

*Szkoła Podstawowa Nr1
im. Andrzeja Knapczyka-Ducha
w Czerwiennem
tel. (18)2658106*

Wędrowki Benforda po krainie liczb

Opiekun szkolny

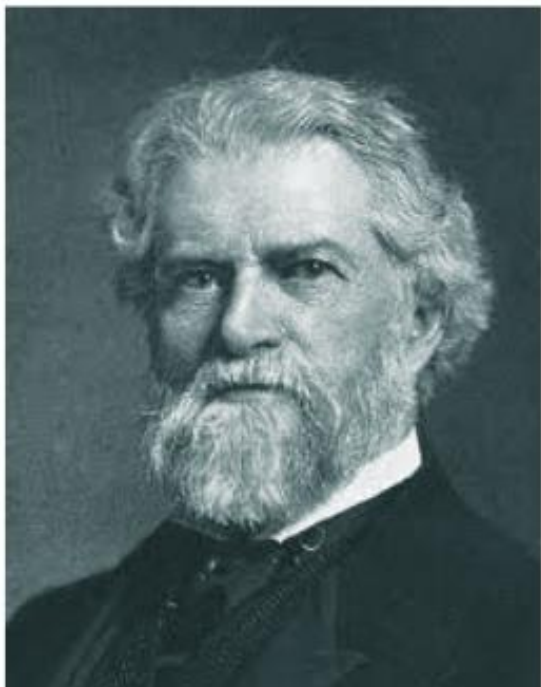
Anna Kutyła

Bartłomiej Staszal

Klasa V

Czerwiennie luty 2024

Odkrywcy prawa Benforda



Simon Newcomb (1835 – 1909)



Frank Benford (1883 – 1948)

Tablice logarytmiczne wraz z suwakiem logarytmicznym były podstawowymi pomocami do obliczeń naukowych, geodezyjnych, astronomicznych i inżynierskich. Za twórcę logarytmów uważa się szkockiego matematyka - **Johna Napier** (lipiec 1614 r.)

Simon Newcomb – astronom i matematyk – do obliczeń astronomicznych korzystał z tablic logarytmicznych, (takie tablice zawierają po prostu listę wartości funkcji logarytmicznej, obliczonych z dużą dokładnością. Główny pożytek z nich był taki, że korzystając z tożsamości $\log(xy) = \log(x) + \log(y)$ przy użyciu takich tablic można łatwo mnożyć). W bibliotece United States Naval Observatory natrafił na egzemplarz którego początkowe strony były bardziej zniszczone niż końcowe, a to oznaczało, że są one częściej używane. A jeśli tak jest, to z jakiegoś powodu korzystający z tablic częściej szukali logarytmów z liczb zaczynających się na jeden niż z liczb zaczynających się na dwa, i tak dalej. W swoim opracowaniu z 1881 r. opublikował nawet częstość występowania poszczególnych cyfr na początku liczb. Stwierdził, że tych zaczynających się od **jedyńki** jest **30,1%**, tych zaczynających się od **dwójki** – **17,6%**, a tych zaczynających się od dziewiątki tylko **4,6%**. Simon Newcomb jako pierwszy wyprowadził wzór , który pozwala wyliczyć prawdopodobieństwo, z jakim dana cyfra będzie pierwszą cyfrą znaczącą.

$$P = \log_{10}\left(1 + \frac{1}{d}\right)$$

gdzie:

d oznacza kolejną cyfrę.

Wzór ten pozwala wyliczyć prawdopodobieństwo występowania kolejnych cyfr na pierwszym miejscu liczb z danego zbioru danych. Opracowanie, które liczyło dwie strony, nie znalazło uznania i musiało poczekać ponad pół wieku, a Simon zajął się przygotowaniem tablic astronomicznych. Nazwiskiem Simona Newcomba została nazwana planetoida i krater na księżycu.

Frank Benford pracował jako inżynier w firmie General Electric. W 1930 roku zauważył coś niezwykłego przeglądając tablice logarytmiczne. Pierwsze strony tych tablic były bardziej przybrudzone niż pozostałe. Stwierdził, że przeglądający je naukowcy szukali liczb zaczynających się od cyfry 1 częściej , niż tych zaczynających się od 2, 3 i tak dalej. niż cyfry Benford spędził kilka lat na zbieraniu różnych danych statycznych dotyczących

różnorodnych zjawisk. W 1938 roku opublikował ponad 20 000 różnych wartości takich jak wielkości zbiorników wodnych, długości rzek, populacje państw, ceny akcji, stałe fizyczne. Okazało się, że ta zasada potwierdza się dla większości z nich. Swoje spostrzeżenia opublikował w czasopiśmie Proceedings of the American Philosophical Society swój artykuł pt. „The Law of Anomalous Numbers”. Opisał w nim zaobserwowane zjawisko dotyczące częstotliwości występowania cyfr na pozycjach znaczących w liczbach pochodzenia naturalnego (zaobserwowanych w przyrodzie - nie generowanych sztucznie). Zjawisko to nazwał prawem pierwszej cyfr (ang. First Digit Law).

Prawo Benforda oznacza fakt częstego występowania charakterystycznego rozkładu w dużych zbiorach danych, gdzie nie wszystkie liczby występują równie często. Nazwane jest również rozkładem Benforda lub prawem pierwszych (znaczących) cyfr.

Zgodnie z tym prawem około 30% liczb w zestawie danych powinno rozpoczynać się od cyfry 1, a tylko 5% – od cyfry 9. Ilustruje to załączony histogram w kolorze granatowym, w którym pokazano funkcję gęstości prawdopodobieństwa rozkładu Benforda $P(k)$ opisanego wzorem:

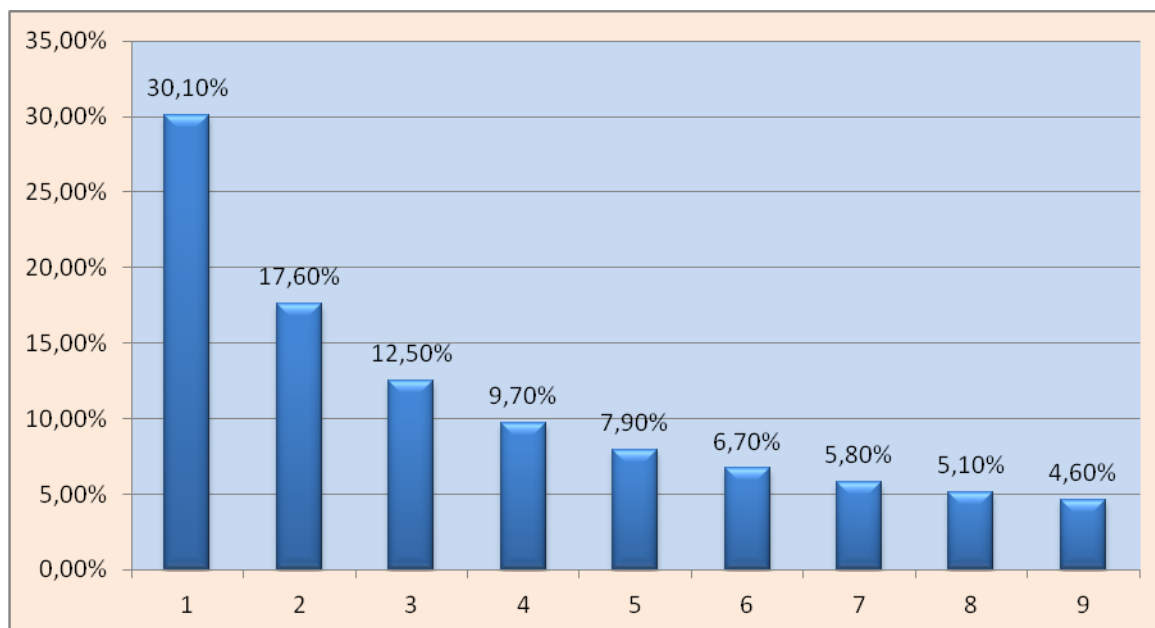
$$P(k) = \log(1 + 1/k)$$

gdzie:

„k” oznacza pierwszą znaczącą cyfrę (k=1, 2, 3, 4, 5, 6, 7, 8, 9) dla liczb ze zbioru danych.

Zakres danych nie może być za bardzo zawężany i musi obejmować kilka rzędów wielkości.

Rysunek 1. Rozkład Benforda dla cyfr od 1 do 9.



źródło: opracowanie własne

Podobną własność odkrył w 1881 roku Simon Newcomb. Dowód prawa Benforda został przeprowadzony dopiero w roku 1995 przez amerykańskiego matematyka Theodora Hilla z Uniwersytetu Kalifornijskiego w Berkeley.

Od lat siedemdziesiątych informatycy zaczęli się zastanawiać nad tym, jak można wykorzystać to prawo przy optymalizacji architektury komputerów. Ponieważ liczby wykorzystywane w obliczeniach nie są rozłożone jednolicie, ale zgodnie prawem Benforda, można by konstruować komputery ze zminimalizowanym miejscem do magazynowania danych. Jednak największe zastosowanie prawa znajdziemy w ekonomii. W latach sześćdziesiątych ubiegłego wieku zauważono, że wiele danych finansowych podlega temu prawu. W następnym dziesięcioleciu zasugerowano, że można tę wiedzę wykorzystać do wykrywania różnego rodzaju oszustw finansowych. Od lat dziewięćdziesiątych korzysta się z prawa Benforda przy wykrywaniu na przykład oszustw podatkowych. Sprawdza się przy tym, czy dane księgowo, które powinny podlegać rozkładowi Benforda, rzeczywiście mu podlegają. Z powodu nieznaności tego prawa „kreatywni” księgowi poprawiają te dane tak, że na przykład jedynka występuje na pierwszym miejscu za rzadko, a czwórka – za często.

Rysunek 3

FRANK BENFORD

The frequency of first digits thus follows closely the logarithmic relation

$$F_a = \log\left(\frac{a+1}{a}\right), \quad (1)$$

where F_a is the frequency of the digit a in the first place of used numbers.

TABLE II
OBSERVED AND COMPUTED FREQUENCIES

Natural Number	Number Interval	Observed Frequency	Logarithm Interval	Observed - Computed	Prob. Error of Mean
1	1 to 2	0.306	0.301	+0.005	±0.008
2	2 to 3	0.185	0.176	+0.009	±0.004
3	3 to 4	0.124	0.125	-0.001	±0.004
4	4 to 5	0.094	0.097	-0.003	±0.003
5	5 to 6	0.080	0.079	+0.001	±0.002
6	6 to 7	0.064	0.067	-0.003	±0.002
7	7 to 8	0.051	0.058	-0.007	±0.002
8	8 to 9	0.049	0.051	-0.002	±0.002
9	9 to 10	0.047	0.046	+0.001	±0.003

Simple Explanation of Benford's Law | Rob Gonsalves | Towards Data Science

Obrazy mogą być objęte prawami autorskimi. Więcej informacji

źródło: wyszukiwarka google.pl

Rysunek 2

TABLE I
PERCENTAGE OF TIMES THE NATURAL NUMBERS 1 TO 9 ARE USED AS FIRST DIGITS IN NUMBERS, AS DETERMINED BY 20,229 OBSERVATIONS

Group	Title	First Digit									Count
		1	2	3	4	5	6	7	8	9	
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Spec. Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	n^{-1}, \sqrt{n}, \dots	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	Digest	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray Volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Black Body	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
S	$n^4, n^2 \dots n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death Rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average		30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Error		±0.8	±0.4	±0.4	±0.3	±0.2	±0.2	±0.2	±0.2	±0.3	—

Benford's Law Explained with Examples - Statistics By Jim

Obrazy mogą być objęte prawami autorskimi. Więcej informacji

źródło: wyszukiwarka google.pl

Zastosowanie rozkładu Benforda

W naukach ekonomicznych do:

- wykrywania fałszywych lub błędnych w księgowości, rozpoznawania oszustw podatkowych,
- analizy danych z giełdy papierów wartościowych,
- analizy cen wylicytowanych na aukcjach internetowych,
- szacowania wysokości odszkodowań w firmach ubezpieczeniowych,
- oceny najbardziej dostosowanych wysokości kar finansowych orzekanych w procesach sądowych.

W naukach informatyczno-technicznych rozkład ten jest wykorzystywany do:

- projektowania architektury pamięci masowych,
- rozpoznawania autentyczności fotografii zapisanych cyfrowo,
- testowania rozmiarów plików komputerowych,
- analizy wydajności algorytmów numerycznych.

W zakresie nauk przyrodniczych do:

- oceny skuteczności działania leków,
- do badania autentyczności komunikatów o poziomach zanieczyszczeń,

W pozostałych naukach jest wykorzystywany do:

- badania wielkości dotacji dla rządzących partii,
- liczby odnośników do przeczytanych publikacji,
- liczby ofiar wypadków,
- liczby zakupionych produktów,
- do statystyk sportowych.

Oszustwo Jamesa Nelsona

Zastosowanie rozkładu Benforda pomogło w wykryciu fałszerstw dokonanych przez Jamesa Nelsona, głównego księgowego i zarządzającego Arizona State Treasurer. W 1992 roku w miasteczku Wayne (Arizona, USA) został uznany za winnego zdefraudowania 1 878 687,58 dolarów. Dokonał tego wystawiając 23 czeki. Oszustwo rozpoczęło się małą kwotą (najmniejszą w całej procedurze), a kolejne kwoty fałszywych czeków stopniowo rosły. Większość czeków wystawiono na kwotę poniżej 100 000 dolarów. Wykrycie przestępstwa umożliwił rozkład pierwszych cyfr poszczególnych kwot.

Wybory prezydenckie w Iranie

Boudewijn F.Roukema z toruńskiego Uniwersytetu Mikołaja Kopernika zaobserwował, iż rozkład pierwszych cyfr liczby głosów oddanych w wyborach w Iranie w 2009 roku w poszczególnych okręgach na kandydata małej partii opozycyjnej Mehdiego Karroubiego nie zgadza się z rozkładem Benforda - wśród pierwszych cyfr dwukrotnie częściej niż powinna pojawia się siódemka (poziom istotności 0,007). Anomalie występują w trzech z sześciu największych okręgów i są to właśnie okręgi w których aktualny prezydent Mahmud Ahmadineżad miał proporcjonalnie wyższe poparcie niż w reszcie kraju.

Prawo Benforda nie dotyczy:

- **istnienia maksymalnych i minimalnych** progów wartości, np. oceny z egzaminu, ilorazy inteligencji, wiek absolwentów szkoły średniej,
- **danych indentyfikacyjnych**, np. numerów rejestracyjne pojazdów, dowodów osobistych, telefonów, kodów pocztowych, kodów kreskowych, numerów rachunków bankowych, numerów rejestrów PESEL, NIP, REGON,
- **tych samych danych powtarzanych wielokrotnie**,
- **danych pochodzących z generatorów liczb losowych**, np. wyniki loterii,
- **wzrostu ludzi w centymetrach**,
- **ilorazu inteligencji**.

Oprócz prawa Benforda istnieje też analogiczne prawo, które dotyczy kolejnych cyfr znaczących. Wzór wygląda podobnie do poprzedniego:

$$P(a, b, c) = \log_{10} \frac{1}{a \times b \times c}$$

gdzie:

a – pierwsza cyfra w liczbie,

b – druga cyfra w liczbie,

c – trzecia cyfra w liczbie.

Można więc badać nie tylko rozkład pierwszej cyfry, ale też drugiej i trzeciej.

Praktyczne zastosowanie
prawa Benforda

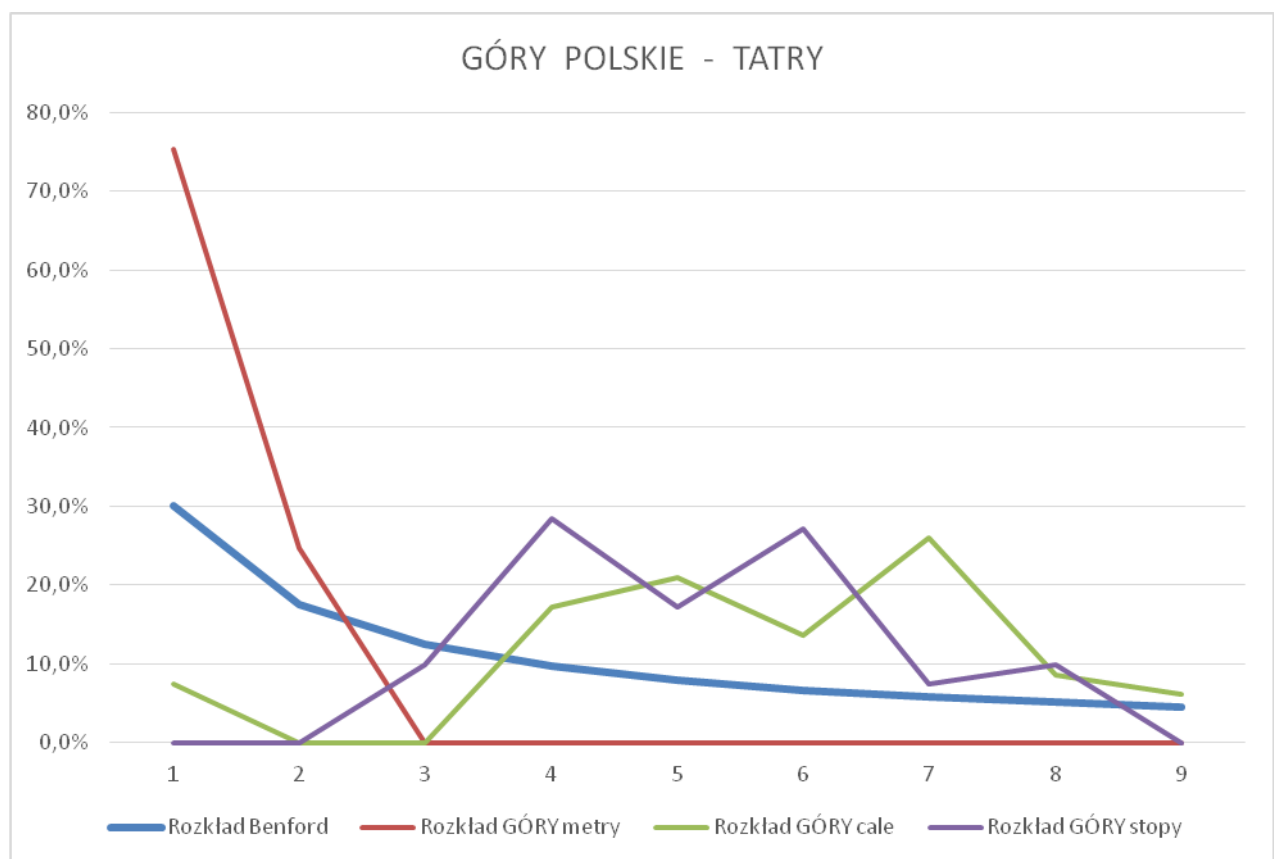
GÓRY POLSKIE - TATRY

W tabeli zostały zebrane zostały wysokości gór. Dane liczbowe zostały wpisane do Microsoft Excel, metry przeliczone zostały na cale i stopy. Wskazana została również pierwsza cyfra. Częstotliwość występowania cyfr w tabeli porównano z Benfordem, na podstawie danych sporządzono wykres.

Góra	metry	cale	stopy	pierwsza cyfra metry	pierwsza cyfra cale	pierwsza cyfra stopy
Hruby Regiel	1339	52716,5	4393,04	1	5	4
Pośredni Wierszyk	1262	49685,0	4140,42	1	4	4
Łysanki	1447	56968,5	4747,38	1	5	4
Jatki	1378	54252,0	4521,00	1	5	4
Mały Giewont	1728	68031,5	5669,29	1	6	5
Suchy Wierch	1539	60590,6	5049,21	1	6	5
Igła	1207	47519,7	3959,97	1	4	3
Kazalnica	1034	40708,7	3392,39	1	4	3
Żar Smerczyński	1566	61653,5	5137,80	1	6	5
Wysoka Turnia	1643	64685,0	5390,42	1	6	5
Krokiew	1618	63700,8	5308,40	1	6	5
Kopa Kalacka	1380	54330,7	4527,56	1	5	4
Kalacka Turnia	1383	54448,8	4537,40	1	5	4
Jaworowa Kopa	1791	70511,8	5875,98	1	7	5
Suchy Wierch Kondracki	1890	74409,4	6200,79	1	7	6
Goryczkowa Czuba	1913	75315,0	6276,25	1	7	6
Pośredni Wierch Goryczkowy	1874	73779,5	6148,29	1	7	6
Salatin	2048	80629,9	6719,16	2	8	6
Brestowa	1934	76141,7	6345,14	1	7	6
Zuberec	1753	69015,7	5751,31	1	6	5
Wierch Kuca	1305	51378,0	4281,50	1	5	4
Wierch Spalenisko	1324	52126,0	4343,83	1	5	4
Diabliniec	1241	48858,3	4071,52	1	4	4
Zadnia Rococho	1721	67755,9	5646,33	1	6	5
Skrajna Rococho	1262	49685,0	4140,42	1	4	4
Cisowa Turnia	1112	43779,5	3648,29	1	4	3
Siwiańskie Turnie	1079	42480,3	3540,03	1	4	3
Zadnia Kopa	1333	52480,3	4373,36	1	5	4
Pośrednia Kopa	1305	51378,0	4281,50	1	5	4
Przednia Kopa	1113	43818,9	3651,57	1	4	3
Ciemniak	2096	82519,7	6876,64	2	8	6
Suchy Tomanowy Wierch	1860	73228,3	6102,36	1	7	6
Tomanowy Wierch Polski	1977	77834,6	6486,22	1	7	6
Tomanowa Kopa	1892	74488,2	6207,35	1	7	6
Chuda Turnia	1858	73149,6	6095,80	1	7	6
Uptązińska Kopa	1794	70629,9	5885,83	1	7	5
Uptązińska Kopka	1457	57362,2	4780,18	1	5	4
Uptąziński Wierszyk	1203	47362,2	3946,85	1	4	3

Czerwony Gronik	1266	49842,5	4153,54	1	4	4
Jaworzynka Miętusia	1294	50944,9	4245,41	1	5	4
Mały Regiel	1142	44960,6	3746,72	1	4	3
Biełańska Kopa	1835	72244,1	6020,34	1	7	6
Belasa	2284	89921,3	7493,44	2	8	7
Kołowy Szczyt	2418	95196,9	7933,07	2	9	7
Baranie Rogi	2526	99448,8	8287,40	2	9	8
Kierzmarski Szczyt	2556	100629,9	8385,83	2	1	8
Łomnica	2634	103700,8	8641,73	2	1	8
Wielka Łomnicka Baszta	2215	87204,7	7267,06	2	8	7
Lodowy Szczyt	2627	103425,2	8618,77	2	1	8
Jaworowy Szczyt	2641	103976,4	8664,70	2	1	8
Ostry Szczyt	2367	93189,0	7765,75	2	9	7
Litworowy Szczyt	2424	95433,1	7952,76	2	9	7
Wysoka	2559	100748,0	8395,67	2	1	8
Gerlach	2655	104527,6	8710,63	2	1	8
Rysy	2499	98385,8	8198,82	2	9	8
Bystra	2248	88503,9	7375,33	2	8	7
Kopa Kondracka	2004	78897,6	6574,80	2	7	6
Małolącziak	2069	81456,7	6788,06	2	8	6
Krzesanica	2122	83543,3	6961,94	2	8	6
Giewont	1894	74566,9	6213,91	1	7	6
Rakoń	1879	73976,4	6164,70	1	7	6
Grześ	1655	65157,5	5429,79	1	6	5
Bobrowiec	1663	65472,4	5456,04	1	6	5
Furkaska	1491	58700,8	4891,73	1	5	4
Jamborowy Wierch	1565	61614,2	5134,51	1	6	5
Małą Furkaska	1133	44606,3	3717,19	1	4	3
Ornak	1854	72992,1	6082,68	1	7	6
Trzydniowiański Wierch	1758	69212,6	5767,72	1	6	5
Kończysty Wierch	2002	78818,9	6568,24	2	7	6
Siwy Zwornik	1959	77126,0	6427,17	1	7	6
Kotłowa Czuba	1840	72440,9	6036,75	1	7	6
Zadni Ornak	1867	73503,9	6125,33	1	7	6
Suchy Wierch Ornaczański	1818	71574,8	5964,57	1	7	5
Kominiarski Wierch	1829	72007,9	6000,66	1	7	6
Mały Kopieniec	1328	52283,5	4356,96	1	5	4
Wielki Kopieniec	1257	49488,2	4124,02	1	4	4
Wielki Opalony Wierch	1485	58464,6	4872,05	1	5	4
Mały Opalony Wierch	1448	57007,9	4750,66	1	5	4
Spalona Czuba	1257	49488,2	4124,02	1	4	4
Stoły	1428	56220,5	4685,04	1	5	4
Kliniowa Czuba	1276	50236,2	4186,35	1	5	4

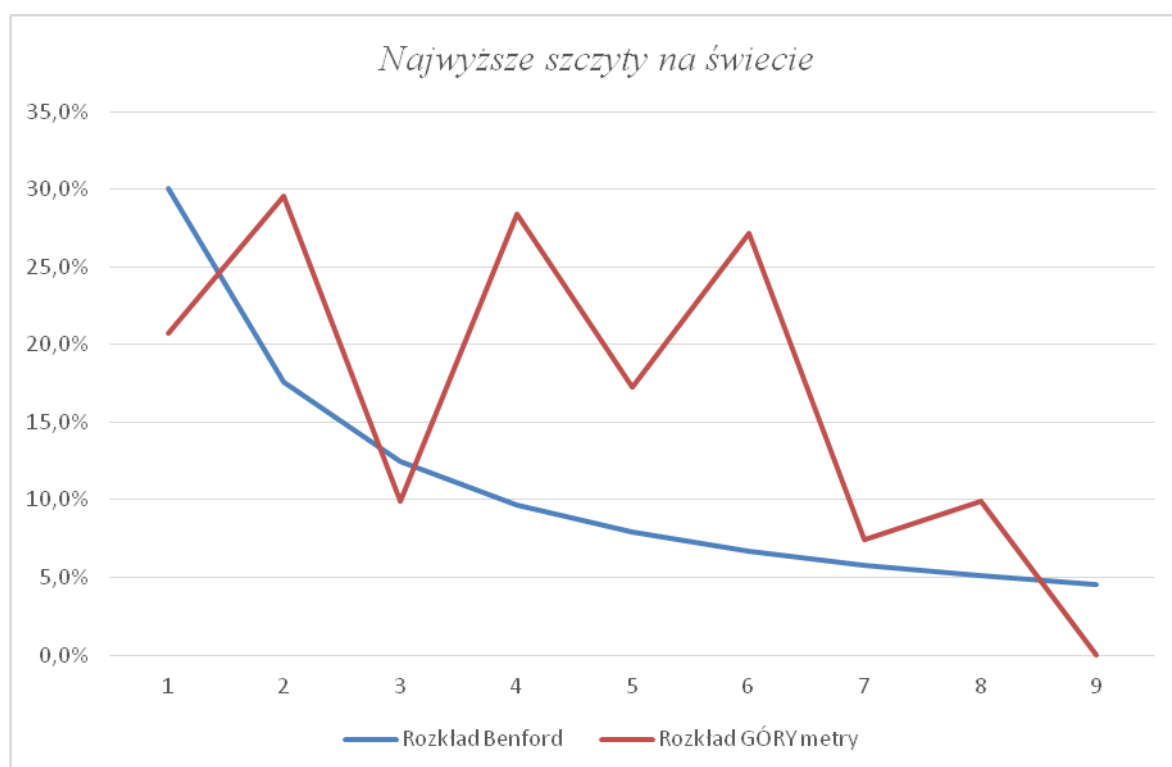
Jednostki miary	Cyfra znacząca:	1	2	3	4	5	6	7	8	9
METRY	Ilość cyfr w zbiorze	61	20	0	0	0	0	0	0	0
	W ilu % występuje dana cyfra	75,31%	24,69%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
CALE	Ilość cyfr w zbiorze	6	0	0	14	17	11	21	7	5
	W ilu % występuje dana cyfra	7,41%	0,00%	0,00%	17,28%	20,99%	13,58%	25,93%	8,64%	6,17%
STOPY	Ilość cyfr w zbiorze	0	0	8	23	14	22	6	8	0
	W ilu % występuje dana cyfra	0,00%	0,00%	9,88%	28,40%	17,28%	27,16%	7,41%	9,88%	0,00%
Rozkład Benforda		30,10%	17,61%	12,49%	9,69%	7,92%	6,69%	5,80%	5,12%	4,58%



Z wykresu wynika, że pierwsze cyfry znaczące nie są zgodne z rozkładem Benforda. Każda z linii reprezentujących góry podana jest w innej jednostce.

Do kolejnego sprawdzenia prawdziwości prawa Benforda wypisałem najwyższe szczyty gór z państw świata (193 państwa, dane pochodziły z portalu „Wikipedia”). Tak jak poprzednio sporządziłem najpierw tabelę, a później wykres posługując się Microsoft Excel.

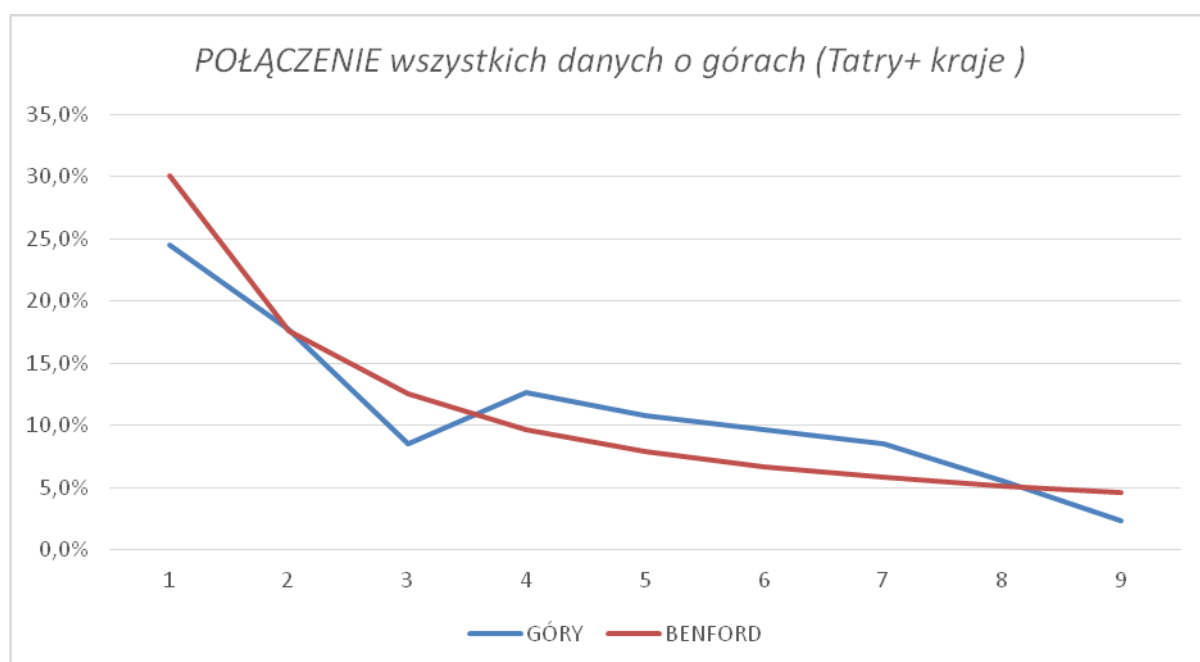
	1	2	3	4	5	6	7	8	9
Ilość cyfr w zbiorze	40	57	29	18	16	9	10	9	5
W ilu % występuje dana cyfra	20,73%	29,53%	9,88%	28,40%	17,28%	27,16%	7,41%	9,88%	0,00%
Rozkład Benforda	30,10%	17,61%	12,49%	9,69%	7,92%	6,69%	5,80%	5,12%	4,58%



Tutaj tylko cyfra 5 i 7 jest najbliższa wartościom odpowiadającym rozkładowi Benforda.

Po połączeniu wartości odpowiadającym wysokościami Tatr i gór świata otrzymałem tabelę i sporządziłem wykres.

POŁĄCZENIE wszystkich danych o górach (metry + cale + stopy + kraje)									
	1	2	3	4	5	6	7	8	9
	107	77	37	55	47	42	37	24	10
GÓRY	24,5%	17,7%	8,5%	12,6%	10,8%	9,6%	8,5%	5,5%	2,3%
BENFORD	30,1%	17,6%	12,5%	9,7%	7,9%	6,7%	5,8%	5,1%	4,6%



Po połączeniu wszystkich wartości liczbowych otrzymałem wykres który można stwierdzić, że jest zgodny z prawem Benforda. Różnice wartości procentowych dla poszczególnych cyfr znaczących nie są tak bardzo duże.

KOSTKA RUBIKA

Ernő Rubik to węgierski architekt i rzeźbiarz który w 1974 wynalazł zabawkę logiczną nazwaną od jego nazwiska kostka Rubika (węg.bűvös kocka, magiczna kostka). W 1976 roku skonstruowana i opatentowana w Japonii przez inżyniera Terutosiego Ishige. Wynalazca kostki Ernő Rubik, układał kostkę po raz pierwszy przez miesiąc.

Zabawa kostką polega na takim ułożeniu kwadratów, aby na każdej ścianie wszystkie posiadały jeden kolor. Składa się ona z 26 sześciątów i przegubu umieszczonego w środku. Przegub ten umożliwia każdej z zewnętrznych warstw kostki obrót wokół osi prostopadłej do danej warstwy i przechodzącej przez środek kostki. Światowy Dzień Kostki Rubika obchodzony jest 17 kwietnia.

Kostka Rubika



Rodzaje kostek:

Kostki sześciennie:	Kostki o innych kształtach:
<ul style="list-style-type: none"> • 2×2×2 (Rubik's Mini Cube) • 3×3×3 (oryginalna Rubik's Cube) • 4×4×4 (Rubik's Revenge) • 5×5×5 (Rubik's Professor Cube) • 6×6×6 • 7×7×7 • 8×8×8 • 9×9×9 • 10×10×10 • 11×11×11 • 13×13×13 • 17×17×17 • Square-1 • Fisher Cube (łamiągówka skonstruowana przez <u>Tony'ego Fishera</u>) • Void Cube (kostka 3×3×3 bez elementów centrowych) • Mirror Cube (kostka, w której każdy element posiada inny kształt) • <u>Skewb</u> (modyfikacja kostki <u>pyraminx</u> do formatu sześciennego) • Gear Cube (kostka posiadająca zębatki) 	<ul style="list-style-type: none"> • Rubik's Twist – układanka w kształcie węża, z której można ułożyć różne figury 3D i 2D • Rubik's Magic – łamiągówka w kształcie prostokątnej płytki • Clock – łamiągówka z zegarami • Floppy Cube – kostka 1×3×3. • 1×2×2 • 1×2×3 • 1×3×3 • 2×2×4 • 2×3×3 • 2×3×4 • 3×3×4 • 3×3×5 • 3×3×6 • 3×3×7 • 3×3×8 • 3×3×9 • 3×4×5 • 4×4×2 • 4×4×5 • 5×5×4 • Pentahedron • Dino's cube (kostka z ruchomymi centrami) • Tower cube 2×2×3

Rodzaje kostek:

Kostki dwunastościenne:	Kostki ośmiościenne:	Kostki czworościenne:
<ul style="list-style-type: none"> • Kilominx – najmniejsza z dwunastościanów • <u>Megaminx</u> • Starminx • Gigaminx • Teraminx • Petaminx – największa z powszechnie dostępnych dwunastościanów, składa się z 218 elementów, a całkowita liczba kombinacji ułożeń to około $1,95 \times 10^{160}$, czyli 19,5 <u>sekwicyliardów</u> 	<ul style="list-style-type: none"> • Diamond • Octahedron • Rubik's UFO 	<ul style="list-style-type: none"> • Pyraminx – łamigłówka w kształcie czworościanu • Tetraminx – kostka w kształcie czworościanu bez wierzchołków

Kilka ciekawostek ze świata Kostki Rubika:

- Rubik stworzył swoją łamigłówkę, aby wyjaśnić studentom zasady przestrzenności i trójwymiarowości.
- Inspiracją dla tego pomysłu miały być gładkie kamienie znad Dunaju. Rubik twierdził, że to one sprawiły, że pomyślał o cylindrycznych kształtach i sześciacie.
- **Najmniejsza Kostka Rubika na świecie** ma niecałe 6 mm wysokości. **Największa** ma bok o długości 1.56m i waży ponad 100 kg, dzięki czemu jest o 20 000 razy większa od wersji zaprojektowanej przez Węgra.
- Gdyby zebrać tyle Kostek Rubika, by każda z nich miała inny układ swoich 26 klocków i ustawić je jedna na drugiej, taka wieża miałaby wysokość... ponad 246 lat świetlnych. To 60 razy dalej niż do Proximy Centauri,
- Gdyby każdemu człowiekowi na Ziemi, niezależnie od rasy, wyznania, wieku czy płci, dać dzisiaj do ręki kostkę i każdy z tych 6 miliardów ludzi wykonywałby przypadkowe obroty kostką w średnim tempie jednego obrotu na sekundę, to średnio co 229 lat jedna kostka byłaby ułożona.
- Liczba kombinacji różnych ułożeń kostki 3x3x3 to ponad 43 tryliony.

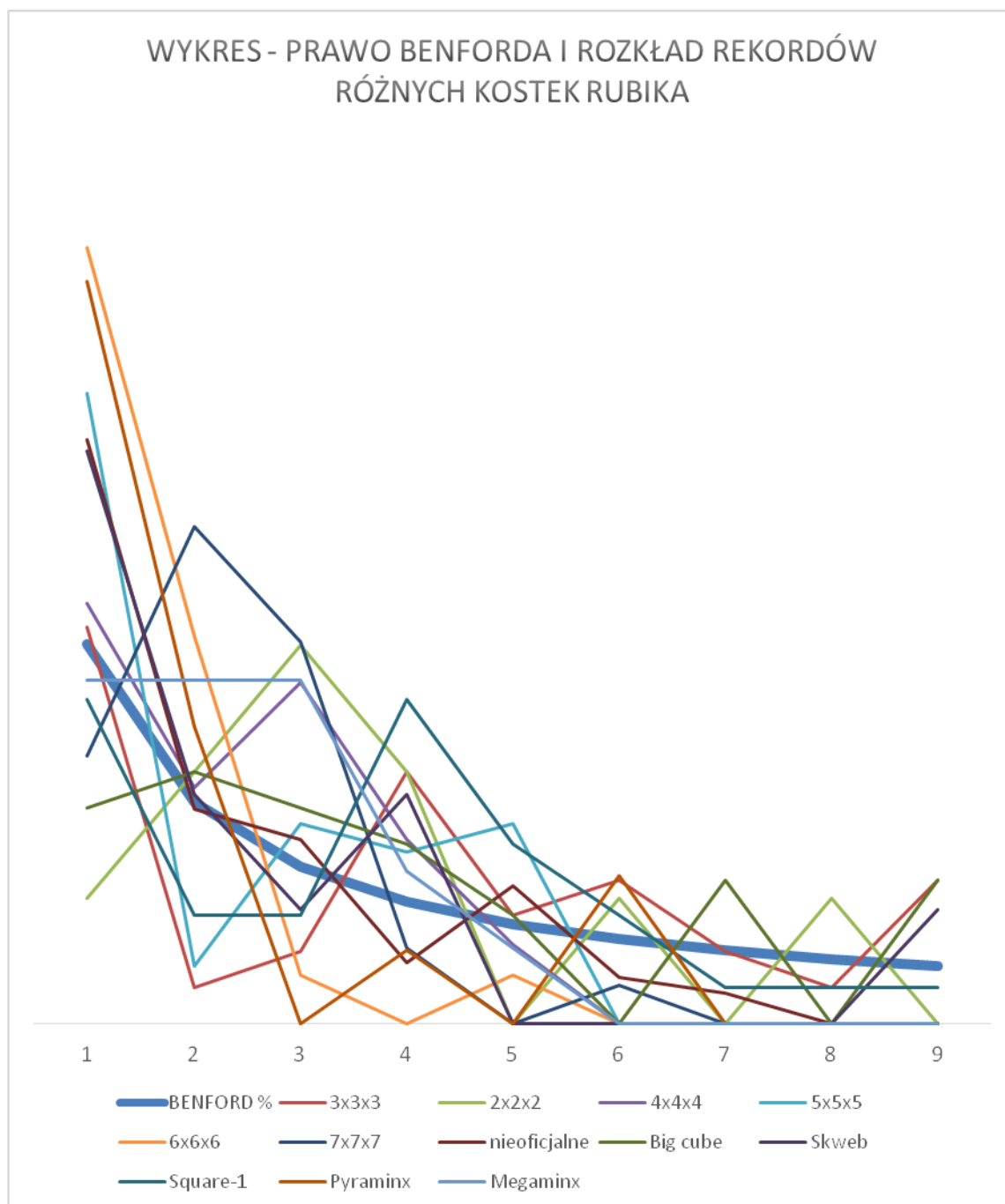
Kostka Rubika podobnie jak i góry stały się kolejnym zadaniem jakim się zająłem. Zebrałem rekordy tradycyjnej kostki 3 x 3 x 3 ale również i innych. Sposób opracowania wyników był taki sam jak i przy górach.

Tabele z wynikami kostek.

	Cyfra znacząca	1	2	3	4	5	6	7	8	9
	BENFORD %	0,30	0,18	0,12	0,10	0,08	0,07	0,06	0,05	0,05
KOSTKA 3x3x3	Ilość cyfr w zbiorze	11	1	2	7	3	4	2	1	4
	W ilu % występuje ta cyfra	0,31	0,03	0,06	0,20	0,09	0,11	0,06	0,03	0,11
KOSTKA 2x2x2	Ilość cyfr w zbiorze	1	2	3	2	0	1	0	1	0
	W ilu % występuje ta cyfra	0,10	0,20	0,30	0,20	0,00	0,10	0,00	0,10	0,00
KOSTKA 4x4x4	Ilość cyfr w zbiorze	16	9	13	7	3	0	0	0	0
	W ilu % występuje ta cyfra	0,33	0,19	0,27	0,15	0,06	0,00	0,00	0,00	0,00
KOSTKA 5x5x5	Ilość cyfr w zbiorze	22	2	7	6	7	0	0	0	0
	W ilu % występuje ta cyfra	0,50	0,05	0,16	0,14	0,16	0	0	0	0
KOSTKA 6x6x6	Ilość cyfr w zbiorze	16	8	1	0	1	0	0	0	0
	W ilu % występuje ta cyfra	0,62	0,31	0,04	0,00	0,04	0	0	0	0
KOSTKA 7x7x7	Ilość cyfr w zbiorze	7	13	10	2	0	1	0	0	0
	W ilu % występuje ta cyfra	0,21	0,39	0,30	0,06	0,00	0,03	0	0	0
REKORDY nieoficjalne	Ilość cyfr w zbiorze	38	14	12	4	9	3	2	0	0
	W ilu % występuje ta cyfra	0,46	0,17	0,15	0,05	0,11	0,04	0,02	0	0
KOSTKA Big cube	Ilość cyfr w zbiorze	6	7	6	5	3	0	4	0	4
	W ilu % występuje ta cyfra	0,17	0,20	0,17	0,14	0,09	0,00	0,11	0,00	0,11
KOSTKA Skweb	Ilość cyfr w zbiorze	5	2	1	2	0	0	0	0	1
	W ilu % występuje ta cyfra	0,45	0,18	0,09	0,18	0,00	0,00	0,00	0,00	0,09
KOSTKA Square-1	Ilość cyfr w zbiorze	9	3	3	9	5	3	1	1	1
	W ilu % występuje ta cyfra	0,26	0,09	0,09	0,26	0,14	0,09	0,03	0,03	0,03
KOSTKA Pyraminx	Ilość cyfr w zbiorze	10	4	0	1	0	2	0	0	0
	W ilu % występuje ta cyfra	0,59	0,24	0,00	0,06	0,00	0,12	0	0	0

KOSTKA Megaminx	Ilość cyfr w zbiorze	9	9	9	4	2	0	0	0	0
	W ilu % występuje ta cyfra	0,27	0,27	0,27	0,12	0,06	0	0	0	0

Na podstawie tabeli otrzymałem wykres.



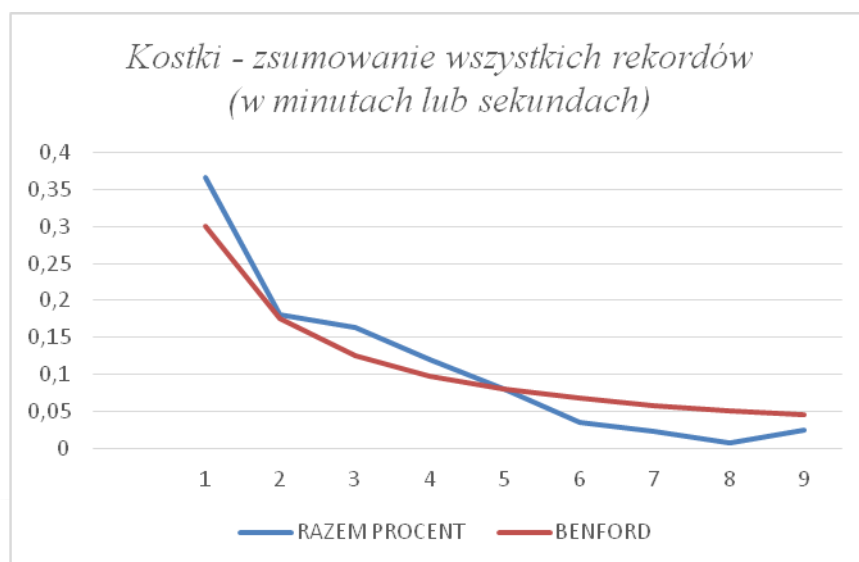
Spróbowałem zsumować wszystkie rekordy a ich czas przeliczyć na przeliczyć sekundy i milisekundy.

Przykład przeliczenia czasów.

	A	B	C	D	E	F
1	22,95	22950				
2	16,71	16710				
3	16,53	16530				
4	15,07	15070				
5	14,76	14760				
6	13,93	13930				
7	13,89	13890				
8	12,11	12110				
9	11,75	11750				
10	11,13	11130				
11	10,48	10480				
12	10,36	10360				
13	9,86	9860				
14	9,77	9770				
15	9,55	9550				
16	9,18	9180				
17	8,72	8720				
18	7,08	7080				
19	7,03	7030				
20	6,77	6770				
21	6,65	6650				
22	6,24	6240				
23	6,18	6180				
24	5,66	5660				
25	5,55	5550				
26	5,25	5250				
27	4,9	4900				
28	4,74	4740				

Po zsumowaniu wszystkich rekordów bez zamiany na sekundy (czyli czasami są podane minuty, a czasami sekundy) wyniki wyszły następujące:

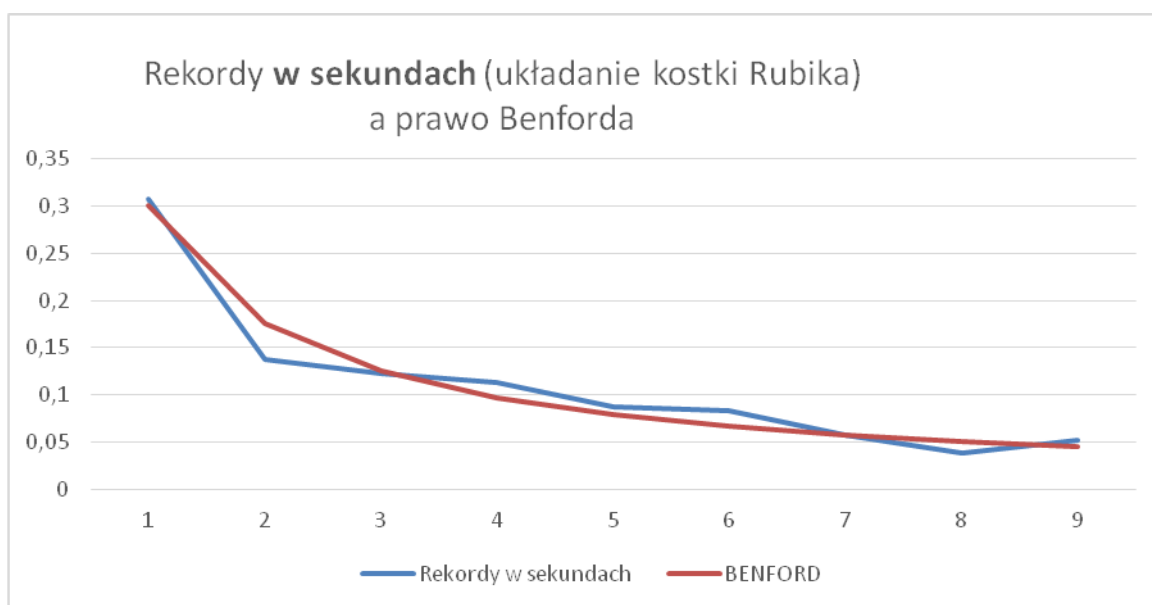
	1	2	3	4	5	6	7	8	9
Ile razy występuje cyfra.	150	74	67	49	33	14	9	3	10
W ilu % występuje cyfra	36,7%	18,1%	16,4%	12,0%	8,1%	3,4%	2,2%	0,7%	2,4%
Rozkład Benforda	30,1%	17,6%	12,5%	9,7%	7,9%	6,7%	5,8%	5,1%	4,6%



TEST - PRAWO BENFORDA SEKUNDY

Rekordy kostki w przeliczeniu na sekundy.

	1	2	3	4	5	6	7	8	9
Ile razy występuje cyfra.	130	58	52	48	37	35	24	16	22
W ilu % występuje cyfra	0,31	0,14	0,12	0,11	0,09	0,08	0,06	0,04	0,05
Rozkład Benforda	0,30	0,18	0,12	0,10	0,08	0,07	0,06	0,05	0,05

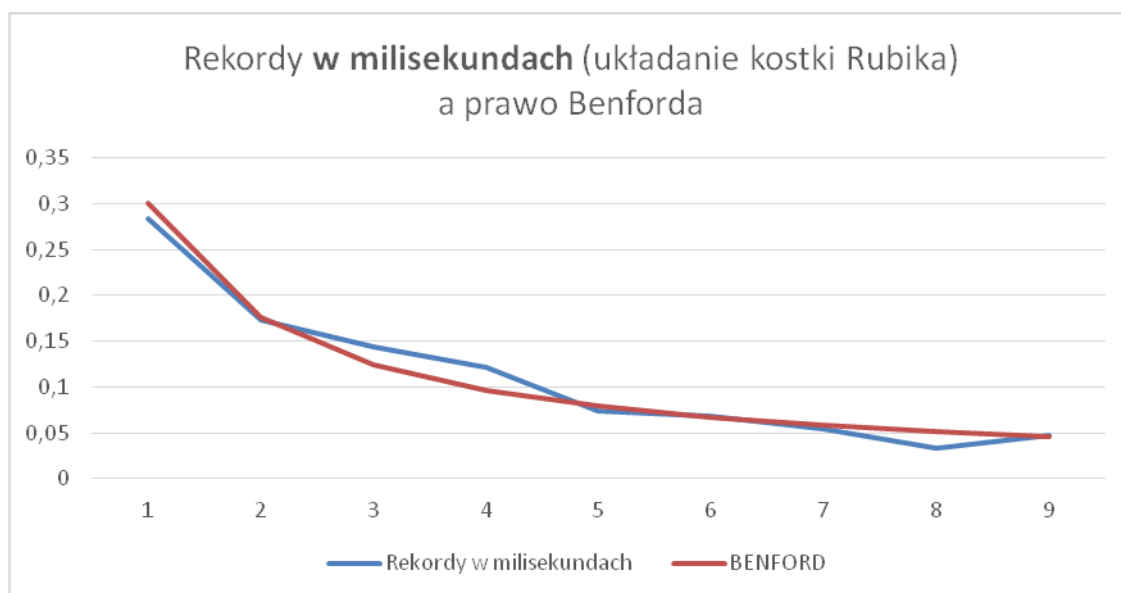


Na wykresie widać, że jest duża zgodność wyników.

TEST - PRAWO BENFORDA MILISEKUNDY

Rekordy kostki w przeliczeniu na milisekundy. Jedna sekunda to 1000 milisekund.

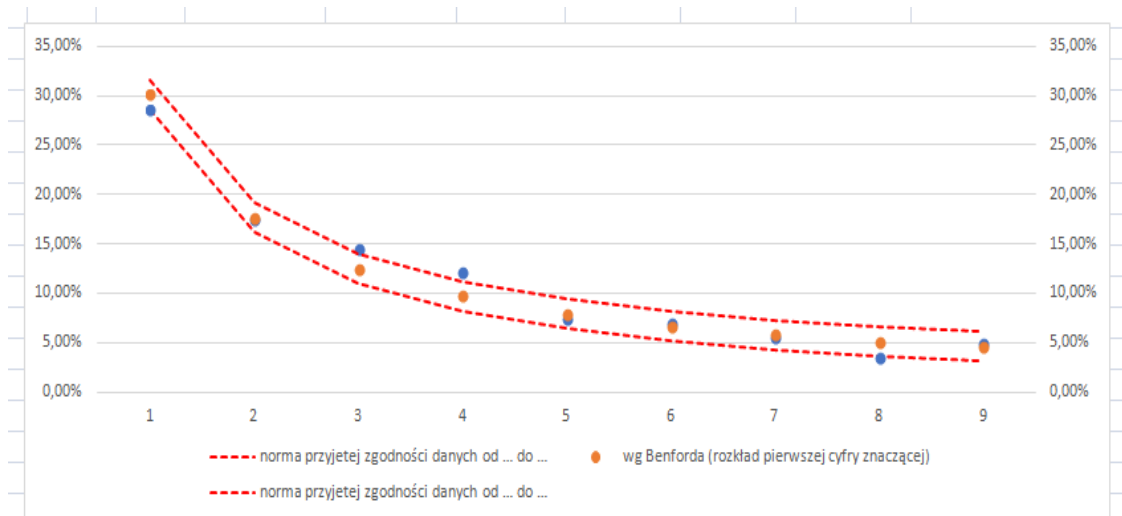
	1	2	3	4	5	6	7	8	9
Ile razy występuje cyfra.	120	73	61	51	31	29	23	14	20
W ilu % występuje cyfra	0,28	0,17	0,14	0,12	0,07	0,07	0,05	0,03	0,05
Rozkład Benforda	0,30	0,18	0,12	0,10	0,08	0,07	0,06	0,05	0,05



Otrzymana linia jest prawie taka jak linia przedstawiająca rozkład Benforda dla pierwszej cyfry znaczącej.

Dane wprowadziłem do pliku excel pobranego z internetu, w którym można również wyliczyć zgodność z prawem Benforda oraz otrzymać inny wykres. Wyniki były następujące:

TEST BENFORDA													
Rodzaj analizowanych danych:	#ADR!												
#ADR!													
Pierwsza cyfra	Ile cyfr w zbiorowości	procentowe występowanie w zestawieniu (rozkład pierwszej cyfry znaczącej)	wg Benforda (rozkład pierwszej cyfry znaczącej)	Różnica	średnie bezwzględne odchylenie (%)	norma przyjętej zgodności danych od ... do ...	norma przyjętej zgodności danych od ... do ...	norma przyjętej zgodności danych (*)	norma przyjętej zgodności danych w zależności od średniego bezwzględnego odchylenia (*)				
1	120	28,44%	30,10%	-1,67%	1,67%	31,60%	28,60%	brak zgodności	od	do			
2	73	17,30%	17,61%	-0,31%	0,31%	19,11%	16,11%	duża zgodność	0,0%	0,6%	duża zgodność		
3	61	14,45%	12,49%	1,96%	1,96%	13,99%	10,99%	brak zgodności	0,6%	0,1%	akceptowalna zgodność		
4	51	12,09%	9,69%	2,39%	2,39%	11,19%	8,19%	brak zgodności	1,2%	1,5%	minimalnie akceptowalna zgodność		
5	31	7,35%	7,92%	-0,57%	0,57%	9,42%	6,42%	duża zgodność	1,5%		brak zgodności		
6	29	6,87%	6,69%	0,18%	0,18%	8,19%	5,19%	duża zgodność	(*) Na podstawie: Nigrini, 2012 s. 160 "Benford's Law: Application for Forensic Accounting, Auditing and Fraud Detection".				
7	23	5,45%	5,80%	-0,35%	0,35%	7,30%	4,30%	duża zgodność					
8	14	3,32%	5,12%	-1,80%	1,80%	6,62%	3,62%	brak zgodności					
9	20	4,74%	4,58%	0,16%	0,16%	6,08%	3,08%	duża zgodność					
	422	100%	100%	0%									

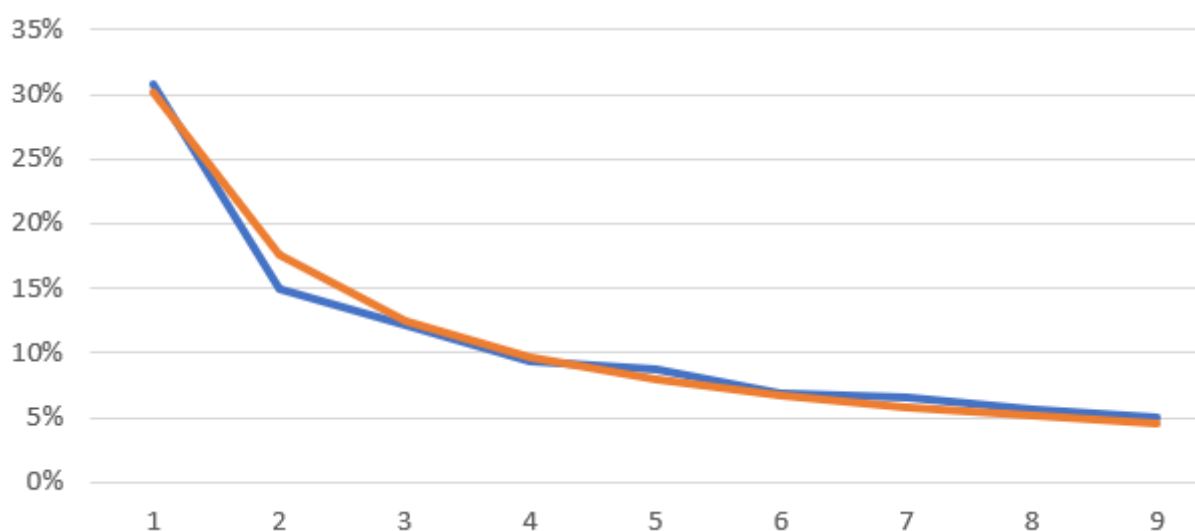


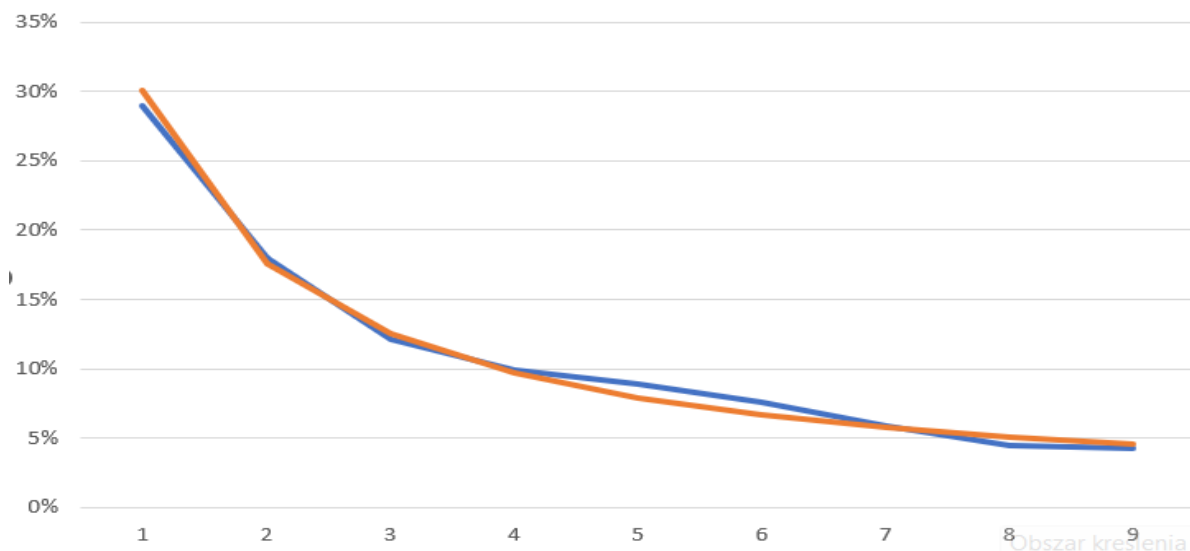
Stosowanie prawa Benforda w wykrywaniu nieprawidłowości w danych finansowych.

W internecie można przeczytać, że Prawo Benforda jest wykorzystywane w wykrywaniu nieprawidłowości w danych księgowych. Na przykład urzędy skarbowe korzystają z oprogramowania, które testuje wysyłane przez podatników deklaracje podatkowe i inne pliki i sprawdza, czy spełnione jest prawo Benforda. Jeżeli nie jest, nie oznacza to od razu, że dokonano jakiegoś oszustwa, jednak jest jakby „czerwoną lampką” i mówi, że tym dokumentom należy się bliżej przyjrzeć.

Również biegli rewidenci badając sprawozdania finansowe wykorzystują Prawo Benforda. Poprosiłem biegłego rewidenta o informacje, czy to prawo działa i sprawdza się w ich pracy.

Biegli rewidenci mają dostęp do nawet tajnych danych, i są często bardzo duże zbiory liczb (np. kwoty z kilkunastu tysięcy faktur). Biegły rewident opowiedział mi, że na tak dużych zbiorach danych prawo Benforda sprawdza się bardzo dobrze. Pozwolono mi zamieścić w pracy dwa zdjęcia pokazujące jak wygląda rozkład Benforda sprawdzający prawdziwe dane finansowe przedsiębiorstwa. Na obu zdjęciach pomarańczowa linia oznacza rozkład Benforda, a linia niebieska to rozkład ze sprawdzanych dokumentów.





Niestety, rozkłady danych z wysokości gór polskich lub z pojedynczych kategorii układanie kostek Rubika nie układały się w tak ładny wzór.

Gdy połączyłem te dane w duże grupy (czyli dane z wysokościami gór z Polski i ze świata, oraz połączyłem wszystkie rekordy z układania różnych rodzajów kostek Rubika) dopiero wtedy wykresy tych danych stały się podobne do kreski z prawem Benforda.

Dane księgowo, których jest bardzo dużo i są to też bardzo różne liczby (od kilku złotych do milionów złotych) tworzą wykres który najbardziej pokrywa się z wykresem z Prawem Benforda. Być może dlatego, że tych liczb jest bardzo dużo i nie są one ograniczone.

Wysokości gór nigdy nie przekroczą wielkości 8848 metrów (Mount Everest), a w przypadku gór polskich 2499 metrów (Rysy), dlatego też wykres z górami najmniej odpowiadał prawu Benforda.

Bibliografia

1. Anna Staszal, *Wykrywanie przestępstw gospodarczych z wykorzystaniem prawa Benforda*.
2. Prawo Benforda – Benford.pl
3. Jack Murtagh, Prawo Benforda. Fenomen, który rządzi liczbami wokół nas, Świat Nauki 1.2024
4. Frank Benford, Państwa – Góry, *Wikipedia*
5. Marzena Farbaniec, Tadeusz Grabiński, Bartłomiej Zabłocki, Waław Zając, *Wykorzystanie prawa Benforda w analizie poprawności danych finansowych na przykładzie informacji o obrocie towarowym*.
6. Marzena Farbaniec, Tadeusz Grabiński, Marta Woźniak-Zapór, Waław Zając, *Wykorzystanie prawa Benforda w analizie poprawności danych finansowych na przykładzie informacji o obrocie towarowym*, Krakow 2016
7. Artur Sierant, *Wrózenie z poóólklych logarytmów*, styczeń 2017
8. Mikołaj Karpiński, *Prawo Benforda*, Matematyka w szkole nr 31, styczeń/luty 2008
9. Joanna Krawiec, *Prawo Benforda jako narzędzie wykrywania manipulacji finansowych*, 2012 r.
10. Metody wykrywania oszustw finansowych oparte na prawie Benforda, Mateusz Baryła, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, 2021 r.
11. Poznański Portal Matematyczny, Marek Kaluba, *Prawo Benforda czyli o niezmienniczości rozkładów*.
12. Kostka Rubika.net
13. Kostka Rubika. Wikipedia
14. https://pl.wikipedia.org/wiki/Rekord_%C5%9Bwiata_w_pojedynczym_u%C5%82o%C5%BCeniu_kostki_Rubika
15. <https://www.worldcubeassociation.org/results/records>
16. <https://www.speedcubing.pl/rekordy-oficjalne>