

# Kurven in der statistischen Darstellung

Von Marie Reidemeister (Gesellschafts- und Wirtschaftsmuseum in Wien)

Ähnlich wie in der Mathematik werden auch in der Statistik die Kurven benützt, um sich ändernde Größen anschaulich darzustellen, zum Beispiel Bevölkerungsbewegungen, Geburten und Sterbefälle in verschiedenen Jahren, Jahresproduktionen und anderes mehr. Neben zeitlich veränderlichen Zahlen sind auch zum Beispiel Größen durch Kurven darstellbar, die sich mit der geographischen Breite der Erde, ebenso Größen, die sich mit der Höhe über dem Meeresspiegel ändern usw. Zeit, geographische Breite und Höhe haben miteinander gemeinsam, daß sie selbst durch Zahlen ersetzt werden können. Soll eine statistische Tabelle durch eine Kurve darstellbar sein, so muß sie durch eine Tabelle von Zahlenpaaren ersetzt werden können, denn nur Zahlenpaare können in der Koordinatenebene durch Punkte einer Kurve dargestellt werden. Bei der Statistik „Eisenproduktion in den Vereinigten Staaten, England, Deutschland, Rußland usw.“ könnte man zwar auch die Ländernamen durch Zahlen ersetzen, aber nur willkürlich; ändert man die Zahlen, so ändert sich sofort die Kurve, man kann ihr auf diese Weise jede beliebige Form geben, das heißt sie besagt gar nichts. Eine solche Kurve könnte aber nützlich sein, wenn man zum Beispiel für die Produktionen verschiedener Stoffe dieselbe Anordnung der Länder beibehielte, wodurch die Produkte ihrer Herkunft nach verglichen werden könnten. Im allgemeinen werden Kurven dort verwendet, wo eine solche Anordnung nicht erst künstlich geschaffen werden muß, also zum Beispiel bei „Eisenproduktion in verschiedenen Jahren“, „Mitgliederbewegung“ usw.

Besteht eine statistische Tabelle aus mehreren Zahlenpaaren, so ist durch sie für die graphische Darstellung eine Anzahl von getrennt liegenden Punkten der Koordinatenebene gegeben. Werden diese Punkte durch eine gebrochene oder gekrümmte Linie verbunden, so würde dies besagen, daß die getrennt liegenden Punkte zu einer unendlich großen Mannigfaltigkeit stetig aufeinanderfolgender Punkte gehören, oder daß die Zahlen der Statistik zu einer stetigen Mannigfaltigkeit von Zahlenpaaren gehören. Ist das nun immer der Fall?

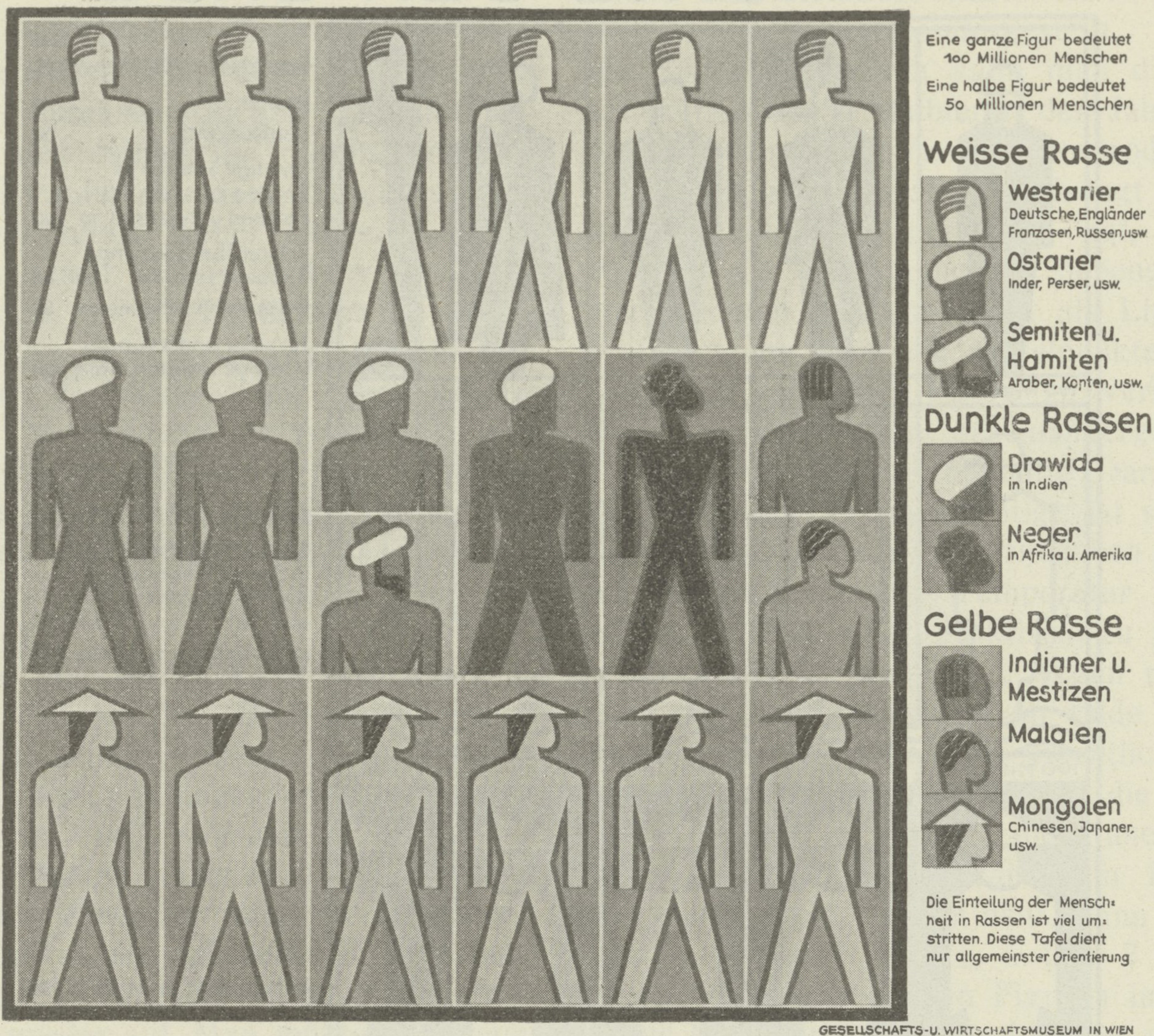
Es handle sich um die Anzahl der Mitglieder einer Krankenkasse in verschiedenen Jahren. Die Stichtage seien der 1. Jänner der Jahre 1910,

1911, 1912, 1913 usw. Diese Stichtage geben für die graphische Darstellung die Abszissen  $a_1, a_2, a_3$ , die dazugehörigen Mitgliederstände die Ordinaten  $o_1, o_2, o_3$ . Verbinden wir die so erhaltenen Punkte  $(a_1, o_1), (a_2, o_2), (a_3, o_3)$  durch gerade Linien, so erhalten wir die erste Annäherung der genauen Kurve der Mitgliederbewegung. Die Zwischenpunkte auf der Abszissenachse bedeuten die übrigen Zeitpunkte des Jahres, die dazugehörigen Ordinaten die Mitgliederstände in diesen Zeitpunkten. Die Kurve hat also denselben Charakter wie eine Fieberkurve: jedem Zeitpunkt ist ein Zustand zugeordnet, jedem Punkt der Abszissenachse eine Ordinate. Die Zustände in zwei getrennten Augenblicken werden stetig ineinander übergeführt.

Wie steht es aber mit der Kurve „Jahresproduktion von Eisen“? Wenn die Jahre als Punkte der Abszissenachse aufgetragen werden, entsprechen den Produktionen Ordinaten. Würde man die zugehörige Tabelle, die etwa die Jahre 1910, 1911, 1912, 1913 enthalten möge, ergänzen, so würde man zu vorhergehenden und nachfolgenden Jahren die Zahlen anfügen, links und rechts von den vorhandenen Punkten noch neue Punkte zeichnen müssen, niemals aber zwischen ihnen! Das heißt, bei noch so großer Genauigkeit kommt man immer nur zu getrennt liegenden Punkten. Es hat hier gar keinen Sinn, die Punkte, wie dies oft geschieht, durch eine Kurve zu verbinden, weil die zwischenliegenden Punkte keine Bedeutung haben. Während bei der Fieberkurve der Zeitablauf durch die ganze Abszissenachse dargestellt wird, haben hier ja nur die getrennt liegenden Punkte der Abszissenachse eine Bedeutung: die den Jahren 1910, 1911 usw. entsprechen. Ganze Zeitabschnitte sind durch Punkte dargestellt. Es gibt kein Individuum derselben Art, das zwischen den Jahren 1910 und 1911 liegt.

Dieser Mangel kann beseitigt werden, wenn man Zeitpunkte zu Punkten der Abszissenachse macht und etwa an Stelle von „1910“ „24 Uhr, 31. Dezember 1910“, usw. setzt. Die zugehörigen Ordinaten sollen dann die Eisenproduktion des vergangenen Jahres bedeuten. Jeder Punkt der Abszissenachse stellt einen Zeitpunkt eines Jahres dar, seine zugehörige Ordinate die Eisenproduktion in dem Jahre, das in diesem Augenblick beendet ist, also zum Beispiel an dem Punkt 24. Juni 1911,

# MENSCHENRASSEN



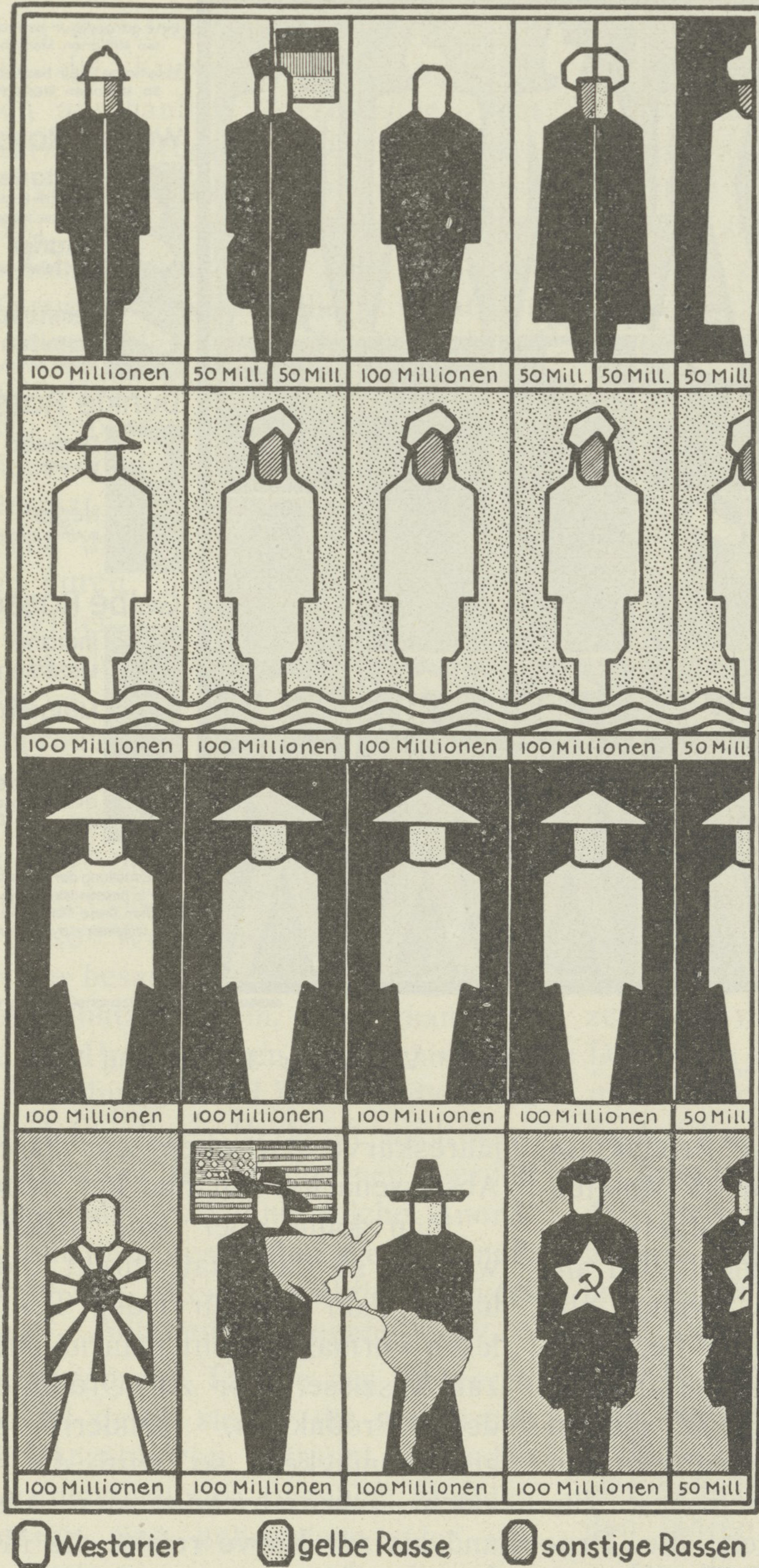
Aus der Zeitschrift „Das Bild“, Deutscher Verlag für Jugend und Volk

3 Uhr 45 Minuten ist als Ordinate die Produktion in der Zeit vom 24. Juni 1910, 3 Uhr 45 Minuten bis 24. Juni 1911, 3 Uhr 45 Minuten aufgetragen. Die Kurve zeigt mit ihrem Steigen und Fallen an, wann die Jahresproduktion zu-, wann abgenommen hat. Das Steigen der Kurve bedeutet, daß in dem Augenblick, den die Abszisse anzeigt, mehr produziert wird, als ein Jahr vor diesem Augenblick; die Größe dieser Steigung gibt ein Maß für die Größe dieses Unterschiedes. Die Kurve gibt aber keine Möglichkeit, die Produktionsmengen für Teile des Jahres bestimmt anzugeben, es sei denn, daß ein Anfangswert gegeben ist. Ist zum Beispiel eine Halbjahresproduktion bekannt, so können auch die übrigen ermittelt werden. Diese Halbjahresmengen sind aber natürlich nicht auf der Jahresproduktionskurve zu finden, sondern ihre Ordinaten sind im allgemeinen etwa halb so groß. Es sei denn, daß nur ein halbes Jahr gearbeitet wird, das zweite halbe Jahr gefeiert; in diesem Falle würde die Halb-

jahreskurve zwischen der Jahreskurve und der Abszissenachse hin und her laufen.

Man kann aber eine andere Darstellungsart wählen, so daß man direkt die Mengen für ein halbes Jahr und für jeden Teil des Jahres abzulesen vermag. Macht man nämlich wieder die Zeit zur Abszisse, aber zur Ordinate nicht die Menge der Produktion, sondern die Produktionsgeschwindigkeit, so wird die Fläche, die über einem Jahresabschnitt der Abszissenachse liegt und bis zur Kurve reicht, die Menge angeben, die in einem Jahre erzeugt wird. Halbiert man die Jahreseinheiten, so geben die Flächenstücke über diesen Hälften direkt die Mengen, die in den halben Jahren erzeugt werden. Durch weitere Aufteilung der Abszissenachse und der Fläche bekommt man die Mengen, die in kleineren Zeitabschnitten erzeugt werden. Dabei leistet die neue Kurve nicht weniger als die frühere: es ist auch ablesbar, wie groß die Produktionen in den Jahren sind, die in einem beliebigen Augenblick beginnen. Aber im

# DIE MÄCHTE DER ERDE



Mill.  
 Frankreich m. Kolonien 100  
 Italien 40  
 Deutschland 60  
 Sonstige Staaten in Europa ausser England u. Russland ca 100  
 Kolonialbevölkerung ausser Malaien ca 50  
 Malaien in Kolonien ca 50  
 Türkei u. andere unabhängige Staaten unter 50

Britisches Reich 450  
 davon Westarier in Europa, Kanada, Australien, Afrika u. s. w. unter 100  
 Inder, Neger u. s. w. über 350

China  
 Kämpfende Gruppen ca 450

Japan mit Korea unter 100  
 Vereinigte Staaten von Amerika über 100  
 Staaten in Mittel- und Südamerika davon 50 Millionen Mestizen unter 100  
 Union der Sowjetrepubliken ca 150  
 davon 50 Mill. gelbe Rasse

Zusammen Mill. 1800

GESELLSCHAFTS- U. WIRTSCHAFTSMUSEUM IN WIEN

Aus der Zeitschrift „Das Bild“, Deutscher Verlag für Jugend und Volk

allgemeinen interessiert dies ja weniger. Sie leistet andererseits viel mehr: man sieht, in welchen Augenblicken die Produktionsgeschwindigkeit am größten ist, wann sie am meisten im Anwachsen begriffen ist usw.

Man könnte einwenden: Wie soll man eine solche Kurve anfertigen, sind uns doch die

Produktionsgeschwindigkeiten im allgemeinen nicht bekannt. Das ist richtig. Aber es ist möglich, mit Hilfe der Jahresmengen eine Annäherung herzustellen. Dazu ist nur nötig, daß man diese Jahresproduktionen nicht als Längen, sondern als Flächen aufträgt! Gleichbreite, rechteckige Streifen von verschiedener Länge, auf eine Achse

aufgestellt, geben mit ihren oberen Enden eine Treppenkurve: die erste Annäherung zur genauen Kurve.

Eine nächste Annäherung würde es sein, wenn man die Jahresstreifen der Länge nach halbiert und statt der durchschnittlichen Produktionsgeschwindigkeit des Jahres die durchschnittliche Produktionsgeschwindigkeit der Halbjahre aufträgt. Durch Anbringung immer kleinerer Stufen nähert man sich immer mehr der Kurve.

Handelt es sich um die Produktion an einem Ort oder in einem Lande mit gleichgehenden Uhren, so wird die Kurve der Produktionsgeschwindigkeit, je genauer sie ist, um so unruhiger und zwischen 0 und einem Höchstwert hin und her schwanken, soweit zum Beispiel in der Nacht nicht gearbeitet wird. Bei größeren Zeitabschnitten verläuft die Kurve in der Höhe eines Mittelwertes.

Wie für Teile von Jahren kann man auch für Vielfache von Jahren die Produktionsmengen direkt ablesen. Ein solches Teilen und Zusammenfügen ist bei nebeneinanderliegenden Flächen ohne weiteres möglich, bei nebeneinanderliegenden Ordinaten nicht. Wenn Augenblickszustände geschildert werden, wie bei der Mitgliederbewegung, wird man die Mengen niemals zu addieren haben: gewisse Individuen tauchen jahrelang hintereinander immer wieder auf. Man müßte sozusagen eine Reihe von Menschen durch die Jahre weitermarschieren lassen; gewisse verschwinden, neue kommen hinzu, andere bleiben. Diese Mitglieder dürfen keine Fläche einnehmen, sondern nur eine Linie; man müßte die Möglichkeit haben, diese Linien immer dichter zu legen — für die zwischenliegenden Tage. In diesem Falle dürfen die Mengen in einer Ebene nicht durch Flächen, sondern durch Linien dargestellt werden.

Geht man dazu über, statistische Mengen durch Mengen von Zeichen abzubilden, so haben sie streng genommen im Sinne obiger Darstellung verschiedene Bedeutung. Symbolisieren wir produzierte Eisenmengen durch kleine Waggons, so sind diese als Flächenteile von bestimmter Grund-

linie und bestimmter Höhe aufzufassen. Übereinandergestellt füllen sie eine rechteckige Jahresfläche aus. Mehrere Jahresflächen aus Waggons bestehend, nebeneinandergestellt, bilden selbst die Stufenkurve. Nun kann man die Halbjahresproduktion durch Halbieren der Jahresflächen erhalten, ebenso die Vierteljahresproduktion usw.

Die bildhafte Darstellung drängt diese Art der Veranschaulichung geradezu auf, man kommt kaum in die Versuchung, die nebeneinander dicht angeordneten Waggonsäulen als Linien zu interpretieren, um dann die Mittelpunkte der obersten Waggons mit einer Kurve zu verbinden, deren Widersinnigkeit oben auseinandergesetzt wurde!

Während hier die bildhafte Darstellung wirkliche Schwierigkeiten beseitigt, ist sie den Fällen, in welchen Zustände wiedergegeben werden sollen, wie Mitgliederstände, Temperatur usw. weniger angepaßt, ohne aber praktisch zu irgendwelchen Irrtümern Anlaß zu geben. Wenn man zum Beispiel den Mitgliederstand am Ende verschiedener Jahre abbildet, sollte man eigentlich die Figuren als dünne Scheiben normal auf die Zeichenebene aufsetzen, um sie beliebig dicht aneinanderrücken zu können: Mitgliederstand am 1. Jänner, am 2. Jänner, am 3. Jänner usw., dann am 1. Jänner 6 Uhr, 6 Uhr 1 Minute, 6 Uhr 2 Minuten usw. Wenn man dennoch den Figuren in der Zeichenebene Höhe einräumt, geschieht es unter der stillschweigenden Annahme, daß man beliebig viele solcher Reihen einschieben dürfte. Statistisch hat auf diesem Bilde nur die Länge der Grundlinie eine Bedeutung, nicht aber die Höhe der Figuren. In diesem Falle wäre die Verbindung der Endpunkte durch eine Linie gestattet, was aber die Darstellungsweise nicht nahelegt.

Die Kurven treten mit einem viel stärkeren Anspruch auf Exaktheit auf als die bildlichen Darstellungen. Um so peinlicher ist es, wenn sie diese Ansprüche nicht erfüllen. Jede Statistik, die mit abzählbaren Elementen arbeitet, wie die Bilderstatistik, gibt kaum Anlaß zu Irrtümern. Sie hat nicht nur pädagogische, sondern auch logische Vorteile.