

# Lab 3

David Hou

March 20, 2018

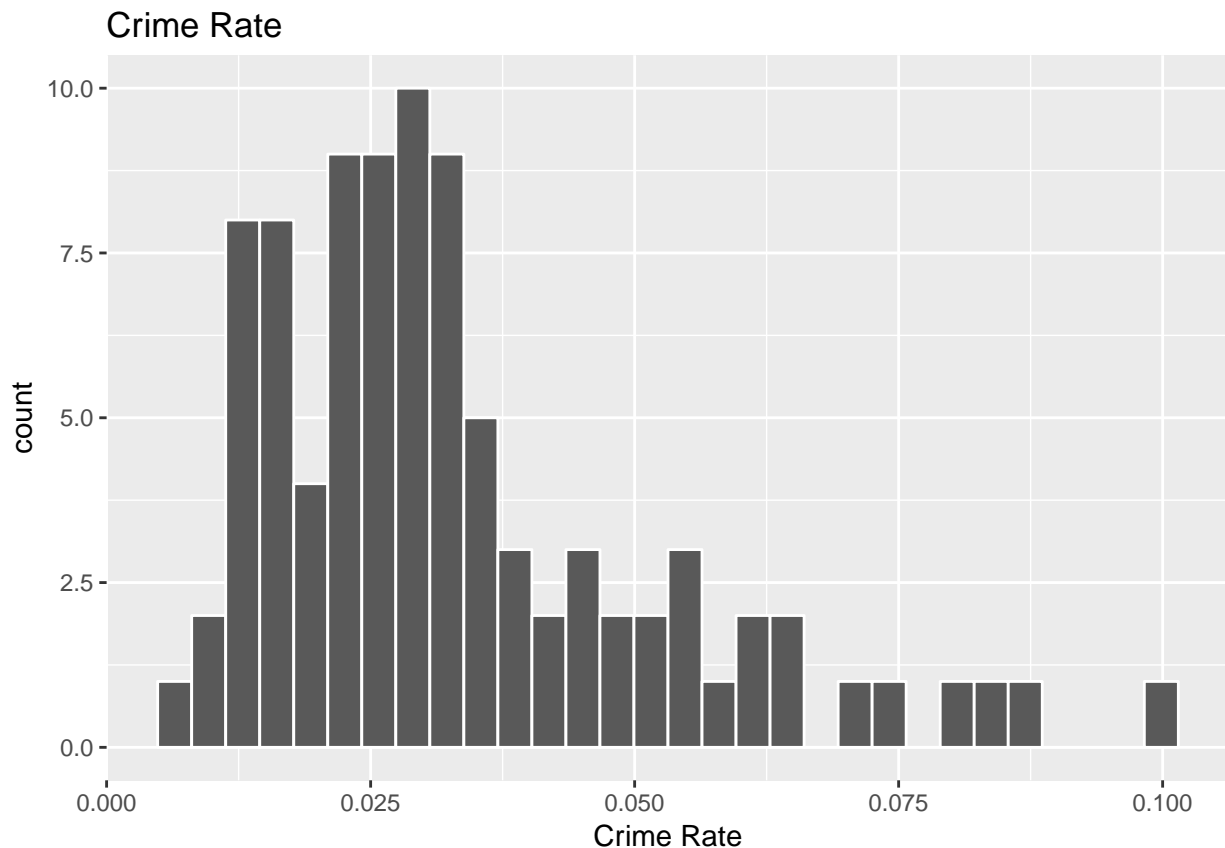
First we clean the data, using the dplyr package for its nice verbs. We remove NAs, change prbconv to numeric, and change all integer columns to factors.

```
raw = as_tibble(read.csv('crime_v2.csv'))
t = raw %>%
  filter(!is.na(county)) %>%
  mutate(prbconv = as.numeric(prbconv) / 100) %>%
  mutate_if(is.integer, as.factor)
levels(t$west) = c('East', 'West')
t$west = relevel(t$west, 'West') # Put West first so it appears on the left on facet plots
levels(t$central) = c('Outer', 'Central')
levels(t$urban) = c('Non-urban', 'Urban')

t = t %>% mutate(wage = wcon + wtuc + wtrd + wfir + wser + wmfg + wfed + wsta + wloc)
```

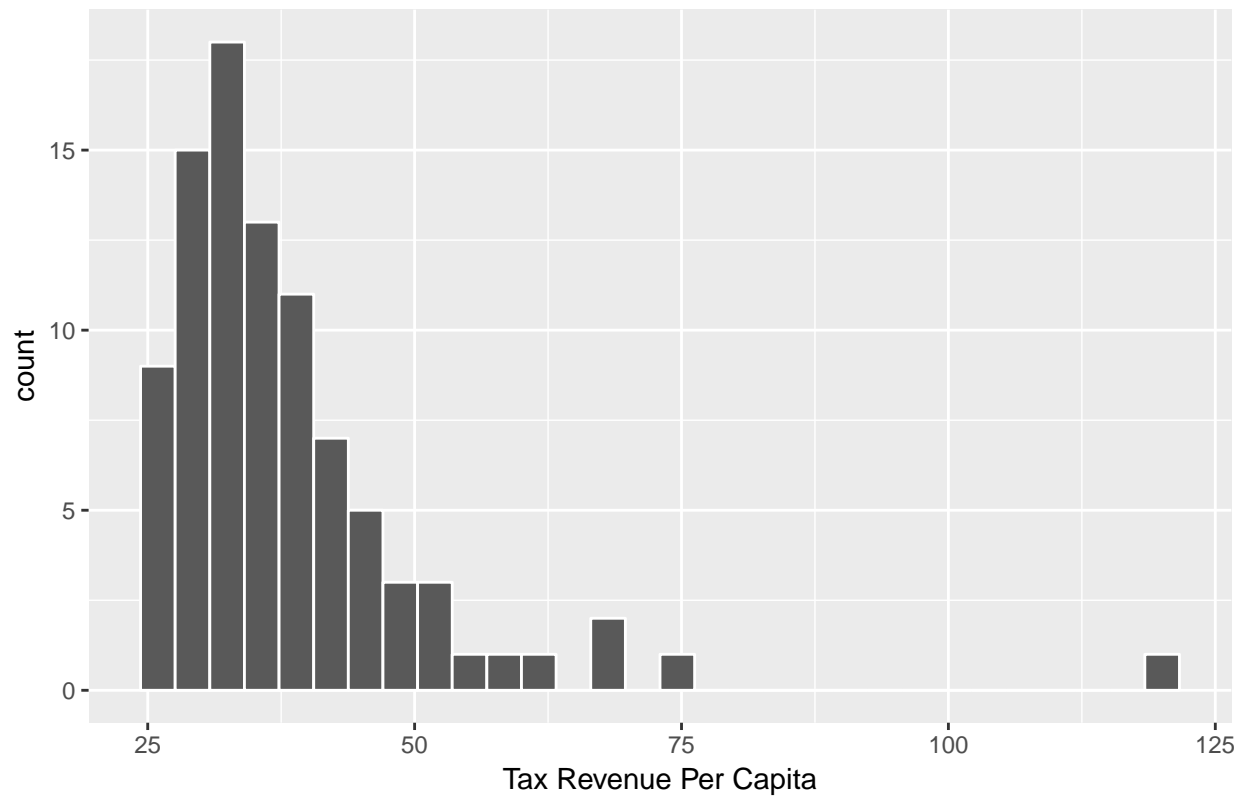
Variables of interest to a politician: 1. Crime rate 2. Tax revenue per capita 3. Wages

```
qplot(t$crmrte, geom = 'histogram', col = I('white'), main = 'Crime Rate', xlab = 'Crime Rate')
```

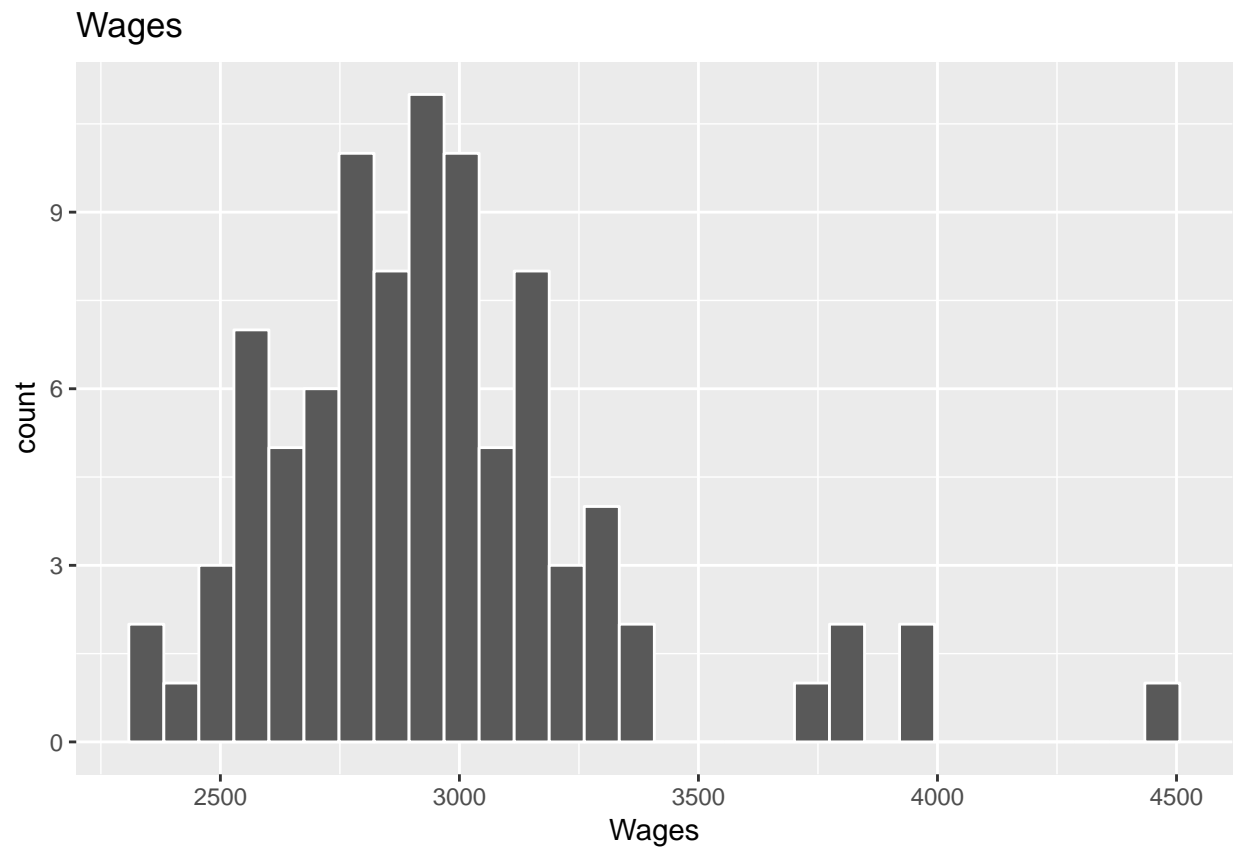


```
qplot(t$taxpc, geom = 'histogram', col = I('white'), main = 'Tax Revenue Per Capita', xlab = 'Tax Revenue')
```

Tax Revenue Per Capita

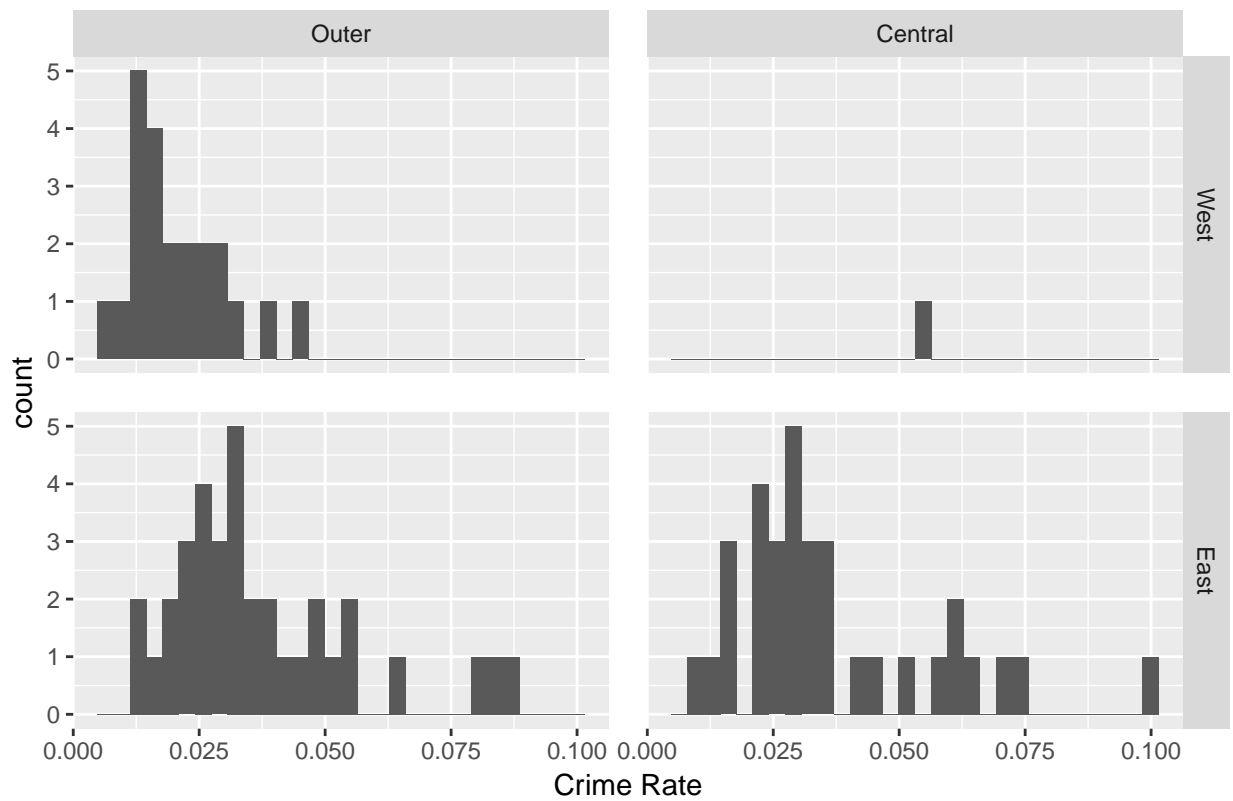


```
qplot(t$wage, geom = 'histogram', col = I('white'), main = 'Wages', xlab = 'Wages')
```

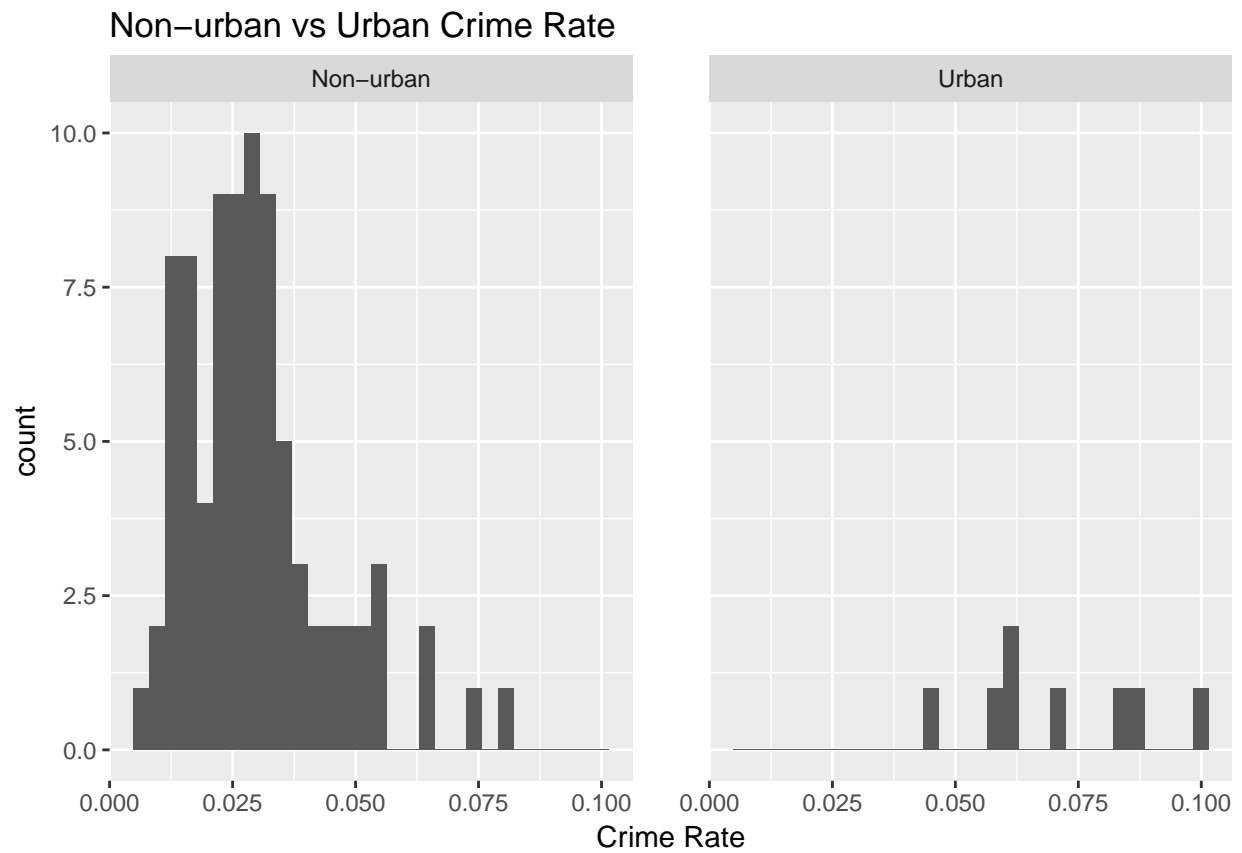


```
ggplot(t, aes(crmrte)) +  
  geom_histogram() +  
  facet_grid(west ~ central) +  
  theme(panel.spacing = unit(1, "lines")) +  
  labs(title = 'Crime Rate by Region', x = 'Crime Rate')
```

# Crime Rate by Region

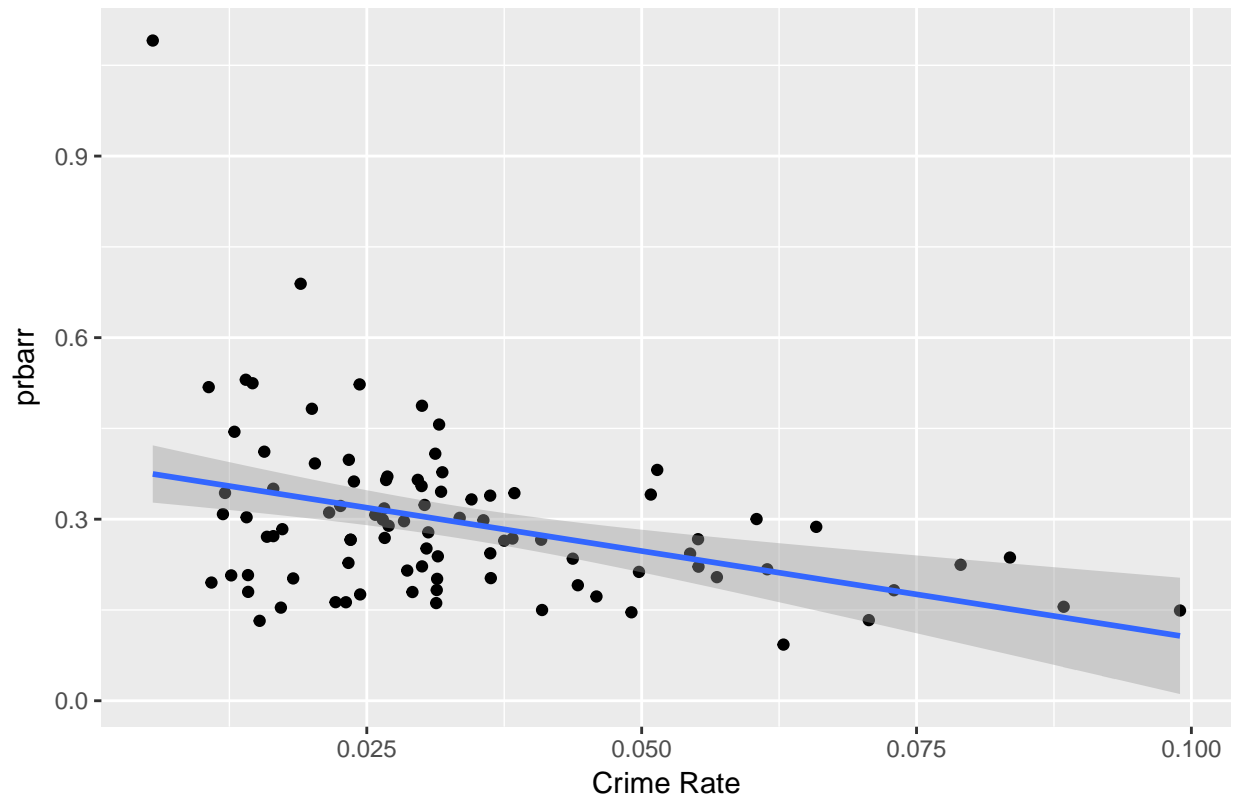


```
ggplot(t, aes(crmrte)) +
  geom_histogram() +
  facet_grid(. ~ urban) +
  theme(panel.spacing = unit(2, "lines")) +
  labs(title = 'Non-urban vs Urban Crime Rate', x = 'Crime Rate')
```



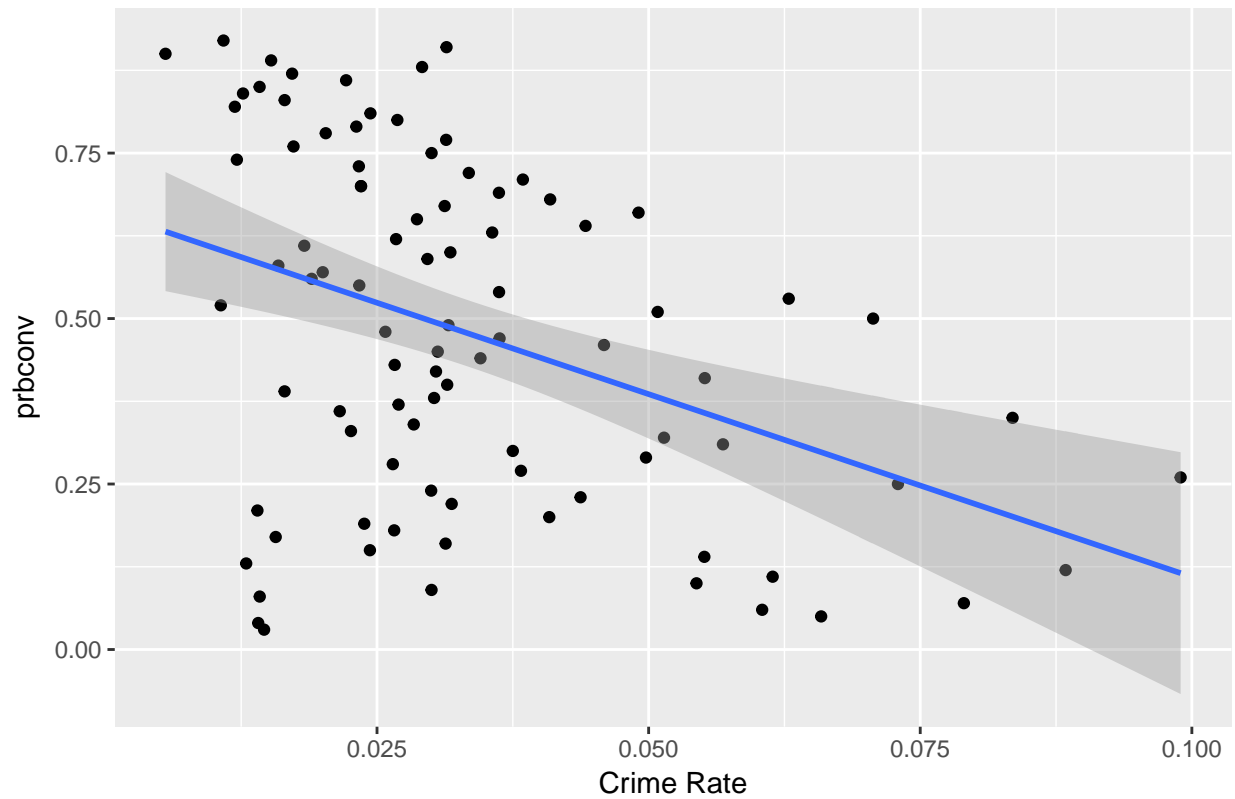
```
ggplot(t, aes(crmrte, prbarr)) +  
  geom_point() +  
  geom_smooth(method = 'lm') +  
  labs(title = 'Crime Rate vs Arrest Probability', x = 'Crime Rate')
```

Crime Rate vs Arrest Probability



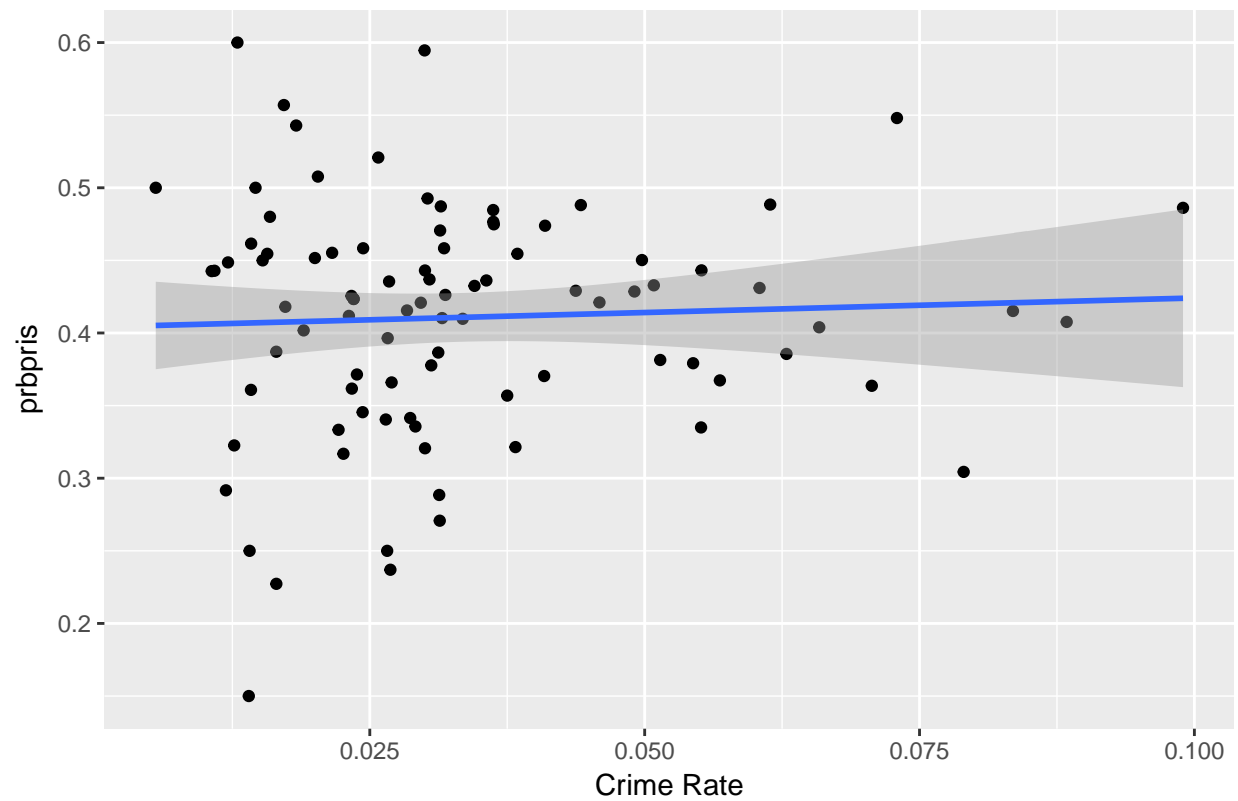
```
ggplot(t, aes(crmrte, prbconv)) +  
  geom_point() +  
  geom_smooth(method = 'lm') +  
  labs(title = 'Crime Rate vs Conviction Probability', x = 'Crime Rate')
```

Crime Rate vs Conviction Probability



```
ggplot(t, aes(crmrte, prbpris)) +  
  geom_point() +  
  geom_smooth(method = 'lm') +  
  labs(title = 'Crime Rate vs Prison Probability', x = 'Crime Rate')
```

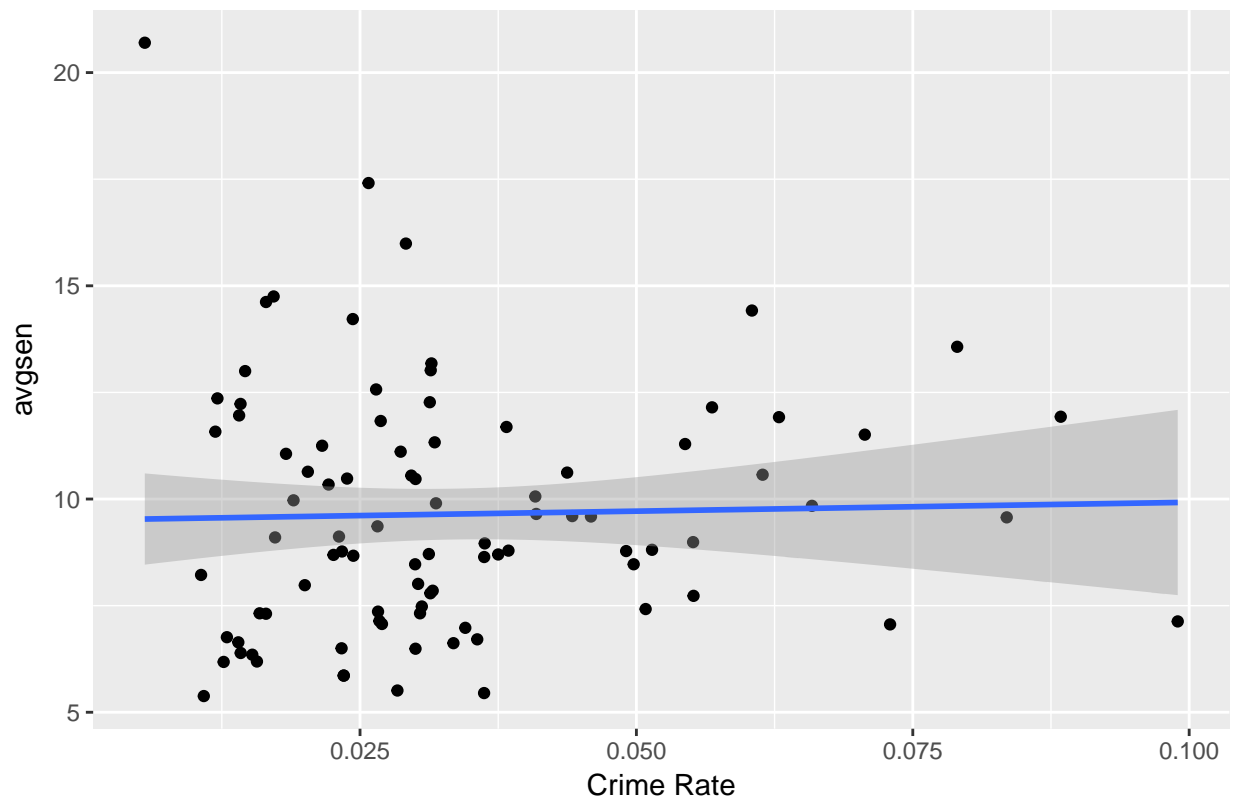
Crime Rate vs Prison Probability



```
ggplot(t, aes(crmrte, avgpsen)) +  
  geom_point() +  
  geom_smooth(method = 'lm') +  
  labs(title = 'Crime Rate vs Average Prison Sentence', x = 'Crime Rate')
```



Crime Rate vs Average Prison Sentence



```
ggplot(t, aes(crmrte, polpc)) +  
  geom_point() +  
  geom_smooth(method = 'lm') +  
  labs(title = 'Crime Rate vs Police Per Capita', x = 'Crime Rate')
```

