

Informative Grafiken und Grafikskripte spielerisch erstellen

Bernd Heinen
SAS Institute
In der Neckarhelle 162
69118 Heidelberg
Bernd.heinen@jmp.com

Zusammenfassung

Grafiken dienen der Information – auch der eigenen - und der Kommunikation. Längst sind Grafiken nicht mehr schmückendes Beiwerk sondern zentrales Element von analytischen Berichten. Gerade auch bei der Einarbeitung in neue Datensätze zeigen Grafiken recht schnell, wo man am Ehesten auf Zusammenhänge oder Abweichungen stößt. Mit den Analyse- und Grafikkomponenten von JMP werden interaktiv Darstellungen erzeugt, die Informationen präzise transportieren. Gerade, wenn es um die Zusammenhänge zwischen mehreren Variablen geht, bewähren sich auch Grafiken, die interaktiv verändert oder simultan gesteuert werden. Einerseits lassen sich solche Grafiken leicht per Maussteuerung erstellen, andererseits erzeugen einmal erstellte Grafiken aber auch ihr eigenes Skript, so dass sie jederzeit mit aktualisierten Daten neu erstellt werden können. Und mit JMPlive gibt es einen Webserver, auf dem solche Grafikskripte gespeichert werden können, so dass autorisierte Nutzer die Interaktivität auch über ein Webinterface nutzen können, ohne über eine eigene JMP Installation zu verfügen.

Schlüsselwörter: Grafik, Balkendiagramm, Likert-Skala, Regression, Dynamische Grafik, JMP, Grafik Praxisempfehlungen

1 Einleitung

Die Bedeutung von Grafiken wird wohl nicht bestritten, aber es ist noch lange nicht so, dass die Möglichkeit, ad-hoc Grafiken zu erstellen, an jedem Arbeitsplatz zur Verfügung steht. Ich sehe – auch hier auf der KSFE – immer wieder Vorträge zur Programmierung von Grafiken, die dann zwar leicht abgerufen werden können, aber eben immer der einmal erstellten Form folgen. Dabei gehören grafische Analysen unbedingt in die Hand der Fachleute eines Unternehmens, denn hier entstehen die relevanten Fragestellungen oder Lösungsansätze und der interaktive Umgang mit Grafiken oder grafischen Datenanalysen schafft die Voraussetzung für effizientes Arbeiten. Ein Beispiel ist hier der Bericht eines Qualitätsmanagers, der sagte, dass er nur durch das Kombinieren mehrerer Faktoren in einer Grafik Bedingungen erkennen konnte, unter denen sich Ausschuss häufte, die man leicht vermeiden konnte. Dadurch konnte er dem Unternehmen 200.000 EUR jährlich sparen.

Mit JMP bietet SAS ein Statistikprogramm an, dass in großem Umfang statistische Analysen auf dem Desktop ermöglicht, sich aber besonders durch seinen innovativen Umgang mit Grafiken auszeichnet. Gleichzeitig sind die angebotenen Grafikmöglichkeiten

auf die Analyse von Daten ausgerichtet, so dass man stets informative Darstellungen erhält. Alle Grafiken, die zu derselben Datei gehören, sind untereinander verknüpft, so dass Datenpunkte, die in einer Grafik ausgewählt werden, auch in anderen Grafiken ausgezeichnet werden. Das schließt auch Analysegrafiken in den statistischen Auswertungen ein, z. B. bei der Modellierung (Abbildung 1)

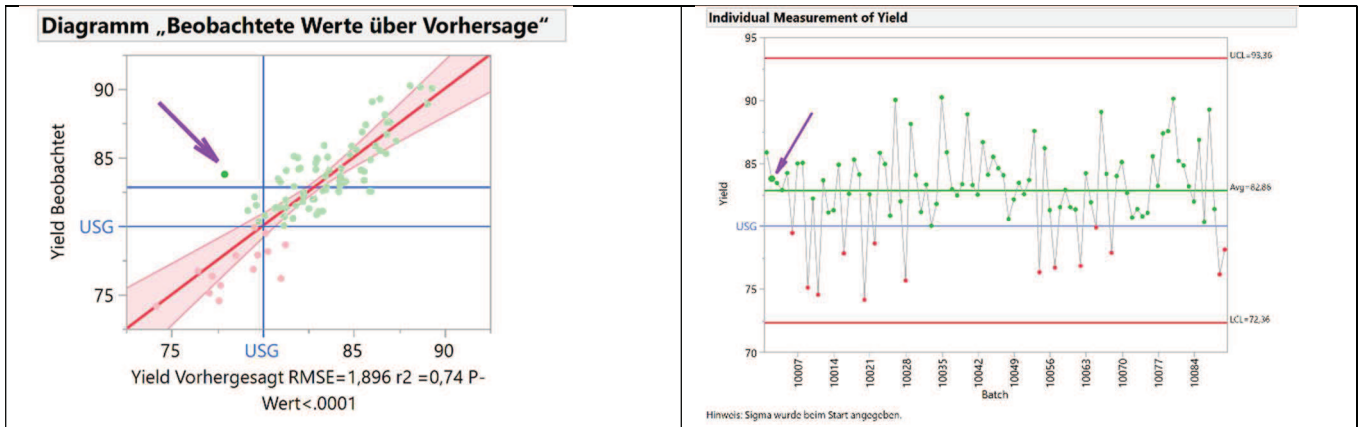


Abbildung 1: Zwei unterschiedliche Grafiken

In der Gegenüberstellung von beobachteten Werten mit den Prognosen eines linearen Modells (links) ist ein Ausreißer klar zu erkennen. Die Regelkarte (rechts) zeigt denselben Punkt als unauffällig bezüglich der Zielgröße „Ertrag“

Auf diese grundlegende Eigenschaft, die es erlaubt, einen Datensatz gleichzeitig aus verschiedenen Blickwinkeln zu betrachten, wird im Folgenden nicht mehr eingegangen. Hier werden in loser Folge Grafiken gezeigt, die über einfache Punktwolken oder Balken hinaus zusätzliche Anforderungen erfüllen oder mehrere Variablen gleichzeitig darstellen. Im Vortrag wurde die Erstellung der Grafiken vorgeführt, für den Druck werden nur die Ergebnisse und wesentliche Eigenschaften zusammengestellt. Der interessierte Leser kann aber gerne Videos zur Erstellung solcher Grafiken ansehen, die im Internet verfügbar sind [1] [2].

Ein Gestaltungselement steht dem Anwender bereits zur Verfügung, bevor er die erste Grafik erstellt. Sehr häufig befindet sich in einem Datensatz eine Spalte, die eine besonders wichtige Information beinhaltet. Das kann eine Gruppierungsvariable sein, eine besonders wichtige Zielgröße oder eine Variable, die das Erfüllen oder Nichterfüllen eines Kriteriums anzeigt. In JMP kann man alle Zeilen nach den Ausprägungen einer Variablen einfärben, so dass die zugehörigen Beobachtungen immer in derselben Farbe dargestellt werden. Es sein denn, in der Grafik selbst wird explizit eine andere Farbgebung gewählt. Auch kann man einzelne Beobachtungen manuell durch spezielle Farben oder Symbole kennzeichnen, so dass sie in allen Darstellungen leicht identifiziert werden können (Abbildung 2).

	Batch	Financial Performance	Yield	Vessel Size	Water Content	Reaction Temperature	
●	1	10001	Accept	85.85	750	1.60	105.
◇	2	10002	Accept	83.77	2000	1.44	100.
●	3	10003	Accept	83.43	750	1.45	100.
●	4	10004	Accept	82.86	500	1.44	98.
●	5	10005	Accept	84.22	2000	1.61	97.
●	6	10006	Reiect	79.46	750	1.29	94.
●	7	10007	Accept	84.98	750	1.36	104.
●	8	10008	Accept	85.03	750	1.51	94.
●	9	10009	Reiect	75.11	2000	1.41	97.
●	10	10010	Accept	82.19	2000	1.51	97.
*	11	10011	Reiect	74.55	2000	1.52	103.
●	12	10012	Accept	83.64	750	1.17	101.
●	13	10013	Accept	81.08	750	1.56	93.

Abbildung 2: Tabellenausschnitt mit Farbmarkierung nach den Werten der Spalte „Financial Performance“ und individuell vergebenen Symbolen.

2 Spezifische vorgefertigte Grafiken

Bevor im Folgenden auf die individuelle Erstellung von Grafiken eingegangen wird, möchte ich empfehlen, einmal „über den Tellerrand“ hinauszuschauen, und Grafiken zu nutzen, die für spezielle Zwecke geschaffen wurden. Meist führen diese ein Schattendasein und sind nur Fachleuten aus entsprechenden Anwendungsgebieten bekannt. Da wir aber hier Software einsetzen, deren Darstellungen immer auf der gleichen Art von Datentabellen aufbauen, ist es leicht, hier auf Entdeckungsreisen zu gehen.

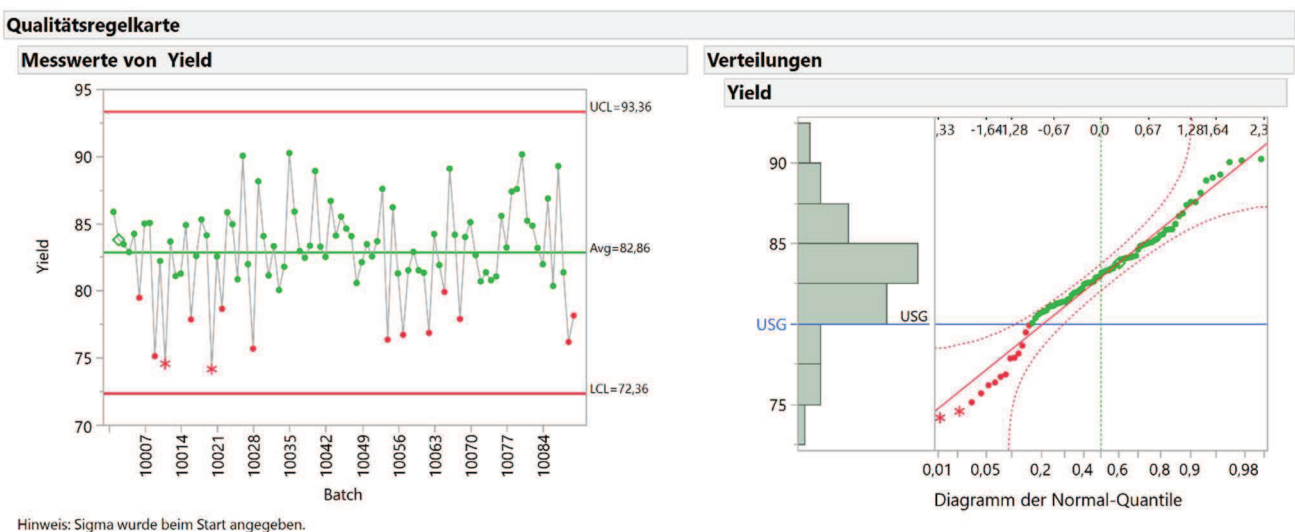


Abbildung 3: Qualitätsregelkarte mit Verteilung und Diagramm der Normal-Quantile

Als Beispiel soll hier eine Regelkarte dienen, die für numerische Werte geeignet ist und die Werte einfach in der Reihenfolge ihres Auftretens in der Datei darstellt. Dazu werden Mittelwert und $\pm 3\sigma$ Grenzen angezeigt (Abbildung 3). Leicht lassen sich die Verteilung und Ausreißer erkennen, und auch unerwartete Trends ausmachen, falls die Daten nicht ohnehin eine zeitliche Abfolge darstellen. Hauptkomponentenanalyse, Entscheidungsbäume oder Textanalyse sind weitere Anwendungen, die mit eigenen grafischen Darstellungen aufwarten, die auch ohne die eigentliche statistische Analyse sehr erhellend sein können.

3 Einfache Kombination mehrerer Variablen

Alle folgenden Darstellungen werden mit der Grafikfunktion „Grafik erstellen“ erstellt. Diese bietet ein Arbeitsfeld an, in dem rund um das eigentliche Grafikfeld Funktionsbereiche angeordnet sind. Eine Variablenliste stellt alle Variablen der aktuellen Tabelle zur Auswahl zur Verfügung. Zieht man eine Variable in ein Funktionsfeld, so werden die Werte der Variablen entsprechend dieser Funktion genutzt (Abbildung 4).

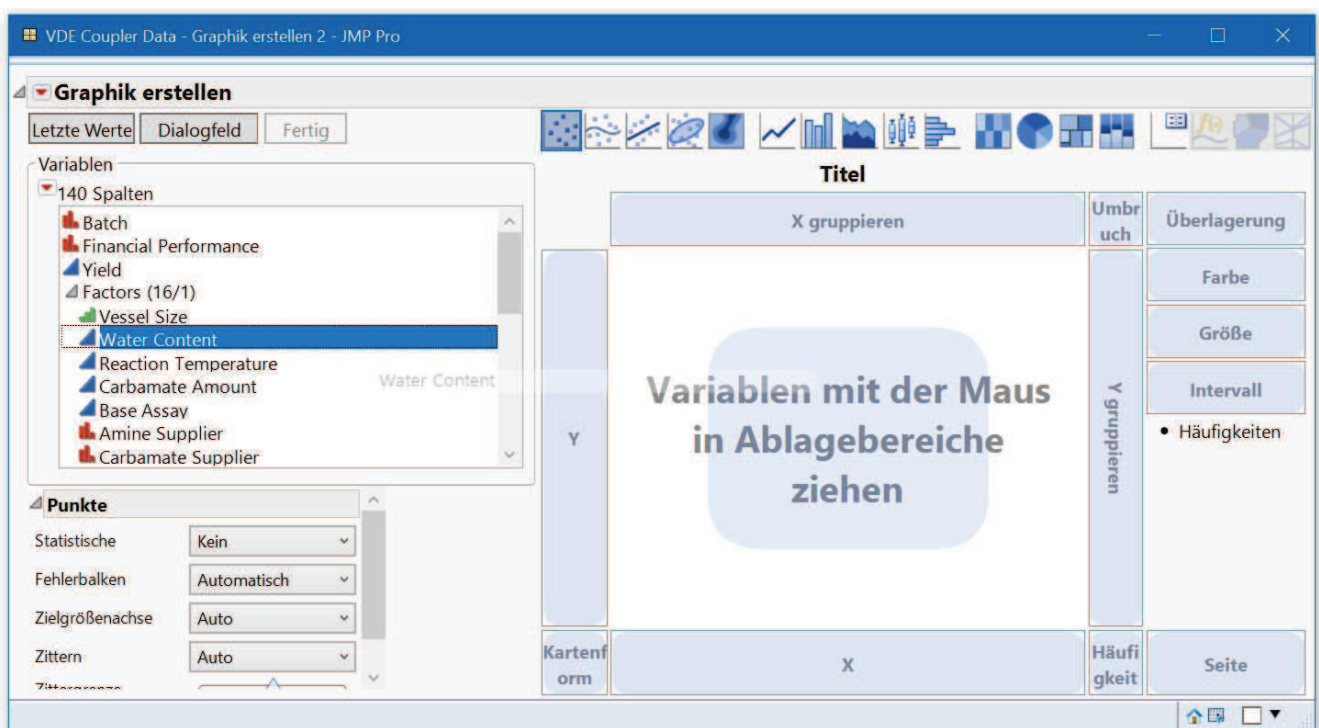


Abbildung 4: Grafikelement, Funktionsbereiche und Variablenliste in JMP „Grafik Erstellen“.

Im einfachsten Fall zieht man je eine Variable auf die X- und die Y-Achse und erhält das entsprechende Punktediagramm. Die unterhalb der Titelleiste angeordneten Symbolbilder stehen für verschiedenen Grafiktypen, die man einzeln auswählen oder beliebig kombinieren kann. Eine Kombination aus Punktediagramm, Regressionslinie und Gruppierungsvariable gibt schon tiefere Einblicke in die Zusammenhänge (Abbildung 5).

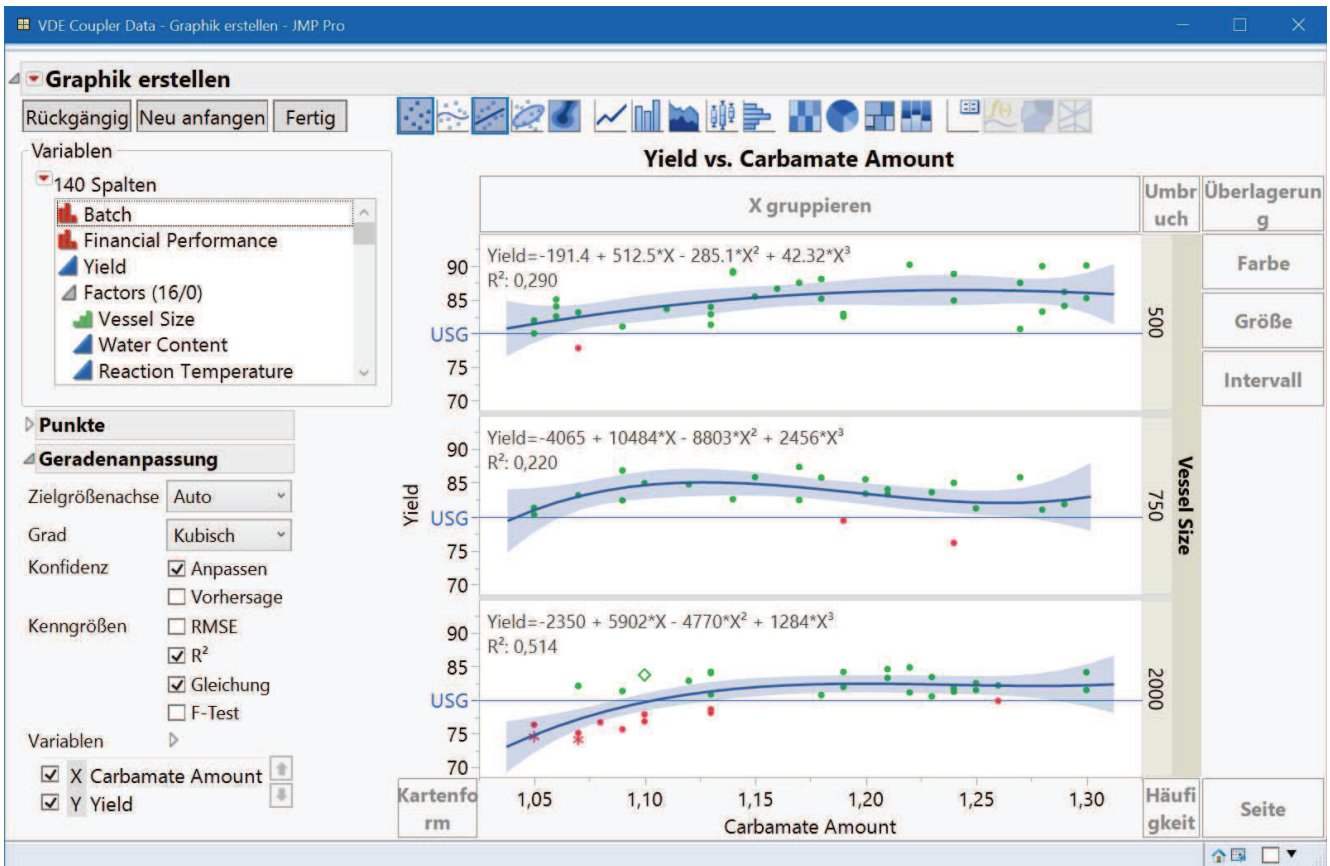


Abbildung 5: einfache kombinierte Grafik mit Regressionseinstellungen

Jedes Grafikelement verfügt über ein Einstellungs Menü, das auf der linken Fensterseite eingeblendet wird und das detaillierte Anpassungen erlaubt. Wichtig ist noch der „Fertig“ Schalter im oberen Bereich links, der alle Steuerelemente ausblendet und nur die Grafik anzeigt. Das ist dann auch die richtige Ansicht für das Speichern oder den Export in andere Programme.

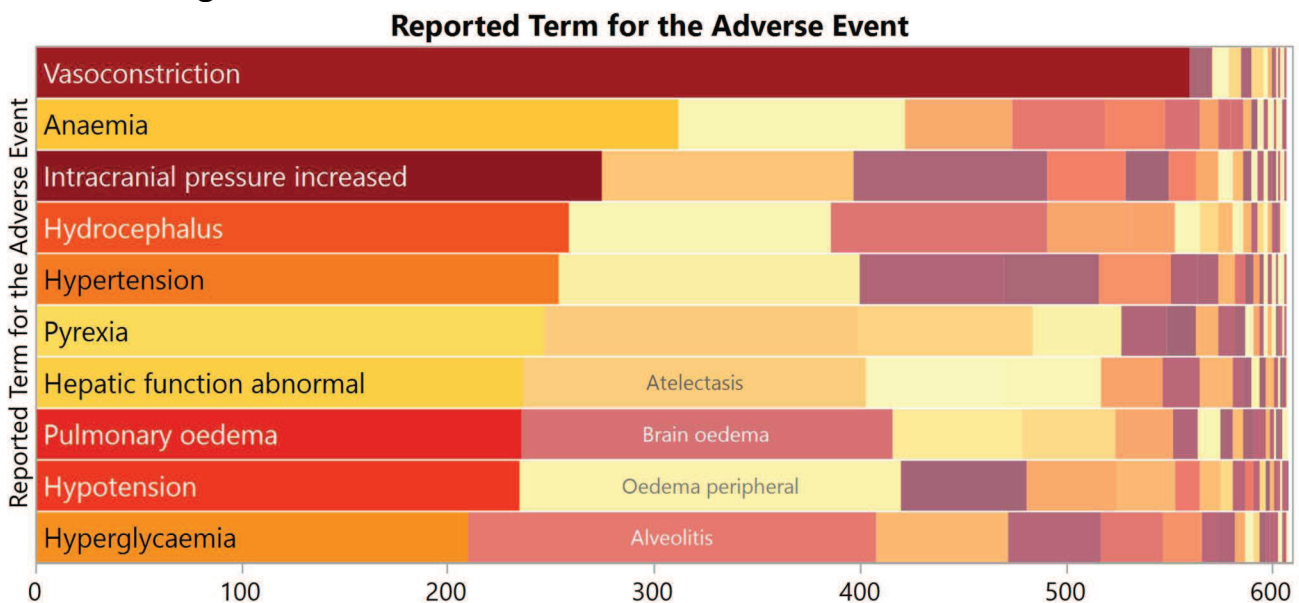


Abbildung 6: gepackte Balkendiagramme

Eine spezielle, relativ neue Darstellung von Häufigkeiten kategorialer Größen sind die gepackten Balkendiagramme (Abbildung 6). Gemäß dem Motto „Fokus und Kontext“ zeigen sie deutlich, welches die häufigsten Kategorien sind. Alle Balken sind im gleichen Maßstab abgebildet, so dass die Längenunterschiede auch korrekt interpretierbar sind. Die weniger häufigen Kategorien sind nicht besonders ausgezeichnet, werden aber auch nicht unterschlagen. So zeigt die Grafik deutlich, dass es sehr viele Ausprägungen gibt, aber nur wenige davon besonders oft auftreten. Diese werden gleich erkannt und lassen sich auch in ihren relativen Häufigkeiten gut vergleichen.

4 Überlagerungen, Kombinationen auf Achsen

Die in Abbildung 5 eingeblendete Regressionsfunktion ist mit Konfidenzbereich dargestellt. Was aber, wenn Konfidenz- oder sonstige Referenzbereiche in den Daten enthalten sind?

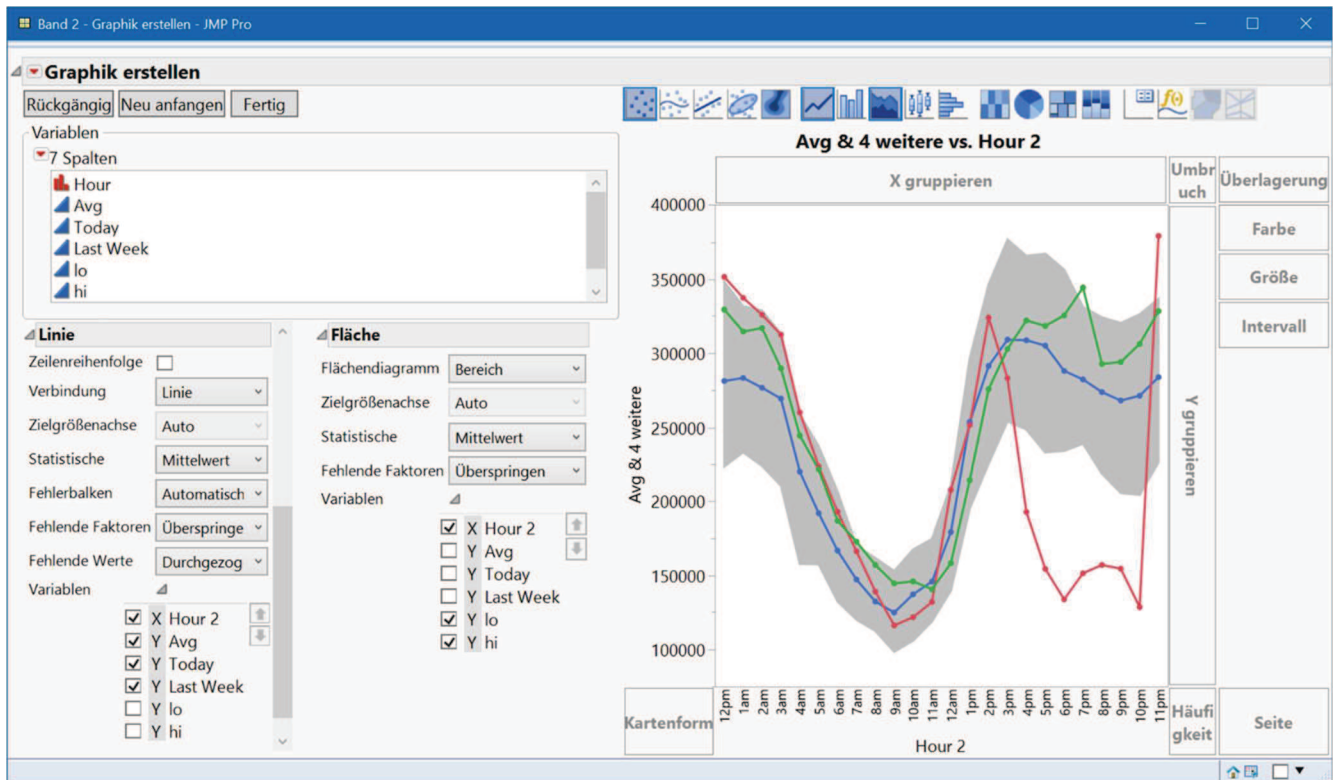


Abbildung 7: Überlagerung mehrerer Variablen und Grafikelemente mit separater Variablenwahl

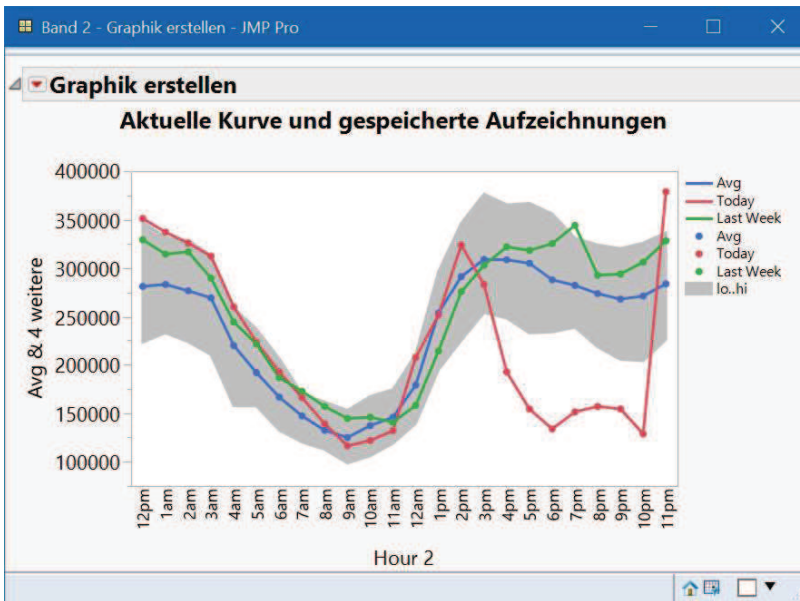


Abbildung 7 zeigt eine solche Darstellung, bei der fünf Variablen überlagernd auf der Y-Achse angeordnet sind. Außerdem sind eine Linien- und eine Flächendarstellung kombiniert worden. In den jeweils zugehörigen Steuerdialogen ist ausgewählt, auf welche der Variablen das jeweilige Grafikelement angewandt werden soll. Die fertige Grafik ohne Steuerelemente ist natürlich wesentlich übersichtlicher (nebenstehend).

In JMP kann auch eine zweite Y-Achse auf der rechten Seite der Grafik eingeblendet werden. Diese kann ebenfalls sehr praktisch zur Überlagerung mehrerer unterschiedlicher Elemente genutzt werden. Man könnte also die obige Grafik auch so erstellen, dass man z. B. die Variablen mit den Minima und Maxima für den schattierten Bereich auf der rechten Achse darstellt. In Abbildung 8 wird nur der Mittelwert auf der rechten Achse angezeigt, aber statt der Liniendarstellung wurde eine Glättung der Kurve vorgenommen.

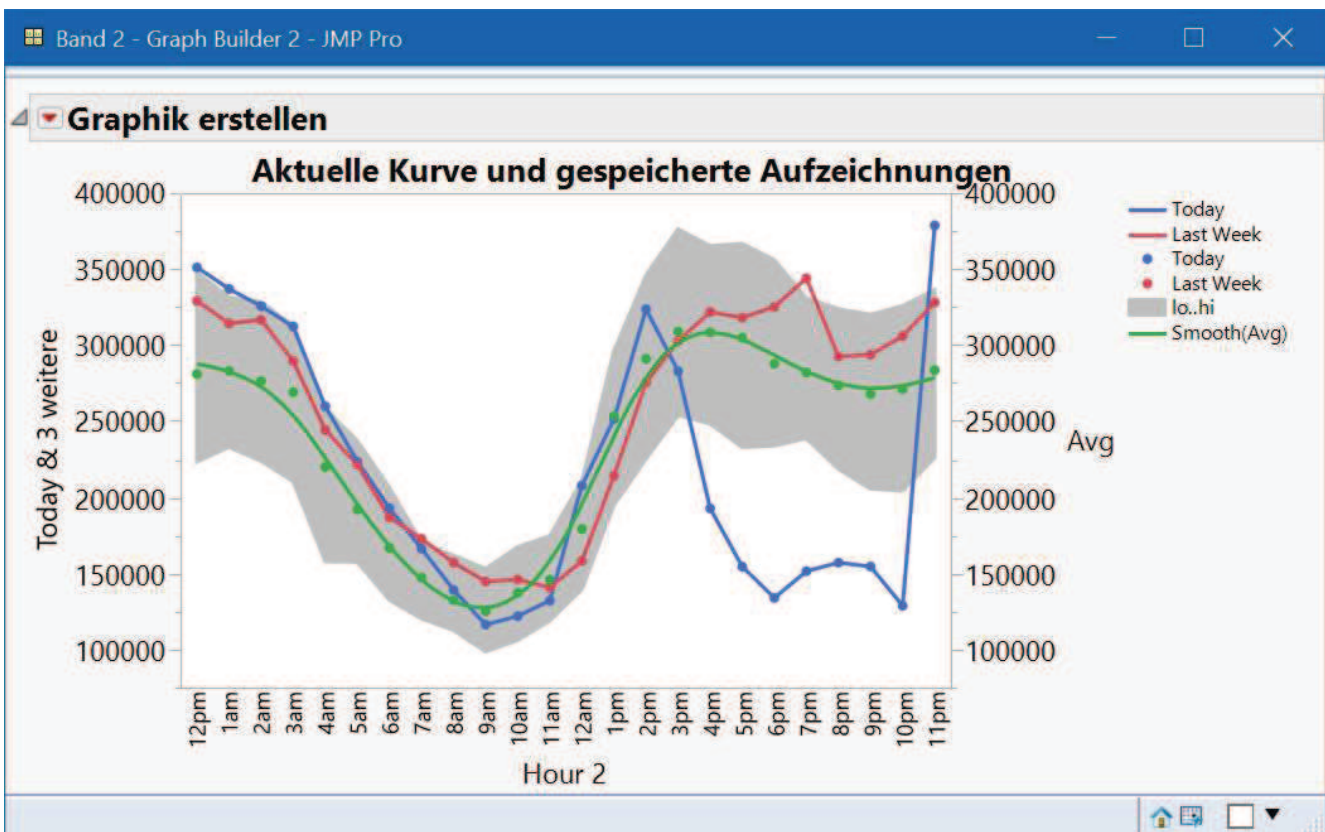


Abbildung 8: Darstellung mit zwei Achsen.

In diesem Fall sind beide Achsen vollkommen identisch aufgebaut. Natürlich könnte man auf der rechten Y-Achse auch eine völlig abweichende Maßeinheit wählen, um Variablen vergleichend darzustellen, die in unterschiedlichen Skalen gemessen wurden. Das ist aber häufig eher verwirrend und sollte gut überlegt sein, wie Stephen Few [3] anschaulich dargelegt hat.

Wenn Variablen in unterschiedlichen Skalen oder Größenordnungen gemessen werden, kann eine sinnvollere Darstellung darin bestehen, sie in getrennten Bereichen einer Grafik darzustellen (Abbildung 9). Auch das lässt sich ohne Programmierung per Mausbewegung einrichten. Ebenso wie die grau hinterlegten Referenzbereiche einfach durch Angaben in einem Dialog zur Achsenskalierung angegeben werden können.

Auf dieselbe Art und Weise kann man erreichen, dass die Achse für eine Variable gebrochen wird. Das ist dann sinnvoll, wenn sich die Werte nicht gleichmäßig über einen bestimmten Bereich verteilen, sondern in einen Bereich sehr großer und einen Bereich sehr kleiner Werte zerfallen. Für die Variable werden einfach zwei Y-Achsen angelegt und unterschiedlich skaliert. Abbildung 10 zeigt, dass somit auch der Trend, der in den niedrigen Werten steckt, deutlich sichtbar wird.

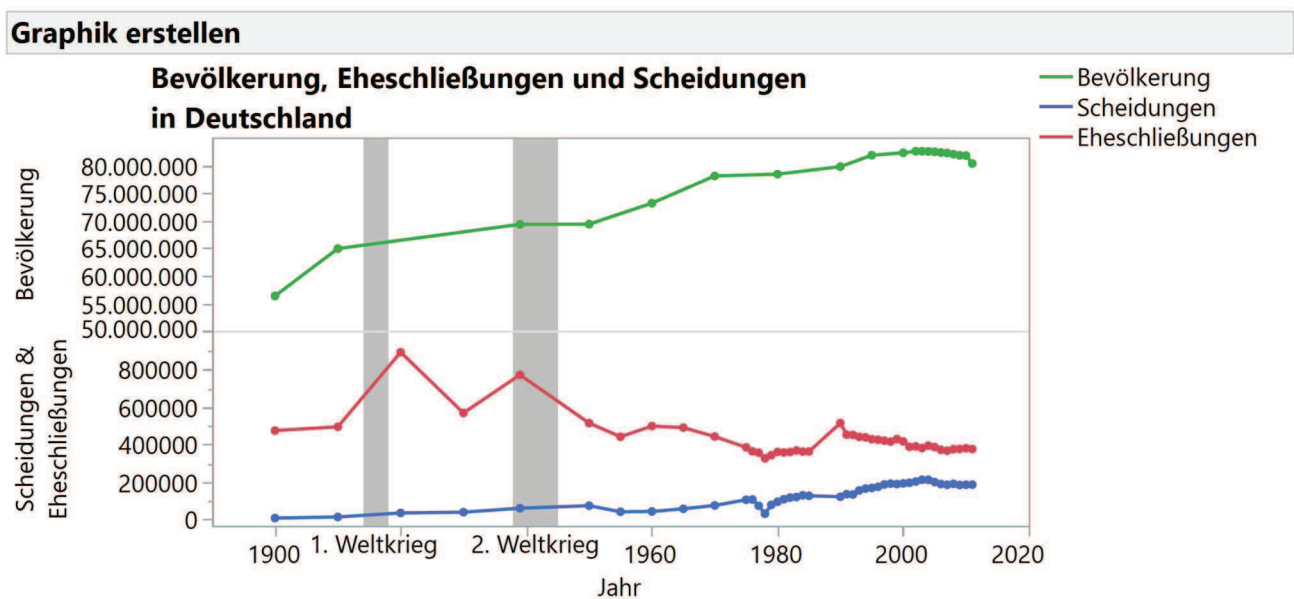


Abbildung 9: Getrennte Y-Achsen und Referenzbereiche

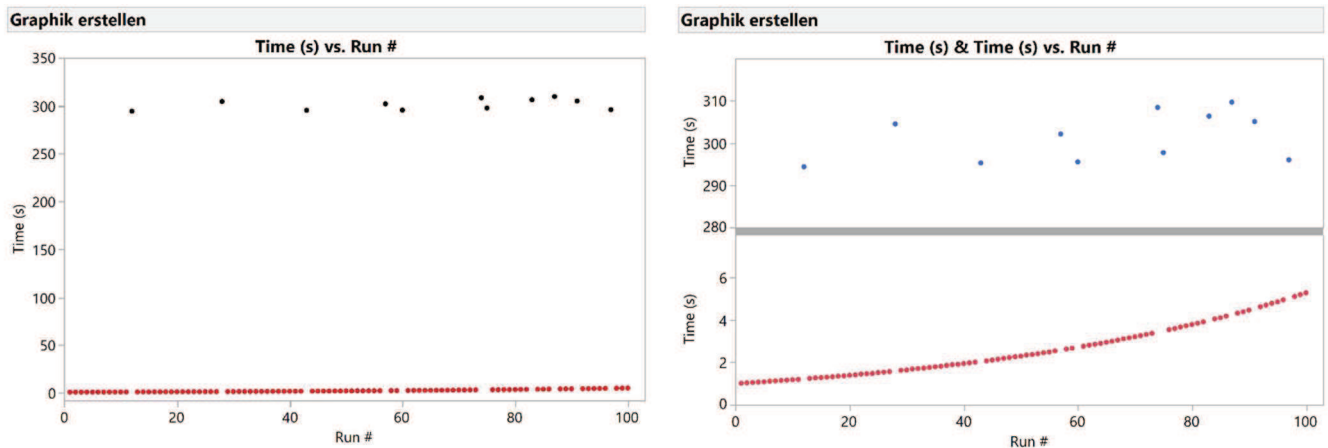


Abbildung 10: einfach skalierte (links) und getrennt skalierte Werte einer Variablen

5 Manche Grafiken brauchen etwas Datenaufbereitung

Bisher konnten alle Grafiken direkt aus den Variablen, wie sie üblicherweise erfasst werden, erstellt werden. Manchmal kommt man nicht umhin, die Daten extra aufzubereiten, um eine informative, oder auch fachspezifische Darstellung zu erhalten. In JMP geschieht das durch Einfügen einer oder mehrerer Formelspalte(n) in die Datentabelle. Nehmen wir als Beispiel eine Tabelle mit Daten (Abbildung 11), die Umfrageergebnisse in Form einer Likert-Skala enthält.

	Question	no answer	strongly disagree	disagree	neutral	agree	strongly agree
1	Q1	15	1052	2841	2266	4032	2950
2	Q10	59	1796	3661	2156	3927	1557
3	Q11	52	5863	4545	1636	725	335
4	Q12	74	922	2457	1683	4182	3838
5	Q13	32	1774	3837	1731	3506	2276
6	Q14	38	3128	5664	2636	1402	288
7	Q15	77	1124	2438	2840	4575	2102
8	Q16	73	2712	4074	1027	3105	1170

Abbildung 11: Erhebungsergebnisse nach einer Likert-Skala (Ausschnitt)

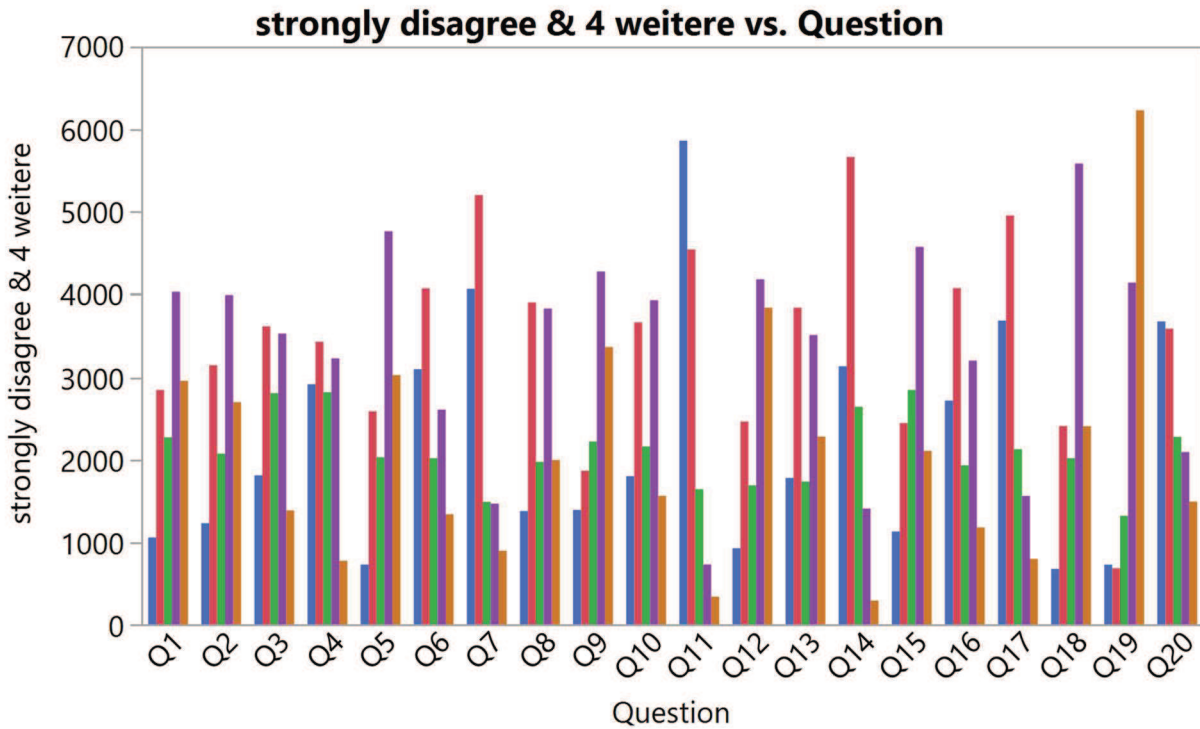


Abbildung 12: Likert-Skala, einfache Balkendarstellung

Ein Balkendiagramm ist durchaus die geeignete Darstellungsform, aber einfach die Balken nebeneinander zu zeichnen führt nicht zu einem übersichtlichen, informativen Ergebnis. Das Prinzip einer Likert Skala ist, dass von einer neutralen Mittelposition mit einer gleichen Zahl von Schritten jeweils in Richtung Zustimmung (rechts) oder Ablehnung (links) gegangen wird. Diese zentrierte Sicht möchte man auch in der Darstellung der Ergebnisse erkennen können. In den Beispieldaten standen jeweils zwei Kategorien rechts und links der mittleren Klasse zur Wahl. Wenn man die Ergebnisse zentrieren will, muss man die Mitte der neutralen Werte finden, hier soll die gedachte Achse verlaufen, und die nach links gehenden Klassen der Ablehnung mit negativen Vorzeichen versehen. In JMP fügt man dazu vier Spalten ein, die diese Formeln enthalten:

		-neutral	
-strongly disagree	-disagree	2	--neutral

Bitte beachten Sie die Vorzeichen.

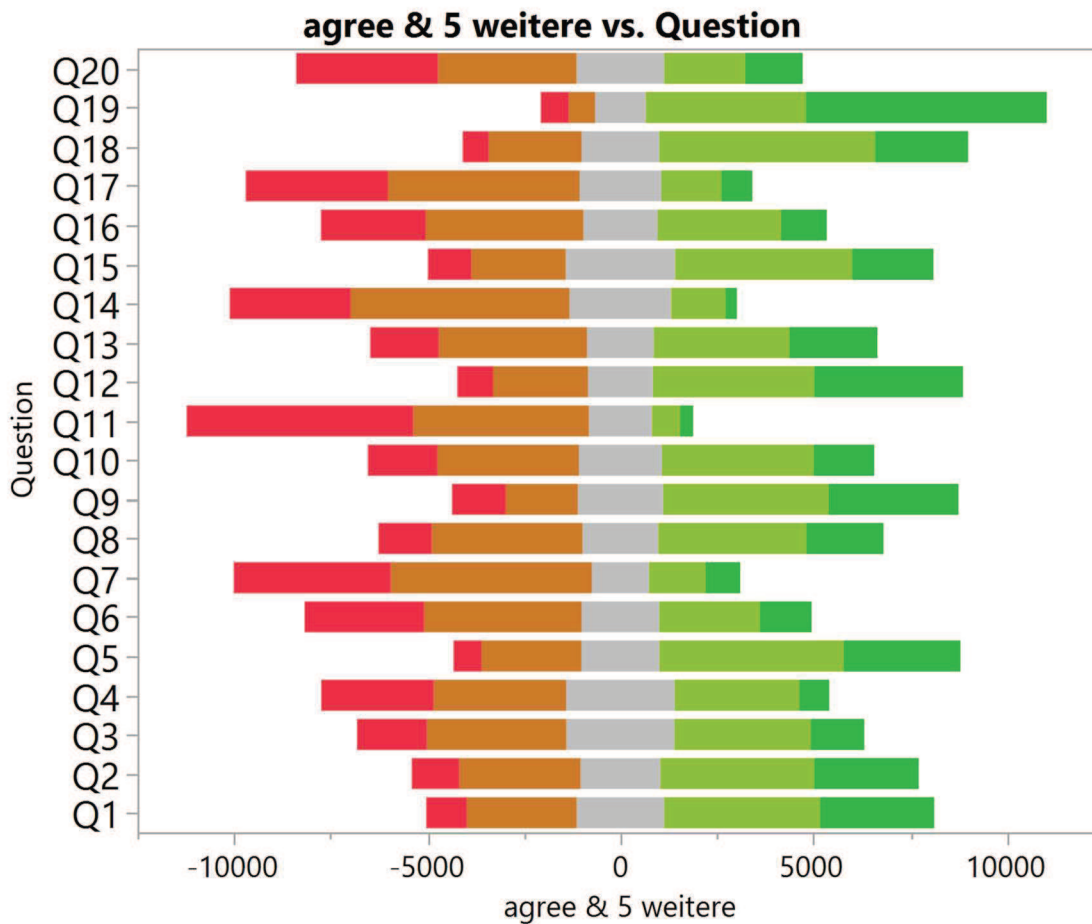


Abbildung 13: Likert Skala, zentrierte Balken

Durch einfache Kombination der Variablen, diesmal auf der X-Achse, entsteht die gut interpretierbare Darstellung aus Abbildung 13.

6 Zielgrößen und Sehgewohnheiten

Hier sollen einige Überlegungen zur sinnvollen Gestaltung einer Grafik dargelegt werden, die einem Seminarvortrag des bekannten Autors und Datenanalysten Kaiser Fung [4], entstammen.

Es wird der Frage nachgegangen, wie sich die Ungleichheit der Einkommen in den Vereinigten Staaten entwickelt hat. Dazu liegt eine recht einfache Tabelle vor, in der die mittleren Einkommen des obersten Prozentes und der unteren 90% für die Jahre 1917 bis 2015 gegenübergestellt werden (Abbildung 14).

	Year	Top one Percent Income	Bottom 90% Income
1	1917	314340,89992	11714,821089
2	1918	267361,89414	11147,045635
3	1919	272652,98703	11017,907652
4	1920	219744,68643	10040,823937
5	1921	202382,81289	8170,4391235
6	1922	252130,11044	9242,8571518
7	1923	257355,99463	10701,120335
8	1924	285304,58782	10114,888164
9	1925	344187,2106	10134,020597
10	1926	338287,8879	10249,666323

Abbildung 14: Mittlere Einkommen des oberen und der unteren 90% US amerikanischer Einkommen (Auszug)

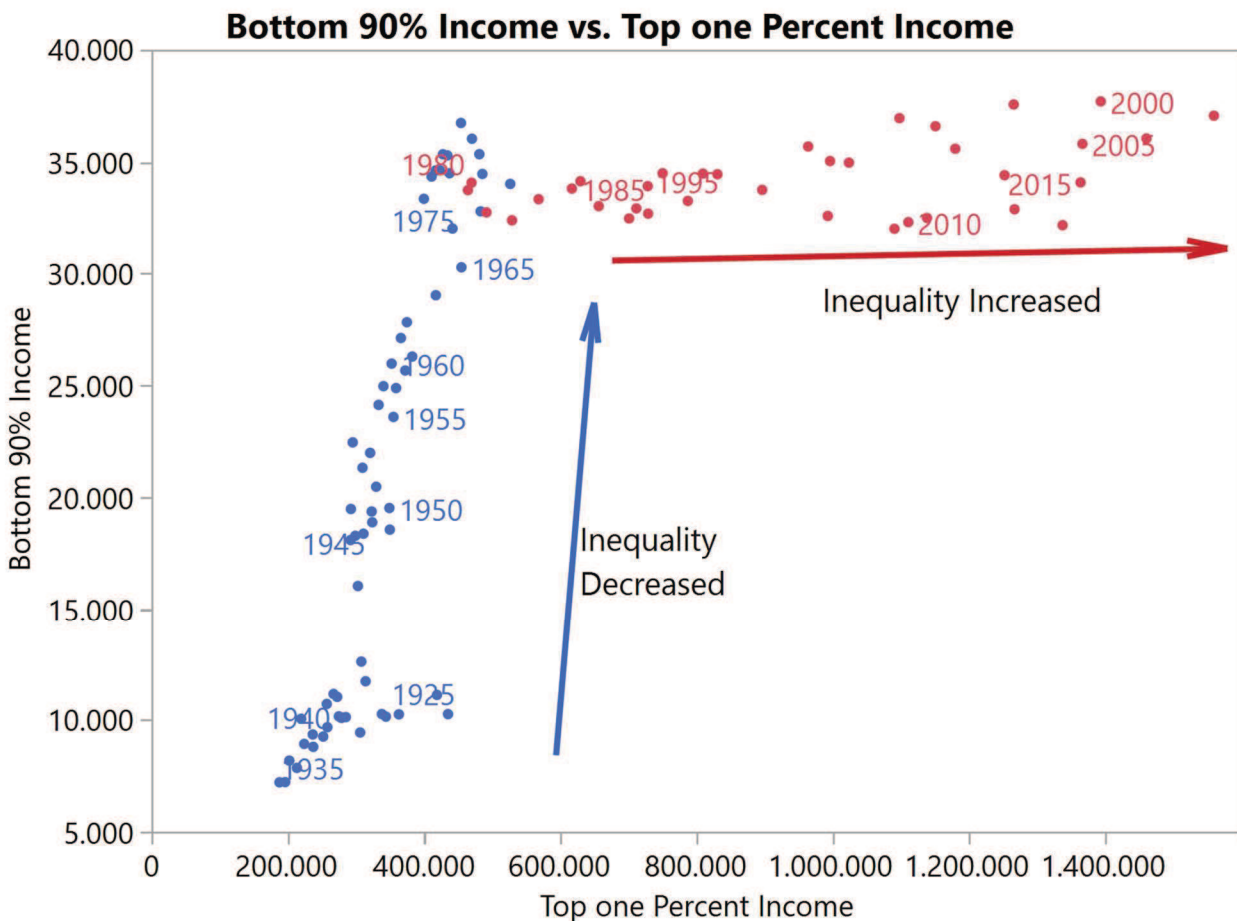


Abbildung 15: Kombination der Einkommen (A)

Stellt man die Punktwolke wie in Abbildung 15 dar, kann man die Trends zwar erkennen, aber sie entsprechen nicht unseren Sehgewohnheiten: da wo die Ungleichheit geringer wird, streben die Punkte nach oben, wo sie wächst, bleiben die Punkte auf einem

Niveau. Alleine die beiden Achsen zu vertauschen, behebt dieses Problem, aber es bleiben weitere, wie Abbildung 16 zeigt.

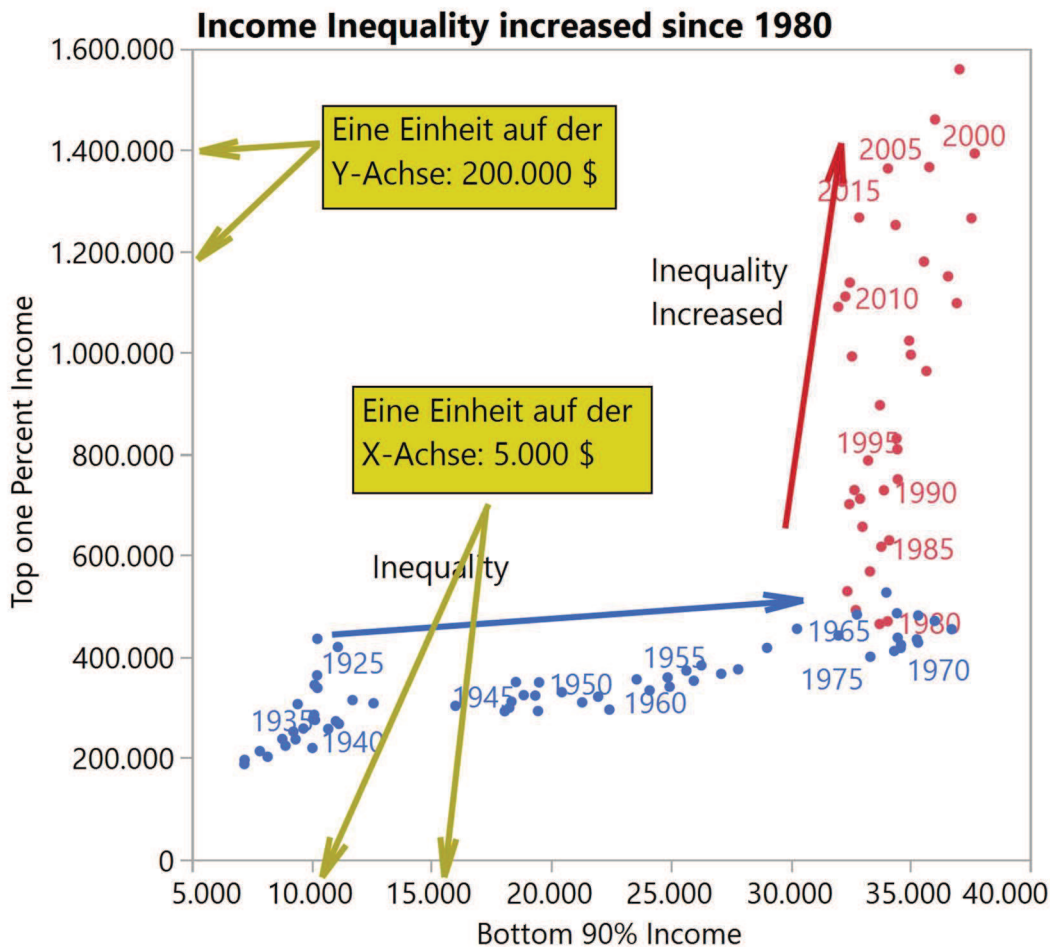


Abbildung 16: Kombination der Einkommen (B)

Würde man die beiden Achsen genau gleich skalieren, würde die Grafik einen nicht mehr lesbaren, schmalen vertikalen Streifen abgeben. Eine wesentlich bessere Darstellung erhält man, indem man den Verlauf der mittleren Einkommen beider Gruppen über die Zeit darstellt, wie in Abbildung 17.

Hier zeigt sich bereits ein ganz anderes Bild, die Ungleichheit hat nur relativ kurz abgenommen, ist dann aber wieder gestiegen und hat ungefähr ab 1980 massiv zugenommen. Einen kleinen Nachteil hat das Bild noch, es überlässt es dem Zuschauer, die Differenz der mittleren Einkommen bei der Betrachtung zu bilden. Es ist aber wesentlich besser, die wirklich interessierende Größe zu berechnen und direkt darzustellen. Was hier wirklich interessiert, ist das Verhältnis der mittleren Einkommen:

$$\frac{\text{Top one Percent Income}}{\text{Bottom 90\% Income}}$$

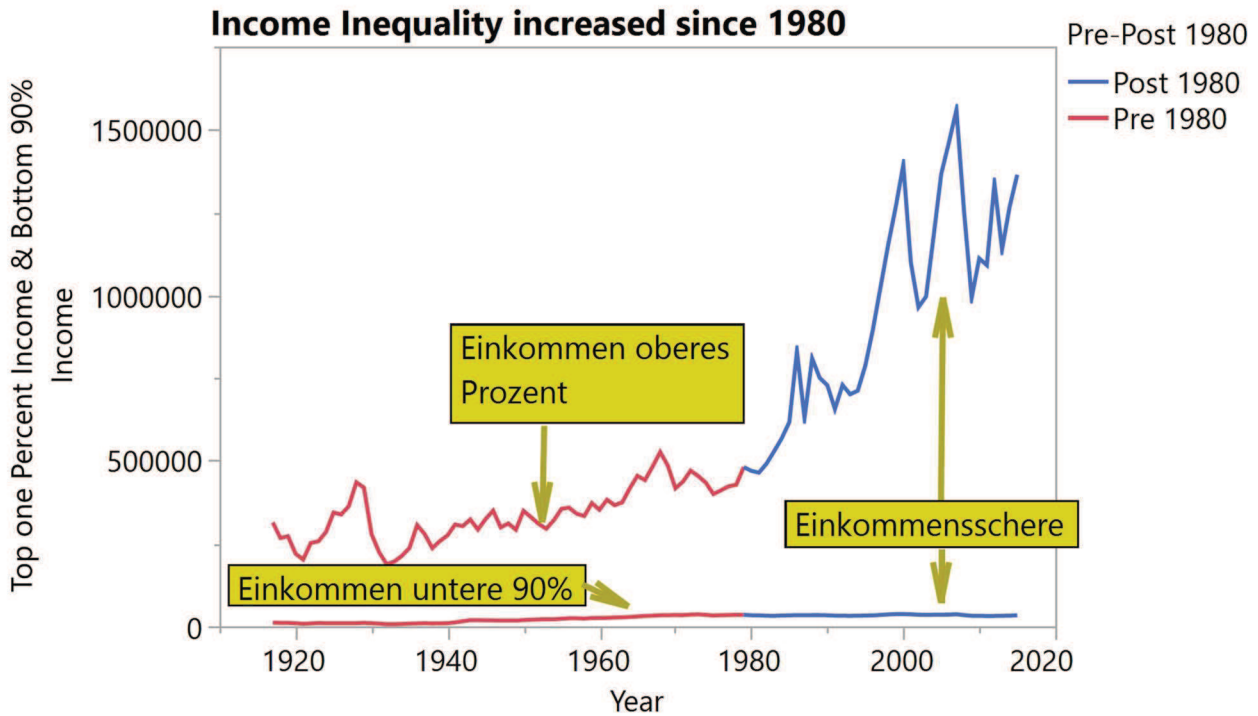


Abbildung 17: Kombination der Einkommen (C)

So sieht man in Abbildung 18 direkt die interessierende Größe, kann den Verlauf besser beurteilen und sieht, dass die Einkommensungleichheit 50 Jahre lang abgenommen hat, dann aber in 40 Jahren wieder auf das alte Niveau gestiegen ist. Nebenbei zeigt diese Grafik auch eine Funktion in JMP, die es erlaubt, Zusatzinformationen aus der Datentabelle punktbezogen in der Grafik anzuzeigen.

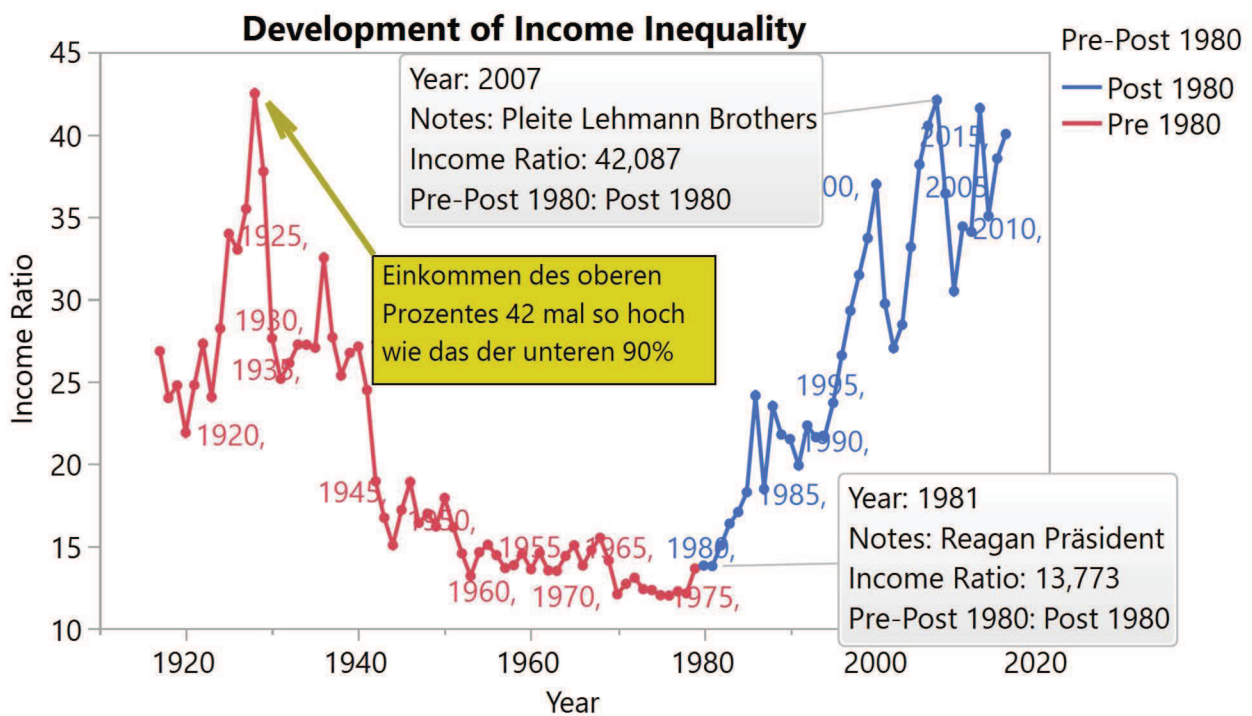
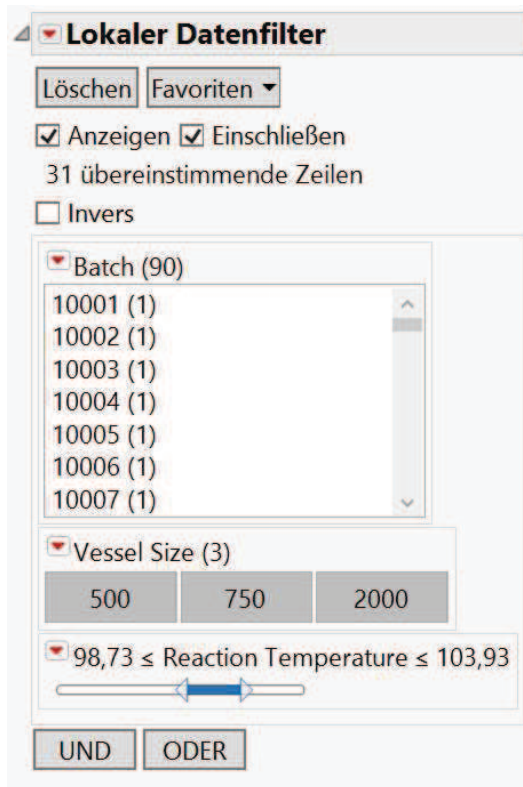


Abbildung 18: Kombination der Einkommen, Änderung des Verhältnisses über die Zeit

7 Interaktive Grafiken

So weit luden alle vorgestellten Grafiken zum Betrachten ein, aber nicht zu einer interaktiven Einflussnahme. Den Autoren von Grafiken stehen drei verschiedene Möglichkeiten offen, sich selbst, oder Betrachtern der Grafik, Möglichkeiten zur Interaktion mit der Grafik zu bieten.

7.1 Der Datenfilter



Eine beliebige Kombination aus Spalten der Datentabelle kann bereitgestellt werden, um die Daten, die der Grafik zu Grunde liegen, zu filtern. Je nach Zahl der Kategorien werden kategoriale Größen dabei als Liste oder Schaltfläche dargestellt, für stetige Größen gibt es einen Schieberegler, dessen Spannweite beliebig eingestellt und verändert werden kann und der dann über den gesamten Wertebereich dieser Variablen bewegt werden kann. Nutzt man Variable als Filter, die nicht an der Konstruktion der Grafik beteiligt waren, so repräsentiert die komplette Grafik genau die gewählten Teilmengen der Daten. Werden auch Variablen genutzt, die zum Aufbau der Grafik herangezogen wurden, wirkt der Datenfilter wie eine Zoomfunktion, die Teile der Grafik prominent hervorhebt.

7.2 Der Spaltenwechsler

Oft möchte man eine informative Darstellung einfach auf andere Variable übertragen, um einen vergleichbaren Sachverhalt zu überprüfen. Der Spaltenwechsler ist die Lösung für diese Aufgabe. Nach Fertigstellung der ersten Grafik wählt man in einem Dialog die Variable aus, deren Rolle in der Grafik man durch eine oder mehrere andere Variable (gleichen Typs) ersetzen möchte. Dies könnte z.B. die Variable sein, die auf der X-Achse angeordnet ist. Ferner wählt man aus einer Liste diejenigen Variablen aus, die eben diese Rolle auch übernehmen sollen. Dabei sollte die Ausgangsvariable auch wieder in dieser Liste erscheinen. Ein Listenelement erlaubt es nun, einfach durch Anklicken mit der Maus, den entsprechenden Austausch der Variablen vorzunehmen. So kann man sehr schnell unterschiedliche Variable in derselben Ansicht betrachten. Abbildungen 19 und 20 zeigen dieselbe Grafik mit jeweils einer anderen Variablen auf der X-Achse.

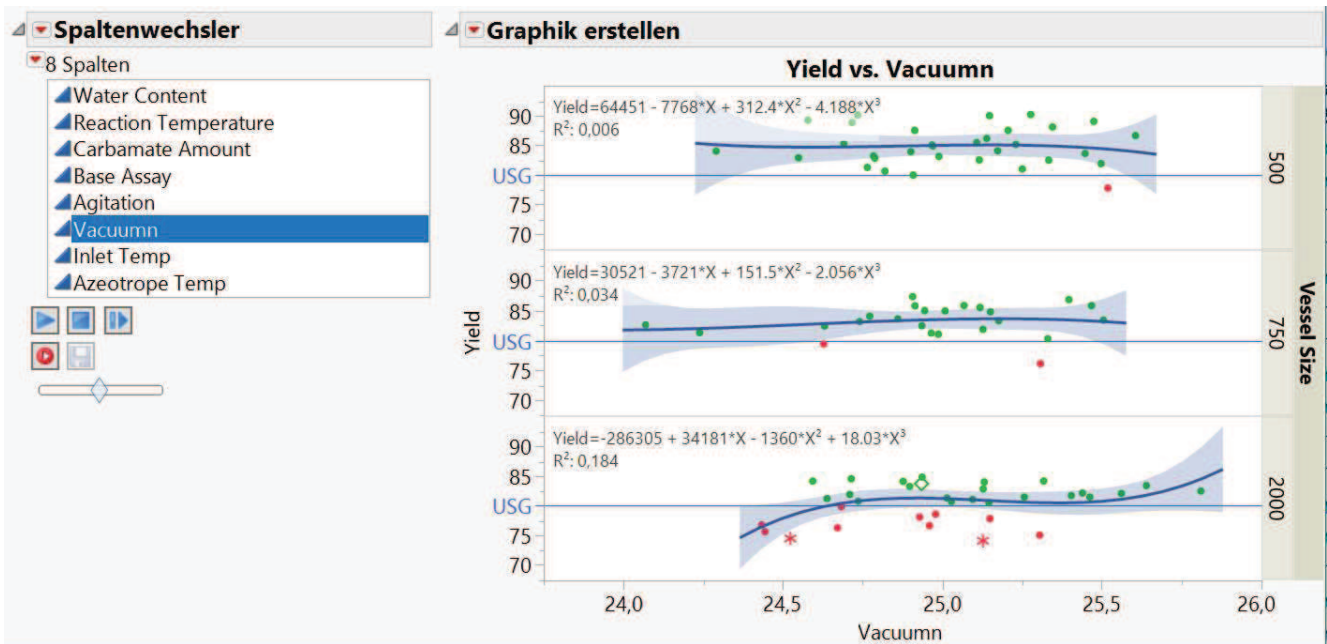


Abbildung 19: Spaltenwechsler mit Variablenauswahl

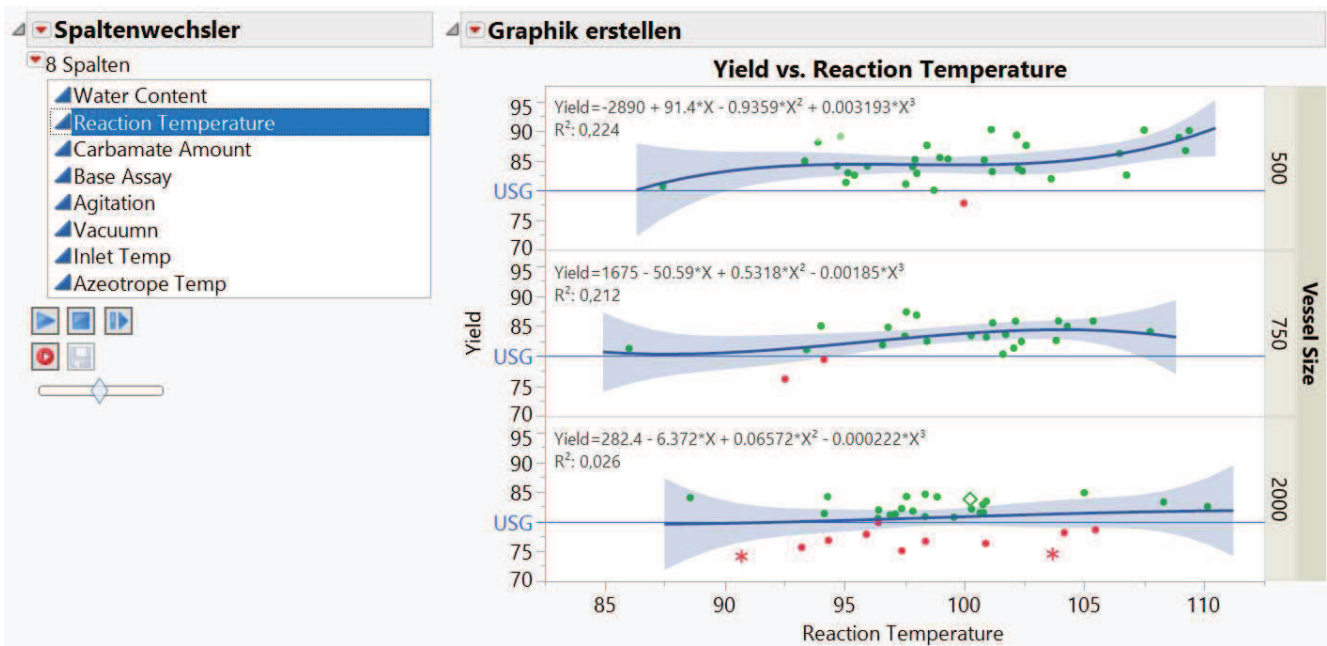


Abbildung 20: Spaltenwechsler mit alternativer Variablenauswahl

7.3 Das Analysediagramm

Die meisten Grafiken so weit wurden mit einer JMP Plattform erstellt, mit „Grafik erstellen“. Eine ganz andere Form der Interaktivität bietet das Analysediagramm. Es ist darauf ausgelegt, die Zusammenhänge darzustellen, die zwischen mehreren Faktoren und einer oder mehreren Zielgrößen bestehen. Dazu muss es eine oder mehrere Formelspalten geben, die diese funktionale Abhängigkeit beschreiben und die für die Darstellung ausgewählt werden. Für jede Eingangsgröße, die in den Formeln gefunden wird, wird ein Profildiagramm erzeugt, das den Einfluss dieser Stellgröße auf die Ergebnisvariable darstellt. In Abbildung 21 sieht man deutlich, dass einige Faktoren (unten, in X-

Richtung aufgetragen) einen linearen Einfluss auf manche Zielgrößen (links, in Y-Richtung aufgetragen) ausüben und einen nichtlinearen Einfluss auf andere. Die vertikalen gestrichelten Linien zeigen die aktuellen Werte an, auf die die Faktoren eingestellt wurden, und deren Werte unten, oberhalb der Faktornamen, angezeigt werden. Horizontale gestrichelte Linien repräsentieren die Ergebnisse der Berechnungen, die auch links neben den Namen der Ergebnisgrößen angegeben werden, in diesem Beispiel ergänzt durch die 95% Konfidenzintervalle. Die vertikalen gestrichelten Linien können mit der Maus verschoben werden, wodurch die Faktoren auf neue Werte eingestellt werden und die Berechnung der Ergebnisse angepasst wird.

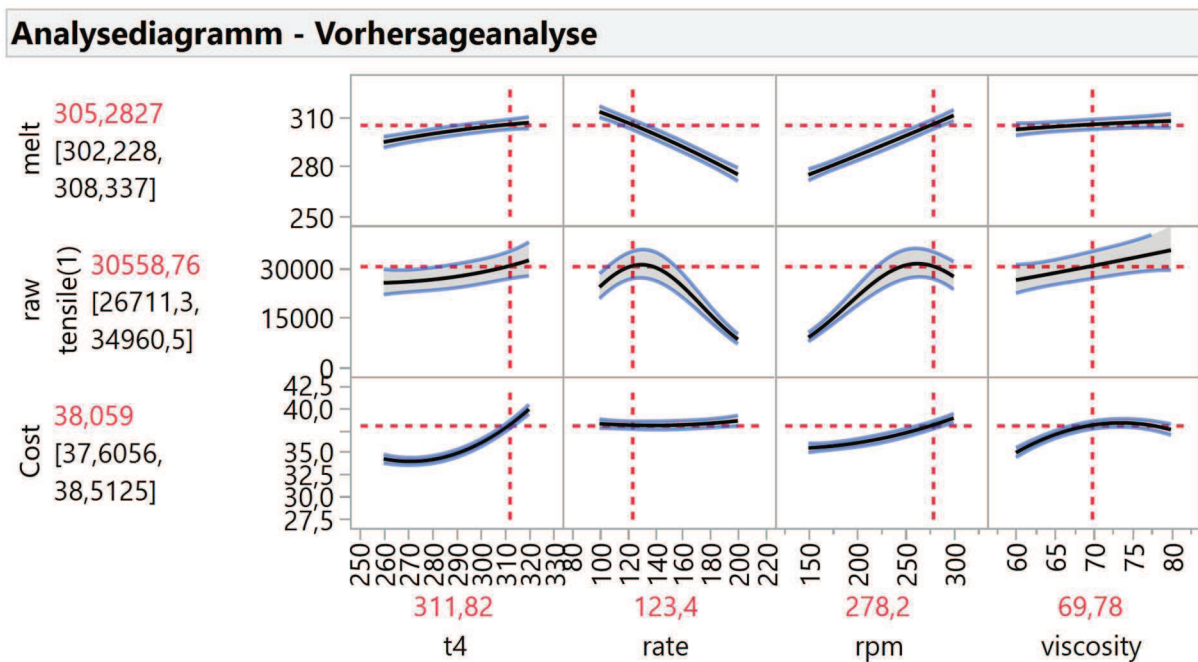


Abbildung 21: Analysediagramm, Einstellung A

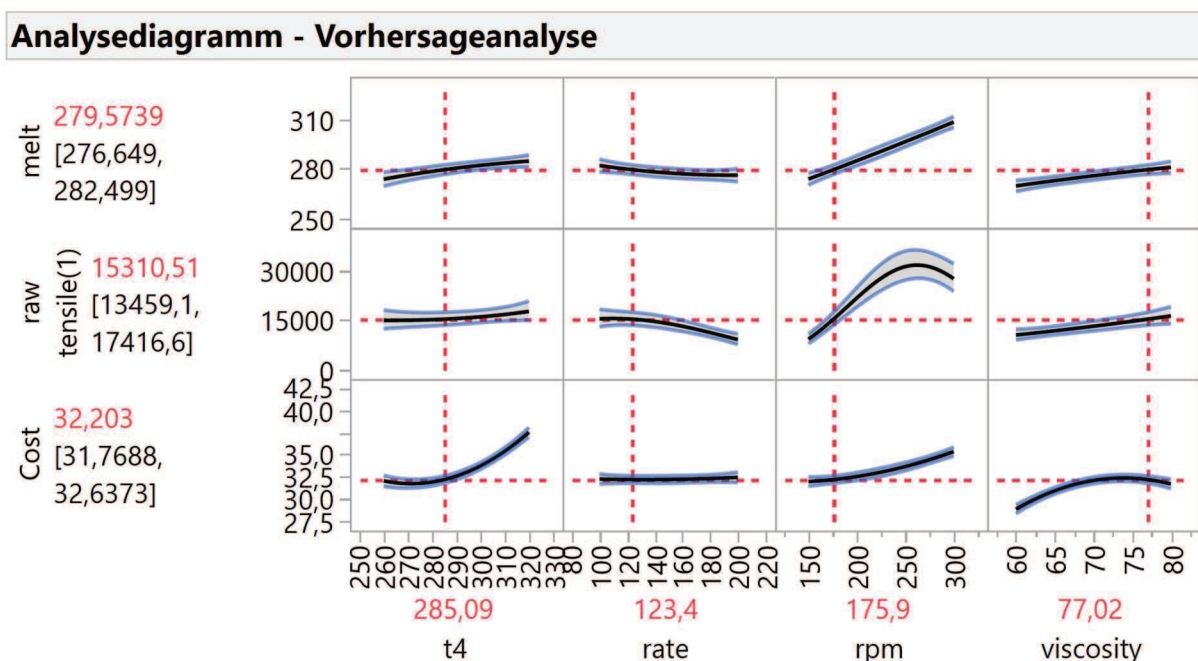


Abbildung 22: Analysediagramm, Einstellung B

In Abbildung 22 ist dasselbe Analysediagramm angezeigt, aber diesmal mit veränderten Werten der Faktoren. Nicht nur sind die Werte der Ergebnisgrößen angepasst, auch haben sich die Profile mancher Faktoren stark verändert. Vergleichen Sie die Funktion von „rate“ zu „raw tensile(1)“ in beiden Darstellungen. Das veränderte Profil zeigt deutlich, dass eine Wechselwirkung mit (einem) anderen Parameter(n) vorliegen muss. Verschieben der vertikalen Linie von „rate“ zeigt, dass „rpm“ der wechselwirkende Faktor ist. So lassen sich nicht nur die Ergebnisse komplizierter Berechnungen schnell und grafisch ermitteln, es wird auch visuell erfahrbar wo und wie stark Wechselwirkungen auftreten.

8 Speichern, Wiederholen, Veröffentlichen

So attraktiv es ja ist, auch komplexe Grafiken einfach und schnell zu erstellen, so ermüdend ist es doch, denselben Aufbau immer wieder zu wiederholen. Außerdem möchte man die tolle Grafik vielleicht auch mal Kollegen zur Verfügung stellen. Kein Problem. Aus dem Hauptmenü der fertigen Grafik (rotes Dreieck oben links) den Punkt „Skript speichern“ wählen und das fertige Grafiksript abspeichern. So lässt sich die Grafik immer wieder mit neuen Daten erstellen. Natürlich setzt die Anwendung voraus, dass der Empfänger auch über eine JMP Installation an seinem Arbeitsplatz verfügt.

Was aber, wenn die Grafik auch Personen zur Verfügung gestellt werden soll, die keine eigene JMP Version haben? Statische Versionen der Grafik können in beliebige Grafikformate oder nach Powerpoint exportiert werden. HTML oder dynamisches HTML bieten ebenfalls Verteilungsmöglichkeiten, die auch in geringem Umfang etwas von der Interaktivität der Grafiken erhalten. Ab der JMP Version 15, die im Oktober 2019 erscheint, wird auch ein Server angeboten, auf dem man alle JMP Analysen, also auch die Grafiken veröffentlichen kann. Hat der Autor der Grafiken diese mit Spaltenwechslern oder Datenfiltern versehen, bleibt auch diese Funktionalität in vollem Umfang erhalten. Heute kann man diese Funktionalität schon auf <https://public.jmp.com/> sehen und ausprobieren. Mit JMP kann man beliebige Reports dort veröffentlichen (und auch wieder löschen) und so bereits einen Eindruck von der zu erwartenden Funktionalität erhalten.

9 Zusammenfassung

Informative Grafiken zu erstellen, ist technisch leicht und für jedermann schnell umzusetzen. Egal für welchen Zweck, ob ad-hoc, um sich schnell einen Einblick in einen Datensatz zu verschaffen, oder etwas elaborierter, um einen Sachverhalt ansprechend zu präsentieren, JMP ist ein interaktives Werkzeug, das diese Arbeit erleichtert. Selbst komplexe Grafiken, die die Informationen aus vielen Variablen kombinieren, lassen sich nur durch Mausbewegungen erstellen, gänzlich ohne Programmierung. Lust auf Ausprobieren bekommen? www.jmp.com/trial

Literatur

- [1] <https://www.youtube.com/user/JMPSoftwareFromSAS>, Kurzanleitungen zu vielen Funktionen in JMP
- [2] https://www.jmp.com/en_us/events/ondemand/mastering-jmp/basics-for-using-graph-builder.html ausführlich Einweisung in die Grafikkomponente von JMP.
- [3] Stephen Few, Dual-Scaled Axes in Graphs, Visual Business Intelligence Newsletter, March 2008, Perceptual Edge
(http://www.perceptualedge.com/articles/visual_business_intelligence/dual-scaled_axes.pdf)
- [4] Kaiser Fung, Mit explorativer Datenanalyse Verborgenes in Ihren Daten aufdecken, SAS Institute GmbH,
https://www.jmp.com/de_de/events/ondemand/local/explorers-guide-eda-kaiser-fung.html, (SAS-Profil erforderlich)