

Data Exploration

Jenn Havens

November 8, 2017

```
require(dplyr)
```

```
## Loading required package: dplyr
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
require(ggplot2)
```

```
## Loading required package: ggplot2
TestPred <- read.csv("./DengAI_Predicting_Disease_Spread_-_Test_Data_Features.csv")

TrainPred <- read.csv("./DengAI_Predicting_Disease_Spread_-_Training_Data_Features.csv")
TrainResp <- read.csv("./DengAI_Predicting_Disease_Spread_-_Training_Data_Labels.csv")

TrainFull <- full_join(TrainResp, TrainPred, by = c("city", "year", "weekofyear"))

enviroPredNames <- colnames(TrainFull)[6:25]

for (variable in enviroPredNames){
  ggplot(data = TrainFull) + geom_point(aes(x=variable, y=total_cases))
}
```