We masked and adjusted our data for the purposes of this exercise. As you work through the questions below, please make any assumptions you need to. Please plan to share your assumptions as part of the session.

#### Links:

- [Kinsa Overview and Data Description](#_knzyz1uznam8)

- [Data](https://drive.google.com/drive/folders/1dWQ77DAh3WUJiOTWy6yygsFEK9enzFOm?usp=sharing)

#### Part 1: Exploration:

Summarize changing demographic trends (age, region, gender, etc) in the Kinsa data set over time.

1. Identify and **show** the significant inflection points and other significant trends in the readings.

#### Part 2: Create illness signals that estimate the % of the broader population that have a fever. Adjustments can be made for any anomalies, as you see fit.

* National signal across all states and age groups
* Adult (13 and up) and Pediatric (12 and less) national signal across all states
* Create County Level signals for California and Texas across all age groups (adult, pediatric, and overall)

1. How did you arrive at the denominator?
2. Create a year over year chart with the X axis starting in July and ending in June. Plot the Signals you created so the illness seasons can be visually compared
3. Did you see any anomalies? Did you make any adjustments?

#### Part 3: Forecasting Discussion:

1. Based on the signals created, describe how you would start developing a forecast for this signal (No need to actually build a forecast)

# Kinsa overview

**Who is Kinsa?**

Kinsa's solutions predict, and help prevent, the impact of infectious illness on healthcare organizations and families. We use advanced technology and epidemiological techniques to analyze unique illness signals, gathered from millions of households and other sources, to track and forecast the spread of illness earlier than other systems and get ahead of the next surge.

**How do we collect this data?**

Kinsa sells and freely distributes FDA-cleared, Bluetooth-enabled smart thermometers. Kinsa thermometers pair with an associated app and record temperatures and symptoms reported by users. When recording symptoms, users can assign symptoms to multiple family members within a single household. Users can opt in to share location of their readings. Kinsa anonymizes and aggregates these readings to produce illness signals for different regions across the US. Kinsa households represent a diverse sample of the US population with 2.5 million households across all 50 states.

**How to interpret the data?**

The data provided in this repository are estimates of the percent of a region’s population experiencing illness.

Since people reach for their thermometer at the first signs of a fever, Kinsa captures data at symptom onset. Other surveillance systems will not start to collect data until a patient has become sick enough to interact with the health system. Thus, Kinsa’s data can capture mild respiratory illness earlier and more robustly than traditional public health surveillance.

**Limitations**

As with all disease surveillance, no single source can capture all illness trends across the country. Kinsa users must have access to a smartphone to use the app. Additionally, some rural states may have a low number of active users and therefore have noisier data for less common symptoms. In rural areas, we’ll pull from surrounding areas to get sufficient signal strength