A00275664 RMarkdown

Érin Kennedy

2023-11-30

RMarkdown Document

Question 1

Importing and Analysing the Dataset

```r
PropertyDataset1 <- read.csv("C:/Users/erink/OneDrive/Desktop/Interpretation of
↪  Data/PropertyDataset1.csv",
    header = TRUE, sep = ",")

names(PropertyDataset1)
```

```
##  [1] "geohash"                  "lat"
##  [3] "lon"                      "sale_date"
##  [5] "price"                    "formatted_address"
##  [7] "raw_address"              "post_code"
##  [9] "county"                   "num_bedrooms"
## [11] "num_bathrooms"            "property_description"
## [13] "property_size_description" "formatted_description"
## [15] "not_full_market_price"    "vat_exclusive"
## [17] "parking"                  "parking1"
```

(a)

```r
PropertyDataset1$sale_date<- dmy(PropertyDataset1$sale_date)
PropertyDataset1$sale_year<- year(PropertyDataset1$sale_date)


house2017_3bed_2btrm <- PropertyDataset1 %>%
  filter(num_bedrooms == "3 Bedrooms",
         num_bathrooms == "2 Bathrooms",
         sale_year == 2017)
```

(b)

```r
Cork_Galway_125Sq <- PropertyDataset1 %>%
  filter(county %in% c("Galway", "Cork"),
         property_size_description == "greater than 125 sq metres") %>%
  arrange (desc(price))
```

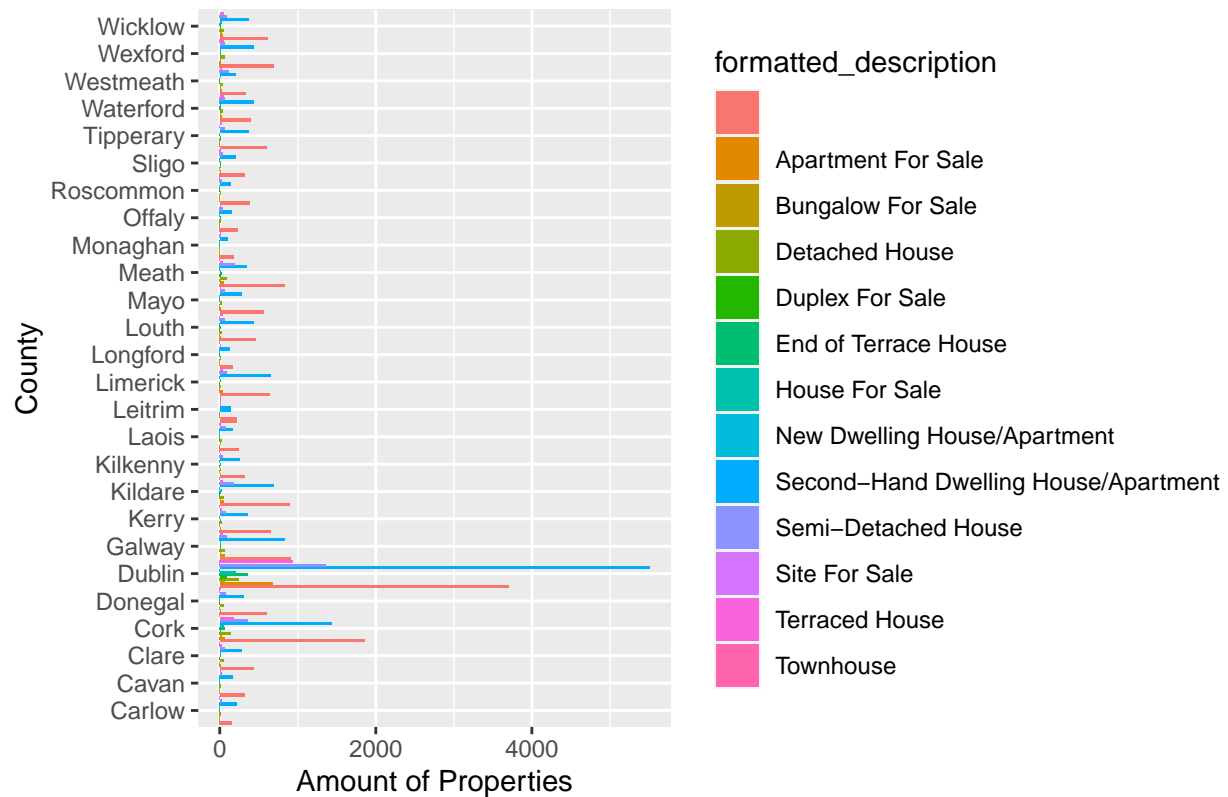**(c)**

```r
Avg_Max_Min_Dwell_Apt<- PropertyDataset1 %>%
  filter(post_code %in% c("Dublin 18", "Dublin 16", "Dublin 15", "Dublin 8", "Dublin"),
         property_description == "New Dwelling house /Apartment")

summ_stats<-Avg_Max_Min_Dwell_Apt %>%
  dplyr::summarise(Avg_Price = mean(price),
                   Max_Price=max(price),
                   Min_Price=min(price))
print(summ_stats)
```

```
##   Avg_Price Max_Price Min_Price
## 1 384030.7   1220000      8000
```

(d)

```r
ggplot(PropertyDataset1, aes(x = county, fill = formatted_description)) +
  geom_bar(width=1, position = "dodge") +
  labs(title = "Property Type Sold per County in Ireland",
       x = "County",
       y = "Amount of Properties")+
  coord_flip()
```

Property Type Sold per County in Ireland

(e)

```
Propdata2<- PropertyDataset1 %>%
  filter(sale_year >= 2015 & sale_year <= 2017) %>%
  group_by(sale_year, formatted_description) %>%
  summarise(Count = n())
```

```
## `summarise()` has grouped output by 'sale_year'. You can override using the
## `.groups` argument.
```

```
ggplot(Propdata2, aes(x = factor(sale_year), y = Count, fill = formatted_description)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Number of Different Types of Houses Sold (2015-2017)",
       x = "Year Sold",
       y = "Amount")
```

Number of Different Types of Houses Sold (2015–2017)

(f)

```
PropData3 <- PropertyDataset1 %>%
  filter((county %in% c("Dublin", "Cork", "Galway")) & formatted_description ==
  ↪  "Semi-Detached House") %>%
  select(-num_bedrooms:-not_full_market_price) %>%
  arrange(price)
```

(g)

```
PropData4<- PropertyDataset1 %>%
  filter(county %in% c("Galway", "Roscommon", "Mayo", "Sligo"),
         price > 200000)
ggplot(PropData4, aes(x=sale_year, y=price, colour= county)) +
geom_line(linewidth=1.5) +
  labs(title = "House Prices in Galway, Mayo, Roscommon & Sligo",
       x="Year Sold",
       y="House Price")
```

House Prices in Galway, Mayo, Roscommon & Sligo