1 Analyzing All Data

Here we analyze all of the data.

First we load the data and view a portion of it. Some more details.

These are the necessary packages.

```
require(useful)
require(plyr)
require(ggplot2)
```

```
load("../../data/pakistan/pak.rdata")
source("../../R/distFuncs.r")
corner(pak, c = 15)
```

```
  New_ID Age  Sex     Date Province District Tehsil
1   1288  26 Male 29082010      KPK  Shangla Besham
2   1290  30 Male 29082010      KPK  Shangla Besham
3   1370  54 Male 28082010      KPK  Shangla Besham
4   1372  53 Male 28082010      KPK  Shangla Besham
5   1371  64 Male 28082010      KPK  Shangla Besham
         Village Latitude Longitude Total Urban Rural
1 abaseen colony    34.94     72.88  90.6     -  90.6
2 abaseen colony    34.94     72.88  90.6     -  90.6
3 abaseen colony    34.94     72.88  90.6     -  90.6
4 abaseen colony    34.94     72.88  90.6     -  90.6
5 abaseen colony    34.94     72.88  90.6     -  90.6
                                 Accommodation
1 Collective centers (school/Public building)
2                                  Host family
3          On the site of the house (Damaged)
4          On the site of the house (Damaged)
5          On the site of the house (Damaged)
  StagnantWater
1           Few
2           Few
3           Few
4          None
5          None
```

Now we build a distribution for all the data and visualize it in Figure 1 with the code here:.

```
ricePerc <- build.dist(data = pak, lhs = "New_ID", group = "Province",
    question = "RiceLost")
ricePerc$Size <- "All"
```

```
ggplot(ricePerc, aes(x = RiceLost, y = Percent)) + geom_bar(stat = "identity") +
    facet_wrap(~Province) + opts(axis.text.x = theme_text(angle = 90))
```

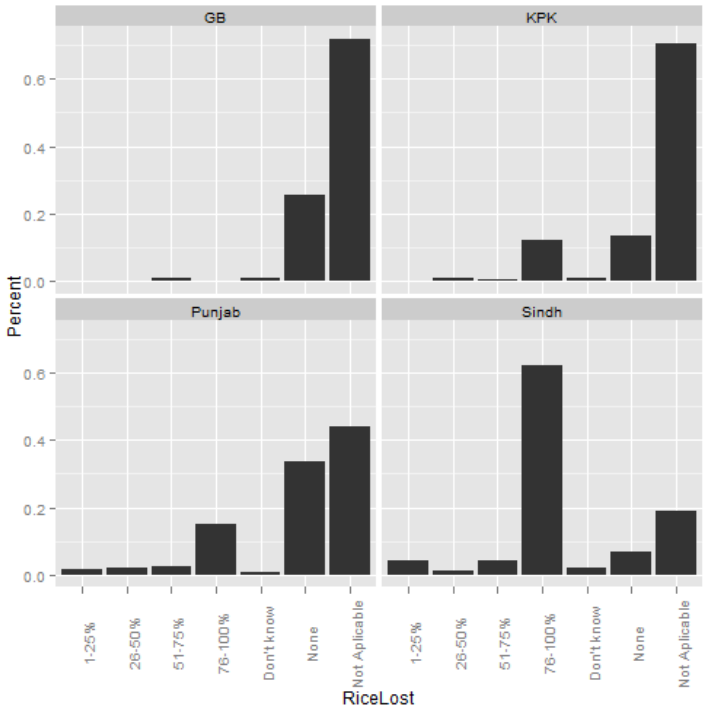

Figure 1: Graphical view of the distribution of responses for all the data.

In Section **??** we analyze the distribution of responses for samples of fewer Tehsils.