

Hinweise zur Nutzung von Stata mit der Job Submission Application (JoSuA) des Forschungsdatenzentrums im Bundesinstitut für Berufsbildung (BIBB-FDZ)

Stand: November 2021

Version 1 – November 2021

Falls Ihnen während der Arbeit mit JoSuA auffällt, dass Ihnen noch weitere Informationen zur Verwendung von Stata mit JoSuA fehlen, die auch anderen Wissenschaftler/-innen helfen würden, freuen wir uns über Ihre Anregungen zur Weiterentwicklung dieses Dokuments. Bitte senden Sie diese per E-Mail an fdz@bibb.de.

Inhaltsverzeichnis

1.	Datennutzung mit JoSuA	4
2.	Einrichten einer Testumgebung auf dem eigenen Gerät	6
3.	Gestaltung von Auswertungscode	7
3.1.	Datenschutzprüfung	7
3.2.	Nachvollziehbarkeit des Codes	7
4.	Nutzung von Stata mit JoSuA	8
4.1.	Input- und Output-Formate	8
4.2.	Master Do-File	9
4.3.	Einrichtung der Testumgebung – profile.do	9
4.4.	Gestaltung der Do-Files	9
4.5.	Verwendung von Ado-Files	11
4.6.	Besonderheiten ausgewählter Befehle	12
	Das Wichtigste in Kürze	Fehler! Textmarke nicht definiert.

1. Datennutzung mit JoSuA

Das Forschungsdatenzentrum im Bundesinstitut für Berufsbildung (BIBB-FDZ) setzt die webbasierte Software JoSuA (Job Submission Application) für die Datenfernverarbeitung (DFV) und das Hochladen von Syntaxen vor und nach Gastwissenschaftler/-innen-Aufenthalten (GWA) ein. Mit JoSuA können Forschende sensible Datenbestände des BIBB-FDZ unter Einhaltung der derzeit gültigen Datenschutzbestimmungen über einen Remote-Zugriff auswerten. Mit JoSuA wird den Forschenden ermöglicht vergleichsweise unabhängig Betriebsdatensätze und Sondervariablen des BIBB-FDZ für ihre wissenschaftliche Arbeit am eigenen Arbeitsplatz oder im Rahmen eines GWA zu nutzen.

In diesem Dokument werden die wichtigsten Hinweise und Informationen zur Erarbeitung von lauffähigen Stata-Auswertungscodes zum Upload in JoSuA gegeben. Die einzelnen Funktionen und die Nutzeroberfläche in JoSuA werden in einem gesonderten Dokument (Daten- und Methodenbericht) auf der [Website des BIBB-FDZ](#) beschrieben und erklärt. Für eine erfolgreiche Nutzung mit JoSuA empfehlen wir ausdrücklich die Lektüre dieses Dokuments.

1. Datennutzungsvertrag

Nach Abschluss eines [Nutzungsvertrags](#) übermittelt das BIBB-FDZ Ihnen Ihre persönlichen Zugangsdaten für JoSuA. Sie erhalten zwei verschiedene Kennungen und Passwörter, jeweils für die DFV und für einen GWA. Das BIBB-FDZ richtet standardmäßig für alle Nutzenden diese beiden Kennungen ein, auch wenn kein GWA geplant ist. Achten Sie darauf, die beiden Kennungen nicht zu vertauschen. Die Weitergabe der Zugangsdaten an Dritte ist strengstens untersagt.

Wurde die Datennutzung für mehrere Forschende in einem Projekt beantragt, bekommen alle Nutzenden jeweils eigene Zugangsdaten für JoSuA. Alle Nutzenden sind in JoSuA dem gleichen Projekt zugeordnet, sodass eine Zusammenarbeit in der Datenfernverarbeitung möglich ist. Alle Forschenden eines Projekts können in JoSuA sehen, was von anderen Teammitgliedern hochgeladen wurde und auf von anderen Nutzenden im Projekt gespeicherte Arbeitsdatensätze zugreifen.

Sowohl die DFV als auch ein GWA folgen drei weiteren Schritten:

2. Testen mit Spieldaten

Mit Strukturdatensätze (Spieldaten), die in Aufbau und Merkmalsausprägungen den originalen Forschungsdaten entsprechen wird in einer Testumgebung auf dem Gerät der Auswertungscodes entwickelt, bis dieser fehlerfrei durchläuft. Die Spieldaten stehen im [Metadatenportal](#) des BIBB-FDZ zur Verfügung. Für den Fall, dass im Metadatenportal beim entsprechenden Datensatz keine Spieldaten verfügbar sein sollten, werden diese auf Anfrage zur Verfügung gestellt.

In einem GWA wird der Auswertungscodes an einem Computer im BIBB-FDZ mit den Originaldaten (weiter)entwickelt, sodass die Ergebnisse bereits während der Code-Entwicklung einsehbar sind. Alternativ können aber auch Auswertungscodes in der eigenen Testumgebung entwickelt werden und anschließend während des GWA genutzt werden.

3. JoSuA – Internal Mode

Danach wird der Auswertungscodes im **Internal Mode** in JoSuA hochgeladen und abgeschickt. Im Internal Mode läuft der hochgeladene Auswertungscodes automatisiert durch und die Ergebnisse werden mit einem Wasserzeichen versehen angezeigt. Datenschutzrechtlich problematische Auswertungen werden dabei vom System automatisch anonymisiert. Die automatische Datenschutzprüfung greift bei geringen Zellbesetzungen (≤ 3). Diese vorläufigen Ergebnisse des Internal Modes dürfen weder von Dritten eingesehen, noch kopiert oder weiterverwendet werden.

Die Nutzung des Internal Modes hat im Wesentlichen zwei Funktionen. Zum einen dient sie der vorläufigen Sichtung der Auswertungsergebnisse und zum anderen wird mit der Nutzung des Internal Modes sichergestellt, dass der Code auch in JoSuA fehlerfrei durchläuft. Es gibt keine Begrenzung zur Nutzung des Internal Modes. Das heißt, es dürfen unbegrenzt viele Jobs in einem beliebigen Zeitraum abgeschickt werden.

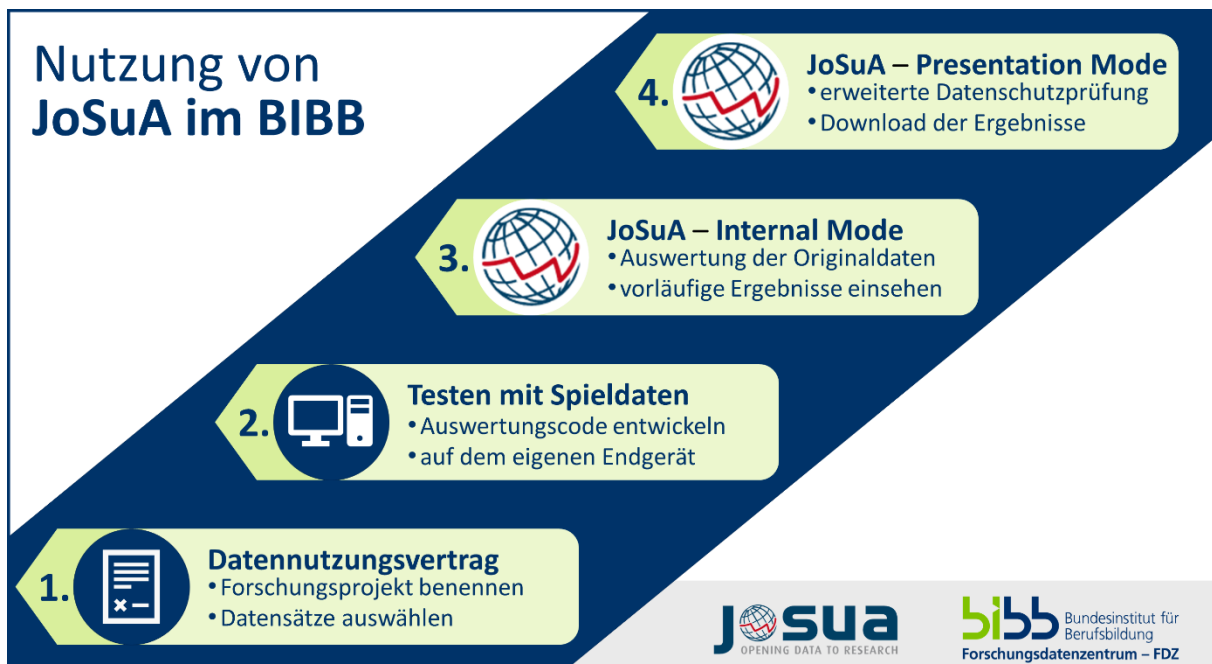
4. JoSuA – Presentation Mode

Wurde der Programmcode erfolgreich in der Testumgebung getestet und wurden im Internal Mode die vorläufigen Ergebnisse gesichtet, muss der Programmcode nun noch einmal im **Presentation-Mode** durchlaufen.

Die Nutzung des Presentation Mode hat im Wesentlichen die Funktion downloadbare Ausgabedateien zur digitalen Weiterverarbeitung in wissenschaftlichen Publikationen zu erzeugen. Die Ergebnisse des Presentation-Mode werden daher nicht nur automatisiert mit JoSuA, sondern auch manuell vom BIBB-FDZ, datenschutzrechtlich geprüft, bevor sie freigegeben werden. Die Prüfung der Ergebnisse erfolgt in der Regel **binnen fünf Werktagen**. Es wird eine Benachrichtigungs-E-Mail verschickt, wenn der Output heruntergeladen werden kann.

Während der Prüfung der Ergebnisse durch das BIBB-FDZ kann kein weiterer Job im Presentation Mode abgeschickt werden.

Abbildung 1 Datenfernverarbeitung mit JoSuA



Quelle: Eigene Grafik.

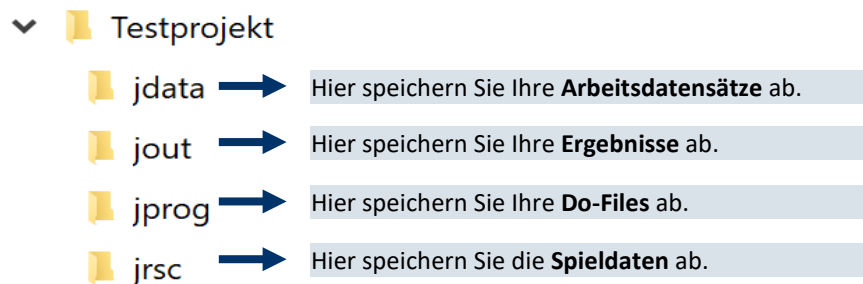
2. Einrichten einer Testumgebung auf dem eigenen Gerät

In der Testumgebung auf dem eigenen Computer wird der Programmcode mit Spieldaten entwickelt und geprüft, sodass die Auswertungen unterbrechungsfrei durchlaufen und keine Syntaxfehler mehr enthalten sind. Die Spieldaten entsprechen in ihrer Struktur den Originaldaten, enthalten aber zufällige, inhaltlich nicht interpretierbare Werte. Spieldatensätze zu den Erhebungen können im [Metadatenportal](#) heruntergeladen werden. Sollten zu dem von Ihnen genutzten Datensatz noch keine Spieldaten verfügbar sein, können Sie uns eine formlose E-Mail an fdz@bibb.de senden und wir erstellen für Sie die entsprechenden Spieldaten.

Wichtig: Benennen Sie die Spieldaten in Ihrer Testumgebung um, sodass die Datensätze genauso benannt sind, wie in der Job Submission Application (JoSuA). In der Regel enden die Bezeichnungen der Spieldaten mit „_spiel“ und die Originaldaten mit „_gwa“. Die Programmcode-Templates enthalten Befehle mit denen Sie eine Liste aller Dateien ausgeben können, die in JoSuA für Sie abgelegt wurden.

Für ein erfolgreiches und übersichtliches Arbeiten mit JoSuA sollte die vom BIBB-FDZ implementierte Ordnerstruktur auf dem eigenen Gerät übernommen werden. Dazu muss ein Ordner mit den folgenden Unterordnern angelegt werden:

Abbildung 2: Ordnerstruktur der Testumgebung



Das BIBB-FDZ hat auf dem JoSuA-Datenserver äquivalente Verzeichnisse mit diesen Funktionen angelegt, die bei der Datenfernverarbeitung entsprechend genutzt werden. Wichtig ist, dass innerhalb der vorgegebenen Ordner keine weiteren Unterordner erstellt werden, denn diese sind dann entsprechend auf dem JoSuA-Datenserver nicht angelegt.

Im Unterschied zum eigenen Endgerät haben Sie auf dem JoSuA-Datenserver, wie unten aufgeführt, nicht für jedes der Verzeichnisse alle Zugriffsrechte, dies sollten Sie bei der Nutzung Ihrer Testumgebung berücksichtigen.

- jdata:** In diesem Ordner können in JoSuA Arbeitsdatensätze gespeichert und wieder eingelesen werden.
- jout:** In diesem Verzeichnis legen Sie in JoSuA Output-Dateien wie Log-Files oder Grafiken ab. Der Inhalt dieses Ordners wird im Internal Mode und im Presentation Mode ausgegeben, sofern die entsprechenden Dateiformate freigegeben sind.
- jprog:** In diesem Verzeichnis werden automatisch Ihre hochgeladenen Do-Files abgelegt. Dieser Ordner ist das Standardarbeitsverzeichnis (default working directory) in JoSuA.
- jrsc:** In diesem Verzeichnis befinden sich die Originaldaten auf dem Server. Für diesen Ordner haben Sie nur Leserechte.

Hinweis zur Verwendung von externen Daten: Grundsätzlich gibt es die Möglichkeit externe Datensätze (z. B. aus der amtlichen Statistik) an die von Ihnen beantragten BIBB-FDZ-Daten zuzuspielen. Diese Datensätze müssen vorab dem BIBB-FDZ zur Verfügung gestellt und geprüft werden. In Ihrer Testumgebung und in JoSuA werden diese Daten im Unterordner jrsc abgelegt. Um das Zuspieren externer Datensätze zu beantragen, schreiben Sie bitte eine E-Mail an fdz@bibb.de.

3. Gestaltung von Auswertungscode

Damit in der Datenfernverarbeitung alles reibungslos funktioniert und die Ergebnisse Ihrer Auswertungen datenschutzkonform übermittelt werden können, gibt es einige Punkte die bei der Gestaltung des Auswertungscode zu beachten sind.

Wichtig: Laden Sie nur Do-Files in JoSuA hoch, die in Ihrer Testumgebung fehlerfrei durchgelaufen sind und erfolgreich einen Output produziert haben.

Nur wenn Ihr Programmcode auf Ihrem Computer erfolgreich getestet wurde, kann dieser auch erfolgreich in JoSuA durchlaufen. Nicht funktionsfähige Jobs blockieren möglicherweise JoSuA-Rechenkapazitäten für andere Forschende.

3.1. Datenschutzprüfung

Ihre Auswertungen durchlaufen verschiedene Schritte der Datenschutzprüfung, welche zum Teil automatisiert sind und zum Teil durch die Mitarbeiter/-innen des BIBB-FDZ erfolgen. Dabei werden alle Ausprägungen kleiner gleich drei sowie sämtliche Werte, die auf bis zu drei Beobachtungen (Personen und/oder Betrieben) basieren, gelöscht. Um das Rückrechnen von gelöschten Werten zu verhindern, werden ggf. weitere Werte gelöscht.

Je nach Anwendungsfall ist es daher ratsam, bereits bei der Programmierung zu berücksichtigen, dass Anteile und Quoten immer mindestens auf vier Beobachtungen basieren müssen. Beispielsweise kann durch eine Wenn-Dann-Bedingung festgelegt werden, dass diese nur berechnet werden, wenn mindestens vier Beobachtungen vorliegen.

Bei Abbildungen können die Fallzahlen entweder direkt in die Grafik oder in einer Tabelle unmittelbar vor oder nach der Grafik aufgeführt werden. Bei Variablen mit Quoten oder Anteilen, muss die Anzahl der gültigen Werte in den einzelnen Subgruppen zusätzlich ausgewiesen werden.

Wichtig: Wenn Stata-Befehle entsprechende Optionen zur Ausgabe der Fallzahlen nicht enthalten oder Optionen zum Ausblenden der Fallzahlen verwendet wurden, werden diese Befehle nicht ausgeführt und die Ergebnisse sind nicht im Log-File enthalten (siehe unten).

Sollen gewichtete Statistiken in den Ergebnissen enthalten sein, müssen die ungewichteten Fallzahlen ausgewiesen werden, sodass die Datenschutzprüfung korrekt durchgeführt werden kann.

Wichtig: Nur, wenn unmittelbar vor einer gewichteten Statistik, der gleiche Befehl mit einer ungewichteten Statistik aufgerufen wurde, wird diese ausgegeben. Unmittelbar bedeutet in diesem Kontext, dass zwischen dem Befehl mit Gewichtung und dem Befehl ohne Gewichtung maximal Leerzeilen oder ein Kommentar im Programmcode enthalten sein dürfen. Wird diese Regel nicht eingehalten, enthält der Log-File an Stelle der gewichteten Statistik den Kommentar „For data protection reasons, this table has been removed.“.

Aus Datenschutzgründen ist es nicht erlaubt, externe Daten im Programmcode selbst einzubinden, wie beispielsweise durch das Aufführen der Werte im Do-File. Wenn externe Datensätze mit den Daten des BIBB-FDZ verknüpft werden sollen, müssen diese über die RESOURCES-Funktion in JoSuA hochgeladen und separat vom BIBB-FDZ freigegeben werden.

3.2. Nachvollziehbarkeit des Codes

Zur effizienten Durchführung der Datenschutzprüfung achten Sie bitte besonders auf die Nachvollziehbarkeit Ihrer Auswertungen. Zielführend sind dabei:

- Kommentare und Überschriften zu den verschiedenen Auswertungsschritten einfügen (z. B. "Aufbereitung der Variable X"),
- selbst generierte Variablen vor Auszählungen kurz beschreiben,
- sprechende Variablennamen verwenden,
- Variablenlabels und Wertelabels vergeben,
- Inhalte von Abbildungen und komplexen Tabellen in einem Kommentar beschreiben,
- die Inhalte und Funktion von selbst programmierten Schleifen und deren Ausgabe beschreiben (u. a. durch automatisierte Tabellenüberschriften als Teil der Schleife),
- Deskription der einzelnen verwendeten Variablen vor multivariaten Auswertungen.

4. Nutzung von Stata mit JoSuA

In der Job Submission Application (JoSuA) ist aktuell Stata 16 installiert. Damit es nicht zu versionsspezifischen Fehlermeldungen kommt, empfehlen wir auch in der Testumgebung Stata 16 zu nutzen. Falls sie eine neuere Version von Stata für die Testumgebung nutzen, prüfen Sie beim Erarbeiten Ihres Do-Files kritisch, ob es zu Kompatibilitätsproblemen kommen kann. In der Regel ist Stata abwärtskompatibel und ermöglicht ein reibungsloses Durchlaufen Ihrer Syntax in JoSuA, wenn Sie Ihre Syntax mit einer Vorgängerversion erstellt haben.

Für die Nutzung von Stata stehen Ihnen auf der Website des BIBB-FDZ einige Beispiel-Do-Files (Templates) zur Verfügung, an denen Sie sich orientieren sollten. Im Folgenden werden die wichtigsten Abschnitte dieser Beispielsyntax erklärt.

4.1. Input- und Output-Formate

Aus Datenschutzgründen sind nicht alle Dateiformate für den Upload oder die Speicherung von Ergebnissen mit Stata zugelassen. Folgende Dateiformate dürfen genutzt werden:

Input (Do-Files):	.do	.ado					
Output Internal Mode:	.log	.txt.	.png				
Output Presentation Mode:	.log	.txt	.png	.csv	.doc	.docx	.eps
	.gph	.pdf	.svg	.wmf	.xls	.xlsx	

Wichtig: Zugelassen sind nur Log-Files im reinen Text-Format, smcl ist nicht erlaubt.

Wird im Internal Mode eine Datei in einem der nur im Presentation Mode zugelassenen Output-Formate abgespeichert, dann wird diese unter Umständen in JoSuA als Textdatei interpretiert und entsprechend falsch dargestellt. Von diesem Problem ist der Output im Presentation Mode unberührt. Nach erfolgreicher Datenschutzprüfung können die Dateien in diesem Modus heruntergeladen und geöffnet werden.

Wichtig: Es können maximal 20 Output-Dateien erzeugt werden. Diese dürfen jeweils die Größe von 2 MB nicht überschreiten.

Aktuell werden im Output des Internal Modes bestimmte Zeichen nicht richtig dargestellt (z. B. ä, ö, ü, ß). Dieses Problem ist bekannt und wird mit einem der nächsten Updates in JoSuA behoben. Im Output des Presentation Mode werden diese Zeichen richtig dargestellt.

Hinweis für Nutzende der Software LaTeX: Eine direkte Ausgabe von Output im .tex-Format ist in JoSuA nicht möglich. Allerdings lässt sich bei Verwendung des `esttab`-Befehls in Kombination mit der Option `tex` oder `booktabs` ein Output im .txt-Format erzeugen, der sich für in LaTeX verfasste Dokumente nutzen lässt.

Wichtig: Der `esttab` –Befehl muss für die Datenschutzprüfung jeweils einmal ohne und einmal mit der Option `tex/booktabs` ausgeführt werden. Ansonsten wird der LaTeX-Code nicht freigegeben.

4.2. Master Do-File

Grundsätzlich führt JoSuA nur Do-Files mit dem Namen master.do aus. Über entsprechende Befehle in diesem Master Do-File werden weitere Do-Files ausgeführt, die ebenfalls hochgeladen werden müssen. Dahinter steckt eine verbreitete Praktik zur Gestaltung von Programmen. Programmcodes für verschiedene, klar abgrenzbare Analyseschritte (z.B. Variablenaufbereitung, Deskription, Regression) werden in verschiedenen Dateien gespeichert, welche dann der Reihe nach von der Master-Datei aus aufgerufen und ausgeführt werden. Dieses Vorgehen dient dazu, den Code in übersichtliche Abschnitte einzuteilen und möglicherweise einzelne bereits durchgeführte, fehlerfreie Analyseschritte überspringen zu können. Zudem können globale Einstellungen übersichtlich in der Master-Datei notiert werden. Ob und in welchem Umfang bei den Auswertungen dieser Logik gefolgt wird, ist nicht vorgegeben.

Wichtig: Es muss immer eine Datei mit der Bezeichnung master.do hochgeladen werden, damit der Job abgeschickt werden kann.

4.3. Einrichtung der Testumgebung – profile.do

Bevor Do-Files in JoSuA hochgeladen werden sollten, muss in der Testumgebung auf dem eigenen Endgerät überprüft werden, ob diese fehlerfrei durchlaufen (s. o.).

In Stata gibt es die Möglichkeit die dazu angelegten Verzeichnisse (jdata, jout, jprog, jrsc; s. o.) über globale Makros zu definieren, sodass sie angesprochen werden können ohne den Dateipfad jedes Mal eingeben zu müssen oder ihn überhaupt zu kennen. Auf diese Weise kann auf dem JoSuA-Server auf die Verzeichnisse zugegriffen werden, ohne Datei-Pfade im Programmcode anpassen zu müssen.

Damit Do-Files aus der Testumgebung änderungsfrei in JoSuA durchlaufen, müssen daher globale Makros für den Zugriff auf die Verzeichnisse definiert werden. Das funktioniert mit dem Befehl `global`, wie unten aufgeführt. Tauschen Sie „C:\Testprojekt“ entsprechend gegen den Pfad zu dem von Ihnen angelegten Order aus.

```
global jrsc "C:\Testprojekt\jrsc\  
global jout "C:\Testprojekt\jout\  
global jprog "C:\Testprojekt\jprog\  
global jdata "C:\Testprojekt\jdata\  

```

Nachdem die Makros angelegt wurden, können die Verzeichnisse mit `$jrsc` usw. aufgerufen werden. Das Öffnen des Datensatzes „Qualifizierungspanel_2017_gwa.dta“ im Ordner jrsc würde dann beispielsweise so funktionieren:

```
use "$jrsc/Qualifizierungspanel_2017_gwa.dta", clear
```

Es gibt die Möglichkeit, dass Stata beim Öffnen automatisch diese Befehle ausführt. Dafür muss ein Do-File mit dem Namen „profile.do“ angelegt werden, welcher die entsprechenden Befehle enthält. Wir empfehlen in der Testumgebung ein solches Do-File zur Festlegung der globalen Makros zu verwenden.

Stata sucht beim Programmstart an verschiedenen Speicherorten nach profile.do. Wo Sie Ihr profile.do mit der Festlegung der globalen Makros ablegen, liegt bei Ihnen. Im [Stata-Manual](#) wird empfohlen, profile.do unter „~/Documents/Stata“ abzulegen. Wurde profile.do erfolgreich erkannt und ausgeführt, wird die folgende Meldung beim Start von Stata angezeigt.

```
running profile.do
```

4.4. Gestaltung der Do-Files

Jeder Stata-Job muss in JoSuA ein Do-File mit der Bezeichnung master.do enthalten. Diese Datei wird automatisch ausgeführt, nachdem der Job abgeschickt wurde. JoSuA öffnet automatisch eine master.log-Datei, daher darf im Code nicht noch eine master.log-Datei geöffnet werden.

Wichtig: Öffnen Sie keinen master.log, dieser wird in JoSuA automatisch gestartet. Falls Sie versuchen einen master.log zu öffnen, bricht der Job ab und es erscheint die Fehlermeldung "log-file already open".

Zu Beginn des master.do sind einige allgemeine Informationen vermerkt und es werden allgemeine Einstellungen für die folgenden Auswertungen festgelegt:

```
*Informationen und Einstellungen zum Do-File
**Autor*in:  Maxi Muster
**Projekt:   Testprojekt
**Datum:    04.05.2050
**Dateiname: master.do

version 16          //Aktuell in JoSuA installierte Stata Version
clear
set more off       // notwendig damit das Programm durchlauft

set linesize 120  // Max.Zeilenlaenge Internal Mode
*set linesize 255 // Max.Zeilenlaenge Presentation Mode
```

Wir empfehlen im Internal Mode die Zeilenlänge mit dem Befehl set linesize auf 120 zu setzen, da dies der Breite des Output-Fensters in JoSuA entspricht.

Wichtig: Verwenden Sie an dieser Stelle im master.do noch nicht den Befehl capture log close. Andernfalls werden alle Fehler, die vor dem ersten "log open"-Befehl aufgetreten sind, nicht angezeigt. Nutzen Sie diesen Befehl nur direkt vor dem Öffnen einer neuen Log-Datei!

Nach diesen Grundeinstellungen sollen die Do-Files mit den verschiedenen Auswertungsschritten ausgeführt werden. Nummerieren Sie die Bezeichnungen Ihrer Do-Files in der Reihenfolge, in der Sie ausgeführt werden sollen und kommentieren Sie kurz deren Funktion.

```
**Aufrufen der Do-Files mit den Auswertungsschritten

do "$jprog/01_teildatensatz.do" // bilden eigener Variablen
do "$jprog/02_deskriptiv.do"   // deskriptive Statistiken berechnen
do "$jprog/03_analysen.do"     // multivariate Regression
do "$jprog/04_plot.do"        // Abbildungen
```

Die Einteilung des Programmcodes in verschiedene Analyseschritte hat den Vorteil, dass Sie schnell ersichtlich wird an welcher Stelle der Auswertungen mögliche Fehlermeldungen auftreten. Zudem müssen Analyseschritte, die bereits erfolgreich durchgelaufen sind, nicht bei jedem Job erneut berechnet werden, sondern können einfach auskommentiert werden.

Zu Beginn jedes Do-Files muss ein Log-File geöffnet werden, welches genauso benannt ist wie das Do-File. Die ersten Zeilen des 01_teildatensatz.do sehen demnach so aus:

```
cap log close      // immer vor dem Öffnen eines neuen Log-File verwenden
log using "$jout/01_teildatensatz.log", replace

*Informationen und Einstellungen zum Do-File
**Autor*in:       Maxi Muster
**Datum:          04.05.2050
**Dateiname:     01_teildatensatz.do

**Datensatz öffnen
use "$jrsc/Qualifizierungspanel_2017_gwa.dta", clear
```

So wird zu jedem Do-File ein entsprechendes Log-File erstellt, welches entsprechend separat im JoSuA - Internal Mode angezeigt wird und im Presentation Mode heruntergeladen werden kann.

Im master.do empfehlen wir das Anlegen eines Log-Files, welches eine Dateiliste zu den vier Verzeichnissen ausgibt. So ist jederzeit klar, welche Dateien mit welchen Bezeichnungen existieren und welche Datensätze und Auswertungsprogramme in Ihrem Projekt verwendet werden.

```
log using "$jout/filelist.log", replace // Log-File für die Dateiliste öffnen

**Ausgabe einer Dateiliste

/*Hinweis:
   Die Dateiliste beinhaltet alle Dateinamen.
   Es sind auch die genauen Bezeichnungen der für Sie im Verzeichnis $jrsc
   abgelegten Originaldaten enthalten.
   Gleichen Sie diese Dateinamen mit den Bezeichnungen der Spieldaten in Ihrer
   Testumgebung ab
   und benennen Sie die Dateien in Ihrer Testumgebung um, sodass die Datensätze
   **genau gleich** benannt sind.
   Andernfalls werden Ihre Do-Files auf JoSuA nicht durchlaufen. */

dir "$jprog/"
dir "$jout/"
dir "$jdata/"
dir "$jrsc/"

log close
```

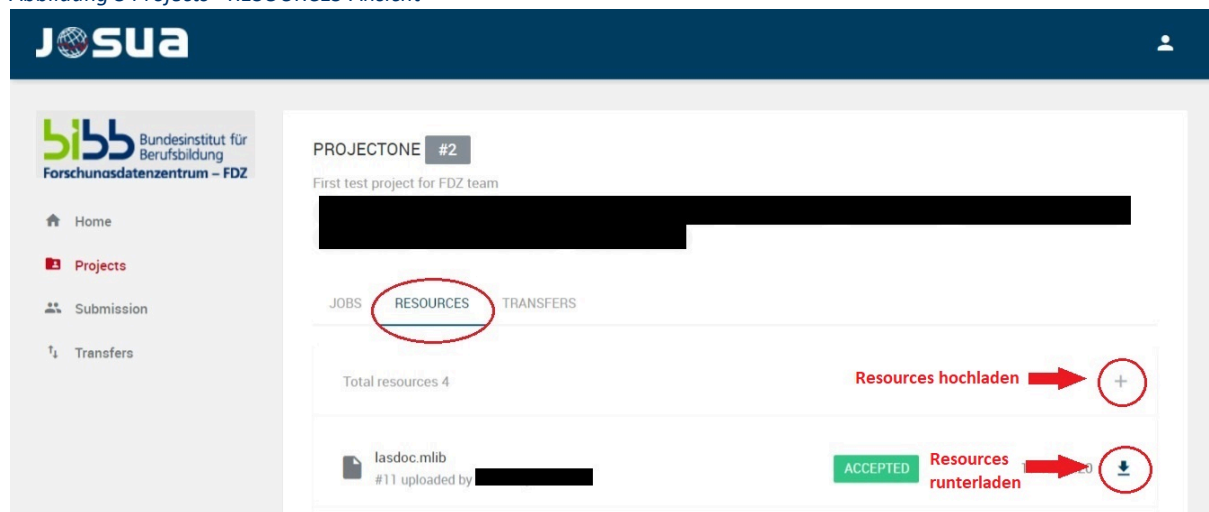
4.5. Verwendung von Ado-Files

Werden Ado-Files (Befehle, die nicht standardmäßig in Stata implementiert sind) benötigt, gibt es zwei verschiedene Möglichkeiten diese zu nutzen.

Temporäre Nutzung: Upload des Ado-Files gemeinsam mit den Do-Files unter dem Reiter „Submission“. Das Ado-File steht dann nur für diesen Job zur Verfügung.

Permanente Nutzung: Upload des Ado-Files für das gesamte Projekt. Dazu öffnet sich unter dem Reiter „Resources“ mit einem Klick auf das Plusymbol (+) rechts ein Uploadfenster (siehe Abbildung). Alle hochgeladenen Dateien müssen danach vom BIBB-FDZ freigegeben werden und stehen danach Projektmitgliedern zur Verfügung.

Abbildung 3 Projects - RESOURCES-Ansicht



Im Unterschied zur temporären Nutzung ist es bei der permanenten Nutzung über die Resources-Funktion auch möglich Ados mit anderen Dateiformaten als .ado zu nutzen (z. B. .mlib; .sthlp). Für Ados, die diese Dateiformate beinhalten ist daher nur die permanente Nutzung möglich.

Die meisten und häufig verwendeten Stata-Ados werden im [Statistical Software Components Archive](#) geführt und können dort auch heruntergeladen werden. Auf dieses Archiv wird auch mit dem Befehl `ssc install` zugegriffen, welcher in JoSuA nicht genutzt werden kann. Es ist in Planung etablierte Ados zukünftig in JoSuA automatisiert zur Verfügung zu stellen.

Wichtig: Der Befehl `ssc install` steht aktuell noch nicht zur Verfügung, laden Sie die Ados einzeln aus herunter und wie beschrieben bei JoSuA hoch. Achten Sie darauf Ados mit den richtigen Dateieinstellungen abzuspeichern (z. B. `.ado`), nachdem Sie Dateien aus dem SSC Archiv heruntergeladen haben.

Um den Überblick zu behalten, welche Ados in den Auswertungen verwendet werden und entsprechend in JoSuA zur Verfügung stehen, ist es empfehlenswert in der Testumgebung einen neuen Adopath anzulegen, der auf den Order `jprog` verweist. Wenn alle anderen Adopath zusätzlich ausgeschaltet werden, findet Stata nur noch Ados in `jprog`.

Der entsprechende Befehl sollte im `profile.do` aufgeführt werden, damit Stata die Ados auch immer in der Testumgebung findet.

```
** Optional: adopath setzen *** //
adopath ++ "$jprog"
adopath - PERSONAL
adopath - PLUS
adopath - OLDPLACE
adopath - "."
exit
```

4.6. Besonderheiten ausgewählter Befehle

Aus Datenschutz- und Sicherheitsgründen stehen einige Stata-Befehle in JoSuA nicht zur Verfügung oder sie sind nur eingeschränkt nutzbar. Nachfolgend werden diese Befehle aufgeführt:

#delimit

Der Befehl `#delimit` darf nicht verwendet werden, da es für die automatische Datenschutzprüfung notwendig ist, dass der Zeilenumbruch das Ende des Befehls markiert. Wenn innerhalb von Befehlen Zeilenumbrüche verwendet werden sollen, dann müssen diese Umbrüche mit `///` erstellt werden.

cd

In der JoSuA-Umgebung können nur die vordefinierten Verzeichnisse (`jdata`, `jout`, `jprog`, `jrsc`) verwendet werden. Eine Änderung der Verzeichnisse mit dem Befehl `cd` ist daher nicht möglich.

exit

Der Befehl `exit` beendet das Programm und kann daher in der JoSuA-Umgebung nicht verwendet werden.

global

In der JoSuA-Umgebung können keine globalen Macros neudefiniert oder selbst angelegt werden.

graph save

Der Befehl `graph save` zur Speicherung von `.gph` Dateien darf nur mit der Option `asis` verwendet werden. Sollen mehrere Grafiken kombiniert werden, muss dies ohne Zwischenspeichern der Grafiken geschehen indem die Grafiken innerhalb eines Jobs mit der Option `name()` benannt und entsprechend kombiniert werden.

list

Der Befehl `list` ist aus Datenschutzgründen gesperrt.

log

Bei dem Befehl `log` muss die Option `txt` mit angegeben werden. Die Ausgabe von Log-Files im `smcl`-Format ist nicht zulässig. Darüber hinaus können Log-Files nur in dem Verzeichnis `jout` abgespeichert werden.

save

Mit dem `save`-Befehl können nur Dateien in den unter 4.1. genannten Output-Formaten ausgegeben und im Verzeichnis `jout` abgespeichert werden.

shell

Da bei dem Befehl `shell` direkt mit dem Betriebssystem interagiert wird, darf dieser in der JoSuA-Umgebung nicht genutzt werden. Gleiches gilt für weitere Befehle, die unmittelbar auf das Betriebssystem zugreifen, z. B. `ashell`.

ssc install

Der Befehl steht aktuell aus technischen Gründen noch nicht zur Verfügung, soll aber zukünftig zur Verfügung gestellt werden.

table

In den Optionen muss immer `n`, `count` oder `freq` angegeben werden. Nur, wenn unmittelbar vor einer gewichteten Tabelle, der gleiche Befehl mit einer ungewichteten Tabelle aufgerufen wurde, wird diese ausgegeben (s. oben).

tabstat

In den Optionen muss immer `n`, `count` oder `freq` angegeben werden. Nur, wenn unmittelbar vor einer gewichteten Tabelle, der gleiche Befehl mit einer ungewichteten Tabelle aufgerufen wurde, wird diese ausgegeben (s. oben).

tabulate

Die Option `nofreq` darf nicht genutzt werden. Nur, wenn unmittelbar vor einer gewichteten Tabelle, der gleiche Befehl mit einer ungewichteten Tabelle aufgerufen wurde, wird diese ausgegeben (s. oben).

use

Mit dem Befehl `use` können nur die im Verzeichnis `jrcs` befindlichen Originaldaten sowie selbst erstellte, unter `jdata` abgelegte Teildaten genutzt werden.

Wichtig: Wird eine dieser Regeln irrtümlicherweise nicht eingehalten, dann wird der Befehl nicht ausgeführt und es erscheint die Meldung `*** CENSORED ***` im Log-File.