INTERPRETING FASHION WITH BENFORD LAW: a probabilistic approach

Author

Deep Sagar Verma,
Assistant Professor, Department of Fashion Management Studies National Institute of Fashion Technology, Panchkula (India)

Corresponding Author: deepsagar.verma@nift.ac.in

Key Words

Benford's law, Probability theory, Mathematical Phenomenon, Fashion data, Blockchain Management and Metaverse

Abstract

Introduction-

Fashion is generally defined as one of the purest expressions of art, which is a manifestation of human endeavor, insight, perception, communication and self-realization. However, the underlying fact is that fashion is also a unique blend of science , mathematics and technology. Probability, a branch of mathematics concerning numerical descriptions of how likely an event is to occur, is highly helpful in understanding and predicting the uncertainties of the real world. Benford's law is an observation of the frequency distribution of leading digits in real life sets of numerical data. The law has been widely used in the field of accounting, geology, law, economics, psephology, science but probably not in the generous world of fashion.

Literature review -

There are plenty of literatures defining and explaining Benford's law which are somehow more usable and readable by staunch mathematicians and statisticians. Originally discovered by an American astronomer Newcomb in 1881 on the basis of repetitive distribution law in logarithmic tables, the law became famous after physicist Frank Benford demonstrated it in his research paper "The law of Anomalous Numbers" in 1938. Kruger and Yadavalli described Benford's Law as "counter-intuitive, difficult to explain in simple terms, and has suffered from being described variously as 'a numerical aberration', 'an oddity', 'a mystery' - but also as 'a mathematical gem''' (2017). Mark Nigrini, a professor and author of Forensics Analytics works majorly on Benford's law application.

Methodology-

An exploratory research design is expected to take place on the basis of secondary and syndicated data collection. Internal and external data would be collected from existing customer databases, social media, business/non-government, syndicated services and data warehousing and mining. Furthermore, collection of big data can be analyzed procured from retail/export firms. Designer's data may be also used to analyze the pattern of purchase and consumption.

Findings and Analysis-

The paper would establish a relationship between Benford's law and data generated from various Fashion houses including retailers and exporters. The analysis of trends and patterns would express the mathematical relationship of a pure art like fashion, which may be used for predictive calculations. It will open a new area of research for the fashion industry in sync with probability distribution. Upon successful completion of the establishment of Benford's law, trends studies would become substantially easier to interpret.

Discussion –

New approach of fashion studies could be ingrained and application of numerous untold concepts of diverse disciplines may be initiated.

Conclusion -

Interpretation of Fashion with Benford's law would be a paradigm shift in studying trends, patterns, business practices and social responsibilities.

Introduction

Fashion is generally defined as one of the purest expressions of art, which is a manifestation of human endeavour, insight, perception, communication and self-realization. However, the underlying fact is that fashion is also a unique blend of art, science, mathematics and technology. It also touches on the ways things are conceptualised and executed. As Shakeshpere wrote, "The fashion wears out more apparel than the man". Presently fashion is omnipresent in various forms. And so is the volume of business associated with it. With countless fashion houses, export and retail units and a huge unorganized sector, it attracts billion dollars businesses. The cut throat competition and changing technology make it more complex and challenging to comprehend. Big data analytics and implementation of statistical and mathematical tools try to define the business in a systematic way with proof. However the complexities of real time business are significant in nature.

Literature review

There are plenty of literatures defining and explaining Benford's law which are somehow more usable and readable by staunch mathematicians and statisticians. Originally discovered by an American astronomer Newcomb in 1881 on the basis of repetitive distribution law in logarithmic tables, the law became famous after physicist Frank Benford demonstrated it in his research paper "The law of Anomalous Numbers" in 1938. Kruger and Yadavalli described Benford's Law as "counter-intuitive, difficult to explain in simple terms, and has suffered from being described variously as 'a numerical aberration', 'an oddity', 'a mystery' - but also as 'a mathematical gem'''' (2017). Mark Nigrini, a professor and author of Forensics Analytics works majorly on Benford's law application.

Why digits?

Probability, a branch of mathematics concerning numerical descriptions of how likely an event is to occur, is highly helpful in understanding and predicting the uncertainties of the real world. Many sub topics of probability have the capacity to define and create a model which could lower down the risks associated with business decisions or trend analysis. Fashion being highly volatile and riskier in nature requires a detailed understanding of art and science. Generally, fashion is always associated with the non-mathematical side; however, it is actually needed a lot to implement higher numerical calculations.

Benford's Law

While retail and fashion houses use various mathematical and statistical tools to understand, anticipate and deliver business, a very least known theory famously known as **Benford's Law** comes into the picture. It defines a pattern of numbers around everything in the world. It seems trivial at the first glance but as one goes deep into it, the astonishing results start coming in. It

shows the regularity in digits and explains the hidden science of so many theories including fraud in a system.

Also known as the l**aw of anomalous numbers**, or the **first-digit law, the credit of discovery** of it goes to a Canadian-American mathematician Simon Newcomb in 1881, when he found that only first few pages of the logarithm table are used and rest are less or not used. Later in the year of 1938, Frank Benford, a physicist, took twenty domains and multiple data sets to find a relationship with numbers. Further, it was forwarded by Ted Hill in 1995 with many more new experiments and the relationship between Benford's law and real life phenomena. Benford's law is an observation of the frequency distribution of leading digits in real life sets of numerical data. The law has been widely used in the field of accounting, geology, law, economics, psephology, science but probably not in the generous world of fashion.

A set of numbers is said to satisfy Benford's law if the leading digit *d* (d ∈ {1, ..., 9}) occurs with probability

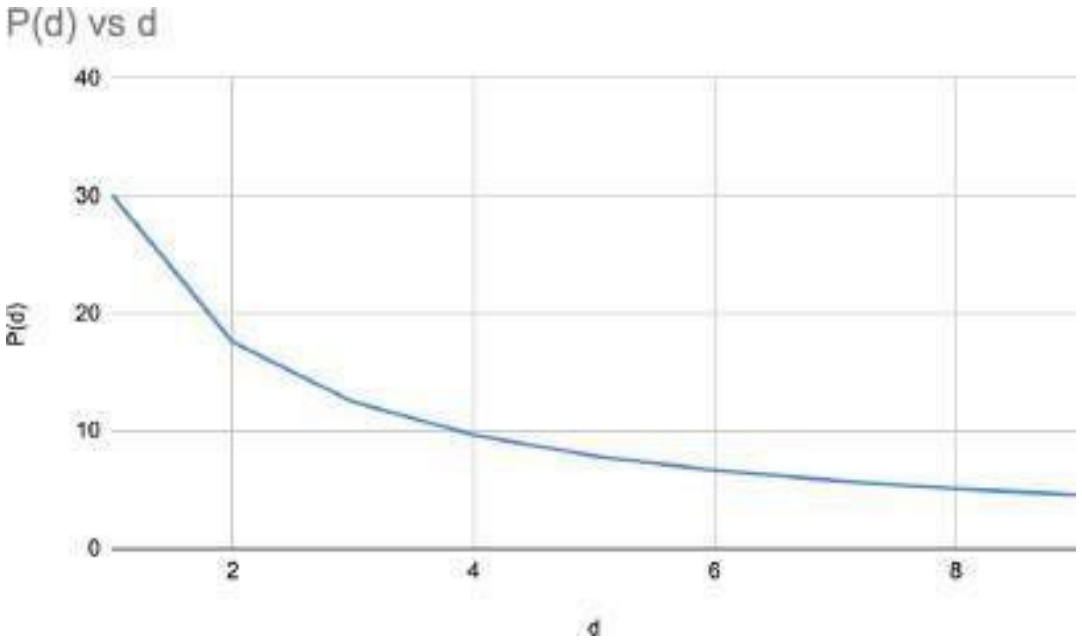
$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}[(d+1)/2] = \log_{10}(1 + 1/d)$$

Title	1	2	3	4	5	6	7	8	9	Samples
Rivers(Area)	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
Newspaper items	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	7.3
Mol Weight	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
Cost data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
n1,n2,n3,...n!	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
Death rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average of all parameters Including some natural phenomena	30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Errors (+ -)	0.8	0.4	0.4	0.3	0.2	0.2	0.2	0.2	0.3	

The leading digits in such sets have following distribution

d	P (d)	Relative Size of P (d)
1	30.1	
2	17.6	
3	12.5	
4	9.7	
5	7.9	
6	6.7	
7	5.8	
8	5.1	
9	4.6	

The graphical representation of the Benford's law is shown below:-



Application of Benford's law

It is a systematic digital analysis technology. This law also obeys Fibonacci series, one of the possible trade areas theories. Some of the following fields are widely applying Benford's law

a. Accounting Fraud detection - Many nations use it to detect Tax evasion and accounting fraud.

b. Analysis of Election Data - Anomaly election data detection is studied.

c. Economic Data

d. Pricing research

e. Genome data

f. Geological data

Methodology

Distributions that can be expected to obey Benford's law

- When the mean is greater than the median and the skew is positive
- Numbers that result from mathematical combination of numbers: e.g. quantity × price
- Transaction level data: e.g. disbursements, sales

Distributions that would not be expected to obey Benford's law

- Where numbers are assigned sequentially: e.g. check numbers, invoice numbers
- Where numbers are influenced by human thought: e.g. prices set by psychological thresholds ($1.99)
- Accounts with a large number of firm-specific numbers: e.g. accounts set up to record $100 refunds

A distribution is said to satisfy Benford's law, if its corresponding random variable does.

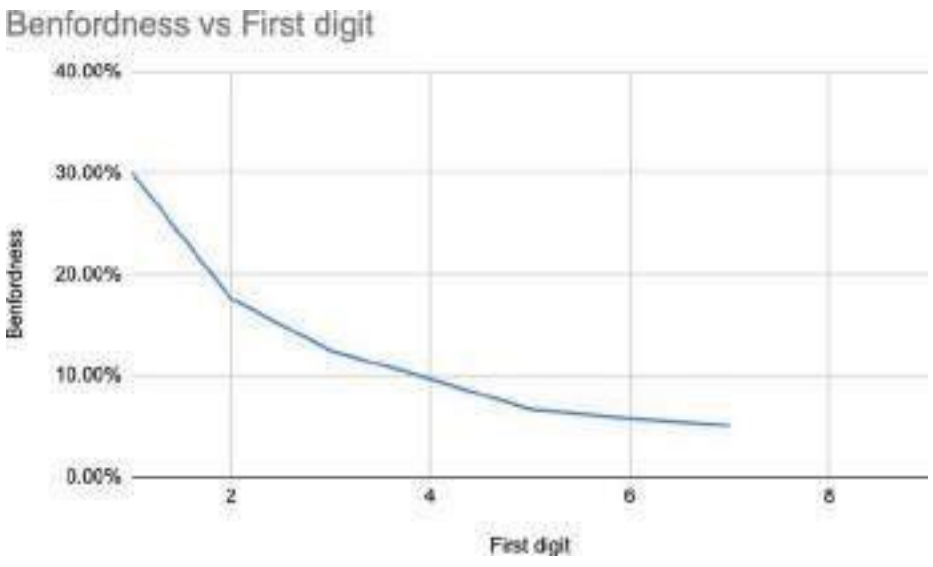
Inverse functions , Scale invariance, Mantissa distribution, Base invariance, Random probability distributions, Convergence etc, Fibonacci series etc.

As per Mark Nigrini's formulation of Benford's Law, the first and second digit tests are significant tests of rationality and are used to determine whether the data set appears reasonable.

Example 1 - Taking a sample data set of 10000 entries of the invoices of a retail store*, the first digit test appears like –

First digit	Benfordness
1	30.10%
2	17.61%
3	12.49%
4	9.69%
5	6.70%
6	5.80%
7	5.12%
8	4.58 %
9	3.02 %

*Fashion retail store of India (Name not disclosed)

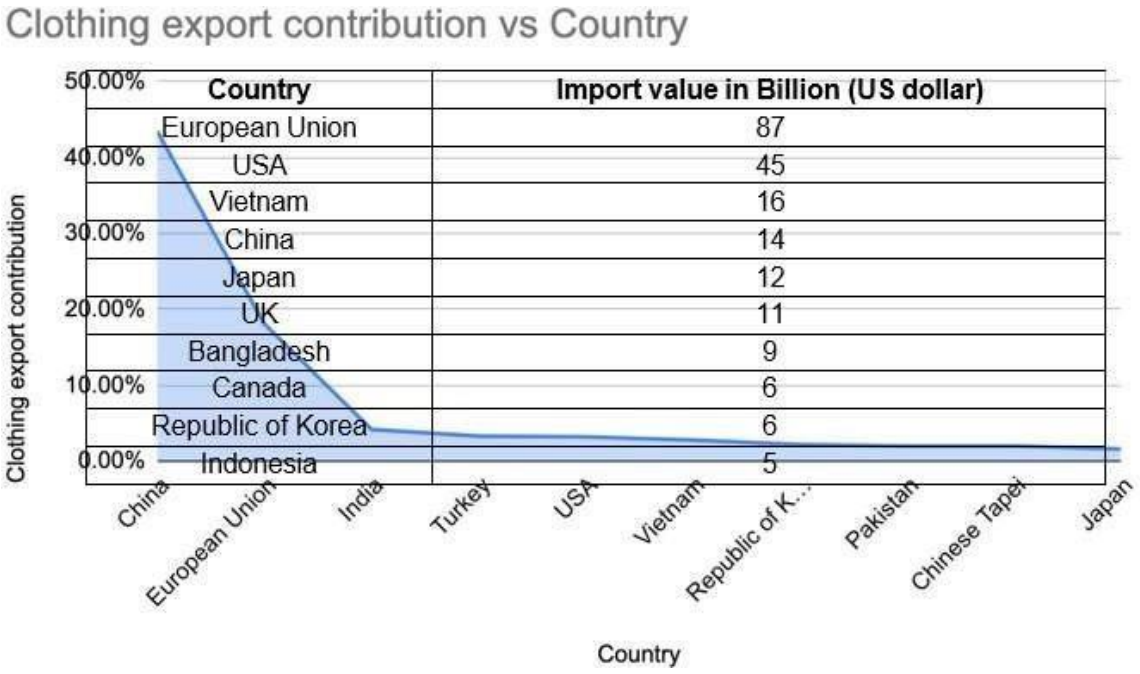


As per the graphical representation, the law of first, second and so on digits follow the law. It proves that the entries are in the right direction and no anomaly is found.

Example 2 - Share in world exports of the leading clothing exporters in 2020, by country

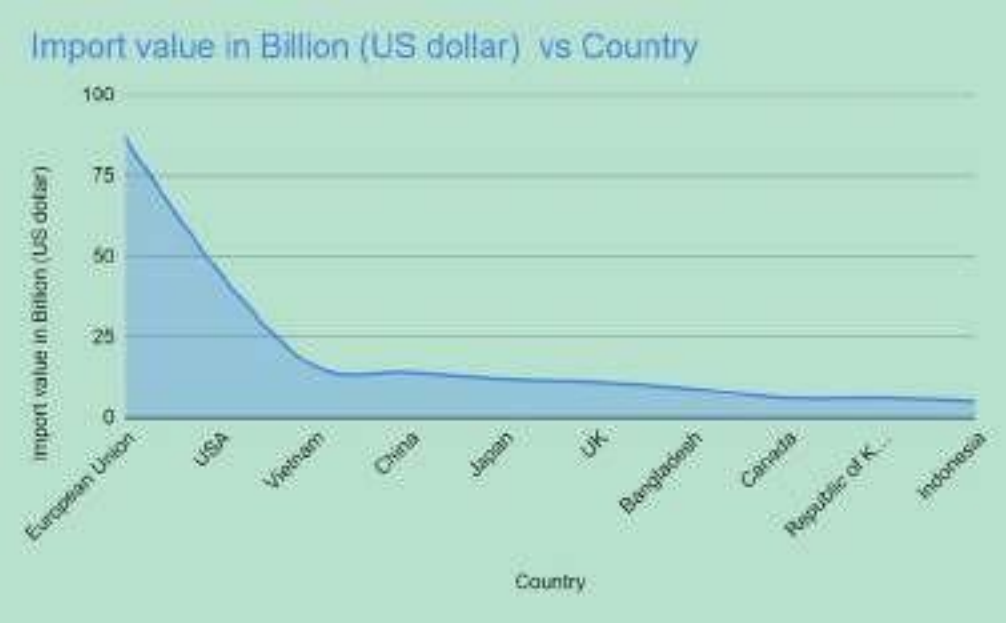
As per the statistica.com, following is the data of 2020, which also follows Benford's law-

Country	Clothing export contribution
China	43.50%
European Union	18.10%
India	4.20%
Turkey	3.30%
USA	3.20%
Vietnam	2.80%
Republic of Korea	2.20%
Pakistan	2%
Chinese Tapei	2%
Japan	1.60%



Country	Import value in Billion (US dollar)
European Union	87
USA	45
Vietnam	16
China	14
Japan	12
UK	11
Bangladesh	9
Canada	6
Republic of Korea	6
Indonesia	5

Example 3 Leading importing countries of apparel worldwide by value in 2020



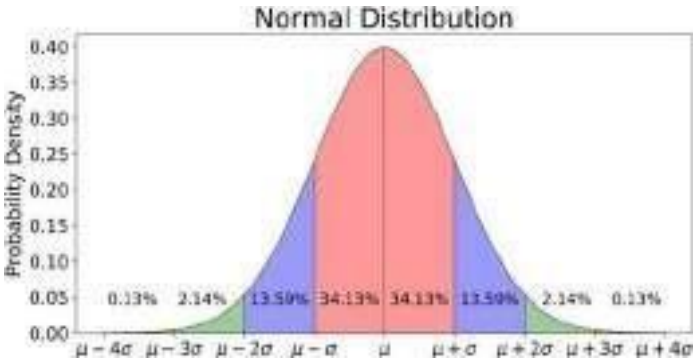
- Every single example with proper data of Fashion field follows Benford's law.
- A distribution is said to satisfy Benford's law, if its corresponding random variable does.
- Inverse functions , Scale invariance, Mantissa distribution, Base invariance, Random probability distributions, Convergence etc, Fibonacci series etc.

Benford's Law and contemporary Technology -

- Big data analytics - Consumer data, Trend analysis, Retail sales data, Supply Chain Management etc.
- Blockchain management - Decentralized system of a shared database. [Crypto currencies]
- Fashion Metaverse and its Crypto currencies

Big Data Analytics -

Normal distribution and other statistical tools are widely used in the area of predictive analysis of fashion big data analysis. The ubiquity of normal distribution applies in all kind of natural phenomena or observations.



Normal (Gaussian) Distribution

- Ubiquity in all kind of natural phenomena
- A probability distribution
- Symmetric
- Unimodal
- Analysis of data set

Benford's law

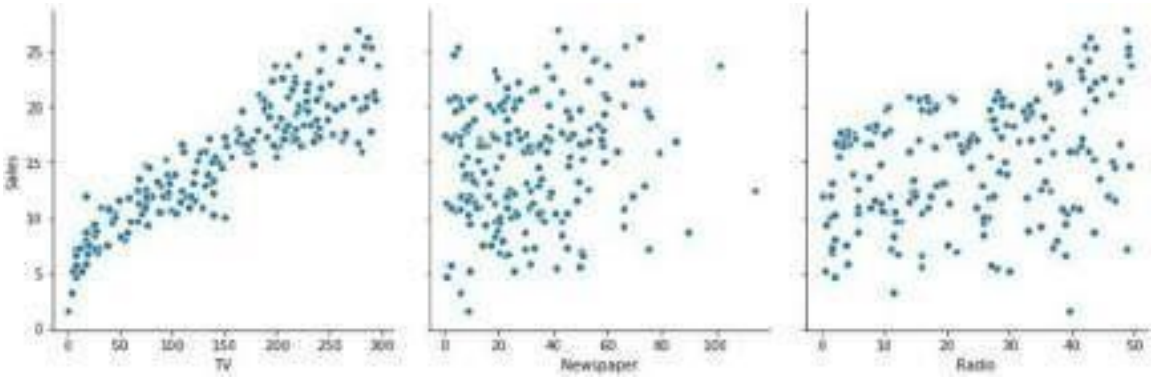
- Phenomenon of significant digits
- Detects patterns
- Catches anomalies or fraud detection

Regression Model Anomalies -

- Regression is used to study the relationship between two variables.
- We can use simple regression if both the dependent variable (DV) and the independent variable (IV) are numerical.
- If the DV is numerical but the IV is categorical, it is best to use ANOVA.

The dependent Variable is Sales and independent variables are cost of advertisement on TV, Newspaper and Radio. From the scatter plot it is clear that there is a positive correlation of dependent variable sales with TV.

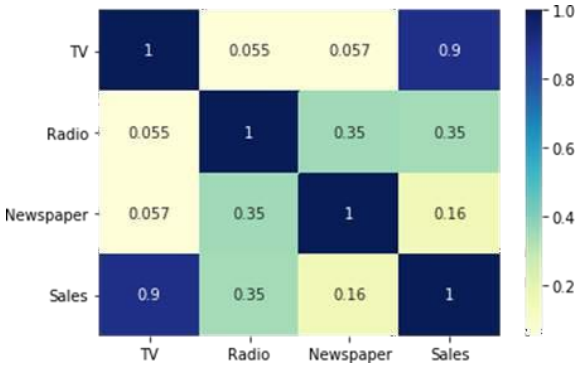
	TV	Radio	Newspaper	Sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9



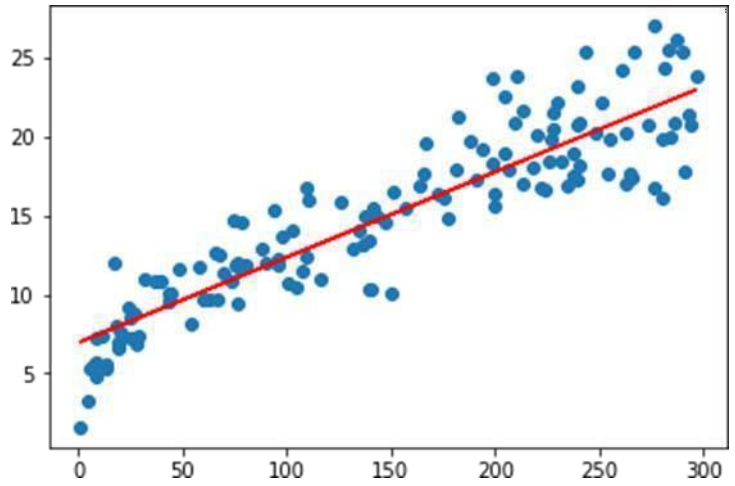
It is visible from the pair plot and the heatmap, the variable TV seems to be most correlated with Sales. So let's go ahead and perform simple linear regression using TV as our feature variable in our case:

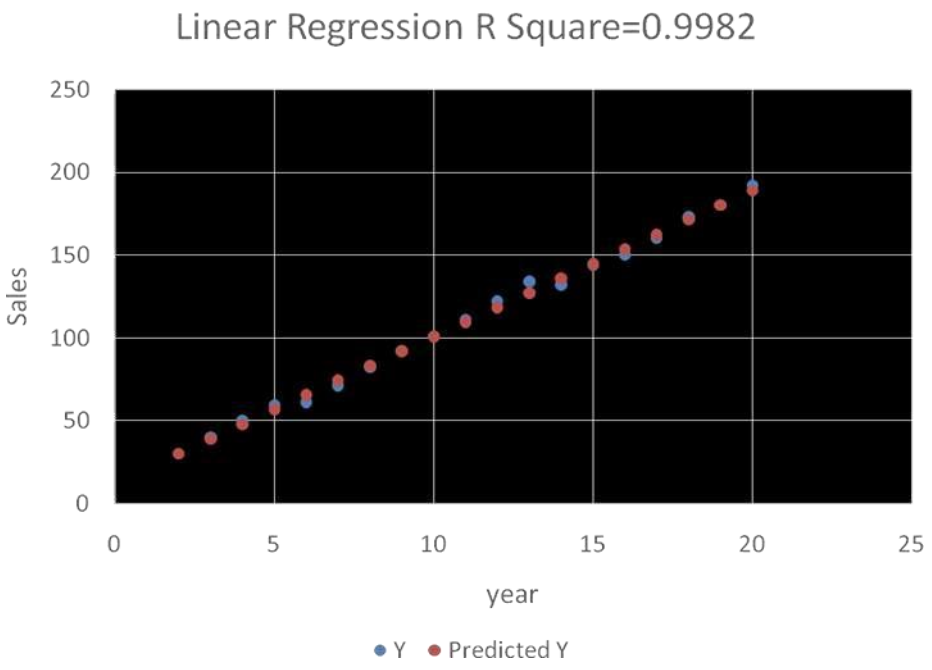
y=c+m1×TV C= constant m1 =Slope

From the regression model Sales=6.948+0.054×TV



Sales vs TV: Line of Best Fit





It is clear from the scatter plot and line of best fit that due to outlier there is a reduction of R squared with such a great extent. Only two outliers are reducing R square from 0.99 to 0.12. So scatter plot can easily identify the anomaly.

Fashion Blockchain Management

- Supply Chain Management
- Reduce Counterfeiting
- Increasing efficiency in production
- Entry to Fashion metaverse
- Fashion NFTs
- Cryptocurrency
- Virtual lands

All the big blockchain platforms by market capitalization that were not biased by any big scandal or lawsuit and that are still functioning three years after the observation time frame, such as Bitcoin (BTC), Ethereum (ETH), or OmiseGo (OMG), conform to Benford's law.

Findings and Analysis

- Establishment of relationship between Benford's law and data generated from various points
- New area of probability distribution could be explored
- Data Fraud detection

- Anomalies can be found
- Implementation in new technology like Blockchain management, Metaverse andrespective crypto world
- Transparency and robust model of business can be adopted.

References

Mark J. Nigrini, *Benford's Law: Applications for Forensic Accounting, Auditing, and FraudDetection* (Hoboken, NJ: Wiley, 2012), pp. 132–35

Pimbley, J.M. (2014). "Benford's Law as a Logarithmic Transformation" (PDF). *MaxwellConsulting, LLC.*

Frank Benford (March 1938). "The law of anomalous numbers". *Proc. Am. Philos. Soc.* **78** (4): 551–
572. JSTOR 984802.

Articles from research gate(accessed on 30/08/2021)

https://en.wikipedia.org/wiki/Benford%27s_law (accessed

on 01/12/2021)Used data from kaggle.com

(01/12/2021)

Statistica.com/statistics/1094515/share-of-the-leading-global-textile-clothing-by-country