*Thank you for applying to the Analytics team at TuneIn! Please take up to **4 hours** to complete this exercise. It doesn't have to be perfect, or even complete—if you can't get to it all, tell us what you would do if you had the time. The purpose of this exercise is to get a sense of how you think about analysis using imperfect data, and to evaluate your technical skills.*

Data

In this drive folder you'll find two additional items:

- A **dataset** with the locations of businesses that pay taxes to the City and County of San Francisco. Each registered business may have multiple locations and each location is a single row. The Treasurer & Tax Collector's Office collects this data through business registration applications, account update/closure forms, and taxpayer filings. [source: *DataSF*]

- A **data dictionary** to help define fields and codes.

We've cleaned the business category data for columns <NAICS Code> and <NAICS Code Description> by reclassifying businesses with multiple <NAICS Code> entries. We took the first code and added the corresponding NAICS Code Description.

Assignment

We ask that you use python or R to complete this assignment, providing your script so we may run it as well. Please be thorough in your explanations and explicit as you make analysis decisions—we want to know why as well as how you reach your conclusions.

1. Identify pockets in San Francisco with a high concentration of active businesses. Please provide visualization(s) to support your analysis.

2. To promote growth in less popular industries in San Francisco, the county is offering a tax break for new businesses in these categories to encourage them set up shop in San Francisco. Identify the NAICS code and description of businesses that would benefit from this scheme.

3. What are the different industry types that have emerged in San Francisco over the years. Are there any trends you can identify?

4. If you had more time, what other conclusions might you be able to draw from this data? What interesting stories might you be able to tell?