AICP Exam Review
Planning Methods Blitz

Bill Drummond
City and Regional Planning Program
Georgia Institute of Technology

http://drummond.gatech.edu/aicpexam.ppt

Basic methods bibliography

- Klosterman, R. E. (1990). Community analysis and planning techniques. Savage, Md.: Rowman & Littlefield. (Technical but good)
- McLean, M. (1992). Understanding your economy : using analysis to guide local strategic planning (2nd ed.). Chicago, Ill.: Planners Press, American Planning Association. (Very clearly written)
- Meier, K. J., & Brudney, J. L. (1997). Applied statistics for public administration (4th ed.). Fort Worth: Harcourt Brace College Publishers. (Many editions; any edition is fine)
- Patton, C. V., & Sawicki, D. S. (1993). Basic methods of policy analysis and planning (2nd ed.). Englewood Cliffs, NJ: Prentice Hall. (Excellent overview of fundamental methods and terms)
- Smith, S. K., Tayman, J., & Swanson, D. A. (2001). State and local population projections : methodology and analysis. New York: Kluwer Academic/Plenum Publishers. (Best resource on local projections)

Session Outline

Introduction (5 min)

A. Descriptive statistics, graphs, tables (5 min)
B. Inferential statistics (10 min)
C. Forecasting methods (10 min)
D. Population analysis and projection (5 min)
E. Economic analysis (5 min)
F. Benefit cost analysis (5 min)

A. Descriptive statistics
Types of data

- Four types of measurement scales
  - Nominal
  - Ordinal
  - Interval
  - Ratio
- Primary data vs. secondary data
- Enumeration or census vs. sample

Measures of central tendency

- Mean
  - Sum of items / Count of items
- Median
  - Sort items high to low
  - Select middle item, or average of two middle items
- Mode
  - What value occurs most often?
  - Bimodal distributions

Measures of dispersion

- Range
  - High value minus low value
- Variance
  - Subtract the mean from each value
  - Square each difference
  - Sum the squares of the differences and divide by the number of cases
- Standard deviation
  - Take the square root of the variance
  - Can relate to original units

Using Tables to Investigate Association

SEX OF STUDENT	FINAL EXAM PERFORMANCE		
	Fail	Pass	Total
Female	34%	66%	100% (205)
Male	54%	46%	100% (130)

Gamma = 0.39

Figure 3-18 Student Test Performance in a Policy Analysis Class

Figure 3–19 Controlling for Whether the Test Taker Attended the Review Session

DID ATTEND REVIEW SESSION	Exam Performance	
	Fail	Pass
Female	40	120
Male	10	30

Gamma = 0.0

DID NOT ATTEND REVIEW SESSION	Exam Performance	
	Fail	Pass
Female	30	15
Male	60	30

Gamma = 0.0

ATTENDANCE AT REVIEW SESSION	FINAL EXAM PERFORMANCE		
	Fail	Pass	Total
Yes	25%	75%	100% (200)
No	67%	33%	100% (135)

Gamma = 0.72

Figure 3–20 Final Exam Performance and Review Session Attendance

Types of Graphs

Figure 3–4 Altman's Typology of Graphic Displays

COMPARISON TYPE	*KEY WORDS*	*GRAPHICAL FORM*	*TYPICAL COMPARISONS*
1. Component	• Contribution • Share • Proportion • Percentage of total	Pie chart	Proportion of tax revenue by major sources Share of municipal budget by operating department Percentage of population in urban, suburban, rural areas
2. Item	• (Item) A more (less) than B • Differences • Rank A is greater (less) than B	Bar chart	Number of employees by department Tax revenue from major sources Operating costs for different field offices of public service organizations
3. Frequency distribution	• Variation • Distribution • Concentration • Relative frequency	Histogram Dot diagram	Number of families in different income classes Distribution of county governments by property tax rate Variation in the population of counties in a given state

Figure 3–4 Altman's Typology of Graphic Displays

COMPARISON TYPE	*KEY WORDS*	*GRAPHICAL FORM*	*TYPICAL COMPARISONS*
4. Co-relationships	• A is related to B • A increases (decreases) as B • A does not increase (decrease) with B	Scatter diagram	Hospital respiratory admissions related to air pollution index. Number of state employees related to size of state population Consumer expenditures increase with disposable income
5. Time series	• Trends • Since • From (date) to (date) • Verbs, such as Fluctuate Change • Nouns, such as Rise Decline Growth	Time series—plotting a curve which shows how the quantity of an item varies with time	Changes in annual municipal budget Trends in amount of refuse collected Seasonal variations in unemployment

SOURCE: Stanley M. Altman, "Teaching Data Analysis to Individuals Entering the Public Service," *Journal of Urban Analysis* 3, no. 2 (October 1976), 217.

Figure 3-5 Pie Chart: Sun City Neighborhood Populations, 1980

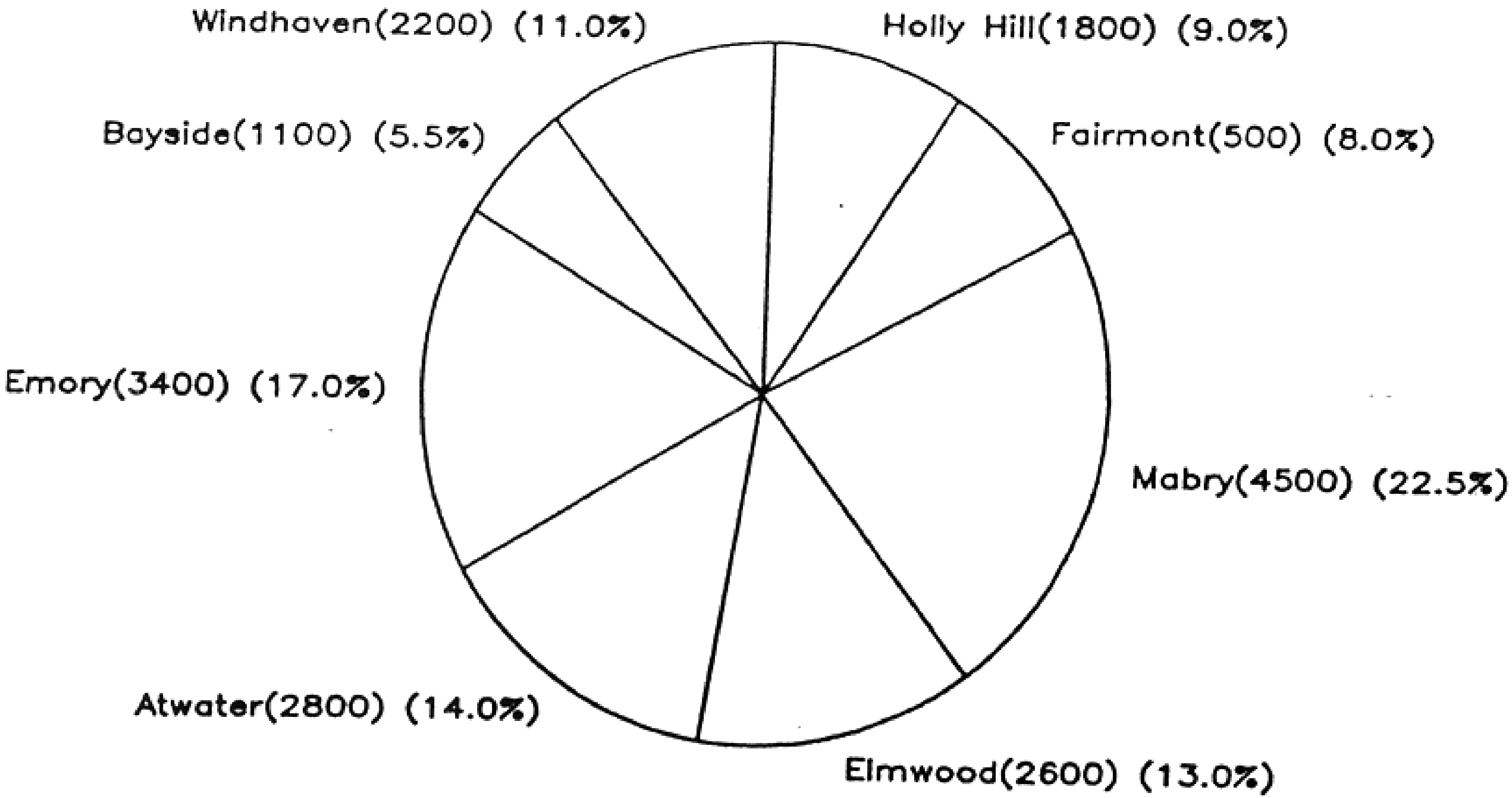


Figure 3-6 Bar Chart: Number of Households by Race, Holly Hill, 1980

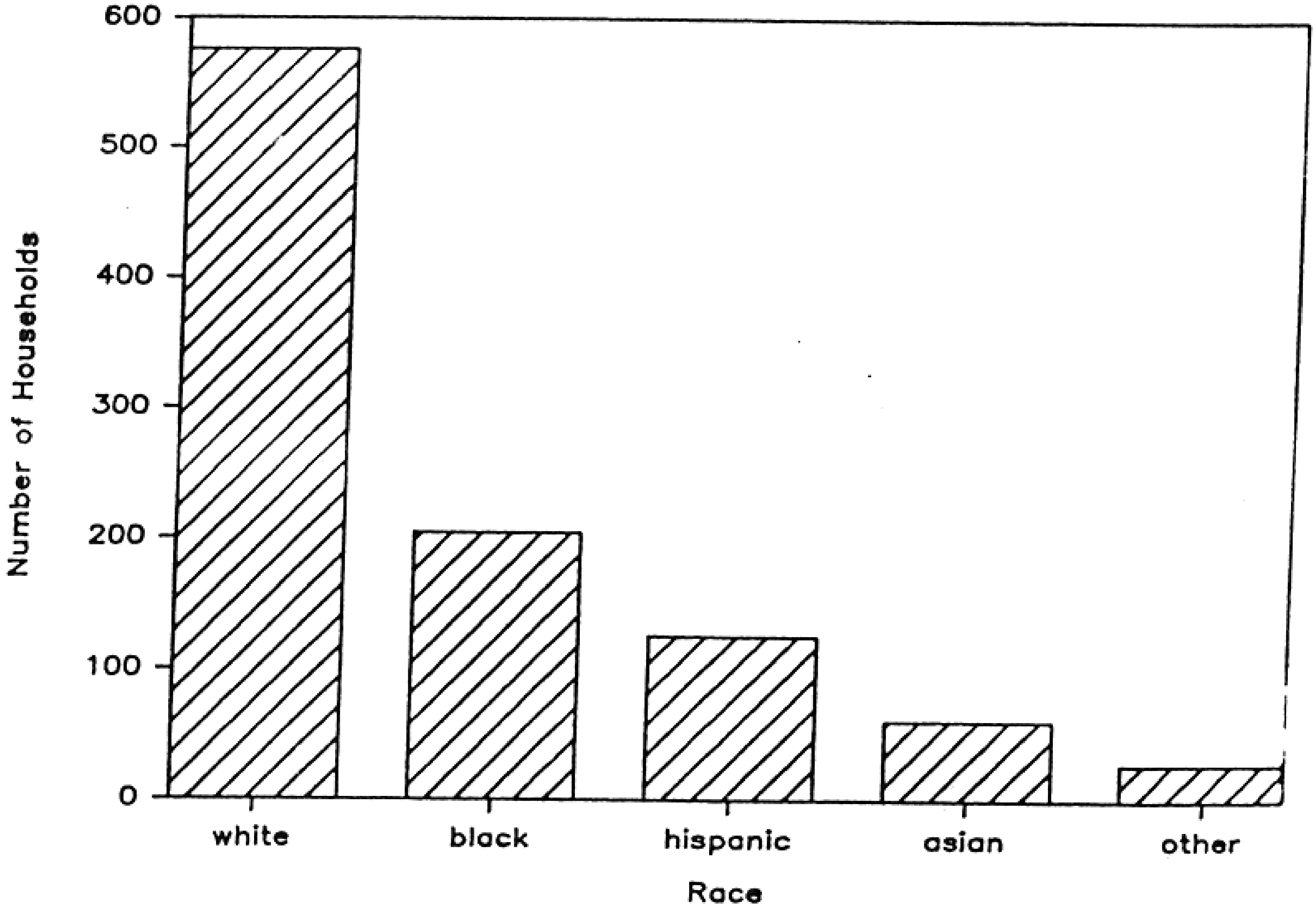


Figure 3-7 Histogram: Number of Households by Income, Bayside, 1980

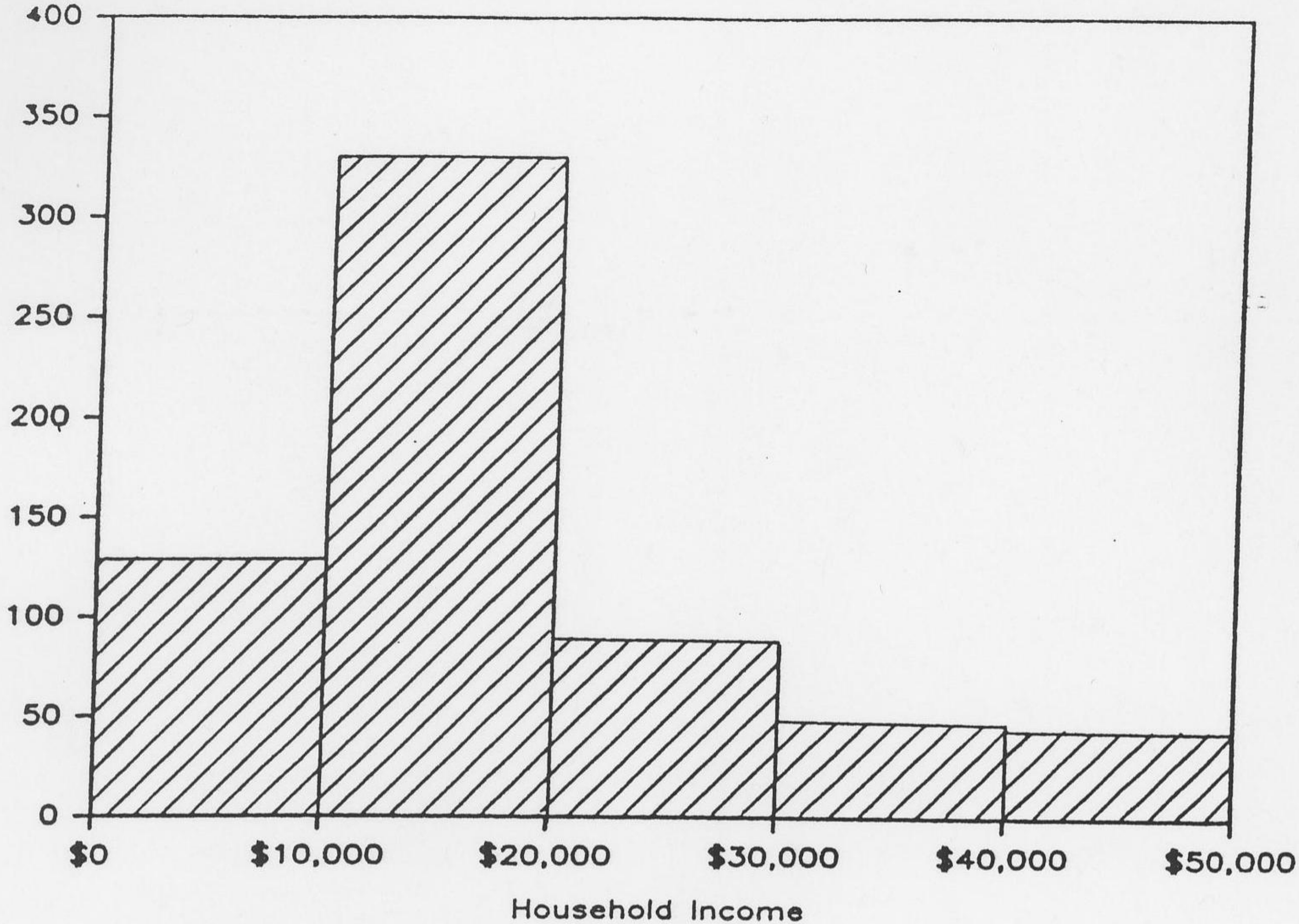


Figure 3-8 Dot Diagram: Number of Households by Income, Bayside, 1980

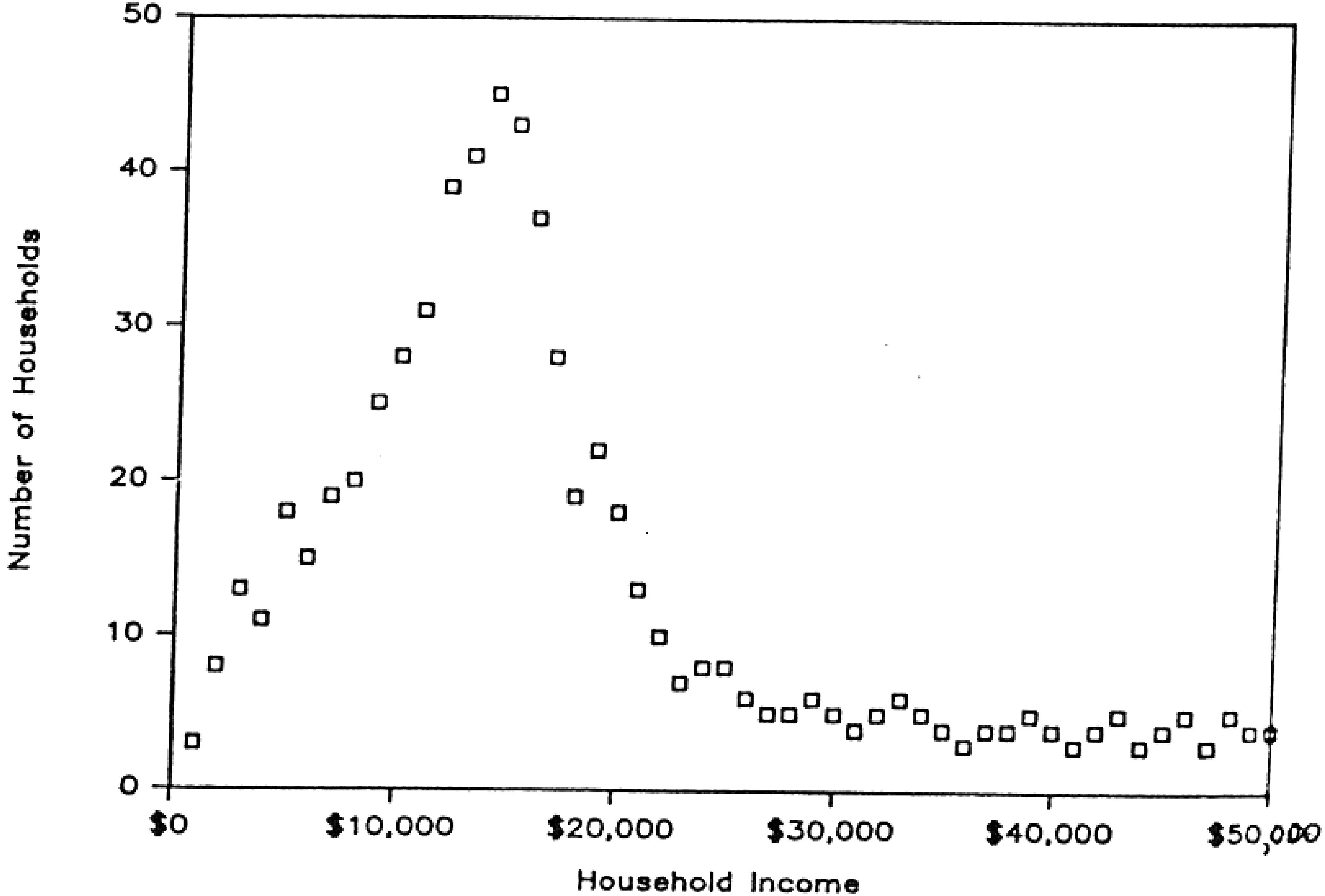


Figure 3-9 Scatter Diagram: Rent versus Household Income, Fairmont, 1980

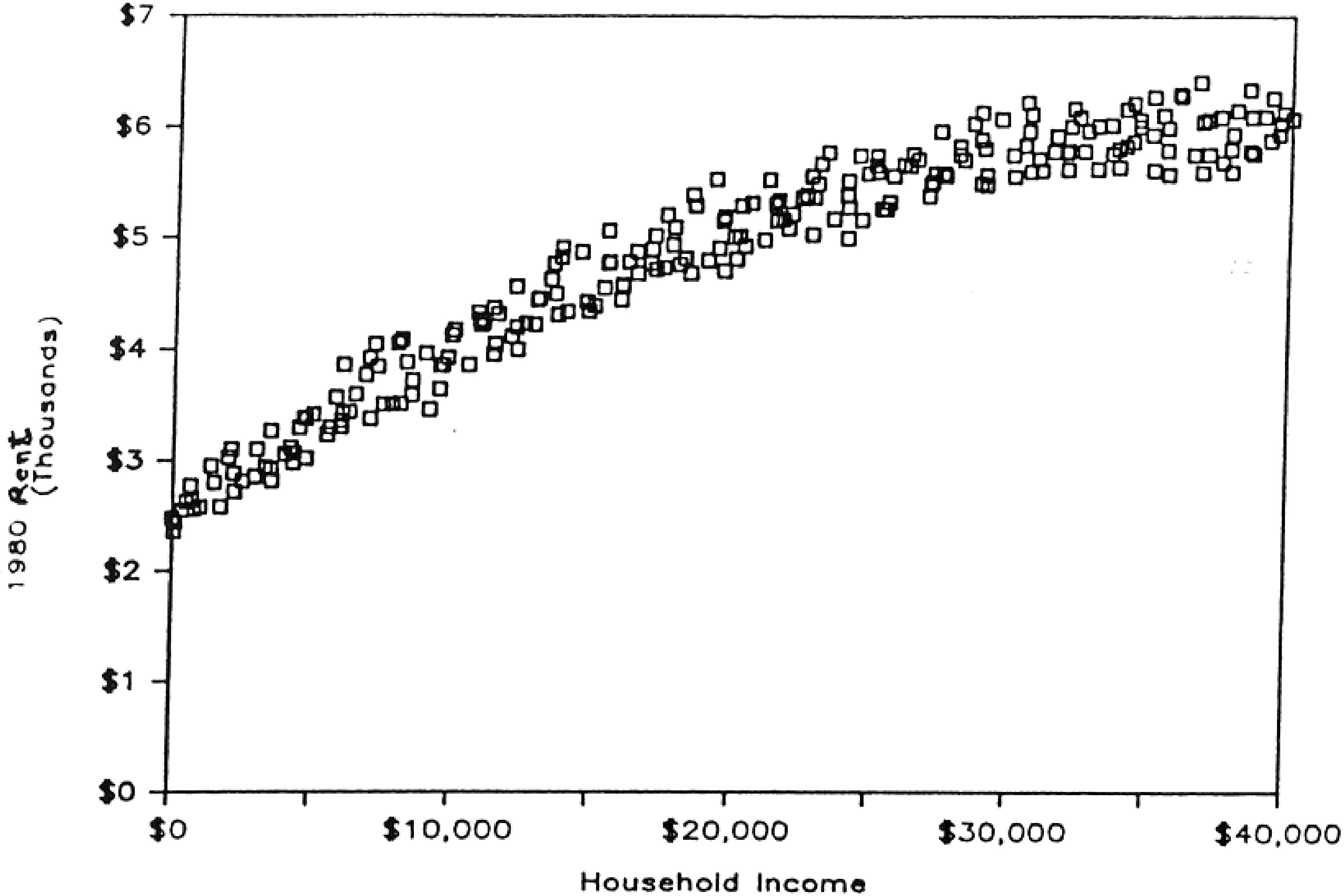
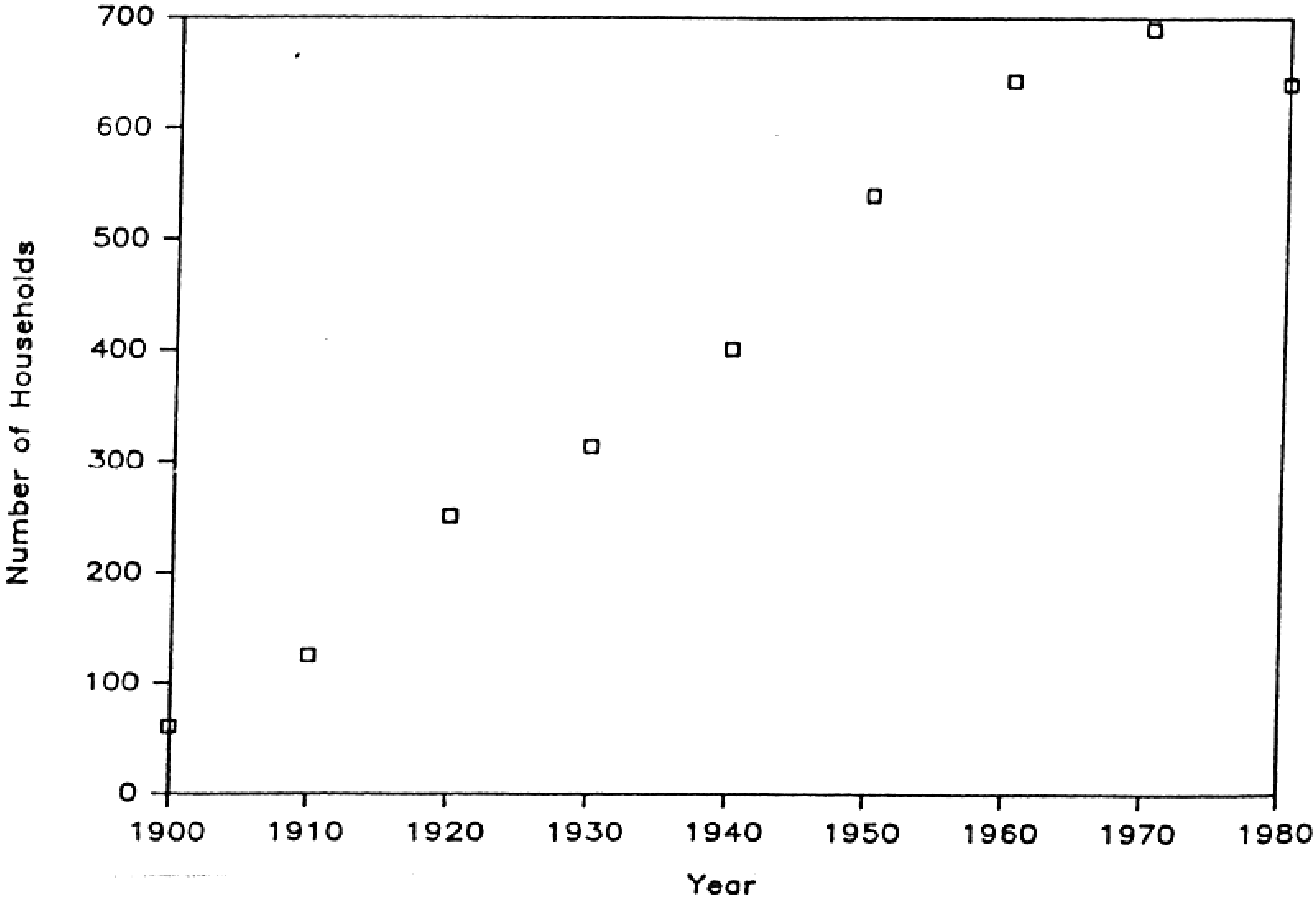


Figure 3-10 Time-Series Diagram: Number of Households Residing in Bayside, 1900–1980

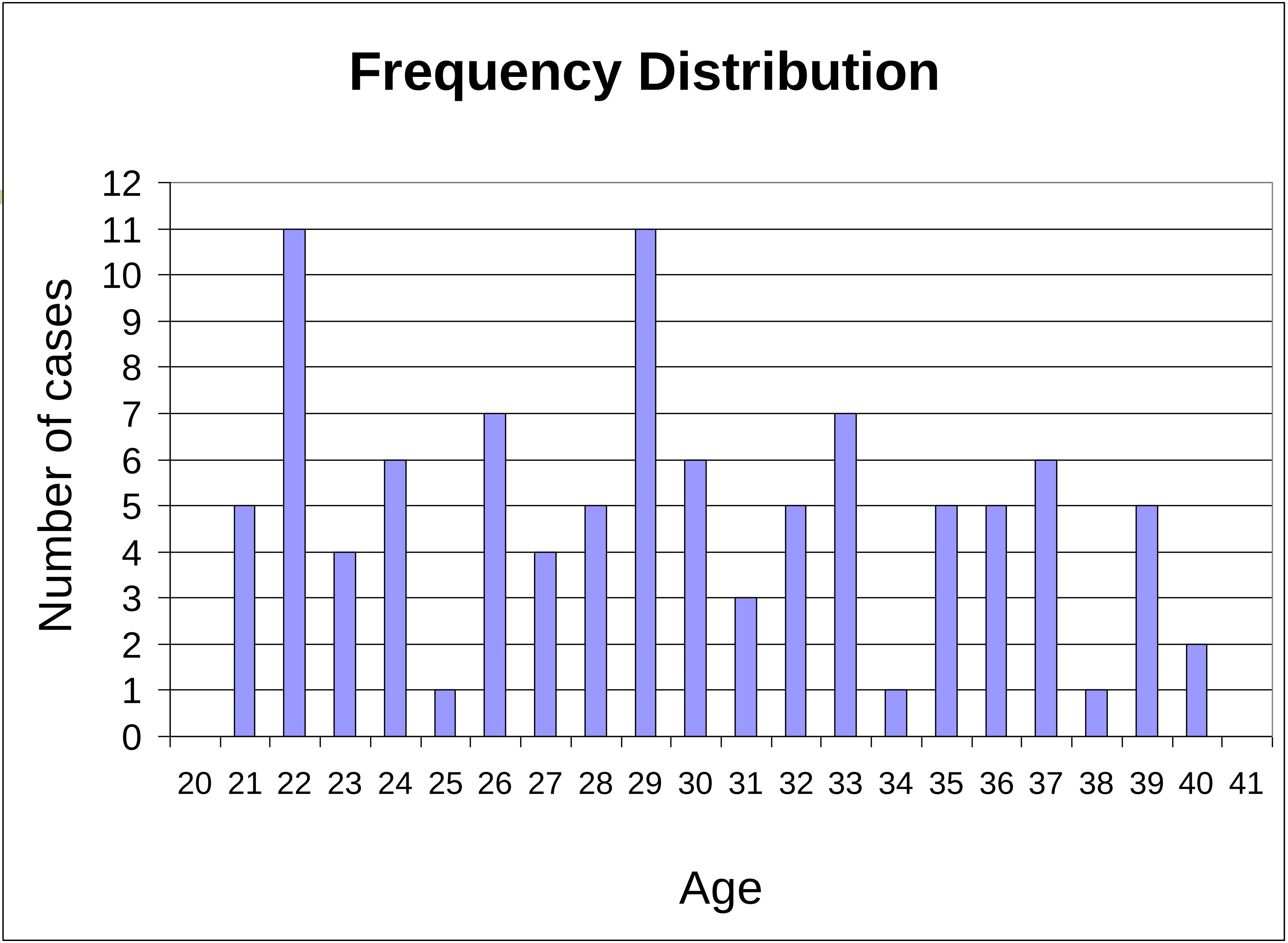


B. Inferential statistics

- What can we infer about a population given a sample size and a sample statistic?
- A population parameter is a (usually unknown) summary measure of a characteristic of a full population
- A sample statistic is a corresponding summary measure of a sample characteristic (usually known or calculated).

Let's say these are the ages of the people now in this room.

Case	Age	Case	Age	Case	Age	Case	Age	Case	Age
1	**37**	21	**28**	41	**35**	61	**37**	81	**37**
2	**23**	22	**29**	42	**24**	62	**31**	82	**27**
3	**26**	23	**38**	43	**33**	63	**32**	83	**29**
4	**35**	24	**24**	44	**30**	64	**22**	84	**40**
5	**33**	25	**37**	45	**28**	65	**25**	85	**22**
6	**28**	26	**22**	46	**24**	66	**24**	86	**27**
7	**26**	27	**32**	47	**27**	67	**29**	87	**23**
8	**21**	28	**34**	48	**29**	68	**30**	88	**30**
9	**39**	29	**29**	49	**26**	69	**32**	89	**22**
10	**27**	30	**30**	50	**35**	70	**26**	90	**33**
11	**29**	31	**36**	51	**31**	71	**24**	91	**21**
12	**37**	32	**39**	52	**26**	72	**30**	92	**23**
13	**39**	33	**29**	53	**37**	73	**22**	93	**33**
14	**36**	34	**29**	54	**21**	74	**22**	94	**26**
15	**35**	35	**35**	55	**29**	75	**21**	95	**31**
16	**39**	36	**30**	56	**36**	76	**36**	96	**23**
17	**22**	37	**28**	57	**22**	77	**26**	97	**24**
18	**40**	38	**32**	58	**39**	78	**32**	98	**33**
19	**22**	39	**22**	59	**33**	79	**29**	99	**36**
20	**40**	40	**22**	60	**28**	80	**21**	100	**33**



Frequency Distribution
Number of cases
0
1
2
3
4
5
6
7
8
9
10
11
12
20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41
Age

Basic calculations:

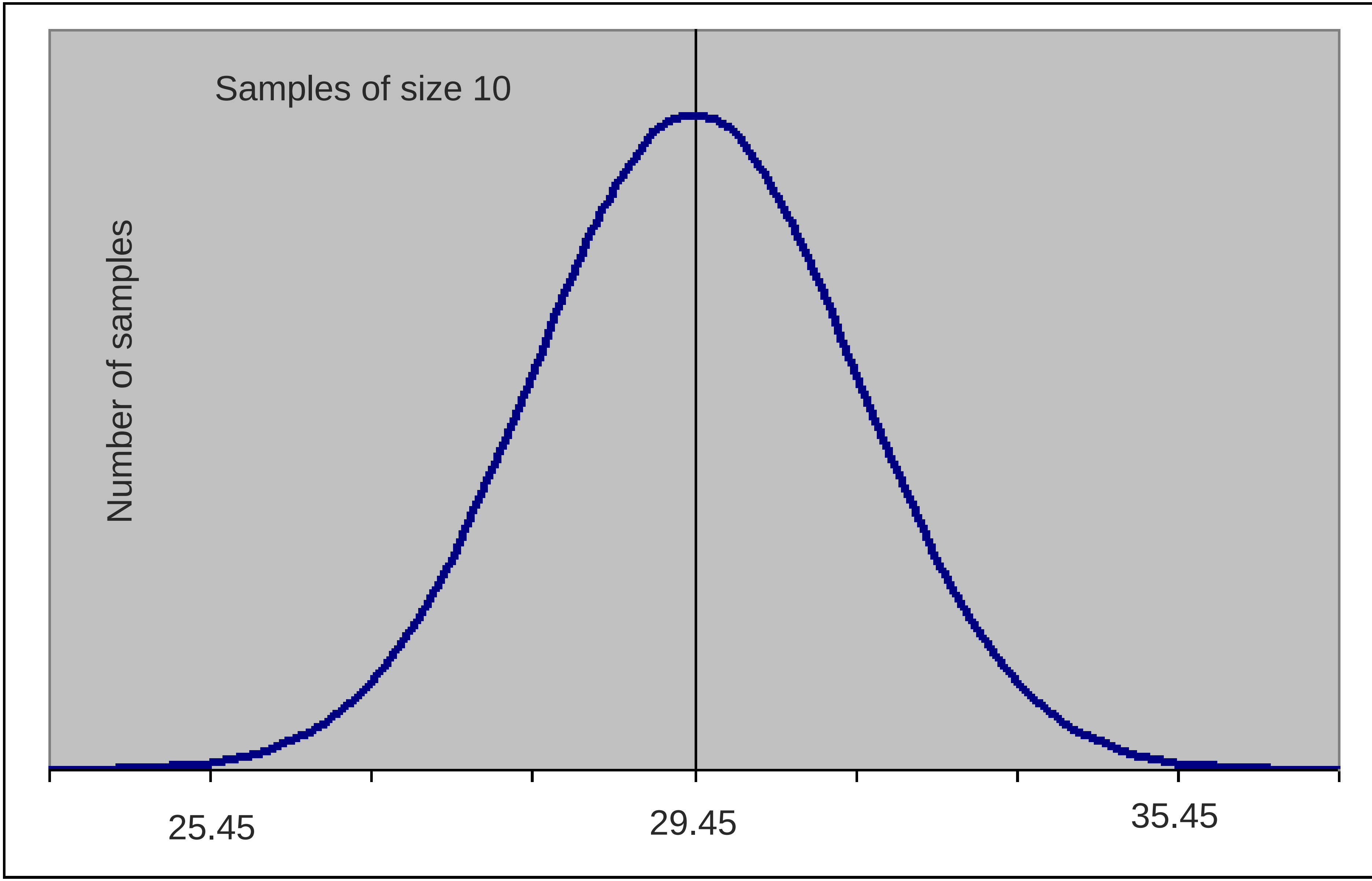
- The range is 40-21 = 19
- The average is 2945 / 100 = 29.45
- The variance is
  - 37 – 29.45 = 7.55 (difference)
  - 7.55 squared is 57.0025 (difference squared)
  - Sum all 100 differences squared and divide by 100 = 30.96
- The standard deviation is the square root of the variance = 5.56
- The cases are bimodal. 11 people are 22 and another 11 are 29.

Now, let's take a random sample of 10 cases

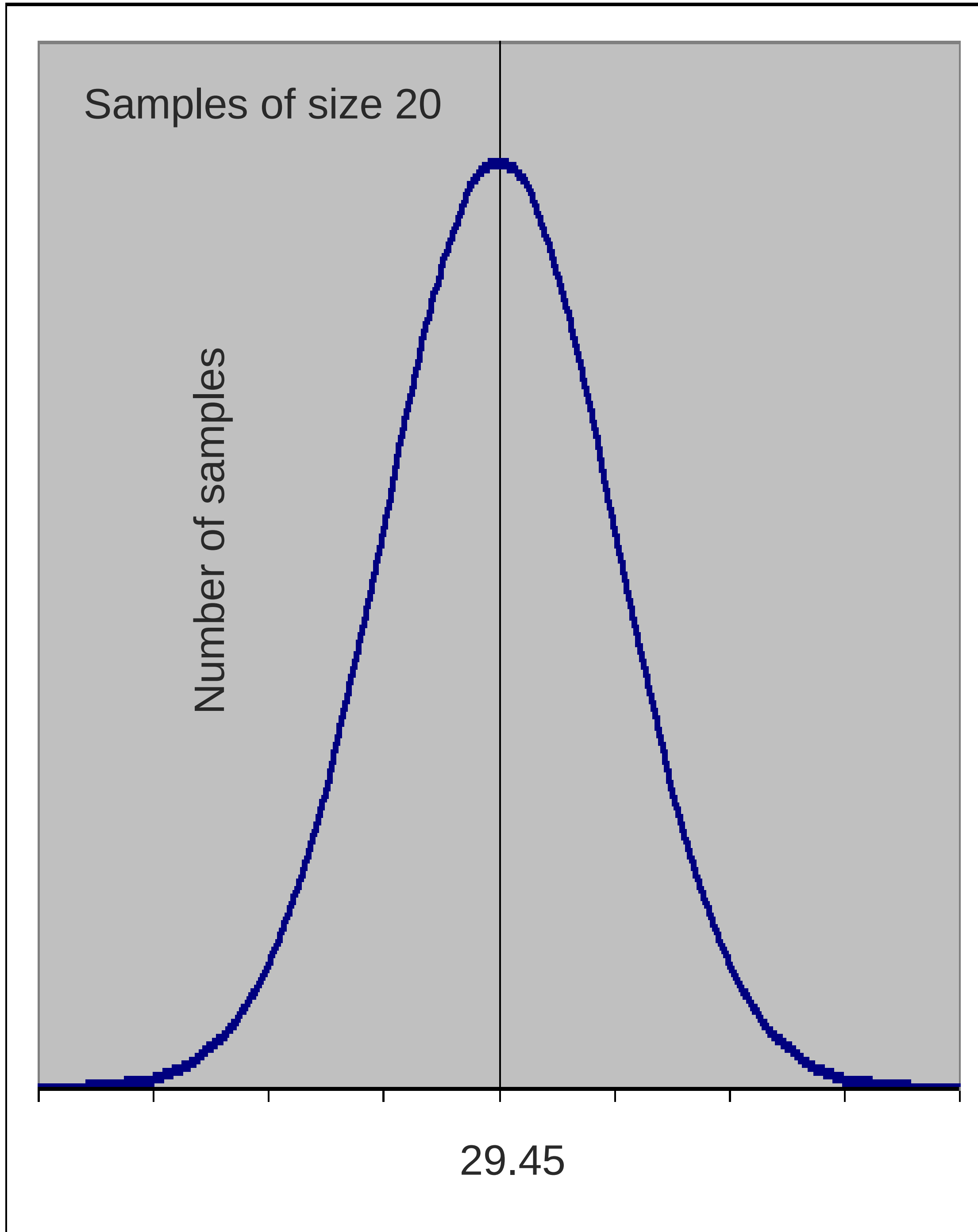
- Cases: 28, 70, 11, 81, 54, 66, 5, 6, 63, 37
- Ages: 34, 26, 29, 37, 21, 24, 33, 28, 32, 28
- The mean of these 10 cases is 29.20 but our population mean was 29.45.
- Inferential statistics help us understand how reliably a (known) sample statistic represents a (usually unknown) population parameter.

Now let's take another sample of 10, and another, and another, and …

- If we took many, many samples of 10, most would have means near 29.45, with a few much lower and a few much higher.
- Over many samples, the mean of all the samples would come closer and closer to the population mean. This is the central limit theorm.
- We can graph a frequency distribution of the mean over many samples, which is called a sampling distribution.


Samples of size 10
Number of samples
25.45
29.45
35.45

If we took samples of 20, the curve would be narrower and higher. More samples would be closer to the real population mean, and fewer would be much lower or much higher.



Sample size and confidence limits

- The standard error of the mean depends on the standard deviation of the population and the size of the sample.
  - The smaller the SD of the population, the smaller the error.
  - The larger the sample size, the smaller the error.
- Choosing an adequate sample size depends on the two factors listed above.
- You may want to be 90% certain that the mean of the sample will be within one year of the mean of the population.

Hypothesis testing

- A sample of 500 voters might show that 52% will vote for candidate X.
- That 52% could result from either
  - Random sampling fluctuation, or
  - Over 50% of all voters will really vote for candidate X
- Hypothesis testing allows us to conclude with 95% certainty, that over 50% of voters support candidate X.

C. Forecasting methods

- Intuitive methods
  - Delphi
  - Scenario writing
- Extrapolation methods
  - Assume future change of same amount added or subtracted per year (or decade)
  - Assume future change of same percentage increase (or decrease) per year (or decade, or any period)

Theoretical methods

- Dependent variable or y variable: the variable being predicted
- Independent variable(s) or x variable(s): the variable(s) used to predict
- Three methods
  - Bivariate regression (one x variable)
  - Multiple regression (two or more x variables)
  - Gravity models

Bivariate regression

- Assumes a straight line can be used to describe the relationship between the independent (x) variable and the dependent (y) variable.
- y = a + b*x
- a is the line's y intercept
- b is the line's slope
- R^2 measures how well the line fits the data and ranges from 0.0 to 1.0

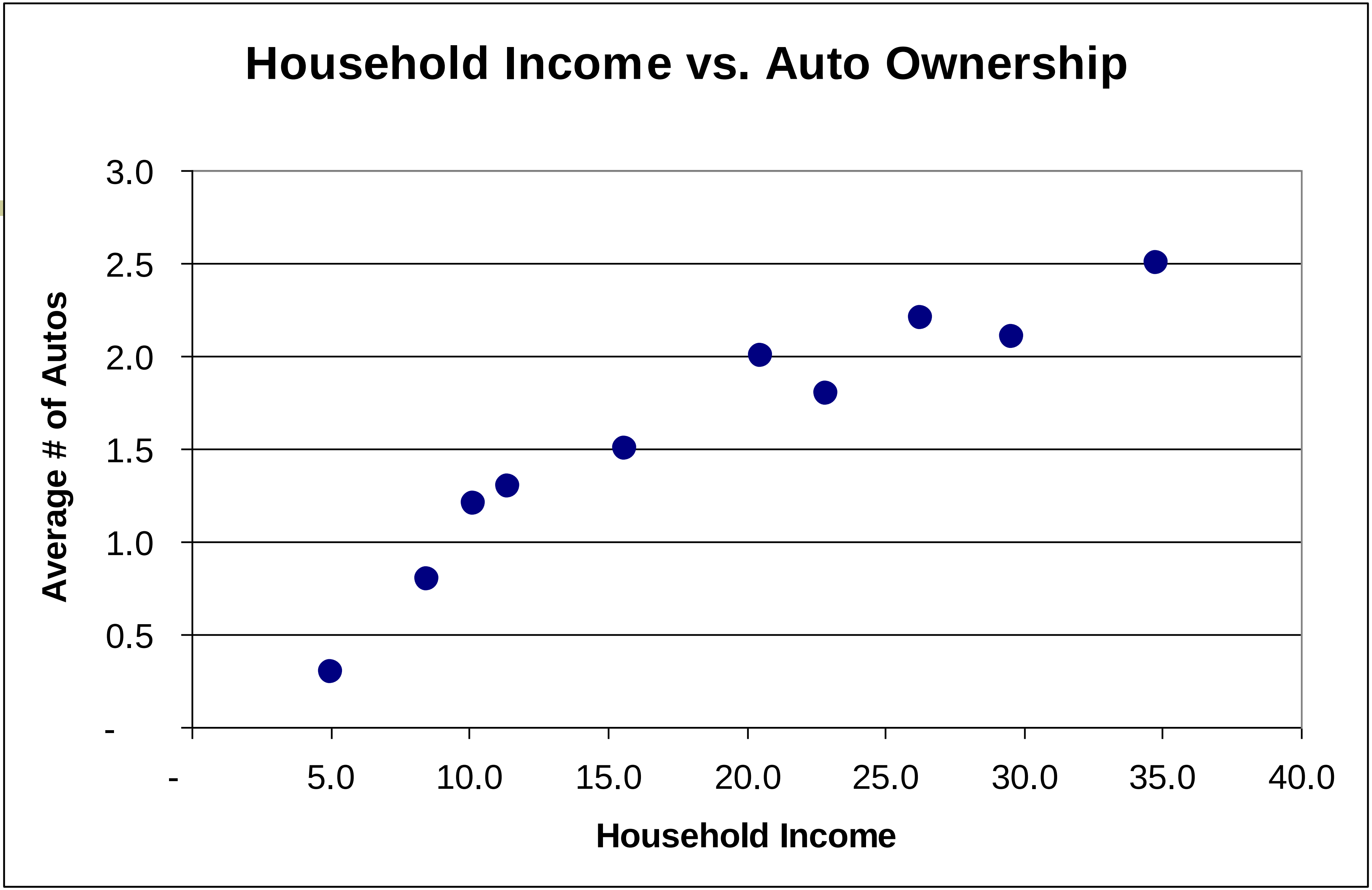
Bivariate regression

We want to predict the number of autos per household.

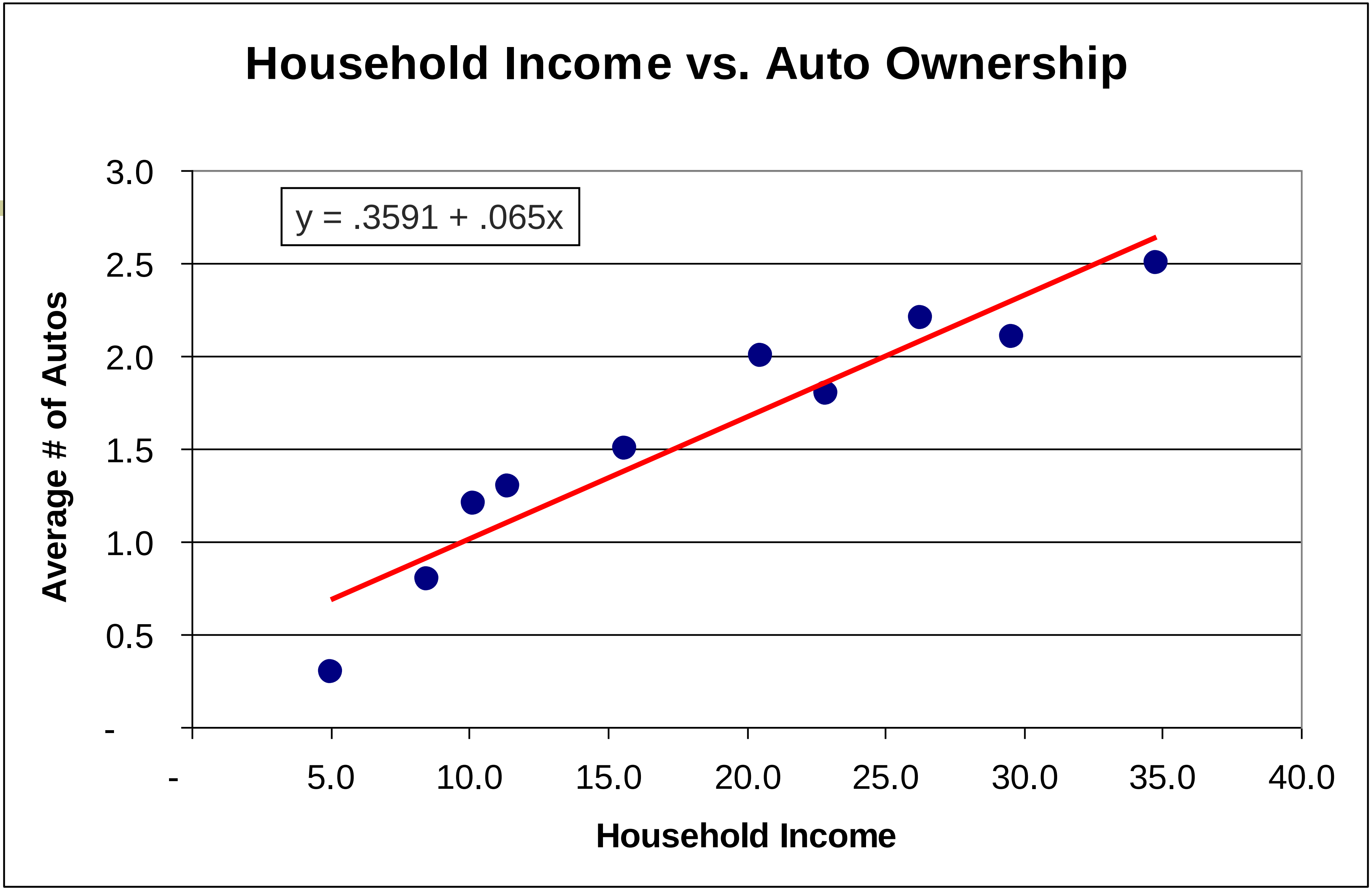
This is our data for 10 census tracts.

Income is listed in thousands of dollars.

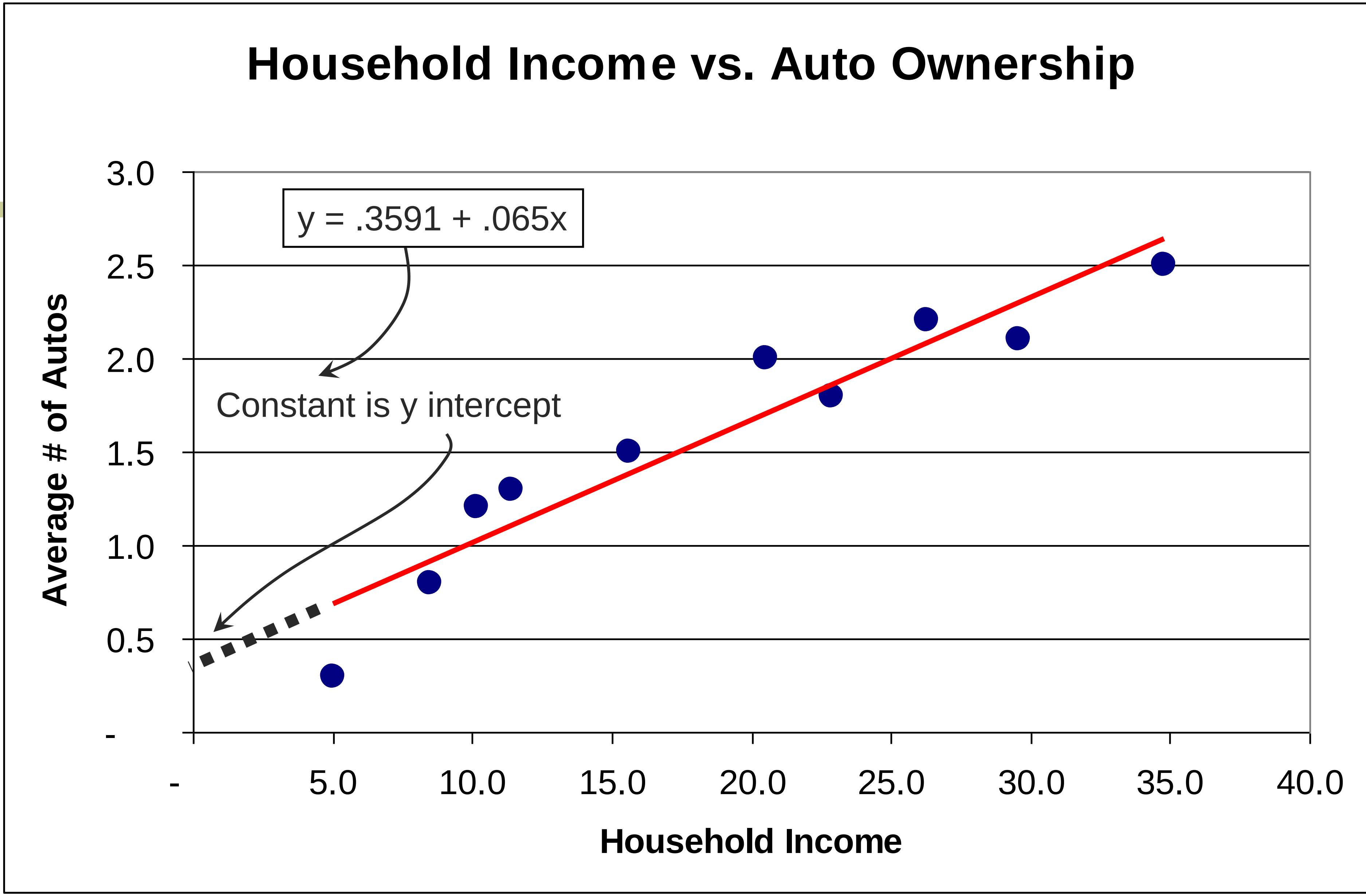
Tract	Avg HH Income	Avg # of Autos per HH
1	5.0	0.3
2	8.5	0.8
3	10.2	1.2
4	11.4	1.3
5	15.6	1.5
6	20.5	2.0
7	22.9	1.8
8	26.3	2.2
9	29.6	2.1
10	34.8	2.5

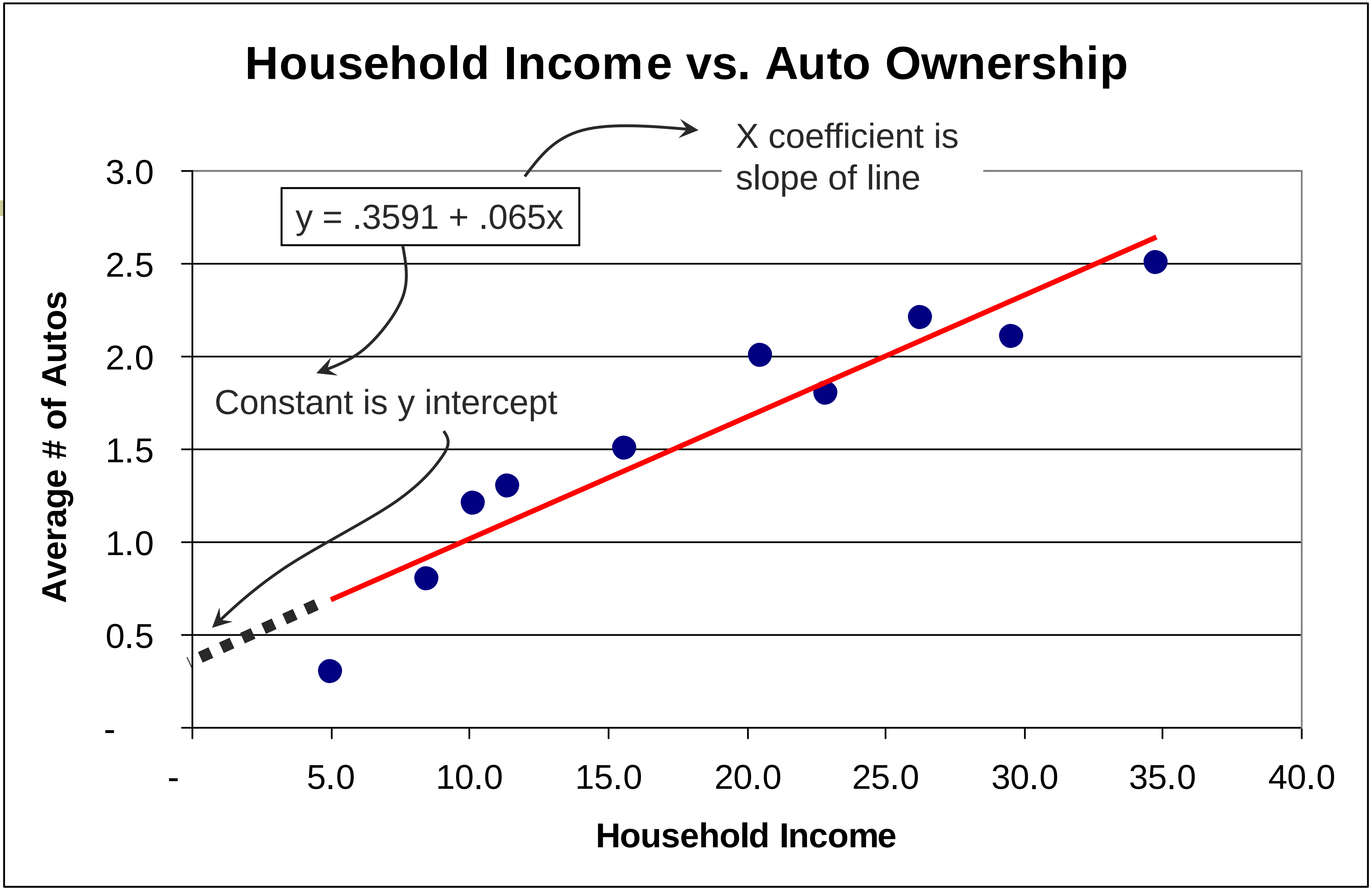


Household Income vs. Auto Ownership
Average # of Autos
3.0
2.5
2.0
1.5
1.0
0.5
-
-
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
Household Income

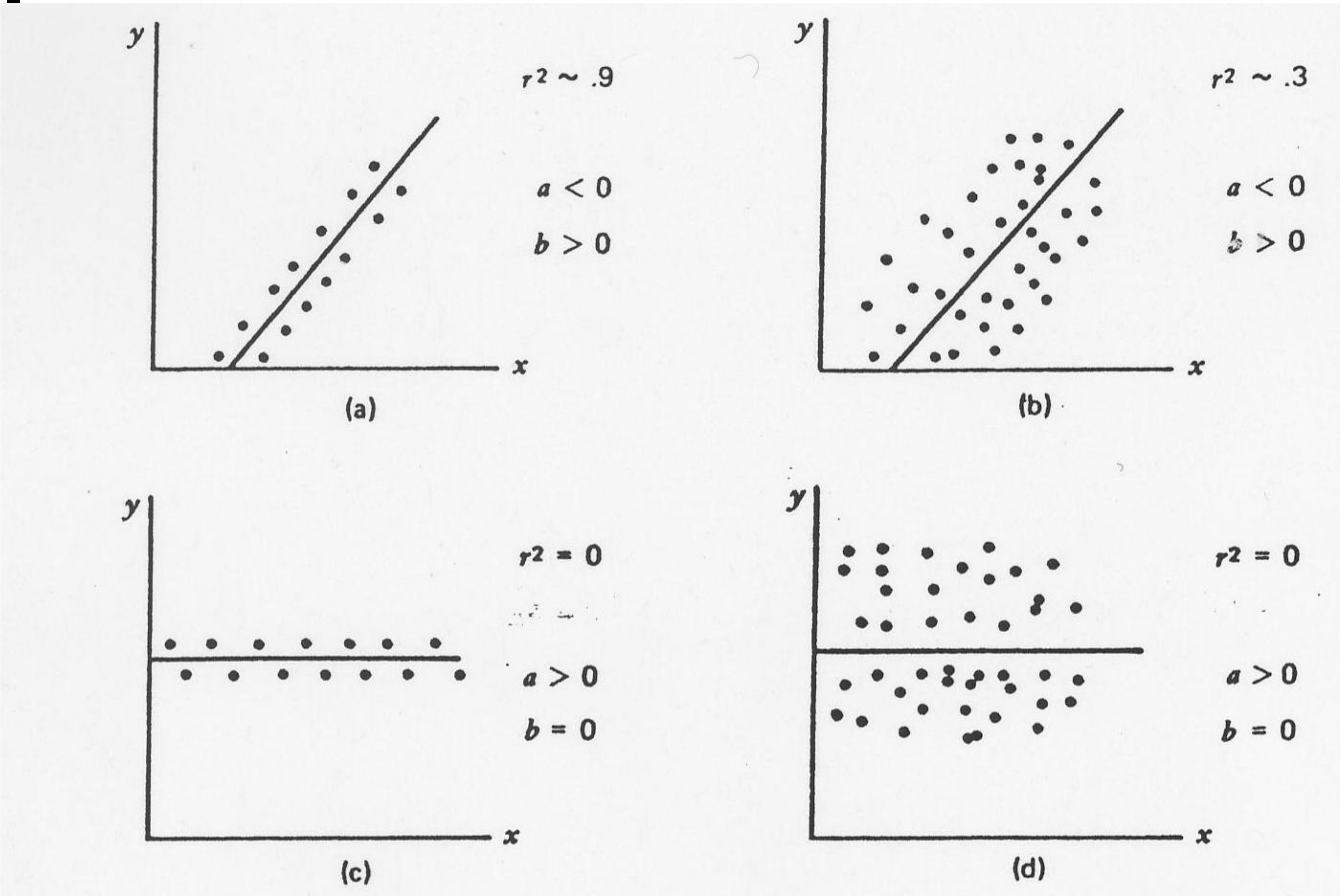


Household Income vs. Auto Ownership
y = .3591 + .065x
Average # of Autos
3.0
2.5
2.0
1.5
1.0
0.5
-
-
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
Household Income

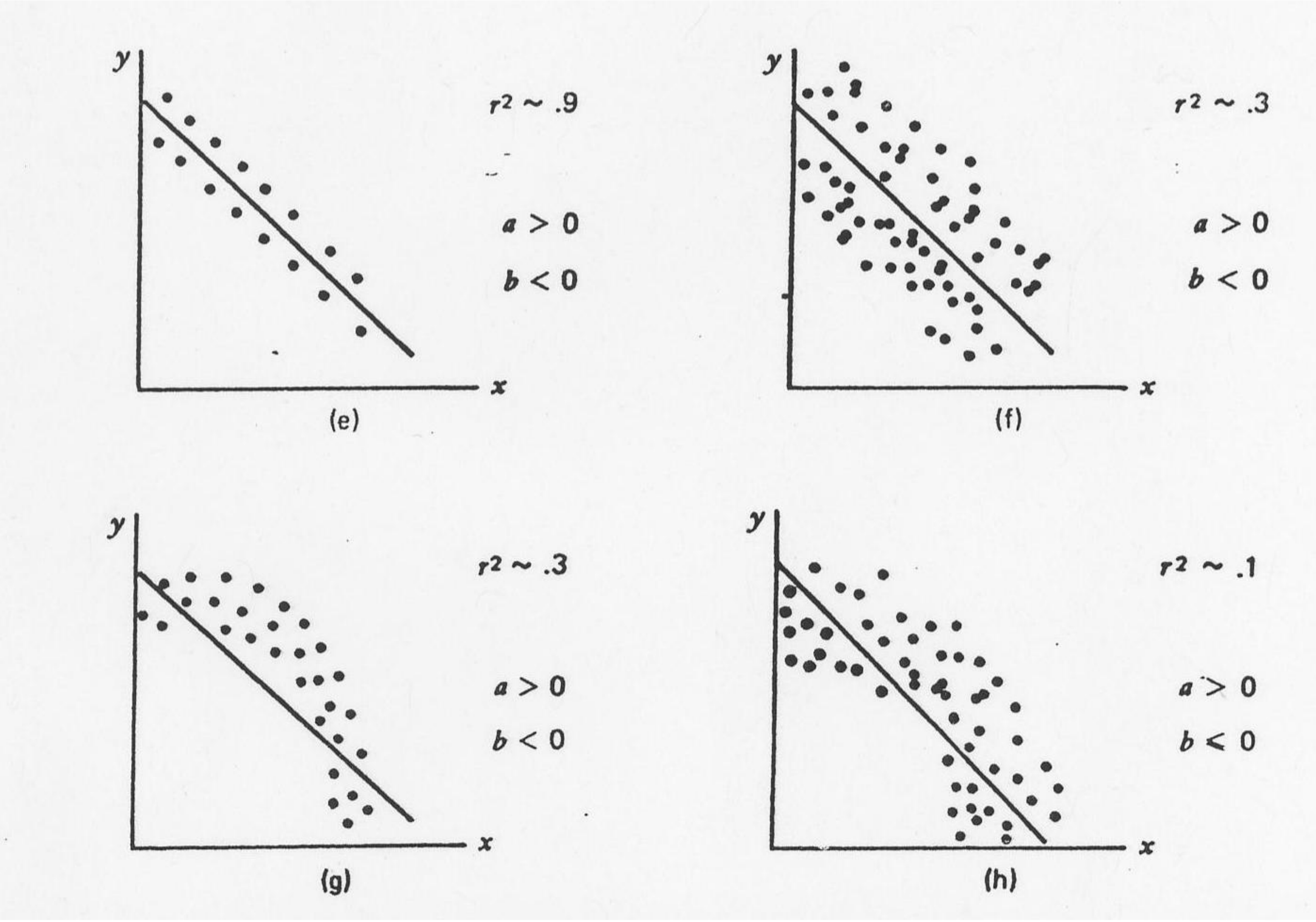

Household Income vs. Auto Ownership
y = .3591 + .065x
Constant is y intercept
Average # of Autos
Household Income
3.0
2.5
2.0
1.5
1.0
0.5
-
-
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0


Household Income vs. Auto Ownership
X coefficient is
slope of line
y = .3591 + .065x
Constant is y intercept
3.0
2.5
2.0
1.5
1.0
0.5
-
Average # of Autos
-
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
Household Income

Results of fitting regression lines to different datasets



Results of fitting regression lines to different datasets



Multiple regression uses more than one x variable

- y (house sale price) =
  x_1 * Square footage +
  x_2 * Number of bedrooms +
  x_3 * Number of bathrooms +
  x_4 * Accessibility to employment +
  x_5 * Location in historic district
- When an x coefficient is positive, higher values of x lead to higher values of y; when negative, lower

Gravity Models

$$T_{ij} = \frac{K * O_i * D_j}{F_{ij}^2}$$

- Trips from zone i to zone j =
  - A constant (K) times
  - An origin push force (population) times
  - A destination pull force (employment) divided by
  - A friction component (travel time) raised to a power (often squared)
- Total trips to one zone (j) are then the sum of trips from all origins (O_i)

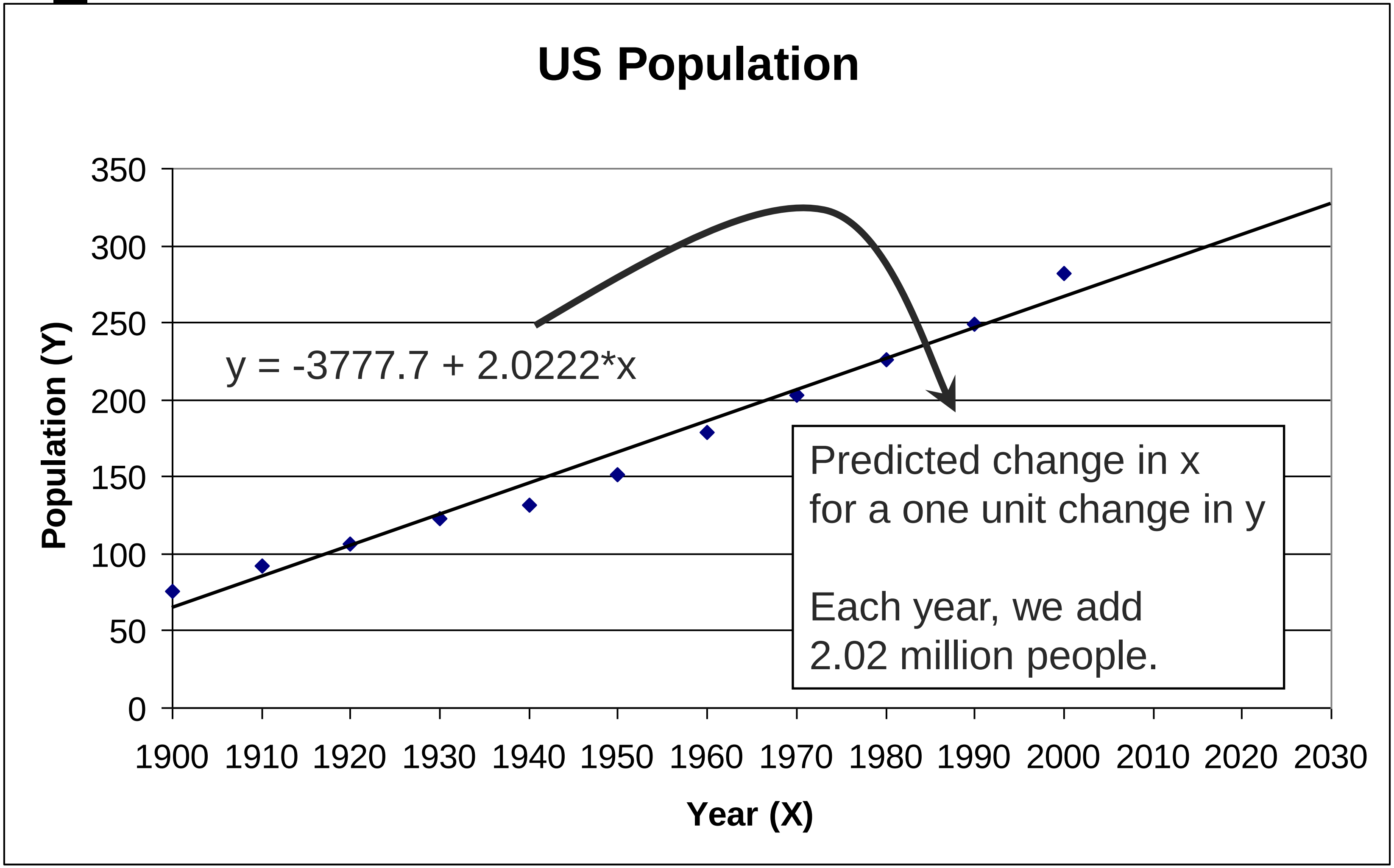
D. Population analysis and projection

- An estimate is an indirect measure of a present or past condition that can not be directly measured.
- A projection (or prediction) is a conditional statement about the future.
- A forecast is a judgmental statement of what the analyst believes to be the most likely future.

Non-component projection methods

- Extrapolation with graphs
- Time series regression, with time (year) as the independent (x) variable
- Ratio methods comparing to similar areas
- Share methods using proportions of regional or state projections

Time series regression to project US population



Cohort component models

- We divide the population into cohorts by age (five years), sex, and race/ethnicity.
- Population change is subdivided into three components: births, deaths, migrants
- Calculate birth rates, survival rates, and migration rates for a recent period
- Extend those rates into the future, possibly adjusting them upward or downward
- Birth and death data is readily available; migration data is difficult, apart from Census years.

Migration notes

- Migration can be projected as a function of changes in employment.
- Net migration = Inmigration - outmigration
- Net migration can estimated by the residual method:

  1990 population: 100,000

  2000 population: 120,000

  1990 to 2000 births: 5,000

  1990 to 2000 deaths 3,000

  How many 1990 to 2000 inmigrants? (18,000)

E: Economic analysis Economic base theory

- Assumes two kinds of industry
  - Basic or export: sells to customers outside the area of analysis
  - Service or non-basic: sells to customers within the area
- Economic base multiplier
  - Total employment / basic employment
  - A multiplier of 4.0 says that 4 total jobs are created for every additional basic job

Location quotients

- LQs compare the local concentration of employment in an industry to the national employment in that industry
- LQ_i =

$$\frac{\left(\dfrac{\text{Local employment in industry I}}{\text{Total local employment in all industries}}\right)}{\left(\dfrac{\text{National employment in industry I}}{\text{Total national employment in all industries}}\right)}$$

More on location quotients

- Alternate formula: LQ_i =

$$\frac{\text{Local percent of employment in industry i}}{\text{National percent of employment in industry I}}$$

- Interpreting LQs
  - If LQ_i is greater than 1.0 we can assume an export or basic industry
  - If LQ_i is less than 1.0 we can assume we import some goods or services
  - If LQ_i = 1.0, the region produces just enough to serve the region, and no more

Shift share analysis

- Shift share analysis interprets changes in an industry's local employment (over a period of x years) in terms of three components:
  - National share: how much would local industry employment have changed if it mirrored changes in total national employment
  - Industry mix: how much additional would it have changed if it mirrored national industry employment
  - Local shift: how many additional jobs did the local industry gain or lose, presumably due to local competitive advantage or disadvantage.

F. Project analysis and benefit cost analysis

- Many public projects have high initial costs, then produce benefits for many years.
- $1,000 of benefits in 10 years is less valuable than $1,000 of benefits this year, because we could invest today’s $1,000 and earn 10 years worth of interest.
- Discounting reduces benefits (and costs) in future years to account for the time value of money.

Since many projects are evaluated by assessing their costs and benefits over some long time period, it is essential to understand both the mechanics of discounting and the major assumptions that underlie its use. Many projects will be characterized by a series of benefits (B) and costs (C) over time:

$$B_t + B_{t+1} + B_{t+2} + B_{t+3} + \ldots + B_n$$

and

$$C_t + C_{t+1} + C_{t+2} + C_{t+3} + \ldots + C_n$$

On occasion it is simply easier to speak about *annual net benefits,* which are really yearly benefits minus yearly costs:

$$(B_t - C_t) + (B_{t+1} - C_{t+1}) + (B_{t+2} - C_{t+2}) + (B_{t+3} - C_{t+3}) + \ldots + (B_n - C_n)$$

Any one of these yearly figures could be negative as well as positive. A typical public investment project ordinarily shows higher costs than benefits in the early years and higher benefits than costs in later years—or negative net benefits early and positive net benefits later. The one instance when you would not wish to compute net annual benefits before discounting would be if you wished to calculate a benefit-cost ratio. To do that, benefits and costs must be discounted separately. We will show this below.

-

Annual benefits

Figure 7–7 Basic Data for the Discounting Example

	YEAR					
	0	1	2	3	4	5
Benefits	0	\$4000	\$4000	\$4000	\$4000	\$4000
Costs	\$15,000	0	0	\$1223	0	0
Discount rate (r)	4%	4%	4%	4%	4%	4%
Discount factor $1/(1 + r)^n$	1.0	.9615	.9246	.8890	.8548	.8219

Initial construction cost

Year 3 maintenance cost

The most commonly encountered measure of the efficiency of a public investment project is its present worth, or net present value (NPV). NPV is the sum of all discounted benefits and discounted costs for the duration of the project. Thus:

$$NPV = (B_t - C_t) + \frac{B_{t+1} - C_{t+1}}{(1 + r)^1} + \frac{B_{t+2} - C_{t+2}}{(1 + r)^2} + \ldots + \frac{B_n - C_n}{(1 + r)^n}$$

Let us illustrate the discounting procedure with an example. A project costs \$15,000 to implement immediately (we won't discount, and we will call this time zero). The project returns \$4000 per year in benefits every year thereafter for the next five years and has only one remaining cost, a maintenance charge of \$1223 in year three. The pattern of costs and benefits is shown in Figure 7-7

Using, for example, a 4% discount rate, we can compute the net present value several ways. We will do it first by treating the cost and benefit streams separately, computing the discount factors using the formula $1/(1 + r)^n$, which yields a discount factor of .9615 for year one and .9246 for year two, e.g. $1/(1 + .04)^1$ and $1/(1 + .04)^2$.

Discounted benefits (DB)
$= 0(1.0) + 4000(.9615) + 4000(.9246) + 4000(.8890) + 4000(.8548) + 4000(.8219)$
$= 4000(4.4518) = \$17,807.20$

Discounted costs (DC)
$= -15,000(1.0) + 0(.9615) + 0(.9246) + (-1223)(.8890) + 0(.8548) + 0(.8219)$
$= -15,000 + (-1223)(.8890) = -15,000 + (-1087.25) = -\$16,087.25$

Net present value thus is the simple sum of discounted benefits and costs. Note that for bookkeeping purposes, costs are expressed negatively. This convention can be helpful with larger problems that involve more extensive sets of figures.

$$
\begin{aligned}
NPV = DB + DC &= \$17{,}807.20 + (-\$16{,}087.25) \\
&= +\$1{,}719.95
\end{aligned}
$$

Another method used to compute net present value first calculates the net annual benefits (benefits minus costs) and then discounts this number. In our example above this would be:

$$
\begin{aligned}
NPV = {} & (0 - 15{,}000)(1.0) + (4000 - 0)(.9615) \\
& + (4000 - 0)(.9246) + (4000 - 1223)(.8890) \\
& + (4000 - 0)(.8548) + (4000 - 0)(.8219) \\
= {} & -15{,}000 + 4000(.9615) + 4000(.9246) + 2777(.8890) + 4000(.8548) \\
& + 4000(.8219) \\
= {} & -15{,}000 + 3846.00 + 3698.40 + 2468.75 + 3419.20 + 3287.60 \\
= {} & -15{,}000 + 16{,}719.95 = +\$1719.95
\end{aligned}
$$

1. If NPV is positive, we should undertake the project.
2. Benefit cost ratio = 17,807.20 / 16,087.25 = **1.107**
   Begin with the projects with the highest BC ratios.

AICP Exam Review
Planning Methods Blitz

Study hard, and

Good luck on the exam!

http://drummond.gatech.edu/aicpexam.ppt