

**Stata-Übungsaufgabe zu Thomas K. Bauer, Stefan Bender, Jörg Heining
und Christoph M. Schmidt (2010):**

**“Brother Sun, Sister Moon: The Lunar Cycle, Sunspots and the
Frequency of Births”**

(Ruhr Economic Papers #225)

1. Öffnen Sie den Datensatz „aufgabe1.dta“.
2. Die Variable „n_bday“ gibt an, wieviele Personen an einem bestimmten Tag („gebdat“) Geburtstag haben. Wieviele Personen werden insgesamt erfasst?
3. Wieviele Leute haben durchschnittlich an einem Tag Geburtstag? Bestimmen Sie zusätzlich die minimale Anzahl, die maximale Anzahl sowie den Median. Speichern Sie den Mittelwert als lokales Makro (`local`).
4. Reproduzieren Sie das Streudiagramm (`scatter`) in Abbildung 2 auf Seite 5, welches die Anzahl der Geburten pro Tag darstellt.
5. Ergänzen Sie eine horizontale Linie, die die durchschnittliche Anzahl von Geburten pro Tag zeigt (siehe Aufgabe 3).
6. Speichern Sie dieses Diagramm als PNG- oder EPS-Datei.
(Diese Formate sind u.a. zur späteren Einbindung in Latex-Dokumente geeignet.)
7. Um den Einfluss der Mond- und Sonnenphasen auf die Häufigkeiten von Geburten an einem bestimmten Tag zu analysieren, schätzen die Autoren ein multivariates Regressionsmodell. Hierbei wird die logarithmierte Anzahl der Geburten pro Tag auf den Logarithmus der Anzahl der Sonnenflecken, Dummyvariablen für die verschiedenen Mondphasen (Neumond, zunehmender Mond, Vollmond) sowie eine Reihe von Kontrollvariablen regressiert. Zu diesen gehören u.a. Jahr, Monat und Wochentag.
Die Variable „gebdat“, die die entsprechenden Informationen enthält, ist im Datumsformat (`%td`) gespeichert. Dieses erlaubt es mit den Datumsfunktionen `year()`, `month()` und `dow()` die entsprechenden Informationen zu extrahieren. Generieren Sie die Variablen „year“, „month“ und „day“. Vergeben Sie Variablenlabels und erstellen Sie für die Variable „day“ zusätzlich Wertelabels (0 = „Sunday“, 1 = „Monday“, ...).
8. Erstellen Sie Jahres-, Monats- und Tagesdummies.
9. Generieren Sie eine Variable, die die Anzahl der Sonnenflecken durch 100 dividiert.
10. Versuchen Sie nun mit Hilfe der generierten Variablen die Schätzergebnisse aus Tabelle 1(1) zu reproduzieren.

11. Testen Sie mit Hilfe eines F-Tests die gemeinsame Signifikanz der Mondphasen-Variablen.
12. Betrachten Sie die Interpretation des Koeffizienten der „Number of Sunspots“ auf Seite 8. Wie berechnet sich die Anzahl von 5 zusätzlichen Geburten durch den Anstieg der Sonnenflecken um eine Standardabweichung?