Group 6

Appendices

Appendix A – Correlation Heatmap

Figure A1 Correlation Heatmap
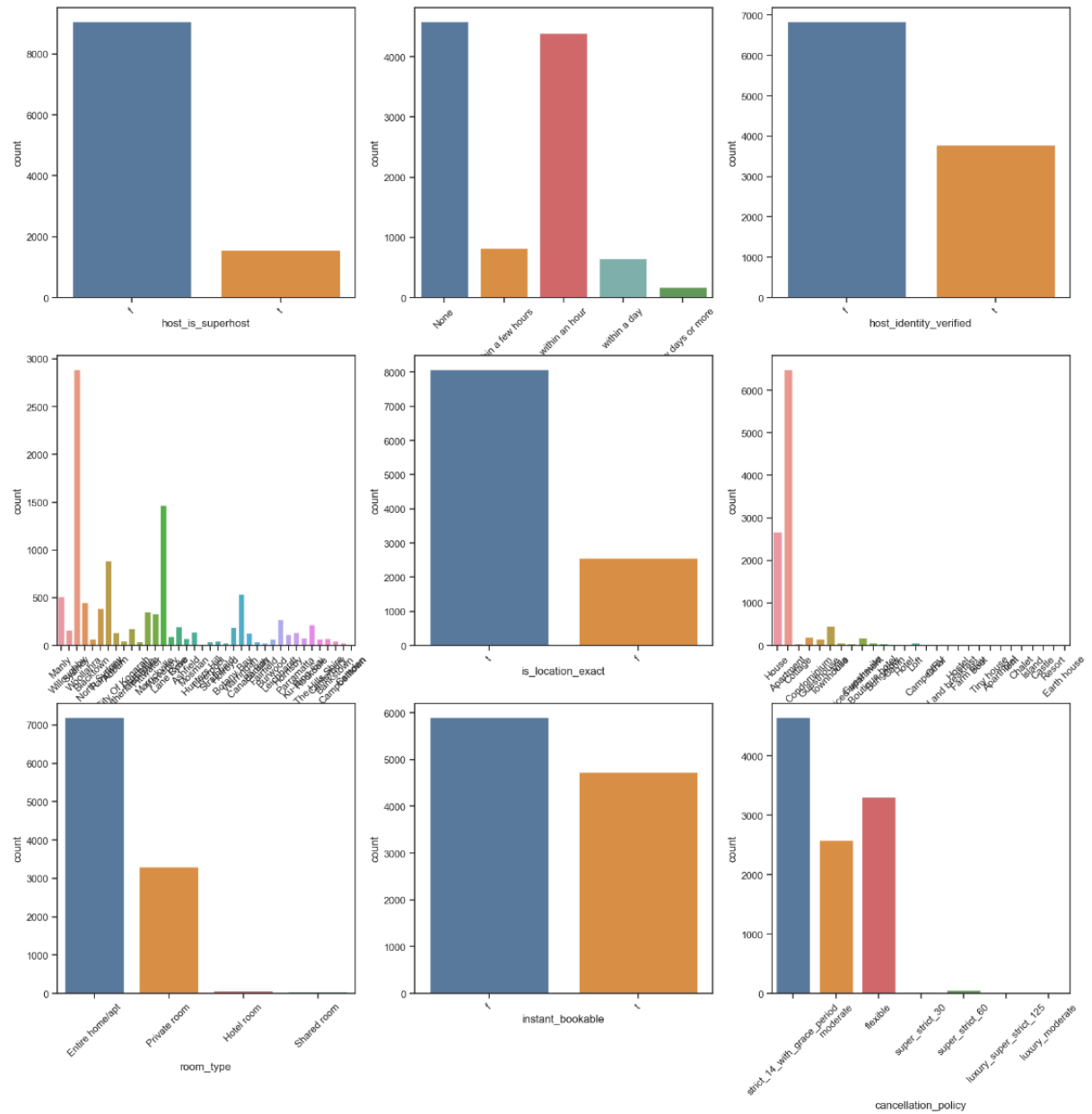
Appendix B – Categorical Variables

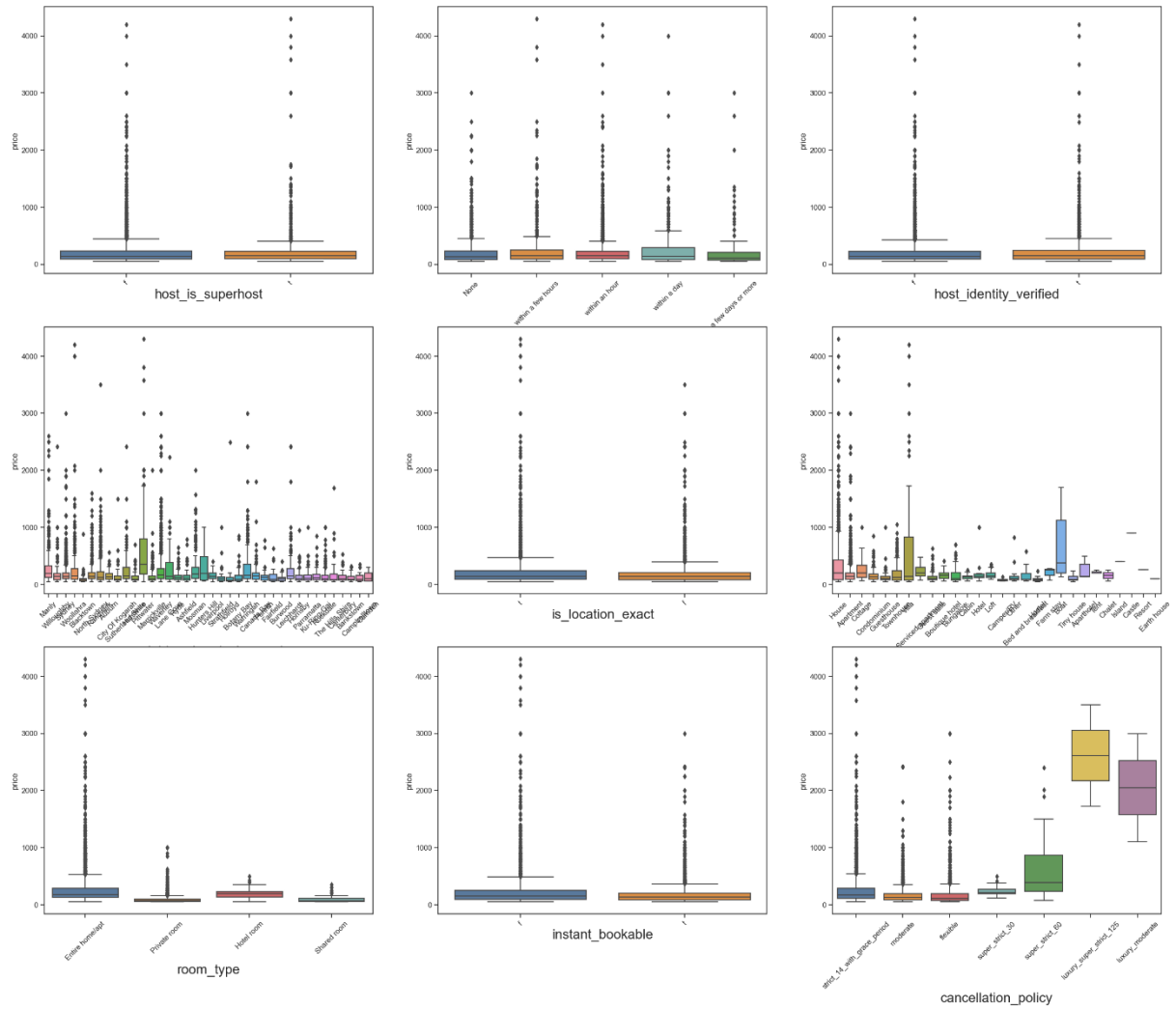*Figure B1 Count Plots of Categorical Variables*

Figure B2 Boxplots of Categorical Variables

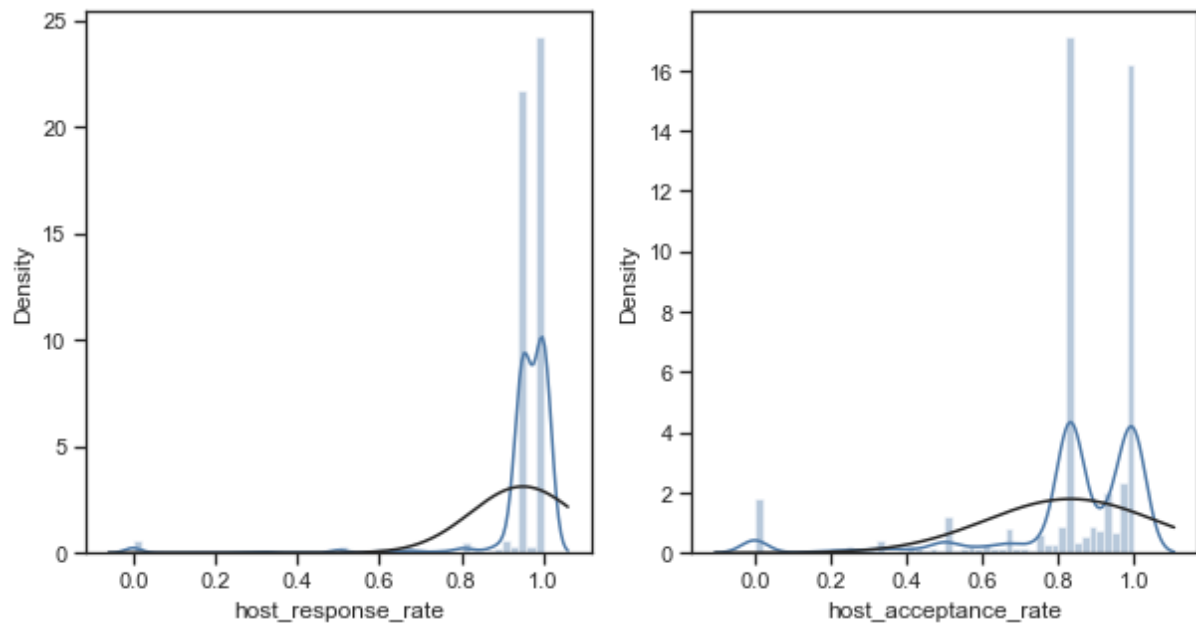Appendix C – Host response rate & host acceptance rate

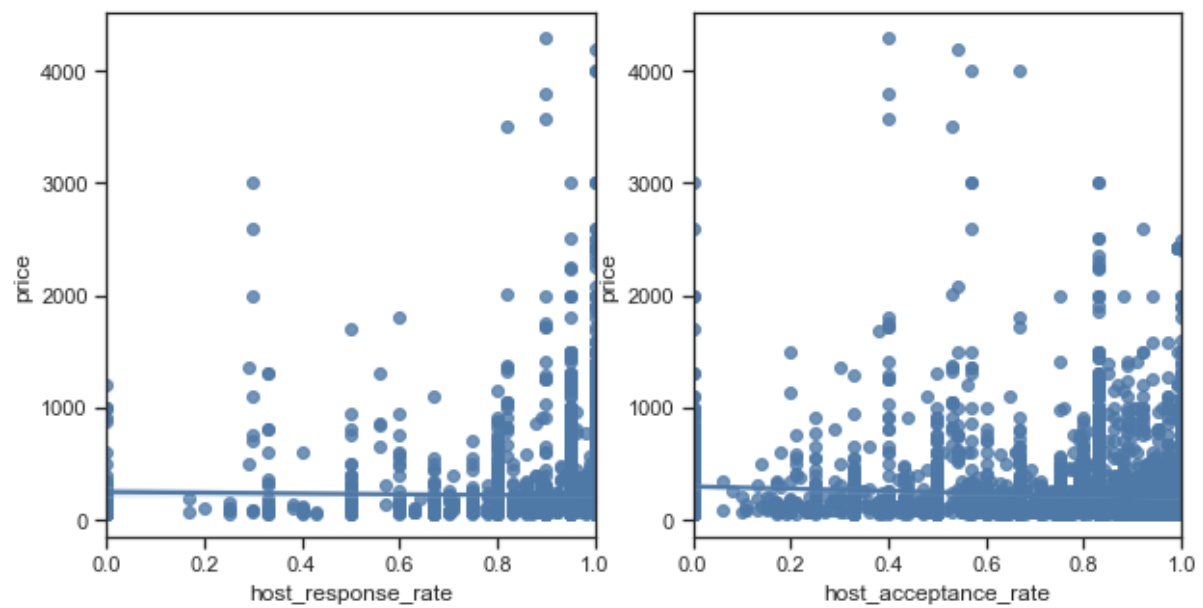*Figure C1 Distribution Plots of host_response_rate & host_acceptance_rate*

Figure C2 Regression Plots of host_response_rate & host_acceptance_rate
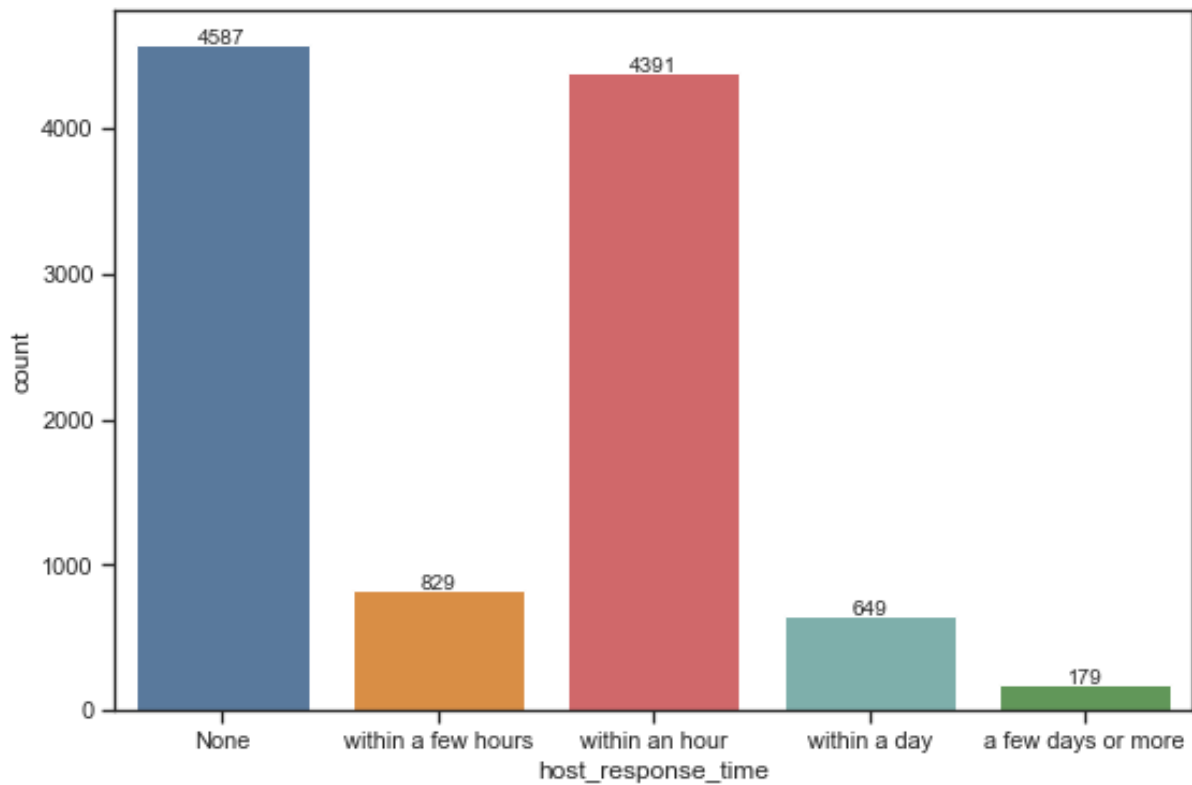
Appendix D – Host response time

Figure D1 Count Plot of host_response_time

	None	A few days or more	Within a day	Within a few days	Within an hour
Count	4587.00	179.00	649.00	829.00	4391.00
Mean	198.9780	249.8380	266.9384	254.6888	208.7725
std	206.6150	400.3514	367.7830	370.6425	245.5139
Min	51.00	51.00	51.00	51.00	51.00
25%	85.00	72.50	85.00	89.00	100.00
50%	129.00	110.00	140.00	150.00	150.00
75%	231.00	205.00	290.00	250.00	220.00
Max	2999.00	2999.00	4000.00	4300.00	4200.00

Table D1 Descriptive Statistics of host_response_time
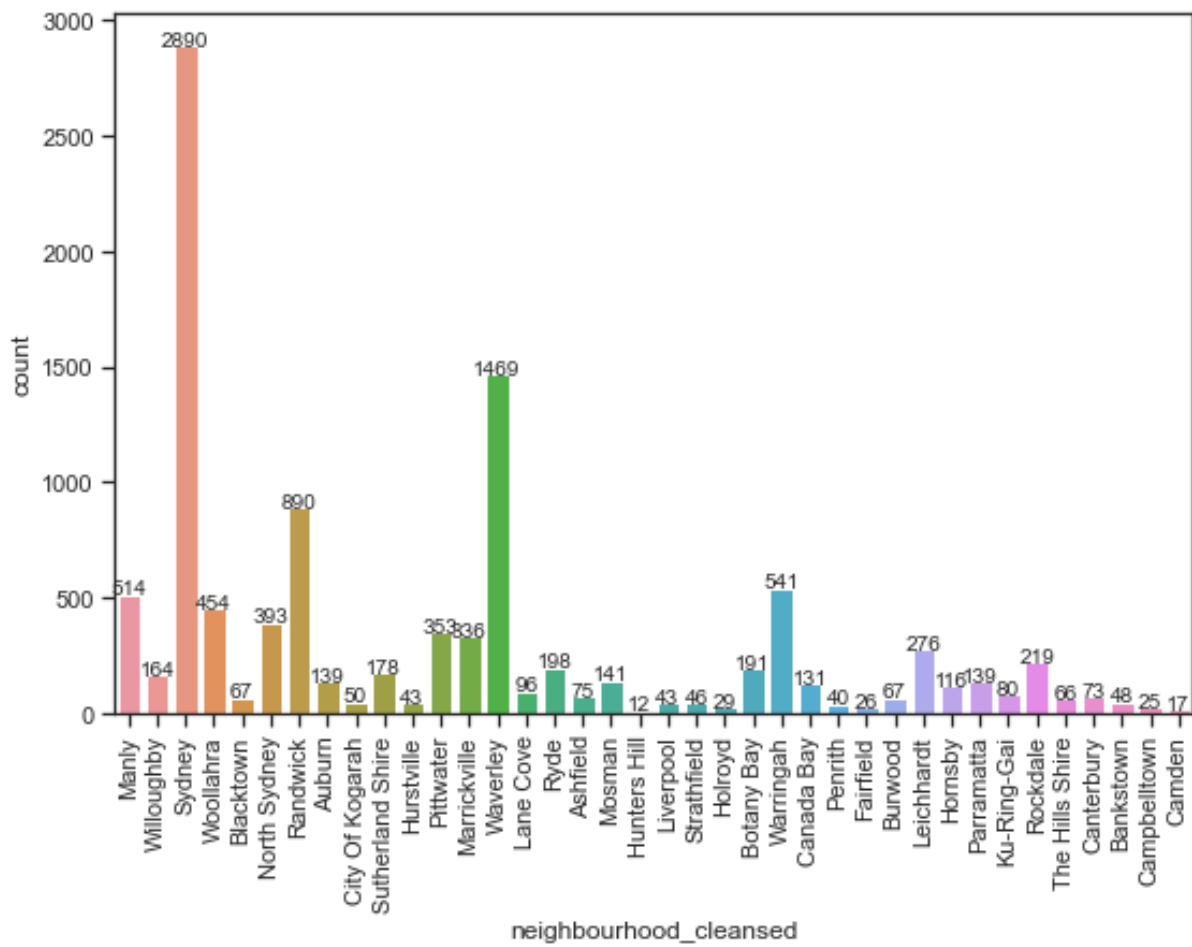
Appendix E – Neighborhood Cleansed

Figure E1 Count Plot of neighbourhood_cleansed

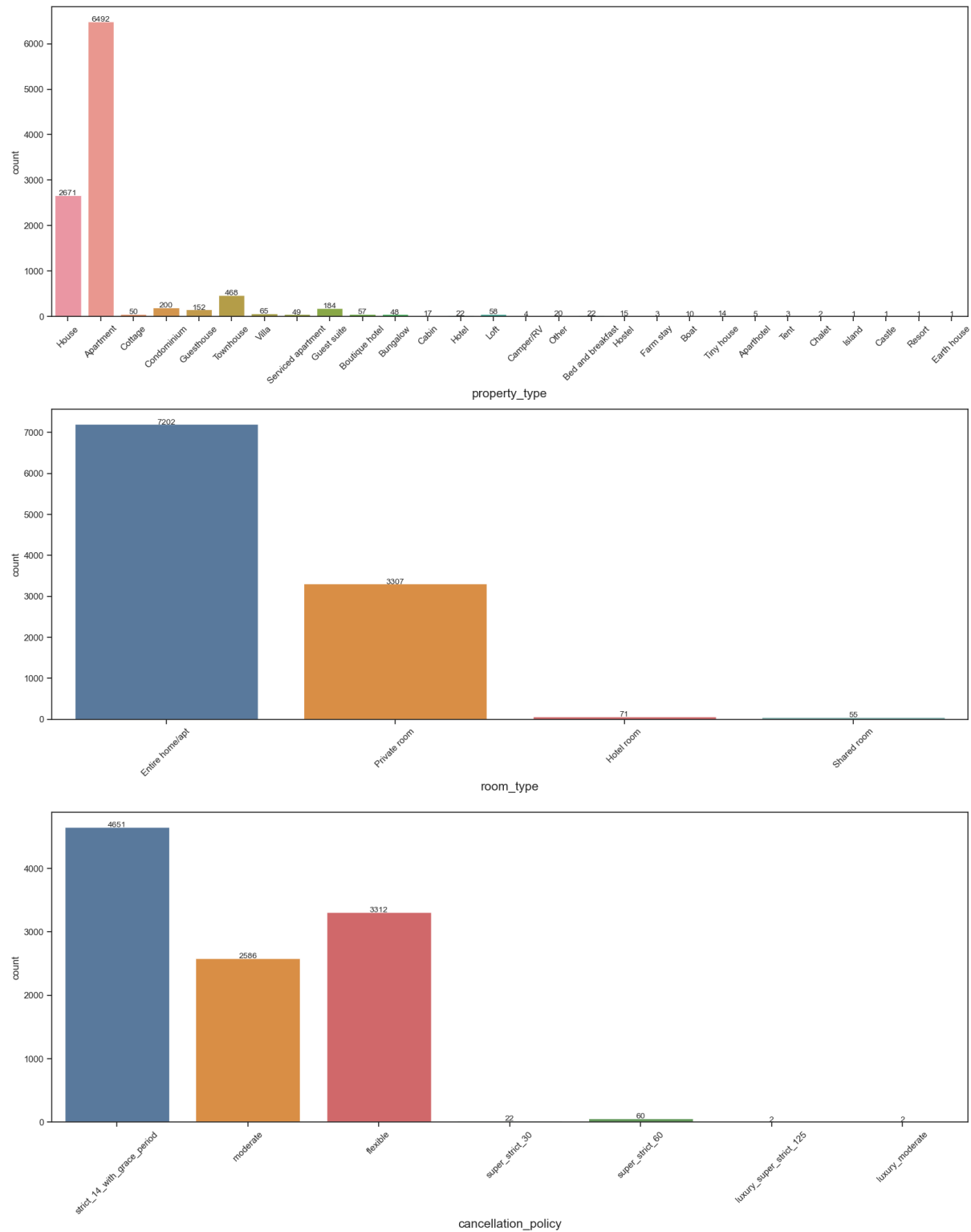Appendix F – Property Type & Room Type & Cancellation Policy

Figure F1 Count Plots of property_type, room_type & cancellation_policy

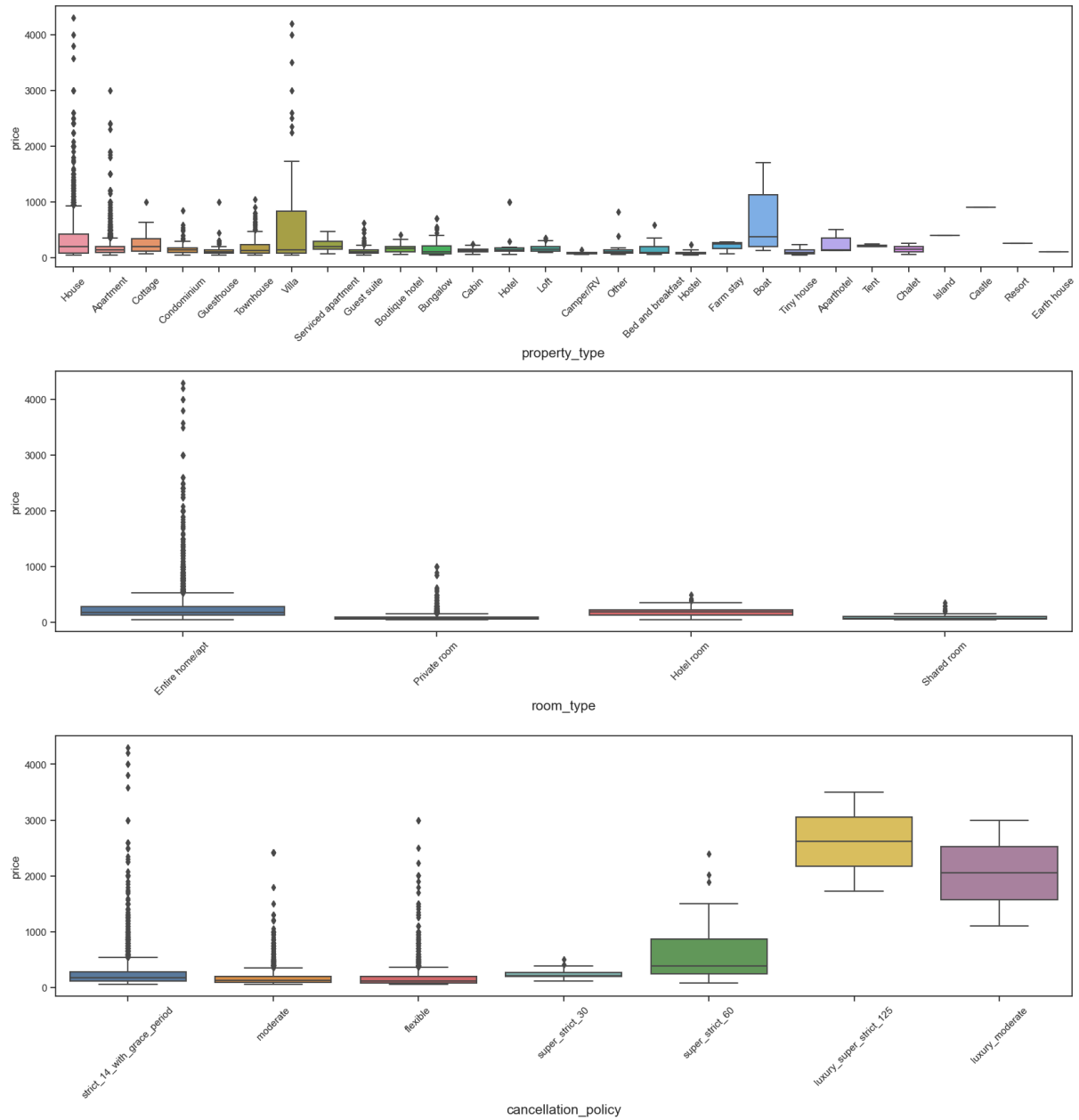*Figure F2 Boxplots of property_type, room_type & cancellation_policy*
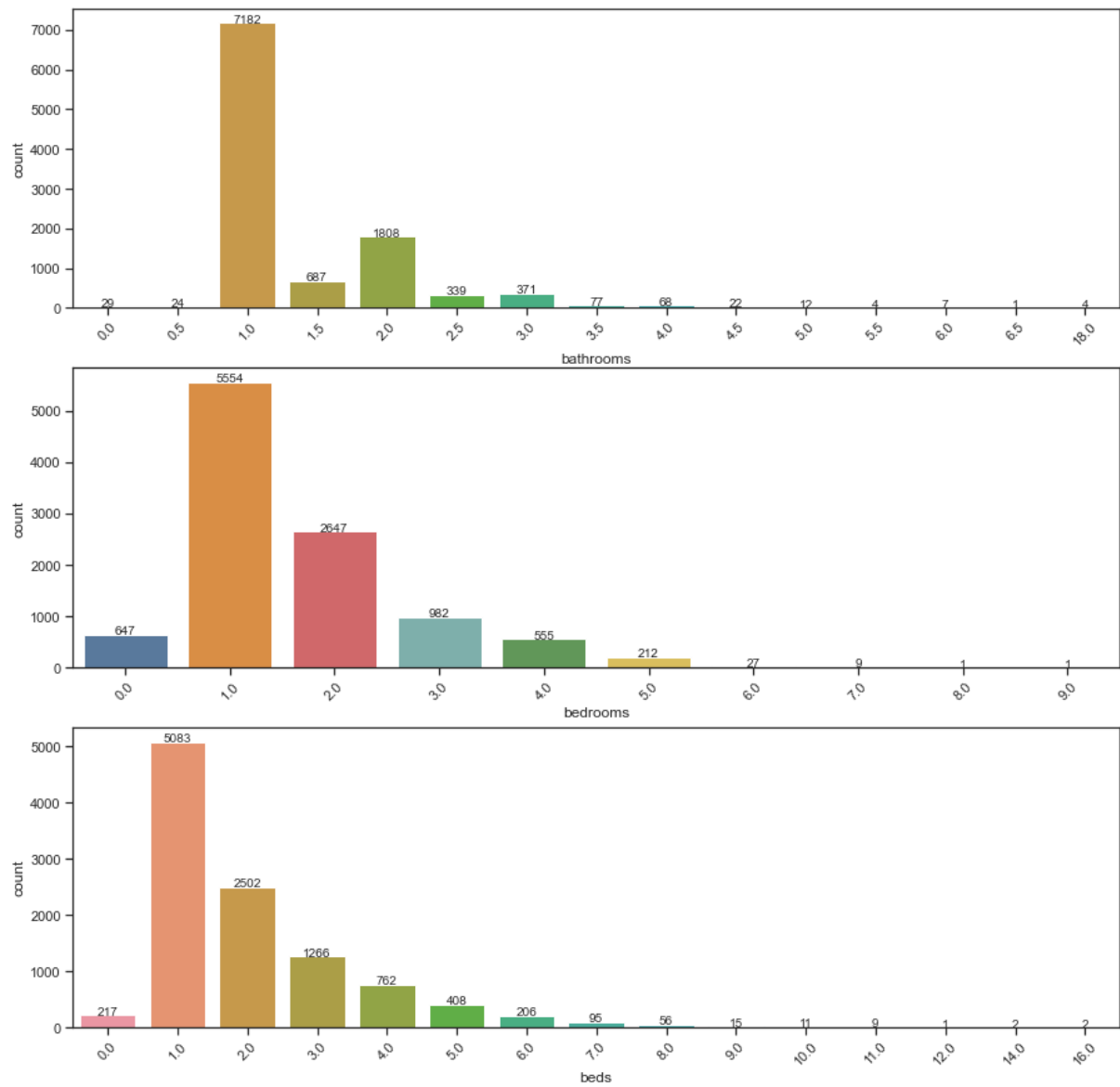
Appendix G – Bathrooms, Bedrooms & Beds

Figure G1 Count Plots of Bathrooms, Bedrooms & Beds

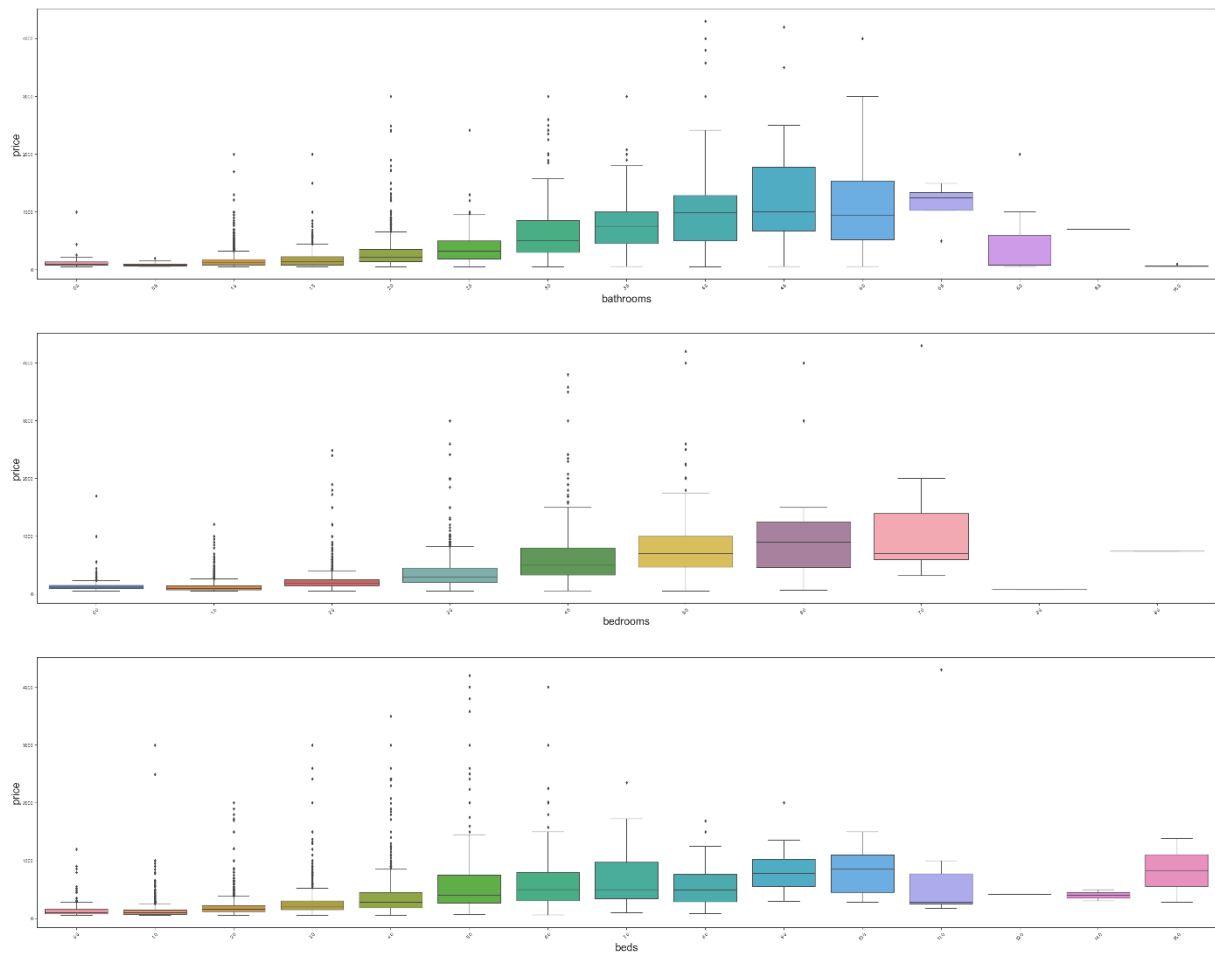*Figure G2 Boxplots of Bathrooms, Bedrooms & Beds*

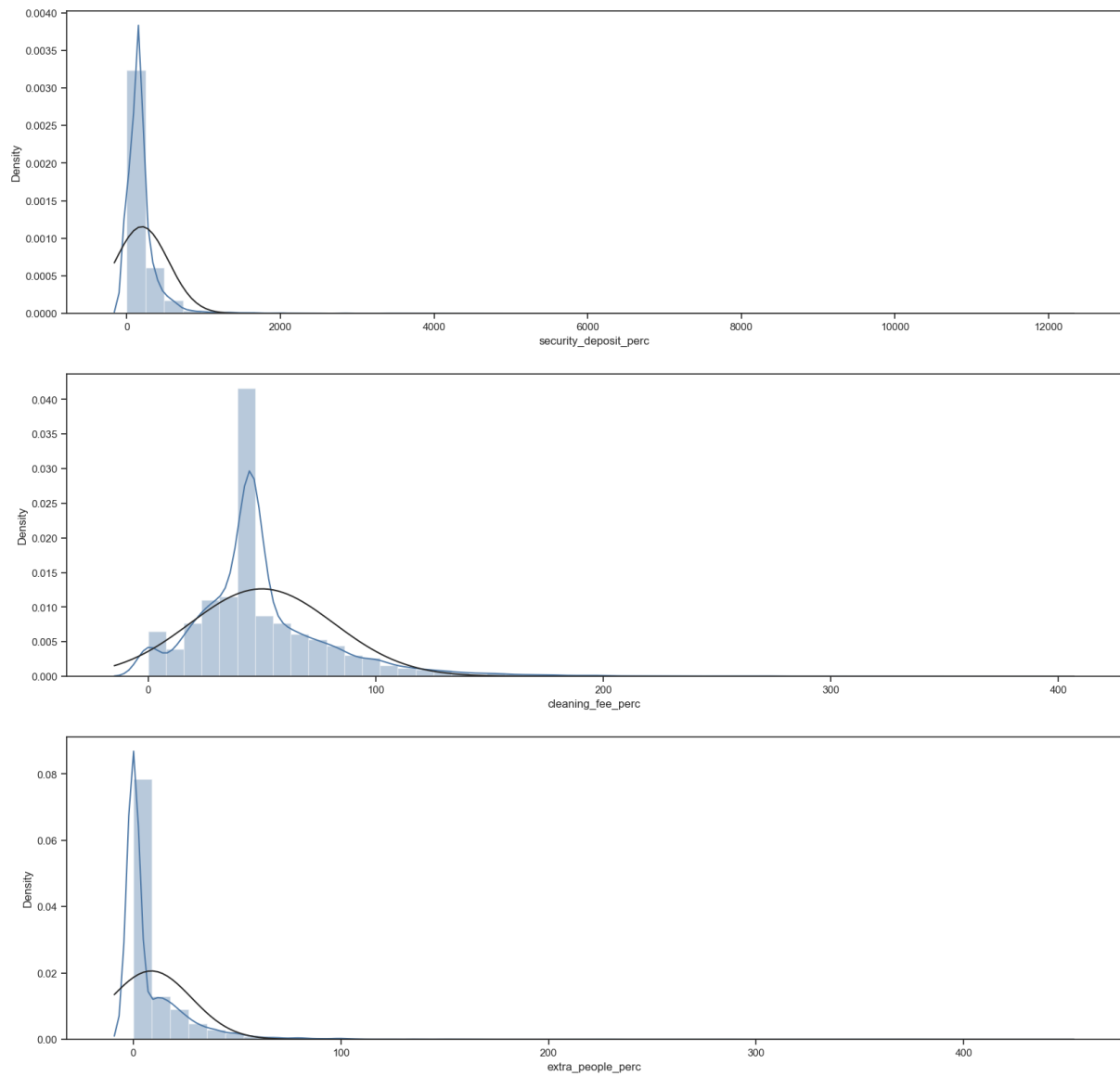Appendix H - Security_deposit_perc, Cleaning_fee_perc & Extra_people_perc

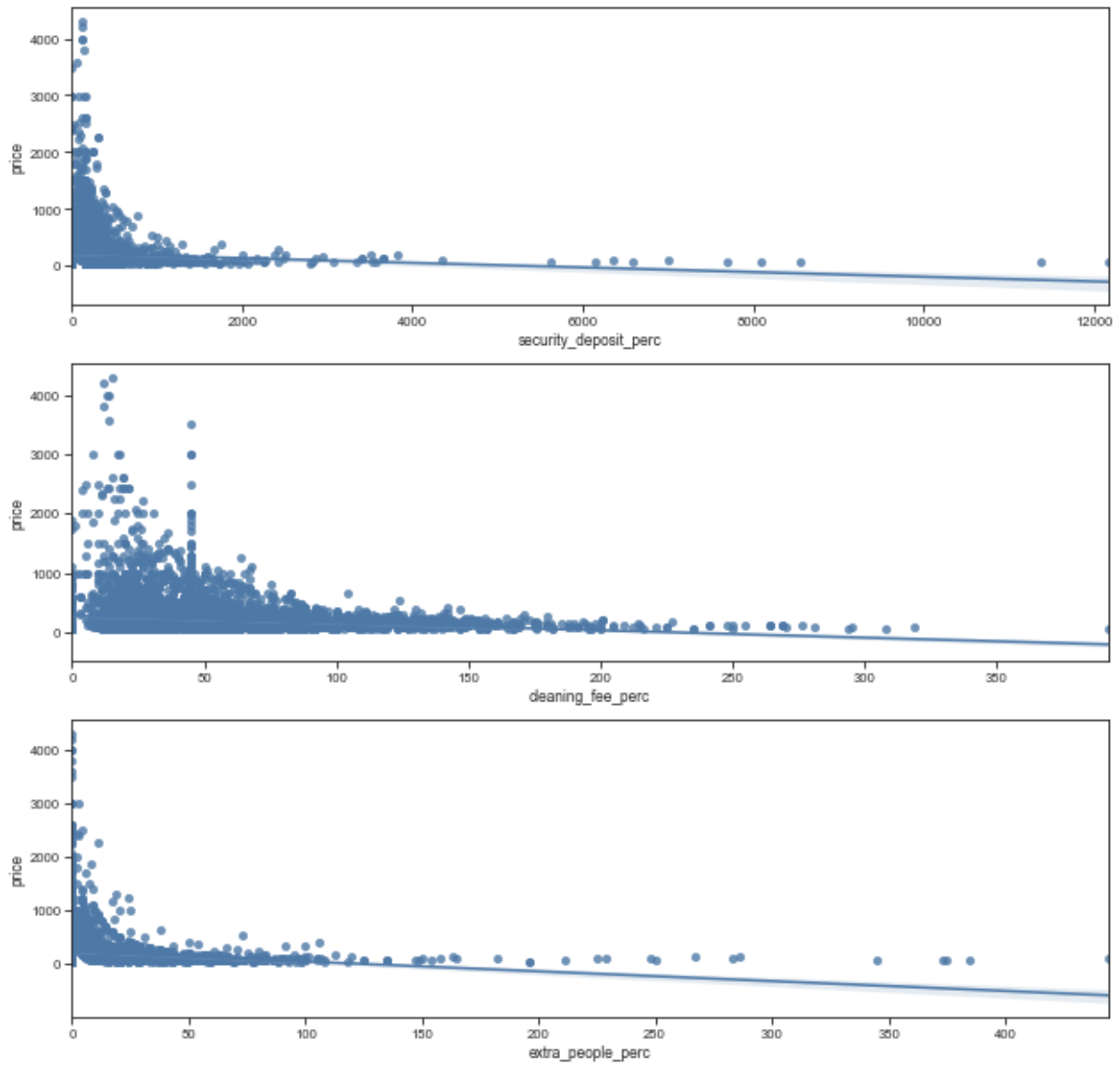*Figure H1 Distribution Plots of Security_deposit_perc, Cleaning_fee_perc & Extra_people_perc*

Figure H2 Regression Plots of Security_deposit_perc, Cleaning_fee_perc & Extra_people_perc

Appendix I – Guest included, minimum_nights_avg_ntm, maximum_nights_avg_ntm & availability_30
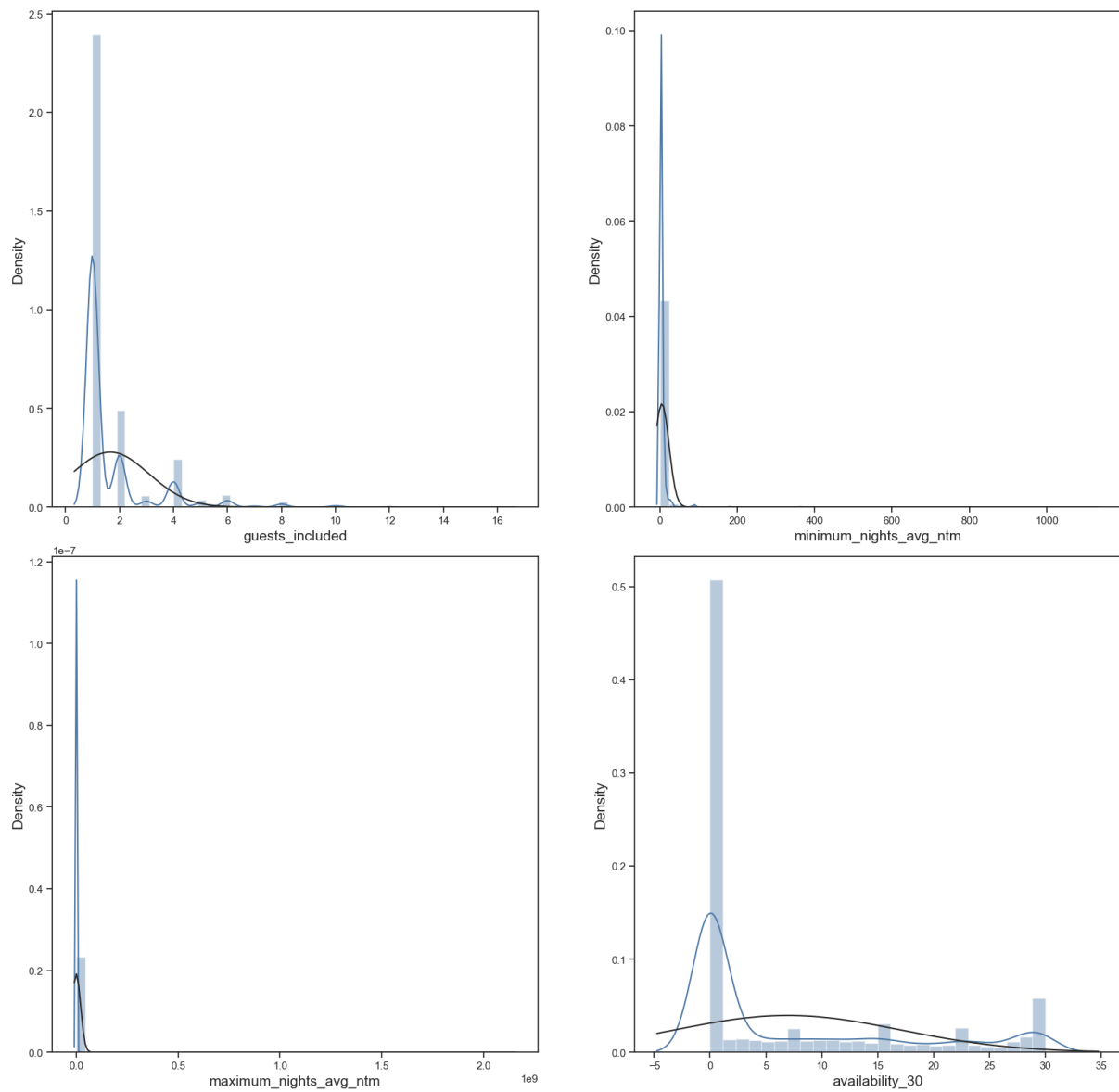

Figure I1 Distribution Plots of guests_included, minimum_nights_avg_ntm, maximum_nights_avg_ntm & availability_30
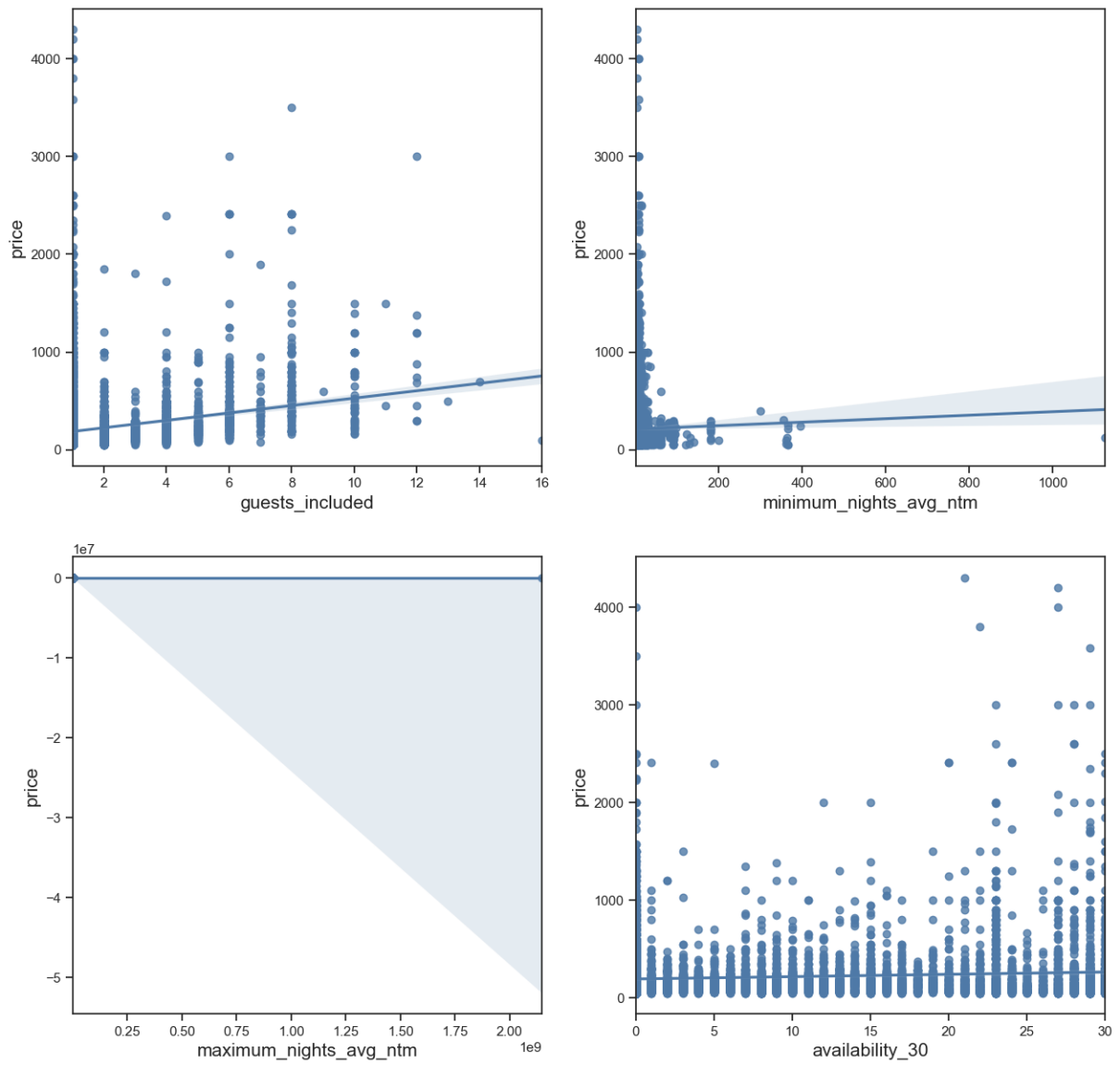
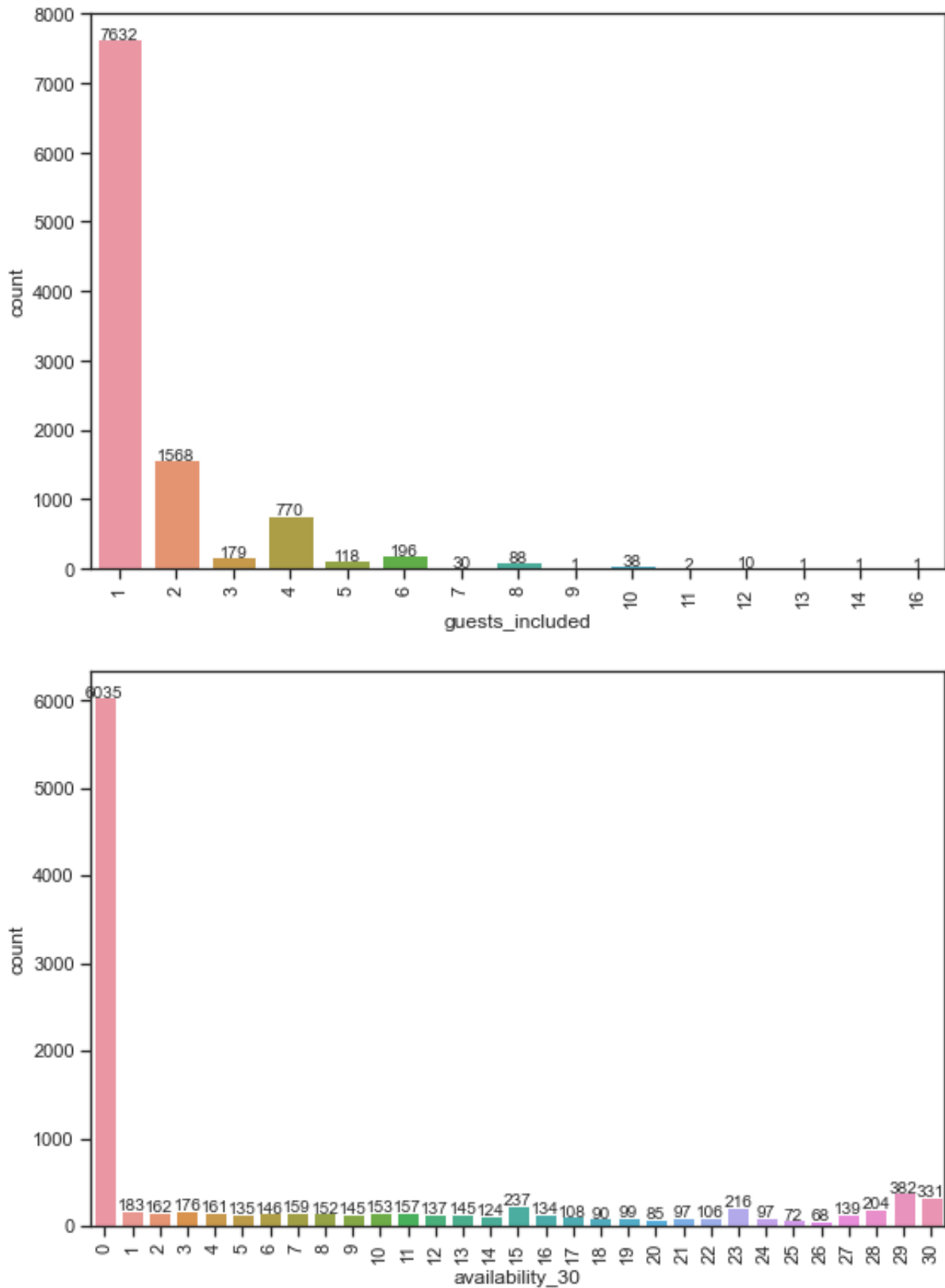*Figure I2 Regression Plots of guests_included, minimum_nights_avg_ntm, maximum_nights_avg_ntm & availability_30*

Figure I3 Count Plots of guests_included & availability_30
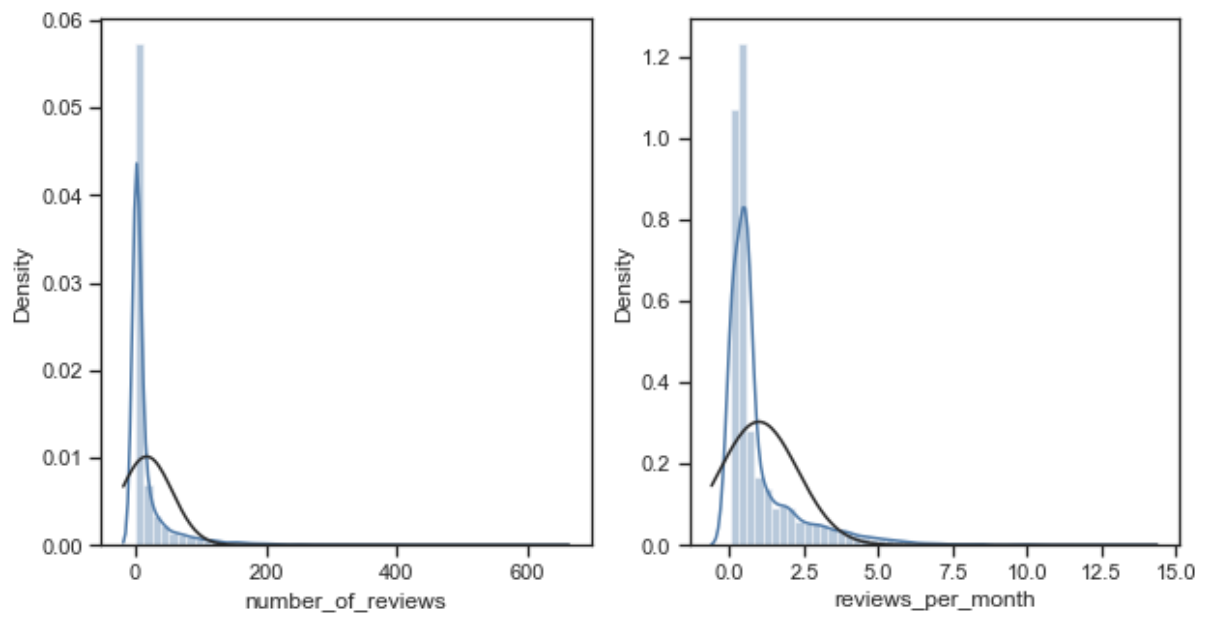
Appendix J – Number of reviews, reviews_per_month

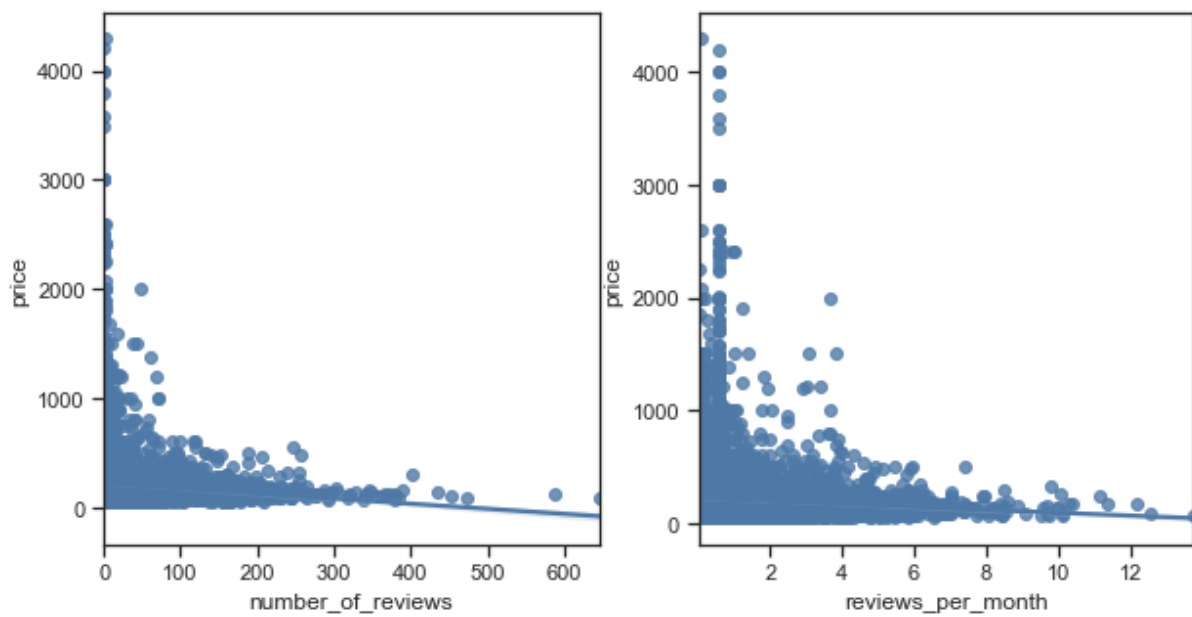*Figure J1 Distribution Plots of number_of_reviews & reviews_per_month*

Figure J2 Regression Plots of number_of_reviews & reviews_per_month
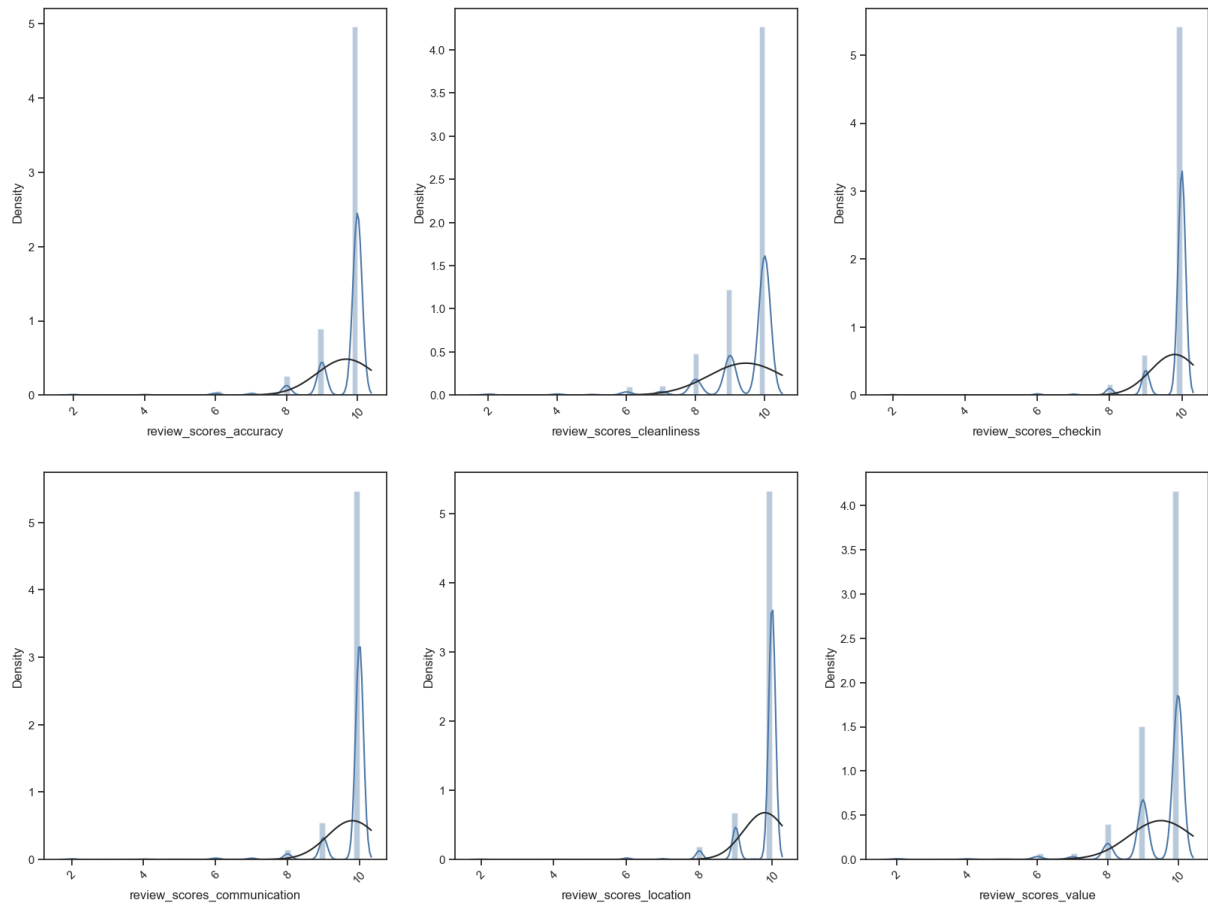
Appendix K – Review Scores
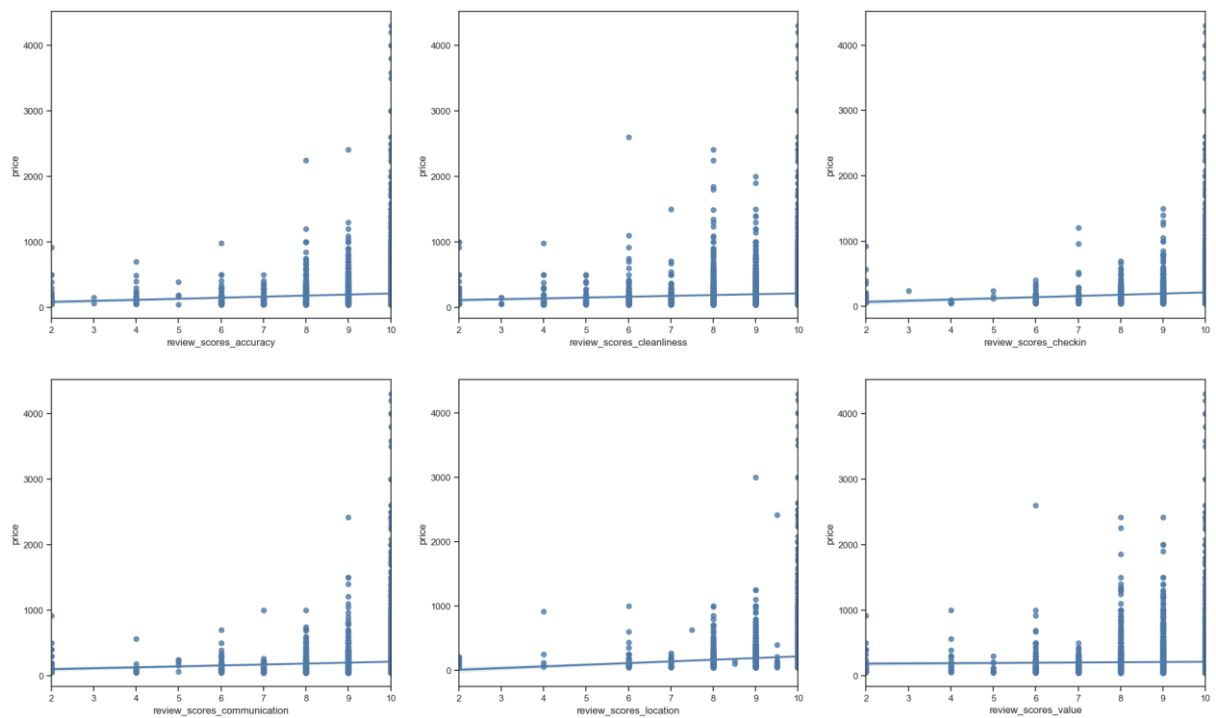

Figure K1 Distribution Plots of Review Scores

Figure K2 Regression Plots of Review Scores

Appendix L – calculated_host_listings_count_entire_homes &
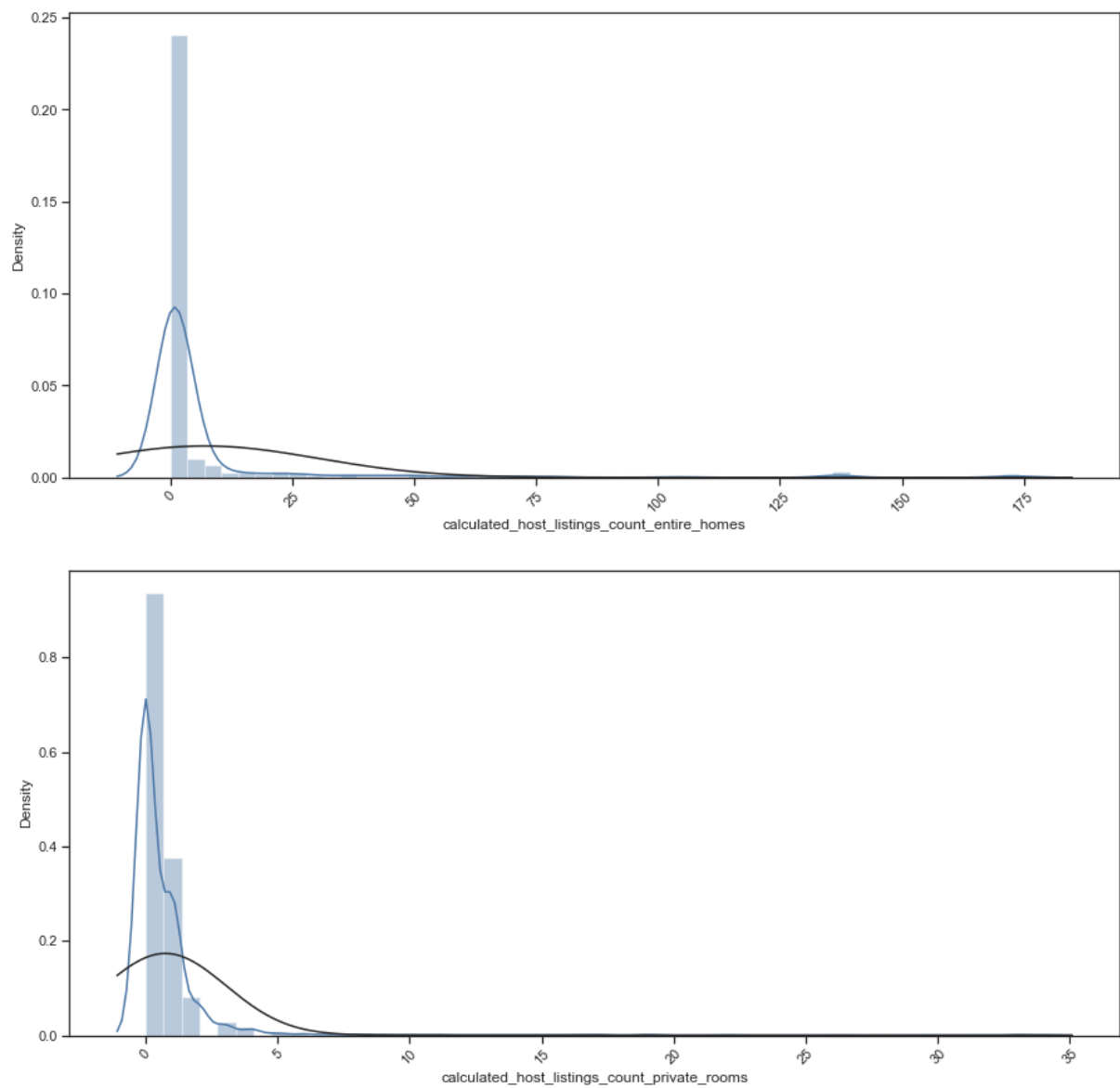calculated_host_listings_count_private_roomes

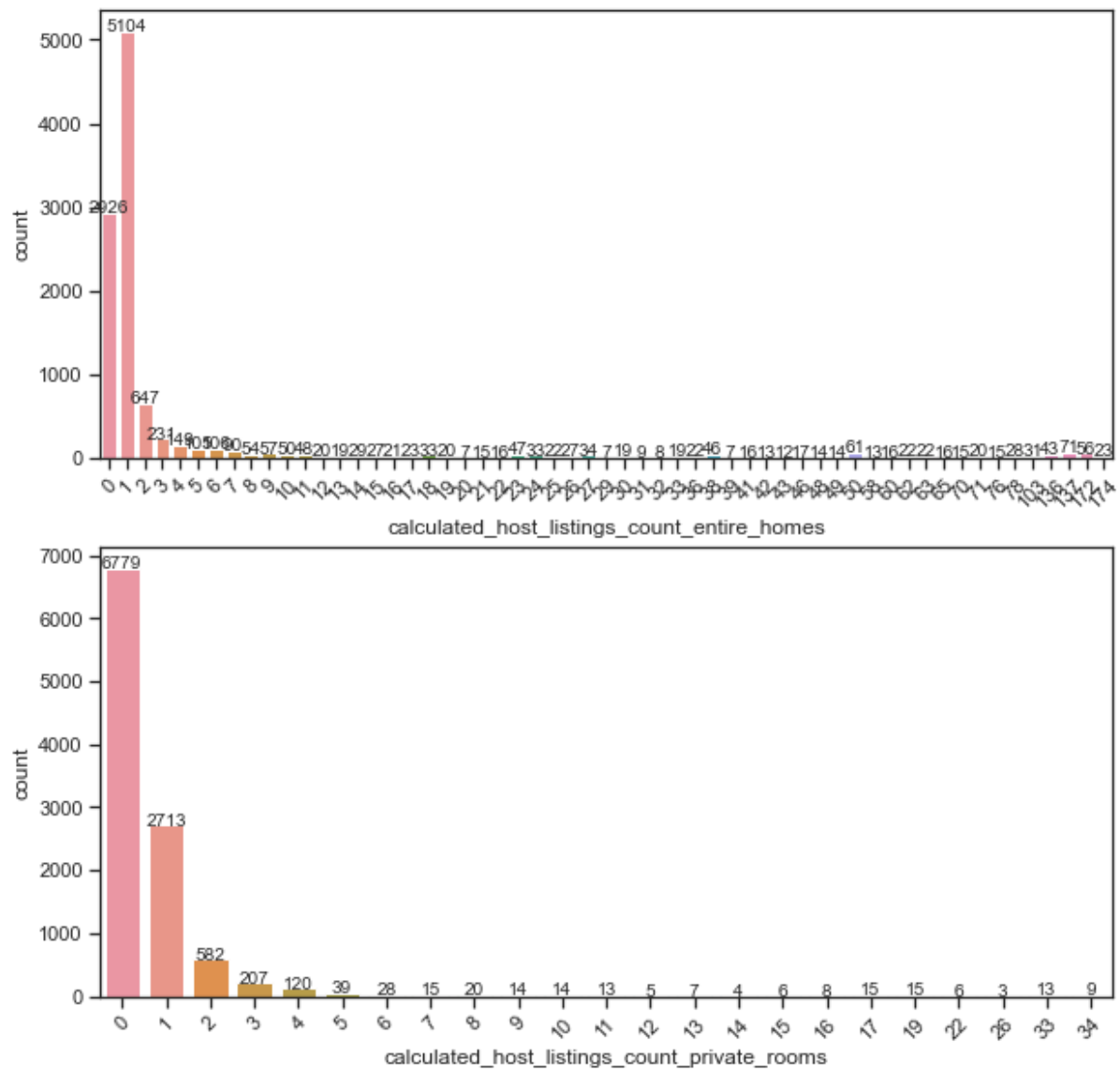*Figure L1 Distribution Plots of host_listings_count*

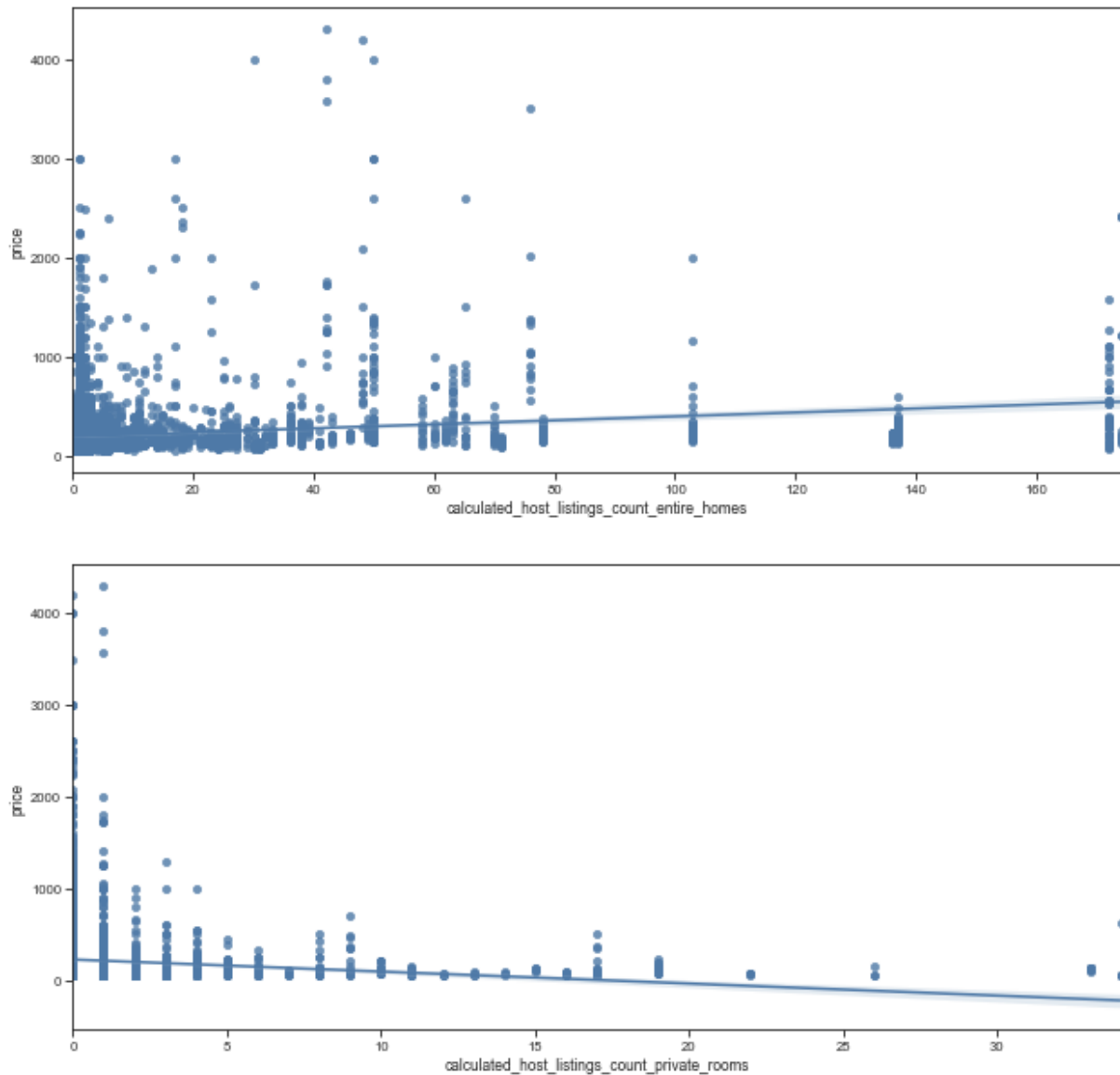*Figure L2 Count Plots of host_listings_count*

Figure L3 Regression Plots of host_listings_count

Appendix M – Algorithm of Random Forest

Figure M1 Algorithm of Random Forest

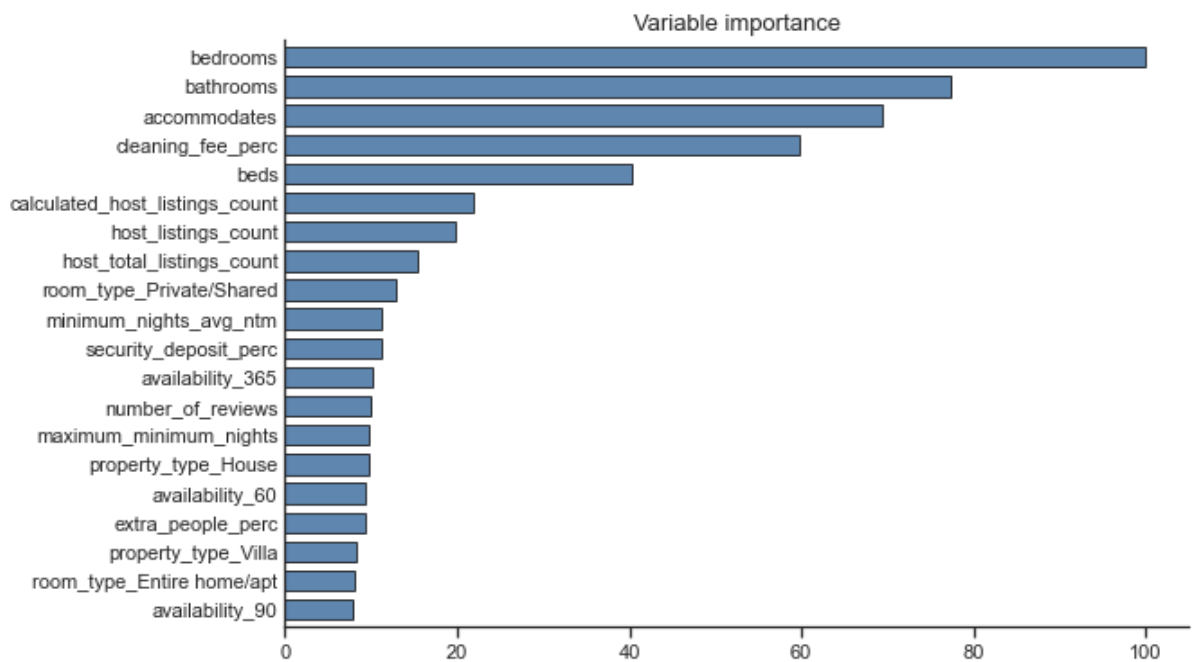Appendix N – Variance Importance

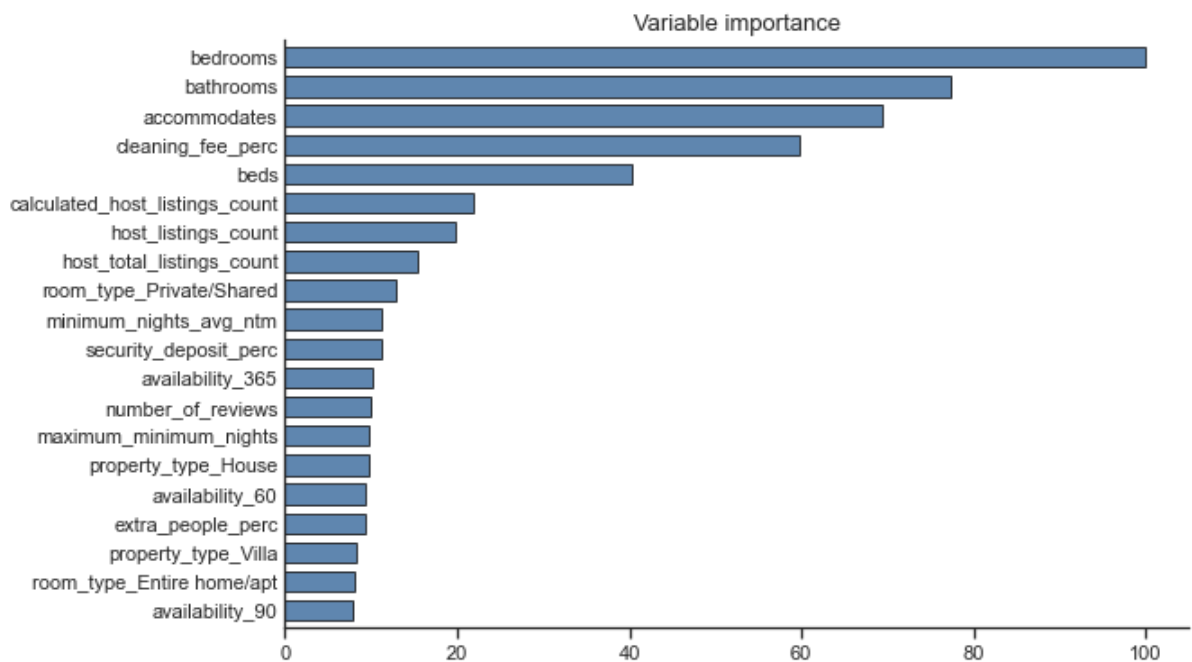
Figure L1 Variable Importance of Ridge

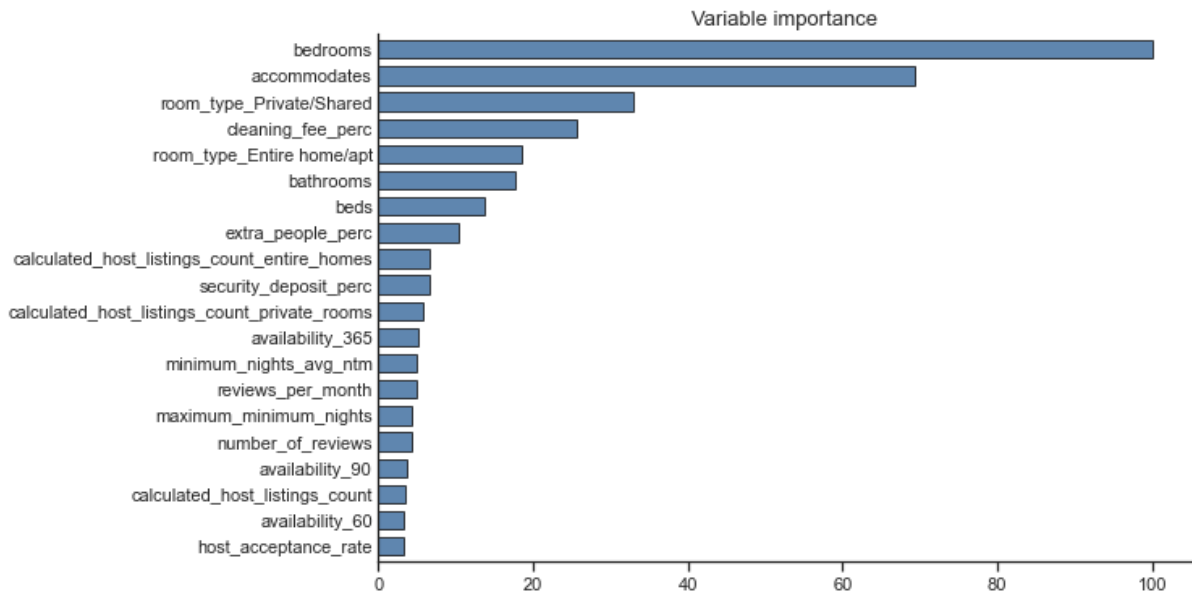*Figure L2 Variable Importance of Random Forest (without hyperparameter tuning)*

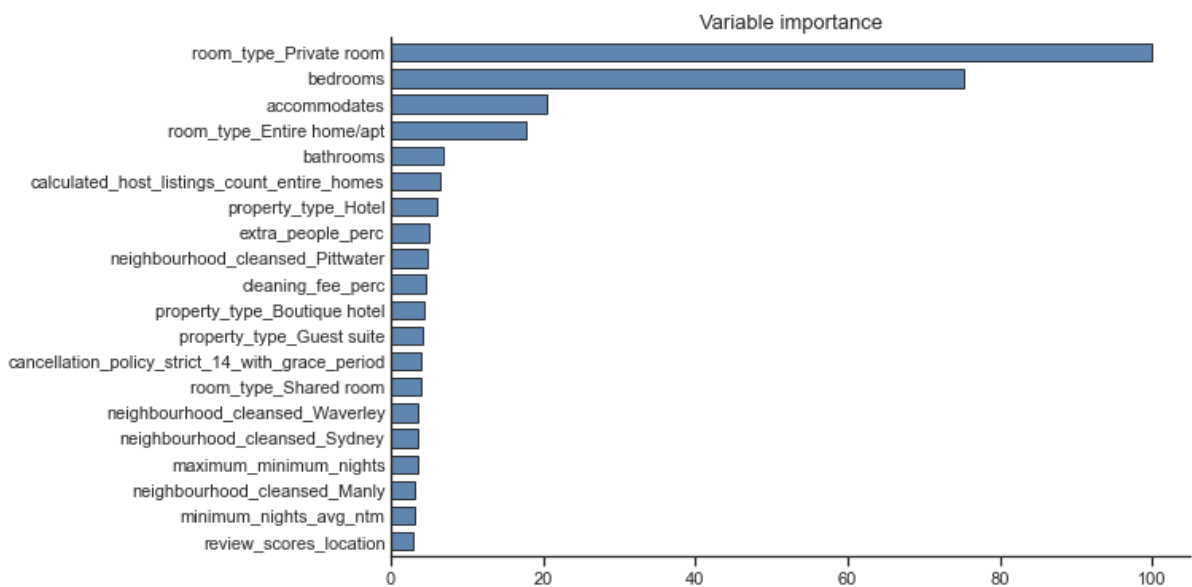*Figure L3 Variable Importance of Random Forest (after hyperparameter tuning)*

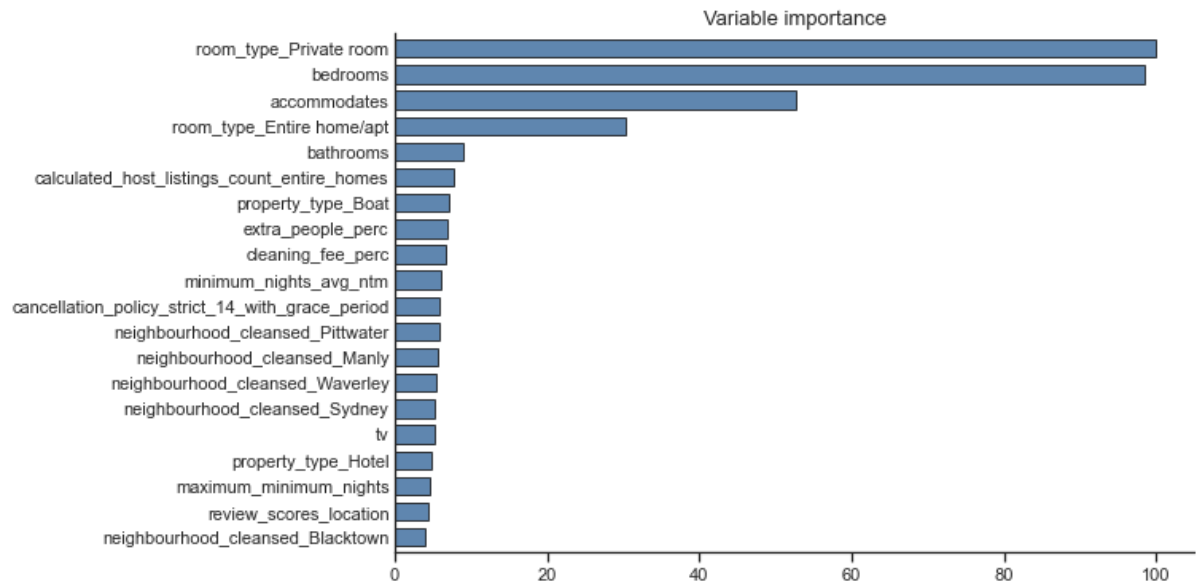*Figure L4 Variable Importance of XGBoost (without hyperparameter tuning)*

Figure L5 Variable Importance of XGBoost (after hyperparameter tuning)

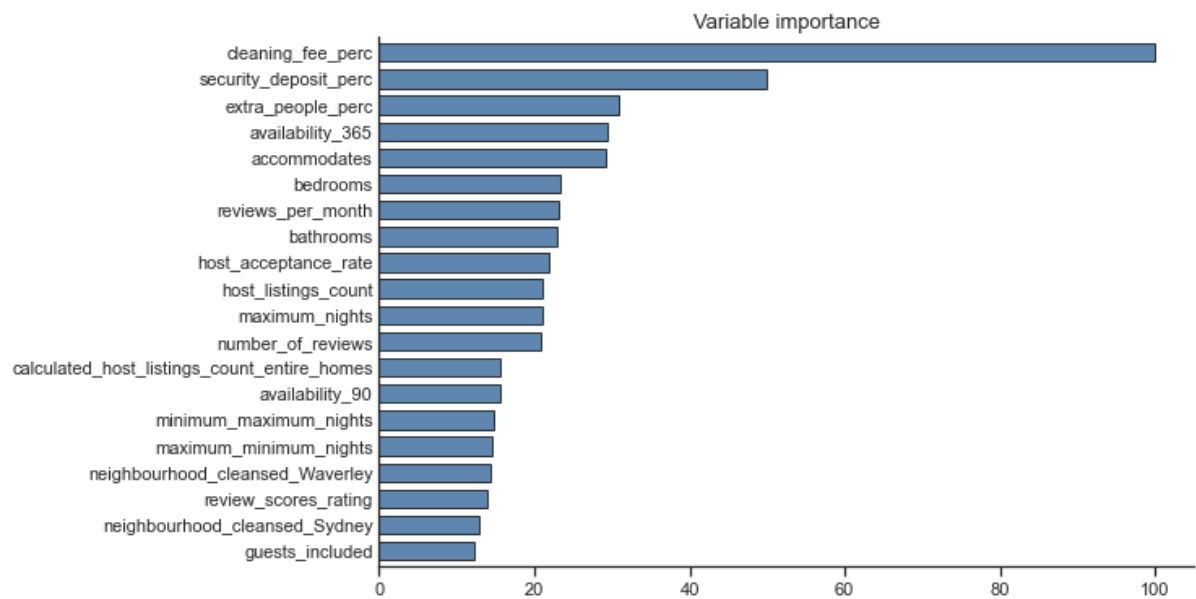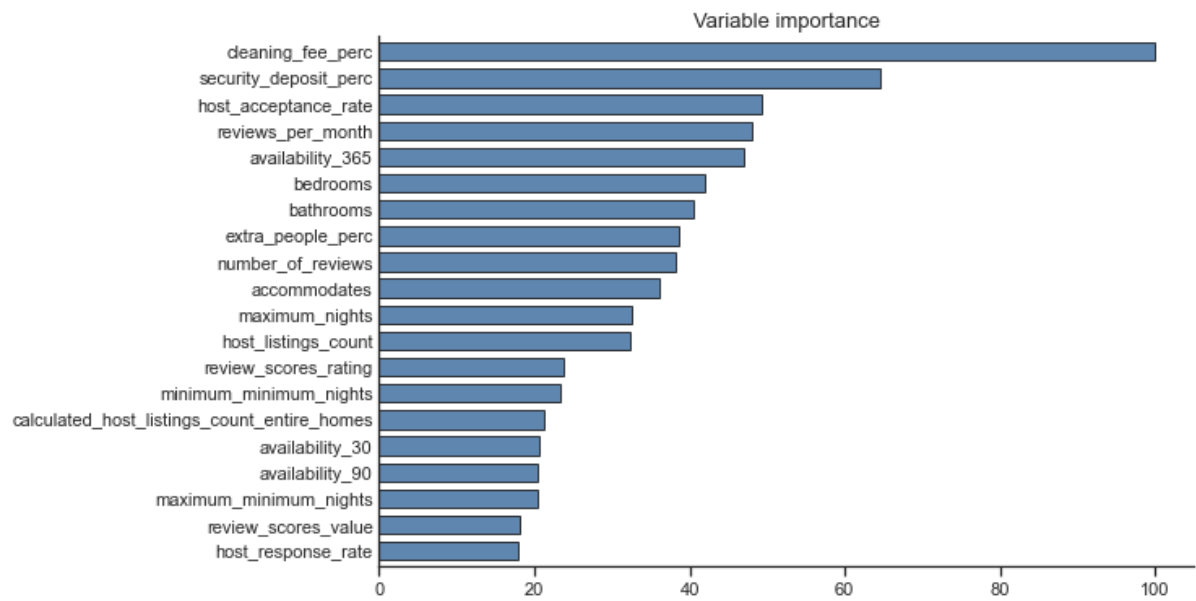*Figure L6 Variable Importance of LightGBM (without hyperparameter tuning)*

Figure L7 Variable Importance of LightGBM (after hyperparameter tuning)