Exploration of County-level Variation in Sleep Deprivation

Michael Geden

Abstract

This document provides an introduction to R Markdown, argues for its...

Introduction

Approximately 1/3 of the population of the United States reports having less sleep on average than the recommended 7 hours a night (Disease Control & Prevention (2016); Liu (2016)]. Chronic sleep deprivation is considered a serious public health issue, as insufficient sleep is associated with increased workplace related accidents (Dinges, 1995; Rosekind et al., 2010), obesity (Gangwisch, Malaspina, Boden-Albala, & Heymsfield, 2004; Knutson, Spiegel, Penev, & Van Cauter, 2004), drowsy driving (Horne & Louise, 1995; Howard et al., 2004), cardiovascular disease (Ayas et al., 2003; Mullington, Hacch, Toth, Serrador, & Meier-Ewert, 2009), and a variety of other risks/conditions.

This distinction does not take into account whether the sleep deprived days were contingent (Gardener, 2015), an improtant factor in chronic partial sleep deprivation, adding additional uncertainty on the interpretation of moderate values in which either continuguous or non-continguous days would be possible. A concern with this method is that it is reliant on a relatively long-term retrospective self-report, in which participants must remember the number of days that they slept well within that duration. This is additionally confounded with the fact that the number of hours an individual needs to feel well-rested depents on a large number of factors, so that we are combining the accuracy of the self-report, the

Data

BRFSS 2009

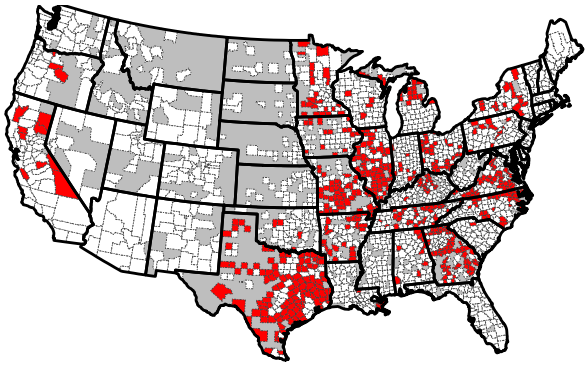
- "During the past 30 days, for about how many days have you felt you did not get enough rest or sleep? (number of days)"

BRFSS 2016

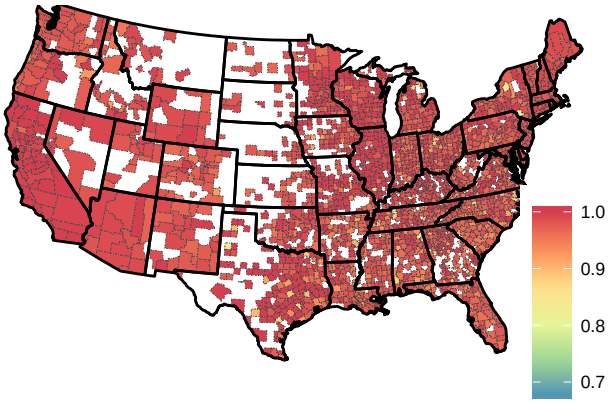
- "On average, how many hours of sleep do you get in a 24-hour period?"

Response Rates

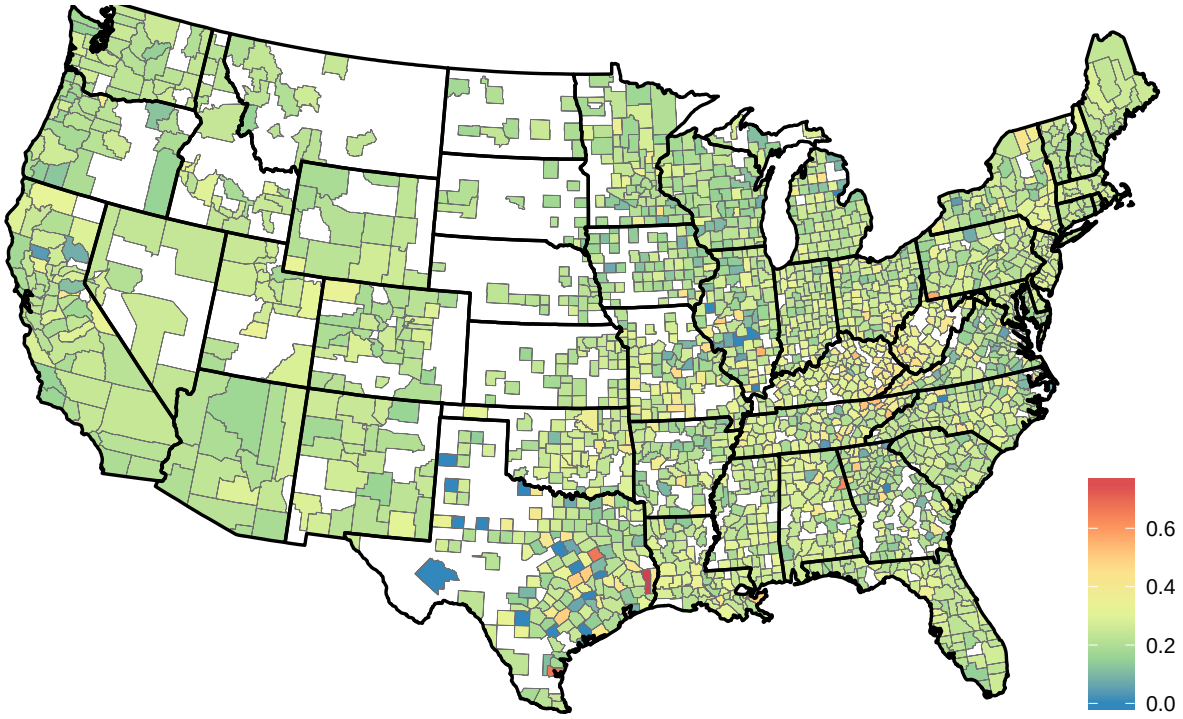
Frequency of Responses



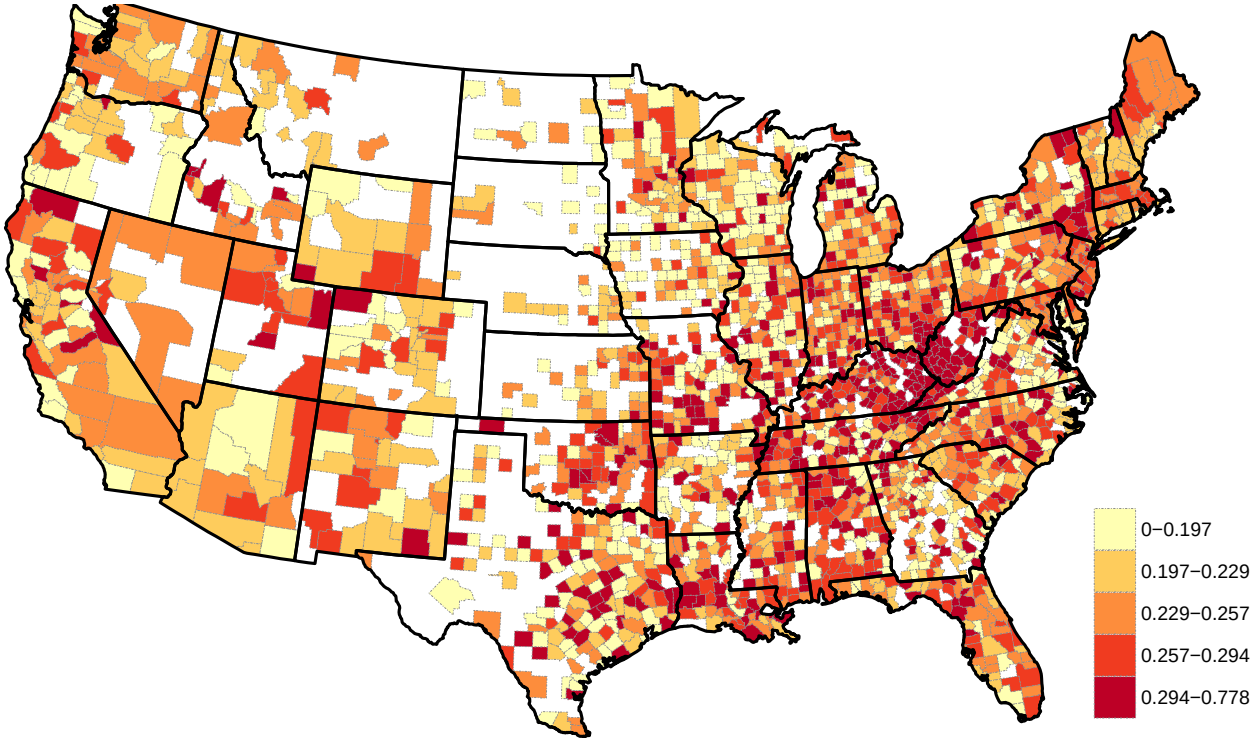
Proportion of Responses



County: Continuous Proportion of Sleep Deprivation



Quintiles of Proportion of Sleep Deprivation



Limitations

Age-Adjusted Proportions

State level could use asmyptotic normal confidence intervals, but the smaller sample sizes in the county level analysis would need to use an exact test or bootstrap estimate (provided moderate n, ~15-30). Given the presence of numerous other variables not controlled for, and the differences in sampling error for older populations vs. younger populations, simply using age-adjusted estimates would likely not be sufficients. Instead using a multilevel approach such as the one outlined in Zhang et al. (2014)

Missing Data

For the purposes of plotting here, very little data has been to account for the meaning behind the different kinds of missing data (lack of responses, didn't know the answer, refused to respond). These are important factors, and should definitely be investigated, but with the simple intent here of visualizing general trends for exploratory analyses this was considered sufficient.

Survey Stratification

One important elemenet of these large scale surveyts is that the survey appropriately targets the samples of interest, whether that is the a specific subgroup or all relevant subgroups. Consideriing the proportion of missing data and relatively limited data, jumping into subgroup differences within the county level does not seem prudent. Instead, an analysis controling for differences between states/counties while looking at the reslationship of interest would be more appropriate. This is not covered here for brevity.

Code

```
BRFSS_geoextraction <- function(filepath, id = "county", position,
                                extraction, contiguous=TRUE,latlong=TRUE,
                                func = function(x) {mean(x, na.rm=T)}) {
    # Error Catching
    if(!grepl("state", tolower(position[1]))){stop("State must be first position.")}
    if(!is.character(filepath)) {stop("Filepath must be character")}
    if(!is.character(position)) {stop("Position must be character")}
    if(!is.character(extraction)) {stop("Extraction variable must be character name.")}
    if(length(extraction)!=1) {stop("Function currently only accepts one extracted var.")}
    if(!id %in% c("county", "state")) {stop("ID must be county or state.")}
    if(!is.function(func)){stop("Func must be a function.")}
    if(id == "county" & length(position) != 2) {
        stop("County level requires position to have State and County varialbe names. ")
    }
    # Setup
    require(SASxport)
    require(dplyr)
    require(maps)
    require(stringr)
    require(ggplot2)
    strsplit2 <- function(x, char, index){
        unlist(strsplit(x, char))[index]
    }
    clear_labels <- function(x) {
        if(is.list(x)) {
            for(i in 1 : length(x)) class(x[[i]]) <- setdiff(class(x[[i]]), 'labelled')
            for(i in 1 : length(x)) attr(x[[i]],"label") <- NULL
        }
        else {
            class(x) <- setdiff(class(x), "labelled")
            attr(x, "label") <- NULL
        }
        return(x)
    }
    # Read Data
    data <- read.xport(filepath) %>%
        clear_labels(.)
    # Subset
    keep <- c(position, extraction)
    keep.states <- if(contiguous) {
        c(15, 2, 66, 72, 78)
    } else {c()}
    # Remove Excess
    data2 <- data %>%
        select(one_of(keep)) %>% # Select columns of interest
```

```
        filter(!(!!rlang::sym(position[1])) %in% keep.states) %>% # Select States
        filter(!(is.na(!!rlang::sym(position[1])))) # remove missing states
    if(id == "county"){
        data2 <- data2 %>%
            filter(!(!!rlang::sym(position[2])) %in% c(777,999)) %>%
            filter(!(is.na(!!rlang::sym(position[2])))) %>%
            mutate(FIPS = as.numeric(paste0(
                !!rlang::sym(position[1]),
                str_pad(!!rlang::sym(position[2]),3,pad = "0"),
                sep = "")
            ))
    } else {
        data2 <- data2 %>%
            mutate(FIPS = !!rlang::sym(position[1]))
    }
    # Get FIPS Codes
    if(id == "county") {
        data(county.fips)
        county.fips$polyname <- sapply(county.fips$polyname, function(x) {
            unlist(strsplit(x, ":"))[1]
        })
        county.fips <- unique(county.fips)
        fips.codes <- county.fips
    } else {
        data(state.fips)
        state.fips$polyname <- sapply(state.fips$polyname, function(x) {
            unlist(strsplit(x, ":"))[1]
        })
        state.fips <- unique(state.fips)
        fips.codes <- as.data.frame(state.fips)
    }
    varname <- paste0(extraction, ".f", sep = "")
    # Merge By FIPS
    data3 <- data2 %>%
        left_join(fips.codes, by = c("FIPS" = "fips")) %>% # Get location string
        group_by(polyname,FIPS) %>% # State/County-wise operations
        summarize(!!varname := func(!!rlang::sym(extraction)), # Function
                  n = sum(!is.na(!!rlang::sym(extraction))), # Frequency of response
                  prop.responded = sum(!is.na(!!rlang::sym(extraction)))/n()) %>%
        filter(!is.na(!!rlang::sym(varname))) %>% # Remove Missing
        mutate(state = strsplit2(polyname, ",", 1)) # Extract state name
    if(id == "county") {
        data3$county <- strsplit2(data3$polyname, ",",2)
    }
    # Merge for coordinates
    if(id == "county"){
        map <- map_data('county') %>%
```

```
            mutate(polyname = paste0(region, ",", subregion, sep = "")) %>%
            select(-c(group, order, region, subregion))
    } else {
        map <- map_data('state') %>%
            mutate(polyname = paste0(region, ",", subregion, sep = "")) %>%
            select(-c(group, order, region, subregion))
    }
    if(latlong) {data3 <- left_join(data3, map, by = "polyname")}
    # Output
    data3
}
```

References

Ayas, N., White, D., Manson, J., Stampfer, M., Speizer, F., Malhotra, A., & Hu, F. (2003). Archives of internal medicine. *Progress in Cardiovascular Diseases*, *2*(163), 205–2009.

Dinges, D. (1995). An overview of sleepiness and accidents. *Journal of Sleep Research*, *14*(4), 4–14.

Disease Control, C. for, & Prevention. (2016). Effect of short sleep duration on daily activities - united states, 2005-2008. *Morbidity and Mortality Weekly Report*, (60), 239.

Gangwisch, J., Malaspina, D., Boden-Albala, B., & Heymsfield, S. (2004). Inadequate sleep as a risk factor for obesity; analysis of the nhanes 1. *Sleep*, *10*(28), 1289–1296.

Horne, J., & Louise, R. (1995). Sleep related vehicle accidents. *British Medical Journal*, *6979*(310), 565–567.

Howard, M., Desai, A., Grunstein, R., Hukins, C., Armstrong, J., Joffe, D., ... Pierce, R. (2004). Sleepiness, sleep-disordered breathing, and accident risk factors in commercial vehicle drivers. *American Journal of Respiratory and Critical Care Medicine*, *9*(170), 1014–1021.

Knutson, K., Spiegel, K., Penev, P., & Van Cauter, E. (2004). The metabolic consequences of sleep deprivation. *Sleep Medicine Reviews*, *3*(11), 163–178.

Liu, Y. (2016). Prevalence of healthy sleep duration among adults- united states, 2014. *Morbidity and Mortality Weekly Report*, (65), 1025–1033.

Mullington, J., Hacch, M., Toth, M., Serrador, J., & Meier-Ewert, H. (2009). Progress in cardiovascular diseases. *Progress in Cardiovascular Diseases*, *4*(51), 294–302.

Rosekind, M., Gregoy, K., Mallis, Melissa, Brandt, S., Seal, B., & Lerner, D. (2010). An overview of sleepiness and accidents. *American College of Occupational and Environmental Medicine*, *52*(1), 91–98.

Zhang, X., Holt, J., Lu, H., Wheaton, A., Ford, Earl, Greenlund, K., & Croft, J. (2014). Mulilevel regression and poststratification for small-area estimation of population health outcomes: A case study of crhonic obstructive pulmonary disease prevelance using the behavioral risk factor surveillance system. *American Journal of Epidemiology*, *179*(8), 1025–1033.