

Anwendungs-Programmsystem DAPROA

Verarbeitungsprogramm Baustein SELSATZ

Version 5 .1

Stand: 31.10.2011

Leistung:

Selektieren(Löschen) von Datensätzen aus der Datenmatrix eines Projektes und Bildung der Standard- bzw. Basisdateien eines neuen Unterprojektes ohne diese gelöschten(selektierten) Datensätze.

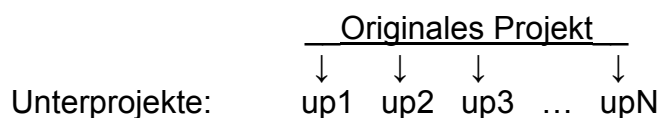
So zum Beispiel wenn bei der Analyse der Daten Variable nur in bestimmten Wertebereichen untersucht, oder Sätze mit Fehldaten ausgeschlossen werden sollen.

Das Löschen geschieht dadurch, dass für ausgewählte Variablen ein Wertebereich festgelegt werden kann. Liegt eine dieser Variablen in diesem Bereich, wird der jeweilige Satz gelöscht.

Sollen bei einer ausgewählten Variablen fehlende Werte, markiert durch -1.D50, als Datensatz-Löschkriterium dienen, dann werden die Bereichsgrenzen durch einen Stern * gekennzeichnet.

Für den Fall, dass bei einer Variablen die Selektion(Löschung) bezüglich mehrerer Wertebereiche erfolgen soll, muss der Vorgang wiederholt werden. Das entstandene neue Unterprojekt wird Eingabeprojekt, die neue Datenmatrix gehört zu einem weiteren neuen Unterprojekt.

Im folgenden wird im Zusammenhang mit Dateinamen immer von " unterprojekt" gesprochen, auch wenn es sich um Dateien des Originalprojektes handelt In der Verarbeitung wird das Originalprojekt wie ein Unterprojekt behandelt.



Eingabedateien:

- **projekt_nqd.txt** geordnete normierte Datei im Textformat
- **projekt_mtx.txt** als Datenmatrix.
- **projekt_std.txt** Statistik-Steuerdatei mit berechneten statistische Kenngrößen des Datenbestandes:
Mittelwert, Streuung, kleinster und größter Wert der Variablen,
einfache Korrelationskoeffizienten zwischen den Variablen.'

Abarbeitung:

SELSATZ liest die Projektarbeitungsdatei **projekt_abd.txt ein** und startet die Abarbeitung.

Sie befindet sich im Verzeichnis **..\daten\ projekt**

In dieser Datei mit dem Bausteinmerkmal <selsatz > befinden sich die Sätze mit den Bausteinsteuerdaten.

Aufbau der Datei **projekt_abd.txt** für den Baustein SELSATZ:

Steuermerkmal	Spalte	Typ	Parameter	Inhalt des Parameters
<Steuerparameter >	01-03	Int	ix	Index einer Selektionsvariablen
	05-17	Real	xu	untere Wertebereichsgrenze für den Selektionsbereich
	19-31	Real	xo	obere Wertebereichsgrenze für den Selektionsbereich
	41-43	Int	ix	Index einer Selektionsvariablen
	45-57	Real	xu	untere Wertebereichsgrenze für den Selektionsbereich
	59-71	Real	xo	obere Wertebereichsgrenze für den Selektionsbereich
<Selektionsupname >	01-08	Char	uprneu	Name des entstehenden neuen Unterprojektes
	10-17	Char	uprnam	Name des Eingabe- Unterprojektes

Das Bausteinmerkmal <selsatz > muss genau 10 Zeichen lang sein (spitze Klammern mitgezählt) .

Die Steuermerkmale müssen genau 21 Zeichen lang sein (spitze Klammern mitgezählt) .

Sie müssen, wie in Beispiel und Aufbaubeschreibung dargestellt, geschrieben, also auch mit der entspr. Anzahl von Leerzeichen aufgefüllt werden.

Parameter, Grenzen:

- Es werden maximal 100 Variable/Messgrößen verarbeitet.
- Es können maximal 20 Variable für die Selektionskriterien ausgewählt werden.
- Die Anzahl der Datensätze ist faktisch unbegrenzt.

Ergebnisdateien:

- **neues_unterprojekt_nqd.txt** geordnete normierte Datei im Textformat
- **neues_unterprojekt_mtx.txt** als Datenmatrix.
- **neues_unterprojekt_std.txt** Statistik-Steuerdatei mit berechneten statistische Kenngrößen des Datenbestandes:
Mittelwert, Streuung, kleinster und größter Wert der Variablen, einfache Korrelationskoeffizienten zwischen den Variablen.'
- **neues_unterprojekt_feh.txt** fehlende Variablen- oder Messwerte
- **selsatz_mld.txt** Programmnachrichten(Ablauf, Fehler, ...)

Dialogbaustein SELSTZ_e

In dem Dialogbaustein werden die Parameter für die Projektarbeitungsdatei **projekt_abd.txt** abgefragt, eingegeben und geprüft.

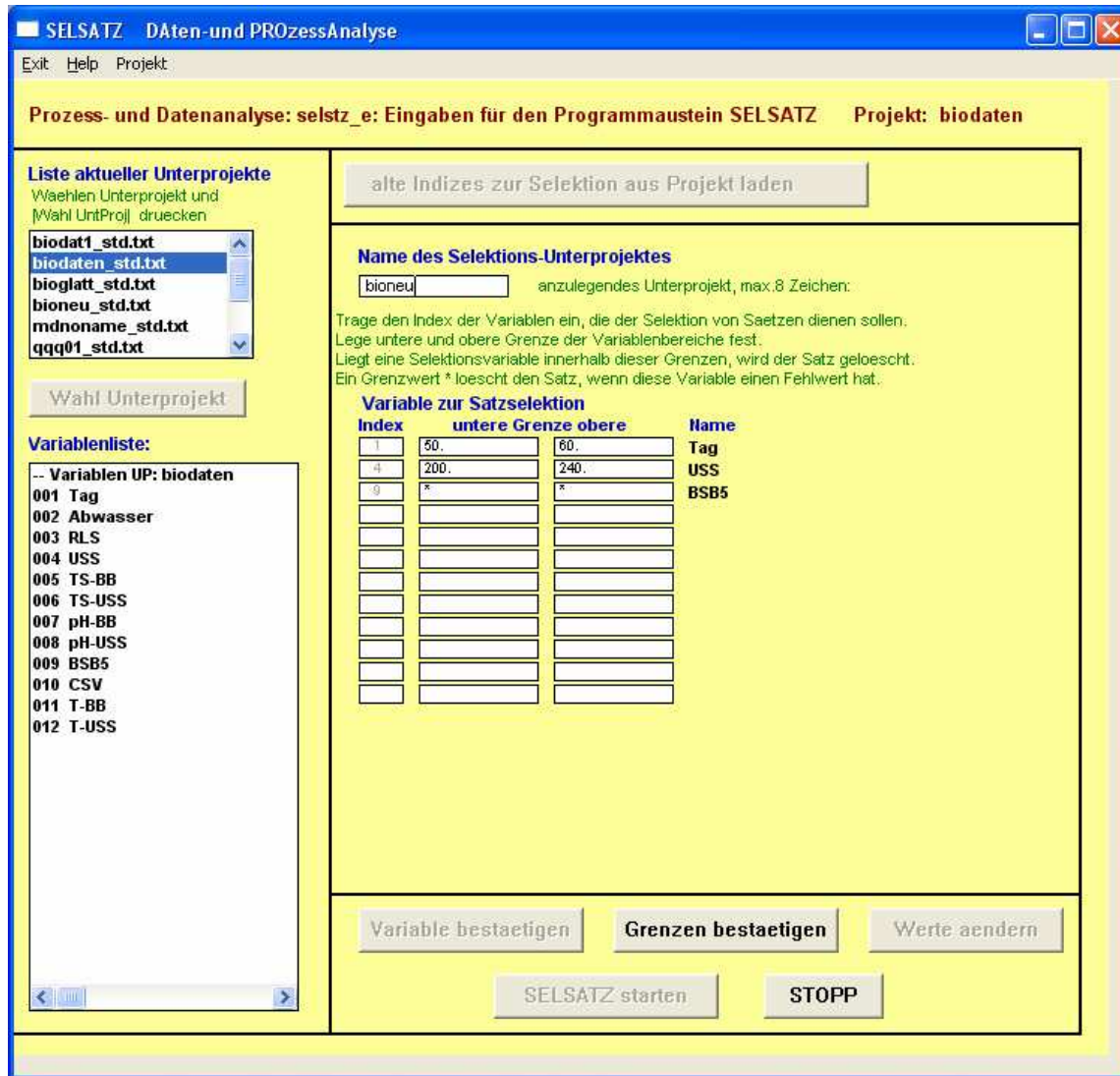


Bild 1:
Dialogfenster von SELSTZ_e ,
Eingabeunterprojekt **biodaten**
Ergebnisunterprojekt: **bioneu**

Mit dem Dialogbaustein SELSTZ_e werden die Selektionssteuersätze aufgebaut. Es ist sehr wichtig sich den "Selektionsplan" vorher zurecht zu legen und auf Widerspruchsfreiheit zu prüfen.

Man kann auch mit einem ersten Teil der Selektion(Löschung) beginnen und nach der Verarbeitung dieser Anweisungen zu einem neuen Unterprojekt mit einem weiteren Teil der Selektion auf die veränderten Dateien zu einem weiteren neuen Unterprojekt fortsetzen.

Wahl des Unterprojektes

Der Dialogbaustein zeigt die Liste **|Liste aktueller Unterprojekte|** der aktuell verfügbaren Unterprojekte bzw. des originären Hauptprojektes an. Aus dieser Liste kann das Projekt gewählt, dessen Daten als Ausgangsgrößen für die Selektion herangezogen werden. Bestätigen dieser Auswahl mit der Taste **Wahl Unterprojekt**.

Die Prüfung, ob alle notwendigen Unterprojekt-Eingabedateien bereitstehen und fachlich richtig sind, obliegt dem Anwender. Ist dies nicht der Fall, bricht das Programm ab.

In der **|Variablenliste:|** werden die Variablen des Unterprojektes angezeigt. Das originäre Hauptprojekt wird beim Start automatisch als gewählt gesetzt.

Im Eingabefeld **[Name des Selektions-Unterprojektes]** wird den Ergebnisdateien der Selektion ein Unterprojektnamen gegeben. Sie werden im originalen Projektverzeichnis angelegt. Dieses Feld muss ausgefüllt werden. Der Name darf nicht gleich dem Namen des vorher ausgewählten Unterprojektes (Eingabedateien) sein.

Festlegen der Variablen, deren Wertebereiche als Selektions(Lösch) -kriterium dienen

In der Eingabespalte **[Index]** werden die Indizes als ganze Zahlen der festzulegenden Variablen eingetragen, die Indizes werden der **|Variablenliste:|** entnommen.

Mit der Taste **alte Indizes zur Selektion aus Projekt laden** können auch die alten Selektionssteuersätze geladen und in den Dialogmasken vorgelegt werden, soweit welche zum ausgewählten Eingabeunterprojekt vorhanden sind.

Mit Taste **Variable bestaetigen** werden die ausgewählten Variablen festgelegt. SELSTZ_e prüft nun die eingetragenen Indizes auf Plausibilität und legt für die unteren und oberen Selektions(Lösch) –bereichsgrenzen Werte vor, soweit der Variablenindex neu eingetragen wurde.

Beachte: Bei der automatischen Vorlage der Selektions(Lösch) –bereichsgrenzen für eine neu eingetragene Variable werden **immer** der Minimal- und der Maximalwert der jeweiligen Variablen eingetragen. Diese Werte **müssen** im nächsten Schritt auf die gewünschten Bereichsgrenzen geändert werden.

Nun können in den Eingabespalten [untere Grenze obere] die Werte für die Selektions(Lösch) –bereichsgrenzen als Dezimalzahlen eingetragen werden. Wenn in einem dieser Felder für eine Variable ein Stern * eingetragen wird, legt man fest, dass alle Datensätze gelöscht werden, in denen diese Variable einen Fehlwert besitzt. Ein solcher Fehlwert ist in der Datenmatrix mit -1D50 markiert. Mit Taste **Grenzen bestaetigen** werden die ausgewählten Selektions(Lösch) –bereichsgrenzen festgelegt. SELSTZ_e prüft nun die eingetragenen Grenzen auf Plausibilität.

Mit Taste **Werte aendern** bietet sich die Möglichkeit alle Eingaben zu ändern, beginnend mit der Auswahl der Variablen.

Nun kann der Baustein SELSATZ durch Betätigung der Taste **SELSATZ starten** gestartet werden. SELSATZ gibt eine Nachricht über Erfolg oder Misserfolg der Verarbeitung.

Wenn als Ergebnis der Selektion(Löschung) weniger als **drei** Datensätze überbleiben, wird kein neues Unterprojekt angelegt.

Nach Beendigung von SELSATZ kehrt die Steuerung zum Dialogbaustein SELSTZ_e zurück.

Mit der Taste **STOPP** wird der Dialog SELSATZ_e beendet.

Beispiel

Steuersätze für <selsatz > in der Ablaufsteuerdatei **biodaten_abd.txt**

```
<selsatz ><Steuerparameter      > 1  0.500000D+02  0.600000D+02          4  0.200000D+03  0.240000D+03
<selsatz ><Steuerparameter      > 9  *                    *
<selsatz ><Selektionsupname     >bioneu  biodaten
```

Die Steuerdatei des Eingabeunterprojektes **biodaten_std.txt** hat folgenden Inhalt :

```
biodaten
  12      60      32
Tag      0.100000D+01 0.600000D+02 0.305000D+02 0.174642D+02      60
Abwasser 0.710000D+04 0.800000D+04 0.751833D+04 0.202937D+03      60
RLS      0.380000D+04 0.467100D+04 0.392518D+04 0.236532D+03      60
USS      0.285000D+02 0.240000D+03 0.105107D+03 0.381208D+02      60
TS-BB    0.288000D+01 0.715000D+01 0.480931D+01 0.897724D+00      58
TS-USS   0.829000D+01 0.175200D+02 0.128890D+02 0.239951D+01      60
pH-BB    0.640000D+01 0.740000D+01 0.701404D+01 0.219934D+00      57
pH-USS   0.620000D+01 0.750000D+01 0.676500D+01 0.221570D+00      60
```

DAPROA: Daten- und Prozessanalyse
Bausteine SELSATZ / SELSTZ_e

```

BSB5          0.385000D+03 0.140000D+04 0.688769D+03 0.173310D+03      39
CSV           0.210000D+03 0.430000D+03 0.316241D+03 0.478835D+02      54
T-BB         0.220000D+02 0.270000D+02 0.243500D+02 0.110966D+01      60
T-USS        0.195000D+02 0.235000D+02 0.216750D+02 0.942522D+00      60
 1.000000    0.478946    1.000000    0.208663    0.061262    1.000000   -0.078056   -0.030711
-0.290549    1.000000    0.274532    0.221913    0.121810   -0.453690    1.000000    0.281149
 0.444312   -0.050569   -0.378615    0.770943    1.000000    0.155454   -0.033310    0.239741
-0.035314    0.063487   -0.055651    1.000000    0.316465    0.274603    0.129648   -0.088566
 0.101285    0.037934    0.544190    1.000000   -0.250939   -0.177598   -0.264207    0.202229
 0.000942   -0.099353   -0.332086   -0.252167    1.000000   -0.193790   -0.310499   -0.226678
 0.323604   -0.387531   -0.533428   -0.008396   -0.112191    0.507535    1.000000   -0.485837
-0.187034    0.058644   -0.048057   -0.164136   -0.103211   -0.018756   -0.183714   -0.404482
 0.011276    1.000000   -0.296808   -0.256311    0.165288    0.048060   -0.111043   -0.185332
 0.094291   -0.140611   -0.277891    0.067094    0.775032    1.000000

```

Nach der Verarbeitung mit obigen Steuersätzen durch SELSATZ, sie entsprechen den Dialogeingaben, erscheint die Meldung



Die Steuerdatei des Ergebnisunterprojektes **bioneu_std.txt** hat nun folgenden Inhalt :

```

biodaten
 12      33      3
Tag      0.400000D+01 0.490000D+02 0.264848D+02 0.140937D+02      33
Abwasser 0.710000D+04 0.800000D+04 0.750303D+04 0.222886D+03      33
RLS      0.380000D+04 0.451200D+04 0.396282D+04 0.241656D+03      33
USS      0.285000D+02 0.175000D+03 0.952121D+02 0.269939D+02      33

```

DAPROA: Daten- und Prozessanalyse
Bausteine SELSATZ / SELSTZ_e

TS-BB	0.364000D+01	0.715000D+01	0.497969D+01	0.884579D+00				32
TS-USS	0.831000D+01	0.175200D+02	0.130761D+02	0.237653D+01				33
pH-BB	0.640000D+01	0.730000D+01	0.697576D+01	0.192078D+00				33
pH-USS	0.620000D+01	0.750000D+01	0.674242D+01	0.261044D+00				33
BSB5	0.385000D+03	0.140000D+04	0.708818D+03	0.177061D+03				33
CSV	0.210000D+03	0.425000D+03	0.313871D+03	0.478440D+02				31
T-BB	0.220000D+02	0.260000D+02	0.243788D+02	0.102340D+01				33
T-USS	0.195000D+02	0.235000D+02	0.216667D+02	0.940966D+00				33
1.000000	0.594417	1.000000	0.360632	0.120806	1.000000	-0.502405	-0.306816	
-0.324955	1.000000	0.378082	0.266711	0.201412	-0.527427	1.000000	0.412192	
0.559484	-0.095704	-0.561125	0.691377	1.000000	0.056424	-0.158818	0.451114	
-0.191181	0.089001	-0.182795	1.000000	0.452909	0.287755	0.233451	-0.257691	
0.175915	0.019167	0.507284	1.000000	-0.098130	-0.183537	-0.317792	0.403373	
-0.079516	-0.143646	-0.360970	-0.237546	1.000000	-0.291774	-0.421780	-0.124225	
0.378993	-0.436625	-0.570358	-0.097509	-0.147912	0.482837	1.000000	-0.378203	
-0.135340	0.154825	-0.032524	-0.308302	-0.218888	0.342276	-0.044486	-0.468519	
0.027534	1.000000	-0.364458	-0.285588	0.258436	0.084316	-0.322907	-0.440100	
0.429372	-0.074213	-0.254339	0.198171	0.743673	1.000000			

KURZANLEITUNG

- (1) Übernahme des vorgelegten Unterprojektes (dazu ist keine Aktion nötig)
oder
Auswahl eines anderen Unterprojektes aus der Liste
|Liste aktueller Unterprojekte| und bestätigen mit der Taste **Wahl Unterprojekt**
- (2) In das Eingabefeld **Name des Selektions-Unterprojektes** Eintragen eines Unterprojektnamens. Dieser Name muss sich vom gewählten Eingabeunterprojekt unterscheiden.
- (3) In der Eingabespalte **Index** werden die Indizes als ganze Zahlen der festzulegenden Variablen eingetragen, die Indizes werden der **Variablenliste:** entnommen.
- (4) Mit Taste **Variable bestaetigen** werden die ausgewählten Variablen festgelegt.
- (5) In den Eingabespalten **untere Grenze obere** werden die Werte für die Selektions(Lösch) –bereichsgrenzen als Dezimalzahlen eingetragen. Wenn in einem dieser Felder für eine Variable ein Stern * eingetragen wird, legt man fest, dass alle Datensätze gelöscht werden, in denen diese Variable einen Fehlwert besitzt.
- (6) Mit Taste **Grenzen bestaetigen** werden die ausgewählten Selektions(Lösch) –bereichsgrenzen festgelegt
- (7) Start des Bausteins SELSATZ mit der Taste **SELSATZ starten**