

: THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

The Data Consists of 100 Cases of Hypothetical Data to
Demonstrate Approval of Loans by a Bank

Name

Course

Subject

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO DEMONSTRATE APPROVAL OF LOANS BY A BANK

Introduction

There has been witnessed an alarming trend in the number of people denied loans. Some have been blacklisted by the credit rating bureaus across the country and declared un-creditworthy. Therefore implying that approximately 25% of the population has no access to credit, hence stalling investments and in extension it has severely slowed the growth of the economy over the last decade. It is in this view that I wished to undertake an in-depth research to establish the main factors which are; age, level of education, household income, and years that employees stick to their current employer.

Method

The study required an in-depth investigation and hence was only appropriate to use a case study. This is because a case study method involved an in detail, examination of a single illustration or result and thereby providing a systematic way of collecting, analyzing, and finally reporting the results. It will enable the bank to make a general inference to the whole population based on the sample. The data comprised of 100 cases of hypothetical data to demonstrate approval of loans by the bank.

Sample size:

Data was got from 100 random consumers by use of a probabilistic method of simple random sampling. Therefore, the data be free from bias..

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

Data Analysis

Data analysis is basically, a process of obtaining meaningful information from raw and unmeaningful data. The data collected was analyzed using SPSS to determine the descriptive statistics such as measures of central dispersion, frequencies and measures of central tendencies. It was also used to determine inferential statistics of the study.

Results

Data Results of the study were depicted in the form of tables and charts, where necessary. And a detailed explanation of the values was given.

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

DESCRIPTIVE Statistics

	Age of respondents	Literacy levels	Years at current job	Income	Proportion of debt to income	Plastic money debt	Additional debt
Mean	34.38	1.74	8.18	45.0400	10.7660	1.9212	2.9174
Median	34.00	1.00	6.00	35.5000	9.7500	1.0175	2.0751
Mode	41	1	0	21.00 ^a	9.30	.03 ^a	.10 ^a
Std. Deviation	7.969	.917	6.654	29.24640	7.27707	2.59574	2.66407

a. Multiple modes exist. The smallest value is shown

From the study of 100 respondents, the maximum age was 53 years with the minimum age of 20 years. The average, age of the respondents was 34years with a standard deviation of 7.969. Majority of the respondents were aged 41 years. With regards to the Level of education, on average, the respondents were High school degree holders with a mean score of 1.74 which is approximately 2) with a standard deviation of 0.917 whereas many of the respondents did not complete High school. On the debt to income ratio, it was found that the average debt to income ratio stood at 10.766% which is approximately 11% with a standard deviation of 7.27707. With regards to Household income in thousands, the minimum income per household was 14 thousand whereas the highest income per household was at 176, with an average of 45.0400 thousand with

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

a standard deviation of 29.24640. Finally on average it was found that, the maximum years stayed with their current employers was at 26 years with many of the respondents only sticking to their current employers at eight years with a standard deviation of 6.654 which is approximately seven years.

Test for normality

Here we test whether the data we collected follows a normal distribution or not. I evaluated both the numeric analysis using the Shapiro-Wilk test because the data set was less than 2000 units. A graphical representation of normal Q-Q plots was also generated to depict a pictorial representation so to give a good judgment in order to assess normality. Besides, this was essential especially where the numerical test may be over or under sensitive. From the data, the ages of the respondents were normally distributed, because the p value of Shapiro – Wilk Test is greater than 0.05. If it was less than 0.05, then the data we conclude that it significantly deviates from a normal distribution. The table below represents this.

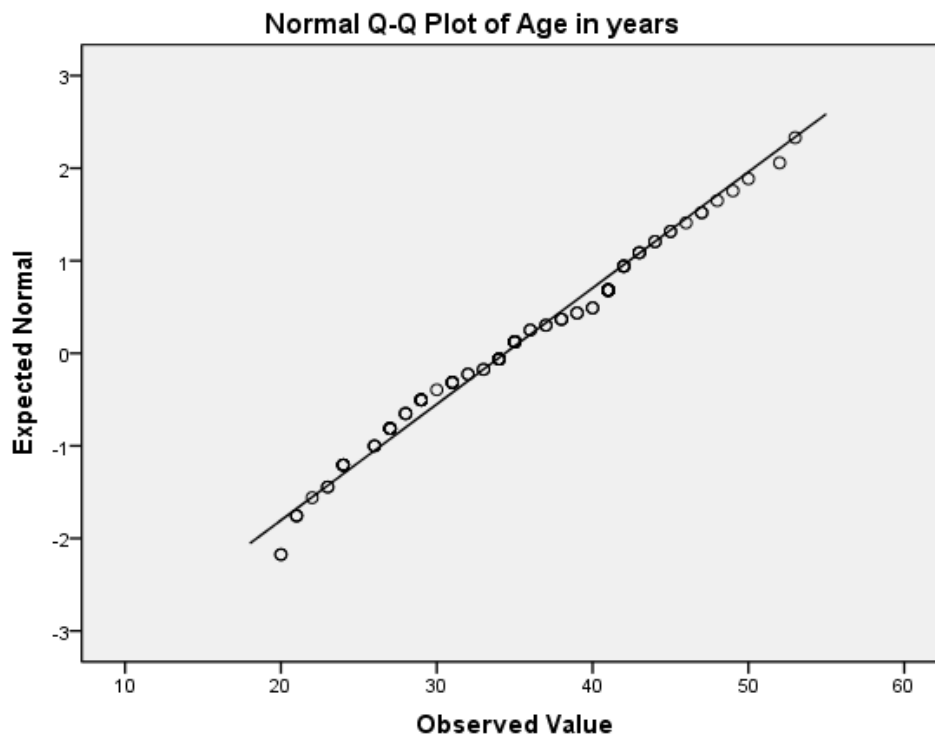
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Age in years	.097	100	.022	.976	100	.069

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO DEMONSTRATE APPROVAL OF LOANS BY A BANK

Normal Q-Q plots

To graphically determine normality, we can use the output of a normal Q-Q plot. For data to be rated as normally distributed, the data points would be aligned to the line of fit. And if they lie far away from the line of best fit then the data does not follow some normal distribution. As shown below, many of the points were along the diagonal and therefore we conclude that the data follows a normal distribution.



Inferential statistics

Regression is mainly used to measure the degree of relationship between two or more variables. After which we are able to fit a linear model of the relationship between the independent variables and the dependent variables. In inferential statistics, data must satisfy normality as we have already illustrated above in order to draw Valid and credible conclusions.

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

The independent variable was if the respondents have ever defaulted on the loan and the independent being the years with current employer, Level of Education, Age, and household income. The R-square was 0.135 with an adjusted R-square at 0.092 while R value was 0.368.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.368 ^a	.135	.092	.408

a. Predictors: (Constant), Years with current employer, Level of education, Age in years, Household income in thousands

When the Nova model of the regression model was fitted, the significance level was 0.020, with 4 degrees of freedom with a mean square of 0.516. This is as shown in the table below.

ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	2.063	4	.516	3.093	.020 ^b
	Residual	13.175	79	.167		
	Total	15.238	83			

a. Dependent Variable: Previously defaulted

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

b. Predictors: (Constant), Years with current employer, Level of education, Age in years,
Household income in thousands

Coefficients

Model	Unstandardized		Standardized	t	Sig.
	Coefficients		Coefficients		
	B	Std. Error	Beta		
(Constant)	.618	.213		2.897	.005
Age in years	-.016	.007	-.307	-2.384	.020
Level of education	.041	.053	.090	.775	.441
Household income in thousands	.006	.002	.390	2.359	.021
Years with current employer	-.017	.010	-.277	-1.664	.100

a. Dependent Variable: Previously defaulted

The coefficients of the model are as shown in the table above. The constant β is 0.618. And this represents the amount of loans defaulted regardless of the influence of the independent variables. For variable age, β_1 is -0.307. This means that as the age of the respondents change by day, the level of loan default decreases. However this is not quite significant at 0.05. For as to whether the

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO DEMONSTRATE APPROVAL OF LOANS BY A BANK

level of education had influence on loan default, the coefficient was 0.90, with a level of significance of 0.441 which was significant at 0.05. As to whether the household income influenced loan default, β_2 was found to be 0.390, implying that as income increased with one thousand, and then the respondent was likely to default on the loan up to 39%. For as to if years the respondents spent with current employer led to loan default, the coefficient was -0.277 implying that, as the number of years the customer stuck to their current employer increased by one year, the risk of loan default also decreased by 0.227. Hence a negative relationship.

DISCUSSIONS

From the descriptive statistics, it was found that a majority were at 40 years while the average age was 34 years. Majorly this was because greater percentages were reaching a climax of their careers with a view of building up long-term investments. This could take them through their future days. In essence this was basically, a speculative motive drive which can be supported by the Ando- Modigliani model of consumption in the famous Life cycle hypothesis (Modigliani, 1986).

On a closer look of the education standard, on average, the respondents were High school graduates with many of them barely able to finish their school. Mainly this arose due to a limited number of higher learning institutions in the country, with many opting out after their high school degree to search for a job.

Upon a look at the number of years that employees spent with their current employers; it was found that on average, employees spent up to 8 years before opting out. This level of turnover is

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO DEMONSTRATE APPROVAL OF LOANS BY A BANK

quite moderate since most of them did not want to lose their current jobs, hence taking a more precautionary motive of job uptake.

On average, it was established that the income per household in thousand was at 45.0400 with a majority lying below average at 21 thousand. Mainly, this was brought about by the low-productivity levels given the level of education and expertise, and the aging level of the respondents hence becoming more inefficient.

The debt to income ratio was determined to be 10.76 thousand on average. This is almost 50% for the majority of who fell below the average of 45.0400. This brought about an impression that a significant number of employees serviced their loans using the income they earn up to about 50% tied to it.

Normality Test

A test on Normality of data was performed with age being a factor to be considered. This is because age as a factor played a major role in establishing the ability to obtain credit. Shapiro-Wilk test was most preferred since the data set was less than 2000 units. The p value was found to be 0.069 which was higher than the hypothesized value of 0.05, hence implying that our data was normal, without any bias. It therefore meant that the results found in the sample of 100 were sufficient to be inferred to the larger population on the factors that affect loan repayment in banking sector.

Discussion on Inferential Analysis

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

A linear regression fit model was run to determine the overall relationship of loan defaulted among employees in the manufacturing sector. Age, level of education, household income, and the number of years in current employment constituted of the independent variables with the dependent variable being if they have ever defaulted on loan repayment. The researcher centered on this four which were seen to be the main factors leading high rate of loan default.

The model is being represented in the form of equation $y = f(x)$. Where y is the dependent variable, which is a function of the independent variables. The independent variables shall take the form of β .

Hence the overall equation takes the form of, $y = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \dots + \varepsilon$, where β_0 represents amount of loan defaulted regardless of the independent variables in the model, β_1 represents the age being the level of education β_3 is the household income β_4 being the number of years the employee stays with the current employer and finally ε is the error term.

Therefore, the regression equation is given as follows;

$y = 0.618 - 0.307\beta_1 + 0.90\beta_2 + 0.390\beta_3 - 0.277\beta_4$. The equation can be explained as

follows; Autonomous loan default is at 61.8%, which means that it is not necessarily that it take customers to default on repayment alone, meaning that other than customers, there are other factors that do not repay back the loans. This can be long term projects which have seized to progress. Equally, as the respondents grew older day by day, it then led to a decrease in risk of default by 0.307. Whereas an increase in the level of education by one level, led to an increase in loan default of 90%.this is because major funds used to educate oneself were mainly from loans without a guarantee of income increment. Also, as income per household rose by 1000, loan

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

risk of loan default also rose by 39%.this is due to money illusion theory. As with the number of years spent with the current employer rose by 1 year, risk of loan default decreased by 27.7%.this is majorly a precautionary motive owing to age and employability of the respondent.

Overall the regression model was not the best, because the r square was 0.135. implying that only 13.5% of the loan default was explained by the age, income, level of education and years with the current employer, with about 86.5% being explained by other factors like inflation, productivity, Financial deepening and international market forces.

THE DATA CONSISTS OF 100 CASES OF HYPOTHETICAL DATA TO
DEMONSTRATE APPROVAL OF LOANS BY A BANK

REFERENCES

Modiglian, F. (1986, june). Life cycle, Individual Thrift, and the wealth of Nations. *American Economic review* , pp. 297-313.

Property of:

<https://mypaperhub.com/>

801 North 34th Street, 3rd floor

Seattle, WA 98103

DISCLAIMER

This Sample paper should be used with proper reference