High Level Elements

1. What is the problem you want to solve?
   1. What are top health trends in Northern California (vs. USA)?
   2. What are factors of health trends in Northern California (vs. USA)?
   3. What can we expect in the future (i.e., predictive model)?
2. Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn’t have otherwise?
   1. Client ~ Typical healthcare organization in Northern California
   2. Decision(s) ~ Should we/organization staff up, down or keep same levels (ie, MD, PA, MA) for a particular type of care (ie, OBGYN) and/or for a particular location (ie, Berkeley)?
3. What data are you going to use for this? How will you acquire this data?
   1. <https://www.census.gov/popclock/>
   2. <https://www.healthdata.gov/search/type/dataset>
   3. <https://www.cdc.gov/nchs/>
   4. <https://www.medicaid.gov/medicaid/index.html>
4. In brief, outline your approach to solving this problem (knowing that this might change later).
   1. Explore trends and secondary sources that may be unstructured or corrupted
   2. Explore structured datasets in population health/demographics from open sources
   3. Import, manipulate, merge and clean datasets
   4. Perform statistics, derive insights and visualize results
5. What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.
   1. Code ~
   2. Slides ~
   3. Report ~