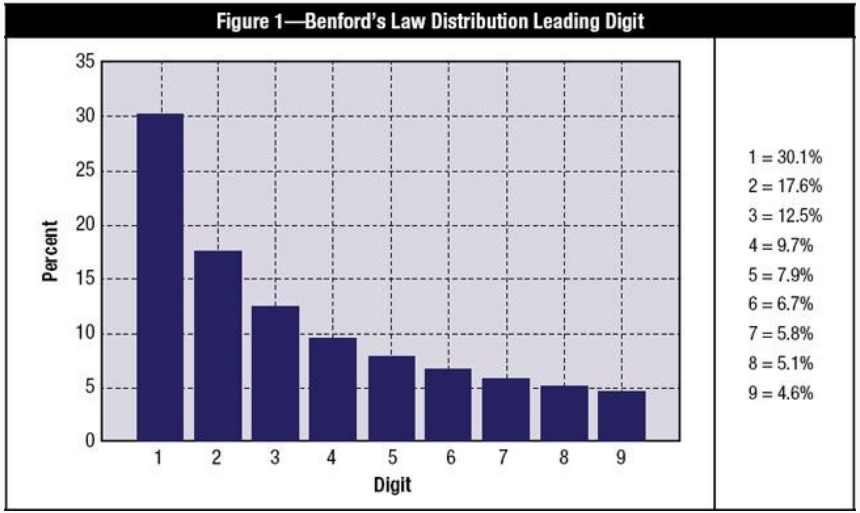
The First Digit Law, A Phenomenon All Around Us

Ishana Karanwal

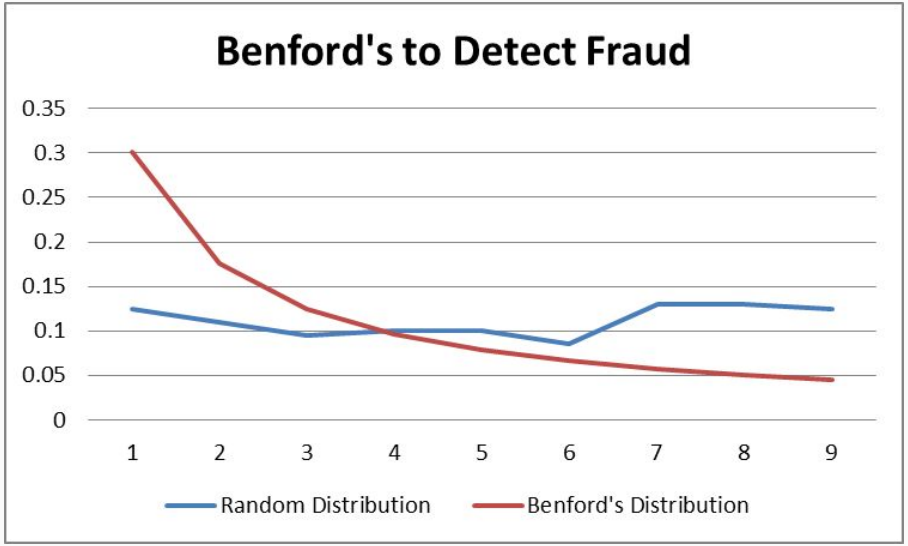
Background Information

Benford's law, which is also referred to as the First digit law, is the theory that the leading digit in numbers taken from random data sets will have the first digit of one appear 30.1% more than any other number.



Background Information

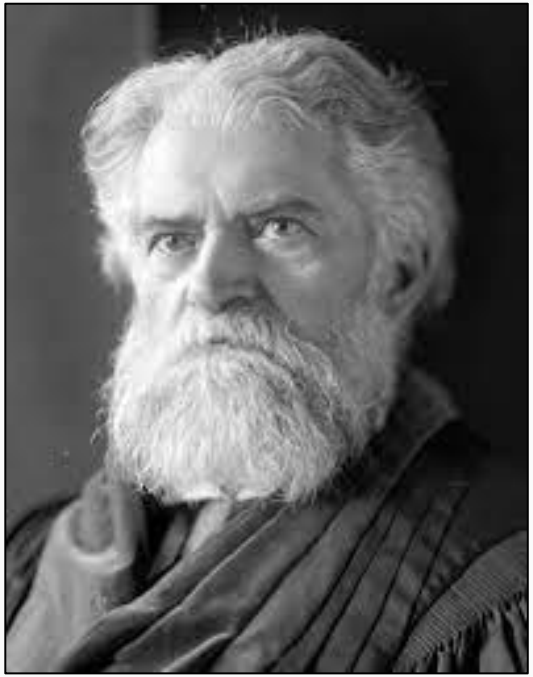
The expected probability of one, or any different number, being the first digit is 11.1%. The likelihood of two being the first digit is 17.6%, the probability of three is 12.5%. The chance of a number being as the first digit is in descending order until nine. The satisfying results of Benford's law form a perfect graph in most scenarios. The formula for this law is log10(1+1/n).



My Hypothesis

By observing the numbers in a data set, it is possible to see if the digits follow the nature of Benford's law. These observations help to verify if the data set is valid or fraudulent. If Benford's law is applied more frequently in the right circumstances, it will be efficient in distinguishing errors and scams.

History of Benford’s Law - Simon Newcomb



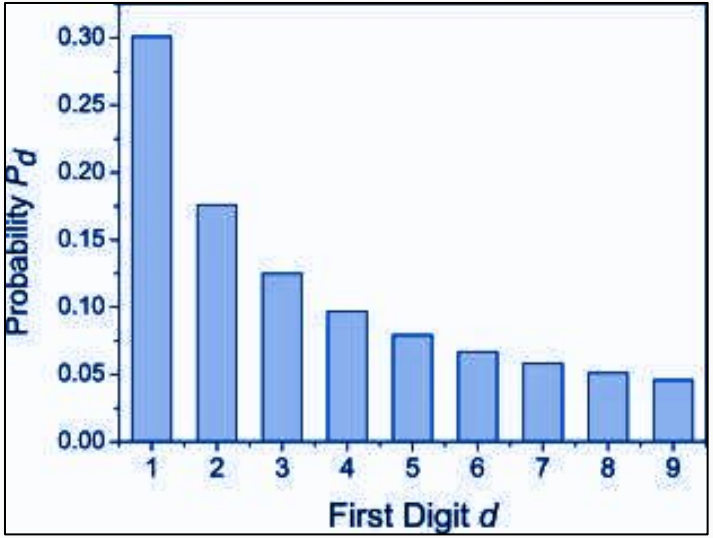
The theory of Benford's law started back in 1881. The Astronomer Simon Newcomb noticed that the book's pages at the start of a book on logarithm tables were more worn than the pages towards the end of the book.

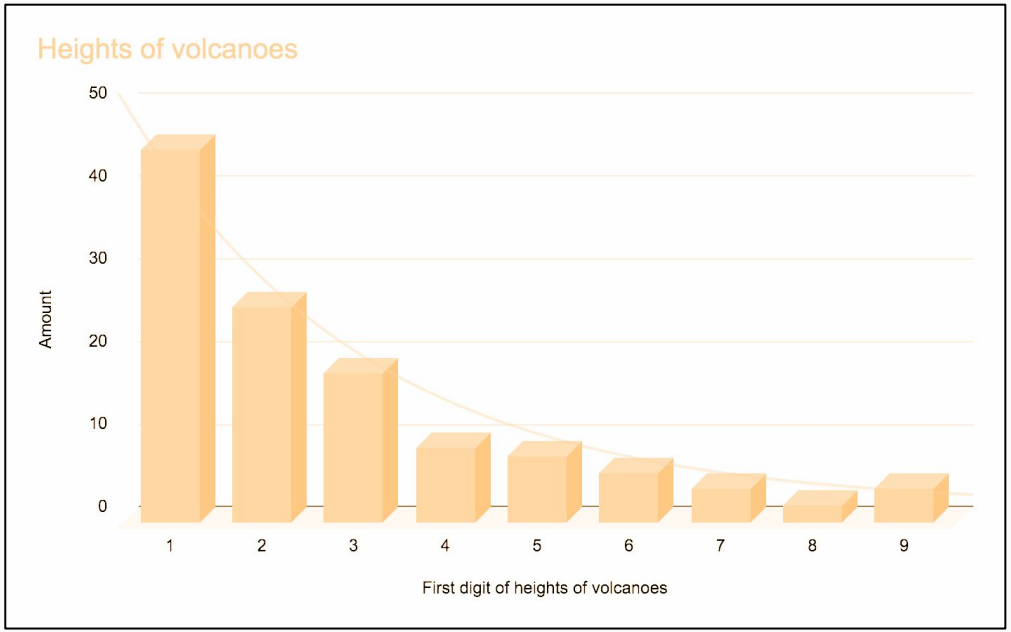
History of Benford’s Law - Frank Benford

Fifty-seven years later, in 1938, Frank Benford tested Newcomb's hypothesis against 20 different data sets. He published a scholarly paper verifying the verification of Benford's law. Newcomb was the scientist to realize the pattern of the law. However, Benford put the law down on paper which gave him the most credit for this fascinating principle.

Examples of Graphs that Obey Benford's Law

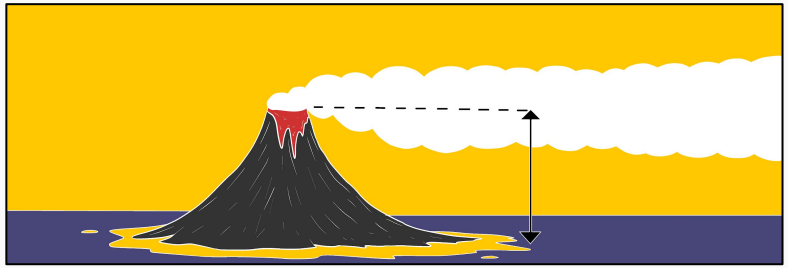
I tested Benford's law against different data sets and put my results in a bar graph to see if the numbers had a satisfactory outcome.

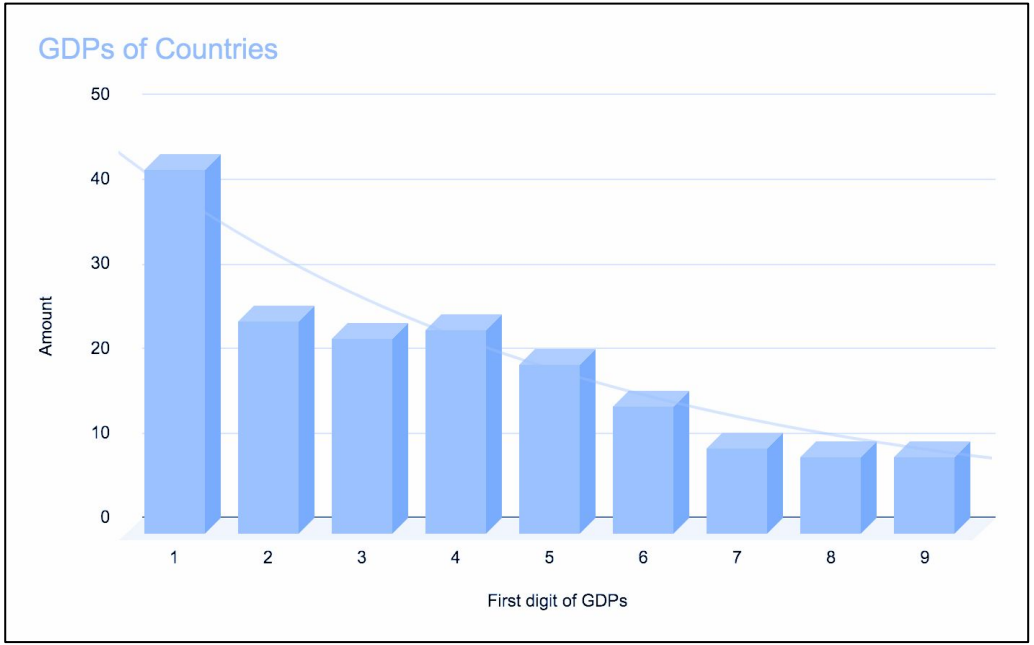




Heights of volcanoes	
1	45
2	26
3	18
4	9
5	8
6	6
7	4
8	2
9	4

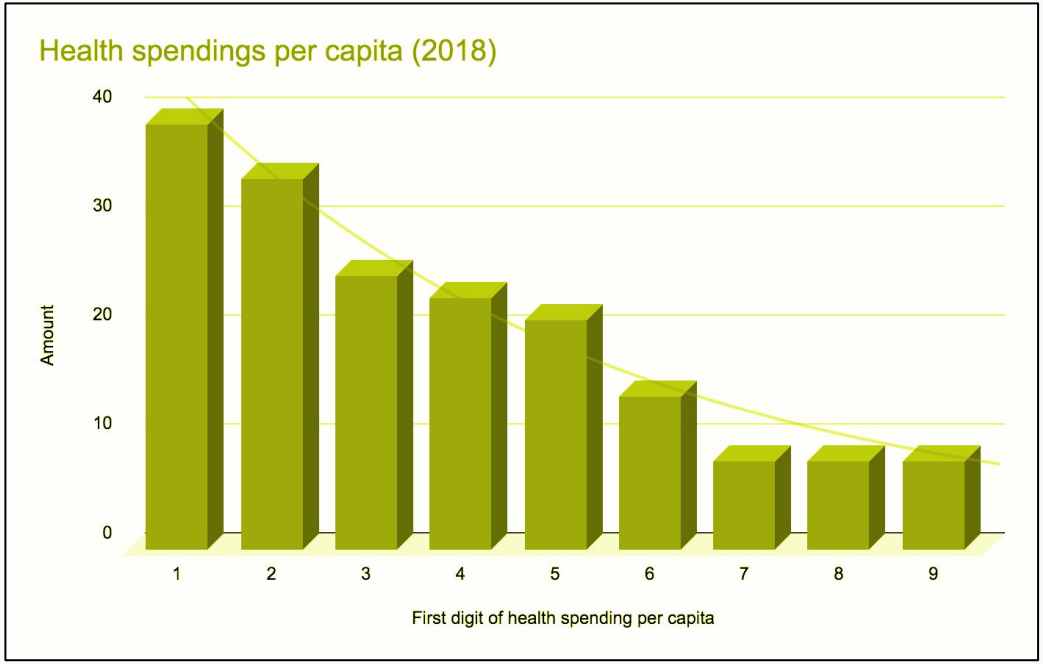
I made sure to add a trendline of how a graph, that is following Benford's law perfectly, would look like. The number one appears 36% more than any other number. The heights of volcanoes follow Benford's law almost perfectly.





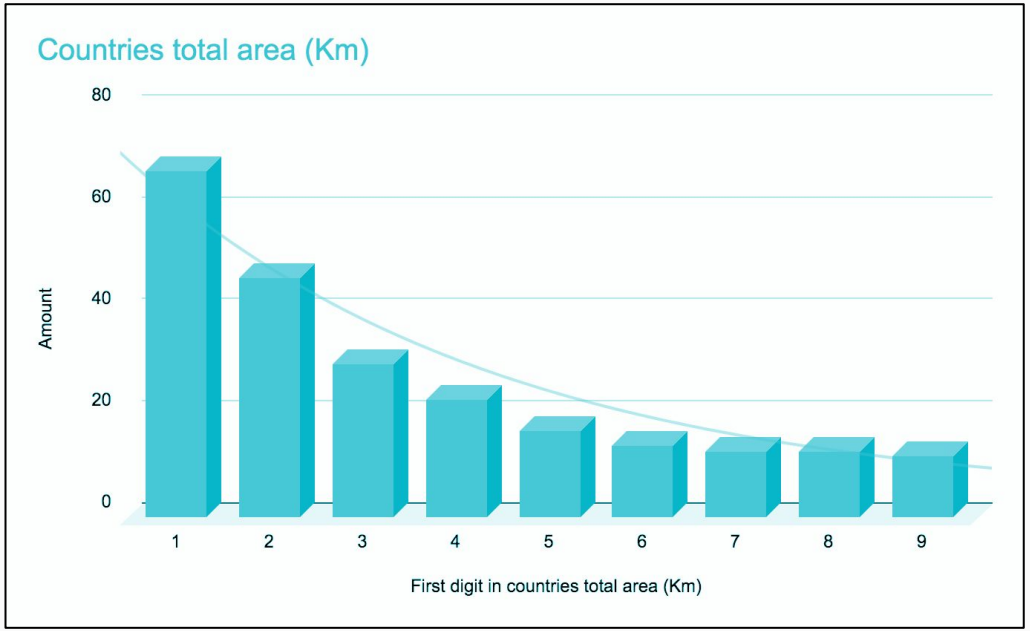
GDPs of Countries	
1	43
2	25
3	23
4	24
5	20
6	15
7	10
8	9
9	9

The GDP's of countries do not follow Benford's law as perfectly as the volcanoes. The trendline does not match up with the numbers flawlessly. The number 1 appears 36% of the time.



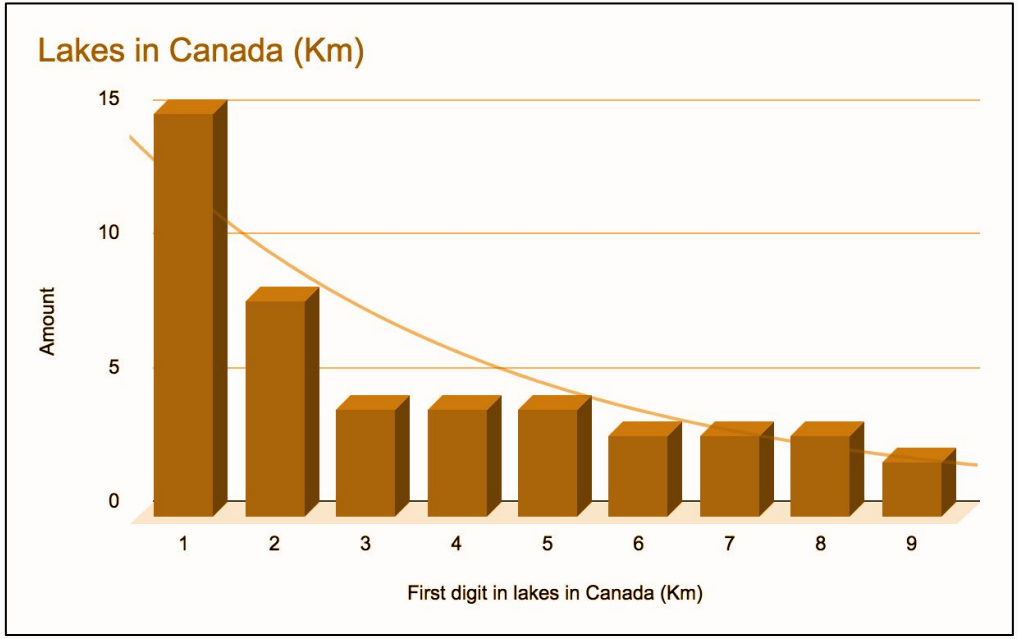
Health spending per capita (2018)	
1	39
2	34
3	25
4	23
5	21
6	14
7	8
8	8
9	8

For the graph of health spendings as of 2018 the number one appears 21% of the time.



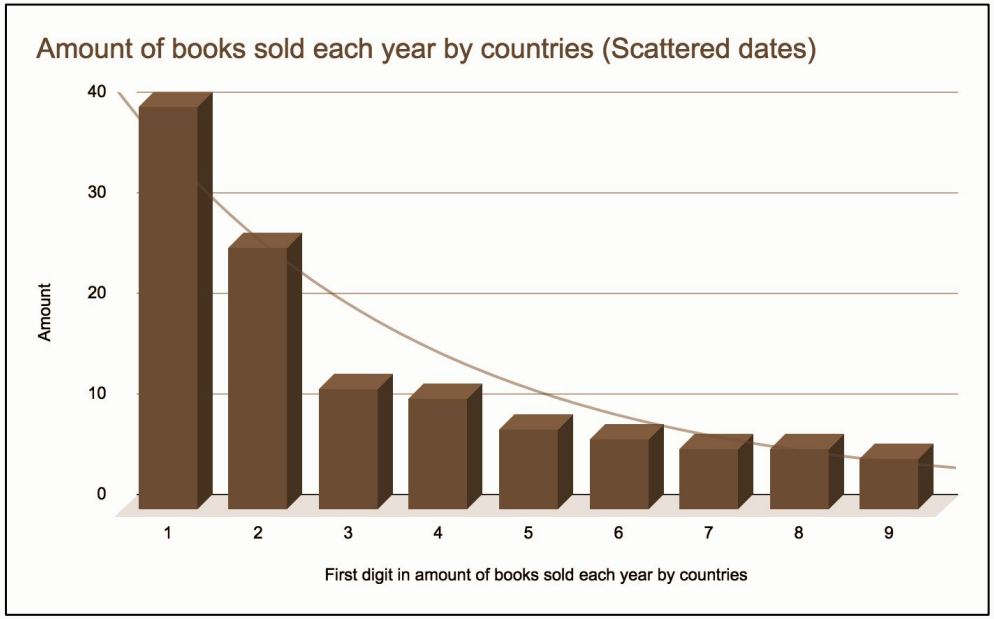
Counties total area (Km squared)	
1	68
2	47
3	30
4	23
5	17
6	14
7	13
8	13
9	12

The total areas of countries was another geographical related graph and followed Benford's law almost perfectly. The number 1 appeared 28% of the time.



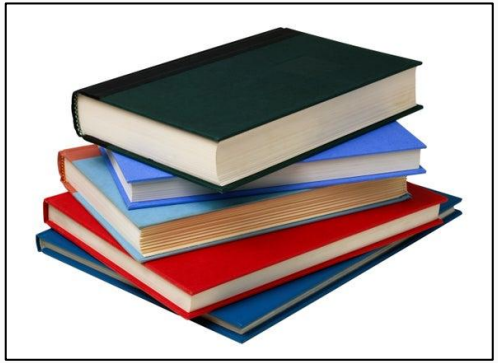
Lakes in Canada (Km)	
1	15
2	8
3	4
4	4
5	4
6	3
7	3
8	3
9	2

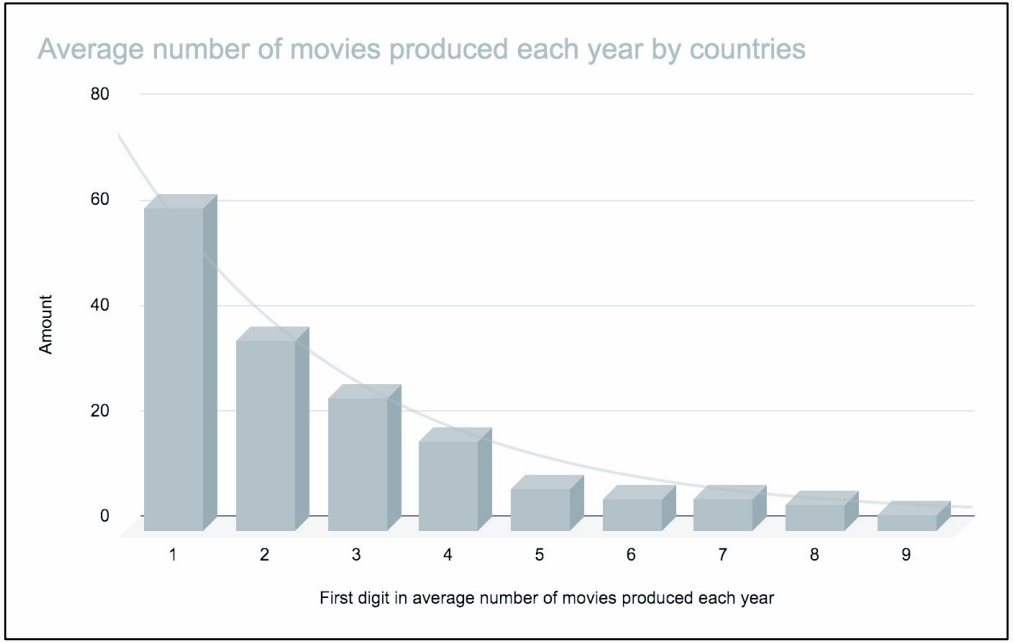
During my tests on random data sets, I noticed that most of the time, Benford's law always worked on geographical related data. In the lakes in Canada graph the number one appears %32 of the time.



Amount of books sold each year by countries(Scattered dates)	
1	40
2	26
3	12
4	11
5	8
6	7
7	6
8	6
9	5

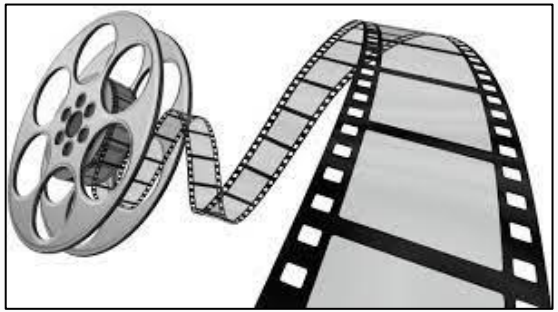
In this graph about the amounts of books sold each year, the number one appears %33 of the time

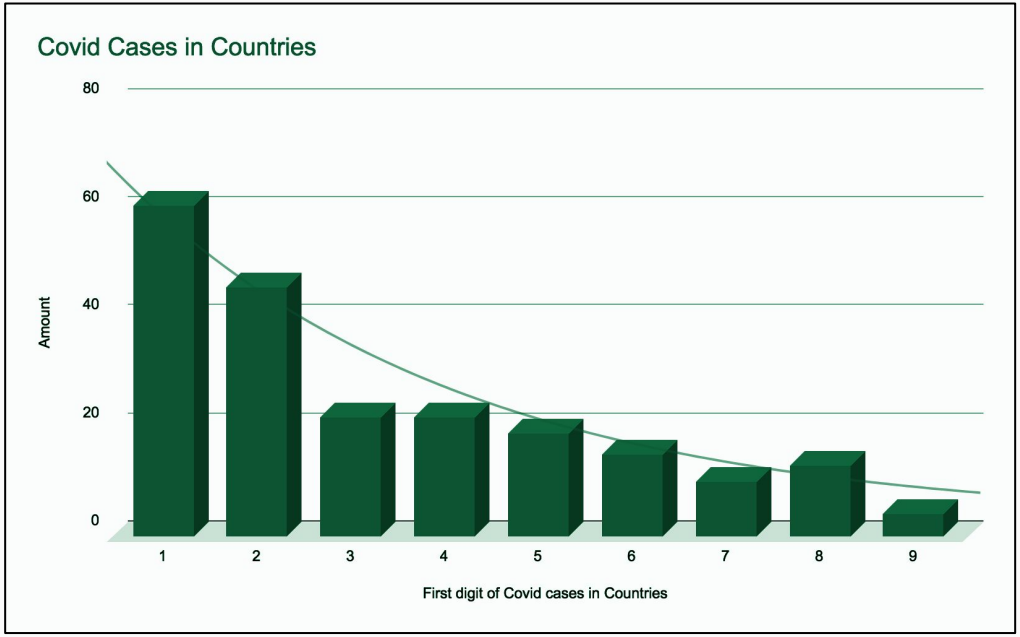




Average number of movies produced each year by countries	
1	61
2	36
3	25
4	17
5	8
6	6
7	6
8	5
9	3

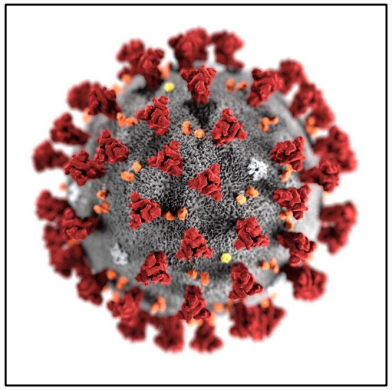
I did not expect for this graph to follow Benford's law so perfectly. The number one appeared %36 of the time.





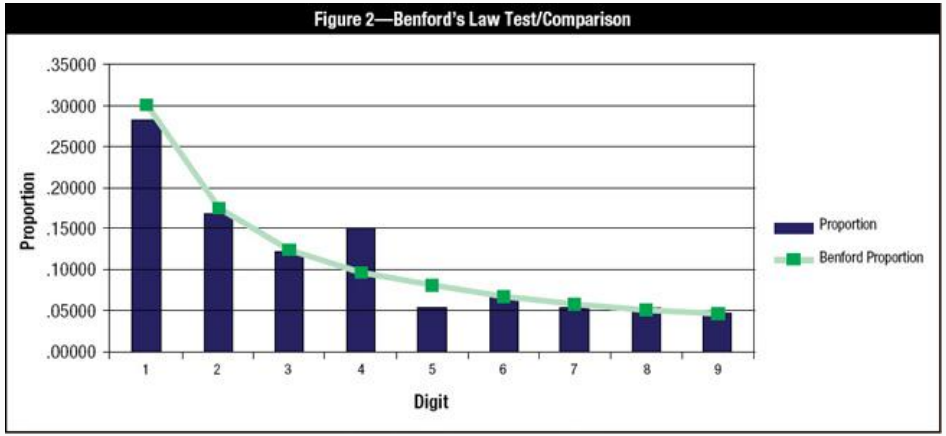
Covid Cases in Countries	
1	61
2	46
3	22
4	22
5	19
6	15
7	10
8	13
9	4

In the Covid Cases graph, the number one appeared %28 of the time.



What are the Correct Circumstances for Applying Benford's Law?

Since the time it was discovered, Benford's law has been proven to be an efficient tool in fraud detection. An example given by Mark Nigrini showed that Benford's law might be used as an indicator for accounting fraud. A fraudster wrote multiple cheques to himself worth around US $100,000. The digits 7, 8, and 9 had percentages that did not match the Benford's law analysis.

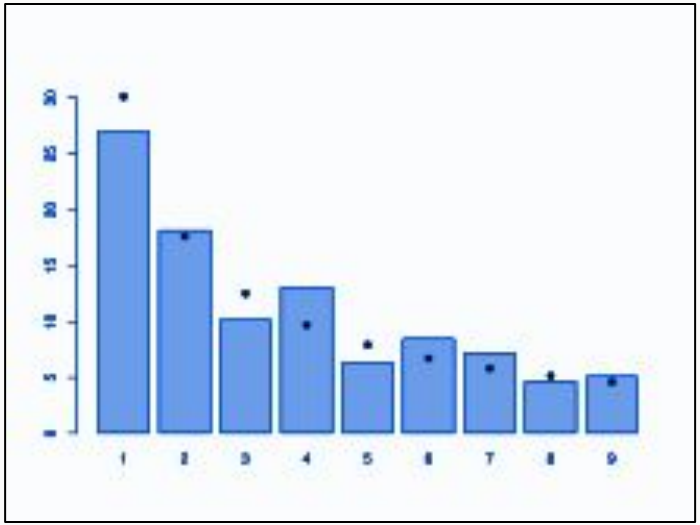


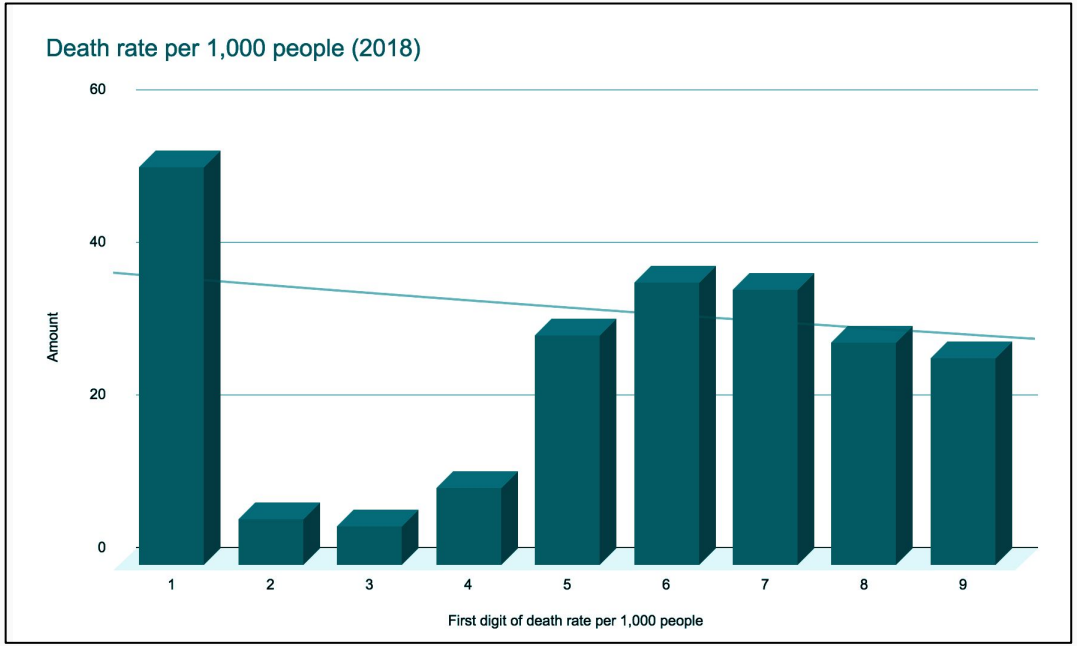
What are the Correct Circumstances for Applying Benford's Law?

Benford's is a tool that has potential in the online world. For instance, Benford's law was used as evidence for voter fraud in the 2009 Iranian election. This law is legally admissible as evidence for criminal cases in the US at federal, state, and local levels. Benford's law shows many other examples to be successful: purchase orders, loan data, stock prices, inventory prices, customer refunds, telephone numbers, and passengers in a plane.

Examples of Graphs that Disobey Benford’s Law

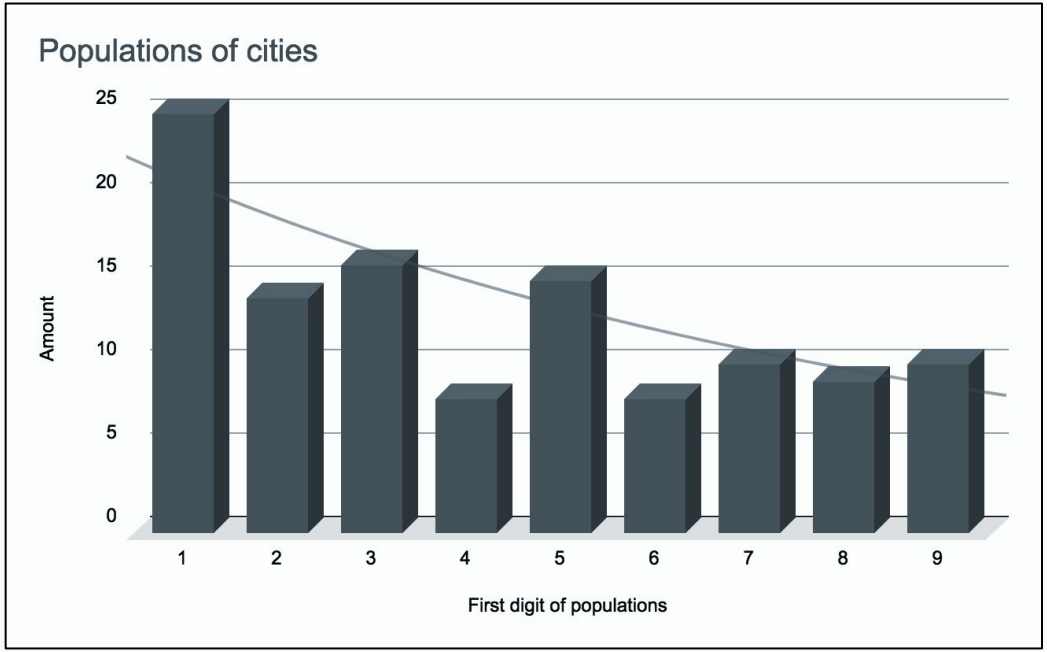
While I was collecting data, I realized that Benford’s law does not work on everything. The following 3 examples are data sets that do not follow Benford’s law.





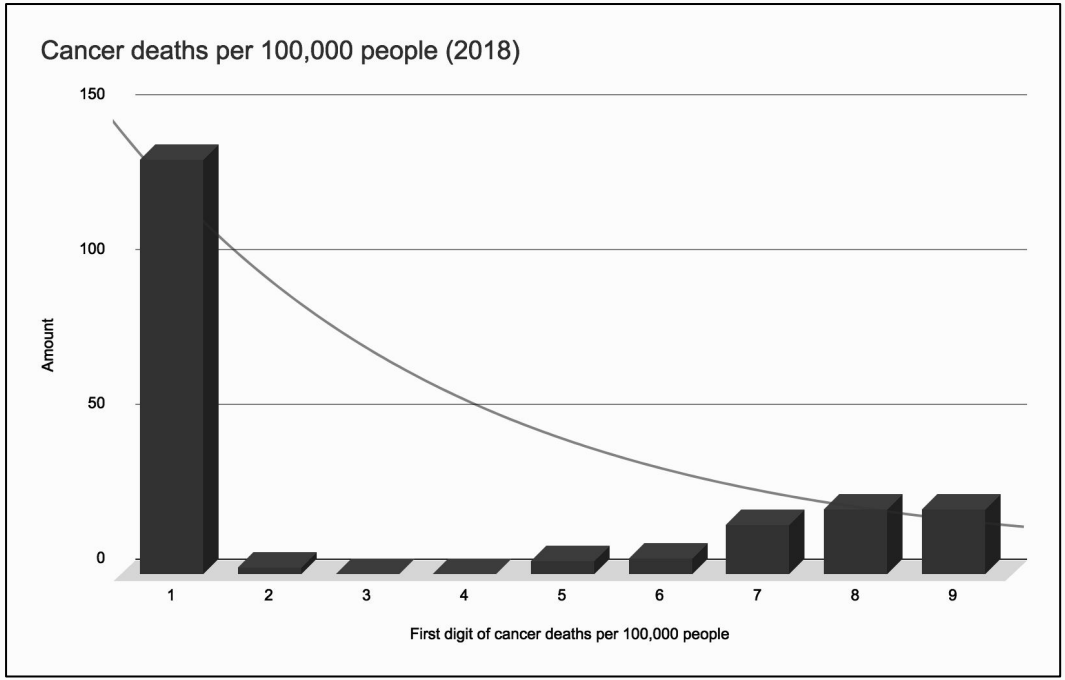
Death rate per 1,000 people (2018)	
1	52
2	6
3	5
4	10
5	30
6	37
7	36
8	29
9	27

The death rate per 1,000 people did not follow Benford's law. However, the number one still shows up the most. The number one shows up 24% of the time.



Populations of cities	
1	25
2	14
3	16
4	8
5	15
6	8
7	10
8	9
9	10

The specific set of data that I took my number did not follow Benford's law. According to Mark Nigini the populations of cities sometimes will follow Benford's law and sometimes it will not. The number one appeared %22 of the time.



Cancer deaths per 100,000 (2018)	
1	134
2	2
3	0
4	0
5	4
6	5
7	16
8	21
9	21

There was an abundance of ones and no threes and fours.
I was surprised that this graph was not more scattered.
The number one appeared %66 percent of the time!

Where will Benford's Law Fail to be Successful?

Although Benford's law is valuable in certain circumstances, there are limitations to where this law can be applied. The law will not work when applied to data sets that are not random and have ordered and data sets with minimums and maximums.

Where will Benford's Law Fail to be Successful?

For example, hourly wages will not follow Benford's law because there are minimum and possibly maximum wages given. Telephone numbers are not random numbers since they have specific numbers at the beginning based on the area. After that, the three to four digits are also not random since they have been given out in order, so telephone numbers will not follow Benford's law.

Interesting Facts about Benford’s Law!

-In sports, Benford's law will follow its nature of how many passes and goals there are. In songs, Benford's law follows how long each note is played in a piece.

-It will help catch tax cheats. Based on the numbers you fill out on the form, they can tell if you are lying just by looking at it.

-Former IRS (Internal Revenue Service) officers and current IRS officers cannot disclose information on whether they are using Benford's law or not.

-European Union is open to talking about Benford's law, and they even have a Benford's law conference.

-FBI data of crimes follow Benford's law almost perfectly.

Interesting Facts about Benford’s Law!

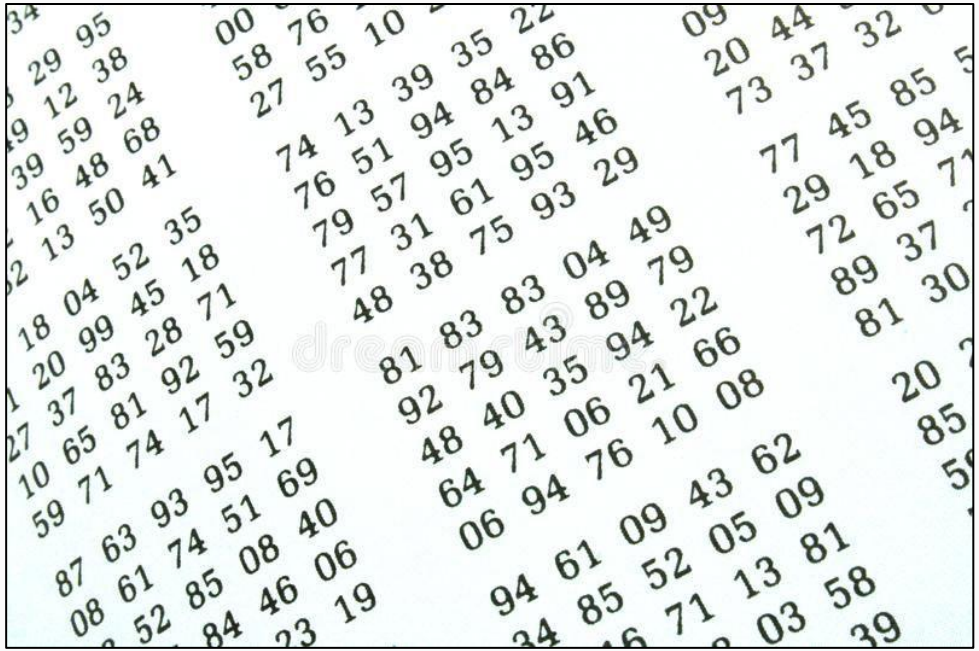
- If everyone voted their first choice in elections, it would follow Benford's law. If they voted their second or third choice, it would not follow Benford's law.

-When you go on social media and look at your followers and your friends' followers, the numbers will follow Benford's law. The number of followers that did not follow Benford's law turned out to be Russian Bots for a researcher.

-When you manipulate an image, you save it more times, the pixels have been played with, and it will not follow Benford's law.

-The distances of galaxies and planets from Earth will follow Benford's law despite the unit you measure the length in.

Conclusion



Benford's law is an effective tool for fraud detection and errors when applied in the right circumstances, such as geographical features, GDPs, movies, and more. This law will not work in cases where there are no random numbers or some limitations. The phenomenon and the probabilities that come from it are based on the mathematical algorithms of digits' appearance in randomly generated numbers in a large data set.

Citations - APA

-GDP by Country - Worldometer. (2021). Retrieved 12 March 2021, from https://www.worldometers.info/gdp/gdp-by-country/

-Volcanoes of the World: alphabetic list from A-Z - find by name. (2021). Retrieved 12 March 2021, from https://www.volcanodiscovery.com/volcanoes/alphabetical-list/a-z.html

-Using Excel and Benford's Law to detect fraud. (2021). Retrieved 12 March 2021, from https://www.journalofaccountancy.com/issues/2017/apr/excel-and-benfords-law-to-detect-fraud.html

-Coronavirus Update (Live): 119,113,952 Cases and 2,641,762 Deaths from COVID-19 Virus Pandemic - Worldometer. (2021). Retrieved 12 March 2021, from https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1

-List of the 100 largest municipalities in Canada by population. (2021). Retrieved 12 March 2021, from https://en.wikipedia.org/wiki/List_of_the_100_largest_municipalities_in_Canada_by_population

-Benford's Law. (2021). Retrieved 12 March 2021, from https://www.londonsoda.com/articles/benfords-law

-Death rate, crude (per 1,000 people) | Data. (2021). Retrieved 12 March 2021, from https://data.worldbank.org/indicator/SP.DYN.CDRT.IN

Citations - APA

-Health spending per capita by country, around the world | TheGlobalEconomy.com. (2021). Retrieved 13 March 2021, from https://www.theglobaleconomy.com/rankings/Health_spending_per_capita/

-ALL CANCERS DEATH RATE BY COUNTRY. (2021). Retrieved 13 March 2021, from https://www.worldlifeexpectancy.com/cause-of-death/all-cancers/by-country/

-Before you continue to Google Search. (2021). Retrieved 13 March 2021, from https://www.google.com/search?q=lakes+in+canada&oq=lakes+in+canada&aqs=chrome..69i57j0l5j46j69i60.2072j0j7&sourceid=chrome&ie=UTF-8&safe=active&ssui=on

-Book Statistics - Worldometer. (2021). Retrieved 13 March 2021, from https://www.worldometers.info/books/

-Using Excel and Benford's Law to detect fraud. (2021). Retrieved 13 March 2021, from https://www.journalofaccountancy.com/issues/2017/apr/excel-and-benfords-law-to-detect-fraud.html#:~:text=The%20story%20of%20Benford's%20Law,the%20end%20of%20the%20book.

-Benford's Law -- from Wolfram MathWorld. (2021). Retrieved 13 March 2021, from https://mathworld.wolfram.com/BenfordsLaw.html

-Understanding and Applying Benfords Law. (2021). Retrieved 13 March 2021, from https://www.isaca.org/resources/isaca-journal/past-issues/2011/understanding-and-applying-benfords-law#:~:text=Benford's%20Law%20can%20recognize%20the,numbers%20in%20large%20data%20sets.

Citations - APA

-Random numbers. A random list of random numbers , #Aff, #numbers, #Random, #list, #random #ad | Image, Numbers, Stock photos. (2021). Retrieved 13 March 2021, from https://www.pinterest.ca/pin/714102084649508223/

-Amin, S. (2021). 3. Benford’s Law. Retrieved 13 March 2021, from https://seprianhidayatamin.wordpress.com/2016/04/04/3-benford-analysis/

-Brigden, N. (2020). *Digits* [Video]. Connected: Hidden science behind everything.

-Simon Newcomb. (2021). Retrieved 16 March 2021, from https://en.wikipedia.org/wiki/Simon_Newcomb

-Coleman, D., & Coleman, D. (2021). Benford's Law and Financial Statements - Audit Analytics. Retrieved 16 March 2021, from https://blog.auditanalytics.com/benfords-law-and-financial-statements/

-Extending Classroom Management Online. (2021). Retrieved 16 March 2021, from https://www.edutopia.org/article/extending-classroom-management-online

Citations - APA

-Benford's law. (2021). Retrieved 16 March 2021, from https://en.wikipedia.org/wiki/Benford%27s_law

-Height - Volcano Top Trumps. (2021). Retrieved 16 March 2021, from https://volcanoestoptrumps.org/our-categories/height/

-Real Gross Domestic Product (Real GDP) Definition. (2021). Retrieved 17 March 2021, from https://www.investopedia.com/terms/r/realgdp.asp

-.health domains are now available. (2021). Retrieved 17 March 2021, from https://get.health/

-Lindholm, J., Bodette, M., Lindholm, J., Bodette, M., Lindholm, J., & Bodette, M. et al. (2021). Why Is The World Split Into Countries?. Retrieved 17 March 2021, from https://www.vpr.org/post/why-world-split-countries#stream/0

-Free Film Reel Png, Download Free Clip Art, Free Clip Art on Clipart Library. (2021). Retrieved 17 March 2021, from http://clipart-library.com/free/film-reel-png.html

-What do you call the disease caused by the novel coronavirus? Covid-19. (2021). Retrieved 17 March 2021, from https://www.statnews.com/2020/02/11/disease-caused-by-the-novel-coronavirus-has-name-covid-19/