Supplemental information

Diet during the COVID-19 pandemic: An analysis of Twitter data

Mark A. Hernandez, Shagun Modi, Kanisha Mittal, Pallavi Dwivedi, Quynh C. Nguyen, Nina L. Cesare, and Elaine O. Nsoesie
### Table S1: List of 52 Terms Indicating References to Alcohol

| Alcohol Terms | beer | rum | pinot grigio | petite sirah |
|---------------|------|-----|--------------|-------------|
| wine          | mimosa | sauvignon blanc | gewurztraminer |
| alcohol       | martini | cabernet sauvignon | lychee martini |
| liquor        | brews | zinfandel | cocktail mix |
| tequila       | blood mary | syrah | rice wine |
| ipa           | daiquiri | tequila sunrise | pinot blanc |
| beers         | chardonnay | sangiovese | malt beverage |
| whiskey       | eggnog | muscat | malt beverages |
| vodka         | merlot | claret | maraschino |
| champagne     | pinot | cabernet franc | fume blanc |
| mimosas       | pinot noir | gamay | rice sake |
| gin           | pina colada | pinot gris | tavern |
| booze         | riesling | chenin blanc | taverns |

Three annotators identified 52 alcohol terms from the full list of unique food terms in the initial dataset of 11,445,868 food tweets collected in the United States from May 15, 2019 to January 31, 2020 and from May 15, 2020 to January 31, 2021.
Geotagged food tweets were extracted from 1% of random tweets collected during the pre-pandemic period (May 15, 2019, to January 31, 2020) and the pandemic period (May 15, 2020, and January 31, 2021) using the Twitter streaming application programming interface. The analytic sample consisted of 458,419 geotagged food tweets collected from 1,258 counties in the United States.
### Table S2: Odds of Geotagged Food Tweet Referencing Healthy Food, Fast Food, or Alcohol by County-level Factors

| County-level Factor                  | Healthy Food |                                                                 | Fast Food |                                                                 | Alcohol |
|-------------------------------------|--------------|----------------------------------------------------------------|-----------|----------------------------------------------------------------|---------|
|                                     | Odds Ratio (95% CI) | P-value              | Odds Ratio (95% CI) | P-value            | Odds Ratio (95% CI) | P-value |
| **Population Mobility**             |              |                      |                      |                    |                     |         |
| % Change in Time Spent in Places of Residence | 1.019 (1.011, 1.027) | < 0.001              | 0.967 (0.956, 0.977) | < 0.001            | 0.999 (0.993, 1.006) | 0.849   |
| **Built Environment**               |              |                      |                      |                    |                     |         |
| Grocery Stores per 10,000           | 1.019 (1.007, 1.032) | 0.002                | 0.967 (0.949, 0.986) | 0.001              | -                    | -       |
| Restaurants per 10,000              | 0.999 (0.997, 1.000) | 0.055                | 0.993 (0.991, 0.995) | < 0.001            | -                    | -       |
| Bars per 10,000                     | -            | -                    | -                    | -                  | 0.989 (0.978, 1.000) | 0.047   |
| Liquor Stores per 10,000            | -            | -                    | -                    | -                  | 1.066 (1.046, 1.087) | < 0.001 |
| **Population Characteristics**      |              |                      |                      |                    |                     |         |
| % Age 10-24 years                   | 1.000 (0.995, 1.004) | 0.906                | 0.980 (0.974, 0.986) | < 0.001            | 1.004 (1.001, 1.008) | 0.020   |
| % Age 65+ years                     | 1.006 (1.001, 1.010) | 0.012                | 0.979 (0.973, 0.985) | < 0.001            | 0.986 (0.982, 0.989) | < 0.001 |
| % Low Income                        | 1.005 (1.001, 1.010) | 0.022                | 0.984 (0.978, 0.990) | < 0.001            | 0.991 (0.987, 0.994) | < 0.001 |
| % High Income                       | 0.998 (0.995, 1.001) | 0.210                | 1.002 (0.998, 1.006) | 0.254              | 0.996 (0.993, 0.998) | < 0.001 |
| % Non-Hispanic Black                | 1.003 (1.001, 1.006) | 0.015                | 0.991 (0.988, 0.995) | < 0.001            | 0.998 (0.996, 1.001) | 0.152   |
| % Hispanic                          | 1.002 (0.999, 1.004) | 0.125                | 0.990 (0.987, 0.993) | < 0.001            | 0.998 (0.996, 1.000) | 0.096   |
| % Non-Hispanic White                | 1.002 (0.999, 1.005) | 0.137                | 0.986 (0.982, 0.989) | < 0.001            | 1.004 (1.002, 1.007) | < 0.001 |
| **Baseline Food Category References** |              |                      |                      |                    |                     |         |
| % Food Tweets Referencing Healthy Food Pre-Pandemic | 1.041 (1.038, 1.043) | < 0.001              | -                    | -                  | -                    | -       |
| % Food Tweets Referencing Fast Food Pre-Pandemic | -            | -                    | 1.077 (1.075, 1.080) | < 0.001            | -                    | -       |
| % Food Tweets Referencing Alcohol Pre-Pandemic | -            | -                    | -                    | -                  | 1.039 (1.038, 1.040) | < 0.001 |

Geotagged food tweets were extracted from 1% of random tweets collected during the pre-pandemic period (May 15, 2019, to January 31, 2020) and the pandemic period (May 15, 2020, and January 31, 2021) using the Twitter streaming application programming interface. The sample used for the regression model consisted of 458,419 geotagged food tweets collected from 1,258 counties during the pandemic period in the US. Data on population mobility were imported from the Google Covid-19 Community Mobility Reports. Built environment variables were derived from two data sources: the 2017-2018 Community Business Patterns database and the 2015-2019 American Community Survey. Population characteristics data were imported from 2015-2019 American Community Survey. Covariates for baseline food category references were derived from the pre-pandemic sample of tweets.