

Pakistan Police Project: Data Exploration Notes

November 24, 2022

This document provides notes on our exploration of the Complaints and First Impression Reports data.

1 Dataset Overview

The numbers of unique observations for each of the FIR and Complaint datasets are the following:

Unique Complaints cases → 154058

Unique FIR cases → 2255694

The merged FIR + Complaint dataset looks like this. The “In both” category shows the unique matched observations between the two datasets.

Table 1: Match between the FIR and Complaint Data

	Frequency	Percent
Data source		
FIR	2,112,426	93.20
Complaints (Lahore)	10,790	0.48
In both	143,268	6.32
Total	2,266,484	100.00

2 Response Speed

These are summary statistics on police response speed.

One thing to note is that there are quite a bit of negative values.

Table 2: Response Speed

	count	min	max	p50	mean
Complaint-1st Contact Time (hours)	143268	-.0197222	2726.877	10.16653	39.7748
Complaint-FIR conversion time (hours)	143268	-26347.93	16247.58	-.3427778	1.866467
Observations	143268				

Without negative values, the summary statistics look like this.

Here are the distributions. I cut off some outliers so that we can see some patterns.

A question: Why do we see some waves?

Table 3: Response Speed

	count	min	max	p50	mean
Complaint-1st Contact Time (hours)	68871	.0075	2726.877	50.05167	71.81699
Complaint-FIR conversion time (hours)	68871	6.78e-11	16247.58	34.64444	60.28819
Observations	68871				

Figure 1: Estimated Effects: All Treatment Categories

