

# Estimating rental housing prices using kriging interpolation

Geography 580 Advanced GIS | Spring 2012

Kelly Foley<sup>1</sup>, James Graham<sup>2</sup>, Mark Edwards<sup>3</sup>, Jarrod Olson<sup>4</sup>, Maria Lewis<sup>1</sup>, Lauren Dennis<sup>1</sup>

(1) Water Resources Policy and Management, Oregon State University (2) Geography, Oregon State University (3) Sociology, Oregon State University (4) Public Policy, Oregon State University



## INTRODUCTION

The rate of population growth in Corvallis has exceeded the city's ability to house everyone, creating what has become popularly known as "The Corvallis Housing Crisis." With Oregon State University growing, many organizations across the city are seeking to better understand the housing market. Willamette Neighborhood Housing Services (WNHS) is one of these organizations seeking to understand the rental market to better help low income families in need of housing. The purpose of this study was to understand how prices for rental housing in Corvallis vary geographically. This will help WNHS to better address the future needs of low income families in Corvallis in the face of population

Alexander Court Affordable Housing by Willamette Neighborhood Housing Services



growth. Specifically, WNHS hopes to utilize this information to place low-income families in suitable housing as well as plan for future development projects like Alexander Court.

## METHODS

### Data Collection

A number of online resources were harvested in an attempt to generate a comprehensive list of all vacant rental properties in March of 2012. The following data was gathered (as much as it was available): rental prices, rental addresses, number of bedrooms, and number of bathrooms. It would have been ideal to also gather square footage, but it was not available from most sources.

### Data Cleaning

Data was entered into a pre-existing shapefile of all addresses in the city. Data was normalized by bedroom rather than square foot because of data availability. Due to this imperfect normalization, rental values with the same geographic location had to be averaged. A portion of the data could not be georeferenced due to a number of factors. 124 addresses were entered.

### Processing

Data was visually inspected for outliers. Moran's I was calculated to test for spatial autocorrelation. Data frequencies, semivariogram, and trend analysis were explored. Following a literature review and data exploration, ordinary kriging with a spherical model and second order polynomial trend were fitted to the data to create a continuous rental price surface.

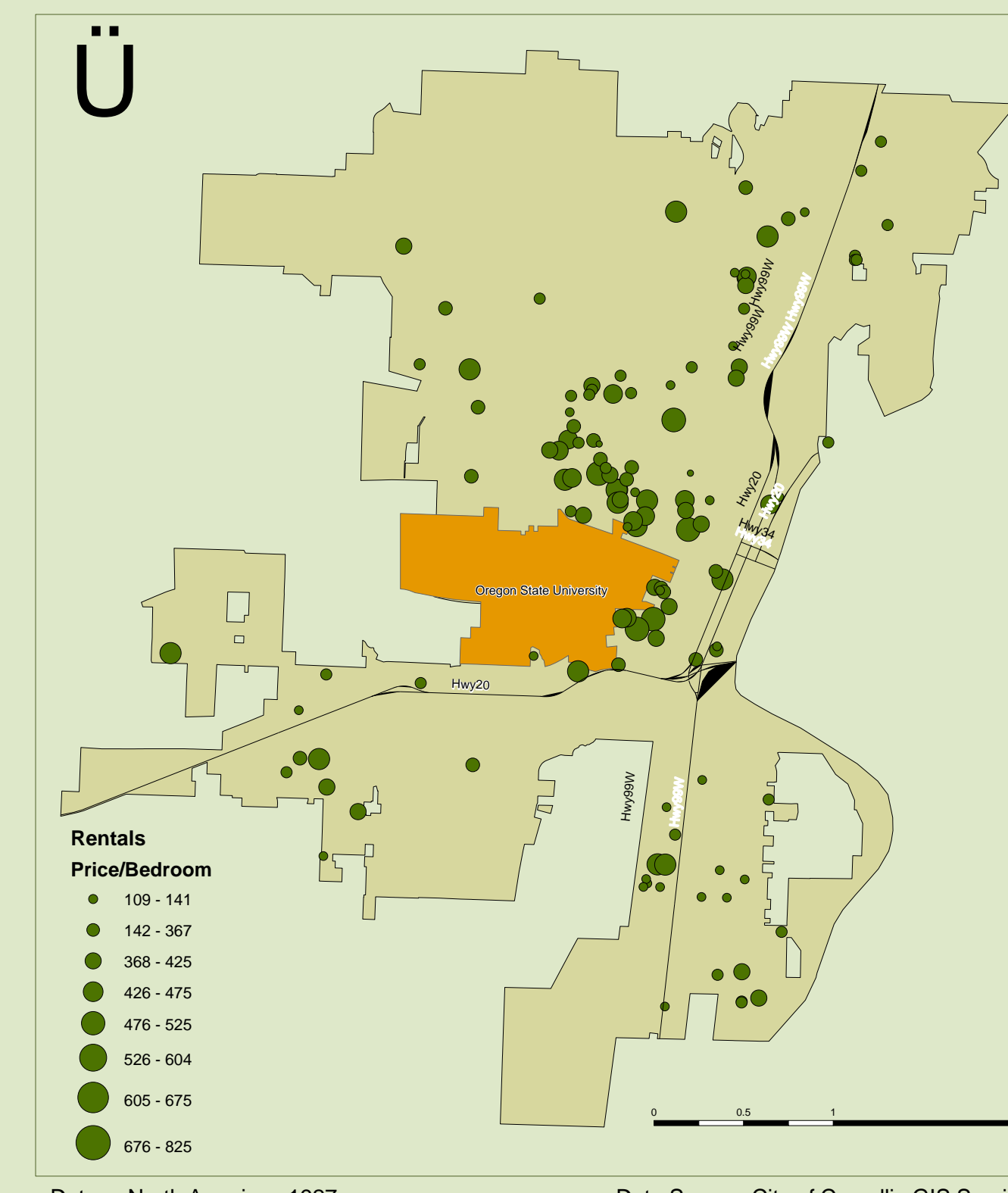
### Craigslist



STREET	Rent bed	V_COORD	X_COORD
202 NE 2ND ST	400	340712.0291	120804.804
25 NW ANLUN DR	465	351214.7033	1281173.3125
2190 NE POWDERHORN PL	417	352586.89743	1285153.89378
956 NW BROWNWOOD AVE	695	348777.376	1281818.1225
2265 NW HUNTINGBRO DR	496	348492.18785	1271103.8519
2125 NW HIGHLAND DR	338	348710.53285	1288948.58974
2130 NW JAMESON ST	328	348687.39283	1281163.94711
981 NW CIRCLE BLVD	600	348561.94181	1281189.35693
2935 NW CIRCLE BLVD	383	347946.5878	1275105
2410 NW JAMESON DR	487	347662	1272335.375
1380 NW FOREST DR	687	345882.77647	1275042.04336
1318 NW 23RD ST	483	345338.85628	127842.9825
1216 NW 23RD ST	388	345118.54783	1276565.20291
2809 NW GRANT AVE	410	345881.8125	1276029.9375
1525 NW 36TH ST	450	344745.5	1273288.8881
1824 NW 26TH ST	368.68899	344881.29718	1275881.82349
1800 NW FILLMORE AVE	804	343794.29831	1275843.19849
NW 26TH ST	388	343779.18927	1275248.3
2300 E.E. 1A	464	343329.12481	1275248.3

### Accuracy

The mean prediction error, root mean square, and root-mean-square standardized were all calculated for the predicted model surface. Additionally, cross validation was conducted using a subset of 30 random points from the 124 total points.



Datum: North American 1927  
Projection: Oregon Lambert Conformal Conic  
Spatial distribution of online rental data and price of data normalized to price per bedroom. Larger points represent higher priced units. The majority of data points are located in the center of the city.

## DISCUSSION

There were a number of sources of error contributing to the uncertainty surrounding the accuracy of the rental price predictions:

- The rental data was collected from a number of online databases, which may not have been accurate
- Data was normalized to bedrooms and not square footage
- Rental listings were not just for March, but for rental openings in June and into the summer. There may be error in that some locations may reflect summer prices rather than spring prices
- There may have been human error in data entry and data cleaning
- There may have been error in the pre-existing address database for the city of Corvallis because it was created by piecing together data from multiple agencies with unknown accuracy
- While kriging interpolation has been used for a number of different housing price studies, there have not been a lot of studies that utilize kriging for rental housing analysis to reference here

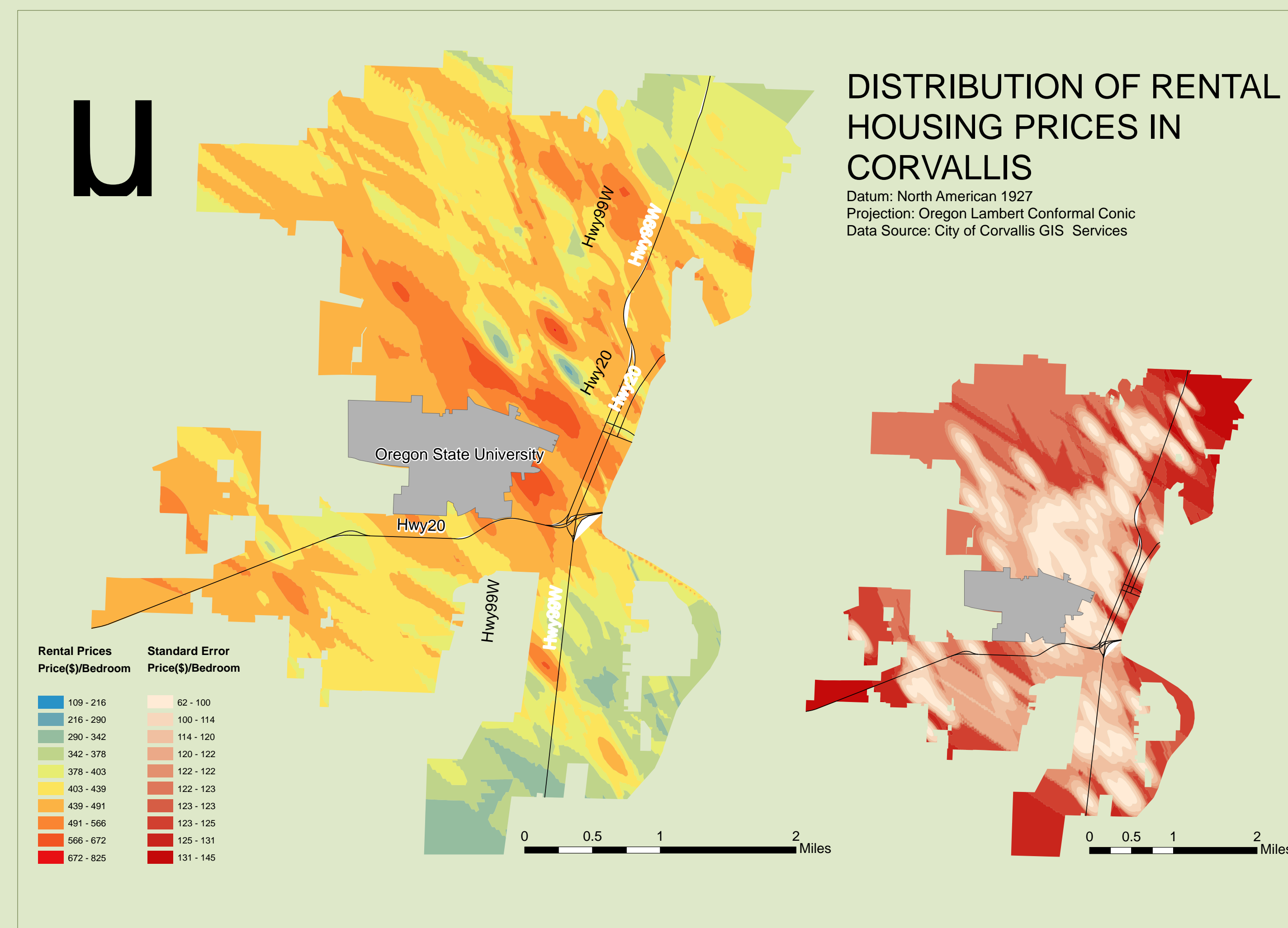
Overall, the accuracy and uncertainty associated with this interpolation is largely unknown. When using this interpolated surface, caution should be taken when considering different times during the year and different locations within the city.

## RESEARCH QUESTIONS

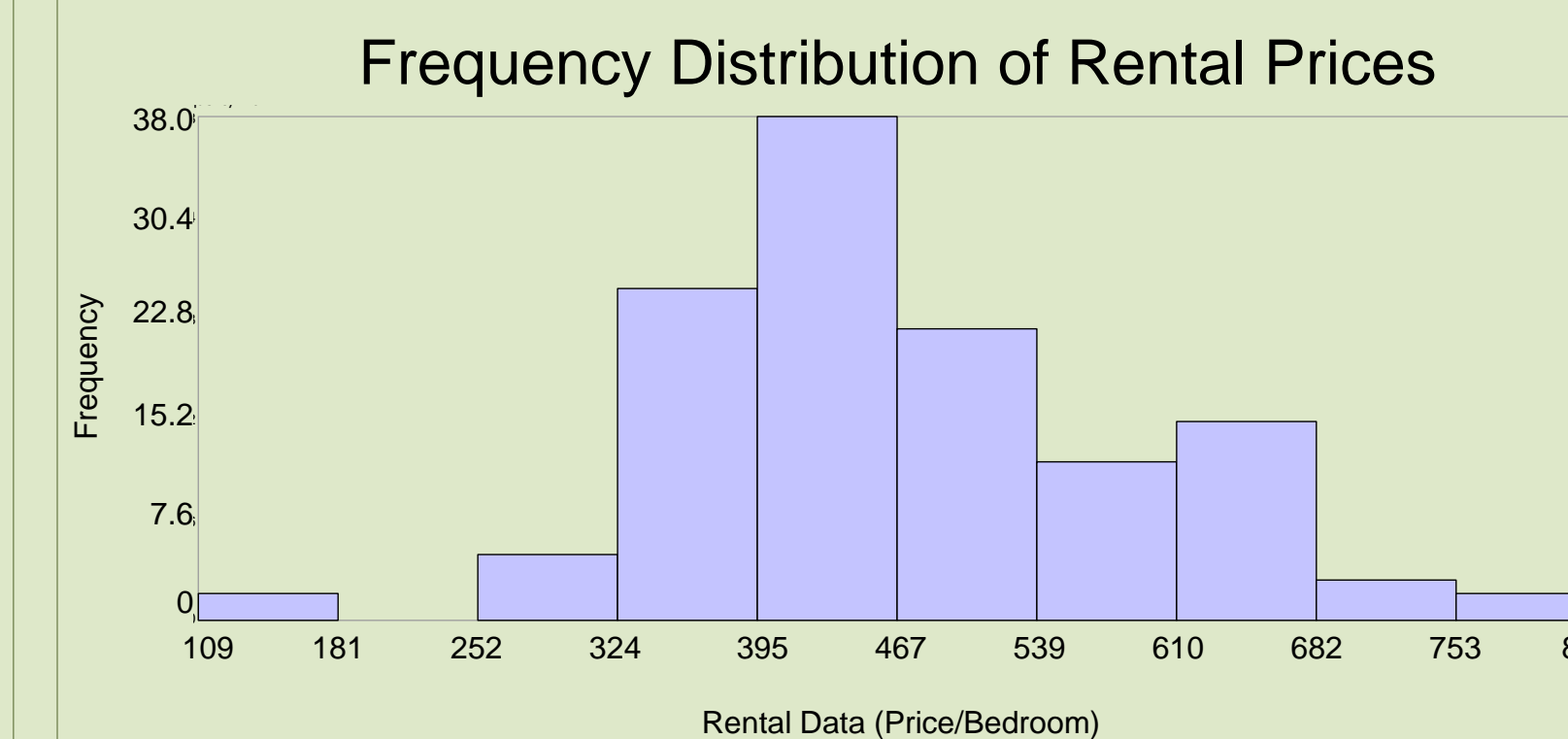
How do rental housing prices vary across the city of Corvallis?

- Is there spatial autocorrelation associated with the rental prices in Corvallis?
- What interpolation method could best estimate a continuous surface of rental prices in Corvallis?
- How accurate is an interpolated estimation of housing prices in Corvallis?

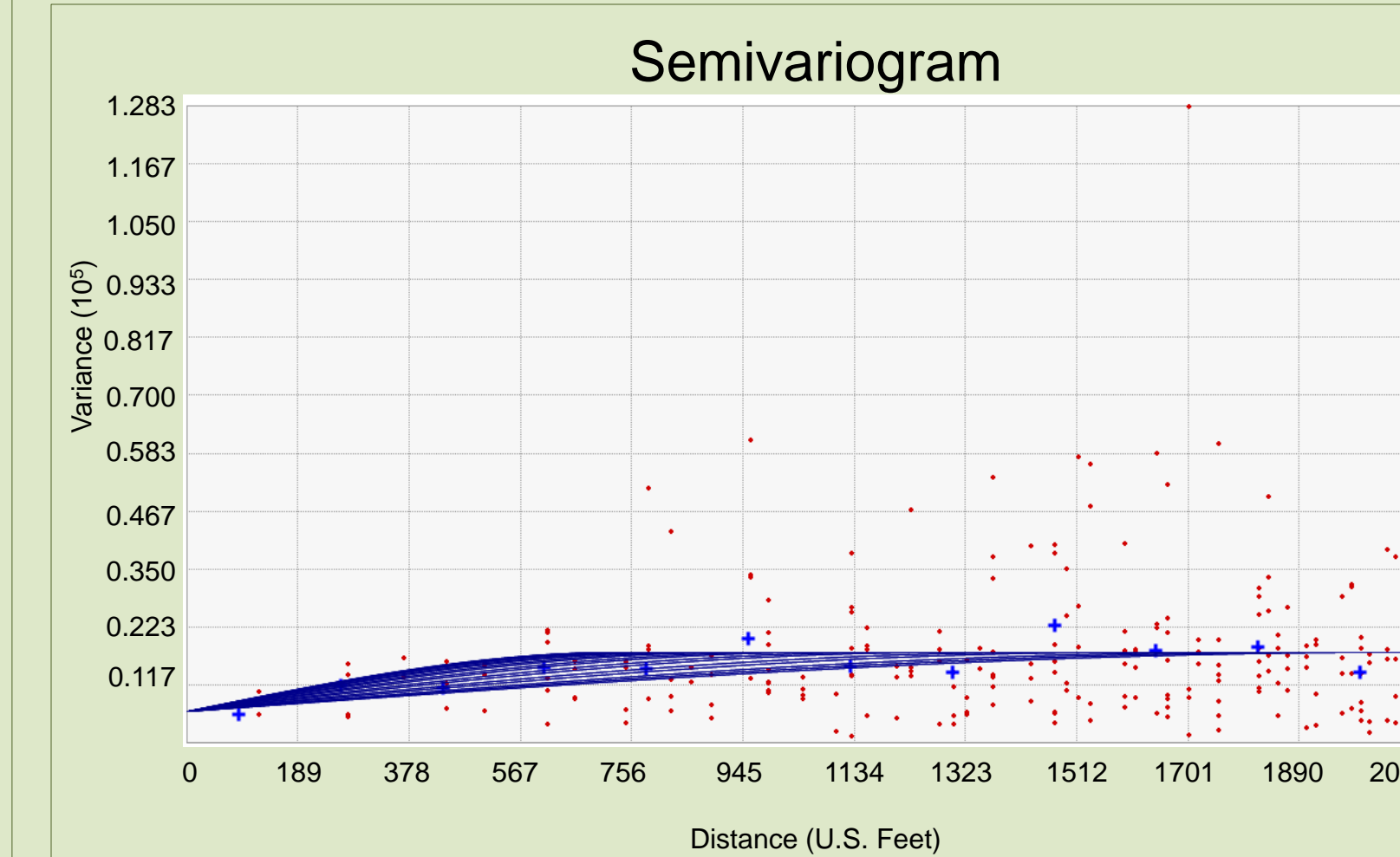
## RESULTS



DISTRIBUTION OF RENTAL HOUSING PRICES IN CORVALLIS  
Datum: North American 1927  
Projection: Oregon Lambert Conformal Conic  
Data Source: City of Corvallis GIS Services



The data is relatively normally distributed. Normal distribution was assumed for the analysis of rental data.

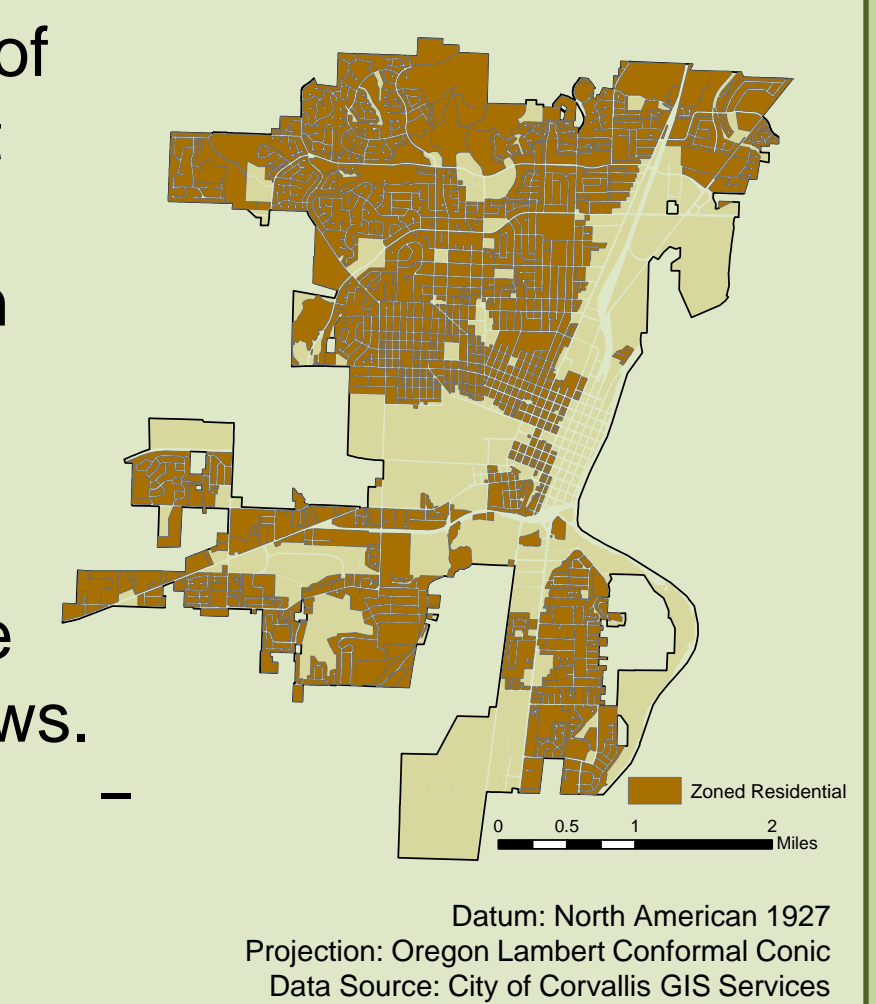


An ordinary kriging model with a spherical trend and a second-order polynomial trend adjustment was applied to the data.

## FUTHER STUDY

While the findings from the predicted surface seem reasonable given observations about housing in Corvallis, without further and repeated studies it is difficult to draw conclusions about the strength of the model. Future studies might include:

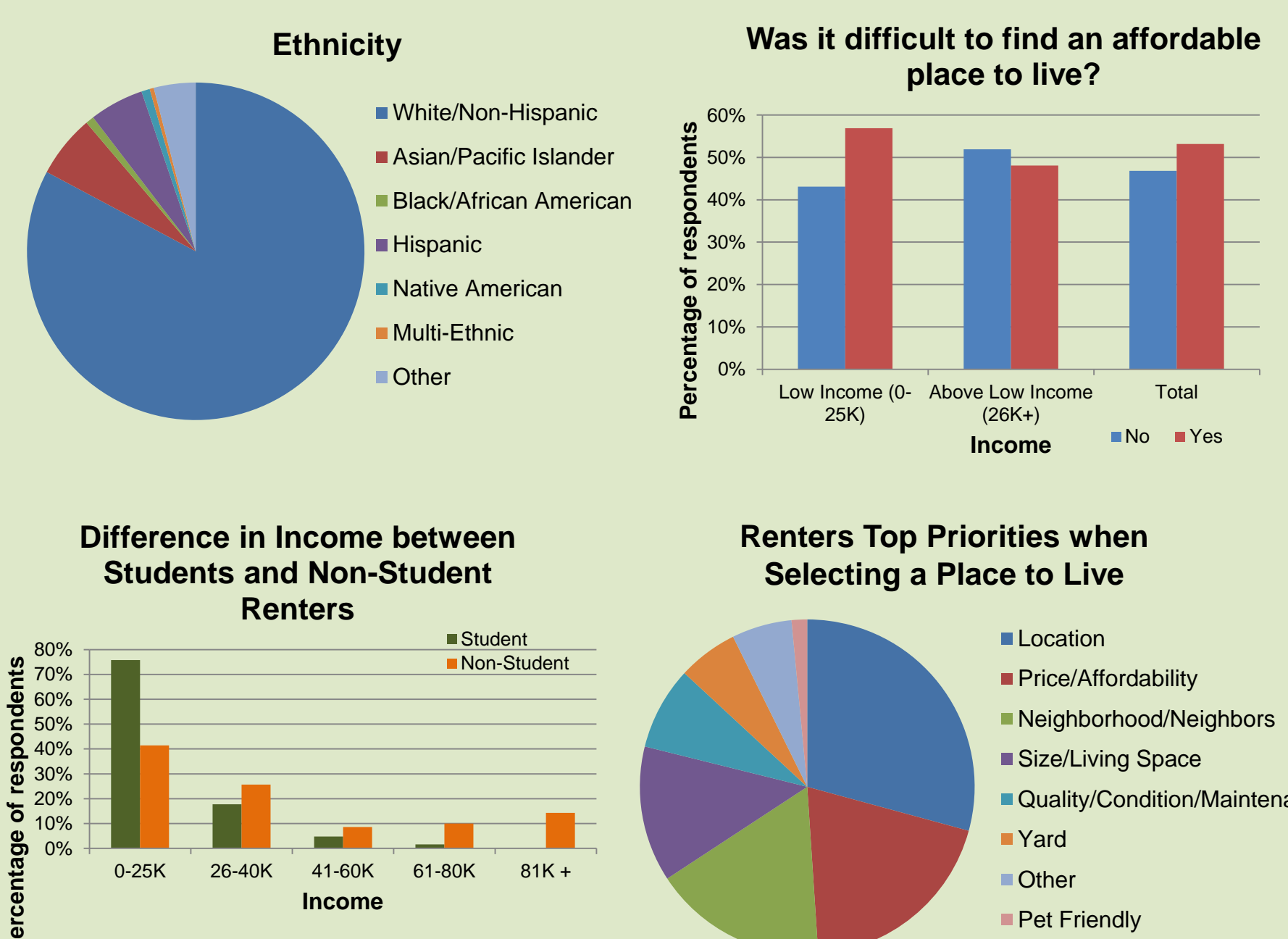
- Analyzing seasonal trends in rental prices (which may be important with the influx of students in the Fall)
- Considering trends over time as the university system grows.
- Utilizing different kriging methods



Datum: North American 1927  
Projection: Oregon Lambert Conformal Conic  
Data Source: City of Corvallis GIS Services

## HOUSING SURVEY RESULTS

A survey of Corvallis renters and homeowners was conducted by students at Oregon State University. 834 doors were knocked on, 270 people participated. 47% of respondents were homeowners and 53% were renters.



### Key Findings

- Rental price data was relatively normally distributed
- Moran's Index was 0.19,  $p=0.004$ , significantly positively autocorrelated
- Rental prices are highest to the north and northwest of Oregon State University and lowest in the southeast portion of the city
- Price per bedroom ranges from \$109-\$825, with an average price of \$467
- The standard error is lower (prediction surface more accurate) in the center of the city and higher in the corners of the city (southwest and northeast)
- The standard error is lowest where the points are located
- The standard error measurements indicate that the error in the prediction surface is relatively low.
- The cross validation indicates that the average error on the model was low, but the range was large. This is likely indicative of outliers.

### Error Measurements for prediction surface

	Standard Error		Cross Validation	
Mean Prediction Error	Root Mean Square	Root-Mean-Square Standardized	Mean Error	Range of Error
Should be close to zero	Smaller is better	Closest to 1 is best	Smaller is better	Narrow is better
-3.33	123.56	0.96	6.22	-318.93 - 351.61

## ACKNOWLEDGEMENTS

I would like to thank Jim Graham, Mark Edwards, Jen Bauer, Jim Moorefield and Willamette Neighborhood Housing Services, as well as everyone in Sociology 519 who helped collect the data to make this project possible.

## REFERENCES

Dubin, R. A. 1992. Spatial autocorrelation quality and neighborhood. *Regional Science and Urban Economics* 22:433-452.  
 Martinez, M. G., J. M. M. Lorenzo, and N. G. Rubio. 2000. Kriging Methodology for Regional Economic Analysis: Estimating the Housing Price in Albacete. *International Advances in Economic Research* 6 (3):438-450.  
 McCluskey, W. J., W. G. Deddis, and I. G. Lamont. 1999. The Application of Surface Generated Interpolation Models for the Prediction of Residential Property Values. *Journal of Property Investment & Finance* 18 (2):162-176.  
 Montero, J. M., and B. Larraz. 2011. Interpolation methods for geographical data: Housing and commercial establishment markets. *Journal of Real Estate Research* 33 (2):233-244.  
 Olmo, J. C. 1995. Spatial Estimation of Housing Prices and Locational Rents. *Urban Studies* 32 (8):1331-1344.  
 Pagourtzi, E., V. Assimakopoulos, T. Hatzichristos, and N. French. 2003. Real estate appraisal: a review of valuation methods. *Journal of Property Investment & Finance* 21 (4):383-401.