Bakkalaureatsarbeit
zum Thema

Volkswirtschaftliche Gesamtrechnung
mit Input-Output-Tabellen
unter Berücksichtigung der Komplexität von Arbeit

unter Anleitung von o.Univ.Prof.i.R. Univ.Doz. Dipl.Ing. Dr. Peter Fleissner

von

Schlegel Michael, 0401831
&
Szolarz Christian, 0325374
Inhaltsverzeichnis

1 Berechnung von volkswirtschaftlichen Verflechtungen ........................................... 5
   1.1 Kreislauf à la Quesnay ....................................................................................... 5
   1.2 Kreislaufbild von Karl Marx ........................................................................... 6
   1.3 Entwicklung hin zur Input-Output-Analyse ...................................................... 8
2 Wirtschaft als Kreislauf ......................................................................................... 10
3 Idee der Input-Output-Tabelle ............................................................................... 11
4 Theoretische Grundlagen zur Input-Output-Tabelle .............................................. 12
   4.1 Aggregation zu Sektoren .................................................................................. 12
   4.2 Maß der ökonomischen Größen ....................................................................... 12
   4.3 Einheitenmaße .................................................................................................. 15
5 Darstellung eines Kreislaufes ............................................................................... 16
   5.1 Grafische Darstellung ....................................................................................... 17
   5.2 Kontenform ....................................................................................................... 18
   5.3 Algebraisch ........................................................................................................ 19
   5.4 Matrix-Form ...................................................................................................... 20
6 Die Input-Output-Tabelle ....................................................................................... 21
   6.1 Ausgangsbasis Matrix-Form ............................................................................. 21
   6.2 Struktur einer Input-Output-Tabelle ................................................................. 23
   6.3 Eigenschaften einer Input-Output-Tabelle ......................................................... 25
      6.3.1 Semantisch .................................................................................................... 25
      6.3.2 Mathematisch ................................................................................................ 26
7 Verwendetes Datenmaterial .................................................................................... 28
   7.1 Klassifikationen ................................................................................................. 28
      7.1.1 ISIC ............................................................................................................... 28
      7.1.2 NACE ............................................................................................................ 30
      7.1.3 ÖNACE .......................................................................................................... 31
      7.1.3.1 Gliederung der Daten nach ÖNACE 2003 .................................................. 31
      7.1.3.2 ÖNACE 2003 vs. ÖNACE 2008 ................................................................. 41
   7.2 Erwerbspersonen nach Bildung und ÖNACE 2001 ............................................. 43
      7.2.1 Einteilung der Bildungsstufen ..................................................................... 44
         7.2.1.1 Österreichische Bildungsebenen ............................................................... 44
            7.2.1.1.1 Universitäten und Hochschulen sowie Fachhochschulen ............... 45
            7.2.1.1.2 Berufs- und lehrerbildende Akademie ............................................. 45
            7.2.1.1.3 Kollegs und Abiturientenlehrgänge .................................................. 46
8 Berechnungen ........................................................................................................ 51

8.1 Leontief-Modell .................................................................................................. 51

8.2 Regression ........................................................................................................... 54

8.2.1 Lineare Regression ........................................................................................... 56

8.2.1.1 Einfache lineare Regression ....................................................................... 56

8.2.1.2 Multiple lineare Regression ....................................................................... 63

8.2.2 Gütekriterien der Regression ........................................................................... 66

8.2.2.1 Regressionsstatistik .................................................................................... 66

8.2.2.1.1 Korrelationskoeffizient .......................................................................... 66

8.2.2.1.2 Bestimmtheitsmaß .................................................................................. 67

8.2.2.1.3 Standardfehler der Regression ............................................................... 67

8.2.2.1.4 Beobachtungen ....................................................................................... 68

8.2.2.2 ANOVA (analysis of variance) ..................................................................... 68

8.2.2.2.1 Freiheitsgrade .......................................................................................... 68

8.2.2.2.2 Quadratsummen ...................................................................................... 69

8.2.2.2.3 Mittlere Quadratsummen ....................................................................... 72

8.2.2.2.4 F-Statistik ............................................................................................... 72

8.2.2.3 Regressionskoeffizienten ........................................................................... 75

8.2.2.4 Standardfehler der Regressionskoeffizienten ........................................... 75

8.2.2.4.1 T-Statistik .............................................................................................. 76

8.2.2.4.2 Konfidenzintervall ............................................................................... 79

8.2.3 Anmerkungen .................................................................................................. 80

9 Dokumentation der Berechnungen in Excel .......................................................... 80

9.1 InstitutionelleI/O.xls ......................................................................................... 80

9.1.1 Ausgangsmaterial .............................................................................................. 81

9.1.1.1 Heimische Produktion ................................................................................. 82

9.1.1.2 Use-Matrix, Vorleistungen ......................................................................... 82

9.1.1.3 Endnachfrage .............................................................................................. 82

9.1.1.4 Wertschöpfung ......................................................................................... 83

9.1.2 Tabellenblätter ................................................................................................. 83

9.1.3 Erläuterungen zu den Berechnungen .............................................................. 83
9.2 Arbeitswerte.xls ................................................................................................................................. 85
9.2.1 Ausgangsmaterial ............................................................................................................................. 85
9.2.2 Tabellenblätter .............................................................................................................................. 85
9.2.3 Erläuterungen zu den Berechnungen .......................................................................................... 86

10 Interpretationen der berechneten Daten ......................................................................................... 89
10.1 Überblick über die Regressionsergebnisse .................................................................................. 92
10.2 Interpretation der aussagekräftigen Resultate .............................................................................. 93

11 Literaturverzeichnis .......................................................................................................................... 96
11.1 Druckwerke ..................................................................................................................................... 96
11.2 Internetseiten .................................................................................................................................. 97
1 Berechnung von volkswirtschaftlichen Verflechtungen


1.1 Kreislauf à la Quesnay


Zahlungsströme im Kreislauffbild von Quesnay
In diesem „Tableau économique“ wurde nur zwischen drei verschiedene Klassen unterschieden: Die produktive Klasse, welche durch Landwirte und Pächter repräsentiert wurde, die distributive Klasse, also die Klasse der Grundbesitzer, welche sich durch den Adel sowie den Klerus zusammensetzten, und die sterile Klasse, zu der Händler und Handwerker zählten. Die produktive Klasse produziert landwirtschaftliche Produkte in Höhe von fünf Milliarden Livres. Zwei Milliarden davon verbraucht die Klasse selbst für Saatgut, Futtermittel, etc. Eine weitere Milliarde wird für Handelserzeugnisse, welche bei der sterilen Klasse erworben werden, aufgewendet. Die zusätzlichen zwei Milliarden Livres werden als Pacht an die Grundeigentümer abgeführt. Diese wiederum geben eine Milliarde an die produktive Klasse weiter um Nahrungsmittel zu kaufen und verwenden die andere Milliarde um Handelserzeugnisse zu erwerben (bei der sterilen Klasse). 2 Milliarden Livres werden für landwirtschaftliche Erzeugnisse von der sterilen Klasse an die produktive Klasse gezahlt. Das Modell von Quesnay kann auch als folgende Matrix dargestellt werden:

<table>
<thead>
<tr>
<th>von Klasse</th>
<th>Produktive Klasse</th>
<th>Sterile Klasse</th>
<th>Distributive Klasse</th>
<th>∑</th>
</tr>
</thead>
<tbody>
<tr>
<td>Produktive Klasse</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Sterile Klasse</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Distributive Klasse</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>∑</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

Das Modell geht von einem geschlossenen Kreislauf aus, da die Summenspalte und Summenzeile gleich sind. Trotzdem das Modell simple ist, kann es mit heutigen Input-Output-Tabellen verglichen werden.

1.2 Kreislaufbild von Karl Marx
Erst in der Mitte des 19. Jahrhunderts wurde die Idee, die Wirtschaft als einen Kreislauf zu betrachten, wiederaufgenommen. Marx wollte zeigen, dass vor Hintergrund der sich verändernden Gesellschaft und der immer bedeutenderen
Ökonomie durch den Faktor Arbeit ein Mehrwert geschaffen wird, der den Kapitalisten zukommt. Sein Hauptfokus lag auf dem Verteilungsproblem.

Kreislauf erweiterter Reproduktion nach Karl Marx


1.3 Entwicklung hin zur Input-Output-Analyse


Demzufolge liefert es einen Rahmen für relativ genau Beschreibung der Wirtschaftsstruktur und es erlaubt außerdem Prognosen über die Auswirkungen wirtschaftspolitischer Eingriffe in diese Struktur. Demgemäß handelt es sich bei der Input-Output-Methode keineswegs um ein theoretisches Modell, wie bei den meisten Vorgängern, sondern um einen Versuch die reale wirtschaftliche Situation abzubilden.


2 Wirtschaft als Kreislauf


der heutigen Wirtschaft steckt und welche Rolle der Preis von Gütern beziehungsweise die Vergütung von Arbeitszeit, damit Güter hergestellt werden können, spielt. So kann man sich fragen, ob der Preis einer Ware oder Dienstleistung dem Arbeitsaufwand, der für die Produktion notwendig ist, entspricht und ob durch die Berechnung mit Hilfe der enthaltenen Arbeitswerte die reale wirtschaftliche Situation angenähert werden kann. Überdies stellt sich die Frage, inwieweit der Arbeitswert abhängig von der Ausbildung ist, also wie viel größer der Output eines besser ausgebildeten Arbeiters bei gleichem Arbeitsvolumen ist.

3 Idee der Input-Output-Tabelle


Die Idee des Input-Output-Schemas ist es nun die gesamte Wirtschaft als einen Kreislauf zu sehen – analog zu Quesnays Vorstellung des Blutkreislaufes. Outputs
entstehen durch den Verbrauch von Inputs. Dazu wirken eine Menge an Entitäten (Produzenten, Händler, Verkäufer, Konsumenten, etc.) mit und eine Vielzahl an Relationen zwischen diesen Elementen (Einkauf, Verkauf, Lieferungen, Transaktionen, etc.) kommt zu tragen. Dadurch lassen sich die volkswirtschaftlichen Verflechtungen sehr genau und nahezu widerspruchsfrei abbilden und mit solcher Hilfe lassen sich wirtschaftliche Interventionen prognostizieren beziehungsweise Simulationsmodelle für bestimmte ökonomische Konstellationen erstellen.

4 Theoretische Grundlagen zur Input-Output-Tabelle

4.1 Aggregation zu Sektoren


Trotz dieser Schwierigkeiten gibt es für Input-Output-Tabellen, die in der volkswirtschaftlichen Gesamtrechnung verwendet werden, schon entsprechende Normen. So zum Beispiel die ÖNACE, welche wir später noch genauer behandeln werden und folglich in dieser Arbeit auch Verwendung finden wird (siehe Kapitel 7.1.3).

4.2 Maß der ökonomischen Größen
Durch An- und Verkauf wandern nun Waren von einem Sektor zu einem Anderen. Dies wird als Strom- oder auch Flußgröße bezeichnet und ist in der Input-Output-

Es sollte erwähnt werden, dass es alternativ zu den Stromgrößen auch noch die *Bestandgrößen* gibt. Diese kennzeichnen sich dadurch, dass sie zeitpunktbezogen sind, also zu einem bestimmten Zeitpunkt erfasst werden. So lässt sich zum Beispiel das Ergebnis einer Volkszählung klassifizieren, da an einem bestimmten Tag die Anzahl der Menschen erhoben wird. Andere Beispiele sind der Lagerbestände einer Firma zur Zeit der Inventur oder die Anzahl der Beschäftigten zu einem bestimmten Stichtag.

da sie im Allgemeinen keine direkt erkennbare ökonomische Gegenleistung haben. Zur Veranschaulichung dient folgende vereinfachte Grafik:\(^1\):

**Ströme: reale vs. monetäre**

Wenn in einem Kreislaufschema alle Sektoren direkt oder indirekt miteinander verbunden sind und jeder dieser Sektoren durch zumindest einen zu- als auch einen abfließenden Strom verbunden sind, so spricht man von einem **offenen** Kreislauf. Dabei wird jede Stromgröße einmal als Zu- und einmal als Abfluss dargestellt und infolgedessen ist die Gesamtsumme der Zuflüsse gleich groß wie die Gesamtsumme der Abflüsse. Sind nun für jeden Sektor die beiden genannten Gesamtsummen wertentsprechend, so liegt ein **geschlossener** Kreislauf vor. Eine ökonomische Interpretation der Gesamtsummengleichheit geht davon aus, dass in geschlossenen Kreisläufen ausnahmslos gleichwertige Tauschhandel abgebildet werden. Das wiederum impliziert die vorgeschriebene Bedingung, dass für den

\(^1\) Die Abbildung zeigt eine geschlossene Wirtschaft ohne Staat und ohne Vermögensbildung
Zufluss jedes Sektors ein gleichwertiger Abfluss existieren muss. Ein offener Kreislauf kann durch Hinzufügen eines oder mehrerer Sektoren in einen geschlossenen Kreislauf umgewandelt werden.

In dieser Arbeit werden nur Input-Output-Tabellen mit einem geschlossenen Kreislaufschema zur Anwendung kommen.

4.3 Einheitenmaße

Güterströme werden aufgrund der besseren Vergleichsmöglichkeit üblicherweise in Preisen bewertet. Dabei ist jedoch zu beachten, dass es je nach konkreter Fragestellung verschiedene Geldeinheiten verwendet werden können. So können unter anderem die Herstellerpreise, die Preise der Zwischenhändler oder die Kosten für den Endverbraucher aufschlussreich sein.


Ebenfalls interessant für die Analyse von Umweltproblemen kann die Bewertung der Güterströme in Energieeinheiten sein. Damit kann der notwendige Energieaufwand für die Produktion, beispielsweise in Joule, angeführt und verglichen werden.

Die Angabe in Arbeitszeit ist die für uns bedeutungsvollste Möglichkeit der Güterstrombezeichnung. Sie wird auch in weiterer Folge noch Verwendung finden. Dabei wird der Wert einer produzierten Waren nach seinem tatsächlichen Arbeitsvolumen berechnet. Das heißt es wird nur die reale Arbeitszeit berücksichtigt, also die Zeit, die zum Beispiel ein Auto benötigt, bis es fertig zusammengebaut ist. Wobei natürlich die Zeit, die zur Fertigstellung aller Einzelteile notwendig ist, ebenfalls berücksichtigt wird. Diese Darstellung kommt vor allem in der Wirtschaftspolitik zum Einsatz. Es können damit die Auswirkungen von Investitionen auf den Arbeitsmarkt beziehungsweise auf die Beschäftigungssituation prognostiziert werden. In unserem Fall nutzen wir diese Ansicht um die Wertigkeit von Arbeit hinsichtlich der höchsten abgeschlossenen Schulbildung zu erhalten.
5 Darstellung eines Kreislaufes

Das Kreislaufschema kann nun auf mehrere Arten visualisiert werden. Folgende vier Formen werden dann auch im weiteren Verlauf genauer beschrieben:

- Grafisch als Blockschaltdiagramm
- Kontenform (z.B. Güterkonto, Produktionskonto, Vermögenskonto)
- Algebraisch als Gleichungssysteme
- In Matrix-Form (z.B. Input-Output-Tabellen)

Hierbei ist festzuhalten, dass die Darstellung in Matrix-Form im Grunde eine spezielle Art der algebraischen Schreibweise und daher im eigentlichen Sinne keine echte, zusätzliche Darstellungsform ist. Dennoch muss sie hier als eigene Form erwähnt werden, da im Weiteren lediglich mit dieser Darstellungsart gearbeitet wird und daher wird dieser Form noch genauer behandelt werden.

Im Weiteren werden folgende Abkürzungen verwendet, wobei meistens, dort wo es möglich ist, auch die vollen Namen angegeben werden:

<table>
<thead>
<tr>
<th>Sektoren</th>
<th>Transaktionen</th>
</tr>
</thead>
<tbody>
<tr>
<td>H Haushalte</td>
<td>Y Einkommen</td>
</tr>
<tr>
<td>U Unternehmen</td>
<td>TR Transferzahlungen</td>
</tr>
<tr>
<td>St Staat</td>
<td>I Investitionen</td>
</tr>
<tr>
<td>A Ausland</td>
<td>Ex Exporte</td>
</tr>
<tr>
<td>V Vermögensveränderungssektor</td>
<td>VL Vorleistungen</td>
</tr>
</tbody>
</table>

Die Flussrichtung ist ebenfalls codiert. Sie wird als Index nach der jeweiligen Transaktion angegeben, wobei die erste Angabe den Sektor des Abflusses bezeichnet und der durch einen Schrägstrich (Slash) getrennten zweiten Wert den Sektor des Zuflusses beschreibt. \( Y_{U/H} \) würde somit den Strom der Einkommen von den Unternehmen zu den Haushalten bedeuten.
5.1 Grafische Darstellung

Einen Wirtschaftskreislauf grafisch abzubilden bringt besonders gute Anschaulichkeit:

Wirtschaftssubjekte werden als Kästchen gezeichnet (Unternehmen, Staat, etc.) und die Relationen zwischen ihnen als gerichtete Pfeile (Einkommen, Steuern, Investitionen, etc.). Dadurch lassen sich Ströme und deren Zusammenhänge schnell...

5.2 Kontenform
Die Darstellung eines wirtschaftlichen Kreislaufes in Kontenform ist sehr einfach zu verstehen, wie die nachfolgende Tabelle zeigt:

<table>
<thead>
<tr>
<th>Zuströme</th>
<th>Abströme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Einkommen von Unternehmen</td>
<td>$Y_U/U$</td>
</tr>
<tr>
<td>Einkommen vom Staat</td>
<td>$Y_{St/U}$</td>
</tr>
<tr>
<td>Transferzahlungen vom Staat</td>
<td>$TR_{St/U}$</td>
</tr>
<tr>
<td>Transferzahlungen vom Ausland</td>
<td>$TR_{A/U}$</td>
</tr>
<tr>
<td>Zuströme gesamt</td>
<td>$\sum$</td>
</tr>
<tr>
<td>Abströme gesamt</td>
<td>$\sum$</td>
</tr>
</tbody>
</table>

5.3 Algebraisch

Die Anschaulichkeit ist, wie man deutlich erkennen kann, nicht in dem Maße gegeben wie bei den anderen Formen. Jedoch finden sich auch hier sämtliche Posten der Zu- und Abflüsse wieder, wie ein simpler Vergleich der folgenden Gleichungssysteme mit der Abbildung des österreichischen Kreislaufes zeigt:

<table>
<thead>
<tr>
<th>Sektor</th>
<th>Zuströme = Abströme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private Haushalte</td>
<td>( Y_{H/H} + Y_{ST/H} + TR_{ST/H} + TR_{A/H} )</td>
</tr>
<tr>
<td>(H)</td>
<td>( = )</td>
</tr>
<tr>
<td></td>
<td>( C_{H/U} + T^{dir}<em>{H/ST} + TR</em>{H/A} + S_{H/V} )</td>
</tr>
<tr>
<td>Unternehmen</td>
<td>( C_{H/U} + VL_{ST/U} + I^b_{ST/U} + Z_{ST/U} + Ex_{A/U} + I^v_{V/U} )</td>
</tr>
<tr>
<td>(U)</td>
<td>( = )</td>
</tr>
<tr>
<td></td>
<td>( Y_{U/U} + Y_{U/ST} + T^{dir}<em>{U/ST} + T^{ind}</em>{U/ST} + Im_{U/A} + S_{U/V} )</td>
</tr>
<tr>
<td>Staat</td>
<td>( T^{dir}<em>{ST/H} + Y</em>{ST/H} + T^{dir}<em>{ST/ST} + T^{ind}</em>{ST/ST} + TR_{A/ST} + I^n_{V/ST} )</td>
</tr>
<tr>
<td>(St)</td>
<td>( = )</td>
</tr>
<tr>
<td></td>
<td>( Y_{ST/H} + TR_{ST/H} + VL_{ST/U} + I^b_{ST/U} + Z_{ST/U} + TR_{ST/A} + S_{ST/V} )</td>
</tr>
<tr>
<td>Ausland</td>
<td>( TR_{A/H} + Im_{U/A} + TR_{ST/A} + (Ex - Im)_{V/A} )</td>
</tr>
<tr>
<td>(A)</td>
<td>( = )</td>
</tr>
<tr>
<td></td>
<td>( TR_{A/H} + Ex_{A/U} + TR_{A/ST} )</td>
</tr>
<tr>
<td>Vermögensveränderungssektor</td>
<td>( S_{H/V} + S_{U/V} + S_{ST/V} )</td>
</tr>
<tr>
<td>(V)</td>
<td>( = )</td>
</tr>
<tr>
<td></td>
<td>( I^n_{V/U} + I^n_{V/ST} + (Ex - Im)_{V/A} )</td>
</tr>
</tbody>
</table>

### 5.4 Matrix-Form

Die Darstellung in Form einer Matrix ist eine spezielle Anordnung der algebraischen Ausdrücke:

<table>
<thead>
<tr>
<th>von</th>
<th>an</th>
<th>Private Haushalte (H)</th>
<th>Unternehmen (U)</th>
<th>Staat (St)</th>
<th>Ausland (A)</th>
<th>Vermögensveränderungssektor (V)</th>
<th>( \sum ) Abflüsse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private Haushalte (H)</td>
<td>-</td>
<td>( C_{H/U} )</td>
<td>( T_{\text{dir}}_{H/\text{St}} )</td>
<td>( TR_{H/A} )</td>
<td>( S_{H/V} )</td>
<td>( C_{H/U} + T_{\text{dir}}<em>{H/\text{St}} + TR</em>{H/A} + S_{H/V} )</td>
<td></td>
</tr>
<tr>
<td>Unternehmen (U)</td>
<td>( Y_{U/H} )</td>
<td>-</td>
<td>( Y_{U/\text{St}} + T_{\text{dir}}<em>{U/\text{St}} + T</em>{\text{ind}}_{U/\text{St}} )</td>
<td>( \text{Im}_{U/A} )</td>
<td>( S_{U/V} )</td>
<td>( Y_{U/H} + Y_{U/\text{St}} + T_{\text{dir}}<em>{U/\text{St}} + T</em>{\text{ind}}<em>{U/\text{St}} + \text{Im}</em>{U/A} + S_{U/V} )</td>
<td></td>
</tr>
<tr>
<td>Staat (St)</td>
<td>( Y_{\text{St/H}} + TR_{\text{St/H}} )</td>
<td>( VL_{\text{St/U}} + I^b_{\text{St}/U} + Z_{\text{St}/U} )</td>
<td>-</td>
<td>( TR_{\text{St}/A} )</td>
<td>( S_{\text{St}/V} )</td>
<td>( Y_{\text{St/H}} + TR_{\text{St/H}} + VL_{\text{St/U}} + I^b_{\text{St}/U} + Z_{\text{St}/U} + TR_{\text{St}/A} + S_{\text{St}/V} )</td>
<td></td>
</tr>
<tr>
<td>Ausland (A)</td>
<td>( TR_{A/H} )</td>
<td>( Ex_{A/U} )</td>
<td>( TR_{A/\text{St}} )</td>
<td>-</td>
<td>-</td>
<td>( TR_{A/H} + Ex_{A/U} + TR_{A/\text{St}} )</td>
<td></td>
</tr>
<tr>
<td>Vermögensveränderungssektor (V)</td>
<td>-</td>
<td>( I^n_{V/U} )</td>
<td>( I^n_{V/St} )</td>
<td>( (Ex - \text{Im})_V/A )</td>
<td>-</td>
<td>( 1^n_{V/U} + 1^n_{V/St} + (Ex - \text{Im})_V/A )</td>
<td></td>
</tr>
<tr>
<td>( \sum ) Zuflüssen</td>
<td>( Y_{U/H} + Y_{\text{St/H}} + TR_{\text{St/H}} + TR_{A/H} )</td>
<td>( C_{H/U} + VL_{\text{St}/U} + I^b_{\text{St}/U} + Z_{\text{St}/U} + Ex_{A/U} + I^n_{V/U} )</td>
<td>( T_{\text{dir}}<em>{H/\text{St}} + Y</em>{U/\text{St}} + T_{\text{dir}}<em>{U/\text{St}} + T</em>{\text{ind}}<em>{U/\text{St}} + TR</em>{A/\text{St}} + I^n_{V/St} )</td>
<td>( TR_{H/A} + \text{Im}<em>{U/A} + TR</em>{\text{St}/A} + (Ex - \text{Im})_V/A )</td>
<td>( S_{H/V} + S_{U/V} + S_{\text{St}/V} )</td>
<td>Abflüsse = Zuflüssen</td>
<td></td>
</tr>
</tbody>
</table>
Die verwendete Matrix ist im Allgemeinen quadratisch, sie enthält also gleich viele Zeilen (horizontal Reihen) wie Spalten (vertikale Reihen). Das ergibt sich daraus, dass sowohl in den Spalten als auch in den Zeilen, sämtliche Wirtschaftssektoren in der gleichen Reihenfolge aufgetragen sind. Dabei enthält dann jedes Element der Matrix den Wert einer Transaktion zwischen zwei Sektoren. Außerdem handelt es sich offensichtlich um einen geschlossenen Kreislauf, da die Bedingung der Gleichheit der Summe der Abflüsse und der Summe der Zuflüsse erfüllt ist.


6 Die Input-Output-Tabelle

6.1 Ausgangsbasis Matrix-Form

Hier sind ebenfalls in den Zeilen die Abströme und in den Spalten die Zuströme abgebildet. Genauso ergibt sich ein geschlossener Kreislauf aufgrund der entsprechend gleichen Werte in der Summenspalte und der Summenzeile.

Die Tabelle stellt in dieser Form den gesamten volkswirtschaftlichen Kreislauf dar. Für eine Input-Output-Tabelle ist jedoch nur jener Part interessant, der das gesamte Güteraufkommen einer Volkswirtschaft als auch dessen Verwendung ausdrückt. Jene Teile, die nicht direkt mit der Herstellung und Verteilung der volkswirtschaftlichen produzierten Güter zu tun haben, werden normalerweise in einer Input-Output-Tabelle nicht beschrieben. In der obenstehenden Abbildung sind die uns interessierenden Teile rot eingerahmt. Man sieht recht deutlich, dass vor allem jene Ströme, welche die Einkommensverteilung, also die direkten und indirekten Steuern sowie die Transferzahlungen, als auch die Erhöhung des Vermögensveränderungskonto, was das Sparen und den Außenbeitrag betrifft, nicht betrachtet werden.
6.2 Struktur einer Input-Output-Tabelle


<table>
<thead>
<tr>
<th>Inputs</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vorleistungen (einheimisch)</td>
<td>Vorleistungsverkäufe</td>
</tr>
<tr>
<td>von</td>
<td>an</td>
</tr>
<tr>
<td>Sektor 1</td>
<td>$X_{11}^B$</td>
</tr>
<tr>
<td>Sektor 2</td>
<td>$X_{21}^B$</td>
</tr>
<tr>
<td>Sektor 3</td>
<td>$X_{31}^B$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Sektor n</td>
<td>$X_{n1}^M$</td>
</tr>
<tr>
<td>Vorleistungen (importiert)</td>
<td>Konsumgüter</td>
</tr>
<tr>
<td>$X_{1}^M$</td>
<td>$C_1$</td>
</tr>
<tr>
<td>Indirekte Steuern abzgl. Subventionen</td>
<td>$T_{1}^{ind} - Z_1$</td>
</tr>
<tr>
<td>$I_{1}^b$</td>
<td>$I_{1}^b$</td>
</tr>
<tr>
<td>Abschreibungen</td>
<td>Exporte</td>
</tr>
<tr>
<td>$D_{1}$</td>
<td>$E_{x1}$</td>
</tr>
<tr>
<td>Löhne</td>
<td>$L_{1}$</td>
</tr>
<tr>
<td>Gewinne</td>
<td>$G_{1}$</td>
</tr>
<tr>
<td>Gesamt</td>
<td>Gesamt</td>
</tr>
</tbody>
</table>

Zwei auffällige Angelegenheiten sind wichtig. Das erste Merkmal ist die in dieser Darstellung abermals obligatorisch Übereinstimmung der beiden endgültigen...

Die zweite Eigenschaft ist um einiges interessanter. So sind nämlich im Produktionskonto jedes Sektors einmal die Vorleistungseinkäufe vom eigenen Sektor und die Verkäufe von Vorleistungen an den eigenen Sektor gleich. Dies gibt nichts anders als den Eigenverbrauch des Sektors wieder. Wie man deutlich sehen kann, sind im Produktionskonto für den Wirtschaftssektor 1 die Werte der ersten Zeile (Sektor 1, $X^D_{11}$) bei den heimische Vorleistungen und den Vorleistungsverkäufe wertmäßig gleich. Dies ist insbesondere dann bedeutsam, wenn nun auf die charakteristische Input-Output-Darstellung geschlossen werden soll. Um auf die typische Struktur zu kommen, muss die rechte Seite des Produktionskonto (Outputs) um 90° gekippt werden, sodass sich eine Matrix aufspannt. Diese muss so organisiert sein, dass im Kreuzprodukt der Input- beziehungsweise Output-Seite jedes Sektors der Eigenverbrauch steht. In der untenstehenden Abbildung ist dies für den Sektor 1 beim Wert $X^D_{11}$ der Fall. Für den zweiten Sektor würde dies dann im Feld $X^D_{22}$ zutreffen. Dieser Umstand ist im Weiteren auch bei der Betrachtung der allgemeinen Input-Output-Tabelle interessant.

\[
\begin{array}{cccccccc}
  X^D_{11} & X^D_{12} & X^D_{13} & \cdots & X^D_{1n} & C_1 & l^b_1 & E_{X_1} \\
  X^D_{21} \\
  X^D_{31} \\
  \cdots \\
  X^M_1 \\
  T^{ind}_1 - Z_1 \\
  D_1 \\
  L_1 \\
  G_1 \\
  X^M_1 
\end{array}
\]
### 6.3 Eigenschaften einer Input-Output-Tabelle

#### 6.3.1 Semantisch

Wenn nun systematisch die Gestalt der Output-Seite der Produktionskonten aller Sektoren geändert wird (nach dem obenstehenden Prinzip, 90°-Drehung), dann ergibt sich der grundsätzliche Aufbau einer Input-Output-Tabelle, welcher der Darstellung in Matrix-Form sehr ähnelt (siehe rot eingerahmte Teil der Tabelle aus Kapitel 6.1).

\[
\begin{array}{cccccccccc}
X_{11}^{D} & X_{12}^{D} & X_{13}^{D} & \cdots & X_{1n}^{D} & C_1 & I_{1}^{b} & E_{1} & X_1 \\
X_{21}^{D} & X_{22}^{D} & X_{23}^{D} & \cdots & X_{2n}^{D} & C_2 & I_{2}^{b} & E_{2} & X_2 \\
X_{31}^{D} & X_{32}^{D} & X_{33}^{D} & \cdots & X_{3n}^{D} & C_3 & I_{3}^{b} & E_{3} & X_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
X_{n1}^{D} & X_{n2}^{D} & X_{n3}^{D} & \cdots & X_{nn}^{D} & C_n & I_{n}^{b} & E_{n} & X_n \\
\end{array}
\]

Im ersten Quadranten (rot eingefärbt) sind die Vorleistungströme zwischen den n Sektoren \(X_{ij}^{D}\) abgebildet. Häufig wird die Matrix der Sekundärinputs auch etwas treffender mit Vorleistungsmatrix bezeichnet. Hier sind, wie schon bei der Matrix-Form, die Zeilen als Lieferungen der Vorleistungen an andere Sektoren zu verstehen. In den Spalten wiederum sind benötigen Vorleistungen verzeichnet. In der Tabelle wurde als letzte Zeilen beziehungsweise Spalte die Bruttoproduktion der

Der zweite Quadrant (blau eingefärbt) beschreibt die Matrix der Endnachfrage. Darin sind die Lieferungen eines Sektors für den Endverbrauch ausgewiesen. Hier sind der Konsum \((C_i)\), die (Brutto-)Investitionen \((I^b_i)\) sowie der Export \((E x_i)\) registriert. Falls es notwendig ist, kann, genauso wie es bei den Unternehmenssektoren der Fall ist, die Endnachfrage tiefer desaggregiert werden, sprich in mehrere Unterklassen aufgefächert werden.


### 6.3.2 Mathematisch

Die Zeilen der Input-Output-Tabelle geben die Verwendung des Güteraufkommens jedes Sektors an. Einerseits werden die produzierten Waren als Input für andere Sektoren aufgewendet, andererseits wandern die Erzeugnisse direkt in den Endverbrauch. Mathematisch kann das folgendermaßen definiert werden:

\[
\sum_{j=1}^{n} X^D_{ij} + C_i + I^b_i + E x_i = X_i \\
\]

\[i, j = 1, 2, 3, ..., n\]
Folgerichtig zeigen sämtliche Summen über die Sekundärinput- und die Endnachfragematrix (quasi als eine zusätzliche Summenzeile vorzustellen) die Höhe der Verwendung des Aufkommens aller Sektoren an. Folgende Gleichung kann dafür angeschrieben werden:

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij}^D + \sum_{i=1}^{n} C_i + \sum_{i=1}^{n} I_{i}^b + \sum_{i=1}^{n} E_{i} = \sum_{i=1}^{n} X_i
\]

\[i, j = 1, 2, 3, ..., n\]  

[6.2]

In den Spalten ist die Aufkommensstruktur aller Sektoren angeführt. So sind dort die volkswirtschaftlichen Einkäufe ebenso wie die Primäraufwendungen jedes Sektors verzeichnet. Die mathematische Repräsentation sieht dann so aus:

\[
\sum_{i=1}^{n} X_{ij}^D + X_{j}^M + (T_{ind}^j - Z_j) + D_j + L_j + G_j = X_j
\]

\[i, j = 1, 2, 3, ..., n\]  

[6.3]

Alle Summen über den ersten und dritten Quadranten (als Summenspalte zu sehen) spiegeln die Produktion des Aufkommens aller Sektoren wider. Nachstehende Formel zeigt dies in mathematischer Schreibweise:

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij}^D + \sum_{j=1}^{n} X_{j}^M + \sum_{j=1}^{n} (T_{ind}^j - Z_j) + \sum_{j=1}^{n} D_j + \sum_{j=1}^{n} L_j + \sum_{j=1}^{n} G_j = \sum_{j=1}^{n} X_j
\]

\[i, j = 1, 2, 3, ..., n\]  

[6.4]

Da schon bei der Betrachtung der Input-Output-Tabelle festgestellt wurde, dass ein geschlossener Kreislauf dann vorliegt, wenn der Wert der Summer einer Zeile gleich
dem Wert der entsprechenden Spaltensumme ist, muss auch hier folgende Regel gelten:

\[ X_i = X_j \quad \forall i = j \]  

[6.5]

7 Verwendetes Datenmaterial

7.1 Klassifikationen

Die Harmonisierung von verwendeten Systematiken ist naturgemäß notwendig um international vergleichbare Statistiken erstellen und nutzen zu können. Einer der vorrangigsten Aufgaben der internationalen statistischen Stellen ist es entsprechende Nomenklaturen zu erarbeiten sowie bestehende Klassifikationen zu revidieren. Vor allem die wirtschaftlichen Tätigkeiten eines Landes werden nahezu immer in Statistiken repräsentiert. Um nun diese Statistiken vergleichen zu können, ist es notwendig einheitliche Maßstäbe anzuwenden, so können durch die Anwendung identischer Klassifikationen die zu kategorisierenden Elemente geordnet und sind somit besser gegenübergestellt werden. Die überwiegende Mehrheit der Schemata verwenden alphabetische oder numerische Codes für die Einteilung.

7.1.1 ISIC

Die Klassifikation der Vereinten Nationen (UN) „international standard industrial classification of all economic activities“ (ISIC) wird auf internationaler Ebene zur Erfassung der wirtschaftlichen Tätigkeiten angewandt.


Die internationale Norm hat einen 4-stelligen Code und ist in mehrere Ebenen unterteilt. Damit diese Klassifikation weltweit angewendet werden kann, ist nicht sehr detailliert. So zum Beispiel bedeutet der Code "9" immer "Sonstige", also die Kategorie, die alle Einheiten zusammenfasst, die nicht den anderen Kategorien
derselben Ebene zugeteilt werden können. Die nachfolgende Tabelle zeigt die wirtschaftlichen Hauptgruppen, wobei jede Branche aus dem primären, sekundären und tertiärer Wirtschaftssektor einer der 17 Hauptgruppen zugeordnet wird, welche dann wiederum in Untergruppen aufgeteilt sind. Trotzdem in dieser Arbeit nicht direkt auf die ISIC-Norm zurückgegriffen wird, ist die Übereinstimmung dennoch sehr hoch, wobei dies nicht verwunderlich ist, da die europäische NACE und die davon abgeleitete österreichische Norm in starker Anlehnung an den Standard der Vereinten Nationen entwickelt wurden.

Eine vollständige Auflistung findet sich auf der Homepage der Vereinten Nationen (Literaturverzeichnis [29])

<table>
<thead>
<tr>
<th>Hauptgruppe</th>
<th>Bedeutung</th>
<th>Übersetzung</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>agriculture, hunting and forestry</td>
<td>Landwirtschaft, Jagd- und Forstwirtschaft</td>
</tr>
<tr>
<td>B</td>
<td>fishing</td>
<td>Fischerei und Fischzucht</td>
</tr>
<tr>
<td>C</td>
<td>mining and quarrying</td>
<td>Bergbau und Gewinnung von Steinen und Erden</td>
</tr>
<tr>
<td>D</td>
<td>manufacturing</td>
<td>Verarbeitendes Gewerbe</td>
</tr>
<tr>
<td>E</td>
<td>electricity, gas and water supply</td>
<td>Energie- und Wasserversorgung</td>
</tr>
<tr>
<td>F</td>
<td>construction</td>
<td>Baugewerbe</td>
</tr>
<tr>
<td>G</td>
<td>wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods</td>
<td>Groß- und Einzelhandel; Werkstätten für Kraftfahrzeuge, Krafträder, Haushaltswaren und Güter des täglichen Bedarfs</td>
</tr>
<tr>
<td>H</td>
<td>hotels and restaurants</td>
<td>Gastgewerbe = Beherbergung und Gastronomie</td>
</tr>
<tr>
<td>I</td>
<td>transport, storage and communications</td>
<td>Verkehr, Lagerhaltung und Nachrichtenwesen</td>
</tr>
<tr>
<td>J</td>
<td>financial intermediation</td>
<td>Kredit- und Finanzwesen</td>
</tr>
<tr>
<td>K</td>
<td>real estate, renting and business activities</td>
<td>Immobilienwirtschaft, Mietwesen und Betriebswirtschaft</td>
</tr>
<tr>
<td>L</td>
<td>public administration, defence and compulsory social security</td>
<td>Öffentliche Verwaltung, Verteidigung und Sozialversicherungswesen</td>
</tr>
<tr>
<td>M</td>
<td>education</td>
<td>Erziehung</td>
</tr>
</tbody>
</table>
7.1.2 NACE

Die europäische Klassifikation der wirtschaftlichen Tätigkeiten „nomenclature européenne des activités économiques“ (NACE) ist ein System zur Einordnung von Wirtschaftszweigen, welches von der Europäischen Union (EU) entworfen wurde.


| N | health and social work | Gesundheit und Pflege |
| O | other community, social and personal service activities | Sonstige öffentliche, soziale und persönliche Dienstleistungen |
| P | private households with employed persons | Private Haushalte mit Hauspersonal |
| Q | extra-territorial organizations and bodies | Exterritoriale Organisationen und Körperschaften |

7.1.3 ÖNACE

Auf nationaler Ebene wird der österreichische Ableger der EU-Norm, die ÖNACE angewandt. Sie wird vom österreichischen statistischen Amt, im weiteren Verlauf der Arbeit nur mehr „Statistik Austria“ genannt, erarbeitet und umgesetzt. Die österreichische Aktivitätsklassifikation untergliedert die europäische NACE mittels Unterklassen noch detaillierter, sodass heimische Spezifika berücksichtigt werden können.


7.1.3.1 Gliederung der Daten nach ÖNACE 2003

Dieser Abschnitt erläutert die Unterteilung der 57 ÖNACE Wirtschaftsabschnitte in 15 neue Bereiche. Da sämtliche Daten von der „Statistik Austria“ stammen, ist es notwendig, die Einteilung der Wirtschaftsdaten, welche in 57 Kategorien vorliegt, auf die Aufschlüsselung der Bildungsstufen, welche nur 15 Kategorien umfasst, zu reduzieren. Dazu wird eine kurze Beschreibung der einzelnen Sektoren gegeben. Die Erläuterungen stammen ebenfalls aus der Dokumentation der „Statistik Austria“.

Fischerei ist definiert als Nutzung der Fischereiressourcen aus dem Meer oder Binnengewässern und umfasst Fang oder Sammeln von Fischen, Krebstieren, Weichtieren und anderen Meereserzeugnissen (z.B. Perlen, Schwämme usw.).

Abschnitt B umfasst auch Fischzucht und andere Aquakulturen, die ähnliche Erzeugnisse hervorbringen, einschließlich Dienstleistungen für die Fischerei und Fischzucht.


<table>
<thead>
<tr>
<th>C</th>
<th>10</th>
<th>Kohle und Torf</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11</td>
<td>Erdöl und Erdgas, Erze</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Steine und Erden</td>
</tr>
</tbody>
</table>

Dieser Abschnitt umfasst die Gewinnung natürlich vorkommender fester (Kohle und Erze), flüssiger (Erdöl) und gasförmiger (Erdgas) mineralischer Rohstoffe. Die Gewinnung kann im Untertage- und Tagebau oder durch Bohrungen erfolgen.


Die Abteilungen, Gruppen und Klassen dieses Abschnitts sind nach dem hauptsächlich gewonnenen Rohstoff gegliedert.

Dieser Abschnitt umfasst nicht:
Verarbeitung der gewonnenen Rohstoffe (s. Abschnitt D)
Gewinnung und Flaschenabfüllung von natürlichen Quell- und Mineralwasser an Quellen und Bohrungen
Nicht im Zusammenhang mit dem Bergbau und der Gewinnung von Steinen und Erden durchgeführtes Zerkleinern, Schleifen oder anderweitiges Behandeln bestimmter Gesteine, Minerale und Erden
Gewinnung, Reinigung und Verteilung von Wasser
Erschließung von Lagerstätten
Auffinden von Erdöl-, Erdgas- und Erzlagerstätten sowie Grundwasservorkommen

<table>
<thead>
<tr>
<th></th>
<th>Nahrungs- und Futtermittel sowie Getränke</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Tabakerzeugnisse</td>
</tr>
<tr>
<td>16</td>
<td>Textilien</td>
</tr>
<tr>
<td>17</td>
<td>Bekleidung</td>
</tr>
<tr>
<td>18</td>
<td>Leder und Lederware</td>
</tr>
<tr>
<td>19</td>
<td>Holz sowie Holz-, Kork- und Flechtware</td>
</tr>
<tr>
<td>20</td>
<td>Papier, Pappe und Waren daraus</td>
</tr>
<tr>
<td>21</td>
<td>Verlags- und Druckerzeugnisse</td>
</tr>
<tr>
<td>22</td>
<td>Mineralölerzeugnisse</td>
</tr>
<tr>
<td>23</td>
<td>Chemische Erzeugnisse</td>
</tr>
<tr>
<td>24</td>
<td>Gummi- und Kunststoffware</td>
</tr>
<tr>
<td>25</td>
<td>Glas, Keramik, bearbeitete Steine und Erden</td>
</tr>
<tr>
<td>26</td>
<td>Metalle und Halbzeug daraus</td>
</tr>
<tr>
<td>27</td>
<td>Metallerzeugnisse</td>
</tr>
<tr>
<td>28</td>
<td>Maschinen</td>
</tr>
<tr>
<td>29</td>
<td>Büromaschinen, EDV-Geräte und -Einrichtungen</td>
</tr>
<tr>
<td>30</td>
<td>Geräte der Elektrizitätserzeugung und -verteilung</td>
</tr>
<tr>
<td>31</td>
<td>Nachrtechn., Rundfunk- u. FS-Geräte, elektr. Bauteile</td>
</tr>
<tr>
<td>32</td>
<td>Medizinisch-, mess-, regeltechnische u. opt. Erz.; Uhren</td>
</tr>
<tr>
<td>33</td>
<td>Kraftwagen und Kraftwagenteile</td>
</tr>
<tr>
<td>34</td>
<td>Sonstige Fahrzeuge</td>
</tr>
<tr>
<td>35</td>
<td>Möbel, Schmuck, Musikinstrumente, Sportgeräte u.a.</td>
</tr>
<tr>
<td>36</td>
<td>Dienstleistungen der Rückgewinnung</td>
</tr>
</tbody>
</table>


Bei den hier eingeordneten Einheiten handelt es sich häufig um Fabriken, Werke und Anlagen, die typischerweise Kraftmaschinen und kraftbetriebenes Förderzeug einsetzen.
Eingeschlossen sind auch Einheiten, die Stoffe manuell oder in Heimarbeiten verarbeiten, und Einheiten, die ihre Waren direkt im Betrieb verkaufen (z.B. Bäckereien, Maßschneidereien, usw.).

Waren herstellende Einheiten können Stoffe entweder selbst verarbeiten oder andere Einheiten mit der Verarbeitung ihrer Stoffe beauftragen. In beiden Fällen werden die Einheiten unter Sachgütererzeugung erfasst.

Man unterscheidet zwischen Fertigwaren für den Gebrauch oder Verbrauch und Halbwaren zur weiteren Be- oder Verarbeitung. Beispiel: Das Erzeugnis der Aluminiumraffination ist Einsatzgut für die Primärerzeugung von Aluminium, Primäraluminium ist Einsatzgut für Drahtziehereien und Aluminiumdraht ist Einsatzgut für die Herstellung von Fertigdraht.

Der Zusammenbau von Warenteilen fällt auch unter Sachgütererzeugung, sofern die Tätigkeit nicht in Abteilung 45 (Bauwesen) einzuordnen ist.


<table>
<thead>
<tr>
<th>E</th>
<th>40</th>
<th>Energie und DL der Energieversorgung</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41</td>
<td>Wasser und DL der Wasserversorgung</td>
</tr>
</tbody>
</table>


Bauinstallation umfasst die Installation aller Arten von Anlagen der Versorgungstechnik, die für die Nutzung eines Gebäudes erforderlich sind. Diese Tätigkeiten werden im Allgemeinen auf dem Baugelände ausgeführt, zum Teil können sie jedoch auch in der Werkstatt vorgenommen werden.

Sonstiger Ausbau beinhaltet Tätigkeiten, die für den Ausbau und die Fertigstellung eines Gebäudes erforderlich sind. Dazu zählen Glaserarbeiten, Putzarbeiten, Maler- und Dekorationsarbeiten, Verlegen von Bodenbelägen wie Fliesen, [...].


Verkauf ohne Weiterverarbeitung umfasst die im Handel üblichen Tätigkeiten (bzw.
Behandlungen) wie Sortieren, Klassieren und Zusammenstellen von Waren, Mischen von Waren, Abfüllen in Flaschen, [...].


**H 55 Beherbergungs- und Gaststättendienstleistungen**


Es gibt eine Überlappung der Tätigkeiten der Abteilung 55 insoweit, als der Verkauf von Getränken zum einen eine gesonderte Tätigkeit ist, zum anderen aber auch zu den Tätigkeiten von Gaststätten gehört. In gleicher Weise kann die Wirtschaftstätigkeit eines Restaurants eine gesonderte Tätigkeit darstellen oder in Verbindung mit der Beherbergung stehen.


Restaurants bieten komplett Kalte Mahlzeiten zum sofortigen Verzehr an. Dabei kann es sich um herkömmliche Restaurants, Selbstbedienungsrestaurants oder Restaurants handeln, die Speisen und Getränke zum Mitnehmen verkaufen, oder aber um feste oder mobile
Würstelstände u.Ä. mit oder ohne Sitzgelegenheiten. Entscheidend ist die Tatsache, dass Mahlzeiten zum sofortigen Verzehr angeboten werden, und nicht die Art der Einrichtung, von der sie angeboten werden.

Nicht hierunter fällt die Produktion von Mahlzeiten, die nicht zum sofortigen Verzehr geeignet oder bestimmt sind, oder von Nahrungsmitteln, die nicht als Mahlzeit angesehen werden. (siehe Abteilung 15)

| 60 | Landverkehrs- u. Transportleist. in Rohrfernleitungen |
| 61 | Schifffahrtsleistungen |
| 62 | Luftfahrtleistungen |
| 63 | DL bezüglich Hilfs- u. Nebentätigkeiten für den Verkehr |
| 64 | Nachrichtenübermittlungsdienstleistungen |

Dieser Unterabschnitt umfasst:

- Tätigkeiten im Zusammenhang mit der Personen- und Güterbeförderung im Linien- oder Gelegenheitsverkehr auf Schienen und Straßen, zu Wasser und in der Luft sowie Transport in Rohrfernleitungen
- Hilfs- und Nebentätigkeiten im Zusammenhang mit Bahnhöfen, Häfen und Flughäfen, Parkplätzen und Parkhäusern sowie Frachtumschlag, Lagerei usw.
- Post- und Fernmeldewesen
- Vermietung von Fahrzeugen mit Fahrer oder Bedienungspersonal

Dieser Unterabschnitt umfasst nicht:

- Größere Reparaturen oder Umbau von Beförderungsmitteln außer Kraftfahrzeugen (siehe Abteilung 35)
- Bau, Unterhaltung und Erneuerung von Straßen, Schienenstrecken, Häfen, Flugplätzen (siehe Abteilung 45)
- Instandhaltung und Reparatur von Kraftwagen (siehe Abteilung 50.20)
- Vermietung von Fahrzeugen ohne Fahrer oder Bedienungspersonal (siehe Abteilung 71.1 und 71.2)
Diese Abteilung umfasst:

- Hereinnahme und Ausleihung von Finanzmitteln sowie die Durchführung von anderen Bank- und Finanzgeschäften (ohne Versicherungen, Pensions- und Sterbekassen, Sozialversicherung)

Anmerkung: Nationale institutionelle Regelungen können bei der Klassifizierung der Einheiten innerhalb dieser Abteilung eine wichtige Rolle spielen.

- Lang- und kurzfristige Risikostreuung mit oder ohne Sparkomponente

- Erbringung von Dienstleistungen, die in engem Zusammenhang mit dem Kredit- und Versicherungswesen stehen, jedoch nicht von Kredit- oder Versicherungsinstituten erbracht werden


Dieser Abschnitt umfasst die Tätigkeiten, die normalerweise von der öffentlichen Verwaltung ausgeführt werden. Dabei ist der rechtliche oder institutionelle Status der Verwaltung per se nicht entscheidend. Dieser Abschnitt betrifft Einheiten, die innerhalb der zentralen oder lokalen öffentlichen Strukturen das reibungslose Funktionieren der Verwaltung des...
Gemeinwesens ermöglichen.

Somit fallen in diesen Abschnitt folgende Tätigkeiten:

Verteidigung, Rechtspflege, öffentliche Sicherheit und Ordnung, auswärtige Angelegenheiten usw.

Tätigkeiten der allgemeinen öffentlichen Verwaltung (z.B. der exekutiven und legislativen Organe, der Finanzverwaltung usw. auf zentraler, regionaler und lokaler Ebene) oder Aufsichtstätigkeiten im wirtschaftlichen und sozialen Bereich

die Verwaltung der gesetzlichen Sozialversicherung


**M 80 | Erziehungs- und Unterrichtsdienstleistungen**

Diese Abteilung umfasst:

- Öffentliches und privates Bildungswesen auf allen Stufen und für alle Berufe, auch über Rundfunk und Fernsehen
- Sonstigen Unterricht, z.B. in Fahrschulen

Sie umfasst sowohl den Unterricht in den verschiedenen Lehranstalten des regulären Schulsystems auf den verschiedenen Stufen (erster Bildungsweg) als auch Erwachsenenbildung, Alphabetisierungsprogramme usw.

Die Klassen umfassen auf jeder Stufe des ersten Bildungsweges auch den Sonderunterricht für körperlich oder geistig behinderte Schüler.

Erwachsenenbildung, die sich inhaltlich an den auf einer spezifischen Stufe vermittelten Lehrstoff anlehnt, wird dieser Stufe zugeordnet.
Der Unterricht kann in Klassenräumen, über Rundfunk und Fernsehen, Internet oder als Fernkurs erteilt werden.

**Tabelle 85**

<table>
<thead>
<tr>
<th>N</th>
<th>DL des Gesundheits-, Veterinär- und Sozialwesens</th>
</tr>
</thead>
</table>

Veterinärmedizin umfasst:

- Medizinische Versorgung und Kontrolluntersuchungen von Nutztieren
- Medizinische Versorgung und Kontrolluntersuchungen von sonstigen Haustieren
- Transport kranker Tiere

Diese Leistungen werden von qualifizierten Tierärzten in Tierarztpraxen und Tierkliniken [...] erbracht.

Sozialwesen wiederrum umfasst Sozial-, Beratungs-, Fürsorge-, Flüchtlingsbetreuungs-, Weitervermittlungs- und ähnliche Tätigkeiten, [...].

Des Weiteren zählen zum Sozialwesen auch Altersheime und sonstige Heime.

**Tabelle 90**

<table>
<thead>
<tr>
<th>O</th>
<th>Abwasser-, Abfallbeseitigungs- u. so. Entsorgungsleist.</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>DL v. Interessenvertretungen, Kirchen u.a.</td>
</tr>
<tr>
<td>92</td>
<td>Kultur-, Sport- und Unterhaltungs-DL</td>
</tr>
<tr>
<td>93</td>
<td>Sonstige Dienstleistungen</td>
</tr>
</tbody>
</table>

Diese Abteilung umfasst:

- Sammlung und Behandlung von Abfällen nicht zur Weiterverwendung in einem industriellen Fertigungsprozess, sondern mit dem Ziel der Entsorgung
- Sonstige Tätigkeiten wie Straßenreinigung, Schneebeseitigung usw.
- Tätigkeiten von Organisationen, die sich im Wesentlichen mit der Entwicklung und Förderung eines bestimmten Wirtschafts- oder Handelszweigs, einschließlich der Landwirtschaft, oder - ohne Berücksichtigung von Wirtschaftszweigen - mit dem
Wirtschaftswachstum und der wirtschaftlichen Situation eines bestimmten geografischen Gebiets oder einer bestimmten Gebietskörperschaft befassen. Hier eingeordnet sind auch Vereinigungen von solchen Verbänden. Die erbrachten Dienstleistungen betreffen vor allem die Informationsverbreitung, die Vertretung vor staatlichen Stellen, die Öffentlichkeitsarbeit und Tarifverhandlungen:
  
  o Tätigkeiten von Wirtschaftskammern, Fachverbänden oder ähnlichen Organisationen

- Film- und Videoherstellung, -verleih und -vertrieb, Lichtspieltheater (Kinos)
- Hörfunk- und Fernsehanstalten, Herstellung von Hörfunk- und Fernsehprogrammen
- Erbringung von sonstigen kulturellen und unterhaltenden Leistungen (Künstlerische und schriftstellerische Tätigkeiten und Darbietungen, Schaustellergewerbe und Vergnügungsparks), Korrespondenz- und Nachrichtenbüros sowie selbstständige Journalisten, Bibliotheken, Archive, Museen, botanische und zoologische Gärten, Sport (Betrieb von Sportanlagen, Schwimmbäder und Schwimmstadien), Erholung und Freizeit (Spiel-, Wett- und Lotteriewesen)
- Erbringung von sonstigen Dienstleistungen (Wäscherei und chemische Reinigung, Frisöre, Kosmetiker und Fußpfleger, Bestattungswesen, usw.)


Das selbst verbrauchte Produkt wird als nicht marktbestimmt erachtet und in den Volkswirtschaftlichen Gesamtrechnungen anhand der Personalkosten bewertet. Diese Dienstleistungen können nicht von Unternehmen erbracht werden.

7.1.3.2 ÖNACE 2003 vs. ÖNACE 2008

Nachdem auch in Österreich die Realität ständigen Erneuerungen unterworfen ist, ändern sich die wirtschaftlichen Tätigkeiten ebenfalls im Wandel der Zeit. So zum Beispiel führen technische Entwicklungen zu geänderten Herstellungsprozessen und

<table>
<thead>
<tr>
<th>Gliederung</th>
<th>ÖNACE 2003</th>
<th>ÖNACE 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anzahl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abschnitt</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>Unterabschnitt</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>Abteilung</td>
<td>62</td>
<td>88</td>
</tr>
<tr>
<td>Gruppe</td>
<td>224</td>
<td>272</td>
</tr>
<tr>
<td>Klasse</td>
<td>514</td>
<td>615</td>
</tr>
<tr>
<td>Unterklasse</td>
<td>722</td>
<td>701</td>
</tr>
</tbody>
</table>


Der neue Code ist fünfstellig und kann dadurch nun kompakter dargestellt werden. Dadurch ergibt sich auch der weitere Vorteil, dass eine eindeutige Abgrenzung zu

Da in dieser Arbeit das verwendete Datenmaterial aus dem Jahre 2001 beziehungsweise 2003 stammt, wird in dieser Arbeit die alte ÖNACE-Norm aus dem Jahre 2003 benutzt (siehe Kapitel 7.3).

7.2 Erwerbspersonen nach Bildung und ÖNACE 2001


Wir führen diesen Unterschied zwischen einfacher und komplizierter Arbeit, wie sie von Marx definiert wurde, nun auf die Arbeitskraft, gewissermaßen auf die hinter der Arbeit stehende Person, zurück. Dabei wollen wir herausfinden, welche Rolle das (Vor-)Wissen eines Arbeiters spielt, wobei hier nur die (Aus-)Bildung in Betracht gezogen werden kann, da das Wissen eines Menschen bekanntermaßen nur sehr schwer in einer Statistik dokumentiert werden kann. Dadurch sind wir auf das vorhandene statistische Datenmaterial angewiesen.
7.2.1 Einteilung der Bildungsstufen


7.2.1.1 Österreichische Bildungsebenen

Die Zuordnung geht auf die im regulären Bildungswesen erworbenen Abschlüsse zurück (Literaturverzeichnis [27]) Diese sind im Schulorganisationsgesetz geregelt und gliedern sich dabei in drei grobe Bildungsebenen, nämlich in die primäre, sekundäre und tertiäre Stufe (siehe Grafik in Kapitel 7.2.1.2). Dabei steht die primäre Stufe für die Pflichtschulausbildung, die beginnend bei sechs Jahre bis zum 15. Lebensjahr geht. In der Abbildung wurden der Beginn und das Ende der allgemeinen Schulpflicht auch durch die beiden rot markierten Ziffer in der Spalte „Mindestalter“ gekennzeichnet. Die sekundäre Stufe spiegelt die Ausbildung wider, die im Allgemeinen bis zur Volljährigkeit (bei regulärem Absolvieren) dauert. Unter der
tertiären Stufe, die auch postsekundäre genannt wird, versteht man weiterführende Ausbildung. Folgende neun Ausbildungsebenen wurden bei der Volkszählung unterschieden:

- Universitäten und Hochschulen
- Fachhochschulen
- Berufs- und lehrerbildende Akademien (hochschulverwandte Lehranstalten)
- Kollegs und Abiturientenlehrgänge
- Berufsbildende höhere Schule
- Allgemein bildende höhere Schule
- Berufsbildende mittlere Schule (Fachschule)
- Lehrlingsausbildung (Lehre)
- Allgemein bildende Pflichtschule

In der nachfolgenden Auflistung sind die Voraussetzungen für die jeweiligen (obenstehenden) Kategorien definiert, welche aus der Dokumentation der „Statistik Austria“ entnommen sind:

7.2.1.1.1 Universitäten und Hochschulen sowie Fachhochschulen
Diese Bildungsebene schließt alle Personen ein, die ein Studium an einer Universität, Hochschule oder Fachhochschule mit zumindest dem für die jeweilige Studienrichtung vorgesehenen Erstabschluss (akademischen Grad) abgeschlossen haben einschließlich Absolventen von Kurzstudien (Versicherungsmathematik, Datentechnik, Übersetzerausbildung, Kunstuniversitäten).

7.2.1.1.2 Berufs- und lehrerbildende Akademie
Ein Abschluss an einer dieser nun dreijährig geführten Bildungseinrichtungen führt nicht zu einem akademischen Grad. Akademien sind aber ebenso Teil des tertiären Bildungswesens wie die Universitäten und Fachhochschulen.
7.2.1.3 **Kollegs und Abiturientenlehrgänge**


7.2.1.4 **Berufsbildende höhere Schule**

Dabei handelt es sich um eine fünfjährige berufliche Ausbildung, die mit einer Reife- und Diplomprüfung abgeschlossen wird und sowohl eine bestimmte berufliche Qualifikation vermittelt als auch zum Erwerb der allgemeinen Hochschulreife führt. Dazu zählen auch Aufbaulehrgänge, die mit der Qualifikation einer berufsbildenden mittleren Schule besucht werden können. Diese Bildungsebene beinhaltet auch Berufsreifeprüfungen.

7.2.1.5 **Allgemein bildende höhere Schule**

Wie schon aus der Bezeichnung hervorgeht, vermittelt dieser Schultyp Allgemeinbildung. Abgeschlossen wird mit einer Reifeprüfung. Geführt werden Allgemeinbildende höhere Schulen (Gymnasien) entweder als achtjährige Formen (Schulstufe 5 bis 12) oder als Oberstufenformen. Eine „Beamtenmatura“ führt nicht zur allgemeinen Hochschulreife und ist daher nicht enthalten.

7.2.1.6 **Berufsbildende mittlere Schule (Fachschule)**


7.2.1.7 **Lehrlingsausbildung**

aufgelassene Lehrberufe (Pfeifenschneider, Fahrradmechaniker) als Abschlüsse dieser Ebene gewertet.

7.2.1.1.8 Allgemein bildende Pflichtschule
Diese Ebene enthält alle Personen, die keine der bisher genannten Ausbildungen absolviert haben. Sie schließt somit auch Personen ein, die außer der Pflichtschule inner- und außerbetriebliche Lehrgänge besucht haben, die aber im Schulorganisationsgesetz nicht genannt sind. Des weiteren enthält diese Kategorie auch Personen, die die Pflichtschule (Volks-, Haupt-, Sonderschule und Polytechnische Schule) nicht formal abgeschlossen haben.

7.2.1.2 ISCED


Österreichisches Bildungssystem nach ISCED
7.3 Anmerkungen

Es müssen zu den vorher erläuterten Tabellen noch einige generelle Bemerkungen gemacht werden. So wurden bei der Auswertung der Volkszählung Personen, die zu ihrer Ausbildung keine korrekten Angaben machten oder wo die Angaben fehlten, zu der Kategorie „Allgemeine Pflichtschule“ zugeordnet. Hierzu ein Auszug aus der Dokumentation der „Statistik Austria“:

„Die Daten wurden auf Konsistenz und Widerspruchsfreiheit geprüft, fehlende Angaben (4% der 15- und mehrjährigen Bevölkerung) wurden nicht nach dem Muster der anderen Personen aufgeschätzt, sondern der Kategorie „Allgemeinbildende Pflichtschule“ zugeordnet.“

Außerdem wurden mittels des Schulorganisationsgesetzes reglementiert, in welche Bildungskategorie die Erwerbstätigen bei der Auswertung der Volkszählung fallen. Hierzu ebenfalls ein Hinweis aus der Dokumentation der „Statistik Austria“:

_Ausbildungen, die nicht im Rahmen der Schulorganisationsgesetze geregelt sind, werden bei der Volkszählung nicht als „höchste abgeschlossene Ausbildung“ erfasst. Daher bilden die Daten zum Bildungsstand der Bevölkerung auch nicht jene Qualifikationen, die Personen in Kursen und Lehrgängen der beruflichen Weiterbildung, am Arbeitsplatz oder im Selbststudium erworben haben, ab. Personen, die zu ihrer Ausbildung keine Angaben gemacht haben, wurden der Kategorie „Pflichtschule“ zugeordnet._

Nähere Informationen zur Auswertung von Volkszählungen sind auf der Homepage der „Statistik Austria“ (Literaturverzeichnis [23]) und auf der Internetseite des „Bildungssystem Österreich“ (Literaturverzeichnis [27]) zu finden.

Es stellt sich die berechtigte Frage, warum in dieser Arbeit nicht die neuesten Klassifikationen verwendet werden. Dies betrifft hauptsächlich die eben eingeführte ÖNACE Revision aus dem Jahre 2008. Hierbei ist anzumerken, dass die Umsetzung dieser Norm gegenwärtig noch vollendet wird. Es existieren freilich schon Statistiken
Verwendetes Datenmaterial


Letztlich noch eine Notiz bezüglich der beiden verwendeten Ausgangstabellen (Input-Output sowie Erwerbstätige nach Ausbildung und ÖNACE). Leider passen die Daten der Bildung, die aus dem Jahre 2001 stammen, nicht zu jenen aus der verwendeten Input-Output-Tafel, die aus dem Jahre 2003 ist. Nach Rücksprache mit den zuständigen Personen bei der „Statistik Austria“ mussten wir feststellen, dass (passende) neuere Daten für die Bildung nur mit Hilfe von Hochrechnungen

8 Berechnungen

8.1 Leontief-Modell

\[
X_{ij} = a_{ij} * X_j
\]

[8.1]

Der Faktor \(a_{ij}\), der die Proportionalität widerspiegelt, zeigt die benötigten Vorleistungen eines Sektors \((X_i)\) für eine Einheit Output eines anderen Sektors \((X_j)\) an. Die Größe hat diverse Namen und wird in weiterer Folge nur mehr als **direkter Inputkoeffizient** bezeichnet – Synonyme dafür wären Materialeinsatzkoeffizient, direkte Verflechtungskoeffizienten oder auch Sekundärinputkoeffizienten. Nun lässt sich die Gleichung [6.1] folgendermaßen umformen:
\[
\sum_{j=1}^{n} a_{ij} \cdot X_j + Y_i = X_i \quad \text{[8.2]}
\]

Dabei werden die Vorleistungsströme \(X_{ij}^D\) durch die direkten Inputkoeffizienten \((a_{ij})\) und die Summen der entsprechenden Spalte \((X_j)\) substituiert:

\[
\sum_{j=1}^{n} X_{ij}^D = \sum_{j=1}^{n} a_{ij} \cdot X_j \quad \text{[8.3]}
\]

Die Endnachfrage \((Y_i)\) setzt sich aus dem Konsum \((C_i)\), den (Brutto-)Investitionen \((I_{bi})\) sowie dem Export \((E X_i)\) zusammen:

\[
C_i + I_{bi} + E X_i = Y_i \quad \text{[8.4]}
\]


\[
a_{ij} = \frac{X_{ij}}{X_j} \quad \text{[8.5]}
\]
Sind schlussendlich alle Koeffizienten berechnet, sieht die Matrix, welche mit $A$ bezeichnet wird, wie folgt aus:

$A = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} & & a_{1n} \\
  a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
  a_{31} & a_{32} & a_{33} & & a_{3n} \\
  \vdots & & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}
\end{bmatrix}$

Mit Hilfe der berechneten Matrix $A$ lässt sich somit die Gleichung [8.2] in folgender Weise anschreiben:

$A \cdot x + y = x$  \ [8.7]

Wenn die Matrix $A$ mit Sicherheit bestimmt werden kann, so können im Weiteren zwei grundsätzliche Aufgabenstellungen unterschieden werden:

1. Bei gegebenem Vektor der Endnachfrage soll der Output der Sektoren bestimmt werden.
2. Bei gegebenem Output soll die Endnachfrage der Sektoren bestimmt werden.
Für beide Fälle lässt sich die Gleichung [8.7] so umformen, dass ein lineares Gleichungssystem mit n Gleichungen und n Unbekannten zu lösen ist:

1. Für die erste Aufgabe ergibt sich somit:

   \[ x = y \cdot (E - A)^{-1} \]  \[ \text{[8.8]} \]

2. Die zweite Aufgabe lässt sich so lösen:

   \[ y = (E - A) \cdot x \]  \[ \text{[8.9]} \]

Die Einheits- oder auch Identitätsmatrix \( E \) stellt eine quadratische Matrix dar, deren Hauptdiagonale nur aus Einen besteht und deren andere Elemente Null sind.

Das Gleichungssystem [8.8] wird als **offenes statistisches Leontief-Modell** bezeichnet. Die darin enthaltene Matrix \( (E - A) \) beziehungsweise \( (E - A)^{-1} \) wird gewöhnlich **Leontief-Matrix** sowie **Leontief-Inverse** genannt.

### 8.2 Regression


Die Regressionsanalyse kann auf verschiedene Weisen eingesetzt werden. Die wichtigsten sind im Folgenden näher erklärt:

- **Erkennen von Zusammenhängen**

  Es wird ein funktionaler Zusammenhang zwischen zwei bisher nicht untersuchten Merkmalen gefunden. So zum Beispiel möchte man wissen, ob und wie sehr sich die Einführung eines neuen Produktionsablaufs auf die Zufriedenheit der Mitarbeiter auswirkt.
• Nachweis von Zusammenhängen

Die Mutmaßung, dass sich mit dem Anstieg des Alters auch die menschliche Reaktion vergrößert, kann mittels Regression nachgewiesen und außerdem mit konkreten Ergebnissen gestützt werden.

• Schätzung der Art und Größe von Zusammenhängen

Mit Hilfe der Regressionsanalyse lässt sich die Art und die Größe eines bekannten Zusammenhangs schätzen. So zum Beispiel ist jedem bewusst, dass sich bei einer höheren Geschwindigkeit auch der Bremsweg verlängert – für die Berechnung des Verhältnisses kann die Regression herangezogen werden.

In der Arbeit ist vor allem dieser Punkt interessant, da die berechneten Ergebnisse den gleichen semantischen Wert widerspiegeln und somit ausnahmslos in Zusammenhang zum originalen Wert stehen.

• Prognose fehlender oder künftiger Werte

Für viele Wissenschaftler, die mit Simulationsmodelle arbeiten, besteht ein großes Interesse daran, die vorhandenen Daten hochzurechnen um damit nicht existente Werte vorherzusagen. Dies gilt vor allem für die Deutung der Zukunft. So kann die Entwicklung der Wirtschaft oder der Bevölkerung basierend auf vorhandenen Daten geschätzt werden.
8.2.1 Lineare Regression
Ein linearer Zusammenhang wird dadurch erkannt, dass der Zusammenhang zwischen den beiden Merkmalen durch eine gerade Linie beschrieben werden kann.

Das lineare Regressionsmodell besteht grundsätzlich aus drei Komponenten:

1. einem Merkmal $X$ – der unabhängigen Variable
2. einem Merkmal $Y$ – der abhängigen Variable
3. der Annahme $Y = f(X) + \varepsilon$ – der vermutete funktionale Zusammenhang zwischen den beiden Merkmalen mit dem Fehlerterm $\varepsilon$

8.2.1.1 Einfache lineare Regression
Für die Berechnung der einfachen linearen Regression muss die obigen Annahme $f(X)$ so ersetzt werden, dass alle Datenpunkte durch eine Gerade mit der bekannten Form

$$Y = a + b \times X + \varepsilon$$

[8.10]

approximiert werden können, wobei $a$ den Schnittpunkt, $b$ die Steigung und $\varepsilon$ wiederum den Fehlerterm bezeichnet. Der Fehlerterm ist notwendig, da die einzelnen Datenpunkte im Grunde nie genau auf der Geraden liegen und dadurch eine Streuung um den Funktionsgraphen auftritt. Diese Streuung wird mit Hilfe des Fehlerterms ausgeglichen.
Für die Regression ist es nun entscheidend wie die Gleichung der Geraden gefunden werden kann. Es lässt sich nachweisen, dass die optimale Gerade folgende Bedingung hat:

\[
\min \sum \varepsilon_i^2 = \min \sum (y_i - \hat{y}_i)^2 = \min \sum (y_i - (a + b \cdot x_i))^2 \quad \text{[8.11]}
\]

Diese Formulierung löst die meisten Probleme, die bei anderen Ansätzen auftreten und stellte sich daher als die brauchbarste Lösung heraus. Wichtig ist unter anderem, dass die Gerade sich eindeutig berechnen lässt, also genau eine Linie als Lösung existiert. Außerdem verläuft die Gerade in einer Weise, sodass sie möglichst nahe an allen Punkten liegt, was aus der Ausgangsbedingung hervorgeht. Eine
weitere interessante Eigenschaft ist, dass sie stets durch den Schwerpunkt \((\bar{X}, \bar{Y})\) geht, wobei \(\bar{X}\) und \(\bar{Y}\) das arithmetische Mittel (den Mittelwert, vgl. [8.47]) der jeweiligen Merkmale ausdrückt. Zudem ist die Summe der Abstände aller Punkte von der Regressionsgeraden immer Null und dadurch wird sie auch „fehlerausgleichende Gerade“ genannt. Mathematisch ist diese Charakteristik derart definiert:

\[
\sum_{i=1}^{n} \varepsilon_i = 0 \quad [8.12]
\]

Um die Bedingung [8.11] zu erfüllen, wird herkömmlicherweise die Minimum-Quadrat-Methode – oder auch Methode der kleinsten Quadrate – eingesetzt. Mathematisch lässt sich das Minimierungsproblem auf folgende Weise formulieren:

\[
L(\hat{a}, \hat{b}) = \min(L(a, b): a \in \mathbb{R}, b \in \mathbb{R}) \quad [8.13]
\]

Der erste Schritt wird durch partielles Differenzieren und Nullsetzen der Ableitungen erster Ordnung berechnet:

\[
\frac{\partial L}{\partial a}(\hat{a}, \hat{b}) = -2 \sum_{i=1}^{n} \left( y_i - (\hat{a} + \hat{b} \cdot x_i) \right) = 0 \quad [8.14]
\]

\[
\frac{\partial L}{\partial b}(\hat{a}, \hat{b}) = -2 \sum_{i=1}^{n} x_i \cdot \left( y_i - (\hat{a} + \hat{b} \cdot x_i) \right) = 0 \quad [8.15]
\]
Nun wird durch Ausrechnen der Klammern und anderes Anordnen der Terme daraus ein System von Normalgleichungen aufgestellt:

\[
\sum_{i=1}^{n} y_i = \hat{a} \cdot n + \hat{b} \cdot \sum_{i=1}^{n} x_i
\]  

[8.16]

\[
\sum_{i=1}^{n} x_i y_i = \hat{a} \cdot \sum_{i=1}^{n} x_i + \hat{b} \cdot \sum_{i=1}^{n} x_i^2
\]  

[8.17]

Die Lösung für die beiden linearen Gleichungssysteme sieht dann folgendermaßen aus, wobei die letzten beiden Formeln [8.19] komplett gleich sind und nur für die Lesbarkeit verschiedenartig dargestellt sind:

\[
\hat{a} = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} y_i\right) - \hat{b} \cdot \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right)
\]  

[8.18]

\[
\hat{a} = \bar{y} - \hat{b} \cdot \bar{x}
\]

\[
\hat{b} = n \cdot \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i \ igg/ \ n \cdot \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2
\]  

[8.19]

Die beiden Gleichungen aus [8.18] sind äquivalent, da die beiden Variablen \(\bar{y}\) und \(\bar{x}\) sowie die geklammerten Summenterme das Selbe ausdrücken, nämlich das jeweilige arithmetische Mittel (den Mittelwert, vgl. [8.47]) – wie man anhand der Substitution der Klammern in der Formel leicht erkennen kann.
Das Ergebnis der Regression ist die geschätzte Regressionsgerade mit der Form:

$$y = \hat{a} + \hat{b} \cdot x$$  \[8.20\]

Die einzelnen prognostizierten Werte ($\hat{y}_i$) lassen sich dann durch die unten angeführte Ausdruck ermitteln:

$$\hat{y}_i = \hat{a} + \hat{b} \cdot x_i$$  \[8.21\]

Der Fehler eines Datenpunktes ($\varepsilon_i$) heißt Residuum und ergibt sich aus der Differenz des tatsächlichen ($y_i$) und des geschätzten Wertes ($\hat{y}_i$) (vgl. [8.11]):

$$\varepsilon_i = y_i - \hat{y}_i$$  \[8.22\]

Mit diesen beiden Gleichungen würden sich Abhängigkeiten bereits hervorragend schätzen lassen. In dieser Arbeit jedoch findet vorwiegende Berechnungen mit Matrizen und Vektoren Verwendung. Somit ist es ein logischer Schritt die Formeln umzubauen und für eine bessere numerische Handhabung in Matrix-Form zu bringen. Dazu werden als erstes die Variablen als Vektoren beziehungsweise Matrizen angeschrieben, wobei $\alpha$ die Faktoren der linearen Regression, also den Schnittpunkt $a$ und die Steigung $b$, repräsentiert:
Die Relation von Daten zu Modell aus [8.10] kann dann in Matrixform wie folgt notiert werden:

\[
y = X \alpha + \varepsilon \equiv \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix} \]

Um in weiterer Folge auf die Normalgleichung zu kommen, werden vorerst die beiden Gleichungen [8.16] und [8.17] zeilenweise in einen Matrix eingetragen:

\[
\begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \]

[8.25]
Die linke Seite der Matrix kann mit Hilfe von $y$ und $X$ aus den Vorgaben von [8.23] erzeugt werden, wobei $X^T$ die transponierte Matrix darstellt:

$$X^T \cdot y =$$

\[
\begin{bmatrix}
1 & 1 & 1 & \ldots & 1 \\
x_1 & x_2 & x_3 & \ldots & x_n
\end{bmatrix}^T \cdot 
\begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
\vdots \\
y_n
\end{bmatrix} = 
\begin{bmatrix}
\sum_{i=1}^n y_i \\
\sum_{i=1}^n x_i y_i
\end{bmatrix} \tag{8.26}
\]

Demgegenüber kann der unbekannte Teil der rechten Seite sogar nur mit $X$ aus den Vorgaben von [8.23] generiert werden, wobei $X^T$ wiederum die transponierte Matrix darstellt:

$$X^T \cdot X =$$

\[
\begin{bmatrix}
1 & 1 & 1 & \ldots & 1 \\
x_1 & x_2 & x_3 & \ldots & x_n
\end{bmatrix}^T \cdot 
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
1 & x_3 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix} = 
\begin{bmatrix}
n \sum_{i=1}^n x_i \\
n \sum_{i=1}^n x_i \sum_{i=1}^n x_i
\end{bmatrix} \tag{8.27}
\]
Mit diesen beiden Ersetzungen kann schließlich die Normalgleichung für die Matrix-Schreibweise konstruiert werden:

\[ X^T * y = X^T * X * \hat{\alpha} \]  \[8.28\]

Daraus folgt nachstehende Formel zeigt die Lösung der Normalgleichung. Sie ist durch simples Umformen zu erreichen, wobei \((...)^{-1}\) die inverse Matrix darstellt:

\[ \hat{\alpha} = (X^T * X)^{-1} * (X^T * y) \]  \[8.29\]

Die prognostizierten Werte \((\hat{y})\) und die Residuen \((\varepsilon)\) sind genauso zu kalkulieren wie bei der Berechnung ohne Matrizen:

\[ \hat{y} = X * \hat{\alpha} \]  \[8.30\]
\[ \varepsilon = y - \hat{y} \]  \[8.31\]

### 8.2.1.2 Multiple lineare Regression

Die multiple lineare Regression erweitert nun das Modell der einfachen linearen Regression und wird auch mehrfache oder multivariate Regression genannt. Sie wird in dieser Arbeit vor allem zur Berechnung mit unterschiedlichen Bildungsstufen dienen. Dabei ist das abhängige Merkmal \(Y\) nicht nur von einer sondern von mehreren Regressionsvariablen abhängig. Der funktionale Zusammenhang soll in der Annahme \(Y = f(X) + \varepsilon\) wiederum in einer Art ersetzt werden, sodass alle Datenpunkte durch eine Gerade angenähert werden können.
Die Form der Geraden sieht bei der multiplen linearen Regression wie folgt aus:

\[ Y = \alpha_0 + \alpha_1 \cdot X_1 + \alpha_2 \cdot X_2 + \alpha_3 \cdot X_3 + \ldots + \alpha_n \cdot X_n + \varepsilon \]  

[8.32]

Die Lösung der geschätzten Koeffizienten wird ebenfalls durch die Minimum-Quadrate-Methode gewonnen:

\[ L(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n) = \min \sum_{i=1}^{n} \varepsilon_i^2 \]  

[8.33]

Dafür werden die Matrix der Ausgangsdaten \( X \) und der Vektor der Koeffizienten \( \alpha \) aus [8.23] folgendermaßen erweitert (der Ergebnisvektor \( y \) und die Residuen \( \varepsilon \) bleiben gleich):

\[
y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}; \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_n \end{pmatrix}; \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \]  

[8.34]
Das Modell zu den Daten bleibt gleich wie bei der einfachen linearen Regression ([8.24]). Die Koeffizienten mit der besten Anpassung werden ebenso in identischer Weise ermittelt ([8.29]) wie es für die Werte der Prognose ([8.30]) und der Residuen ([8.31]) zutrifft, wobei $X^T$ die transponierte und $(...)^{-1}$ die inverse Matrix ist:

$$y = X \alpha + \epsilon \quad [8.35]$$

$$\hat{\alpha} = (X^T \cdot X)^{-1} \cdot (X^T \cdot y) \quad [8.36]$$

$$\hat{y} = X \cdot \hat{\alpha} \quad [8.37]$$

$$\varepsilon = y - \hat{y} \quad [8.38]$$

Abschließend kann festgehalten werden, dass die einfache und die multiple lineare Regression sich bei der Berechnung durch Matrizen nicht unterscheidet. Dies ist ein großer Vorteil, da in dieser Arbeit eigentlich nur mit Matrizen gerechnet wird und somit keine Umstellung zwischen den diversen Berechnungen vorgenommen werden muss.
8.2.2 Gütekriterien der Regression

8.2.2.1 Regressionsstatistik
8.2.2.1.1 Korrelationskoeffizient
Der erste Parameter wird meistens nur zusätzlich angegeben, da er sich recht leicht aus dem Bestimmtheitsmaß, welches aber mehr Aussagekraft besitzt, berechnen lässt:

\[ R = \sqrt{R^2} \]  \[8.39\]

Dabei bezeichnet \( R \) den Absolutbetrags des empirischen Korrelationskoeffizienten von \( X \) und \( Y \). Die Berechnung in dieser Weise ist aber nicht glücklich gewählt, da die Wurzel immer positiv ist, der Korrelationskoeffizient aber laut Definition Werte zwischen 1 und -1 annehmen kann. Aufgrund dessen wird der Faktor eigentlich immer direkt kalkuliert:

\[ R(X,Y) = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2}} \]  \[8.40\]
Ist der Faktor \( R(X, Y) = 1 \) (bzw. \(-1\)) besteht ein vollständiger positiver (oder negativer) linearer Zusammenhang zwischen den beiden beobachteten Merkmalen. Falls der Korrelationskoeffizient Null ergibt, dann ist überhaupt kein linearer Zusammenhang gegeben. Für die Bestimmung der allgemeinen Güte der Regression wird aber eher das Bestimmtheitsmaß herangezogen.

### 8.2.2.1.2 Bestimmtheitsmaß

Nachdem der Wert für den Korrelationskoeffizienten ermittelt worden ist, kann in weiterer Folge die geforderte Güte des linearen Modells erforscht werden. Eines der wichtigsten Maße der Regression ist das **Bestimmtheitsmaß** \( R^2 \). Es kann einerseits durch den quadrierten Korrelationskoeffizienten \( R \) ausgedrückt werden. Andererseits ist es jedoch (semantisch gesehen) das Verhältnis von erklärter zu gesamter Streuung (Quadratsumme der Regression zu Gesamtvariabilität) und wird daher im Allgemeinen mittels der Quadratsummen (siehe Kapitel 8.2.2.2.2) berechnet:

\[
R^2 = \frac{SS_R}{S_{yy}} \quad [8.41]
\]

Je näher die Division bei Eins liegt, desto geringer ist der unerklärte Fehler und desto besser ist die Anpassung der Geraden an die Punktwolke der Daten. Der Grenzfall tritt ein, wenn die Geraden exakt angepasst ist und durch sämtliche Datenpunkt verläuft. In diesem Fall wäre kein Fehler vorhanden \( (SS_E = 0) \) und somit der erklärte Anteil der Regression 100 Prozent \( (SS_R = 1) \), was dann ebenfalls einem 100-prozentigen Bestimmtheitsmaß \( (R^2 = 1) \) entspräche.

### 8.2.2.1.3 Standardfehler der Regression

Das Regressionsmodell, das aus einer Anzahl an Stichproben ermittelt wird (in dieser Arbeit werden immer sämtliche gemessenen Werte verwendet), ist höchst wahrscheinlich nicht das optimale Modell für den gesamten Wertebereich und somit für auch nicht für prognostizierte Werte. Bei anderen Stickprobenmessungen, zum
Beispiel durch Hinzunehmen von weiteren Messwerten oder auch Verwenden von anderen Stichproben, würde das Ergebnis ein anderes Modell liefern. Die Abweichungen zwischen den Stichprobenwerten und dem Stichprobenmodell lässt sich berechnen. Die Abweichung zwischen Stichprobenwerten und dem Modell für den gesamten Wertebereich hingegen lässt sich nur schätzen. Der **Standardfehler der Regression** ($S_{\text{Regression}}$) ist nun die Schätzung der Streuung der Stichprobenwerte um das vorhergesagte Modell der gesamten Werte. Die Berechnung sieht folgendermaßen aus:

$$
S_{\text{Regression}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{(n - m)}}
$$

Dabei spiegelt $m$ die Anzahl der Koeffizienten aus der Gleichung der Geraden [8.32] und $n$ die Anzahl der Stichproben (in dieser Arbeit immer die Zahl der gemessenen Werte) wider (siehe auch „Hinweis“ im Kapitel 8.2.2.2.1).

### 8.2.2.1.4 Beobachtungen
Unter diesem Punkt wird die Anzahl der tatsächlichen Beobachtungen oder der gemessenen Werte ausgewiesen. Dieser Wert wird unter anderem bei der Berechnung der Freiheitsgrade (siehe Kapitel 8.2.2.2.1) benötigt.

### 8.2.2.2 ANOVA (analysis of variance)
#### 8.2.2.2.1 Freiheitsgrade
Die Freiheitsgrade der Regression ($df = \text{degree of freedom}$) werden zur Berechnung andere Werte verwendet, so zum Beispiel für die f-Statistik (siehe Kapitel 8.2.2.4) genauso wie für die t-Statistik (siehe Kapitel 8.2.4.1).
Die verschiedenen Freiheitsgrade werden mit den nachstehenden Formeln berechnet:

\[ df_{Regression} = m - 1 \] \hspace{1cm} [8.43]

\[ df_{Residuen} = n - m \] \hspace{1cm} [8.44]

\[ df_{gesamt} = n - 1 \] \hspace{1cm} [8.45]

Dabei spiegelt \( m \) die Anzahl der Koeffizienten aus der Gleichung der Geraden [8.32] und \( n \) die Anzahl der Stichproben (in dieser Arbeit immer die Zahl der gemessenen Werte) wider.

**Hinweis:**

Oft werden die Freiheitsgrade als Werte im Nenner verwendet. So zum Beispiel bei der Berechnung der Standardfehler \((n - m)\) oder der Prüfgröße F \((m - 1\) sowie \(n - m\)). Der Freiheitsgrad bringt dort Zahl der Koeffizienten \((m)\) aus der Gleichung der Geraden [8.32] mit in die Berechnung ein. Dies ist wichtig, da bei einer kleinen Stichprobe \((n)\) die Repräsentation für die gesamten Schätzdaten schlechter ausfällt und auch der geschätzte Fehler gegenüber dem Fehlermittel der Stichprobe dementsprechend größer ist. Indessen fällt dann die Modellkomplexität \((m)\) stärker ins Gewicht.

### 8.2.2.2 Quadratsummen

Ein weiterer sehr wichtiger Parameter ist die **totale quadratische Abweichung** oder **Gesamtvariabilität** \((S_{y|y})\). Sie zeigt die allgemeine Qualität der gemessenen Daten und dadurch auch die richtige Auswahl der Messwerte. Sie wird durch die Abweichung der Messwerte vom entsprechenden arithmetischen Mittel \((\bar{y})\) berechnet.
Die Formel sieht dann wie folgt aus:

\[ S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 \]  
\[ \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \]

Damit kann zwar festgestellt werden ob eine Analyse mittels Regression generell sinnvoll ist und ob überhaupt ein funktionaler Zusammenhang besteht, aber über die Güte der geschätzten und prognostizierten Werte kann weiterhin keine vernünftige Aussage getroffen werden. Um diese Angelegenheit zu lösen, unterteilt man die Gesamtvariabilität in zwei Komponenten:

\[ S_{yy} = SS_R + SS_E \]

\[ \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

Der Beweis für die Gültigkeit dieser Aufteilung ist der stochastischen Fachliteratur zu entnehmen.
Grafisch sieht die Aufteilung jedenfalls wie folgt aus:

[Diagramm]

Dabei wird die gesamte Streuung einerseits in den Teil „Quadratsumme der Regression“ ($SS_R = \text{sum of squares - regression}$) zerlegt. Dies ist jener Teil, der durch die Regression erklärt wird, also die beschriebene Datenvariabilität, und sollte idealerweise den Großteil der Abweichungen ausmachen. Dabei wird der geschätzte Wert ($\hat{y}_i$) mit dem arithmetischen Mittel (dem Mittelwert, vgl. [8.47]) der Messung ($\bar{y}$) verglichen:

$$SS_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$  [8.49]
Der andere Teil wird „Quadratsumme der Fehler“ genannt ($SS_E = \text{sum of squares - error}$) und ist der eigentlich Ausgangspunkt um die Regressionsgerade aufzustellen (vgl. [8.11]). Deshalb spiegelt das Residuum ($\epsilon_i$) auch den unbekannten Fehlerterm wider, was wiederum bedeutet, dass diese Abweichung nicht durch die Regressionsanalyse erklärt werden kann. Unter optimalen Bedingungen sollte die Restvariabilität gegen Null streben. Das Ergebnis erhält man, indem die Differenz des gemessenen und des geschätzten Wertes berechnet wird:

$$SS_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$ [8.50]

### 8.2.2.2.3 Mittlere Quadratsummen

Die mittlere Quadratsumme ($MS$) soll den durchschnittliche Fehler ausdrücken und wir sowohl für die erklärte als auch für die nicht erklärte Abweichung ($SS$) berechnet. Es wird als Quotient der Mittelung über die Freiheitsgrade ($df$) dargestellt:

$$MS = \frac{SS}{df}$$ [8.51]

### 8.2.2.4 F-Statistik

Mit Hilfe der F-Statistik kann geprüft werden, ob die Werte für das Bestimmtheitsmaß ($R^2$) rein zufällig sind oder der angegebene funktionale Zusammenhang mit einer gewissen Sicherheit zutrifft. Um die Signifikanz des gesamten Modells zu überprüfen, muss zwischen zwei Hypothesen unterschieden werden:

1. $H_0$: Die durch die Regression beschriebene Beziehung besteht nur zufällig in dieser einen Stichprobe, gilt aber nicht für die gesamte Schätzung.
2. $H_1$: Zwischen den geschätzten und gemessenen Werten besteht die durch die Regression beschriebene Beziehung.
Für den Signifikanztest eines Regressionsmodells wird die F-Statistik gebraucht. Das Ergebnis dieses F-Tests wird als die Prüfgröße $F$ bezeichnet und wie folgt berechnet:

\[
F = \frac{\frac{R^2}{(m - 1)}}{\frac{1 - R^2}{(n - m)}} = \frac{R^2 * (n - m)}{(m - 1) * (1 - R^2)} \quad [8.52]
\]

Dabei spiegelt $m$ die Anzahl der Koeffizienten aus der Gleichung der Geraden [8.32] und $n$ die Anzahl der Stichproben (in dieser Arbeit immer die Zahl der gemessenen Werte) wider (siehe auch „Hinweis“ im Kapitel 8.2.2.2.1).

Für die einfache lineare Regression, also dem Fall, dass $Y$ außer vom Schnittpunkt nur von einem Parameter abhängt, lässt sich der F-Wert auch anhand der T-Statistik berechnen (siehe Kapitel 8.2.2.4.1):

\[
F = T_b^2 = \left(\frac{b}{S_b}\right)^2 \quad [8.53]
\]

Dabei ist $T_b$ und $S_b$ der T-Wert sowie der Standardfehler des einzigen Koeffizienten.

Der F-Wert alleine reicht aber nicht aus um eine Aussage über die Regression zu machen. Deshalb hat man nun zwei Möglichkeiten. Entweder berechnet man sich passenden zum F-Wert den P-Wert ($P_F$) oder vergleicht den F-Wert mit einer selbst gewählten Grenzschwelle, dem kritischen F-Wert ($F_{kritisch}$). Die F-Verteilung würde folgendermaßen aussehen:
Der **P-Wert** gibt die Wahrscheinlichkeit an, einen Wert, der größer als die Prüfgröße $F$ ist, aus den Stichprobendaten zu erhalten. Der Wert wird mit Hilfe des absoluten F-Wertes und der Freiheitsgrade bestimmt und in der Praxis meistens aus einer statistischen Tabelle ausgelesen (siehe auch 8.2.3). Er lässt sich aber auch als Fläche unter der Dichtefunktion $D$ darstellen, wobei dann der kritische und der berechnete F-Wert zusammenfallen:

$$P_F = \int_{F}^{\infty} D \quad [8.54]$$

Die Gegenwahrscheinlichkeit $1 - P_F$ stellt das Siginifikanzniveau dar. Je näher die berechnete Wahrscheinlichkeit (die Fläche des P-Wertes) zu Null tendiert, desto geringer ist die Chance ein falsches $H_0$ zu akzeptieren und desto gewisser kann daraus geschlossen werden, dass die Höhe des F-Wertes ziemlich sicher nicht
zufällig in diesem Maße ausfällt. Diese Aussage stützt damit folglich auch die Hypothese $H_1$ und infolgedessen werden auch das Bestimmtheitsmaß sowie der generelle funktionale Zusammenhang weiter untermauert.

Um den kritischen Wert von $F$ zu bestimmen, wird ein Signifikanzniveau festgelegt, ab dem die Hypothese $H_1$ akzeptiert wird. Die meisten Berechnungen gehen von einem $\alpha$-Quantil von 0,05 aus, was einer Sicherheit von 95% ($1-\alpha$) entspricht. Das Restrisiko von 5% sagt aus, dass der gefundene F-Wert auch aus der abgelehnten, aber vermeintlich wahren Hypothese $H_0$ stammt und somit eigentlich kein Zusammenhang zwischen den Merkmalen $X$ und $Y$ besteht. Dies kann angesichts der geringen Fehlerwahrscheinlichkeit in Kauf genommen werden. Zudem wird in dieser Arbeit das $\alpha$-Quantil sogar auf unter 0,01 gedrückt und das Signifikanzniveau somit fast auf eine Stufe gänzlicher Sicherheit gebracht. Der kritische F-Wert wird in der Praxis unter Zuhilfenahme der Freiheitsgrade der Regression und der Residuen ebenfalls aus einer F-Verteilungstabelle ausgesogen (siehe auch 8.2.3). Ist nun das Ergebnis des F-Tests größer als der ausgelesene Grenzwert, kann gefolgt werden, dass die Hypothese $H_0$ – dass kein Zusammenhang in der gesamten Schätzung besteht – ungültig ist und verworfen werden kann. Übersteigt der F-Wert das kritische Niveau von $F$ (bei entsprechend hoher Wahrscheinlichkeit), dann weist dies höchst zuverlässig auf einen funktionalen Zusammenhang hin. Normalerweise wird der P-Wert berechnet, da er eine anschaulichere Aussagekraft besitzt. Jedoch kann mit dem kritischen F-Wert eine bessere Vergleichbarkeit erzielt werden, da bei verschiedenen Berechnungen der F-Wert jeweils gegen die gleiche Wahrscheinlichkeit getestet werden kann.

8.2.2.3 Regressionskoeffizienten
8.2.2.4 Standardfehler der Regressionskoeffizienten

Der Standardfehler der Regression (siehe Kapitel 8.2.2.1.3) beschäftigt sich mit der Abweichung aller Werte im gesamten Vorhersagemodell. Es lässt sich aber auch für jeden einzelnen Koeffizienten ein Standardfehler berechnen. Der **Standardfehler der Regressionskoeffizienten** ($S_{\hat{\alpha}_i}$) ist nun die Schätzung der Streuung jedes einzelnen Stichprobenwertes um den vorhergesagten Wert im Gesamtmodell.
Die Berechnung sieht folgendermaßen aus:

\[ S_{\alpha_i} = S_{Regression} \sqrt{x_{ii}} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n - m)} \cdot \sqrt{x_{ii}} \quad \text{[8.55]} \]

Dabei sind \( x_{ii} \) die Elemente der Hauptdiagonal der Matrix \( \hat{\alpha} \) aus der Gleichung [8.29] oder [8.36]. Außerdem sind in der Standardabweichung der Regression (\( S_{Regression} \)) noch die Anzahl der Koeffizienten (\( m \)) aus der Gleichung der Geraden [8.32] und die Anzahl der Stichproben (\( n \)) (in dieser Arbeit immer die Zahl der gemessenen Werte) enthalten (siehe auch „Hinweis“ im Kapitel 8.2.2.2.1).

Eine wichtiger Zusammenhang ist noch festzuhalten: Je größer der Standardfehler eines Regressionskoeffizienten ist, desto breiter wird auch das entsprechende Konfidenzintervall ausfallen (siehe auch 8.2.2.4.2).

### 8.2.2.4.1 T-Statistik

Nachdem festgestellt wurde, dass das Regressionsmodell signifikant ist, können die einzelnen Koeffizienten auf ihre Signifikanz getestet werden. Analog zur F-Statistik muss auch hier eine Unterscheidung zwischen zwei Hypothesen gemacht werden:

1. \( H_0 \): Die von \( \alpha_i \) gewichtete Variable spielt für den Zusammenhang der Gesamtschätzung keine Rolle.
2. \( H_1 \): Die von \( \alpha_i \) gewichtete Variable ist wichtig für den Zusammenhang in der Gesamtschätzung.

Für den Signifikanztest der einzelnen Koeffizienten wird die T-Statistik verwendet. Das Ergebnis dieses T-Tests wird als die Prüfgröße \( T \) bezeichnet und spiegelt das Verhältnis zwischen dem Koeffizienten (\( \alpha_i \)) und dem entsprechenden Standardfehler (\( S_{\alpha_i} \)) wider.
Als Formel drückt sich dieser Sachverhalt so aus:

\[ T_{\alpha_i} = \frac{\alpha_i}{s_{\alpha_i}} \]  

[8.56]

Wenn der Betrag des Prüfwertes \( T \) hoch genug ist, kann geschlussfolgert werden, dass der Koeffizient für die Berechnung des geschätzten Wertes hilfreich ist. Die übliche Interpretation geht davon aus, dass ab einem Quotient von mindestens 1,96 – der Koeffizienten also doppelt so groß wie der Standardfehler ist – die Abweichungen so gering sind, dass man den Koeffizienten als signifikant betrachtet.

Obwohl der T-Wert schon eine gute Bewertung vornimmt, kann trotzdem wie bei der F-Statistik ein P-Wert \( (P_T) \) oder alternativ ein kritischer T-Wert \( (T_{kritisch}) \) in Betracht gezogen werden. Die T-Verteilung würde folgende Form haben:
Der **P-Wert** der T-Statistik gibt wie das Pendant in der F-Statistik die Wahrscheinlichkeit an, einen Wert, der größer als die Prüfgröße $T$ ist, aus den Stichprobenwerten zu erhalten. Mittels des absoluten T-Wertes und der Freiheitsgrade der Residuen lässt sich der Wert $P_T$ bestimmen. Auch dieser P-Wert wird im Allgemeinen aus einer bereits berechneten T-Verteilungstabelle entnommen (siehe auch 8.2.3). Da sich die Verteilungsfunktion der Dichte $D$ diesmal anders manifestiert, muss auch die Berechnung der Flächenstücke angepasst werden, wobei sich ebenfalls wieder der kritische und der berechnete T-Wert (oberer/positiver sowie unterer/negativer) als der gleiche Grenzwert darstellen:

$$P_T = \int_T^\infty D + \int_{-\infty}^{-T} D = 2 \cdot \int_T^\infty D = 2 \cdot \int_{-\infty}^{-T} D$$  \[8.57\]

Das Signifikanzniveau wird wiederum durch die Gegenwahrscheinlichkeit $1 - P_T$ errechnet. Je näher die berechnete Wahrscheinlichkeit (die gesamte Fläche des T-Wertes) zu Null tendiert, desto geringer ist die Chance ein falsches $H_0$ zu akzeptieren und desto sicherer kann die berechnete Größe des T-Wertes bestätigt werden. Diese Aussage stützt infolgedessen auch die Hypothese $H_1$. Wie man sich leicht vorstellen kann, wird die ganze Regressionsrechnung immer stabiler, je mehr Variablen durch den P-Wert als hilfreich eingestuft werden.

Die Berechnung des **kritischen Werts von $T$** erfolgt in gleicher Weise wie bei der des F-Wertes. Es wird ein Signifikanzniveau festgelegt, ab dem die Hypothese $H_1$ akzeptiert wird. Auch hier gehen die meisten Berechnungen von einem $\alpha$-Quantil von $0,05$ aus, was sich in einer Sicherheit von $95\%$ ($1-\alpha$) kundtut. Das besagte Restrisiko von $5\%$ ist ebenfalls für die zutreffende, aber als falsch angenommene Hypothese $H_0$ zuständig. Man nimmt in Kauf, dass zu $5\%$ eigentlich kein Zusammenhang zwischen den Merkmalen $X$ und $Y$ besteht. Dies wird aber in der Praxis angesichts der geringen Fehlerwahrscheinlichkeit als signifikant beglaubigt. Außerdem wird in dieser Arbeit sogar ein Signifikanzniveau von über $99,99\%$ erzielt, was im Grunde einen Zusammenhang garantiert. Der kritische T-Wert wird wie auch schon der kritische F-Wert einfach in einer T-Verteilungstabelle ausgelesen (siehe auch 8.2.3). Wenn der
T-Wert größer als der ausgelesene Grenzwert ist, dann kann daraus geschlossen werden, dass die Variable, die von dem betrachteten $\alpha_i$ gewichtet wird, der Regression dienlich ist. Somit kann die Hypothese $H_0$ verworfen und die Alternativhypothese ($H_1$) akzeptiert werden. Auch bei der T-Statistik wird im Normalfall der P-Wert berechnet, da er sich als eine leichter erkennbare Größe präsentiert. Allerdings kann für eine bessere Vergleichbarkeit der kritische T-Wert eher helfen, da bei verschiedenen Berechnungen der T-Wert jeweils gegen die gleiche Wahrscheinlichkeit getestet werden kann.

8.2.2.4.2 Konfidenzintervall

Das Konfidenzintervall ($I$) gibt an, innerhalb welcher Grenzen der wahre Parameter ($\alpha$) mit einer selbst ausgesuchten Wahrscheinlichkeit ($\alpha$-Quantil) um den eigenen Schätzwert ($\hat{\alpha}$) liegt. Die Berechnung sieht dann wie folgt aus:

$$I_{\alpha_i} = \alpha_i \pm T_{kritisch} \alpha_i * s_{\alpha_i} \quad [8.58]$$

Dabei ist $\alpha$ der berechnete Koeffizient und $T_{kritisch}$ der kritische Wert der T-Statistik, der üblicherweise aus einer statistischen T-Verteilungstabelle entnommen wird (siehe auch 8.2.3).

Man könnte das Intervall auch als Signifikanzniveau des P-Wertes der T-Verteilung in absolute Zahlen interpretieren. Dadurch kann man aus der T-Statistik schon erahnen, wie breit das Intervall ausfallen wird. Je größer das Verhältnis zwischen Koeffizient und Standardfehler, das durch den T-Wert beschrieben wird, beziehungsweise je kleiner die Brauchbarkeitswahrscheinlichkeit, die durch den P-Wert angegeben wird, wird, desto kleiner wird das Intervall für diesen Koeffizienten. Standardmäßig wird ein Intervall mit einer Genauigkeit von 95% getestet. Die restlichen fünf Prozent sagen wiederum aus, dass der gefundene Wert mit dieser Wahrscheinlichkeit außerhalb des Intervalls liegt, was jedoch recht gering ist und daher im Allgemeinen akzeptiert wird.
8.2.3 Anmerkungen


Die kritischen Werte von $F$ und $T$ werden in dieser Arbeit nicht betrachtet, sie lassen sich bei Bedarf aber auch einfach in Excel mit den Funktionen „FINV“ und „TINV“ anzeigen. In beiden Funktionen muss als erstes das $\alpha$-Quantil eingegeben werden, wobei der Wert „1 – gewünschte Wahrscheinlichkeit“ annehmen muss. Bei „FINV“ werden im Weiteren die beiden Freiheitsgrade der Regression und der Residuen gefordert. Um die „TINV“-Funktion zu benutzen müssen nur die Freiheitsgrade der Residuen verwendet werden. Diese Funktion wird vor allem für die Berechnung der Intervalle der Koeffizienten ($l_{\alpha,i}$) gebraucht, wobei Excel bei der Berechnung der Regression standardmäßig ein $\alpha$-Quantil von 0,05 (95% Sicherheit) benutzt und man erst als weitere Option eine selbst gewählte Wahrscheinlichkeit angeben kann.

9 Dokumentation der Berechnungen in Excel

9.1 InstitutionelleIO.xls
In dieser Excel Datei wird eine institutionelle Input-Output-Tafel erstellt.

Die Statistik Austria publiziert Input-Output-Tabellen ausschließlich in der Form Güter x Aktivitäten. Um aus einer solchen Tabelle möglichst realistische Arbeitswerte
errechnen zu können, benötigt man die Daten jedoch in der Form Aktivitäten x Aktivitäten.

Auf der Homepage (Literaturverzeichnis [23]) stellt die Statistik Austria sämtliche zur Bestimmung dieser Tabelle notwendige Tabellen zur Verfügung.

### 9.1.1 Ausgangsmaterial

Um eine Input-Output-Tabelle in der Form Aktivitäten x Aktivitäten zu erstellen benötigen wir folgende Ausgangstabellen:

- Heimische Produktion, zu Herstellungspreisen (Make-Matrix)
- Vorleistungen, zu Herstellungspreisen (Use-Matrix)
- Endnachfrage, zu Anschaffungspreisen
- Wertschöpfung, zu Anschaffungspreisen

Das Aktivitäten-/Güteraufkommen bzw. die Aktivitäten-/Güterverwendung kann nach zwei Konzepten bewertet werden:

1. zu Anschaffungspreisen
2. zu Herstellungspreisen

Anschaffungspreise können nach folgendem Schema in Herstellungspreise übergeleitet werden:

- Anschaffungspreise
  - Nicht abzugsfähige Mehrwertsteuer
  - Importabgaben
  - Sonstige Gütersteuern
  + Gütersubventionen
  - Handelsspannen
  - Transportspannen
  - Herstellungspreise
9.1.1.1 Heimische Produktion

Obwohl per Konvention eine Make-Matrix in der Form Aktivitäten x Güter angeschrieben wird, publiziert die Statistik Austria diese in transponierter Form (Güter x Aktivitäten).

9.1.1.2 Use-Matrix, Vorleistungen
In der Use-Matrix kann die Güterverwendung zu Anschaffungspreisen oder zu Herstellungspreisen ausgewiesen werden. Die Bewertung zu Anschaffungspreisen entspricht dem vom Käufer bezahlten Betrag, abzüglich der abziehbaren, aber inklusiv der nicht abziehbaren Mehrwertsteuer. Sonstige Gütersteuern und Importabgaben sind ebenfalls inkludiert, Gütersubventionen nicht. Im Anschaffungspreis sind auch Handels- und Transportspannen enthalten; dies gilt auch für die auf Importen liegenden im Inland erbrachten Handels- und Transportleistungen. Exporte sind fob bewertet, sie enthalten alle Handels- und Transportspannen bis zur Ausfuhrgrenze sowie die auf den Exporten liegenden Gütersteuern abzüglich der Gütersubventionen.

In der Berechnung einer Input-Output-Tabelle der Form Aktivitäten x Aktivitäten wird die Use-Matrix zu Herstellungspreisen benötigt.

9.1.1.3 Endnachfrage
Die Endnachfrage verhält sich bezüglich Anschaffungs- und Herstellungspreisen wie die Vorleistungen. Auch sie wird sowohl zu Herstellungspreisen, als auch zu Anschaffungspreisen publiziert.

In der Berechnung einer Input-Output-Tabelle der Form Aktivitäten x Aktivitäten wird die Use-Matrix zu Anschaffungspreisen benötigt.
9.1.1.4 Wertschöpfung
Die Wertschöpfung wird zu Herstellungspreisen ausgewiesen. Sie errechnet sich als Differenz zwischen dem Produktionswert zu Herstellungspreisen und den Vorleistungen zu Anschaffungspreisen.

9.1.2 Tabellenblätter
Die Namen der Arbeitsblätter in der Datei „InstitutionelleIO.xls“ werden wie folgt benannt:

<table>
<thead>
<tr>
<th>Tabellenblattname</th>
<th>Semantische Bedeutung</th>
</tr>
</thead>
<tbody>
<tr>
<td>V trans.</td>
<td>Heimische Produktion (Make-Matrix), transponiert</td>
</tr>
<tr>
<td>W</td>
<td>Wertschöpfung</td>
</tr>
<tr>
<td>U</td>
<td>Vorleistungen (Use-Matrix)</td>
</tr>
<tr>
<td>F</td>
<td>Endnachfrage</td>
</tr>
<tr>
<td>g</td>
<td>Produktionswert der Aktivitäten</td>
</tr>
<tr>
<td>q trans.</td>
<td>Verwendung heimischer Güter, transponiert</td>
</tr>
<tr>
<td>B</td>
<td>Vorleistungskoeffizienten</td>
</tr>
<tr>
<td>D</td>
<td>Market-Shares-Matrix</td>
</tr>
<tr>
<td>Vorl. abs.</td>
<td>Vorleistungen, absolute Werte</td>
</tr>
<tr>
<td>IO 57x57</td>
<td>Input-Output-Tabelle, 57 Aktivitäten x 57 Aktivitäten</td>
</tr>
<tr>
<td>IO 15x15</td>
<td>Input-Output-Tabelle, 15 Aktivitäten x 15 Aktivitäten</td>
</tr>
</tbody>
</table>

9.1.3 Erläuterungen zu den Berechnungen
Die Daten für die Tabellen „V trans“, „W“, „U“ und „F“ werden von der Statistik Austria publiziert.

Die Summen der Vorleistungskoeffizienten (Tabellenblatt „U“) werden nach folgender Formel berechnet:

\[ B = U \cdot \text{diag}(g)^{-1} \]  \hspace{1cm} [9.1]

Dabei steht diag(g) für die Diagonalmatrix von g.

Anschließend kann die Market-Shares-Matrix (Tabellenblatt „D“) errechnet werden:

\[ D = V \text{ trans} \cdot \text{diag}(q \text{ trans})^{-1} \]  \hspace{1cm} [9.2]

Mit Hilfe der eben errechneten Werte für B und D können nun die Summen der absoluten Werte der Vorleistungen bestimmt werden. Dies geschieht mit der Formel:

\[ \text{Vorl abs} = B \cdot D \cdot \text{diag}(g) \]  \hspace{1cm} [9.3]


Da für nachfolgende Berechnungen eine Gliederung von 15x15 Aktivitäten von Nöten ist, werden im Tabellenblatt „IO 15x15“ die Daten aus „IO 57x57“ entsprechend Kapitel 13 „Gliederung der ÖNACE 2003“ zusammengefasst.
9.2 Arbeitswerte.xls
In dieser Excel-Datei werden die Arbeitswerte für die einzelnen Sektoren, abhängig von der Bildung der Beschäftigten, errechnet und abgebildet.

9.2.1 Ausgangsmaterial
Zur Berechnung der Arbeitswerte auf Basis der Bildungsstufe der Beschäftigten werden folgende Tabellen benötigt:

- Input-Output-Tabelle (15x15 Aktivitäten)
- Produktionswert zu Herstellungspreisen
- Erwerbspersonen nach höchster abgeschlossener Ausbildung und ÖNACE-Abschnitten

Mit Hilfe dieser Daten können alle, zur Berechnung des Arbeitswertes notwendigen Rechnungen, durchgeführt werden.

9.2.2 Tabellenblätter
Die Namen der Arbeitsblätter in der Datei „InstitutionelleIO.xls“ werden wie folgt benannt:

<table>
<thead>
<tr>
<th>Tabellenblattname</th>
<th>Semantische Bedeutung</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zusammenfassung</td>
<td>Übersicht über alle Regressionsergebnisse</td>
</tr>
<tr>
<td>Regressionsbewertung</td>
<td>Grafische Darstellung aller p-Werte der Regressionen</td>
</tr>
<tr>
<td>InputOutput</td>
<td>Input-Output-Tabelle (15x15 Aktivitäten)</td>
</tr>
<tr>
<td>q</td>
<td>Produktionswert zu Herstellungspreisen</td>
</tr>
<tr>
<td>E</td>
<td>Einheitsmatrix (15x15)</td>
</tr>
<tr>
<td>Ausbildung</td>
<td>Erwerbspersonen nach höchster abgeschlossener Ausbildung und ÖNACE-Abschnitten</td>
</tr>
<tr>
<td>Bildungssegmentierung</td>
<td>Liste aller Bildungssegmentierungen</td>
</tr>
<tr>
<td>Seg Summe – Seg E3</td>
<td>Errechnete Arbeitswerte für die entsprechenden Bildungssegmentierungen, sowie deren Regressionskoeffizienten</td>
</tr>
</tbody>
</table>
9.2.3 Erläuterungen zu den Berechnungen

Zu Beginn werden die Vorleistungen aus der Tabelle „InputOutput“ in das Namenfeld „_Z“ gespeichert. Im Tabellenblatt „q“ stehen die transponierten Produktionswerte zur Herstellungs-preisen.

Mit Hilfe der nachfolgenden Formel wird die A-Matrix gebildet, welche im gleichnamigen Tabellenblatt abgebildet ist:

\[ A = \text{diag}(q^{\text{trans}})^{-1} \times Z \]  

[9.4]

Die Einheitsmatrix im Blatt „E“ wird für nachfolgende Berechnungen benötigt.


In „Bildungssegmentierung“ wird die Zahl der Beschäftigten nach insgesamt 27 verschiedenen Segmentierungen eingeteilt. Bei der Benennung der Segmentierungen wird folgendes Schema angewandt:

<table>
<thead>
<tr>
<th>Segmente</th>
<th>Semantische Bedeutung</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seg A1 – A8</td>
<td>2 Bildungssegmente: Niedrige und Hohe Bildung</td>
</tr>
<tr>
<td>Seg B1 – B7</td>
<td>3 Bildungssegmente: Niedrige, Mittlere und Hohe Bildung</td>
</tr>
<tr>
<td></td>
<td>Hohe Bildung ist immer Fachrichtung 9</td>
</tr>
<tr>
<td>Seg C1 – C5</td>
<td>Bildungssegmente: Niedrige, Mittlere und Hohe Bildung</td>
</tr>
<tr>
<td></td>
<td>Hohe Bildung sind immer Fachrichtungen 9 und 8</td>
</tr>
<tr>
<td>Seg D1 – D4</td>
<td>3 Bildungssegmente: Niedrige, Mittlere und Hohe Bildung</td>
</tr>
<tr>
<td></td>
<td>Hohe Bildung sind immer Fachrichtungen 9, 8 und 7</td>
</tr>
<tr>
<td>Seg E1 – E3</td>
<td>3 Bildungssegmente: Niedrige, Mittlere und Hohe Bildung</td>
</tr>
<tr>
<td></td>
<td>Hohe Bildung sind immer Fachrichtungen 9, 8, 7 und 6</td>
</tr>
</tbody>
</table>
Auf den diversen Tabellenblättern der Segmentierungen (zum Beispiel „Seg A1“) werden nun Arbeitswerte in Abhängigkeit der von der spezifischen Ausbildungssegmentierung berechnet.

Der ‚AWT nach Bildung (in Personen)‘ wird durch folgende Formel bestimmt:

\[ AWT = Personen \times (E - A)^{-1} \]  

Dabei steht ‚Personen‘ für die Anzahl der Personen in der jeweiligen Ausbildungsstufe und ist in der Datei im Namenfeld „_segXX“ zu finden.

Neben der Spalte für die Summe der Arbeitswerte in Personen, wird der Brutto-Produktionswert gelistet. Dieser wird aus der Input-Output-Tafel aus dem Blatt „InputOutput“ ausgelesen.


Über den AWT in Personen und BPW in Personen kann nun die Regression berechnet werden. Ihre Ergebnisse werden auf jedem „Seg XX“-Blatt unter der Überschrift ‚Regressionsstatistik‘ angeführt. Von besonderem Interesse sind hier die jeweiligen Regressionskoeffizienten für die Bildungssegmentierungen. Mit ihnen wird nun ein geschätzter AWT (wieder in Personen) berechnet.


Mit Hilfe der Regressionsstatistiken kann die Güte des Ergebnisses abgeschätzt werden. Vor allem der P-Wert eine Regression ist von großer Bedeutung (näheres
dazu siehe Kapitel 8.2.2.2.4). Um diese Werte zu verdeutlichen sind im Blatt „Regressionsgüte“ die P-Werte aller Regressionen in einer Grafik abgebildet. So lässt sich auf einen Blick erkennen, ob eine Regression brauchbar ist oder nicht. Sind alle beiden (oder alle 3) Punkt einer Segmentierung unterhalb der 20% Marke spricht man von einem aussagekräftigen Regressionsergebnis.

Im letzten Blatt „Zusammenfassung“ werden nun die wichtigsten Ergebnisse aller Regressionen verdeutlicht. Die einzelnen Ausbildungssegmentierungen sind farblich hervorgehoben (orange steht zum Beispiel für Niedrige Ausbildung) und beinhalten den entsprechenden Koeffizienten und darunter den P-Wert. Daneben befinden sich 3 Spalten, die weitere Gütekriterien der Regression verdeutlichen sollen:

- das Bestimmtheitsmaß $r^2$
- der Standardfehler
- der Schnittpunkt der Regression
10 Interpretationen der berechneten Daten

### Legende

- **Hohe Ausbildung**
- **Mittlere Ausbildung**
- **Niedrige Ausbildung**

### Regressionskoeffizient

<table>
<thead>
<tr>
<th>r²</th>
<th>Standardfehler</th>
<th>Schnittpunkt</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6056</td>
<td>23002.06</td>
<td>2857.69</td>
</tr>
<tr>
<td>0.9560</td>
<td>116648.81</td>
<td>-19996.80</td>
</tr>
<tr>
<td>0.9564</td>
<td>116034.06</td>
<td>-20392.19</td>
</tr>
<tr>
<td>0.9505</td>
<td>123702.24</td>
<td>-4799.56</td>
</tr>
<tr>
<td>0.9508</td>
<td>123330.29</td>
<td>-6272.60</td>
</tr>
<tr>
<td>0.9565</td>
<td>115899.27</td>
<td>-19360.57</td>
</tr>
<tr>
<td>0.9581</td>
<td>113868.78</td>
<td>-25880.11</td>
</tr>
<tr>
<td>0.9520</td>
<td>121864.33</td>
<td>-14826.01</td>
</tr>
<tr>
<td>0.9502</td>
<td>124103.10</td>
<td>-1541.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segmentierung</th>
<th>Ausbildung</th>
<th>Regressionsstatistik</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universitäten und Hochschulen</td>
<td>Fachhochschulen</td>
<td>Kollegs und Abiturientenlehrgänge</td>
</tr>
<tr>
<td>Summe</td>
<td></td>
<td>Hohe Ausbildung</td>
</tr>
<tr>
<td>A1</td>
<td></td>
<td>2.95</td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td>3.00</td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td>1.42</td>
</tr>
<tr>
<td>A4</td>
<td></td>
<td>1.48</td>
</tr>
<tr>
<td>A5</td>
<td></td>
<td>1.97</td>
</tr>
<tr>
<td>A6</td>
<td></td>
<td>1.92</td>
</tr>
<tr>
<td>A7</td>
<td></td>
<td>1.35</td>
</tr>
<tr>
<td>A8</td>
<td></td>
<td>1.18</td>
</tr>
<tr>
<td>Segmentierung</td>
<td>Universitäten und Hochschulen</td>
<td>Fachhochschulen</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Legende</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B4</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B5</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B6</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B7</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C3</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legende: Hohe Ausbildung, Mittlere Ausbildung, Niedrige Ausbildung, Regressionskoeffizient, p-Wert (in %)
<table>
<thead>
<tr>
<th>Segmentierung</th>
<th>Universitäten und Hochschulen</th>
<th>Fachhochschulen</th>
<th>Berufs- und lehrerbildende Akademie</th>
<th>Kollegs und Abituriertenlehrgänge</th>
<th>Berufsbildende höhere Schule</th>
<th>Allgemein bildende höhere Schule</th>
<th>Berufsbildende mittlere Schule</th>
<th>Lehrlingsausbildung</th>
<th>Allgemein bildende Pflichtschule</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>6.71 (4.1%)</td>
<td></td>
<td></td>
<td></td>
<td>-1.20 (42.2%)</td>
<td>0.48 (49.7%)</td>
<td>1.26 (0.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>3.38 (6.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.22 (13.0%)</td>
<td>1.96 (29.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>0.03 (95.5%)</td>
<td></td>
<td>9.61 (0.0%)</td>
<td></td>
<td>0.79 (14.1%)</td>
<td>0.22 (13.0%)</td>
<td>0.77 (2.1%)</td>
<td>0.79 (69.2%)</td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>-0.01 (99.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.38 (3.7%)</td>
<td>0.77 (2.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>0.93 (44.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77 (2.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>1.41 (17.6%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77 (2.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>0.06 (93.8%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.38 (3.8%)</td>
<td>0.38 (17.3%)</td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td>1.08 (36.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.81 (1.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td>1.47 (15.0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.90 (65.2%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legende**

- **Hohe Ausbildung**
- **Mittlere Ausbildung**
- **Niedrige Ausbildung**

<table>
<thead>
<tr>
<th>Regressionsstatistik</th>
<th>$r^2$</th>
<th>Standardfehler</th>
<th>Schnittpunkt</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>0.9635</td>
<td>110914.13</td>
<td>7219.66</td>
</tr>
<tr>
<td>C5</td>
<td>0.9579</td>
<td>119220.39</td>
<td>-17883.31</td>
</tr>
<tr>
<td>D1</td>
<td>0.9884</td>
<td>62491.43</td>
<td>20.84</td>
</tr>
<tr>
<td>D2</td>
<td>0.9793</td>
<td>83492.79</td>
<td>-26240.75</td>
</tr>
<tr>
<td>D3</td>
<td>0.9526</td>
<td>126376.27</td>
<td>-17091.81</td>
</tr>
<tr>
<td>D4</td>
<td>0.9505</td>
<td>129151.26</td>
<td>-5196.07</td>
</tr>
<tr>
<td>E1</td>
<td>0.9788</td>
<td>84593.04</td>
<td>-26481.41</td>
</tr>
<tr>
<td>E2</td>
<td>0.9523</td>
<td>126889.50</td>
<td>-16320.10</td>
</tr>
<tr>
<td>E3</td>
<td>0.9508</td>
<td>128806.57</td>
<td>-6420.80</td>
</tr>
</tbody>
</table>
10 Interpretationen der berechneten Daten

10.1 Überblick über die Regressionsergebnisse


Beinhaltet eine Linie nun mindestens einen Punkt über der 20%-Marke, so ist die entsprechende Bildungssegmentierung zu vernachlässigten.

Betrachtet man die Grafiken im Blatt „Regressionsgüte“ ist zu erkennen, dass lediglich die Regressionen für die Bildungssegmentierungen A1, A2, A3, A4, A5, A6, A7, B2, B3 sowie C1 ein zufriedenstellendes Resultat liefern.


Aussagekräftige Resultate liefern folglich nur die Regressionen für die Bildungssegmentierungen A1, A2, A3, A4, A5, A6 und A7.
10.2 Interpretation der aussagekräftigen Resultate

Aussagekräftige Bildungssegmentierungen

In Anbetracht der oben abgebildeten Grafik ("aussagekräftige Bildungssegmentierung") lässt sich der Schluss ziehen, dass eine Person mit höherer abgeschlossener Ausbildung nachweisbar mehr zur Wirtschaft eines Landes (in diesem Falle Österreichs) beiträgt, als eine in derselben Branche beschäftigte Person mit Niedrigerer Ausbildung.

<table>
<thead>
<tr>
<th>Segmentierung A1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Universitäten und Hochschulen</td>
<td>Hohe Ausbildung</td>
</tr>
<tr>
<td>8 Fachhochschulen</td>
<td>Niedrige Ausbildung</td>
</tr>
<tr>
<td>7 Berufs- und lehrerbildende Akademien</td>
<td></td>
</tr>
<tr>
<td>6 Kollegs und Abiturientenlehrgänge</td>
<td></td>
</tr>
<tr>
<td>5 Berufsbildende höhere Schule</td>
<td></td>
</tr>
<tr>
<td>4 Allgemein bildende höhere Schule</td>
<td></td>
</tr>
<tr>
<td>3 Berufsbildende mittlere Schule</td>
<td></td>
</tr>
<tr>
<td>2 Lehrlingsausbildung</td>
<td></td>
</tr>
<tr>
<td>1 Allgemein bildende Pflichtschule</td>
<td></td>
</tr>
</tbody>
</table>

Regressionskoeffizienten

- Hohe Ausbildung
- Niedrige Ausbildung
<table>
<thead>
<tr>
<th>Segmentierung A2</th>
<th>Segmentierung A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Universitäten und Hochschulen</td>
<td>Universitäten und Hochschulen</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Fachhochschulen</td>
<td>Fachhochschulen</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Berufs- und lehrerbildende Akademien</td>
<td>Berufs- und lehrerbildende Akademien</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kollegs und Abiturientenlehrgänge</td>
<td>Kollegs und Abiturientenlehrgänge</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Berufsbildende höhere Schule</td>
<td>Berufsbildende höhere Schule</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Allgemein bildende höhere Schule</td>
<td>Allgemein bildende höhere Schule</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Berufsbildende mittlere Schule</td>
<td>Berufsbildende mittlere Schule</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Lehrlingsausbildung</td>
<td>Lehrlingsausbildung</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Allgemein bildende Pflichtschule</td>
<td>Allgemein bildende Pflichtschule</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segmentierung A4</th>
<th>Segmentierung A5</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Universitäten und Hochschulen</td>
<td>Universitäten und Hochschulen</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Fachhochschulen</td>
<td>Fachhochschulen</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Berufs- und lehrerbildende Akademien</td>
<td>Berufs- und lehrerbildende Akademien</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kollegs und Abiturientenlehrgänge</td>
<td>Kollegs und Abiturientenlehrgänge</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Berufsbildende höhere Schule</td>
<td>Berufsbildende höhere Schule</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Allgemein bildende höhere Schule</td>
<td>Allgemein bildende höhere Schule</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Berufsbildende mittlere Schule</td>
<td>Berufsbildende mittlere Schule</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Lehrlingsausbildung</td>
<td>Lehrlingsausbildung</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Allgemein bildende Pflichtschule</td>
<td>Allgemein bildende Pflichtschule</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segmentierung A6</th>
<th>Segmentierung A7</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Universitäten und Hochschulen</td>
<td>Universitäten und Hochschulen</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Fachhochschulen</td>
<td>Fachhochschulen</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Berufs- und lehrerbildende Akademien</td>
<td>Berufs- und lehrerbildende Akademien</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kollegs und Abiturientenlehrgänge</td>
<td>Kollegs und Abiturientenlehrgänge</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Berufsbildende höhere Schule</td>
<td>Berufsbildende höhere Schule</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Allgemein bildende höhere Schule</td>
<td>Allgemein bildende höhere Schule</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Berufsbildende mittlere Schule</td>
<td>Berufsbildende mittlere Schule</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Lehrlingsausbildung</td>
<td>Lehrlingsausbildung</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Allgemein bildende Pflichtschule</td>
<td>Allgemein bildende Pflichtschule</td>
</tr>
</tbody>
</table>
Die Abbildung der Bildungssegmentierung von A1 bis A7 zeigt, dass die Grenze zwischen Niedriger und Hoher Ausbildung nie unter die 'Lehrlingsausbildung' gesetzt wird.

Summiert man nun die Koeffizienten für Hohe Ausbildung erhält man ein Total von ca. 14,09. Die Division dieses Wertes durch die Summe aller Koeffizienten für Niedrigen Ausbildung (ca. 6,25) liefert ein einen gemittelten Faktor von rund 2,25 den eine 'gebildeter' Person im Vergleich zu einer 'ungebildeten' Person an Mehrwert produziert.

Oder salopp formuliert: Vergleicht man den wirtschaftlichen Output einer Person mit Lehrlingsausbildung oder Pflichtschulabschluss und einer Person mit höherem Abschluss, so ist die Person mit höherem Abschluss etwas mehr als doppelt so produktiv als die Person mit Lehrlingsausbildung/Pflichtschulabschluss.
11 Literaturverzeichnis

11.1 Druckwerke

[1] BACKHAUS, Klaus; Erichson, Bernd; Plinke, Wulff: Multivariate Analysemethoden. – Berlin: Springer Verlag, 2000


11.2 Internetseiten
[27] http://www.bildungssystem.at
[28] http://www.esds.ac.uk

Alle angeführten Internetseiten wurden letztmals im März 2008 auf Aktualität und Korrektheit der entnommenen Informationen überprüft.