# HW1

Evan Cranmer

2024-09-03

## Homework 1 - Evan Cranmer

### Question 1

```
hprice$homeprice<-exp(hprice$narsp)*1000

mean(hprice$homeprice)
```

```
## [1] 94411.42
```

```
var(hprice$homeprice)
```

```
## [1] 1583110349
```

The mean home price is this dataset is 94,411.42. This is the average home price in USD. The variance of homeprice is 1583110349. This number is representative of the spread of home prices within the dataset. Specifically, if we take the observed sample values compared to the mean, square this number, and then divide that by (n-1).

### Question 2

```
## [1] 1.967336
```

```
## [1] 90062.70 98760.14
```

```
##
##  One Sample t-test
##
## data:  hprice$homeprice
## t = 42.711, df = 323, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  90062.70 98760.14
## sample estimates:
## mean of x
##  94411.42
```

Lower bound 95% CI: 90,062.7 Upper bound 95% CI: 98,760.1

The confidence interval implies that if we were to take repeated samples from the population, the true mean of the average home price lies within 90,062 and 98.760 dollars 95% of the time.

### Question 3

```
library(plotrix)
hprice%>%
  group_by(ajwtr)%>% #grouped by new binary variable
  summarize(m=mean(homeprice),
            se=std.error(homeprice))
```

```
## # A tibble: 2 x 3
##   ajwtr       m    se
##   <fct>   <dbl> <dbl>
## 1 0      82389. 1229.
## 2 1     111243. 4656.
```

### Question 4

To compare means of home prices between coastline and non-coastline houses, I will use a two-sample T-test.

```
coast <- filter(hprice, ajwtr==1) ## Subset Coastal data to only include coastline homes
noncoast <- filter(hprice, ajwtr!=1) ## Subset noncoast data to only include noncoastal

#two-sample t test
t.test(coast$homeprice, noncoast$homeprice, var.equal=F)
```

```
##
##  Welch Two Sample t-test
##
## data:  coast$homeprice and noncoast$homeprice
## t = 5.9922, df = 152.79, p-value = 1.43e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  19340.96 38367.19
## sample estimates:
## mean of x mean of y
## 111242.96  82388.89
```

After conducting a two sample t-test we reject the null which states that there is no difference between the sample means. We fail to reject the alternative hypothesis which states the true difference in mean home prices is not equal to 0.

### Question 5 and 6

```
cor(hprice$homeprice, hprice$ypc)
```

```
## [1] 0.7437474
```

```
cor.test(hprice$homeprice, hprice$ypc)
```

```
##
##  Pearson's product-moment correlation
##
## data:  hprice$homeprice and hprice$ypc
## t = 19.965, df = 322, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6907661 0.7887854
## sample estimates:
##       cor
## 0.7437474
```

Estimated coefficient = 0.7437474 Since the p-value is less than .05, we can reject the null hypothesis that states the correlation coefficient we calculated is equal to 0. We fail to reject the alternative hypothesis that states the true correlation coefficient between homeprice and per capita income of the MSA is not equal to 0.

### Question 7

No, because the pearson's test is correlational and not causal.

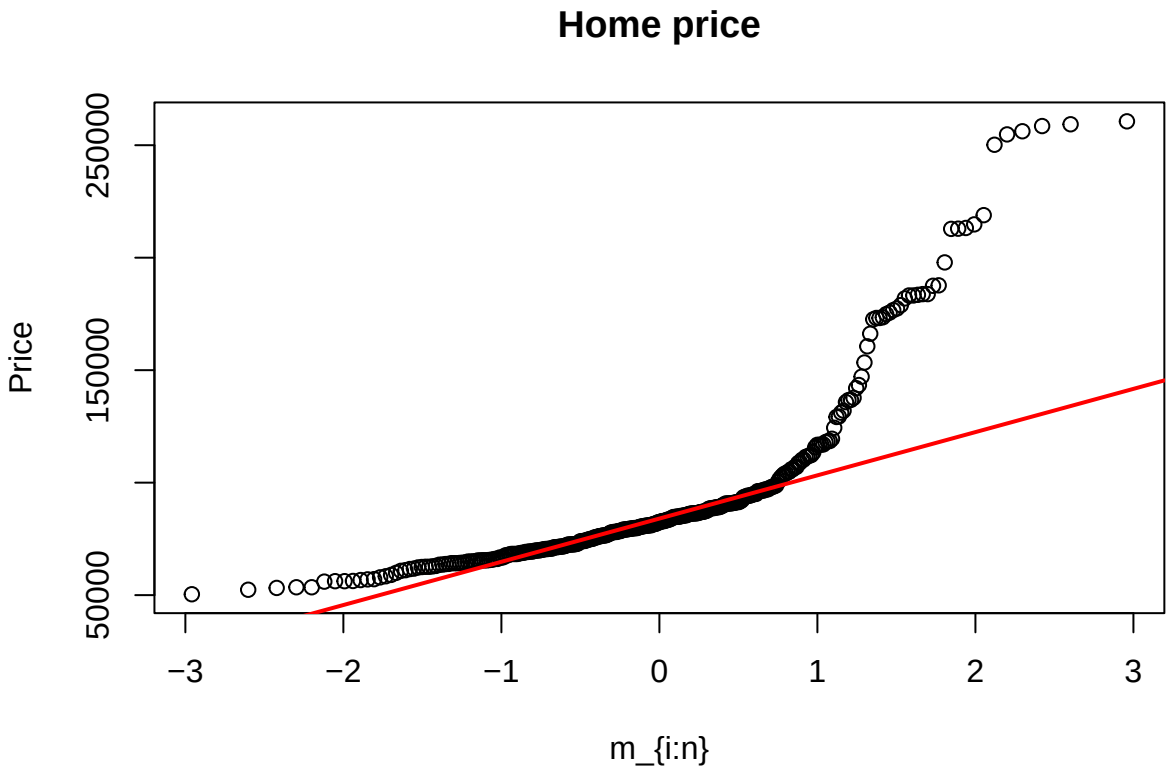
### Question 8

```
shapiro.test(hprice$homeprice)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  hprice$homeprice
## W = 0.7327, p-value < 2.2e-16
```

```
#graphing
qqnorm(hprice$homeprice, main="Home price", ylab="Price", xlab="m_{i:n}")+
qqline(hprice$homeprice, col="red",lwd=2)
```

```
## Error in qqnorm(hprice$homeprice, main = "Home price", ylab = "Price", : non-numeric argument to bina
```




According to the Shapiro-Wilk normality test of homeprice, we can reject the null hypothesis that states homeprice is normally distributed. If we look at the Q-Q plot we can see the distribution is skewed-right. This result does not change my answer to #7 because this test is looking at the normality of distribution and not a hypothesized causal relationship that per capita income has on house price.