Case study for AstraZeneca Oncology Machine Learning & AI interview

The U.S. Food and Drug Administration (FDA) regulates over-the-counter and prescription drugs in the United States, including biological therapeutics and generic drugs. This work covers more than just medicines. For example, fluoride toothpaste, antiperspirants, dandruff shampoos and sunscreens are all considered drugs.

An adverse event is submitted to the FDA to report any undesirable experience associated with the use of a medical product in a patient. For drugs, this includes serious drug side effects, product use errors, product quality problems, and therapeutic failures for prescription or over-the-counter medicines and medicines administered to hospital patients or at outpatient infusion centers.

The FDA's database of adverse event reports is made available through a web API at https://open.fda.gov/apis/drug/event/ . Each report contains general information about the report, patient information, a list of the drugs that the patient is taking, and a list of the patient reactions. It is possible to use these data in many ways: your brief is to explore these data and to see what might be learned from them. As a guide, you might consider a practical solution to one of the following questions:

- Are different adverse events reported in different countries?
- What are the different adverse events associated with different disease areas?
- What drugs tend to be taken together?

You should publish any code you write to your personal github repository, and send a link two days before interview. At interview you should expect to discuss your code, any statistics or visualizations you may have used or think should be used, limitations of the underlying data, and how your solution or approach could be generalized and used in practice.