Die Gesetze der Mittelwerte: Teil 3, Der durchschnittliche Durchschnitt

geschrieben von Andreas Demmig | 31. August 2017

Durchschnittliches

Sowohl das Wort als auch das Konzept „Durchschnitt“ sind in der breiten Öffentlichkeit sehr viel Verwirrung und Missverständnis unterworfen und sowohl als Wort als auch im Kontext ist eine überwältigende Menge an „lockerem Gebrauch“ auch in wissenschaftlichen Kreisen zu finden, ohne die Peer-Reviewed-Artikel in Zeitschriften und wissenschaftlichen Pressemitteilungen auszuschließen.

In Teil 1 dieser Serie lasen Sie meine Auffrischung über die Begriffsinhalte von Durchschnitten, Mittelwerten und Beispiele dazu. Entspricht Ihr Hintergrundwissen bezüglich Mathematik oder Wissenschaft dem großen Durchschnitt der Normalbürger, schlage ich vor, dass Sie einen schnellen Blick auf die Grundlagen in Teil 1 [Teil 1 hier übersetzt] und dann Teil 2 [Teil 2 hier übersetzt] werfen, bevor Sie fortfahren.

Warum ist es eine mathematische Sünde, eine Reihe von Durchschnittswerten zu mitteln?

„Der Umgang mit Daten kann manchmal Verwirrung verursachen. Ein oft gemachter Fehler ist es, Durchschnittswerte zu mitteln. Dies ist oft zu sehen, wenn versucht wird, regionale Werte aus Landes- oder Kreisdaten abzuleiten“. – Was man mit Daten nicht machen darf: Durchschnittswerte zu mitteln.

„Ein Kunde hat mich heute gebeten, einen“ Durchschnitt der Durchschnittswerte „zu einigen seiner Leistungsberichten hinzuzufügen. Ich gebe offen zu, dass ein nervöses und hörbares Stöhnen meinen Lippen entrang, als ich mich in Gefahr fühlte, hilflos in die fünfte Dimension von „Simpsons Paradox“ zu stürzen – Anders ausgedrückt: Kann die Mittelung der Mittelwerte der verschiedenen Populationen den Durchschnitt der zusammengefassten Bevölkerung darstellen?“- Ist der Durchschnitt der Durchschnittswerte korrekt? (Hinweis: NEIN!)

„**Simpsons Paradoxon …** ist ein Phänomen der Wahrscheinlichkeiten und Statistik, in dem ein Trend in verschiedenen Datengruppen zu erkennen ist, aber verschwindet oder sich umkehrt, wenn diese Gruppen kombiniert werden. Es wird manchmal auch als **Umkehrparadox** oder **Verschmelzungsparadox** angegeben. „- siehe Wiki.de “Simpsons Paradox“

Durchschnittliche Mittelwerte sind nur gültig, wenn die Sätze von Datengruppen, Kohorten, Anzahl der Messungen – alle gleich groß sind (oder fast gleich groß) und die gleiche Anzahl von Elementen enthalten, denselben Bereich repräsentieren, die gleiche Lautstärke, die gleiche Anzahl von Patienten, die gleiche Anzahl von Meinungen und, wie bei allen Durchschnittswerten, sind die Daten selbst physisch und logisch homogen (nicht heterogen) und physisch und logisch kommensurabel (nicht inkommensurabel). [wem dies unklar ist, bitte schauen Sie noch mal Teil 1]

Zum Beispiel, hat man vier Klassen der 6. Klasse, zu denen jeweils genau 30 Schüler gehören und wollte nun die durchschnittliche Größe der Schüler finden, könnte man über es zwei Wege berechnen: 1) Durchschnitt jeder Klasse durch die Summierung der Größen der Schüler Dann finden Sie den Durchschnitt durch Division mit 30, dann summieren Sie die Mittelwerte und teilen sie durch vier, um den Gesamtdurchschnitt zu bekommen – ein Durchschnitt der Mittelwerte oder 2), Sie kombinieren alle vier Klassen zusammen für einem Satz von 120 Studenten, Summieren die Größen und teilen durch 120. Die Ergebnisse sind gleich.

Das gegenteilige Beispiel sind vier Klassen von 6. Klasse Schülern, die alle unterschiedlich groß sind und jeweils unterschiedliche Klassenstärken haben: 30, 40, 20, 60 Schüler. Die Suche nach den Mittelwerten der vier Klassen- und dann Mittelung der Mittelwerte ergibt eine Antwort – ganz anders als die Antwort, als wenn man die Größe von allen 150 Schülern summiert und dann durch 150 teilt.

Warum? Das liegt daran, dass die einzelnen Schüler der Klasse mit nur 20 Schülern und die einzelnen Schüler der Klasse von 60 Schülern unterschiedliche, ungleiche Auswirkungen auf den Gesamtdurchschnitt haben. Um für den Durchschnitt gültig zu sein, sollte jeder Schüler den 0,66ten Anteil des Gesamtdurchschnitts repräsentieren [0,66 = 1/150].

Wenn nach Klasse gemittelt wird, steht jede Klasse für 25% des Gesamtdurchschnitts. So würde *jeder Schüler* in der Klasse von 20 Schülern für 25% / 20 = 1,25% des Gesamtdurchschnitts zählen, während jeder Schüler in der Klasse von 60 jeweils nur für 25% / 60 = 0,416% des Gesamtdurchschnitts steht. Ebenso zählen die Schüler in den Klassen von 30 und 40 jeweils für 0,83% und 0,625%. Jeder Schüler in der kleinsten Klasse würde den Gesamtdurchschnitt doppelt so stark beeinflussen wie jeder Schüler in der größten Klasse – im Gegensatz zu dem Ideal eines jeden Schülers, der *im Durchschnitt gleichberechtigt* ist.

Es gibt Beispiele dafür in den ersten beiden Links für die Quoten, die in diesen Abschnitt vorgestellt wurden. (Hier und hier)

Für unsere Leser in Indiana (das ist einer der Staaten in den USA), konnten wir das beim Pro-Kopf Einkommen im Indianapolis Metropol Bereich erkennen:

Table 2: Indianapolis-Carmel, IN Metro Area PCPI, 2011
Per Capita Personal Income

Area	PCPI
Boone County	$52,975
Hamilton County	$51,824
Hancock County	$43,714
Marion County	$38,309
Morgan County	$37,634
Johnson County	$36,570
Hendricks County	$36,188
Brown County	$35,863
Shelby County	$35,376
Putnam County	$31,817
What You Get If You Average County PCPIs	**$40,027**
True Metro PCPI	**$40,572**

Source: IBRC, using Bureau of Economic Analysis data http://www.incontext.indiana.edu/2013/mar-apr/article3.asp

Informationen zum Pro-Kopf-Einkommen vom *Indiana Business Research Center* und hier deren Artikel mit dem Titel: „Was man mit Daten nicht machen darf: Durchschnittswerte zu mitteln.

Wie Sie sehen können, mittelt man die Durchschnittswerte der Landkreise, bekommt man ein Pro-Kopf Einkommen von $ 40,027, aber zuerst kumulieren und dann mitteln gibt die wahre (wahrere) Zahl von $ 40.527. Dieses Ergebnis hat einen Unterschied – einen Fehler – von 1,36%. Von Interesse für diejenigen in Indiana, nur die Top drei der Verdiener in den Landkreisen haben ein Pro-Kopf-Einkommen, dass höher ist als der Landesdurchschnitt nach jedem Rechenweg und acht Landkreise liegen unter dem Durchschnitt.

Das mag trivial für Sie sein, aber bedenken Sie, dass verschiedene Behauptungen von „auffallenden neuen medizinischen Entdeckungen“ und „dem heißesten Jahr überhaupt“ nur diese Art von Unterschieden in den Rechenwegen basieren, die im Bereich der einstelligen oder sogar Bruchteilen von Prozentpunkten liegen, oder ein Zehntel oder ein Hundertstel Grad.

Um das der Klimatologie zu vergleichen, so reichen die veröffentlichten Anomalien aus dem 30-jährigen Klima- Referenzzeitraum (1981-2011) für den Monat Juni 2017 von 0,38 ° C (ECMWF) bis 0,21 ° C (UAH), mit dem Tokyo Climate Center mit einem mittlerem Wert von 0,36 ° C. Der Bereich (0,17 ° C) beträgt fast 25% des gesamten Temperaturanstiegs für das letzte Jahrhundert. (0,71ºC). **Selbst bei Betrachtung nur den beiden höchsten Werten: 0,38 ° C und 0,36 ° C, ergibt die Differenz von 0,02 ° C bereits 5% der Gesamtanomalie.** [Kann man diesen Unterschied merken? –

der Übersetzer]

Wie genau diese Mittelwerte im Ergebnis dargestellt werden, ist völlig unerheblich. Es kommt überhaupt nicht darauf an, ob man absolut Werte oder Anomalien schätzt – die Größe des induzierten Fehlers kann riesig sein

Ähnlich, aber nicht identisch, das ist Simpsons Paradox.

Simpsons Paradox

Simpsons Paradox, oder korrekter der Simpson-Yule-Effekt, ist ein Phänomen, das in Statistiken und Wahrscheinlichkeiten (und damit mit Mittelwerten) auftritt, oft in medizinischen Studien und verschiedenen Zweigen der Sozialwissenschaften zu sehen, in denen ein Ergebnis (z. Bsp. ein Trend- oder Auswirkungs-Unterschied) dargestellt wird, der beim Vergleich von Datengruppen verschwindet oder sich umkehrt, wenn die Gruppen (von Daten) kombiniert werden.

Einige Beispiele mit Simpsons Paradox sind berühmt. Eine mit Auswirkungen auf die heutigen heißen Themen behauptete Grundlage über die Anteile von Männer und Frauen bei den zum Studium an der UC Berkeley zugelassenen Bewerbern.

[Die 1868 gegründete University of California, Berkeley, gehört zu den renommiertesten Universitäten der Welt.]

Hier erklärt es einer der Autoren:

„Im Jahr 1973 wurde UC Berkeley für geschlechtsspezifische Regeln verklagt, weil ihre Zulassungszahlen zum Studium offensichtliche Nachteile gegen Frauen zeigen.

UCB	**Bewerber** Applicants	**Zugelassen** Admitted
Männer	**8442**	**44%**
Frauen	**4321**	**35%**

Statistischer Beweis für die Anklage (Tabelle abgetippt)

Männer waren viel erfolgreicher bei der Zulassung zum Studium als Frauen und damit führte es Berkeley zu einer der ersten Universitäten, die für sexuelle Diskriminierung verklagt werden sollten. Die Klage scheiterte jedoch, als Statistiker jede Abteilung getrennt untersuchten. Die Fachbereiche haben voneinander unabhängige Zulassungssysteme, daher macht es Sinn, sie separat zu überprüfen – und wenn Sie das tun, dann sie scheinen eine Vorliebe für Frauen zu haben.“

Department	Men		Women	
	Applicants	Admitted	Applicants	Admitted
A	825	62%	108	**82%**
B	560	63%	25	**68%**
C	325	**37%**	593	34%
D	417	33%	375	**35%**
E	191	**28%**	393	24%
F	373	6%	341	**7%**

In diesem Fall gaben die kombinierten (amalgamierten) Daten über alle Fachbereiche hinweg keine informative Sicht auf die Situation.

Natürlich, wie viele berühmte Beispiele, **ist die UC Berkeley Geschichte eine wissenschaftliche Legende** – die Zahlen und die mathematischen Phänomen sind wahr, aber **es gab nie eine Klage** wegen Benachteiligung von Geschlechtern. Die wahre Geschichte finden sie hier.

Ein weiteres berühmtes Beispiel für Simpsons Paradox wurde (mehr oder weniger korrekt) auf der langlaufenden TV-Serie Numb3rs vorgestellt. (Ich habe alle Episoden dieser Serie über die Jahre hinweg gesehen, einige öfter). Ich habe gehört, dass einige Leute Sportstatistiken mögen, also ist dieses etwas für Sie. Es „beinhaltet die durchschnittlichen Leistungswerte des Schlägers von Spielern im Profi-Baseball. Es ist möglich, dass ein Spieler in einer Reihe von Jahren einen höheren Durchschnitt der Abschlagwerte hat als ein anderer Spieler, obwohl er in all diesen Jahren nur einen niedrigeren Durchschnitt erreichte.“

Diese Tabelle zeigt das Paradox:

	1995		1996		1997		Combined	
Derek Jeter	12/48	.250	183/582	.314	190/654	.291	385/1284	**.300**
David Justice	104/411	**.253**	45/140	**.321**	163/495	**.329**	312/1046	.298

In jedem einzelnem Jahr erreichte David Justice einen etwas besseren Abschlag – Durchschnitt, aber wenn die drei Jahre kombiniert werden, hat Derek Jeter [auch tatsächlich ein Baseballspieler, Jg. 1974] die etwas bessere Statistik. Dies ist das Paradox von Simpson, die Ergebnisse werden umgekehrt, je nachdem ob mehrere Gruppen von Daten separat oder kumuliert betrachtet werden.

Klimatologie

In der Klimatologie tendieren die verschiedenen Gruppen die Daten über längere Zeiträume zu betrachten, um die Nachteile der Mittelungsdurchschnitte zu vermeiden. Wie wir in den Kommentaren sehen, werden verschiedene Vertreter die verschiedenen Methoden gewichten und ihre Methoden verteidigen.

Eine Gruppe behauptet, dass sie überhaupt nicht mitteln – sie engagieren sich in „räumlicher Vorhersage“, die irgendwie magisch eine Vorhersage hervorbringt, die sie dann einfach als die globale durchschnittliche Oberflächentemperatur markieren (während sie gleichzeitig bestreiten, eine Mittelung durchgeführt zu haben). Sie fangen natürlich mit täglichen, monatlichen und jährlichen Durchschnitten an – aber das sind keine echten Mittelwerte…. mehr dazu später.

Ein anderer Experte mag behaupten, dass sie definitiv keine Durchschnittswerte der Temperaturen bilden – nur den Durchschnitt von Anomalien. Das heißt, sie berechnen erst die Anomalien und dann mitteln sie diese. Wenn sie energisch genug befragt werden, dann wird diese Fraktion zugeben, dass die Mittelung längst durchgeführt wurde, die lokalen Stationsdaten – die tägliche durchschnittliche „trockene“ [dry-bulb] -Temperatur – wird wiederholt gemittelt, um monatliche Mittelwerte zu erreichen, dann Jahresdurchschnitte. Manchmal werden mehrere Stationen gemittelt, um einen „Messzellen-Durchschnitt“ zu erreichen und dann werden diese jährlichen oder klimatischen Mittel von dem gegenwärtigen absoluten Temperaturdurchschnitt subtrahiert (monatlich oder jährlich, je nach Prozess), um einen Rest zu erhalten, der dann als sogenannte „Anomalie“ bezeichnet wird – oh, dann sind die Durchschnittswerte der Anomalien „gedurchschnittet“ (gemittelt).

Die Anomalien können oder können nicht, je nach Berechnungssystem, tatsächlich gleiche Flächen der Erdoberfläche darstellen. [Siehe den ersten Abschnitt für den Fehler bei der Mittelung von Durchschnittswerten, die nicht den gleichen Bruchteil des kumulierten Ganzen darstellen (~präsentieren)]. Diese Gruppe von Experten und fast alle anderen, verlassen sich auf „nicht echte Durchschnittswerte“ an der Wurzel ihrer Berechnungen.

Die Klimatologie hat ein Mittelungsproblem, aber das echte Problem ist nicht so sehr das, was oben diskutiert wurde. In der Klimatologie ist die tägliche Durchschnittstemperatur, die bei Berechnungen verwendet wird, kein Durchschnitt der Lufttemperaturen, die während der letzten 24-Stunden-Periode bei der Wetterstation erlebt oder aufgezeichnet wurden. Es ist das arithmetische Mittel der niedrigsten und höchsten aufgezeichneten Temperaturen (Lo und Hi, das Min Max) für den 24-Stunden-Zeitraum. Es ist zum Beispiel nicht der Durchschnitt aller stündlichen Temperaturaufzeichnungen, auch wenn sie aufgezeichnet und berichtet werden. Egal wie viele Messungen aufgezeichnet werden, **der Tagesdurchschnitt wird berechnet, indem man den Lo und den Hi summiert**

und durch zwei teilt.

Macht das einen Unterschied? Das ist eine heikle Frage.

Die Temperaturen wurden als hoch und niedrig (Min-Max) für 150 Jahre oder mehr aufgezeichnet. Das ist genau so, wie es gemacht wurde und um konsequent zu bleiben, so macht man es auch heute noch.

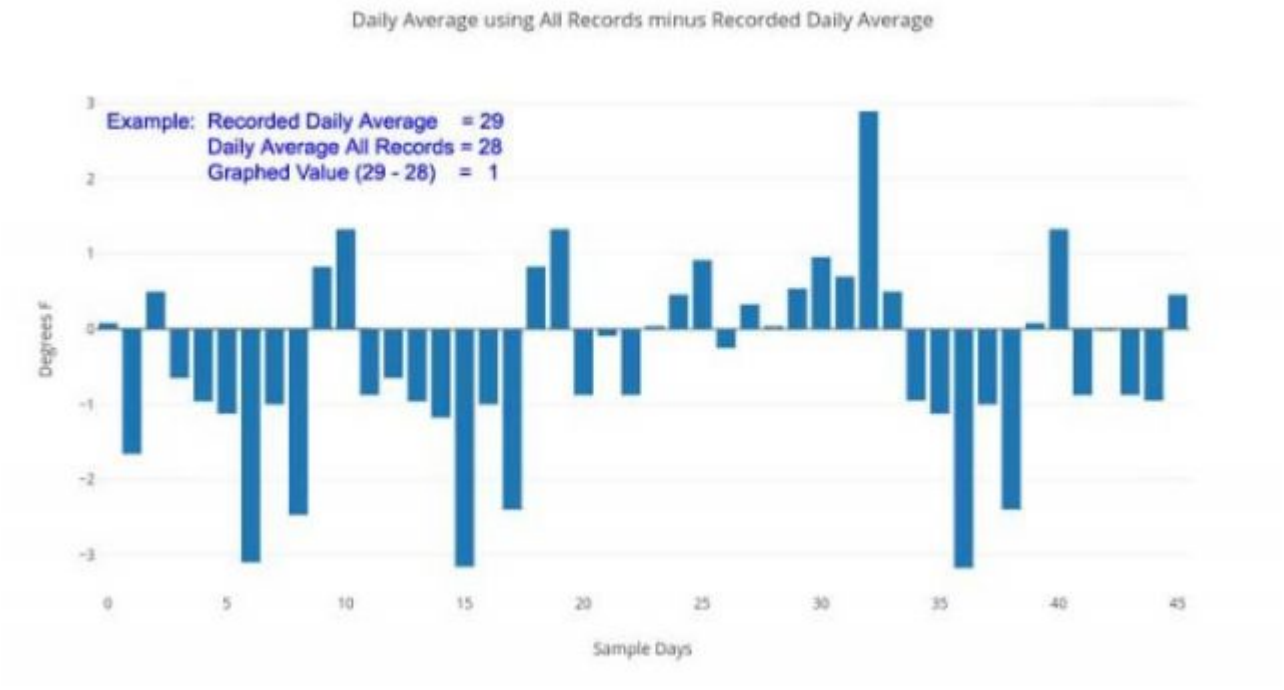
Ein Daten-Download von Temperaturaufzeichnungen für die Wetterstation WBAN: 64756, Millbrook, NY, für Dezember 2015 bis Februar 2016 enthält alle fünf Minuten eine Temperaturablesung. Der Datensatz enthält Werte für „DAILYMaximumDryBulbTemp“ und „DAILYMinimumDryBulbTemp“, gefolgt von „DAILYAverageDryBulbTemp“, alles in Grad Fahrenheit. Die „DAILYAverageDryBulbTemp“ durchschnittliche trockene Temperatur ist das arithmetische Mittel der beiden vorhergehenden Werte (Max und Min). Dieser letzte Wert wird in der Klimatologie als die **tägliche durchschnittliche Temperatur** verwendet. Für einen typischen Dezembertag sehen die aufgezeichneten Werte so aus:

Täglich Max 43 F – Täglich Min 34 F – Täglicher Durchschnitt 38F (das arithmetische Mittel ist eigentlich 38,5, aber der Algorithmus rundet x,5 ab auf x)

Allerdings ist der Tagesdurchschnitt **aller aufgezeichneten Temperaturen**: 37.3 F… Die Differenzen für diesen einen Tag:

> Unterschied zwischen dem berichteten Tagesdurchschnitt von Hi-Lo und dem tatsächlichen Durchschnitt der aufgezeichneten Hi-Lo-Zahlen ist = 0,5 ° F aufgrund des Rundungsalgorithmus.
>
> **Unterschied** zwischen dem angegebenen Tagesdurchschnitt und dem korrekteren Tagesdurchschnitt unter Verwendung aller aufgezeichneten Temperaturen = **0,667 ° F**

Andere Tage im Januar und Februar zeigen einen Differenzbereich zwischen dem gemeldeten täglichen Durchschnitt und dem Durchschnitt aller aufgezeichneten Temperaturen von 0,1 ° F bis 1,25 ° F bis zu einem hohen Wert von 3,17 ° F am 5. Januar 2016.



Täglicher Durchschnittswert aller Aufzeichnungen, korrekte Rechnung minus des aufgezeichneten Durchschnitts.

Dies ist kein wissenschaftliches Sampling – aber es ist eine schnelle Fallstudie, die zeigt, dass die Zahlen von Anfang an gemittelt werden – als tägliche Durchschnittstemperaturen, die offiziell an den Oberflächenstationen aufgezeichnet wurden. Die unmodifizierten Grunddaten selbst, werden in keinem Fall für die Genauigkeit oder Präzision überhaupt berechnet [zur Berechnung herangezogen] – aber eher werden berechnet „wie wir das schon immer getan haben“ – der Mittelwert zwischen den höchsten und niedrigsten Temperaturen in einer 24-Stunden-Periode – das gibt uns nicht einmal, was wir normalerweise als „Durchschnitt“ erwarten würden für die Temperatur an diesem Tag „- aber irgendeinen anderen Wert – ein einfaches Mittel zwischen dem Daily Lo und dem Daily Hi, das, wie die obige Grafik offenbart, ganz anders zu sein scheint. Der durchschnittliche Abstand von Null [Abweichung, d.h. wie es korrekt wäre] für die zweimonatige Probe beträgt 1,3 ° F. Der Durchschnitt aller Unterschiede, unter Beachtung des Vorzeichens beträgt 0,39 ° F [also in Richtung wärmer].

Die Größe dieser täglichen Unterschiede? Bis zu oder [sogar] größer als die gemeinhin gemeldeten klimatischen jährlichen globalen Temperaturanomalien. Es spielt keine Rolle, ob die Unterschiede nach oben oder unten zählen – es kommt darauf an, dass sie implizieren, dass die Zahlen, die verwendet werden, um politische Entscheidungen zu beeinflussen, nicht genau berechnete, grundlegende täglichen Temperaturen von einzelnen Wetterstationen sind. Ungenaue Daten erzeugen niemals genaue Ergebnisse. Persönlich glaube ich nicht, dass dieses Problem bei der Verwendung von „nur Anomalien“ verschwindet (was einige Kommentatoren behaupten) – die grundlegenden Daten der ersten Werteebene

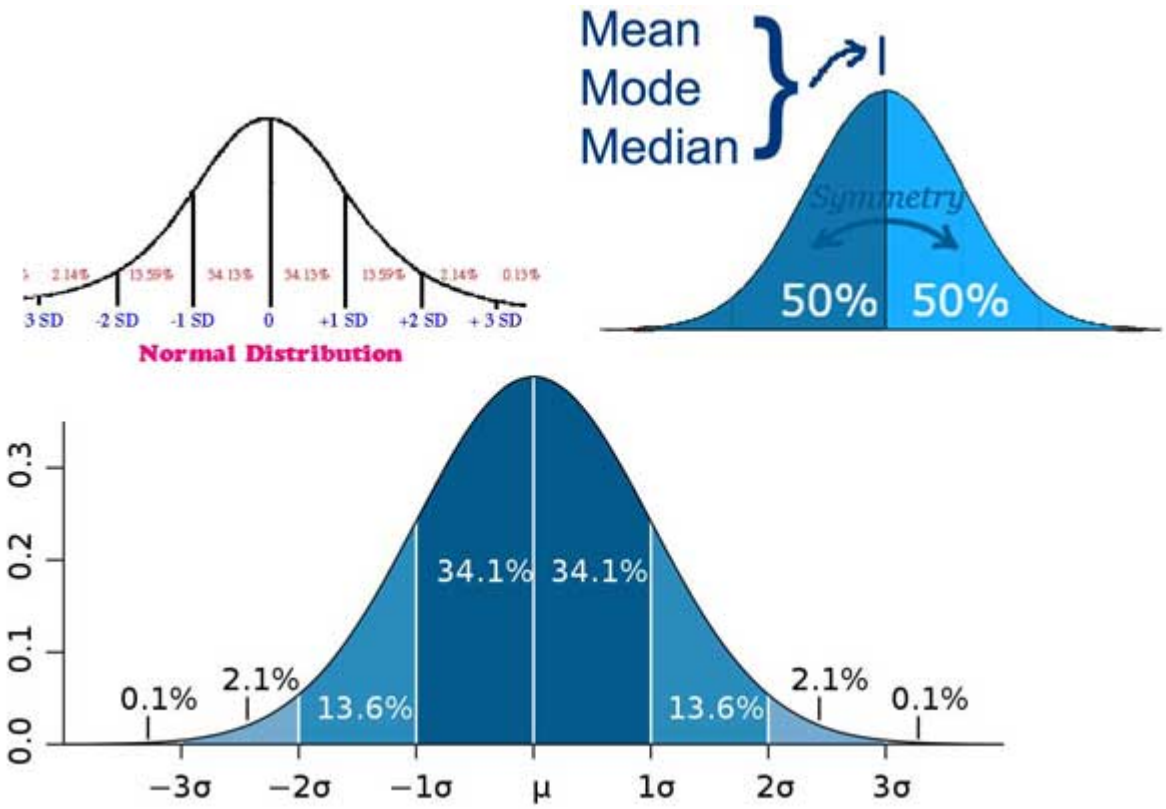
sind falsch, unpräzise, ungenau berechnet.

Aber, aber, aber … Ich weiß, ich kann die Maulerei bereits jetzt hören. Der übliche Chor von:

1. Am Ende gleicht sich alles aus (tut es nicht)
2. Aber was ist mit dem Gesetz der großen Zahlen? (Magisches Denken)
3. Wir sind nicht durch absolute Werte besorgt, nur durch Anomalien.

Die beiden ersten Argumente sind fadenscheinig.

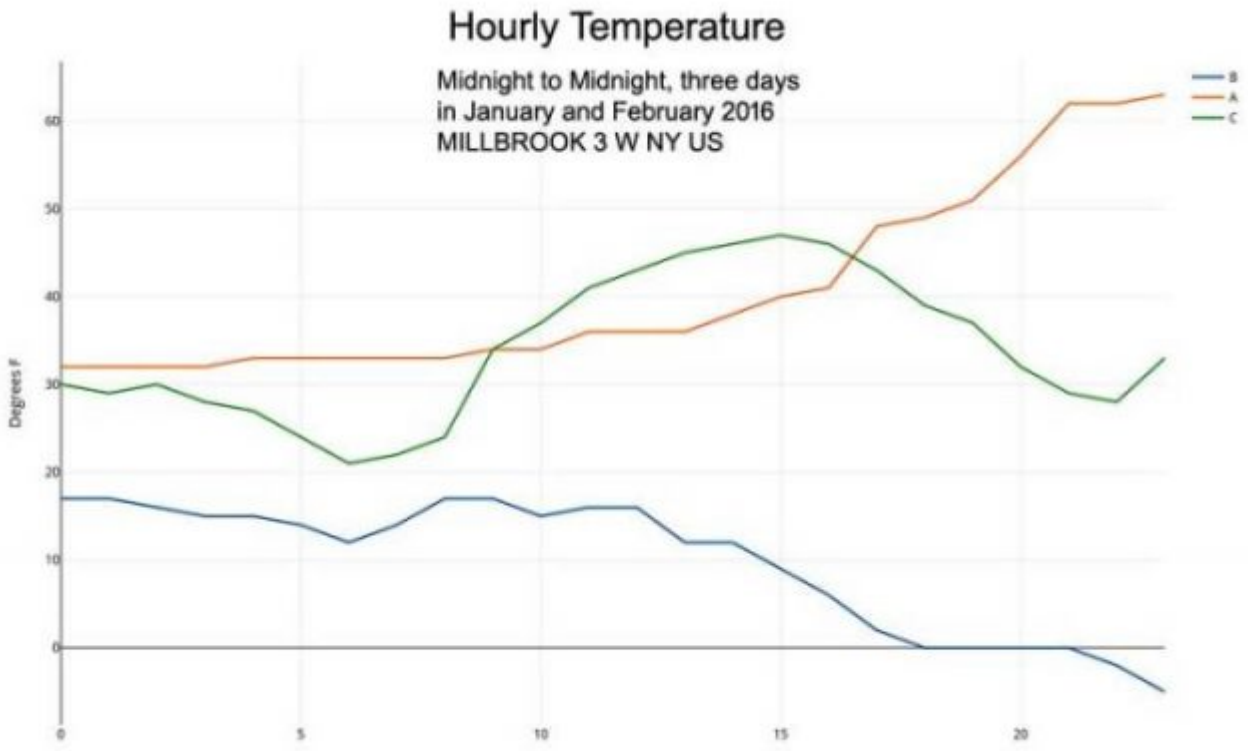
Das letzte will ich ansprechen. Die Antwort liegt in dem „Warum“ der oben beschriebenen Unterschiede. Der Grund für den Unterschied (außer der einfachen Auf- und Abwärtsbewegung von Bruchteilen in ganzem Grad) ist, dass die Lufttemperatur an einer beliebigen Wetterstation nicht normal verteilt wird … das heißt, von Minute zu Minute oder Stunde zu Stunde, keiner würde eine „normal Verteilung“ sehen, die sich so darstellt:



Normal-oder-Standard-Verteilung

Wenn die Lufttemperatur normalerweise so über den Tag verteilt wäre, dann wäre die aktuell verwendete tägliche durchschnittliche trockene Temperatur – das arithmetische Mittel zwischen dem Hi und Lo – korrekt und würde sich nicht vom täglichen Durchschnitt aller aufgezeichneten Temperaturen für den Tag unterscheiden.

Aber echte Oberflächentemperaturen der Luft sehen viel mehr aus, wie diese drei Tage von Januar und Februar 2016 in Millbrook, NY:



Reale stündliche Temperaturen

Die Lufttemperatur an einer Wetterstation startet nicht am Lo Wert – um gleichmäßig und stetig zum Hi aufzusteigen und sich dann gleichmäßig zum nächsten Lo zurück zu schleichen. Das ist ein Mythos – jeder der sich draußen aufhält (Jäger, Seemann, Camper, Forscher, sogar Jogger) kennt diese Tatsache. Doch in der Klimatologie werden die tägliche Durchschnittstemperatur – **und konsequent alle nachfolgenden wöchentlichen, monatlichen, jährlichen Durchschnitte** – auf der Grundlage dieser falschen Idee berechnet.

Zuerst nutzten die Wetterstationen Min-Max-Aufnahmethermometer und wurden oft nur einmal pro Tag überprüft und die Aufnahme-Tabs zu diesem Zeitpunkt zurückgesetzt – und nun so weitergeführt aus Respekt für Konvention und Konsistenz.

Wir können nicht zurückkehren und die Fakten rückgängig machen – aber wir müssen erkennen, dass die täglichen Mittelwerte aus diesen Min-Max / Hi-Lo-Messwerten nicht die tatsächliche tägliche Durchschnittstemperatur darstellen – weder in Genauigkeit noch in Präzision. Dieses beharren auf Konsistenz bedeutet, dass die Fehlerbereiche, die in dem obigen Beispiel dargestellt sind, alle globalen durchschnittlichen Oberflächentemperaturberechnungen beeinflussen, die Stationsdaten als Quelle verwenden.

Anmerkung: Das hier verwendete Beispiel ist von Wintertagen in einem gemäßigten Klima. Die Situation ist repräsentativ, aber nicht unbedingt quantitativ – sowohl die Werte als auch die Größen der Effekte werden für verschiedene Klimazonen [… Gegenden], verschiedene Stationen, verschiedene Jahreszeiten unterschiedlich sein. Der Effekt kann durch statistische Manipulation oder durch Reduzierung der Stationsdaten zu Anomalien nicht vermieden werden.

Alle Anomalien, die durch Subtrahieren von klimatischen Durchschnittswerten von aktuellen Temperaturen abgeleitet werden, werden uns nicht mitteilen, ob die durchschnittliche absolute Temperatur an einer Station steigt oder fällt (oder um wie viel). Es wird uns nur sagen, dass der Mittelwert zwischen den täglichen Hochtemperaturen steigt oder fällt – was etwas ganz anders ist. Tage mit sehr niedrigen Tiefs für eine Stunde oder zwei am frühen Morgen gefolgt von hohen Temperaturen für die meiste Zeit des restlichen Tages haben die gleichen Max-Min Mittelwerte wie Tage mit sehr niedrigen Tiefs für 12 Stunden und eine kurze heiße Spitze am Nachmittag. Diese beiden Arten von Tagen, haben nicht die gleiche tatsächliche durchschnittliche Temperatur. Anomalien können den Unterschied nicht erhellen. Ein Klimawandel von einem zum anderen wird in Anomalien nicht auftauchen. Weder noch würde die Umwelt von einer solchen Verschiebung stark betroffen sein.

Können wir etwas erfahren, aus der Nutzung dieser ungenauen „täglichen Durchschnittswerte“ und alle anderen von ihnen abgeleiteten Zahlen?

Es gibt einige die in Frage stellen, dass es tatsächliche eine globale durchschnittliche Oberflächentemperatur gibt. (Siehe „Gibt es eine globale Temperatur?“)

Auf der anderen Seite, formulierte Steven Mosher in seinem Kommentar kürzlich so treffend:

> *Sind abgeleitete Temperaturen bedeutungslos?*
>
> *Ich denke, dass es schlecht ist, wenn Alarmisten versuchen, die kleine Eiszeit und die Mittelalterliche Warmzeit zu löschen … WUWT wird die ganze Geschichte leugnen wollen. Die globale Temperatur ist vorhanden. Es hat eine genaue physikalische Bedeutung. Es ist diese Bedeutung, die uns erlaubt zu sagen … Die kleine Eiszeit war kühler als heute … es ist die Bedeutung, die es uns erlaubt, zu sagen, dass die die Tagseite des Planeten wärmer ist als die Nachtseite … die gleiche Bedeutung, die uns erlaubt zu sagen, dass Pluto kühler ist als die Erde und dass Merkur wärmer ist.*

Was ein solcher globaler Mittelwert, basierend auf einen fragwürdigen abgeleiteten „Tagesdurchschnitt“, uns nicht sagen kann, ist oder war es in diesem Jahr ein Bruchteil eines Grades wärmer oder kühler?

Der Berechnungsfehler – der Messfehler – der am häufigsten verwendeten Station der täglichen durchschnittlichen Durchschnittstemperatur ist in der Größe gleich (oder nahezu gleich groß) wie die langfristige globale Temperaturänderung. Die historische Temperaturaufzeichnung kann nicht für diesen Fehler korrigiert werden. Moderne digitale Aufzeichnungen würde eine Neuberechnung der Tagesmittelwerte von Grund auf neu erfordern. Selbst dann würden die beiden Datensätze quantitativ nicht vergleichbar sein – möglicherweise nicht einmal qualitativ.

Also, „Ja, es macht was aus“

Es ist sehr wichtig, wie und was man mittelt. Es macht was aus, den ganzen Weg nach oben und unten durch das prächtige mathematische Wunderland, was die Computerprogramme darstellen, die diese grundlegenden digitalen Aufzeichnungen von Tausenden von Wetterstationen auf der ganzen Welt lesen und sie zu einer einzigen Zahl verwandeln.

Es macht vor allem dann etwas aus, wenn diese einzelne Zahl dann später als Argument verwendet wird, um die breite Öffentlichkeit zu treffen und unsere politischen Führer zu bestimmten gewünschten politischen Lösungen zu bringen (~ zu zwingen), die große – und viele glauben negative – Auswirkungen auf die Gesellschaft haben werden.

Schlussfolgerung:

Es reicht nicht aus, den Durchschnitt eines Datensatzes korrekt mathematisch zu berechnen.

Es reicht nicht aus, die Methoden zu verteidigen, die Ihr Team verwendet, um die [oft-mehr-missbrauchten-als-nicht] globalen Mittelwerte von Datensätzen zu berechnen.

Auch wenn diese Mittelwerte von homogenen Daten und Objekten sind und physisch und logisch korrekt sind, ein Mittelwert ergibt eine einzelne Zahl und kann nur fälschlicherweise als summarische oder gerechte Darstellung des ganzen Satzes, der ganzen Information angenommen werden.

Durchschnittswerte, in jedem und allen Fällen, geben natürlicherweise nur einen sehr eingeschränkten Blick auf die Informationen in einem Datensatz – und wenn sie als Repräsentation des Ganzen akzeptiert werden, wird sie als Mittel der Verschleierung fungieren, die den Großteil verdecken und die Information verbergen. Daher, anstatt uns zu einem besseren Verständnis zu führen, können sie unser Verständnis des zu untersuchenden Themas reduzieren.

In der Klimatologie wurden und werden tägliche Durchschnittstemperaturen ungenau und unpräzise aus den täglichen minimalen und maximalen

Temperaturen berechnet, die damit Zweifel an den veröffentlichten globalen durchschnittlichen Oberflächentemperaturen hervorrufen.

Durchschnitte sind gute Werkzeuge, aber wie Hämmer oder Sägen müssen sie korrekt verwendet werden, um wertvolle und nützliche Ergebnisse zu produzieren. Durch den Missbrauch von Durchschnittswerten verringert sich das Verständnis des Themas eher, als das es die Realität abbildet.

UPDATE:

> Diejenigen, die mehr Informationen bekommen wollen, über die Unterschiede zwischen Tmean (das Mittel zwischen Täglichem Min und Max) und Taverage (das arithmetische Mittel aller aufgezeichneten 24 stündlichen Temperaturen – einige verwenden T24 dafür) – sowohl quantitativ als auch in jährlichen Trends beziehen sich beide auf die *Spatiotemporal Divergence of the Warming Hiatus over Land Based on Different Definitions of Mean Temperature* von Chunlüe Zhou & Kaicun Wang [Nature Scientific Reports | 6:31789 | DOI: 10.1038/srep31789].*[~Räumliche und zeitliche Abweichungen der Erwärmungspause über Land auf der Grundlage verschiedener Definitionen der mittleren Temperatur]* von Chunlüe Zhou & Kaicun Wang [Nature Scientific Reports | 6: 31789 | DOI: 10.1038 / srep31789]. Im Gegensatz zu Behauptungen in den Kommentaren, dass Trends dieser unterschiedlich definierten „durchschnittlichen“ Temperaturen gleich sind, zeigen Zhou und Wang diese Figur und Kation: (h / t David Fair)

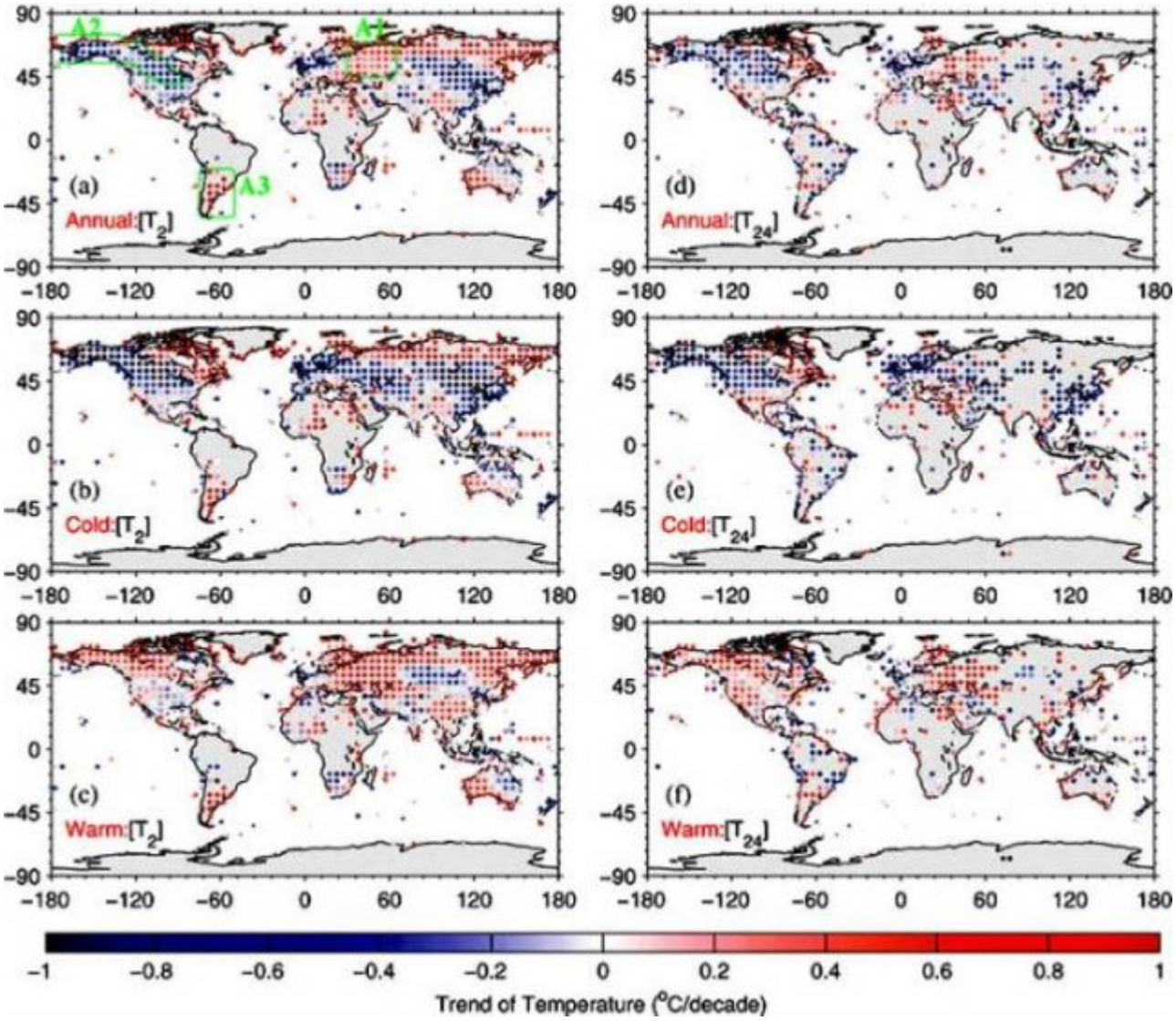


Abbildung 4. Die (a, d) jährliche, (b, e) Kälte und (c, f) warme saisonale Temperaturtrends (Einheit: ° C / Jahrzehnt) aus dem Global Historical Climatology Network-Daily Version 3.2 (GHCN-D , [T2]) und die „Integrated Surface Database-Hourly“ (ISD-H, [T24]) sind für 1998-2013 gezeigt. Die GHCN-D ist eine integrierte Datenbank der täglichen Klimazusammenfassungen von Land-Oberflächenstationen auf der ganzen Welt, die Tmax und Tmin Werte von rund 10.400 Stationen von 1998 bis 2013 zur Verfügung stellt. Die ISD-H besteht aus globalen stündlichen und synoptischen Beobachtungen bei etwa 3400 Stationen aus über 100 Originaldatenquellen. Die Regionen A1, A2 und A3 (innerhalb der grünen Regionen, die in der oberen linken Teilfigur gezeigt sind) werden in dieser Studie ausgewählt.

[Klicken Sie hier für ein Bild in voller Größe] in Nature

Erschienen auf WUWT am 24.07.2017

Übersetzt durch Andreas Demmig

https://wattsupwiththat.com/2017/07/24/the-laws-of-averages-part-3-the-average-average/