Regression Analysis of Ship Characteristics

Occasional Paper

This Paper presents the results of statistical analyses of ship characteristics which have been undertaken to provide input to models of ship costs and operations in particular trades. Standard least squares regressions were performed on the data to relate particular ship characteristics to deadweight. Deadweight was selected as the common denominator for the regressions because of its universal acceptance as a measure of ship size and because of its wide use in the reporting of statistical information.

Regression Analysis of Ship Characteristics

G. P. Piko

AUSTRALIAN GOVERNMENT PUBLISHING SERVICE
CANBERRA 1980

ISBN 0 642 05820 2

Printed by C. J. THOMPSON, Commonwealth Government Printer, Canberra

FOREWORD

This paper has been published to disseminate information on statistical analyses of ship characteristics undertaken by the Bureau of Transport Economics. While information of a similar nature has been published, it is considered that the comprehensiveness of the results presented in this paper will be of general interest.

The analyses were undertaken to provide input to supply oriented studies into port infrastructure and port and shipping operations. Because of this application the characteristics:

- length;
- breadth;
- draught; and
- container capacity;

which all influence port infrastructure and operating costs, have each been modelled as a function of deadweight tonnes; the nearest measure to the economic usefulness of a ship. In addition, the following characteristics have each been regressed against deadweight:

- gross registered tons;
- net registered tons;
- age;
- power; and
- speed.

Standard minimum least squares regression techniques were used to generate models for each of the above characteristics for each of the following ship types:

- container ships;
- roll on-roll off ships (ro-ro);
- bulk carriers;

. ore carriers;
. tankers;
. general cargo ships; and
. passenger ships.

Computer data tapes of Lloyd's Register of Shipping (current at May 1977) were used as the data source for the analyses. Hence, the ships on which the models were based are representative of the world fleet. Details of the data extraction, which required considerable computing effort, are also described in the paper.

The analyses were performed and the report prepared by Mr G.P. Piko of the Planning and Technology Branch under the general supervision of Mr C.R. Sayers.

R.W.L. Wyers
Assistant Director
Planning and Technology

Bureau of Transport Economics
CANBERRA
December 1980

CONTENTS

TABLES

<u>FIGURES</u>

CHAPTER 1 - INTRODUCTION

The Bureau of Transport Economics (BTE) is currently undertaking a number of shipping and port facility studies which require the development of models of ship costs and operations in particular trades. This paper presents the results of statistical analyses of ship characteristics which have been undertaken to provide input to these models.

The source of the data used in this project was Lloyd's Register of Shipping, which consists of detailed information on over 60 000 ships and is considered to be representative of the world fleet. Lloyd's Register of Shipping records a comprehensive range of data items. However, for some ships information is not presented for every data item. The Register of Shipping was obtained in the form of magnetic tape for computer analyses: the version used for this project was current at May 1977.

Standard least squares regressions were performed on the data to relate particular ship characteristics to deadweight. Deadweight was selected as the common denominator for the regressions because of its universal acceptance as a measure of ship size and because of its wide use in the reporting of statistical information. The ship characteristics which were each regressed against deadweight were:

- length;
- breadth;
- draught;
- gross registered tons (GRT);
- net registered tons (NRT);
- age;
- power;
- speed; and
- container capacity.

The length, breadth, draught and container capacity of ships are of interest because they each influence the port facilities required to handle a given ship. The power and speed of ships affect such factors as fuel consumption and travel time, and therefore are of interest in many shipping studies.

Regressions on age are included to provide an indication of the trend over time in the size of ships built. However, ships are deleted from Lloyd's Register as they are withdrawn from service, hence, the data consist of the ships still in service at May 1977 rather than all the ships built before that time. This means that the data do not necessarily provide an accurate description of the range of sizes of ships built in any given year.

Deadweight is a more common measure of the carrying capacity of a ship than either gross registered tons (GRT) or net registered tons (NRT), hence, deadweight has been used as the independent variable. However, regressions have been carried out which provide an indication of the relationship between deadweight and GRT, and deadweight and NRT.

Regressions on the ship characteristics described above were performed separately for seven ship types:

- container (fully cellular);
- roll on-roll off (ro-ro);
- bulk carrier;
- ore carrier;
- tanker;
- general cargo; and
- passenger.

The variation in the characteristics of ships from these different categories is sufficiently great to justify analysing each ship type separately. Also, most port studies, including those of the BTE, need to consider separately the facilities for various ship types.

While the major aim of this project was to provide a data base which satisfied the requirements of the dependent BTE shipping and port projects, an endeavour has been made to present the results in a format suitable for use by others working in related fields. The presentation of results includes tabulation of the regression coefficients and statistics for each model. However, to further assist interpretation of the results a series of figures is included which shows each regression model, together with its 95 per cent confidence interval, superimposed on a plot of the sample data. These figures not only illustrate the relationship derived between each variable and deadweight, but also provide other valuable information including:

- the spread of the sample data on which the regression model was based;
- the range of the variables over which the regression is valid;
- the width of the 95 per cent prediction confidence intervals; and
- the upper limit of the variable that occurs in the sample data.

The information presented in this paper is intended to assist those examining investment in shipping and port infrastructure. For example, for a given increase in the depth of a channel a regression equation will provide an indication of the maximum deadweight of ships (of a given type) that would then be able to use the channel. The plot of the sample data would indicate whether there are, in fact, many ships of that size in service, and therefore the likely demand for additional channel depth.

CHAPTER 2 - DESCRIPTION OF THE DATA AND OUTLINE OF THE ANALYSIS

Lloyd's Register of Shipping has been published for many years. It is constantly updated, newly built ships being added and ships withdrawn from service being deleted, in order to maintain an accurate description of the vessels in service around the world. The version of Lloyd's Register purchased by the BTE was current at May 1977 and is now held by the Department of Transport as part of the sea transport information system.

Lloyd's Register attempts to provide extremely comprehensive data on each ship. The full range of specifications sought for inclusion in the Register is shown in Annex A, however, the information on each ship falls broadly into the following categories:

- name;
- owner and manager's name;
- detailed specification of the hull;
- detailed specification of the equipment and propulsion systems;
- ship type;
- ship builder;
- engine builder; and
- country of registry.

This information is stored as a series of computer card images which are each identified by a card type code and a sequence number. The card type code indicates the ship characteristics that are to be found on a particular card. In the event of the information for a given ship characteristic filling a number of cards, each card is distinguished by a sequence number. Annex A shows the characteristics recorded in the Register and the code number of the card on which each is to be found. The file is structured so that all the information for the first ship is recorded on a series of these cards which is then followed by all the cards for the second ship, and so on.

Most data items are recorded in a clearly defined numeric or alphanumeric field, which makes computer analysis convenient. An example of a typical fixed format card of data is shown in Figure 1. Characters 4-10 contain a Lloyd's Register number which identifies the ship, and the data items are recorded in fixed locations after the card type code. Occasionally, however, an item of data is recorded in a field without fixed content or format. An example of a variable format field is shown in Figure 2. It is difficult to extract the data from a field of this type because the information does not always occur at the same position on the card.

The ship characteristics extracted from the file for analysis in this project were:

- deadweight (DWT)- the weight in tonnes of cargo, stores, fuel, passengers and crew carried by the ship when loaded to the summer loadline;

- length - the extreme length of the ship;

- breadth - the extreme breadth, which is the maximum breadth; to the outside of the ship's structure

- draught - in most cases this represents the summer loadline draught amidships, but in some ships the maximum draught is at the aft end and then this figure is recorded;

- gross registered tons (GRT) - the capacity in cubic feet of the spaces within the hull and of the enclosed spaces above the deck, all divided by one hundred;

- net registered tons (NRT) - derived from the gross tonnage by deducting spaces used for the accommodation of the master, officers, crew, navigation, propelling machinery and fuel;

- age - years since the ship was built (to the time of publication of the data, 1977);

- power - total brake or shaft horsepower depending on propulsion type (it was assumed shaft horsepower equals brake horsepower);

- speed - the speed (in knots) that the ship is stated to be capable of maintaining at sea in normal weather, and at normal service draught; and

- container capacity - capacity is measured in twenty foot equivalent units (TEUs).

The degree of accuracy for individual items of information depends on the availability of the data. Certain items of information are recorded for all ships, while other items are only available for some ships. In general, the more fundamental the item the better the coverage. Lloyd's Register provides a scale which indicates the overall quality of the data recorded for each item of information. There may be variations within ship types and within year of build, but the scale gives an indication of the overall reliability of an item of information. The graduations of the scale are:

- very good;
- good;
- acceptable for analysis purposes;
- should be treated with caution if used in analysis; and
- poor.

The data for deadweight and speed are described as being 'acceptable for analysis purposes'. The data for all other quantities used in this project were described as either 'good' or 'very good'. An attempt was made to further ensure the reliability of the data used by only including a ship in the analysis if figures were available for all the ten characteristics taken from the Register (except TEUs which are only relevant to container ships), i.e. ships with incomplete data were excluded. By using only ships for which all the relevant information was available, the sample was restricted to ships for which the data were likely to be most reliable.

The sample was further restricted by excluding ships which belong to more than one of the ship type categories listed above, i.e. only ships which operate in purely one fashion were considered in the analysis. This restriction is intended to normalise the data to consistent design criteria. It would be unreasonable to expect, for example, that fully cellular container ships exhibit the same characteristics as ships which handle both general cargo and containers. These latter ships are likely to be general cargo ships, designed to different criteria, which enable them to carry some containers.

The sample sizes which remained for analysis were:

289	container ships
107	ro-ro ships
2462	bulk carriers
277	ore carriers
3014	tankers
4146	general cargo ships
39	passenger ships
10334	total

The nature of the original data set and the culling criteria used to derive the analysis data must affect the relationships obtained. Account should be taken of the extent to which the composition of the sample data is likely to affect its relevance to any particular application. In summary, the principal features of the sample data are:

. representative of the world fleet;

- contains those ships in service at May 1977;

- contains only ships which operate in purely one fashion; and

- ships which do not contain all relevant data items have been excluded.

The data files were very large, therefore, the first step of the analysis was to delete all the ships that were not relevant so that the subsequent data preparation was performed on much smaller files which were simpler and cheaper to manipulate. The data preparation was performed using three computer programs: DATAPREP1, DATAPREP2 and DATAPREP3. Figures 3(a), 3(b) and 3(c) illustrate the major functions of these programs in diagrammatic form.

Broadly, the first program locates those ships that are to be included in the sample and creates seven output files which each contain the information for one ship type. The second program locates the specific information to be analysed for each ship and performs some checks on the data. The third program carries out further checks on the data, converts it to a form suitable for regression analysis and creates seven files for input to the statistical package. These three programs and the analysis performed are described in more detail in Annex B.

Standard linear regression techniques were then used to determine the line of best fit for each set of data. All regressions were performed on the sample sizes shown above.

The output presented for each regression consists of a point plot of the sample data with the regression relationship and 95 per cent prediction confidence interval superimposed. These plots are shown in Figures 4-67. The point plots are generated such that an asterisk signifies one data point; a '2' denotes two coincident data points etc.; and a '9' denotes nine or more coincident data points.

Tables 1-9 contain information on the form of the regression model recommended in each instance, together with the regression estimates of the coefficients and the 't' and 'R^2' statistics. For large samples, if the absolute value of the t-statistic is greater than 1.96 then the regression coefficient is significantly different from zero, i.e. that term in the regression is making a significant contribution in explaining the variance of the sample data. The R^2 statistic represents the proportion of the variance in the data that can be explained by the regression model. However, the R^2 statistic cannot be compared from one regression model to another.

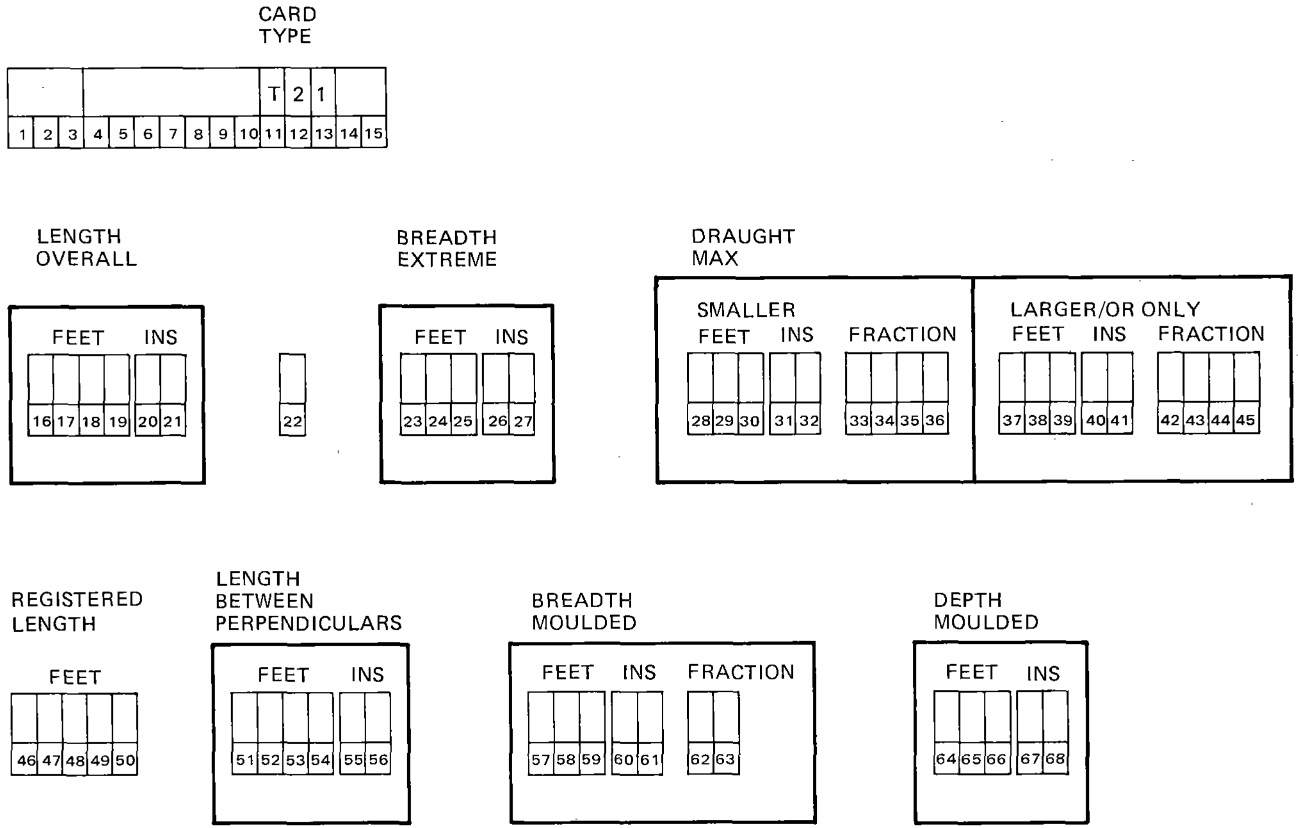


FIGURE 1 FIXED FORMAT CARD IMAGE

SOURCE: Documentation of Lloyd's Register Book File of Shipping.

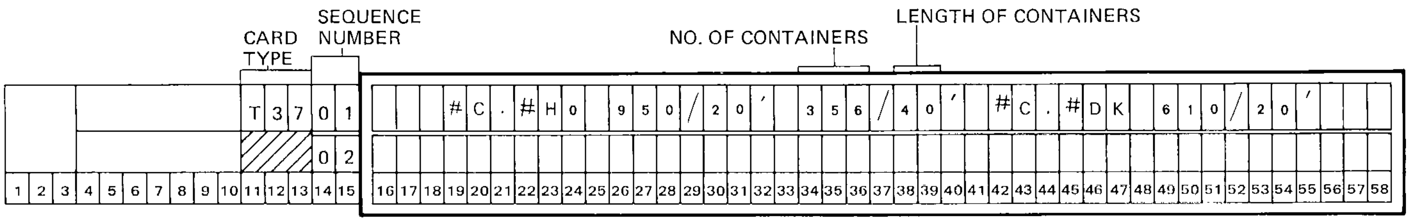


FIGURE 2 VARIABLE FORMAT CARD IMAGE

SOURCE: Documentation of Lloyd's Register Book File of Shipping.

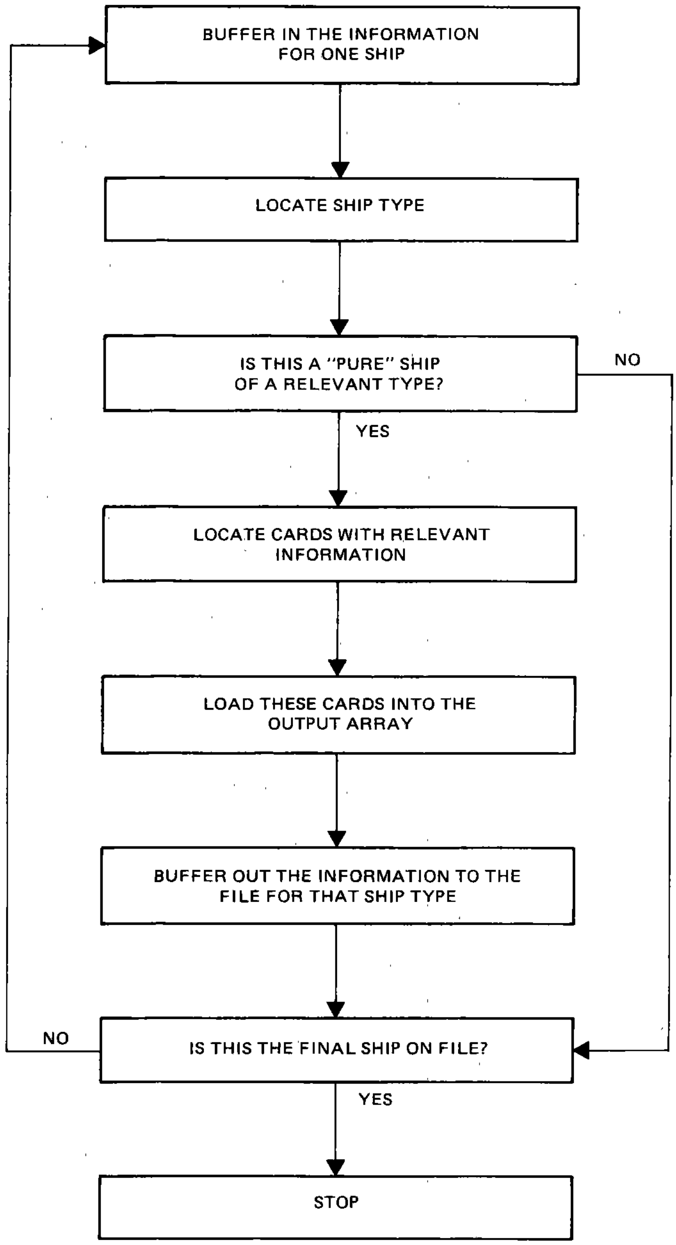


FIGURE 3(a) PROGRAM DATAPREP 1

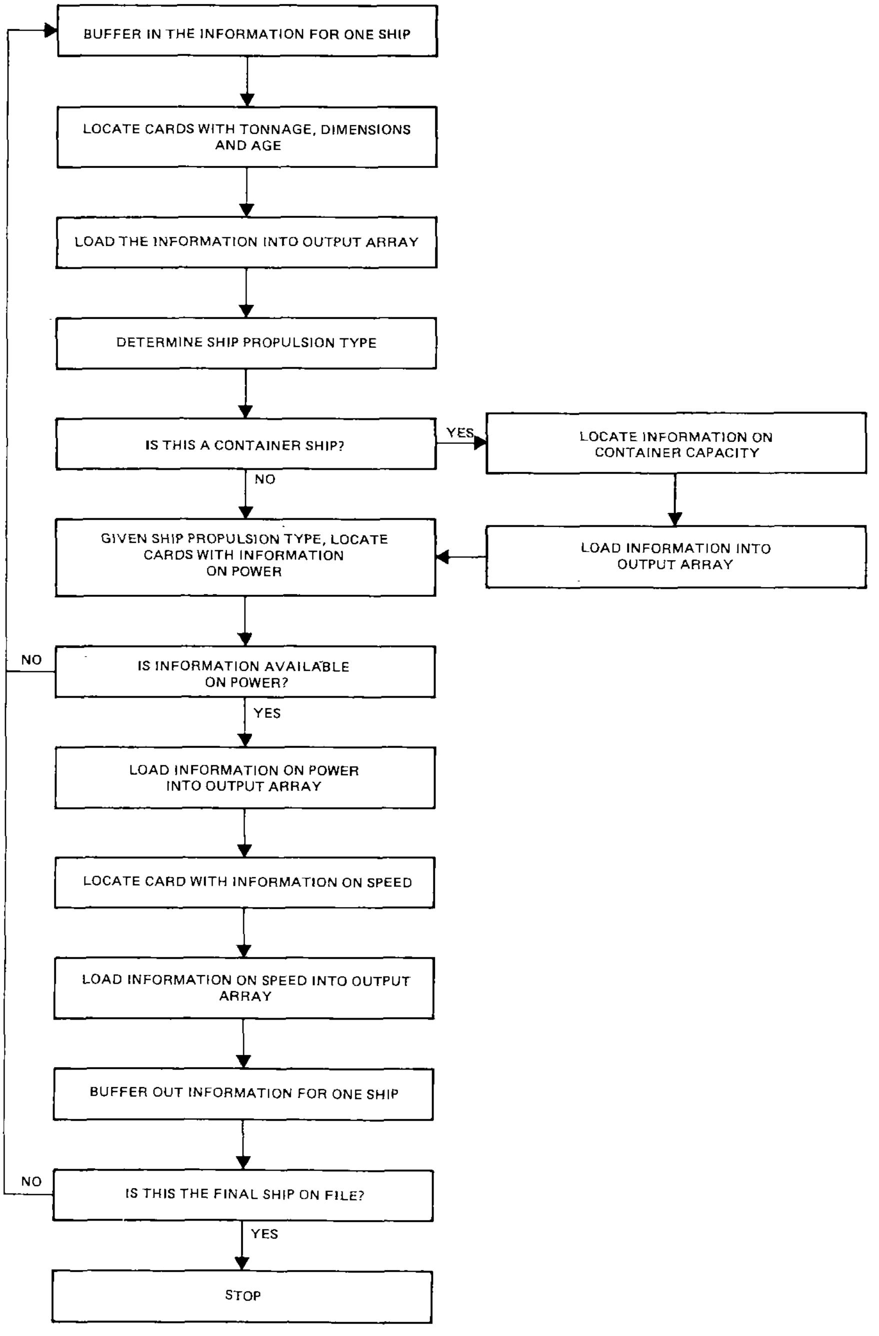


FIGURE 3(b) PROGRAM DATAPREP 2

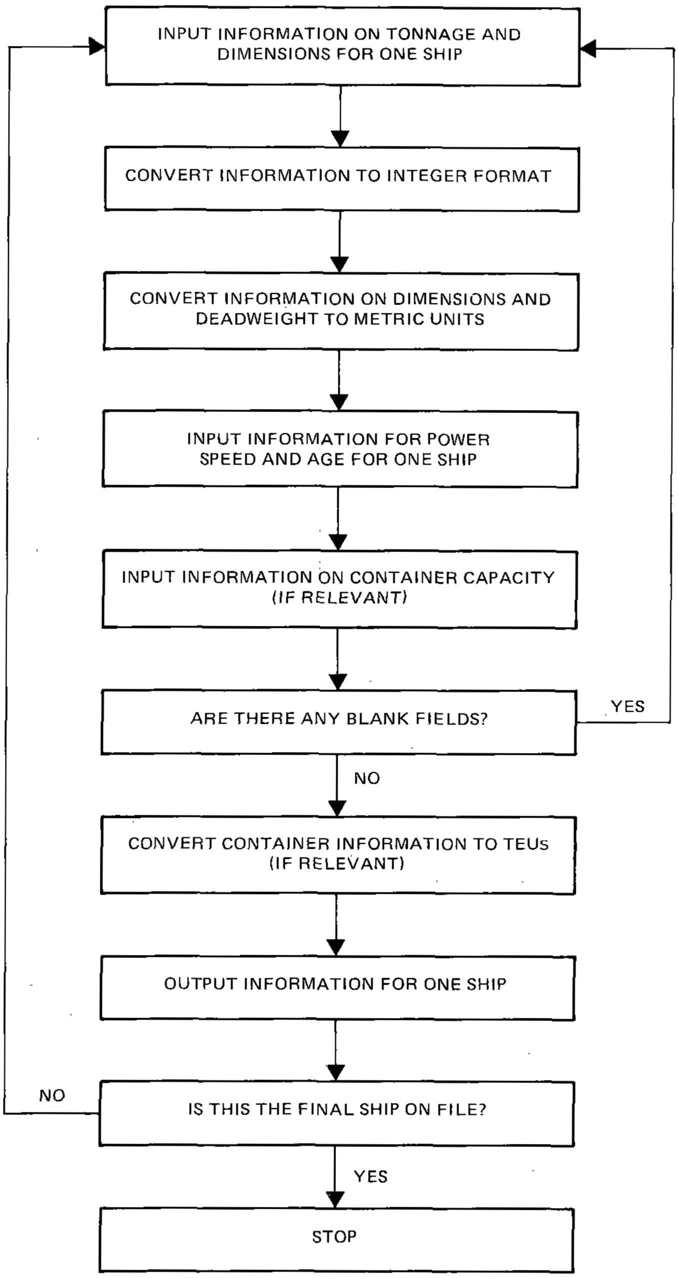


FIGURE 3(c) PROGRAM DATAPREP 3

CHAPTER 3 - DISCUSSION OF RESULTS

LENGTH, BREADTH AND DRAUGHT

Figures 4-24 show length, breadth and draught plotted against deadweight for each of the seven ship types. Table 1-3 present the regression coefficients and statistics for the models:

- Length $= \alpha.(DWT)^{\beta}.\varepsilon$
- Breadth $= \alpha.(DWT)^{\beta}.\varepsilon$
- Draught $= \alpha.(DWT)^{\beta}.\varepsilon$

The summary statistics in Tables 1-3 show that these models do, in general, indicate a strong correlation between each dimension and deadweight for each of the ship types considered. Although it was clear from the graphs of the sample data that a model of the form shown above should fit the data, it was also possible that a quadratic function may be appropriate. Hence, models of the following form were also regressed:

- Length $= \alpha + \beta.(DWT) + \gamma.(DWT)^2 + \varepsilon$
- breadth $= \alpha + \beta.(DWT) + \gamma.(DWT)^2 + \varepsilon$
- Draught $= \alpha + \beta.(DWT) + \gamma.(DWT)^2 + \varepsilon$

Although the t-statistics suggested the coefficients were statistically significant (at the 95 per cent confidence level) they were very small and had little influence on the estimates of length, breadth and draught. It was considered that the former functions provided accurate estimates of the parameters over a wider range of deadweight and, therefore, they are the recommended relationships. (The regression coefficients and statistics of the quadratic functions are recorded in Tables C.1 to C.3 in Annex C).

There are a number of physical relationships which determine the length/breadth and length/draught ratios for a ship which minimise the resistive force a ship experiences as it progresses through the water. However, the relationships derived between deadweight and length, deadweight and breadth, and deadweight and draught do not reflect any such physical relationship. Rather, they are the result of designing the lowest cost ship to handle a particular trade given that length, breadth and draught are constrained by factors such as the physical dimensions of the ports at which the ship is intended to operate. Some of the effects of these constraints on the sample data are readily apparent in Figures 4-24. The most obvious of these effects is the clear upper limit of thirty-two metres for the breadth of a container ship (Figure 11). This obviously corresponds to the 'third generation' container ships which are constrained by the width of the Panama Canal. A number of other graphs (for example Figures 6, 8, 16, 22 and 23) also seem to be bounded by an upper limit. It is apparent, therefore, that the relationships derived in this paper reflect the combined influence of the many criteria that the designer of a ship must consider, for example:

- minimum cost;
- maximum carrying capacity; and
- constraints on length, breadth, draught.

Tables 1-3 show the relationships which describe the line of best fit for the plots of length, breadth and draught against deadweight. However, a fuller explanation of the distribution of the sample data is obtained if these equations are considered in conjunction with the 95 per cent confidence intervals shown in Figures 4-24. Inspection of these graphs will show that the width of the confidence interval does vary considerably with ship type. This should be taken into account in any study which aims to generalise across ship types.

TONNAGE

The volume of a ship is expressed in volumetric tons and is referred to as its tonnage (1 ton = 100 ft^3). The charges for berthing and docking a ship, for passage through canals and locks and for many other facilities are usually based on a ship's tonnage. The two tonnage measures of prime interest are the gross registered tons (GRT) and net registered tons (NRT). Figures 25-38 show GRT and NRT plotted against deadweight for each of the seven ship types.

As tonnage is broadly a measure of the volume of a ship's cargo, and deadweight is broadly a measure of the weight of a ship's cargo, one would expect a positive correlation to exist between these variables. Examination of Figures 25-38 shows that the data do, in fact, exhibit such a correlation. A linear regression model was therefore used for each of GRT and NRT.

The resulting regression coefficients and statistics are presented in Tables 4 and 5. The t and R^2 statistics verify the fact that a correlation exists between tonnage and deadweight. As a check, power functions of the form GRT = $\alpha.(DWT)^{\beta}.\varepsilon$ were also fitted to the data and the resultant regression coefficients and statistics are presented in Tables C.4 and C.5. The regression estimates of the exponent, β, are all close to unity which further suggests that the linear model is appropriate.

It should be noted that it is not statistically valid to use the information in this paper to derive, by substitution, relationships between tonnage and any of the other ship characteristics. The questionable reliability of relationships derived in this manner should be taken into account in quantitative analysis.

AGE

The age of a ship in years, at 1977, was plotted against deadweight for each of the seven ship types and is shown in Figures 39-45. These graphs do, in general, illustrate the expected trend; namely, that the size of ships has been increasing with time. Although larger ships are being built with time, a variety of smaller ships is still being built to service that trade which does not justify the use of larger ships. This trend to build larger ships is exhibited most clearly by bulk carriers, ore carriers and tankers. For the remaining ship types it is still seen that the larger ships tend to be of newer construction.

The fact that ships are deleted from the Register as they are withdrawn from service suggests that the sample on which the regressions were performed probably provides a close approximation to the ships constructed in recent years, but for earlier years a smaller proportion of ships actually constructed would remain in the sample.

Due to the fact that there is a large range of ship sizes built in any one year, it is not reasonable to generate an equation which will describe age for a ship type simply as a function of deadweight. Therefore, a linear model was fitted to the date purely to check whether or not there is a trend for the size of ships being built to increase with time. Table 6 shows the regression coefficients and statistics for the model

$$\text{Age} = \alpha + \beta.(\text{DWT}) + \varepsilon$$

An F-test of the joint null hypothesis that $\alpha=0$ and $\beta=0$ was carried out at the 95 per cent confidence level to test for a relationship between age and deadweight. The test indicated that a significant correlation existed between age and deadweight for five of the seven ship types. The analysis found no significant relationship between deadweight and age for either ro-ro or passenger ships while for all other ship types the regression was significant and showed a negative correlation between age and deadweight, i.e. the larger ships tend to be newer.

These results really only provide a statistical verification of what can be seen by inspection of the graphs. There is, in general terms, a significant trend toward larger vessels for bulk carriers, ore carriers, tankers, general cargo ships and container ships, while no such trend is apparent for ro-ro or passenger ships.

A model of the form

$$\text{Age} = \alpha + \beta.(\text{DWT})^{-1} + \varepsilon$$

was also fitted to the data, and the regression coefficients and statistics are tabulated in Table C.6. The form of the graphs of sample data suggested this model may be able to explain more of the variation in deadweight with age, however, it provided no better results than did the simple linear model. Hence, all that can be claimed is that for five of the seven ship types there is some correlation between age and deadweight.

POWER

Figures 46-52 show total power plotted against deadweight for each of the seven ship types. It is evident from these plots that the data exhibit considerable scatter for each ship type. There does, however, appear to be a strong correlation between power and deadweight for ore carriers, tankers and to a lesser extent bulk carriers. The following model was fitted to the data

$$\text{Power} = \alpha.(\text{DWT})^{\beta}.\varepsilon$$

The regression coefficients and statistics are presented in Table 7.1. The regression coefficients indicate a marked difference between the data for container ships and the data for the remaining ships, i.e. the exponent, β, is in excess of unity for container ships but less than unity for the remaining ship types. A number of container ships have been designed to travel at 25-30 knots and these, having high power, would tend to move the regression line up, giving it an increasing slope with deadweight. Given current fuel prices, however, these ships have not been proving economic at such high speeds. Hence, it is thought analysis of the container ships currently being designed may well show a significantly lower exponent for a regression of power on deadweight.

These variations in the form of the regression models are, in fact, due to the differing economic considerations present over time or between ship types. It should be noted, therefore, that the regression models presented in this paper are the result of interaction between all the factors that must be considered in the design of a ship. They do not represent fundamental physical relationships in their own right.

Although the regression models presented in Table 7.1 are significant for each ship type, it is clear that individual data points are sometimes scattered far from the line of best fit such that the confidence intervals are wide and diverge at large values of deadweight. In an attempt to determine a relationship which gives a better explanation of the variance in the data, alternative models were tested. The regression coefficients and statistics for the following model are shown in Table C.7:

$$\text{Power} = \alpha + \beta.(\text{DWT}) + \varepsilon$$

None of these regressions provided any better explanation of the data than did the previous model. Figures 53-59 show the regression line for a model which includes the term 'speed' (V) superimposed on a plot of the sample data contained in Figures 46-52. Table 7.2 shows the regression coefficients and statistics for this model, which is of the form

$$\text{Power} = \alpha . (\text{DWT})^{\beta} . (\text{V})^{\gamma} . \varepsilon$$

The relationships obtained for bulk carriers, ore carriers and tankers using this model are very similar to those derived using the previous model. This result is due to the fact that all ships, within each of these ship types, travel at almost the same speed and therefore the speed term in the regression virtually becomes a constant. (Figures 62-64 show the speed versus deadweight plots for these ship types.) On the other hand, various container ships have been designed for different operating speeds. Figure 53 shows that the variation in power can be more fully explained by including speed as a further explanatory variable. Considering this figure in conjunction with the plot of speed against deadweight (Figure 60), it is clear that the ships designed to travel at around sixteen knots tend to lie close to the power regression generated with speed equal to sixteen knots. Similarly, the ships designed for twenty-three knots fall close to the regression with speed equal to twenty-three, and twenty-eight knot container ships lie close to the twenty-eight knot regression. Thus, as one would expect, speed is seen to be an important factor in explaining the variation in power for container ships.

Figure 61 indicates that a number of ro-ro ships travel at about seventeen knots while a number of others travel at about twenty-two knots. The ships from each of these two groups tend to fall close to the regression using the corresponding value of speed (Figure 54).

General cargo ships travel at a range of speeds from ten to twenty-three knots (Figure 65), however, a number of these ships travel at about fifteen knots and tend to fall around the regression with the corresponding value of speed (Figure 58).

25474/80—3

The sample is small for passenger ships. However, there are a number of ships that travel at twenty to twenty-one knots and these tend to lie just above the regression with speed equal to twenty knots (Figure 59).

In general, it is clear that the variation in speed for a given ship type is a major factor in explaining the variation in power. However, the design speed for bulk carriers, ore carriers and tankers is virtually a constant and, hence, omission of the speed term does not have a major effect on the regressions for these ship types. On the other hand, container ships show the largest systematic variation in speed, and this results in the speed term being very significant in explaining the variation in power.

The 95 per cent confidence intervals of power modelled on deadweight alone are quite wide and divergent at extreme values of deadweight for all ship types. However, when power is modelled on deadweight and speed, we find that the confidence intervals for all ship types are improved.

SPEED

Figures 60-66 show speed plotted against deadweight for each of the seven ship types. As discussed in the previous section, these graphs indicate that there is very little variation in operating speed for bulk carriers, ore carriers and tankers which means that the rate at which these ships move freight may be quite accurately predicted.

Table 8 shows the regression coefficients and statistics for the model

$$\text{Speed} = \alpha + \beta.(\text{DWT}) + \varepsilon.$$

A positive correlation was found between speed and deadweight for all ship types. However, the slopes of the bulk carrier, ore carrier and tanker regressions are so small that for many practical purposes they may be considered negligible. The regression for container ships shows a pronounced positive slope. The data exhibit a distinctly linear relationship between 5000 deadweight tonne ships which travel at fifteen knots and 35 000 deadweight tonne ships which travel at twenty-five knots. For ro-ro ships (see Figure 61) the regression line shows a positive slope, but the data do not show a steady trend for larger ships to operate at a higher speed. Rather, there appears to be a quantum jump at 14 000 deadweight tonnes : ships less than 14 000 DWT travel at about seventeen knots, while ships greater than 14 000 DWT travel at about twenty-two knots. Such irregularities in the data illustrate the need to consider the plot of the sample in conjunction with the regression coefficients and statistics in order to obtain a fuller understanding of the relationship between the two variables. The regression model for cargo ships (Figure 65) provides only a broad indication that a positive correlation exists. The data, in fact, are scattered over a quite wide range of speeds. The data for passenger ships (Figure 66) shows that most ships in excess of 1500 deadweight tonnes travel at about twenty knots. The implication of the regression model, that the larger the passenger ship the higher will be its operating speed is not, in fact, true. Notwithstanding this fact, examination of Figure 66 will show that the regression model still provides an indication of the operating speed that could be expected of a passenger ship of a given deadweight.

An alternative model was examined to see whether it provided a better explanation of the data. The form of the model was

$$\text{Speed} = \alpha . (\text{DWT})^{\beta} . \varepsilon .$$

The regression coefficients and statistics for this model are shown in Table C.8, but it provided no better correlations than the simple linear model.

The confidence intervals for bulk carriers, ore carriers and tankers are quite narrow and do not diverge at large values of deadweight. The confidence intervals for the other ship types are not encouraging: they are all wide and the passenger ship confidence limits diverge at extreme values of deadweight.

CONTAINER CAPACITY

The plot of container capacity against deadweight shown in Figure 67 suggests that there is a linear relationship between the number of containers carried by a ship and its deadweight. Hence, the following model was fitted to the data:

$$TEU = \alpha + \beta.(DWT) + \varepsilon,$$

where TEU = twenty foot equivalent units.

The resulting regression coefficients and statistics are shown in Table 9. These indicate the presence of a strong positive correlation between TEUs and deadweight. It is reasonable that there would be a linear relationship between these two variables as deadweight gives an indication of the carrying capacity of a ship, and the number of TEUs is the carrying capacity of a container ship.

An alternative model

$$TEU = \alpha .(DWT)^{\beta}.\varepsilon$$

was also regressed and the regression coefficients and statistics are presented in Table C.9. The exponent, β, resulting from this regression is very close to unity which reinforces the hypothesis that there is a linear relationship between container capacity and deadweight.

The confidence interval shows that although the data do show a direct correlation between the variables, there is still a degree of scatter about the line of best fit.

This relationship, together with those for length, breadth and draught, can be used to determine what size container ships are likely to be able to enter a port, berth and have their containers handled and stored efficiently. For example, the draught will limit the size of ship that can use the channels, (when fully loaded), length will affect the berthing of the ship, breadth will affect the handling of the cargo and the container capacity will determine the adequacy of storage facilities. The relationships may then be used to determine the effect of a given change to the port's operating characteristics.

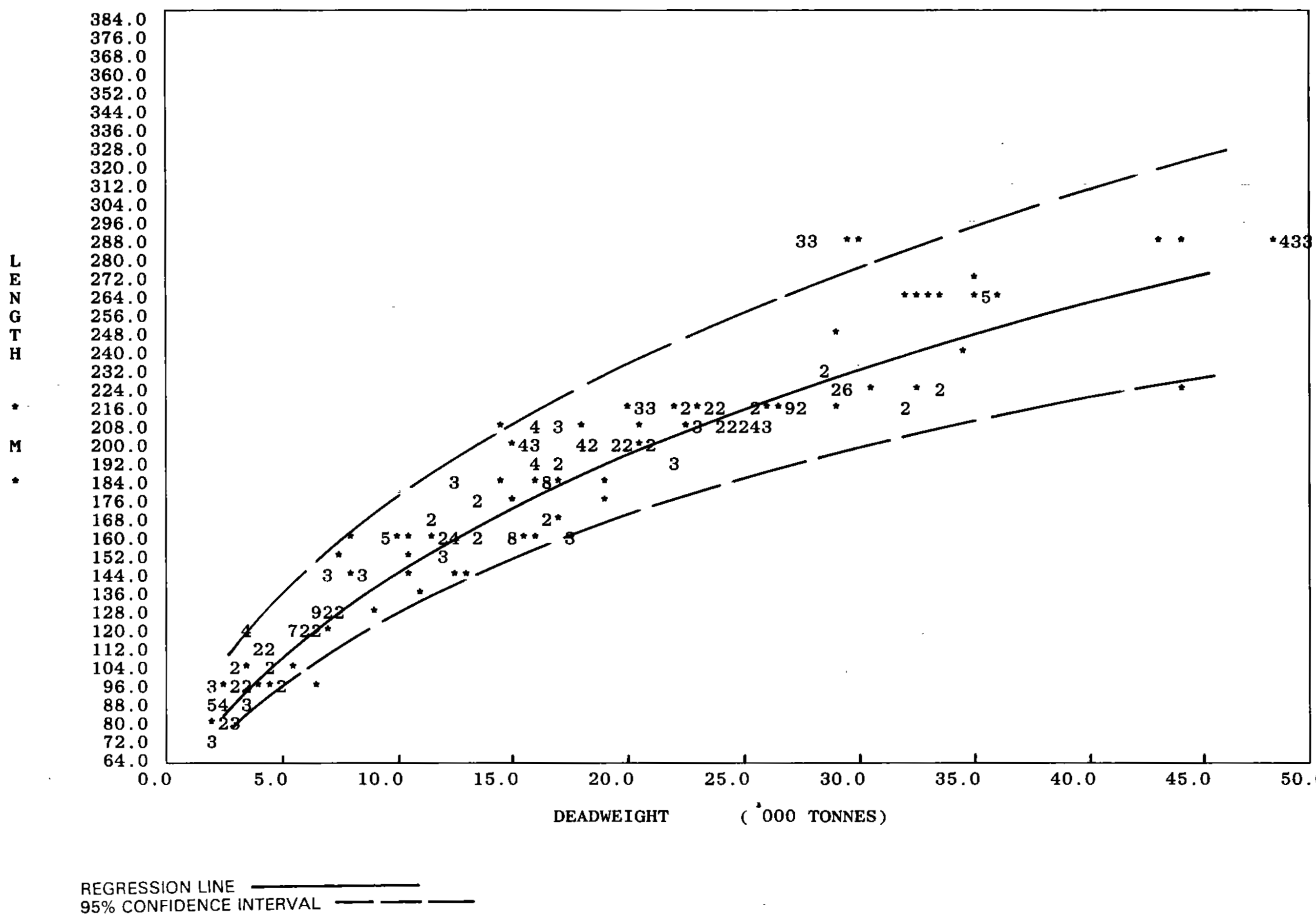


FIGURE 4 LENGTH VS DEADWEIGHT (CONTAINER SHIPS)

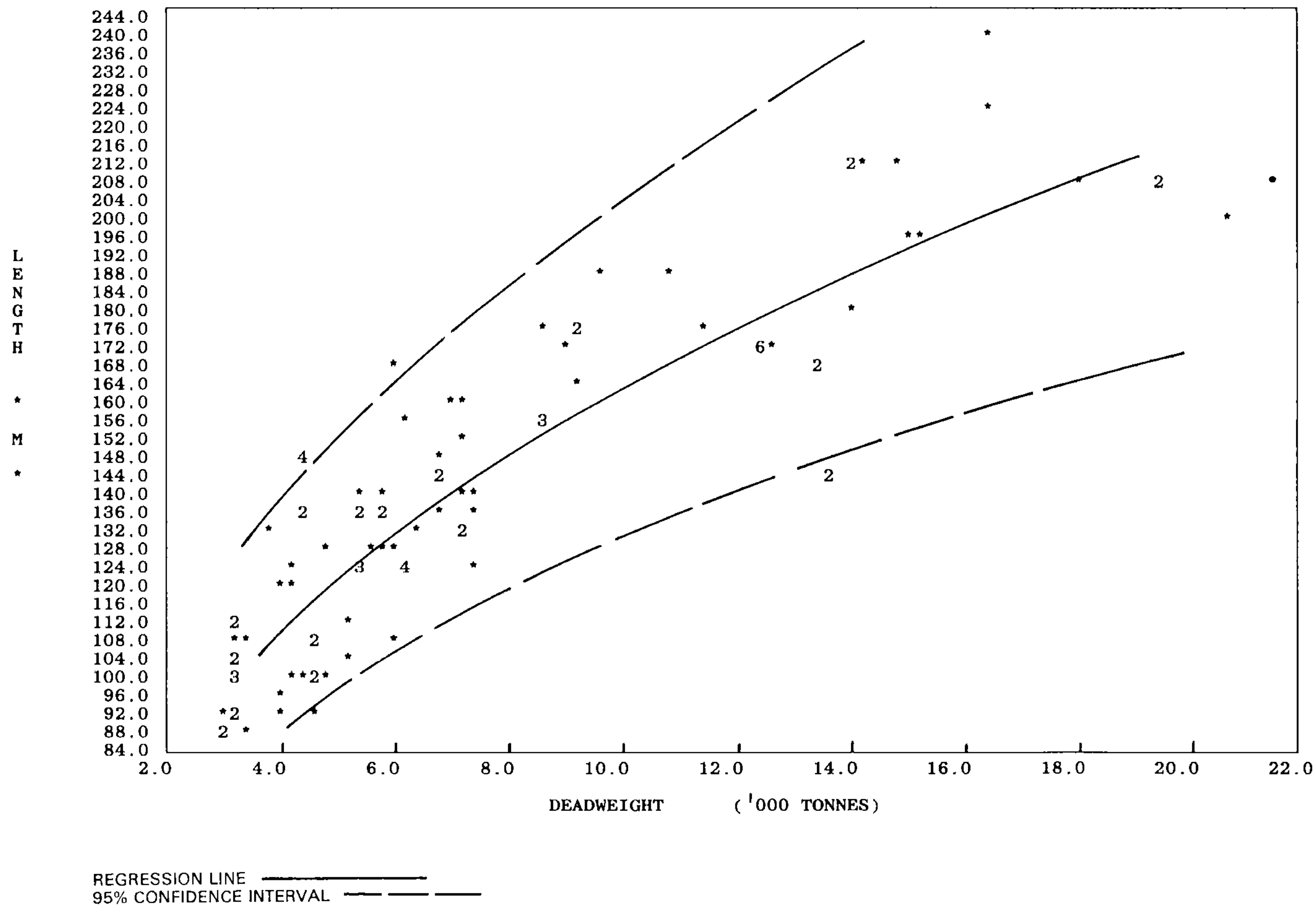


FIGURE 5 LENGTH VS DEADWEIGHT (RO. RO. SHIPS)

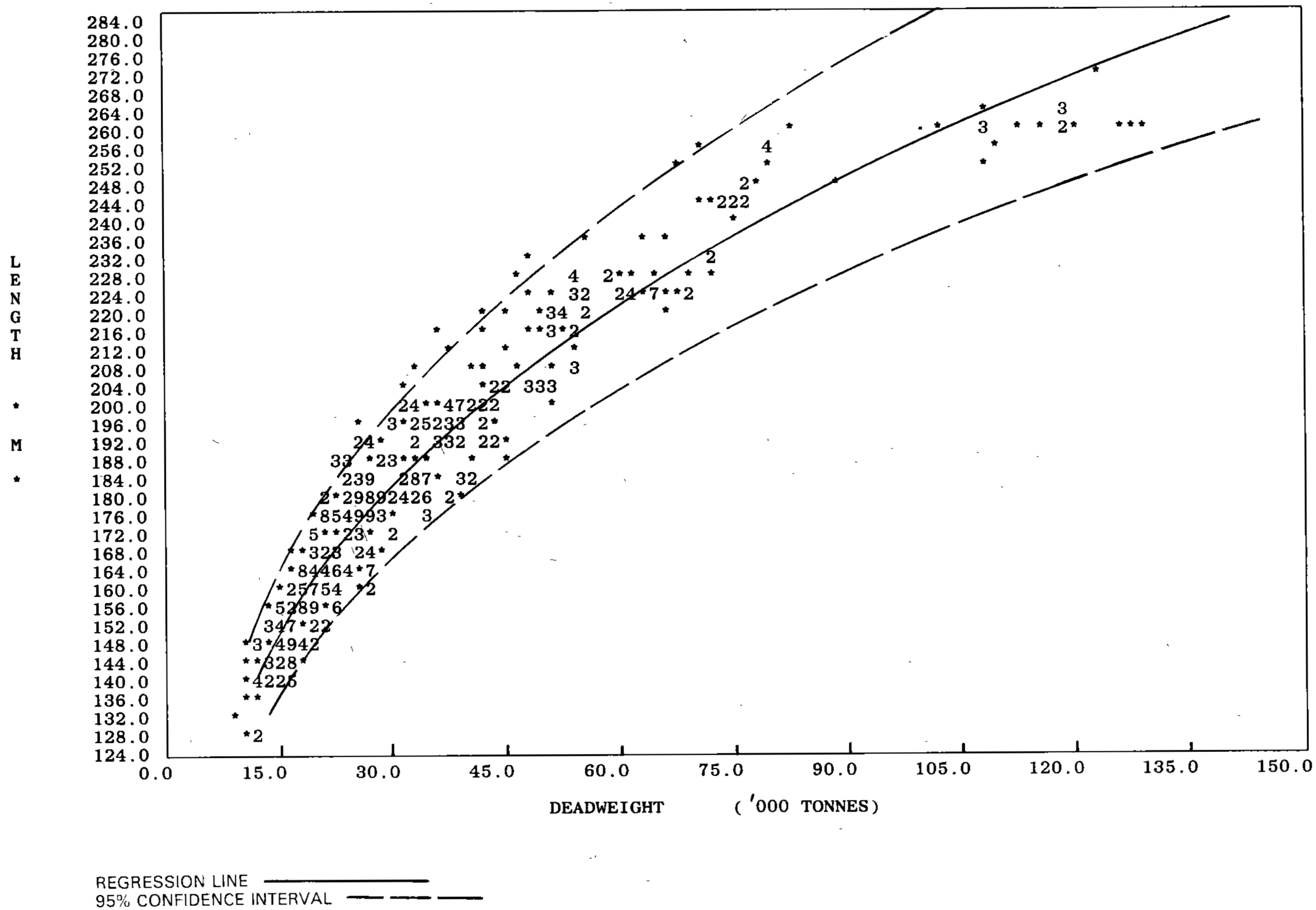


FIGURE 6 LENGTH VS DEADWEIGHT (BULK CARRIERS)

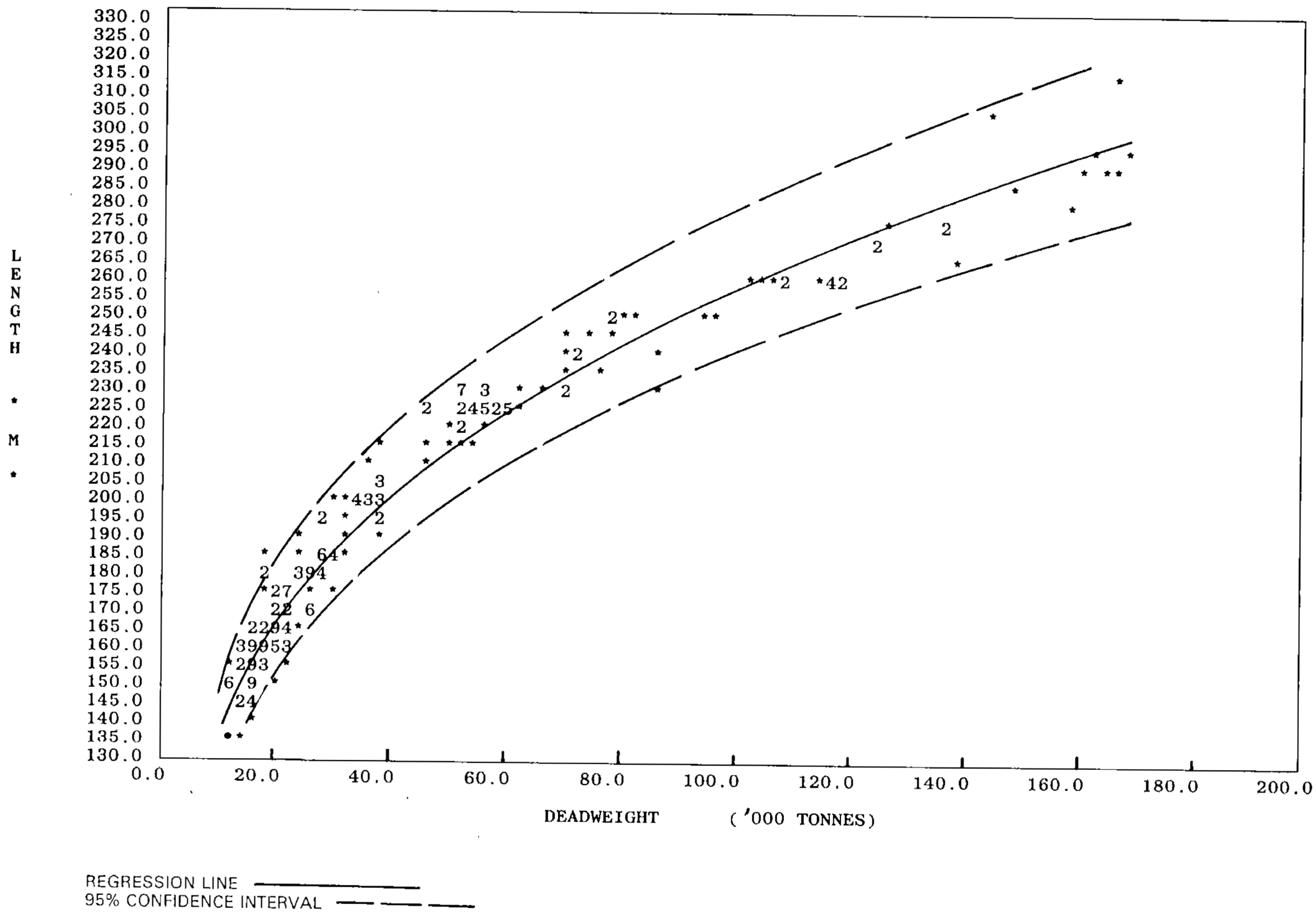


FIGURE 7 LENGTH VS DEADWEIGHT (ORE CARRIERS)

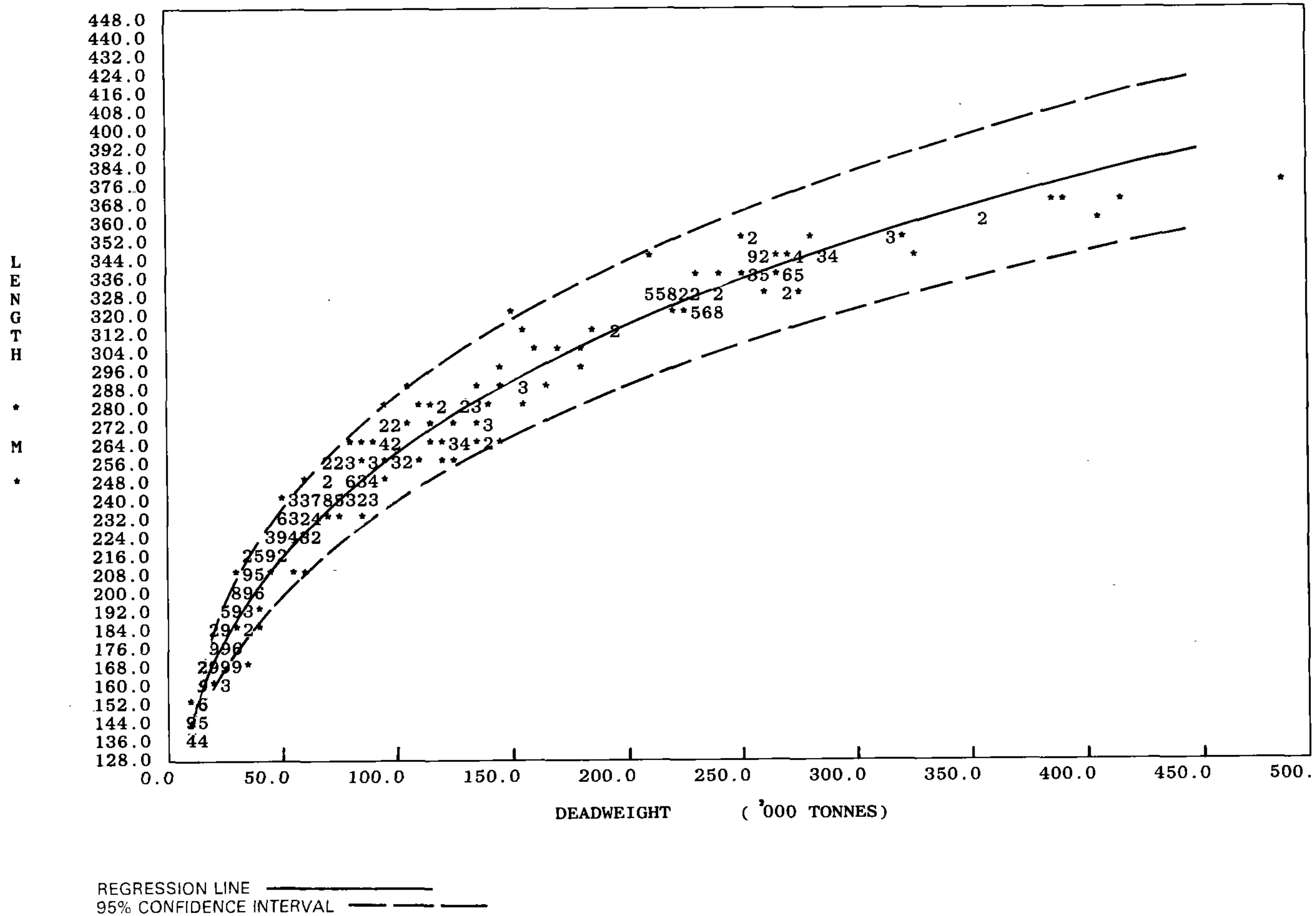
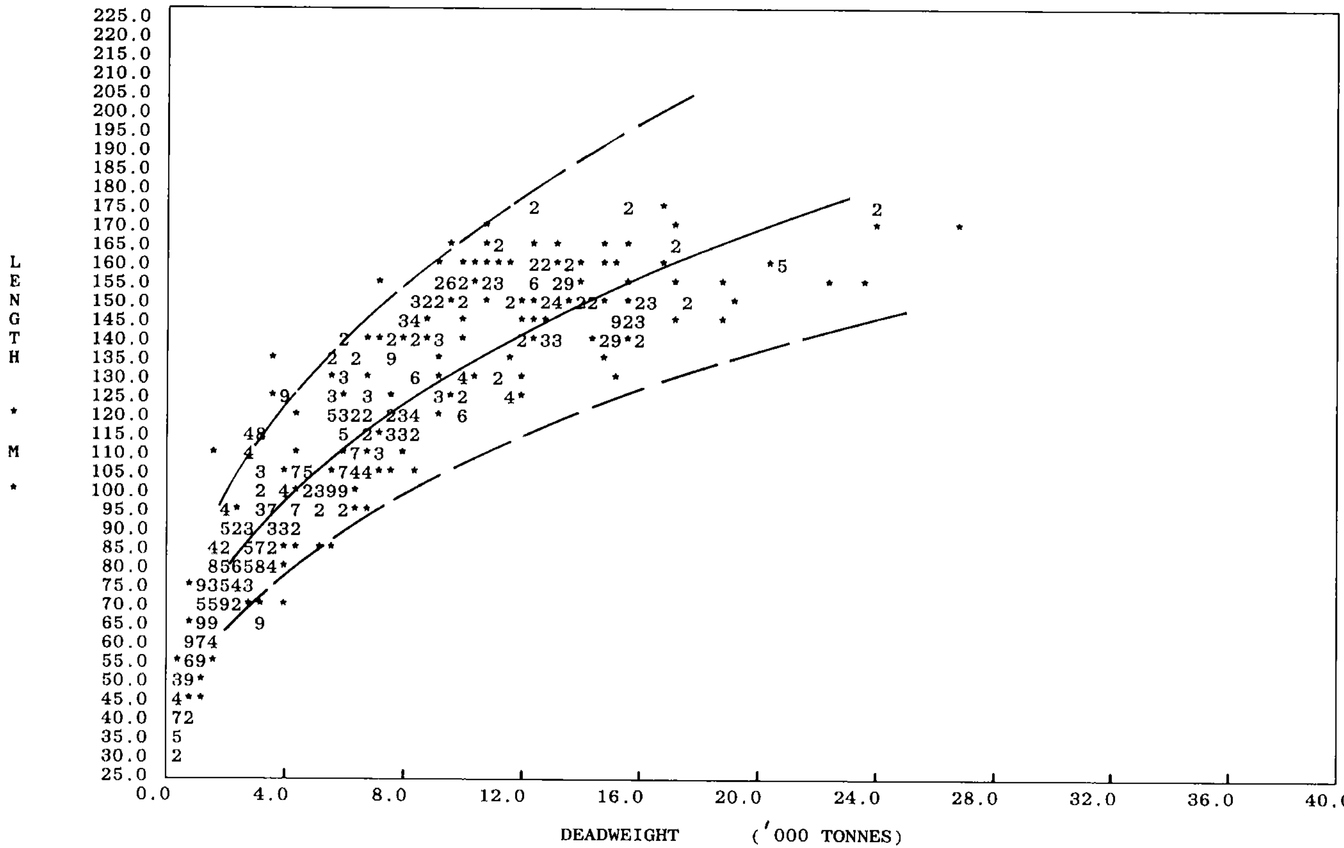


FIGURE 8 LENGTH VS DEADWEIGHT (TANKERS)



REGRESSION LINE ———
95% CONFIDENCE INTERVAL — — —

FIGURE 9 LENGTH VS DEADWEIGHT (GENERAL CARGO SHIPS)

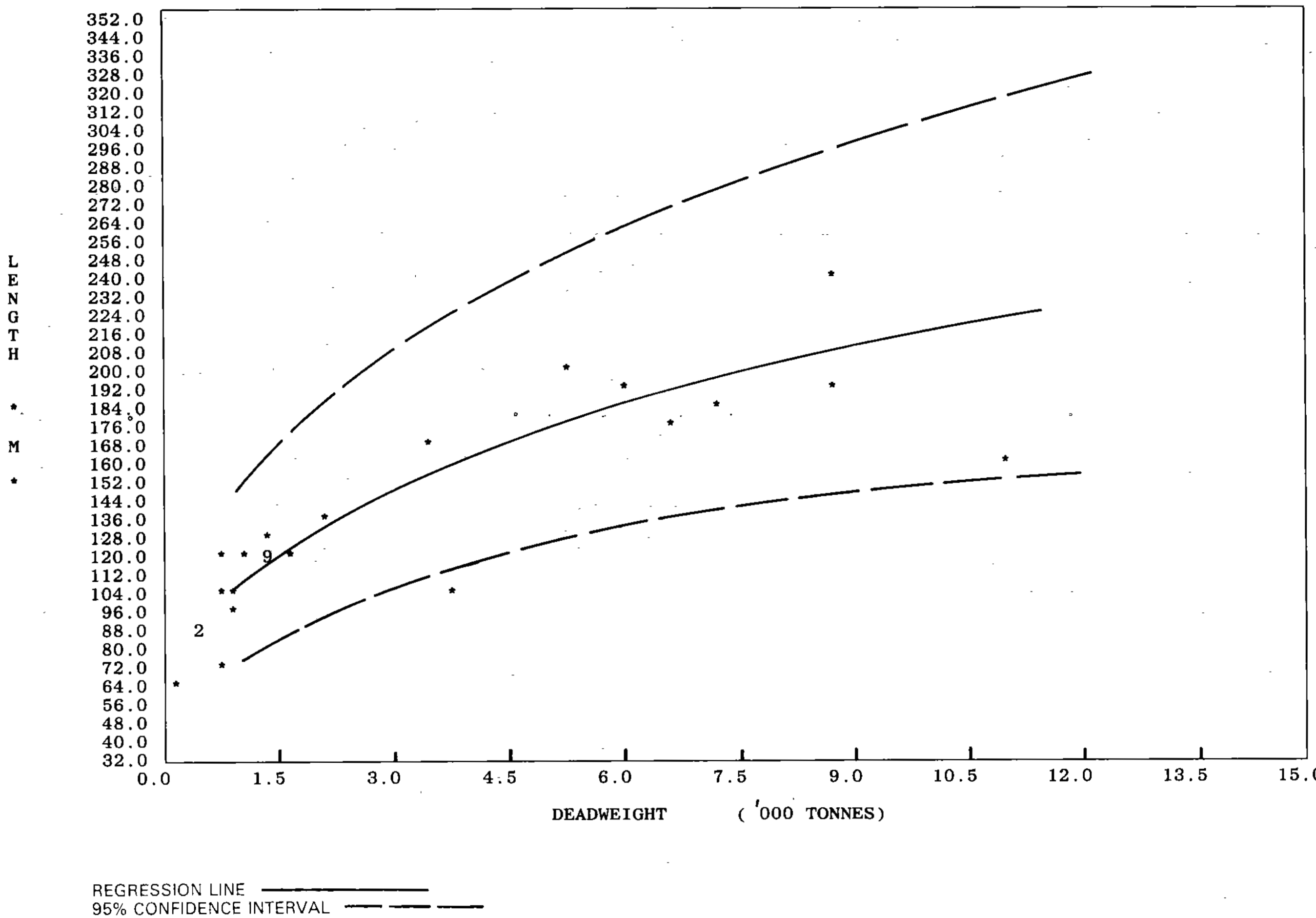


FIGURE 10 LENGTH VS DEADWEIGHT (PASSENGER SHIPS)

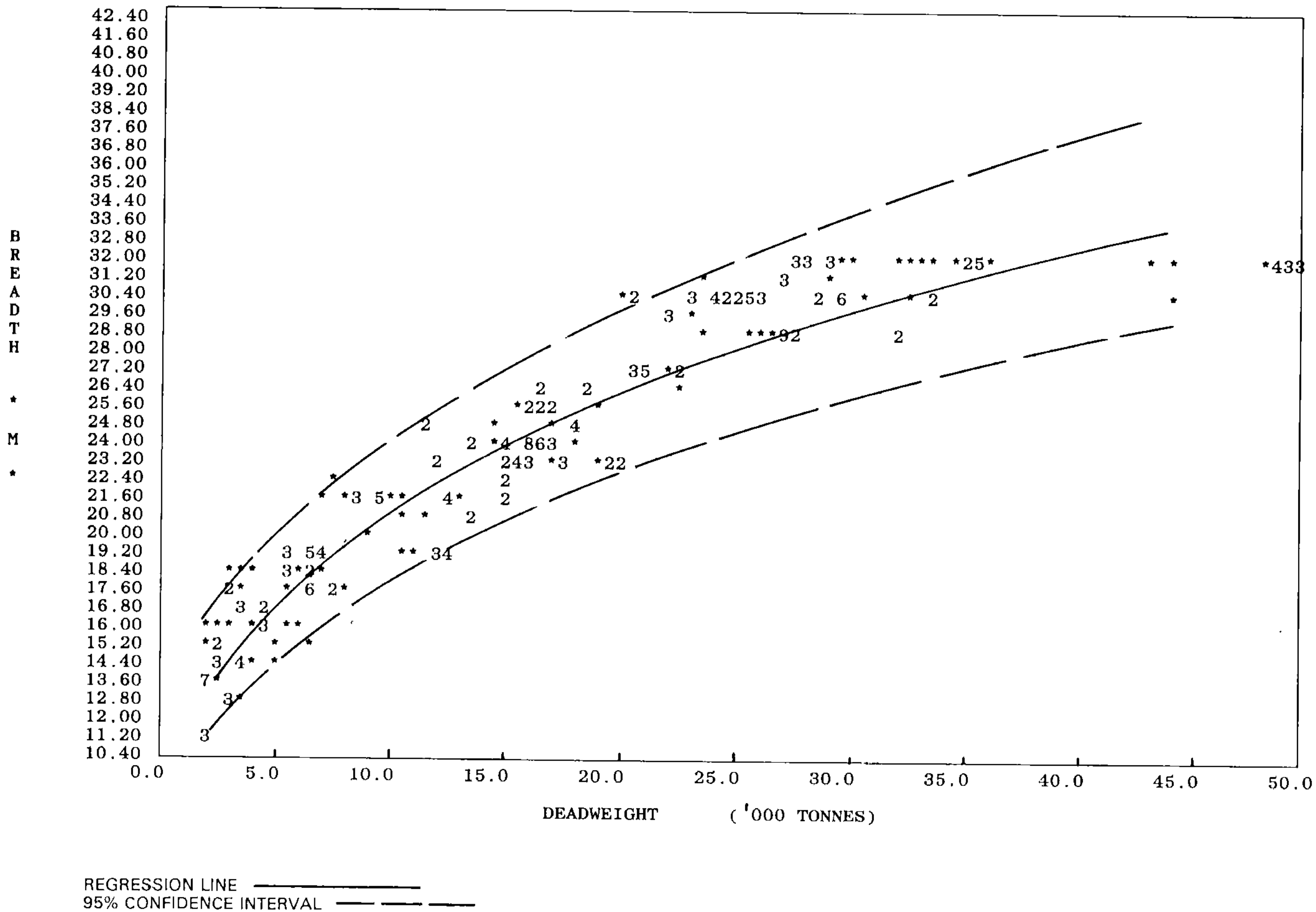


FIGURE 11 BREADTH VS DEADWEIGHT (CONTAINER SHIPS)

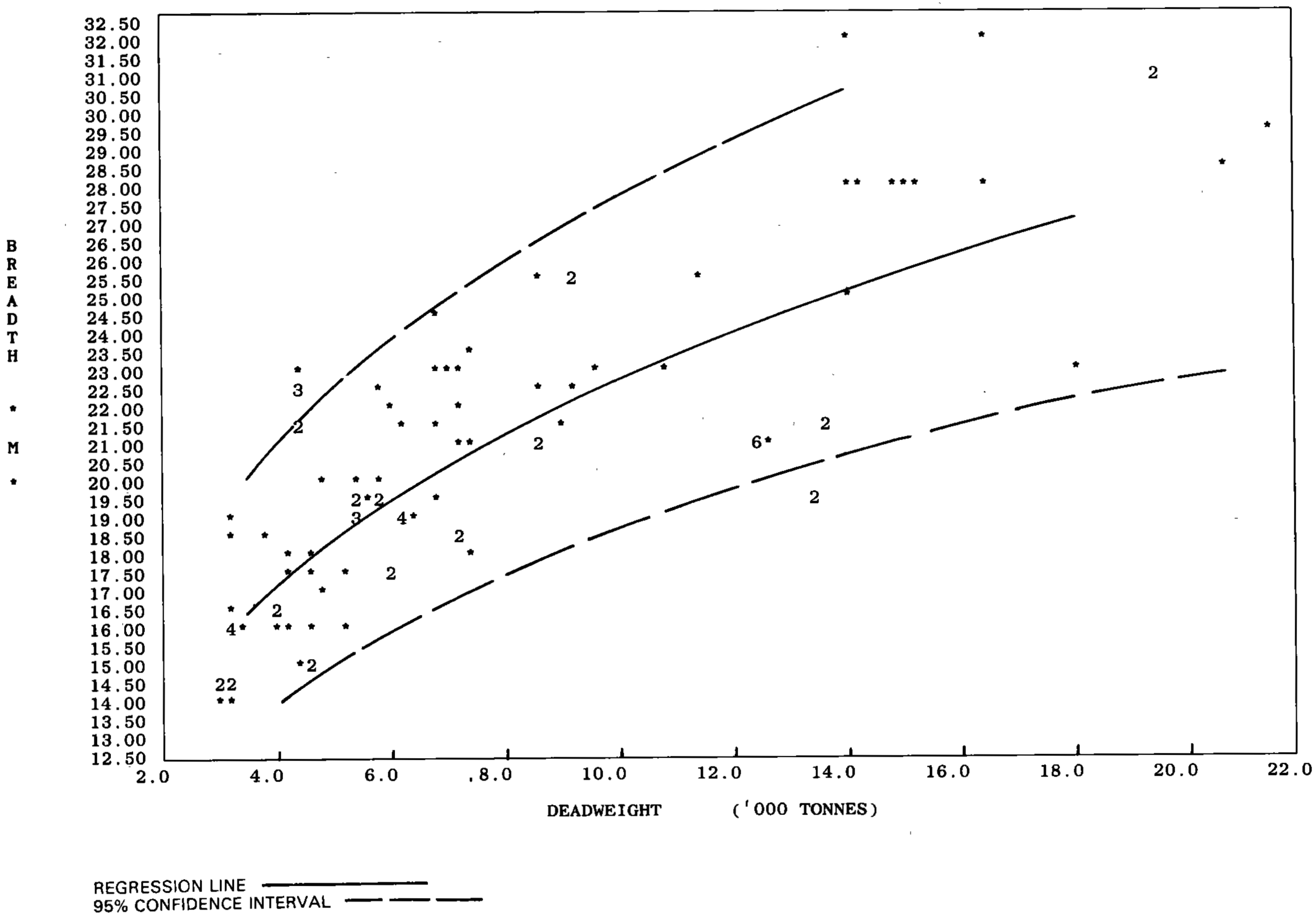


FIGURE 12 BREADTH VS DEADWEIGHT (RO.RO SHIPS)

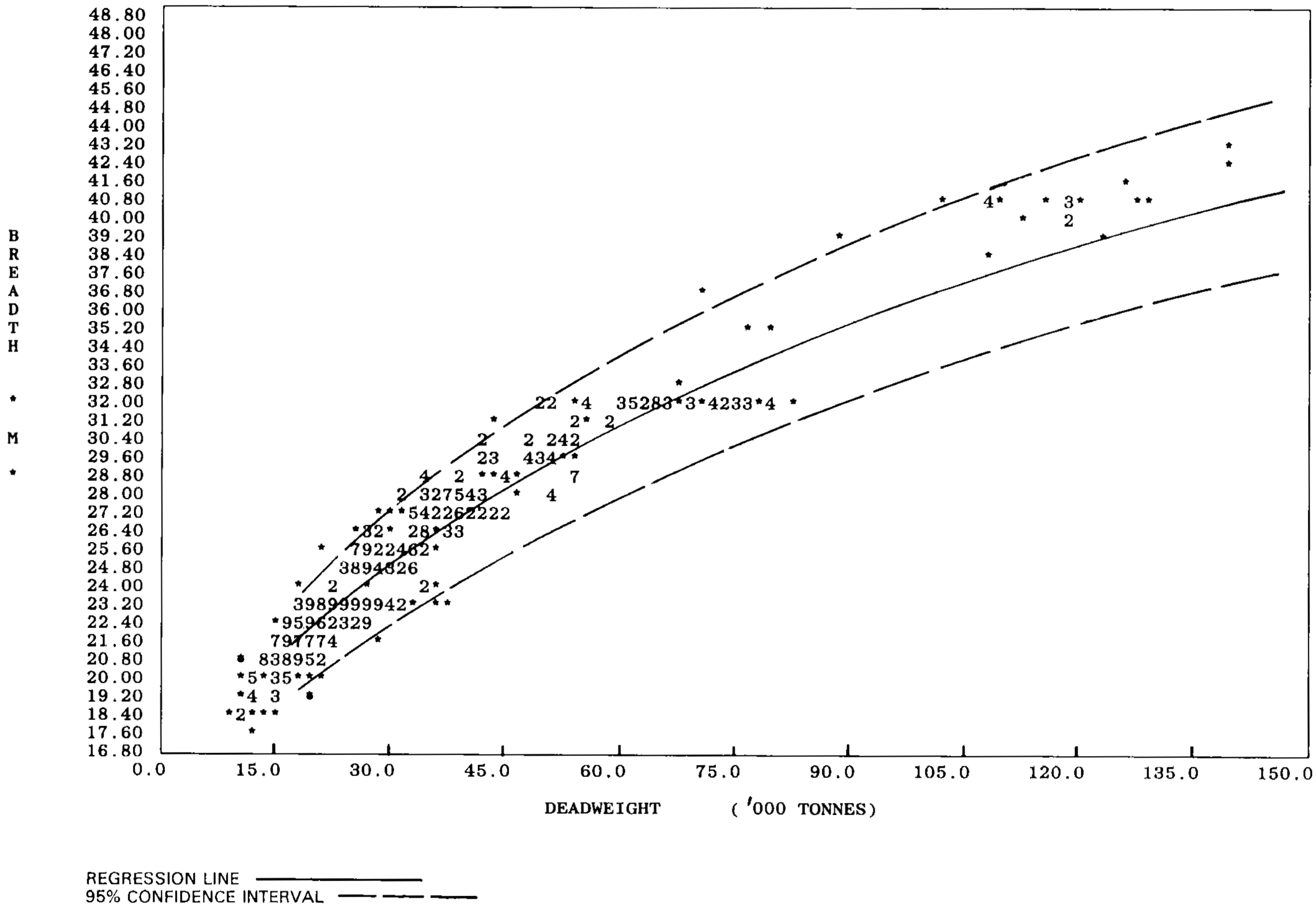


FIGURE 13 BREADTH VS DEADWEIGHT (BULK CARRIERS)

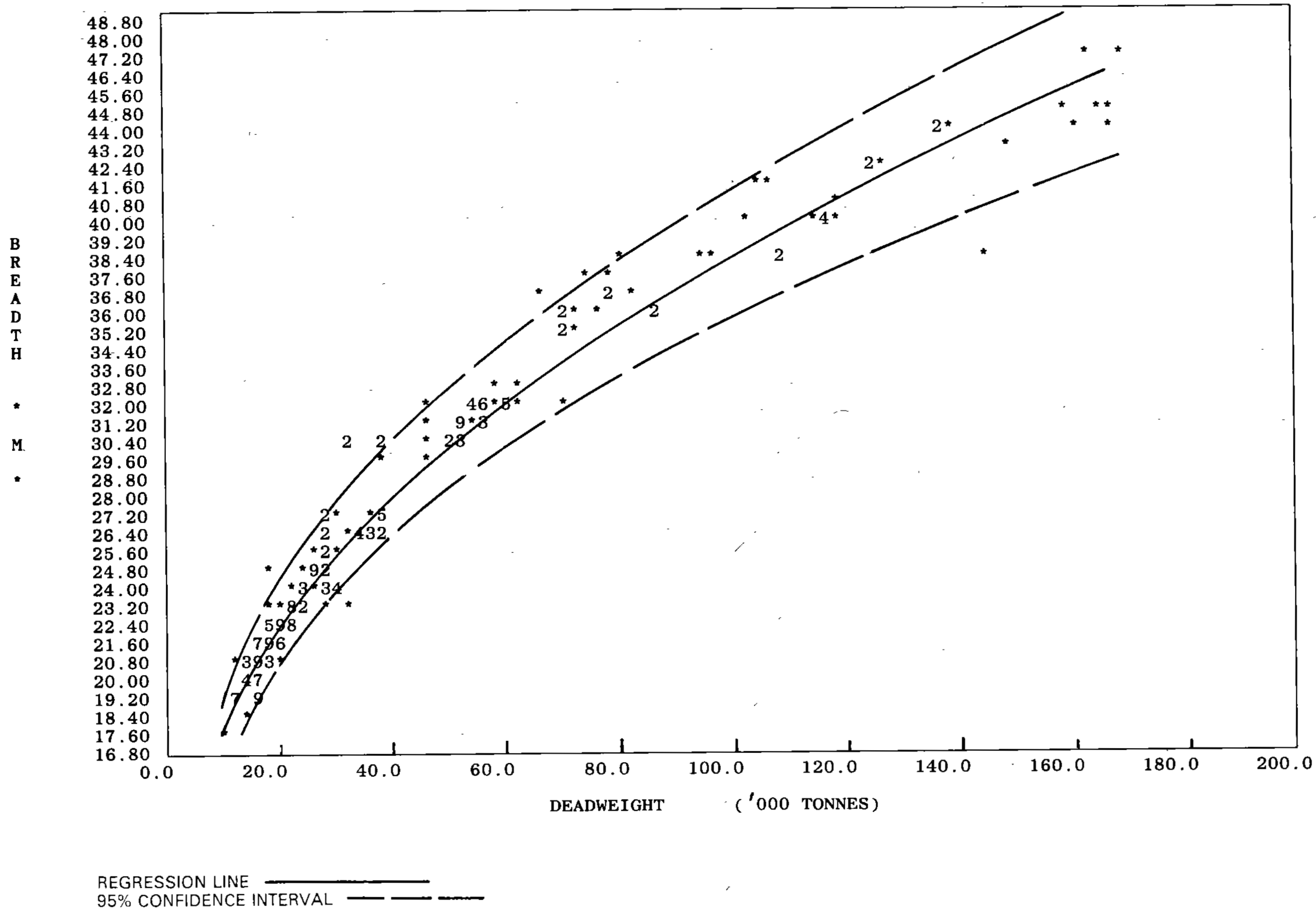


FIGURE 14 BREADTH VS DEADWEIGHT (ORE CARRIERS)

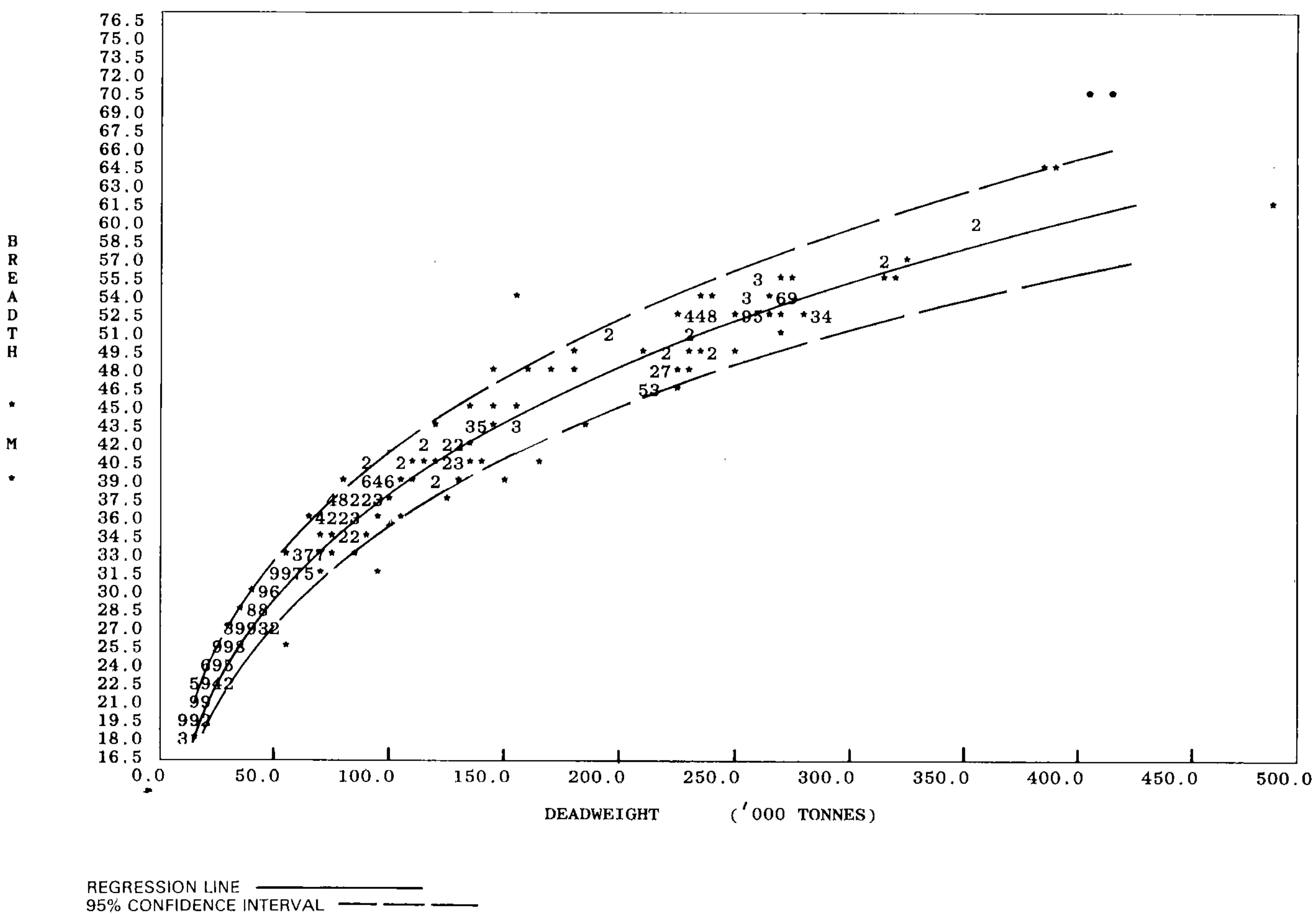


FIGURE 15 BREADTH VS DEADWEIGHT (TANKERS)

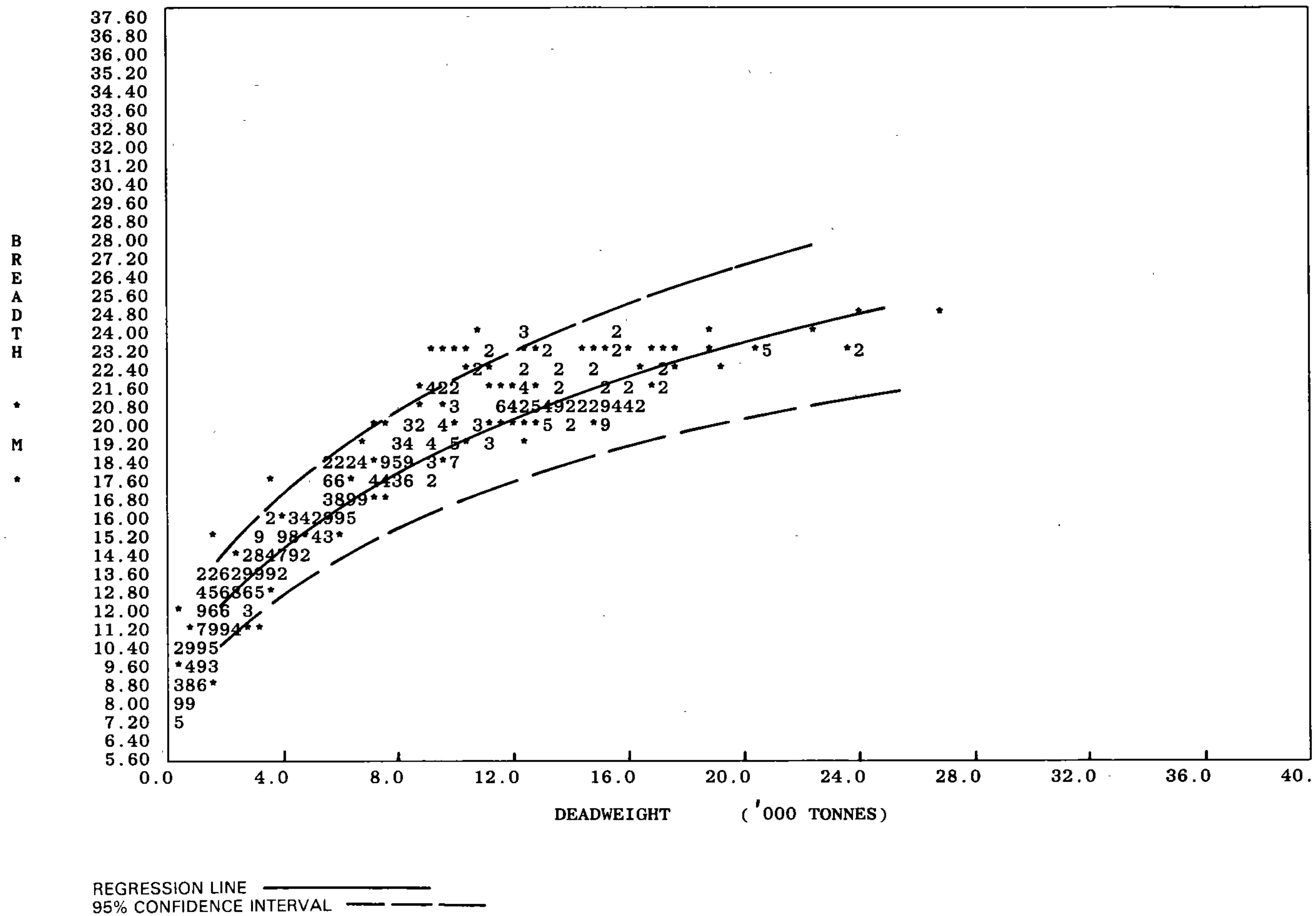


FIGURE 16 BREADTH VS DEADWEIGHT (GENERAL CARGO SHIPS)

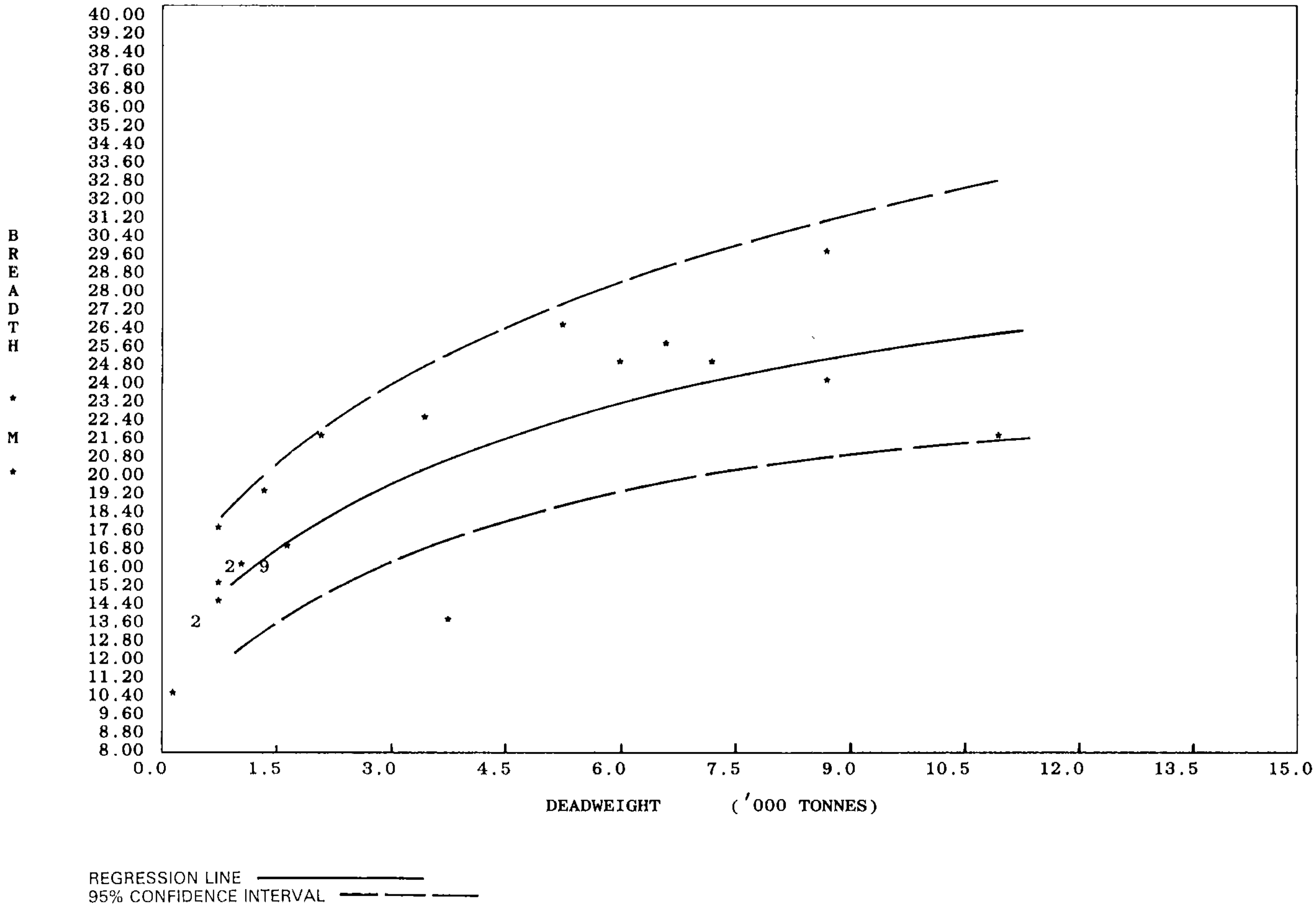


FIGURE 17 BREADTH VS DEADWEIGHT (PASSENGER SHIPS)

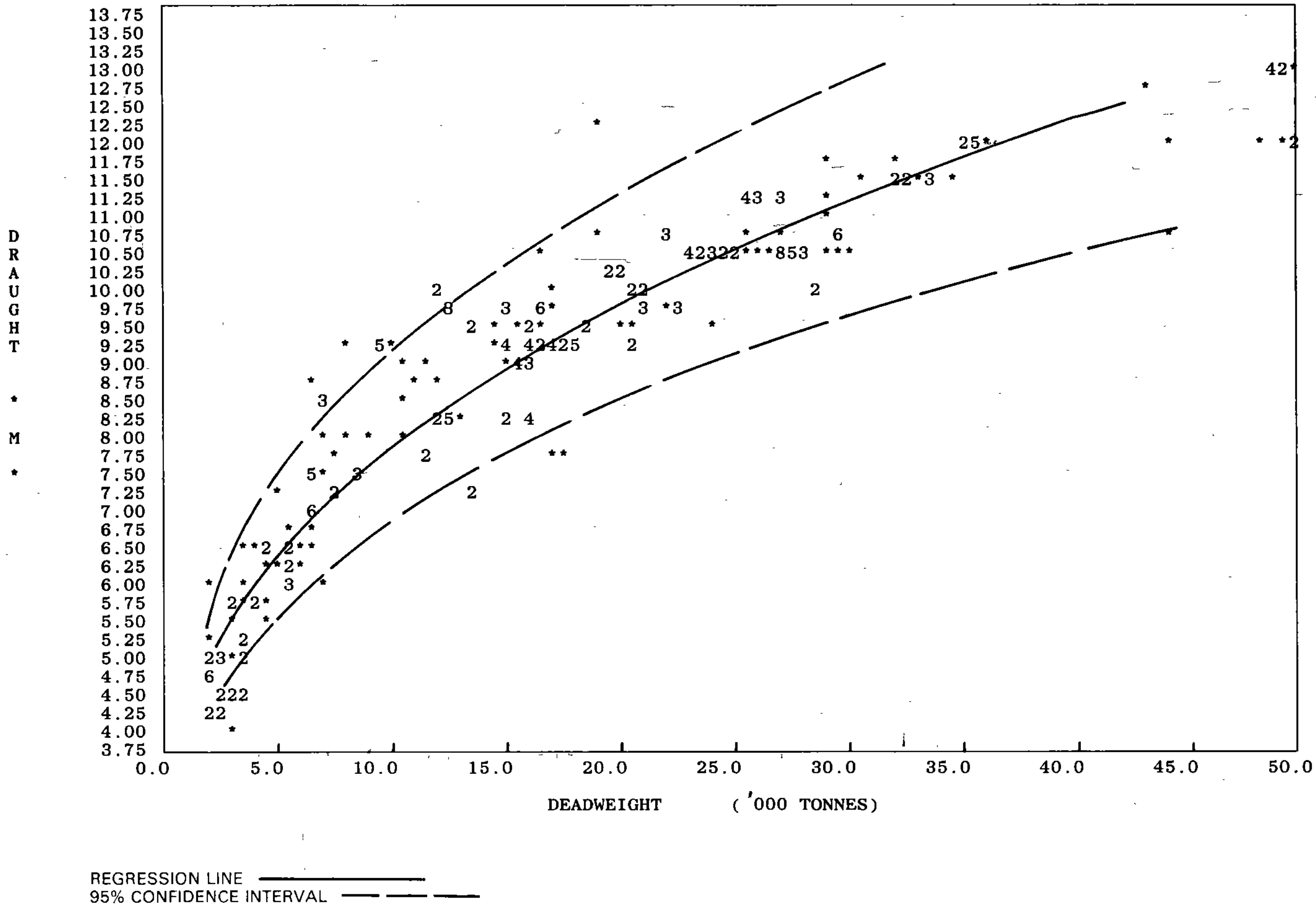


FIGURE 18 DRAUGHT VS DEADWEIGHT (CONTAINER SHIPS)

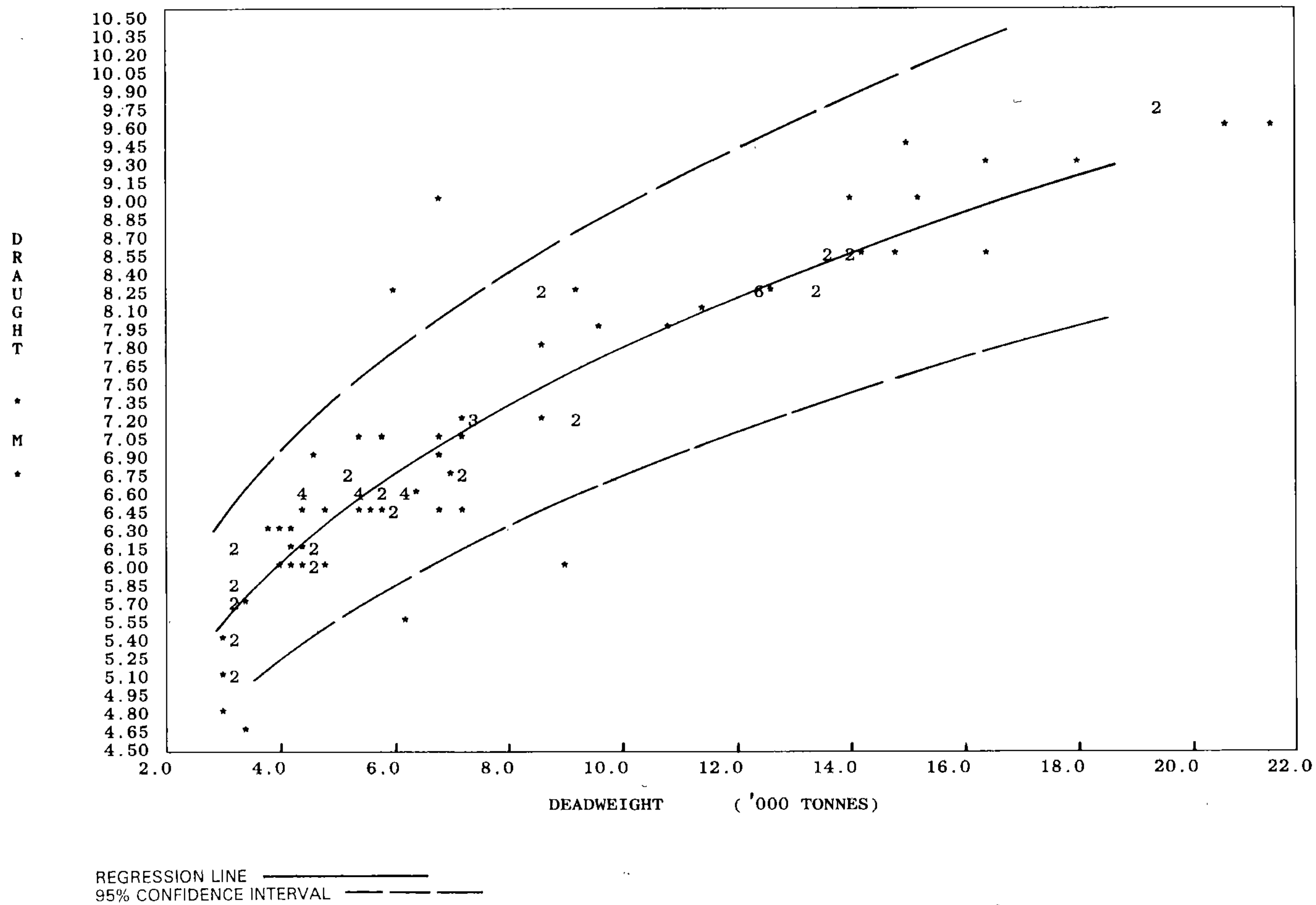


FIGURE 19 DRAUGHT VS DEADWEIGHT (RO. RO SHIPS)

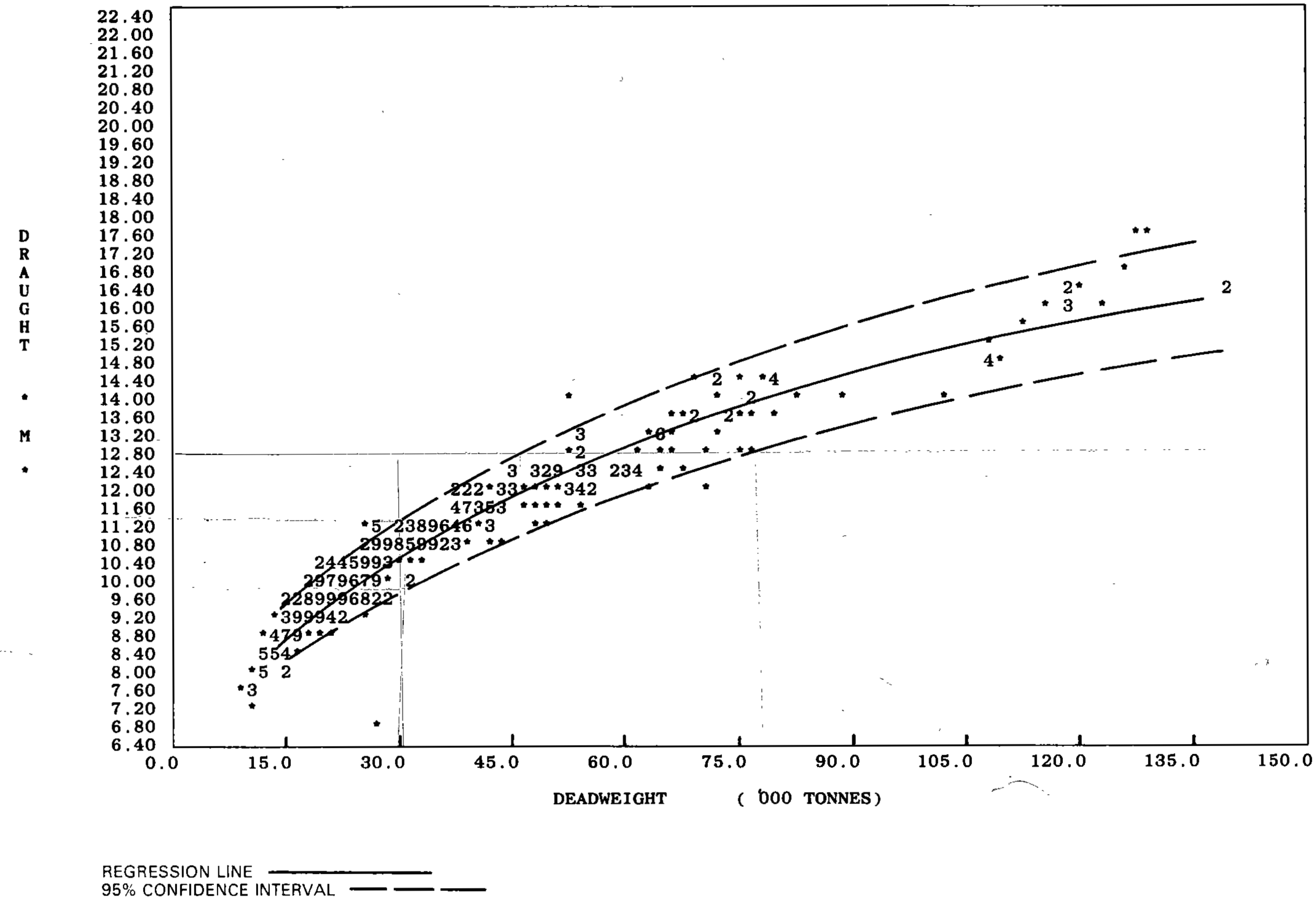


FIGURE 20 DRAUGHT VS DEADWEIGHT (BULK CARRIERS)

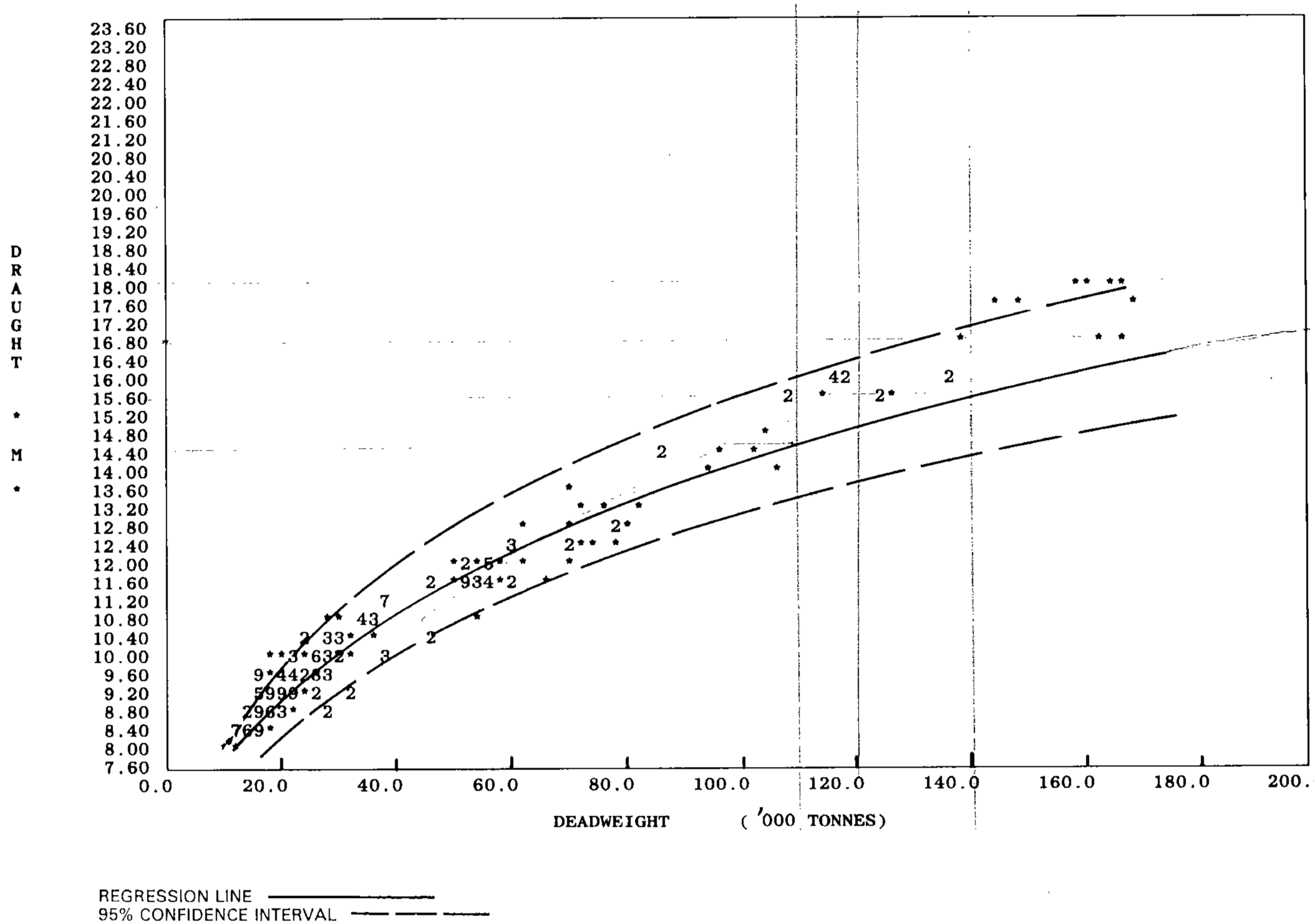
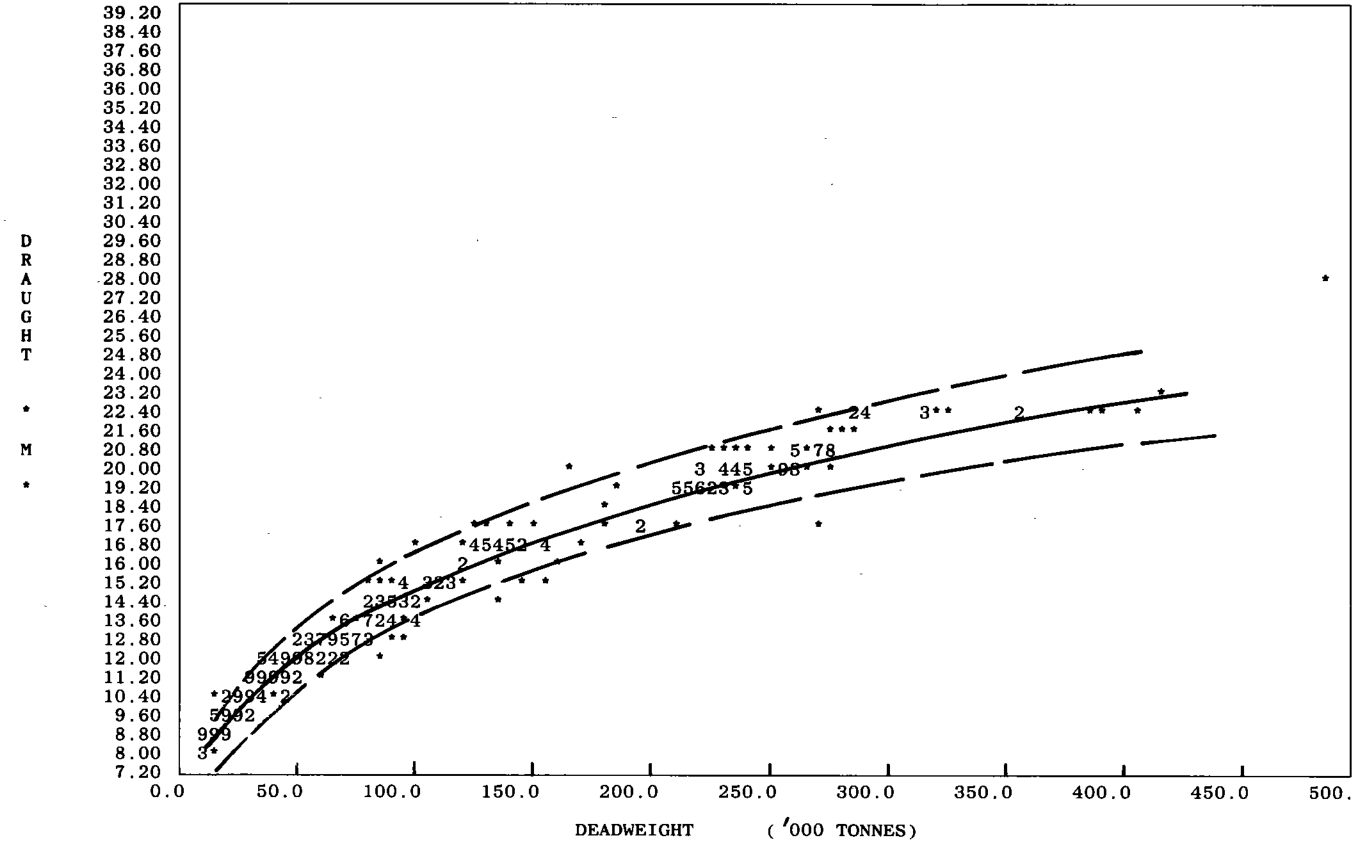


FIGURE 21 DRAUGHT VS DEADWEIGHT (ORE CARRIERS)



REGRESSION LINE ————————
95% CONFIDENCE INTERVAL — — — —

FIGURE 22 DRAUGHT VS DEADWEIGHT (TANKERS)

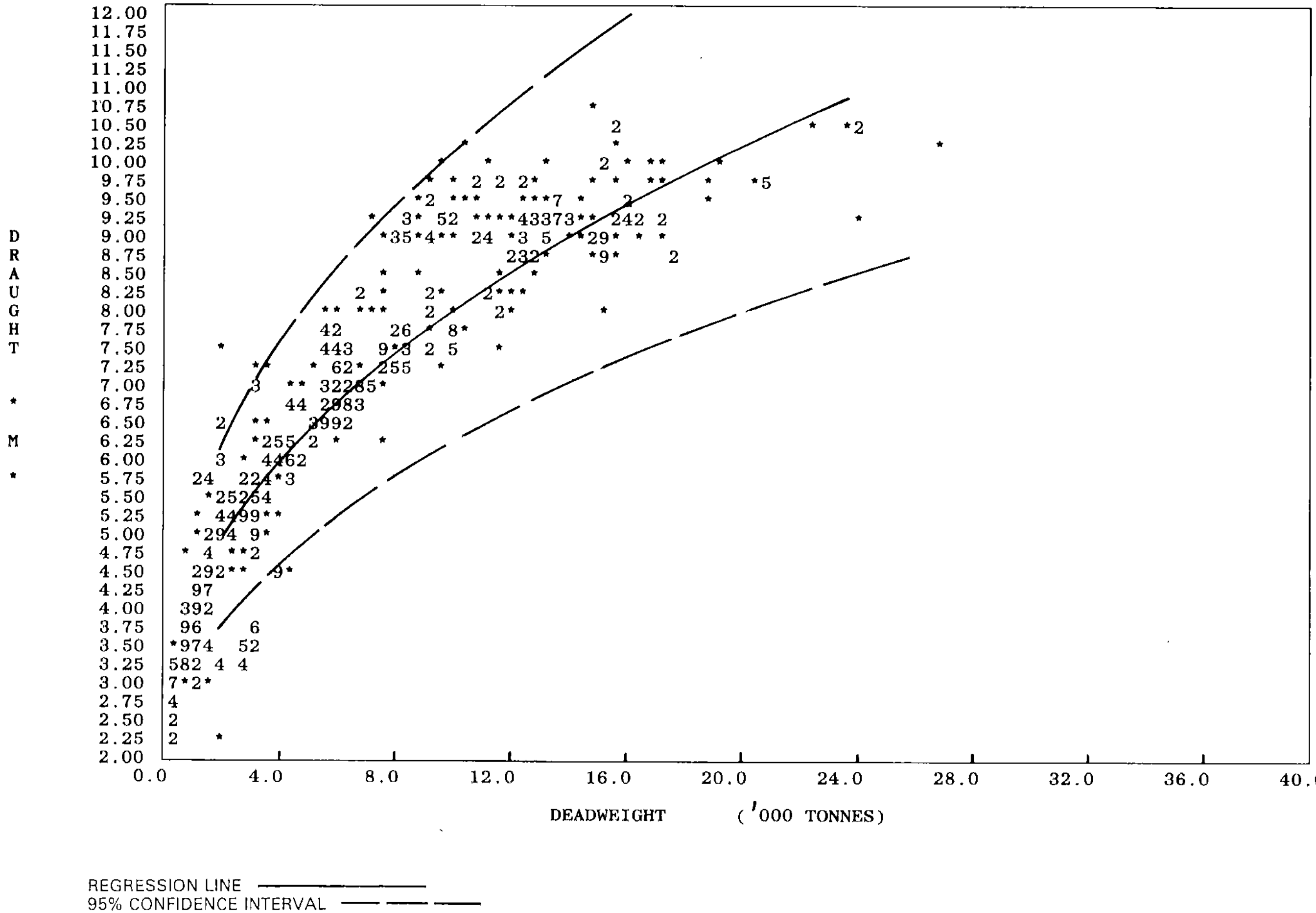


FIGURE 23 DRAUGHT VS DEADWEIGHT (GENERAL CARGO SHIPS)

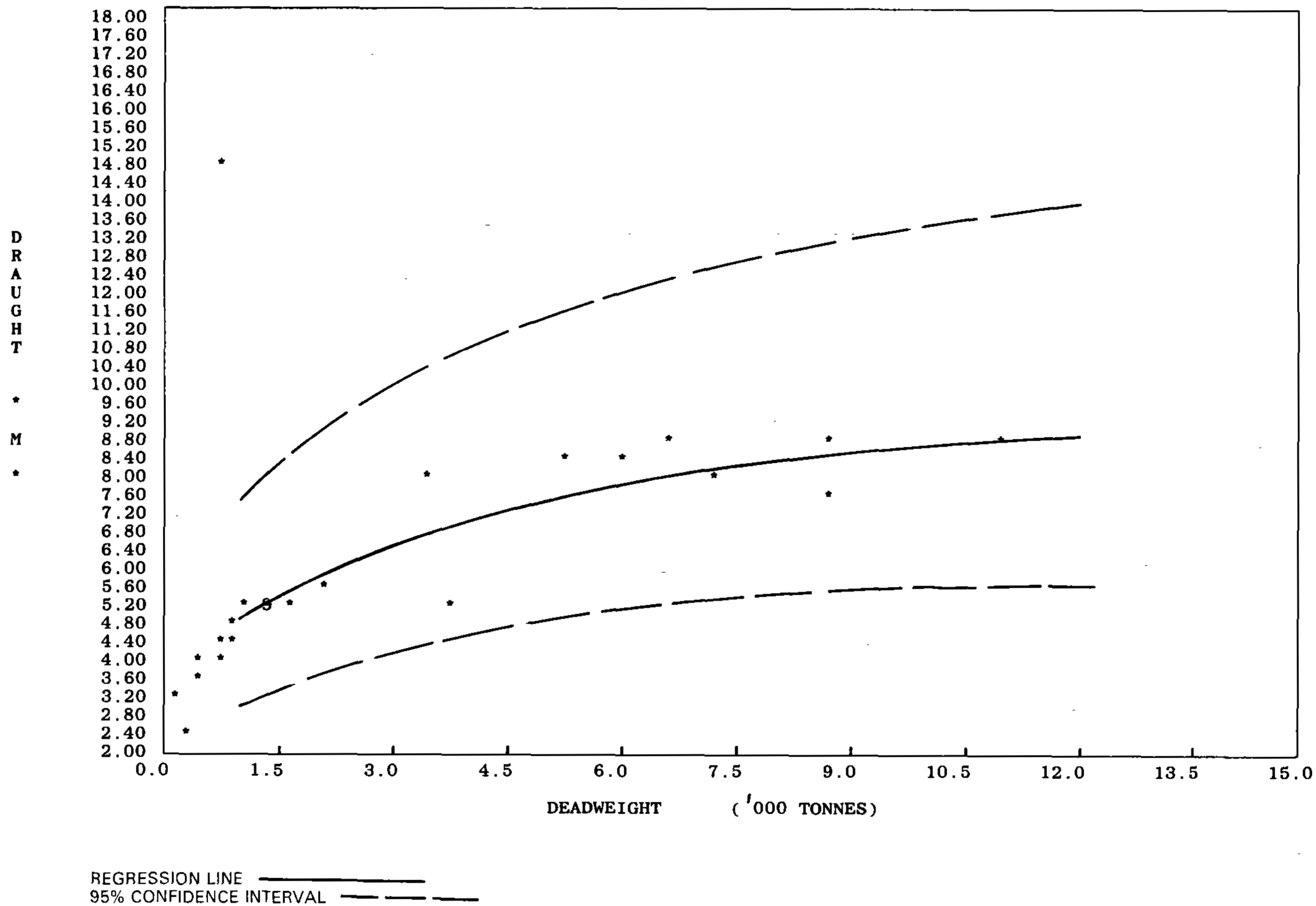


FIGURE 24 DRAUGHT VS DEADWEIGHT (PASSENGER SHIPS)

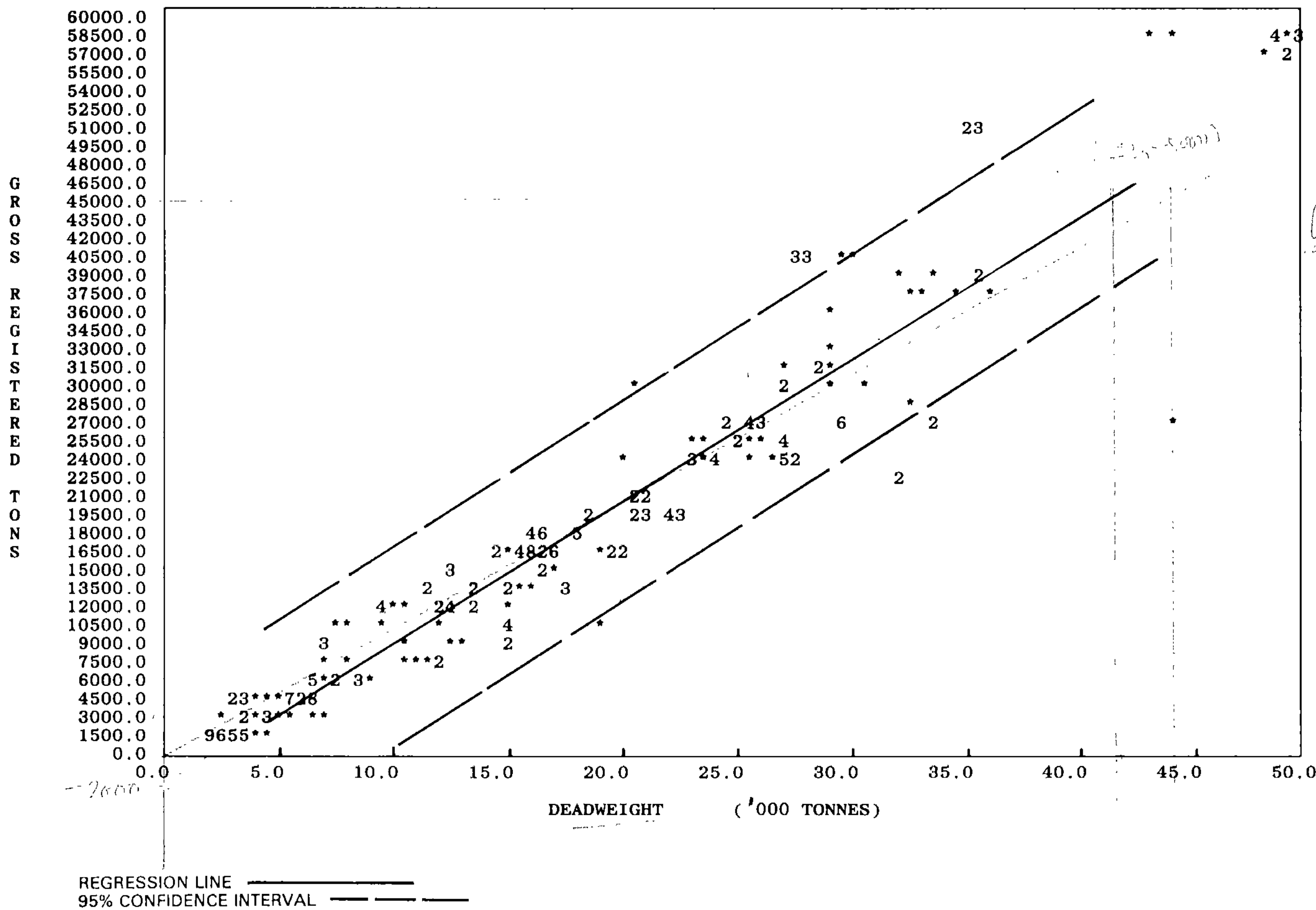


FIGURE 25 GROSS REGISTERED TONS VS DEADWEIGHT (CONTAINER SHIPS)

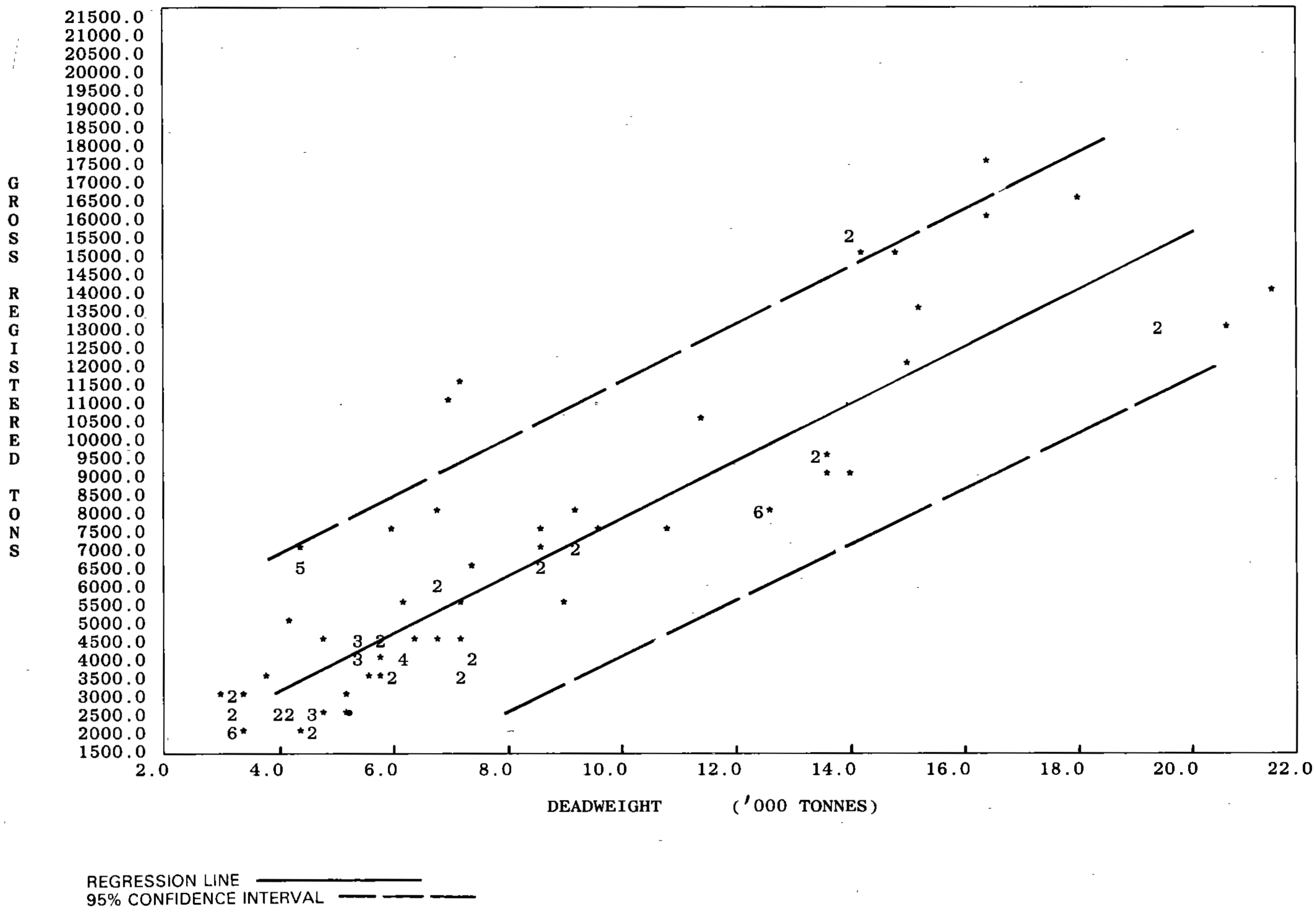


FIGURE 26 GROSS REGISTERED TONS VS DEADWEIGHT (RO. RO SHIPS)

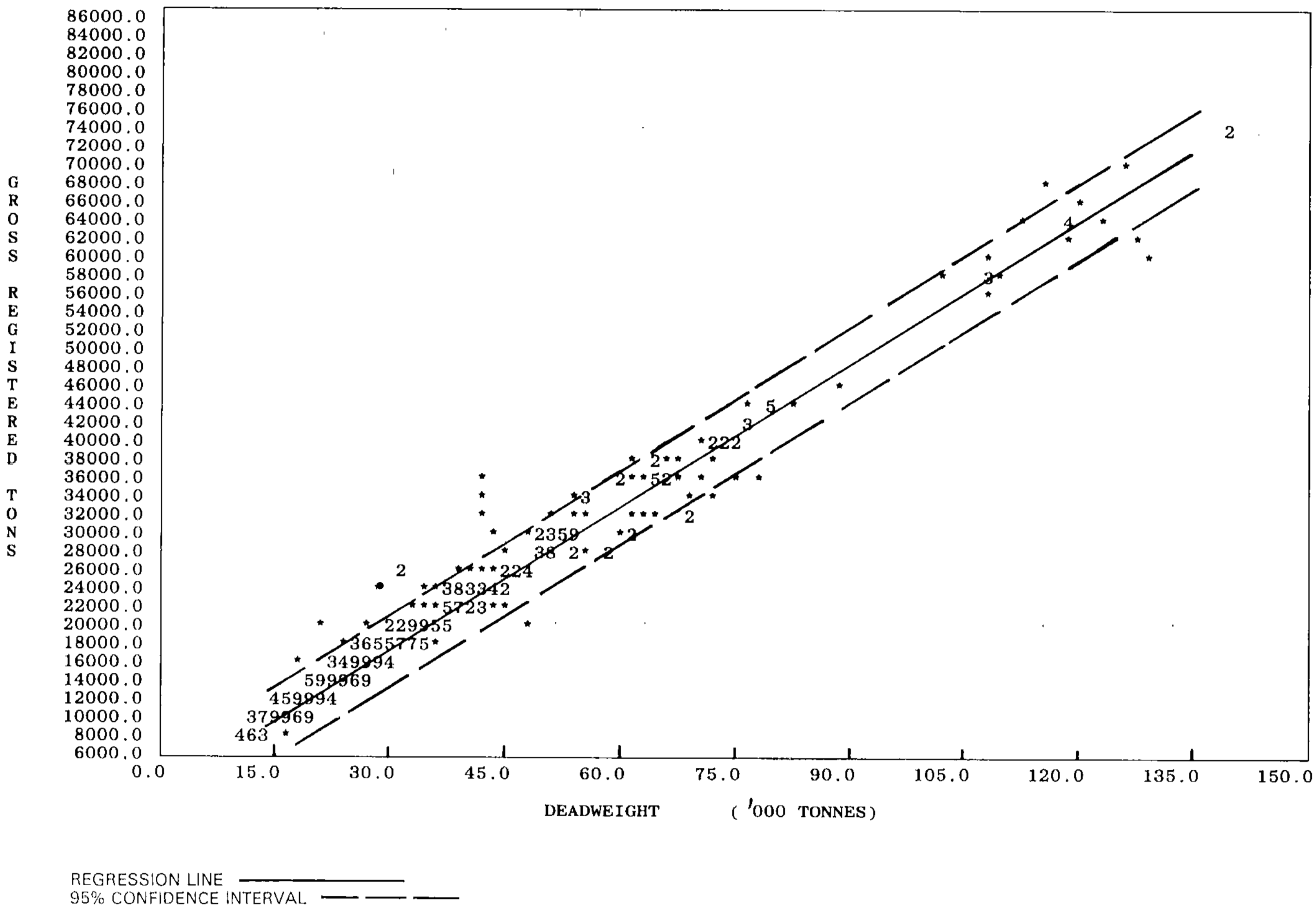


FIGURE 27 GROSS REGISTERED TONS VS DEADWEIGHT (BULK CARRIERS)

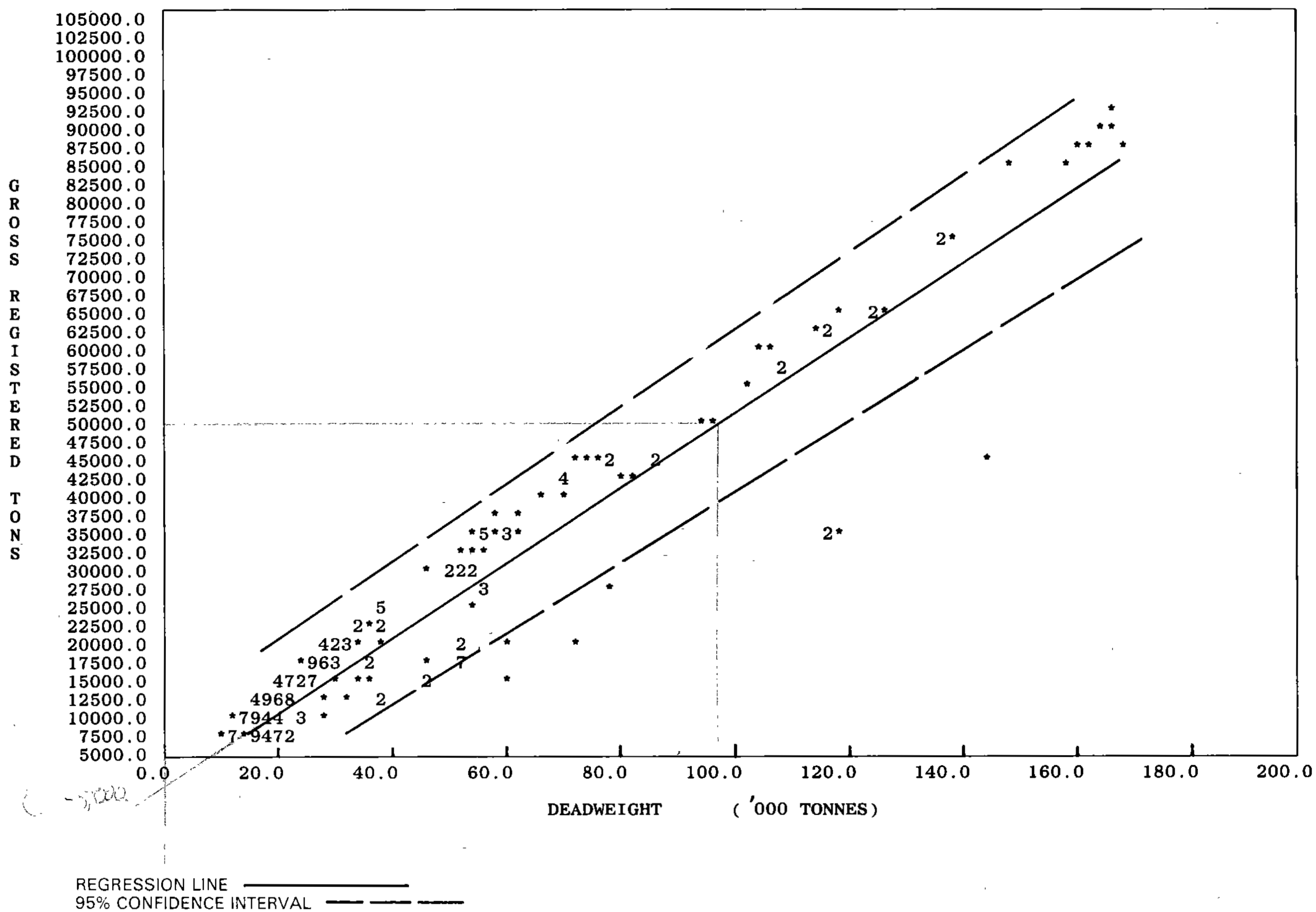


FIGURE 28 GROSS REGISTERED TONS VS DEADWEIGHT (ORE CARRIERS)

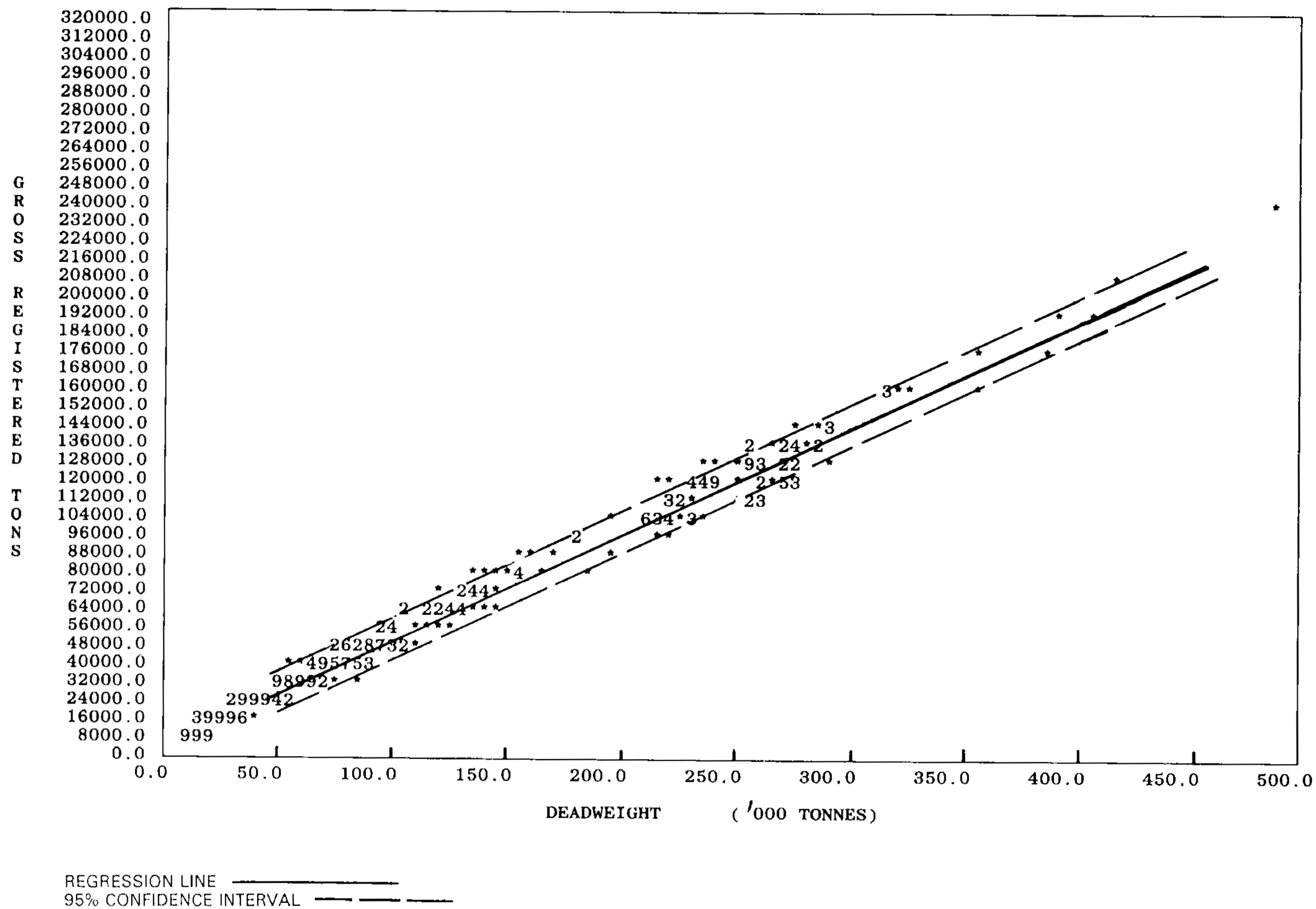


FIGURE 29 GROSS REGISTERED TONS VS DEADWEIGHT (TANKERS)

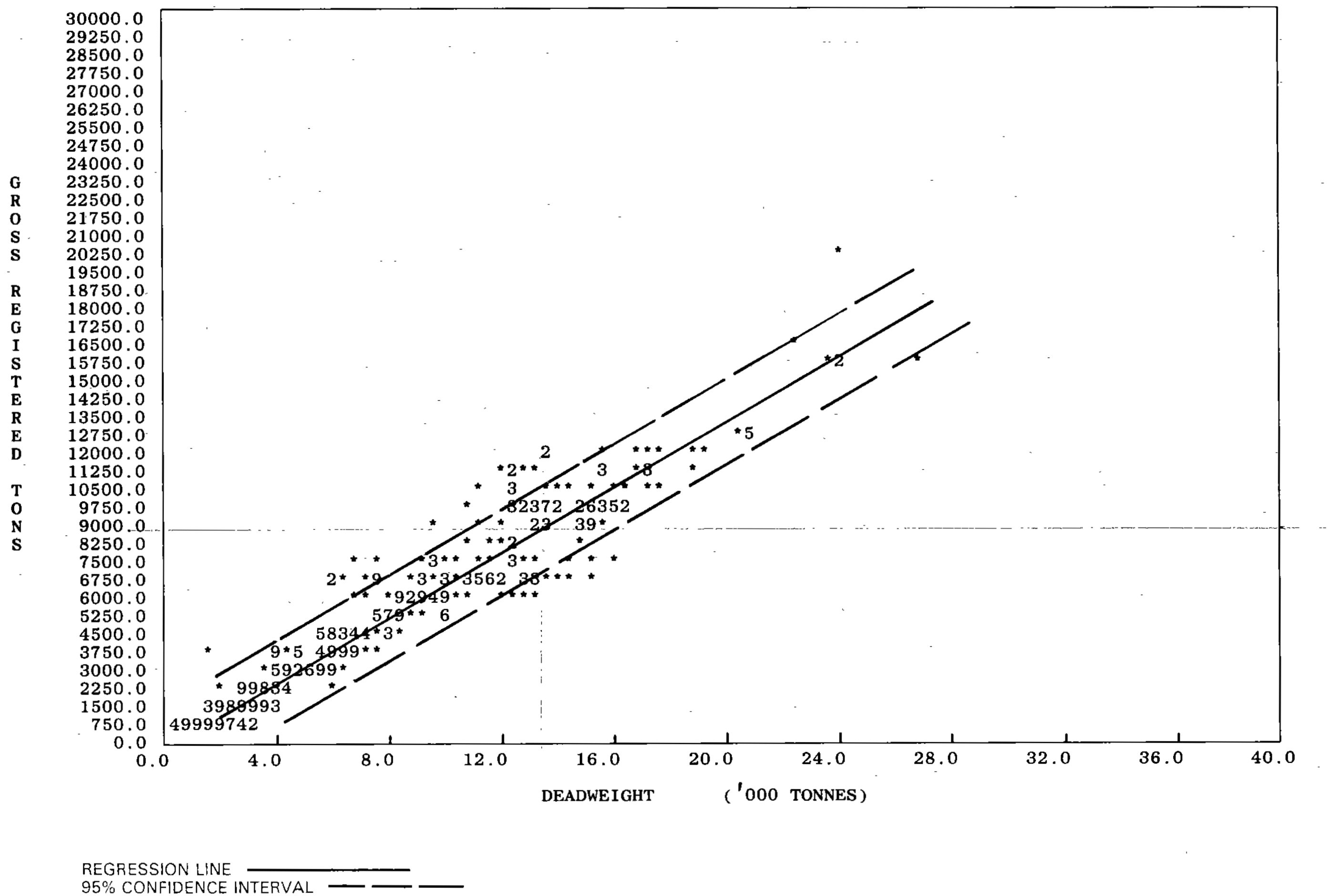


FIGURE 30 GROSS REGISTERED TONS VS DEADWEIGHT (GENERAL CARGO SHIPS)

25474/80—5

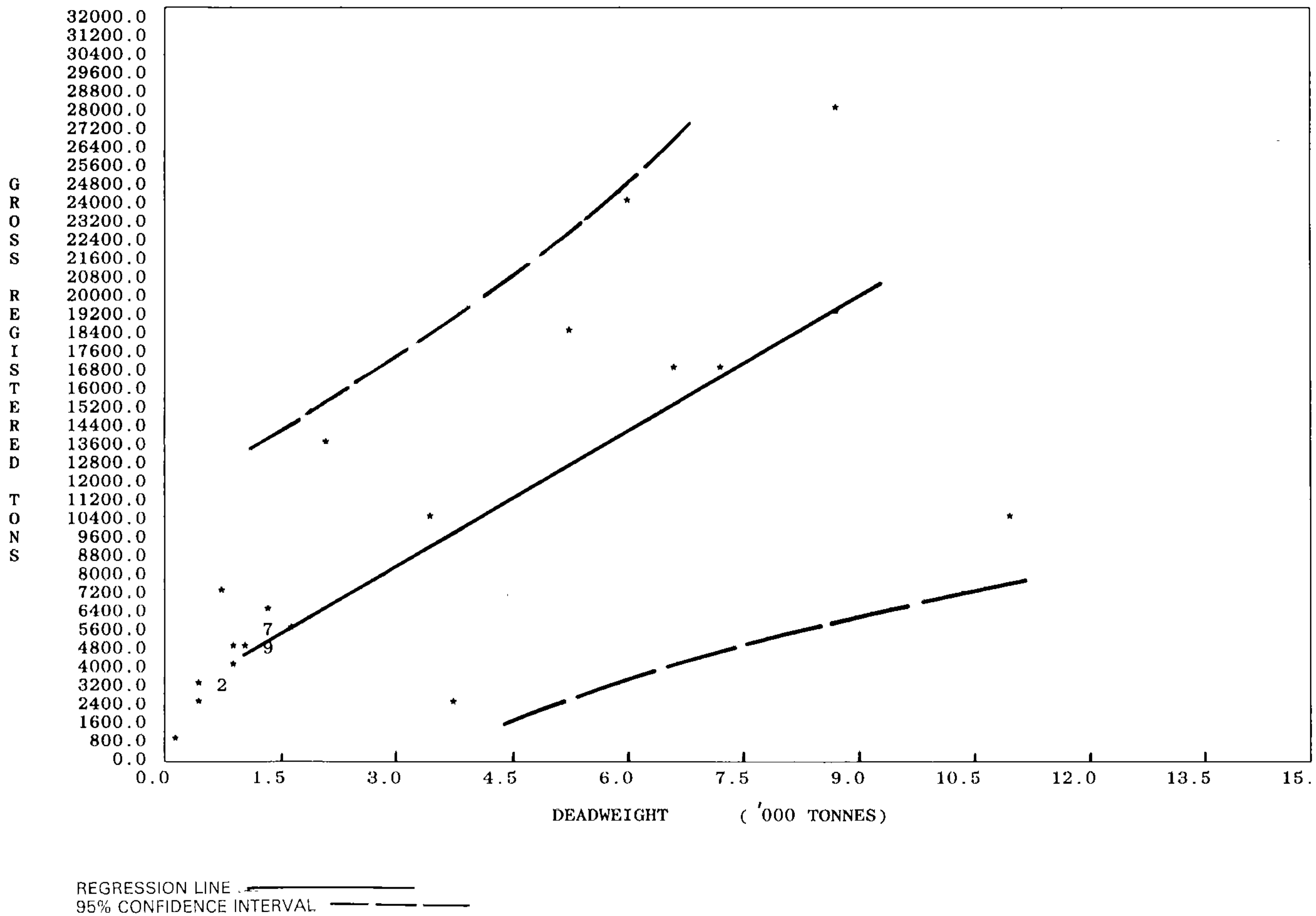


FIGURE 31 GROSS REGISTERED TONS VS DEADWEIGHT (PASSENGER SHIPS)

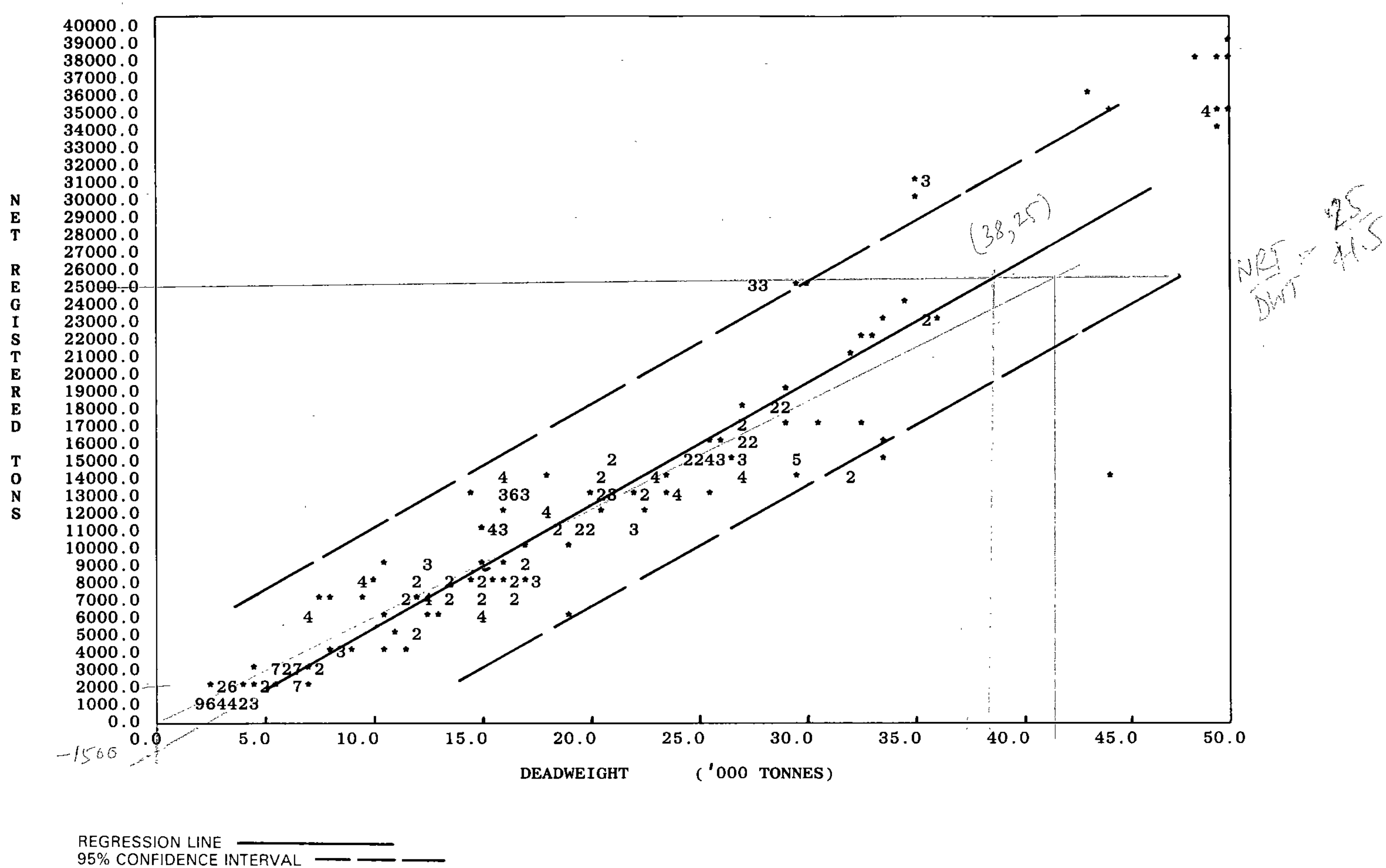


FIGURE 32 NET REGISTERED TONS VS DEADWEIGHT (CONTAINER SHIPS)

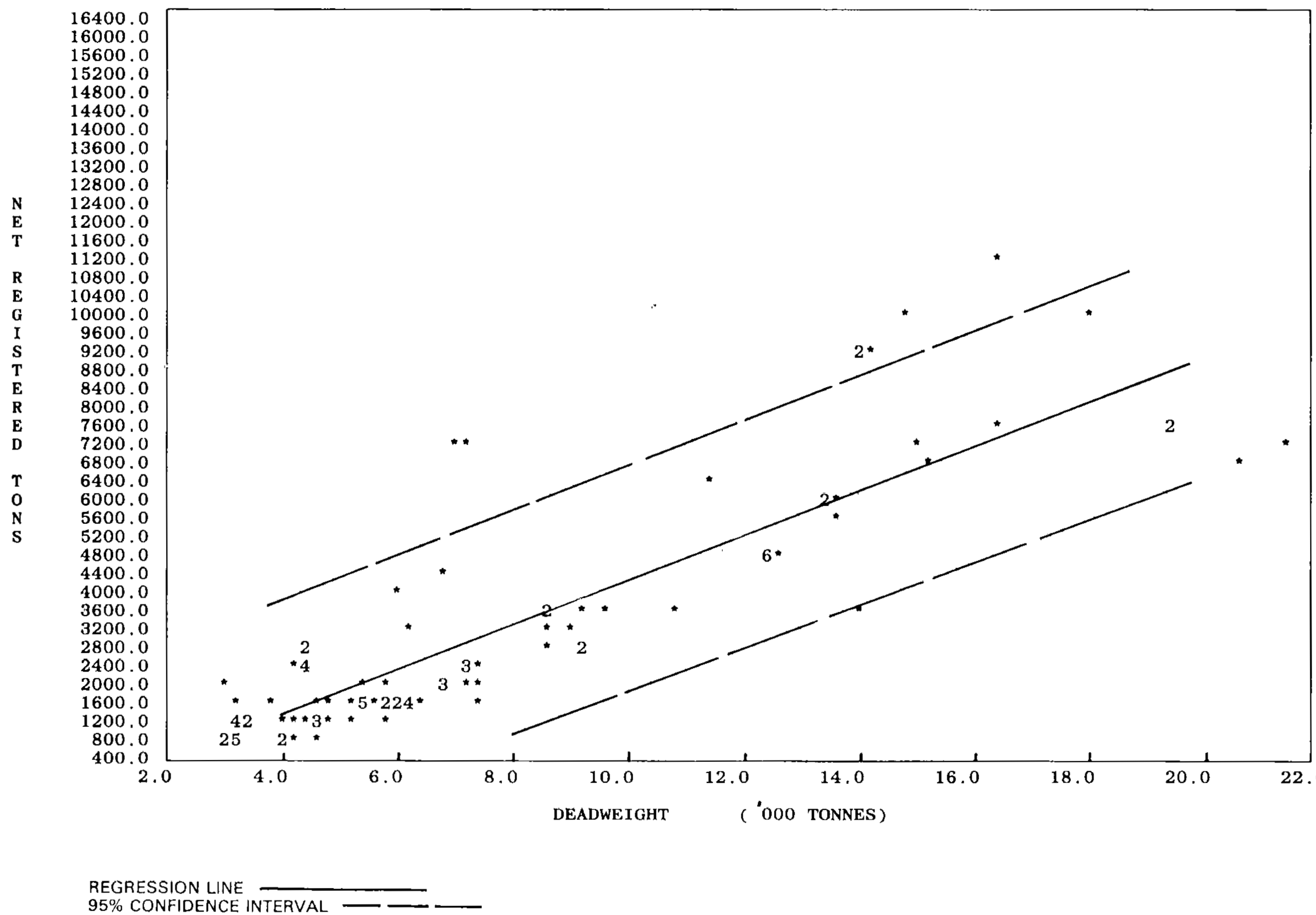
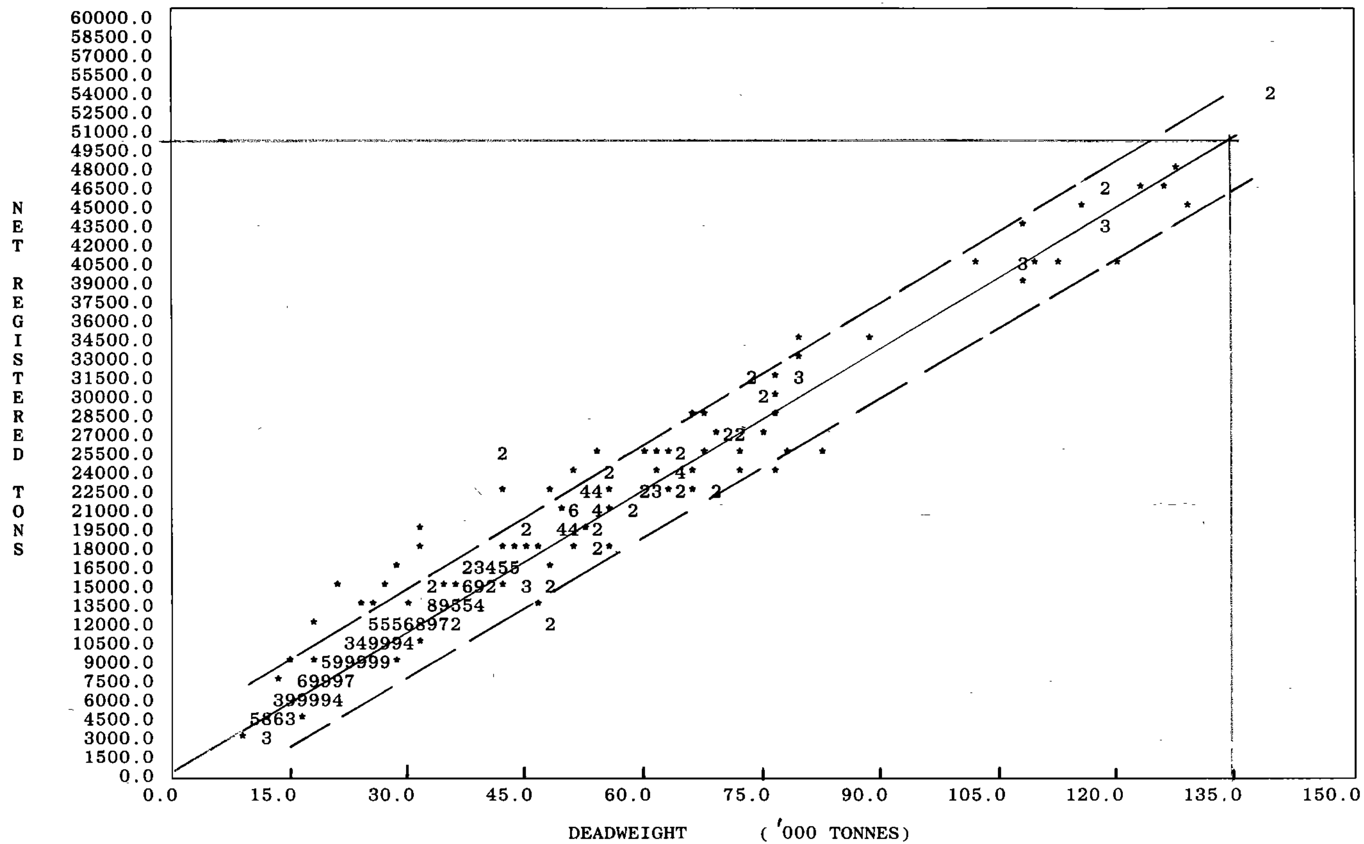


FIGURE 33 NET REGISTERED TONS VS DEADWEIGHT (RO. RO SHIPS)



REGRESSION LINE ————————
95% CONFIDENCE INTERVAL — — —

NRT = (10/27) DWT

FIGURE 34 NET REGISTERED TONS VS DEADWEIGHT (BULK CARRIERS)

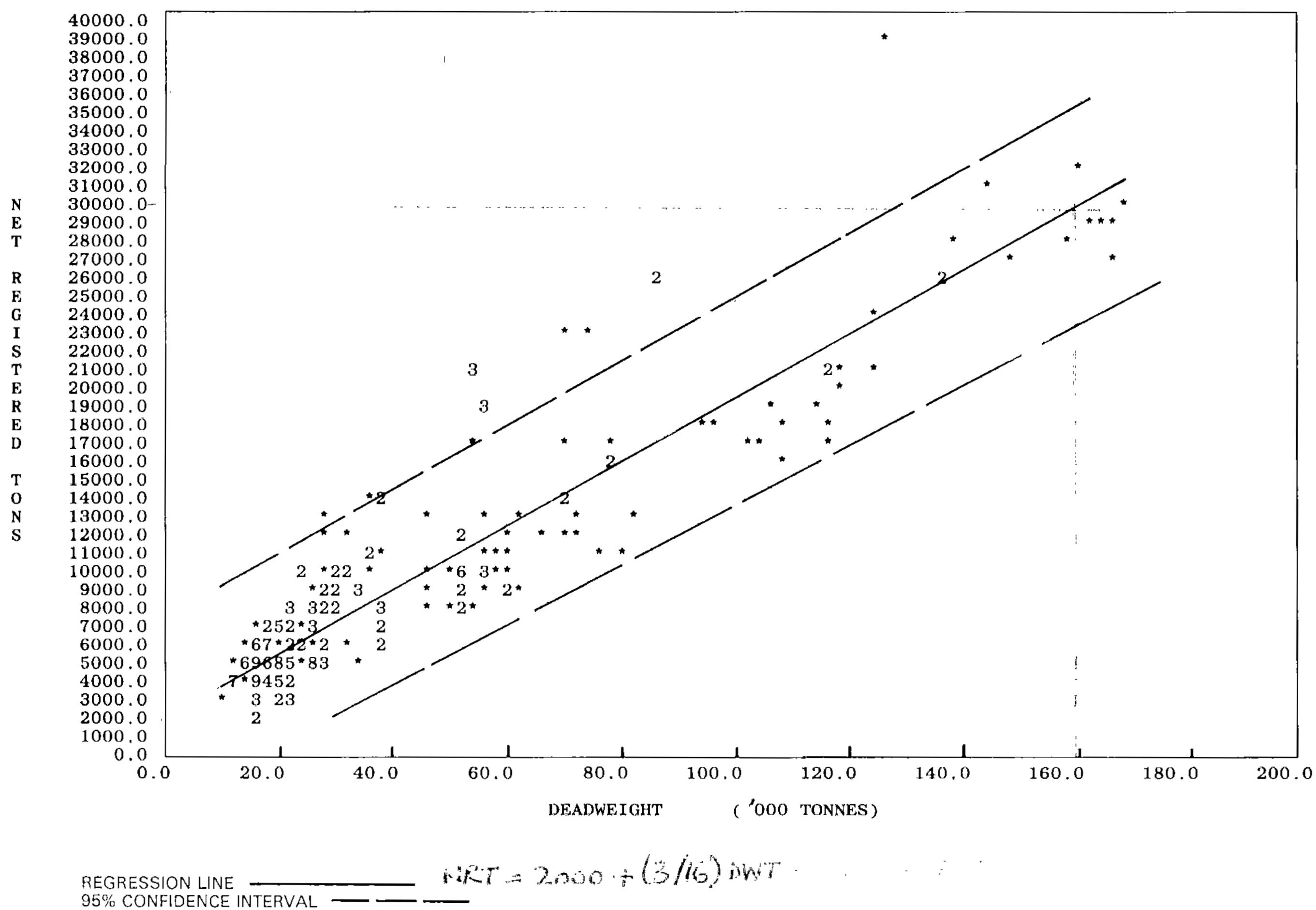


FIGURE 35 NET REGISTERED TONS VS DEADWEIGHT (ORE CARRIERS)

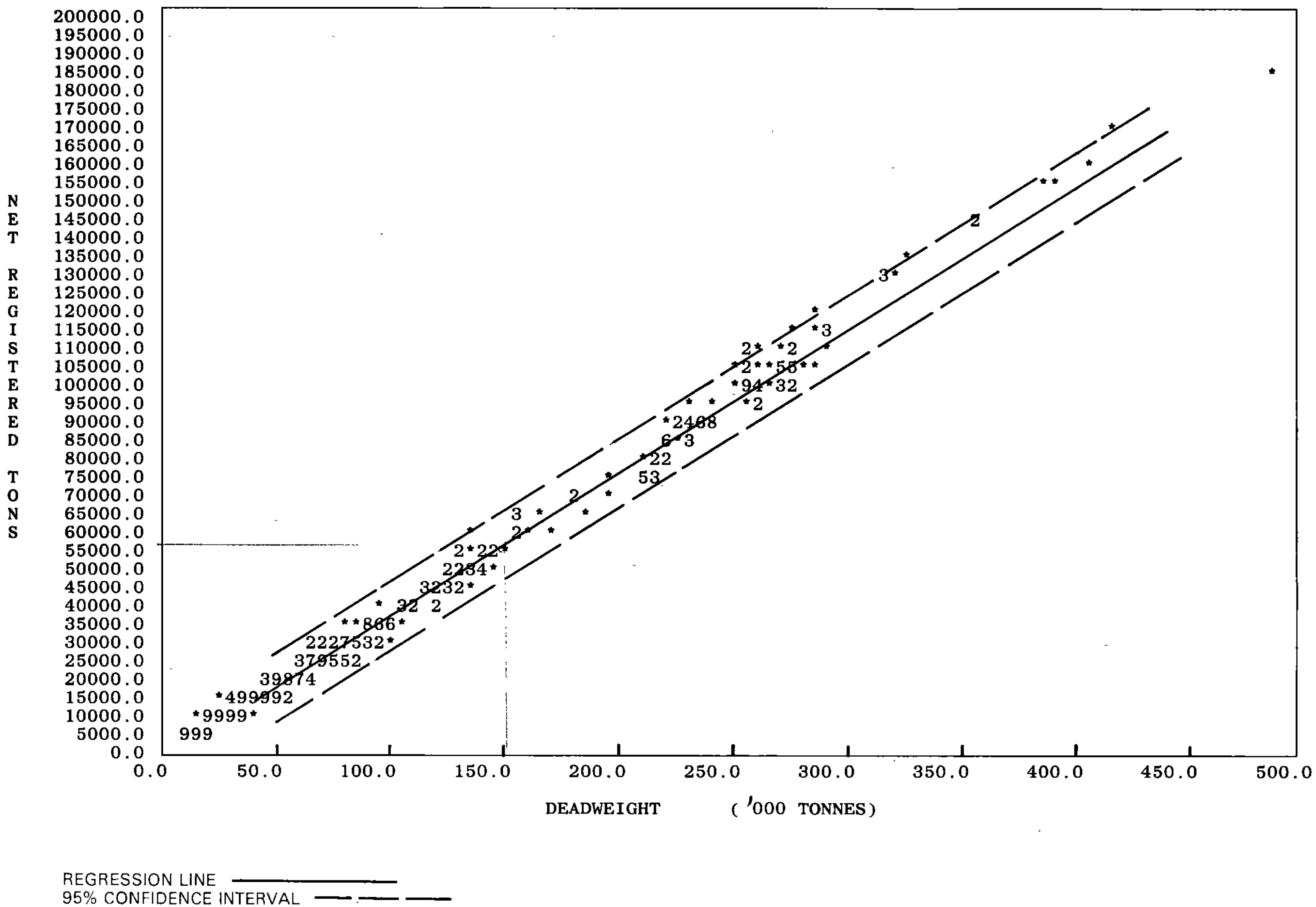


FIGURE 36 NET REGISTERED TONS VS DEADWEIGHT (TANKERS)

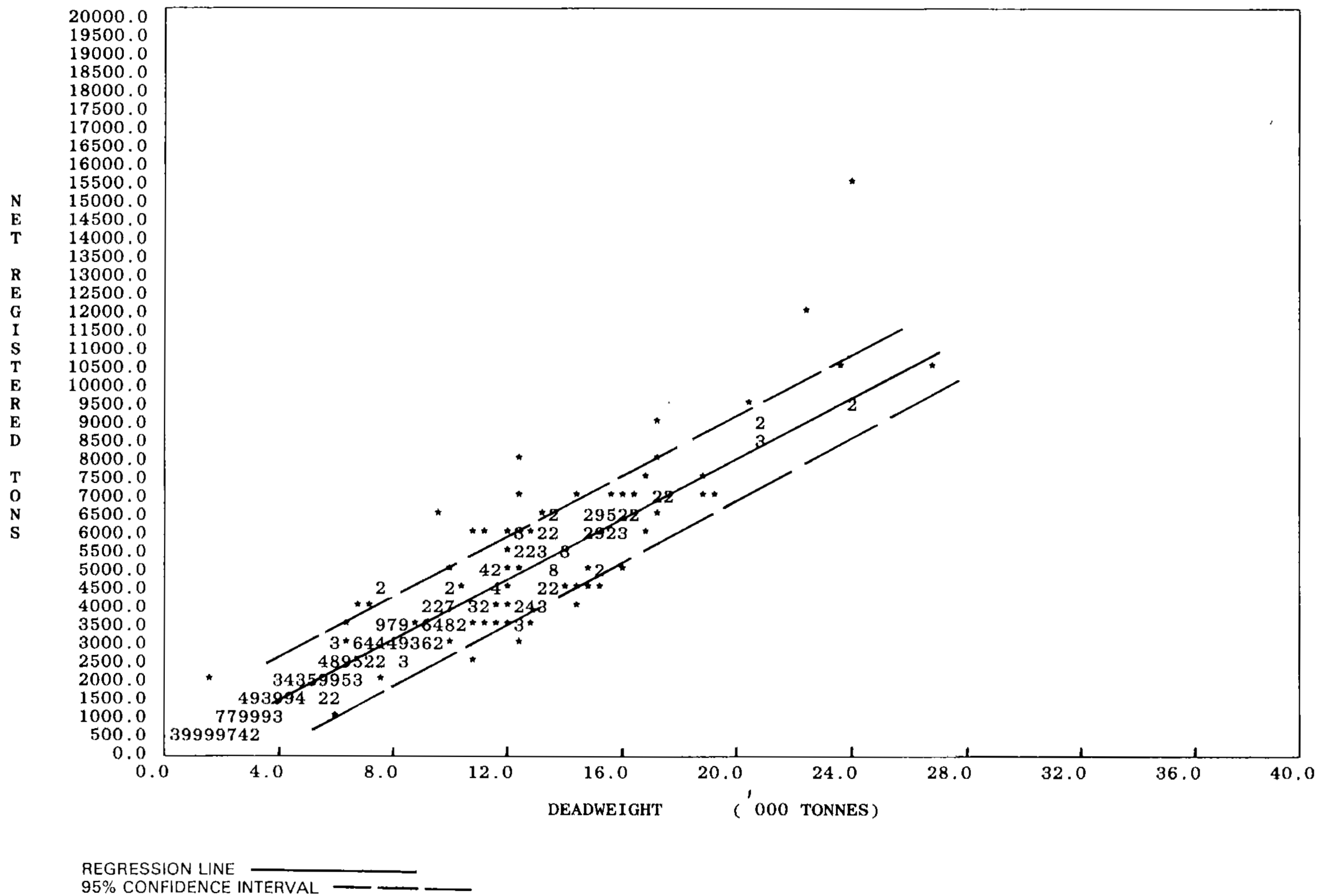


FIGURE 37 NET REGISTERED TONS VS DEADWEIGHT (GENERAL CARGO SHIPS)

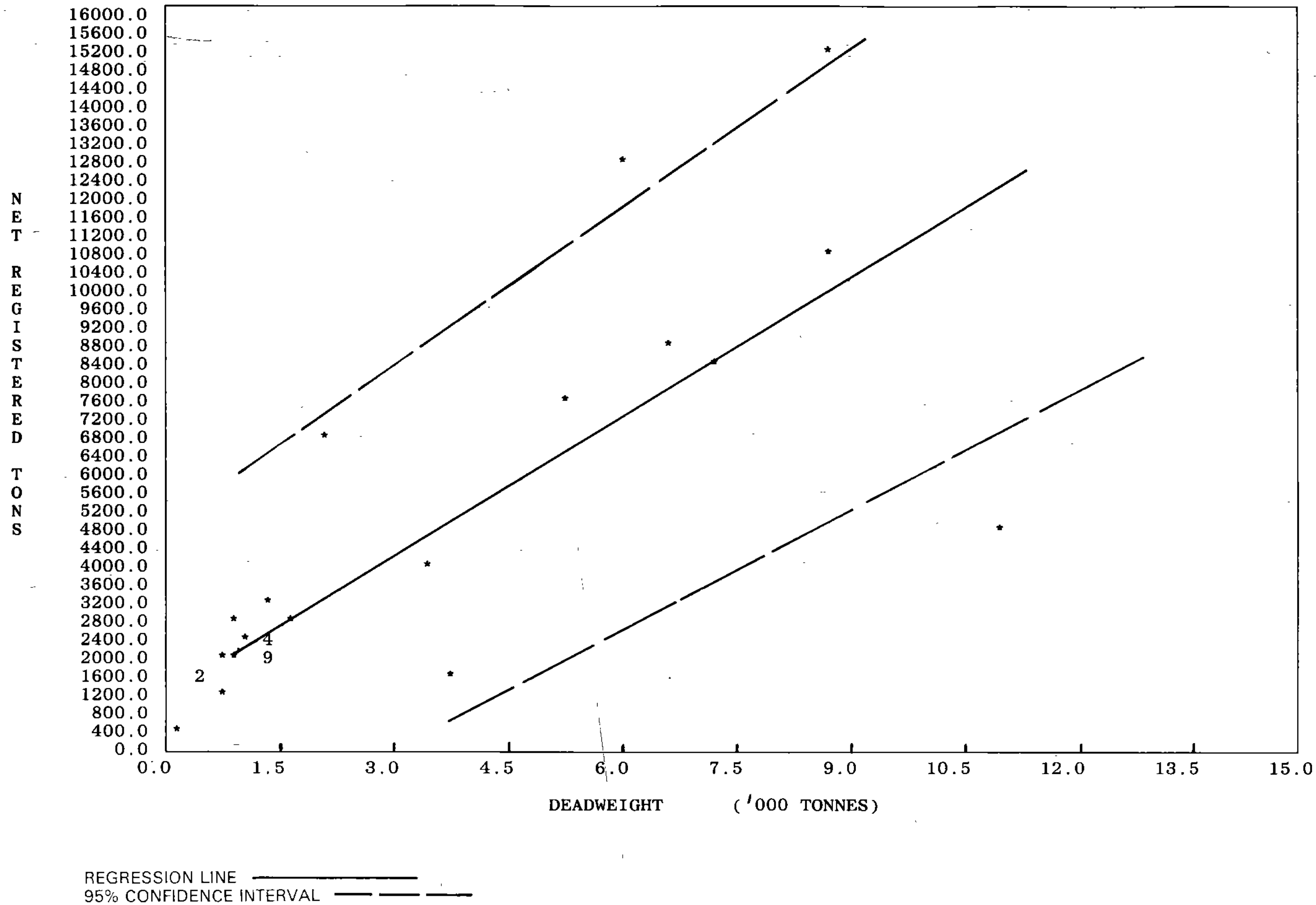


FIGURE 38 NET REGISTERED TONS VS DEADWEIGHT (PASSENGER SHIPS)

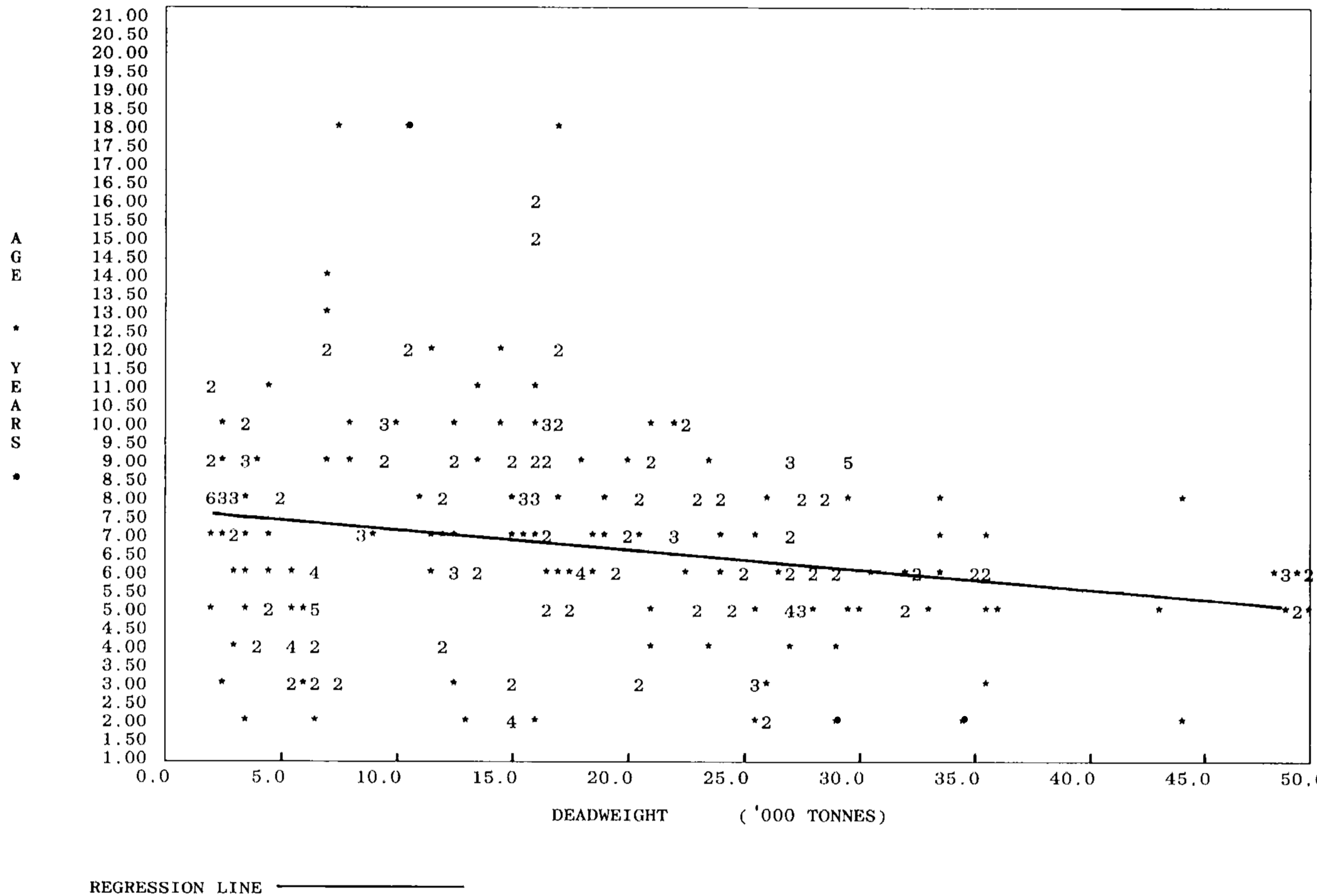


FIGURE 39 AGE VS DEADWEIGHT (CONTAINER SHIPS)

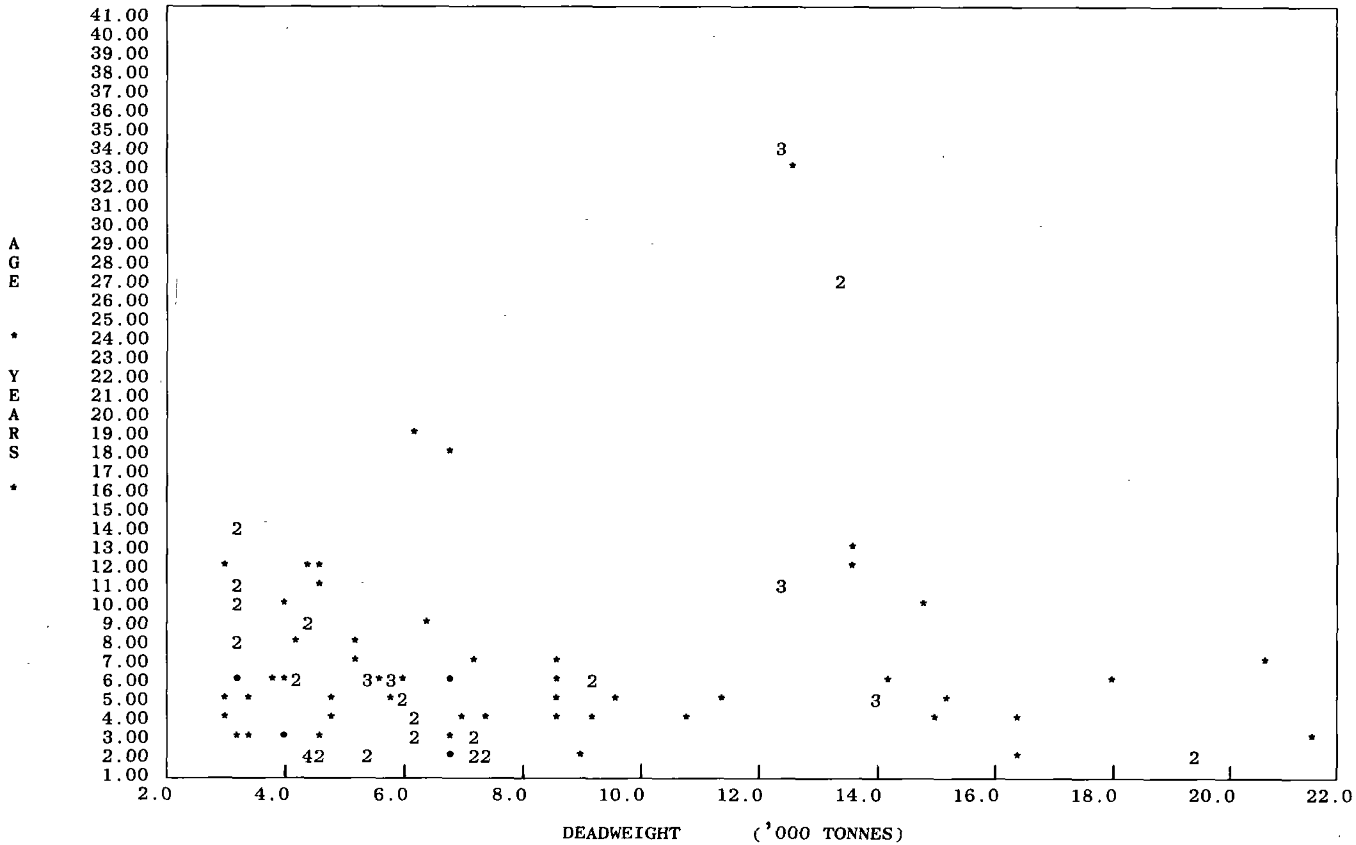


FIGURE 40 AGE VS DEADWEIGHT (RO. RO SHIPS)

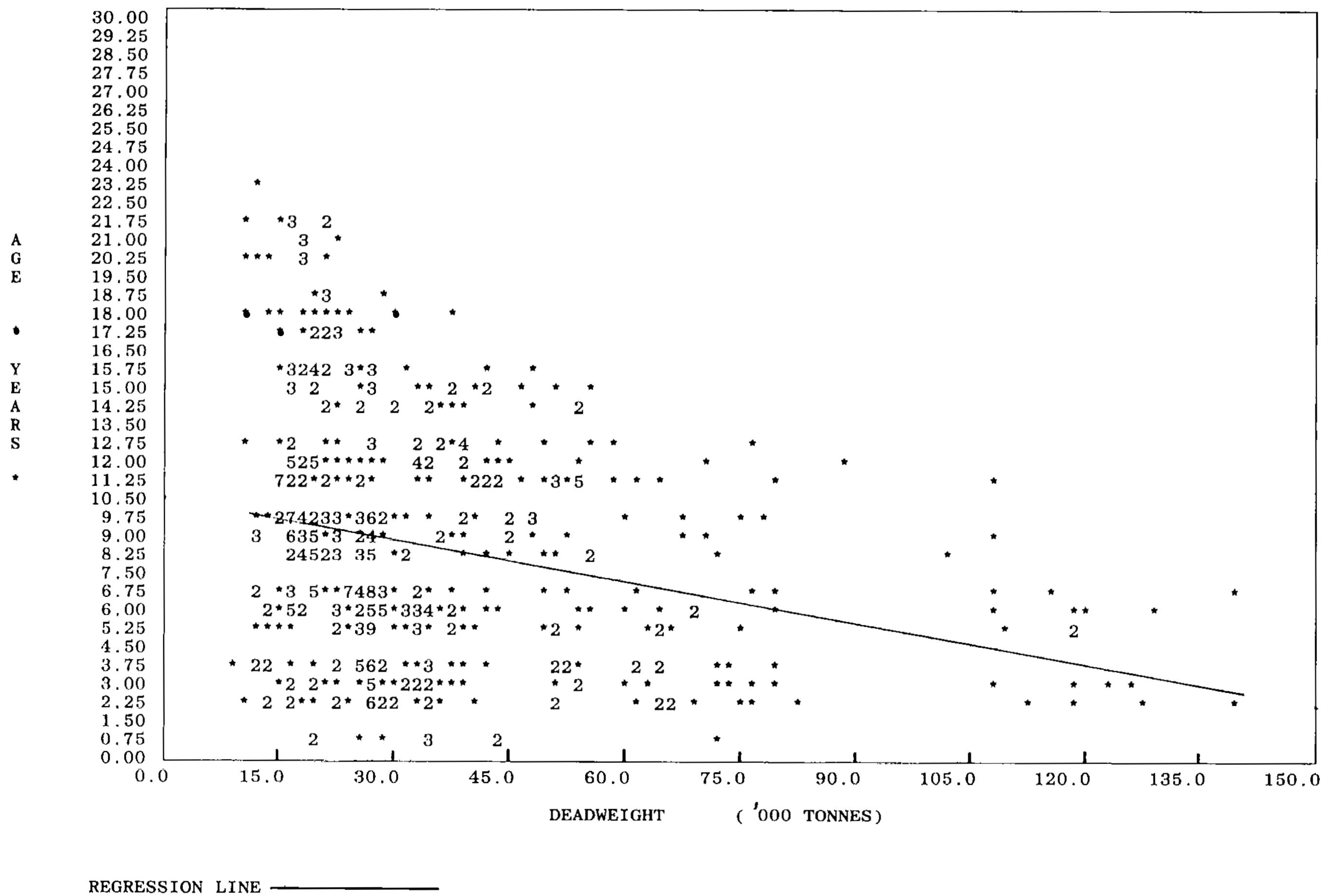


FIGURE 41 AGE VS DEADWEIGHT (BULK CARRIERS)

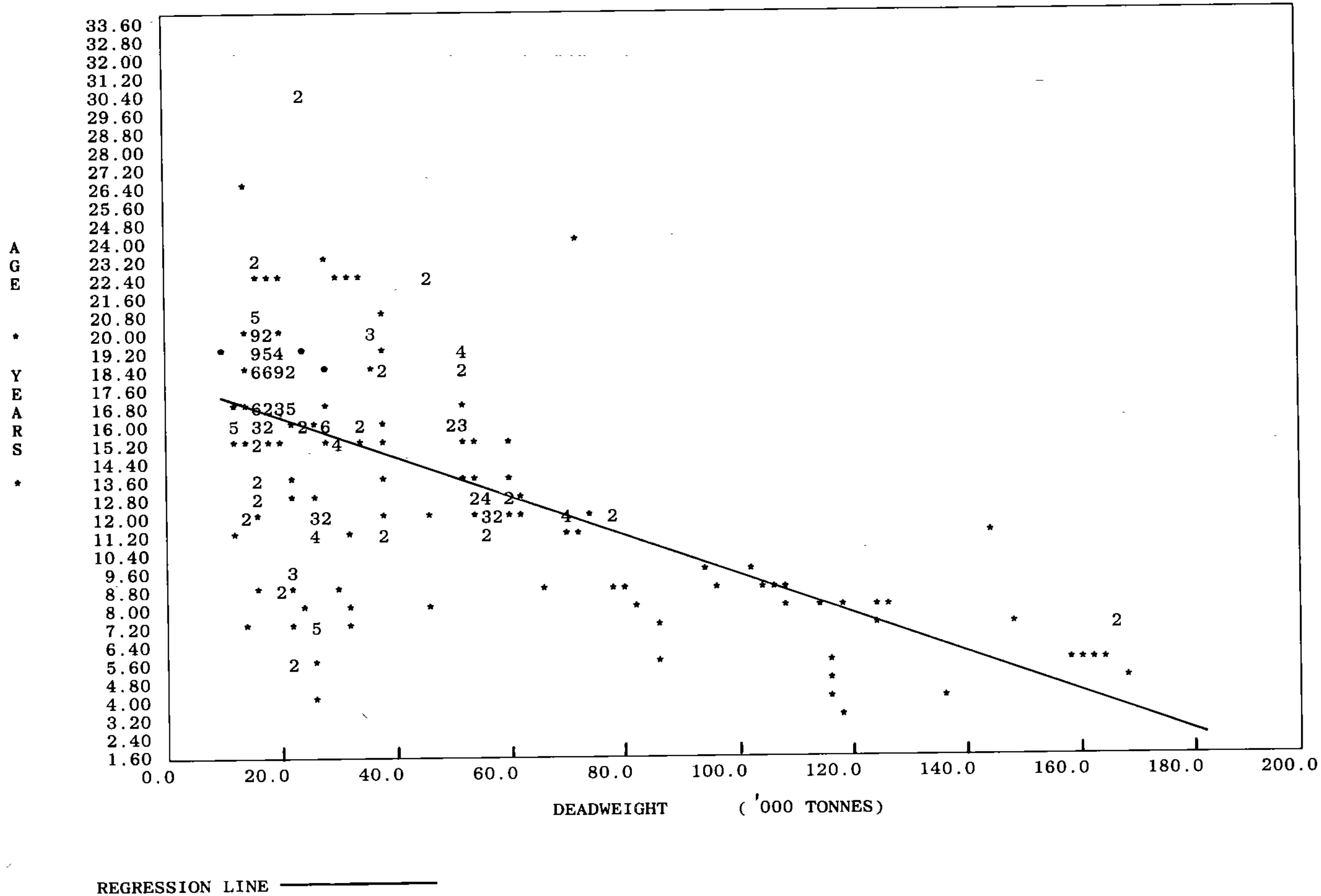


FIGURE 42 AGE VS DEADWEIGHT (ORE CARRIERS)

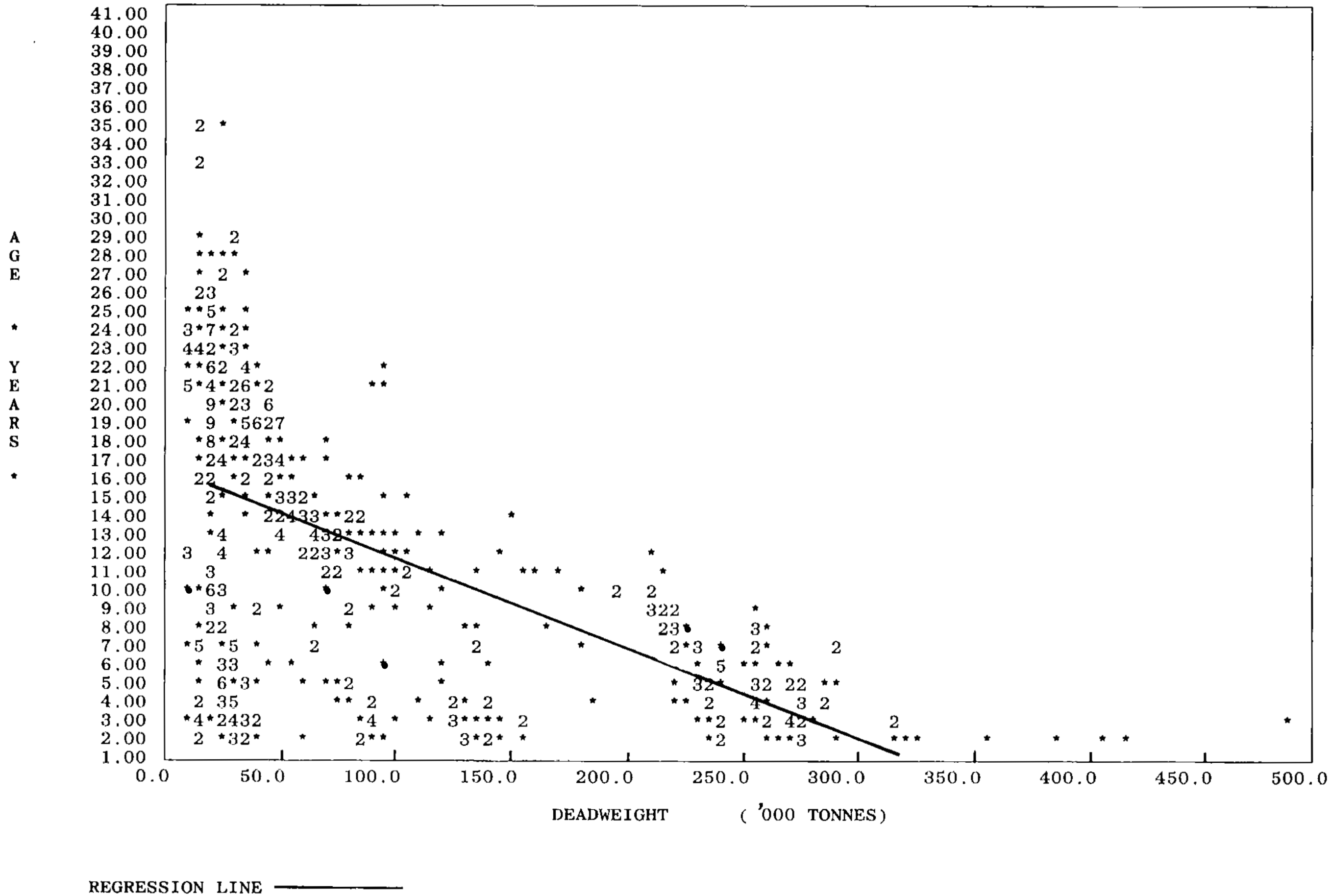


FIGURE 43 AGE VS DEADWEIGHT (TANKERS)

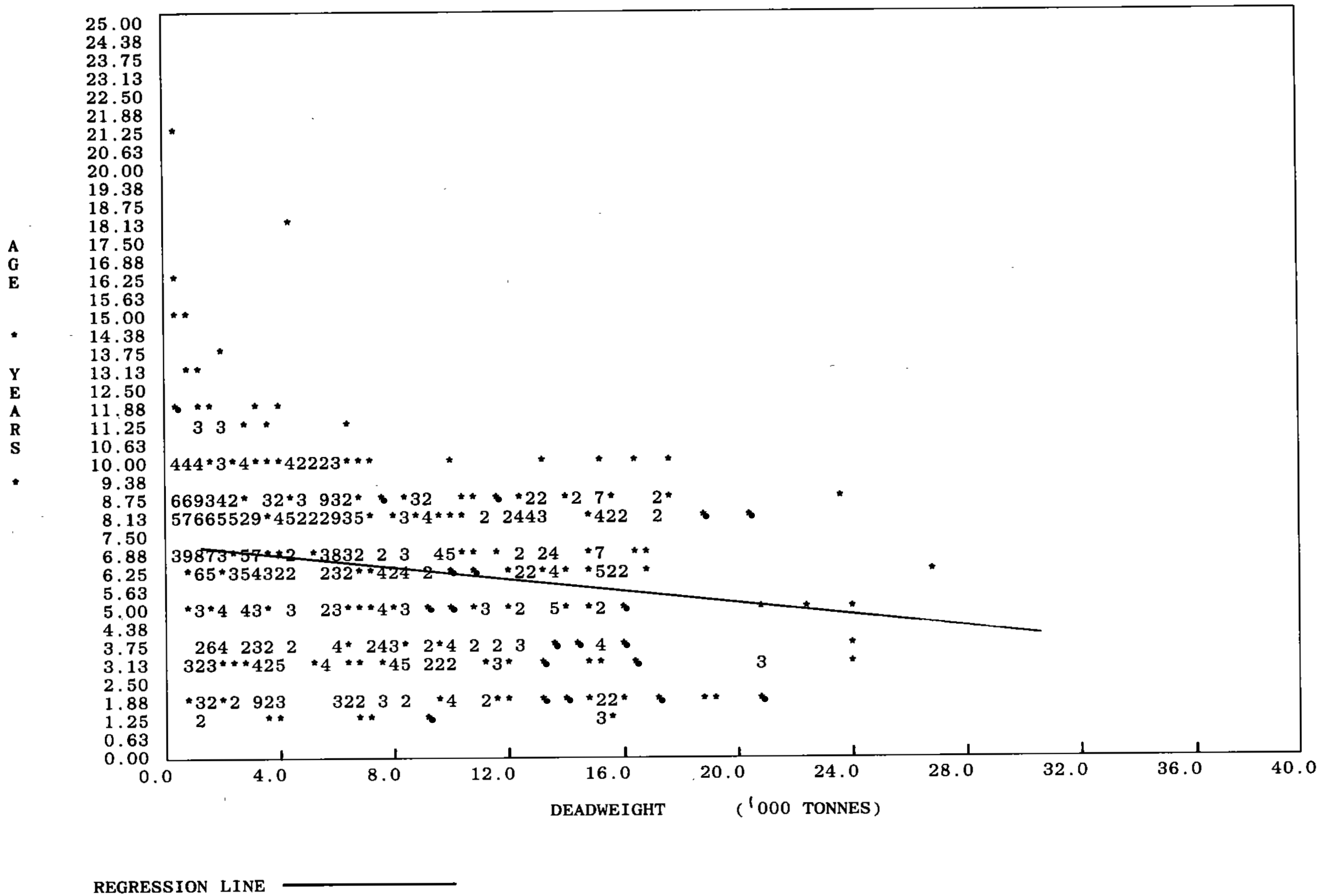


FIGURE 44 AGE VS DEADWEIGHT (GENERAL CARGO SHIPS)

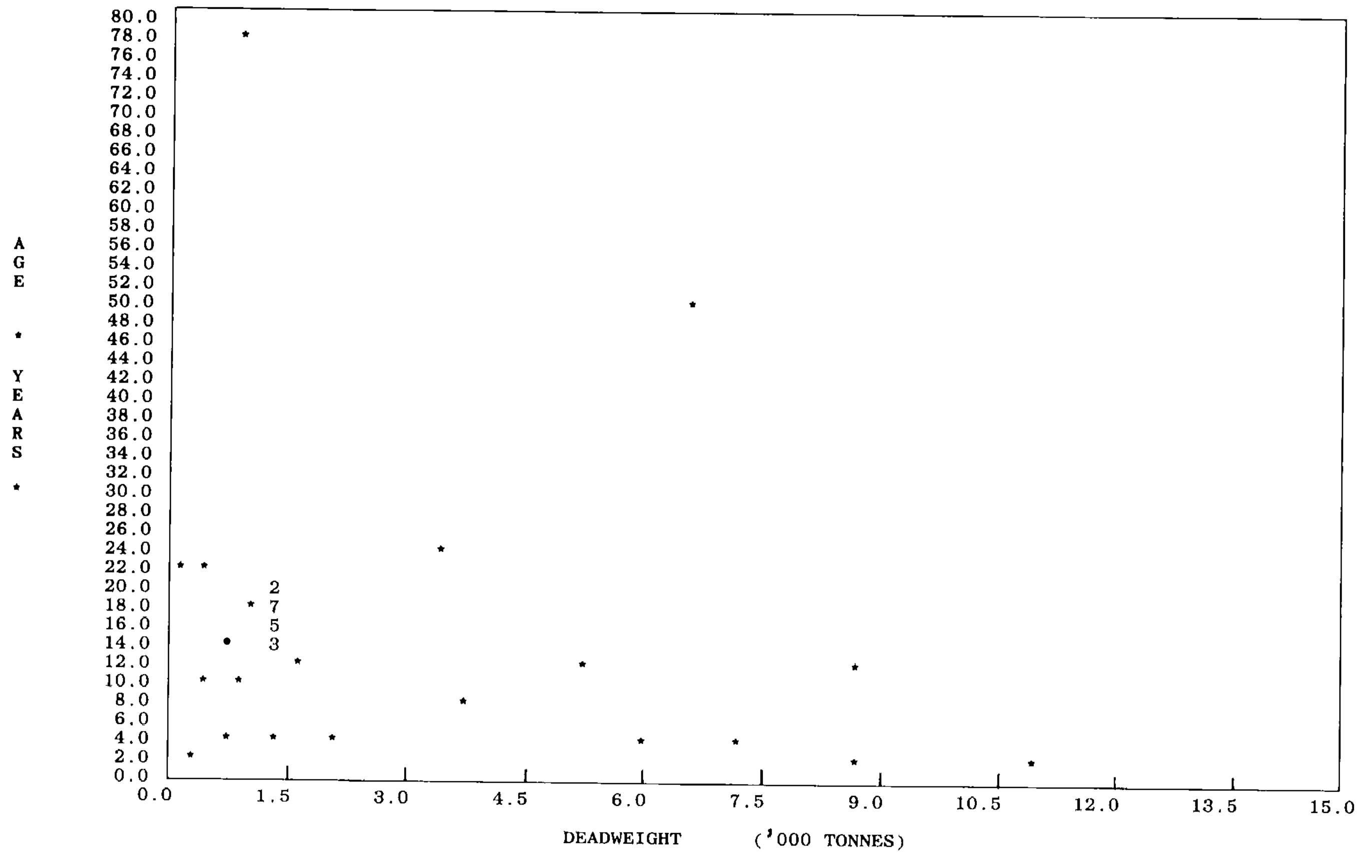


FIGURE 45 AGE VS DEADWEIGHT (PASSENGER SHIPS)

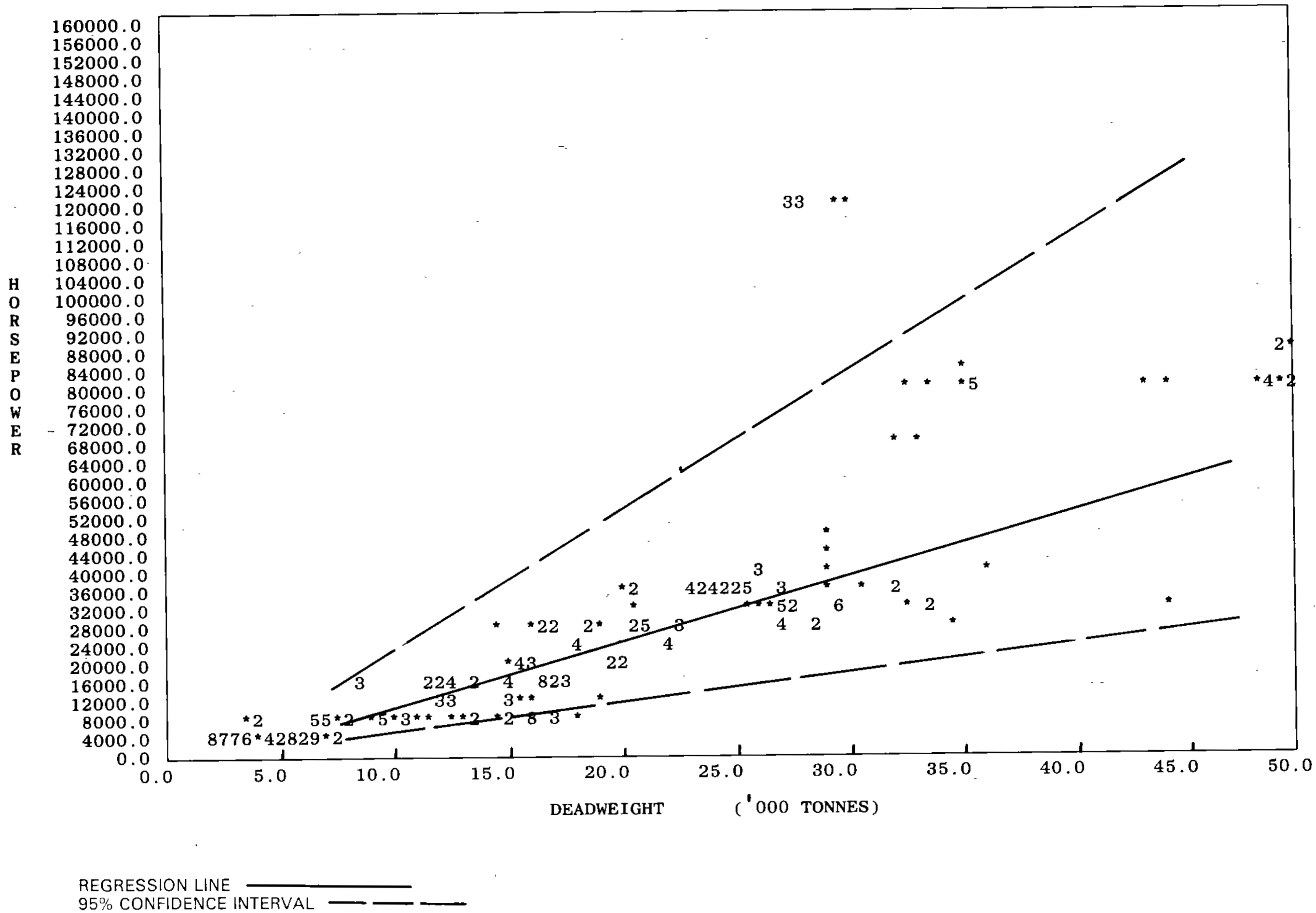


FIGURE 46 HORSEPOWER VS DEADWEIGHT (CONTAINER SHIPS)

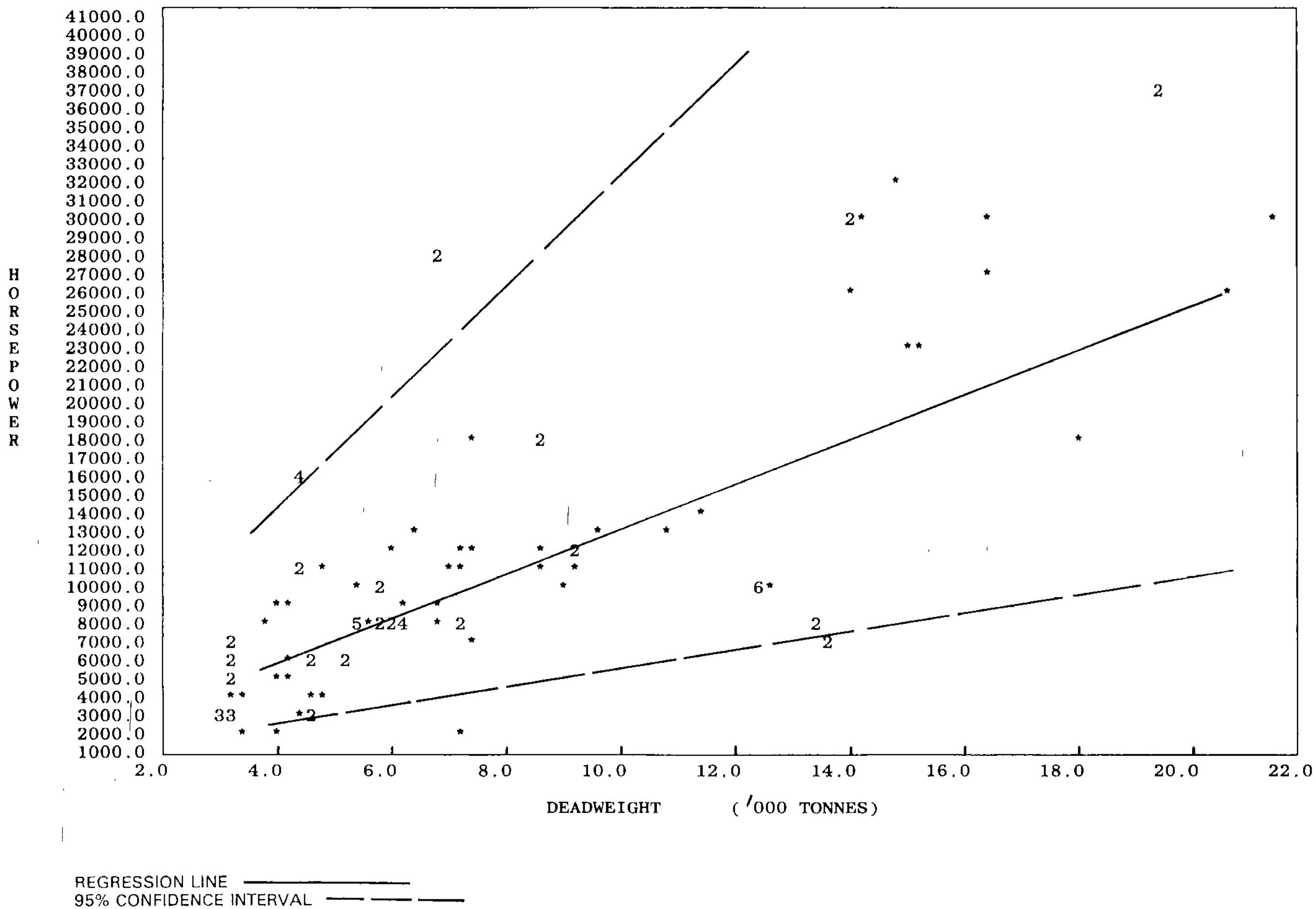


FIGURE 47 HORSEPOWER VS DEADWEIGHT (RO. RO SHIPS)

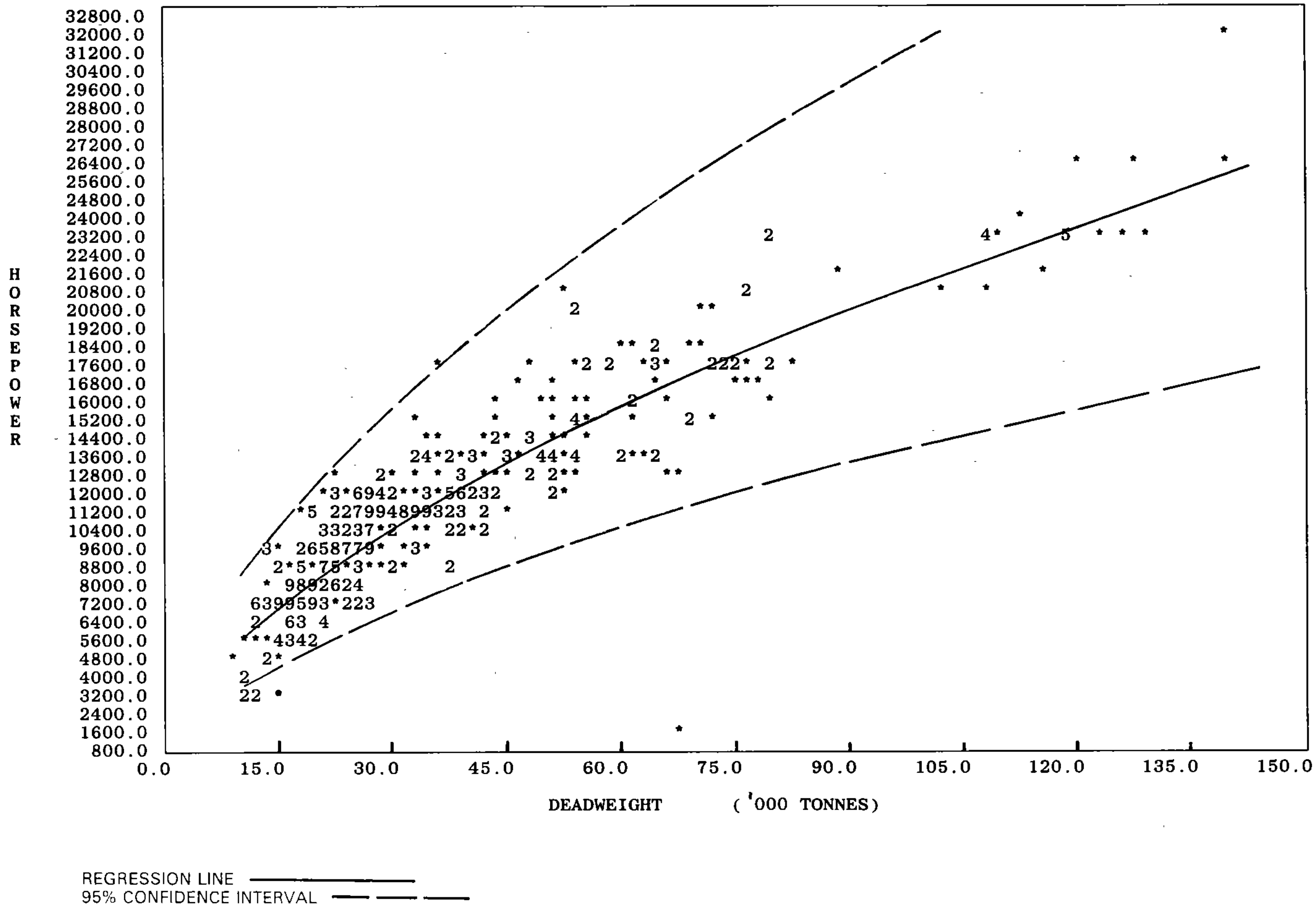


FIGURE 48 HORSEPOWER VS DEADWEIGHT (BULK CARRIERS)

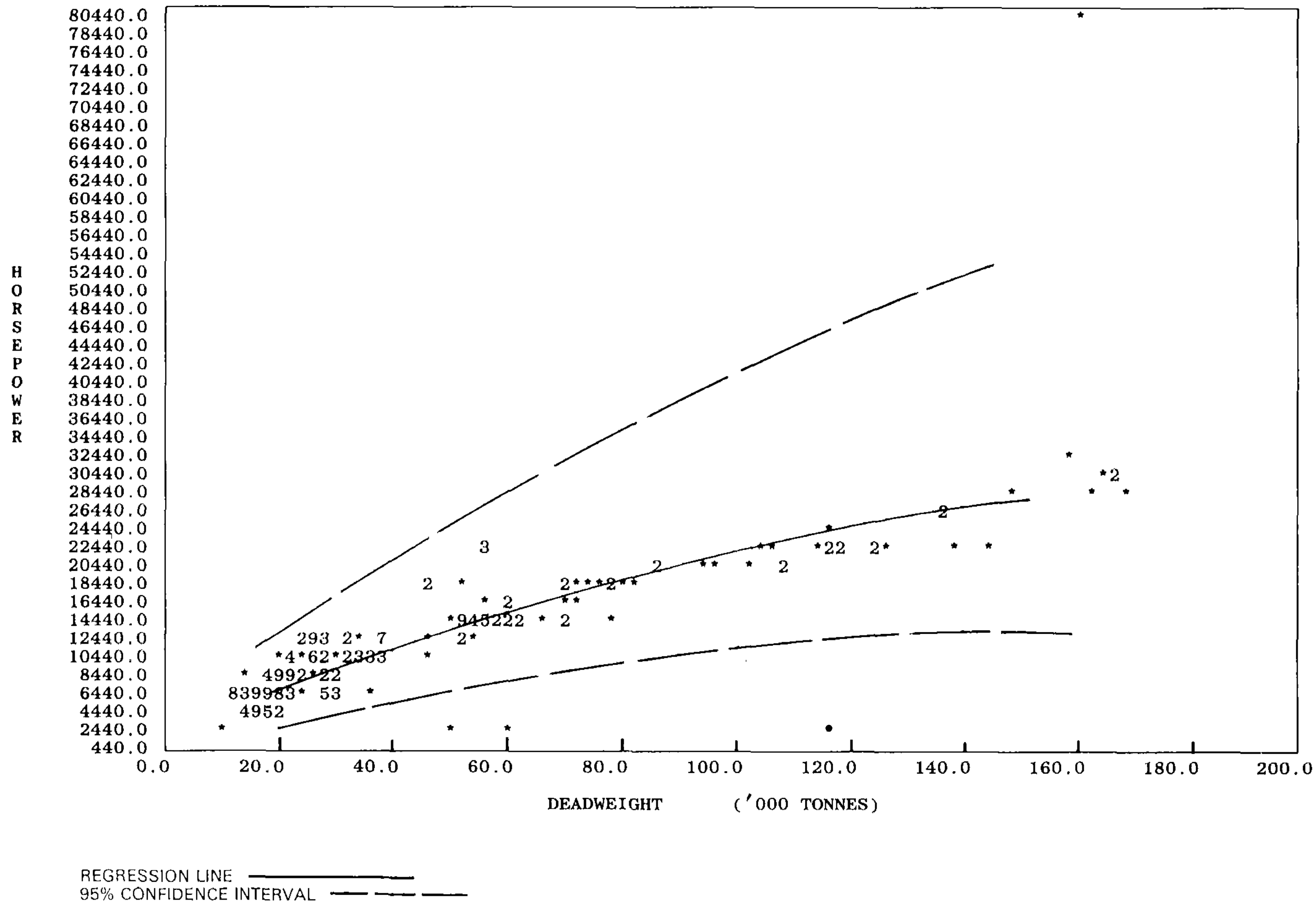
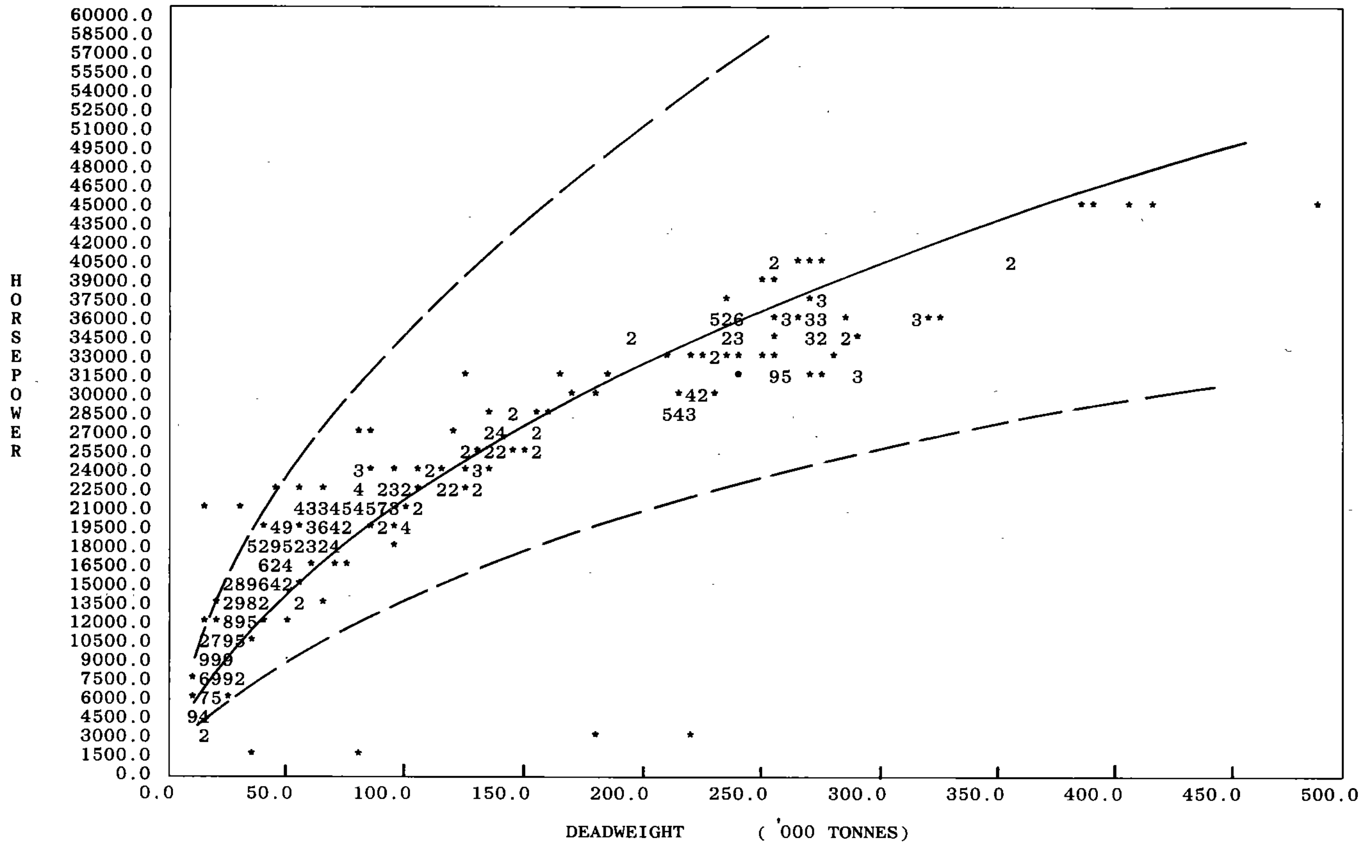


FIGURE 49 HORSEPOWER VS DEADWEIGHT (ORE CARRIERS)



REGRESSION LINE
95% CONFIDENCE INTERVAL

FIGURE 50 HORSEPOWER VS DEADWEIGHT (TANKERS)

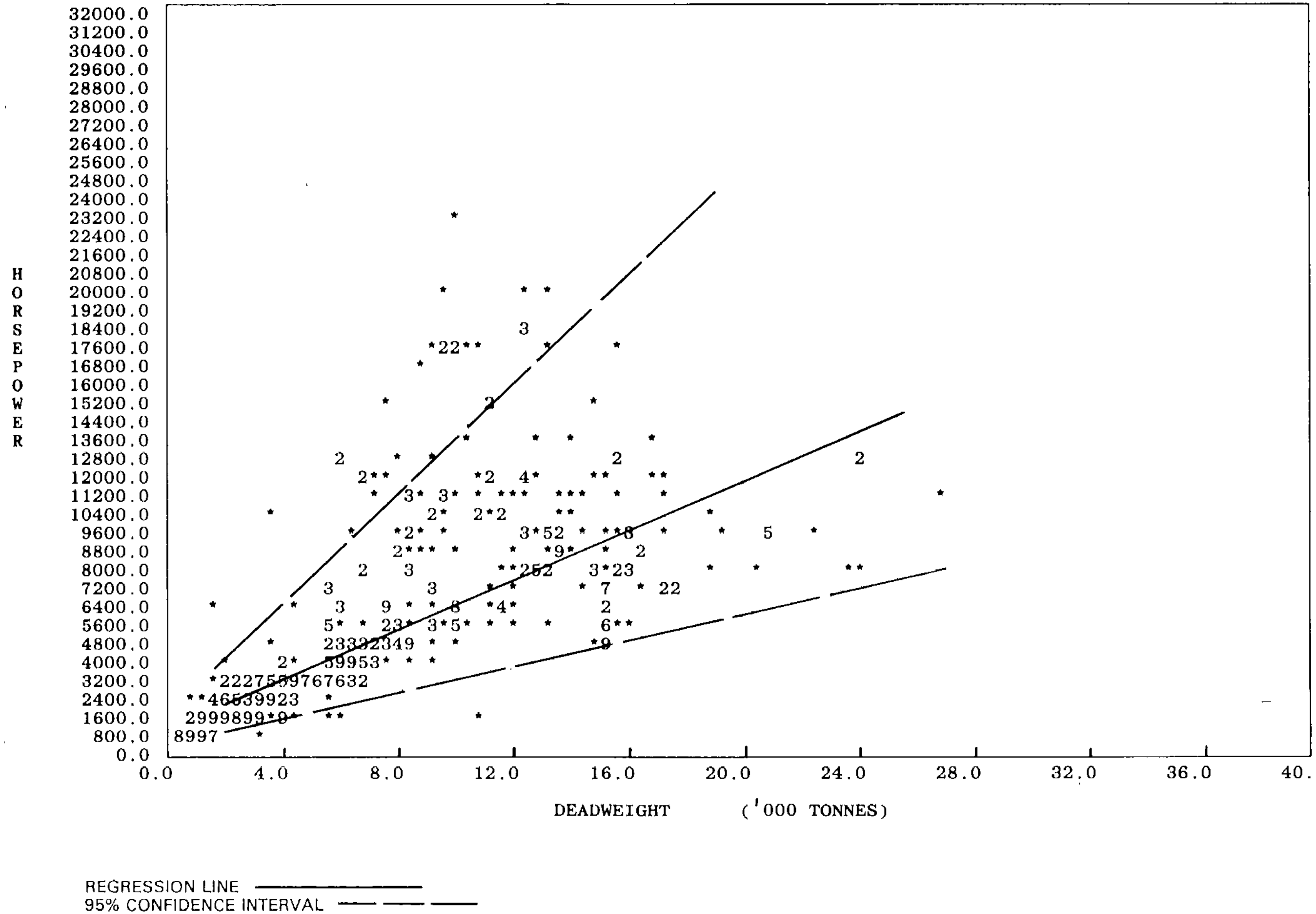


FIGURE 51 HORSEPOWER VS DEADWEIGHT (GENERAL CARGO SHIPS)

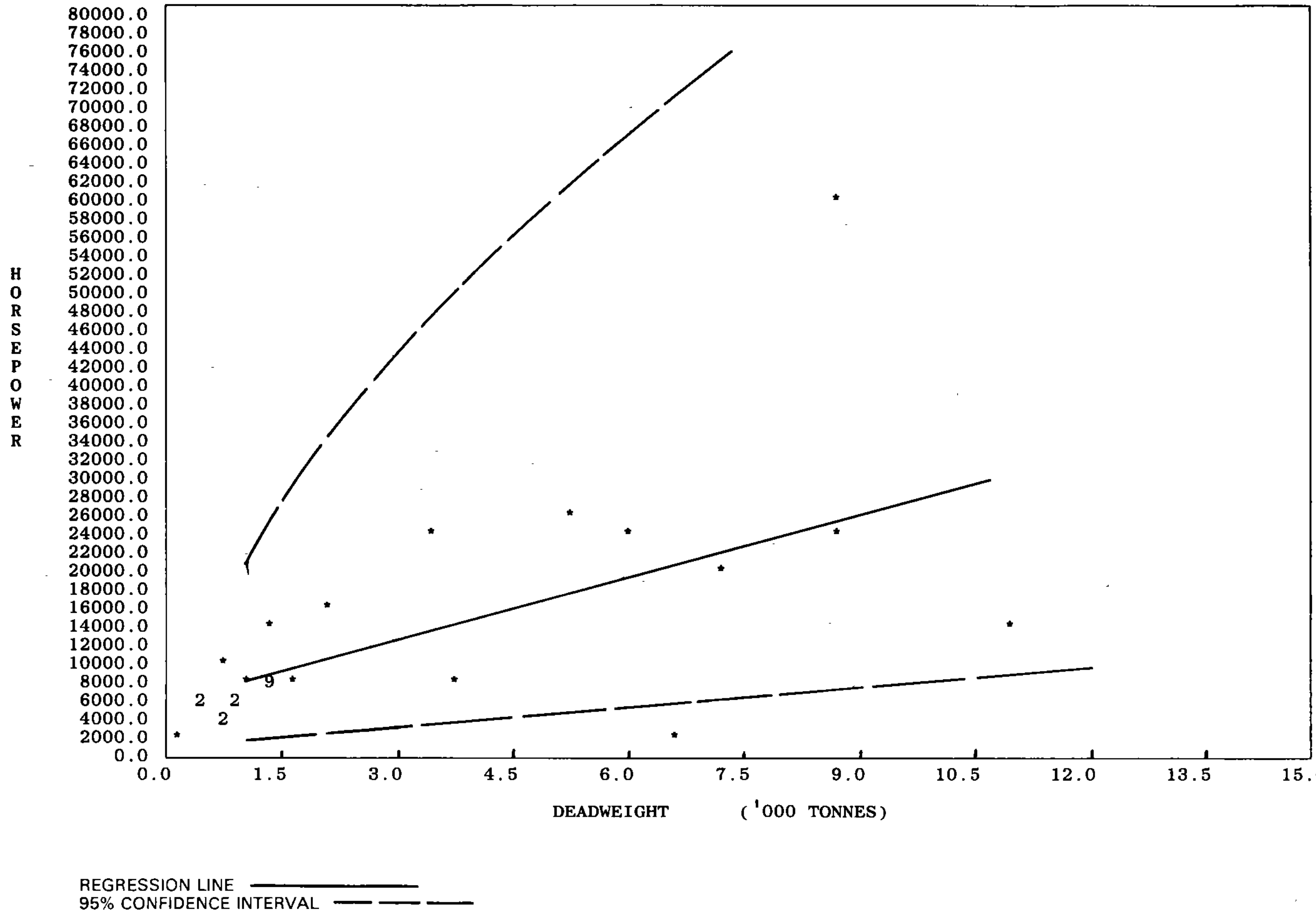


FIGURE 52 HORSEPOWER VS DEADWEIGHT (PASSENGER SHIPS)

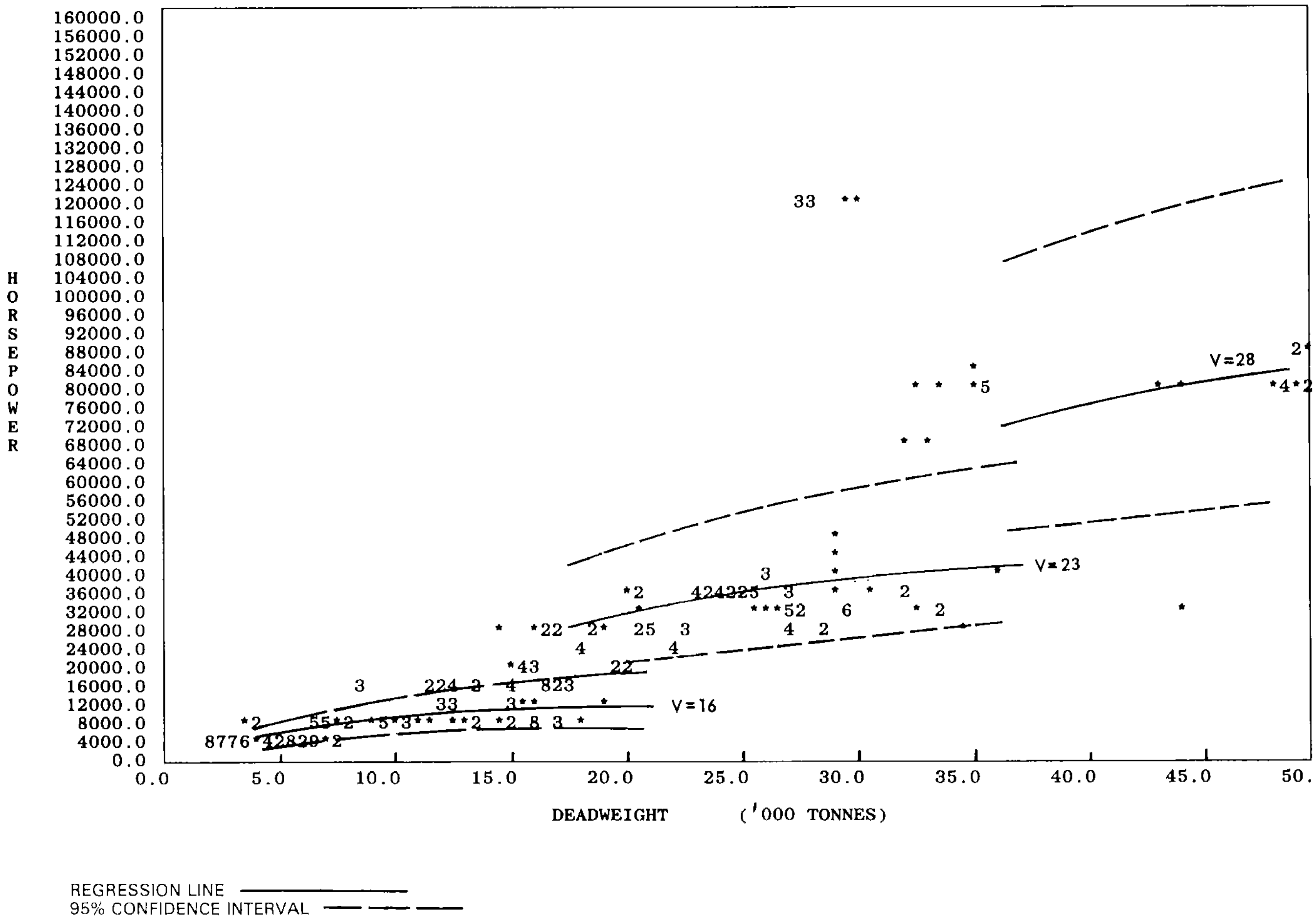


FIGURE 53 HORSEPOWER VS DEADWEIGHT (CONTAINER SHIPS)

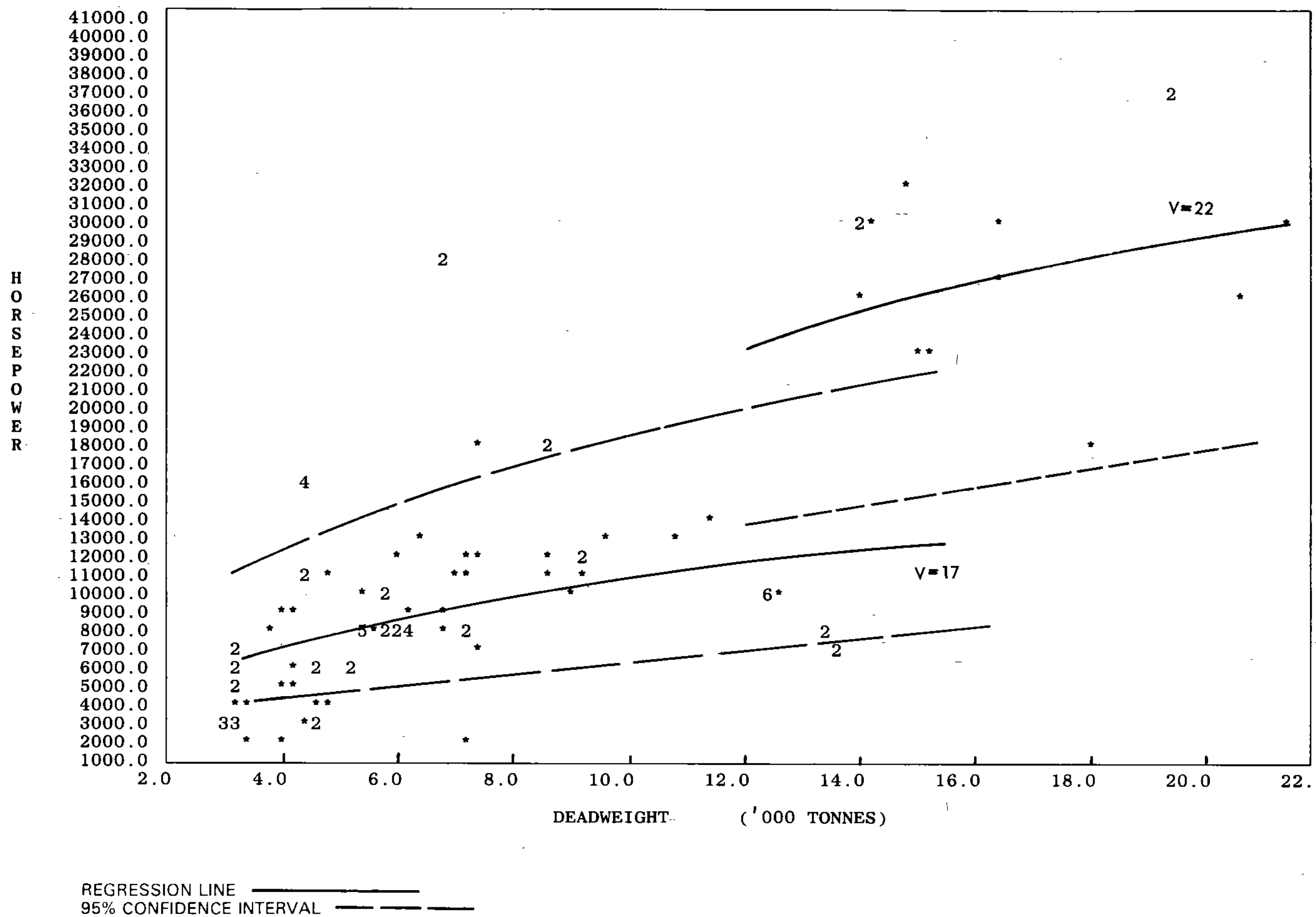


FIGURE 54 HORSEPOWER VS DEADWEIGHT (RO. RO SHIPS)

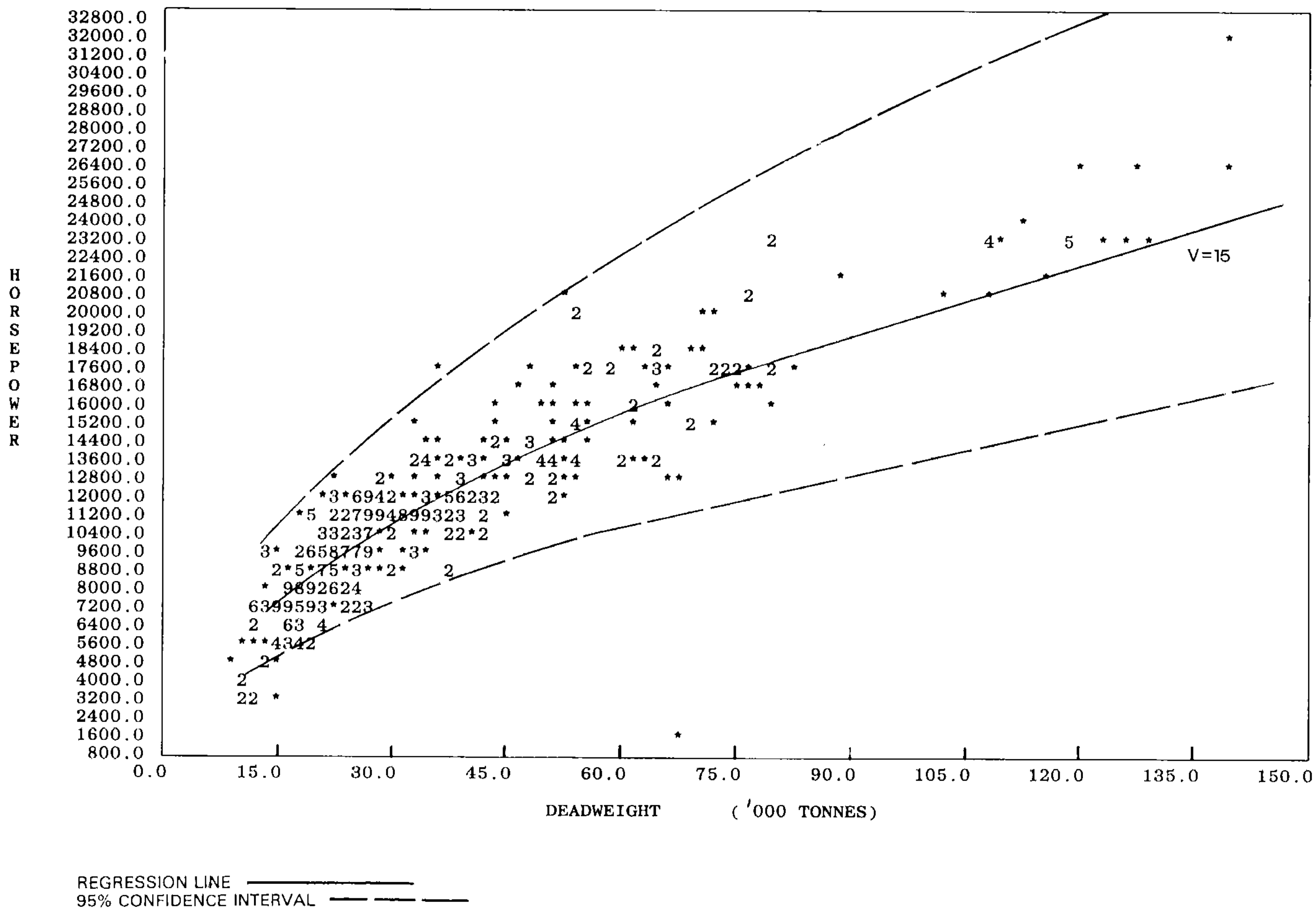


FIGURE 55 HORSEPOWER VS DEADWEIGHT (BULK CARRIERS)

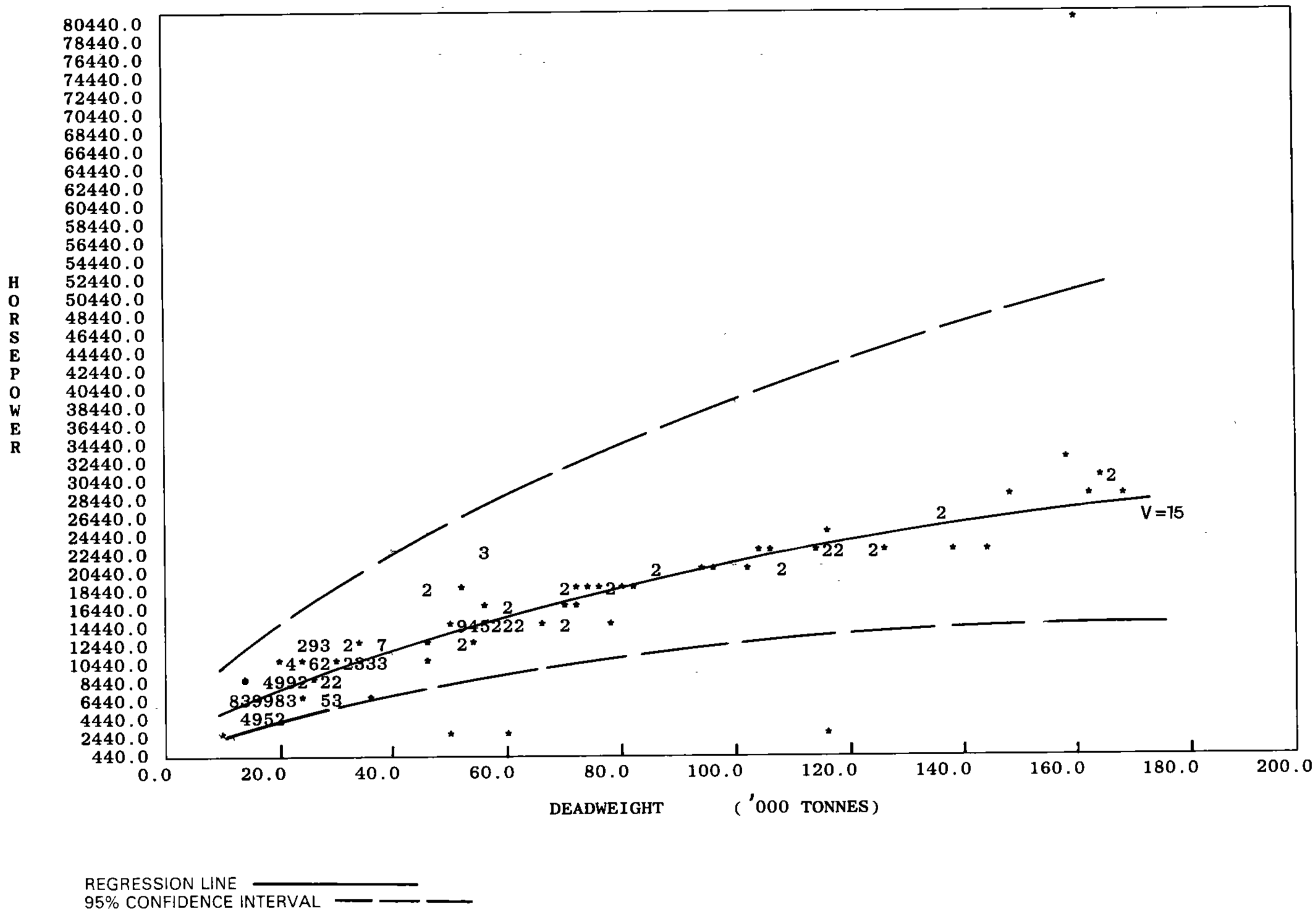


FIGURE 56 HORSEPOWER VS DEADWEIGHT (ORE CARRIERS)

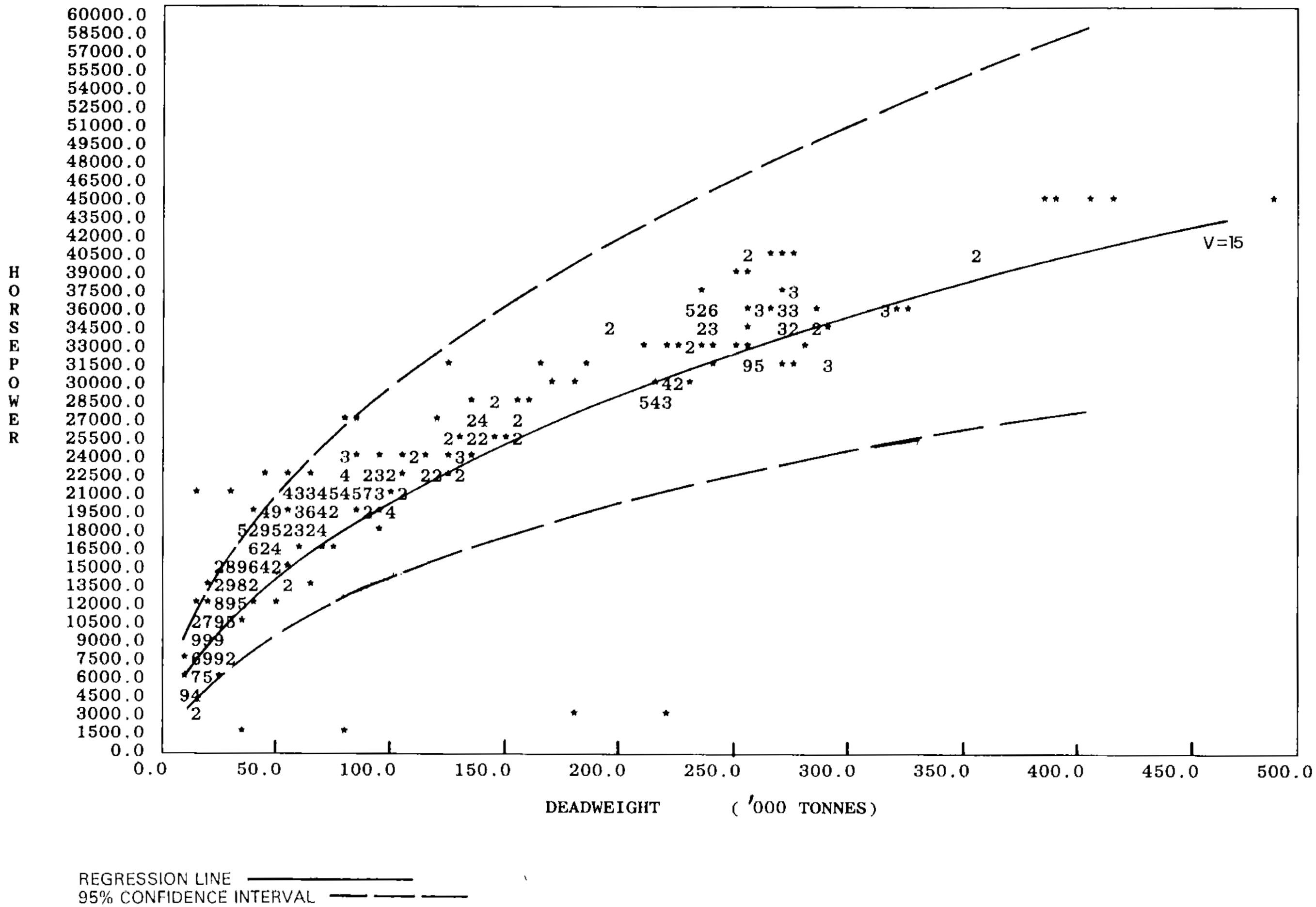


FIGURE 57 HORSEPOWER VS DEADWEIGHT (TANKERS)

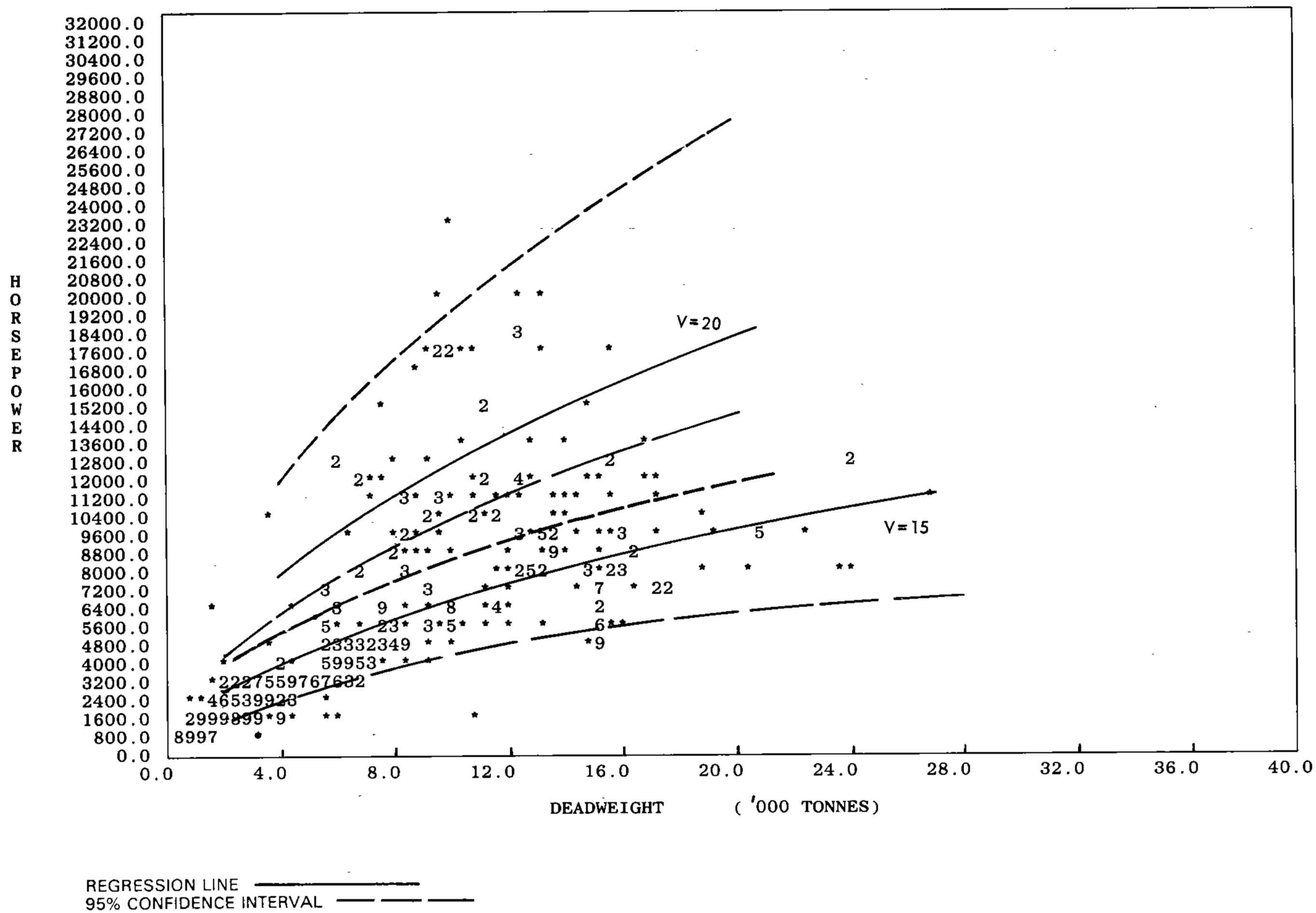


FIGURE 58 HORSEPOWER VS DEADWEIGHT (GENERAL CARGO SHIPS)

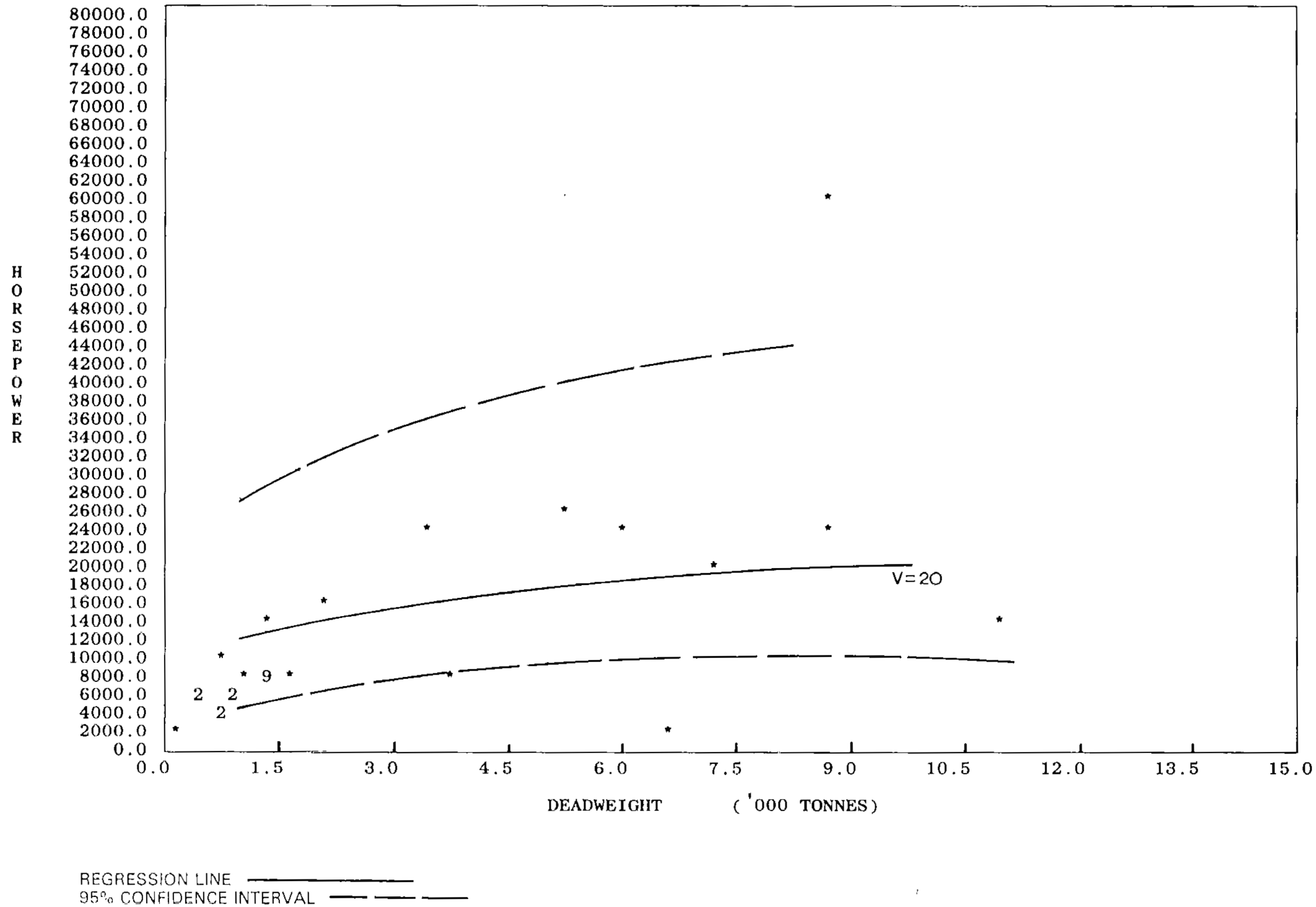


FIGURE 59 HORSEPOWER VS DEADWEIGHT (PASSENGER SHIPS)

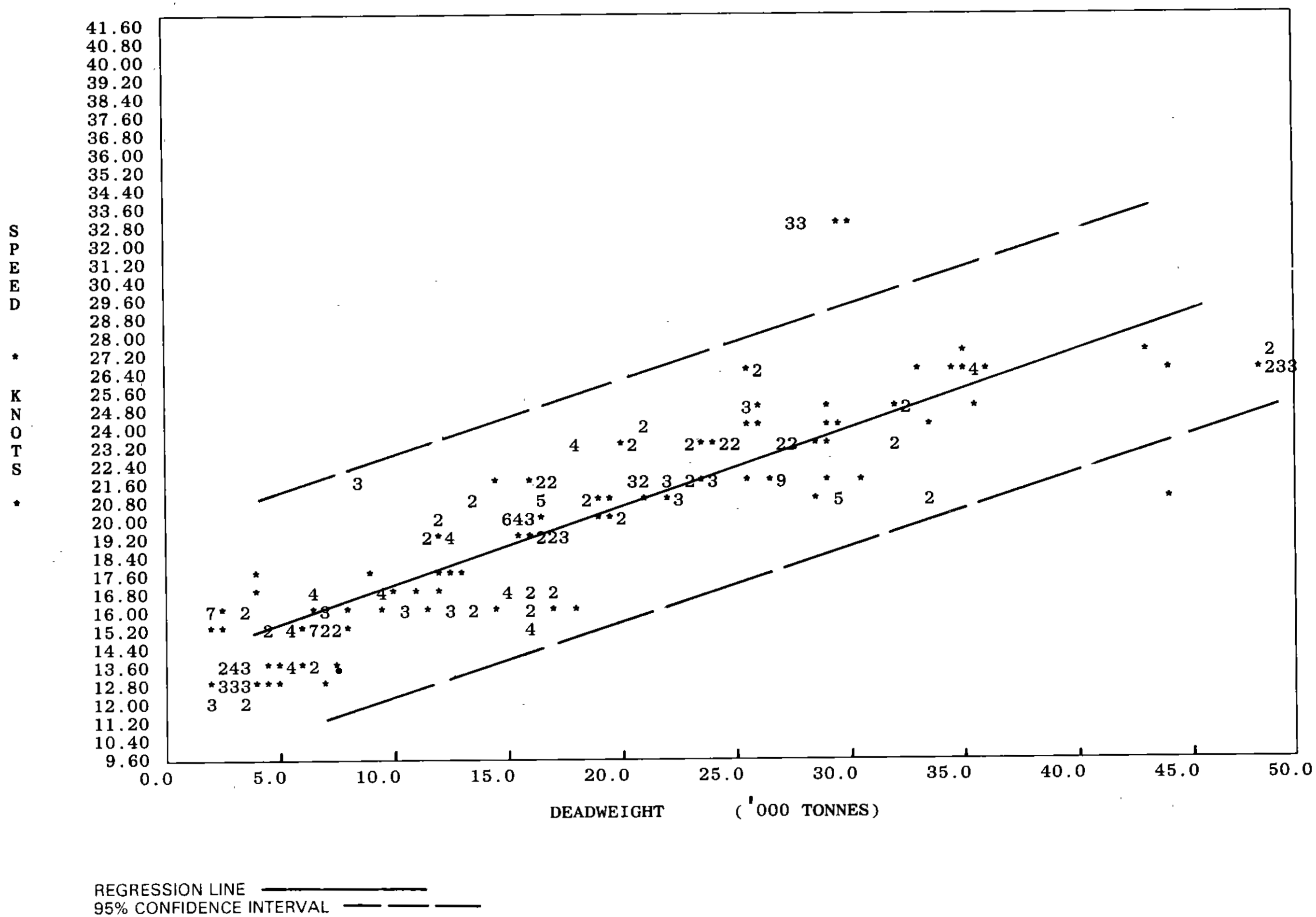


FIGURE 60 SPEED VS DEADWEIGHT (CONTAINER SHIPS)

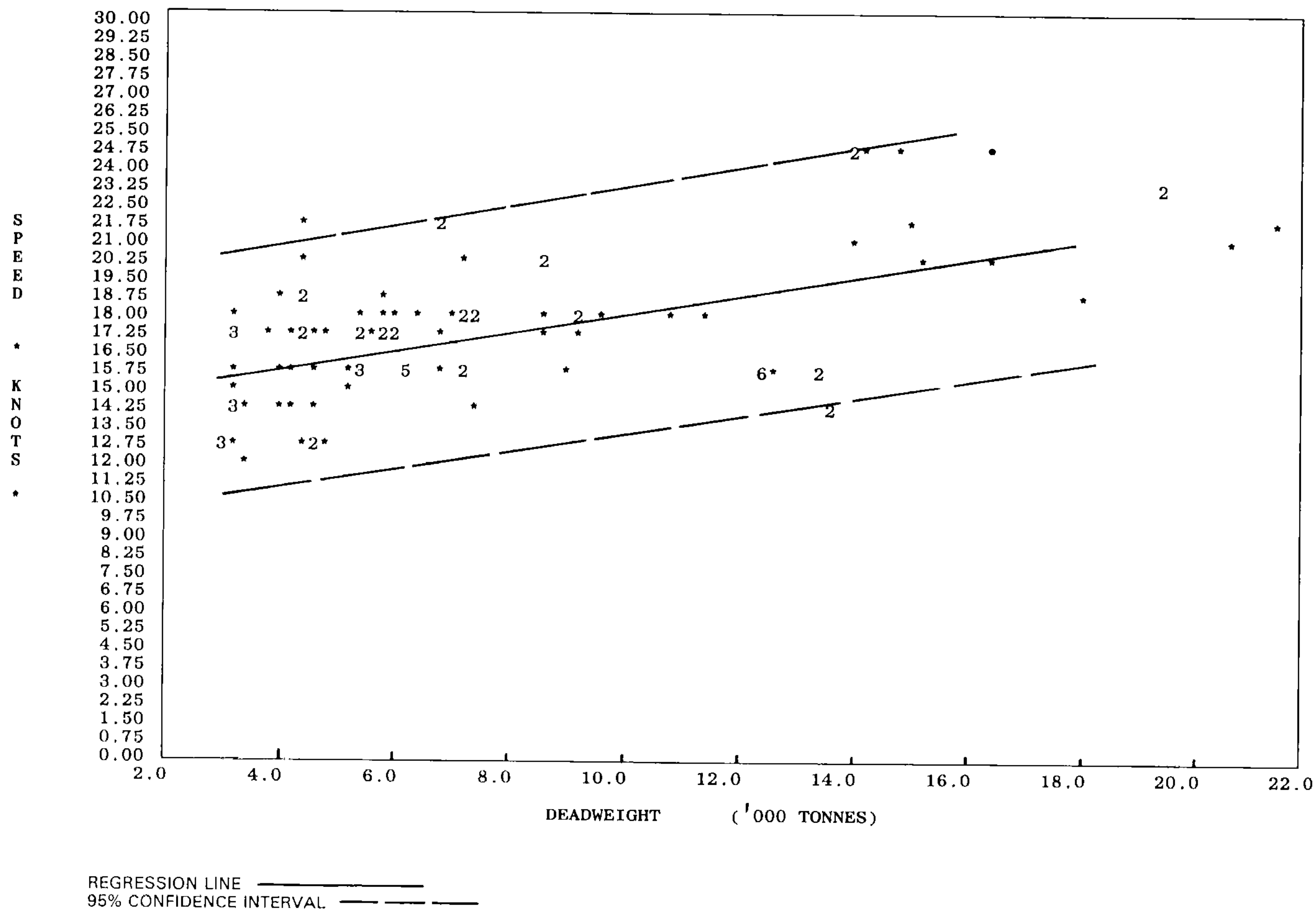
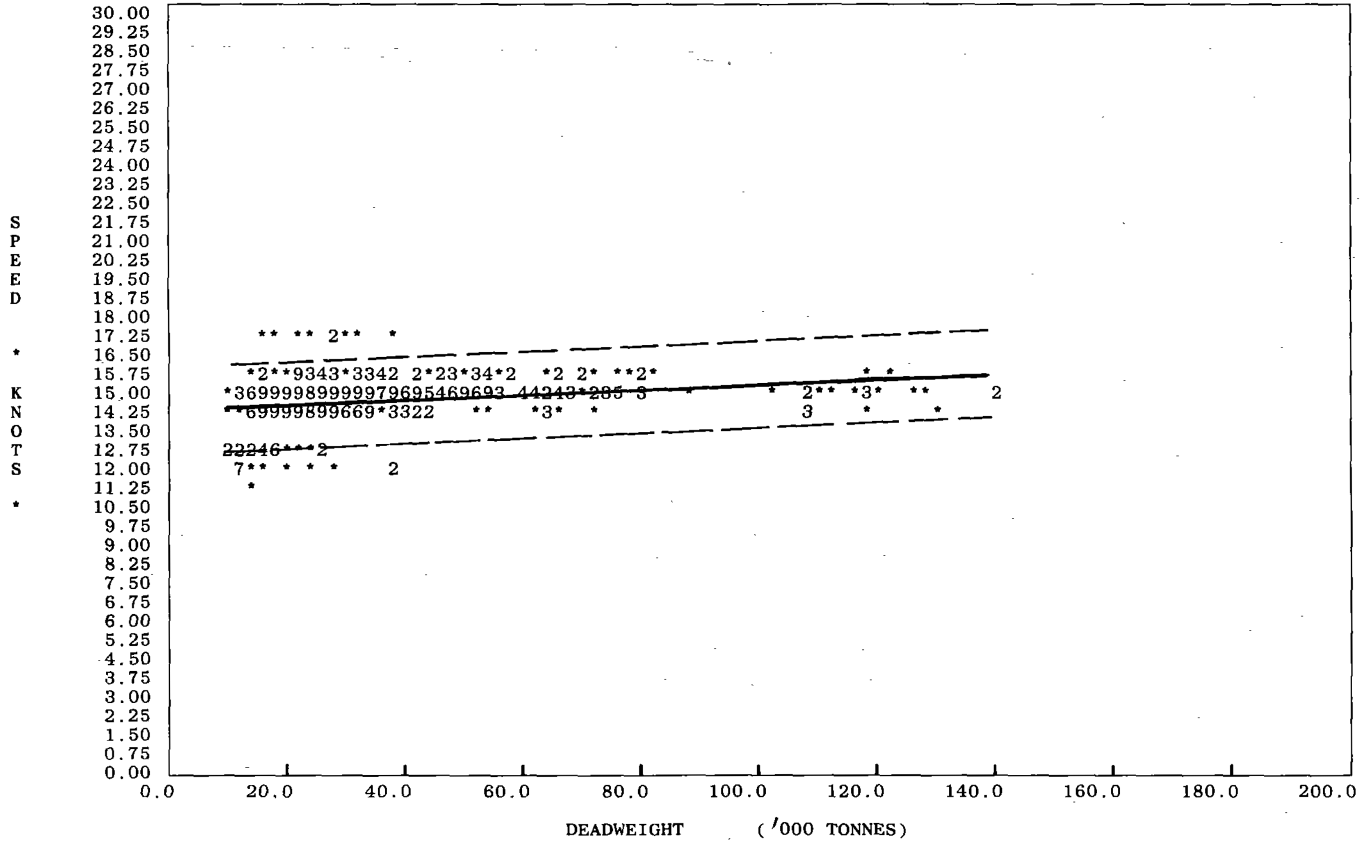


FIGURE 61 SPEED VS DEADWEIGHT (RO. RO SHIPS)



REGRESSION LINE
95% CONFIDENCE INTERVAL

FIGURE 62 SPEED VS DEADWEIGHT (BULK CARRIERS)

25474/80—7

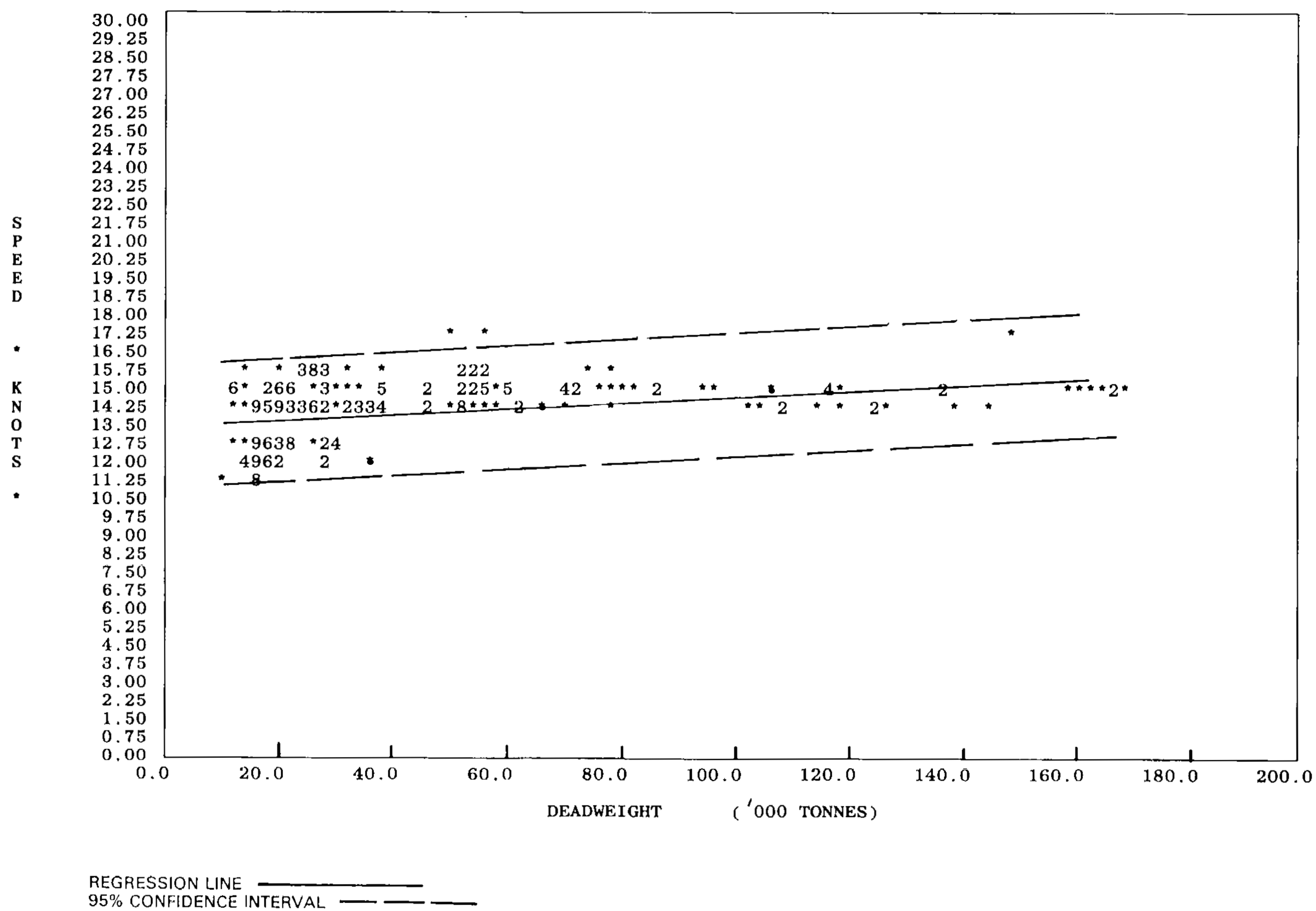
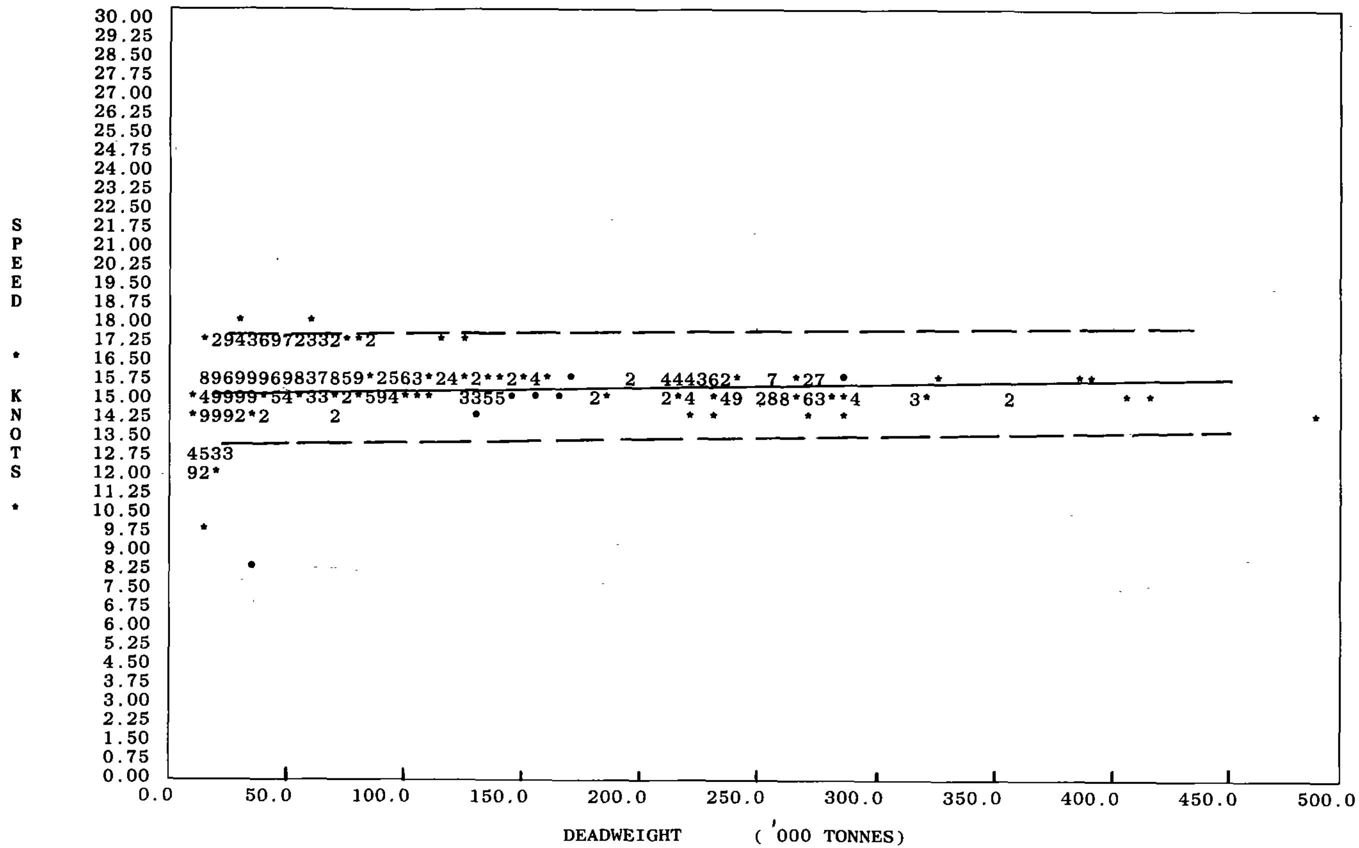


FIGURE 63 SPEED VS DEADWEIGHT (ORE CARRIERS)



REGRESSION LINE
95% CONFIDENCE INTERVAL

FIGURE 64 SPEED VS DEADWEIGHT (TANKERS)

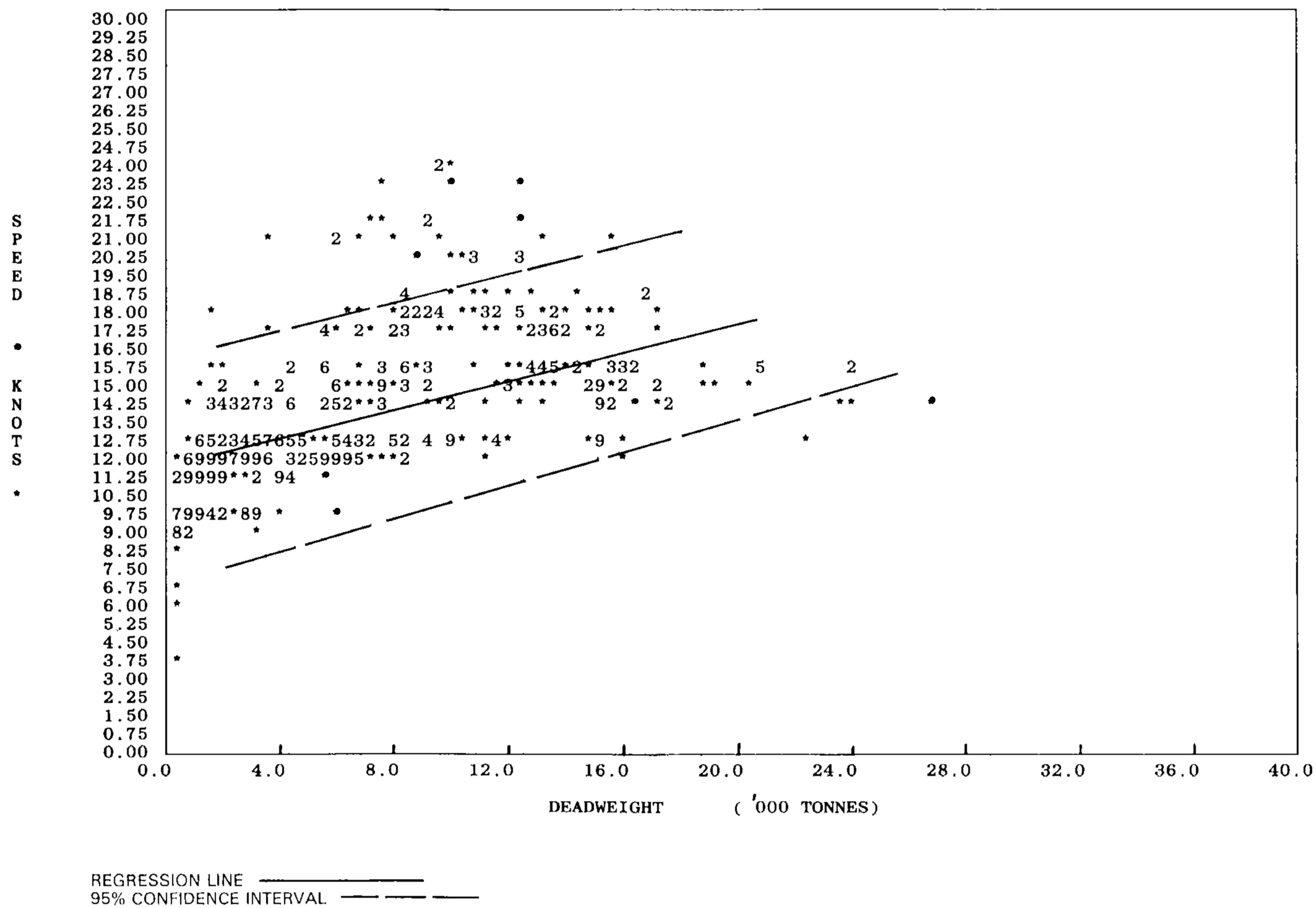
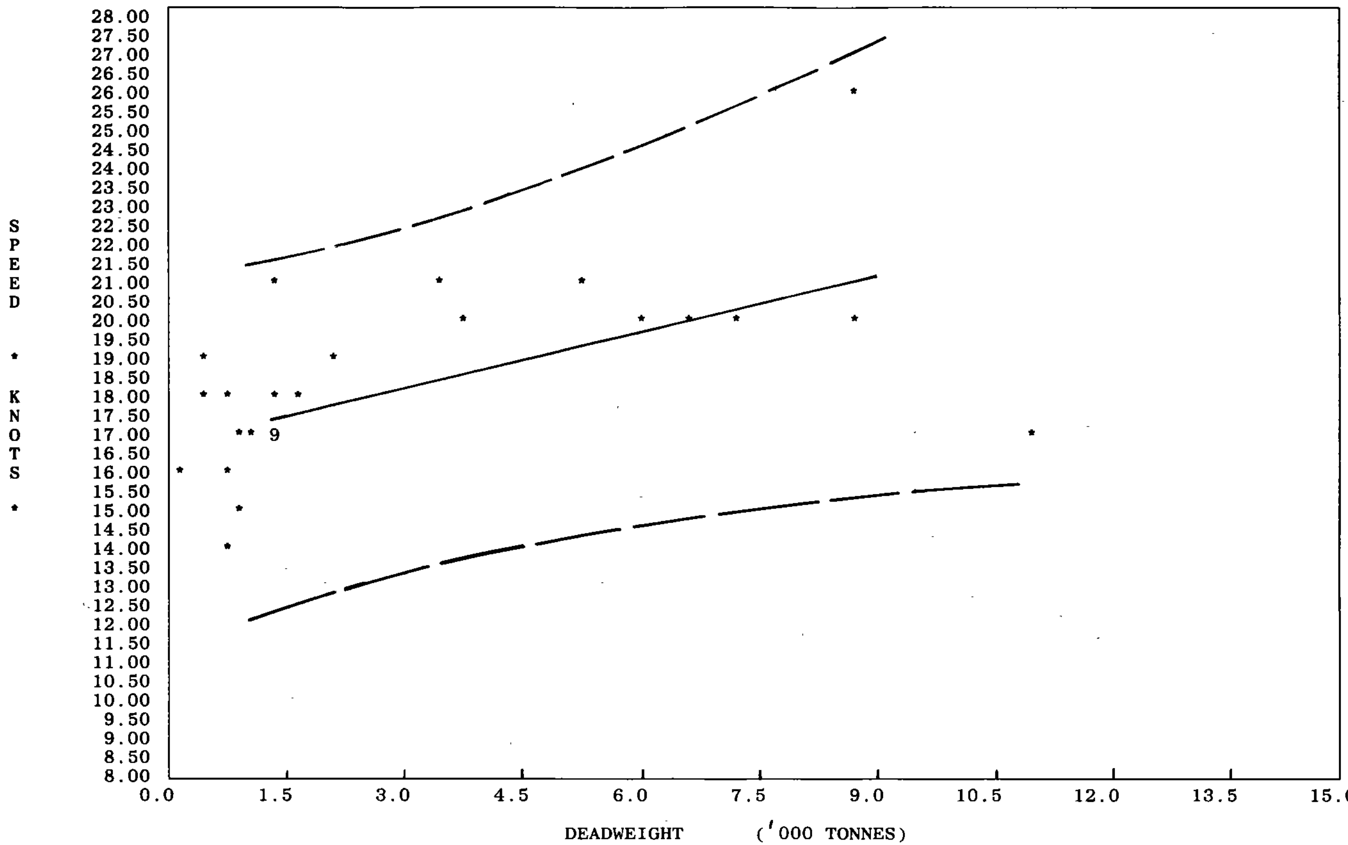
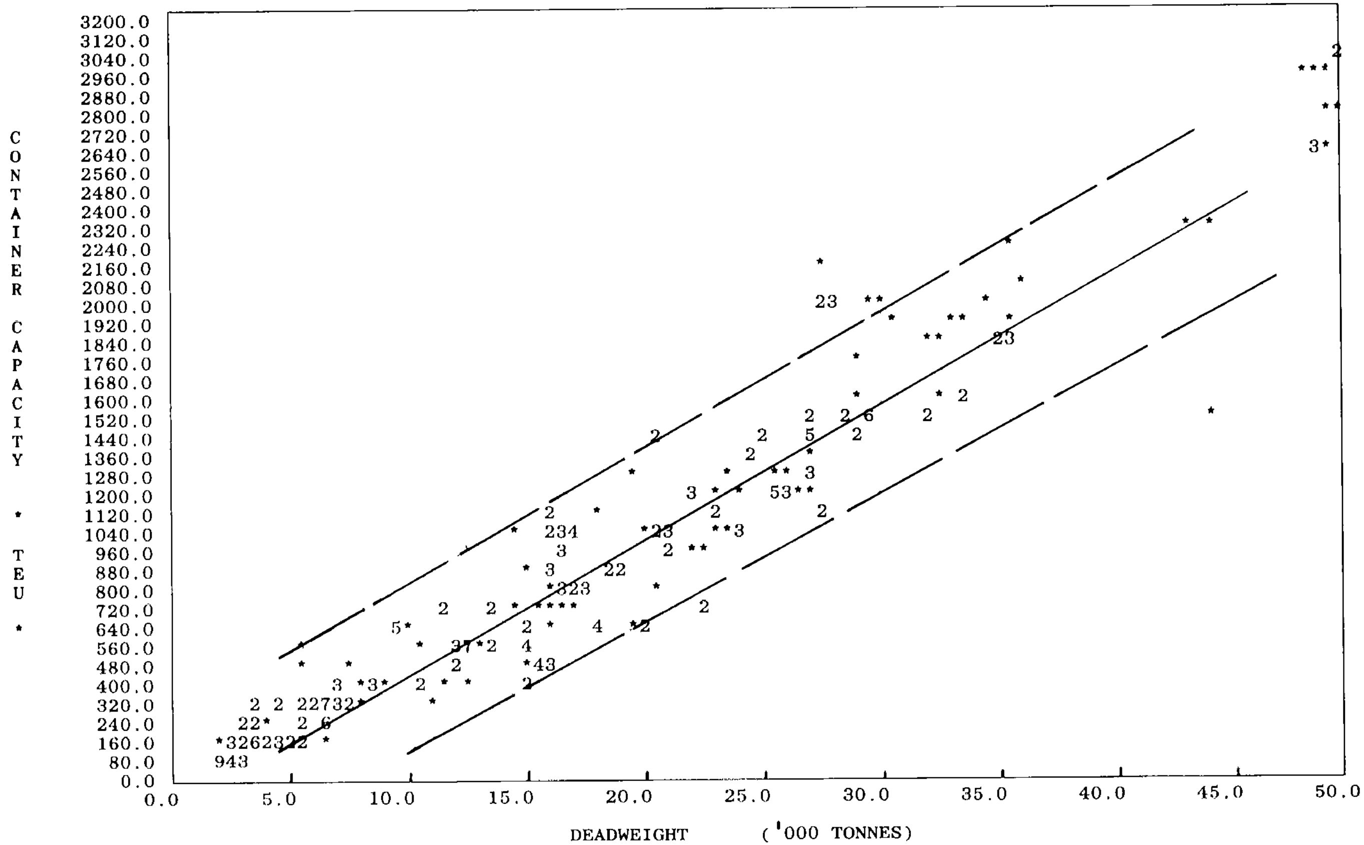


FIGURE 65 SPEED VS DEADWEIGHT (GENERAL CARGO SHIPS)



REGRESSION LINE
95% CONFIDENCE INTERVAL

FIGURE 66 SPEED VS DEADWEIGHT (PASSENGER SHIPS)



REGRESSION LINE
95% CONFIDENCE INTERVAL

FIGURE 67 CONTAINER CAPACITY VS DEADWEIGHT

CHAPTER 4 - CONCLUDING REMARKS

GENERAL

Standard regression techniques have been used to determine a regression model for a number of data sets. The form of these data sets is the result of a variety of interacting factors and, hence, does not illustrate any single, fundamental principle. Therefore, the regression relationships provide a description of the existing world fleet, but they do not illustrate the physical principles applied to the design of ships.

The relationships presented in Tables 1-9 provide an indication of the trends that are evident in the data. However, the additional information available by also examining the plots of the sample data and by considering the prediction confidence intervals will be found worthwhile when applying these relationships. The information provided in this paper as a whole should prove to be a valuable tool to those studying shipping and port infrastructure.

LIMITATIONS OF THE RESULTS

When applying the results presented in this paper consideration should always be given to the characteristics of the data from which the results have been derived. The relevance of these relationships, in any application, is dependent on the relevance of the data from which they have been determined.

As mentioned in Chapter 1, Lloyd's Register is very extensive and is expected to be representative of the world fleet. Hence, the results should draw together any trends evident in individual ship building nations and present an aggregate estimate of the relationships between the ship characteristics investigated.

Any arbitrary selection of data will always affect the regression model derived. Therefore, it should be remembered that ships with incomplete data have been excluded from the analysis, as have those ships that operate in more than one fashion, e.g. general cargo/container ships (see Chapter 2). These facts may affect the applicability of the results depending on the type of trade expected at a given port and the accuracy desired of the estimates of the ship characteristics. However, due to the comprehensive nature of the information in Lloyd's Register it is expected that such effects would only be of minor significance.

As mentioned in Chapter 1, ships are deleted from Lloyd's Register of Shipping as they are withdrawn from service, and the version of the Register used for this analysis was current at May 1977. About 1.5-2.0 per cent of ships on the Register are withdrawn from service each year, and these constitute about 1.0-1.5 per cent of the gross tonnage of the world fleet. Assuming that this rate of withdrawal from service continues, after ten years one would expect the data used in this analysis to represent 80-85 per cent of the ships still in service and these ships would constitute 85-90 per cent of the gross tonnage of the world fleet. The accuracy required for a particular study will determine whether these results are still of use in a given situation, but for many applications the information in this paper will be relevant for ten years or more. However, if in the future it becomes obvious that a large number of new ships, of a particular ship type, have been constructed to a significantly different design then this should be taken into account before applying the relationships presented in this paper.

It should be noted, of course, that the regression models provide no justification for any prediction outside the range of values encountered in the samples.

TABLE 1 - LENGTH[(a)]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$L = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	60.5 (269)	0.399 (70.4)	0.94
Ro-ro	61.7 (99.7)	0.423 (20.3)	0.80
Bulk Carrier	68.7 (695)	0.288 (162)	0.92
Ore Carrier	72.8 (390)	0.276 (87.9)	0.97
Tanker	75.8 (1290)	0.268 (333)	0.97
General Cargo	60.0 (1310)	0.349 (197)	0.90
Passenger	105 (145)	0.317 (9.97)	0.72

(a) Length measured in metres.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 2 - BREADTH[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$B = \alpha.(DWT)^{\beta}.\varepsilon$$

Ship Type	α	β	R^2
Container	10.4 (169)	0.311 (60.4)	0.92
Ro-ro	11.3 (57.3)	0.303 (14.2)	0.65
Bulk Carrier	8.69 (340)	0.313 (169)	0.92
Ore Carrier	8.05 (188)	0.341 (107)	0.98
Tanker	8.10 (671)	0.337 (450)	0.99
General Cargo	10.1 (1080)	0.281 (232)	0.93
Passenger	15.3 (124)	0.234 (10.8)	0.75

(a) Breadth measured in metres.

t - statistics shown in brackets

DWT measured in '000 tonnes.

TABLE 3 - DRAUGHT[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$D = \alpha.(DWT)^{\beta}.\varepsilon$$

Ship Type	α	β	R^2
Container	3.78 (87.6)	0.320 (56.9)	0.91
Ro-ro	3.94 (53.3)	0.297 (22.9)	0.83
Bulk Carrier	4.18 (275)	0.275 (182)	0.93
Ore Carrier	4.07 (101)	0.274 (69.1)	0.95
Tanker	3.81 (386)	0.299 (360)	0.98
General Cargo	3.77 (390)	0.334 (173)	0.88
Passenger	4.89 (40.4)	0.263 (6.74)	0.54

(a) Draught measured in metres.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 4 - GROSS REGISTERED TONS: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$GRT = \alpha + \beta . (DWT) + \varepsilon$$

Ship Type	α	β	R^2
Container	-2510 (-6.33)	1160 (60.3)	0.92
Ro-ro	24.6 (0.07) (a)	780 (19.6)	0.78
Bulk Carrier	2310 (31.7)	514 (290)	0.97
Ore Carrier	1860 (4.06)	499 (59.1)	0.93
Tanker	4180 (39.9)	469 (586)	0.99
General Cargo	-129 (-6.30)	659 (269)	0.95
Passenger	2840 (3.47)	1910 (9.38)	0.65

(a) t - statistic not significantly different from zero at the 0.05 level of significance.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 5 - NET REGISTERED TONS: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$NRT = \alpha + \beta . (DWT) + \varepsilon$$

Ship Type	α	β	R^2
Container	-1440 (-5.07)	699 (50.8)	0.89
Ro-ro	-618 (-2.57)	485 (18.2)	0.76
Bulk Carrier	361 (5.31)	372 (225)	0.95
Ore Carrier	2240 (9.28)	172 (38.6)	0.84
Tanker	-868 (-6.72)	387 (392)	0.98
General Cargo	-226 (-16.3)	415 (252)	0.94
Passenger	980 (2.24)	1040 (8.52)	0.65

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 6 - AGE: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL

$$AGE = \alpha + \beta . (DWT) + \varepsilon$$

Ship Type	α	β	R^2
Container	7.75 (27.6)	-0.0538 (-3.95)	0.04
Ro-ro	no significant regression		
Bulk Carrier	10.6 (57.4)	-0.0558 (-12.4)	0.06
Ore Carrier	18.3 (50.2)	-0.0889 (-13.2)	0.39
Tanker	16.5 (98.5)	-0.0474 (-37.0)	0.31
General Cargo	7.19 (94.8)	-0.106 (-11.7)	0.03
Passenger	no significant regression		

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 7.1 - HORSEPOWER[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$HP = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	759 (95.7)	1.16 (45.0)	0.86
Ro-ro	1640 (43.9)	0.905 (10.6)	0.51
Bulk Carrier	1480 (253)	0.579 (68.9)	0.66
Ore Carrier	830 (63.6)	0.701 (23.3)	0.66
Tanker	1640 (392)	0.561 (124)	0.84
General Cargo	969 (670)	0.836 (144)	0.83
Passenger	6260 (77.3)	0.638 (5.69)	0.45

(a) Power has been regressed in units of horsepower because of the continued usage of this unit by the shipping industry.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

1 hp = 0.746 kW .

TABLE 7.2 - HORSEPOWER[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$HP = \alpha . (DWT)^{\beta} . (v)^{\gamma} . \varepsilon$$

Ship Type	α	β	γ	R^2
Container	1.52 (1.97)	0.526 (20.7)	2.66 (29.4)	0.96
Ro-ro	1.46 (0.73) [b]	0.422 (6.86)	2.80 (13.9)	0.83
Bulk Carrier	38.9 (20.6)	0.521 (63.3)	1.43 (20.7)	0.71
Ore Carrier	12.4 (4.39)	0.586 (18.5)	1.74 (7.43)	0.72
Tanker	6.70 (14.3)	0.504 (131)	2.10 (41.7)	0.90
General Cargo	5.49 (29.7)	0.539 (115)	2.17 (90.7)	0.94
Passenger	0.289 (-0.86) [c]	0.230 (2.43)	3.55 (6.96)	0.76

(a) Power has been regressed in units of horsepower because of the continued usage of this unit by the shipping industry.

(b), (c) t - statistics not significantly different from one at the 0.05 level of significance.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

1 hp = 0.746 kW.

TABLE 8 - SPEED[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$V = \alpha+\beta.(DWT)+\varepsilon$$

Ship Type	α	β	R^2
Container	14.3 (48.4)	0.321 (24.4)	0.70
Ro-ro	14.4 (31.0)	0.373 (7.24)	0.33
Bulk Carrier	14.4 (441)	0.00989 (12.4)	0.06
Ore Carrier	13.5 (123)	0.0140 (6.95)	0.15
Tanker	15.1 (544)	0.00172 (8.07)	0.02
General Cargo	11.5 (211)	0.314 (48.3)	0.36
Passenger	16.4 (34.0)	0.555 (4.13)	0.30

(a) Speed measured in knots.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE 9 - CONTAINER CAPACITY[a]: REGRESSION COEFFICIENTS AND STATISTICS FOR MODEL,

$$TEU = \alpha + \beta . (DWT) + \varepsilon$$

Ship Type	α	β	R^2
Container	-79.1 (-4.00)	55.7 (59.3)	0.92

(a) Container capacity measured in Twenty foot Equivalent Units.

t - statistics shown in brackets.

DWT measured in '000 tonnes.

ANNEX A

MAIN CONTENTS OF LLOYD'S REGISTER OF SHIPPING

CARD TYPE	MAIN CONTENTS
T00	Lloyd's Register number Ship's name
T01	Call sign Official number Navigational aids
T03	Year of change of name Former name
T04	Owner
T05	Manager
T06	Flag Port
T08	Gross tons Net tons Deadweight
T10	Classification society, other than Lloyd's Register (LR)
T11	LR hull classification symbols
T12	Classification notation of ship (LR)
T13	Machinery classification (LR)

T20	Date of build Shipbuilder and place of build Yard number
T21	Dimension of hull Length overall Extreme breadth Draught Registered length Length between perpendiculars Moulded breadth Moulded depth
T22	Superstructures
T23	Number of decks Type of decks
T24	Number of complete decks (including shelter decks) Rise of floor Keel type Keel length
T25	Information on keel
T26	Cargo battens Bulkheads Water ballast
T27	Type of alterations Date of alterations
T29	Conversions

T30	Ship type and sub-types Propulstion type Number of screws Number of passengers Material of ship
T31	Special features of ship
T33	Number, length and type of: holds tanks combined holds/tanks between deck space wing holds wing tanks
T36	Grain capacity Bale capacity Insulated capacity Liquid capacity Heating coils
T37	Number and size of containers carried Number of lighters carried
T38	Number, material, length and breadth of centreline hatchways
T39	Number, material, length and breadth of wing side hatchways
T40	Number of winches Number of cranes Safe working load
T41	Number of derricks Safe working load

T42	Number, material, type, shape and position of special tanks

NB: Cards T50 through to T59 relate to oil prime movers

T50	Number of engines Position Cylinder layout Number of cylinders with bore and stroke dimensions
T51	Information on gearing and coupling
T52	Total horsepower (bhp) Engine design code
T53	Generators driven by oil engines Number of generators Kilowatt Volts Alternating/direct current Indication of secondary propulsion
T54	Electric motors driven by generators Number of engines Position Shaft horsepower Type of system
T55	Diesel electric motors
T56	Emergency or secondary propulsion
T58	Second or third oil engine group Number of engines Position

Cylinder layout
Number of cylinders with bore and stroke dimensions

T59	Total horsepower (bhp), second or third engine group Engine design code
T60	Steam reciprocating engine (dimensions in imperial units) Type of reciprocating engine Number of engines Bore and stroke Position
T61	Steam reciprocating engine (dimensions in metric units) Type of reciprocating engine Number of engines Bore and stroke Position
T62	Information on gearing and coupling of low pressure turbines when combined with reciprocating engines
T63	Horsepower, reciprocating engines (ihp) Engine design code
T70	Steam turbines Number of steam turbines Information on gearing and coupling
T71	Total shaft horsepower Turbine design code

T75	Turbo-electric engines Number of steam turbines Total shaft horsepower Kilowatt, volts of generators Shaft horsepower of motors Type of steam
T76	Information on gearing
T77	Turbine design code
T80	Gas turbine Total shaft horsepower
T85	Engine dates Year when engine was made, fitted, refitted or added
T86	Fuel bunkers Capacity and type of fuel bunkers
T87	Engine builder(s) and where made
T89	Boilers (LR classed ships only) Number, type and position of boilers
T90	Primary and secondary pressure and working temperature and pressure
T95	Auxiliary generators Number of generators Kilowatt, volts of generators Alternating/direct current Frequency in Hz

T96 Special propellors and speed
Number of special propellors
Type of special propellors
Position of special propellors
Speed

T99 Cross reference ship's name.

ANNEX B

DATA PREPARATION AND ANALYSIS

Data from Lloyd's Register of Shipping is input to DATAPREP1 which reads the information for the first ship, and determines whether the ship type is one to be analysed. It also determines whether the ship operates in purely one fashion, i.e. is it a 'pure' ship type. If the ship is eligible for inclusion in the sample, the program uses the card type code to locate all cards with the particular items of data to be examined. These cards are loaded into the output array and output to the file for that ship type. Then the information for the next ship is input, and the procedure repeated. Seven files are thereby created; one for each ship type. In summary, program DATAPREP1 reduces the information being processed to the relevant cards for the relevant ships and stores this data in a separate file for each ship type.

DATAPREP2 reads the data for the first ship from one of the seven files created by the previous program. It then locates the 'words' (strings of ten characters) that contain the information on deadweight, gross registered tons, net registered tons, length, breadth, draught and age. These 'words' are loaded into the output array, and the ship propulsion type is located and decoded. It must be decoded because the form of propulsion determines on which cards the data for power is located. The 'words' containing informtion on container capacity are then located and loaded into the output array. Next the data on power are located and decoded. If they are blank the program drops that ship and reads the information for the next ship. If the data on power is present it is loaded into the output array. Finally, the 'word' containing the information on speed is located and loaded into the output array. The information for that ship is then output the information for the next ship read in and the process repeated. Therefore, DATAPREP1 produces files which contain the relevant 'cards' of data for ships of the relevant types, while DATAPREP2 produces files which contain the relevant 'words' of data.

DATAPREP3 reads the information on tonnage and dimensions for the first ship and converts it to integer format. The ship dimensions and deadweight are then converted to metric units. The information for power, speed, age and container capacity (if relevant) is read and all data are checked for missing items. If any data are missing the ship is dropped from the sample. Otherwise, the container capacity (if relevant) is converted to TEUs, after which the information for that ship is output in a fixed format record, and the procedure repeated for the next ship. This fixed format record consists of fifty-seven characters for each ship type, except container ships, for which it consists of sixty-two characters. The information on each record contains all nine characteristics listed in Chapter 2 (ten characteristics for container ships).

The statistical analysis of the data output by DATAPREP3 was performed using the GENSTAT[1] statistical package. Firstly, GENSTAT was used to plot the sample data for each set of variables that was to be regressed. The exact samples on which the regressions were to be performed (described in Chapter 2) were plotted for all ship types except general cargo. The sample for this ship type was so large (4146 ships) that GENSTAT was unable to cope with all the data. A random selection of the sample was therefore plotted for this ship type.

Standard linear regression techniques were then used to determine the regression model for each set of data. All regressions, including those for general cargo ships, were performed on the

---

(1) GENSTAT: A General Statistical Program. The Statistics Department. Rothamstead Experiment Station. 1977.

sample of ships described in Chapter 2. The regression models described in this paper generally have one of the following forms:

$$Y = \alpha + \beta . X + \varepsilon \quad (1)$$

$$Y = \alpha . X^{\beta} . e^{\varepsilon}. \quad (2)$$

The first model simply determines the straight line of best fit for the data, where α is the intercept on the dependent variable axis and β is the slope of the line. There is an unpredictable randomness in all data which is described by the stochastic error term, ε. This term accounts for error from two sources. The first is the fact that when framing a regression model one does not claim to have included all the variables which influence the relationship and so there will be specification error in the equation. The second source of error is in the measurement or recording of the data.

Before regression, the second model is linearised by log transformation, to the form of the first model, i.e.

$$\log Y = A + \beta . \log X + \varepsilon, \quad (3)$$

where $\log \alpha = A$, and
$\log e^{\varepsilon} = \varepsilon$

A standard linear regression is then performed on this data. The α term in regression model 2 is determined by taking the antilogarithm of the estimate of the regression coefficient, A, while the β term in regression model 2 is the same as the coefficient β in regression model 3.

An additional computer program was written which was used to calculate the 95 per cent prediction confidence interval for each regression model. The confidence interval indicates that, in the long run, one would expect ninety-five out of a hundred new observations to fall between the confidence limits. The limits were calculated using the formula

$$\hat{Y} - t_{(1-\frac{\alpha}{2})} \cdot s\,(Y_{new}) \leq Y_{new} \leq \hat{Y} + t_{(1-\frac{\alpha}{2})} \cdot S\,(Y_{new}),$$

where

$S\,(Y_{new}) = \sqrt{MSE\ (1\ X'\,(X'X)^{-1}X)}$,

$\hat{Y}$ is the regression estimate of Y_{new}

$t_{(1-\frac{\alpha}{2})}$ is the t-statistic, and

MSE is the error mean square or residual mean square.

ANNEX C

ALTERNATIVE MODELS: REGRESSION COEFFICIENTS AND STATISTICS

TABLE C.1

$$LENGTH = \alpha + \beta.(DWT) + \gamma.(DWT)^2 + \varepsilon$$

Ship Type	α	β	γ	R^2
Container	77.2 (31.7)	7.65 (30.7)	-0.0709 (-13.1)	0.92
Ro-ro	62.0 (9.44)	13.6 (8.95)	-0.314 (-4.40)	0.81
Bulk Carrier	119 (203)	23.6 (93.3)	-0.00941 (-46.5)	0.92
Ore Carrier	127 (102)	2.07 (42.0)	-0.00688 (-21.5)	0.96
Tanker	153 (337)	1.20 (135)	-0.00179 (-64.6)	0.96
General Cargo	59.5 (145)	8.88 (105)	-0.153 (-42.0)	0.84
Passenger	75.7 (12.0)	32.7 (7.17)	-2.19 (-4.75)	0.77

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.2

$$BREADTH = \alpha + \beta.(DWT) + \gamma.(DWT)^2 + \varepsilon$$

Ship Type	α	β	γ	R^2
Container	12.2 (53.3)	0.952 (40.6)	-0.0110 (-21.5)	0.94
Ro-ro	13.5 (13.0)	1.07 (4.41)	-0.0151 (-1.33)	0.64
Bulk Carrier	16.3 (196)	0.313 (86.9)	-0.000965 (-33.5)	0.93
Ore Carrier	15.7 (92.7)	0.349 (51.9)	-0.00109 (-24.9)	0.98
Tanker	19.4 (249)	0.202 (132)	-0.000270 (-56.5)	0.96
General Cargo	10.0 (217)	1.09 (115)	-0.0178 (-43.5)	0.87
Passenger	11.7 (16.3)	3.66 (7.05)	-0.243 (-4.63)	0.77

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.3

$$DRAUGHT = \alpha + \beta . (DWT) + \gamma . (DWT)^2 + \varepsilon$$

Ship Type	α	β	γ	R^2
Container	4.73 (42.8)	0.334 (29.5)	-0.00372 (-15.1)	0.89
Ro-ro	4.42 (22.5)	0.423 (9.29)	-0.00844 (-3.95)	0.85
Bulk Carrier	7.50 (250)	0.109 (84.2)	-0.000303 (-29.3)	0.93
Ore Carrier	7.56 (117)	0.0849 (33.1)	-0.000141 (-8.49)	0.97
Tanker	8.35 (382)	0.0703 (164)	-0.0000842 (-63.1)	0.98
General Cargo	3.62 (161)	0.552 (119)	-0.00974 (-48.7)	0.87
Passenger	4.15 (6.87)	0.953 (2.19)	-0.0512 (-1.16)	0.32

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.4

$$GRT = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	588 (159)	1.17 (78.5)	0.95
Ro-ro	668 (59.4)	1.06 (19.1)	0.77
Bulk Carrier	844 (597)	0.898 (273)	0.97
Ore Carrier	833 (110)	0.886 (50.8)	0.90
Tanker	855 (1320)	0.896 (727)	0.99
General Cargo	504 (793)	1.10 (247)	0.94
Passenger	3900 (88.7)	0.766 (8.28)	0.64

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.5

$$NRT = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	332 (70.5)	1.18 (38.7)	0.82
Ro-ro	235 (45.0)	1.22 (19.9)	0.79
Bulk Carrier	359 (83.2)	1.01 (48.9)	0.49
Ore Carrier	486 (46.5)	0.796 (21.0)	0.61
Tanker	361 (149)	0.999 (105)	0.79
General Cargo	283 (450)	1.11 (156)	0.86
Passenger	1490 (45.7)	0.914 (5.76)	0.46

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.6

$$AGE = \alpha + \beta . (DWT)^{-1} + \varepsilon$$

Ship Type	α	β	R^2
Container	no significant regression		
Ro-ro	no significant regression		
Bulk Carrier	5.74 (24.8)	76.3 (13.8)	0.07
Ore Carrier	9.69 (17.7)	126 (10.3)	0.28
Tanker	6.80 (35.5)	205 (35.7)	0.30
General Cargo	6.01 (105)	1.19 (15.9)	0.06
Passenger	no significant regression		

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.7

$$HORSEPOWER = \alpha + \beta.(DWT) + \varepsilon$$

Ship Type	α	β	R^2
Container	-6840 (-4.77)	1800 (25.9)	0.66
Ro-ro	1010 (0.95)	1340 (11.4)	0.55
Bulk Carrier	5740 (61.3)	161 (70.6)	0.67
Ore Carrier	3650 (10.1)	180 (27.1)	0.73
Tanker	9610 (69.2)	98.5 (92.8)	0.74
General Cargo	1090 (15.1)	571 (66.6)	0.52
Passenger	5140 (3.02)	2480 (5.22)	0.41

t - statistics shown in brackets.

DWT measured in '000 tonnes.

1 hp = 0.746 kW.

TABLE C.8

$$SPEED = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	10.3 (105)	0.237 (28.9)	0.71
Ro-ro	12.3 (51.8)	0.173 (7.06)	0.32
Bulk Carrier	12.9 (328)	0.0404 (17.8)	0.11
Ore Carrier	11.1 (97.2)	0.0662 (9.35)	0.24
Tanker	13.7 (482)	0.0269 (20.6)	0.12
General Cargo	10.9 (618)	0.137 (62.6)	0.49
Passenger	16.6 (117)	0.115 (4.80)	0.37

t - statistics shown in brackets.

DWT measured in '000 tonnes.

TABLE C.9

$$CONTAINER\ CAPACITY = \alpha . (DWT)^{\beta} . \varepsilon$$

Ship Type	α	β	R^2
Container	43.8 (80.7)	1.04 (60.3)	0.93

t - statistics shown in brackets.

DWT measured in '000 tonnes.

FIGURES 4-67

'*' denotes one data point

'2' denotes two coincident data points

'9' denotes nine or more coincident data points.