

## ReadMe

The sample data include three data sets:

1. **HoustonWeaponUse.dbf** – a file of 3,709 robberies that took place in Houston, TX. The data set is for use in modeling weapon use during the robbery using the multinomial logit model described in Chapters 21 and 22. The coordinate system is projected with the distance units being feet. The data are simulated in that random temporal and spatial error has been added to the coordinates to protect confidentiality. The data set should be used only for tutorial purposes, not for research.

The variables are as follows:

### *ID and Coordinates*

- A. ID – the ID field for the robbery
- B. XCoord – the X coordinate
- C. YCoord – the Y coordinate

### *Dependent variable*

- D. WEAPON – type of weapon used. This is an alphanumeric variable and takes the values Firearm, Knife, Bodily Force, Threat, and Other Weapon.

### *Independent variables*

- E. AGE – age of offender
- F. FEMALE – gender (1 – female; 0 – male)
- G. BLACK – race (1 - African-American; 0 - non-African-American)
- H. HISPANIC – Hispanic ethnicity (1 – Hispanic; 0 – non-Hispanic)
- I. NUMSUSPCTS – number of suspects committing robbery
- J. COMMERCIAL – type of building where robbery committed (1 – commercial; 0 – non-commercial)
- K. NIGHT (1 – robbery committed between midnight and 6 am; 0 – robbery committed at other time)
- L. MORNING (1 – robbery committed between 6 am and noon; 0 – robbery committed at other time)
- M. AFTERNOON (1 – robbery committed between noon and 6 pm; 0 – robbery committed at other time)

- N. CL\_MEDINC – year 2000 median household income of traffic analysis zone where robbery was committed
- O. DISTANCE – distance of robbery location from downtown Houston (miles)

2. **TheHagueBurglars.dbf** – a file of 548 cleared burglaries that took place in The Hague, The Netherlands in the years 1996-2001. The data set is for use as a **case file** in creating a data set for the conditional logit model, as described in Chapters 21 and 22. The coordinate system is projected with the distance units being meters. The data are an exact copy of the data analyzed by Bernasco and Nieuwbeerta (2005).

The variables are as follows:

*ID and Coordinates*

- A. CASE – the ID field for the burglary (offence)
- B. PERSON\_ID – the ID field for the burglar (person)
- C. XRESID – the X coordinate of the centroid of the burglar’s neighborhood of residence
- D. YRESID – the Y coordinate of the centroid of the burglar’s neighborhood of residence

*Neighborhood identifiers*

- E. NHOODBUR – the ID field of the neighborhood where the burglary was committed
- F. NHOODRES – the ID field of the neighborhood where the burglar lived at the time the burglary was committed

*Individual characteristics*

- G. B\_MINOR –minor burglar (0 – burglar 18+ years; 1 – burglar 12-17 years)
- H. B\_ADULT – adult burglar (1 – burglar 18+ years; 0 – burglar 12-17 years)
- I. B\_NATIVE – native burglar (1 – native Dutch; 0 – foreign-born)
- J. B\_FOREIGN – nonnative burglar (0 – native Dutch; 1 – foreign-born)

3. **TheHagueNeighborhoods.dbf** – a file of 89 neighborhoods in The Hague for the years 1996-2001. The data set is for use as an **alternatives file** in creating a data set for the conditional logit model, as described in Chapters 21 and 22. The coordinate system is projected with the distance units being meters. The data are an almost exact copy of the data analyzed by Bernasco and Nieuwbeerta (2005), the difference being that in TheHagueNeighborhoods.dbf the neighborhood characteristics are averages over the period 1996-2001, while the original research used annually measured neighborhood characteristics.

The variables are as follows:

*ID and Coordinates*

- A. NHOODID – the ID field for the neighborhood (values correspond to NHOODBUR and NHOODRES in **TheHagueBurglars.dbf**)
- B. X – the X coordinate of the centroid of the neighborhood
- C. Y – the Y coordinate of the centroid of the neighborhood

*Neighborhood characteristics*

- D. RESUNITS – number of residential units (x 1000)
- E. PROPVAL – mean property value (x 100,000 euro)
- F. HOMEOWN –percentage home ownership (x 10%)
- G. ETNHETERO –Herfindahl index of ethnic heterogeneity (x 100)
- H. RESMOBIL – percentage residential mobility (x 10%)
- I. SINGFAM – percentage of properties that are single family dwellings (x 10%)
- J. PROXCITY – negative distance to The Hague city center (x 1 km)
- K. SURFACE – neighborhood surface in square kilometers

4. **TheHagueNeighborhoodsXBurglars.dbf** – this file is the result of matching the file *TheHagueBurglars.dbf* with *TheHagueNeighborhoods.dbf* using the Create dataset for conditional discrete choice model routine under Discrete Choice I (in the Spatial Modeling II section). The user should match the two files using the routine. Then, the matched data set can be compared with this data set to validate the results. The matched file is used in the conditional logit discrete choice model.