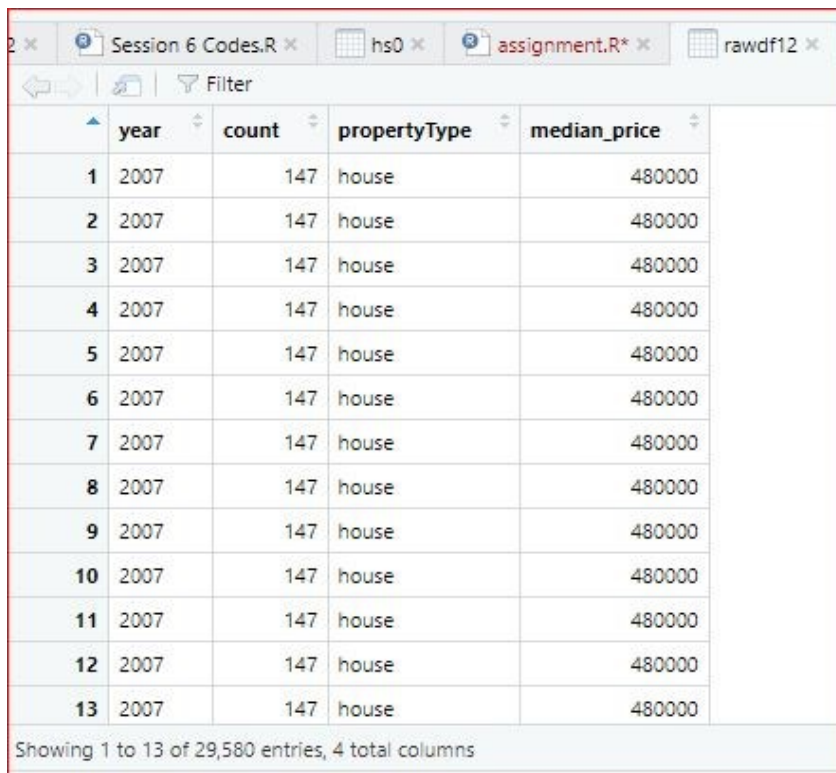**Time series**

https://www.kaggle.com/datasets/htagholdings/property-sales?select=raw_sales.csv

1. Yearly median housing Price

2. Housing Price Mean/Median, Standard Deviation/(1 st Quartile and 3 rd Quartile) by

number of bedrooms.

3. Housing Price Mean per year where the years will be in the columns and each row

will represent the number of bedrooms.

Query-1:

Solution:

```
rawdf12 <- rawdf %>%
  select (datesold, price, propertyType) %>%
  mutate (year = format(as.Date(datesold), "%Y")) %>%
  group_by(year) %>%
  summarise (count = n(), propertyType, median_price = median(price, na.rm = TRUE)   )
print(rawdf12)
```


	year	count	propertyType	median_price
1	2007	147	house	480000
2	2007	147	house	480000
3	2007	147	house	480000
4	2007	147	house	480000
5	2007	147	house	480000
6	2007	147	house	480000
7	2007	147	house	480000
8	2007	147	house	480000
9	2007	147	house	480000
10	2007	147	house	480000
11	2007	147	house	480000
12	2007	147	house	480000
13	2007	147	house	480000

Showing 1 to 13 of 29,580 entries, 4 total columns
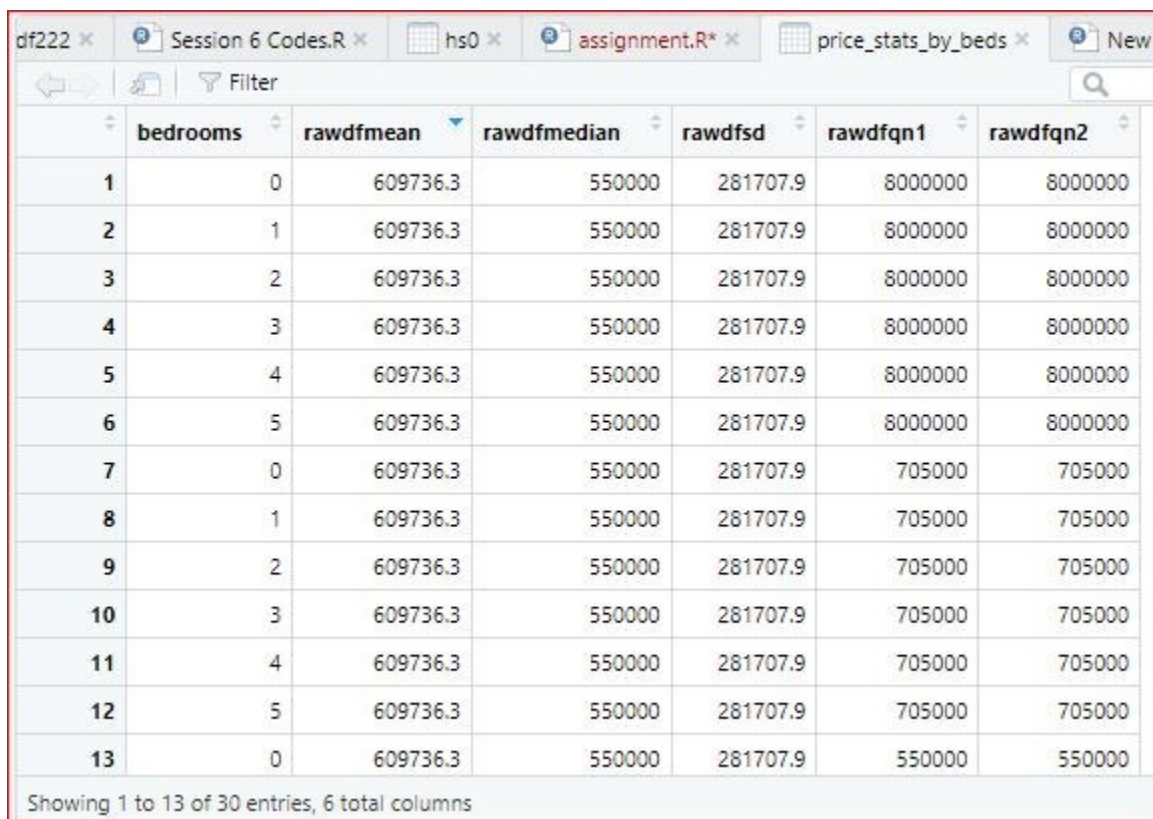
Query-2:

Solution:

Step-1:

rawdfmean<-mean(rawdf$price,na.rm=T)

rawdfmedian<-median(rawdf$price,na.rm=T)

rawdfsd<-sd(rawdf$price,na.rm=T)

rawdfqn1<-quantile(rawdf$price,pa=0.25,na.rm=T)

rawdfqn2<-quantile(rawdf$price,pa=0.75,na.rm=T)

Step-2:

price_stats_by_beds <- rawdf %>%

 group_by(bedrooms) %>%

 summarise(rawdfmean,rawdfmedian,rawdfsd,rawdfqn1,rawdfqn2)

print(price_stats_by_beds)
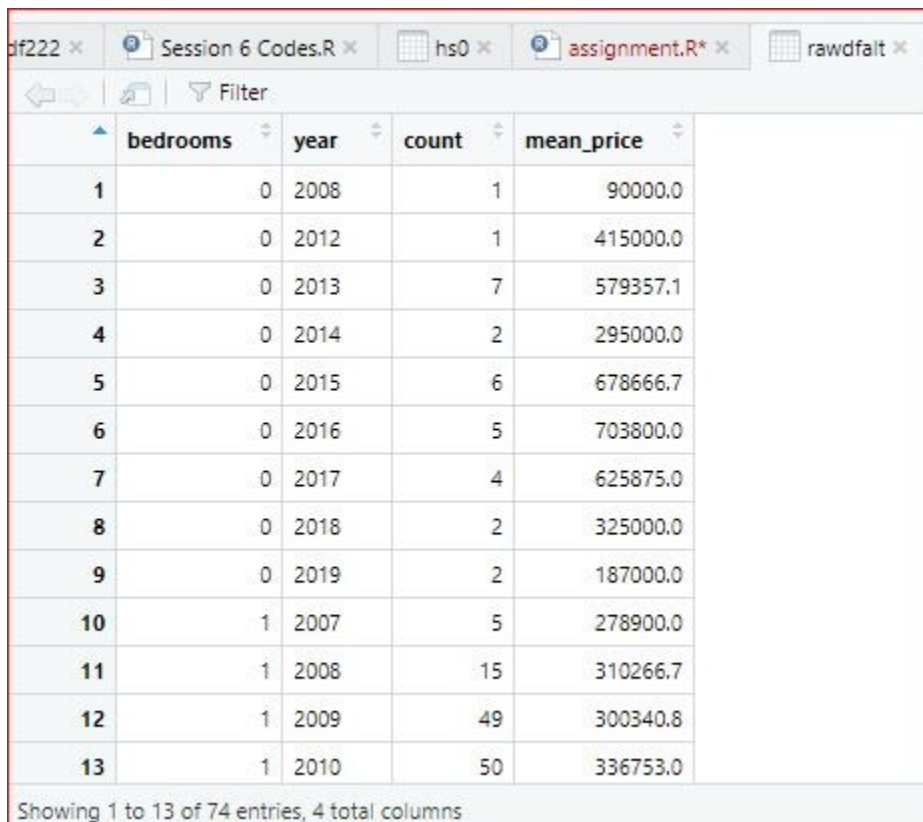
	bedrooms	rawdfmean	rawdfmedian	rawdfsd	rawdfqn1	rawdfqn2
1	0	609736.3	550000	281707.9	8000000	8000000
2	1	609736.3	550000	281707.9	8000000	8000000
3	2	609736.3	550000	281707.9	8000000	8000000
4	3	609736.3	550000	281707.9	8000000	8000000
5	4	609736.3	550000	281707.9	8000000	8000000
6	5	609736.3	550000	281707.9	8000000	8000000
7	0	609736.3	550000	281707.9	705000	705000
8	1	609736.3	550000	281707.9	705000	705000
9	2	609736.3	550000	281707.9	705000	705000
10	3	609736.3	550000	281707.9	705000	705000
11	4	609736.3	550000	281707.9	705000	705000
12	5	609736.3	550000	281707.9	705000	705000
13	0	609736.3	550000	281707.9	550000	550000

Showing 1 to 13 of 30 entries, 6 total columns

```
rawdfalt <- rawdf %>%

  select(price, bedrooms) %>%

  mutate() %>%  # Extract the year

  group_by(bedrooms, year) %>%

  summarise(count = n(), mean_price = mean(price, na.rm = TRUE)

  ) %>%

  ungroup()
```

print(rawdfalt)

	bedrooms	year	count	mean_price
1	0	2008	1	90000.0
2	0	2012	1	415000.0
3	0	2013	7	579357.1
4	0	2014	2	295000.0
5	0	2015	6	678666.7
6	0	2016	5	703800.0
7	0	2017	4	625875.0
8	0	2018	2	325000.0
9	0	2019	2	187000.0
10	1	2007	5	278900.0
11	1	2008	15	310266.7
12	1	2009	49	300340.8
13	1	2010	50	336753.0

Showing 1 to 13 of 74 entries, 4 total columns