After scraping everyone’s Stormfront profiles, with the location field called Location, here is the R code to get states.

NOTE: For the final data, I just used people who named the state. In other words, I did not include people who named a city. I attempted to do this for a large set of cities (change 134 to 104 in the code below) and the results were almost identical.

data <- data[which(data$Location!=""),]

data$Location <- tolower(data$Location)

data$Location <- paste("",data$Location,";;",sep="")

data$Location <- gsub("\\,","\\ ",data$Location)

data$Location <- gsub("\\.","\\ ",data$Location)

data$State <- NA

v1 <- c("california","california")

v2 <- c("kjljfdfdaf","california")

v3 <- c("mississippi","mississippi")

v4 <- c("alabama","alabama")

v5 <- c("arizona","arizona")

v6 <- c("arkansas","arkansas")

v7 <- c("colorado","colorado")

v8 <- c("connecticut","connecticut")

v9 <- c("delaware","delaware")

v10 <- c("florida","florida")

v11 <- c("georgia","georgia")

v12 <- c("hawaii","hawaii")

v13 <- c("idaho","idaho")

v14 <- c("illinois","illinois")

v15 <- c("indiana","indiana")

v16 <- c("iowa","iowa")

v17 <- c("kansas","kansas")

v18 <- c("kentucky","kentucky")

v19 <- c("louisiana","louisiana")

v20 <- c("maine","maine")

v21 <- c("maryland","maryland")

v22 <- c("massachusetts","massachusetts")

v23 <- c("michigan","michigan")

v24 <- c("minnesota","minnesota")

v25 <- c("montana","montana")

v26 <- c("nebraska","nebraska")

v27 <- c("nevada","nevada")

v28 <- c("new hampshire","new hampshire")

v29 <- c("new jersey","new jersey")

v30 <- c("new mexico","new mexico")

v31 <- c("new york","new york")

v32 <- c("north carolina","north carolina")

v33 <- c("north dakota","north dakota")

v34 <- c("ohio","ohio")

v35 <- c("oklahoma","oklahoma")

v36 <- c("oregon","oregon")

v37 <- c("pennsylvania","pennsylvania")

v38 <- c("rhode island","rhode island")

v39 <- c("south carolina","south carolina")

v40 <- c("tennessee","tennessee")

v41 <- c("texas","texas")

v42 <- c("utah","utah")

v43 <- c("vermont","vermont")

v44 <- c("west virginia","west virginia")

v45 <- c("w. virginia","west virginia")

v46 <- c("virginia","virginia")

v47 <- c("washington dc","district of columbia")

v48 <- c("washington","washington")

v49 <- c("wisconsin","wisconsin")

v50 <- c("wyoming","wyoming")

v51 <- c("alaska","alaska")

v52 <- c("missouri","missouri")

v53 <- c("south dakota","south dakota")

v54 <- c(" al;;","alabama")

v55 <- c(" ak;;","alaska")

v56 <- c(" az;;","arizona")

v57 <- c(" ar;;","arkansas")

v58 <- c(" ca;;","california")

v59 <- c(" co;;","colorado")

v60 <- c(" ct;;","connecticut")

v61 <- c(" de;;","delaware")

v62 <- c(" fl;;","florida")

v63 <- c(" ga;;","georgia")

v64 <- c(" hi;;","hawaii")

v65 <- c(" id;;","idaho")

v66 <- c(" il;;","illinois")

v67 <- c(" in;;","indiana")

v68 <- c(" dc;;","district of columbia")

v69 <- c(" ia;;","iowa")

v70 <- c(" ks;;","kansas")

v71 <- c(" ky;;","kentucky")

v72 <- c(" la;;","louisiana")

v73 <- c(" me;;","maine")

v74 <- c(" md;;","maryland")

v75 <- c(" ma;;","massachusetts")

v76 <- c(" mi;;","michigan")

v77 <- c(" mn;;","minnesota")

v78 <- c(" ms;;","mississippi")

v79 <- c(" mo;;","missouri")

v80 <- c(" mt;;","montana")

v81 <- c(" ne;;","nebraska")

v82 <- c(" nv;;","nevada")

v83 <- c(" nh;;","new hampshire")

v84 <- c(" nj;;","new jersey")

v85 <- c(" nm;;","new mexico")

v86 <- c(" ny;;","new york")

v87 <- c(" nc;;","north carolina")

v88 <- c(" nd;;","north dakota")

v89 <- c(" oh;;","ohio")

v90 <- c(" ok;;","oklahoma")

v91 <- c(" or;;","oregon")

v92 <- c(" pa;;","pennsylvania")

v93 <- c(" ri;;","rhode island")

v94 <- c(" sc;;","south carolina")

v95 <- c(" sd;;","south dakota")

v96 <- c(" tn;;","tennessee")

v97 <- c(" tx;;","texas")

v98 <- c(" ut;;","utah")

v99 <- c(" vt;;","vermont")

v100 <- c(" va;;","virginia")

v101 <- c(" wa;;","washington")

v102 <- c(" wv;;","west virginia")

v103 <- c(" wi;;","wisconsin")

v104 <- c(" wy;;","wyoming")

v105 <- c("atlanta","georgia")

v106 <- c("anchorage","alaska")

v107 <- c("philadelphia","pennsylvania")

v108 <- c("pittsburg","pennsylvania")

v109 <- c("brooklyn","new york")

v110 <- c("san francisco","california")

v111 <- c("los angeles","california")

v112 <- c("boston","massachusetts")

v113 <- c("baltimore","maryland")

v114 <- c("chicago","illinois")

v115 <- c("new orleans","louisiana")

v116 <- c("vegas","nevada")

v117 <- c("dallas","texas")

v118 <- c("houston","texas")

v119 <- c("miami","florida")

v120 <- c("orlando","florida")

v121 <- c("tampa","florida")

v122 <- c("jacksonville","florida")

v123 <- c("denver","colorado")

v124 <- c("tulsa","oklahoma")

v125 <- c("louisville","kentucky")

v126 <- c("memphis","tennessee")

v127 <- c("nashville","tennessee")

v128 <- c("phoenix","arizona")

v129 <- c("seattle","washington")

v130 <- c("albuquerque","new mexico")

v131 <- c("omaha","nebraska")

v132 <- c("honolulu","hawaii")

v133 <- c("portland","oregon")

for (i in 1:104) {

gg <- eval(parse(text=paste("v",i,sep="")))

data$State[which(grepl(gg[1],data$Location)==TRUE & is.na(data$State))] <- gg[2]

}