAL LOCAL	METROPOLITAN	TOTAL	CENSUS TRACT OR WARD & BLOCK		ASSESSED VALUATION NET DENSITY (FOR TRACT)	AVERAGE RESIDENTIAL RENT (FOR TRACT)	AVERAGE RESIDENTIAL RENT (FOR BLOCK)
	2	5	5	5	4		8	4	4
	1940	1940	1940	1940			1935	1940	1940
	10^3 p	10^3p/mi	10^3p/mi	10^3p/mi			$10^3$$/ac.	$/mo.	$/mo.
BOSTON	770.8								
CHARLESTOWN	25.6	139	536	675	C3	12	88	22.2	22.9
E BOSTON	56.9	197	459	656	A6	10	24	19.7	19.6
N END	19.7	193	561	754	F4	1	600	19.4	17.6
W END	27.3	189	568	757	H1	17	304	23.0	29.5
S END	27.3	157	588	745	I1	16	289	18.9	15.3
S BOSTON	64.4	187	472	659	O3	17	106	19.3	19.9
BRIGHTON	63.4	121	439	560	Y3B	58	113	35.9	59.1
BACK BAY	123.5	212	512	724	S1	2	123	35.3	21.5
N DORCHESTER	110.9	219	462	682	T5A	4	57	29.0	28.8
S DORCHESTER	86.3	144	379	523	X4B	16	57	37.6	38.0
ROXBURY	89.0	180	505	684	V1	26	64	28.8	47.8
W ROXBURY	53.6	103	336	438	W5	24	43	37.3	30.4
HYDE PARK	20.1	49	295	343	Z1B	80	15	30.6	20.8
HARBOR	2.7	3	287	290	B6	3	-	46.1	60.3
CAMBRIDGE	110.9								
E	68.6	189	522	711	10	624		38.0	30.2
W	42.3	100	491	591	25	214		61.2	147.1
SOMERVILLE	102.2			656	4	374		32.8	34.8
LYNN	98.1			369	5	452		28.8	31.6
QUINCY	75.8			330	5B	194		47.2	34.9
NEWTON	69.9			366	2	604		59.6	105.2
MEDFORD	63.1			498	4	382		36.2	21.0
MALDEN	58.0			504	7	381		26.6	20.0

TOTAL POPULATION: 1888 000 p

19	20	21	22	23	24
POPULATION POTENTIAL	TOTAL	TRAFFIC RT. 2	FLOW RT. 9	RT C1 (NORTH)	TRAFFIC INTENSITY
5	9	9	9	9	9
1940	1941	1941	1941	1941	1941
10^3 p/mi	10^3 v/da	10^3 v/da	10^3 v/da	10^3 v/da	10^3 v/da/mi
200	213	8.3	18.2		3.45
300	259	10.0	20.0		6.48
400	244	12.0	21.8	16.0	8.09
500	294	14.8	23.8	17.5	13.01
600	292		25.0	18.7	18.18
700	282			20.0	28.22

POTENTIAL CORRELATION	5	6	7	8	9	10	20	24	16	17	18
COEFFICIENT r	0.85	0.91	−0.23	−0.12	0.82	−0.03	0.83	0.96	0.57	−0.57	−0.26
r^2	0.73	0.83	0.05	0.01	0.68	0.00	0.70	0.92	0.32	0.32	0.07
APPROX. NORMAL STD. UNIT t	7.7	9.4	1.4	1.6	6.6	0.3	2.1	3.4	2.0	2.8	1.2

Data Sources

1. *16th U. S. Census 1940*, Population, Vol. I, Number of Inhabitants, Massachusetts.

2. *16th U. S. Census 1940*, Population and Housing, Statistics for Census Tracts.

3. *16th U. S. Census 1940*, Housing, First Series, Data for Small Areas, Massachusetts.

4. *16th U. S. Census* 1940, Housing, Supplement to the First Series, *Housing Bulletin* for Massachusetts, Block Statistics.

5. Massachusetts State Planning Board, *Boston Metropolitan District Spot Map*, 1944.

6. Massachusetts State Planning Board, *Statistics of Massachusetts Cities and Towns by Planning Regions*, Boston, The Board, n. d., Folded maps and charts.

7. Massachusetts State Planning Board, *Development of Retail Trade in Massachusetts Cities and Towns 1929-1939*, Boston, The Board, Nov. 1941, 27p., 2 folded maps, tables.

8. Boston City Planning Board, *Report on the Income and Cost Survey of the City of Boston 1935*, E. R. A. Project No. X2235-F2-U46, Boston, The Board, 1935, 69p., 2 folded maps, tables, diagrams.

9. Massachusetts Department of Public Works, *Traffic Flow Map of the Boston Metropolitan District, 1941*.