Home Appraisals

Josh Christensen, Riley Millar and Greg Paulukaitis

4/16/2021

Section 1: Introduction and Problem Background

Professionally trained home appraisers determine the fair market value of a home based on multiple factors (i.e., square footage, number of bathrooms, number of bedrooms, etc.). We have access to a dataset consisting of various characteristics of 517 homes in Ames, Iowa. Additionally, we have access to the sale prices of 465 of these homes. Through our analysis of this data, we hope to determine if the characteristics of a home are a good indicator of sale price, and if so, which characteristics have the greatest impact on increasing sale price. We also hope to determine if the variability of sale price increases with the size of a home's living area. Finally, we wish to predict the sale prices for the 52 homes in our data set that have not yet been sold.

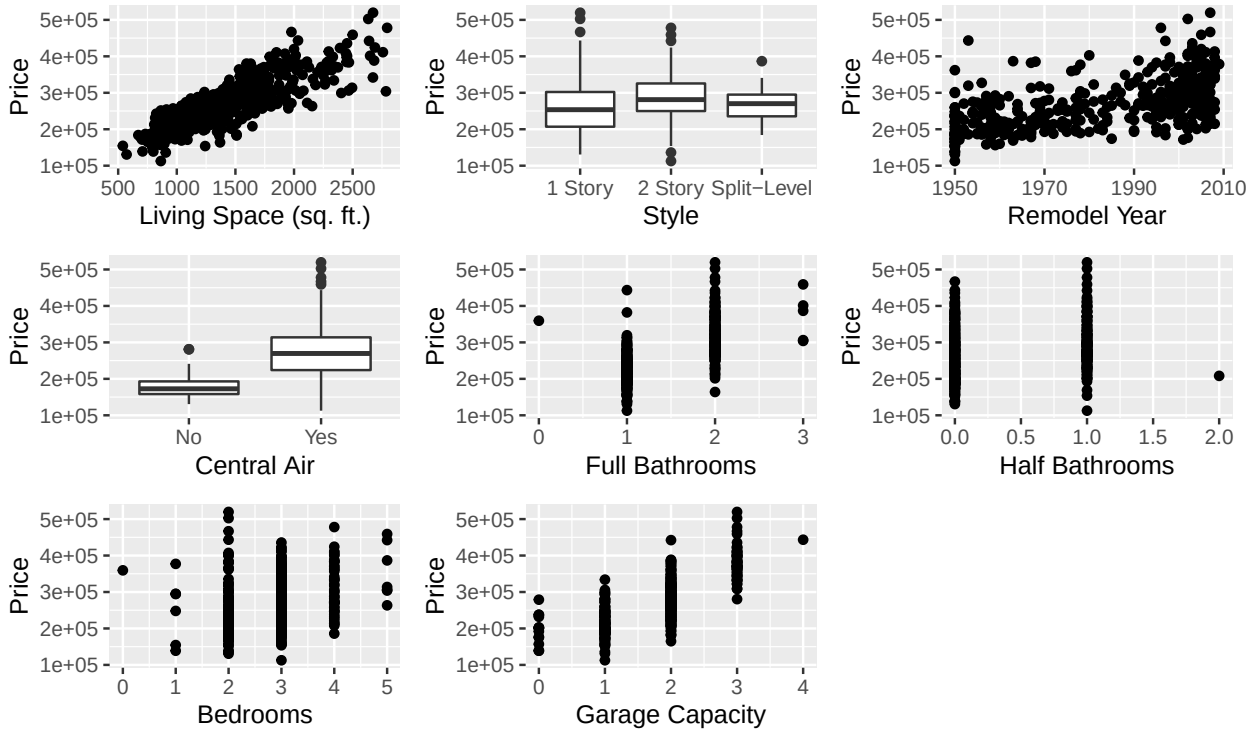

Figure 1: Exploratory Plots

Our dataset has two potential issues. One is that there is spatial correlation between house locations

(i.e., houses that are closer together tend to be similar in price). Another is the unequal variability in sale price with the size of the home. If we ignored these issues, our standard error calculations would be incorrect. This would in turn cause problems in our confidence intervals and prediction intervals for individual houses. We account for the spatial correlation issue by using a spatial correlation structure, and we account for the heteroskedasticity problem by using a separate exponential variance function. By addressing these issues, we will improve our ability to make inference and to accurately portray the variation of the data.

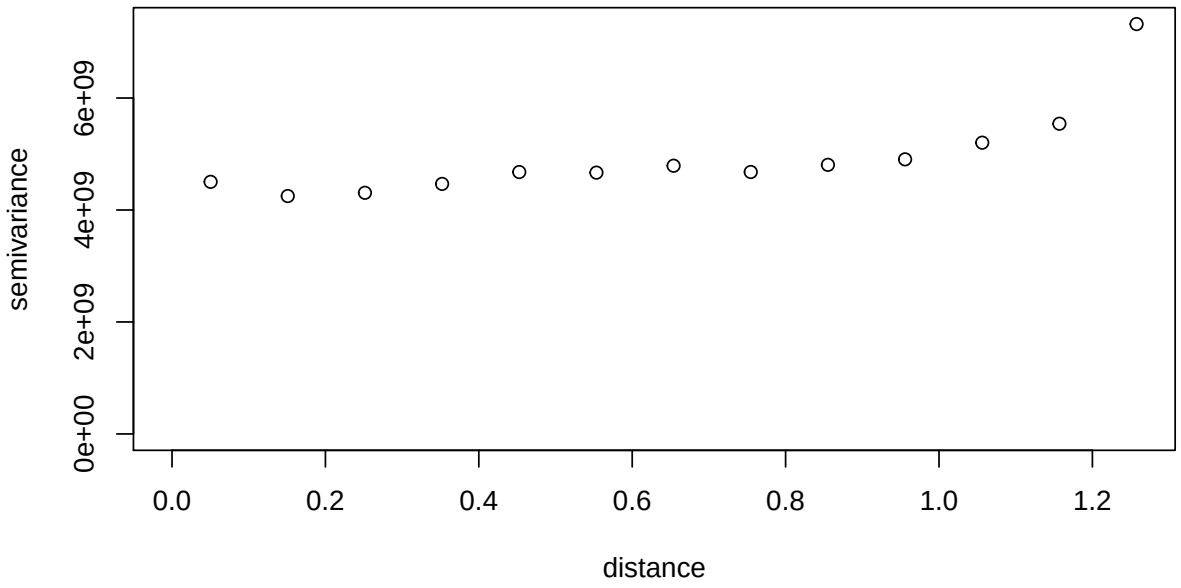

Figure 2: Price Variogram

Due to the spatial correlation of the houses, we will compare spherical, Gaussian, and exponential models to find which one has the best fit. To this we will add an exponential variance function to account for the heteroskedasticity. From there, we will check our model assumptions and verify the validity of our model using metrics such as pseudo R^2 and root mean square error (RMSE). We will then be able to use our final model to determine which house characteristics increase sale price and whether or not variability of sale price increases with overall living area size. We will also be able to predict the sale price for the 52 homes without a listed sale price.

Section 2: Statistical Model

We will use a heteroskedastic spatial multiple linear regression model to analyze the home appraisal data. We define the model for our data as $\boldsymbol{y} \sim MVN(\boldsymbol{X\beta}, \sigma^2\Sigma)$ where $\boldsymbol{\Sigma} = \boldsymbol{DRD}$. In our model, $\boldsymbol{y}$ represents the sale price of the home. $\boldsymbol{X}$ represents the design matrix which includes our explanatory variables as well as a column of ones, representing the intercept. We included above-ground living area (sq. ft.), house style (1 story, 2 stories, split level), year of remodel or construction, central air status (yes/no), number of full bathrooms above ground, number of half bathrooms above ground, number of bedrooms above ground

and the car capacity of the garage. Our parameters in this model are $\boldsymbol{\beta}$, σ^2, and the parameters of our covariance matrix: θ, ϕ. $\boldsymbol{\beta}$ is a vector of coefficients that quantifies the relationship between each respective explanatory variable and the response variable, in addition to setting the intercept. σ^2 represents the variance of the residuals from our model. The θ parameter quantifies the heteroskedasticity associated with house size and defines, through the exponential variance function, the diagonal weights of the $\boldsymbol{D}$ matrix in our covariance matrix decomposition. ϕ defines the range of the spatial correlation. It is also used to calculate individual correlations within the spherical correlation function within the $\boldsymbol{R}$ matrix in our covariance matrix decomposition.

Our model depends on four assumptions. The first is that all quantitative variables have a linear relationship with the response variable. For our model the quantitative variables include living space, year remodeled, number of full and half bathrooms, number of bedrooms and car capacity. The second is that each sale price is independent of the other sale prices, after accounting for spatial correlation. In other words, our residuals must be independent. While we do not assume spatial independence between sale prices, the validity of our model does depend on capturing the spatial correlation with our correlation function. The third assumption is that the residuals are normally distributed. The last assumption is that our $\boldsymbol{\Sigma}$ covariance matrix multiplied by σ^2 accurately represents the variability of the data. In other words, while our response is not assumed to have constant variance, we would like our covariance matrix to account for the changes in variance so that the underlying σ^2 is still constant. This is equivalent to constant variance of the residuals.

Section 3: Model Validation

We tested the linearity assumption using added-variable plots, which regress both the response and the explanatory variable against all other variables in the model and then display the resulting fitted values plotted against one another. This allows us to evaluate the relationship between the explanatory variable and the response, while accounting for the effects of all other variables in the model. The added-variable plots displayed below show no clear deviations from a linear relationship.

We believe the assumption of independence to be reasonable due to the lack of connection between houses. Intuitively, one house having two full bathrooms does not dictate how many full bathrooms the house across town will have. A potential breach of independence could be if multiple houses were sold by the same realtor or appraised by the same appraiser and there were differences between realtors and appraisers. We do not expect significant effects from either of these sources, in part because of our large sample size. Most other effects, such as neighborhood, proximity to entertainment or medical care or quality of local schools should be captured by our spatial correlation. Additionally, while we do not assume independence between sale prices of homes that are close together, we do assume that the spherical correlation structure captures the correlation. This is evident in the variogram of the decorrelated residuals shown below.

We checked the assumption of normally distributed residuals by creating a histogram of the standardized decorrelated residuals. The histogram appears to show a normal shape, unimodal with most points within 3 deviations from the 0. Additionally, there are no evident outliers. We also performed a Kolmogorov-Smirnov hypothesis test for normality. Given our p-value of 0.9844 we failed to reject the null hypothesis, therefore concluding that the standardized residuals came from a normal distribution.

The final assumption of constant variance in our residuals was checked by plotting the standardized residuals against the fitted values from our model. The scatterplot of standardized residuals against the fitted

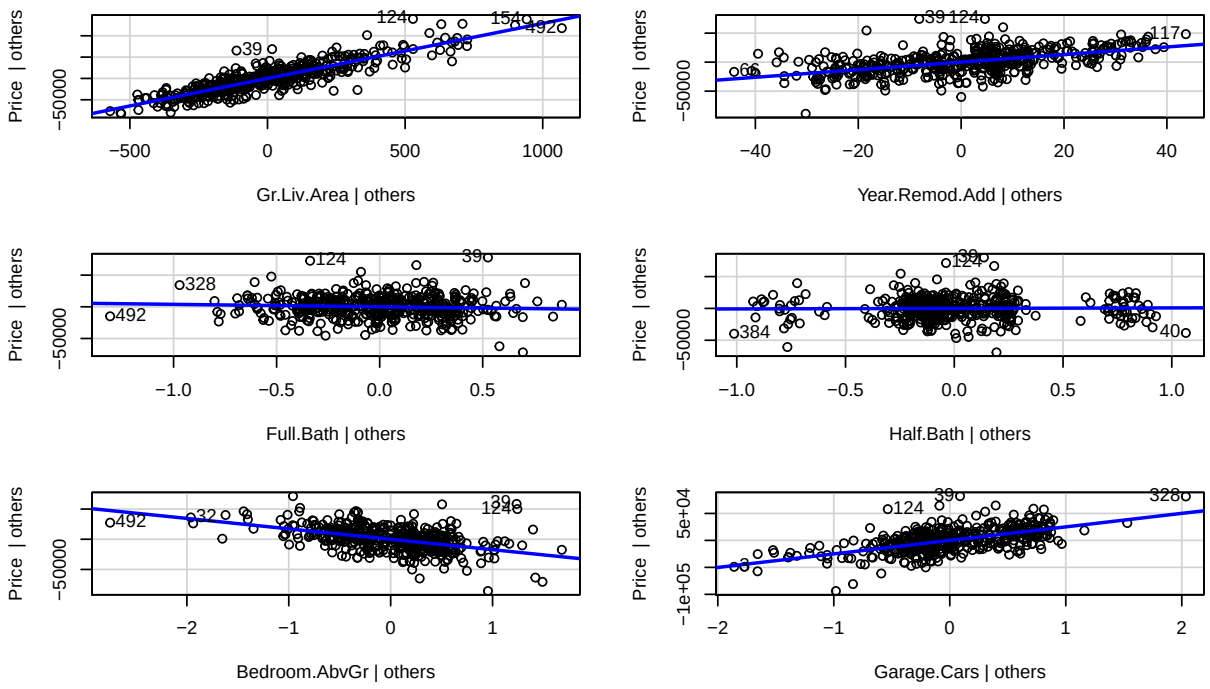

Figure 3: Added-Variable Plots

values shows constant variance.

We calculated pseudo R^2 as 0.93, indicating that roughly 93% of the variation in sale price of homes can be explained using the explanatory variables included in our model. The root mean square error is $17,769.39. This is moderately small relative to our data, which confirms that our model fits the data well.

We evaluate our predictions using root prediction mean square error (RPMSE), coverage, prediction interval width and bias. We evaluate these metrics using Monte Carlo cross-validation. The average of each metric is reported in the table below along with the histograms of RPMSE and interval width.

Table 1: Prediction Diagnostics

	Means
Bias	227.077
RPMSE	13974.718
Coverage	0.960
Width	53474.297

Given our negligible bias and relatively low RPMSE we find our predictions to be very accurate. Our coverage is also very close to the expected 0.95. Our prediction interval width indicates that we will generally construct intervals with a margin of error smaller than $13,974.72. This allows us to predict appraisals with a fair amount of accuracy.

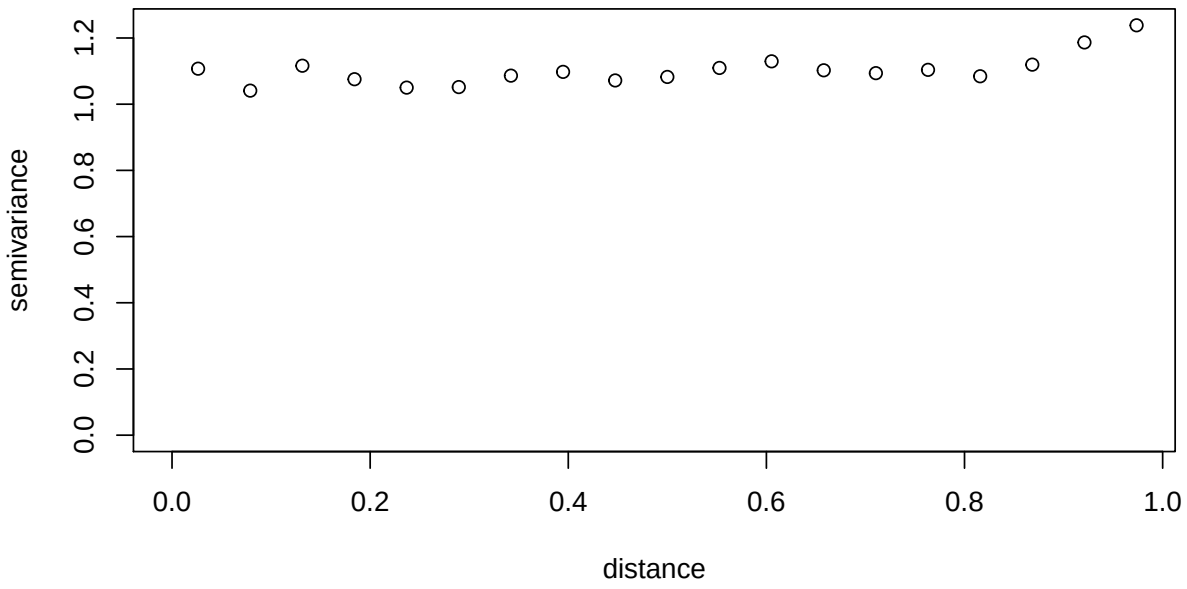

Figure 4: Residual Variogram

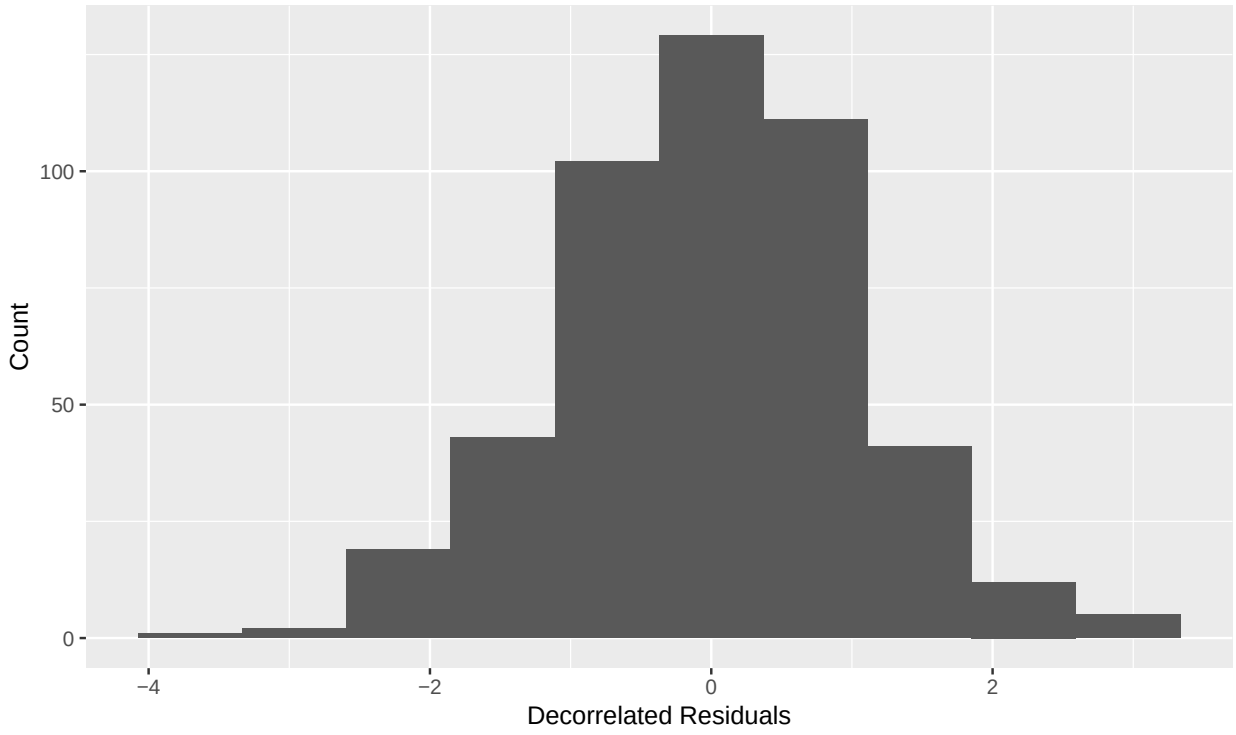

Figure 5: Decorrelated Residual Histogram

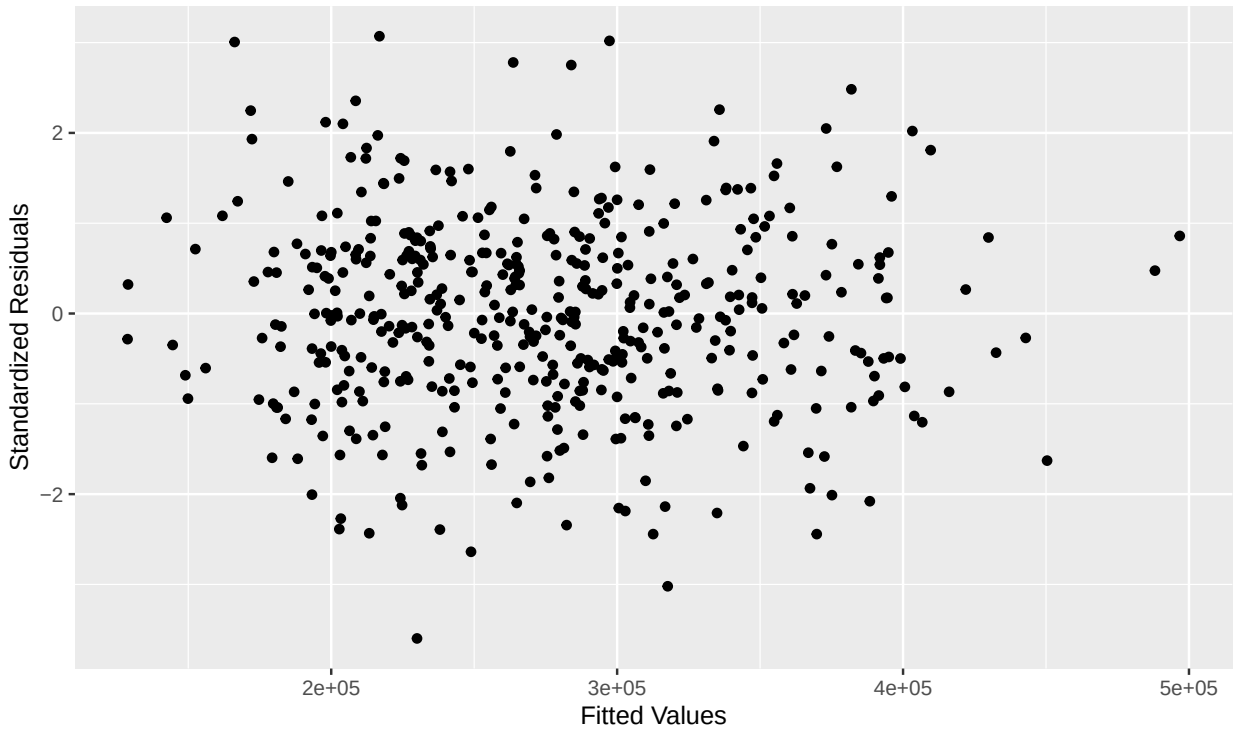

Figure 6: Standardized Residuals vs. Fitted Values

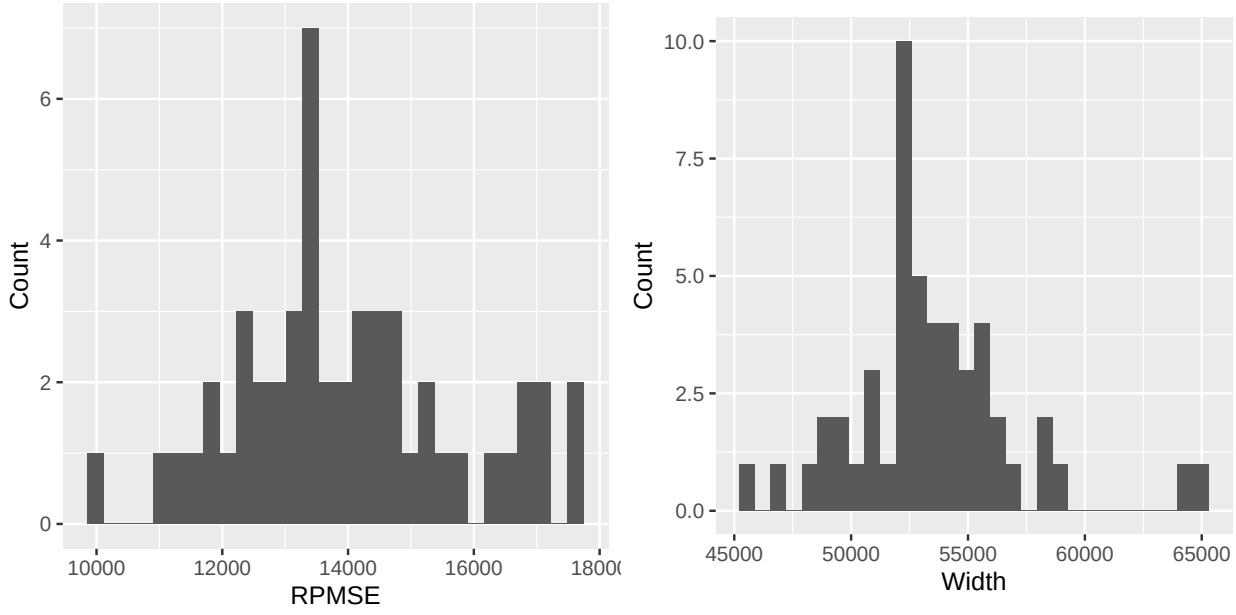

Figure 7: Prediction Evaluation Plots

Section 4: Analysis Results

The resulting pseudo R^2 value from our model was 0.93 which means that about 93% of the variability in the sale prices can be explained by our data's given house characteristics. Our average RPMSE value was 1.397472×10^4 which indicates that the average difference between our appraisal and the actual sale price is \$13,974.72. Our high pseudo R^2 and relatively low RPMSE means that home characteristics do a decent job at explaining sale price.

Since our model proved to be a good fit of the data, we were able to conclude which factors increase a house's sale price. From our analysis, we found that homes with more above-ground living space, recent construction or remodeling dates, larger garage capacities, and the presence of central air conditioning will lead to greater house values.

Table 2: Coefficient Table

	Estimate	Lower	Upper
Intercept	\$-1,324,753.83	\$-1,425,805.17	\$-1,223,702.48
Living Area	\$124.51	\$117.26	\$131.75
2 Story	\$-43,102.17	\$-46,483.08	\$-39,721.26
Split-Level	\$716.84	\$-2,998.10	\$4,431.78
Construction/Remodel Year	\$714.27	\$662.24	\$766.30
Central Air	\$21,555.52	\$17,404.75	\$25,706.29
Full Bath	\$-2,334.69	\$-5,572.54	\$903.17
Half Bath	\$461.76	\$-2,474.01	\$3,397.54
Bedrooms	\$-15,440.29	\$-17,217.71	\$-13,662.86
Garage Capacity	\$22,866.23	\$21,096.83	\$24,635.63

Although our model indicated that larger houses will have higher sales prices, we wanted to know if the variability in prices was consistent across all house sizes. We are 95% confident that the θ variance function parameter is between 6×10^{-4} and 8×10^{-4}. Since this interval is positive, we can conclude that there is greater variability in sale price with larger house sizes.

We used our model to predict the selling price of all of the houses in our dataset that did not have a selling price listed. This predictions are displayed in the heat map below, as well as the following table.

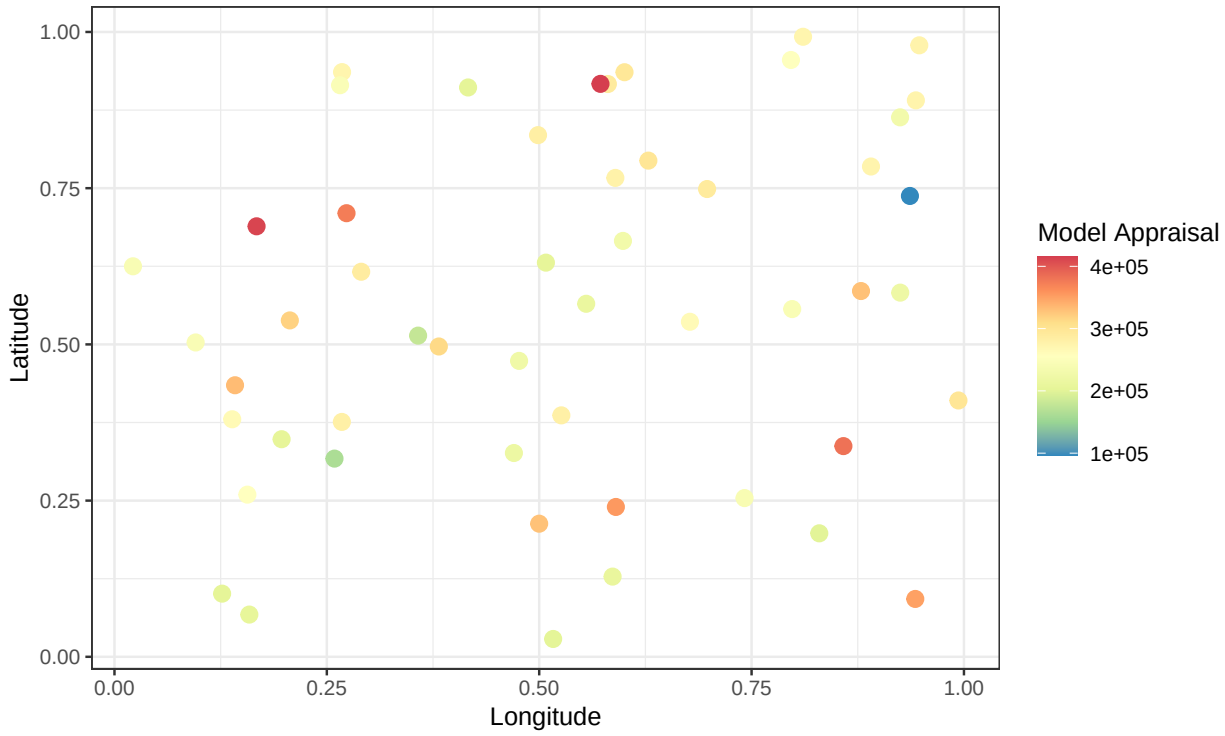

Figure 8: Model Appraisals

Table 3: Model Appraisals with Home Characteristics

Living Area	Style	Remodel Year	Central Air	Full Bath	Half Bath	Bedrooms	Garage Capacity	Model Appraisal
1144	1Story	1960	Y	1	0	3	1	$228,127.95
2855	2Story	2000	Y	2	1	4	3	$411,535.69
1114	1Story	2004	Y	1	1	3	0	$209,994.50
1576	SLvl	1961	Y	1	0	4	2	$264,610.26
1478	1Story	1992	Y	2	0	3	2	$272,994.82
1483	1Story	2001	Y	1	1	1	2	$332,654.77
1099	1Story	2006	Y	1	1	3	1	$241,238.36
1762	2Story	2002	Y	2	1	3	2	$287,504.30
936	1Story	1969	Y	1	0	2	1	$200,351.66
1437	1Story	1987	Y	1	1	3	2	$298,148.03
1732	2Story	1985	Y	2	1	3	2	$295,499.51
1839	2Story	1998	Y	2	1	4	2	$280,868.38
2554	2Story	1998	Y	1	1	3	2	$381,985.20
1264	1Story	1960	Y	1	0	3	2	$240,010.01
1610	1Story	2001	Y	2	0	3	2	$313,352.83
985	SLvl	1977	Y	2	0	3	1	$202,055.94
1294	1Story	1991	Y	2	0	3	2	$281,300.55
1086	2Story	1950	N	1	0	2	2	$161,786.71

Living Area	Style	Remodel Year	Central Air	Full Bath	Half Bath	Bedrooms	Garage Capacity	Model Appraisal
1478	1Story	1957	Y	1	1	3	2	$277,870.83
808	1Story	1995	N	1	0	1	1	$202,255.20
1877	2Story	1974	Y	2	1	4	2	$244,353.73
960	1Story	1958	Y	1	0	3	2	$212,498.76
1172	1Story	1998	Y	1	0	3	1	$254,032.58
1114	1Story	2004	Y	1	1	3	2	$274,498.54
1177	1Story	1984	Y	2	0	3	2	$264,280.60
1004	1Story	1970	Y	1	0	2	2	$242,704.94
1690	2Story	2001	Y	2	1	3	2	$297,911.08
1214	1Story	1965	Y	1	0	2	2	$273,666.28
1288	2Story	2007	Y	1	0	3	1	$227,631.65
1624	1Story	1995	Y	2	0	2	3	$350,245.04
1103	1Story	1980	Y	1	0	2	2	$275,035.85
1736	1Story	2008	Y	2	0	3	3	$354,878.38
1686	1Story	1980	Y	2	0	3	2	$329,487.03
1098	1Story	1955	Y	1	0	3	1	$203,704.05
1050	1Story	1956	Y	1	0	2	1	$205,886.44
2495	2Story	1993	Y	2	1	4	2	$328,734.70
2064	2Story	2005	Y	2	1	4	2	$318,004.05
1032	1Story	1950	N	1	0	2	1	$178,117.60
1344	2Story	1997	Y	1	0	3	1	$222,840.00
1285	1Story	1977	Y	1	1	3	2	$276,230.28
1638	2Story	1998	Y	2	1	3	2	$290,128.16
864	1Story	1972	Y	1	0	3	2	$220,330.85
2372	2Story	1965	N	2	0	4	1	$283,680.30
334	1Story	1950	N	1	0	1	0	$96,046.05
1001	1Story	2007	Y	1	0	2	1	$241,940.01
912	1Story	1995	Y	1	0	2	1	$217,639.11
1122	1Story	2005	Y	1	0	2	2	$289,277.93
1080	1Story	2005	Y	1	0	3	0	$203,021.78
2020	1Story	2008	Y	2	0	3	3	$415,584.99
2156	1Story	2003	Y	2	0	3	2	$374,035.12
1536	2Story	2005	Y	2	1	3	2	$256,070.18
1138	SLvl	1958	Y	1	0	3	1	$203,846.17

Section 5: Conclusions

From our analysis, we found that the characteristics provided by our dataset do a good job of explaining house price. Of these characteristics the ones that are most closely tied to higher house values are larger above-ground living space, more recent constructions or remodels, larger garage capacities and the presence

of central air conditioning. We also found that sale prices were more variable for larger homes. These findings are valuable because they would allow sellers to estimate the value of their home without having to pay for a formal appraisal.

We suggest that this analysis be generalized using data outside of Ames, Iowa. This would allow us to determine if the relationships between variables that we observed are consistent in other locations. We also believe that an understanding of the type of neighborhood (e.g., suburban, urban, rural, etc.) would help us to more accurately assess the value of a house.

Appendix A: Analysis Code

```
# Include packages
library(ggplot2)
library(geoR)
library(nlme)
library(car)
library(magrittr)
library(gridExtra)
library(knitr)
library(formattable)

# Include standardized gls residuals function
source("stdres.gls.R")

# Define function predictgls
source("predictgls.R")

# Read in the data
home1 <- read.csv("https://mheaton.byu.edu/docs/files/Stat469/Topics/3%20-%20SpatialCorrelation/3%20-%2
    header = TRUE)

# No NAs
home <- na.omit(home1)

# Convert factor variables to factor type
home$Central.Air <- factor(home$Central.Air)
home$House.Style <- factor(home$House.Style)

# Exploratory plots of every explanatory variable's relationship
# to the response variable
p1 <- ggplot(home, aes(Gr.Liv.Area, Price)) + geom_point() + labs(x = "Living Space (sq. ft.)")
p2 <- ggplot(home, aes(House.Style, Price)) + geom_boxplot() + labs(x = "Style") +
    scale_x_discrete(labels = c("1 Story", "2 Story", "Split-Level"))
p3 <- ggplot(home, aes(Year.Remod.Add, Price)) + geom_point() + labs(x = "Remodel Year") +
    scale_x_continuous(n.breaks = 4)
p4 <- ggplot(home, aes(Central.Air, Price)) + geom_boxplot() + labs(x = "Central Air") +
    scale_x_discrete(labels = c("No", "Yes"))
p5 <- ggplot(home, aes(Full.Bath, Price)) + geom_point() + labs(x = "Full Bathrooms")
p6 <- ggplot(home, aes(Half.Bath, Price)) + geom_point() + labs(x = "Half Bathrooms")
p7 <- ggplot(home, aes(Bedroom.AbvGr, Price)) + geom_point() + labs(x = "Bedrooms")
p8 <- ggplot(home, aes(Garage.Cars, Price)) + geom_point() + labs(x = "Garage Capacity")
```

```r
# Arrange plots in a grid
grid.arrange(p1, p2, p3, p4, p5, p6, p7, p8, ncol = 3, nrow = 3)

# Using variog from library(geoR)
variogram <- variog(coords = home[, 2:3], data = home$Price, messages = FALSE)

# Plot the variogram
plot(variogram)

# Independent MLR
home.lm <- lm(Price ~ . - Lon - Lat, data = home)

# Linearity
avPlots(home.lm, terms = ~. - House.Style - Central.Air)

# Refit the best spatial model and add the weights
home.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style + Year.Remod.Add +
    Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr + Garage.Cars,
    data = home, correlation = corSpher(form = ~Lon + Lat, nugget = TRUE),
    weights = varExp(form = ~Gr.Liv.Area), method = "ML")

# Store the decorrelated residuals
sres <- stdres.gls(home.gls)

## Using variog from library(geoR)
variogram2 <- variog(coords = home[, 2:3], data = sres, breaks = seq(0,
    1, length.out = 20), max.dist = 1, messages = FALSE)

# Plot the variogram
plot(variogram2)

# Draw a histogram of the decorrelated residuals
ggplot() + geom_histogram(mapping = aes(x = sres), bins = 10) + xlab("Decorrelated Residuals") +
    ylab("Count")

# Run ks-test for normality
ks <- ks.test(sres, "pnorm")

# Scatterplot of the fitted values vs. decorrelated residuals
ggplot(mapping = aes(fitted(home.gls), sres)) + geom_point() + xlab("Fitted Values") +
    ylab("Standardized Residuals")

# Calculate Pseudo R2
```

```r
R2 <- cor(home$Price, home.gls$fitted)^2

# Calculate the RMSE
rmse <- (home$Price - home.gls$fitted)^2 %>% mean() %>% sqrt()

n.cv <- 50  #Number of CV studies to run
n.test <- round(0.2 * nrow(home), 0)
rpmse <- rep(x = NA, times = n.cv)
bias <- rep(x = NA, times = n.cv)
wid <- rep(x = NA, times = n.cv)
cvg <- rep(x = NA, times = n.cv)
for (cv in 1:n.cv) {
    ## Select test observations
    test.obs <- sample(x = 1:nrow(home), size = n.test)

    ## Split into test and training sets
    test.set <- home[test.obs, ]
    train.set <- home[-test.obs, ]

    ## Fit a gls() using the training data
    train.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style + Year.Remod.Add +
        Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr + Garage.Cars,
        data = train.set, correlation = corSpher(form = ~Lon + Lat,
            nugget = TRUE), weights = varExp(form = ~Gr.Liv.Area),
        method = "ML")

    ## Generate predictions for the test set
    my.preds <- predictgls(train.gls, newdframe = test.set)

    ## Calculate bias
    bias[cv] <- (my.preds[, "Prediction"] - test.set[["Price"]]) %>%
        mean()

    ## Calculate RPMSE
    rpmse[cv] <- (test.set[["Price"]] - my.preds[, "Prediction"])^2 %>%
        mean() %>% sqrt()

    ## Calculate Coverage
    cvg[cv] <- ((test.set[["Price"]] > my.preds[, "lwr"]) & (test.set[["Price"]] <
        my.preds[, "upr"])) %>% mean()

    ## Calculate Width
    wid[cv] <- (my.preds[, "upr"] - my.preds[, "lwr"]) %>% mean()
```

```
}

# Table of cross-validation prediction diagnostics
kable(data.frame(Means = c(mean(bias), mean(rpmse), mean(cvg), mean(wid)),
    row.names = c("Bias", "RPMSE", "Coverage", "Width")), caption = "Prediction Diagnostics",
    digits = 3)

# RPMSE histogram
p1 <- ggplot(mapping = aes(rpmse)) + geom_histogram(bins = 30) + xlab("RPMSE") +
    ylab("Count") + theme(aspect.ratio = 1)

# Width of prediction intervals histogram
p2 <- ggplot(mapping = aes(wid)) + geom_histogram(bins = 30) + xlab("Width") +
    ylab("Count") + theme(aspect.ratio = 1)

# Combine plots into a grid
grid.arrange(p1, p2, ncol = 2)

# Store intervals as an object
ints <- intervals(home.gls)

# Reorganize columns and store as a dataframe
frame <- as.data.frame(ints$coef)[c(2, 1, 3)]

# Add desired column names
colnames(frame) <- c("Estimate", "Lower", "Upper")

# Add desired row names
rownames(frame) <- c("Intercept", "Living Area", "2 Story", "Split-Level",
    "Construction/Remodel Year", "Central Air", "Full Bath", "Half Bath",
    "Bedrooms", "Garage Capacity")

# Reformat as dollars
frame$Estimate <- currency(frame$Estimate)
frame$Lower <- currency(frame$Lower)
frame$Upper <- currency(frame$Upper)

# Display in kable
kable(frame, caption = "Coefficient Table", digits = 2)

# Estimate and confidence interval for theta
thta <- intervals(home.gls)$varStruct[1, c(2, 1, 3)]
```

```r
# Create prediction data.frame
preddf <- home1[is.na(home1$Price), ]

# Predict
preds <- predictgls(home.gls, newdframe = preddf)

# Plot all of the predictions and original
ggplot() + geom_point(data = preds, aes(x = Lon, y = Lat, color = Prediction),
    size = 3) + scale_color_distiller(palette = "Spectral", na.value = NA) +
    theme_bw() + labs(x = "Longitude", y = "Latitude", color = "Model Appraisal")

# Produce a table of the predicted values along with the
# respective covariates
preds[, "Prediction"] <- currency(preds[, "Prediction"])

kable(preds[, -c(1:3, 13:15)], row.names = FALSE, col.names = c("Living Area",
    "Style", "Remodel Year", "Central Air", "Full Bath", "Half Bath",
    "Bedrooms", "Garage Capacity", "Model Appraisal"), caption = "Model Appraisals with Home Characteris
```

Appendix B: Exploratory and Model Selection Code

```
# Independent MLR
home.lm <- lm(Price ~ . - Lon - Lat, data = home)

# Linearity
avPlots(home.lm)

# Spatial correlation Pull the residuals
resids <- resid(home.lm)

# Heat map of the independent linear model residuals
ggplot(data = home, aes(x = Lon, y = Lat, color = resids)) + geom_point() +
    scale_color_distiller(palette = "Spectral", na.value = NA)

# Using variog from library(geoR) for the independent MLR
variogram <- variog(coords = home[, 2:3], data = home$Price, breaks = seq(0,
    1, length.out = 20), max.dist = 1)

# Plot the variogram
plot(variogram)

# Equal variance assumption on independent MLR
ggplot(mapping = aes(fitted(home.lm), resid(home.lm))) + geom_point() +
    xlab("Fitted Values") + ylab("Standardized Residuals")

# Fit an exponential spatial corelation model
homeexp.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style + Year.Remod.Add +
    Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr + Garage.Cars,
    data = home, correlation = corExp(form = ~Lon + Lat, nugget = TRUE),
    method = "ML")

# Fit a Gaussian spatial correlation model
homegaus.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style + Year.Remod.Add +
    Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr + Garage.Cars,
    data = home, correlation = corGaus(form = ~Lon + Lat, nugget = TRUE),
    method = "ML")

# Fit a spherical spatial correlation model
homesphere.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style +
    Year.Remod.Add + Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr +
    Garage.Cars, data = home, correlation = corSpher(form = ~Lon +
    Lat, nugget = TRUE), method = "ML")
```

```
AIC(homeexp.gls, homegaus.gls, homesphere.gls)

# Refit the best spatial model and add the weights
home.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style + Year.Remod.Add +
    Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr + Garage.Cars,
    data = home, correlation = corSpher(form = ~Lon + Lat, nugget = TRUE),
    weights = varExp(form = ~Gr.Liv.Area), method = "ML")

AIC(home.gls, homesphere.gls)

homeinteractions.gls <- gls(model = Price ~ Gr.Liv.Area + House.Style +
    Year.Remod.Add + Central.Air + Full.Bath + Half.Bath + Bedroom.AbvGr +
    Garage.Cars + Bedroom.AbvGr:Full.Bath + Bedroom.AbvGr:Gr.Liv.Area +
    Full.Bath:Gr.Liv.Area + Gr.Liv.Area:House.Style + Garage.Cars:Gr.Liv.Area,
    data = home, correlation = corSpher(form = ~Lon + Lat, nugget = TRUE),
    weights = varExp(form = ~Gr.Liv.Area), method = "ML")

AIC(homeinteractions.gls, home.gls)

coef(home.gls)
confint(home.gls)

# Colinearity exploration
cormat <- cor(home[, c(4, 6, 8:11)])
corrplot(cormat, method = "number")

# Plot all of the predictions and original
ggplot() + geom_point(data = home, aes(x = Lon, y = Lat, color = Price)) +
    geom_point(data = preds, aes(x = Lon, y = Lat, color = Prediction)) +
    scale_color_distiller(palette = "Spectral", na.value = NA)

## Colinearity fiddling Refit the best spatial model with different
## covariates to see if we can make bedroom positive
homefiddle.gls <- gls(model = Price ~ House.Style + Year.Remod.Add +
    Central.Air + Bedroom.AbvGr + Garage.Cars, data = home, correlation = corExp(form = ~Lon +
    Lat, nugget = TRUE), weights = varExp(form = ~Gr.Liv.Area), method = "ML")

summary(homefiddle.gls)

# Plot of every variable against each other to explor colinearity
plot(~Price + Gr.Liv.Area + Year.Remod.Add + Full.Bath + Half.Bath +
    Bedroom.AbvGr + Garage.Cars, data = home)
```