Distance from home

Dominik Glandorf

2022-12-06

```
loc_data = bg[,c('distance_from_home','city_residence_app','state_residence_app','country_residence_app
```

```
head(loc_data[!is.na(loc_data$distance_from_home)&loc_data$country_residence_app=="USA",], 100)
```

```
## # A tibble: 100 x 5
##    distance_from_home city_residence_app state_residence_app country_r~1 zip_c~2
##                 <dbl> <chr>              <chr>                <chr>       <chr>
##  1                 27 LA PUENTE, CA      CA                   USA         91744-~
##  2                360 FREMONT, CA        CA                   USA         94555
##  3                391 RICHMOND, CA       CA                   USA         94806
##  4                 36 LOS ANGELES, CA    CA                   USA         90032
##  5                 50 HIGHLAND, CA       CA                   USA         92346
##  6                 12 WESTMINSTER, CA    CA                   USA         92683
##  7                 12 MISSION VIEJO, CA  CA                   USA         92692
##  8                 52 TUJUNGA, CA        CA                   USA         91042
##  9                381 SAN FRANCISCO, CA  CA                   USA         94134
## 10                  5 COSTA MESA, CA     CA                   USA         92626
## # ... with 90 more rows, and abbreviated variable names
## #   1: country_residence_app, 2: zip_code_app
```

It is most probably the straight line distance in miles.

```
aggregate(is.na(distance_from_home) ~ country_residence_app, loc_data, FUN=mean)
```

```
##   country_residence_app is.na(distance_from_home)
## 1               Austria               0.000000000
## 2                Canada               0.000000000
## 3                 China               0.004494382
## 4                 India               0.000000000
## 5             SINGAPORE               0.000000000
## 6           South Korea               0.055555556
## 7                Taiwan               0.000000000
## 8                   USA               0.230720098
```

Overall distance from home is missing in 28% of the cases.