Chapter 5

Appendix: Tables & Figures

Table 5.1: Multiple Linear Regression Models for Listing Prices on Willhaben. Snapshot sample size: n = 254.

Dependent Variable	Willhaben Ad Listing Price		
	logarithm (log)	Euros (€)	Euros (€)
Size Category 2	0.463***	122.719***	122.633***
	(0.053)	(10.196)	(10.700)
Size Category 3	0.540***	147.148***	148.611***
	(0.049)	(8.411)	(8.939)
Size Category 4	0.739***	222.415***	223.786***
	(0.051)	(11.200)	(11.562)
Size Category 5	0.881***	284.006***	285.935***
	(0.059)	(16.447)	(16.956)
Size Category 6	0.816***	250.610***	252.040***
	(0.066)	(13.771)	(13.706)
Size Category 7	-0.024	-3.934	-3.934
	(0.126)	(8.863)	(9.351)
Size Category 8	0.491***	134.110***	133.110***
	(0.095)	(49.283)	(48.851)
Good Condition	0.102*	36.055***	36.684***
	(0.052)	(11.513)	(11.620)
Used Condition	-0.080***	-26.981***	-26.152***
	(0.018)	(5.291)	(5.344)
Dealer i	-0.018	-22.672*	-22.672*
	(0.034)	(13.094)	(12.979)
Last 48 Hours i	-0.045*	-15.366*	-17.010**
	(0.023)	(8.057)	(8.168)
Psychological Pricing i	0.034*	13.648*	13.648*
	(0.020)	(7.747)	(7.836)
Logistic Costs i	/	/	0.651*
			(0.585)
Constant	5.335***	210.243***	206.266***
	(0.049)	(9.022)	(9.702)
Observations	254	254	254
R-squared	0.646	0.647	0.651
Adjusted R-squared	0.628	0.628	0.632

Notes: t-statistics are calculated using Huber-White robust standard errors. Significance levels: * p≤0.1; ** p≤0.05; *** p≤0.01.
Sample size n = 254 out of approximately 1,500 observations due to missing data. This snapshot was obtained by web-scraping the following URL: https://www.willhaben.at/iad/kaufen-und-verkaufen/marktplatz/fahrraeder/kinderfahrraeder-4558?keyword=woom
Variable calculation:

$$\text{logistic\_costs} = \begin{cases} 0 & \text{if total\_count} = 0 \text{ (Note: potential bias due to missing imputation for this variable).} \\ \frac{1}{\text{weighted\_sum}} & \text{otherwise} \end{cases}$$

where total_count = Number of Products within 0-10 km (i) + Number of Products within 10-30 km (i) + Number of Products within 30-60 km (i) and weighted_sum = Number of Products within 0-10 km (i) $\cdot \frac{1}{10}$ + Number of Products within 10-30 km (i) $\cdot \frac{1}{30}$ + Number of Products within 30-60 km (i) $\cdot \frac{1}{60}$.

Table 5.2: Stepwise AIC-Best Regression Model. Sample size: $n = 826$.

	log(price)
Size Category 2	0.472^{***}
	(0.018)
Size Category 3	0.543^{***}
	(0.016)
Size Category 4	0.745^{***}
	(0.017)
Size Category 5	0.850^{***}
	(0.027)
Size Category 6	0.837^{***}
	(0.020)
Size Category 7	-0.056^{*}
	(0.029)
Good Condition	0.142^{***}
	(0.030)
Used Condition	-0.098^{***}
	(0.009)
Color: Blue	-0.080^{***}
	(0.026)
Color: Yellow	-0.072^{***}
	(0.027)
Color: Green	-0.065^{**}
	(0.026)
Color: Orange	-0.161^{***}
	(0.045)
Color: Red	-0.074^{***}
	(0.026)
Color: Violet	-0.033
	(0.026)
Dealer i	0.012
	(0.008)
Constant	5.400^{***}
	(0.030)
Observations	826
R-squared	0.718
Adjusted R-squared	0.713
Residual Std. Error	0.106 (df = 810)
F Statistic	137.695^{***} (df = 15; 810)

Notes: t-statistics are based on Huber-White standard errors.
Significance levels: $^{*}p \leq 0.1$; $^{**}p \leq 0.05$; $^{***}p \leq 0.01$.

Table 5.3: Multiple Linear Regression Model. Sample size: $n = 826$.

	Price Parsed
Size Category 2	121.165***
	(4.809)
Size Category 3	144.205***
	(3.767)
Size Category 4	219.943***
	(4.680)
Size Category 5	277.963***
	(8.130)
Size Category 6	261.302***
	(8.164)
Size Category 7	5.253
	(5.037)
Good Condition	53.938***
	(16.337)
Used Condition	−35.406***
	(3.414)
Dealer i	5.066
	(3.407)
Last48Hours i	−0.337
	(5.065)
HasPsychologicalPricing i	2.865
	(4.616)
Number of Same Size in 0-10 km Radius i	0.044
	(0.044)
Number of Same Size in 10-30 km Radius i	−0.027
	(0.025)
Number of Same Size in 30-60 km Radius i	−0.001
	(0.015)
Constant	217.381***
	(4.857)
Observations	826
R-squared	0.680
Adjusted R-squared	0.675

Notes: t-statistics are based on Huber-White robust standard errors.
Significance levels: *$p \leq 0.1$; **$p \leq 0.05$; ***$p \leq 0.01$.

Table 5.4: Multiple Linear Regression Model. Sample size: $n = 826$.

	Price Parsed
Size Category 2	120.533***
	(4.880)
Size Category 3	143.698***
	(3.814)
Size Category 4	219.195***
	(4.700)
Size Category 5	276.562***
	(8.106)
Size Category 6	260.235***
	(8.200)
Size Category 7	4.943
	(4.784)
Good Condition	54.776***
	(16.184)
Used Condition	−35.241***
	(3.412)
Dealer *i*	4.538*
	(2.742)
Last48Hours *i*	−0.505
	(5.089)
HasPsychologicalPricing *i*	2.895
	(4.633)
Logistic Costs	−0.504
	(1.949)
Constant	217.430***
	(4.509)
Observations	826
R-squared	0.679
Adjusted R-squared	0.675

Notes: t-statistics are based on Huber-White robust standard errors.
Significance levels: *$p \leq 0.1$; **$p \leq 0.05$; ***$p \leq 0.01$.

Table 5.5: Fixed Effects Model for *hasPsychologicalPricing_i*

	Dependent Variable:
	log_price
hasPsychologicalPricing_i	3.841
	(0.115)
Fixed Effects:	
Size	✓
Color	✓
Condition	✓
Observations	826
RMSE	37.3
Adjusted R-squared	0.689

Note: p-Value in parentheses. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Moran's I Test under Randomisation for Psychological Pricing	
Statistic	Value
Moran I statistic standard deviate	1.4564
p-value	0.07264
Expectation	-0.001212121
Variance	0.001543322

Table 5.6: Moran's I Test Results for Psychological Pricing. Two snapshots: n = 826.

Moran's I Test under Randomisation for Log Price	
Statistic	Value
Moran I statistic standard deviate	-0.62634
p-value	0.7345
Expectation	-0.001212121
Variance	0.001544525

Table 5.7: Moran's I Test Results for Log Price. Two snapshots: n = 826.

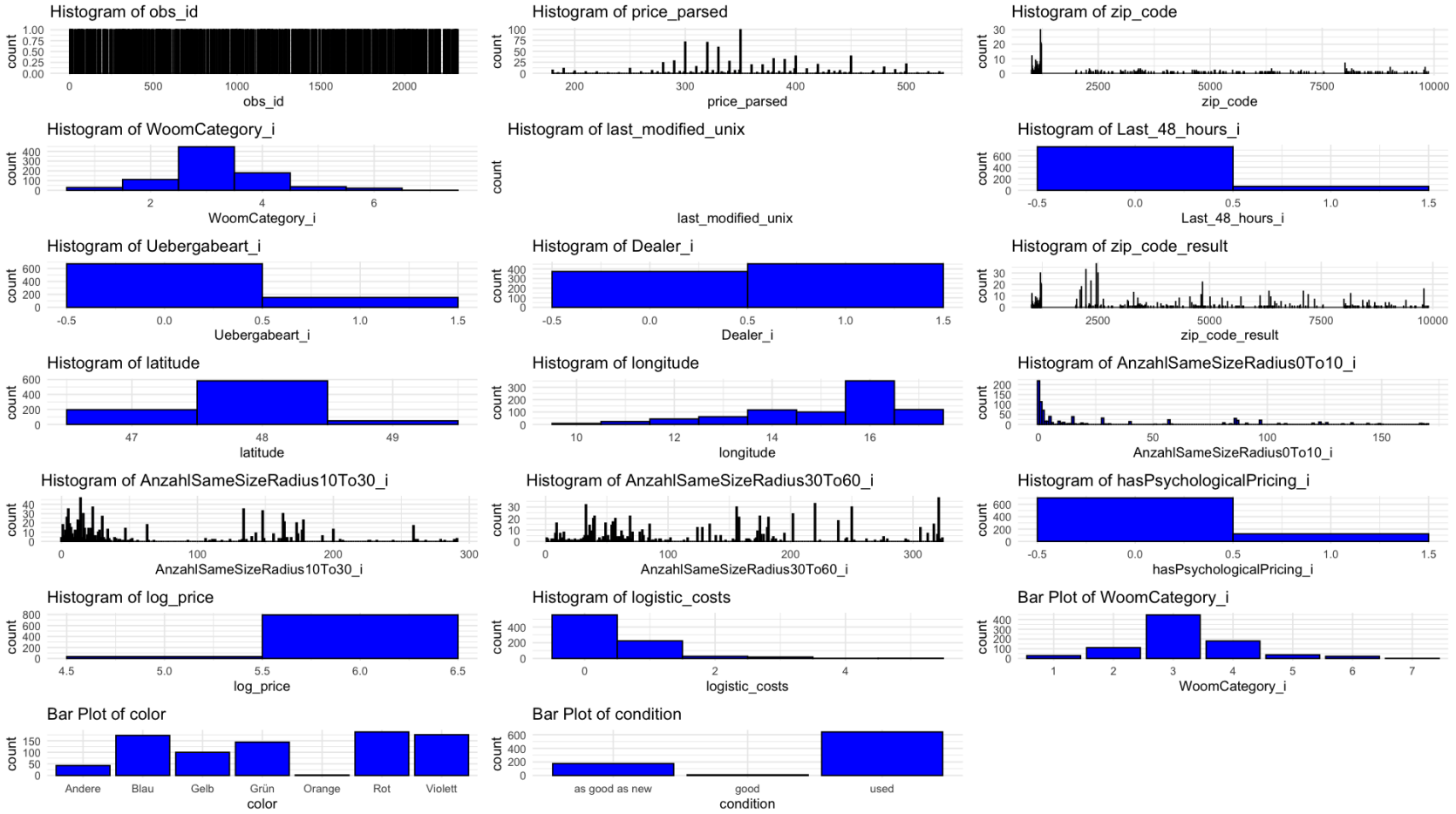

Figure 5.1: Histograms of Key Variables. This figure provides a view of the distribution of most variables, including categorical variables across the dataset.

Note: Notably, price outliers that fall outside the 1.5 IQR (Interquartile Range) have already been removed. Two snapshots: n = 826.

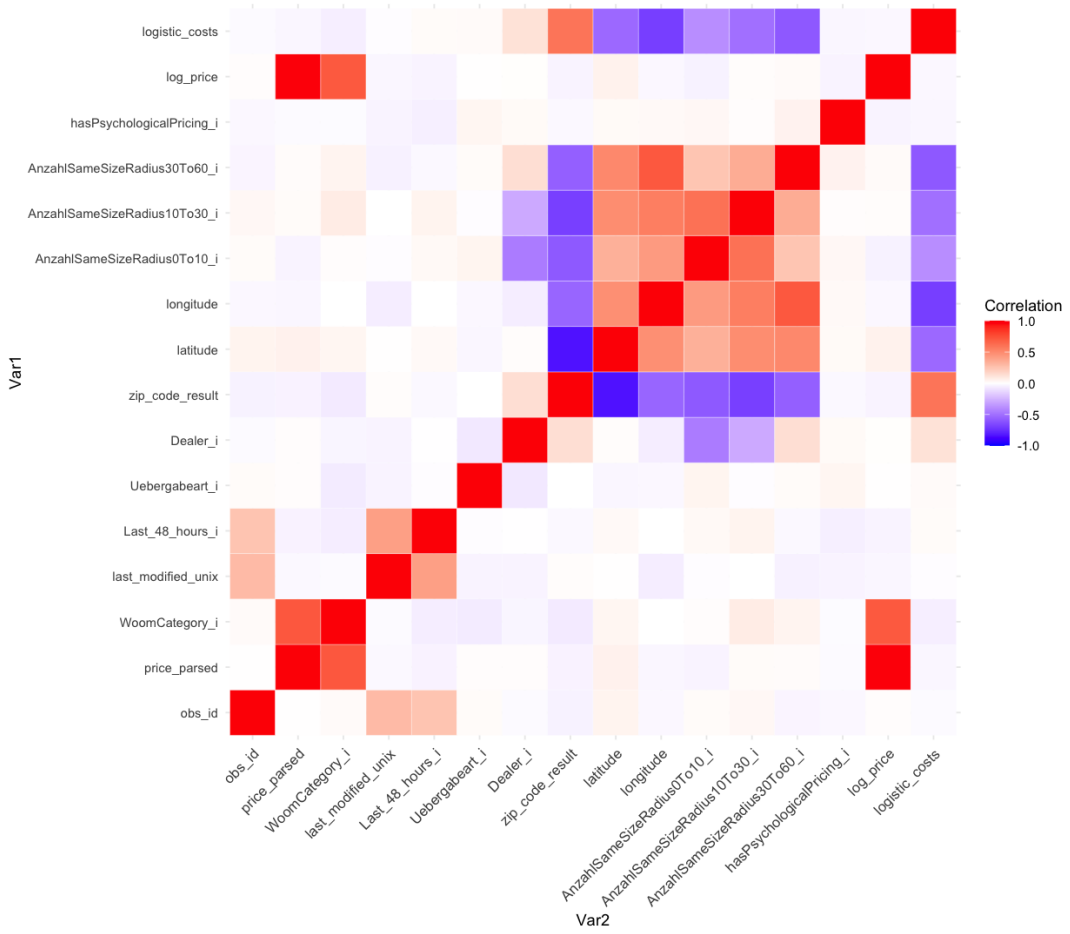

Figure 5.2: Correlation Heatmap. This heatmap visualizes the pairwise correlations between numeric variables in the dataset. *Note:* The heatmap helps in identifying patterns of multicollinearity and relationships between variables. Strong correlations (both positive and negative) are visible as colored blocks, while weak or no correlations are depicted in white. Two snapshots: n = 826.

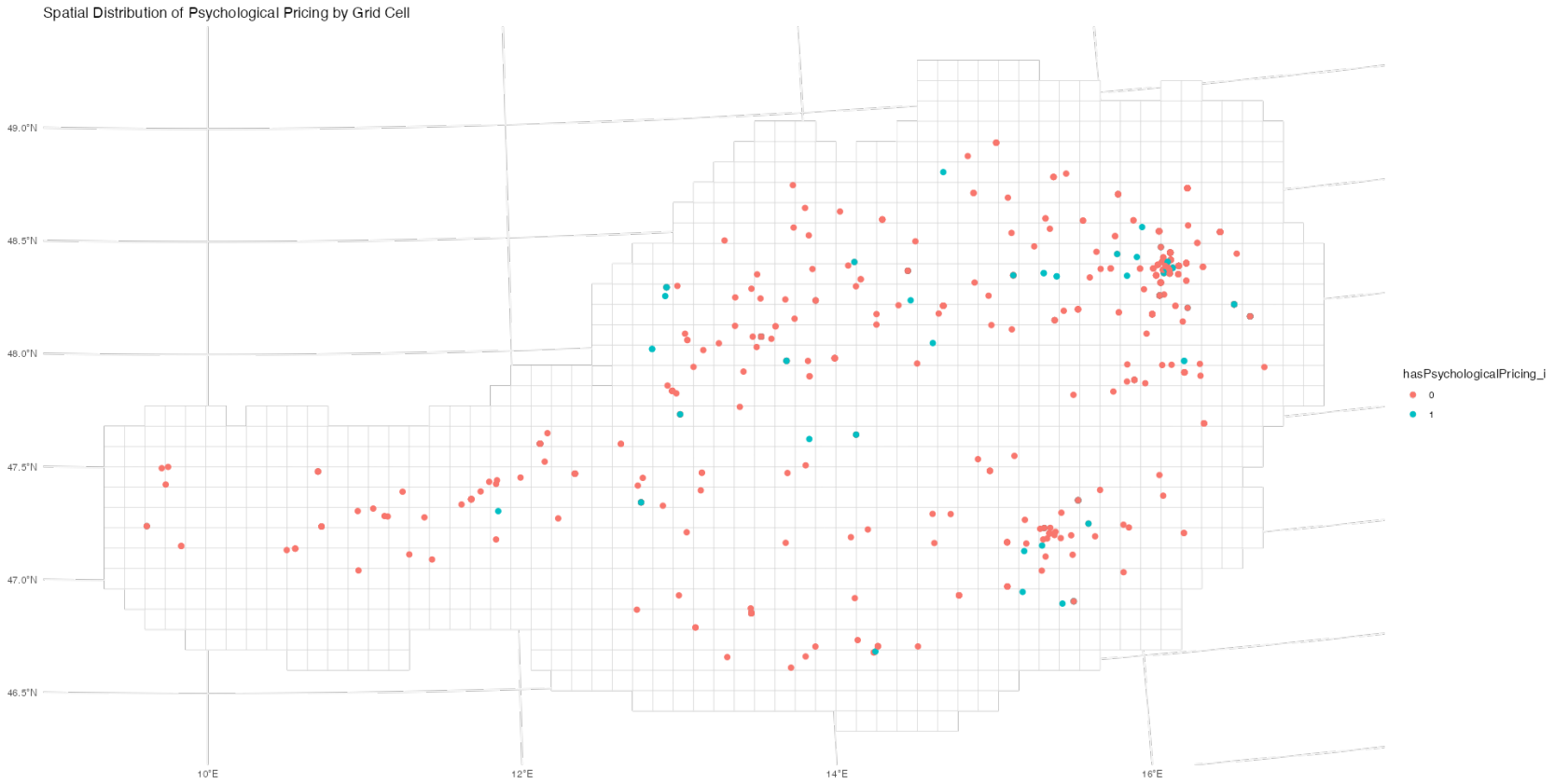

Figure 5.3: Spatial Distribution of Psychological Pricing. This figure illustrates the geographical distribution of the psychological pricing binary dummy variable across Austria.

Note: The plot overlays the spatial distribution of bike data points onto a shapefile of Austria, with different colors representing the presence or absence of psychological pricing. Two snapshots: n = 826.

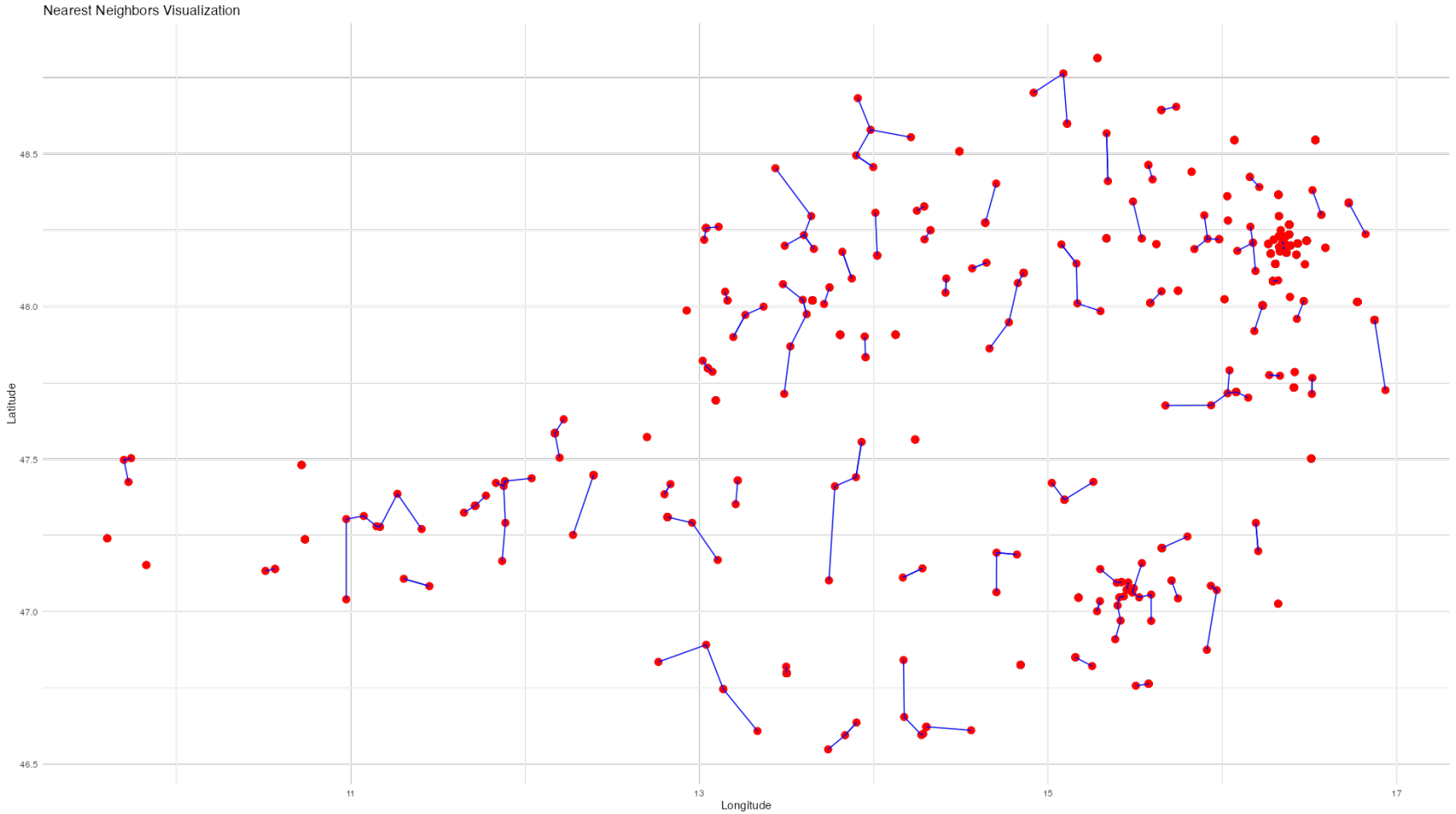

Figure 5.4: This figure depicts the nearest neighbor relationships among bike data points using K-Nearest Neighbors (KNN) with $K = 1$. *Note:* Each data point is connected to its nearest neighbor, highlighting the spatial connections and proximity between observations. The red points represent the bike data locations, while the blue lines indicate the connections to the nearest neighbors. Two snapshots: n = 826.

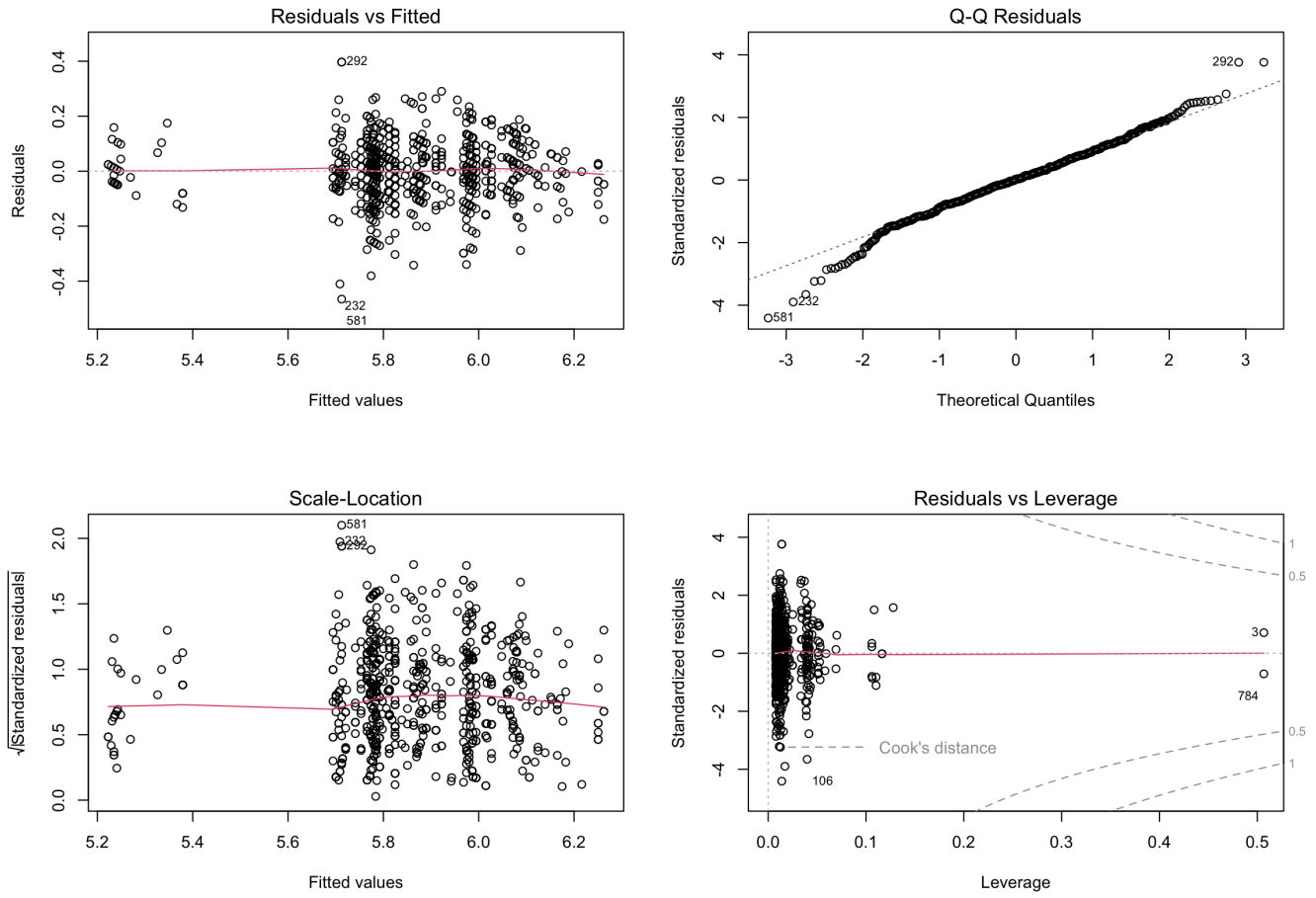

Figure 5.5: Residual Analysis/Q-Q Plot.

Note: The Q-Q plot is a diagnostic tool for checking the normality assumption in regression analysis. Points deviating from the 45-degree line may indicate issues with the residuals. Two snapshots: n = 826.