

**Sesame Street Study**

**Project 1**

**STAT 470W**

**Context**

Client: Dr. Education

Dr. Education just started working for production company Children’s Television Workshop.  Prior to her hiring, the company hired Educational Testing Services (ETS) to collect data on the effectiveness of their show *Sesame Street.* ETS was responsible for the design and collection of the data.  CTW also thought they would analyze the data, unfortunately they just handed over a spreadsheet.  Now, **Dr. Education has been asked to summarize the findings and present the results at a board meeting next month**. Can you help her?  There are many objectives that CTW has for its programming and thus to assess in this initial study, but the first board meeting will prioritize answering the following questions.

1.  Does our programming improve children’s knowledge of **letters**, **numbers**, and **body parts**?

2.  What, if any, area should we focus on for improvement?  E.g. are we better at teaching letters than we are at numbers?

Client background: She is good with spreadsheets but has never fit a linear model and does not know what a p-value is.

\* The Pretest and Posttest occurred 6 months apart.

\* We do not know much of anything about how the test was administered or how participants were recruited.

**Study Details**

The data are part of a larger data set that evaluated the impact of the first year of the Sesame Street television series. Sesame Street was concerned mainly with teaching preschool related skills to children in the 3-5-year age range, with special emphasis on reaching 4-year-old disadvantaged children.

The format of the show was designed to hold young children’s attention through action oriented, short duration presentations teaching specific preschool cognitive skills and some social skills. Each show was one hour long and involved much repetition of concepts within and across shows.

A main concern for the evaluation, which was carried out at the Educational Testing Service, was that it would permit generalization to the populations of children of most interest to the producers of the program (the Children’s Television Workshop). Five population were of interest:

1. Three to five-year-old disadvantaged children from inner city areas in various parts of the country.
2. Four-year-old advantaged suburban children.
3. Advantaged rural children.
4. Disadvantaged rural children.
5. Disadvantaged Spanish speaking children.

Children representative of these populations were sampled from five different sites in the United States.

Both before and after viewing the series, the children were tested on a variety of cognitive variables (variables 8 through 19 in the data set), including knowledge of body parts, knowledge about letters, knowledge about numbers, etc.

**Variables in Data Set**

Here are some notes, in addition to the table provided below.

The specific levels for VIEWCAT are:

1 = rarely watched Sesame Street

2 = watched Sesame Street 1-2 times per week

3 = watched Sesame Street 3-5 times per week

4 = watched Sesame Street >5 times per week

AGECAT is an indicator for AGE > 51.5, the mean age.

ENCOUR is a re-coding of VIEWENC from 1 & 2 to 1 & 0, respectively.

REGULAR is an indicator for VIEWCAT > 1.  We presume that REGULAR in this context means a child is a regular viewer of Sesame Street.

The “maximum score” values provided in the table below are misleading; they are not necessarily the maximum for the pre and post scores for a given knowledge area.

See the next page for details regarding the variables in the data set.

