A Statistical Analysis of the Airbnb Prices in Istanbul

Group-4
2365718 BERAT MERT KAYACAN
2364764 EMRE YILDIRIM
2482635 EREN ÜNAL
2365606 MAHMUT ESAT DİLİK

Abstract: This study analyzes Airbnb listings in Istanbul. The objective is to examine the distribution of the listing prices and summarize the dataset through descriptive and inferential statistical methods. Statistical measures such as the mean, median, standard deviation, confidence intervals, and hypothesis tests were used to evaluate the data. The normality was tested using graphical and statistical techniques. The results indicate that Airbnb prices are highly variable and do not follow a normal distribution due to the presence of extreme price values. These findings provide insight into the characteristics of Istanbul's short-term rental market while also highlighting the limitations of comparing prices across a long time period without accounting for economic changes. Overall, this study whows how statistical methods and tests can be applied to real-world datasets.

Presentation Link: https://www.youtube.com/watch?v=5Xn1vAfE820

1 INTRODUCTION

1.1 Research Question, Its Importance and Background Information

In recent years, platforms for short-term rental houses such as Airbnb have become increasingly popular in the tourism industry. Airbnb is one of the largest platforms that allows the property owners to rent their homes and rooms to tourists or travelers for short periods. Because of Istanbul being a major tourist destination in the world, it has a large and active Airbnb market. Airbnb prices may vary depending on different factors such as its neighborhood, room type and review numbers. Research question in this sense is: What factors have a significant impact on Airbnb listing prices in Istanbul?

Understanding the factors that influence Airbnb prices in Istanbul is important for both hosts and customers. The findings of this research may help hosts to make better pricing decisions and also help customers to use this information to understand the relationship between the Airbnb listing features and the Airbnb prices.

1.2 Objectives of the Study

The aim of this study is to:

- Examine the relationship between listing characteristics and pricing using statistical analysis.
- Identify the factors that significantly affect the Airbnb listing prices in Istanbul.
- Compare Airbnb prices across different features of the data.

1.3 The Data

The dataset used in this study is a single comma seperated values (CSV) file that consists of Airbnb listings in Istanbul obtained from InsideAirbnb, a database that provides publicly available Airbnb data. There are 15 columns and 22865 rows in the cleaned version of the dataset. The variables include id, host_id, neighborhood, latitude, longitude, room_type, price, minimum_nights, number_of_reviews, last_review, reviews_per_month, calculated_host_listings_count, availability_365, number_of_reviews_ltm and side. Our target variable "price" is measured in Turkish Liras (₺).

2 DATA COLLECTION & DESCRIPTION

2.1 Data Collection

The dataset used in this study is collected from InsideAirbnb, a secondary data source that provides detailed information about the listings of Airbnb across different cities in the world. It contains 22867 Airbnb listings, which is a large subset of all Airbnb listings in Istanbul. Therefore, this dataset can be considered as a sample dataset rather than the population because listings are continuously updated, and this dataset is lastly updated on September 29, 2025.

2.2 Summary Statistics

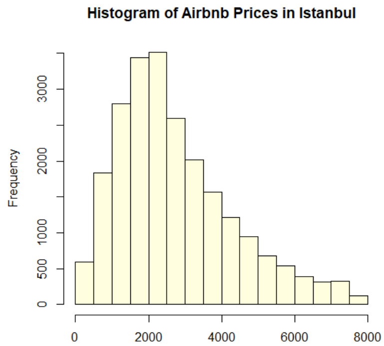
The summary statistics of the variable "price" is shown in **Table 2.1**.

Table 2.2.1 Summary Statistics

STATISTIC	VALUE
Mean	2752.049
Median	2412
Mode	2000
Variance	2531678
Standard Deviation	1591.125
Minimum	80
1st Quartile	1581
3rd Quartile	3602
Maximum	7787
Range	7707
Interquartile Range	2021
Coefficient of Variation	0.5782

2.3 Data Visualizations

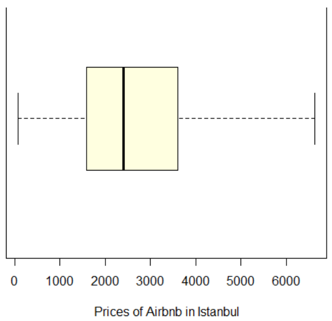
The histogram and boxplot are shown along with the necessary R code in **Figure 2.3.1**, **Figure 2.3.2**, **Figure 2.3.3** and **Figure 2.3.4**.




```
# Histogram
hist(df$price,
     col = "lightyellow",
     main = "Histogram of Airbnb Prices in Istanbul",
     xlab = "Prices of Airbnb in Istanbul")
```

Figure 2.3.3 R code for histogram

Figure 2.3.1 Histogram




```
# Boxplot
boxplot(df$price,
        col = "lightyellow",
        main = "Boxplot of Airbnb Prices in Istanbul",
        xlab = "Prices of Airbnb in Istanbul",
        outline = FALSE,
        horizontal = TRUE)
```

Figure 2.3.4 R code for boxplot

3 FREQUENCY DISTRIBUTION

3.1 Range, Width, Class Limits and Boundaries

The range of the prices are 7707₺. The class limits and class boundaries are given in the table below with the necessary R code.

Table 3.1.1.1 Class Boundaries

FREQUENCY	CLASS BOUNDARIES
288	[0,400)
1343	[400,800)
1738	[800,1200)
2522	[1200,1600)
2780	[1600,2000)
2701	[2000,2400)
2352	[2400,2800)
1886	[2800,3200)
1533	[3200,3600)
1208	[3600,4000)
976	[4000,4400)
808	[4400,4800)
664	[4800,5200)
485	[5200,5600)
431	[5600,6000)
342	[6000,6400)
251	[6400,6800)
262	[6800,7200)
208	[7200,7600)
87	[7600,8000)

```r
# class width and boundaries
width <- range_price / 20
print(width) # 385.35, round up to 400
width <- 400

breaks <- seq(0,8000,400)
class_boundaries <- cut(df$price,
                        breaks = breaks,
                        right = FALSE,
                        dig.lab = 5
                        )
frequencies <- hist(df$price,
                    plot = FALSE,
                    breaks = breaks)$counts

class_price <- data.frame(
  "Frequency" = frequencies,
  "Classes" = levels(class_boundaries)
)

class_price
# output example
#
#  Frequency     Classes
#1       288       [0,400)
#2      1343     [400,800)
#3      1738    [800,1200)
```

Figure 3.1.1 R Code for Class Boundaries

3.2 Frequency Distribution Graphs

The frequency polygon and ogive and plots are given in the figures below with the corresponding R code.

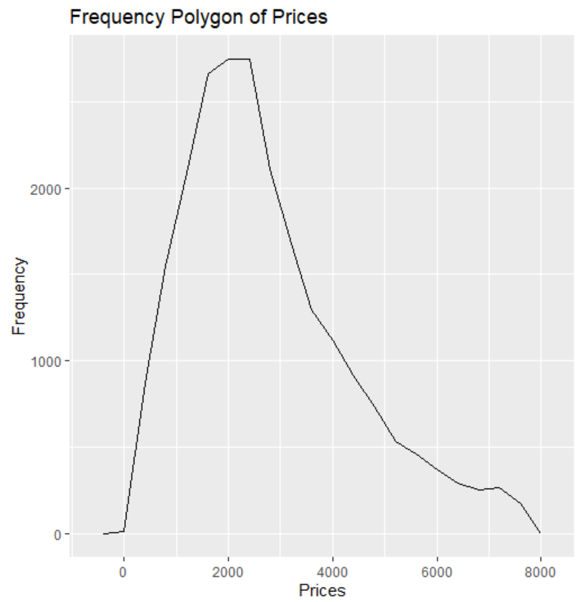


Figure 3.2.1 Frequency Polygon

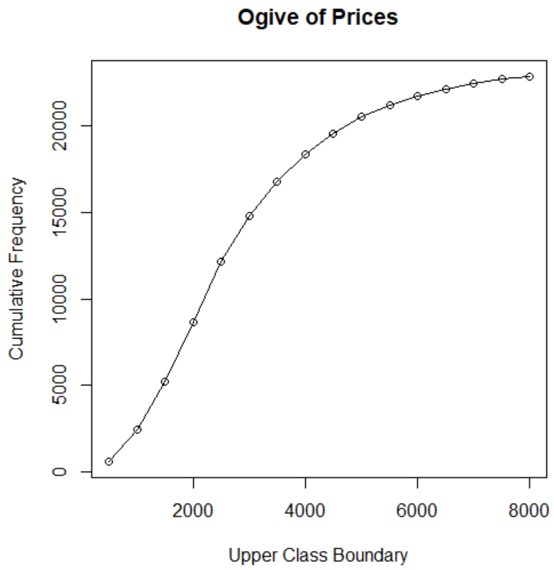


Figure 3.2.2 Ogive

```
# frequency polygon
ggplot(df, mapping = aes(x = price)) +
  geom_freqpoly(binwidth = 400) +
  labs(title = "Frequency Polygon of Prices",
       x = "Prices",
       y = "Frequency")

# o-give
h <- hist(df$price, plot = FALSE)

plot(
  h$breaks[-1],
  cumsum(h$counts),
  type = "o",
  xlab = "Upper Class Boundary",
  ylab = "Cumulative Frequency",
  main = "Ogive of Prices"
)
```

Figure 3.2.3 R Code

3.3 Shape of the Distribution

The visualizations of the data show that the distribuiton deviates from normality due to the presence of long right tail and we are able to conculde that the prices of the Airbnb houses in Istanbul are heavily right (positive) skewed. The quantitative skewness value is 0.9082 (R code is shown in **Figure 3.3.1**).

According to the histogram and the frequency polygon, the density of the prices are in the 1000₺-3000₺ range, the mode being 2000₺. The skewness shows us the presence of the high-priced luxury properties or outliers that extend the right tail of the distribuiton and the short left tail suggests that the Airbnb prices in Istanbul are mostly concentrated at lower to moderate prices rather than extremely low prices.

```
# skewness
library(e1071)
skewness(df$price)
# 0.9081838
```

Figure 3.3.1 R Code for Skewness

4 MEASURES OF DATA

4.1 Measures of Central Tendency

The calculated measures of central tendency of the "price" column is shown in the table below with the R code.

Table 4.1.1 Mesaure of Central Tendency

MEASURE	VALUE
Mean	2752.049
Median	2412
Mode	2000
Midrange	2021
Weighted Mean (w = number of reviews ltm)	3099.944
Skewness	0.9081
Kurtosis	0.4048

```
> # measures of central tendency
> mean(df$price)
[1] 2752.049
> median(df$price)
[1] 2412
> get_mode(df$price)
[1] 2000
> IQR <- q3 - q1
> IQR
[1] 2021
> weighted.mean(df$price, w = df$number_of_reviews_ltm)
[1] 3099.944
> library(e1071)
> skewness(df$price)
[1] 0.9081838
> kurtosis(df$price)
[1] 0.4048928
```

Figure 4.1.1 R Code for Central Tendencies

4.2 Measures of Spread

The calculated measures of variation of the "price" column is shown in the table below with the R code.

Table 4.2.1 Measures of Dispersion

MEASURE	VALUE
Range	7707
Variance	2531678
Standard Deviation	1591.125
Coefficient of Variation	57.816

```
> # measures of variation
> diff(range(df$price))
[1] 7707
> var(df$price)
[1] 2531678
> sd(df$price)
[1] 1591.125
> coef_of_var <- (sd(df$price)/mean(df$price) * 100)
> coef_of_var
[1] 57.81601
```

Figure 4.2.1 R Code for Measures of Variation

4.3 Measures of Position

Z-score of first 5 rows and quartiles are given in **Figure 4.3.1** and **Figure 4.3.2**.

```
> # first 5 z-score
> scale(df$price)[1:5]
[1] -0.2903912 -1.0376614  0.4260831  0.2677046
[5] -0.5606403
```

Figure 4.3.1 First 5 Z-scores of Prices

```
> # percentiles
> min(df$price) # q0
[1] 80
> as.numeric(quantile(df$price,0.25)) # q1
[1] 1581
> median(df$price) # q2
[1] 2412
> as.numeric(quantile(df$price,0.75)) # q3
[1] 3602
> max(df$price) # q4
[1] 7787
```

Figure 4.3.2 Quartiles

The boxplot with outliers and the R code are shown in figures below. Because there is a large number of rows in the price column, there are many values that are outliers. This is why in the boxplot, the outliers seem like a straight line.

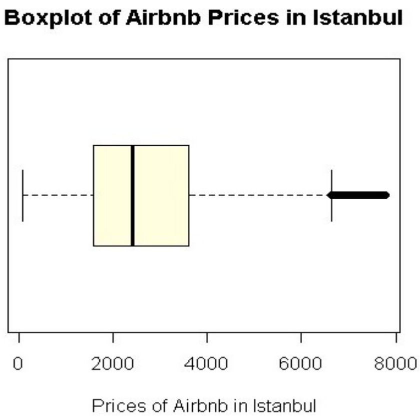


Figure 4.3.3 Boxplot with outliers

```
> # boxplot
> boxplot(
+   df$price,
+   main = "Boxplot of Airbnb Prices in Istanbul",
+   xlab = "Prices of Airbnb in Istanbul",
+   col = "lightyellow",
+   horizontal = TRUE
+ )
```

Figure 4.3.4 R Code for Boxplot

5 NORMAL DISTRIBUTION

5.1 Checking for Normality Using Plots

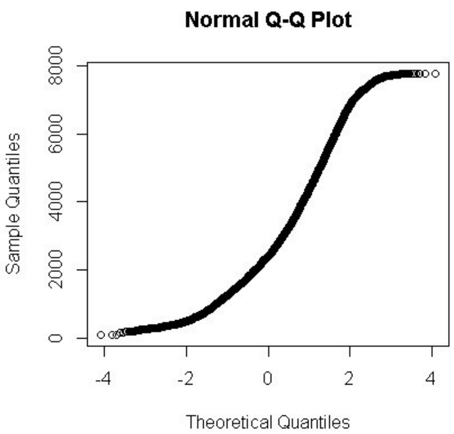


Figure 5.1.2 Q-Q Plot of Prices

```
> # Normal Probability Plot (Q-Q Plot)
> qqnorm(df$price,
+        main = "Normal Q-Q Plot")
```

Figure 5.1.1 R Code for Q-Q Plot

The Q-Q Plot shown in **Figure 5.1.1** suggest that the distribution has a right skew. We can conclude this by looking at the points at the left of the Q-Q Plot, that are non-linear and has a curvature.

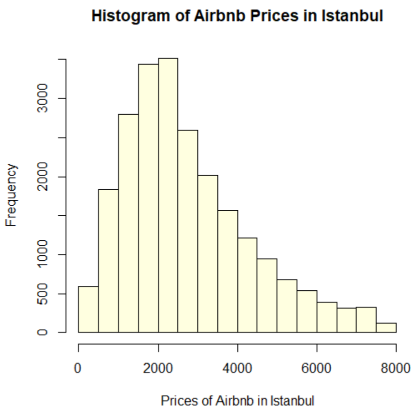


Figure 5.1.3 Histogram of Prices

```
# Histogram
hist(df$price,
     col = "lightyellow",
     main = "Histogram of Airbnb Prices in Istanbul",
     xlab = "Prices of Airbnb in Istanbul")
```

Figure 5.1.4 R Code for Histogram

Looking at the histogram shown in **Figure 5.1.3**, we can say that the distribution is right skewed, same as we inferred from the Q-Q Plot.

5.2 Shapiro-Wilk Test and Kolmogorov-Smirnov Tests

The conducted Shapiro-Wilk Test and Kolmogorov-Smirnov Test are shown in the figures below as an R code.

```
> set.seed(42)
> sample_price <- sample(df$price, 5000)
> # Shapiro-Wilk Test
> shapiro.test(sample_price)

        Shapiro-Wilk normality test

data:  sample_price
W = 0.93519, p-value < 2.2e-16
```

Figure 5.2.1 Shapiro-Wilk Test

```
> # Kolmogorov-Smirnov Test
> ks.test(jitter(df$price),
+         "pnorm",
+         mean(jitter(df$price)),
+         sd(jitter(df$price)))

        Asymptotic one-sample Kolmogorov-Smirnov
        test

data:  jitter(df$price)
D = 0.095736, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Figure 5.2.2 Kolmogorov-Smirnov Test

A sample of 5000 was taken from the data to conductthe Shapiro-Wilk test, because our data has more than 2000 values. The W value of 0.935 tells us that the distribution of the prices are not fairly normal, and the p-value being less than 0.001 confirms this situation.

The Kolmogorov-Smirnov D value of 0.096 tells us that the difference between the normal distribution and the distribution of the prices is at maximum %9.5. The p-value being smaller than 0.001 confirms that the distribution of the prices are not similar to the normal distribution.

5.3 Central Limit Theorem

Altough the distribution of the price column seem to be right skewed, sample size is sufficiently large (≈25000). With the Central Limit Theorem, the sampling distribution of the mean can be considered approximately normal. This allows us to use the parametric tests.

6 CONFIDENCE INTERVALS

6.1 Confidence Intervals for Mean and Proportion

The confidence intervals for the mean and the proportion is shown in the figures below.

```
> # Confidence Interval for the Mean
> std_err <- sd(df$price)/sqrt(length(df$price))
> lower <- mean_price - qt(0.975, df = length(df$price) - 1)*std_err
> upper <- mean_price + qt(0.975, df = length(df$price) - 1)*std_err
> cat(
+   "95% Confidence interval for the mean: [",lower,",",upper,"]"
+ )
95% Confidence interval for the mean: [ 2731.424 , 2772.673 ]
```

Figure 6.1.1 Confidence Interval for the Mean

```
> # Confidence Interval for the Proportion
> prop <- sum(df$price > mean(df$price)) / length(df$price)
> std_err_prop <- sqrt(prop*(1-prop)/length(df$price))
> z <- qnorm(0.975)
> lower_prop <- prop - z*std_err_prop
> upper_prop <- prop + z*std_err_prop
> cat(
+   "95% Confidence interval for the proportion: [",lower_prop,",",upper_prop,"]"
+ )
95% Confidence interval for the proportion: [ 0.4072182 , 0.4199849 ]
```

Figure 6.1.2 Confidence Interval for the Proportion

6.2 Margin of Error, t-Distribution, Degrees-of-Freedom, and Chi-Square Distribution

```
> # margin of error
> std_err <- sd(df$price)/sqrt(length(df$price))
> margin_of_err <- qt(0.975, df = length(df$price))*std_err
> margin_of_err
[1] 20.6248
```

Figure 6.2.1 Margin of Error

The margin of error calculated in **Figure 6.2.1** tells us that our estimate mean is expected to differ ±20.625₺ from the true population mean.

```
> # degrees-of-freedom
> dof <- length(df$price)-1
> dof
[1] 22864
> # t-value
> t <- qt(0.975, df = dof)
> t
[1] 1.960068
```

Figure 6.2.2 Degrees-of-Freedom & t-Value

With the given R code above in **Figure 6.2.2**, we can find the degrees-of-freedom as 22864 and the t-value as 1.960. The t-value indicates that to achieve a 95% confidence interval for the mean price, we have to be 1.960 standard errors away from the estimated sample mean, with 22864 degrees-of-freedom.

```
> # Confidence Interval for the Variance/standard_deviation
> variance_price <- var(df$price)
> chisq_upper <- qchisq(0.975, df = dof)
> chisq_lower <- qchisq(0.025, df = dof)
> lower_bound <- (dof * variance_price)/chisq_upper
> upper_bound <- (dof * variance_price)/chisq_lower
> cat(
+   " 95% Confidence Interval for variance: [",lower_bound,",",upper_bound,"]\n",
+   "95% Confidence Interval for standard deviation: [",sqrt(lower_bound),",",sqrt(upper_bound),"]"
+ )
 95% Confidence Interval for variance: [ 2485903 , 2578735 ]
 95% Confidence Interval for standard deviation: [ 1576.675 , 1605.844 ]
```

Figure 6.2.3 Confidence Interval using Chi-Square Distribution

The %95 confidence interval for the variance and the standard deviation is shown in **Figure 6.2.3**. These values are calculated using the chi-square distribution with degrees-of-freedom 22864.

7 HYPOTHESIS TESTING

7.1 One-Sample t-Test

To determine whether the true mean of the Airbnb prices in Istanbul is equal to 2500₺, a one-sample t-test was conducted. The null and alternative hypotheses are:

H_0 (The null hypothesis) : $\mu = 2500$₺

H_1 (The alternative hypothesis) : $\mu \neq 2500$₺

The conducted one-sample t-test is shown in **Figure 7.1.1** with the R code.

```
> # 7.1 One-Sample t-Test
> OneSample_t <- t.test(airbnb$price, mu = 2500)
> OneSample_t

	One Sample t-test

data:  airbnb$price
t = 23.953, df = 22864, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 2500
95 percent confidence interval:
 2731.424 2772.673
sample estimates:
mean of x
 2752.049

>
> OneSample_t$statistic  # Traditional Method  t value
       t
23.95332
> OneSample_t$p.value    # P-value Method
[1] 2.978868e-125
> OneSample_t$conf.int   # Confidence Interval Method
[1] 2731.424 2772.673
attr(,"conf.level")
[1] 0.95
```

Figure 7.1.1 One-Sample t-Test

The question here is whether the true average Airbnb listing price in Istanbul sits at 2500 ₺. With a sample of 22,865 listings and a sample mean of 2752.05 ₺, the gap from 2500 already looks noticeable — the t-test was used to check whether this gap could simply be due to sampling variation.

The test statistic was $t = 23.95$ with 22,864 degrees of freedom. The critical value at $\alpha = 0.05$ (two-tailed) is 1.96, and 23.95 far exceeds it, so H_0 is rejected. The p-value is essentially zero ($p < 2.2e-16$). The 95% confidence interval for the true mean was [2731.42, 2772.67] ₺ — since 2500 does not fall within this range, the same conclusion holds. The evidence across all three approaches is consistent: the average Istanbul Airbnb price is statistically higher than 2500 ₺, and the difference is not a product of random chance.

7.2 Independent Samples t-Test

Istanbul is divided geographically into two sides, and it is reasonable to ask whether this division also shows up in Airbnb pricing. The European side contains most of the city's major tourist areas, historic districts, and commercial centers, which could push prices higher. To test this, a Welch independent samples t-test was applied, as the two groups had inequal sample sizes. The null and alternative hypotheses for the independent samples t-test are:

H_0: $\mu_{European} = \mu_{Anatolian}$
H_1: $\mu_{European} \neq \mu_{Anatolian}$
The applied independent samples t-test is shown in **Figure 7.2.1** along with the R code.

```
        Welch Two Sample t-test

data:  european and anatolian
t = 7.6534, df = 6483.7, p-value = 2.242e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 153.3593 258.9736
sample estimates:
 mean of European mean of Anatolian
         2791.154          2584.988

>
> IndependentSample_t$statistic  # Traditional Method  t value
       t
7.653401
> IndependentSample_t$p.value    # P-value Method
[1] 2.242396e-14
> IndependentSample_t$conf.int   # Confidence Interval Method
[1] 153.3593 258.9736
attr(,"conf.level")
[1] 0.95
```

Figure 7.2.1 Two Sample t-Test

The European side average is 2791.15₺ across 18,528 listings; the Anatolian side average is 2584.99₺ across 4,337 listings. The test gave t = 7.65 with approximately 6,483 degrees of freedom. Since this exceeds the critical value of 1.96 at the 5% level, H_0 is rejected. The p-value of 2.242e-14 is far below 0.05. The 95% confidence interval for the mean difference was [153.36, 258.97] ₺, which lies entirely above zero. The price gap of roughly 206₺ between the two sides is statistically significant. Geographic location within Istanbul appears to have a real effect on how listings are priced.

7.3 Dependent Samples t-Test

Each listing in the dataset carries two review-related variables: total number of reviews ever received (number_of_reviews), and number of reviews received in the last twelve months (number_of_reviews_ltm). Since these come from the same listings, a paired t-test is the appropriate approach for comparing them. The null and alternative hypotheses are as follows:

H_0: $\mu_{total\text{-}reviews} = \mu_{reviews\text{-}ltm}$

H_1: $\mu_{total\text{-}reviews} \neq \mu_{reviews\text{-}ltm}$

The conducted paired t-test and the R code is shown in the **Figure 7.3.1**.

```
> # 7.3 Paired t-Test
> paired_t <- t.test(airbnb$number_of_reviews, airbnb$number_of_reviews_ltm, paired = TRUE)
> paired_t

        Paired t-test

data:  airbnb$number_of_reviews and airbnb$number_of_reviews_ltm
t = 55.284, df = 22864, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 13.30744 14.28574
sample estimates:
mean difference
       13.79659

>
> paired_t$statistic  # Traditional Method  t value
       t
55.28436
> paired_t$p.value    # P-value Method
[1] 0
> paired_t$conf.int   # Confidence Interval Method
[1] 13.30744 14.28574
attr(,"conf.level")
[1] 0.95
```

Figure 7.3.1 Paired t-Test

The test yielded t = 55.28 at 22864 degrees of freedom, $p < 0.001$. The average difference was 13.80 reviews, with a 95% confidence interval of [13.31, 14.29]. Because the interval does not include zero, H_0 is rejected. Older listings have had more time to collect reviews overall, so it is expected that total counts would run higher than the last-year figures and the data confirms this gap is statistically significant.

7.4 One-Way ANOVA

Room type is one of the more obvious factors that might influence how much a host charges. To test whether average prices actually differ across the four room types in the dataset. Entire home/apt, Private room, Shared room, and Hotel room a one-way ANOVA was used. The hypotheses are:
H_0: $\mu_{Entire} = \mu_{Private} = \mu_{Shared} = \mu_{Hotel}$
H_1: At least one group mean differs from the others.
The conducted ANOVA is shown along with the R code in **Figure 7.4.1**.

```
> # 7.3 ANOVA Result
> anova_model <- aov(price ~ room_type, data = airbnb)
> summary(anova_model)
               Df    Sum Sq   Mean Sq F value Pr(>F)
room_type       3 6.117e+09 2.039e+09   900.4 <2e-16 ***
Residuals   22861 5.177e+10 2.264e+06
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 7.4.1 ANOVA

The result was F = 900.4 with 3 between-groups degrees-of-freedom and 22861 within-groups degrees-of-freedom, $p < 0.001$, which leads to rejection of H_0. Since ANOVA does not identify which specific pairs differ, a Tukey HSD (Honestly Significant Difference) post-hoc test was run.

```
> tukey_result <- TukeyHSD(anova_model)
> tukey_result
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = price ~ room_type, data = airbnb)

$room_type
                                 diff        lwr         upr     p adj
Hotel room-Entire home/apt    -666.15157 -1070.3803  -261.92280 0.0001351
Private room-Entire home/apt -1137.07731 -1193.7191 -1080.43551 0.0000000
Shared room-Entire home/apt  -1112.56077 -1445.4822  -779.63931 0.0000000
Private room-Hotel room       -470.92574  -876.8195   -65.03201 0.0152560
Shared room-Hotel room        -446.40921  -968.3078    75.48937 0.1238778
Shared room-Private room        24.51654  -310.4245   359.45761 0.9976425
```

Figure 7.4.2 Tukey HSD

Entire home/apt listings were significantly more expensive than every other type: roughly 666₺ above Hotel rooms ($p < 0.001$), 1137₺ above Private rooms ($p < 0.001$), and 1113₺ above Shared rooms ($p < 0.001$). Hotel rooms were also significantly pricier than Private rooms by about 471₺ ($p = 0.015$). No significant difference appeared between Shared rooms and Hotel rooms ($p = 0.124$), or between Shared rooms and Private rooms ($p = 0.998$) these two categories occupy a similar price range in the Istanbul market.

7.5 Proportion z-Test

Entire home/apt listings tend to dominate short-term rental markets in major cities. To check whether 70% is a reasonable description of the Istanbul Airbnb market, a one-proportion z-test was carried out. The hypotheses are:
H_0: $p = 0.70$
H_1: $p \neq 0.70$

The z-Test with the R code is given in **Figure 7.5.1**.

```
> # 7.4 Proportion z-Test
> entire_home_count <- sum(airbnb$room_type == "Entire home/apt")
> proportion_t <- prop.test(entire_home_count, n, p = 0.70)
> proportion_t

        1-sample proportions test with continuity correction

data:  entire_home_count out of n, null probability 0.7
X-squared = 1.0498, df = 1, p-value = 0.3055
alternative hypothesis: true p is not equal to 0.7
95 percent confidence interval:
 0.6971494 0.7090362
sample estimates:
        p
0.7031271
```

Figure 7.5.1 Proportion z-Test

Out of 22,865 listings, 16,077 were classified as Entire home/apt, giving an observed proportion of 70.31%. The chi-square statistic was $\chi^2 = 1.0498$ with degrees-of-freedom 1 and p = 0.3055, which is above the 0.05 cutoff, so H_0 is not rejected. The 95% confidence interval for the true proportion was [0.697, 0.709], and 0.70 falls within this range. The observed proportion is close enough to the hypothesized value that the difference cannot be considered statistically meaningful. The data is consistent with the claim that roughly 70% of Istanbul listings are entire homes.

8 TESTING THE DIFFERENCE BETWEEN TWO MEASURES

8.1 Independent Samples t-Test

Comparing Airbnb prices between the European and Anatolian sides of Istanbul reveals a noticeable gap. To determine whether this difference holds statistically, an independent samples t-test was applied using Welch's correction for inequal variances.

```
> IndependentSample_t <- t.test(european, anatolian)
> names(IndependentSample_t$estimate) <- c("mean of European", "mean of Anatolian")
> IndependentSample_t

        Welch Two Sample t-test

data:  european and anatolian
t = 7.6534, df = 6483.7, p-value = 2.242e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 153.3593 258.9736
sample estimates:
 mean of European mean of Anatolian
         2791.154          2584.988
```

Figure 8.1.1 Independent Samples t-Test

The European side (n = 18,528) had a mean price of 2791.15₺ ; the Anatolian side (n = 4,337) averaged 2584.99₺. The test yielded t = 7.65 at approximately 6484 degrees-of-freedom, p = 2.242e-14. The 95% confidence interval for the difference in means was [153.36, 258.97] ₺ entirely above zero, with no overlap at the lower end. The price difference is statistically significant at the 5% level. One likely factor is that the European side concentrates most of Istanbul's tourist destinations, historic sites, and major business districts, making it easier for hosts there to charge higher nightly rates.

8.2 Dependent Samples t-Test

The paired comparison for this section draws on the analysis already carried out in **Section 7.3**. The same listings were compared across two review counts: total reviews and last-twelve-month reviews. The result was t = 55.28 with degrees-of-freedom 22864, $p < 0.001$, with a mean difference of 13.80 and a 95% confidence interval of [13.31, 14.29]. Since zero is not in the interval, H_0 is rejected. The difference between the two measures is statistically significant.

```
> paired_t <- t.test(airbnb$number_of_reviews, airbnb$number_of_reviews_ltm, paired = TRUE)
> paired_t

	Paired t-test

data:  airbnb$number_of_reviews and airbnb$number_of_reviews_ltm
t = 55.284, df = 22864, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 13.30744 14.28574
sample estimates:
mean difference 
       13.79659 
```

Figure 8.2.1 Paired t-Test

8.3 Two Proportions Test

Beyond prices, it is worth asking whether the mix of room types also differs between the two sides of the city. Specifically, whether the share of Entire home/apt listings is different on the European and Anatolian sides.

```
> european_data  <- airbnb[airbnb$side == "European", ]
> anatolian_data <- airbnb[airbnb$side == "Anatolian", ]
>
> europe_entire <- sum(european_data$room_type == "Entire home/apt")
> anatolia_entire <- sum(anatolian_data$room_type == "Entire home/apt")
>
> prop.test(c(europe_entire, anatolia_entire), c(nrow(european_data), nrow(anatolian_data)))

	2-sample test for equality of proportions with continuity correction

data:  c(europe_entire, anatolia_entire) out of c(nrow(european_data), nrow(anatolian_data))
X-squared = 193.71, df = 1, p-value < 2.2e-16
alternative hypothesis: two.sided
95 percent confidence interval:
 0.09142107 0.12338992
sample estimates:
   prop 1    prop 2 
0.7234996 0.6160941 
```

Figure 8.3.1 Two Proportions Test

On the European side, 72.3% of listings fell into this category; on the Anatolian side, 61.6%. A two-proportions test gave $\chi^2 = 193.71$ with 1 degrees-of-freedom, $p < 0.001$, and the 95% confidence interval for the difference in proportions was [0.091, 0.123] which does not include zero. H_0 is rejected. Entire home/apt listings make up a notably larger share of the market on the European side, which may reflect differences in housing stock and property types between the two parts of the city.

8.4 Two Variances F-Test

Even though Section 8.1 found that average prices differ between the two sides, that does not necessarily mean the spread of prices is also different. An F-test was used to directly compare the two variances.

```
> variance_result <- var.test(european, anatolian)
> variance_result

	F test to compare two variances

data:  european and anatolian
F = 0.98429, num df = 18527, denom df = 4336, p-value = 0.503
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.9389946 1.0310584
sample estimates:
ratio of variances 
         0.9842889 
```

Figure 8.4.1 Two Variances F-Test

The result was F(18527, 4336) = 0.984, p = 0.503. The 95% confidence interval for the variance ratio was [0.939, 1.031], which contains 1. Since $p > 0.05$ and the interval includes 1, H_0 is not rejected. Despite the difference in mean prices, the variability in listing prices is not statistically different between the European and Anatolian sides — both groups show a similar spread around their respective averages.

9 CORRELATION AND REGRESSION

9.1 Correlation Analysis

9.1.1 Pearson Correlation

To explore whether there is any relationship between listing prices and how frequently a property receives reviews, a Pearson correlation was computed. The result was r = 0.103, which points to a weak positive association between the two variables. The correlation was statistically significant: t = 15.655 at 22863 degrees-of-freedom, $p < 0.001$, with a 95% confidence interval of [0.090, 0.116]. While the result is significant, the low r value makes clear that the two variables do not move closely together and reviews per month explains very little of the variation in price.

```
> # 9.1.1 Pearson Correlation
> correlation_result <- cor.test(airbnb$price, airbnb$reviews_per_month, method = "pearson")
> correlation_result

	Pearson's product-moment correlation

data:  airbnb$price and airbnb$reviews_per_month
t = 15.655, df = 22863, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.09014321 0.11579200
sample estimates:
      cor 
0.1029847 
```

Figure 9.1.1.1 Pearson Correlation

9.1.2 Hypothesis Testing On Correlation

The hypotheses are:
H_0: r = 0 (There is no correlation between the price and the reviews_per_month)
H_1: $r \neq 0$ (There is correlation between the price and the reviews_per_month)
The Critical Value is calculated in **Figure 9.1.2.1** and used to test the null hypothesis.

```
> # Extract specific parts
> correlation_result$estimate    # 1. Correlation coefficient (r)
      cor 
0.1029847 
```

Figure 9.1.2.1 Critical Value

The t-statistic of 15.655 with 22,863 degrees-of-freedom calculated in **Section 9.1.1** exceeds the critical value of 1.96 at $\alpha = 0.05$. H_0 is rejected. A linear relationship between price and reviews per month is statistically present.

The t-Value is calculated in **Figure 9.1.2.2** and used to test the null hypothesis.

```
> correlation_result$statistic   # 2. Test Statistic (t)
       t
15.65506
```

Figure 9.1.2.2 t-Value

The same t-statistic of 15.655 is used here directly. Since it clears the critical value threshold, the correlation is statistically significant at the 5% level.

The p-Value is calculated in **Figure 9.1.2.3** and used to test the null hypothesis.

```
> correlation_result$p.value     # 3. p-value
[1] 5.921261e-55
```

Figure 9.1.2.3 p-Value

The p-value of 5.92e-55 is far below 0.05. H_0 is rejected on this basis as well.

The confidence interval is calculated in **Figure 9.1.2.4** and used to test the null hypothesis.

```
> correlation_result$conf.int    # 4. Confidence Interval
[1] 0.09014321 0.11579200
attr(,"conf.level")
[1] 0.95
```

Figure 9.1.2.4 Confidence Interval

The 95% confidence interval for the correlation is [0.090, 0.116]. Zero is not in this range, so H_0 is rejected. The interval also shows that the true correlation, while real, is small somewhere between 0.09 and 0.12.

9.1.3 Conclusion

All three methods lead to the same decision: H_0 is rejected. The correlation between listing price and reviews per month is statistically significant ($r = 0.103$, $p < 0.001$), but weak. In practical terms, reviews per month account for a very small share of what drives price differences across Istanbul listings.

9.2 Regression Analysis

9.2.1 Regression Line Construction and Interpretation

A simple linear regression was fitted to model listing price as a function of reviews per month. The estimated regression equation is given below (1) with the necessary R code. $\hat{y}$ is the estimated value of price.

$$\hat{y} = 2648.419 + 127.952 \times \text{reviews\_per\_month} \quad (1)$$

```
> regression_model <- lm(price ~ reviews_per_month, data = airbnb)
> summary(regression_model)

Call:
lm(formula = price ~ reviews_per_month, data = airbnb)

Residuals:
    Min      1Q  Median      3Q     Max 
-3367.6 -1173.3  -336.9   851.6  5138.6 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)       2648.419     12.384  213.85   <2e-16 ***
reviews_per_month  127.952      8.173   15.65   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1583 on 22863 degrees of freedom
Multiple R-squared:  0.01061,	Adjusted R-squared:  0.01056 
F-statistic: 245.1 on 1 and 22863 DF,  p-value: < 2.2e-16
```

Figure 9.2.1.1 R Code for the Regression Model

The intercept of 2648.42₺ gives the baseline estimated price for a listing with no reviews per month. The slope of 127.95 means that each additional review per month is associated with roughly a 128₺ increase in estimated price. The slope was statistically significant at $t = 15.65$, $p < 0.001$ but as the R^2 figure will show, statistical significance here does not translate into a useful predictive model.

9.2.2 Scatter Plot and Regression Line

The scatter plot with the regression line and the R code is shown in **Figure 9.2.2.1** and **Figure 9.2.2.2**.

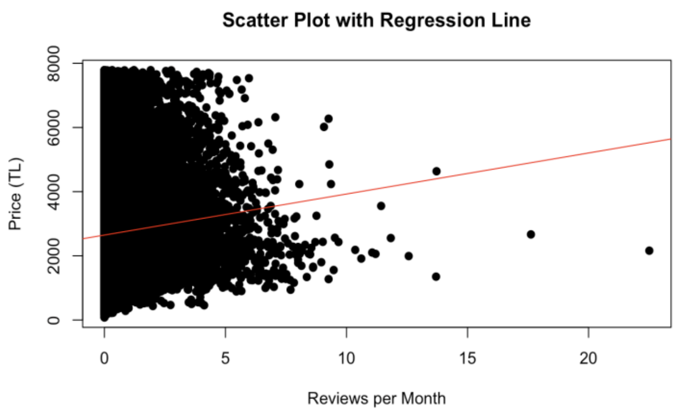


Figure 9.2.2.1 Scatter Plot with Regression Line

```
> plot(airbnb$reviews_per_month, airbnb$price,
+      main = "Scatter Plot with Regression Line",
+      xlab = "Reviews per Month",
+      ylab = "Price (TL)",
+      pch = 19)
> abline(regression_model, col = "red")
```

Figure 9.2.2.2 R Code for the Scatter Plot

The scatter plot shows that most listings cluster at low review counts, with a wide range of prices at those levels. The regression line has a slight upward slope, reflecting the positive relationship found in the correlation analysis ($r = 0.103$), but the scatter around the line is wide. The plot makes it visually clear that knowing a listing's review rate does not give much information about its price.

9.2.3 Coefficient of Determination (R^2), Adjusted R^2 and Standard Error of Estimate

The R^2 of the model is 0.0106 reviews per month accounts for roughly 1.06% of the variation in listing prices. The adjusted R^2 is nearly identical at 0.01056, which is expected with only one predictor. These figures confirm that the model has very limited explanatory power.

The standard error of estimate is 1583₺. On average, the model's price predictions are off by about 1583₺. Although the relationship is statistically significant, the prediction error remains high factors other than review activity likely play a much larger role in determining what hosts charge in Istanbul. The R code for these values is given below.

```
> summary(regression_model)$r.squared # R²
[1] 0.01060585
> summary(regression_model)$adj.r.squared # Adjusted R²
[1] 0.01056258
> summary(regression_model)$sigma #Standart error of estimate
[1] 1582.699
```

Figure 9.2.3.1 R Code for the R^2 and Standard Error of Estimate

9.2.4 Residual Plot

The residual plot with the corresponding R code is given in **Figure 9.2.4.1** and **Figure 9.2.4.2**.

Figure 9.2.4.1 Residual Plot

```
> plot(regression_model$fitted.values, regression_model$residuals,
+      main = "Residual Plot of Regression Model",
+      xlab = "Fitted Values",
+      ylab = "Residuals",
+      pch = 16)
> abline(h = 0, col = "red") # add horizontal reference line at 0
```

Figure 9.2.4.2 R Code for the Residual Plot

The residual plot shows that errors are not evenly spread around zero. Most data points fall on the left side of the plot, and the spread of residuals narrows as fitted values increase. This uneven pattern. Suggests the model does not perform equally well across all price levels. For lower-priced listings, the errors are much larger relative to the fit. This is not surprising given the R^2 of 1.06%: a model that explains almost none of the variation will naturally leave large, irregular residuals behind.

9.2.5 Prediction Interval

The prediction interval is given in the **Figure 9.2.5.1**.

```
> predict(regression_model, newdata = data.frame(reviews_per_month = 3), interval = "prediction")
       fit       lwr      upr
1 3032.277 -70.18691 6134.741
```

Figure 9.2.5.1 Prediction Interval

Using the regression model, a listing that receives 3 reviews per month has an estimated price of approximately 3032₺. The 95% prediction interval for this estimate runs from -70.19 to 6134.74₺. The negative lower bound is not meaningful in practice a price cannot be negative so it would be treated as 0.

The interval spans over 6000₺, which shows how imprecise individual predictions are when the model explains only about 1% of price variation. The point estimate of 3032₺ gives a rough figure, but the actual price for any given listing could fall almost anywhere within that range.

10 OTHER CHI-SQUARE TESTS

10.1 Goodness-of-Fit Test

A chi-square goodness-of-fit test was conducted to test the following hypotheses:

H_0: The listings are uniformly distributed among different neighborhoods.

H_1: The listings are not uniformly distributed among different neighborhoods.
The result is given in the **Figure 10.1.1**.

```
> # GOODNES-OF-FIT TEST
> observed_frequencies <- table(df$neighbourhood)
> expected_proportion <- 1/length(observed_frequencies)
> gof_test <- chisq.test(observed_frequencies,
+                        p = rep(expected_proportion, length(observed_frequencies)))
> gof_test

        Chi-squared test for given probabilities

data:  observed_frequencies
X-squared = 61571, df = 27, p-value < 2.2e-16
```

Figure 10.1.1 Chi-Square Goodness-of-Fit Test with the Necessary R Code

The analysis resulted in a χ^2 value of 61571 with 27 degrees-of-freedom. The p-value being close to 0 means that the p-value is statistically significant enough to reject the null hypothesis. From this test we can conclude that the observed frequencies of the listings significantly differ from an uniform distribution among different neighborhoods.

10.2 Independence Test & Homogeneity Test

An independence test was applied to determine if the room_type and neighborhood variables depend on each other or not.
H_0: The room_type and neighborhood variables depend on each other.
H_1: The room_type and neighborhood variables do not depend on each other.
The conducted chi-square test is given in **Figure 10.2.1**.

```
> # INDEPENDENCE TEST
> contingency <- table(df$neighbourhood, df$room_type)
> independence_test <- chisq.test(contingency,
+                                 simulate.p.value = TRUE)
> independence_test

        Pearson's Chi-squared test with simulated p-value (based on 2000 replicates)

data:  contingency
X-squared = 1633.4, df = NA, p-value = 0.0004998
```

Figure 10.2.1 Chi-Square Independence & Homogeneity Test with the Necessary R Code

The contingency table of neighborhoods and room types has many cell values lower than expected, this means that the R function “chisq.test()” may not give the correct approximation of the χ^2 value. In order to obtain the correct approximation, 2000 (by default) Monte Carlo simulated p-values were used. The only limitation of this method are the values may differ in a minimal scale. This method also does not give the degree-freedom, but we can calculate it as:

$$(row\ count - 1) \times (column\ count - 1)$$

With the formula (2), we can calculate the degree-of-freedom as (28 - 1)×(4 - 1) = 81.

The χ^2 value is approximately 1633.4 with 81 degree-of-freedom. The approximated p-value being smaller than 0.001, we are able to reject the null hypothesis, thus concluding that there is no significant dependence between listed room types and neighborhood.
To determine the homogeneity between thesevariables, we can apply the same test. The hypotheses for the homogeneity test are:

H_0: The room types are distributed uniformly across different neighborhoods.
H_1: The distribution of room types are not uniform across different neighborhoods.

From the statistics calculated in Figure 10.2.1 the same conclusion as the independence test applies for the homogeneity test. The p-value being smaller than 0.001 means that there are statistically significant evidence for us to reject the null hypothesis. Thus we can say that the room types differ significantly across different neighborhoods.

10.3 Shapiro-Wilk Normality Test

The Shapiro-Wilk test is used to determine if the Airbnb listing prices follow a normal distribution or not. The hypotheses are as follows:
H_0: The distribution of the prices follow a normal distribution.
H_1: The distribution of the prices does not follow a normal distribution.
The Shapiro-Wilk test is given in **Figure 10.3.1** along with the R code. Because of the dataset consisting more than 5000 listings, we have to take a sample in order to apply the Shapiro-Wilk test for normality.

```
> # SHAPIRO-WILK TEST
> set.seed(2026)
> price_sample <- sample(df$price, size = 5000)
> shapiro.test(price_sample)

        Shapiro-Wilk normality test

data:  price_sample
W = 0.93796, p-value < 2.2e-16
```

Figure 10.3.1 Shapiro-Wilk Test for Normality

The p-value being significantly small, we are able to reject the null hypothesis. This suggests that the distribution of the Airbnb prices are not normally distributed.

With the W value of 0.94, we can draw the conclusion that the distribution is not normal, but also it is not far from being a normal distribution.

11 ANALYSIS OF VARIANCE (ANOVA)

11.1 One-Way ANOVA

The One-Way ANOVA test was used to determine if the price means of different continent sides differ from each other.
H_0: The group means are equal to each other.
H_1: The group means are different from each other.
The conducted test is shown in **Figure 11.1.1**.

```
> # ONE-WAY ANOVA
> anova_price <- aov(data = df, price ~ side)
> summary(anova_price)
               Df    Sum Sq   Mean Sq F value   Pr(>F)
side            1 1.494e+08 149376637   59.15 1.52e-14
Residuals   22863 5.773e+10   2525255

side         ***
Residuals
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 11.1.1 One-Way ANOVA with the R Code

The calculated F-value of 59.15 with 1 degree-of-freedom suggests that there is a significant difference between the means of the prices of each continent side. Thus, we reject the null hypothesis.

The between-groups sum of squares and within-groups sum of squares are calculated as 1.494×10^8 and 5.773×10^{10}, respectively. These values show that the variation within group means contribute more to the total variation than the variation between group means.

11.2 Tukey HSD Test

To obtain the comparison between the two group means, $\mu_{Anatolia}$ and μ_{Europe}, a Tukey HSD test was performed on the ANOVA model used in **Section 11.1**. The conducted test is shown below.

```
> # TUKEY HSD
> hsd <- TukeyHSD(anova_price)
> hsd
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = price ~ side, data = df)

$side
                    diff      lwr      upr p adj
Europe-Anatolia 206.1664 153.6253 258.7076     0
```

Figure 11.2.1 Tukey HSD Test

The result shows that the difference between the group means is approximately 206.17 with [153.63, 258.71] confidence interval and p-value is practically 0. Since this interval does not include 0, the difference is statistically significant.

11.3 Two-Way ANOVA

A Two-Way ANOVA was applied to evaluate the individual effects and interactions of room types and continent side on Airbnb prices. The conducted analysis is given in **Figure 11.3.1**.

```
> # TWO-WAY ANOVA
> tw_anova <- aov(data = df, price ~ room_type*side)
> summary(tw_anova)
                  Df    Sum Sq  Mean Sq F value   Pr(>F)    
room_type          3 6.117e+09 2.039e+09  902.37  < 2e-16 ***
side               1 2.373e+07 2.373e+07   10.50  0.00119 ** 
room_type:side     2 9.560e+07 4.780e+07   21.16 6.62e-10 ***
Residuals      22858 5.165e+10 2.260e+06                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 11.3.1 Two-Way ANOVA

The summary statistics of room_type show that the F-value is 902.37 with 3 degree-of-freedom, the p-value is smaller than 0.001 and these values suggest that there is a significant difference between the mean prices of room types.

The effect of the continent side has p-value smaller than 0.01 with the F-value being equal to 10.5 on 1 degree-of-freedom, meaning that the mean prices of the Europe and Anatolia sides also differ significantly from each other, although being less significant than room_type.

The groupwise interaction of room_type and side results in an F-value of 21.16 with 2 degrees-of-freedom and a p-value smaller than 0.001. These statistics show that the interaction between the room_type and side variables are statistically meaningful. In other words, the prices of the room types are not consistent across two sides.

12 NON-PARAMETRIC STATISTICS

The non-parametric statistics include the Sign Test and the Wilcoxon Signed Rank Test. Both of these tests require us to use paired before-and-after observations. Since our dataset does not contain any before-and-after variables, we are not able to perform these tests.

Another non-parametric test is the Wald-Wolfowitz Runs Test, that is used to assess whether the sequence of observations are randomly ordered or not. The dataset used in this study does not contain any time-ordered observation and the implementation of this test on our dataset would not provide meaningful results.

12.1 Wilcoxon Rank Sum Test (Mann-Whitney)

The Wilcoxon Rank Sum Test is used to determine if there is a significant difference between the distributions of the prices among different continent sides. The hypotheses are:
H_0: The distributions of price in Europe and Anatolia sides are equal. (Location shift = 0)
H_1: The distributions of price in Europe and Anatolia sides are different. (Location shift ≠ 0)
The performed test is shown in **Figure 12.1.1** along with the R code.

```
> # WILCOXON RANK SUM TEST
> wilcox.test(data = df, price ~ side)

        Wilcoxon rank sum test with continuity correction

data:  price by side
W = 36928935, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

Figure 12.1.1 Wilcoxon Rank Sum Test

The test gives us a W-value of 36928935, and a p-value smaller than 0.001, which indicates that the distribution of Airbnb prices differs significantly between the European and Anatolian sides of Istanbul. Therefore, we are able to reject the null hypothesis.

12.2 Kruskal-Wallis Test (H-Test)

The Kruskal-Wallis test is applied in order to determine if there is a significant difference between the distribution of prices across different room types. The difference between this test and one-way ANOVA is that this test is used in the absence of normality, which is the case for our data.

The hypotheses are:
H_0: The distribution of prices is the same across all room types.
H_1: The distribution of prices differs for at least one room type.
The Kruskal-Wallis test is shown in the figure below.

```
> # KRUSKAL-WALLIS TEST
> kruskal.test(data = df, price ~ room_type)

        Kruskal-Wallis rank sum test

data:  price by room_type
Kruskal-Wallis chi-squared = 3163, df = 3, p-value < 2.2e-16
```

Figure 12.2.1 Kruskal-Wallis Test

The χ^2 value is 3163 with 3 degree-of-freedom. The p-value is much lower than 0.001, allowing us to reject the null hypothesis. Therefore, there are statistically significant differences between the distribution of prices across the room types.

12.3 Spearman Rank Correlation Coefficient

To test the monotonic relationship between the total number of reviews and price, a Spearman Rank Correlation Test was carried out. The null and alternative hypotheses are given as:
H_0: There is no monotonic relationship between number of reviews and price. ($\rho = 0$)
H_1: There is a monotonic relationship between number of reviews and price. ($\rho \neq 0$)
The test is give with the R code in **Figure 12.3.1**.

```
> # SPEARMAN RANK CORRELATION TEST
> cor.test(df$price, df$number_of_reviews, method = "spearman")

        Spearman's rank correlation rho

data:  df$price and df$number_of_reviews
S = 1.6377e+12, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1780076

Warning message:
In cor.test.default(df$price, df$number_of_reviews, method = "spearman") :
  Cannot compute exact p-value with ties
```

Figure 12.3.1 Spearman Rank Correlation Test

The test yielded a p-value that is lower than 0.001, and $\rho = 0.178$. Therefore, we can make the inference that the price tends to increase as the number of reviews increases, rejecting the null hypothesis. Although from the ρ value we are able to tell that the relationship between these variables are very weak.

While performing the test, R console has given a warning message indicating that the exact p-value couldn't be computed due to the large amount of repeating values. This is expected as many of the Airbnb listings in this dataset have 0 reviews.

13 RESULTS AND INTERPRETATION

13.1 Summary

To sum up our findings, we can say that the mean value of the Airbnb prices in Istanbul is likely between the [2731, 2772] confidence interval. This value alsoholds for our hypothesis tests for the mean, which we have found that it is not equal to 2500 and larger than 2500. Although there are many listing prices that are clustered around this interval, the variation is also very high with 2531678 and a standard deviation of 1591.125. The very high Airbnbs in Istanbul enables the presence of many outliers.

We have also tested if the distribution of the prices are normal or not. From the histogram and the QQ-plot of the prices, we could see that the distribution of the Airbnb prices are right skewed. To find statistically meaningful values, we have used Shapiro-Wilk and Kolmogorov-Smirnov tests in order to make a conclusion. With these tests, we found statistical evidences that tells us the distribution of the Airbnb prices in Istanbul is far from normal and right skewed.

In order to find the relationship between certain variables and the Airbnb prices, we have conducted correlation and regression analysis. We found that there is a statistically meaningful correlation between reviews per month and price features, but this correlation is weak.

Based on the constructed regression line between reviews_per_month and price, we are able to state that the reviews_per_month variable almost explains none of the variation in the Airbnb prices, and this is supported with the R^2 value of 1.06%.

To determine if the mean Airbnb prices differr from each other across the two continent sides is Istanbul, Europe and Anatolia, we have conducted a one-way ANOVA. Based on theresults we can make the interpretation that there is a significant difference between the mean Airbnb price of Europe and Anatolia. We have also carried a Tukey HSD post-hoc test to see this difference, and the difference turned out to be about 206.17₺, with the confidence interval [153.63, 258.71], the Airbnb prices in Europe being more expensive than Anatolia.

The two-way ANOVA was done to see the individual and interaction effects of the room_type and side variables. We could observe that the mean price values of different room types also differed significantly from each other. Based on the p-value being smaller than 0.001, we could make the conclusion that the prices of different room types are not consistent across the two continent sides in Istanbul.

Wilcoxon Rank Sum test was used to determine if the distribution of the price is equal on both continents. The p-value of this test being smaller than 0.001 rejects this hypothesis. The Kruskal-Wallis test was conducted if the distribution of the price differs across room types, and this test also gave a p-value that is smaller than 0.001. We can also make the same conclusion as we did previously. The Spearman Rank Correlation Coefficient was calculated to test the relationship between number of reviews and the Airbnb prices. This test gave a ρ value that is not equal to 0, meaning that there is a statistically significant correlation between these features, although it is a weak correlation.

13.2 Biases, Limitations, and Potential Errors

Our dataset may be subject to several sources of bias. Because of the dataset consisting only the listings in Istanbul, it may contain a geographic bias. These finding may not be generalized on the different cities' Airbnb listings. Another source of bias may be a survivorship bias, that the listings those stay in the Airbnb market for a short amount of time may not be truly represented.

One of the limitations of this dataset is that the dataset consists of listings that span from 2012 to 2025. Because of this study being done in 2026, the change in the economy like inflation may have significantly affected the Airbnb listing prices in Istanbul; and therefore not representing true prices today. Extremely high prices and incorrect data may also influence the mean and other statistical measures, distorting the results.

The dataset used in this study also did not include paired before-and-after data and we could not conduct the Sign test and the Wilcoxon Signed Rank non-parametric tests. Our dataset is also not time ordered and we could not use the Wald-Wolfowitz Runs test.

14 CONCLUSION

This study analyzed Airbnb listings in Istanbul using statistical methods to better understand the characteristics of the dataset and the distribution of the prices. The analysis showed a high variation in Airbnb prices, with the presence of extreme values causing the distribution to be skewed than normal. As a result, the median of the prices may provide a more representative summary of typical listing prices than the mean.

These finding show the importance of considering the data distribution when interpreting the Airbnb prices. The findings can be useful for hosts, travelers and also researchers who want to understand the pricing patterns of the Airbnb prices in Istanbul.

According to the limitations in section **13.2**, future research could compare multiple cities and analyze a data that has a consistent time period, or adjust prices for inflation to give more accurate results.

15 REFERENCES

Inside Airbnb. (2025). *Inside Airbnb: Adding data to the debate*. insideairbnb.com