



Thank you for taking your time to do the Fifth Eye data analytics assessment. This project should take between 4 and 8 hours depending upon your skills and the depth of your analysis.

You will work with two data files of a single patient from [MIMIC](#), a large, freely-available database comprising deidentified health-related data from the critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. The database includes information such as demographics, vital signs, laboratory test results, procedures, medications, caregiver notes, imaging reports, and mortality (both in and out of hospital).

One of the two data files (Numerics) is a set of vital signs for the patient over time. The other data file (InputsMV) is a set of medical interventions including fluids and medications over time. We are asking you to explore the data to look for relationships between vital signs and interventions. As part of working at Fifth Eye, you would need to align data from multiple sources prior to analysis. Our tasks are written to reflect this workflow.

Some of the tasks are intentionally open-ended to allow you to highlight your interpretation skills as well as your analytical skills. Keep the time window of 4-8 hours in mind as you consider how far to take the open-ended tasks.

Using your favorite language perform the following tasks. As you execute the tasks, please ensure that the code is written in such a way that the tasks 1-5 (and possibly 6) can be executed on new patient data without requiring much change.

- 1) In the Numerics table, create a timeline by converting epoch time to date-time. Note that as part of deidentifying the data it has been time-shifted to the year 2184.
- 2) Using the converted timeline from Task 1, plot the vital signs: HR (heart rate), NBPMean (blood pressure mean), and RESP (respiration rate) along time.
- 3) Using the 'starttime' column, identify all rows in the InputsMV table that fall within the range of the Numerics timeline. A description of the column heading meanings can be found [here](https://mimic.physionet.org/mimictables/inputevents_mv/) (https://mimic.physionet.org/mimictables/inputevents_mv/)
- 4) From the subset of the InputsMV table generated in task 3, group (by 'label') and plot time versus each of the grouped entries that have 'ordercategorydescription' = 'Continuous Med'.
- 5) Healthcare data, like many real-world sources, is messy. Using your time-aligned plots, summarize your observations and/or questions. Provide an explanation if you can infer any relationship(s) between continuous meds administered and vital signs.
- 6) Finally, as a stretch goal if you have spare time (not mandatory), use the entire InputsMV table to create a summary table of the different data elements (grouped by 'label'). This step is to understand your approach to the statistical summarization of a complex table.

Please feel free to ask clarifying questions. After you are done, email me your outputs and code. (Jeff Basch: jbasch@fiftheyeye.com)