NAME: Benign Breast Disease 1-3 Matched Case-Control Study (BBDM13.DAT)

SIZE: 200 observations, 14 variables

KEYWORDS: Logistic Regression

TYPE: Matched Case-Control

SOURCE: These data come from Hosmer and Lemeshow (2000) Applied Logistic

Regression: Second Edition, page 245.

These data are copyrighted by John Wiley & Sons Inc. and must be

acknowledged and used accordingly.

DESCRIPTIVE ABSTRACT:

These data contain information from a 1 - 3 matched design studying the

risk factors associated with benign breast disease. The data are a subset

from a hospital based case-control study designed to examine the

epidemiology of fibrocystic breast disease. Data are provided on 50 women

who were diagnosed as having benign breast disease and 150 age matched

controls, with three controls per case. Matching was based on the age of

the subject at the time of interview.

Cases included women with a biopsy-confirmed diagnosis of fibrocystic

breast disease identified through two hospitals in New Haven, Connecticut.

Controls were selected from among patients admitted to the general surgery,

orthopedic, or otolaryngologic services at the same two hospitals. Trained

interviewers administered a standardized structured questionnaire to collect

information from each subject [see Pastides et al. (1983) and Pastides,

et al. (1985)].

LIST OF VARIABLES:

Variable Description Codes/Values Name

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1 Stratum 1 - 50 STR

2 Observation within Stratum 1 = Case, OBS

2 - 4 = Control

3 Age at Interview Years AGMT

4 Final Diagnosis 1 = Case, 0 = Control FNDX

5 Highest Grade in School 5 - 20 HIGD

6 Degree 0 = None DEG

1 = High School

2 = Jr. College

3 = College

4 = Masters

5 = Doctoral

7 Regular Medical Check-ups 1 = Yes, 2 = No CHK

8 Age at First Pregnancy Years AGP1

10 No. of Stillbirths, 0 - 7 NLV

Miscarriages etc.

11 Number of Live Births 0 - 11 LIV

12 Weight of the Subject Pounds WT

At Interview

14 Marital Status 1 = Married MST

2 = Divorced

3 = Separated

4 = Widowed

5 = Never Married

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Missing values are denoted by "." in the listing of the data.

ANSWER THE FOLLOWING:

1. Summarize the variable, DEG, NLV , MST and WT
2. Check if there is statistically significant relationship between Educational qualification(DEG) and Marital status (MST)