SHIELD ILLINOIS

SHIELD Illinois, (also referred to as “SHIELD”), is the University of Illinois System’s initiative to make the innovative saliva-based covid SHIELD COVID-19 test available to K-12 schools, colleges, universities, companies, and the public across the state of Illinois. SHIELD began testing in the Fall of 2020. SHIELD moved away from the cotton swab test process and adapted a saliva-based covid test.

**The Testing Process**

The saliva was collected in test tubes at collection/testing centers. According to our data, these are called specimen. The specimen were then transported to the labs. At the labs, a process of sample preparation occurred. Since PCR test could not be conducted directly on saliva, the specimen was treated. Sample preparation involves heating saliva then mixing it with a combination of buffer and detergent​. This is sample. Samples are then transferred to plates (Plate 96). Each plate has 96 wells, and each well would contain one sample. The plates were then inserted into testing equipment (PCR machines) that would conduct the test. The results of the test would then be delivered back to the individual.

**Transport and Tracking**

Manifest: Manifests were used to track the specimen from each testing center. A manifest was created when testing began at a specific location, and it was closed when testing ended for that day. Manifest is not something physical, rather it was just used to track the specimen. The size of any manifest would depend on the testing volume at that location. There are manifests in the data with only one specimen as well. Multiple manifests were also sent out from one location in a single day. Each test center had different criteria for transporting the manifests. Some would do it at a specific time of the day, some would do it at the end of the day.

From the test centers, the manifests (specimen from different locations) would be transported to labs. However, not all specimens went directly to the labs. Specimens were transported to the Depot where they would get sorted and then sent to specific labs. The sorting was a manual process and was done to achieve quicker results and to manage load balance at the labs as well. The Depot would decide which specimen would be transported to which lab.

**Labs and Testing Centers**

SHIELD had 12 diagnostic laboratories. We have the location and last operating date for each lab. Not all labs were in Illinois, one of the labs was in Kentucky and one in Wisconsin as well. When testing volume increased, SHIELD T3 labs were used including the two in Kentucky and Wisconsin. SHIELD would use existing labs at different locations as building a new lab would take an ample amount of time, for e.g. the lab at Loyola University.

All the labs were of different sizes and would handle different volumes of tests. Similarly, the labs had different numbers of technicians and equipment. The number of plates and techs at each lab would depend on the size of the lab and the volume of testing being done there.

There were close to 2,000 testing centers for SHIELD. We have the locations for most of the centers in our data. All the testing centers were located in Illinois. SHIELD was not allowed to have customers outside of Illinois. SHIELD T3 collected outside Illinois.

**Data**

We have around 10 datasets.

1. Equipment
2. Plate\_96
3. Plate\_96\_Scan
4. Sample
5. Sample\_Plate\_96
6. Specimen
7. Specimen\_Rejected
8. T\_Covid\_Specimen
9. Accounts
10. Opportunities

The specimen and t\_covid datasets are the largest of the ten.

**Equipment**

There are 386 observations of 6 variables in this dataset. The variables are Barcode, ID, Name, Serial, Type and LabID. This dataset gives us information about the equipment used in the labs. A lot of data is missing and seems incorrect in the Equipment name and type columns. Anna from SHIELD did not have much insight to provide regarding these.

**Plate**

Plate\_96 and Plate\_96\_Scan datasets range from February 2021 to June 2023. Plate\_96 has 91,140 observations and 7 variables. The variables are BufferMix\_ID, Equipment\_ID, Lab\_ID, Plate\_Barcode, Plate\_ID, Tech\_ID and Timestamp. Plate\_96\_Scan has 80,802 observations and 5 variables; Equipment\_ID, Plate\_ID, Plate\_Scan\_ID, Tech\_ID, and Timestamp. The common variables between the two datasets are Equipment\_ID, Plate\_ID, and Tech\_ID. 38.45% of data (Equipment\_ID) is missing from Plate\_96\_Scan​. From the data, the highest number of plates processed was on 01/20/22.

There is a difference in the Plate and Plate Scan datasets. Are all the plates not scanned? Or is there missing data in the Scan dataset?​ All labs have plates that are not present in the Plate\_Scan dataset. Reason?

As per Anna, each plate would be used on just one equipment for the PCR testing. The plate might be used on another equipment for other processes.

**Sample**

Sample dataset has 6,598,721 observations and 6 variables; Lab\_ID, Sample\_Barcode, Sample\_ID, Tech\_ID, Timestamp and Specimen\_Count. The data ranges from February 2021 to June 2023. Sample\_Plate dataset has 6,876,649 observations and 4 variables; Plate\_Scan\_ID, Sample\_ID, Plate\_ID, and Plate\_Well. October 2022 onwards, there is data for only Lab 5, 8 and 13. The rest of the labs stopped operating.

The column Specimen\_Count shows a value of 1 for each sample ID. This reiterates the idea that a sample is prepared from a specimen. One specimen per sample. The specimen taken from an individual is treated to become a sample, which is then used for testing. However, Anna mentioned that the terms sample and specimen were often used interchangeably.

**Specimen**

The specimen dataset has 6,599,595 observations and 21 variables/columns. Three different time stamps, TS, TS.Lab1, and TS.TestResult1. According to Anna, the TS indicates the time the manifest was sent/closed. TS.Lab1 indicates when the manifest reached the lab and TS.TestResult1 tells us when the result was generated. As per Anna, the data might not be too accurate here.

Specimen\_Rejected has 112,012 observations and 16 variables/columns. Using this dataset, we identified Lab ID 11 as the Depot. ‘Order Cancelled’ was identified as the most common reason for specimen rejection. However, Anna told us that order cancelled did not mean that the specimen was rejected, rather the whole order/manifest was cancelled at the testing location. Sorting of specimens/manifests would happen at the Depot. During this process, some specimens were rejected at the depot. Most specimens were rejected at the labs. If the specimens did not meet the standards at the collection center, they would just be thrown out and not included in the manifest, hence we would not have their data.

**T\_Covid\_Specimen**

The dataset has 2,907,747 observations and 88 variables/columns. We get insights into manifests in this dataset, and we have multiple timestamps in this dataset as well. Out of 2,907,747 we found 2,906,850 unique specimens from specimen barcodes. From the dataset we identified the peak testing periods in Oct 2021 leading up to December 2021 and January 2022. Anna explained that this can be attributed to more than 500 school districts signing up together with SHIELD. Another thing Anna addressed was the difference in result times for manifests. She told us that one manifest contains a number of specimens, and sometimes all specimens from a manifest will be processed on different plates, hence the same manifest can result in different result times. The mean time from collection to result is 19.22 hours as per the dataset. There are 39,794 manifests in the data. There is an average of 5.6 hours between the manifest created and closed timestamps.

**Accounts and Opportunities**

These datasets are two separate pages pulled from SHIELD CRM, as reported by Anna. In the Opportunities dataset there is a column for ‘Stage’. If the stage is shown as ‘Expressed Interest’, SHIELD never tested in these locations. The poll happened near the end of SHIELD hence they could not get a chance to go forward with the testing.

The accounts dataset has location related data. In account type it categorizes the accounts in different categories, like Collection Agency, Collection Sub-Agency, Collection Location, etc. All the testing centers fall under the category of collection location.