

Aufbau und Analyse der Datenbank 'Sozialgeschichte Berlin's von 1650 bis 1799'

Kapelle, Günter; Reymann, Wolfgang; Schwarz, Rainer

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Kapelle, G., Reymann, W., & Schwarz, R. (1988). Aufbau und Analyse der Datenbank 'Sozialgeschichte Berlin's von 1650 bis 1799'. *Historical Social Research*, 13(4), 3-54. <https://doi.org/10.12759/hsr.13.1988.4.3-54>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

Aufbau und Analyse der Datenbank »Sozialgeschichte Berlins von 1650 bis 1799«

*Günter Kapelle, Wolfgang Reymann, Rainer Schwarz**

Abstract: Under the title »Social History of Berlin from 1650 to 1799« a data bank was established at the Academy of Sciences of the GDR between 1981 and 1985. This data bank contains social-historical data of approximately 130.000 persons listed in the parish-registers of the Berlin parishes St. George and St. Nikolai as well as the Huguenot parish in the period between 1650 and 1799. According to the sources the data bank consists of fundamental files TAUFGÄNGEN (baptisms), HEIRATEN (marriages), BEERDIGUNGEN (funerals), HUGHEIRATEN (Huguenot-marriages) and HUGBEERDIGUNGEN (Huguenot-funerals) and of a secondary file JAHRE (ages) that describes the socioeconomical environment. The data bank was used as the empirical basis for the book by Helga Schultz »Berlin 1650 bis 1800 - Sozialgeschichte einer Residenz« (Social History of a Residence), AkademieVerlag Berlin, 1987. In the present article the authors describe the data model, the way in which the data bank was established, the documentation of the features and they outline the strategy for analyzing the data.

* In Konsultation mit: Prof. Dr. H. Schultz (Zentralinstitut für Geschichte der Akademie der Wissenschaften der DDR), Dr. J. Wilke (Institut für Wirtschaftsgeschichte der Akademie der Wissenschaften der DDR), K. Gaede (Zentralinstitut für Geschichte der Akademie der Wissenschaften der DDR).

Address all Communications to: Zentrum für gesellschaftswissenschaftliche Information der Akademie der Wissenschaften der DDR, Abteilung RV, Forschungsgruppe Statistische Datenanalyse, Leipziger Str. 3-4, DDR-1086 Berlin, German Democratic Republic.

0. Vorwort

Im Zusammenhang mit der damals bevorstehenden 750-Jahrfeier der Stadt Berlin wurde im Jahr 1980 die Idee geboren, eine Analyse der Sozialgeschichte Berlins zur Zeit der absolutistischen Herrschaft zu erarbeiten. Das vorliegende Quellenmaterial, das einer statistischen Analyse unterzogen werden sollte, erwies sich sehr bald als derartig umfangreich, daß die Anwendung leistungsfähiger Rechentechnik notwendig wurde. Es lagen Erfahrungen bei der Anwendung mathematischer und statistischer Verfahren in verschiedenen gesellschaftswissenschaftlichen Disziplinen vor und es galt zu klären, ob diese Erfahrungen auf die historischen Wissenschaften übertragbar waren. Im Jahre 1983 konstituierte sich mit H. Schultz und J. Wilke als wissenschaftlichen Hauptautoren, G. Kapelle, W. Reymann und R. Schwarz als Bearbeitern des EDV-Projektes und K. Gae-de für die Codierarbeiten eine interdisziplinäre Arbeitsgruppe. Die zu lösenden Aufgaben waren für alle Beteiligten hinsichtlich des Umfanges, der Komplexität, aber auch des Zeitmaßes neu. Historiker und Mathematiker wurden vor die Aufgabe gestellt, eine gemeinsame Sprache zu finden. Die simultane Behandlung der unterschiedlichsten Aufgaben, d.h. die Entwicklung der historisch orientierten Analysestrategie, der Aufbau der Datenbank und die Ableitung und praktische Anwendung einer adäquaten statistischen Analysemethodologie, war eine Herausforderung für alle Beteiligten.

Die Ergebnisse der Analyse wurden 1987 in dem Buch von H.Schultz »Berlin 1650-1800 Sozialgeschichte einer Residenz«, erschienen im Akademie-Verlag, veröffentlicht.

Die vorliegende Ausarbeitung beschreibt die Entstehung, die Struktur und die Arbeiten zur Analyse der Datenbank »Sozialgeschichte Berlin's 1650 - 1799«. Die Autoren verbinden mit dem Erscheinen dieser Informationen die Hoffnung, daß interessierte Historiker, Mathematiker und andere quantitativ orientiert forschende Gesellschaftswissenschaftler helfen, den wertvollen Datenfonds weiter zu analysieren bzw. quantitativ gestützt, methodische und methodologische Erkenntnisse über die Anwendung der statistischen Datenanalyse in den Geschichtswissenschaften vertiefen zu können. Dabei könnten vergleichende Untersuchungen, Untersuchungen zur Sozialgeschichte der Medizin, vertiefte demographische Untersuchungen und Arbeiten zu den einzelnen Branchen, Berufen, Klassen und Schichten als mögliche Forschungsgebiete dienen, die bisher nur überblicksweise ausgewertet werden konnten.

1. Quellenmaterial

Die Datenbank basiert auf den Kirchenbüchern von

- St. Nikolai/Berlin 1650 - 1799 und
- St. Georgen/Berlin 1689 - 1799 (Tochterparochie von St. Nikolai).

Es wurden folgende Quellen benutzt:

- Kirchenbuchverfilmungen der Zentralstelle für Genealogie Leipzig, Dimitroffplatz 1, DDR-7010 Leipzig,
- Originalkirchenbücher von St. Georgen (Gesamtkirchenbuch bis 1714; Beerdigungsregister bis 1765) im Archiv der Evangelischen Georgen-Parochial-Kirchengemeinde, Waisenstraße 28a, DDR-1020 Berlin,
- Lichtbildkopien der Beerdigungsregister von St. Nikolai (1724-1735) Kirchenbuchstelle Altberlin, Jebenstraße 3, D-1000 Berlin(West) 12.

Die Datenbank für die Hugenottengemeinde basiert auf folgenden Registern der Französischen Gemeinde in Berlin:

- Registres de mariages de l'église française a Berlin:
1696 - 1706 (Vol. 1),
1725 - 1743 (Vol. 2),
1756 - 1762 (Vol. 3),
1790 - 1800 (Vol. 3,8);
- Registres mortuaires de l'église française a Berlin:
1695 - 1706 (Vol. 1),
1746 - 1756 (Vol. 6),
1790 - 1800 (Vol. 7,8)

Die Register stammen aus dem Archiv des Consistoriums der Französischen Gemeinde in Berlin, Platz der Akademie, (Französischer Dom), DDR-1080 Berlin.

2. Datenmodell

Eine Grundforderung an die Erfassung historischer Daten ist die Bewahrung der Originaltreue. Man spricht in diesem Zusammenhang von einer quellennahen Datenerfassung. Optimale Voraussetzung dafür wäre die Anwendung portabler leistungsfähiger Mikrorechenstechnik. Mittels geeigneter Software wäre eine Erfassung der Informationen vor Ort, d.h. auch im Archiv möglich. Für das vorliegende Projekt waren derartige Voraussetzungen nicht gegeben. Dieser Mangel wurde mittels eines erhöhten personellen Aufwandes kompensiert. Als Kompromiß zwischen den Wün-

sehen der Historiker nach möglichst großer Quellennähe und den Restriktionen, die die verfügbare Software und Hardware auferlegten, wurden die primär vorliegenden textlichen Informationen über eine numerische Codetabelle erfaßt. Es ist bekannt, daß dieses Vorgehen international kontrovers diskutiert wird. Im vorliegenden Fall zeigten sich aber auch Vorteile:

- der zu erfassende Datensatz besteht nur aus wenigen Zahlen, hat somit eine minimale Länge, was eine Senkung des Codieraufwandes nach sich zieht und die Fehlersuche vereinfacht,
- der Speicherplatz auf maschinenlesbaren Datenträgern und der Hauptspeicherbedarf bei der Analyse der Daten werden minimiert und
- die Anwendung mathematischer Methoden zur Datenanalyse wird erleichtert.

Andererseits erforderte die Definition von Codelisten für die Merkmalsausprägungen bereits zu Beginn der Arbeiten gewisse endgültige Entscheidungen über die Zuordnung einer Quelleninformation zu einem Code. Da jedoch in der Startphase des Projekts keine geeignete Software zur direkten Quelleneingabe verfügbar oder herstellbar war, ist die Entscheidung pro Codeliste ohne Alternative gewesen.

Nicht alle Eintragungen in den Kirchenbüchern waren sofort einem Code zuordenbar. Beim Entwurf der Datenerfassungsbelege wurde dies durch extra Rubriken für primäre Quelleneinträge berücksichtigt, so z.B. für bestimmte Berufe und Todesursachen. In zunächst unklaren Fällen konnte auf diese Weise die endgültige Entscheidung hinausgeschoben werden, bis die Sachlage hinreichend geklärt war.

Trotz der augenscheinlichen Vorteile der numerischen Codierung für die elektronische Datenverarbeitung sei aber nicht verschwiegen, daß seitens der Historiker eine alphanumerische Erfassung der Quelleninformationen favorisiert wird.

Autorin der Codeliste ist H. Schultz. Auf sie geht auch die Entscheidung zurück, die Ortsnamen nicht zu codieren, sondern nur Region und Ortsgröße in die Datenbank aufzunehmen.

Die schrittweise Erfassung und Codierung wurde dann erforderlich, wenn viele verschiedene Mitarbeiter die Primärerfassung, d.h. die Übertragung der Kirchenbucheinträge auf Datenerfassungsblätter, realisierten. Da eine direkte Codierung durch jeden dieser Mitarbeiter bestimmt sehr fehlerintensiv hätte werden können, wurden im ersten Schritt die Textinformationen auf die Datenerfassungsblätter übernommen und im zweiten Schritt durch zwei spezialisierte Mitarbeiter numerisch codiert. Die Erfassungbelege enthalten mithin quellennähere und umfangreichere Informationen, als die in der Datenbank gespeicherten Daten und wurden deshalb im Zentralinstitut für Geschichte archiviert.

Für jede Ereignisart (Taufe, Hochzeit, Beerdigung) wurde ein spezieller Erfassungsbeleg entworfen, der das Rohdatenformat definiert. Da die verfügbare Analysesoftware grundsätzlich nur rechteckige Datenstrukturen verarbeiten kann, wurden alle Beobachtungen (Personen) einer Ereignisart zu einer Basisdatei zusammengefaßt. Die nachfolgende Darstellung der inhaltlichen Struktur dieser Basisdateien widerspiegelt sowohl die Merkmalsnamen als auch die Merkmalsetiketten.

Merkmalsnamen (z.B. VAR1) werden möglichst kurz gewählt, denn sie dienen zur Steuerung der Datenbehandlung und werden benötigt, um für die Datenanalyseprozeduren Auswahlkriterien zu definieren. Zur übersichtlicheren und eindeutigen Darstellung von Ergebnissen erhalten die Merkmalsnamen Merkmalsetiketten zugeordnet, wodurch jeweils die inhaltliche Bedeutung der Merkmale näher beschrieben wird. Analog zur erklärenden Verbindung von Merkmalen und Merkmalsetiketten, werden die codierten Merkmalsausprägungen durch Wertetiketten verständlich gemacht. Diese Zuordnung wird häufig auch als Codetabelle bezeichnet, in der für jedes Merkmal festgelegt wird, durch welches Etikett der entsprechende Merkmalswert ersetzt werden soll. Da dieser Prozeß eine Formatierung der auszugebenden Merkmalswerte darstellt, spricht man anstelle von Wertetiketten auch von Formaten. Es ist möglich, daß mehrere Merkmale ein identisches Format benutzen. Die Merkmalsetiketten werden direkt in den Basisdateien gespeichert, während die Formate in einer separaten Formatbibliothek abgelegt und den Daten über einen eindeutigen Formatnamen zugeordnet werden.

2.1 Basisdatei T A U F E N

Jede Person, die im folgenden auch Beobachtung oder Fall genannt werden soll, wird durch 41 Merkmale beschrieben. In Tabelle 1 werden die zusammengehörigen Merkmalsnamen und Merkmalsetiketten zusammengestellt.

Merkmalsname	Merkmalsetikett	Formatname
VAR1	LAUFENDE	NUMMER
VAR2	KIRCHSPIEL	KSPIEL
VAR3	JAHR UND MONAT	DATE
VAR4	JAHR	
VAR5	MONAT	MONAT
VAR6	JAHRZEHT	JHZEHT

VAR7	GENERATIONSZEITRAUM	GZEITR
VAR8	FAMILIENSTAND	FAMST
VAR9	GESCHLECHT	GESCHL
VAR10	VW-ZWEIG VATER	VWZWEIG
VAR11	BERUF VATER	BERUFE
VAR12	SOZIALE GRUPPE VATER	SOZGRUP
VAR13	KLASSE-SCHICHT VATER	KLASSE
VAR14	PATEN INSGESAMT	
VAR15	PATEN WEIBLICH	
VAR17	PATEN MAENNLICH	
VAR18	GESCHLECHT PATE1	GESCHL
VAR19	VW-ZWEIG PATE1	VWZWEIG
VAR20	BERUF PATE1	BERUFE
VAR21	SOZIALE GRUPPE PATE1	SOZGRUP
VAR22	KLASSE-SCHICHT PATE1	KLASSE
VAR25	GESCHLECHT PATE2	GESCHL
VAR26	VW-ZWEIG PATE2	VWZWEIG
VAR27	BERUF PATE2	BERUFE
VAR28	SOZIALE GRUPPE PATE2	SOZGRUP
VAR29	KLASSE-SCHICHT PATE2	KLASSE
VAR32	GESCHLECHT PATE3	GESCHL
VAR33	VW-ZWEIG PATE3	VWZWEIG
VAR34	BERUF PATE3	BERUFE
VAR35	SOZIALE GRUPPE PATE3	SOZGRUP
VAR36	KLASSE-SCHICHT PATE3	KLASSE
VAR39	GESCHLECHT PATE4	GESCHL
VAR40	VW-ZWEIG PATE4	VWZWEIG
VAR41	BERUF PATE4	BERUFE
VAR42	SOZIALE GRUPPE PATE4	SOZGRUP
VAR43	KLASSE-SCHICHT PATE4	KLASSE
VAR46	GESCHLECHT PATE5	GESCHL
VAR47	VW-ZWEIG PATE5	VWZWEIG
VAR48	BERUF PATE5	BERUFE
VAR49	SOZIALE GRUPPE PATE5	SOZGRUP
VAR50	KLASSE-SCHICHT PATE5	KLASSE

Tab. 1: Merkmalsetiketten und Formate der Basisdatei TAUFEN

Die Lücken in der Merkmalsindizierung (z.B. VAR43, VAR46) resultieren aus dem Bearbeitungsprozeß während des Aufbaus der Datenbank und die Abkürzung VW steht für Volkswirtschaftszweig. Weiterhin ist zu bemerken, daß nur bis zu 5 Paten, in der Reihenfolge ihrer Nennung im Kirchenbuch, unter Auslassung der Frauen ohne erkennbare soziale Zuordnung, erfaßt worden sind. Dies war die obrigkeitlich verordnete Höchstzahl der Paten. Häufige Überschreitungen gehen aus VAR14 bis VAR17 hervor, sind jedoch für die Analyse der sozialen Beziehungen nicht erheblich.

2.2 Basisdatei HEIRATEN

Merkmals- name	Merkmalsetikett	Formatname
VAR1	LAUFENDE NUMMER	
VAR2	KIRCHSPIEL	KSPIEL
VAR3	JAHR UND MONAT	DATE
VAR4	JAHR	
VAR5	MONAT	MONAT
VAR6	JAHRZEHNT	JHZEHNT
VAR7	GENERATIONSZEITRAUM	GZEITR
VAR8	VOLKSWIRTSCHAFTSZWEIG BRAEUTIGAM	VWZWEIG
VAR9	BERUF BRAEUTIGAM	BERUFE
VAR10	SOZIALE GRUPPE BRAEUTIGAM	SOZGRUP
VAR11	KLASSE - SCHICHT BRAEUTIGAM	KLASSE
VAR12	MONATSGEWICHT	
VAR13	REGION BRAEUTIGAM	REGION
VAR14	ORTSGROESSE BRAEUTIGAM	GROESSE
VAR15	VOLKSWIRTSCHAFTSZWEIG BRAEUTIGAMVATER	VWZWEIG
VAR16	BERUF BRAEUTIGAMVATER	BERUFE
VAR17	SOZIALE GRUPPE BRAEUTIGAMVATER	SOZGRUP
VAR18	KLASSE - SCHICHT BRAEUTIGAMVATER	KLASSE
VAR21	FAMILIENSTAND DER BRAUT	FAMST
VAR22	VOLKSWIRTSCHAFTSZWEIG BRAUTVATER	VWZWEIG
VAR23	BERUF BRAUTVATER	BERUFE
VAR24	SOZIALE GRUPPE BRAUTVATER	SOZGRUP
VAR25	KLASSE - SCHICHT BRAUTVATER	KLASSE
VAR28	REGION BRAUT	REGION
VAR29	ORTSGROESSE BRAUT	GROESSE

Tab. 2: Merkmalsetiketten und Formate der Basisdatei HEIRATEN

In dieser Datei wird jeder Fall durch 25 Merkmale beschrieben, deren Bedeutung aus den Merkmalsetiketten in Tabelle 2 hervorgeht. Zum Ausgleich der unterschiedlichen Monatslängen bei der Analyse saisonaler Schwankungen, wurde für diese Datei das Merkmal MONATSGEWICHT eingeführt. Bei einer Wiederheirat der Braut beziehen sich die Angaben zum Brautvater auf ihren früheren Ehemann.

2.3 Basisdatei BEERDIGUNGEN

Jeder Fall wird durch 17 Merkmale beschrieben.

Merkmalname	Merkmalsetikett	Formatname
VARI	LAUFENDE NUMMER	
VAR2	KIRCHSPIEL	SPIEL
VAR3	JAHR UND MONAT	DATE
VAR4	JAHR	
VAR5	MONAT	MONAT
VAR6	JAHRZEHNT	JHZEHNT
VAR7	GENERATIONSZEITRAUM	GZEITR
VAR8	VOLKSWIRTSCHAFTSZWEIG	VWZWEIG
VAR9	BERUF	BERUFE
VARIO	SOZIALE GRUPPE	SOZGRUP
VAR11	KLASSE - SCHICHT	KLASSE
VARI 2	GESCHLECHT	GESCHL
VARI 3	ALTER	FSTP
VARI 5	ALTERSGRUPPE	ALTGRUP
VARI 6	TODESURSACHE	DIAGNOSE
VARI 7	BEERDIGUNGSGEBUEHREN	FSTP
VAR22	TODESURSACHENGRUPPEN	DIAGRUP

Tab. 3: Merkmalsetiketten und Formate der Basisdatei BEERDIGUNGEN

2.4 Sekundärdatei JAHRE

Die personenorientierten Basisdateien wurden auf jährlicher Basis aggregiert und durch personenunabhängige Informationen über das sozio-ökonomische Umfeld ergänzt. Es handelt sich um folgende Merkmale bzw. Quellen:

- Brotpreise in Berlin:
Brotpreis 1578 - 1700: Stadtarchiv Berlin, Bestand Amtsbücher, HS59: Bäcker- und Fleischerordnungen 16. Jahrhundert; Stadtarchiv Potsdam, Rep. 30, Bln. A, Nr. 483: Kornbuch 1716 - 1763; ebenda, Nr. 448: Semmel und Brodt- Kalkulation; »Wöchentliche Berlinische Frag- und Anzeigungsnachrichten... nebst den marktgängigen wöchentlichen Korn- und Wollpreisen, Bier-, Brodt- und Fleisch-Taxe...«, 1746 - 1763; »Berlinische Intelligenzzettel...«, 1763 - 1781; »Neues Berliner Intelligenzblatt...«, 1786 - 1799;
- Berliner Temperaturen:
in: Otto Behre, Das Klima von Berlin. Eine meteorologisch-hygiene-

- sehe Untersuchung, Berlin 1908;
- EINWOHNERZAHL TOTAL:
Sterbe-, Heirats-, Geburten- und Nettowanderungsrate total 1740 - 1799: Zentrales Staatsarchiv Merseburg, Generaldirektorium Kurmark, Zit. 265, Nr. 1-6; Johann Peter Süßmilch, Der Königlichen Residentz Berlin schneller Wachsthum und Erbauung..., Berlin 1752; als KRIEGSJAHR wurden von Prof. Schultz alle die Jahre eingestuft, in denen sich Brandenburg-Preußen im Kriege befand, nicht nur die, in denen die Kurmark Kriegsschauplatz war,
 - die EINWOHNERZAHL der beiden Untersuchungskirchspiele wurden aus den jährlichen Beerdigungszahlen rekonstruiert (siehe Punkt 7.1.1), wobei weitgehend der Methode von J.P. Süßmilch gefolgt wurde,
 - die korrigierten Zahlen der jährlichen Taufen, Heiraten, Beerdigungen ergeben sich aus der Interpolation fehlender Daten in den Kirchenbüchern, indem bei Fehlen einzelner Blätter die Monatssummen entsprechend dem Anteil dieser Monate an der Ereignisart in vollständig überlieferten Jahren angenommen wurden,
 - die Sterbe-, Heirats- und Geburtenraten (Ereignisse je 1000 Einwohner) wurden aus den korrigierten jährlichen Zahlen der Ereignisse und den extrapolierten Einwohnerzahlen errechnet,
 - die Zahlen der jährlich aufgetretenen Krankheiten ergaben sich direkt aus der Basisdatei,
 - die Raten der Krankheiten können aus den jährlichen Zahlen (ab 1719 erst vollständig) und den extrapolierten Einwohnerzahlen errechnet werden.

Merkmals- name	Merkmalsetikett	Zusatz- information
V1	Jahr	
V2	Jahrzehnt	
V3	Generationszeitraum	
V4	Jährliche Sterbeziffern	
V5	Korrigierte J.Sterbeziffern	
V6	Pocken	
V7	Ritteln etc	
V8	Hitziges Fieber	
V9	Stickfluss	
V10	Ruhr	
V11	Sonstiges Fieber	
V12	Bauch - u. Magenkrämpfe	
V13	Jammer	
V14	Brustkrankheit	
V15	Kriegsjahr 0-Nein 1-Ja	

V16	LOT ROGGENBROT FUER EINEN GROSCHEN	+
V17	MITTLERE JAHRESTEMPERATUR	+
V18	MITTLERE JANUARTEMPERATUR	+
V19	MITTLERE JULITEMPERATUR	+
V20	JAEHRL.HEIRATZIFFER	
V21	KORR. J. HEIRATZIFFER	
V22	EINWOHNERZAHL TOTAL	+
V23	STERBERATE TOTAL	+
V24	GEBURTENRATE TOTAL	+
V25	WACHSTUMSRATE TOTAL	+
V26	NETTOWANDERUNGSRATE TOTAL	+
V27	HEIRATSRATE TOTAL	+
V28	EINWOHNERZAHL	
V29	STERBERATE	
V30	HEIRATSRATE	
V31	STERBEZIFFER TRENDWERT	
V32	WIEDERHEIRATEN DER BRAUT	
V33	JAEHRLICHE TAUFEN	
V34	KORR.J.TAUFEN	
V35	GEBURTENRATE	
V36	GEBURTENUEBERSCHUSS	

Tab. 4: Merkmalsetiketten der Datei JAHRE

In der Tabelle 4 sind alle aggregierten Merkmale und die Zusatzinformationen, die mit '+' gekennzeichnet sind, aufgeführt. Für die Sekundärdatei JAHRE wurden keine Werteetiketten definiert, deshalb entfällt die Angabe über das Format.

2.5 Gesamtdarstellung der Datenbank

Moderne Datenanalysesysteme integrieren Datenmanagement und statistische Analyse. Nutzerfreundlich werden die Beschreibung der Merkmale durch Merkmalsetiketten und die Beschreibung der Merkmalsausprägungen durch Werteetiketten unterstützt.

Die Steuerinformationen zum Aufbau der Dateien, zu den Merkmalsetiketten, zu Transformationen der Dateien und zur Fehlerkontrolle werden in einer Quelltextbibliothek gespeichert.

Die Gesamtdarstellung der Datenbank »Sozialgeschichte Berlin's 1650-1799« ist in Abbildung 1 ersichtlich. Der Teil, der sich auf die Hugentengemeinde mit den Basisdateien HUGHEIRATEN und HUGBEERDIGUNGEN bezieht, wird in Abschnitt 8 beschrieben.

Die Basisdateien und die Sekundärdatei sind rechteckige Datenmatrizen, wobei die Spalten die Merkmale, die Zeilen in den Basisdateien die Personen und in der Sekundärdatei die Jahre repräsentieren. Eine Zeile

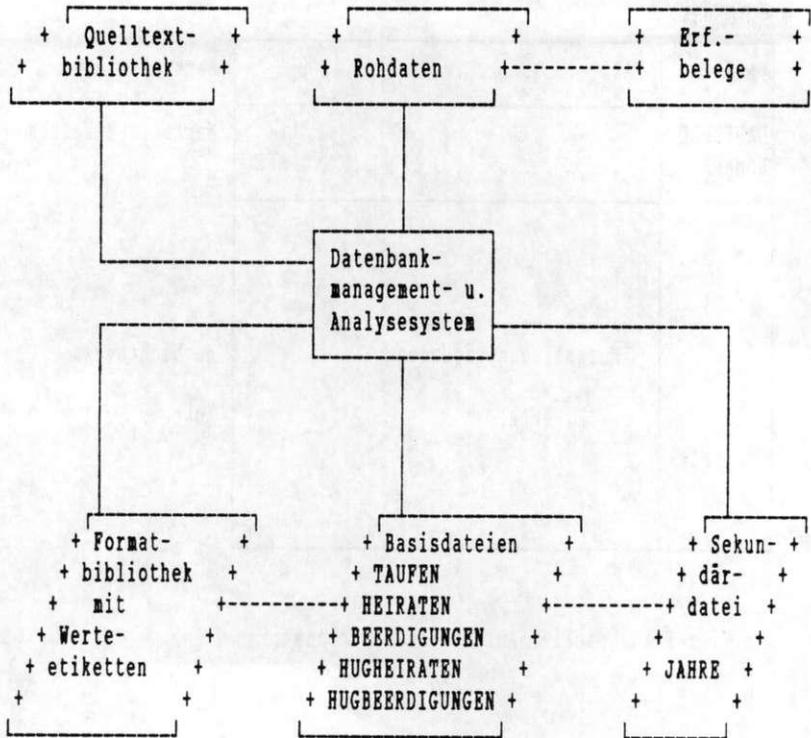


Abb. 1 : Datenbank 'Sozialgeschichte Berlin's 1650-1799'

der Datenmatrix wird oft auch als Datensatz bezeichnet (siehe Abbildung 2).

3. Projektablaufplanung

Projekte vorliegender Größenordnung erfordern eine sorgfältige Ablaufplanung, da die parallele Bearbeitung verschiedener Arbeitsgänge leicht dazu führen kann, daß der Überblick verlorengeht und nichtreparable Fehler oder kostenintensive Fehlerkorrekturen auftreten. Die Projektplanung muß insbesondere die typischen Bedingungen der EDV-Anwendung

Merkmale

VAR1	VAR2	Merkmalsnamen
LAUFERDE NUMMER	Merkmalsetiketten
1 2 3 n	Merkmalsausprägungen (Werte)	Beobachtungen (Datensätze)

Abb. 2 : Prinzipieller Aufbau einer Datematrix

berücksichtigen. Hier sind wirkungsvolle Vorkehrungen mit dem Ziel maximaler Datensicherheit so zu treffen, daß die Daten unverfälscht, nur durch kontrollierte Transformationen und Editionen verändert, von der Datenerfassung zur Analyse gelangen. Durch Sicherheitskopien sind sie jederzeit vor Verlust zu schützen. Der Umfang geeigneter Maßnahmen war so groß, daß sich daraus eine separate Disziplin der statistischen Datenanalyse - die Datenvorverarbeitung - entwickelte. Grob lassen sich drei wesentliche Aufgabenkomplexe des Projektes unterscheiden:

- die Datenerfassung mit manueller Aufnahme und Übertragung auf maschinenlesbare Datenträger,
- die Datenvorverarbeitung mit Datenprüfung, Datenkorrektur und Aufbau von analysefähigen Dateien und
- die statistische Datenanalyse mit der Erstellung von Ergebnistabellen und deren Interpretation.

Der personelle und rechentechnische Zeitaufwand verteilte sich zu 50% auf Komplex 1 und zu jeweils 25% auf die Komplexe 2 und 3. Es sei

angemerkt, daß die Zeitaufwendungen sowohl für Komplex 2 als auch für Komplex 3 hard- und softwareseitig beeinflusst werden können.

Abbildung 3 veranschaulicht grob die zeitliche Aufsplittung der Teilaktivitäten, die mehrfach parallel abgearbeitet werden mußten.

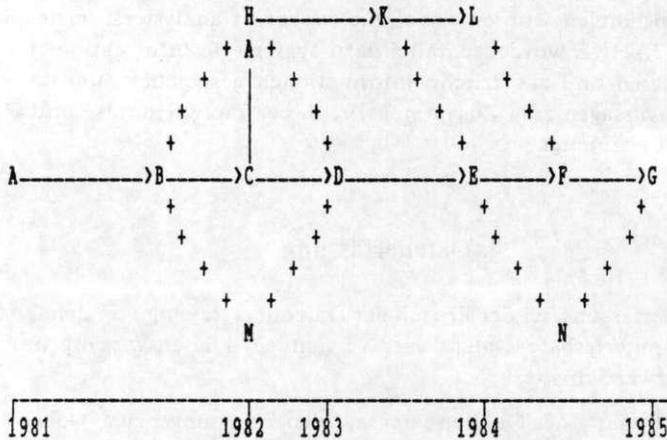


Abb. 3 : Netzwerk des Projektverlaufes

In diesem Netzwerk bedeuten:

- A - B: Entwurf des Datenmodells
(Dazu gehören die Festlegung der zu erfassenden Merkmale, die Entscheidung über die Codetabellen und der Entwurf der Erfassungsbelege).
- B - H: Datenerfassung für die Basisdatei BEERDIGUNGEN
(Hierzu zählen die Übertragung der Daten in die Belege und die maschinelle Erfassung auf Magnetbandkassetten).
- H - K: Datenerfassung für die Basisdatei HEIRATEN
- K - L: Datenerfassung für die Basisdatei TAUFEN
- B - C: Programmierung und Testung der Vorverarbeitungsprozeduren
- B - D: Entwicklung der Analysestrategie
- B - M: Erfassen der Werteetiketten
- M - D: Aufbau der Etikettendatei
- H - D: Vorverarbeitung für die Basisdatei BEERDIGUNGEN
- K - E: Vorverarbeitung für die Basisdatei HEIRATEN
- L - F: Vorverarbeitung für die Basisdatei TAUFEN

- D - E: Statistische Analyse der Basisdatei BEERDIGUNGEN
- E - F: Statistische Analyse der Basisdatei HEIRATEN
- E - N: Erfassung für die Sekundärdatei JAHRE
- F - G: Statistische Analyse der Basisdatei TAUFEN
- N - G: Aufbereitung und Analyse der Sekundärdatei JAHRE

Die drei Basisdateien wurden dabei zuerst separat analysiert. In der Sekundärdatei JAHRE wurden parallel dazu aggregierte Informationen aus den Basisdateien und zusätzliche Informationen gespeichert, um daraus komplexere Aussagen zum Zusammenwirken der Tauf-, Heirats- und Geburtenzahlen abzuleiten.

4. Datenerfassung

Unter Datenerfassung sei der Prozeß der Datenübertragung von den Quellen auf computerlesbare Datenträger verstanden. Dieser Vorgang unterteilt sich in zwei Schritte:

- die Übertragung **der** **Quellendaten** auf den Erfassungsbeleg, wobei die Originaleintragungen nach einer Codiervorschrift in Zahlenwerte transformiert werden und
- die Übertragung der codierten Werte auf einen **maschinenlesbaren** Datenträger.

Wird, wie im Fall des vorliegenden Projektes, der Datenerfassungsbeleg selbst als Arbeitsgegenstand verwendet, auf dem stufenweise die Codierung erfolgt, übt dessen Gestaltung einen starken Einfluß auf die Datensicherheit aus. Von den Erfassungsbelegen wurden die Daten mit Hilfe von Datenerfassungsgeräten mit einem speziell für dieses Projekt erarbeiteten Erfassungsprogramm auf Magnetbandkassetten übertragen. Leider konnten durch die geringe Leistungsfähigkeit des Datenerfassungsgerätes nur sehr eingeschränkte Fehlerkontrollen bei der Datenerfassung direkt durchgeführt werden. Erfassungsfehler wurden dadurch meist erst nach Konvertierung der Magnetbandkassetten im Verlauf der Fehlerkontrollen auf dem Großrechner festgestellt. Doppelerfassungen, Auslassungen, Vertauschung verschiedener Rohdatenbestände usw. konnten daher oft erst nach Wochen korrigiert werden. Dann war aber die Rückkopplung zu den Datenerfassungskräften nicht mehr gegeben, so daß der Aufwand für die Fehlerkorrektur sehr hoch wurde.

5. Datenvorverarbeitung

Die Datenvorverarbeitung umfaßt alle Arbeitsschritte von der Übernahme der erfaßten Daten bis zur Bereitstellung einer analysefähigen Systemdatei. Unter Systemdatei versteht man eine nach inhaltlichen Gesichtspunkten zusammengefaßte Datenmenge, die unter der Steuerung eines speziellen Datenverarbeitungssystems zur effizienten Weiterverarbeitung durch systemeigene Prozeduren, einschließlich aller Vereinbarungen über Etiketten und Fehlstellen abgespeichert wurde. In unserem Projekt stellen die Basisdateien TAUFEN, HEIRATEN, BEERDIGUNGEN und die Sekundärdatei JAHRE jeweils eine solche Systemdatei dar.

Bis zu einer analysefähigen Systemdatei durchlaufen die Daten verschiedene Stadien: vom Rohdatenformat, wie sie durch die Erfassungstation bereitgestellt werden, über eine ungeprüfte, dann korrigierte Systemdatei und nach Transformationen zu einer analysefähigen Systemdatei oder kurz zu einer Analysedatei.

5.1 Dateneingabe

Im ersten Arbeitsschritt müssen die erfaßten Daten von der Rohdatenstruktur in eine verarbeitungsspezifische Form überführt werden. Dies wird für jede Basisdatei gesondert mit einer Dateneingaberoutine realisiert. Im Ergebnis entsteht eine Systemdatei, die das genaue Abbild der Rohdaten ist. Damit sind die Daten in das Analysesystem integriert und können mit den Mitteln des Datenbankmanagement- und Analysesystems weiterverarbeitet werden.

5.2 Datenprüfung

Voraussetzung einer wissenschaftlich hochwertigen Auswertung ist eine streng an den Originalquellen orientierte Datenerfassung, dem Datenmodell entsprechende, logisch widerspruchsfreie Speicherung und unverfälschte Weiterverarbeitung der Daten. Der Umfang des Projektes und die mehrjährige Bearbeitungszeit erzwangen deshalb eine über das normale Maß hinausgehende Konzentration auf die Datenprüfung.

Eine weitgehend computergestützte, d.h. automatisierte Fehlerkontrolle und -korrektur würde das Datenmodell zwar nahelegen, Aufwand und Nutzen stünden aber in ungünstigem Verhältnis. Aus diesem Grunde erfolgte eine Beschränkung derart, daß nur ein gewisser Standard definierter Fehlermöglichkeiten durch den Computer überprüft und gegebenenfalls

zur Korrektur geführt wurden. Zu den Standardfehlermöglichkeiten rechnet man:

- nichtidentifizierbare Datensätze,
- fehlende Datensätze,
- doppelte Datensätze,
- falsche Dateiidentifikation,
- nicht identifizierbare Merkmalsausprägung,
- Merkmalsausprägung außerhalb des Wertebereichs und
- fehlerhafte logische Beziehungen zwischen Merkmalen.

Sowohl die Verschiedenheit der Basisdateien als auch der spezifische Erfassungsprozess erforderten, die Daten merkmals- und beobachtungsweise, d.h. vertikal und horizontal zu durchmustern. Die Datenprüfung erfolgte deshalb in vier Prüfkomplexen, die separat, aber in der angegebenen Reihenfolge durchgeführt wurden:

- Identifikation und Selektierung vermischter Datenbestände durch Überprüfung der zulässigen Ereignisart,
- Prüfung auf Einhaltung definierter Wertebereichsgrenzen und logischer Kombinationen von Merkmalsausprägungen,
- Prüfung anhand von Zufallsstichproben und
- Häufigkeitsanalysen zur Prüfung der Vollständigkeit.

Als Ergebnis einer Datenprüfung entstehen Fehlerlisten. Geeignete Listendarstellungen bilden den Ausgangspunkt für notwendige Korrekturintragungen und die Eingabe der Korrekturdaten. Konsequenterweise muß nach der Eingabe von Korrekturdaten erneut ein Prüfprozeß durchlaufen werden. Diese Vorgehensweise ist auch deshalb notwendig, da die Editionsarbeiten vorrangig im Dialog mit dem Computer erfolgten und so eine weitere Fehlerquelle gegeben ist.

Die Komplexität der zweiten Prüfmaßnahme (Wertebereichsgrenzen und logische Kombinationen) wird durch das in Abbildung 5 dargestellte Struktogramm verdeutlicht.

Um die Datenkorrektur effektiv zu unterstützen, ist die Fehlerliste so aufgebaut, daß sie unmittelbar als Korrekturbeleg verwendet werden kann (vgl. Abbildung 6).

Die Spaltennummerierung über jeder Seite vereinfacht die Fehlersuche besonders bei Verschiebungen. Über die Datensatznummer (REC.-NO.) kann jede fehlerhafte Beobachtung eindeutig identifiziert werden. Ein Fehler wird in diesem Falle vermittels der Angabe von Merkmalsnamen und der entsprechenden Merkmalsausprägung (VAR3= 99) angezeigt. Auf diese Weise kann eine übersichtliche Anordnung der Fehlernachrichten und eine relativ bequeme Korrektur ermöglicht werden. Die Aufspürung mehrfacher Datensätze erfolgt nicht über den Vergleich aller Merk-

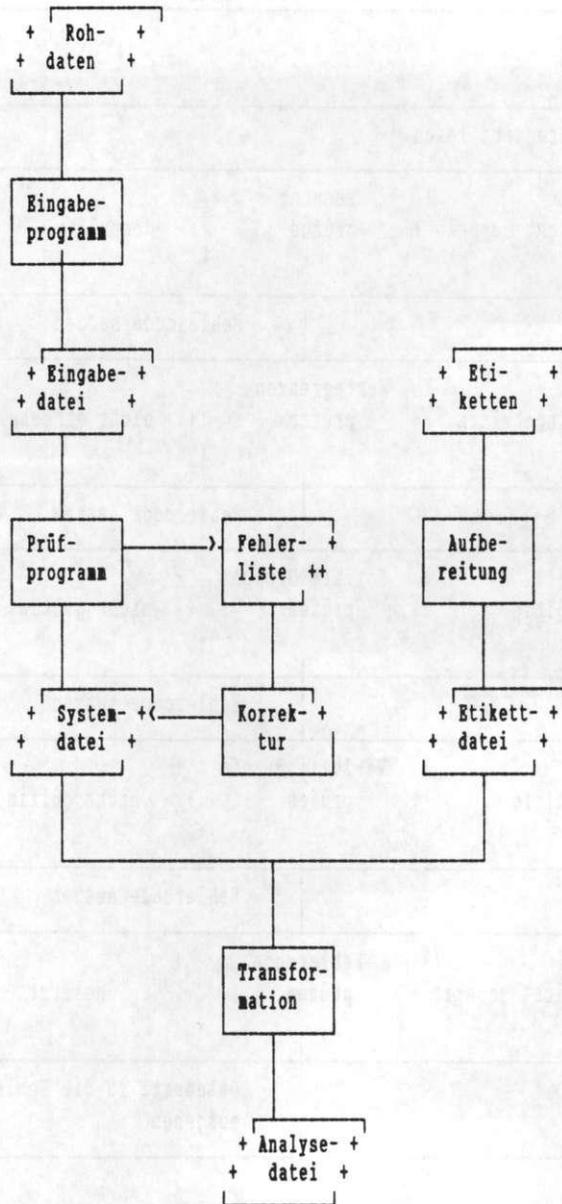


Abb. 4 : Ablauf der Datenvorverarbeitung

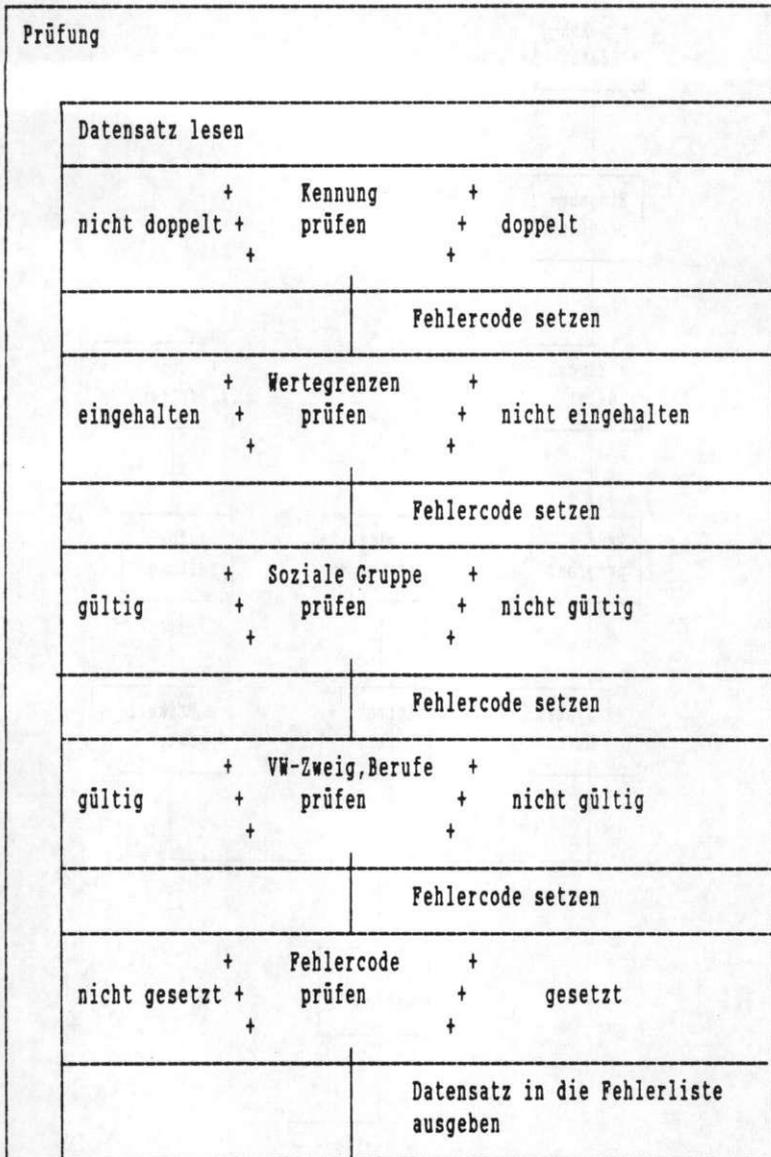


Abb. 5 : Struktogramm zur Fehleranalyse

SPALTE	1234567890123456789012345678901234567890123456789	*
REC.-NO. XXX1 Inhalt des Datensatzes	*
FEHLER	VAR3= 99 VAR8= 8 VAR20= 96	*
REC.-NO. YYY1 Inhalt des Datensatzes	*
REC.-NO. #YYY2 Inhalt des doppelten Datensatzes . . .	*

Abb. 6 : Fehlerlistenaufbau

male, sondern nach geeigneter Vorsortierung über einen Vergleich ausgewählter Merkmale. Doppelt bzw. mehrfach auftretende Datensätze werden mit »#« besonders gekennzeichnet und vor der Elimination dokumentiert.

Als Resultat der Fehleranalyse konnten bei 1,6% der Datensätze Fehler der angegebenen Klassifikation nachgewiesen werden. Insgesamt sind dies ca. 1900 Datensätze. Interessant ist, daß etwa 84% der Fehler als Eingabefehler am Datenerfassungsgerät zu identifizieren waren. Dieser Prozentsatz liegt wesentlich über der Norm. Die Ursache dürfte in der ungünstigen, zu unübersichtlichen Gestaltung der Erfassungsbelege zu suchen sein.

5.3 Datenkorrekturen

Der Modus für die Datenkorrekturen entscheidet wesentlich über den zu planenden Zeitaufwand für die Datenvorverarbeitung. Korrekturen können interaktiv mittels direkter Änderung in der Datei oder über den Aufbau einer Änderungsdatei mit nachträglichem Mischen (updating) beider Dateien realisiert werden. Vorteilhaft ist eine interaktive Editierung. Sie wurde unter TSO (= Teilnehmerunterstützung zur gleichzeitigen und unabhängigen interaktiven Nutzung der Ressourcen des Großrechners EC 1055 M) realisiert und erlaubte die direkte Änderung von Fehlern in den Systemdateien. Mit Hilfe eines Editors konnten die fehlerhaften Da-

tensätze unter Angabe der Recordnummer (REC.-NO.) aufgefunden und anschließend die Merkmalsausprägungen direkt am Bildschirm korrigiert werden.

5.4 Etikettenaufbereitung

Etiketten sind ein wichtiges Hilfsmittel bei der Interpretation von Ergebnislisten. In allen modernen Datenbank- und Analysesystemen werden Etikettierungsmöglichkeiten angeboten, die außerdem eine wichtige Funktion bei der Dokumentation der Datenbank haben. Mit Etiketten ist eine Datenbank selbstdokumentierend und erlaubt besonders bei interaktiver Arbeit einen leichten Umgang mit den Daten. Merkmalsetiketten lassen sich unkompliziert während eines Verarbeitungsschrittes (Transformationen) in eine Systemdatei integrieren.

Die Aufbereitung der Werteetiketten ist wesentlich aufwendiger, da hier allen Merkmalsausprägungen bzw. Wertebereichen ein Etikett zugeordnet werden muß. Jeweils alle Merkmalsausprägungen einer Variablen werden mit den zugeordneten Etiketten unter einem Namen (Formatname) in der Etikettendatei des Systems (Formatbibliothek), unabhängig von den eigentlichen Daten, gespeichert. Eine Verbindung zwischen Daten und Eintragungen in der Formatbibliothek muß dann in einem Verarbeitungsschritt (Transformationen) dem System mitgeteilt werden.

5.5 Datentransformationen

Umfangreiche Transformationen bilden den Abschluß der Datenvorverarbeitung. In diesem letzten Schritt vor der eigentlichen Datenanalyse werden einige Variablen, entsprechend aktualisierten Erkenntnissen, durch Zusammenfassung von Wertestufen oder arithmetische bzw. logische Umformungen geändert, neue Merkmale mittels arithmetischer bzw. logischer Verknüpfungen vorhandener Merkmale berechnet und die Systemkopplung zu den fehlerfrei aufbereiteten Werteetiketten hergestellt.

Für jede Basisdatei wurde ein Transformationsprogramm geschrieben, das aus den korrigierten, aber noch unvollständigen Systemdateien eine jeweils analysefähige Systemdatei erstellte. Alle wichtigen Informationen zum Transformationsprozeß sind unter Abschnitt 6 dargestellt. Im Ergebnis sind auch alle Eintragungen zur Etikettierung und Fehlstellenbehandlung enthalten.

Nach dem Transformationsschritt stellen sich die Basisdateien geordnet nach dem Generationszeitraum (siehe 6.1.7) wie folgt dar:

Generations- zeitraum	Taufen	Heiraten	Beerdi- gungen	Gesamt
1650 BIS 1679	3229	409	1746	5384
1680 BIS 1709	8182	2453	5847	16482
1710 BIS 1739	13444	4140	10809	28393
1740 BIS 1769	12004	4772	13957	30733
1770 BIS 1799	14183	6172	17834	38189
Gesamt	51042	17946	50193	119181

Tab. 5: Absolute Ereignishäufigkeiten in den Basisdateien

6. Merkmalsdokumentation

In den Basisdateien und zwischen den Basisdateien haben einige Merkmale identische Bedeutung, so daß auch dieselbe Werteetikettierung Verwendung findet. Einen Überblick gibt Tabelle 6.

Merkmalsgruppe	Taufen	Heiraten	Beerdigungen
LAUFENDE NUMMER	VAR1	VAR1	VAR1
KIRCHSPIEL	VAR2	VAR2	VAR2
JAHR UND MONAT	VAR3	VAR 3	VAR 3
JAHR	VAR4	VAR4	VAR4
MONAT	VAR5	VAR5	VAR5
JAHRZEHNT	VAR6	VAR6	VAR 6
GENERATIONSZEITRAUM	VAR7	VAR7	VAR7
Merkmale zum Beruf	VAR 11,20,27,34, 41,48	VAR9,16,23	VAR9
Merkmale zum Volkswirtschaftszweig	VAR 10,19,26,33, 40,47	VAR8,15,22	VAR8
Merkmale zur sozialen Gruppe	VAR12,21,28,35 42,49	VAR 10,17,24	VAR10
Merkmale zu Klasse - Schicht	VAR 13,22,29,36, 43,50	VAR11,18,25	VAR 11
Merkmale zum Familienstand	VAR 8	VAR21	---
Merkmale zum Geschlecht	VAR9,18,25,32, 39,46	---	VAR 13
Merkmale zur Herkunftsregion	---	VAR 13,28	---
Merkmale zur Ortsgröße	---	VAR 14,29	---

Tab. 6: Zuordnung der Merkmale zu den Basisdateien

In den folgenden Ausführungen werden die Merkmale nach den oben angegebenen Gruppen mit ihren Wertetiketten aufgelistet. Zur besseren Anschaulichkeit sind an einigen Stellen gleichzeitig Häufigkeitsauszählungen enthalten.

6.1 Allgemeine Merkmale

6.1.1 LAUFENDE NUMMER

Die Beobachtungen werden in der Reihenfolge ihres Auftretens durchnummeriert. Diese Zahl dient als Hilfsmittel zur Identifizierung der einzelnen Personen. Auf die Erfassung der Personennamen wurde aus Gründen des zu hohen Arbeitsaufwandes verzichtet, zumal die Namen nur für eine Familienrekonstitution bedeutsam gewesen wären, die aber nicht beabsichtigt war.

6.1.2 KIRCHSPIEL

Das Kirchspiel besitzt die Merkmalsausprägungen »1« und »2« mit folgender Bedeutung:

KIRCHSPIEL	Taufen	Heiraten	Beerdigungen
1= NIKOLAI	22727	7699	20664
2= GEORGEN	28315	10247	29529

Tab. 7: Wertetiketten und absolute Häufigkeiten für das Merkmal KIRCHSPIEL

6.1.3 JAHR UND MONAT

Dieses Merkmal wird aus den Merkmalen MONAT (VAR5) und JAHR (VAR4) gebildet. Aus kalendertechnischen Gründen wird der Monat jeweils auf den 15. Tag eingestellt. Der Vorteil ist, daß rechnerintern jedes beliebige Datum als Zahl der Tage zwischen diesem und einem speziellen Systemdatum gebildet werden kann, was für mögliche Sortierungen Bedeutung hat. In den Ausdrucken erscheint jedes Datum in einer üblichen und leicht verständlichen Form, z.B. 15APR1650.

6.1.4 JAHR

Das Merkmal JAHR (VAR4) ergibt sich aus in den Kirchenbüchern vorfindbaren Merkmalen Jahrhundert, Jahrzehnt und Jahr.

6.1.5 MONAT

Das Merkmal MONAT ist völlig identisch mit den Urdaten. Tabelle 8 stellt den Werteetiketten die entsprechenden absoluten Häufigkeiten in den Basisdateien gegenüber.

Monat	Taufen	Heiraten	Beerdigungen
1 = JANUAR	4613	1405	4183
2 = FEBRUAR	4193	1105	3912
3 = MÄRZ	4590	761	4301
4 = APRIL	4260	2050	4492
5 = MAI	4150	1532	4401
6 = JUNI	3871	1308	4037
7 = JULI	3935	1624	4159
8 = AUGUST	4287	1228	4653
9 = SEPTEMBER	4466	1429	4231
10 = OKTOBER	4441	2690	4174
11 = NOVEMBER	4142	1797	3758
12 = DEZEMBER	4094	1017	3892

Tab. 8: Werteetiketten und absolute Häufigkeiten für das Merkmal MONAT

6.1.6 JAHRZEHNT

JAHRZEHNT wird aus JAHR (VAR4) gebildet und besitzt folgende Werteetiketten:

1 = 1650 BIS 1659	2 = 1660 BIS 1669
3 = 1670 BIS 1679	4 = 1680 BIS 1689
5 = 1690 BIS 1699	6 = 1700 BIS 1709
7 = 1710 BIS 1719	8 = 1720 BIS 1729
9 = 1730 BIS 1739	10 = 1740 BIS 1749
11 = 1750 BIS 1759	12 = 1760 BIS 1769
13 = 1770 BIS 1779	14 = 1780 BIS 1789
15 = 1790 BIS 1799	

Tab. 9: Werteetiketten für das Merkmal JAHRZEHNT

6.1.7 GENERATIONSZEITRAUM

Dieses Merkmal wird aus JAHR (VAR4) gebildet und besitzt folgende Wertetiketten:

- 1 = 1650 BIS 1679
- 2 = 1680 BIS 1709
- 3 = 1710 BIS 1739
- 4 = 1740 BIS 1769
- 5 = 1770 BIS 1799

Tab. 10: Wertetiketten für das Merkmal GENERATIONSZEITRAUM

6.1.8 Merkmale zum Beruf

Die Codes für die Berufe sind überwiegend direkt aus den Benennungen der Kirchenbücher abgeleitet. Zusätzlich sind in der Literatur publizierte Berufslisten herangezogen worden (Deutsch-Ungarische-Lateinisches Handwerksnamenverzeichnis in: A magyarországi cehes Kezmueresipar forrasanyaganak Katasztere / Quellenmaterialkataster des zünftigen Handwerks in Ungarn, Bd. 1, Budapest 1975, S. 233-248 /). Selten auftretende Tätigkeitsbezeichnungen wurden unter Oberbegriffen subsummiert (z.B. Zollbediente). Bei Doppelbegriffen wurde nur der erste Begriff codiert. Die nachfolgende Tabelle 11 erstreckt sich über mehrere Seiten und enthält alle Wertetiketten für die Berufsangaben. Die erste Zahl des Berufscodes stellt gleichzeitig die Angabe für den Volkswirtschaftszweig dar.

001 = BOMBARDIER	002=CANONIER
003 = CORPORAL	004 = DRAGONER
005 = FAHNENJUNKER	006= FELDWEBEL
007= FELDPREDIGER	008= FEUERWERKER
010=FOURIER	011 = GARDIST
012= GEFREITER	013= GESCHIRRMESTER
014= HAUPTMANN	015 = HOBOIST
016= LEUTNANT	017=MUSKETIER
019=OBRIST	020= OFFIZIER
021 = PAUKER PFEIFFER BEIM HEER	022 = SERGEANT
023 = SOLDAT	024= ABGEDANKTER SOLDAT
025 = TAMBOUR	026= TRABANT
027 = TROMPETER BEIM HEER	028= UNTEROFFIZIER
029= QUARTIERMEISTER	030= AUDITOR
031= STADTHAUPTMANN WACHTMSTR.	032= FELDSCHER
033= RITTMEISTER	034= GENERAL
035 = PLATZMAJOR	036= PULVERMACHER
037=RUESTMEISTER	101 = ACCISEBEAMTE

102 = ACCISEVISITATOR
104 = ADVOKAT
106 = AUFSEHER
108 = BUERGERMEISTER
110 = CREISSBEDIENTE
112 = INSEIGNEUR
114 = FISKAL
116 = KAMMERFOURIER
118 = GERICHTSDIENER
120 = FOERSTER HOLZVERWALTER

122 = HOLZKNECHTE
124 = KAMMERDIENER
126 = KANZLEIDIENER
128 = KOMMISSARIUS
130 = KRIMINALRAT
132 = LAKAI
134 = LOHNLAKAI
136 = MAGISTRATSMITGLIED
138 = MUENZARBEITER
140 = NOTARE
142 = PRAESIDENT
144 = PRAELAT
146 = RATSDIENER
149 = REITKNECHT
151 = RENTMEISTER RENDANT
154 = SCHREIBER
156 = STADTRICHTER
158 = SUPERINTENDENTEN
160 = TAXATORES
162 = TORWAERTER
164 = VOGT
166 = WAISENHAUSBEDIENTE
168 = VERWALTER NICHTAGRARISCH
170 = STALLMEISTER
172 = TRAKTEUR
174 = BILLETEUR
176 = KORREKTOR
179 = BRANDMEISTER
182 = ASSESSOR
184 = KNECHT

186 = KAEMMERER
189 = PROVISOER
191 = KASSIERER
202 = BAECKER
204 = BRANNTWEINBRENNER
207 = BUTTERHAENDLER
210 = FISCHER
213 = FLEISCHER
215 = GRUETZMACHER
217 = GUTSHERR

103 = ACTUARIUS
105 = AMTMANN
107 = BUCHHALTER
109 = CONTROLLEUR
111 = DIENSTKNECHT
113 = DIREKTOR
115 = FORSTMEISTER
117 = GEHEIMRAT
119 = HEIDEBEDIENTE
121 = HOFBEDIENTE ADELIG
123 = KAEMMEREI KNECHTE
125 = KAMMERGERICHTSRAT
127 = KIRCHENKNECHT
129 = KRIEGSRAT
131 = KUTSCHER
133 = LANDREITER
135 = MAGAZINBEDIENTE
137 = MARKTMEISTER
139 = MUENZBEDIENTE
141 = POLIZEIBEDIENTE
143 = PROCURATOR
145 = KOENIGLICHER RAT
148 = REGISTRATOR
150 = RENTSCHREIBER
153 = SCHARFRICHTER
155 = SEKRETAER
157 = STEUEREINNEHMER
159 = SYNDIKUS
161 = TEICHINSPEKTOR
163 = TOTENGRAEBER
165 = WAAGEBEDIENTE
167 = ZOLLBEDIENTE
169 = AMTSBEDIENTE
171 = HAUSVERWALTER
173 = REITER VORREITER
175 = SCHLOSSKNECHT
178 = HOFBRAT
181 = RICHTER
183 = STADTSCHREIBER
185 = HOCHADEL KGL FAMILIE
188 = FAKTOR
190 = VISITATOR
201 = ACKERKNECHT
203 = ACKERMANN
205 = BRAUER
208 = CONDITOR
212 = FISCHHAENDLER
214 = GAERTNER
216 = GUETERVERWALTER
219 = HIRTEN

- 220 = HOEKER
 222 = ITALIENER
 225 = KAESEHAENDLER
 227 = KORKSTOEPSELMACHER
 229 = KORNMESSER
 231 = KOSSAT
 233 = MEHLHAENDLER
 235 = MUEHLENBESCHIEDER
 237 = MUELLER
 239 = PFERDDEBEREITER
 241 = SALZWIRKER
 245 = SCHWEINESCHNEIDER

 248 = TABAKSPINNER

 250 = VIEHHAENDLER
 252 = WEINHAENDLER
 254 = WILDNER
 256 = ARENDATOR
 259 = VICTUALIENHAENDLER
 261 = HAEUSLER
 263 = HEGEMEISTER
 266 = FELDHUETER
 301 = BECKENSCHLAEGER
 303 = BLEISTIFTMACHER
 305 = BOHRENSCHMIED
 307 = DRAHTZIEHER
 309 = FEILENHAUER
 311 = GELBGIESSER
 313 = GOLD SILBERDRAHTZIEHER
 315 = GOLDSCHLAEGER
 317 = HAMMERSCHMIED
 319 = INGENIEUR
 321 = KLEMPNER
 323 = LEDERTAUER
 326 = MESSERSCHMIED
 328 = NAGELSCHMIEDE
 330 = ROTGIESSER
 332 = SCHERENSCHLEIFER
 334 = SCHLOSSER
 336 = SCHRIFTGIESSER
 338 = SCHWERTFEGER
 340 = SENSENTRAEGER
 342 = SPLETTREISSER
 344 = STELLMACHER
 346 = ZINNGIESSER
 348 = KANNENGIESSER
 350 = ZIRKELSCHMIED
 352 = KUNSTGI ESSER
 354 = SCHIRRMACHER
 402 = BARCHENTMACHER

 221 = HOPFENMESSER
 223 = JAEGER
 226 = KOCH
 228 = KORNSHAENDLER
 230 = KORNSCHREIBER
 232 = MALZBEREITER
 234 = MEIER
 236 = MUEHLENSCHREIBER
 238 = PFEFFERKUECHLER
 240 = SALZBEDIENTE
 244 = SCHULZE
 246 = STAERKEPUDERMA-
 CHER
 249 = UNTERTAN VOM LAN-
 DE
 251 = VIEHMAESTER
 253 = WEINMEISTER
 255 = ZUCKERSIEDER
 258 = GARNMEISTER
 260 = SCHAEFER
 262 = WEINGAERTNER
 265 = GEWUERZHAENDLER
 267 = TABAKHAENDLER
 302 = BLATTMACHER
 304 = BLEIWEISSMACHER
 306 = BUECHSENMACHER
 308 = EISENHAENDLER
 310 = FORMENSCHNEIDER
 312 = GLOCKENGIESSER
 314 = GOLDSCHIEDDER
 316 = GOLDSCHMIEDDE
 318 = HUFWAFFENSCHMIED
 320 = KESSELFlickER
 322 = KUPFERSCHMIED
 324 = MECHANIKER
 327 = NADLER
 329 = METALLARBEITER
 331 = SATTLER
 333 = SCHIFFBAUER
 335 = SCHRIEBELMACHER
 337 = SCHROTGIESSER
 339 = SEILER
 341 = SIEBMACHER
 343 = SPORER
 345 = WINDENMACHER
 347 = EISENSCHMELZER
 349 = ZEUGSCHMIED
 351 = KOEHLER
 353 = KOHLENMESSER
 401 = BANDMACHER
 403 = «BAUMWOLLZEUG-
 MACHER

404 = BEUTELTUCHMACHER
 406 = DECKENMACHER
 408 = ETAMINMACHER
 410 = FLANELLWIRKER
 412 = RIETHMACHER
 414 = LAGERHAUSBEDIENTE
 416 = LEINENWEBER
 418 = LEINWANDBEREITER
 420 = LISEUR
 422 = ZEUG RASCHMACHER
 424 = SEGELMACHER

 426 = SEIDENWIRKER
 428 = SPINNER
 430 = TIREUR
 432 = TUCHHAENDLER

 434 = TUCHSCHERER

 436 = WALKMUELLER
 441 = CALAMANGWEBER
 443 = PRESSER
 446 = ZEUGDRUCKER
 448 = CRISETMACHER

 502 = WEISSGERBER BEUTLER
 504 = BRILLENMACHER
 506 = DIAMANTSCHLEIFER
 508 = GOLD U. SILBERSTICKER
 511 = HANDSCHUHMACHER
 514 = TROEDLER
 516 = KUERSCHNER
 518 = LOH ROTGERBER
 520 = MEDAILLONMACHER
 522 = PANTOFFELMACHER
 524 = PERUECKEMACHER
 526 = SCHNEIDER

 529 = SEIDENSTICKER
 601 = BAUMEISTER
 603 = BAUSCHREIBER
 605 = BOETTCHER
 607 = BRETTSCHEIDER
 609 = GIPSBRENNER STUKKATEUR
 611 = GLASMACHER
 613 = GLASSCHNEIDER
 615 = HOLZSETZER
 617 = KLEINBINDER
 619 = LACKIERER
 621 = LEHMGRABER LEHMENTIERER
 624 = MALER
 626 = MAURER

 405 = BLEICHER
 407 = DESIGNEUR
 409 = FAERBER
 411 = HECHELMACHER
 413 = KATTUNDRUCKER
 415 = DAMASTWEBER
 417 = LEINWANDDRUCKER
 419 = LEINWANDHAENDLER
 421 = MANCHESTERMACHER
 423 = SAMTMACHER
 425 = SEIDENSTRUMPFWE-
 BER
 427 = SEIDENWEBER
 429 = STRUMPFWIRKER
 431 = TUCHBEREITER
 433 = TUCHMACHER
 WOLLARBEITER
 435 = WACHSLEINWAND-
 MACHER
 439 = ZELTSCHNEIDER
 442 = KAMMACHER
 444 = WAIDMEISTER
 447 = FAHNENSCHNEIDER
 501 = BERNSTEINSCHNEI-
 DER
 503 = BLUMENFABRIKANT
 505 = CORDUANMACHER
 507 = FISCHBEINREISSER
 510 = GUERTLER
 512 = HUTMACHER
 515 = KNOPFMACHER
 517 = LEDERHAENDLER
 519 = LUMPENSAMMLER
 521 = LEISTENSCHNEIDER
 523 = POSAMENTIERER
 525 = SAFFIANMACHER
 528 = SCHUHMACHER
 FLICKER
 530 = STROHHUTFABRIKANT
 602 = BAUINSPEKTOR
 604 = BESENBINDER
 606 = BRUNNENMACHER
 608 = DRECHSLER
 610 = GLASER
 612 = GLASSCHLEIFER
 614 = HOLZHAENDLER
 616 = KALKBRENNER
 618 = KORBMACHER
 620 = LACKMACHER
 623 = LICHTZIEHER
 625 = MARMORARBEITER
 627 = PORZELLAN ARBEITER

629 = SAEGEMEISTER
 631 = SCHORNSTEINFEGER
 633 = STEINMETZ
 636 = TAPEZIERER
 638 = TOEPFER
 640 = VERGOLDER
 642 = WALLARBEITER
 643 = ZIEGLER
 645 = MULDENMACHER
 647 = PLUMPENMACHER
 649 = SCHNEIDDEMUELLER
 651 = ROEHRENMEISTER
 701 = BOTE
 703 = FUHRMANN
 705 = KARRENFUEHRER
 707 = LATERNENANSTECKER
 709 = POSTBEDIENTE
 712 = SCHIFFER SCHIFFSKNECHTE
 715 = STEUERMANN
 717 = STROMMEISTER
 719 = BRUECKENMEISTER
 721 = PORTECH AISENTRAEGER
 802 = APOTHEKER
 804 = BILDERHAENDLER
 806 = BUCHBINDER
 808 = BUCHHAENDLER
 810 = CANDIDAT
 812 = FEDERSCHNEIDER
 814 = HOSPITALBEDIENTE
 816 = KANTOR
 818 = KOMOEDIANT
 820 = KUPFERSTECHE
 822 = MUSIKINSTRUMENTENMACHER
 825 = PAPIERHAENDLER
 827 = PERGAMENTMACHER
 829 = RECHENMEISTER
 831 = SEIFENSIEDER
 834 = TANZMEISTER
 837 = KUNSTMALER
 839 = HAUSLEHRER
 842 = PROFESSOR
 844 = REKTOR
 846 = DOKTOREN
 848 = BIBLIOTHEKAR
 901 = ARME
 903 = BERGMANN
 905 = EIGENTUEMER
 907 = HOSPITALINSASSEN
 909 = KAUFLEUTE MANUFAKTUR
 911 = LABORANT
 913 = MAKLER
 915 = RENTIER
 630 = SCHIEFERDECKER
 632 = STEINHAENDLER
 635 = TAPETENMACHER
 637 = TISCHLER
 639 = UHRMACHER
 641 = WACHSPASSIERER
 634 = STEINSCHNEIDER
 644 = ZIMMERMANN
 646 = EMAILLEUR
 648 = SPIEGELSCHLEIFER
 650 = STAFFHAUER
 652 = DACHDECKER
 702 = FLOESSER
 704 = GASTWIRT
 706 = LAMPENPUTZER
 708 = PACKHOFSBEDIENTE
 711 = SACKTRAEGER
 714 = STEIN DAMMSETZER
 716 = STRASSENFEGER
 718 = TAGELOEHNER
 720 = WEGEMEISTER
 801 = ARZT
 803 = BARBIER CHIRURG
 805 = BILDHAUER
 807 = BUCHDRUCKER
 809 = CALEANT
 811 = FECHTMEISTER
 813 = HEBAMME
 815 = KAMMERJAEGER
 817 = KARTENMACHER
 819 = KUESTER
 821 = MUSIKANTEN ZIVIL
 823 = ORGANIST
 826 = PAPIERMACHER
 828 = PREDIGER
 830 = SCHULMEISTER
 833 = STUDENT
 835 = ZEICHNER
 838 = VITRIOLFABRIKANT
 840 = TROMPETER ZIVIL
 843 = SPRACHMEISTER
 845 = VICAR
 847 = GYMNASIAST
 849 = DIACON
 902 = BANKIER
 904 = COLONIST
 906 = EINWOHNER
 908 = HUNDEFUETTERER
 910 = KOMMISSIONAER
 912 = LANDPASSATEN
 914 = MATERIALIST
 916 = STADTVERORDNETER

917 = TABULETTKRAEMER
 919= KAUFMANN
 921 = KIRCHENVORSTEHER

918= ARRESTANTEN
 920= FABRIKANT
 922= KAUFMANNSDIENER

Tab. 11: Wertetiketten für die Merkmale zum Beruf

6.1.9 Merkmale zum Volkswirtschaftszweig

Die Angaben zum Volkswirtschaftszweig stammen unverändert aus den Urdaten. Der Volkswirtschaftszweig wurde mit der Codierung des Berufes festgelegt und ist in dessen ersten Ziffer verschlüsselt.

VOLKSWIRTSCHAFTSZWEIG	Beerdigungen
MISSING	8967
0=MILITAER	2579
1 = HOFHALTUNG VERWALTUNG GERICHT DIENSTLSTG	3967
2= LANDWIRTSCHAFT FISCHEREI NAHRUNGSMITTEL	7251
3= METALLGEWERBE WERKZEUG- HERSTELLUNG	2179
4 = TEXTILGEWERBE	6135
5= KLEIDUNG SCHMUCK MODE	5780
6= BAUGEWERBE MOEBEL WOHNUNG HAUSRAT	3743
7= TRANSPORT HERBERGE GASTHAUS TAGELOEHNER	5425
8= BILDUNG KUNST WISSENSCHFT KIRCHE GESHWES	1873
9= SONSTIGE	2294

Tab. 12: Basisdatei BEERDIGUNGEN - absolute Häufigkeiten für das Merkmal VOLKSWIRTSCHAFTSZWEIG

6.1.10 Merkmale zu den sozialen Gruppen

Die Einteilung der Beobachtungen in soziale Gruppen erfolgte entsprechend allgemeinen historischen Erkenntnissen über die Klassen- und Schichtengliederung und die Binnenstruktur von Klassen und Schichten in der spätfeudalen Stadt. Diese Zuordnung (Tab.13) bildete den problematischsten Teil der Codierung und war zugleich wesentlich für die Analyse, die überwiegend auf eben dieser Abstraktionsebene, nicht auf der Basis der Berufe erfolgte.

1= ADEL	2 = OFFIZIERE
3= HANDELS UND MANUF.BOURGEOISIE	4 = HOHE BUERGERLICHE BEAMTE
5= AKADEMIKER	6= KUENSTLER TECHNIKER
7= MITTLERE BEAMTE	8 = AGRARGEWERBE
9=LEHRER	10= SUBALTERNE BEAMTE
11= UNTEROFFIZIERE	12= KLEINHAENDLER
13= GASTWIRTE UND TRANSPORTWESEN	14= SOLDATEN
15= HANDWERKSMEISTER	16 = MANUFAKTURARBEITER
17= DIENSTBOTEN	18 = TAGELOEHNER
19= ARME	20 = GESELLEN
A = BUERGER UND EIGENTUEMER	B = EMIGRANTEN
C"=EINWOHNER	P = ANGABE ZU UNSICHER

Tab. 13: Wertetiketten für die Merkmale zur sozialen Gruppe

Die Zuordnung der Probanden erfolgte vorwiegend auf der Grundlage der Berufsbezeichnung, daneben anhand von Standesbezeichnungen in den Kirchenbüchern («Meister«, »Geselle«, »Kaufmann«, »Bürger« und »Eigentümer«), seltener aufgrund präziser Angaben der Kirchenbücher ('Arbeiter in der 3. kgl. Zuckersiederei/Tuchmacher vom Lagerhaus'), nur in Ausnahmefällen nach dem sozialen Umfeld (Kindsvater ohne Berufsangabe, alle Paten Lohnarbeiter) oder aufgrund von Kenntnissen über die historische Entwicklung Berlins ('Zeugweber' ab 1740, 'Seidenwirker' ab 1760 allgemein als Manufakturarbeiter). In der sozialen Gruppe MANUFAKTURARBEITER sind unterschiedlich abhängige Gruppen dezentralisiert arbeitender Produzenten einschließlich erkennbar verlegter Arbeiter enthalten. Die als 'Fabrikant' bezeichneten Produzenten wurden wegen der großen Bandbreite und Unsicherheit zu den selbständigen Produzenten geschlagen und sind hier gemeinsam mit den Zunftmeistern als HANDWERKSMEISTER etikettiert.

Die Codes 'P A B C' werden bei jenen Beobachtungseinheiten verwendet, für die keine eindeutige Zuordnung zu einer sozialen Gruppe möglich ist.

6.1.11 Merkmale zu Klasse - Schicht

Die Merkmalsausprägungen leiten sich aus den sozialen Gruppen ab und sind in Tabelle 14 zusammengefaßt.

- 1= FEUDALKLASSE
- 2 = HANDELS UND MANUFAKTURBOURGEOISIE

- 3 = WERKTAETIGE INTELLIGENZ UND BEAMTE
- 4 = GEWERBETREIBENDE
- 5 = LOHNARBEITER
- 6 = SOLDATEN
- P = ANGABE ZU UNSICHER

Tab. 14: Merkmalsetiketten für Merkmale zur Klasse-Schicht

Der Code P ist nur für die Basisdatei Beerdigungen wirksam und wurde bei Fehlstellenprozenten größer 75% im betrachteten Monat gesetzt.

Wie die sozialen Gruppen zu Klassen und Schichten zusammengefaßt wurden, ist Tabelle 15 zu entnehmen.

Soziale Gruppe	Klasse-Schicht
1,2,4	1
3	2
5,6,7,9,10	3
8,12,13,15	4
16,17,18,19,20	5
11,14	6
A, B, C	MISSING

Tab. 15: Bildungsvorschrift für die Merkmale zur Klasse-Schicht aus den Merkmalen zur sozialen Gruppe

Klasse - Schicht	Patel	Pate2	Pate3	Pate4	Pate5
MISSING	6895	10858	18287	28623	36505
FEUDALKLASSE	3780	2195	1704	1365	911
HANDELS U. MANU- FAKTURBOURGEOISIE	2755	2112	1389	820	534
WERKTAETIGE INTELLI- GENZ U. BEAMTE	7446	5429	4149	2717	1593
GEWERBETREIBENDE	21206	20504	16360	11048	7308
LOHNARBEITER	7430	8470	7833	5457	3548
SOLDATEN	1530	1474	1320	1012	643

Tab. 16: Basisdatei TAUFEN - absolute Häufigkeiten der Merkmale zur Klasse - Schicht der Paten

Klasse - Schicht	Bräutigam	Bräutigamvater	Brautvater
MISSING	392	12228	1088
FEUDALKLASSE	331	53	497

HANDELS U. MANU- FAKTURBOURGEOISIE	493	56	378
WERKTAETIGE INTELLI GENZ U.BEAMTE	1910	355	1929
GEWERBETREIBENDE	8592	3822	9673
LOHNARBEITER	5461	1204	3146
SOLDATEN	767	228	1235

Tab. 17: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale zur Klasse-Schicht

6.1.12 Merkmale zum Familienstand

Der Familienstand ist im allgemeinen nicht in die Datei aufgenommen worden. Ausnahmen sind die durchgängig klar erkennbaren Tatsachen der Wiederverheiratung einer Witwe und einer unehelichen Geburt.

In der Basisdatei HEIRATEN ist der Familienstand mit den Merkmalsrealisierungen:

- 1 = ERSTHEIRAT
- 2 = WIEDERHEIRAT

gespeichert.

In der Basisdatei TAUFEIN tritt dieses Merkmal als Familienstand der Mutter mit den Werten:

- 1 = EHELICH
- 2 = UNEHELICH

auf.

In den Dateien haben diese Merkmale keine Fehlstellen. Die zugehörigen Häufigkeitsverteilungen enthalten die Tabellen 18 und 19.

FAMILIENSTAND	N	N(KUM)	%
EHELICH	47618	47618	93.29
UNEHELICH	3424	51042	6.71

Tab. 18: Basisdatei TAUFEIN - Häufigkeitsverteilung des Merkmals FAMILIENSTAND

FAMILIENSTAND DER BRAUT	N	N(KUM)	%
ERSTHEIRAT	14654	14654	81.66
WIEDERHEIRAT	3292	17946	18.34

Tab. 19: Basisdatei HEIRATEN - Häufigkeitsverteilung des Merkmals FAMILIENSTAND DER BRAUT

6.1.13 Merkmale zum Geschlecht

Für die Angabe des Geschlechts treten nur die Codes

1 = MAENNLICH

2 = WEIBLICH

auf. Tabelle 20 zeigt die Verteilung der Todesfälle nach dem Geschlecht.

GESCHLECHT	N	N(KUM)	%
MISSING	691	.	.
MAENNLICH	25393	25393	51.30
WEIBLICH	24109	49502	48.70

Tab. 20: Basisdatei BEERDIGUNGEN - Häufigkeitsverteilung des Merkmals GESCHLECHT

6.1.14 Merkmale zur Herkunftsregion

Herkunftsregion	Bräutigam	Braut
0 = BERLIN	12978	11081
1 = MARK BRANDENBURG	2636	4364
2 = OESTLICHE PROVINZEN (SCHLESIEN NACH 1740)	362	362
3 = ANDER WESTL MITTL PROVINZEN BRDBG PREUSS	246	275
4 = MECKLENBURG VORPOMMERN UND HOLSTEIN	94	128
5 - SACHSEN THUERINGEN UND ANHALT	583	422
6 = WESTELBISCHE TERRITORIEN (INCL BAYERN)	285	222
7 = HABSBUURGISCHE LAENDER (SCHLES BIS 1740)	74	61
8 = ANDERE WESTEUROPAEISCHE STAATEN	33	30
9 = ANDERE OSTEUROPAEISCHE STAATEN	40	8
MISSING	615	993

Tab. 21: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale REGION BRAUT und REGION BRAEUTIGAM

Für die Zuordnung der Herkunftsorte zu Herkunftsregionen waren die jeweiligen staatlichen Grenzen maßgebend. In den meisten Fällen bezogen sich die Angaben in den Kirchenbüchern nicht direkt auf die Eheleute, sondern bezeichneten den Wohnort des Vaters. Insgesamt ergibt sich das in Tabelle 21 dargestellte Bild.

6.1.15 Merkmale zur Ortsgröße

In der Basisdatei HEIRATEN wurde der Herkunftsort nach Ortsgrößen klassifiziert. Die Klassifizierung der Städte erfolgte anhand der Angaben der Einwohnerzahl im Deutschen Städtebuch.

Ortsgröße	Bräutigam	Braut
1 = GROSSSTADT AB 10000 EINWOHNER	13409	1026
3 = MITTELSTADT V. 3000 B. 10000 EINWOHNER	723	971
5 = KLEINSTADT BIS 3000 EINWOHNER	1375	2172
7 = DORF	1801	2189
MISSING	648	1026

Tab. 22: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale ORTSGROESSE BRAUT und ORTSGROESSE BRAEUTIGAM

6.2 Spezielle Merkmale der Basisdatei B E E R D I G U N G E N

6.2.1 Merkmale zum Sterbealter

Aus dem Merkmal ALTER (VAR14), das in den Werten 1-106 vertreten ist, wurden das Merkmal ALTERSGRUPPE (VAR15) folgendermaßen abgeleitet:

1 - UNTER 1 JAHR	2 = 1 BIS 4 JAHRE
3 = 5 BIS 14 JAHRE	4 = 15 BIS 34 JAHRE
5 = 35 BIS 49 JAHRE	6 = 50 BIS 69 JAHRE
7 = 70 BIS 84 JAHRE	8 = UEBER 85 JAHRE
9 = KINDER OHNE ALTERSANGABE	MISSING

Tab. 23: Wertetiketten für das Merkmal ALTERSGRUPPE

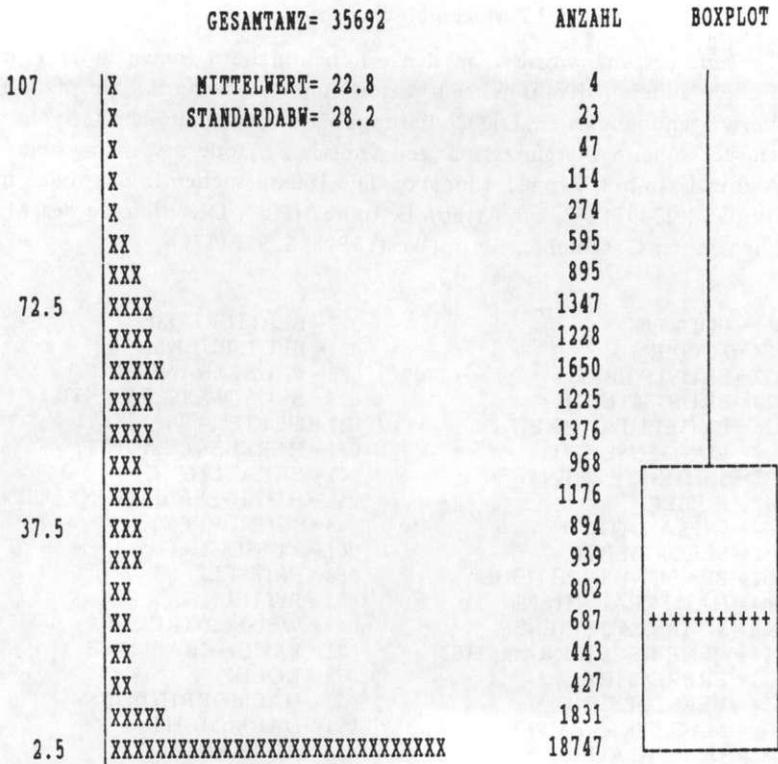


Abb. 7 : Basisdatei BEERDIGUNGEN - Häufigkeitsdiagramm mit Boxplot des Merkmals ALTER

In den Erfassungsbelegen wurden Todesfälle von Kindern ohne Altersangaben separat erfaßt und später in das Merkmal ALTERSGRUPPE als Werteausprägung 9 integriert. In der Annahme, daß auf exakte Altersangaben eher bei ganz jungen Kindern verzichtet wurde, sind sie der Säuglingssterblichkeit, d.h. der Altersgruppe 1 zugerechnet worden.

Das Häufigkeitsdiagramm in Abbildung 7 zeigt die Verzerrung der Altersverteilung durch die hohe Säuglings- bzw. Kindersterblichkeit.

6.2.2 Merkmale zur Todesursache

Die Todesursachenangaben in den Kirchenbüchern folgen einer grundsätzlich anderen Krankheitsauffassung und -Systematik. Sie stammen überwiegend von Laien. Die Codierung folgte sehr genau den Bezeichnungen der Quellen, ergänzt um zeitgenössische Todesursachensystematik (Arthur E. Imhof/Bengt I. Lindskog, Die Todesursachen in Schweden und Finnland 1749-1773, in: Arthur E. Imhof (Hg.), Die Biologie des Menschen in der Geschichte, Berlin(West) 1978, S. 97-177).

011 = POCKEN	012 = BLATTERROSE
013 = MASERN	021 = BLUTSPUCKEN
022 = BRUSTFIEBER	023 = BRUSTKRANKHEIT
024 = BRUSTSCHADEN	025 = BRUSTWASSERSUCHT
026 = ENGBRUESTIGKEIT	027= HUSTEN
028 = LUNGENSUCHT	031 = HERZBEKLEMUNG
032 = SEITENSTECHEN	041 = ERKAELTUNG
042 = FIEBER	043 = HITZIGES FIEBER INFLUENZA
044 = INFLAMATION	045 = SCHNUPFEN
051 = FLECKFIEBER	061 = AUSSCHLAG
062 = BRAEUNE DIPHTERIE	063 = FRIESEL
064 = HAUTKRANKHEIT	065 = RITTELN
066 = SCHARLACHFIEBER	067 = WEISSER TRIESEL
068 = VENERISCHE KRANKHEIT	071 = BAUCHKRAEMPFE
072 = ERBRECHEN	073 = KOLIK
074 = VERSTOPFUNG	075 = HAEMORRHOIDEN
076= MAGENKRAEMPFE	081 = DARMGICHT
082 = DURCHLAUF	083 = ROTE RUHR
091 = GALLENFIEBER	092 = GELBSUCHT
093 = LEBERKRANKHEIT	101 = AUSZEHRUNG
102= SCHWINDSUCHT	111 = MALARIA WECHSELFIEBER
121 = BOESARTIGES FIEBER	122 = FAULFIEBER
123 = FLUSSFIEBER	124 = KALTES FIEBER
125 = NERVENFIEBER	126= NESSELFIEBER
127 = HALSSTARRE	128= VERZEHRENDES FIEBER
131 = GICHTWASSERSUCHT	132 = LEBERVERHAERTUNG
133 = WASSERSUCHT	141 = ALTERSSCHWAECHT
142= MISERERE	151 = HARTE GEBURT
152 = KINDBETT	161 = ENGLISCHE KRANKHEIT
162 = SCHWERE GEBRECHEN	163 = JAMMER KRAEMPFE
SCHWACHHEIT	
164 = NABELBRUCH	165 = SCHWAEMME
166 = WUERMER	167 = VERSTILLUNG
168 = ZAEHNE	171 = KEUCHHUSTEN STICKHUSTEN
181 = PEST	191 = STEINLEIDEN
192 = URINVERSTOPFUNG	201 = BLUTSTURZ
202 = NASENBLUTEN	211 = PODAGRA
212 = ROSE	213 = SKORBUT
221 = BRAND INNEN UND AUSSEN	222= FISTEL
223 = BEINFRASS FUSSSCHAEDEN	224 = GESCHWUER NABEL

225 = BRUSTGESCHWUER	GESCHWUER
227 = GEWAECHS AM KOPF	226= GESCHWULST KREBSSCHADEN
229 = KOFWASSERSUCHT	228= KNOCHENBRAND
232 = GLIEDERKRANKHEIT	231 = GICHT
	233 = KNOCHENKRANKHEIT
	KNOCHENSCHADEN
234 = RHEUMATISMUS	241 = ZERSPRUNGENE ARTERIE
242 = OHNMACHT	243 = SCHLAGFLUSS
244 = STICKFLUSS	245 = PLOETZLICHER TOD
251 = KINDESMORD	252 = ERSTICKTE KINDER
261 = MORD	262 = VERBRECHEN
263 = VERGIFTUNG	264= ERDROSSELT
265 = HALS ABGESCHNITTEN	266= VON EINEM HIEB
271 = HUNGER	281 = ERTRUNKEN
282 = AUF DEM EIS	291 = ERFROREN
EINGEBROCHEN	ERFRORENE FUESSE
301 = KNOCHENBRUECHE	302= STURZ
303 = BRUSTQUETSCHUNG	304= UNFALL VERUNGLUECKT
U EBERFAHREN	
305 = VERBRUEHT	306= STOSS
307 = ARBEITSUNFAELLE	311 = VERBRANNT
312 = RAUCHVERGIFTUNG	321 = HEIMSUCHT
322 = AN AERGERNIS	323 = MELANCHOLIE
324 = ERHAENGT	325= SELBSTMORD
331 = BLESSUREN	332 = KOPFWUNDEN
333 = ERSCHOSSEN	341 = HINGERICHTET
351 = EPILEPSIE FALLSUCHT	352= NERVENKRANKHEIT
353 = SCHLAFSUCHT	354= WAHNSINN
355 = KOPFSCHMERZEN	361 = TOTGEBURT
371 = BRUCHSCHADEN	372 = HALSSCHADEN
373 = INNERLICHE SCHAEDEN	374= SCHWARZE SUCHT
P = 'ANGABE ZU UNSICHER'	

Tab. 24: Werteetiketten für das Merkmal TODESURSACHE

Ein Gruppierungsversuch mittels Clusteranalyse nach dem jahreszeitlichen Auftreten der Krankheiten lieferte lediglich Anhaltspunkte zu der Frage, welche Krankheiten nicht zusammengefaßt werden dürfen. Aufgrund der sehr unterschiedlichen Besetzungen brachte sie gerade für die vielen seltenen Todesarten keine plausible Zuordnung. H. Schultz gruppierte die Todesursachen deshalb aufgrund der Häufigkeitsverteilung nach Alter, Sterbemonat und den groben Anhaltspunkten aus der Bezeichnung, die etwa die Körperregion oder Symptome (wie Fieber) betrafen.

TODESURSACHE	TODESURSACHENGRUPPEN	Anzahl
163,164	1 = JAMMER	4176
101,142,162,167,224,271	2 = AUSZEHRUNG	3688
23	3 = BRUSTKRANKHEIT	3073
11,12	4 = POCKEN	2620
241,242,243,245	5 = SCHLAGFLUSS	2389

165,168,213	6 = ZAEHNE	2357
361	7 = TOTGEBURT	1842
13.62.63,65.66.121,126,128,		
161,212,372	8 = RITTELN MASERN UA	1276
43	9 = HITZIGES FIEBER	1145
226,227	10 = KREBSSCHADEN	957
141	11 = ALTERSSCHWAECHEN	950
21,28,102,201,202	12 = SCHWINDSUCHT	908
244	13 = STICKFLUSS	805
81,82,83	14 = RUHR DURCHLAUF	781
25,133,229	15 = WASSERSUCHT	533
27,171	16 = KEUCHHUSTEN	528
251,252,261,262,266,		
281,291,		
301-307,311,312,		
322,324,325. 331-333,341	17 = GEWALTSAMER TOD	264
22,41,42,44,45,		
51,111,122-124, 181	18 = SONSTIGE FIEBER	437
71-76	19 = BAUCH U MAGEN KRAEMPFE	427
91-93,191,231,232,234	20 = GICHT STEINLEIDEN	384
24,221-225,233,371,373	21 = SCHAEDEN	327
151,152	22 = KINDBETT	
26,31,32	23 = ENGBRUESTIGKEIT	271
125,323,351-355	24 = NERVENKRANKHEITEN	214
alle anderen Codes außer		
MISSING und 'F	25 = SONSTIGE	13685
	P = ANGABE ZU UNSICHER	5959

Tab. 25: Basisdatei BEERDIGUNGEN - Bildungsvorschrift des Merkmals
TODESURSACHENGRUPPEN und absolute Häufigkeiten

6.3 Spezielle Merkmale der Basisdatei TAU FEN

6.3.1 Merkmale zur Anzahl von Paten

Die drei Merkmale (VAR15,VAR16,VAR17) zur Anzahl der Paten sind keine Rückrechnungen aus den Angaben unter Patel bis Patc5, sondern unverändert den Kirchenbüchern als Zusatzinformation entnommen. Die angegebene Anzahl für Paten insgesamt variiert zwischen 0 und 28. Die dazugehörige Häufigkeitsverteilung ist in Tabelle 24 dargestellt:

Anzahl Paten	insgesamt	Anzahl Paten	insgesamt
0	260	15	243
1	47	16	91
2	357	17	115

3.	2985	18	46
4	2737	19	32
5	15375	20	10
6	6796	21	18
7	7599	22	10
8	4330	23	4
9	4233	24	2
10	2090	25	6
11	1921	26	3
12	880	27	-
13	615	28	1
14	263		

Tab. 26: Basisdatei TAUFEN - Häufigkeitsverteilung des Merkmals
PATEN INSGESAMT

Analog wurde im Merkmal PATEN MAENNLICH die Anzahl der männlichen Paten und unter PATEN WEIBLICH die Anzahl weiblicher Paten als Teil der Paten insgesamt erfaßt. Da im Erfassungsbeleg für die Zahl der männlichen und weiblichen Paten jeweils nur eine Spalte vorgesehen war, bedeutet '9', wenn die Summe kleiner als die Gesamtzahl Paten ist, auch 'viele'.

7. Datenanalyse

Die Datenanalyse stellt das eigentliche Ziel der vorangegangenen Datenverarbeitungsschritte dar und beeinflußt somit auch maßgeblich die Struktur der erfassten und gespeicherten Datensätze. In dieser Ausarbeitung wird nicht ausführlich auf die angewandten Analysealgorithmen eingegangen werden, jedoch wird versucht, auf Analysestrategien hinzuweisen und diese anhand von Modellen transparent zu machen.

In den Abbildungen werden diejenigen Merkmale, die durch Primärinformationen belegbar und damit meßbar sind, in eckigen Umrandungen aufgeführt. In runden Umrandungen werden die empirisch bzw. theoretisch erschlossenen Merkmale angegeben. Diese Merkmale können im Sinne der Analyse latenter Strukturen als latente Merkmale bezeichnet werden. Die Relationen zwischen meßbaren Merkmalen und latenten Merkmalen bilden das sogenannte Meßmodell. Untersuchungen dieser Relationen dienen dazu, die Güte der Erklärung der latenten Merkmale durch die tatsächlich vorliegenden Daten zu analysieren. Dies ist deshalb fundamental, da die latenten Merkmale als Elemente der Theoriebildung in der Anwendungsdisziplin, hier in der Geschichtswissenschaft, verwendet werden. Die Beziehungen zwischen den latenten Merkmalen sind Diskus-

sionsgegenstand der Theoriebildung. Statistische Analysen der vorliegenden Daten bilden daher den explorativen Ausgangspunkt, um eine Theorie über die zu untersuchende Gesamtheit abzuleiten.

7.1 Analyse der Sekundärdatei JAHRE

In den drei Basisdateien, die aus den Tauf-, Heirats- und Beerdigungsregistern aufgebaut wurden, entspricht eine Datenzeile jeweils einer Beobachtung dieser Ereignisarten an einer Person. Für jedes Ereignis ist die Zeit als JAHR UND MONAT (VAR3), JAHR (VAR4), MONAT (VAR5), JAHRZEHNT (VAR6) und als GENERATIONSZEITRAUM (VAR7) erfaßt. In der Sekundärdatei JAHRE entspricht jede Beobachtung und damit jede Datenzeile einem Jahr zwischen 1650 und 1799. Damit sind die Daten als historische Daten, die die Dimension Zeit stets beinhalten, charakterisiert. Bei der Analyse dieser Daten blieb nur für globale Aussagen über den gesamten Untersuchungszeitraum die Zeit unberücksichtigt. Das Merkmal JAHR (VAR4) bildete die Grundlage für Aussagen zur Bevölkerungsentwicklung. Zur Untersuchung der Entwicklung der Sozialstruktur und deren Auswirkungen ging die Zeit als GENERATIONSZEITRAUM (VAR7) in die Analyse der Daten ein. Das Merkmal MONAT (VAR3) diente der Analyse saisonaler Phänomene.

Die Bevölkerungsentwicklung wurde getrennt nach Generationszeiträumen auf jährlicher Basis analysiert. Während zur Analyse der natürlichen Bevölkerungsentwicklung die Tauf- und Sterberegister eine Schätzgrundlage lieferten, waren zur Schätzung der Wanderungsgewinne neben den Aussagen des Heiratsregisters weitere Informationen aus anderen Quellen erforderlich. Der folgende Ansatz wurde der Rekonstruktion der natürlichen Bevölkerungsentwicklung zugrundegelegt.

- Bestimmung der Sterblichkeit pro Generationszeitraum aus der mittleren Lebenserwartung pro Generationszeitraum. Dabei konnten die mittleren Lebenserwartungen nur für die letzten drei Generationszeiträume aus den Daten berechnet werden, für die ersten beiden Generationszeiträume lagen dazu zuwenig gültige Altersangaben vor, so daß man für diese Zeiträume die Sterberate nur aus anderen Quellen schätzen konnte.
- Für jeden Generationszeitraum wurde jährlich der lineare Trend der korrigierten Sterbeziffern ermittelt. Gegebenenfalls könnte hier auch ein quadratischer Trendansatz genutzt werden.
- Schätzung der jährlichen Einwohnerzahl:

$$\text{Einwohnerzahl (jährlich)} = \frac{\text{Trend(jährlich)} * 1000}{\text{Sterberate(pro Generationszeitraum)}}$$

- Schätzung der jährlichen Sterbe-, Heirats- und Geburtenrate:

$$\text{Heiratsrate (jährlich)} = \frac{\text{Heiratsziffer(jährlich)} * 1000}{\text{Einwohnerzahl(jährlich)}}$$

$$\text{Geburtenrate (jährlich)} = \frac{\text{Taufziffer(jährlich)} * 1000}{\text{Einwohnerzahl(jährlich)}}$$

$$\text{Sterberate (jährlich)} = \frac{\text{Sterbeziffer(jährlich)} * 1000}{\text{Einwohnerzahl(jährlich)}}$$

Die berechneten Merkmale wurden in der Sekundärdatei JAHRE gespeichert. Dieses Verfahren wurde für die letzten drei Generationszeiträume differenziert auf die verschiedenen sozialen Gruppen angewandt, um die sozial differenzierte natürliche Bevölkerungsbewegung zu erschließen.

Nach der Analyse der Heirats-, Geburten- und Sterberate wurde der Zusammenhang mit den Merkmalen zum Brotpreis und zum Krieg durch Bildung von PEARSON'schen Korrelationskoeffizienten und durch varianzanalytische Ansätze untersucht. Dabei wurde beim Brotpreis der Trend eliminiert. Da zu erwarten war, daß der Zusammenhang zwischen diesen Merkmalen eventuell zeitverzögert auftreten kann, wurden bei der Berechnung der Korrelationskoeffizienten LAG-Verschiebungen im Sinne der Zeitreihenanalyse berücksichtigt. Eigentliche Zeitreihenanalysen wurden bisher nicht durchgeführt, wären aber zur weiteren Analyse der Daten notwendig. Das Merkmal Geburtenüberschuß scheint dafür besonders gut geeignet, da auf dieses Merkmal der absolute Bevölkerungszuwachs keinen Einfluß hat und somit die Zeitreihe nicht von einem Trend bereinigt werden muß.

Die Abbildung 8 gibt eine Übersicht zur Analysestrategie der Sekundärdatei JAHRE.

Abhängigkeiten des Brotpreises von klimatischen Bedingungen in Form von mittleren Jahres-, Januar- und Julitemperaturen wurden durch Korrelationsanalyse untersucht. Die Ergebnisse erwiesen sich für diesen Zweck als ungeeignet, da die klimatischen Bedingungen nicht komplex genug erfaßt werden konnten und genauere Informationen über die Regionen, aus denen Berlin seine Lebensmittel bezog, nicht verfügbar waren.

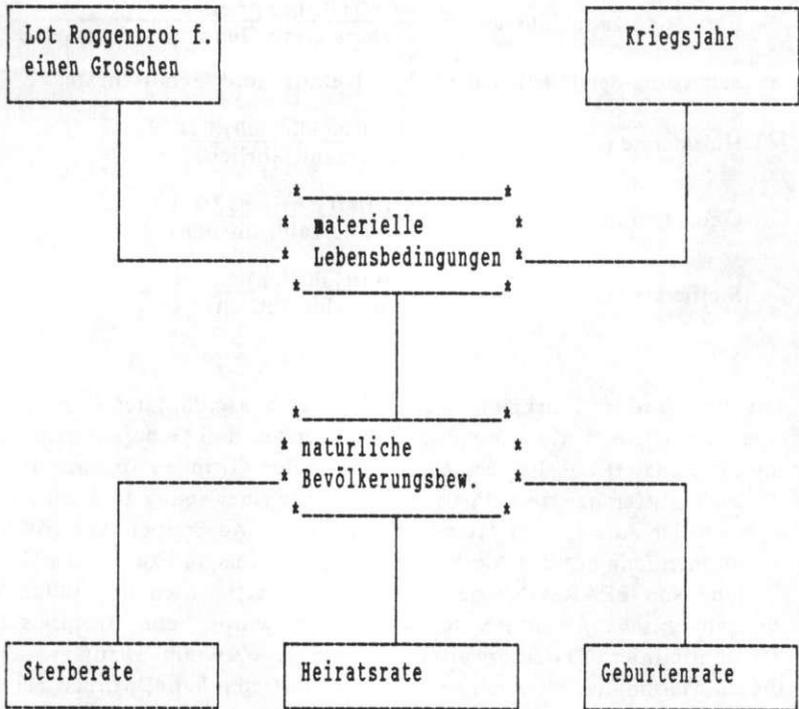


Abb. 8 : Natürliche Bevölkerungsentwicklung

7.2 Analyse der Basisdatei TAUFEN

Schwerpunkt der Analyse waren die sozialen Beziehungen, die sich durch die Patenwahl manifestierten. Die Abbildung 9 stellt die Beziehungen zwischen den Merkmalen der Basisdatei TAUFEN dar und soll als Orientierung zur Wahl der Auswertungsverfahren dienen.

Zunächst wurde der Einfluß der Zeit auf die Merkmale zur Klasse Schicht und Familienstand durch zweidimensionale Kontingenztafeln untersucht. Zur Analyse der unehelichen Geburten wurde eine dreidimensionale Kontingenztafel mit den Merkmalen zu Generationszeitraum, Familienstand und Klasse-Schicht untersucht.

Zur sozialen Beziehung durch Patenwahl wurde jede Wahl des Vaters als ein Ereignis aufgefaßt und danach eine zweidimensionale Kontingenztabelle mit den Merkmalen zu Klasse-Schicht des Vaters und Klasse-Schicht der Paten analysiert. Die Besetzungszahlen außerhalb der Hauptdiagonalen zeigten die soziale 'Migration' durch die Patenwahl. Zur Erschließung der zeitlichen Variabilität dieser Art der Migration wurde dann die Zeit als Generationszeitraum mit in die Kontingenztabelle einbezogen.

7.3 Analyse der Basisdatei HEIRATEN

Die Abbildung 10 zeigt die regionalen und sozialen Beziehungen, die der Auswertung hypothetisch zugrunde gelegt wurden. Ähnlich wie bei der Patenwahl wurden die verschiedensten Arten der Mobilität analysiert. Es wurden, unter Berücksichtigung der Zeit als Generationszeitraum, die Beziehungen der Klasse-Schicht des Bräutigamvaters und des Bräutigams, des Bräutigams und des Brautvaters und die Beziehungen der Herkunftsregion des Bräutigamvaters und des Brautvaters durch dreidimensionale Kontingenztabelle analysiert.

Zur Analyse des Heiratsverhaltens, differenziert nach Klasse-Schicht des Bräutigams unter Berücksichtigung der Zeit, wurden Hypothesen über saisonale Schwankungen ebenfalls durch Kontingenztabelleanalyse bestätigt.

7.4 Analyse der Basisdatei BEERDIGUNGEN

Die Auswertung, der in der Basisdatei BEERDIGUNGEN erfaßten Daten, zielte auf die Aufdeckung von Determinanten der Lebenserwartung. Abbildung 11 veranschaulicht die Auswertungshypothesen. Der Einfluß der Zeit als Generationszeitraum auf die Entwicklung der Klassen und Schichten und die Todesursachen wurde durch zweidimensionale Kontingenztabelle analysiert, der Einfluß der Zeit auf das Sterbealter durch Varianzanalysen getestet. Zur Untersuchung des Zusammenhanges zwischen Alter und Todesursache wurde das Alter in Altersgruppen zusammengefaßt und zweidimensionale Kontingenztabelle analysiert. Da sehr viele verschiedene Todesursachen erfaßt wurden, mußten auch die Todesursachen zu Todesursachengruppen zusammengefaßt werden. Zur Erschließung des Zusammenhanges zwischen Klasse-Schicht, Geschlecht, Zeit und Sterbealter wurde die Varianzanalyse eingesetzt. Dabei wurden die Berechnungen auf alle Probanden und auf Probanden mit einem Sterbealter ab 20 Jahren bezogen, um die Säuglings- bzw. Kindersterblichkeit differenziert betrachten zu können. Zur Vertiefung der Aussagen über die Lebenserwartungen

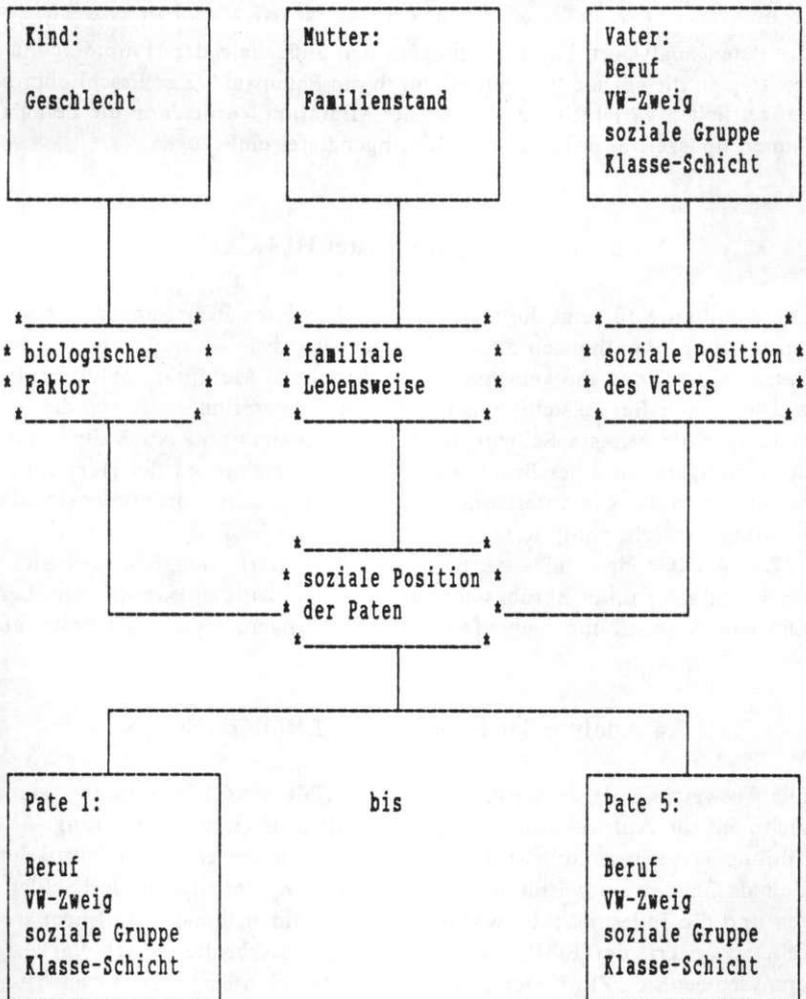


Abb. 9 : Soziale Beziehungen durch Patenwahl

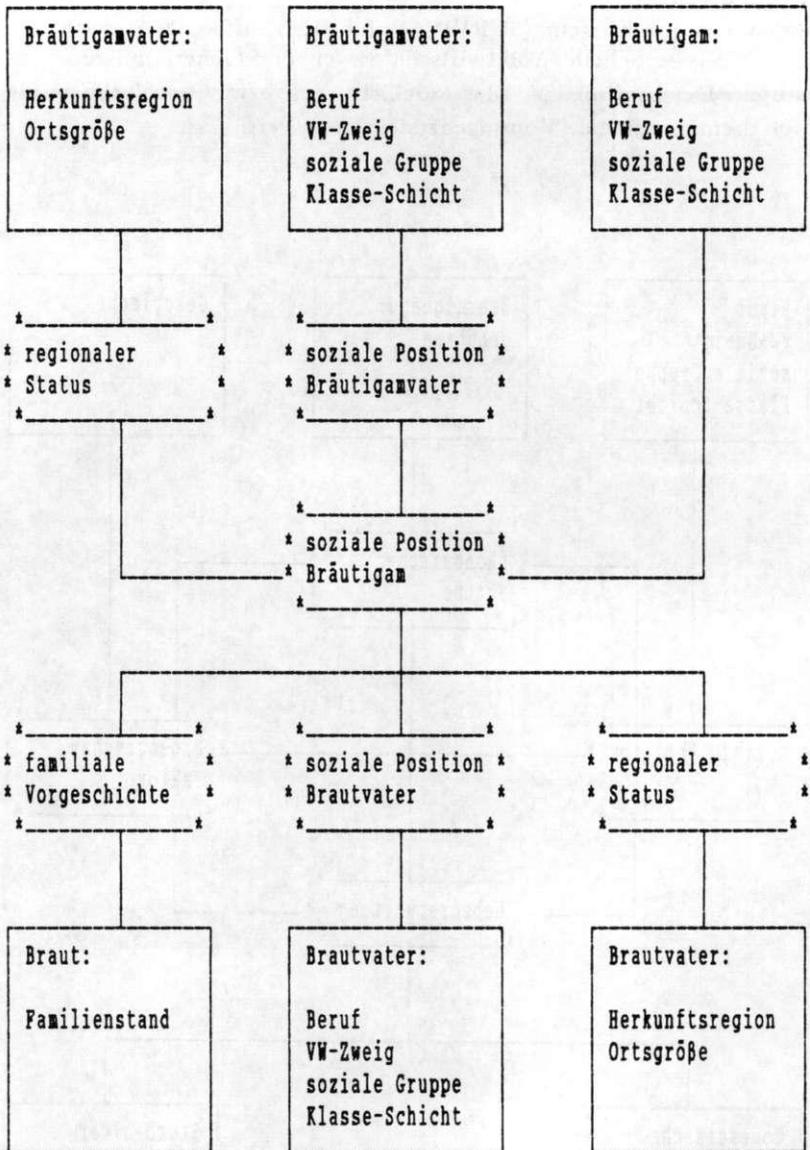


Abb.10 : Regionale und soziale Intergenerations- und Konubiusibilität

wurden Überlebensstafeln (SURVIVAL-Tabellen) differenziert nach Geschlecht, Klasse-Schicht, Volkswirtschaftszweig und Generationszeitraum ermittelt. Der Einfluß von Klasse-Schicht, Zeit bzw. Geschlecht wurde durch dreidimensionale Kontingenztafeln analysiert.

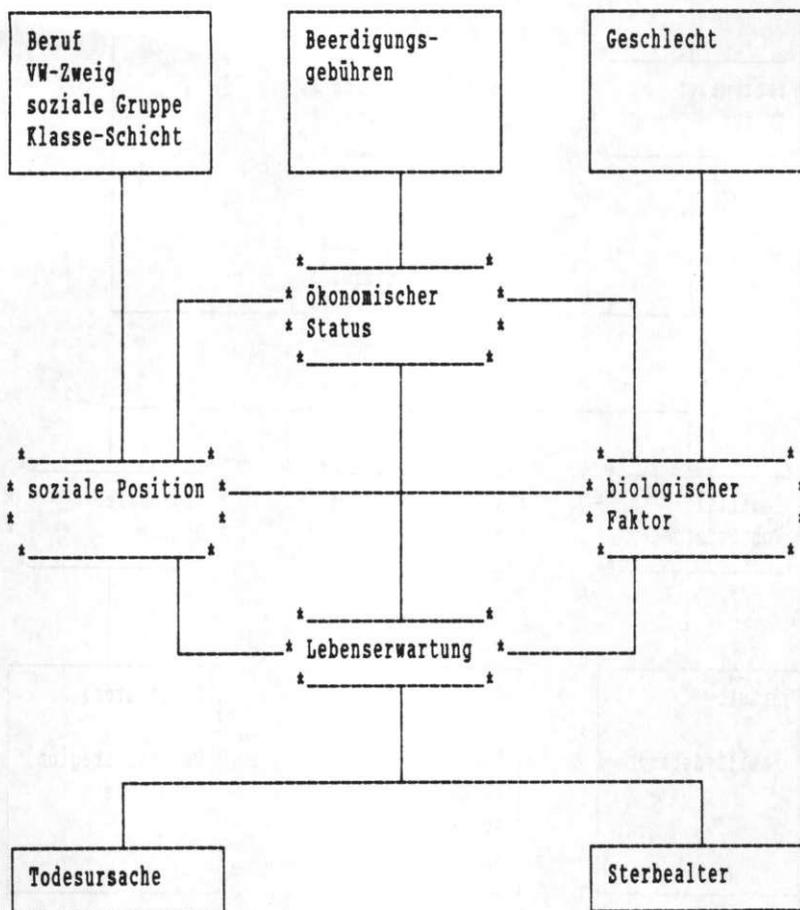


Abb.11 : Determinanten der Lebenserwartung

8. Basisdateien der Hugenottengemeinde

Die Basisdateien der Hugenottengemeinde beinhalten Angaben aus den Heirats- und Sterberegistern der Französischen Gemeinde Berlind. Es wurden nicht alle Ereignisse des Zeitraumes 1650 bis 1799 erfaßt, sondern nur ausgewählter Zeitabschnitte. Zur Datenaufnahme wurden die vorhandenen Belege für die Kirchen St. Nikolai und St. Georgen verwendet, die allerdings in einigen Positionen erweitert wurden. Es ergaben sich deshalb für einige Merkmale abweichende Merkmalsausprägungen, weshalb auch keine Verkettung der Hugenottendateien mit den Basisdateien der Berliner Kirchspiele St. Nikolai und St. Georgen erfolgte.

Die Basisdatei HUGHEIRATEN enthält zusätzlich zu den Merkmalen der Basisdatei HEIRATEN das Merkmal NATIONALITAET (VAR30). Außerdem wurden die Wertetiketten für die Merkmale zum Familienstand und zur Region abgeändert.

Zur Beschreibung der Nationalität der Partner wurden 10 Codes benötigt:

1 = FR /FR	
2 = FR /DT	
3 - DT /FR	Abkürzungserklärung:
4 = JU /FR	
5 = FR /JU	KATH - katholisch
6 = DT /DT	FR - französisch
7 = FR(REF) /KATH	DT - deutsch
8 = KATH /FR(REF)	JU - jüdisch
9 - AUSLAENDER	REF - reformiert

Tab. 27: Wertetiketten für das Merkmal NATIONALITAET

Auch für den Familienstand ergaben sich andere Codierungen:

1 - VERHEIRATET
2 = SIE WITWE
3 = GESCHIEDEN
4 = ER WITWER
5 = BEIDE VERWITWET

Tab. 28: Wertetiketten für das Merkmal FAMILIENSTAND in der Basisdatei HUGHEIRATEN

Die Regionen werden differenzierter als in der Basisdatei HEIRATEN angegeben.

0=BERLIN	1 = REST BRANDENBURG
2=METZ(ERLAND)	3=SEDAN
4= GENF	5=PARIS
6= LAUSANNE	7=PFALZ
8=LANGUEDOC	9= NIMES
10= GRENoble	11 = LOTHRINGEN
12 = DAUPHINE	13= CHAMPAGNE
14= BRIE	15= GUIENNE/GASCOGNE
16= PICARDIE/ARTOIS	17= SEVENNEN/AUVERGNE
18= NORMANDIE	19=VIVARAIS
20= ILE DE FRANCE	21 = FLANDERN
22=POITOU/XAINTONGE	23 = SOITONNE/ANGOUMOIS
24= PROVENCE	25= ORLEANAIS
26= LA BRESSE/GREX	27=ELSASS
28 = BERRY	29= LYONNAIS
30= LYON	31 = NEUFCHATEL/VALLENGIN
32 = FRANCHE COMTE	33= PIEMONT
34= ORANGE	35= FOIX/ROUSSILLON
36= HENNEGAW	37= BRETAGNE
38 = TOURAINE	39=BEARN
40= ANJOU	41 = NIVERNOIS
42= LA MARCHE	43= LIMOUSIN
44 = MAINE	45 = FRANKREICH
46 = SCHWEIZ	47= HOLLAND
48 = IRELAND	49= ENGLAND
50= DAENEMARK	51= SCHWEDEN
52 = STRASBURG (UCKERMARK)	53=PRENZLAU
54=PASEWALK	55= GRAMZOW/SCHWEDT/ BERGHOLZ/VIERRADEN
56= ITALIEN	57= KLEIN/GROSS ZIETHEN
58= EBERSWALDE/CHORIN	59=LOECKNITZ
60= ANGERMUENDE	61 = STARGARDT
62= KAGAR/RHEINSBERG/ RUPPIN/ZEHDENIK	63= MAGDEBURG
64 = HALBERSTADT/CALBE	65 = MANNHEIM
66=KOEPEINICK	67 = BUCHHOLZ/PANKOW
68 = MUENCHEBERG	69= POTSDAM/SPANDAU
70 = BERNAU	71 = FRANKFURT/O
72 = FRANKFURT/M	73 = FUERSTENWALDE
74 = HALLE	75 = BRANDENBURG(STADT)
76 = STENDAL	77= ALTMARK(WITTSTOCK/ NEUSTADT/ PRIEGNITZ)
78 = BURG	79= NEUHALDENSLEBEN
80= DUISBURG/SOEST/MINDEN	81 = KOENIGSBERG
82 = WESEL	83= RAUM COTTBUS
84= LEIPZIG	85= DRESDEN
86 = STETTIN/DANZIG	87= OSTFRIESLAND/KLEVE/ MARK/RAVENSBERG
88 = HAMBURG	89= BREMEN
90= ANDERE OEST PROVINZEN/SCHLESISIEN	91 = MECKLENBURG
92=SACHSEN/THUERINGEN	93= POMMERN/SCHLESWIG HOLSTEIN

94= BAYERN/WESTELB.	95 = HESSEN
94= DT TERRITORIEN	
96= BADEN/WUERTEMBERG/ BAYERN	97= RUSSLAND/POLEN
98=BOEHMEN/MAEHREN/ HABSBURG.L.	99=BOURGOGNE/ANDERE LAEN- DER

Tab. 29: Werteetiketten für die Merkmale zur Region in der Basisdatei HUGHEIRATEN

Eine Gesamtübersicht zu den Basisdateien der Hugenottengemeinde gibt die Tabelle 30.

Generationszeitraum	Heiraten	Beerdigungen	Gesamt
1650 bis 1709	671	1868	2539
1710 bis 1739	960	-	960
1740 bis 1769	473	3084	3557
1770 bis 1799	458	1694	2152
Gesamt	2562	6646	9208

Tab. 30: Absolute Häufigkeiten für die Basisdateien der Hugenottengemeinde

Sowohl Transformationen als auch die Datenanalyse wurden analog zu den Basisdateien der Kirchspiele St. Georgen und St. Nikolai durchgeführt. Insbesondere stellte sich heraus, daß die starke Differenzierung nach Regionen zu geringe Besetzungszahlen ergab, so daß bei den Analysen hauptsächlich mit den nachfolgend aufgeführten Zusammenfassungen der Regionen gearbeitet wurde.

0 = BERLIN	1 = REST BRANDENBURG
2-45 = FRANKREICH	46 = SCHWEIZ
47 = HOLLAND	48 = IRELAND
49 = ENGLAND	50 = DAENEMARK
51 = SCHWEDEN	52-87= KOLONIEN
88 = HAMBURG	89 = BREMEN
90 = ANDERE OEST. PROVINZEN/SCHLESSEN	91 = MECKLENBURG
92 = SACHSEN/THUERINGEN	93 - POMMERN/SCHLESWIG HOLSTEIN
94 = BAYERN/WESTELB. DTTERRITORIEN	95 = HESSEN
96 = BADEN/WUERTEMBERG/ BAYERN	97 = RUSSLAND/POLEN
98 = BOEHMEN/MAEHREN/ HABSBURG.L.	99 = BOURGOGNE/ANDERE LAENDER

Tab. 31: Werteetiketten für die Merkmale zur gruppierten Region

Die beiden Basisdateien der Hugenottengemeinde wurden in der folgenden Form archiviert:

Merkmals- name	Merkmalsetikett	Format
VARI	LAUFENDE NUMMER	
VAR2	KIRCHSPIEL	KSPIEL
VAR3	JAHR UND MONAT	DATE
VAR4	JAHR	
VAR5	MONAT	MONAT
VAR6	JAHRZEHLT	JHZEHLT
VAR7	GENERATIONSZEITRAUM	GZEITR
VAR8	VOLKSWIRTSCHAFTSZWEIG	
	BRAEUTIGAM	VWZWEIG
VAR9	BERUF BRAEUTIGAM	BERUFE
VAR10	SOZIALE GRUPPE	
	BRAEUTIGAM	SOZGRUP
VAR11	KLASSE - SCHICHT	
	BRAEUTIGAM	KLASSE
VAR12	MONATSGEWICHT	
VAR13	REGION BRAEUTIGAM	XREGION
VAR14	ORTSGROESSE	
	BRAEUTIGAM	GROESSE
VAR15	VOLKSWIRTSCHAFTSZWEIG	
	BRAEUTIGAMVATER	VWZWEIG
VAR16	BERUF	
	BRAEUTIGAMVATER	BERUFE
VAR17	SOZIALE GRUPPE	
	BRAEUTIGAMVATER	SOZGRUP
VAR18	KLASSE - SCHICHT	
	BRAEUTIGAMVATER	KLASSE
VAR21	FAMILIENSTAND DER	
	BRAUT	XFAMST
VAR22	VOLKSWIRTSCHAFTSZWEIG	
	BRAUTVATER	VWZWEIG
VAR23	BERUF BRAUTVATER	BERUFE
VAR24	SOZIALE GRUPPE	
	BRAUTVATER	SOZGRUP
VAR25	KLASSE - SCHICHT	
	BRAUTVATER	KLASSE
VAR28	REGION BRAUT	XREGION
VAR29	ORTSGROESSE BRAUT	GROESSE
VAR30	NATIONALITAET	XNATION

Tab. 32: Merkmalsetiketten und Formate der Basisdatei HUGHEIRATEN

Merkmalname	Merkmalsetikett	Format
VARI	LAUFENDE NUMMER	
VAR2	KIRCHSPIEL	KSPIEL
VAR3	JAHR UND MONAT	DATE
VAR4	JAHR	
VAR5	MONAT	MONAT
VAR6	JAHRZEHNT	JHZEHNT
VAR7	GENERATIONSZEITRAUM	GZEITR
VAR8	VOLKSWIRTSCHAFTSZWEIG	VWZWEIG
VAR9	BERUF	BERUFE
VARIO	SOZIALE GRUPPE	SOZGRUP
VAR11	KLASSE - SCHICHT	KLASSE
VARI 2	GESCHLECHT	GESCHL
VARI 3	ALTER	FSTP
VAR15	ALTERSGRUPPE	ALTGRUP
VARI 6	TODESURSACHE	DIAGNOSE
VARI 7	BEERDIGUNGSGEBUEHREN	FSTP
VAR22	TODESURSACHENGRUPPEN	DIAGRUP

Tab. 33: Merkmalsetiketten und Formate der Basisdatei
HUGBEERDIGUNGEN

Verzeichnis der Abbildungen und Tabellen

- Tab. 1: Merkmalsetiketten und Formate der Basisdatei TAUFEN
 Tab. 2: Merkmalsetiketten und Formate der Basisdatei HEIRATEN
 Tab. 3: Merkmalsetiketten und Formate der Basisdatei BEERDIGUNGEN
 Tab. 4: Merkmalsetiketten der Sekundärdatei JAHRE
 Abb. 1: Datenbank 'Sozialgeschichte Berlind 1650 bis 1799'
 Abb. 2: Prinzipieller Aufbau einer Datenmatrix
 Abb. 3: Netzwerk des Projektverlaufes
 Abb. 4: Ablauf der Datenvorverarbeitung
 Abb. 5: Struktogramm zur Fehleranalyse
 Abb. 6: Fehlerlistenaufbau
 Tab. 5: Absolute Ereignishäufigkeiten in den Basisdateien
 Tab. 6: Zuordnung der Merkmale zu den Basisdateien
 Tab. 7: Wertetiketten und absolute Häufigkeiten für das Merkmal KIRCHSPIEL
 Tab. 8: Wertetiketten und absolute Häufigkeiten für das Merkmal MONAT
 Tab. 9: Wertetiketten für das Merkmal JAHRZEHNT
 Tab. 10: Wertetiketten für das Merkmal GENERATIONSZEITRAUM

- Tab. 11: Werteetiketten für die Merkmale zum Beruf
- Tab. 12: Basisdatei BEERDIGUNGEN - absolute Häufigkeiten für das Merkmal VOLKSWIRTSCHAFTSZWEIG
- Tab. 13: Werteetiketten für die Merkmale zu den sozialen Gruppen
- Tab. 14: Werteetiketten für die Merkmale zur Klasse-Schicht
- Tab. 15: Bildungsvorschrift für die Merkmale zur KlasseSchicht aus den Merkmalen zur sozialen Gruppe
- Tab. 16: Basisdatei TAUFFEN - absolute Häufigkeiten der Merkmale zur Klasse-Schicht der Paten
- Tab. 17: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale zur Klasse-Schicht
- Tab. 18: Basisdatei TAUFFEN - Häufigkeitsverteilung des Merkmals FAMILIENSTAND
- Tab. 19: Basisdatei HEIRATEN - Häufigkeitsverteilung des Merkmals FAMILIENSTAND DER BRAUT
- Tab. 20: Basisdatei BEERDIGUNGEN - Häufigkeitsverteilung des Merkmals GESCHLECHT
- Tab. 21: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale REGION BRAUT und REGION BRAEUTIGAM
- Tab. 22: Basisdatei HEIRATEN - absolute Häufigkeiten der Merkmale ORTSGROSSEBRAUT und ORTSGROSSE BRAEUTIGAM
- Tab. 23: Werteetiketten für das Merkmal ALTERSGRUPPE
- Abb. 7: Basisdatei BEERDIGUNGEN - Häufigkeitsdiagramm mit Boxplot des Merkmals ALTER
- Tab. 24: Werteetiketten für das Merkmal TODESURSACHE
- Tab. 25: Basisdatei BEERDIGUNG - Bildungsvorschrift des Merkmals TODESURSACHENGRUPPEN und absolute Häufigkeiten
- Tab. 26: Basisdatei TAUFFEN - Häufigkeitsverteilung des Merkmals PATEN INSGESAMT
- Abb. 8: Natürliche Bevölkerungsbewegung
- Abb. 9: Soziale Beziehungen durch Patenwahl
- Abb. 10: Regionale und soziale Intergenerations- und Konubiumsmobilität
- Abb. 11: Determinanten der Lebenserwartung
- Tab. 27: Werteetiketten für das Merkmal NATIONALITAET
- Tab. 28: Werteetiketten für das Merkmal FAMILIENSTAND in der Basisdatei HUGHEIRATEN
- Tab. 29: Werteetiketten für die Merkmale zur Region in der Basisdatei HUGHEIRATEN
- Tab. 30: Absolute Ereignishäufigkeiten für die Basisdateien der Hugenottengemeinde
- Tab. 31: Werteetiketten für die Merkmale zur gruppierten Region
- Tab. 32: Merkmalsetiketten der Basisdatei HUGHEIRATEN
- Tab. 33: Merkmalsetiketten der Basisdatei HUGBEERDIGUNGEN