Statistics Research Project

Layne Proctor

April 27, 2017

**The Correlation Between Number of Pages and Width of a Book**

Statement of the Problem

On its face, this seems like a question with an obvious answer: do books get thicker as the number of pages increases? Clearly, there is more matter between the covers of a 100-page book than there is between those of a 50-page book, so it would follow that the 100-page book is thicker. What remains interesting, or at the least informative and potentially useful (more on usefulness below), is how strong a relationship exists between number of pages and width. While there is likely to be at minimum a moderate positive correlation, can number of pages alone be a practically accurate indicator of width given that other considerations such as paper stock, cover material, binding, wear, and water or sun damage might also affect how wide a book is (and therefore how much space a given book will take up on a book shelf)? This test will attempt to quantify this relationship.

Methods

The population for this research was all the books currently shelved in the Capital Hill apartment of the author of this study. Since by current count this population comprises 2,007 books, and since the deadline for this project is in less than a week, a random sample was selected using a systematic sampling procedure. The sample should be less than .05 of the total population, so every fiftieth book would yield a sample size of 40 books, which equals .02 of the total (less than .05) but still satisfying the Law of Large Numbers (at least 20 or 30 observations). This would provide a sample that included representatives from various publishers, and thus observations (books) featuring a representative variety of characteristics (paper stock, cover thickness, wear and weathering). Once the sample was obtained, each member of the sample was flipped through to discover the number of pages (including the unnumbered pages that sometimes appear in the back of books—indices, advertisements, order forms, and blank-page filler). These numbers were recorded in an Excel spreadsheet. After determining the number of pages in each book, the width was measured to the millimeter by means of a metric tape measure. These measurements were also recorded in the spreadsheet.

Given that the study was attempting to quantify the correlation between number of pages and book width, it was not necessary to set an alpha-level, nor was it necessary to articulate either an alternate (Ha) or null (H0) hypothesis. That said, it was assumed prior to the study that the relationship would be a positive one, likely r ≥ .5 (the benchmark for a moderate positive linear correlation).

Results

Once the data was collected, descriptive statistics were computed using Excel’s data analysis package. Following this, a regression analysis was performed to obtain the regression statistics. The descriptive statistics are summarized below in Table 1; the regression statistics are provided in Table 2.

Table 1

|  |  |  |
| --- | --- | --- |
|  | *Number of Pages* | *Width in mm* |
|  |  |  |
| Mean | 405.6 | 24.275 |
| Standard Error | 41.22722837 | 1.958867903 |
| Median | 318 | 21.5 |
| Mode | 132 | 21 |
| Standard Deviation | 260.7438865 | 12.38896842 |
| Sample Variance | 67987.37436 | 153.4865385 |
| Kurtosis | 0.339000225 | -0.730829026 |
| Skewness | 1.057613251 | 0.375031498 |
| Range | 986 | 46 |
| Minimum | 101 | 5 |
| Maximum | 1087 | 51 |
| Sum | 16224 | 971 |
| Count | 40 | 40 |

Table 2

|  |  |
| --- | --- |
| *Regression Statistics* | |
| Multiple R | 0.935355351 |
| R Square | 0.874889632 |
| Adjusted R Square | 0.871597254 |
| Standard Error | 4.439379794 |

A scatter plot with trend line illustrating the relationship between number of pages and book width is shown here in Graph 1.

Graph 1

Included in the plot are three “forecasted” widths of hypothetical books of 1100, 1200, and 1300 pages, respectively. These plot points were derived using the regression equation of y = m+bx, where b = intercept and m = slope. The inputs for this formula were derived from the regression coefficients provided by Excel based on the book-measurements data. These are listed in Table 3 below.

Table 3

|  |  |
| --- | --- |
|  | *Coefficients* |
| Intercept | 6.24915665 |
| Number of Pages | 0.044442415 |

As expected, there was at least a moderate positive linear relationship; in fact, the relationship between the two data sets is strong. The “goodness of fit” between the trend line and the data (RSQ or r2) was .875; the correlation coefficient, which describes the strength of the linear relationship between the two paired data sets, was .935. The number of pages in a book has a very strong relationship to book width.

Conclusions

Though the results from this study might be unsurprising, it does appear to establish as minimal other factors that it was assumed contribute to a given book’s width.

As for the usefulness of this study, the data set above and the results derived from the statistical analyses of said data provide a relatively quick and easy method for determining how much shelf space would be needed in order to accommodate a book or set of books totaling x pages. As illustrated by the forecasted book widths in the scatter plot above in Graph 1, the coefficients obtained in this study could have practical use if, say, the author of this study wished to see the impact of introducing a single-volume edition of *À la recherché du temps perdu* or an unabridged set of *The Golden Bough* into his personal library. Listed in Table 4 below are the potential widths of books of various pagination.

Table 4

|  |
| --- |
| Width of a book of 500 pages |
| 28.47 mm |
| Width of a book of 750 pages |
| 39.58 mm |
| Width of a book of 1000 pages |
| 50.69 mm |
| Width of a book of 5000 pages |
| 228.46 mm |

Knowing the number of pages of any book (which can be learned simply by flipping to the back of a book or is information provided by online booksellers such as Amazon) would give a good indication as to how much space that book would take up on a shelf. If the author was so inclined, he could increase the sample size to get a more precise depiction of the correlation between pages and width. He could also perform a separate study involving hardcover books to have a formula for non-paperback books.

References

French, V. (2017). Welcome to week 12 Tues [Powerpoint slides]. Retreived from http://vf-tropi.com/stat1.html.

French, V. (2017). Welcome to week 12 Thurs [Powerpoint slides]. Retreived from http://vf-tropi.com/stat1.html.

Sullivan, M. (2014). Fundamentals of Statistics Fourth Edition [Looseleaf]. Pearson Education,

Inc.

Appendix

Data used in this study

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of Pages | Width in mm | Trend Line | Forecast | Book Title |
| 373 | 21 | 22.82617728 | n/a | *The Angela Y. Davis Reader* |
| 823 | 37 | 42.82526384 | n/a | *Roll, Jordan, Roll: The World The Slaves Made* |
| 538 | 35 | 30.15917569 | n/a | *After Babel: Aspects of Language and Translation* |
| 760 | 40 | 40.02539172 | n/a | *Gravity's Rainbow* |
| 827 | 43 | 43.0030335 | n/a | *Underworld* |
| 883 | 51 | 45.49180872 | n/a | *Dune* |
| 132 | 11 | 12.11555537 | n/a | *How to Read Lacan* |
| 443 | 23 | 25.9371463 | n/a | *The Thing on the Doorstep and Other Weird Stories* |
| 255 | 17 | 17.58197237 | n/a | *Democratic Theory: Essays in Retrieval* |
| 236 | 18 | 16.73756649 | n/a | *The Debt Resisters' Operations Manual* |
| 403 | 29 | 24.15944972 | n/a | *Anti-Oedipus: Capitalism and Schizophrenia* |
| 396 | 29 | 23.84835282 | n/a | *A Confederacy of Dunces* |
| 298 | 20 | 19.49299619 | n/a | *Breakfast of Champions* |
| 294 | 21 | 19.31522653 | n/a | *Something Wicked This Way Comes* |
| 272 | 22 | 18.33749341 | n/a | *Neuromancer* |
| 251 | 22 | 17.40420271 | n/a | *The Wretched of the Earth* |
| 463 | 32 | 26.8259946 | n/a | *Memoirs of My Nervous Illness* |
| 170 | 9 | 13.80436713 | n/a | *The History of Sexuality: Volume 1: An Introduction* |
| 239 | 11 | 16.87089373 | n/a | *Jenseits von Gut und Böse* |
| 594 | 33 | 32.64795091 | n/a | *A Treatise of Human Nature* |
| 1087 | 49 | 54.55806129 | n/a | *Infinite Jest* |
| 384 | 26 | 23.31504385 | n/a | *Reading Lolita in Tehran: A Memoir in Books* |
| 304 | 19 | 19.75965068 | n/a | *Democracy Against Capitalism: Renewing Historical Materialism* |
| 114 | 7 | 11.31559191 | n/a | *The Fire Next Time* |
| 332 | 18 | 21.00403829 | n/a | *Rewiring the Real* |
| 603 | 34 | 33.04793264 | n/a | *Truth and Method* |
| 1019 | 39 | 51.5359771 | n/a | *Kritik der reinen Vernunft* |
| 131 | 5 | 12.07111296 | n/a | *The Diving Bell and the Butterfly* |
| 210 | 16 | 15.58206371 | n/a | *Adult Children of Alcoholics* |
| 216 | 18 | 15.8487182 | n/a | *Copywriting: Successful Writing for Design, Advertising, and Marketing* |
| 419 | 28 | 24.87052836 | n/a | *A Documentary History of the United States* |
| 727 | 44 | 38.55879204 | n/a | *History of Madness* |
| 101 | 7 | 10.73784052 | n/a | *On Existentialism* |
| 416 | 32 | 24.73720111 | n/a | *Let Us Now Praise Famous Men* |
| 574 | 39 | 31.75910261 | n/a | *A Writer's Diary* |
| 132 | 9 | 12.11555537 | n/a | *The Beginning of Philosophy* |
| 136 | 8 | 12.29332503 | n/a | *A Mind That Found Itself* |
| 234 | 17 | 16.64868166 | n/a | *Black Power: The Politics of Liberation* |
| 150 | 11 | 12.91551884 | n/a | *Giving an Account of Oneself* |
| 285 | 21 | 18.9152448 | n/a | *Change the World Without Taking Power: The Meaning of Revolution Today* |
| 1100 |  |  | 55.13581268 | *Imaginary book 1* |
| 1200 |  |  | 59.58005414 | *Imaginary book 2* |
| 1300 |  |  | 64.02429559 | *Imaginary book 3* |