

2.0 Multivariate Visualization Assignment

JD Santos

Problem 1

Use the bivariate boxplot on the scatterplot of pairs of variables ((temp, wind), (temp, precip)) in the air pollution data to identify any outliers. Calculate the correlation between each pair of variables using all the data and the data with any identified outliers removed. Comment on the results.

1. Begin by pulling the variable pairs from USairpollution into objects

```
data("USairpollution", package = "HSAUR2")
temp_wind <- USairpollution[, c("temp", "wind")]
temp_precip <- USairpollution[, c("temp", "precip")]
```

2. Create the bivariate boxplot for temp/wind pair

```
bvbox(temp_wind, xlab = "Temperature", ylab = "Wind (MPH)", type = "n")

text(USairpollution$temp, USairpollution$wind,
     cex = 0.6,
     labels = abbreviate(row.names(USairpollution)))
```

