1. **Import the data**

library(readxl)

IPV\_Table <- read\_excel("C:/Users/Mahsa/Desktop/Assignment #3/IPV Table.xlsx",

col\_types = c("text", "text", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric"))

1. **Create a subset of your dataset with only three variables**

> #subset:

> current<-IPV\_Table[c(1:2083),c(10,11,12)]

> str(current)

Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 2083 obs. of 3 variables:

$ 2009: num 4 0 2 617 1 434 55 1 29 0 ...

$ 2010: num 7 0 4 617 0 364 51 0 33 0 ...

$ 2011: num 8 0 3 548 0 339 53 0 42 0 ...

1. **Describe the central tendency of the variables**

**Mean:**

2009: 33.03

2010: 31.59

2011: 28.13

**Median:**

2009: 1

2010: 1

2011: 1

**Mode:**

2009: 0

2010: 0

2011: 0

1. **Describe the variables through variation**

**Variance:**

2009: 34306.65

2010: 33158

2011: 23148.16

**Range:**

2009: 4294

2010: 3925

2011: 3601

**Standard Deviation:**

2009: 185.2206

2010: 182.0934

2011: 152.1452

1. **Discuss normality. Are your variables approximating normality? What are ways that you can show that they are or are not?**











