Intro_Visual

Connor Mignone

4/18/2022

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.1     v dplyr   1.0.6
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(readxl)
library(readr)
```

```
ZestAccuracy <- read_excel("../data/ZestAccuracy.xlsx",
    col_types = c("text", "text", "text",
        "text", "text", "text", "text"))
View(ZestAccuracy)
```

```
df<- ZestAccuracy %>% select(`Median Error`,`Within 5% of Sale Price`,`Within 10% of Sale Price`,`Withi
df <- data.frame(sapply(df, function(x) as.numeric(gsub("%", "", x))))
```

```
## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion
```

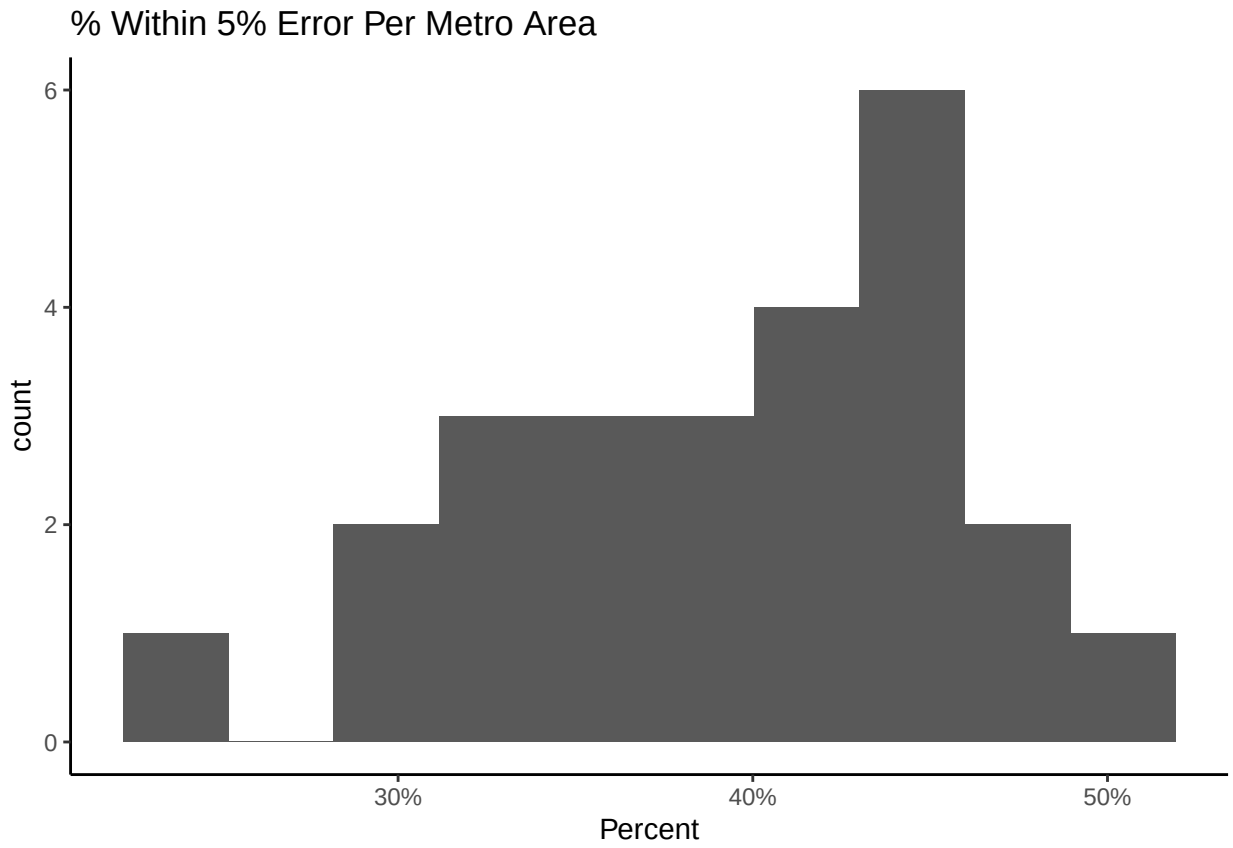
```
df <- cbind(df, ZestAccuracy)
```

```
df1=df %>% filter(Type == 'Off-Market')
ggplot(data = df1)+
  geom_histogram(aes(Median.Error), bins = 10)+
  labs(title='Metropolitan Zestimate Median Error')+
  xlab("Percent Error")+
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_classic()+
  scale_x_continuous(labels = scales::percent_format(scale = 1,suffix = "%"))
```

```
## Warning: Removed 5 rows containing non-finite values (stat_bin).
```


```
ggplot(data = df1)+
  geom_histogram(aes(Within.5..of.Sale.Price), bins = 10)+
  labs(title='% Within 5% Error Per Metro Area')+
  xlab("Percent")+
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_classic()+
  scale_x_continuous(labels = scales::percent_format(scale = 1,suffix = "%"))
```

```
## Warning: Removed 5 rows containing non-finite values (stat_bin).
```


% Within 5% Error Per Metro Area
count
6
4
2
0
30%
40%
50%
Percent