

Unsicherheiten bei Routine-Messungen der Temperatur – Teil 3

geschrieben von Chris Frey | 17. Oktober 2022

Thomas Berger und Geoffrey Sherrington.

Teil 1 [hier](#), in deutscher Übersetzung [hier](#).

Teil 2 [hier](#), in deutscher Übersetzung [hier](#)

Alle Hervorhebungen im Original

Begrüßen Sie bitte den Mitautor Tom Berger, der diese australischen Temperaturdatensätze mehrere Jahre lang mit mathematisch-forensischen Untersuchungsmethoden untersucht hat. Viele davon basieren auf der Statistiksoftware SAS JMP.

Bitte lesen Sie diesen langen Aufsatz bis zum Ende, denn wir sind zuversichtlich, dass er eine Menge neues Material enthält, das Sie veranlassen wird, das gesamte Thema der Qualität der grundlegenden Daten hinter dem Klimawandel-Theater zu überdenken. Ich fand es Augen öffnend, beginnend mit diesem Satz von Temperaturen von der ersten BOM-Site am Sydney Observatory. Die angeblichen „Roh“-Daten sind nicht roh, weil identische Daten von Juli 1914 bis Juli 1915 kopiert und eingefügt wurden:

SYDNEY JULY 2014 DAILY Tmin			SYDNEY JULY 2015 DAILY Tmin		
DATE	minv22	minraw	DATE	minv22	minraw
1914-07-01	12.1	12.3	1915-07-01	12.1	12.3
1914-07-02	12.3	12.7	1915-07-02	12.3	12.7
1914-07-03	12.5	12.9	1915-07-03	12.5	12.9
1914-07-04	11.6	11.2	1915-07-04	11.6	11.2
1914-07-05	12.1	12.2	1915-07-05	12.1	12.2
1914-07-06	11.7	11.3	1915-07-06	11.7	11.3
1914-07-07	11.8	11.6	1915-07-07	11.8	11.6
1914-07-08	8.2	7.8	1915-07-08	8.2	7.8
1914-07-09	8.3	8.0	1915-07-09	8.3	8.0
1914-07-10	8.2	7.7	1915-07-10	8.2	7.7
1914-07-11	7.6	7.2	1915-07-11	7.6	7.2
1914-07-12	7.6	7.2	1915-07-12	7.6	7.2
1914-07-13	6.9	6.3	1915-07-13	6.9	6.3
1914-07-14	6.8	6.2	1915-07-14	6.8	6.2
1914-07-15	8.2	7.7	1915-07-15	8.2	7.7
1914-07-16	4.1	4.1	1915-07-16	4.1	4.1
1914-07-17	6.8	6.2	1915-07-17	6.8	6.2
1914-07-18	8.2	7.7	1915-07-18	8.2	7.7
1914-07-19	7.4	6.9	1915-07-19	7.4	6.9
1914-07-20	8.2	7.7	1915-07-20	8.2	7.7
1914-07-21	9	8.8	1915-07-21	9	8.8
1914-07-22	9.6	9.2	1915-07-22	9.6	9.2
1914-07-23	7	6.6	1915-07-23	7	6.6
1914-07-24	8.2	7.7	1915-07-24	8.2	7.7
1914-07-25	8.2	7.8	1915-07-25	8.2	7.8
1914-07-26	9.2	8.9	1915-07-26	9.2	8.9
1914-07-27	8.2	7.9	1915-07-27	8.2	7.9
1914-07-28	7.3	6.8	1915-07-28	7.3	6.8
1914-07-29	9.7	9.4	1915-07-29	9.7	9.4
1914-07-30	9.7	9.4	1915-07-30	9.7	9.4
1914-07-31	9	8.8	1915-07-31	9	8.8

Dies geschah vor den Augen derjenigen, die die homogenisierte Version 22 von ACORN-SAT erstellt haben, die Ende 2021 veröffentlicht worden war. Damit wurde die Büchse der Pandora geöffnet, denn dies ist nicht das einzige Kopieren und Einfügen in den „Rohdaten“ von Sydney. Siehe auch

Juni 1943 bis Juni 1944 (ganzer Monat), Dezember 1935 bis Dezember 1936 (ganzer Monat, aber 2 im Jahr 1935 fehlende Werte haben Werte im Jahr 1936).

Tom hat mir gegenüber argumentiert, dass...

... „Die Strategie, die Menschen dazu zu bringen, den wissenschaftlichen Weg zu respektieren, ist ein hoffnungsloser Fall, fürchte ich. Die Idee ist, verschiedene Dinge zu zeigen und sie selbst zu einer Schlussfolgerung kommen zu lassen!!! Das ist entscheidend.“

Das geht mir als langjährigem Wissenschaftler zwar gegen den Strich, aber da wir das tun, was aufgeschlossene Wissenschaftler tun sollten, werden wir Toms Arbeit akzeptieren und veröffentlichen. Tom schlägt vor, dass wir der Propaganda stärker entgegentreten, also lassen Sie uns das tun, indem wir mit dem Thema der „rohen“ Temperaturmessungen beginnen. Dies führt dazu, was „bereinigte“ Temperaturen für die intellektuelle und mathematische Reinheit bedeuten können. Dann wird Tom einige Methoden vorstellen, die sicher nur wenige Leser kennen.

Bei diesen Temperaturen handelt es sich bei den „rohen“ Daten um die Aufzeichnungen der Beobachter über ihre täglichen Beobachtungen. Ich habe die Bibliothek des Bureau of Meteorology in Melbourne besucht und mit Erlaubnis fotografiert, was als „Rohdaten“ gezeigt wurde. Hier ist ein solches Blatt, aus Melbourne, August 1860. (Ich habe die Spalten, die von besonderem Interesse sind, leicht eingefärbt):

<http://www.geoffstuff.com/melborig.jpg>

Der nächste Monat, September 1860, hat ein ganz anderes Aussehen, was auf monatliche Aufräumarbeiten hindeutet:

<http://www.geoffstuff.com/nextmelborig.jpg>

Diejenigen, die täglich Messungen vorgenommen haben, könnten den Eindruck haben, dass es sich bei diesen Blättern um Abschriften von früheren Dokumenten handelt. Ich habe ihn. Die Handschrift ändert sich von Monat zu Monat, nicht von Tag zu Tag. Abschriften sind ein fruchtbarer Boden für Korrekturen.

Abgesehen von den Berichtigungen der Rohdaten stammen die in diesem Aufsatz verwendeten „bereinigten“ Daten fast immer aus dem ACORN-SAT-Homogenisierungsverfahren, das von der BOM verwendet wird. Es gibt vier Versionen, die in der Kurzform als V1, V2, v21 und v22 für Version 1, Version 2. Version 2.1 und Version 2.2. Die Tageshöchsttemperaturen und -minima werden als Diagrammbeschriftungen im Stil von minv22 oder maxraw abgekürzt, um nur zwei Beispiele zu nennen. Die meisten Wetterstationen sind nach Ortschaften benannt, wie z.B. die nächstgelegene Stadt, wobei die ACORN-SAT-Stationen in diesem Katalog aufgeführt sind.

<http://www.bom.gov.au/climate/data/acorn-sat/stations/#/23000>

Der ACORN-SAT-Anpassungs-/Homogenisierungsprozess wird in mehreren BOM-Berichten wie diesem und weiteren beschrieben:

http://www.bom.gov.au/climate/change/acorn-sat/documents/About_ACORN-SAT.pdf

http://www.bom.gov.au/climate/data/acorn-sat/documents/ACORN-SAT_Report_No_1_WEB.pdf

Ohne die wirklich grundlegenden, fundamentalen, rohen Beweise wie z. B. Original-Beobachterbögen, die ich erfolglos gesucht habe, wende ich mich nun an Tom, der die verfügbaren Daten studiert, um zu sehen, was bei einer forensischen Untersuchung noch alles gefunden werden kann.

Von der BOM für die Erstellung von Trends verwendete Australische Klimadaten werden analysiert und seziert. Die Ergebnisse zeigen, dass die Daten verzerrt und fehlerhaft sind, bei einigen Stationen sogar bis zum Jahr 2010, so dass sie sich nicht für Vorhersagen oder Trends eignen.

In vielen Fällen handelt es sich bei den Temperatursequenzen der Daten um eine Aneinanderreihung von Duplikaten und wiederholten Sequenzen, die keine Ähnlichkeit mit den beobachteten Temperaturen haben.

Diese Daten würden in vielen Branchen wie der Pharmaindustrie und der industriellen Steuerung verworfen werden. Viele der Methoden zur Verarbeitung von Stücklistendaten sind für die meisten Branchen ungeeignet.

Ungeeignete Datenstationen scheinen in das Netz aufgenommen worden zu sein, um das Argument der Knappheit des Klimanetzes zu entkräften, das gegen das australische Klimanetz vorgebracht wurde. (Modeling And Pricing Weather-Related Risk, Antonis K. Alexandridis et al)

Wir verwenden eine forensische Explorations-Software (SAS JMP), um plausibel gefälschte Sequenzen zu identifizieren, entwickeln aber auch eine einfache Technik, um Cluster von gefälschten Daten aufzuzeigen. Zusammen mit Data-Mining-Techniken wird kausal gezeigt, dass BOM-Anpassungen gefälschte unnatürliche Sequenzen erzeugen, die nicht mehr als Beobachtungs- oder Beweisdaten funktionieren.

„Diese (Klima-)Forschungsergebnisse enthalten einen Zirkelschluss, weil die Hypothese am Ende mit Daten bewiesen wird, aus denen die Hypothese abgeleitet wurde.“

Zirkelschluss in der Forschung zum Klimawandel – Jamal Munshi

BEVOR WIR BEGINNEN – EINE ANOMALIE EINER ANOMALIE

Ein ewig wahrender Mythos:

*„Beachten Sie, dass die Temperatur-Zeitreihen als Anomalien oder Abweichungen vom Durchschnitt 1961-1990 dargestellt werden, da **Temperaturanomalien dazu neigen, in weiten Gebieten konsistenter zu sein als die tatsachlichen Temperaturen.**“* –BOM ([Link](#))

Das ist Blodsinn. Beachten Sie den Terminus „dazu neigen“, das auf der endgultigen NASA-Website nicht vorkommt. Die BOM versucht, die Aussage abzuschwachen, indem sie einen Ausweg anbietet. Wenn Worthulsen wie „vielleicht“, „kann“, „konnte“ oder „tendieren“ verwendet werden, sind dies rote Fahnen, die nutzliche Bereiche fur Untersuchungen liefern.

Die Verwendung eines **willkurlich** gewahlten Durchschnittswerts eines 30-jahrigen Temperaturblocks, eines Offsets, macht diese Gruppe nicht „normal“ und liefert auch nicht mehr Daten, als man bereits hat.

Die Aufzeichnung von Abweichungen von diesem **willkurlich** gewahlten Offset fur ein begrenztes Netz von Stationen bringt keine weiteren Erkenntnisse und bedeutet ganz sicher nicht, dass man die Analyse auf Gebiete ohne Stationen ausdehnen oder Extrapolationen legitimieren kann, wenn man dort keine Messungen vorgenommen hat.

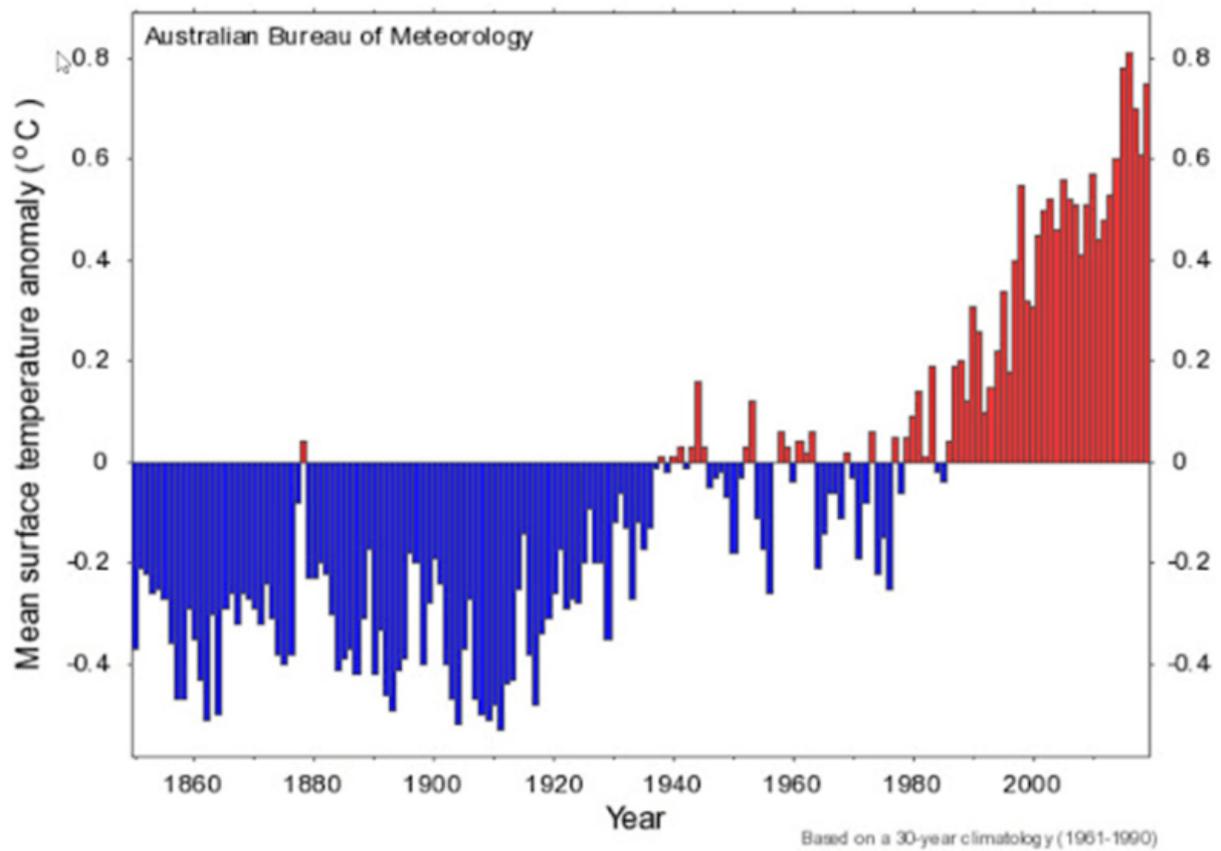
Die Mittelung von Temperaturanomalien *„uber weite Gebiete“*, wenn man nur einige wenige Stationsmesswerte hat, ergibt kein genaueres Bild als die Mittelung der reinen Temperaturen.

GROSS DENKEN, GLOBAL DENKEN:

Betrachtung der jahrlichen globalen Temperaturanomalien: Ist dies die Waffe der Wahl, wenn es darum geht, Panikmache zu betreiben? Sie besteht darin, fast eine Million Temperaturanomalien zu einer einzigen Zahl zusammenzufassen. ([Link](#))

Hier eine zusammenfassende Grafik von der BOM-Website aus dem Jahr 2022 (die Daten reichen eigentlich nur bis 2020):

Annual mean temperature anomaly - Global (1850-2019)



Die Wayback-Website findet die Jahre 2014 und 2010 und 2022 auf der BOM-Website. Es wurde nichts früheres gefunden. Hier das Diagramm aus dem Jahr 2010:

Variable Region Season

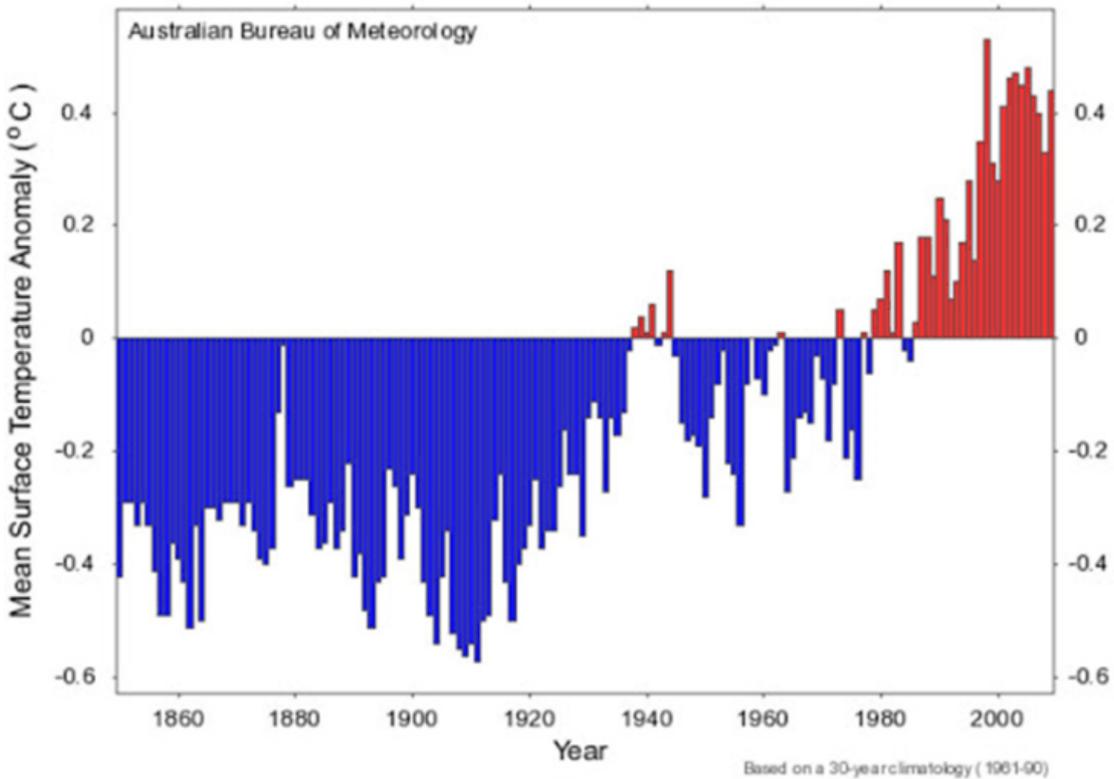
Years of running average (T=linear trend; A=average)

[Trend map](#) | [Average map](#)

Download Raw data set | Sorted data set

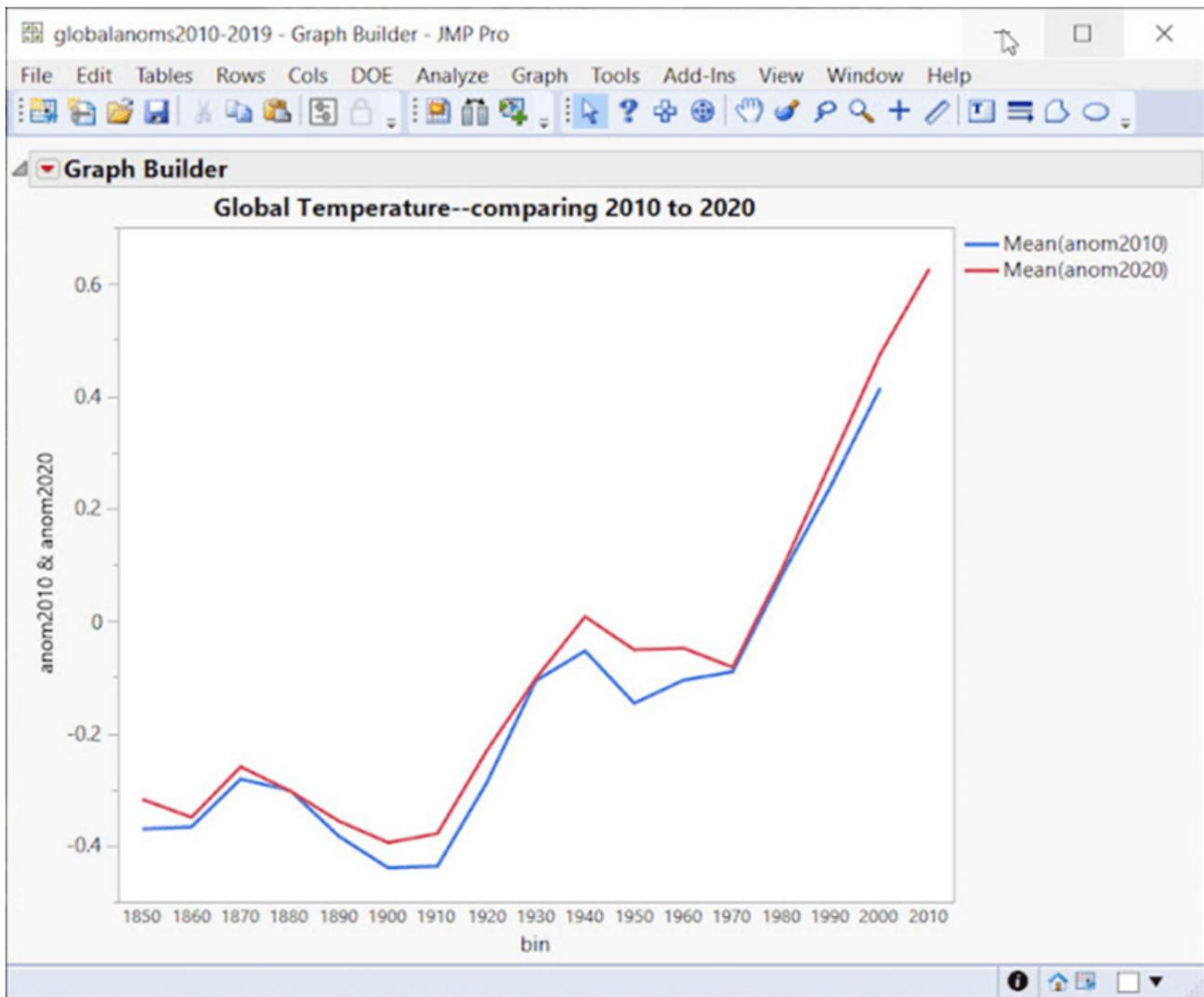
Average (1961-90) 13.97 °C

Annual Mean Surface Temperature Anomaly - Global



Wenn man sich die beiden Diagramme ansieht, kann man Unterschiede erkennen. Diese lassen darauf schließen, dass es eine Erwärmung gegeben hat, aber um wie viel?

Die Überlagerung der Temperaturanomalien aus den Jahren 2010 und 2020 ist hilfreich:



Die BOM betonen, dass ihre Anpassungen und Änderungen gering sind, zum Beispiel:

„Die Unterschiede zwischen den ‚rohen‘ und den ‚homogenisierten‘ Datensätzen sind gering und erfassen die Unsicherheit in den Temperaturschätzungen für Australien.“ – BOM [\(Link\)](#)

Lassen Sie uns dies mit einer Hypothese testen: Nehmen wir an, dass bei jeder Version der Grafik der globalen Temperaturanomalien (2010, 2014, 2020) die Erwärmung auf dem 95 %-Niveau signifikant zugenommen hat (unter Verwendung der gleichen Signifikanzniveaus wie bei BOM).

Die Nullhypothese besagt, dass die Verteilungen gleich sind und keine signifikante Erwärmung zwischen den verschiedenen Plots stattgefunden hat.

Also: 2010 > 2014 < 2020.

Um dies zu testen, verwenden wir:

EINEN NICHTPARAMETRISCHEN KOMBINATIONSTEST

Dabei handelt es sich um einen Permutationstest, der eine genaue Kombination verschiedener Hypothesen ermöglicht. Er macht keine Annahmen, außer dass die Beobachtungen austauschbar sind, funktioniert mit kleinen Stichproben und wird nicht durch fehlende Daten beeinträchtigt.

Pesarin hat den NPC populär gemacht, aber Devin Caughey vom MIT hat die aktuellste und flexibelste Version des Algorithmus, geschrieben in R. ([Link](#)). Er wird auch häufig verwendet, wenn eine große Anzahl von Kontrasten untersucht wird, wie z. B. in Gehirnscan-Labors. ([Link](#))

„Da die NPC auf der Permutationsinferenz basiert, erfordert sie keine Modellierungsannahmen oder asymptotische Begründungen, sondern nur, dass die Beobachtungen unter der globalen Nullhypothese, wonach die Behandlung keine Wirkung hat, austauschbar sind (z. B. zufällig zugeordnet).“ – Devin Caughey, MIT

Nachdem wir NPC in R ausgeführt haben, ist unser wichtigstes Ergebnis:

2010 < 2014 führt zu einem p-Wert von 0,0444

Dieser Wert liegt unter unserem Grenzwert von $p = 0,05$, so dass wir den Nullfall ablehnen und sagen können, dass die globalen Temperaturanomalien in den Diagrammen 2010 und 2014 **eine signifikante Erwärmung erfahren haben und dass die Verteilungen unterschiedlich sind.**

Das Ergebnis $2020 > 2014$ hat einen p-value = 0,1975

Wir verwerfen hier nicht die Null, also ist 2014 nicht signifikant verschieden von 2020.

Wenn wir die p-Werte unter Verwendung der Hypothese ($2010 < 2014 > 2020$, d. h. zunehmende Erwärmung in jeder Version) mit der NPC kombinieren, erhalten wir einen p-Wert von 0,0686. Dieser Wert liegt knapp unter dem Signifikanzniveau von 5 %, so dass wir die Nullhypothese nicht zurückweisen, obwohl die Diagramme eine fortschreitende Erwärmung belegen.

Daraus ergibt sich, dass sich die Grafiken der globalen Temperaturanomalien durch die Erhöhung der Temperaturen zwischen 2010 und 2014 erheblich verändert haben, danach blieben sie im Wesentlichen gleich. Die Frage ist, ob es sich um eine tatsächliche Erwärmung oder um eine synthetische Erwärmung handelt.

ICH SEHE ES, ABER ICH GLAUBE ES NICHT...

„Wenn Sie Durchschnittswerte verwenden, liegen Sie im Durchschnitt falsch.“ – Dr. Sam Savage über den Makel der Durchschnittswerte

Wie unten gezeigt wird, neigt das BOM dazu, Temperatursequenzen zu kopieren/einzufügen oder zu verändern, wodurch Blöcke mit doppelten Temperaturen oder doppelten Sequenzen entstehen, die einige Tage oder Wochen oder sogar einen ganzen Monat dauern. Bei einer so kleinen Stichprobe wie den globalen Temperaturanomalien hätten sie das sicher nicht getan, oder?

year	bin	anom2010	Difference[anom 2010]	diff0_2010	anom2014	anom2020	Difference[anom 2020]	diff0_2020
1	1850	1850	-0.42	0	-0.37	-0.37	0	0
2	1851	1850	-0.29	0.13	0	-0.21	-0.21	0.16
3	1852	1850	-0.29	0	1	-0.22	-0.22	-0.01
4	1853	1850	-0.33	-0.04	0	-0.27	-0.26	-0.04
5	1854	1850	-0.29	0.04	0	-0.25	-0.25	0.01
6	1855	1850	-0.33	-0.04	0	-0.27	-0.27	-0.02
7	1856	1850	-0.41	-0.08	0	-0.36	-0.36	-0.09
8	1857	1850	-0.49	-0.08	0	-0.47	-0.47	-0.11
9	1858	1850	-0.49	0	1	-0.47	-0.47	0
10	1859	1850	-0.36	0.13	0	-0.29	-0.29	0.18
11	1860	1860	-0.39	-0.03	0	-0.35	-0.35	-0.06
12	1861	1860	-0.43	-0.04	0	-0.43	-0.43	-0.08
13	1862	1860	-0.51	-0.08	0	-0.51	-0.51	-0.08
14	1863	1860	-0.33	0.18	0	-0.3	-0.3	0.21
15	1864	1860	-0.5	-0.17	0	-0.5	-0.5	-0.2
16	1865	1860	-0.3	0.2	0	-0.28	-0.29	0.21
17	1866	1860	-0.3	0	1	-0.25	-0.26	0.03
18	1867	1860	-0.32	-0.02	0	-0.32	-0.32	-0.06
19	1868	1860	-0.29	0.03	0	-0.26	-0.26	0.06
20	1869	1860	-0.29	0	1	-0.27	-0.27	-0.01
21	1870	1870	-0.29	0	1	-0.29	-0.29	-0.02
22	1871	1870	-0.33	-0.04	0	-0.32	-0.32	-0.03
23	1872	1870	-0.29	0.04	0	-0.24	-0.24	0.08
24	1873	1870	-0.34	-0.05	0	-0.31	-0.31	-0.07
25	1874	1870	-0.39	-0.05	0	-0.38	-0.38	-0.07
26								

Unglaublicherweise gibt es sogar in dieser kleinen Stichprobe eine doppelte Sequenz. SAS JMP berechnet, dass die Wahrscheinlichkeit, diese Sequenz bei dieser Stichprobengröße und der Anzahl eindeutiger Werte zufällig zu sehen, gleich der Wahrscheinlichkeit ist, dass bei einem Münzwurf 10 Köpfe hintereinander fallen. Mit anderen Worten: unwahrscheinlich, aber möglich. Skeptiker halten es für wahrscheinlicher, dass es sich um die Hypothese der fragwürdigen Daten handelt.

Fragwürdige Sequenzen gibt es in Rohdaten, wenn bis zu 40 Stationen gemittelt werden, sie entstehen auch durch „Anpassungen“.

DER FALL DES HUNDES, DER NICHT GEBELLT HAT

So wie der Hund, der in einer bestimmten Nacht nicht bellt, für Sherlock Holmes bei der Lösung eines Falles von großer Bedeutung war, so ist es für uns wichtig zu wissen, was nicht da ist.

Wir müssen wissen, welche Variablen verschwinden und auch, welche plötzlich wieder auftauchen.

„Eine Studie, die Daten auslässt, ist eine große rote Fahne. Eine Entscheidung, Daten einzubeziehen oder auszuschließen, macht manchmal einen großen Unterschied.“ – Standardabweichungen, fehlerhafte Annahmen, verfälschte Daten und andere Möglichkeiten, mit Statistiken zu lügen, Gary Smith.



The screenshot shows the 'Summary Statistics' dialog box in SPSS. It lists 10 columns with their respective statistics. The table is as follows:

Columns	N	N Missing	Min	Max	Mean	Std Dev
maxv1	36427	2651	0.0	42.8	32.48	2.9815429799967
maxv2	36959	2119	15.3	43.2	32.80	2.8142628052402
maxv21	37319	1759	15.3	43.2	32.80	2.8191394769295
maxv22	37315	1763	15.3	43.2	32.798413506633	2.8190815009825
maxraw	36520	2558	16.1	42.9	32.79	2.8184137899998
minv1	34599	4479	-2.4	30.3	18.42	4.1548091630198
minv2	35170	3908	-3.1	28.9	17.73	4.4739018629163
minv21	35532	3546	-3.1	28.9	17.73	4.4770590738973
minv22	35507	3571	-3.1	28.9	17.733413129805	4.4767094086592
minraw	34777	4301	-0.1	29.0	19.10	4.0364497501255

Eine Zusammenfassung der fehlenden Daten der Wetterstation Palmerville in Nord-Queensland zeigt, wie die Daten von verschiedenen Versionen der BOM-Software gelöscht oder unterstellt werden. Betrachtet man die Mindesttemperaturen, so sind die anfänglichen Daten, mit denen die BOM arbeitet, Rohdaten, so dass minraw 4301 fehlende Temperaturen aufweist. Nach den Anpassungen der Version minv1 fehlen nun 4479 Temperaturen, was einem Verlust von 178 Werten entspricht.

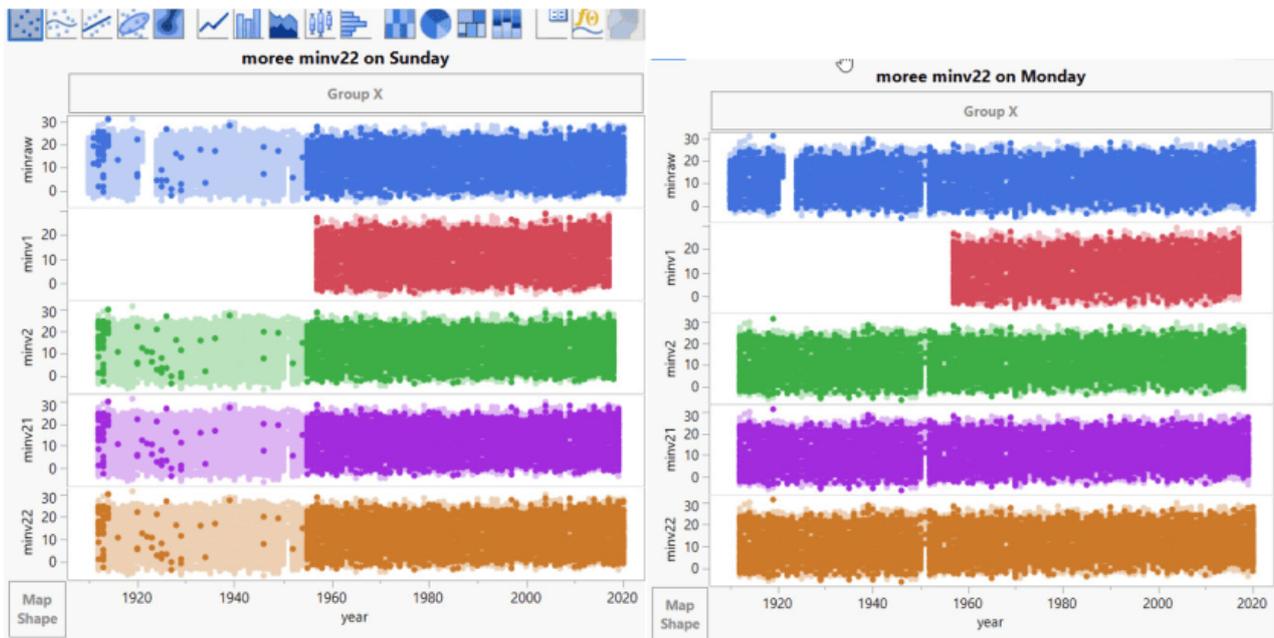
Nach den Anpassungen der Version minv2 fehlen jetzt 3908 Werte, so dass jetzt 571 Werte importiert oder ergänzt wurden.

Einige Jahre später ist die Technologie so weit fortgeschritten, dass das BOM eine neue Version minv21 herausbringt, und nun fehlen 3546 Werte – ein Nettogewinn von 362 Werten, die importiert/ausgefüllt wurden. In der Version minv22 fehlen 3571 Werte, und 25 Werte sind jetzt nicht mehr verfügbar.

Wenn man bedenkt, dass die BOM kein Problem mit dem massiven Datenimport von 37 Jahren hatte, wie in Port Macquarie, wo es keine Daten gab, wurden die Daten zusammen mit Ausreißern und Trends und ausgeprägten Monaten erstellt (zweifellos gemittelt aus der Liste der 40 Stationen, nach denen die BOM-Software sucht). Es ist fast so, als würden die fehlenden/hinzugefügten Werte die Hypothese unterstützen.

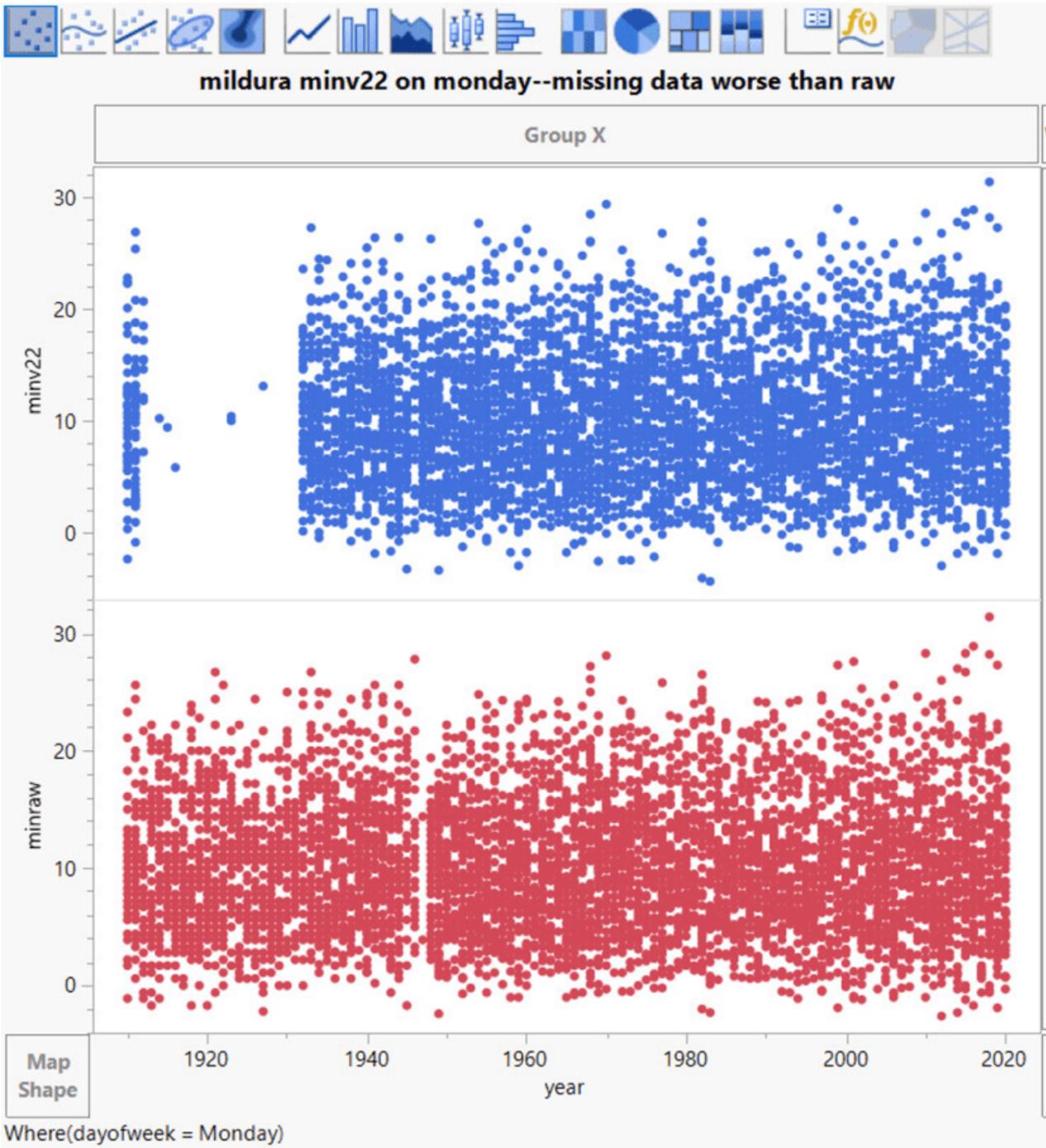
Daten, die NICHT zufällig fehlen, werden gemeinhin als MNAR (missing not at random) bezeichnet und führen zu einer Verzerrung der Daten.

So wie hier von der Station Moree, NSW:

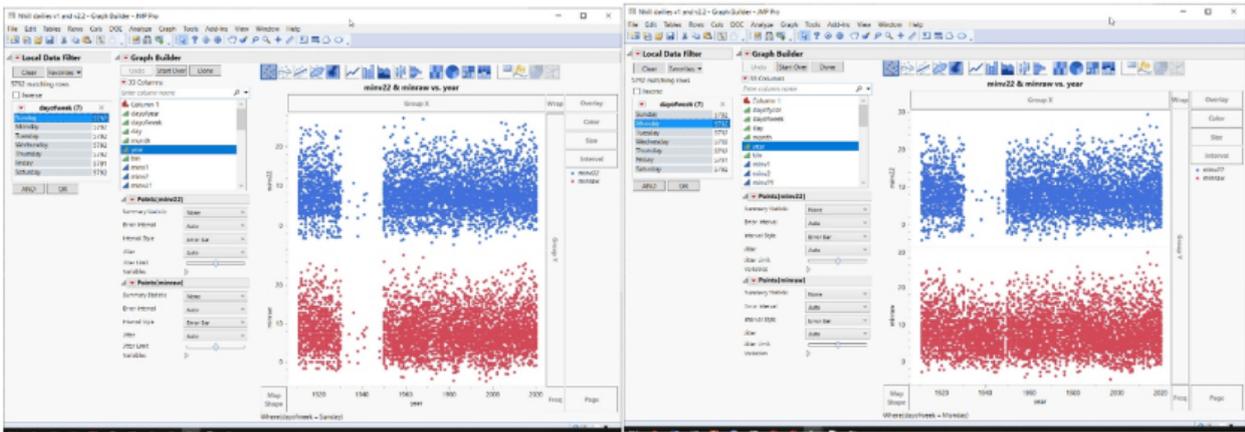


Ein Drittel der Zeitreihe verschwindet an einem Sonntag! Den Rest der Woche sind die Werte wieder da. Die Version minv1 ist dauerhaft gelöscht – offenbar gefiel dem BOM nach der Erstellung der Rohdaten das erste Drittel der Daten für minv1 nicht und es löschte es. Dann haben sie es sich mit allen anderen Versionen anders überlegt und Daten erstellt – außer für den Sonntag natürlich.

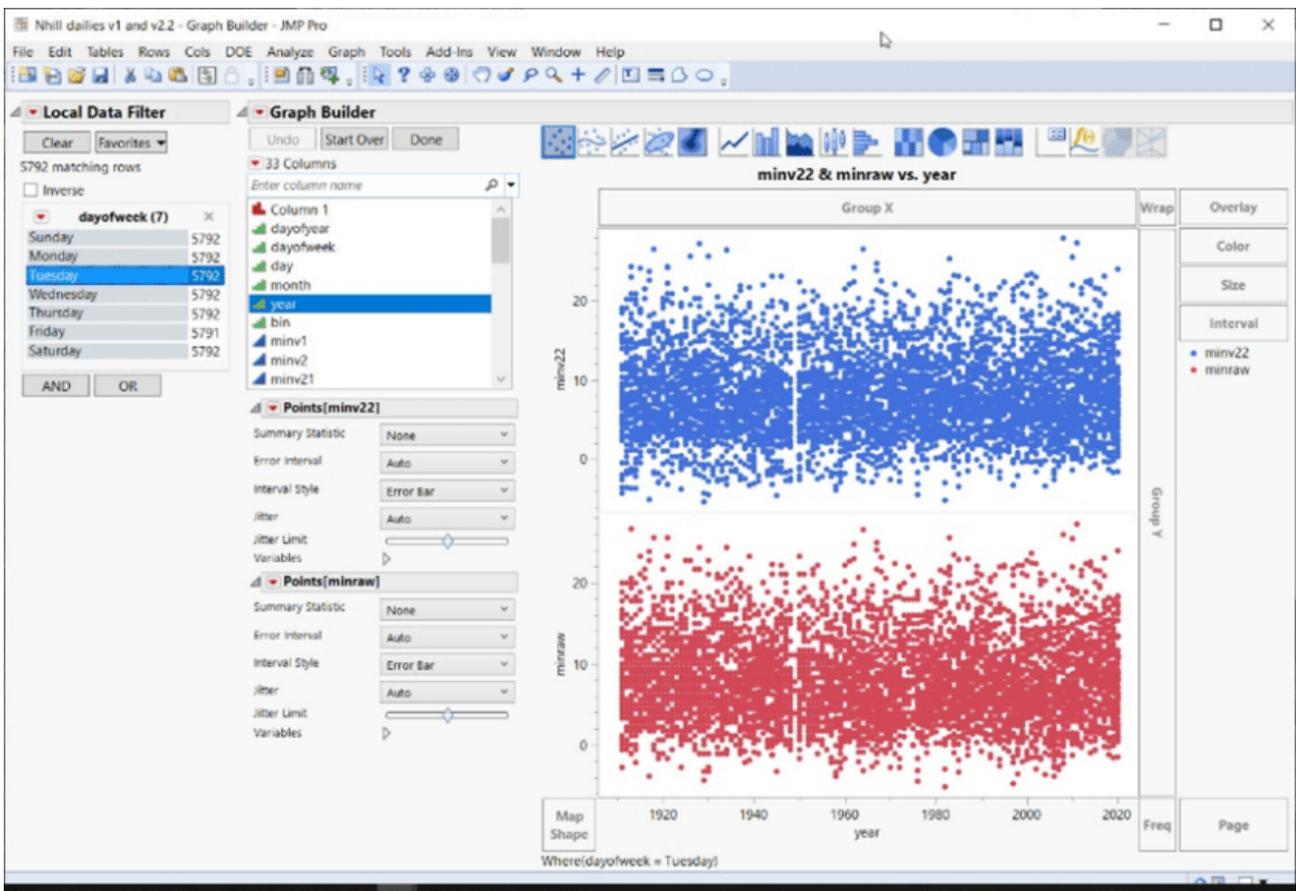
Mildura möchte jedoch anders sein und führt den Montag als den Tag, an dem die Daten auf magische Weise verschwinden:



Nhill an einem Sonntag (unten links) ist anders als an einem Montag (unten rechts) ...:



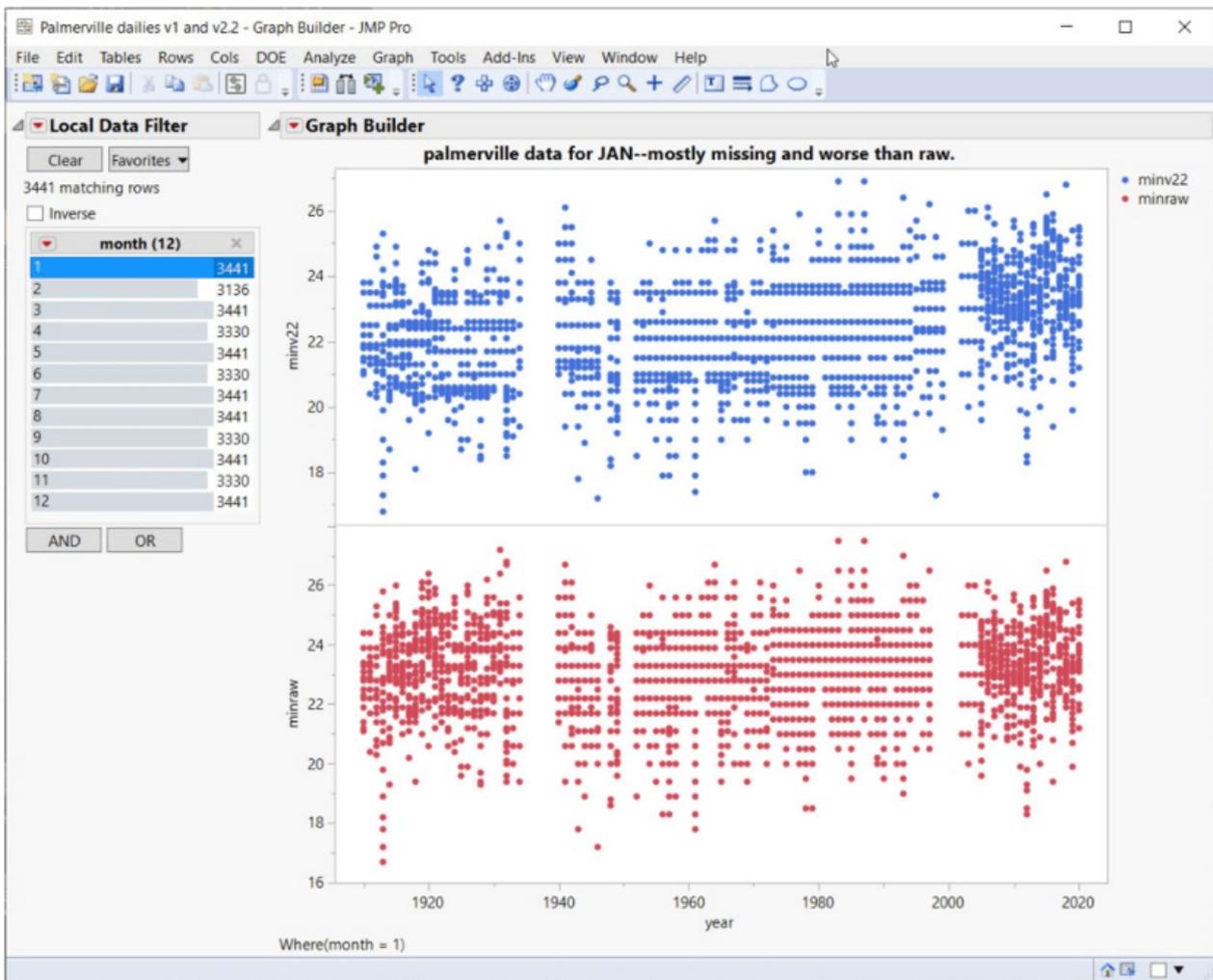
Das ist etwas anderes als Dienstag – Samstag (unten). Wie Sie sehen, gibt es allerdings immer noch eine Lücke bei den Daten um 1950. Aber warum sollte man eine gute Sache mit all diesen Daten-Importen erzwingen; wenn man einige Lücken lässt, wirkt das Ganze authentischer.



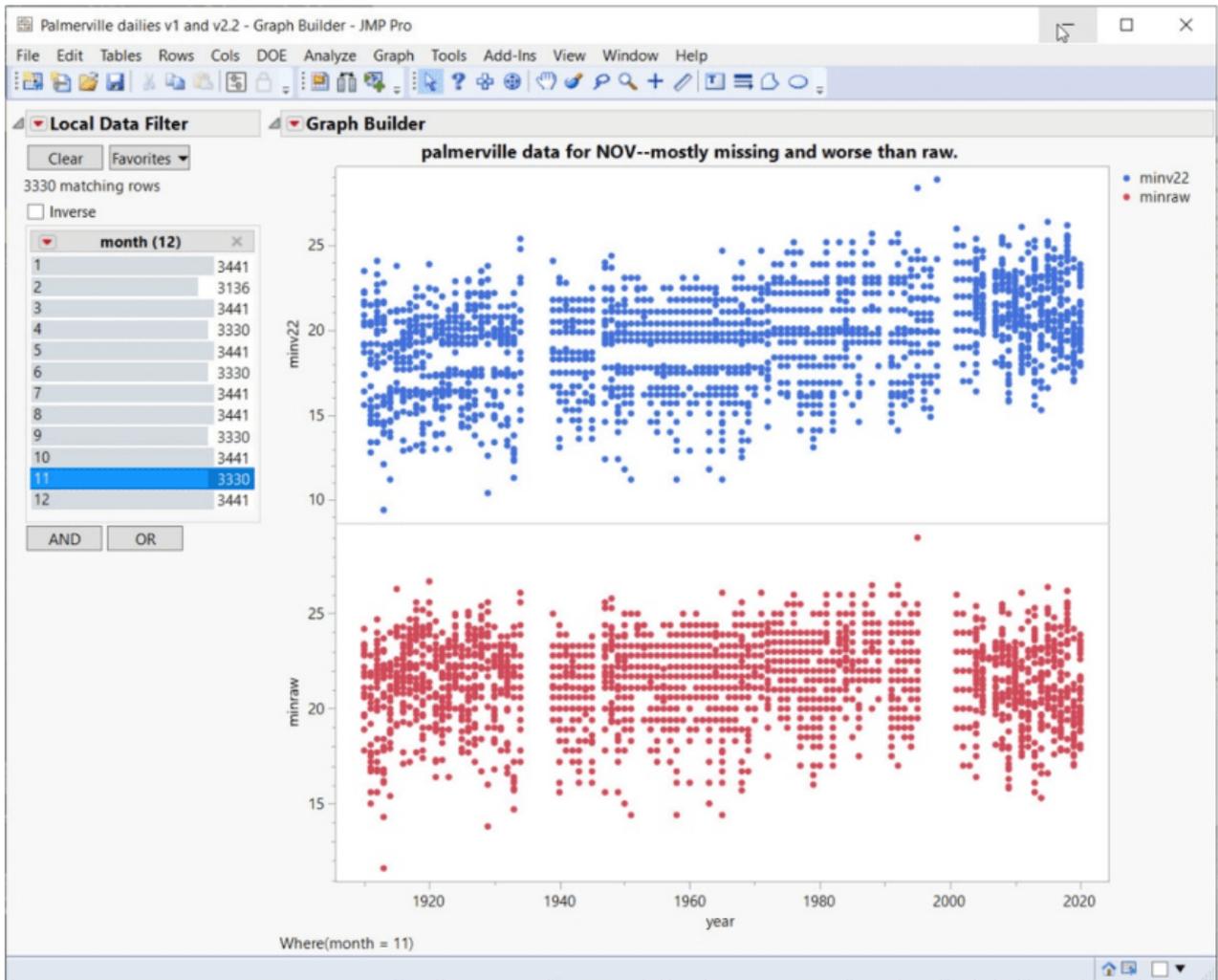
FEHLENDE TEMPERATURBEREICHE

Streudiagramme gehören zu den ersten Dingen, die wir bei der Datenexploration machen. Bei den BOM-Daten fällt als erstes auf, dass je nach Monat, Jahrzehnt oder Wochentag einige Daten fehlen. Komplette Temperaturbereiche können für viele, ja sogar für die meisten der 110-jährigen ACORN-SAT-Zeitreihen fehlen.

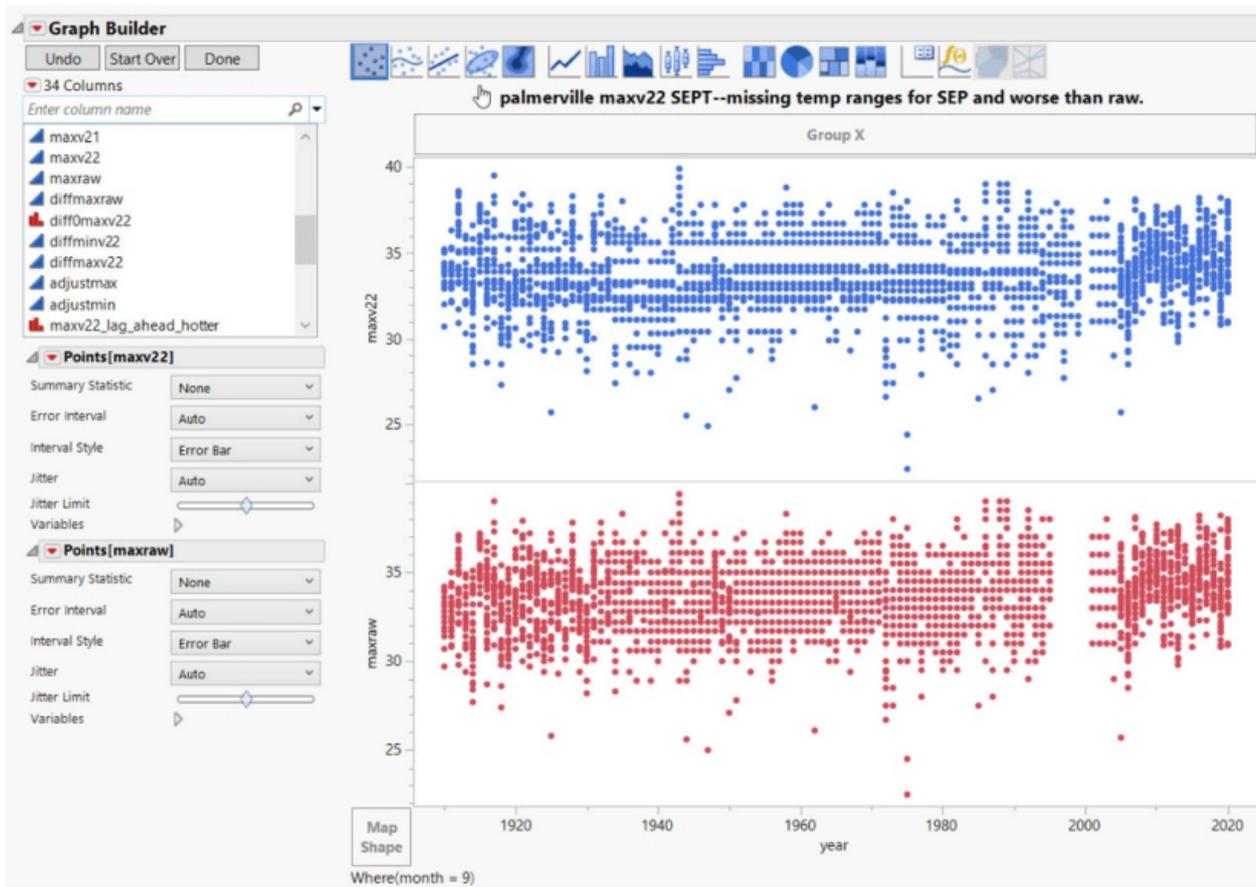
Zum Beispiel Palmerville:



Die langen horizontalen Lücken oder „Korridore“ mit fehlenden Daten zeigen vollständige Temperaturbereiche, die für die meisten Zeitreihen nicht existieren. Hier dauert es bis etwa 2005, bis alle Temperaturbereiche angezeigt werden. Außerdem ist zu beachten, dass Anpassungen, die auf die Rohdaten folgen, die Situation noch verschlimmern:



Gleiches gilt für den November:



Die beunruhigende Tendenz dabei ist, dass Anpassungen die Situation oft noch verschlimmern – Daten fehlen, Temperaturbereiche fehlen, falsche Sequenzen werden eingeführt:

Sydney dailies v1 and v2.2 2 - JMP Pro

	maxv2	maxv21	maxv22	diff10ormore	diffmaxv22	maxraw	adjustmax	a
4907	21.4	21.4	21.3	0	1	20.2	1.1	
4908	19.1	19.1	19.1	0	-2.2	17.9	1.2	
4909	19.1	19.1	19.1	0	0	17.6	1.5	
4910	19.1	19.1	19.1	0	0	17.8	1.3	
4911	19.1	19.1	19.1	0	0	17.8	1.3	
4912	19.1	19.1	19.1	0	0	17.6	1.5	
4913	20.3	20.3	20.3	0	1.2	19.4	0.9	
4914	19.1	19.1	19.1	0	-1.2	17.9	1.2	
4915	14.6	14.6	14	0	-5.1	14.6	-0.6	
4916	19.6	19.6	19.5	0	5.5	18.7	0.8	
4917	19.1	19.1	19.1	0	-0.4	17.9	1.2	

Wenn man sich die zuerst ertellten Rohdaten ansieht, kann man erkennen, wie eine duplizierte Sequenz nach „Anpassungen“ entstanden ist. Das Ergebnis ist eine offensichtlich verfälschte Temperatursequenz.

Im nachstehenden Beispiel aus Charleville, bei dem 1942 mit 1943 verglichen wird, ist zu erkennen, dass ein einzelner Rohwert aus dem Jahr 1942 aus dem Jahr 1943 gelöscht und ein einzelner minv2-Wert zwischen den Jahren geringfügig geändert wurde (...), aber eine lange duplizierte Sequenz wurde so belassen, wie sie war:

day	month	year	maxv2	maxraw	minv2	minraw	day	month	year	maxv2	maxraw	minv2	minraw
1	12	1942	36.3	38.3	19.6	21.1	1	12	1943	23.6	•	19.6	21.1
2	12	1942	34	35.6	18.7	20	2	12	1943	34	35.6	18.7	20
3	12	1942	34	35.6	19.1	20.6	3	12	1943	34	35.6	19.1	20.6
4	12	1942	31.9	33.9	18.7	20	4	12	1943	31.9	33.9	18.7	20
5	12	1942	29.3	31.7	18.3	19.4	5	12	1943	29.3	31.7	18.3	19.4
6	12	1942	28.6	31	18.3	19.4	6	12	1943	28.6	31	18.3	19.4
7	12	1942	27.2	29.7	18.7	20	7	12	1943	27.2	29.7	18.7	20
8	12	1942	27.7	30	17.9	18.9	8	12	1943	27.7	30	17.9	18.9
9	12	1942	21.4	25.3	18.7	20	9	12	1943	21.4	25.3	18.7	20
10	12	1942	25.9	28.9	17.6	18.3	10	12	1943	25.9	28.9	17.6	18.3
11	12	1942	28.3	30.8	18.3	19.4	11	12	1943	28.3	30.8	18.3	19.4
12	12	1942	24.6	27.9	18.3	19.4	12	12	1943	24.6	27.9	18.3	19.4
13	12	1942	25.1	28.3	16	17.2	13	12	1943	25.1	28.3	16	17.2
14	12	1942	•	•	16	17.2	14	12	1943	•	•	16	17.2
15	12	1942	•	33.3	•	•	15	12	1943	•	33.3	•	•
16	12	1942	29.8	32.2	•	17.5	16	12	1943	29.8	32.2	•	17.5
17	12	1942	30.7	32.8	19.3	20.8	17	12	1943	30.7	•	19.5	•
18	12	1942	31.5	33.3	18	19.2	18	12	1943	31.5	•	18	19.2
19	12	1942	32.3	34.2	18.3	19.4	19	12	1943	32.3	34.2	18.3	19.4
20	12	1942	29.8	32.2	19.6	21.1	20	12	1943	29.8	32.2	19.6	21.1
21	12	1942	31.9	33.9	17.9	18.9	21	12	1943	31.9	33.9	17.9	18.9
22	12	1942	31.5	33.3	16.8	17.8	22	12	1943	31.5	33.3	16.8	17.8
23	12	1942	33.8	35.4	18.3	19.4	23	12	1943	33.8	35.4	18.3	19.4
24	12	1942	34.8	36.1	18.3	19.4	24	12	1943	34.8	36.1	18.3	19.4
25	12	1942	36.2	37.8	19.3	20.8	25	12	1943	36.2	37.8	19.3	20.8
26	12	1942	32.4	34.4	21.5	22.8	26	12	1943	32.4	34.4	21.5	22.8
27	12	1942	25.9	28.9	24.6	26.7	27	12	1943	25.9	28.9	24.6	26.7
28	12	1942	30.7	32.8	18.7	20	28	12	1943	30.7	32.8	18.7	20
29	12	1942	34.8	36.1	21	22.2	29	12	1943	34.8	36.1	21	22.2
30	12	1942	29.8	32.2	21.5	22.8	30	12	1943	29.8	32.2	21.5	22.8
31	12	1942	30.7	32.8	21.5	22.8	31	12	1943	30.7	32.8	21.5	22.8

Es gibt zwei Arten von Sequenzen, die hier von Interesse sind:

- 1 – Eine Reihe von duplizierten Temperaturen.
- 2 – Eine Temperatursequenz, die an anderer Stelle dupliziert wird.

Es gibt Hunderte und Aberhunderte dieser Sequenzen über den größten Teil der ACORN-SAT-Zeitreihe. In Sydney wurden sogar zweieinhalb Monate in ein anderes Jahr kopiert/eingefügt, wie wir gesehen haben. Computer wählen nicht „zufällig“ Daten aus, die nur einen Kalendermonat lang sind, immer und immer wieder – das machen Menschen.

Die Wahrscheinlichkeit, dass diese Sequenzen zufällig auftreten, wird von der SAS JMP-Software als unwahrscheinlich bis unmöglich berechnet. Selbst die Simulation von Daten mit Autokorrelation unter Verwendung der Block-Bootstrap-Methode zeigt die Unmöglichkeit der Sequenzen.

Das Ergebnis sind gefälschte Temperaturverläufe, die die meisten Zifferntests für Beobachtungsdaten nicht bestehen, wie z. B. Benford's Law für Anomalien, Simonsohn's Number Bunching Tests (www.datacolada.com) und Ensminger+Leder-Luis Bank of Digit Tests (Measuring Strategic Data Manipulation: Evidence from a World Bank

Project von Jean Ensminger und Jetson Leder-Luis)

Beispiele aus Palmerville finden Sie unten. Die duplizierten Sequenzen auf der linken Seite sind eine sehr hinterhältige Art der Erwärmung – eine wärmere Monatssequenz wird in einen kälteren Monat kopiert.

Das BOM hat offensichtlich Schwierigkeiten, selbst im Jahr 2004 saubere Daten zu erhalten:

Column 1	dayofweek	day	month	year	minv1	minv2	minv21	minraw	maxdiff	mindiff	Column 1	dayofweek	day	month	year	maxv1	maxv2	maxv21		
1	2002-03-19	Tuesday	19	3	2002	20.9	21.0	21.0	21.0	0	0	1	2004-04-21	Wednesday	21	4	2004	31.0	31.0	31.0
2	2002-03-20	Wednesday	20	3	2002	20.9	21.0	21.0	21.0	0	0	2	2004-04-22	Thursday	22	4	2004	31.0	31.0	31.0
3	2002-03-21	Thursday	21	3	2002	18.9	19.0	19.0	19.0	0	0	3	2004-04-23	Friday	23	4	2004	31.0	31.0	31.0
4	2002-03-22	Friday	22	3	2002	18.9	19.0	19.0	19.0	0	0	4	2004-04-24	Saturday	24	4	2004	31.0	31.0	31.0
5	2002-03-23	Saturday	23	3	2002	19.7	20.0	20.0	20.0	0	0	5	2004-04-25	Sunday	25	4	2004	31.0	31.0	31.0
6	2002-03-24	Sunday	24	3	2002	20.9	21.0	21.0	21.0	0	0	6	2004-04-26	Monday	26	4	2004	31.0	31.0	31.0
7	2002-03-25	Monday	25	3	2002	18.9	19.0	19.0	19.0	0	0	7	2004-04-27	Tuesday	27	4	2004	31.0	31.0	31.0
8	2002-03-26	Tuesday	26	3	2002	18.9	19.0	19.0	19.0	0	0	8	2004-04-28	Wednesday	28	4	2004	31.0	31.0	31.0
9	2004-05-05	Wednesday	5	5	2004	21.0	21.0	21.0	21.0	0	0	9	2004-04-29	Thursday	29	4	2004	31.0	31.0	31.0
10	2004-05-06	Thursday	6	5	2004	21.0	21.0	21.0	21.0	0	0	10	2004-04-30	Friday	30	4	2004	31.0	31.0	31.0
11	2004-05-07	Friday	7	5	2004	19.0	19.0	19.0	19.0	0	0	11	2004-05-01	Saturday	1	5	2004	31.0	31.0	31.0
12	2004-05-08	Saturday	8	5	2004	19.0	19.0	19.0	19.0	0	0	12	2004-05-02	Sunday	2	5	2004	31.0	31.0	31.0
13	2004-05-09	Sunday	9	5	2004	20.0	20.0	20.0	20.0	0	0	13	2004-05-03	Monday	3	5	2004	31.0	31.0	31.0
14	2004-05-10	Monday	10	5	2004	21.0	21.0	21.0	21.0	0	0	14	2004-05-04	Tuesday	4	5	2004	31.0	31.0	31.0
15	2004-05-11	Tuesday	11	5	2004	19.0	19.0	19.0	19.0	0	0	15	2004-05-05	Wednesday	5	5	2004	31.0	31.0	31.0
16	2004-05-12	Wednesday	12	5	2004	19.0	19.0	19.0	19.0	0	0	16	2004-05-06	Thursday	6	5	2004	31.0	31.0	31.0
												17	2004-05-07	Friday	7	5	2004	31.0	31.0	31.0

Palmerville min temps.
 8 day sequence in march is duplicated into may.
 Since March is warmer than May, it effectively creates warming.
 Prob of seeing this duplicated sequence at random is equal to a rarity of 36 -- equal to 36 heads in a row in a coin toss.

17 day duplicate sequence at Palmerville.
 Probability of seeing this at random has a rarity of 100 which is equal to 100 heads in a coin tossing.

Palmerville gehört zu den schlechtesten Acorn-Standorten, was die Datenreinheit angeht. Sehr viele der Temperaturbeobachtungen sind ganzzahlig, mit keiner Nachkommastelle. Aus dem Stationskatalog von ACORN-SAT:

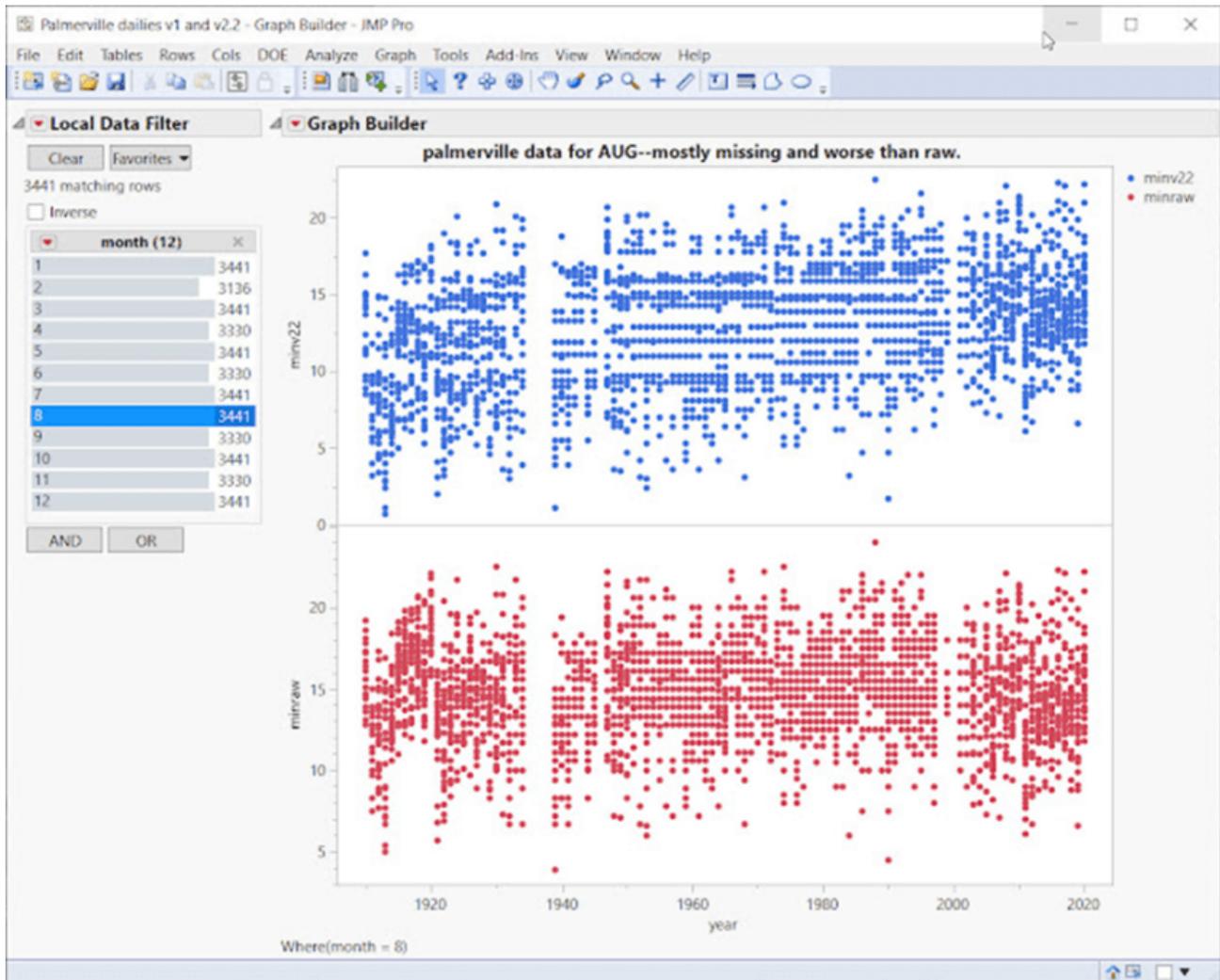
„Der gleiche Beobachter hat 63 Jahre lang, von 1936 bis 1999, Beobachtungen durchgeführt und war während des letzten Teils dieses Zeitraums der einzige verbliebene Bewohner des Gebiets. Nachdem die manuellen Beobachtungen bei ihrem Weggang 1999 eingestellt wurden, wurden die Beobachtungen erst mit der Inbetriebnahme der automatischen Station wieder aufgenommen.“

Die Stationen in Melbourne Regional (ca. 1856-2014) und Sydney Observatory (ca. 1857-Okt 2017) werden hier als die besten der BOM-Qualität angesehen. [\(Link\)](#)

Temperaturbereiche fehlen auch über viele Jahre hinweg, sie erscheinen in vielen Fällen erst in den 2000er Jahren. Und das kann auch nach Anpassungen passieren.

Im Folgenden bleiben wir bei Palmerville für alle Augustmonate von 1910 bis 2020. Dazu verwenden wir die grundlegendste aller Datenanalysegrafiken, das Streudiagramm. Dies ist eine Datendarstellung,

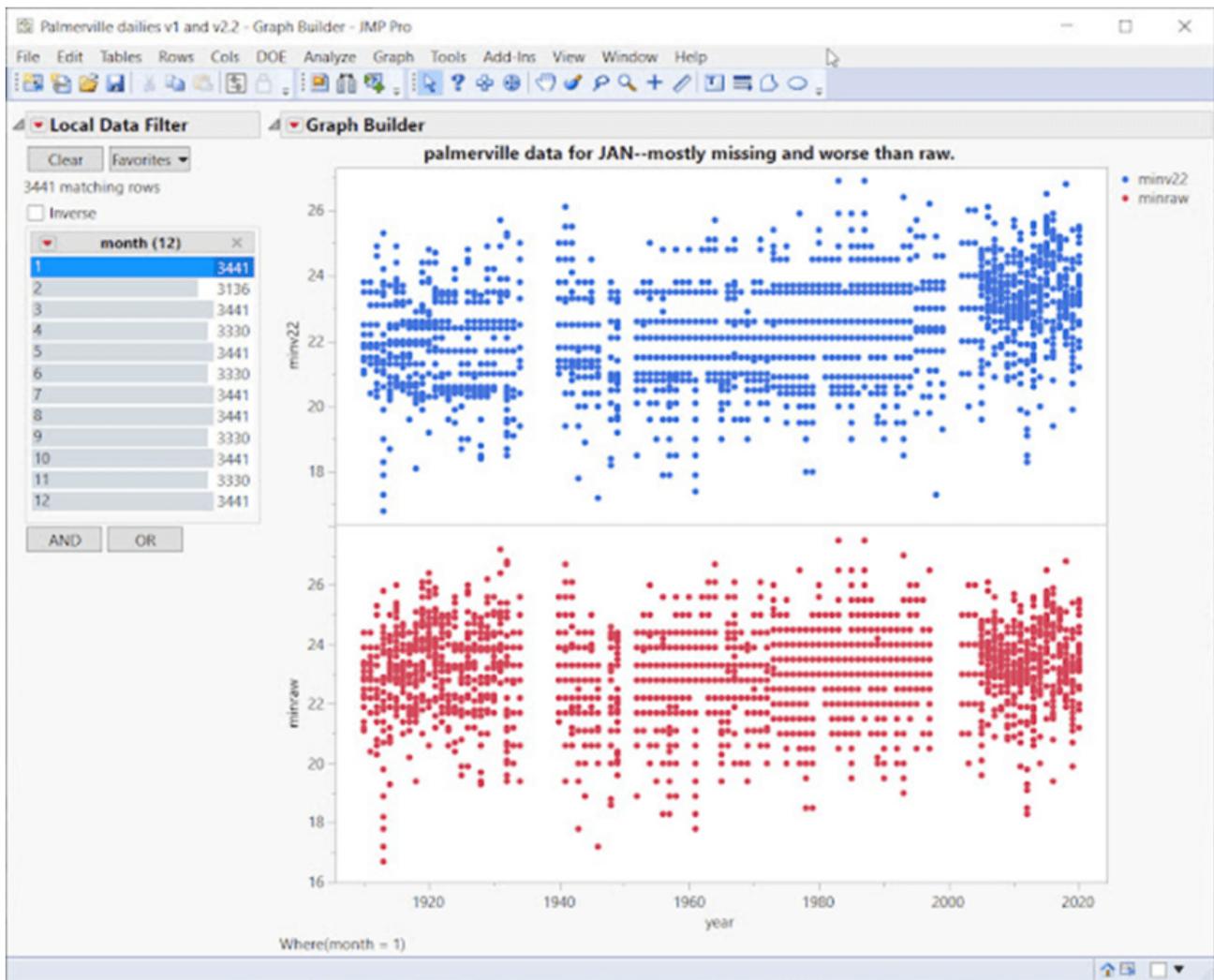
die die Beziehung zwischen zwei numerischen Variablen zeigt:



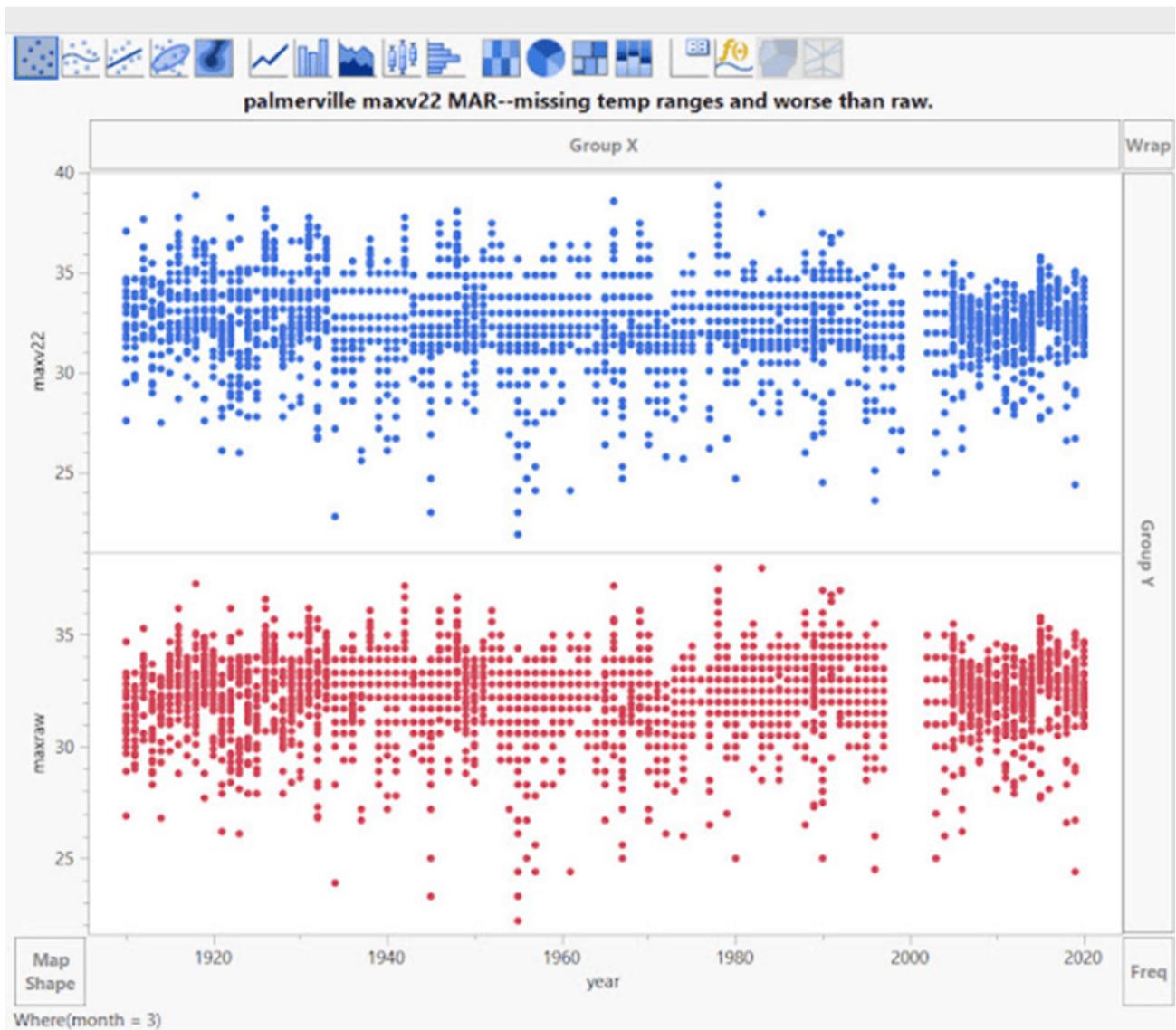
Oben – Dies ist eine vollständige Datenansicht (Streudiagramm) der gesamten Zeitreihe, minraw und minv22. Raw kam zuerst in der Zeit (unten in rot), so dass dies unsere Referenz ist. Nach den Anpassungen von minv22 sind ganze Bereiche verschwunden, die horizontale „Rinnen“ zeigen fehlende Temperaturen, die nie auftauchen. Selbst im Jahr 2000 sieht man horizontale Lücken, in denen Dezimalwerte verschwunden sind, so dass man nur ganzzahlige Temperaturen wie 15°C, 16°C usw. erhält.

Im Katalog der BOM-Station ist vermerkt, dass der Beobachter 1999 abreiste, und für die AWS gilt: „Die automatische Wetterstation wurde Mitte 2000 160 m südöstlich des früheren Standorts installiert, begann aber erst im Juli 2001 mit der Übertragung von Daten.“ Zwischen diesen Daten wurden also offenbar keine Daten gesammelt. Woher stammen die Daten, die im obigen Diagramm rot dargestellt sind?

Die Rohdaten wurden viermal mit vier Versionen der modernsten BOM-Software angepasst, und das ist das Ergebnis – ein schlechteres Ergebnis:

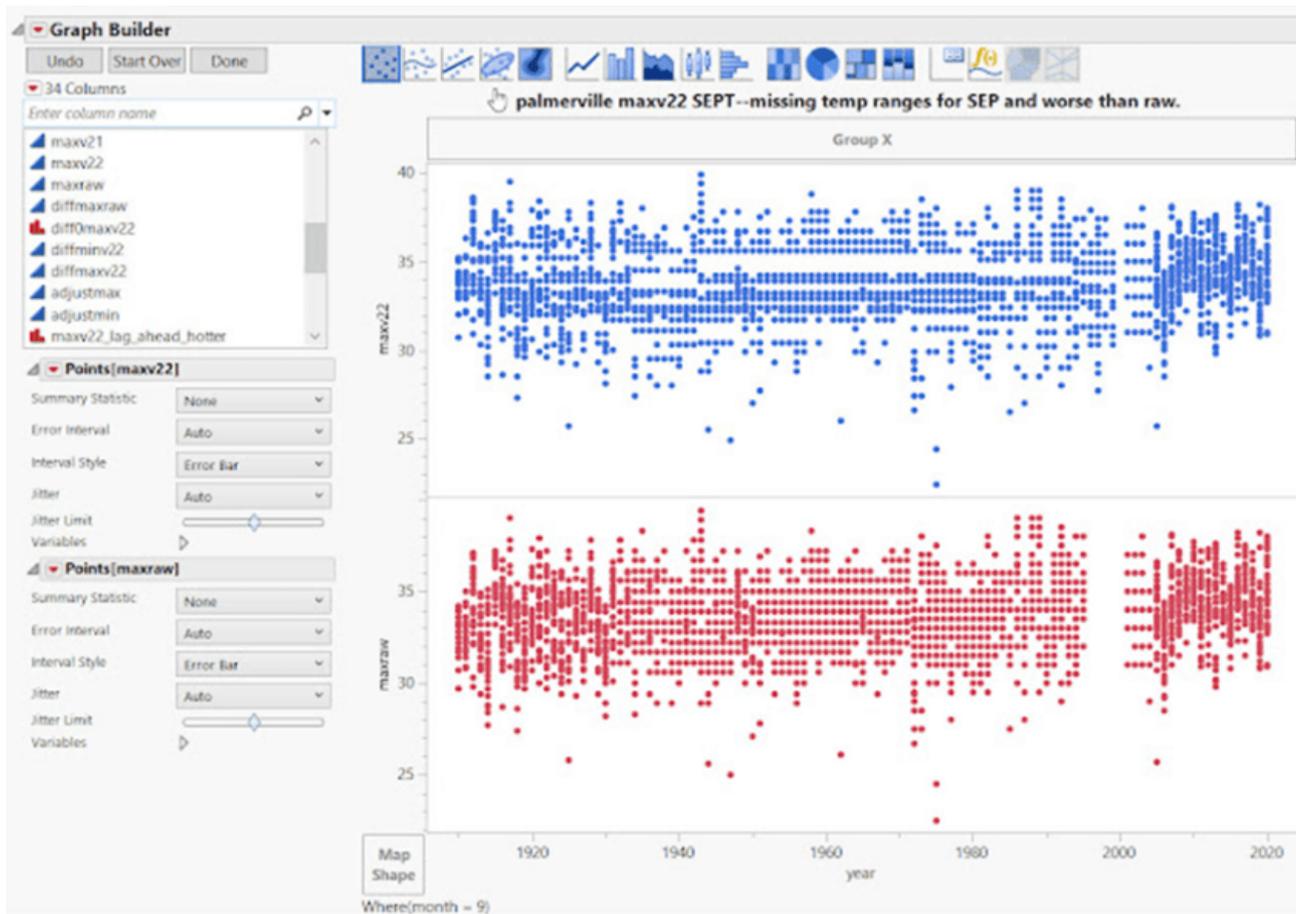


Für den Januar gibt es keine sauberen Daten, massive „Korridore“ mit fehlenden Temperaturbereichen bis 2005 oder so. Auch hier sind die Daten nach den Anpassungen schlechter:



Auch die Märzdaten sind nach den Anpassungen schlechter. Sie hatten ein echtes Problem mit den Temperaturen von etwa 1998-2005.

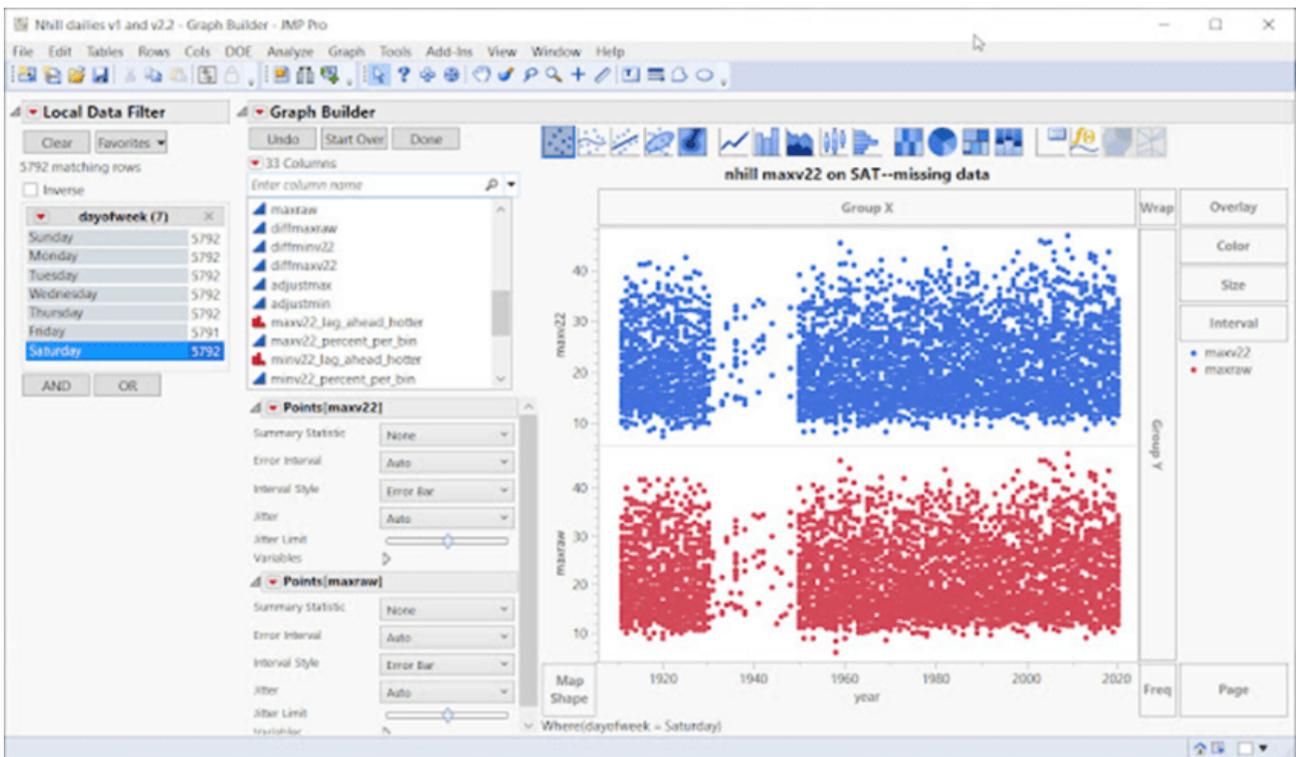
Unten – Sehen Sie sich die Daten *vor und nach den Anpassungen* an. Dies sind sehr schlechte Datenverarbeitungsverfahren, und es ist nicht zufällig, also erwarten Sie nicht, dass diese Art von Manipulation Fehler ausgleicht:



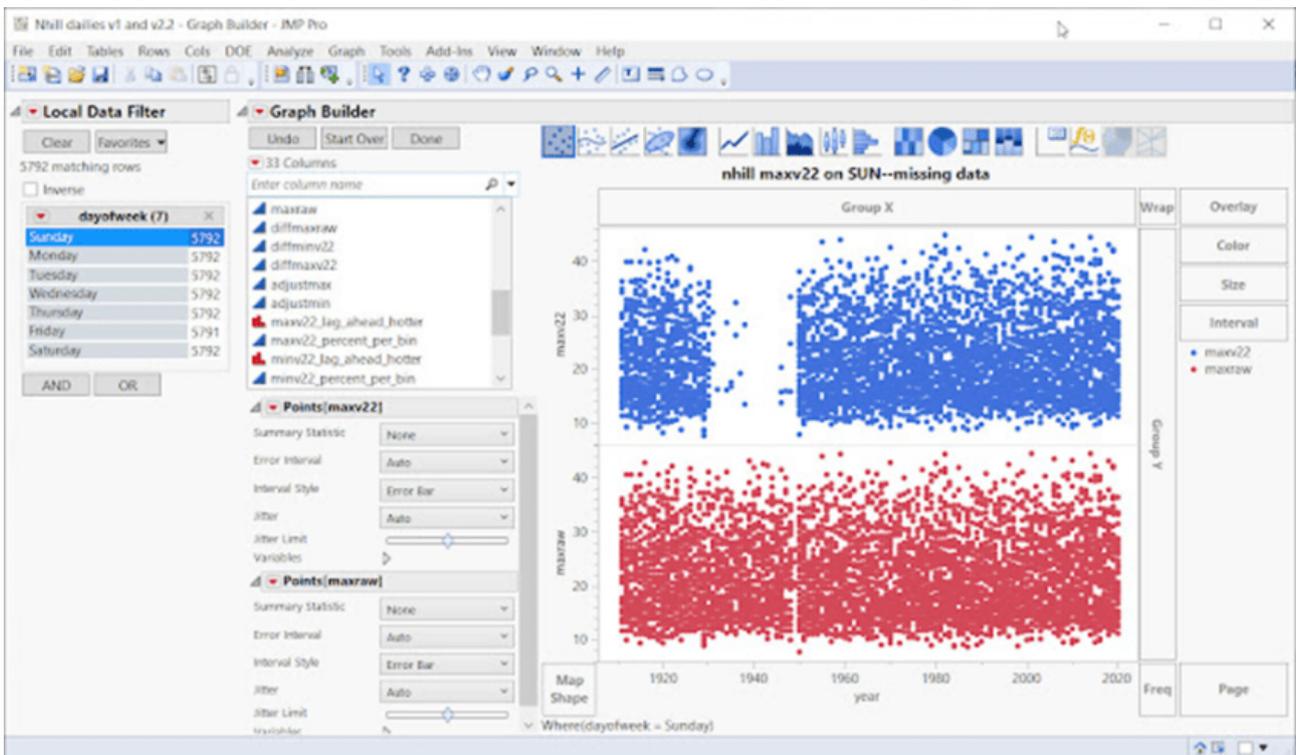
SONNTAG IN NHILL = FEHLENDE DATEN NICHT ZUFÄLLIG

EINE VERZERRUNG ENTSTEHT DURCH NICHT ZUFÄLLIG FEHLENDE DATEN ([Link](#)).

Unten – Nhill an einem Samstag hat einen großen Anteil an fehlenden Daten, sowohl bei den Rohdaten als auch bei den angepassten Daten:



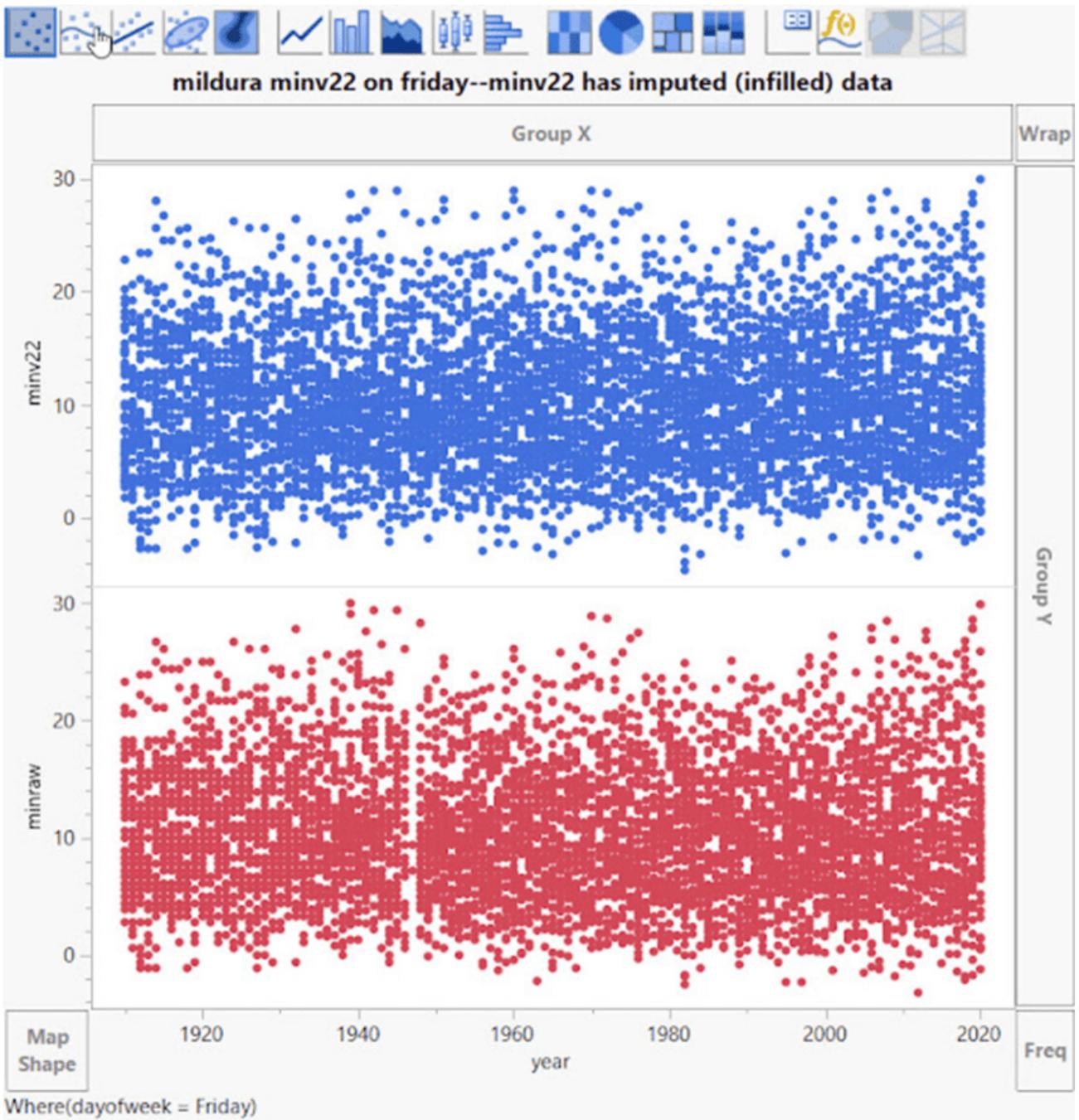
Unten: Am **Sonntag**, voila ... gibt es jetzt Tausende von Rohtemperaturen, aber es fehlen immer noch die angepassten Daten:



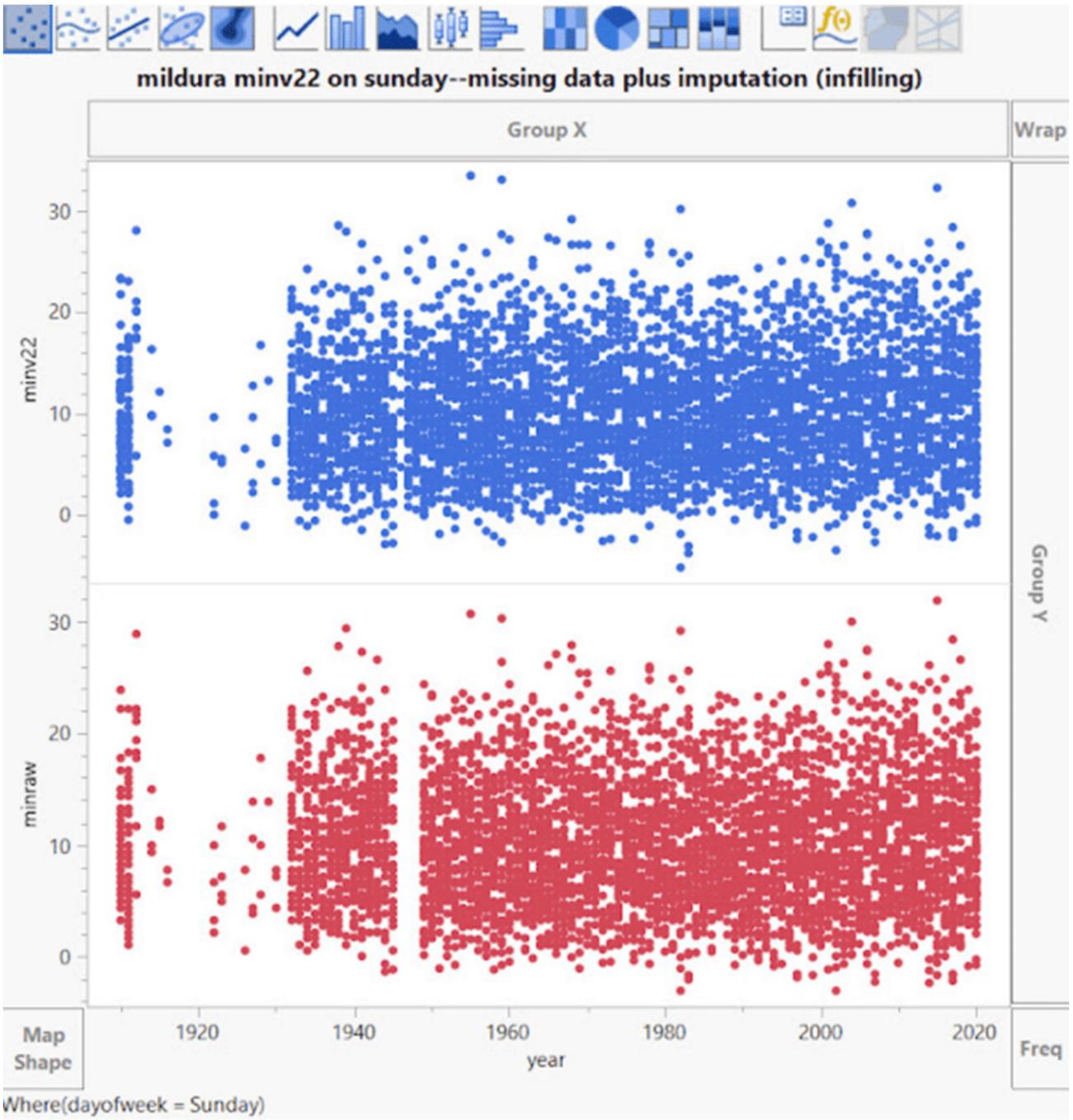
Unten – Moment, da ist noch mehr – jetzt ist es **Montag**, und einfach so erscheinen Tausende von angepassten Temperaturen!



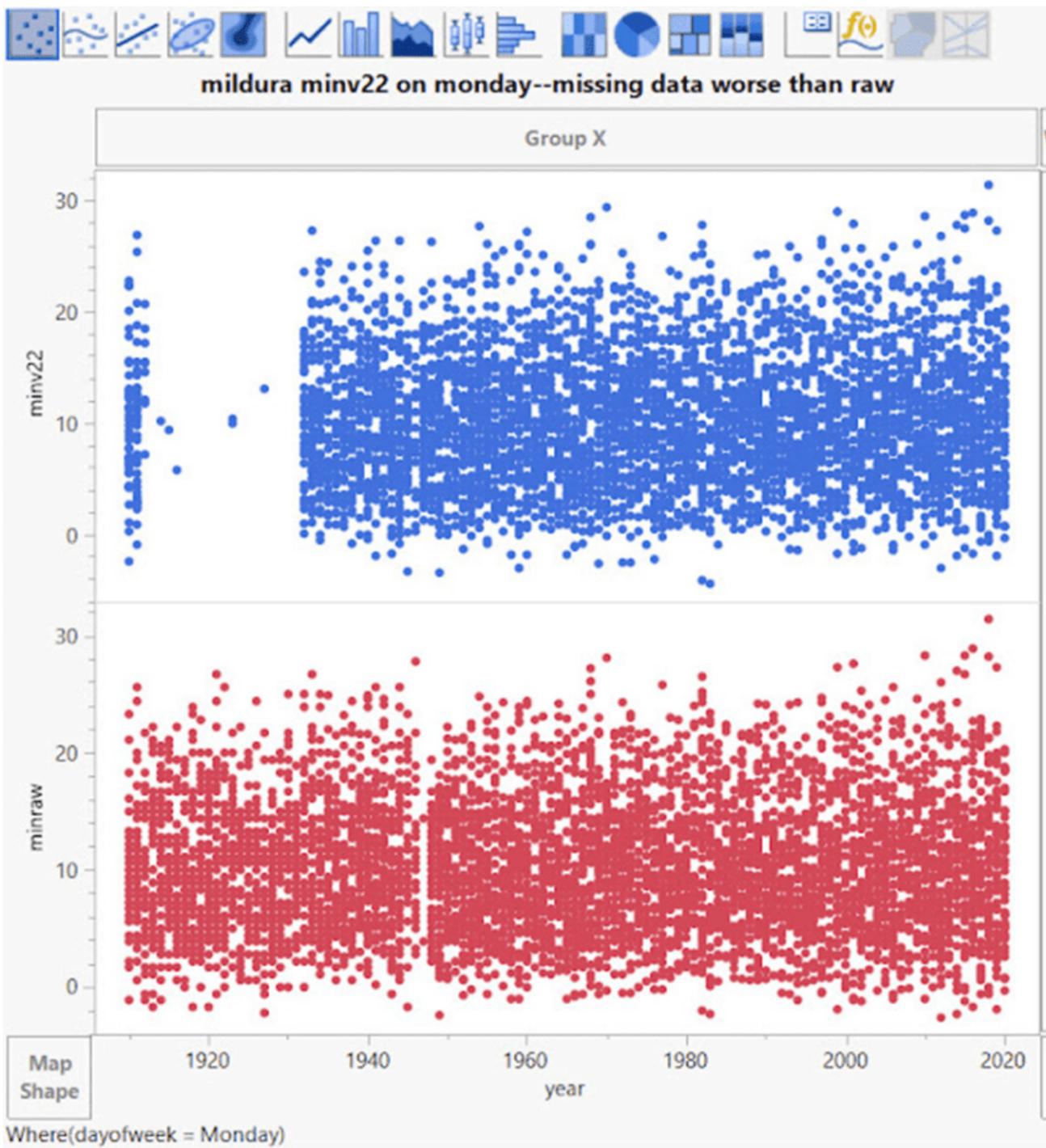
Mildura am **Freitag**:



Mildura am **Sonntag**:



Mildura am Montag:



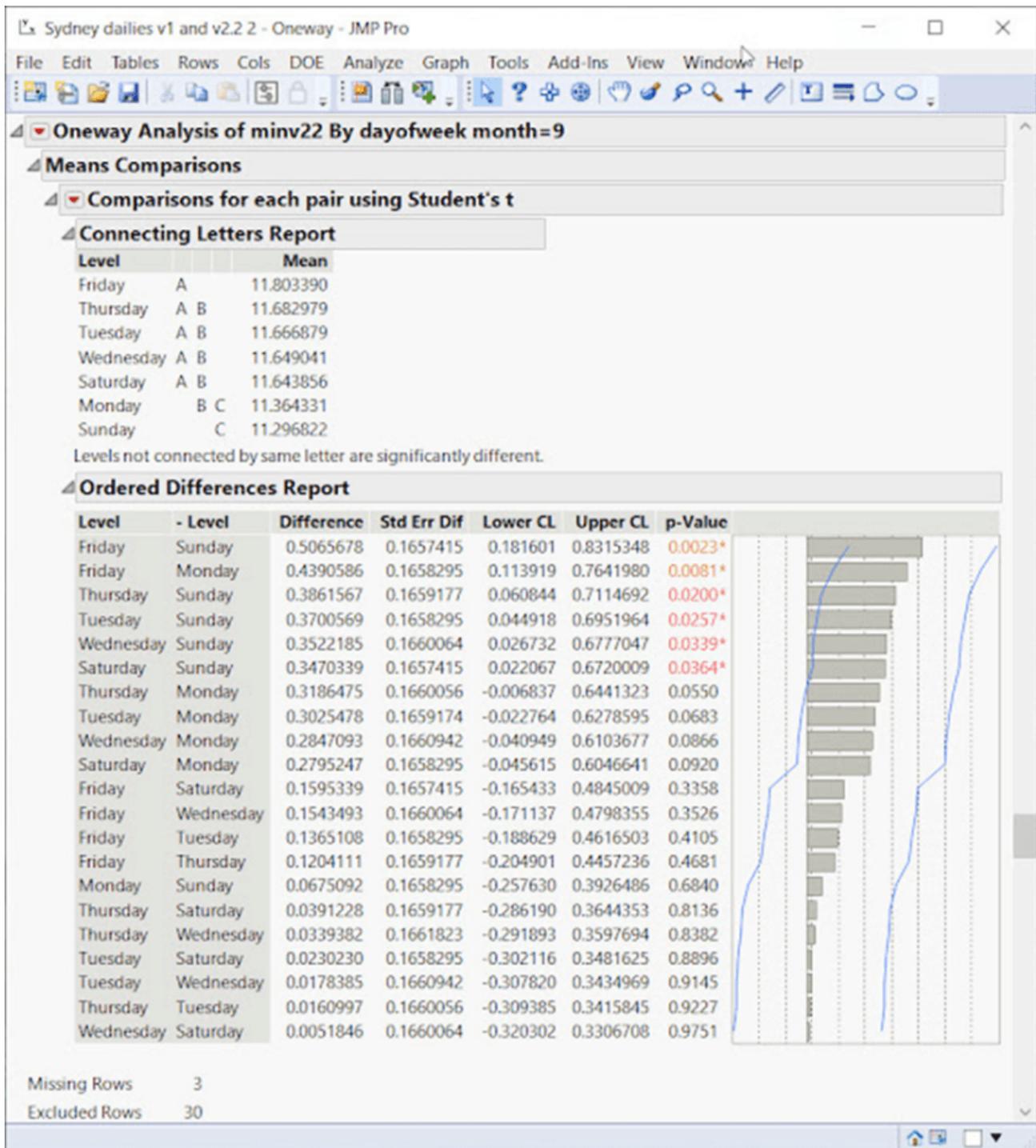
Oben – Am Montag verschwindet ein großes Stück in den angepassten Daten. Seltsamerweise wird der dünne Streifen fehlender Rohdaten bei etwa 1947 in minv2.2 aufgefüllt.

Diese Art der Datenverarbeitung deutet auf viele andere Probleme der Verzerrung hin.

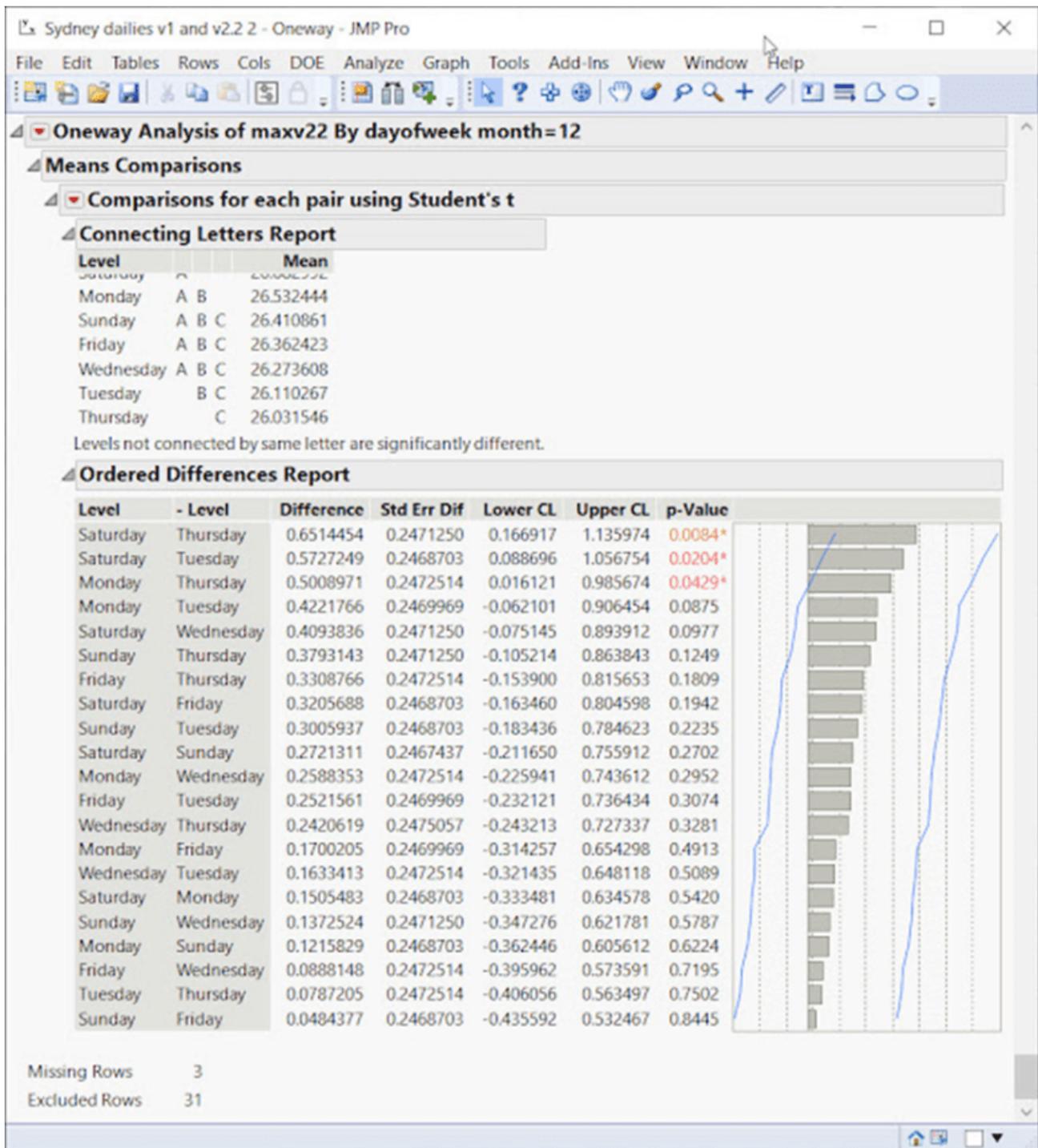
SYDNEY-WOCHENTAG-EFFEKT

Betrachtet man alle Septembermonate der Sydney-Zeitreihe von 1910-2020, so zeigt sich, dass der Freitag eine deutlich andere Temperatur aufweist als der Sonntag und der Montag.

Die Wahrscheinlichkeit, dass dies zufällig auftritt, liegt bei über 1000:1:



Samstag ist ebenfalls wärmer als Donnerstag im Dezember, das ist hoch signifikant:

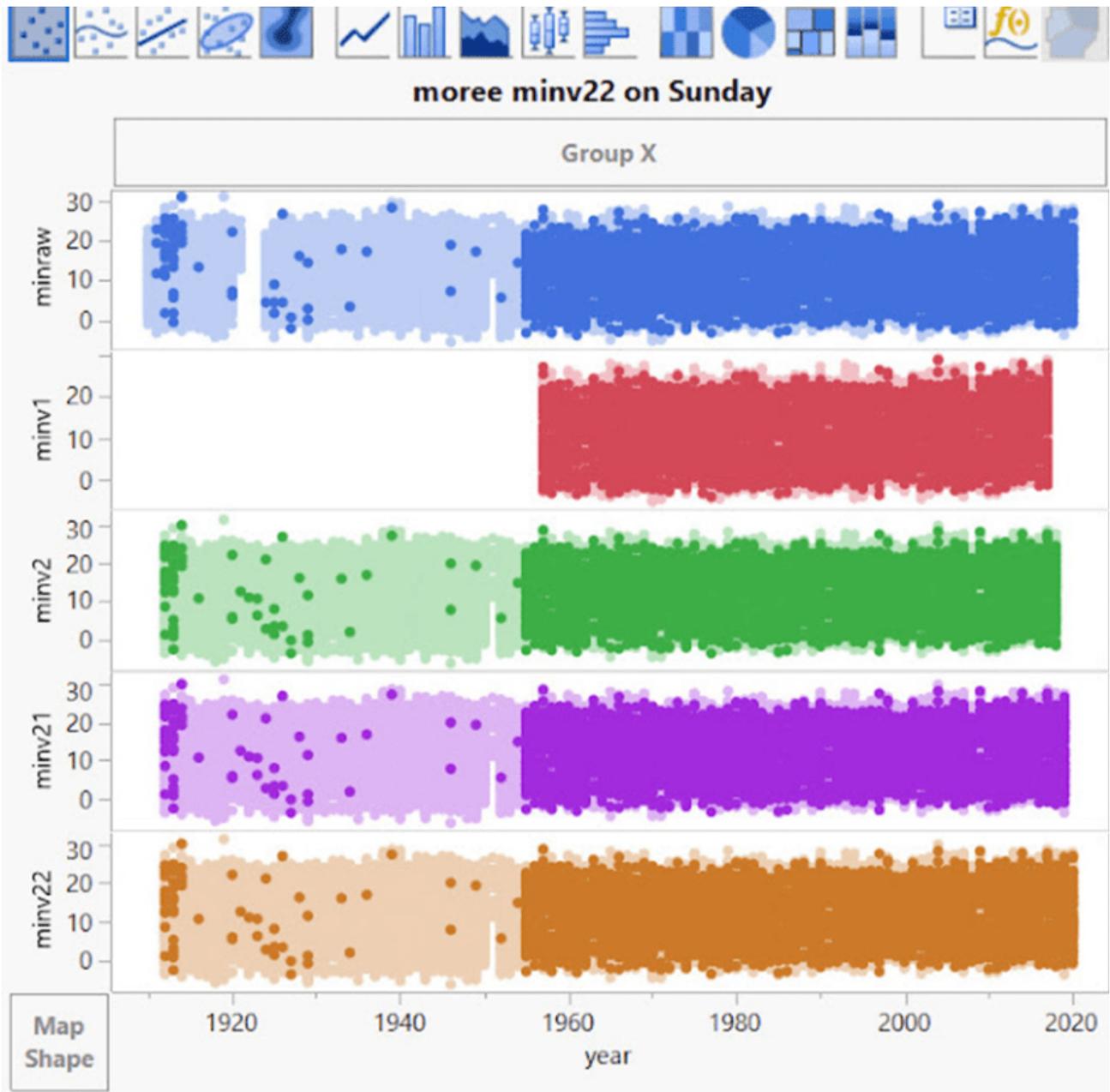


NIEMALS AN EINEM SONNTAG

Moree von Montag bis Donnerstag sieht so aus:



Unten – Aber dann kommt der Sonntag in Moree, und ein Drittel der Daten verschwindet! (bis auf ein paar ungerade Werte):



Ein Drittel der Zeitserie geht am Sonntag verloren! Es scheint, dass die griechische Filmkomödie „Niemals am Sonntag“, in der die griechische Prostituierte Ilya versucht, Homer zu verführen (aber niemals an einem Sonntag), auf Moree übergegriffen hat.

ANPASSUNGEN ERZEUGEN DOPPELTE DATENSEQUENZEN

Unten – Sydney zeigt, wie durch Anpassungen Duplikate erzeugt werden:

Die doppelten Daten werden von der BOM mit ihrer hochmodernen Anpassungssoftware erzeugt, wobei sie zu vergessen scheinen, dass es sich um Beobachtungsdaten handelt. Unterschiedliche Rohwerte werden in maxv22 zu einer Folge von doppelten Werten!

Sydney dailies v1 and v2.2.2 - JMP Pro

File Edit Tables Rows Cols DOE Analyze Graph Tools Add-Ins View Window Help

		maxv2	maxv21	maxv22	diff10ormore	diffmaxv22	maxraw	adjustmax	a
Source	4907	21.4	21.4	21.3	0	1	20.2	1.1	
Columns (33/1)	4908	19.1	19.1	19.1	0	2.2	17.9	1.2	
Column 1	4909	19.1	19.1	19.1	0	0	17.6	1.5	
dayofyear	4910	19.1	19.1	19.1	0	0	17.8	1.3	
dayofweek	4911	19.1	19.1	19.1	0	0	17.8	1.3	
day	4912	19.1	19.1	19.1	0	0	17.6	1.5	
Rows	4913	20.3	20.3	20.3	0	1.2	19.4	0.9	
All rows 40,543	4914	19.1	19.1	19.1	0	-1.2	17.9	1.2	
Selected 5	4915	14.6	14.6	14	0	-5.1	14.6	-0.6	
Excluded 366	4916	19.6	19.6	19.5	0	5.5	18.7	0.8	
Hidden 366	4917	19.1	19.1	19.1	0	-0.4	17.9	1.2	
Labeled 0									

EINE SCHLAUE ART DER ERWÄRMUNG:

Die letzten beiden Beispiele aus Palmerville, von denen eines eine raffinierte Art der Erwärmung durch Kopieren vom März und Einfügen in den Mai zeigt!:

	Column 1	dayofweek	day	month	year	minv1	minv2	minv21	minraw	maxdiff	mindiff
1	2002-03-19	Tuesday	19	3	2002	20.9	21.0	21.0	21.0	0	0
2	2002-03-20	Wednesday	20	3	2002	20.9	21.0	21.0	21.0	0	0
3	2002-03-21	Thursday	21	3	2002	18.9	19.0	19.0	19.0	0	0
4	2002-03-22	Friday	22	3	2002	18.9	19.0	19.0	19.0	0	0
5	2002-03-23	Saturday	23	3	2002	19.7	20.0	20.0	20.0	0	0
6	2002-03-24	Sunday	24	3	2002	20.9	21.0	21.0	21.0	0	0
7	2002-03-25	Monday	25	3	2002	18.9	19.0	19.0	19.0	0	0
8	2002-03-26	Tuesday	26	3	2002	18.9	19.0	19.0	19.0	0	0
9	2004-05-05	Wednesday	5	5	2004	21.0	21.0	21.0	21.0	0	0
10	2004-05-06	Thursday	6	5	2004	21.0	21.0	21.0	21.0	0	0
11	2004-05-07	Friday	7	5	2004	19.0	19.0	19.0	19.0	0	0
12	2004-05-08	Saturday	8	5	2004	19.0	19.0	19.0	19.0	0	0
13	2004-05-09	Sunday	9	5	2004	20.0	20.0	20.0	20.0	0	0
14	2004-05-10	Monday	10	5	2004	21.0	21.0	21.0	21.0	0	0
15	2004-05-11	Tuesday	11	5	2004	19.0	19.0	19.0	19.0	0	0
16	2004-05-12	Wednesday	12	5	2004	19.0	19.0	19.0	19.0	0	0

Palmerville min temps.

8 day sequence in march is duplicated into may.

Since March is warmer than May, it effectively creates warming.

Prob of seeing this duplicated sequence at random is equal to a rarity of 36 -- equal to 36 heads in a row in a coin toss.

id	Column 1	dayofweek	day	month	year	maxv1	maxv2	maxv21
1	2004-04-21	Wednesday	21	4	2004	31.0	31.0	31.0
2	2004-04-22	Thursday	22	4	2004	31.0	31.0	31.0
3	2004-04-23	Friday	23	4	2004	31.0	31.0	31.0
4	2004-04-24	Saturday	24	4	2004	31.0	31.0	31.0
5	2004-04-25	Sunday	25	4	2004	31.0	31.0	31.0
6	2004-04-26	Monday	26	4	2004	31.0	31.0	31.0
7	2004-04-27	Tuesday	27	4	2004	31.0	31.0	31.0
8	2004-04-28	Wednesday	28	4	2004	31.0	31.0	31.0
9	2004-04-29	Thursday	29	4	2004	31.0	31.0	31.0
10	2004-04-30	Friday	30	4	2004	31.0	31.0	31.0
11	2004-05-01	Saturday	1	5	2004	31.0	31.0	31.0
12	2004-05-02	Sunday	2	5	2004	31.0	31.0	31.0
13	2004-05-03	Monday	3	5	2004	31.0	31.0	31.0
14	2004-05-04	Tuesday	4	5	2004	31.0	31.0	31.0
15	2004-05-05	Wednesday	5	5	2004	31.0	31.0	31.0
16	2004-05-06	Thursday	6	5	2004	31.0	31.0	31.0
17	2004-05-07	Friday	7	5	2004	31.0	31.0	31.0

**17 day duplicate sequence at Palmerville.
Probability of seeing this at random has a rarity of
100 which is equal to 100 heads in a coin tossing.**

„Achten Sie auf unnatürliche Gruppierungen von Daten. Bei der eifrigen Suche nach veröffentlichungsfähigen Theorien – ganz gleich, wie unplausibel sie sind – ist es verlockend, die Daten so zu verändern, dass sie die Theorie besser stützen, und es ist ganz natürlich, nicht allzu genau hinzuschauen, wenn ein statistischer Test die erhoffte Antwort liefert“ – Standardabweichungen, fehlerhafte Annahmen, manipulierte Daten und andere Möglichkeiten, mit Statistiken zu lügen, Gary Smith.

„Bei dieser Art von voreingenommener Forschung suchen die Forscher nicht objektiv nach der Wahrheit, was auch immer sie sein mag, sondern versuchen vielmehr, die Wahrheit dessen zu beweisen, was sie bereits wissen, dass es wahr ist oder was wahr sein muss, um den Aktivismus für eine edle Sache zu unterstützen (Nickerson, 1998).“ – Zirkel-

DIE QUALITÄT DER BOM-ROHDATEN

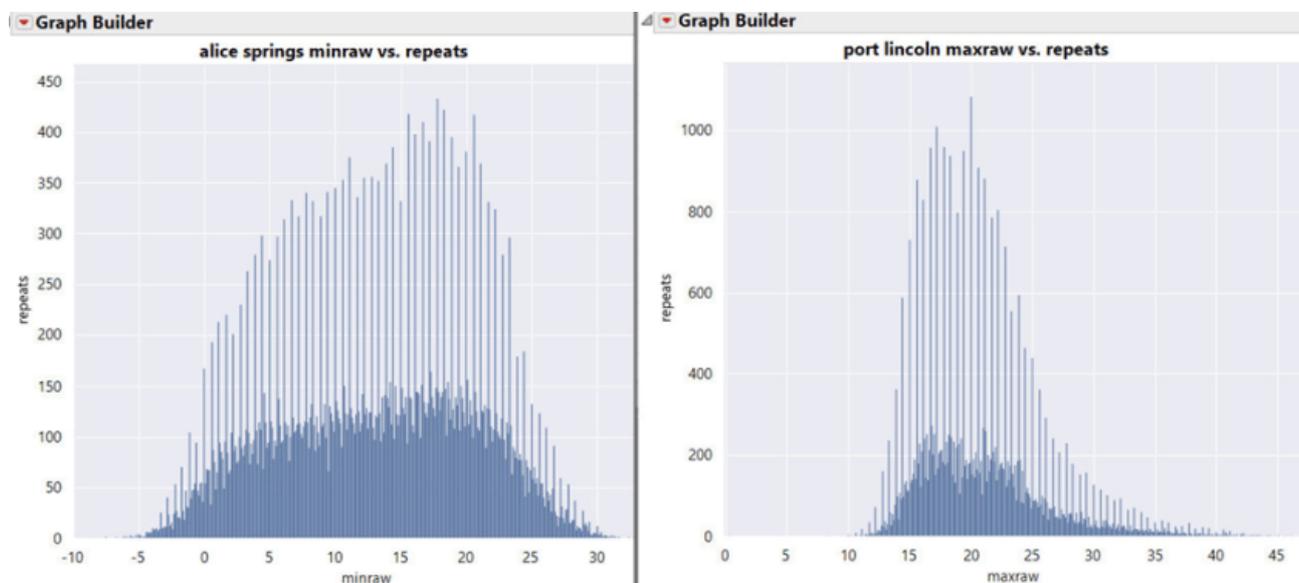
Wir sollten nicht von Rohdaten sprechen, weil dies ein irreführender Begriff ist ...

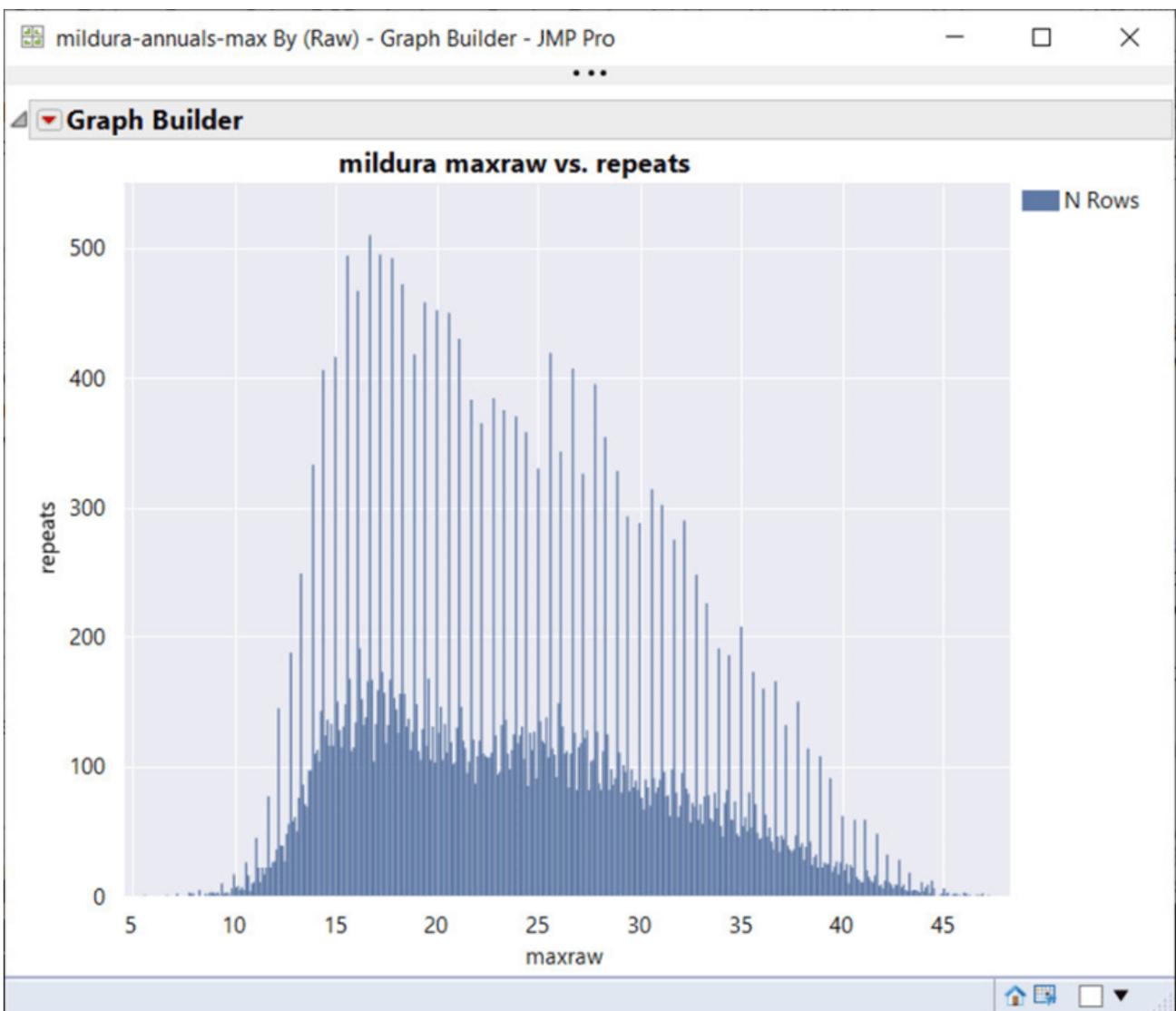
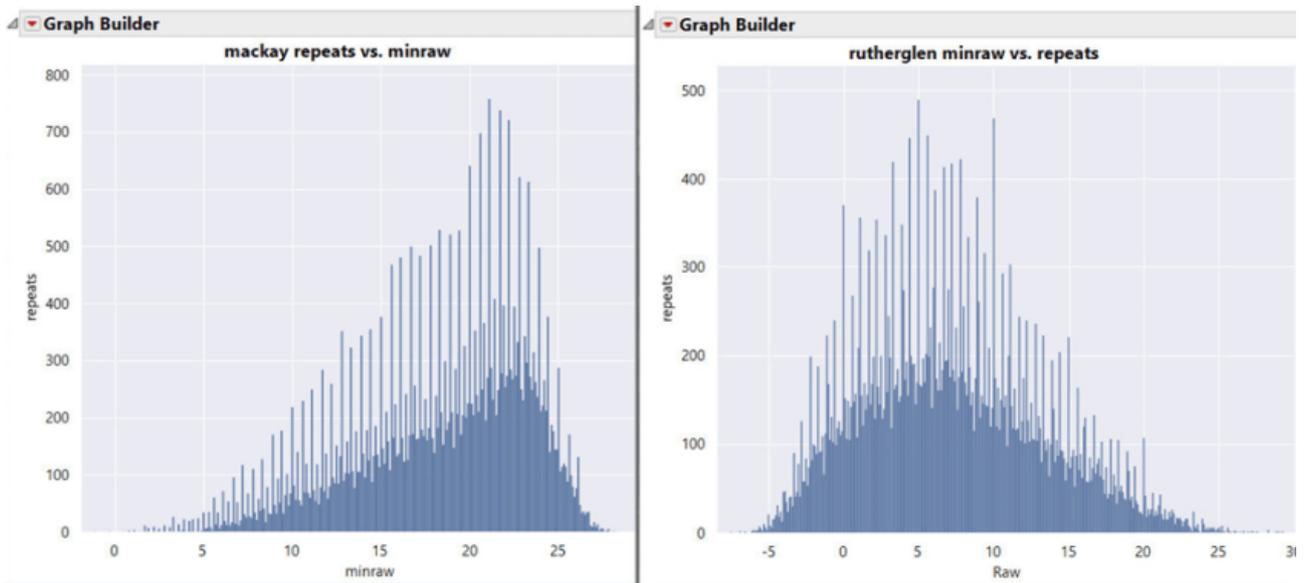
„Der Begriff Rohdaten ist an sich schon irreführend, da er häufig einen vor der Anpassung erstellten Datensatz impliziert, der als reine Aufzeichnung an einem einzigen Stationsstandort betrachtet werden könnte. Für zwei Drittel der ACORN-SAT gibt es keine rohen Temperaturreihen, sondern eine zusammengesetzte Reihe von zwei oder mehr Stationen.“ – BOM

„Die Homogenisierung erhöht nicht die Genauigkeit der Daten – sie kann nicht höher sein als die Genauigkeit der Beobachtungen.“ (M.Syrakova, V.Mateev, 2009)

Der Grund für die Irreführung liegt darin, dass BOM die Daten weiterhin als „roh“ bezeichnet, obwohl es sich um *einen einzigen Durchschnittswert vieler Stationen handelt*; die Standard-Auswahl in der BOM-Software beträgt 40 Stationen. Dies ist die schwache Form des *Flaw of Averages* (Dr. Sam Savage, 2009) [Fehler von Durchschnittswerten], so dass diese einzelne zusammengesetzte Zahl wahrscheinlich falsch ist.

In der Phase der Datenexploration sollte man sich zunächst die Verteilung mit einem Histogramm ansehen. Hier beginnen die Probleme:





Die Wiederholungen/Häufigkeit des Auftretens sind auf der Y-Achse, die Rohtemperaturen sind auf der X-Achse aufgetragen. Das Histogramm zeigt, wie oft bestimmte Temperaturen auftreten. Es ist sofort ersichtlich, dass es ein Problem gibt; es scheinen zwei Histogramme übereinander

gelegt zu sein. Das stachelige Histogramm ist sehr methodisch mit gleichmäßigen Abständen. Es ist schnell klar, dass es sich nicht um saubere Beobachtungsdaten handelt.

Es stellt sich heraus, dass die Ursache für diese Spitzen die doppelte Rundungsungenauigkeit ist, bei der Fahrenheit auf z. B. 1 Grad genau gerundet, dann in Celsius umgewandelt und auf 0,1 C genau gerundet wird, wodurch ein Überschuss an dezimalen 0,0er-Werten und ein Mangel an 0,5er-Werten* (in diesem Beispiel) entsteht; **es gibt verschiedene Rundungsszenarien, bei denen unterschiedliche dezimale Knappheiten und Überschüsse in der gleichen Zeitreihe entstehen!**

[*Am Meteorologischen Institut der Freien Universität Berlin wurde bei der Berechnung der Mitteltemperaturen an der Station Berlin-Dahlem ein anderes Verfahren praktiziert. Bei Werten von x,5 wurde grundsätzlich zur geraden Zahl gerundet. Ein Wert von 2,5 ergibt also 2, einer von 3,5 demnach 4. Auf diese Weise wurde eine Warm-Verzerrung durch ein mathematisch korrektes grundsätzliches Aufrunden in diesen Fällen vermieden. A. d. Übers.]

Die nachstehende Abbildung stammt aus der zitierten Studie und zeigt verschiedene Rundungs-/Umrechnungs-/Rundungsprotokolle. Betrachtet man zum Beispiel die erste Zahlenreihe, 0,5°C und 2,0°F, so bedeutet dieses Szenario, dass Fahrenheit auf 2 Grad genau gerundet und dann auf 0,5°C genau umgerechnet wurde. Alle diese Szenarien wurden in der Studie untersucht, aber was sich speziell auf die BOM-Daten auswirkt, ist das sechste Protokoll von unten – Rundung auf 1,0°F Genauigkeit, Konvertierung und anschließende Rundung auf 0,1°C Genauigkeit. Dies führt zu einem Übermaß an dezimalen 0's und keinen 0,5's

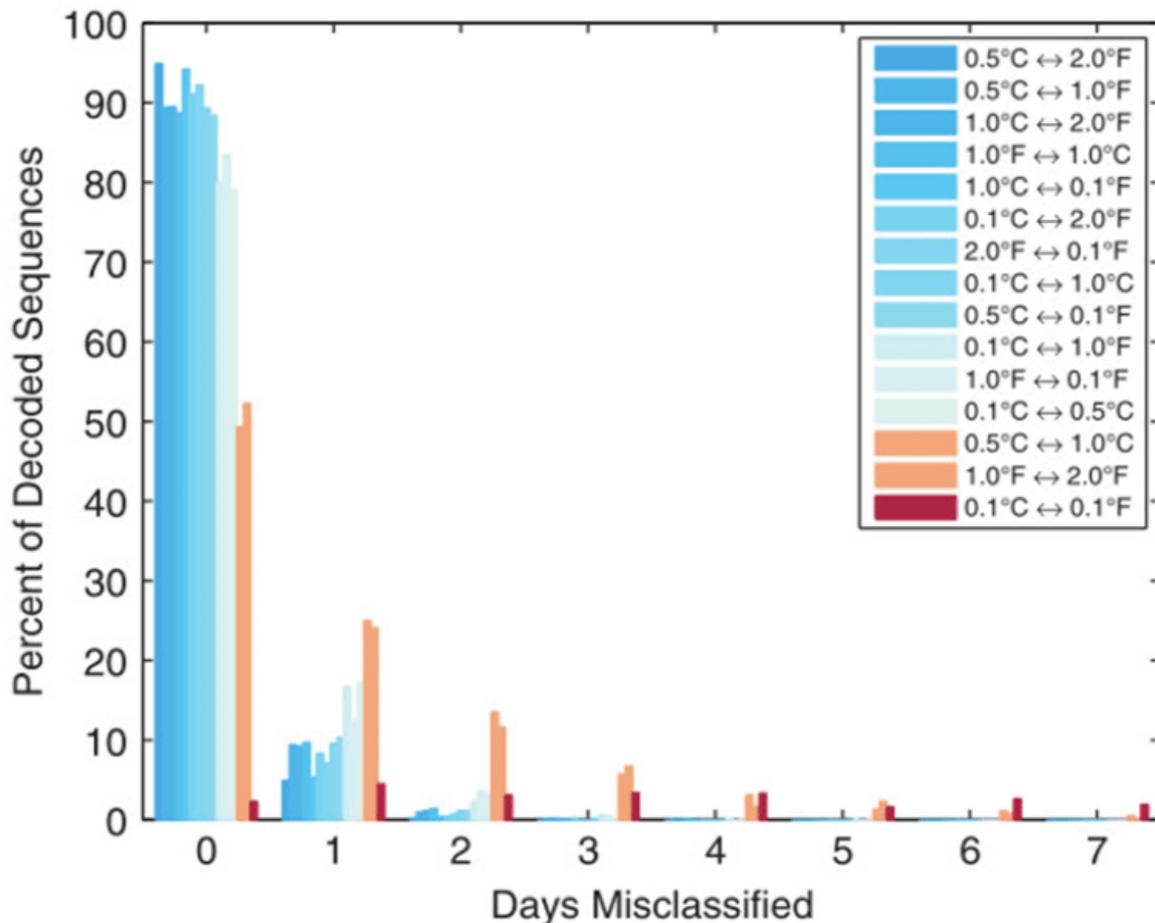
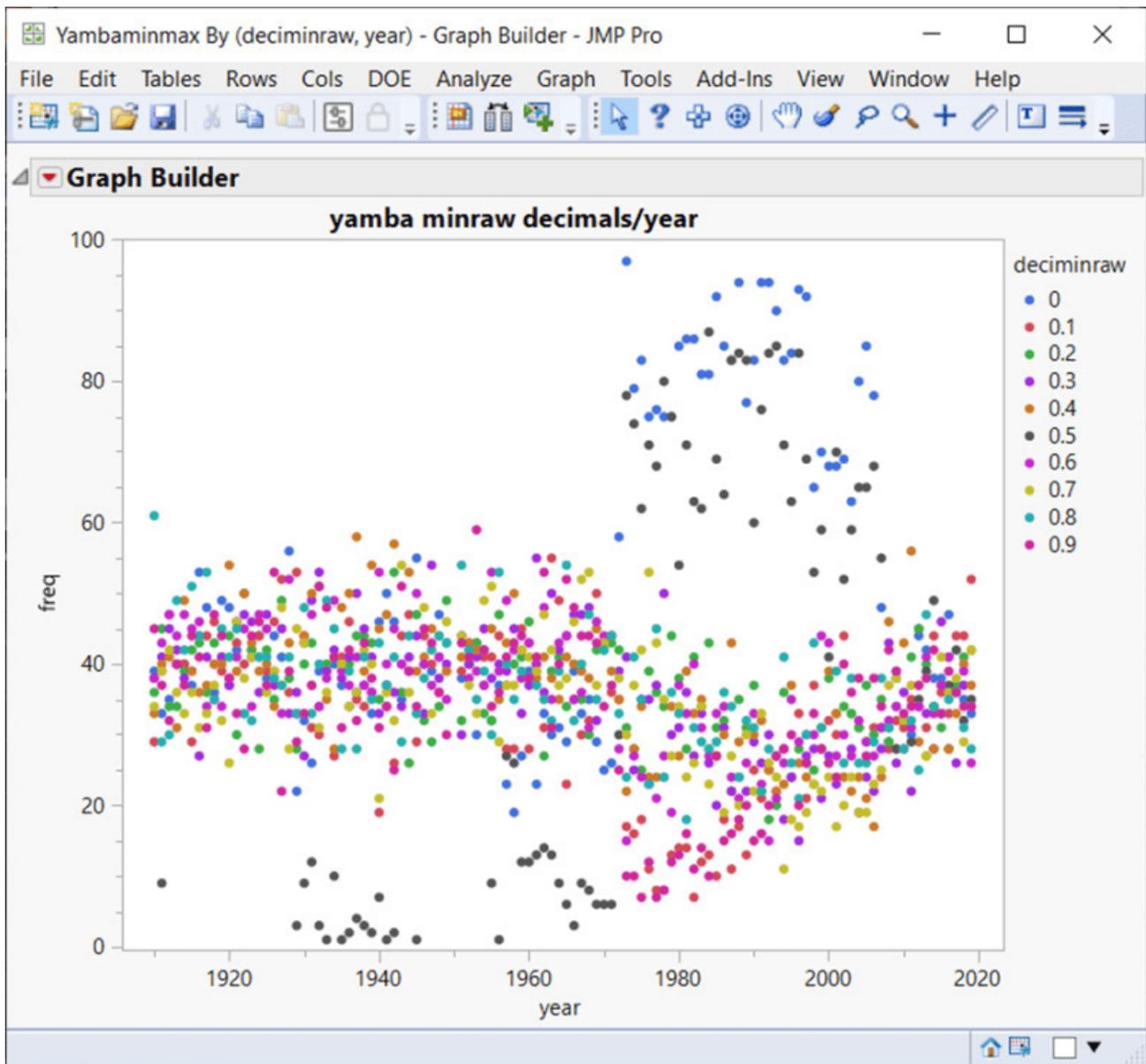


Figure 5. Histograms of the error rate in the decoded precision of synthetic observations. For each pair of precisions, 1000 trials using 20 years of synthetic data are generated, with the observational precision changing after 10 years. Histogram bars are shaded by the Jensen–Shannon divergence on the same colour scale as in Figure 4. Non-zero errors in the 0.1 °F-to-0.1 °C case occur over a much wider range which is not fully shown, with an expected misclassification count of 35 days. Its distribution has a qualitatively similar shape to the other cases.

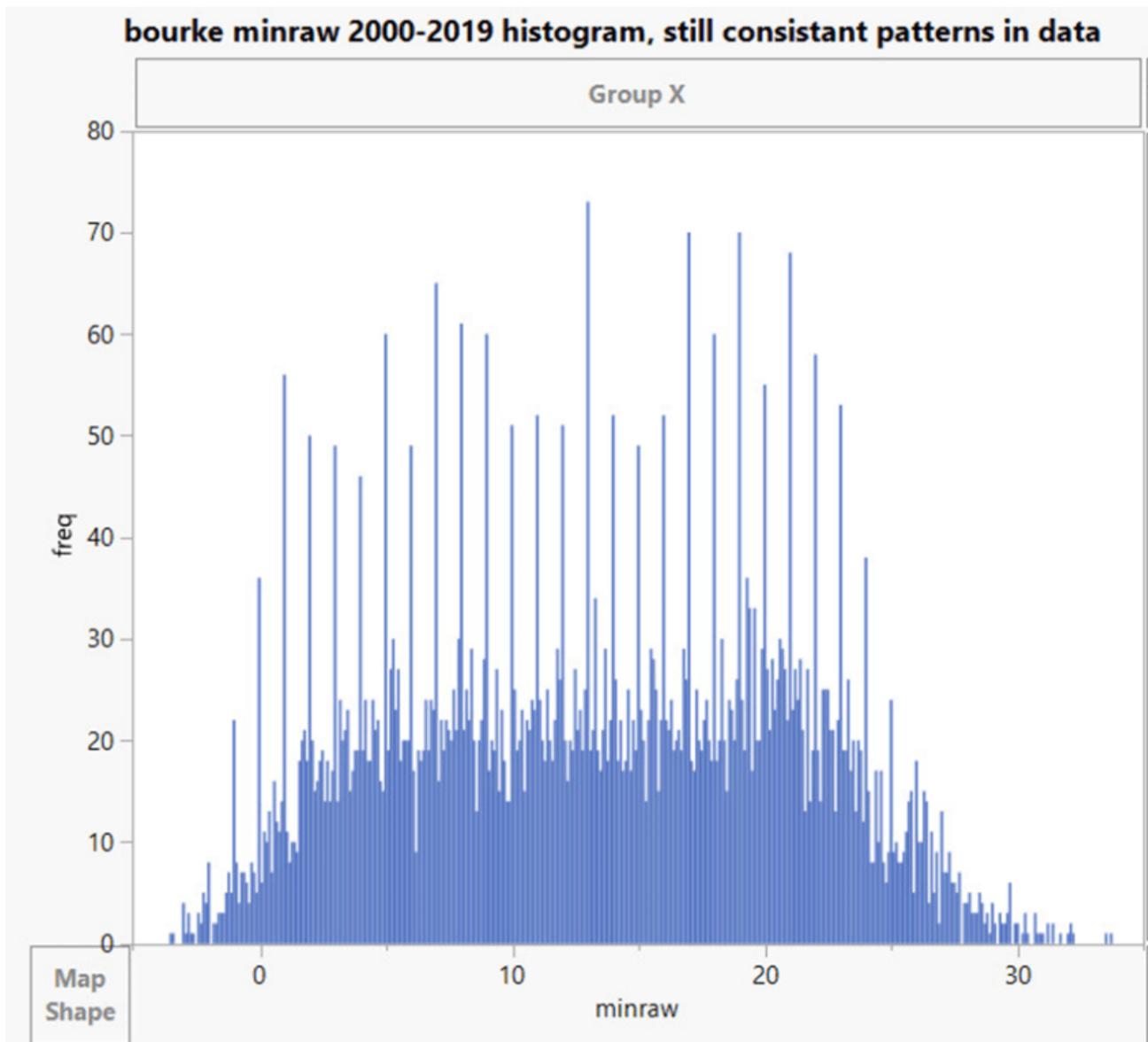
Die Studie: „*Decoding The Precision Of Historical Temperature Observations*“ – Andrew Rhimes, Karen A McKinnon, Peter Hubers. [Etwa: Entschlüsselung der Genauigkeit historischer Temperaturmessungen]

Die BOM-Doppelrundungsfehler, die es gibt, werden durch dieses spezielle Protokoll verursacht, und man kann sie unten leicht erkennen, wenn man die **Dezimalhäufigkeit pro Jahr** mit den Jahren vergleicht. Die Verwendung der Dezimalhäufigkeit pro Jahr ist der Schlüssel zum Verständnis des Gesamtbildes:



Die Abbildung zeigt die Gemeinde Yamba und verdeutlicht den Mangel an 0,5er-Dezimalstellen von 1910 bis 1970 und die übermäßige Verwendung von 0,0 und 0,5 Dezimalstellen von 1980 bis 2007 oder so. Wenn man sich die meisten Landesstationen anschaut, wird deutlich, dass die Probleme mit dem Dezimalsystem noch lange nach der Dezimalisierung in den 70er Jahren bestehen, und in einigen Fällen sogar bis 2010.

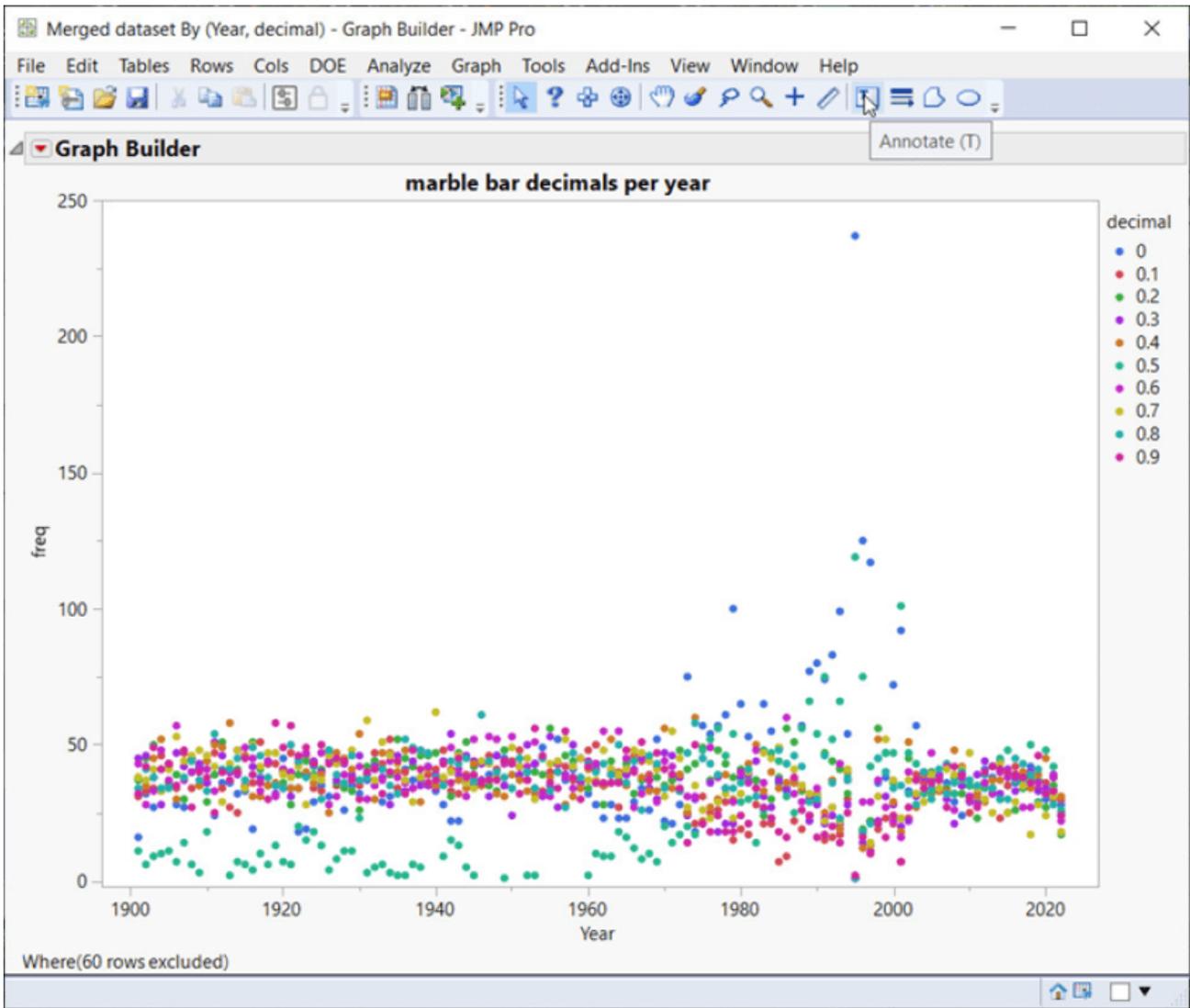
Sehen Sie sich zum Beispiel Bourke unten an:



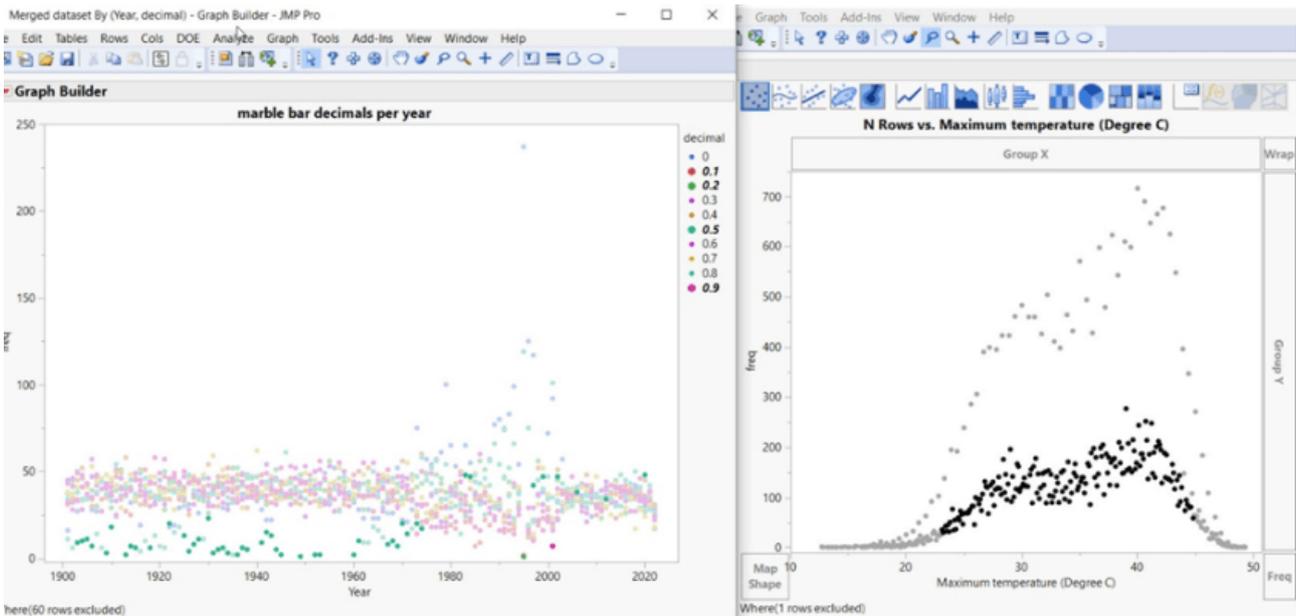
Diese verschiedenen Szenarien mit doppelten Rundungsfehlern lassen in einigen Fällen Zweifel an den Aufzeichnungen aufkommen (siehe oben) und führen dazu, dass die Zeitreihen nicht mehr übereinstimmen. Wie zu sehen ist, hat BOM das Problem des doppelten Rundungsfehlers nicht korrigiert, obwohl eine einfache Matlab-Korrektursoftware in dem oben genannten Dokument vorhanden ist.

Der Marmorbalken unten zeigt genau das gleiche Problem: unkorrigierte doppelte Rundungsfehler, die zu einer Knappheit von 0,5 Dezimalstellen führen. Auch hier gibt es um das Jahr 2000 herum einen Überschuss an 0,0 und 0,5er-Dezimalstellen, die verwendet werden. Aus dem BOM-Stationskatalog für Marble Bar:

„Die automatische Wetterstation wurde im September 2000 eröffnet, mit Instrumenten im gleichen Bildschirm wie die frühere manuelle Station (4020). Der manuelle Standort wurde bis 2006 weitergeführt, wurde aber wegen der sich danach verschlechternden Datenqualität nur noch bis Ende 2002 in ACORN-SAT verwendet.“



Zurück zu den stacheligen Histogrammen: So werden sie gebildet:



Durch die Verwendung von dynamisch verknüpften Datentabellen in SAS JMP werden durch die Auswahl bestimmter Datenpunkte diese mit allen anderen

Datentabellen verknüpft, was zeigt, wie die Spikes gebildet werden.

Die geringe Anzahl von 0,5er-Dezimalstellen führt zu einem Histogramm auf niedrigerem Niveau (siehe obige Darstellung), bei dem die Datenpunkte auf ein niedrigeres Niveau fallen und Lücken und Spitzen hinterlassen. Dies verursacht die seltsam aussehenden Histogramme und ist ein Hinweis darauf, dass die Daten nicht um doppelte Rundungsfehler korrigiert sind. Praktisch alle untersuchten BOM-Landstationen weisen dieses Problem auf.

ANPASSUNGEN ODER KORREKTUREN DER TEMPERATUREN, UM TRENDS ZU VERSTÄRKEN.

„Stellen Sie sich zum Beispiel vor, dass eine Wetterstation in Ihrem Vorort oder Ihrer Stadt wegen eines Bauvorhabens verlegt werden musste. Es ist gut möglich, dass der neue Standort etwas wärmer oder kälter ist als der vorherige. Wenn wir der Allgemeinheit die beste Schätzung des wahren langfristigen Temperaturtrends an diesem Ort liefern wollen, ist es wichtig, dass wir derartige Veränderungen berücksichtigen. Zu diesem Zweck verwenden das Amt und andere große meteorologische Organisationen wie die NASA, die National Oceanic and Atmospheric Administration und das britische Met Office ein wissenschaftliches Verfahren namens Homogenisierung.“ – BOM

Das herablassende BOM-Zitat unter der Überschrift bereitet Sie auf das vor, was kommen wird:

Es wurde eine ganze Industrie geschaffen, um „Verzerrungen“ zu korrigieren. BOM hat den SNHT-Algorithmus mit einer Standardliste von 40 in die engere Wahl gezogenen Stationen verwendet, um daraus eine einzige Zahl zu ermitteln. Bei einem Signifikanzniveau von 95 % bedeutet dies, dass 1 von 20 „Schritten“ oder „Verzerrungen“ ein **falsches Positiv** ist.

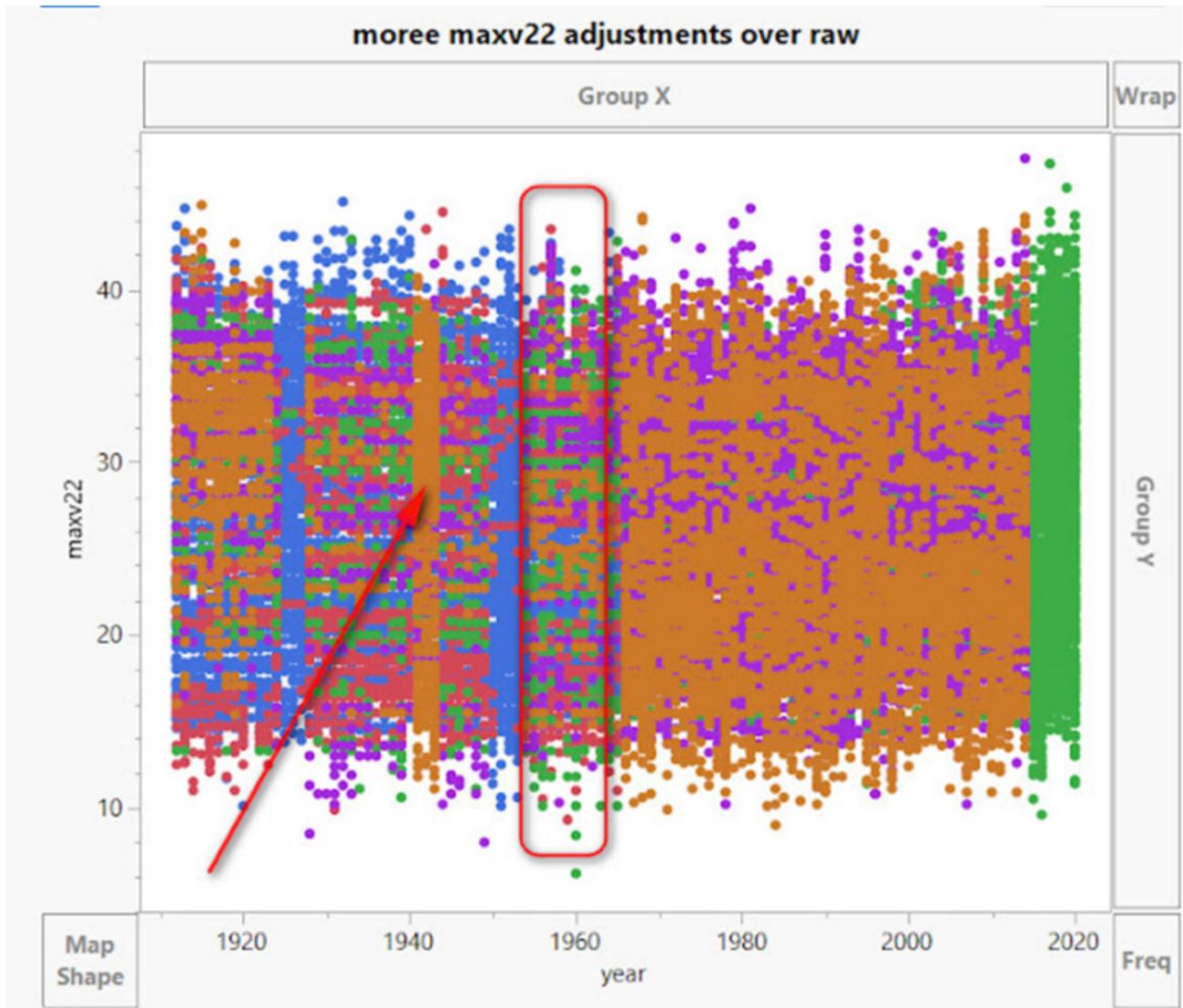
Die iterative Verwendung verschiedener Softwareversionen für die gleichen Daten ohne die Verwendung von Vielfachkorrekturen, um den Zufall bei der Suche nach dieser „Verzerrung“ zu berücksichtigen, führt zu Fehlern und Verzerrungen in den Daten.

„Systematische Verzerrungen haben, solange sie sich nicht ändern, keinen Einfluss auf die Temperaturveränderungen. So führt eine falsche Platzierung der Messstationen zu einer Verzerrung, aber solange sie sich nicht ändert, ist sie unerheblich. Aber jede Änderung der Anzahl und des Standorts der Messstationen könnte den Anschein eines falschen Trends erwecken.“ – Prof. Thayer Watkins, Universität San Jose.

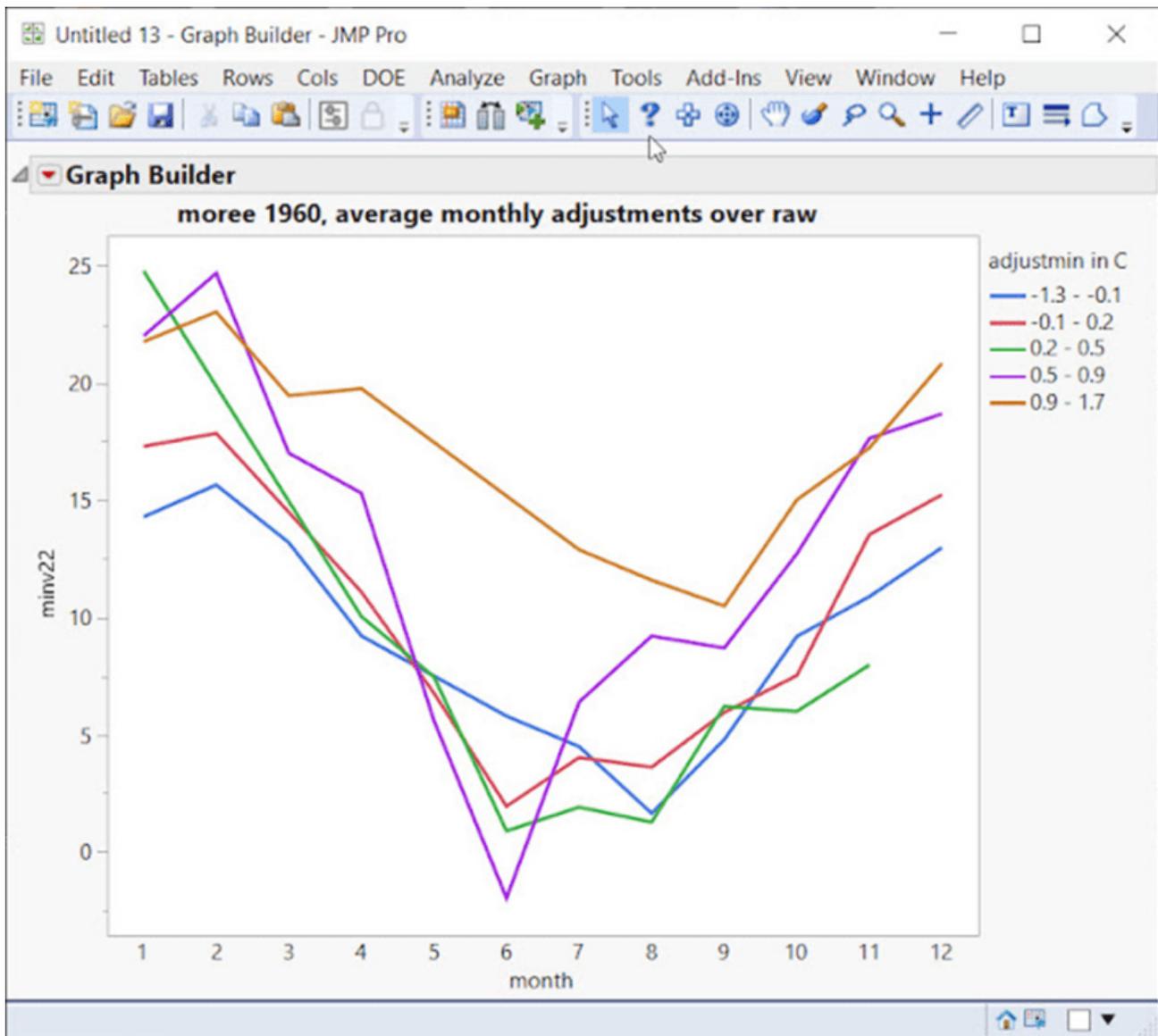
Da es keine Reihen gibt, die eindeutig als homogen eingestuft werden können, ist es nicht möglich, eine vollkommen homogene Referenzreihe zu erstellen.

Außerdem liefert SNHT keine Schätzung des Konfidenzniveaus der Anpassungen.

Wir sollen glauben, dass die Verschiebung einer Station oder das Wachstum der Vegetation oder jede andere Verzerrung, die sich bildet, zu diesen Arten von Anpassungen führt:



Eine „schrittweise Veränderung“ würde wie der Pfeil aussehen, der auf 1940 zeigt. Aber schauen Sie sich die Zeit um 1960 an – es gibt eine Vielzahl von Anpassungen, die einen großen Temperaturbereich abdecken, es ist ein chaotisches Durcheinander von Anpassungen in verschiedenen Bereichen und Jahren. Wenn wir nur das Jahr 1960 in Moree betrachten:

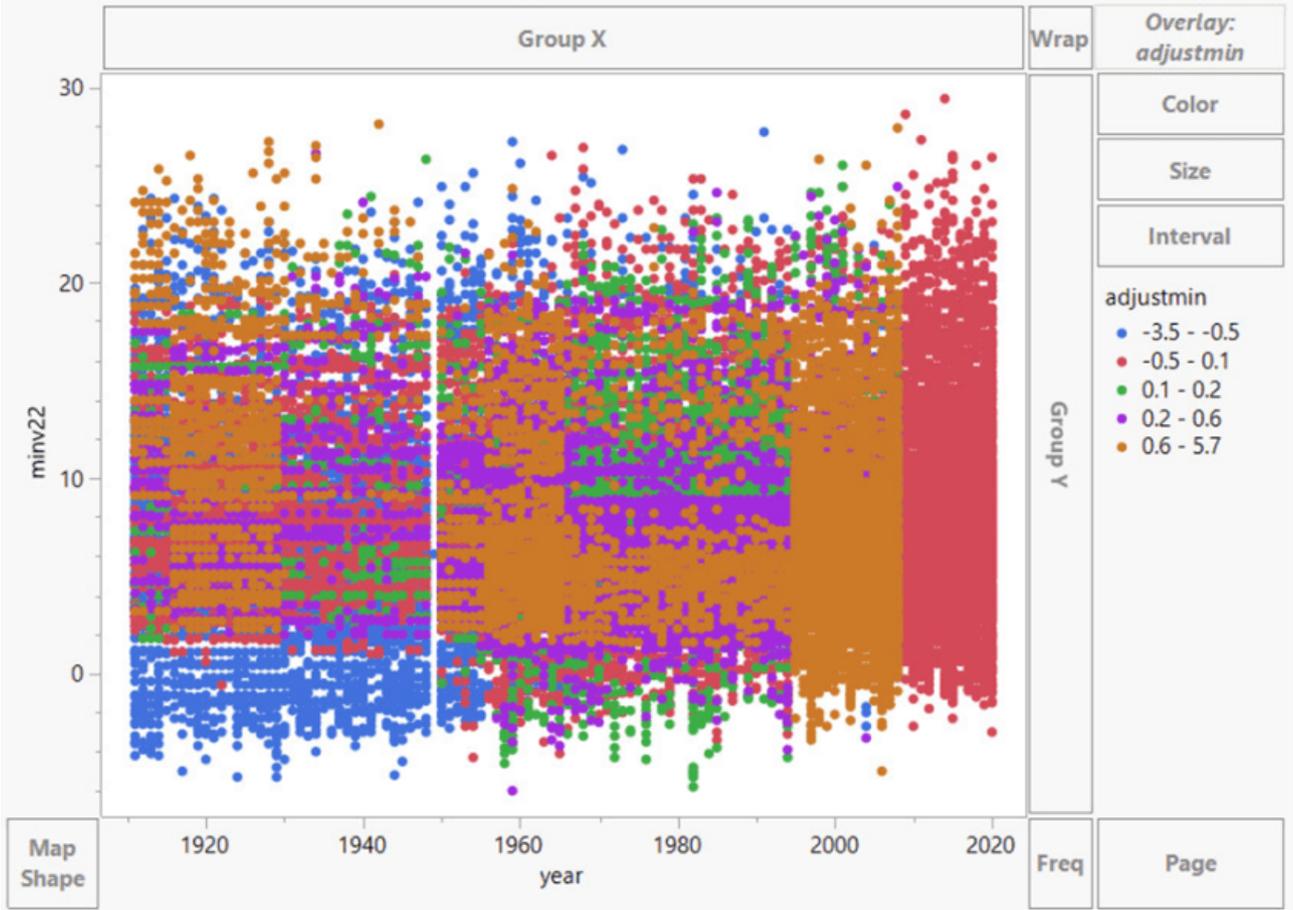


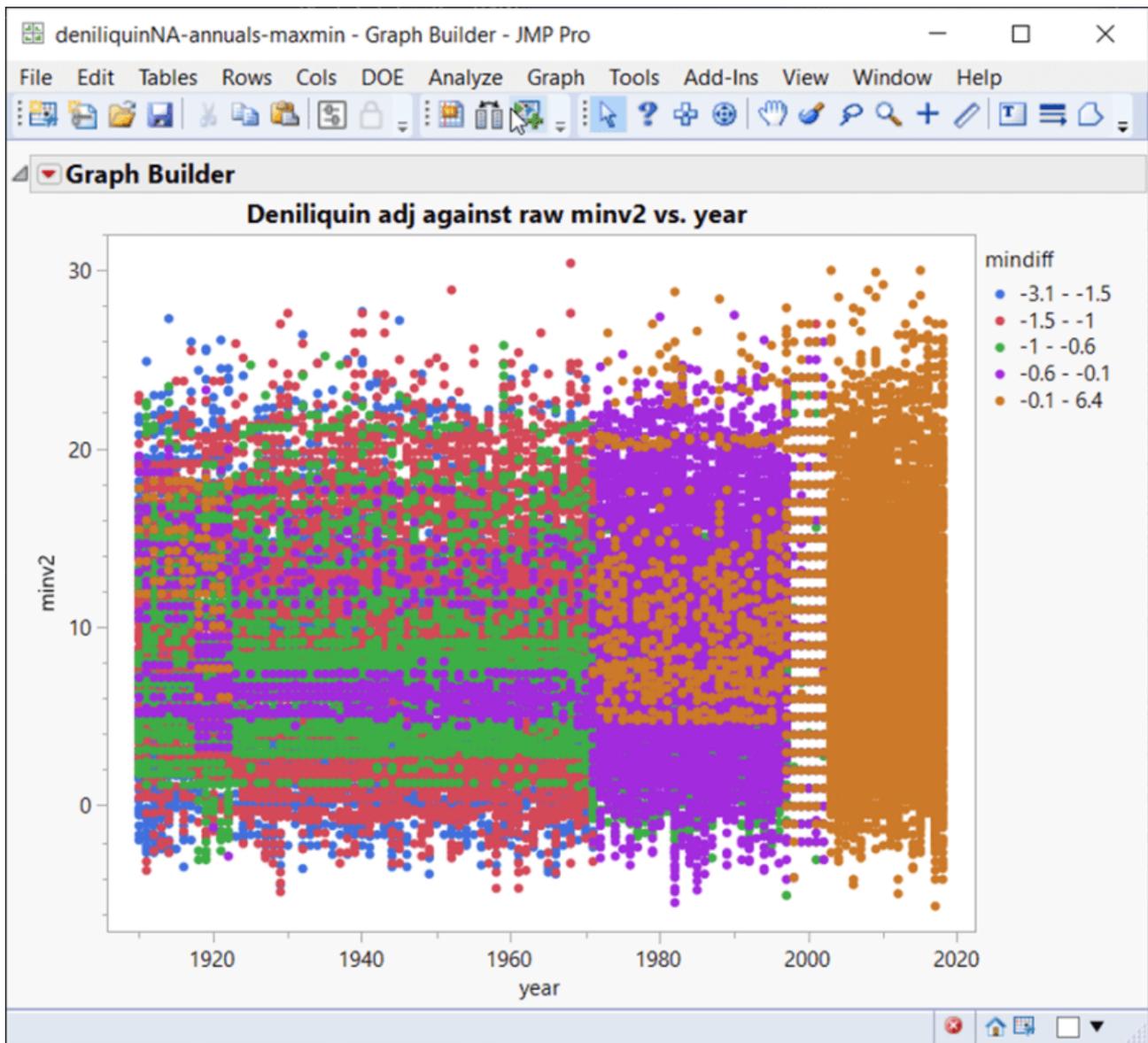
Das BOM möchte uns glauben machen, dass diese chaotischen Anpassungen für nur 1960 in diesem Beispiel exakte und präzise Anpassungen sind, die zur Korrektur von Verzerrungen erforderlich sind.

Weitere Streudiagramme für Anpassungen:



nhill minv22 adjustments over raw





Eine wahrscheinlichere, auf dem Modus Operandi der BOM beruhende Erklärung ist, dass bestimmte Monate und Jahre eine bestimmte Erwärmung und Abkühlung erfahren, um den gewünschten Trend zu verstärken.

Wenn man sich die Streudiagramme ansieht, kann man deutlich erkennen, dass es bei den Anpassungen nicht um die Korrektur von „Schrittwechseln“ und Verzerrungen geht. Wie wir oben gesehen haben, verschlimmern die Anpassungen in den meisten Fällen die Daten, indem sie doppelte Sequenzen und andere Verzerrungen hinzufügen. Das Bedford'sche Gesetz zeigt auch eine *geringere Übereinstimmung mit Anpassungen, die auf Datenprobleme hinweisen*.

Bestimmte Monate werden mit den größten Anpassungen belastet, und dies ist bei vielen Zeitreihen der Fall:

	month	N Rows
1	1	3
2	3	11
3	4	951
4	5	2
5	6	9
6	8	1
7	11	1
8	12	8

bourke minv2.1
adj over 1.5
97% in april

	month	N Rows
1	2	1
2	4	2
3	5	414
4	6	2
5	7	5
6	8	1330
7	9	6
8	11	2
9	12	2

bourke minv2.1
adjust less than or equal -2.7C

	month	N Rows
1	1	36
2	4	1
3	8	76

charleville minv2.1
adj <= 3.1C

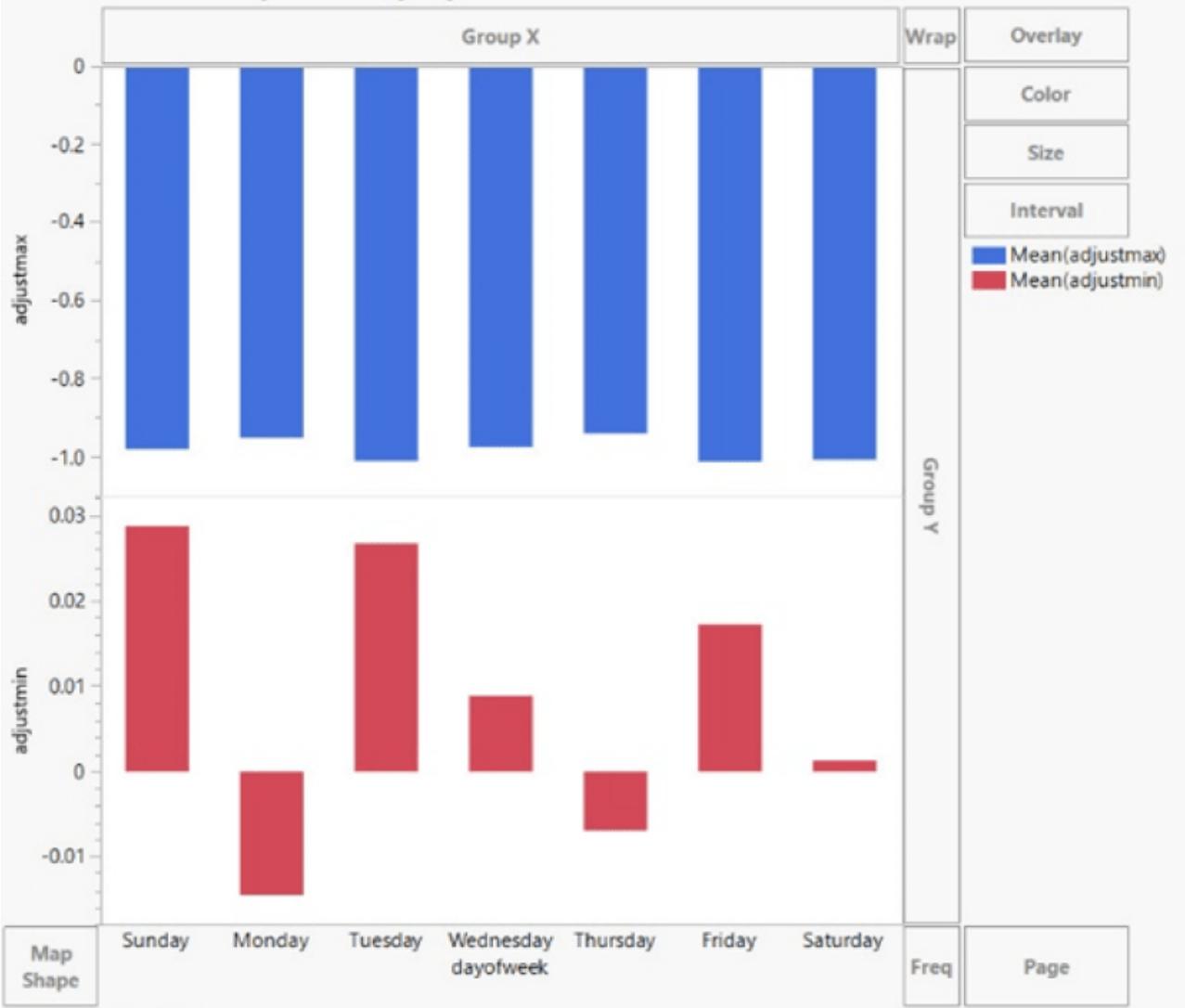
	Month[Column 1]	N Rows
1	1	3
2	2	12
3	3	289
4	4	370
5	5	66
6	6	30
7	7	1
8	8	2
9	9	2
10	10	6
11	12	7

deniliquin minv2.1
adj <= -2.3C

Und einige Anpassungen hängen davon ab, welcher Wochentag gerade ist:

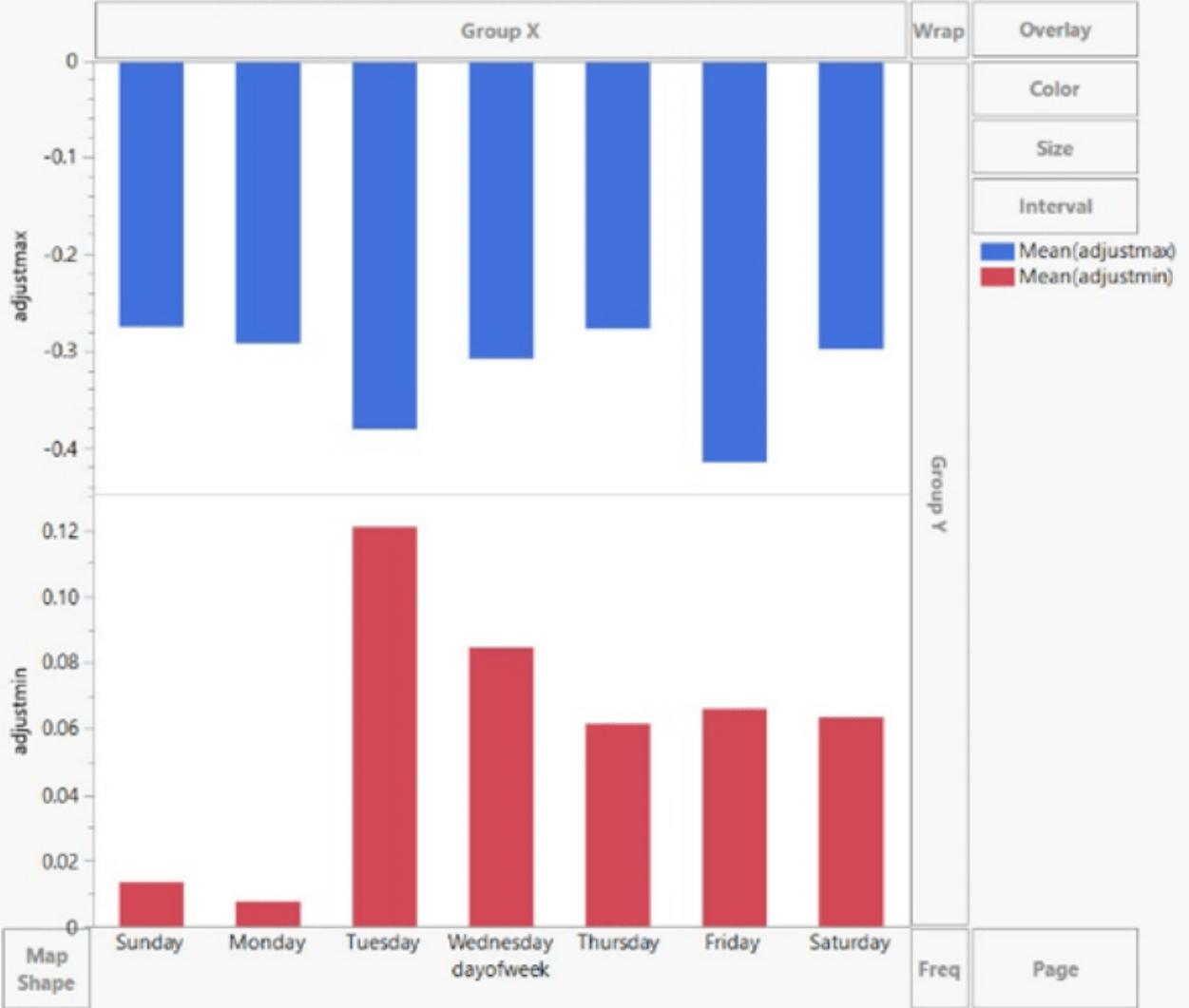


melbourne adjustments by day of week for november 1910-2020

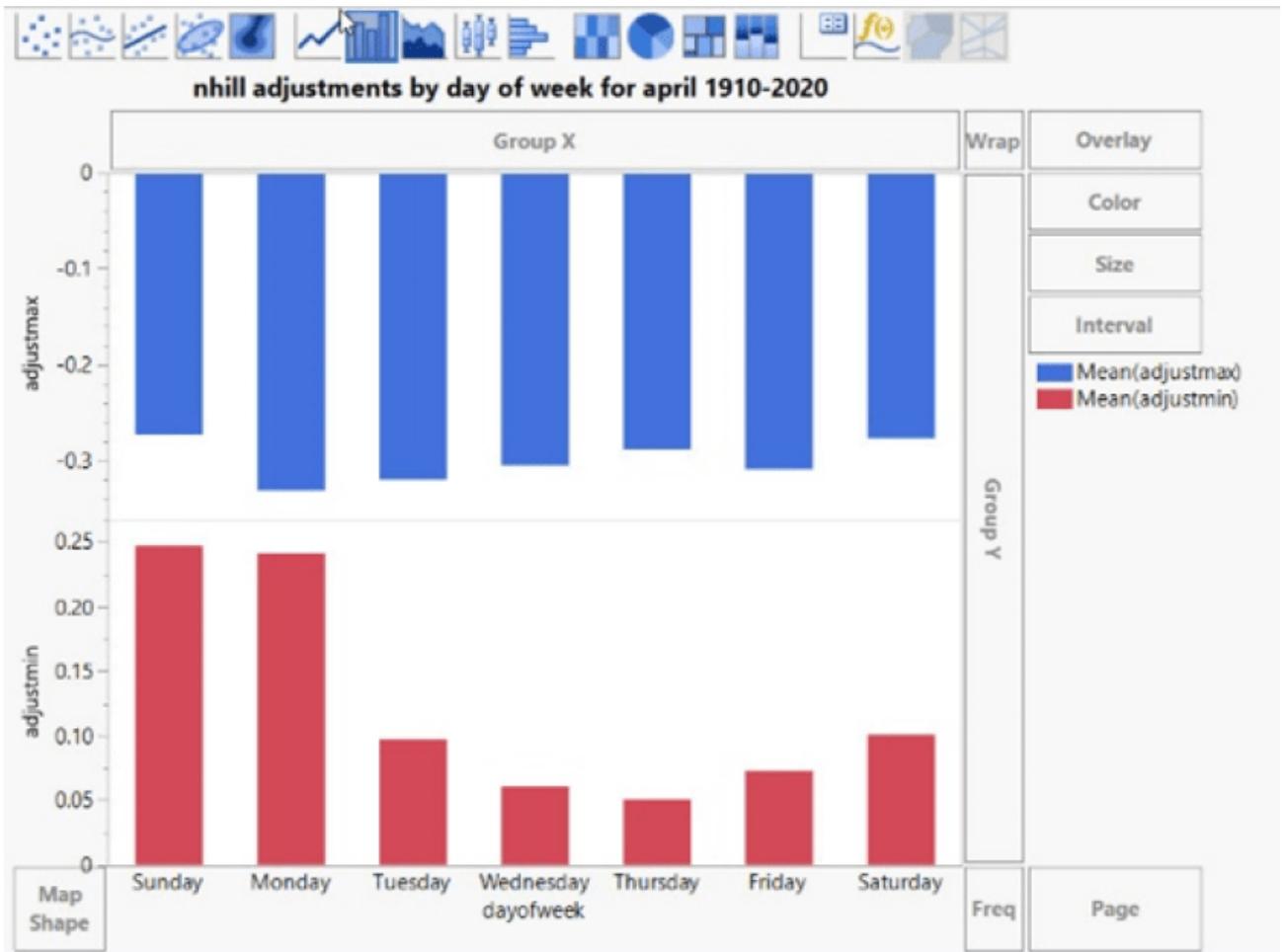
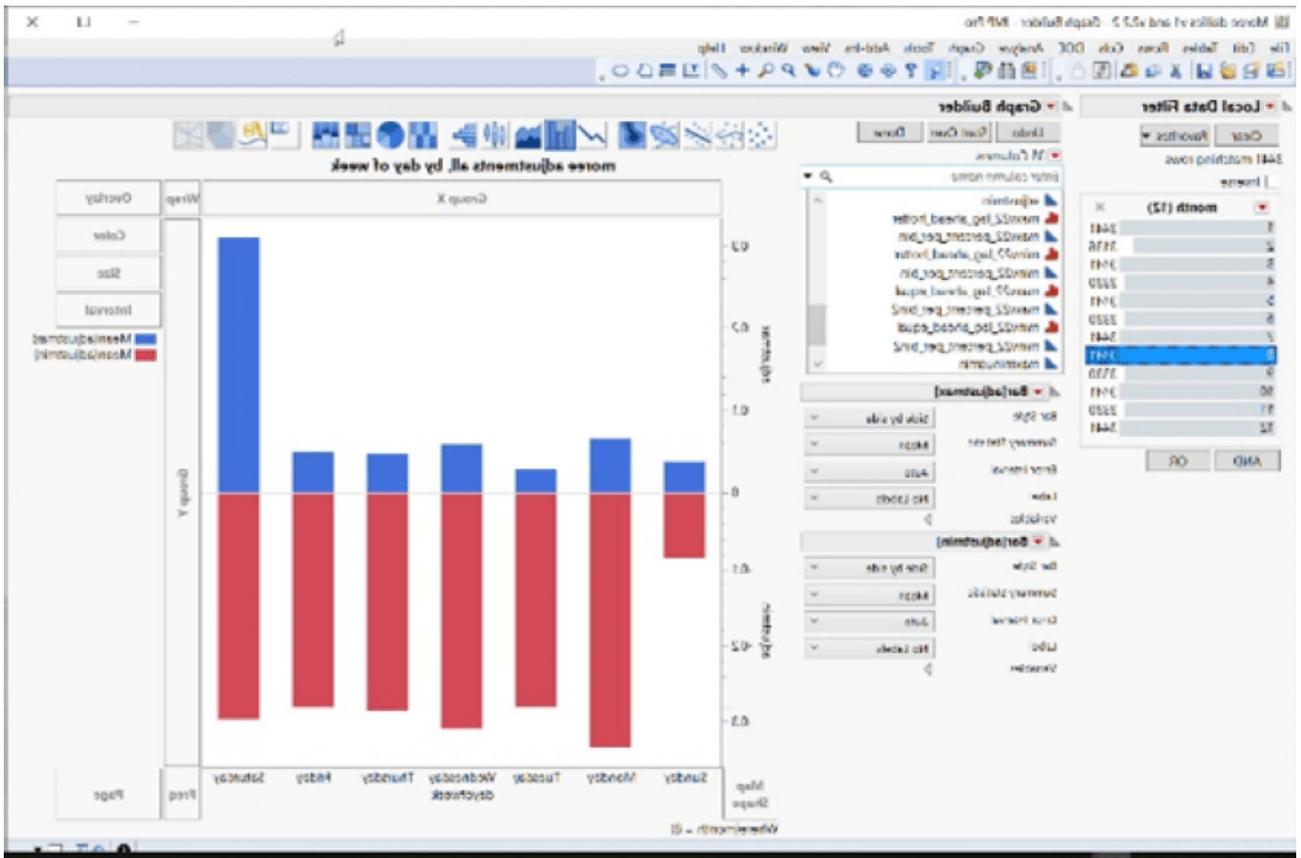


Where(month = 11)

melbourne adjustments by day of week for 1965



Where(year = 1965)



	dayofweek	N Rows	Mean(adjustmin)	Mean(adjustmax)
1	Sunday	5792	-0.043166262	0.4221390374
2	Monday	5792	-0.020779505	0.4402535658
3	Tuesday	5792	-0.134812165	0.4391759777
4	Wednesday	5792	-0.155991479	0.4427201395
5	Thursday	5792	-0.147346939	0.4422564821
6	Friday	5791	-0.133541258	0.4366654892
7	Saturday	5792	-0.137578027	0.4232623805

	Table	dayofweek	Y	Statistic	mode
1	Mildura dailies...	Sunday	maxv22	Mode	17.2
2	Mildura dailies...	Monday	maxv22	Mode	17.3
3	Mildura dailies...	Tuesday	maxv22	Mode	15.4
4	Mildura dailies...	Wednesday	maxv22	Mode	15.4
5	Mildura dailies...	Thursday	maxv22	Mode	15.4
6	Mildura dailies...	Friday	maxv22	Mode	15.7
7	Mildura dailies...	Saturday	maxv22	Mode	15.4

Die Anpassung nach Wochentag bedeutet, dass der Modalwert, die häufigste Temperatur in der Zeitreihe, ebenfalls von dieser Verzerrung betroffen ist – und damit auch die mittlere Temperatur.

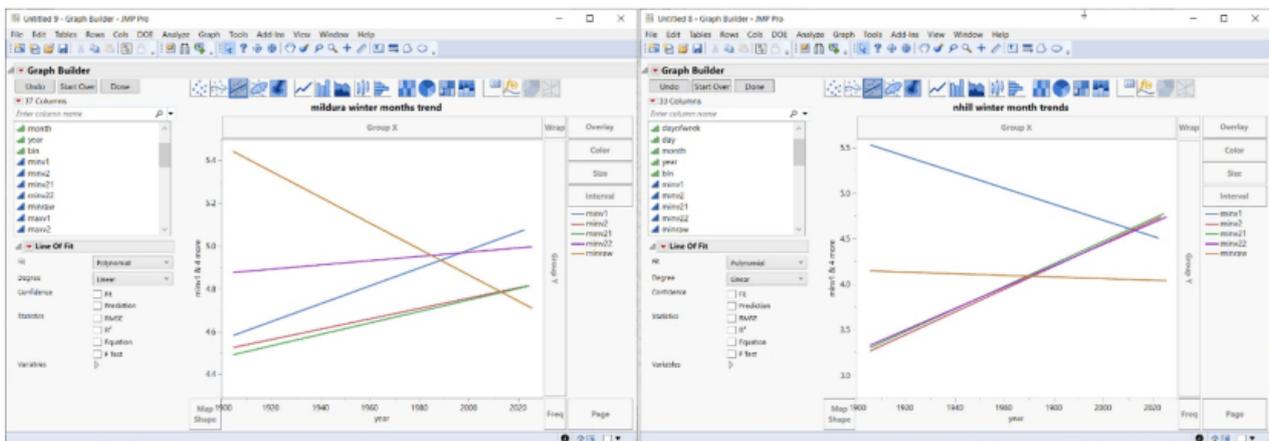
Table	dayofweek	Y	Statistic	mode
1 Mildura dailies v1 and v2.2.2	Sunday	minv22	Mode	12.1
2 Mildura dailies v1 and v2.2.2	Friday	minv22	Mode	10.1
3 Mildura dailies v1 and v2.2.2	Wednesday	minv22	Mode	8.8
4 Mildura dailies v1 and v2.2.2	Monday	minv22	Mode	8.4
5 Mildura dailies v1 and v2.2.2	Tuesday	minv22	Mode	8.4
6 Mildura dailies v1 and v2.2.2	Thursday	minv22	Mode	8.4
7 Mildura dailies v1 and v2.2.2	Saturday	minv22	Mode	8.4

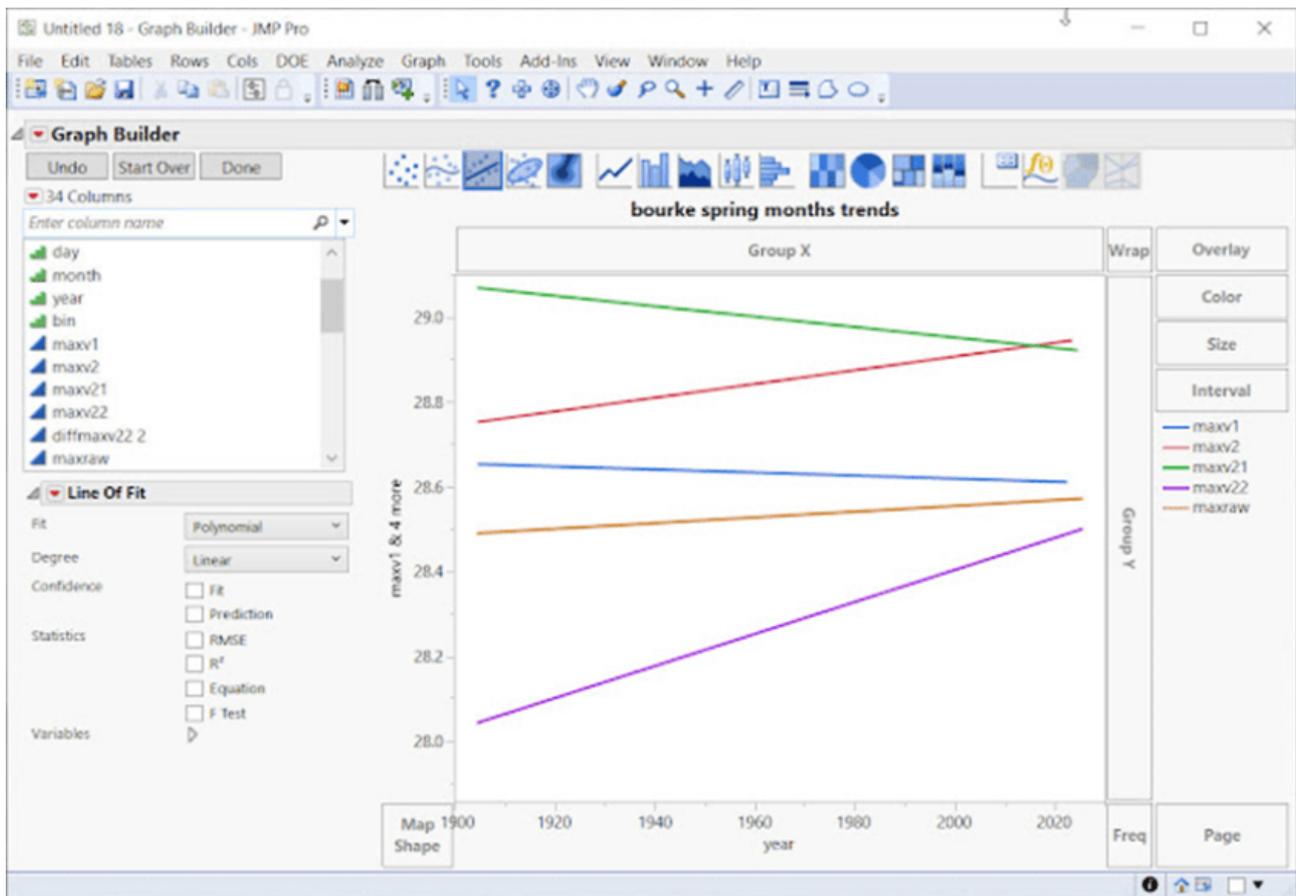
„Die Analyse hat gezeigt, dass die neu vorgenommenen Anpassungen in ACORN-SAT Version 2.2 den geschätzten langfristigen Erwärmungstrend in Australien nicht verändert haben.“ – BOM

„...gepoolte Daten können verschiedene individuelle Manipulations-Signaturen ausgleichen.“ – (Diekmann, 2007)

Es wird angenommen, dass die frühen Temperaturen die gleichen Konfidenzintervalle haben wie die jüngeren Temperaturen, was offensichtlich falsch ist.

Einzelne Stationstrends können sicherlich durch Anpassungen verändert werden, insbesondere bei Jahreszeiten, die die BOM-Software auswertet:

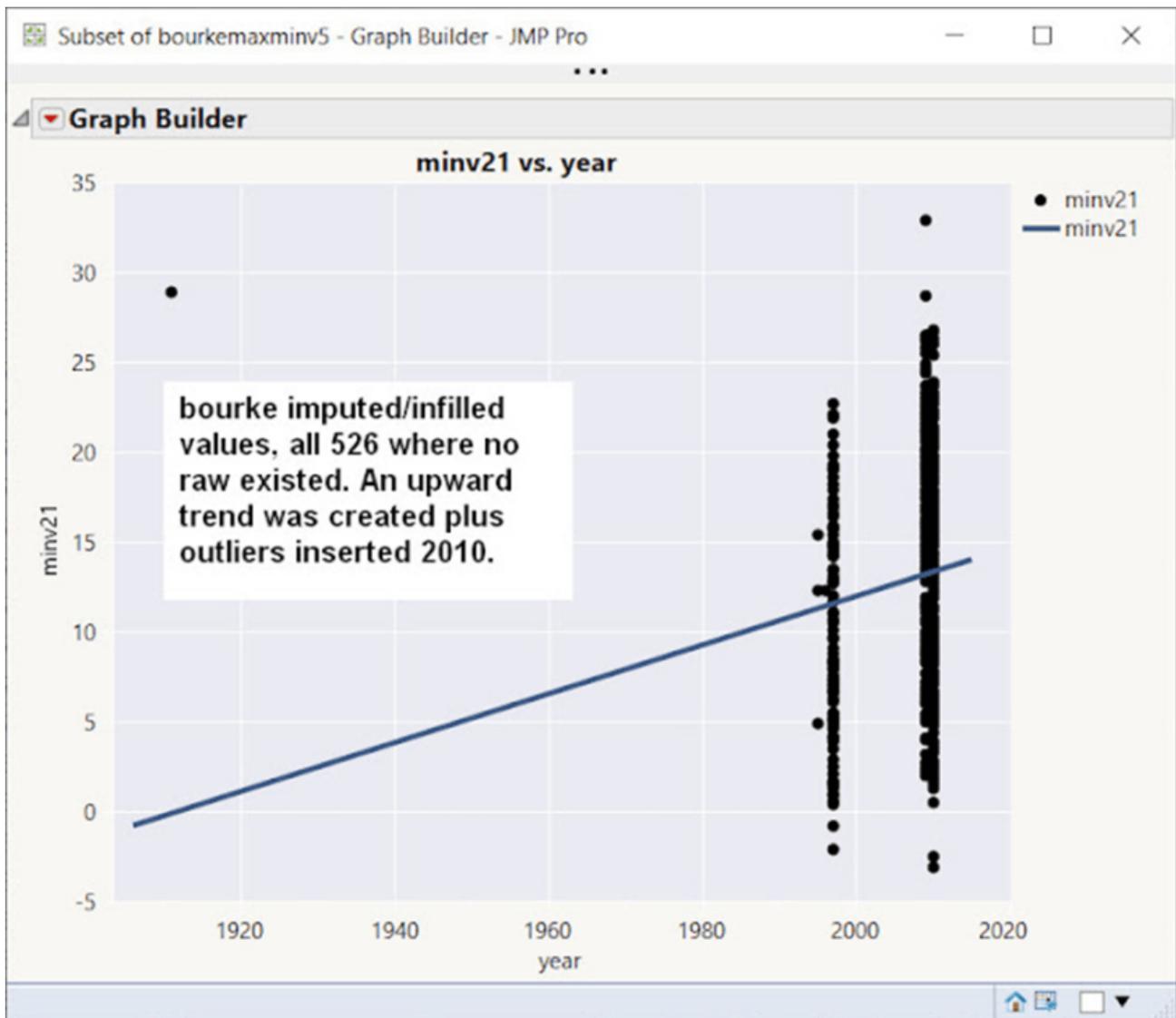




Nachfolgend sehen wir uns Werte an, die in Raw fehlen, aber in Version 2.1 oder 2.2 erscheinen.

Dies zeigt uns, dass es sich um erstellte oder importierte Werte handelt. In diesem Fall fehlen die schwarzen Punkte im Rohdatenmaterial, erscheinen aber jetzt als Ausreißer.

Diese Ausreißer und Werte haben für sich genommen einen Aufwärtstrend. Mit anderen Worten: Die importierten/erstellten Daten weisen einen Erwärmungstrend auf (siehe unten).



Zusammenfassend lässt sich sagen, dass die Anpassungen Läufe mit doppelten Temperaturen erzeugen und auch Läufe mit doppelten Sequenzen, die in verschiedenen Monaten oder Jahren existieren. Dies sind offensichtlich „gefälschte“ Temperaturen – keine Klimadatenbank kann zum Beispiel 17 Tage (oder schlimmer) in Folge mit der gleichen Temperatur rechtfertigen.

Anpassungen werden vor allem in bestimmten Monaten vorgenommen und können davon abhängen, welcher Wochentag gerade ist.

Anpassungen führen auch dazu, dass Daten verschwinden oder wieder auftauchen, ebenso wie Ausreißer. Datenanalyse-Digit-Tests (Simonsohn's oder Benford's oder Luis-Leder Weltbank-Digit-Tests) zeigen, dass die *Rohdaten nach Anpassungen im Allgemeinen schlechter werden!* Außerdem können ganze Temperaturbereiche bis zu 80 Jahren oder mehr verschwinden.

Diese Daten wären für die meisten Branchen ungeeignet.

DAS DEUTSCHE PANZER-PROBLEM

Im Zweiten Weltkrieg wurde jeder in Deutschland hergestellte Panzer oder jedes Waffenteil mit einer Seriennummer versehen. Anhand der Seriennummern von beschädigten oder erbeuteten deutschen Panzern konnten die Alliierten die Gesamtzahl der Panzer und anderer Maschinen im deutschen Arsenal berechnen.

Die Seriennummern enthielten zusätzliche Informationen, in diesem Fall eine Schätzung der Gesamtbevölkerung auf der Grundlage einer begrenzten Stichprobe.

Dies ist ein Beispiel für das, was David Hand als **Dark Data** bezeichnet. Dabei handelt es sich um Daten, die in vielen Branchen vorhanden sind, aber nie genutzt werden, jedoch interessante Informationen preisgeben, die genutzt werden können. [\(Link\)](#)

Dark Data im Zusammenhang mit den australischen Klimadaten würde uns einen zusätzlichen Einblick in das gewähren, was das BOM hinter den Kulissen mit den Daten macht, von denen sie vielleicht nichts weiß. Wenn fragwürdige Arbeit geleistet würde, wären sie sich eines „Informationslecks“ nicht bewusst.

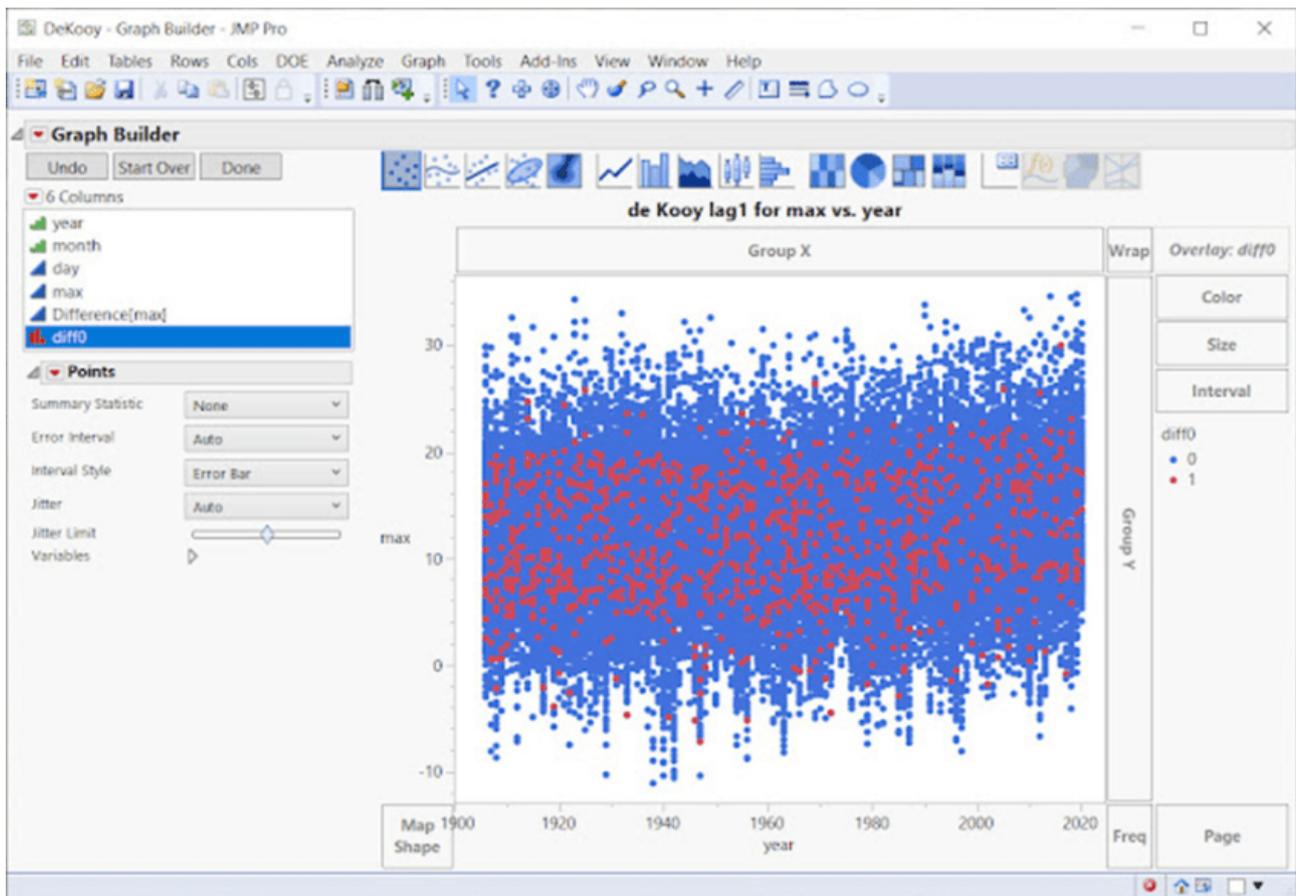
Ein einfaches Dark-Data-Szenario wird hier ganz einfach durch die erste Differenz einer Zeitreihe erstellt. (Die erste Differenzierung ist eine bekannte Methode in der Zeitreihenanalyse).

Ermitteln Sie die Differenz zwischen Temperatur 1 und Temperatur 2, dann die Differenz zwischen Temperatur 2 und Temperatur 3 und so weiter. (unten)

maxv22	diffmaxv22
24	-
24.8	0.8
25.7	0.9
32.8	7.1
32.7	-0.1
26.7	-6
28.9	2.2
30.8	1.9
30.1	-0.7
26.7	-3.4
24.3	-2.4
24.7	0.4

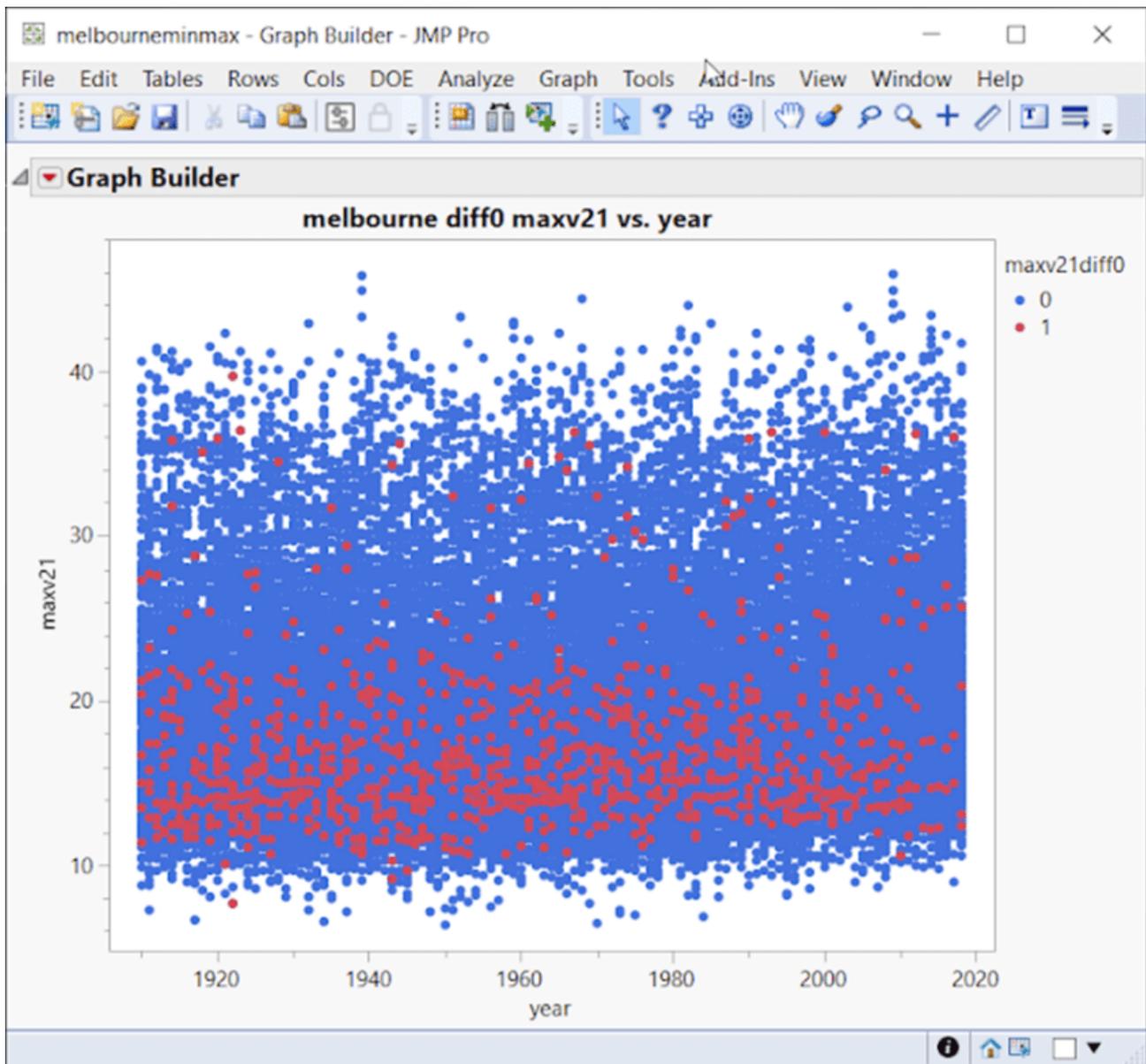
Wenn die Differenz zwischen zwei aufeinanderfolgenden Tagen gleich Null ist, haben die beiden gepaarten Tage die gleiche Temperatur. Auf diese Weise kann man schnell und einfach feststellen, ob die beiden Tage die gleiche Temperatur haben.

Intuitiv würde man eine Zufallsverteilung ohne offensichtliche Häufungen erwarten. Block-Bootstrap-Simulationen, bei denen die Autokorrelation erhalten bleibt, bestätigen dies. Einige der untersuchten europäischen Zeitreihen weisen zum Beispiel eine gleichmäßige Verteilung der gepaarten Tage auf:



Oben ist deKooy in den Niederlanden mit einer ziemlich gleichmäßigen Verteilung zu sehen. Schweden ist sehr ähnlich. Diff0 in der Graphik bezieht sich auf die Tatsache, dass der Unterschied zwischen einem Temperaturpaar bei Anwendung der obigen Technik der ersten Differenz gleich Null ist, was bedeutet, dass die beiden Tage identische Temperaturen aufweisen. Die roten Punkte zeigen die Streuung an.

Schauen wir uns unten Melbourne, Australien, an:

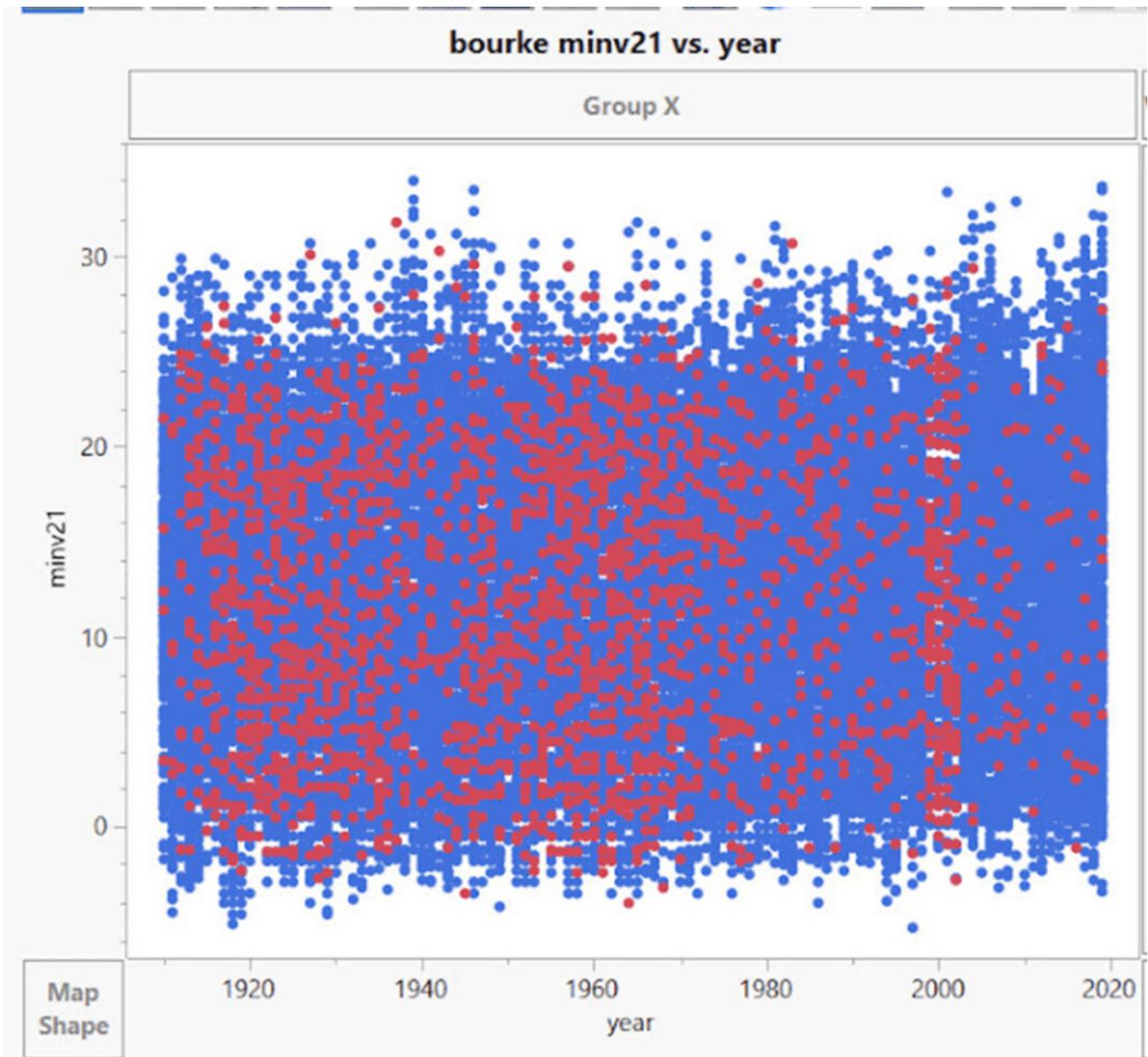


Die gepaarten Tage mit gleichen Temperaturen sind im kühleren Teil des Diagramms zusammengefasst. Sie verjüngen sich nach 2010 oder so (weniger rote Punkte). Die Daten für Melbourne stammen von verschiedenen Standorten, wobei 2014 ein Wechsel vom Standort des BOM Regional Office (86071) zum 2 km entfernten Standort des Olympic Park (86338) erfolgte.

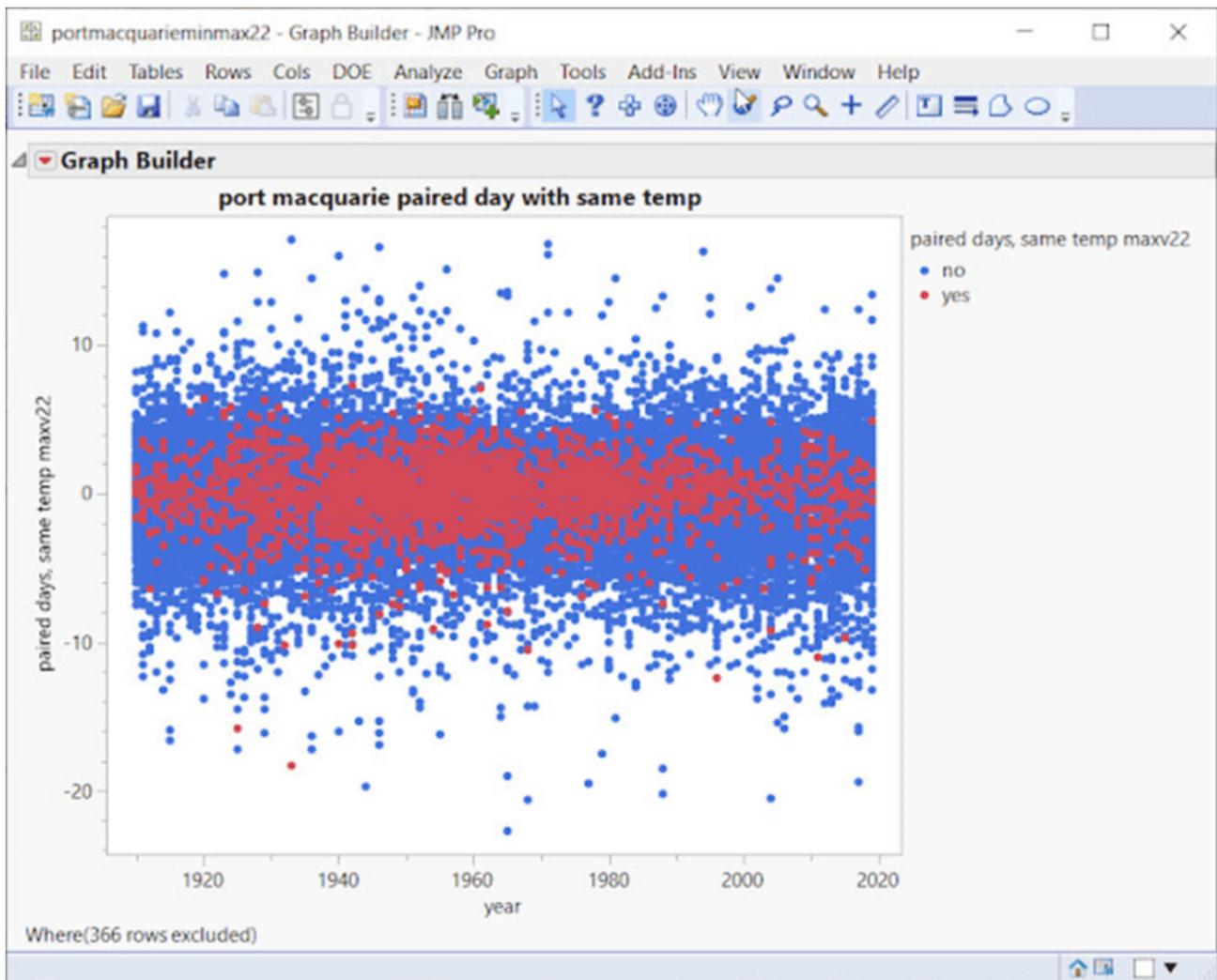
Unten ist Bourke zu sehen, und auch hier sind die Daten mit roten Punkten gebündelt.

Aus dem BOM Stationskatalog:

„Der derzeitige Standort (48245) ist eine automatische Wetterstation auf der Nordseite des Flughafens Bourke ... Er wurde im Dezember 1998 in Betrieb genommen, 700 m nördlich des früheren Flughafenstandortes, aber mit nur minimaler Überlappung. Diese Daten werden seit dem 1. Januar 1999 in ACORN-SAT verwendet.“



Unten ist Port Macquarie zu sehen, wo es zwischen 1940 und 1970 zu einer extrem dichten Häufung kommt:



Diese Daten stammen aus bereinigten ACORN-SAT-Quellen, nicht aus Rohdaten. Sie variieren mit den Anpassungen, in vielen Fällen gibt es sehr große Unterschiede vor und nach den Anpassungen.

Die Hauptstädte variieren um 3-4% der gepaarten Daten. Die Länderstationen können bei einigen Nischengruppen bis zu 20% betragen.

Die Hypothese lautet wie folgt: Die am stärksten geclusterten Datenpunkte sind die am stärksten manipulierten Datenbereiche.

Außerdem können einige der roten Punktcluster visuell mit den im Katalog vermerkten Standortänderungen korreliert werden.

Betrachten wir einen sehr dichten Punkt in Port Macquarie, 1940-1970:

Source	dayofyear	dayofweek	day	month	year	bin	maxv1	maxv2	maxv21	maxv22	diffmaxv22	maxraw	minv1	minv2	minv21	minv22	diffminv22	m
18269	7	Thursday	7	1	1960	1960	26.8	26.3	26.3	26.3	0.6	25.6	13.7	12.9	12.9	12.1	0.6	
18270	8	Friday	8	1	1960	1960	26.8	26.3	26.3	26.3	0	25.6	15.6	14.7	14.7	13.9		1.8
18271	9	Saturday	9	1	1960	1960	26.8	26.3	26.3	26.3	0	25.6	15.6	14.7	14.7	13.9		0
18272	10	Sunday	10	1	1960	1960	26.8	26.3	26.3	26.3	0	25.6	16.5	16.2	16.2	15.9		2
18273	11	Monday	11	1	1960	1960	26.8	26.3	26.3	26.3	0	25.6	17.2	17.2	17.2	16.8		0.9
18274	12	Tuesday	12	1	1960	1960	26.8	26.3	26.3	26.3	0	25.6	17.2	17.2	17.2	16.8		0
18275	13	Wednesday	13	1	1960	1960	27.6	26.8	26.8	26.8	0.5	26.1	18.0	17.8	17.8	17.4		0.6
18276	14	Thursday	14	1	1960	1960	27.6	26.8	26.8	26.8	0	26.1	16.2	15.8	15.8	15.2		-2.2
18277	15	Friday	15	1	1960	1960	27.6	26.8	26.8	26.8	0	26.1	16.2	15.8	15.8	15.2		0
18278	16	Saturday	16	1	1960	1960	27.6	26.8	26.8	26.8	0	26.1	18.0	17.8	17.8	17.4		2.2
18279	17	Sunday	17	1	1960	1960	24.8	24.5	24.5	24.5	-2.3	23.9	17.2	17.2	17.2	16.8		-0.6
18280	18	Monday	18	1	1960	1960	24.3	23.9	23.9	23.9	-0.6	22.8	17.2	17.2	17.2	16.8		0
18281	19	Tuesday	19	1	1960	1960	23.8	23.2	23.2	23.2	-0.7	22.2	17.2	17.2	17.2	16.8		0
18282	20	Wednesday	20	1	1960	1960	21.6	21.0	21.0	21	-2.2	20.0	16.2	15.8	15.8	15.2		-1.6
18283	21	Thursday	21	1	1960	1960	25.8	24.5	24.5	24.5	3.5	23.9	15.6	14.7	14.7	13.9		-1.3
18284	22	Friday	22	1	1960	1960	24.8	24.5	24.5	24.5	0	23.9	17.2	17.2	17.2	16.8		2.9
18285	23	Saturday	23	1	1960	1960	28.9	28.0	28.0	28	3.5	26.7	16.5	16.2	16.2	15.9		-0.9
18286	24	Sunday	24	1	1960	1960	28.9	28.0	28.0	28	0	26.7	12.5	12.3	12.3	11.5		-4.4
18287	25	Monday	25	1	1960	1960	29.8	28.9	28.9	28.9	0.9	27.2	12.5	12.3	12.3	11.5		0
18288	26	Tuesday	26	1	1960	1960	28.9	28.0	28.0	28	-0.9	26.7	19.1	19.0	19.0	19		7.5
18289	27	Wednesday	27	1	1960	1960	28.9	28.0	28.0	28	0	26.7	18.5	18.2	18.2	17.8		-1.2
18290	28	Thursday	28	1	1960	1960	28.9	28.0	28.0	28	0	26.7	20.3	20.4	20.4	20.4		2.6
18291	29	Friday	29	1	1960	1960	28.9	28.0	28.0	28	0	26.7	20.3	20.4	20.4	20.4		0
18292	30	Saturday	30	1	1960	1960	27.6	26.8	26.8	26.8	-1.2	26.1	20.3	20.4	20.4	20.4		0
18293	31	Sunday	31	1	1960	1960	28.9	28.0	28.0	28	1.2	26.7	18.5	18.2	18.2	17.8		-2.6
18294	32	Monday	1	2	1960	1960	28.1	27.5	27.5	27.4	-0.6	26.7	20.0	20.4	20.4	20.4		2.6
18904	33	Tuesday	2	2	1960	1960	33.6	33.6	33.6	33.6	-4.8	33.3	10.7	10.8	10.8	10.6		-1.8

Es ist sofort ersichtlich, dass viele Tage doppelte Sequenzen aufweisen. Auch wenn es sich um kürzere Sequenzen handelt, sind es immer noch mehr, als man zufällig erwarten würde, aber beachten Sie auch die seltsamen systematischen Abstände und Lücken unten.

Mehr über Port Macquarie, BOM-Stationskatalog:

„Es gab eine Verschiebung um 90 m seewärts (die genaue Richtung ist unklar) im Januar 1939 und eine Verschiebung um 18 m nach Südwesten am 4. September 1968... Der derzeitige Standort (60139) ist eine automatische Wetterstation am Flughafen Port Macquarie... Er befindet sich an der südöstlichen Seite der Landebahn des Flughafens. Sie wurde 1995 in Betrieb genommen, wird aber erst seit Januar 2000 im ACORN-SAT-Datensatz verwendet, da es Probleme mit den ersten Daten gab. Im Oktober 2020 wurde auf dem Flughafengelände ein neuer Standort (60168) eingerichtet, der den derzeitigen Standort zu gegebener Zeit ablösen soll“.

Das folgende Beispiel bezieht sich auf die frühen 1950er Jahre:

Source	year	bin	maxv1	maxv2	maxv21	maxv22	diffmaxv22	maxraw	minv1	minv2	minv21	minv22	diffminv22	minraw
15880	1953	1950	19.2	18.4	18.4	18.4	1	4.5	5.2	5.2	4.6	3.5		
15881	1953	1950	19.0	18.3	18.3	18.3	-0.1	4.5	5.2	5.2	4.6	0		
15882	1953	1950	16.5	16.0	16.0	16	-2.3	1.8	2.7	2.7	2.6	-2		
15883	1953	1950	16.8	16.3	16.3	16.3	0.3	3.1	4.8	4.8	4.3	1.7		
15884	1953	1950	17.2	16.6	16.6	16.6	0.3	3.1	4.8	4.8	4.3	0		
15885	1953	1950	19.2	18.4	18.4	18.4	1.8	1.8	2.7	2.7	2.6	-1.7		
15886	1953	1950	17.8	17.4	17.4	17.4	-1	1.0	2.2	2.2	1.8	-0.8		
15887	1953	1950	18.4	17.9	17.9	17.9	0.5	0.1	0.8	0.8	0.4	-1.4		
15888	1953	1950	18.2	17.7	17.7	17.6	-0.3	0.0	0.2	0.2	-0.3	-0.7		
15889	1953	1950	17.5	17.1	17.1	17.1	-0.5	0.0	0.2	0.2	-0.3	0		
15890	1953	1950	17.5	17.1	17.1	17.1	0	0.3	0.5	0.5	0.1	0.4		
15891	1953	1950	18.2	17.7	17.7	17.6	0.5	0.0	0.2	0.2	-0.3	-0.4		
15892	1953	1950	18.8	18.4	18.4	18.3	0.7	0.0	0.2	0.2	-0.3	0		
15893	1953	1950	18.2	17.7	17.7	17.6	-0.7	0.0	0.2	0.2	-0.3	0		
15894	1953	1950	19.9	19.3	19.3	19.3	1.7	-0.3	-0.1	-0.1	-0.5	-0.2		
15895	1953	1950	21.1	20.3	20.3	20.3	1	2.6	2.6	2.6	2.1	2.6		
15896	1953	1950	14.4	14.3	14.3	14.2	-6.1	10.7	10.6	10.6	10.4	8.3		
15897	1953	1950	11.6	11.5	11.5	11.4	-2.8	2.6	2.6	2.6	2.1	-8.3		
15898	1953	1950	14.4	14.3	14.3	14.2	2.8	2.6	2.6	2.6	2.1	0		
15899	1953	1950	15.0	14.8	14.8	14.7	0.5	2.6	2.6	2.6	2.1	0		
15900	1953	1950	16.4	16.0	16.0	15.9	1.2	2.6	2.6	2.6	2.1	0		
15901	1953	1950	18.2	17.7	17.7	17.6	1.7	2.6	2.6	2.6	2.1	0		
15902	1953	1950	22.5	21.3	21.3	21.3	3.7	0.3	0.5	0.5	0.1	-2		
15903	1953	1950	18.2	17.7	17.7	17.6	-3.7	0.3	0.5	0.5	0.1	0		
15904	1953	1950	21.1	20.3	20.3	20.3	2.7	0.3	0.5	0.5	0.1	0		
15905	1953	1950	18.2	17.7	17.7	17.6	-2.7	5.2	5.3	5.3	4.9	4.8		
15906	1953	1950	18.2	17.7	17.7	17.6	0	2.6	2.6	2.6	2.1	-2.8		

Hier haben wir Lücken von 1 und 3 zwischen den Sequenzen.

Unten haben wir Lücken von 8 zwischen den Sequenzen:

	year	bin	maxv1	maxv2	maxv21	maxv22	diffmaxv22	maxraw	minv1	minv2	minv21	minv22	diffminv22	minraw
15902	1953	1950	22.5	21.3	21.3	21.3	3.7	•	0.3	0.5	0.5	0.1		-2
15903	1953	1950	18.2	17.7	17.7	17.6	-3.7	•	0.3	0.5	0.5	0.1		0
15904	1953	1950	21.1	20.3	20.3	20.3	2.7	•	0.3	0.5	0.5	0.1		0
15905	1953	1950	18.2	17.7	17.7	17.6	-2.7	•	5.2	5.3	5.3	4.9		4.8
15906	1953	1950	18.2	17.7	17.7	17.6	0	•	2.6	2.6	2.6	2.1		-2.8
15907	1953	1950	22.5	21.3	21.3	21.3	3.7	•	2.9	2.8	2.8	2.2		0.1
15908	1953	1950	22.8	21.4	21.4	21.4	0.1	•	0.0	0.2	0.2	-0.3		-2.5
15909	1953	1950	19.1	18.7	18.7	18.5	-2.9	•	-0.6	-0.5	-0.5	-0.9		-0.6
15910	1953	1950	22.5	21.3	21.3	21.3	2.8	•	9.0	8.9	8.9	8.7		9.6
15911	1953	1950	17.5	17.1	17.1	17.1	-4.2	•	8.6	8.4	8.4	8		-0.7
15912	1953	1950	18.8	18.4	18.4	18.3	1.2	•	0.0	0.2	0.2	-0.3		-8.3
15913	1953	1950	19.2	18.8	18.8	18.7	0.4	•	3.3	4.1	4.1	3.4		3.7
15914	1953	1950	21.1	20.3	20.3	20.3	1.6	•	3.3	4.1	4.1	3.4		0
15915	1953	1950	18.2	17.7	17.7	17.6	-2.7	•	7.0	7.2	7.2	6.7		3.3
15916	1953	1950	18.2	17.7	17.7	17.6	0	•	5.1	4.7	4.7	4		-2.7
15917	1953	1950	19.2	18.8	18.8	18.7	1.1	•	1.5	1.7	1.7	1.1		-2.9
15918	1953	1950	22.0	21.1	21.1	21.1	2.4	•	0.3	0.5	0.5	0.1		-1
15919	1953	1950	21.4	20.2	20.2	20.3	-0.8	•	1.1	2.0	2.0	1.1		1
15920	1953	1950	20.8	20.0	20.0	20.1	-0.2	•	2.8	3.7	3.7	3		1.9
15921	1953	1950	20.3	19.5	19.5	19.6	-0.5	•	7.9	8.1	8.1	7.6		4.6
15922	1953	1950	20.3	19.5	19.5	19.6	0	•	12.1	12.7	12.7	12.6		5
15923	1953	1950	17.7	17.1	17.1	17.2	-2.4	•	8.7	8.8	8.8	8.4		-4.2
15924	1953	1950	23.2	22.3	22.3	22.4	5.2	•	8.7	8.8	8.8	8.4		0
15925	1953	1950	19.1	18.6	18.6	18.7	-3.7	•	2.8	3.7	3.7	3		-5.4
15926	1953	1950	19.6	19.0	19.0	19.1	0.4	•	2.8	3.7	3.7	3		0
15927	1953	1950	19.6	19.0	19.0	19.1	0	•	-0.7	0.4	0.4	-0.6		-3.6
15928	1953	1950	20.8	20.1	20.1	20.1	1	•	-0.7	0.4	0.4	-0.6		0

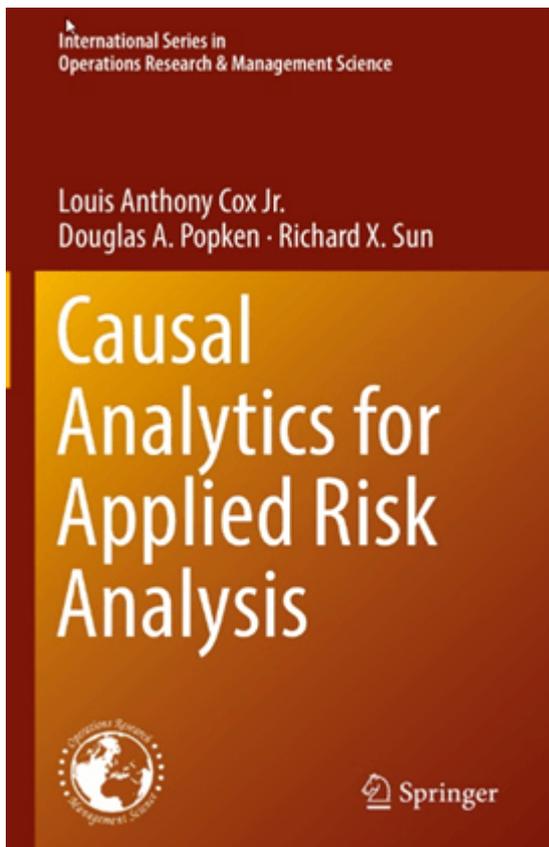
Unten – jetzt haben wir Lücken von 2, dann 3, dann 4, dann 5. Denken Sie daran, dass die meisten Zeitreihen viele dieser „Fake“-Sequenzen haben!

	year	bin	maxv1	maxv2	maxv21	maxv22	diffmaxv22	maxraw	minv1	minv2	minv21	minv22	diffminv22	minraw
15938	1953	1950	18.3	17.8	17.8	17.8	1.3	-2.1	-1.2	-1.2	-2.1	0		
15939	1953	1950	19.1	18.6	18.6	18.7	0.9	6.9	6.6	6.6	6	8.1		
15940	1953	1950	18.3	17.8	17.8	17.8	-0.9	6.9	6.6	6.6	6	0		
15941	1953	1950	19.1	18.6	18.6	18.7	0.9	1.9	3.1	3.1	2.4	-3.6		
15942	1953	1950	19.1	18.6	18.6	18.7	0	0.0	0.8	0.8	0	-2.4		
15943	1953	1950	19.1	18.6	18.6	18.7	0	6.9	6.6	6.6	6	6		
15944	1953	1950	16.8	16.4	16.4	16.5	-2.2	6.9	6.6	6.6	6	0		
15945	1953	1950	20.3	19.5	19.5	19.6	3.1	4.3	5.0	5.0	4.1	-1.9		
15946	1953	1950	18.3	17.8	17.8	17.8	-1.8	6.9	6.6	6.6	6	1.9		
15947	1953	1950	13.9	13.0	13.0	13.1	-4.7	10.4	10.1	10.1	9.7	3.7		
15948	1953	1950	18.3	17.2	17.2	17.3	4.2	11.4	11.2	11.2	10.9	1.2		
15949	1953	1950	25.3	23.1	23.1	23.2	5.9	11.4	11.2	11.2	10.9	0		
15950	1953	1950	32.6	31.1	31.1	31.1	7.9	4.6	4.6	4.6	3	-7.9		
15951	1953	1950	28.4	26.8	26.8	26.8	-4.3	8.1	8.0	8.0	7.1	4.1		
15952	1953	1950	26.5	25.3	25.3	25.3	-1.5	4.3	4.1	4.1	2.5	-4.6		
15953	1953	1950	17.7	17.6	17.6	17.6	-7.7	1.1	2.0	2.0	0.7	-1.8		
15954	1953	1950	21.9	21.6	21.6	21.6	4	0.4	1.4	1.4	0	-0.7		
15955	1953	1950	24.0	22.8	22.8	22.7	1.1	0.4	1.4	1.4	0	0		
15956	1953	1950	24.9	23.9	23.9	23.9	1.2	4.6	4.6	4.6	3	3		
15957	1953	1950	21.9	21.6	21.6	21.6	-2.3	4.3	3.9	3.9	2.3	-0.7		
15958	1953	1950	21.9	21.6	21.6	21.6	0	7.0	7.1	7.1	6	3.7		
15959	1953	1950	21.2	20.4	20.4	20.4	-1.2	15.1	15.3	15.3	14.8	8.8		
15960	1953	1950	19.4	19.0	19.0	19	-1.4	8.8	9.4	9.4	8.8	-6		
15961	1953	1950	17.5	17.1	17.1	17.1	-1.9	1.9	2.4	2.4	1	-7.8		
15962	1953	1950	19.4	19.0	19.0	19	1.9	1.9	2.4	2.4	1	0		
15963	1953	1950	19.4	19.0	19.0	19	0	1.9	2.4	2.4	1	0		
15964	1953	1950	22.0	22.1	22.1	22.1	2.1	4.6	4.6	4.6	3	2		

Nachweis der Kausalität

CO₂ und Erwärmung beruhen auf Korrelation, aber wir alle wissen, dass Korrelation nicht gleichbedeutend mit Kausalität ist.

Tony Cox ist ein führender Experte auf dem Gebiet der Kausalanalyse und verfügt über eine Software zur Prüfung der Kausalität.



Mit Hilfe von CART-Entscheidungsbäumen und einigen neueren Algorithmen testen wir, welche Variablen kausal sind, wenn wir gepaarte Tage mit der gleichen Temperatur betrachten.

Wenn wir die **Bourke-Minimum-Temperaturen** über die gesamte Zeitreihe verwenden und als Ziel „MIN Paired Days, Same Temp“ festlegen, werden von CART Regeln erstellt, um prädiktive kausale Zusammenhänge zu finden:

```
/*Rules for terminal node 3*/  
  
if  
  
MAXpaired days same temp <= 0.06429 &&  
  
MINadjustments > -0.62295 &&  
  
MINadjustments <= -0.59754  
  
terminalNode = 3;
```

Dies bedeutet, dass, wenn das Maximum der gepaarten Tage gleicher Temperatur weniger als 0,06429 beträgt UND die Anpassungen bei den MIN-Temperaturen zwischen -0,62295 und -0,59754 liegen, Knoten 3 zutrifft und ein hoch prädiktiver Cluster von 50 % gefunden wurde.

ANMERKUNG: Gepaarte Tage mit gleicher Temperatur für die MAX-Reihe und Anpassungen durch BOM wurden als kausale Vorhersage erkannt!

Port Macquarie Minimum-Temperatur-Zeitreihe

Ziel: Minimum-gepaarte Tage gleicher Temperaturen

```
/*Rules for terminal node 3*/
```

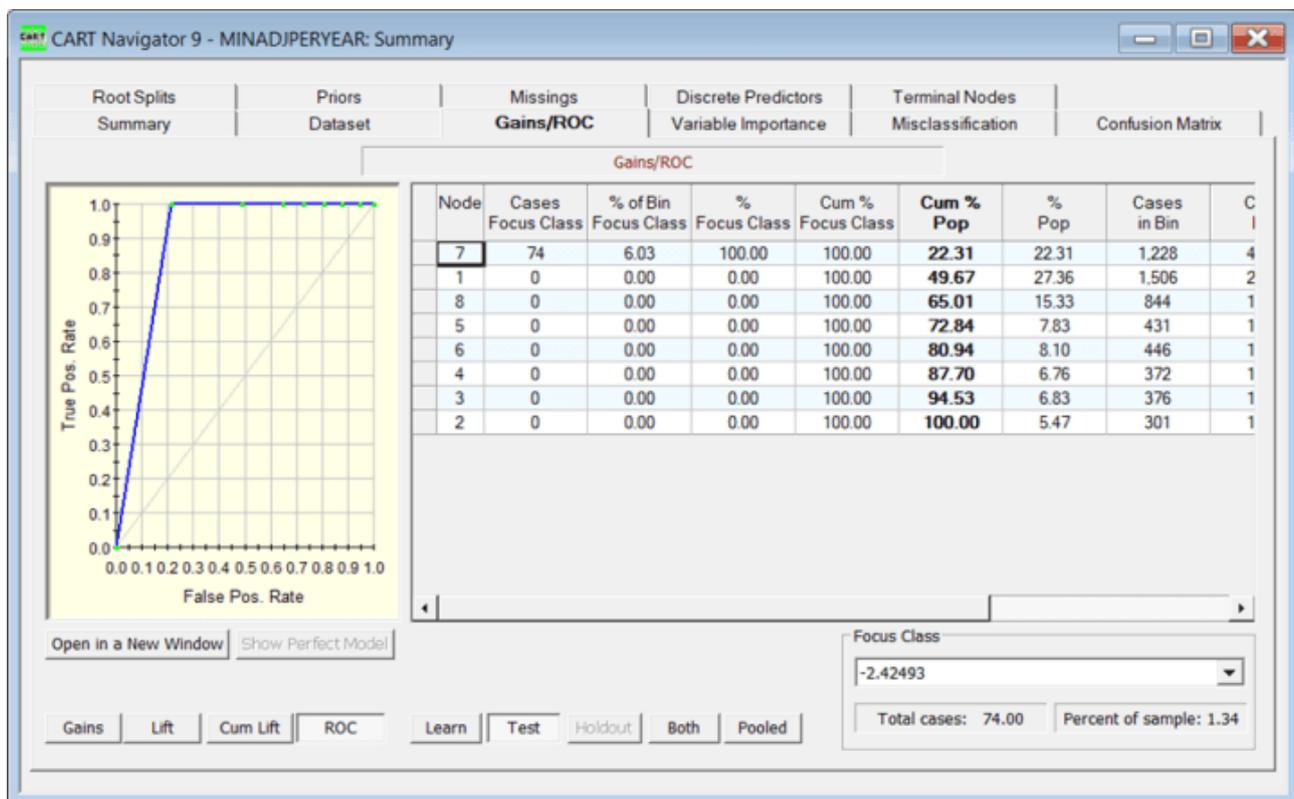
```
if
```

```
MAXadjustments > 0.53663 &&
```

```
MAXpaired days same temps <= 0.02329
```

```
terminalNode = 3;
```

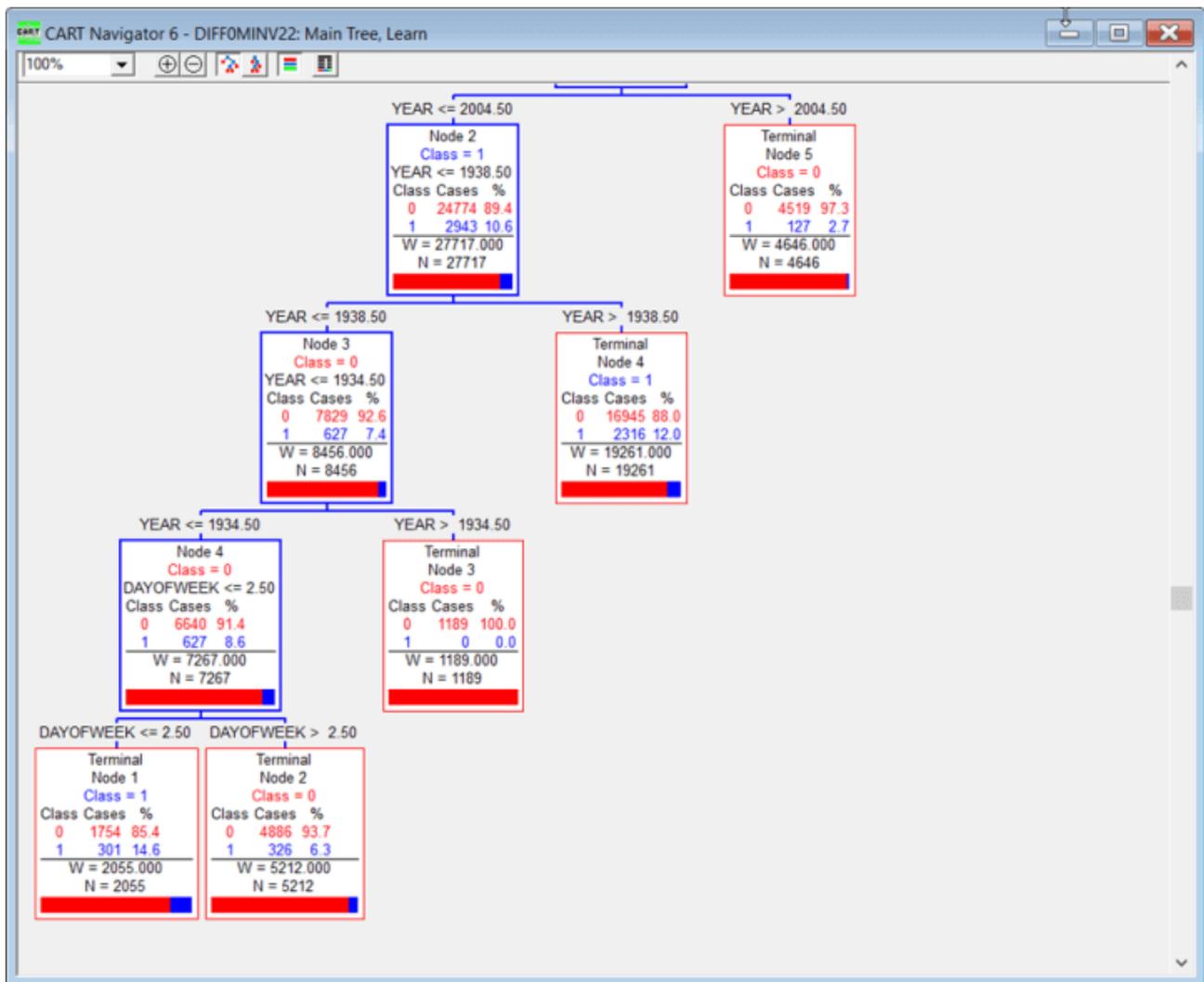
```
class prob= 20%
```



Die obige ROC-Kurve für Port Macquarie zeigt ein sehr prädiktives Modell, wobei die prädiktivste kausale Variable MAX gepaarte Tage, gleiche Temperaturen ist, wobei große Max-Anpassungen kausal für die Vorhersage von MIN gepaarten Tagen, gleiche Temperaturen sind!

Unten ist die CART-Baum-Ausgabe für die Suche nach dem Ziel in Palmerville:

MIN gepaarte Tage, gleiche Temperaturen:



Hier findet es Wochentage und Jahre vorhersagend kausal. Sie lesen den Baum, indem Sie einen Zweig Wenn wahr abwärts gehen und die Wahrscheinlichkeit ablesen, dass der Cluster wahr ist.

In diesen und weiteren getesteten Fällen ist das kausale Vorhersageergebnis des Ziels *Minimum Paired Days With Same Temps* der **Wochentag, das Jahr und das Ausmaß der Anpassungen!**

Erinnern wir uns: Gepaarte Tage mit gleichen Temperaturen waren ein Hinweis auf doppelte oder „gefälschte Läufe“ von Temperaturen. Je höher die Clusterkonzentration, desto mehr Sequenzen wurden gefunden.

Offensichtlich haben die Daten ernsthafte Probleme, wenn der Wochentag bei der Modellierung eine Rolle spielt. BOM-Anpassungen sind ebenfalls ursächlich für die Bildung von Clustern „falscher“ Sequenzen.

Hier sind einige Fragen, die sich aus Toms Arbeit ergeben: Warum mussten wir diese grundlegende Qualitätskontrolle der primären Temperatur-Rohdaten der BOM im Bürgerwissenschaftler-Modus durchführen?

Wusste die BOM bereits von der großen Anzahl von Fällen korrumpierter und/oder verdächtiger Daten, teilte aber weiterhin mit, dass die Rohdaten im Wesentlichen unangetastet blieben, abgesehen von der Entfernung von Tippfehlern und einigen offensichtlichen Ausreißern?

Wie kann man angesichts der Tatsache, dass sowohl die „Rohdaten“ als auch die „ACORN-SAT“-Daten inhärente Probleme aufweisen, behaupten, dass aus diesen Zahlen abgeleitete Verteilungen Gültigkeit haben?

Wie kann man also ein auf Verteilungen basierendes Argument rechtfertigen, um die Anwendung des zentralen Grenzwertsatzes zu ermöglichen, wenn so viele Verteilungen verfälscht sind?

Kann man das Gesetz der großen Zahlen gültig auf Daten anwenden, die von Menschen erfunden wurden und die keine Beobachtungen sind?

Wie schätzt man die Messunsicherheit bei Daten, die von Menschen erfunden wurden?

Wo ist das Handbuch für die Schätzung des Vertrauens in unterstellte Werte? Ist es genau?

Sind Beobachtungen dieser Art überhaupt für den Zweck geeignet?

Warum wurde dem BOM von [Experten](#), die den Wert von ACORN-SAT untersucht haben, ein „Freifahrtschein“ ausgestellt? Warum haben sie keine Unregelmäßigkeiten in den Daten gefunden?

Ist es möglich, diese Millionen von Beobachtungen „nachzujustieren“, um sicherzustellen, dass sie Tests der hier beschriebenen Art bestehen?

Oder sollte die BOM darauf hinweisen, dass in Zukunft nur noch Daten seit (sagen wir) dem 1. November 1996 verwendet werden dürfen? (Das war der Zeitpunkt, an dem viele BOM-Wetterstationen von manueller Thermometrie auf elektronische AWS-Beobachtungen umstellten).

Gibt es auch andere Bereiche der Beobachtungswissenschaft, die ähnliche Probleme haben, oder stellt das BOM einen Sonderfall dar?

Wir mussten diesen langen Aufsatz wirklich kürzen. Hätten die WUWT-Leser gerne einen vierten Teil dieser Serie, der viel mehr über diese Temperaturen zeigt?

Schlussfolgerungen

Wir kritisieren in keiner Weise die vielen Beobachter, die die ursprünglichen Temperaturdaten aufgezeichnet haben. Unsere Sorge gilt den nachträglichen Änderungen der Originaldaten, wobei wir daran erinnern, dass das australische Bureau of Meteorology (BoM) über einen 77 Millionen Dollar teuren Cray XC-40 Supercomputer namens Australis verfügt. Ein einziger Tastendruck an einem Tag auf Australis kann die

geduldige, hingebungsvolle Arbeit vieler Menschen über viele Jahrzehnte hinweg, wie die der ursprünglichen Beobachter, verändern.

Ich bitte Sie ernsthaft um Antworten auf diese Fragen, denn es gibt ein großes Problem. In der Wissenschaft der Metrologie wird häufig beschrieben, dass Messungen auf Primärnormale zurückgeführt werden müssen, wie z. B. auf den 1-Meter-Stab [das Ur-Meter], der in Frankreich für Längenmessungen verwendet wird. In dem von uns untersuchten Zweig der Meteorologie haben wir versucht, die Primärdaten nachzuweisen, sind aber gescheitert. Daher muss der Schwerpunkt auf ungewöhnliche Muster und Ereignisse in den derzeit verwendeten Daten gelegt werden. Genau das hat Tom getan. Es gibt eine sehr große Anzahl von Unregelmäßigkeiten.

Die gegenwärtig verwendeten Daten werden von der BOM verpackt und an globale Zentren geschickt, wo Schätzungen der globalen Temperatur vorgenommen werden. Tom hat zu Beginn dieses Aufsatzes gezeigt, wie das BOM ein globales Erwärmungsmuster dargestellt und vermutlich gebilligt hat, das sich durch Veränderungen im 21. Jahrhundert verstärkt hat.

Ist die angezeigte Erwärmung real oder ein Artefakt der Datenmanipulation?

Diese Frage ist von grundlegender Bedeutung, denn die globale Erwärmung hat inzwischen zu Befürchtungen einer „existenziellen Krise“ und zu Maßnahmen geführt, die darauf abzielen, die Nutzung fossiler Brennstoffe in vielen Bereichen einzustellen. Dies bedeutet enorme Veränderungen für die gesamte Gesellschaft, so dass die Daten, die zu diesen Veränderungen führen, von hoher Qualität sein müssen.

Wir zeigen, dass sie von geringer Qualität sind. Das gilt für Australien. Was ist über Ihre eigenen Länder bekannt?

Für uns Australier ist es wichtig genug, erneut zu fordern, dass unabhängige Untersuchungen, ja sogar eine königliche Kommission, eingesetzt werden, um festzustellen, ob diese Temperaturmessungen mit ihren schwerwiegenden Folgen zweckmäßig sind.

Link:

<https://wattsupwiththat.com/2022/10/14/uncertainty-of-measurement-of-routine-temperatures-part-iii/>

Übersetzt von [Christian Freuer](#) für das EIKE