The data set

* 2 cities
* Originally 17 weeks, 1 data point for each week
* 1998 bacteria signatures 🡪 1998 time series
* Interpolated to make the data set daily 🡪 113 data points

Answer 2 questions:

* Is there currently a significant trend in the bacterial signatures over time?
* Considering the trends, are the most recent data points outliers?

Split into 2 groups to work on these 2 questions:

1. Long, John, Sylvia: 1st question
2. Andy, Dylan, Ojas: 2nd question
3. Special Group: Rays, Andrea, Gloria: prepare for presentation

Steps for each group:

1. Extract the interested measurement from each time series. For 1: change in the time series mean in the last 7/14 days. For 2: error between the data point and the prediction model.
2. Apply step 1 to all bacteria signatures
3. Rank the bacteria signature by interested measurement. Point out the highest/lowest signatures
4. (optional) parallel programming on step 2.

Follow these guides:

Group 1: <https://rpubs.com/richkt/269908> (R)

Group 2: <https://www.statsmodels.org/devel/examples/notebooks/generated/stl_decomposition.html> (python)