# Professor Notes About the "Bivariate EDA - Cat" Homework

- You must provide labeled tables and figures to support your results and refer to these tables in your sentences.
- Sentences cannot begin with a number (e.g., you cannot say "22.2% of ...."). You must reorganize the sentence so that it does not begin with a number.
- Percentages are typically rounded to one decimal place.
- In the last question, you are trying to describe how the response variable (nearsightedness) depends on the explanatory variable (light used). Thus, you want to focus on the table that show the percentages of nearsightedness within (or by or for) each level of light used. In other words, your summary should come from comparing the rows in the row percentages table (assuming that you put the response variable as columns as I have suggested). You definitely do NOT want to compare rows from the total percentage table because that is affected by overall total number of individuals and is not "adjusted for" differences in numbers of individuals in the lights used categories.
- In the last question, do NOT say that light used "caused" the nearsightedness response. You can only attribute cause from a carefully constructed experiment or after many well crafted observational studies. This is only one observational study.

# Lights and Nearsightedness

1. A total of 17 children slept in "no light" conditions and developed nearsightedness (Table 1).

   Table 1. Frequency table for children by lighting conditions during sleep and whether they developed nearsightedness or not.

| | Nearsightedness | | |
|---|---|---|---|
| Light | No | Yes | Sum |
| lamp | 34 | 41 | 75 |
| night light | 153 | 79 | 232 |
| no light | 155 | 17 | 172 |
| Sum | 342 | 137 | 479 |

2. The percentage of children that slept in "no light" conditions that then developed nearsightedness is 9.9% (Table 2).

   Table 2. Percentages of chidren by nearsightedness type within each sleeping condition.

| | Nearsightedness | | |
|---|---|---|---|
| Light | No | Yes | Sum |
| lamp | 45.3 | 54.7 | 100.0 |
| night light | 65.9 | 34.1 | 100.0 |
| no light | 90.1 | 9.9 | 100.0 |

3. The percentage of all children that slept with a "lamp" and developed nearsightedness is 8.6% (Table 3).

   Table 3. Percentages of all students in each sleeping condition and nearsightedness type.

| | Nearsightedness | | |
|---|---|---|---|
| Light | No | Yes | Sum |
| lamp | 7.1 | 8.6 | 15.7 |
| night light | 31.9 | 16.5 | 48.4 |
| no light | 32.4 | 3.5 | 35.9 |
| Sum | 71.4 | 28.6 | 100.0 |

4. The percentage of children that slept with a "night light" that did not develop nearsightedness is 65.9% (Table 2).
5. The percentage of children that developed nearsightedness that slept with a "lamp" is 29.9% (Table

4).

Table 4. Percentage of children by sleeping conditions within each nearsightedness type.

| | Nearsightedness | |
|---|---|---|
| Light | No | Yes |
| lamp | 9.9 | 29.9 |
| night light | 44.7 | 57.7 |
| no light | 45.3 | 12.4 |
| Sum | 99.9 | 100.0 |

6. The percentage of children that developed nearsightedness is 28.6% (Table 3).
7. It appears that the percentage of children that developed nearsightedness is greater when the child slept with some sort of light (either a lamp or a night light), with a somewhat greater prevalence of nearsightedness with the lamp (Table 2).

# R Appendix

```
library(NCStats)
setwd('C:/aaaWork/Books/IntroStats/HW/')
d <- read.csv("nightlight.csv")
( tbl <- xtabs(~Light+Nearsightedness,data=d) )
addmargins(tbl)
( row.tbl <- percTable(tbl,margin=1) )
( col.tbl <- percTable(tbl,margin=2) )
( perc.tbl <- percTable(tbl) )
```