



## **Aufbau und Erprobung des internationalen Datenservicezentrums für arbeitsmarktrelevante Daten am IZA**



**Gesamtziele**

**Voraussetzungen und Vorleistungen**

**Umfang und Bereitstellung internat. Arbeitsmarktdaten**

**Umfang und Bereitstellung deutscher Arbeitsmarktdaten**

**Metainformationsportal**

**Metadatenstandardisierung**

**Umsetzung Metadatenbank beim IZA**

**Integration von Metadaten und virtueller Datenbank**

Metainformationsportal für Sekundäranalysen zur Bereitstellung von Informationen zur Arbeitsmarktforschung und arbeitsmarktrelevanten Daten

---

■ Bestimmung der zwanzig wichtigsten Studien durch Umfrage unter Arbeitsmarktökonominnen (IZA-Fellows) und Schaffung eines indirekten Zugangs durch kontrollierte Datenfernverarbeitung

---

■ ■ Deutsche Daten international zugänglich machen für externe internationale Forscher: Faktisch anonymisierte Scientific Use files (Grundfiles) der amtlichen Statistik, Umfragedaten, prozessgenerierte Daten aus der aml. Verwaltung, Unternehmensdaten etc.

---

■ ■ ■ Erstellung und Standardisierung übergreifend recherchierbarer Metadaten und Dokumentationen in Kooperation mit Forschungsdatenzentren (FDZ) und DSZ

### **Inhaltlich:**

- Innovative und wettbewerbsfähige Arbeitsmarktforschung im internationalen Kontext
  - Überwiegend empirische Forschungstätigkeit auf Grundlage nationaler und internationaler Mikrodaten
  - Enge Kooperation mit ausländischen wissenschaftlichen Instituten der Arbeitsökonomie
- 

### **Organisatorisch:**

- IZA Datenkommission zur Beratung der Forscher in allen Fragen des Datenhandlings und z.T. der Datenanalyse
  - Verantwortung für Bereitstellung aller gängigen Statistik- und Datenanalyseprogramme
  - Kontrolle und Regelung des Datenschutz
  - Datenpflege und -dokumentation
- 

### **Technisch:**

- Leistungsfähige SUN-Solaris-Server, zur Datenvorhaltung und –analyse,
- Gängige Analyse-Programme zur Verfügung (STATA, SAS, SPSS etc.)

# Internat. Datenservicezentrum (DSZ)

## Umfang und Bereitstellung internat. Arbeitsmarktdaten



- Zwanzig wichtigste internationale Studien durch eine Umfrage (Online-Survey, alle zwei Jahre) unter Arbeitsmarktökonominnen (IZA-Fellows) bestimmen
- Soweit rechtlich möglich, Zugang für dezentral verfügbare Mikrodaten durch kontrollierte Datenfernverarbeitung (Schalterstelle, virtuelle Datenbank)
- Ergebnisse der IZA-Fellow-Surveys: Kerndatenbestand der Metadatenbank

### Beispiel für Umfang und Zusammensetzung

Kurzname	Studien-Name
BHPS	British Household Panel Survey
CPS	Current Population Survey, Outgoing Rotation Groups
ECHP	European Community Household Panel
GSOEP	Das Sozio-oekonomische Panel
HILDA	Australian Household, Income and Labour Dynamics
HUS	The Swedish Panel Study Market and Nonmarket Activities
NLSY	National Longitudinal Survey of Youth
ILS	Israeli Longitudinal Survey
JPSC	Japanese Panel Survey on Consumers
LFS	British Labor Force Survey

- Kerndatenbestand aus IZA-Fellow Survey und nach Pilotphase gesamten Datenbestand des IZA integrieren
- Keine Datenweitergabe durch Nutzer-Download etc.

- Faktisch anonymisierte Scientific Use files der amtlichen Statistik (Grundfiles): Mikrozensus, die EVS, die Zeitbudgetstudie, Beschäftigtenstichprobe, den Regionalfile der Beschäftigtenstichprobe und den IZA Datensatz zur Evaluation aktiver Arbeitsmarktpolitik
- Weiterentwicklung bisheriger rechtlicher Zugangsmöglichkeiten von Scientific Use files für externe Forscher aus dem Ausland in Kooperation mit FDZ, DSZ, etc.
- Zentraler Zugang und added value Betreuung von Arbeitsmarktökonominnen für dezentral verfügbare Daten der Scientific und Public Use files
- Einrichtung indirekten Zugangs durch kontrollierte Datenfernverarbeitung (Schalterstelle, virtuelle Datenbank)
- Standardsoftware (STATA, SPSS, SAS)
- Bei rechtlichen Restriktionen Zugangsmöglichkeiten auf Metadaten und IZA-Gastwissenschaftlermodell beschränkt

## Serviceleistungen des Metainformationsportals (Auswahl):

- **Zentrale Anlaufstelle für Datenanfragen der Arbeitsmarktforschung**  
Informationen über relevante Daten und Statistiken, Chancen und Grenzen der Daten, Verweise zu verwandten Studien und weiteren Quellen, theoretisches Studienkonzept, Fragebogen, Variablen Korrespondenzen, Datenhomogenisierungen etc.
- 
- 2) **Hilfestellung für Sekundäranalysen der Arbeitsmarktforschung**  
Studienübergreifend Angaben zu statistischen Details wie Algorithmenbeschreibung, Stichprobenziehung, Filterführung, Gewichtung etc.
- 
- **Instrument zur Wissens-Akkumulation als Datenbank über Arbeitsmarktforschung und zugehörige Daten**  
Einbindung von Hyperlinks auf wiss. Institute, Publikationen, Wissenschaftler, Forschungsprojekte, Kommentare, dokumentierte Analyseroutinen und -modellspezifikation etc.

### **Grundelemente der Metadaten**

Kooperative Entwicklung übergreifend recherchierbarer Dokumentationen und Metadaten durch FDZ und DSZ auf Daten- und Variablenebene

---

#### **DDI / Dublin Core**

- EU FASTER-Initiative: Metadatenbankformat zur Integration von Mikro- und Makrodaten in einheitlicher Systemarchitektur und kontrollierten Attributen
  - DDI-Initiative & Dublin Core (DC): Minimalanforderung zur Dokumentation von Mikro- und Makrodaten. DC enthält bspw. Titel, Urheber, Fundstelle, thesaurusbasierte Schlagworte etc.
- 

#### **DSZ-Studien- und Variablen-Thesaurus**

- Entwicklung eines eigenen DSZ Thesaurus auf Grundlage arbeitsmarktrelevanter Einträge aus bereits verfügbaren Thesauri
- 

#### **Metadaten-Schnittstellendefinition für Im- bzw. Export**

- DDI als XML-Tags für jeden Aspekt Mikro/Makro incl. Mapping zu DC verfügbar
- Empfehlung: XML als geeigneten Schnittstellenstandard zwischen FDZ/DSZ festlegen

# Internat. Datenservicezentrum (DSZ)

## Umsetzung Metadatenbank



- Datenmodell**      ➤    Implementierung des DDI & DC

---

- Thesaurus**      ➤    Entwicklung eines eigenen DSZ Thesaurus auf Grundlage arbeitsmarktrelevanter Thesauri (HASSET, JEL, ILO, OECD etc.) mit hierarchischen Topics und special Topics

---

- Variablen**      ➤    Suchfunktion über sämtliche Attribute der Metadatenbank: DDI/DC, Variablen der Studien etc.  
                         ➤    Variable Correspondence

---

- First Results**   ➤    Derzeitiger Datenbestand des IZA in SQL-DB  
                         ➤    SQL-DB Internet-Abbildung und Recherche erfolgt via PHP

---

- Next Steps**      ➤    Mit Projektstart DSZ Internet-Verfügbarkeit  
                         ➤    Sukzessive Kodierung der IZA-Fellow Survey Datensätze und weitere Pflege / Ergänzung der Metadatenbank  
                         ➤    Klärung Datenschutz ⇔ Einrichtung Schalterstelle

# Internat. Datenservicezentrum (DSZ)

## Metadatenbankmodell Detail: Mikrodaten Variablen

Attribute	Description
-----------	-------------

1:n

<b>variable_meaning</b>	
<i>variable_meaning_id</i>	int (auto)
wave_id	int
question_number	varchar(255)
variable_name	varchar(255)
variable_label	varchar(255)
question_text	text

Primary Key  
 Wave id  
 Number of Questions in Questionnaire  
 Name or Variable in Study  
 Label of Variable in Study  
 Literal Question in Questionnaire

1:n

<b>variable_value</b>	
<i>variable_value_id</i>	int (auto)
variable_meaning_id	int
variable_value_number	int
meaning	text

Primary Key  
 Primary Key of the Table variable\_meaning  
 Variable that reference an Answer of a Question  
 Possible Answers

1:n

<b>variable2topic</b>	
<i>variable2topic_id</i>	int (auto)
variable_meaning_id	int
topic_lu_id	int

Primary Key  
 Primary Key of the Table variable\_meaning  
 References a Topic

# Internat. Datenservicezentrum (DSZ)

## Integration von Metadaten und virtueller Datenbank



### **Generelle Beschränkung eines integrierten Datenzugriffs:**

- Nutzung vorhandener Rohdaten in getrennten Systemen und nur durch separate Software

### **Einschränkungen virtueller Datenbanklösungen (NESSTAR, SDA etc.):**

- Windows oder Internet basiert (Setups, Geschwindigkeit, Stabilität etc.)
- Keine syntaxorientierte Sprache (STATA, SPSS, SAS) zur Analyse-Automatisierung
- Keine „machtvolle“ Analyse (Ökonometrie, Datenfusion, Mergen etc.)

### **Grundelemente zur Integration DSZ**

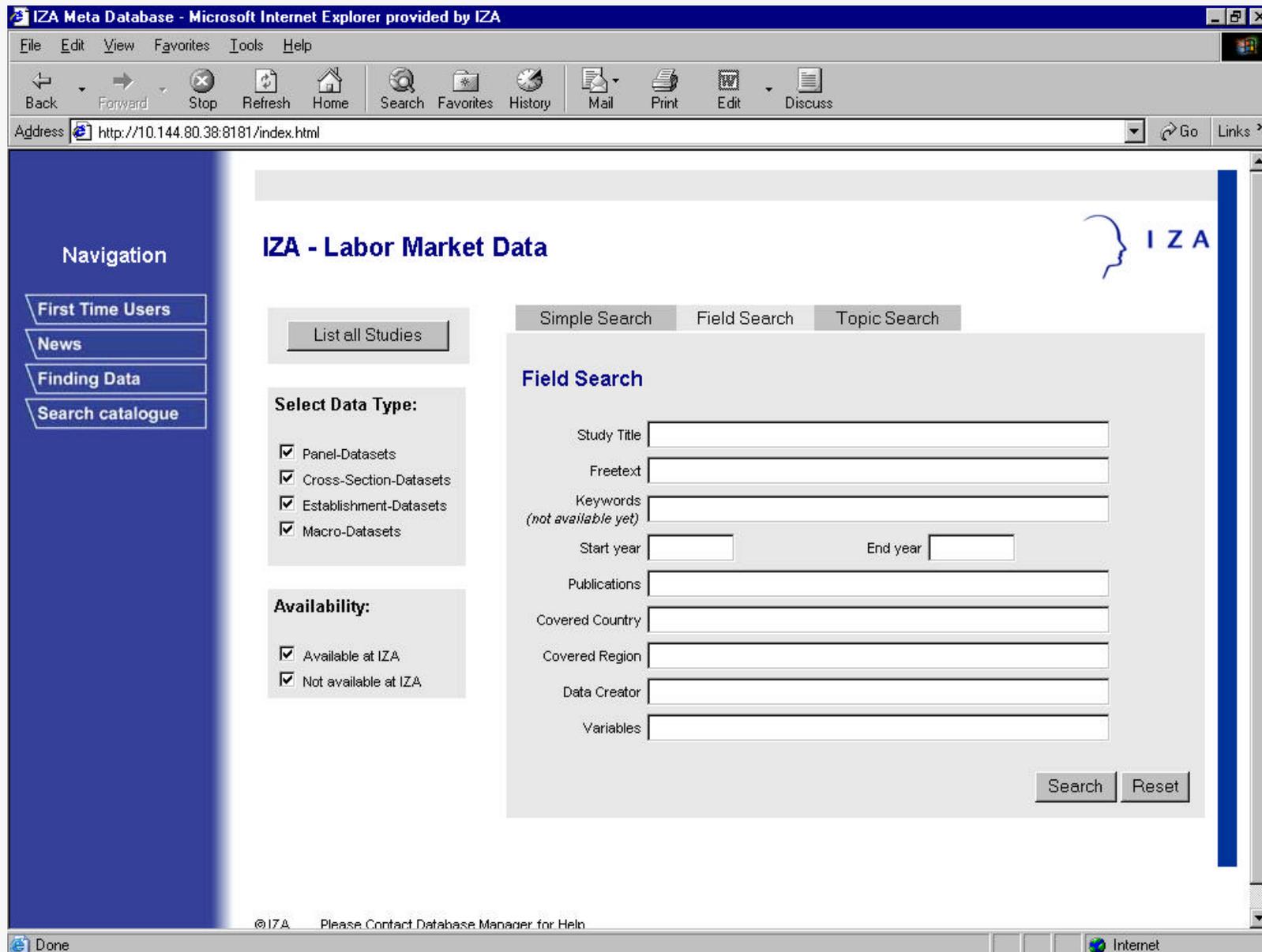
- ⇒ XML-Dokumente auf Basis des DDI/DC aus SQL-Datenbank generieren (Download/Export)
- ⇒ Offen: Abfrage- und Abbildungsort der DDI/XML-Files, etc.
- ⇒ Offen: Überwindung der Einschränkungen bisheriger virt. Datenbanken durch „Neuerfindung“ von Standard-Analysesoftware?



**Vielen Dank !**

# Internat. Datenservicezentrum (DSZ)

## Anhang



IZA Meta Database - Microsoft Internet Explorer provided by IZA

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss

Address <http://10.144.80.38:8181/index.html> Go Links >>

Navigation

- First Time Users
- News
- Finding Data
- Search catalogue

IZA - Labor Market Data

List all Studies

Select Data Type:

- Panel-Datasets
- Cross-Section-Datasets
- Establishment-Datasets
- Macro-Datasets

Availability:

- Available at IZA
- Not available at IZA

Simple Search Field Search Topic Search

Field Search

Study Title

Freetext

Keywords   
*(not available yet)*

Start year  End year

Publications

Covered Country

Covered Region

Data Creator

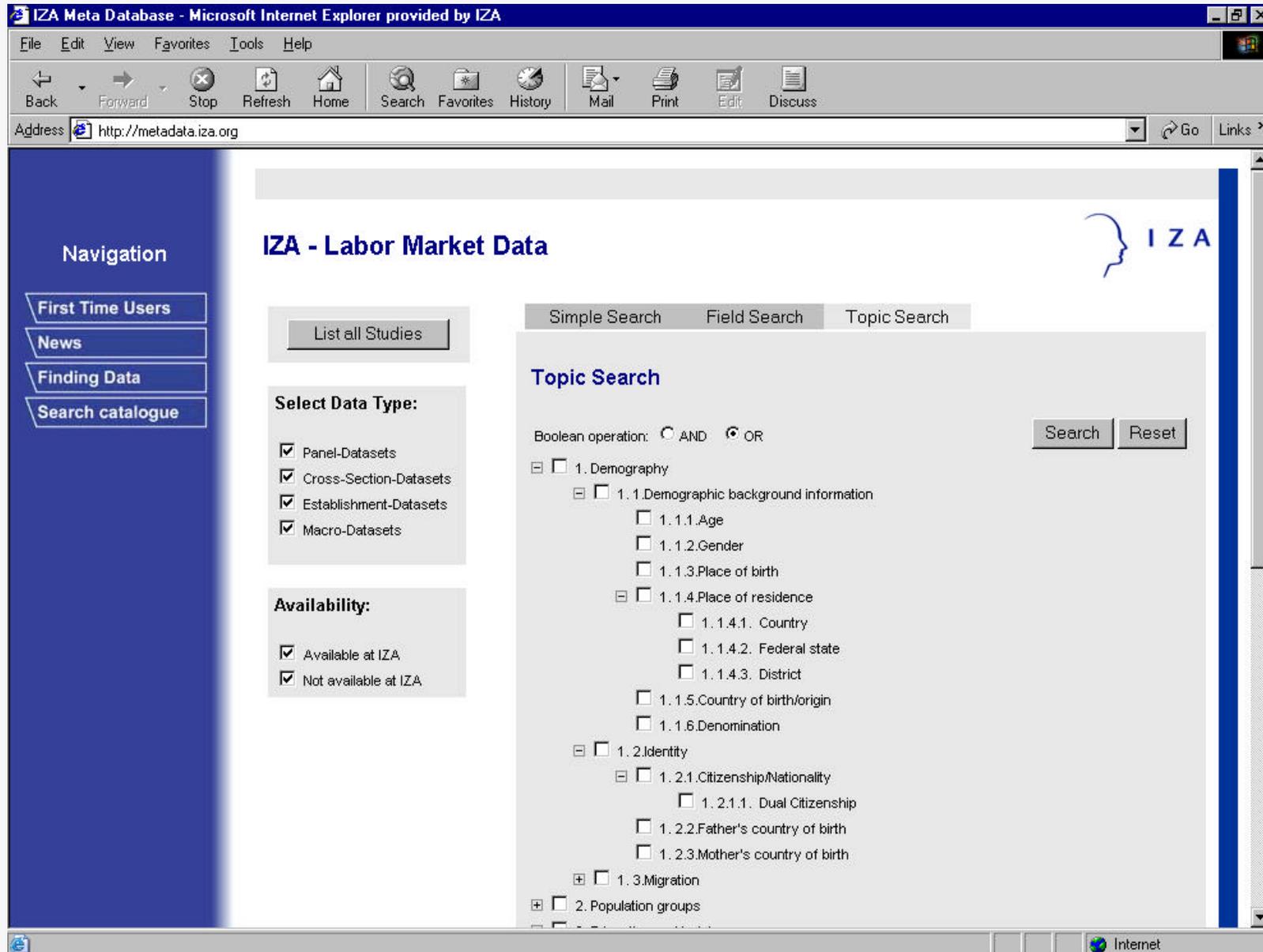
Variables

Search Reset

@ IZA Please Contact Database Manager for Help

# Internat. Datenservicezentrum (DSZ)

## Anhang



IZA Meta Database - Microsoft Internet Explorer provided by IZA

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss

Address <http://metadba.iza.org> Go Links

### Navigation

- First Time Users
- News
- Finding Data
- Search catalogue

## IZA - Labor Market Data

List all Studies

**Select Data Type:**

- Panel-Datasets
- Cross-Section-Datasets
- Establishment-Datasets
- Macro-Datasets

**Availability:**

- Available at IZA
- Not available at IZA

Simple Search Field Search Topic Search

### Topic Search

Boolean operation:  AND  OR

Search

- 1. Demography
  - 1.1. Demographic background information
    - 1.1.1. Age
    - 1.1.2. Gender
    - 1.1.3. Place of birth
    - 1.1.4. Place of residence
      - 1.1.4.1. Country
      - 1.1.4.2. Federal state
      - 1.1.4.3. District
    - 1.1.5. Country of birth/origin
    - 1.1.6. Denomination
  - 1.2. Identity
    - 1.2.1. Citizenship/Nationality
      - 1.2.1.1. Dual Citizenship
    - 1.2.2. Father's country of birth
    - 1.2.3. Mother's country of birth
  - 1.3. Migration
- 2. Population groups

# Internat. Datenservicezentrum (DSZ)

## Anhang

IZA Meta Database - Microsoft Internet Explorer provided by IZA

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss

Address <http://10.144.80.38:8181/index.html> Go Links

Navigation

First Time Users

News

Finding Data

Search catalogue

Browse Studies

Selected Study

List Datasets

Selected Dataset

Dataset Variables

Study Bibliography

Resources

Study Helpdesk

### IZA - Labor Market Data

#### ALLBUS 1982

Click on a variable name to look at th

Variable Name	Label	Variable Remarks	Values	Value Remarks	Topics
V1	ZA-STUDIENNUMMER 1160				
V2	BEFRAGTENNUMMER				
V3	SPLIT-NUMMER		1 FORM A 2 FORM B		20.11. Questions about interview and survey
V4	F1.WOHNDAUER			5.1. Occupancy	
V5	F1.WOHNDAUER ANZ JAHRE			5.1. Occupancy	
V6	F1.WOHNDAUER JAHRE KATEG			5.1. Occupancy	
V7	F2.SEIT WANN IN BRD			1.3.2.2. Duration of stay	
V8	F2.BRD ANZAHL JAHRE			1.3.2.2. Duration of stay	
V9	F2.BRD ANZAHL JAHR KATEG			1.3.2.2. Duration of stay	
V10	F2A.WO VORHER GELEBT			1.1.4.1. Country 1.3.2.1. Immigration	
V11	F3.WOHNSTATUS			5.2.1. Tenure status	
V12	F4.NACHTS NICHT ALLEINE			19.4.1. Residential security	
V13	F4A.STADT UNGERN GEHEN			19.4.1. Residential security	
V14	F5A.WERT FAMILIE			4.3.3. Attitude towards family	
V15	F5B.WERT BERUF			7.4.4. Performance	
V16	F5C.WERT FREIZEIT			17.4. Attitude towards leisure	
V17	F5D.WERT FREUNDE			15.2.1. Social networks	

Done Internet