DallasArrests

Group Project

2024-11-03

Introduction

This report analyzes the relationship between arrest rates and property prices in Dallas by zip code. Using datasets on arrests and property values, we aim to identify any correlations that may exist at the neighborhood level.

Data handling

```
library(ggplot2)
# Load necessary libraries
library(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
```

```
# Load the data
arrests <- read.csv("DallasArrests_cleaned.csv")
zillow <- read.csv("Zillow-dallas.csv")

# Count Arrests per Zip Code
arrests_per_zip <- arrests %>%
  group_by(Zipcode = Arrest.Zipcode) %>%
  summarise(Arrest_Count = n())

# Calculate Average or Median Property Prices per Zip Code
property_price_stats <- zillow %>%
  group_by(Zipcode = addressZipcode) %>%
  summarise(
    Average_Price = mean(price, na.rm = TRUE),
    Median_Price = median(price, na.rm = TRUE)
  )

# Merge the Results
summary_df <- merge(arrests_per_zip, property_price_stats, by = "Zipcode")
```

```r
# Save the Results
print(summary_df)
```

```
##    Zipcode Arrest_Count Average_Price Median_Price
## 1    75104            1      329900.0     329900.0
## 2    75134           19      209900.0     209900.0
## 3    75201         3762     1994333.3     899000.0
## 4    75202         2520      297718.8     284250.0
## 5    75203         1258      329950.0     334900.0
## 6    75204         2540      594946.1     575000.0
## 7    75205          131     1024950.0    1024950.0
## 8    75206         1718      745423.0     650000.0
## 9    75208         1277      818478.5     592500.0
## 10   75209          379     3193186.8    1474499.5
## 11   75210         1184      155833.3     182500.0
## 12   75211         2369      381676.1     325000.0
## 13   75212         1228      409331.1     360000.0
## 14   75214          732      898055.7     747000.0
## 15   75215         2855      292823.3     277000.0
## 16   75216         3490      256529.4     250000.0
## 17   75217         2881      346141.4     280000.0
## 18   75218          588      718922.0     625000.0
## 19   75219         1281      413275.0     391500.0
## 20   75220         3455     1698224.8     969500.0
## 21   75223          787      529030.7     390000.0
## 22   75224         1292      431149.9     366500.0
## 23   75225          721     2524600.0    1570000.0
## 24   75226         1645       85000.0      85000.0
## 25   75227         1516      277029.6     284000.0
## 26   75228         2297      359125.9     349900.0
## 27   75229         1413     2850438.9     599000.0
## 28   75230          281     1516520.6    1099000.0
## 29   75231         2162      455094.2     440000.0
## 30   75232         1124      293549.9     287450.0
## 31   75233          333      393149.7     354999.5
## 32   75234          189      263600.0     259950.0
## 33   75235         1456      430460.0     387000.0
## 34   75236          267      332442.9     290000.0
## 35   75238          717      633026.7     624900.0
## 36   75240         1041      533422.2     450000.0
## 37   75241         1416      224970.6     239900.0
## 38   75243         2769      350712.8     309999.0
## 39   75244          145      835791.7     902400.0
## 40   75246          421      975000.0     975000.0
## 41   75248          439      959107.1     697499.5
## 42   75249          132      315527.2     349900.0
## 43   75252          231      554286.1     594700.0
## 44   75253          331      303005.8     287495.0
## 45   75254          413      683800.0     479900.0
## 46   75287          711      686076.8     530000.0
```

```r
write.csv(summary_df, "Dallas_Arrests_Property_Summary.csv", row.names = FALSE)
```

Data visulization

```
ggplot(summary_df, aes(x = Average_Price, y = Arrest_Count)) +
  geom_point(color = "blue") +
  labs(
    title = "Arrest Counts vs. Average Property Price by Zip Code",
    x = "Average Property Price",
    y = "Total Arrests"
  ) +
  theme_minimal()
```