Place, Policy and Parity: Examining and Visualizing Spatial and Socioeconomic Contributions to Hospital Charge Markup

Rebecca W Stubbs

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Arts

University of Washington

2017

Committee:

John Mark Ellis

Sarah A Elwood-Faustino

Program Authorized to Offer Degree:

Geography

University of Washington

Abstract

Place, Policy and Parity: Examining and Visualizing Spatial and Socioeconomic Contributions to Hospital Charge Markup

Rebecca W Stubbs

Chair of the Supervisory Committee:
Professor John Mark Ellis
Geography

This thesis comprises two components. The first is an analysis of the relationship between hospital "retail" price lists as described in charge master lists, normalized to Medicare payments, and socio-economic and demographic covariates of the counties in which facilities are located. This analysis finds that little of the variation in how much hospitals charge the uninsured for services is due to the proportion of the population in a facility's surrounding county that is uninsured. The second component of the thesis is a description of a data visualization toolkit, the MapSuite package written for R. This code base is designed to facilitate the representation of spatio-temporal data sets. The design philosophy and implementation are described, along with illustrative examples.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to thank my family—for their love, support, and most of all for the love of learning and science that they each taught me from a very young age. I cherish you all, and could not feel more grateful and lucky—I firmly believe that I have some of the world's most exceptional and wonderful people as my relatives.

Next, I want to express my thanks to my advisor, Mark Ellis, and Sarah Elwood, for their comments and insight. Finally, I would like to show my gratitude to the Institute for Health Metrics and Evaluation, for the knowledge and experience I have gained, the mentorship I received, and the life-long friends I have made while under the Institute's wing.

Chapter 1. PLACE, POLICY, AND PARITY: EXAMINING SPATIAL AND SOCIOECONOMIC CONTRIBUTIONS TO HOSPITAL CHARGE MARKUP

Hospitals bill the uninsured prices that far exceed those determined as fair compensation by Medicare. The extent to which the chargemaster (list, or retail) price of a service differs from Medicare's reimbursement varies dramatically over the United States. However, the extent to which this variation is due to facility-level differences, socioeconomic characteristics of the population surrounding the facility, or other state- and county-level influences such as policy or regional pricing dynamics was previously unknown. Through the use of a mixed-effect linear regression model, the associations between socioeconomic factors, including the proportion of the population that is uninsured, and the discrepancy between list price and Medicare payments are explored. Although a number of socioeconomic variables may contribute to hospital pricing structure, it appears that much of the variation is due to exogenous factors at the state and county level.

1.1 INTRODUCTION

For the uninsured and the underinsured in the United States, the choice to seek medical care in a critical health emergency is often one between "life, and debt."[1,2] Out of the 3 trillion dollars spent on health care in the United States in 2015, 338.1 billion were paid out of pocket[3], and medical bills have been implicated as a major cause of bankruptcy.[4,5] Despite the magnitude of health care spending and the universal need for health care, hospital pricing structures are opaque and the way charges are constructed often remains a mystery to patients and their families. Pricing structures generated by hospitals have been described[6] as “chaos behind a veil of secrecy,” in which hospitals determine charges without a rigorous methodology.

From the moment a patient first arrives at a hospital, the facility tracks the services rendered, equipment used, and drugs administered to treat or diagnose a patient’s conditions. A hospital’s chargemaster list (master charge list or charge description master or CDM), describes the full menu of tens of thousands of procedures and goods available at a given facility, along with a price for that treatment or item. The prices for the care and commodities used are added up to generate the total bill for the services rendered. However, even when presented with a bill containing the full details of items included from the chargemaster list, the naming conventions and significance of different line items may be difficult or impossible for consumers to understand. Filled with acronyms, abbreviations, and medical jargon, the list of “services rendered” by the hospital may be unrecognizable to the patient. Different hospitals distribute revenue or profit margin on items used for the same procedure differently, which makes comparisons of the “list price” of individual goods (for example, the cost of a syringe) less meaningful.

Furthermore, most patients and insurance providers receive discounts that are a fraction of the supposed ‘cost’ of treatment. Mediated by insurance and co-pays, the chargemaster price of a hospital stay often does not reflect the amount paid to the hospital. Insurance companies and hospitals use the chargemaster list as a starting point, but negotiate rates far below the full ‘list price’ of the item. For the uninsured, however, the chargemaster price is a more accurate reflection of the bill. Additionally, some hospitals offer discounts for those paying out of pocket, but those unaware or ineligible for such programs may be facing substantial debt for their healthcare, and may also end up paying closer to the chargemaster price. Although the number of uninsured has declined in recent years, nearly one in ten people still do not have health insurance in the United States.[7]

When permanent residents and citizens of the United States turn 65, they become eligible for Medicare, the largest insurance program in the country. Run by the United States government and serving 57 million persons,[8] Medicare follows a different process to determine how much hospitals are paid for services rendered to patients. The Inpatient Prospective Payment System (IPPS) is the foundation for how Medicare determines payments for inpatient stays, which are covered under Medicare Part A. The base price for Medicare Severity Diagnosis Related Group (MS-DRG) is supplemented to account for (i) the area's cost of living, which can require higher staffing costs, (ii) whether the hospital has a medical education program, and (iii) the extent to which the facility is utilized by a disproportionate share of low-income and uninsured patients from the surrounding community. For unusually complex or resource-intensive cases, additional payment is disbursed to the hospital through the use of "outlier payments." Through these adjustments, the Inpatient Prospective Payment System seeks to pay hospitals a fair price that will compensate for both the procedure and the associated overhead of treating the patient. The additional payment granted to hospitals that serve a large share of the impoverished, however, may not completely defray the costs of caring for those who cannot or do not pay for the hospital's services.

1.2 LACK OF INSURANCE AS A DRIVER FOR HIGHER HOSPITAL CHARGES

One of the reasons cited by hospitals for high prices is the theory of 'cost shifting' such that private insurers and self-paying patients bear a higher financial burden in order for a hospital to make ends meet. The Emergency Medical Treatment and Active Labor Act (EMTALA) dictates that hospitals accepting payment from Medicare also provide medical screenings and emergency procedures regardless of the patient's citizenship or ability to pay. This "charity care" provision does not include any reimbursement for these services rendered. These emergency department patients, combined with patients who fail to pay for their non-emergency medical services, accounted for 4.2 percent of total hospital expenses in 2015.[8] Medicare reimburses hospitals who treat a larger share of the community's impoverished at a higher rate to adjust for this overhead—however, some facilities also claim that the relatively low amounts paid by Medicare and Medicaid require higher prices.[9]

The motives for higher charges likely extend beyond cost shifting. Inflated chargemaster pricing structures may give hospitals more power in negotiations with insurance companies, as insurance providers are incentivized to include hospitals with high prices in their insurance networks, in order to avoid costs incurred by patients receiving medical care at an un-discounted facility.[6] Increases in chargemaster price have been recognized as a mechanism for increased hospital revenue, and a study that examined the fifty US hospitals with the highest charge-to-cost ratios found that nearly all of them were for-profit.[10,11]

What drives this variation in hospital charges? Is the amount that hospitals charge linked to the socioeconomic properties of the population within a county, such that the amount charged by hospitals is linked to the patients' ability to pay? If charity care were a main motivator for hospital charge structures, there would be a strong relationship between higher charges and the proportion of the uninsured or unable to pay surrounding a facility. Is there evidence that the proportion of uninsured within a county is associated with higher charges? If so, how might the hospital charges be different if more of the population were insured? This paper addresses these issues. The sections that follow summarize prior work and define terminology, describe the methodology used, discuss the results, and present conclusions.

1.3 UNDERSTANDING THE GAP BETWEEN MEDICARE REIMBURSEMENT AND HOSPITAL CHARGES

Given that some procedures tend to cost more than others, examining price differences in relative terms, rather than absolute dollar amounts, allows for the comparison of charge structures across diseases, procedures, and hospitals. Given that Medicare reimburses hospitals differently given a variety of factors related to both the facility (cost of living and compensation for teaching programs), and its patients (the proportion of uninsured served at the hospital, as well as for "outlier cases" where treatment is more involved than the norm), comparing charges to Medicare payments should represent the additional cost above a negotiated "fair price." The amount that hospitals charge above and beyond what Medicare pays is often referred to by researchers as "markup," and is measured in units of the Medicare payment.[11,12] A procedure with a markup ratio of 1 (loosely described as "a markup of 1") would mean that the hospital charges the uninsured the same amount as what Medicare would pay, while "a markup of 2" corresponds to double what Medicare would pay for the same procedure.

In 1984, hospital charges were 135 percent of Medicare-allowable costs (a markup ratio of 1.35),[9] and the ratio between charges and costs has since risen further: in 2012, US Hospital chargemaster prices were a staggering 3.4 times the Medicare-allowable costs.[10] Hospital charge markup is far from spatially uniform: hospital prices vary markedly across the United States, such that out of the top 50 most expensive hospitals, over half were in the state of Florida.[10] Although the notion that where you live may determine the pricing structure of health services is far from new,[13] the extent to which there are spatial inequities in hospital charges is garnering increasing media attention and academic study.

In response to open data initiatives, and in an effort to make the healthcare marketplace more transparent, the Center for Medicare and Medicaid Services (CMS) published[14] the Provider Utilization and Payment Data Inpatient Public Use File (Inpatient PUF) in May of 2013. An unprecedented data release, the Inpatient PUF contains hospital-specific information on how much hospitals were compensated through Medicare's payment system (payments) compared to the amount billed to Medicare based on chargemaster list pricing (charges) for any disease reporting group with over 10 discharges per facility per year.

Research by Park et al. utilized the Inpatient PUF, and conclude[15] that "hospital charges lacked a relationship with population health indicators." Incorporating information on health outcomes, health behaviors, access to and quality of outpatient care, the physical environment, and demographic factors, Park et al. found that only the percentage of the population that is uninsured had a statistically significant correlation with raised hospital charges. The association seen by Park et al. between insurance levels and higher charges lends evidence to the claim that hospitals utilize cost-shifting and raise charges in order to compensate for the uninsured and those unable to pay for services.

This study laid groundwork for future explorations of this data set, but was limited in many ways. First, the analysis by Park et al. focused on associations between the average charges (the chargemaster list pricing for services) at each hospital without using Medicare payments as a normalization factor. This required the authors to generate a standardized metric for each facility that was not dependent on the types of patients seen at each facility. To achieve this, the authors generated "condition-aggregated" charges for each facility, calculating average charges across six conditions. This approach restricts the data to only certain conditions, which reduces the sample size and may mask some of the variation occurring in hospital pricing structures. Furthermore, the model presented by Park et al. included a wide range of covariates (29 in total), but potentially influential variables--such as the income level, as well as the racial and ethnic composition of the counties--were not included. Finally, the data used by Park et al. were restricted to only 2011. With three more years of data added to the Inpatient Public Use Data File after 2011, and the use of markup measures as a standardized point of comparison across space and disease category, deeper insights into the distribution of hospital price markup are within reach. The Inpatient PUF data presents an opportunity to examine how hospital charges compare to Medicare payments for a multitude of different conditions, spanning across the United States from 2011 to 2014.

1.4 STUDY METHODS

1.4.1 *Data*

The Inpatient Public Use Data File was obtained from the Centers for Medicare and Medicaid Services. This file contains records of the average total Medicare payments, and the total charges (chargemaster list price) for each disease reporting group (DRG) with over 10 discharges per facility per year at each hospital registered with Medicare between 2011 and 2014.[14] The data set is based on information from the Medicare Provider Analysis and Review (MEDPAR) data. Only the top 100 DRGs by number of discharges were included in the data set in 2011-2013, while all DRGs above the discharge threshold were included in 2014. The latitude and longitude, county, and census tract associated with each facility was determined by geolocating each hospital using the Google Maps API to generate coordinate positions based on the facility address.[16] The 29 facilities (out of an initial total of 3,445) that did not contain adequate location information to generate coordinates were excluded from the study, resulting in a total of 3416 facilities in the final data set, and 670,707 DRG cost observations. There was at least one facility located in 1,527 out of the 3,141 counties or county-equivalents within the United States. While more than half

(63%) of these counties contained only one facility, roughly one in ten counties represented in the data set (12%) contained four or more facilities, and a small subset (3%) had over 10 facilities within the county's boundaries. These data represent charges and payments for 27,185,017 inpatient discharges from October 1, 2011 through September 30, 2014. Markup was calculated by dividing the total average charges by the total Medicare payments for each disease reporting group in each facility-year. On average, each facility had 51.2 different DRGs captured within the data set, each with its own markup measure based on the charges and Medicare payments for each disease category. Normalizing each facility's chargemaster prices to their locally-determined Medicare reimbursement (generating a markup measure) should (if the Medicare adjustments for cost of living and underserved population factors were perfect) enhance the signal from other covariate factors.

Information on county-level education, household income, rural status, and race, were obtained[17] from the American Community Survey via the National Historical Geographic Information System. The proportion of the population below the poverty line and the proportion of uninsured within each county were procured from the Small Area Income and Poverty Estimates (SAIPE), and the Small Area Health Insurance Estimates (SAIHE) produced[18,19] by the Census bureau. Information on county-level unemployment, were sourced from the Bureau of Labor Statistics.[20] Age-standardized, all-cause mortality estimates at the county level were taken from estimates by Dwyer-Lindgren et al.[21] These mortality estimates and covariates refer to county or county-equivalents, except in a small number of cases in which historical boundary changes required the merging of counties into temporally stable units using population weighted averages. For covariates unavailable for the year 2014, covariates were extrapolated using the rate of change observed between 2008 and 2013.

1.4.2 *Methods*

The relationships between community socioeconomic factors and hospital markup for each disease reporting group in each facility in the years of 2011-2014 were modeled using a linear mixed-effect regression. Nine county-level covariates were included in the model. The percent of the population without insurance, as well as the proportion that were unemployed (due to the fact that many individuals and families are insured through an employer) were included in the model in order to examine the association between charge markup and the rates of uninsured. To disentangle rates of the uninsured from other socioeconomic traits of the community, demographic information on race (percent Black), ethnicity (percent Hispanic), and education (percent over 25 with a bachelor's degree) were included. All-cause mortality rate was inserted to evaluate whether a less healthy population would result in higher charge markups, which could indicate that the both the base payments and the outlier charges (paid by Medicare) did not adequately address the intensity of illness present in the community, causing the hospital to compensate by raising charges. Household median income, and the percent of individuals below the poverty line were incorporated in order to examine whether the charges levied by the facilities were related to the community's ability to pay for the services rendered. The percent of families that live in rural areas within the county was also included, to explore whether less urban counties were subject to a different markup than facilities in more metropolitan areas.

The model included random effects on state, county, facility, and year to account for the impact of spatial processes and unobserved variables that may influence hospital pricing structures. To reduce the extreme right-skew of the hospital markup data and achieve a normal distribution, the markup was log-transformed. Each of the covariates were transformed to z-scores, where the measurement for each county represented the difference in standard deviations from the mean of all counties from 2011 to 2014. This allows the magnitude of the relationship between each covariate and the outcome of interest (hospital charge markup) to be compared to one another in units of deviation from each covariate's mean—a large coefficient on one covariate compared to another will not be due to the units in which they are measured. The model is specified such that:

$$log(m_{fyg}) = \beta_0 + \beta_1 + X_{cy} + \gamma_s + \gamma_c + \gamma_f + \gamma_y + e_{fyg}$$

$$\gamma_s \sim N(0,\sigma_s)\ \gamma_c \sim N(0,\sigma_c)\quad \gamma_f \sim N(0,\sigma_f)\ \gamma_y \sim N(0,\sigma_y)$$

Where m_{fyg} is the markup in facility *f*, year *y*, and disease reporting group *g*, respectively; *c* represents the county each facility is located in, and *y* is the year in which the data were observed. β_0 is a fixed intercept which represents the expected markup before the contribution of covariates, time, state, county or facility. β_1 is an array of nine fit coefficients, one for each of the covariates described above; X_{cy} is a matrix of covariate values based on the county and year of each markup observation. The random effects represent deviations from the expected markup based on the intercept and covariates, and are represented by γ_s (a random intercept for each state), γ_c (a random intercept for each county), γ_f (a random intercept for each facility), and γ_y (the year-level random intercepts). Each of these random effects are taken as Gaussians with a mean of zero and independent standard deviations. Analysis was conducted in R version 3.3.2, [22] using the Template Model Builder package.[23]

1.5 RESULTS

1.5.1 *Descriptive Statistics*

Average chargemaster markup in relation to Medicare payment at the state level ranged from 1.22 in Maryland to 8.38 in New Jersey, and was 5.00 times the Medicare payment on average within the United States as a whole in 2014 (Table A-2, Table A-3). Markup was highly spatially variable below the state level, with areas of extremely high average markups in counties within New Jersey, Florida, and coastal California (Figure 1-1). Out of the 16 counties with average markups over 10, over half were in Florida and New Jersey alone, with five in Florida (Hernando, Okaloosa, Clay, Charlotte, and Saint Lucie Counties), and another five in New Jersey (Sussex, Somerset, Hudson, Mercer, and Warren counties).

At the hospital level, markup across different diseases tended to have more variation within facilities with a high markup on average across all diseases. The standard deviation of a facility's markup across disease reporting groups was strongly associated with the facility's median markup, producing a Pearson's correlation coefficient of 0.904 (95% UI 0.898-0.911; $p<0.0001$) (see Figure A-1).

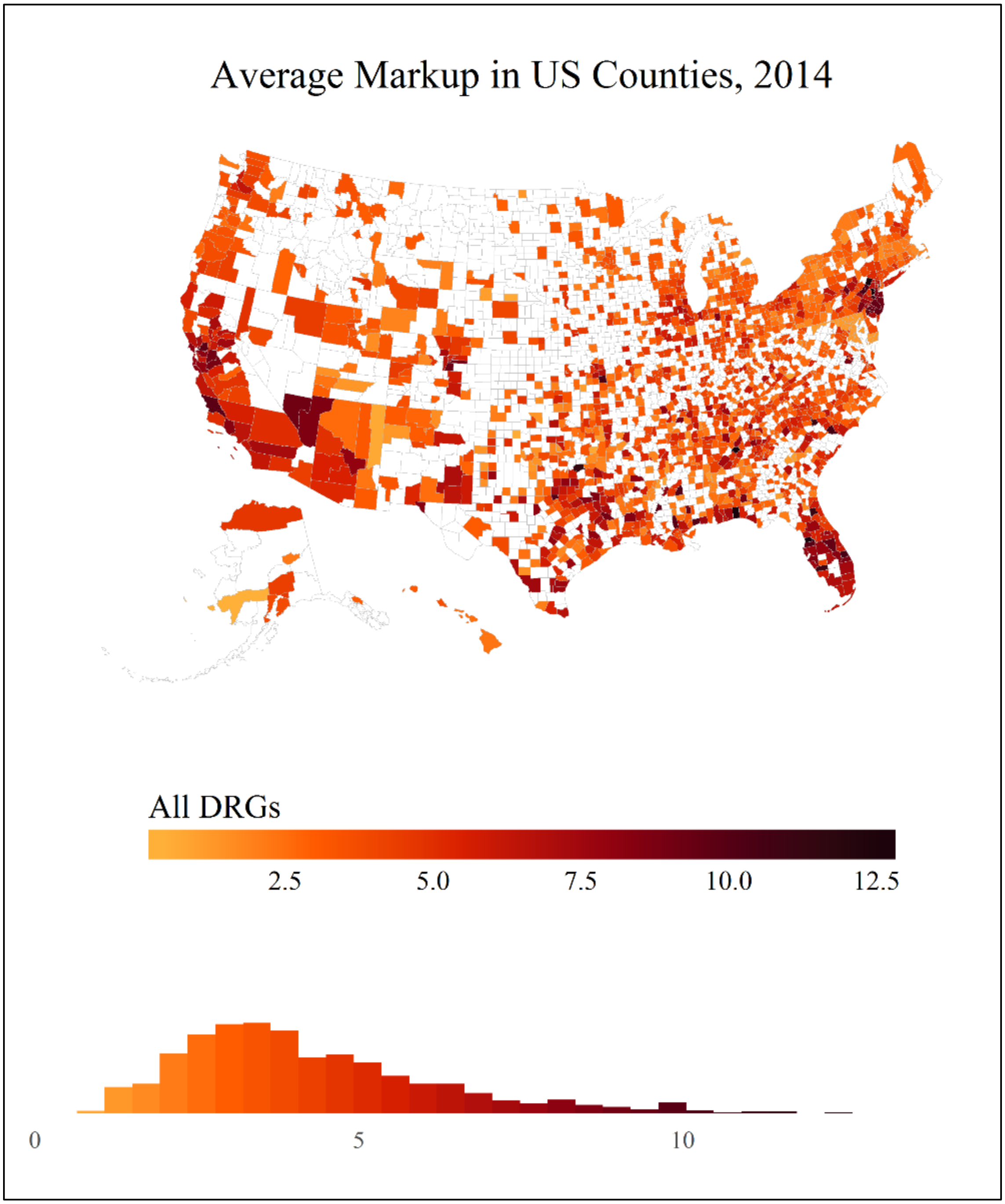


Figure 1-1: Markup on average in US Counties, 2014, for all Disease Reporting Groups. Counties that appear as white had no facility that met the selection thresholds. Histogram below represents the distribution of markup averaged to the county level.

When aggregated to the national level, the top 15 disease reporting groups (in terms of total discharges), markup ranged from 4.01 (Mental Diseases and Disorders), to 5.42 (Diseases and Disorders of the Nervous System) (see Table A-1) in 2014. For all DRGs in the top 15, the mean was consistently higher than the median, but the difference between the two measures was never greater than one. This indicates consistently right-skewed data, where relatively rare high values raised the average compared to the mean. Given that each observation for charges and Medicare payments represents the average of the discharges in the DRG for that year, it is worth noting that individual patients may have experienced markup ratios much higher than what is represented by the mean in any given State, County, Facility, or DRG.

1.5.2 *Regression Analysis*

Out of all of the variables supplied to the regression model, only household median income and the proportion of persons below the poverty line were not statistically significant at an alpha level of 0.005 (Table 1-1). This is potentially due to the method by which Medicare determines hospital pricing. Medicare pays more to hospitals that serve a higher proportion of the community's impoverished, and also pays more to hospitals where a higher cost of living requires additional funding for staff salaries (that could be reflected in a higher household median income). The lack of association between household median income, and the impoverished population could indicate that Medicare is adequately compensating hospitals based on salary requirements and the excess financial burden of serving the indigent. Variables associated with racial (percent Black) and ethnic composition (percent Hispanic), rural status, and unemployment levels were associated with lower markups. The proportion of the county's population with a bachelor's degree, the all-cause mortality rate, and the proportion of the population that is uninsured were all associated with rising hospital charges relative to Medicare payments.

Table 1-1. Fixed-Effect Coefficients from Regression Analysis

Parameter	**Estimate**	**95% Conf. Int.**	**P-Value**
Intercept*	1.123618	(1.03376, 1.21348)	1.210690e-132
Percent rural*	-0.216426	(-0.23544,-0.19741)	3.070644e-110
Percent black*	-0.086216	(-0.10666,-0.06577)	1.393117e-16
Percent Hispanic*	-0.075055	(-0.09502,-0.05509)	1.722108e-13
Mortality rate*	0.036599	(0.02789, 0.04530)	1.710491e-16
Percent bachelor’s degree*	0.022011	(0.00948, 0.03454)	5.731700e-04
Percent uninsured*	0.013190	(0.00976, 0.01662)	4.809692e-14
Percent unemployed*	-0.006358	(-0.00917,-0.00355)	9.339602e-06
Household median income	0.003809	(-0.00209, 0.00971)	2.058491e-01
Percent below poverty line	-0.000669	(-0.00469, 0.00335)	7.444701e-01
Root mean square error: 0.22; R^2=0.82. ** Represents statistically significant below critical value of 0.005.*			

The coefficients can be interpreted such that a 1-standard deviation change in the covariate is associated with a 1% change in markup that is approximately 100 times the beta coefficient. For example, a 1-standard deviation change in the percentage of rural population resulted in a 22% reduction in hospital markup. As a representation of the unobserved county-level and state-level variables that impact hospital markups, the maps of random effect values are also of interest. Evident in the map of random effect values are large spatial patterns present in hospital markup that are unexplained by socioeconomic variables. Of special note is the magnitude of the spatial random effects—while the coefficients of the socioeconomic variables included in the model may be significant, their contribution towards hospital markup is relatively small (usually less than a 10% change per standard deviation) compared to the impact of regional differences. Most of this spatial effect is seen at the state level rather than the county level, indicating that larger-scale processes such as state legislation may play a larger role than local dynamics (Figure 1-2). One pattern of note is that states in the north have a lower expected markup than those in the south of the United States.

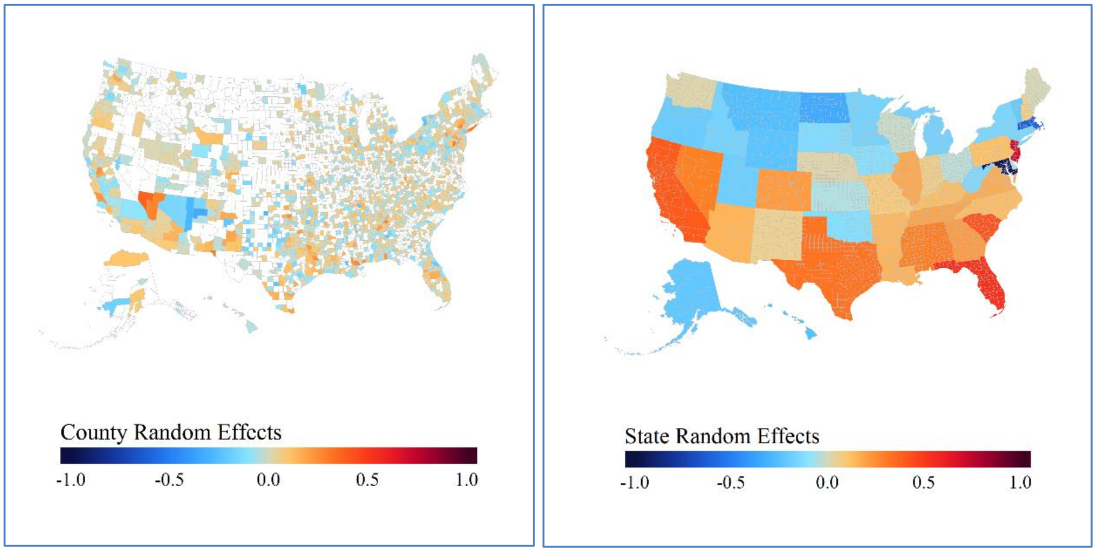


Figure 1-2: Maps of random effect values from regression analysis. These values represent the amount that county- or state-specific factors are associated with hospital markup above and beyond the impact of facility, or the socioeconomic characteristics of the county.

Of special note are states where the random effects produce exceptionally low markup values, such as Maryland and Massachusetts. Maryland is unique in terms of its pricing structure, having adopted an “all-payer rate system” in which hospital prices are set by the state legislature such that private insurers and the uninsured pay the same amount for services. Although the model evaluates the association between county-level insurance rates and hospital markup, state-level policies such as participating in the Medicaid expansion may have an additional effect on hospital pricing that could contribute to the state-level random effect. However, state-level random effects did not neatly align with Medicaid expansion decisions: Although many states that expanded Medicare as of November of 2014 did have state-level effect values that indicated a lower markup on average, there were exceptions, such as California, Nevada, New Mexico, and Colorado, all of which had state-level effects expected to raise markup despite having expanded Medicare.

Although some covariates were associated with lower charges, and others with higher charges, in the majority of cases, the combination of covariate values within each county resulted in an increase in the expected markup. Very low or high markup values were explained not by extreme covariate values, but by combinations of state, county, and facility-level differences that compounded—for example, a facility with a very high markup could be located in a county and state that tended to have higher markups, leading to a very high markup value, despite fairly standard covariate values. In Figure 1-3, 200 random samples were taken from the data set and visualized to show (i) how each component of the model (covariates, state, county, and facility-level effects) contributed to the predicted markup value, and (ii) how it compares to the observed markup value for that disease reporting group within a particular facility.

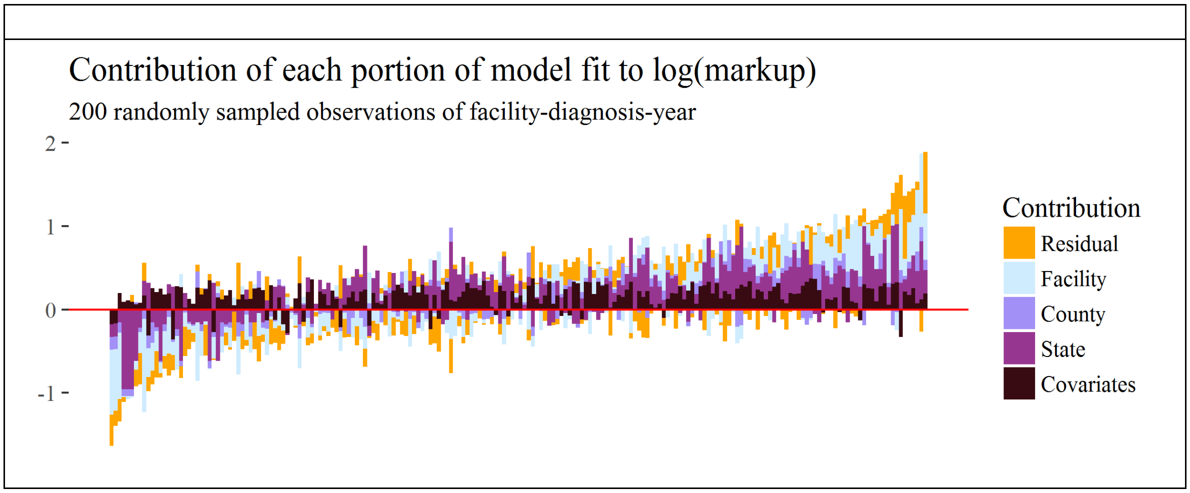


Figure 1-3: Examples of the contribution of each component of the model fit for any given data point. Graphic shows 200 randomly sampled observations from the data set, sorted left to right from lowest markup to highest markup. To generate the contribution of covariates, the covariate values for that county were multiplied by the beta coefficients, and added together. The Facility, County, and State values represent the deviation of the expected value based on the facility and its geographic location. The residual is the difference between the predicted values of markup based on the model, and the observed values seen in the real world. Contributions from each portion of the model are stacked, with segments above zero indicating that the contribution was positive, raising the markup rate, while segments below zero represent contributions that lowered the expected markup. An observation with more negative contributions than positive (as seen on the left) will be below the mean, while observations with a larger bar above zero (as seen on the right) will be above the national mean.

1.5.3 *Counterfactual Scenario*

To measure the potential impact of a better insured population on hospital charges, a hypothetical scenario was constructed where relationships established by the above model were applied to a counterfactual in which every county had uninsured rates of 1% of the population. To construct this scenario, the fraction of every county's population that does not have insurance was artificially set to 1 in 100 persons, while all other socioeconomic factors were held constant. The coefficients, fit parameters, and residuals from the above model, fit on the observed data, were used to generate a set of estimates for hospital charge markup given a far lower population of uninsured than what currently exists in the United States.

A hypothetical scenario in which each county was simulated to have low (1 in 100) rates of the uninsured revealed relatively little reduction in hospital markup. Given the relationship between the percentage of the population that is uninsured and markup as established by the regression model, a reduction in uninsured rates down to 1 per 100 in each county would result in a national average markup of 4.89, down from 5.00 (meaning that hospitals would charge 489% of the amount Medicare pays for the service, rather than 500%). On average, the hospital markup for any given condition treated in any facility in 2014 would decrease by 0.104, or roughly 10% of the Medicare payment. However, the effect of a better-insured population was distributed unevenly across the United States, given that areas with high rates of uninsured have the most room to improve (Figure 1-4).

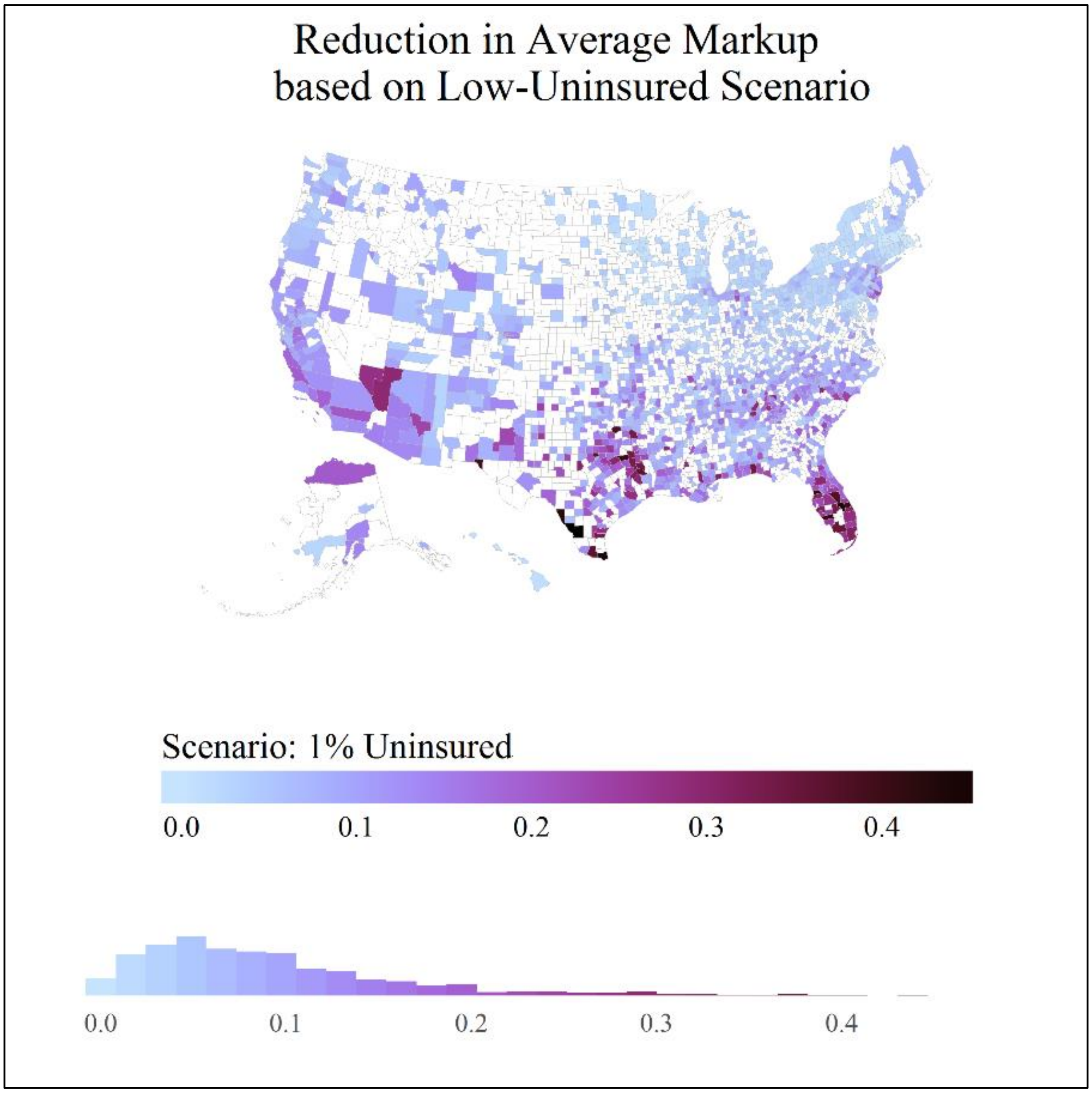


Figure 1-4: The average amount that markup would decrease within each county based on a hypothetical scenario in which the level of uninsured were 1 per 100. The histogram below the map represents the distribution of change in markup at the county level in 2014 based on a scenario in which 1 % of the population in each county in uninsured.

1.6 DISCUSSION

Some of the relationships (or lack thereof) between certain socioeconomic factors and hospital markups may be at least partially explained by Medicare's system of hospital reimbursement. Normalizing the 'retail' chargemaster prices to the Medicare reimbursement received by each facility will suppress dependencies on cost-of-living, since the Medicare system makes an adjustment for this factor. If the adjustment made by Medicare were perfect, we would expect to see no dependence on variables related to the wealth of an area, such as median household income. Medicare calculates a "disproportionate share hospital" (DSH) adjustment to compensate hospitals for serving a large share of the vulnerable community, who may not be able to pay for services. The way Medicare calculates this DSH may give some indication as to why poverty was not significant, yet the proportion of uninsured within the community did seem to contribute to hospital markup. The DHS Patient Percent used to adjust hospital payments is based on the proportion of inpatient bed days attributable to patients eligible for Medicaid, but not Medicare Part A, which means that the uninsured population is never directly captured within Medicare's DSH adjustment.

Other relationships revealed by this analysis are more surprising. The relatively strong influence of the proportion of a county that lives in a rural area challenges the narrative that facilities with few competitors (like those in rural areas) have the market power to charge more for their services.[12] From a social justice standpoint, it is not clear why the racial and ethnic composition of a county should contribute to the price of inpatient procedures. However, given that level of education, race, and ethnicity are associated with likelihood of being insured, these variables were included in the model to disentangle the impact of health insurance alone.

The relatively minor contribution of a county's level of uninsured challenges the notion that cost-shifting practices based on charity-care are major driver of hospital prices. Although reducing the number of uninsured (i) is clearly beneficial for many other reasons, and (ii) would protect a larger share of the population from being subject to the "list prices" for medical care, this analysis suggests that policy makers will need to look elsewhere to control the rising charges presented by hospitals.

1.6.1 *Study and Data Limitations*

A variety of policies and events that may impact the pricing structure of hospitals (both in terms of Medicare payments and charges) occurred preceding and during the time period of this analysis, which may impact the generalizability of these results. The release of the Inpatient Public Use Data File by the Centers for Medicare Services could have had an impact on hospital pricing, as facilities could easily compare themselves to their peers. A variety of changes to insurance requirements resulted from the Affordable Care Act (ACA), which passed in 2010. In addition, ACA enacted some changes in how Medicare reimbursed hospitals for care based on metrics of performance. Starting in fiscal year 2014, hospitals in the worst quartile in terms of hospital-acquired conditions were penalized by 1% as an incentive to reduce hospital-acquired cases of diseases such as MRSA, or clostridium difficile.[24] In addition, for discharges beginning in October of 2012, hospitals with excess readmissions also had their payments reduced up to 3%.[24] Given that hospital charges differ from Medicare payments on the order of multipliers, not percentage points, these effects are likely to be minor. The sample of hospitals within the study was limited to only facilities with over ten discharges for a given disease reporting group, which may have censored the data to exclude hospitals with smaller patient populations, such as rural areas. For counties with only one facility, the model cannot differentiate between county-level and facility-level effects—since a fit parameter is generated for each county, and for each facility, there is no way for the model to discern the contribution of the county as opposed to the facility in these circumstances. Additionally, the disease reporting groups captured within the dataset may not represent the conditions and injuries faced by more youthful populations.

1.6.2 *Conclusions and Broader Implications*

The magnitude of the differences in hospital markup across the United States is evidence for a deep and driving inequity, in which the potential cost of being uninsured is radically different depending on spatial location. It is contrary to the values of a socially just society that the financial penalty for medical treatment is so spatially variable. Unlike other services governed by consumer choice and a free market, patients rarely are able to make informed choices about where they receive medical care - even if information on pricing structures were readily available, a life-threatening emergency is not a convenient time for price shopping. Higher costs of care would be more understandable if paying more equated to better treatment. However, research linking the costs of care and patient success rates have shown the opposite relationship: less expensive facilities often offer better outcomes.[25]

Private insurance companies negotiate their rates with hospitals annually based on some combination of flat rates, adaptations of Medicare DRG payment structures, and discounts from the chargemaster prices.[6] Given that the chargemaster prices for a hospital serve as a starting point for insurance negotiations (which may impact premiums and out-of-pocket costs), understanding what drives hospital charges in relation to costs is to the benefit of the insured and uninsured alike.

Policies to protect the uninsured may mitigate the impact of this inequity, but do not offer complete protection. Section 9007 of the Patient Protection and Affordable Care Act revokes non-profit hospitals' tax-exempt status if they are found to be charging self-pay patients more than the insured.[26] However, the language of the bill is ambiguous and lacks definitive guidelines for which hospitals can be evaluated, and patients who find themselves at for-profit hospitals have little protection from these charges.

This work does not seek to show causality between county-level socioeconomic factors and hospital markup, but rather suggests that future work should further investigate the extent to which state-level policies and a community's properties are significant and contributing factors impacting hospital pricing. The influence of place (independent of poverty, wealth, or other socioeconomic factors) on markup serves to underscore how little we understand what drives the geographic variations in pricing structures.

This study continues the process of disentangling the complicated mosaic of factors that generates a landscape of pricing inequality. Future work on this topic could explore the impact of hospital-level variables, such as hospital type, proximity to other facilities, and wealth of the specific neighborhood in which the facility is located.

MAPSUITE: AN R PACKAGE FOR THEMATIC MAPS

MapSuite is an R package that streamlines the process for making maps of point, polygon, and raster data using the foundation of the ggplot2 graphics library. This paper describes a series of functions designed to facilitate mapping spatial data, from simple representations of the geometries to visualizing a variables over a series, such as time or data subgroup. Also described is how the outputs of these functions can be used to stack geospatial data layers, generating sophisticated maps in R that previously were relegated to lengthy code or desktop mapping software.

2.1 INTRODUCTION TO THE WOODSON MAPPING SUITE

The ability for scientists, students, and citizens to easily visualize and analyze spatial data is more crucial than ever. As the amount of freely available geo-referenced data continues to grow, an ever-increasing community is discovering and utilizing spatial information[27]. A host of commercial and open source software with point-and-click interfaces exist for spatial analysis and visualization-- from ESRI's software suite to QGIS, Google Maps and other online options. [28] For those who wish to move beyond the graphical user interface, open-source languages such as R and Python are no longer restricted to those with academic or computer science backgrounds. Massive online open courses, an extensive blogging community, and online forums such as StackExchange provide instruction and trouble-shooting help from any corner of the world. The barriers to a truly participatory geographic information science (GIS) community are being lowered. This explosion of free resources empowers citizen scientists, grassroots organizations, and institutions in low-resource settings to take advantage of the ever-growing library of open-source toolkits for data science. R, as an open-source programming language used extensively for data analysis and statistics, has a variety of tools designed for spatial analysis and visualization.[22,29]

Working with spatial data is not, however, straightforward. Full of idiosyncrasies such as varying data types, projections, and the threat of invalid geometries, there is an additional overhead to working with spatial data. This poses a challenge: how can users with less experience in programming and GIS be brought into the fold? The ability to present and visualize findings from any analysis is critical, and map-making has been shown to be a key tool. Empowered communities have the ability to present maps to decision makers, leading to policy changes.[30]

The **MapSuite** package joins a community of prior work on mapping in R. In addition to the **sp** package[29], which contains both spatial data types and simple plotting functions, a variety of packages focus on spatial data visualization. Finely-customized graphics of many kinds are possible using **ggplot2**,**[30]** which offers a number of data-driven plotting functions. Mapping-specific packages, such as **ggmap**,**[31]** **ggspatial**,**[32]** **rCarto**,[33] and **choroplethR**,[34] each present ways for R users to visualize spatial data. However, none of the currently existing packages emphasize (i) the ability to map quickly and easily over time or subgroup, or (ii) the ability to layer complex maps to create a more sophisticated cartographic output. The growing popularity of interactive tools, such as **Leaflet**[35] and **Shiny**,[36] have changed the landscape of what it means to visualize multi-dimensional data. However, generating images and PDFs of results is still a key way to share and present results.

Many of the static graphics packages, such as **ggplot** and **ggmap**, utilize a paradigm established in R by Hadley Wickham called the "grammar of graphics."[37] Within the "grammar of graphics," a plot is built in layers, where the data forms the foundation of the plot, and different geometric outputs (points, paths/lines, and polygons) can be layered and manipulated to generate a diagram. This approach to code has many advantages-- the syntax is highly flexible and well-suited to the kinds of data used by and produced from statistical analyses. There is also extensive documentation on how to use **ggplot2**, and how to achieve a variety of effects using this flexible package.

The MapSuite package uses the **ggplot2** library as a launching point for making maps, using three main functions (PolygonMap, PointMap, and RasterMap) as a wrapper for **ggplot2**'s geom_polygon, geom_point and geom_raster geometries. Unlike building a plot from scratch in **ggplot2**, these functions demand relatively few inputs to create a pleasing choropleth map. The default settings of the MapSuite are designed to follow a set of graphical guidelines set forth by E.R. Tufte.[38]

2.1.1 *Maximizing the Data-to-Ink Ratio*

Edward Tufte was well-known for the philosophy that effective graphics maximize information transfer while minimizing distractions. According to Tufte, in his book *The Visual Display of Quantitative Information*, the principles of graphical excellence[38] are as follows:

- Show the data
- Induce the viewer to think about substance rather than about methodology, graphic design, the technology of graphic production, or something else
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data

- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with the statistical and verbal descriptions of a data set

The maps produced by the **MapSuite** package default settings are designed to follow these guidelines while lowering the burden on R users such that time can be spent on the data and analysis, not the plotting code. The sections that follow will introduce functions available in the **MapSuite** package, and discuss how the default settings for these functions are designed to follow Tufte's guidelines. First, examples of basic maps of point, raster, and polygon objects will provide an orientation to the parameters required by the **MapSuite** functions. Next, mapping categorical and numeric variables will be presented. An exploration of the challenges of merging spatial data and data.frames in R will provide context for why the **MapSuite** functions take external data sets as an optional parameter. After discussing the role of color in cartography, and showcasing a new set of easily accessible color ramps, the paper will introduce a function that generates histograms following the same color scheme as the map in order to understand the data's non-spatial distribution properties. Finally, the **MapSuite** functions that facilitate mapping over multiple dimensions (such as subgroup or time), and the layering and stacking of map layers will be revealed.

2.1.2 *Data and Objects Used in Examples*

Example data is used to illustrate the functionalities of the MapSuite package. A shapefile of county borders was obtained from the Census Bureau’s TIGER/Line files[39], and used to create a variety of spatial objects. A listing of these objects can be found in Table 2-1.

Table 2-1. Data and Objects Used in the MapSuite Package

Object Name	**Desc-ription**
county_polygons	A SpatialPolygonsDataFrame of county boundaries, originally derived from the US Census Bureau's Tiger LINE files
county_centroids	A data.table with coordinates for county centroids, generated to be the central point within county_polygons boundaries
county_pixels	A data.table with coordinates on a regular grid with a field that describes the county location of each point
county_simulated	A data.table of simulated data at the county level
county_simulated_time	A data.table of simulated data at the county level, multiple time periods

Each of the polygons, centroids, and pixels objects contain columns with a unique identifier ('cnty'), and columns for state name ('state_name'), and the mean elevation of the county ('elevation'). The data and spatial objects can be subset by state name, allowing for graphics focusing on a particular area (the states of Washington and Colorado are emphasized in these examples). Subsets of these data sets for individual states are also used.

2.2 SHOWING THE DATA: PLOTTING BASIC GEOMETRIES AND VARIABLES

Before mapping a variable, it is sometimes useful to visualize the spatial objects themselves. The required inputs (listed in Table 2-2) to each of the main functions (PolygonMap, PointMap, and RasterMap) are minimal.

Table 2-2. Required Inputs for the Three Mapping Functions in the MapSuite Package

Functions	Input	Description of Required Inputs
PolygonMap	map	A SpatialPolygonsDataFrame with a field in the @data slot that can serve as a unique identifier
PolygonMap, PointMap and RasterMap	id	The name of the unique ID field that identifies the geometry you want to map
PointMap and RasterMap	coords	A data.frame or data.table with a field that can serve as a unique identifier, and a column for the latitude and longitude of the points
PointMap and RasterMap	xcol	The name of the column in the coords data object that represents x or longitude
PointMap and RasterMap	ycol	The name of the column in the coords data object that represents y or latitude

By default, the functions will create maps with gray shapes, as could be used for a background map or preliminary data exploration (Figure 2-1).

```
PolygonMap(map=wa_polygons, id="cnty")
PointMap(coords=wa_centroids, id="cnty", xcol="x_albers", ycol="y_albers")
RasterMap(coords=wa_pixels, id="cnty", xcol="x_albers", ycol="y_albers")
```

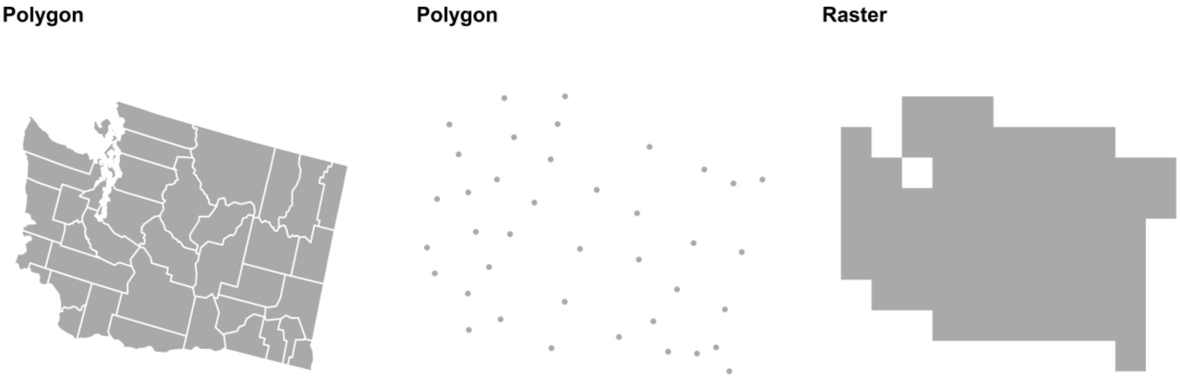


Figure 2-1. Mapping Point, Polygon, and Raster Data.

By adding optional parameters, various map aesthetics can be changed. The colors of the map (fill and outline in the case of polygons, fill alone in the case of points and rasters) can be modified. An outline layer (of class SpatialPolygons) that will lie atop your main map for context can be added and formatted.

```
PolygonMap(map=wa_polygons, id="cnty", map_colors = "lightgreen",
          map_outline_color ="black")
PointMap(coords=wa_centroids,id="cnty", xcol="x_albers", ycol="y_albers",
        map_colors = "red", pointsize = 1,
        outline = wa_polygons, outline_size = 1, outline_color = "black")
```

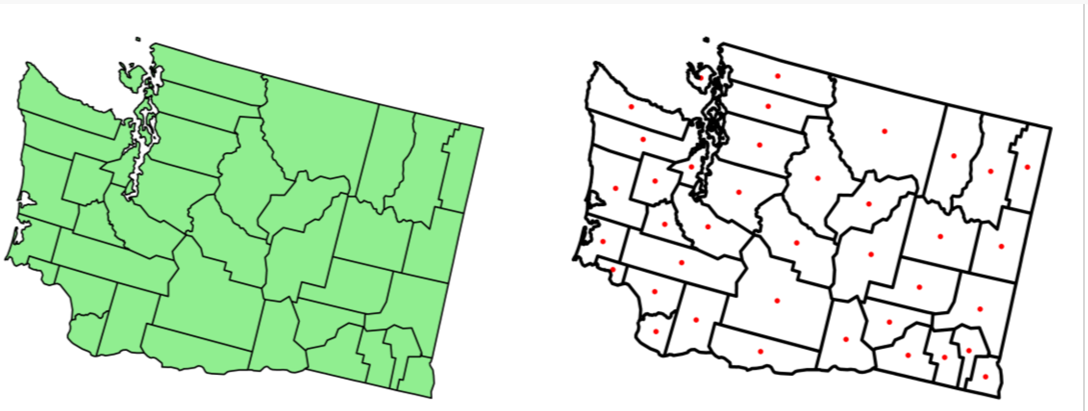


Figure 2-2. Changing basic map aesthetics and adding outline polygons.

To achieve a stronger figure-ground effect where a certain area is visually emphasized, adding an outline of a different shape and color can be utilized by passing a different SpatialPolygonsDataFrame to the outline parameter, or by passing only a subset of the SpatialPolygonsDataFrame used for the main map to the outline parameter (Figure 2-3).

```
Okanagan<-wa_polygons[wa_polygons@data[["cnty"]]==2944,]
PolygonMap(map=wa_polygons, id="cnty", map_colors = "lightblue",
        outline = Okanagan, outline_size = 1, outline_color = "orange",
        map_title="Okanagan County")
```

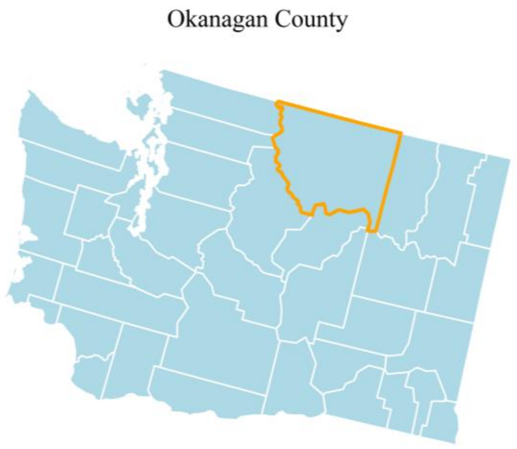


Figure 2-3. Using outlines to add emphasis.

Mapping a variable within the data attributes of the spatial object (within the data.frame that serves as the coordinates for the Point and Raster functions or within the @data slot of the SpatialPolygonsDataFrame in the PolygonMap function) is possible by adding a "variable" parameter. This variable can either be numeric or categorical in nature (Figure 2-4).

```
PolygonMap(map=county_polygons,
        id="cnty", variable="elevation",
        map_title="Mean Elevation")

county_polygons@data[,above_below:=ifelse(county_polygons@data[["elevation"]]>1000,
                              "<1,000 ft",">1,000 ft")]
PolygonMap(map=county_polygons,
        id="cnty", variable="above_below",
        legend_title = "Mean Elevation")
```

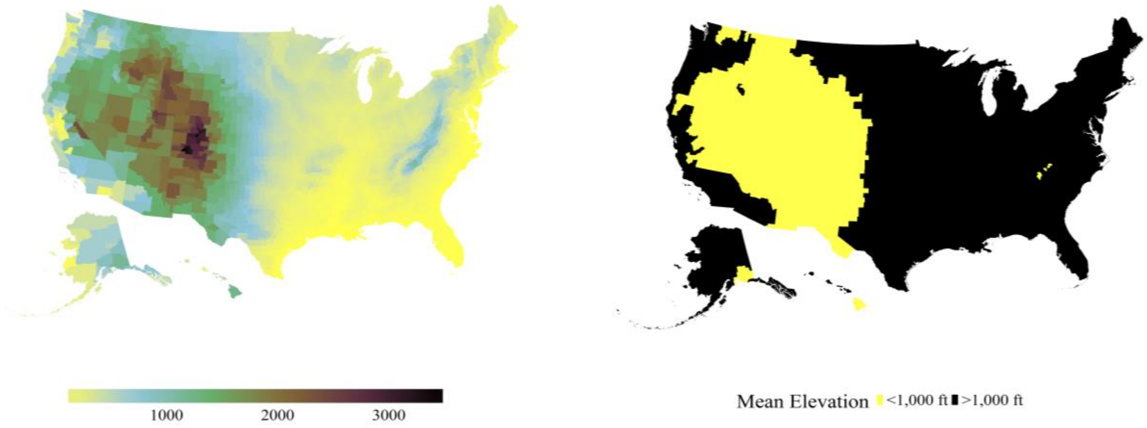


Figure 2-4. Mapping binned and categorical data.

Although providing a color range is not required (a default color scheme will portray the data), it is easy to change. It is also possible to label specific cut-points on the legend with values of interest, and modify font and legend properties (Figure 2-5).

```
PolygonMap(map=county_polygons, id="cnty", variable="elevation",
        map_colors=c("red","yellow","white","blue","violet"),
        map_title="United States", map_title_font_face = "bold", font_size = 12,
        map_subtitle ="Counties", font_family = "sans",
        legend_title="Mean Elevation", legend_font_face = "italic",
        legend_position = "right", legend_orientation="horizontal" ,
        legend_label_values = c("Sea Level","1 mile high"),legend_label_breaks =c(10,3250))
```

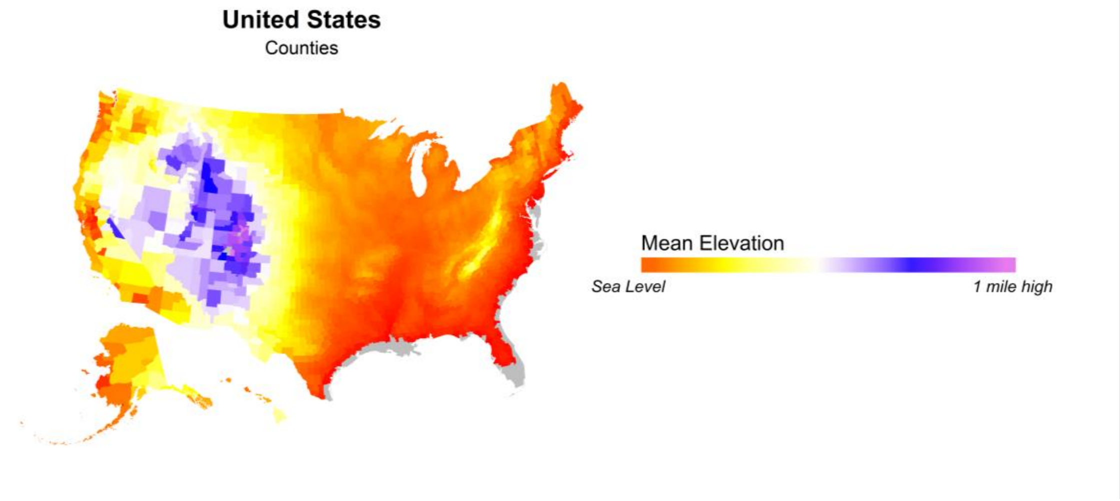


Figure 2-5. Changing color schemes and legend properties.

For point data, a separate variable can be defined as the variable that describes the size of the points (Figure 2-6).

```
PointMap(coords=county_centroids, id="cnty", xcol="x_albers", ycol="y_albers",
        map_colors=wpal("betafish"),
        variable="elevation_range",legend_title = "Elevation Range",
        sizevar = "elevation", sizetitle = "Elevation", font_size = 12,
        legend_position = "right", legend_orientation = "horizontal", legend_stacking="vertical")
```

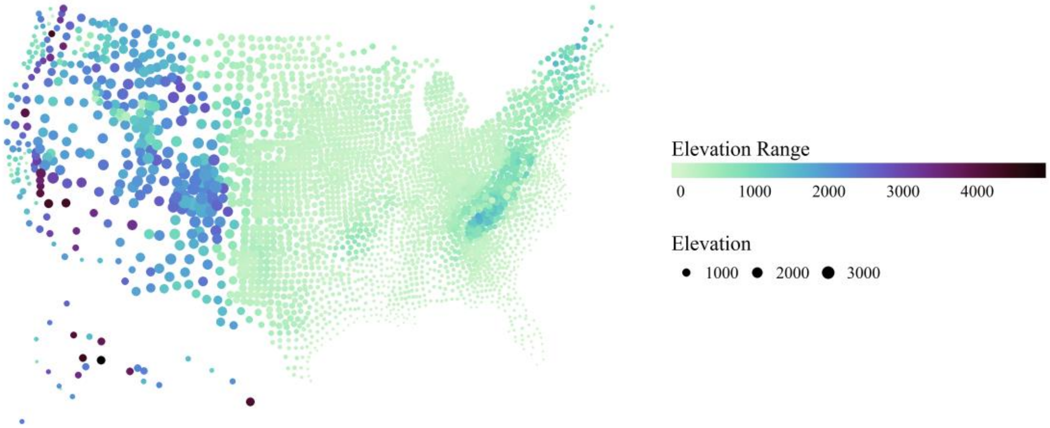


Figure 2-6. Modifying point size based on another variable.

2.3 ALLOW THE PROGRAMMER TO THINK ABOUT SUBSTANCE RATHER THAN ABOUT METHODOLOGY

What if the variable you wish to plot is not already in same data object as the object you are using to plot the geometry's coordinates or boundaries (either a data.frame or SpatialPolygonsDataFrame)? Although a seemingly innocuous solution is to use one of R's base functions such as merge() to join spatial data with other attributes, this can prove to be problematic. An under-appreciated quirk of the SpatialPolygonsDataFrame data type is that the geometry slot (@polygons) is related to the attributes (@data) by a shared ordering, where the first item in the

@polygons slot has the attributes of the first item in the @data slot. If an R user wishes to map a variable not yet in the @data slot, the temptation is to use the merge() function to join the two data sets and carry on with plotting your spatial data. However, this approach is dangerous- unbeknownst to many, the merge() function re-orders the resulting data frame by the ID fields used to join the two data sets. This can lead to maps that look plausible (there are data values and attributes for each polygon), but the ordering of the @data slot no longer matches with the @polygons, leading to a map where attributes are mismatched with their geometries (Figure 2-7).

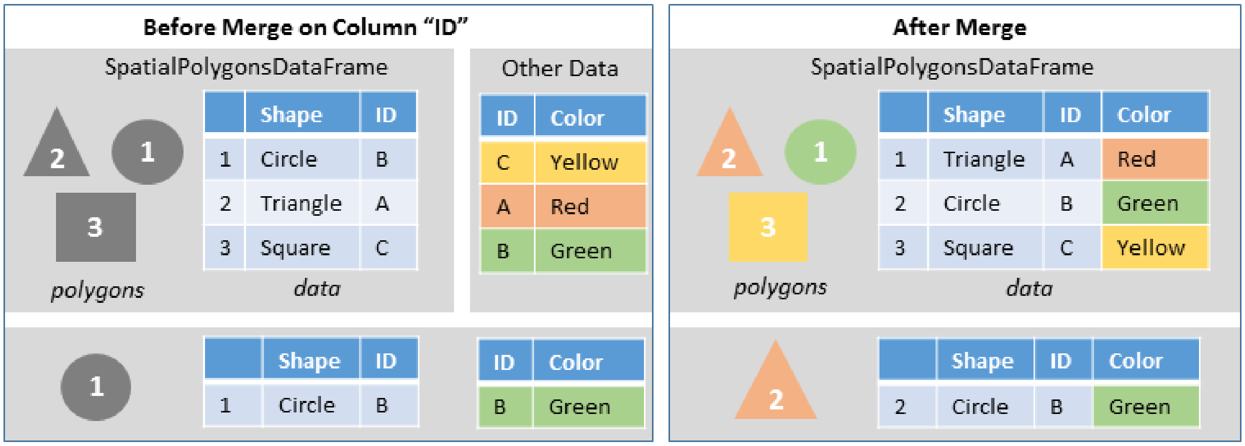


Figure 2-7. Illustration of how polygons become unlinked to their attributes after a merge() operation.

To prevent this, the PolygonMap, PointMap, and RasterMap functions take external data sets as parameters, which can be joined onto the spatial data using a specified field present in both the spatial and tabular data. The merging of the external data and the spatial objects within the PolygonMap function are ensured to preserve the polygon-to-attribute relationships. The polygon object is "fortified" (converted from a SpatialPolygonsDataFrame object into a data.frame with coordinate locations for plotting) within the function, and the additional data set is merged on.

This syntax makes it easy to switch between different variables that exist in an external data set (Figure 2-8).

```
PolygonMap(map=county_polygons, data=county_simulated,
        id="cnty", variable="var1",  map_title="Simulated Variable 1")
PolygonMap(map=county_polygons, data=county_simulated,
        id="cnty", variable="var2",  map_title="Simulated Variable 2")
```

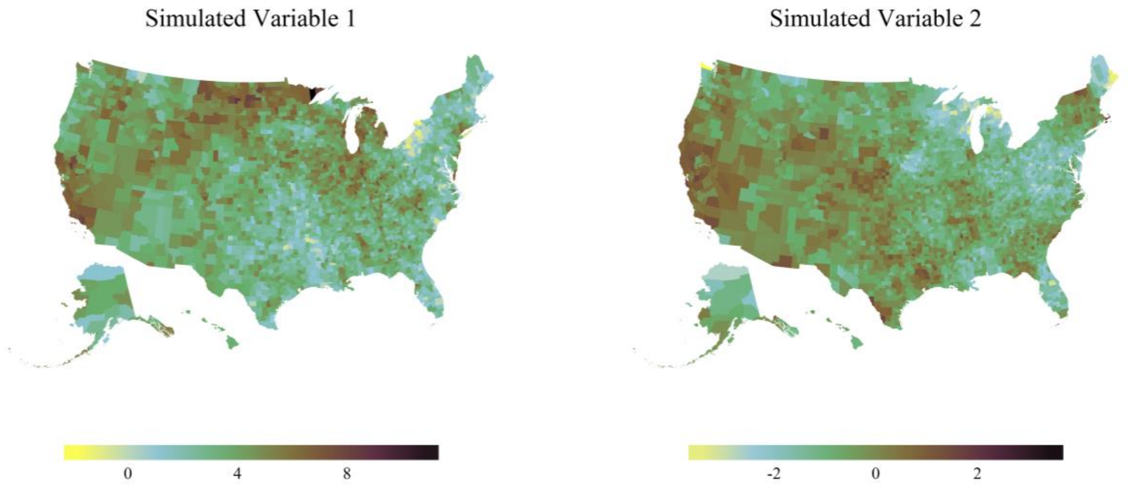


Figure 2-8. Mapping two different variables from an external data frame.

2.4 Avoiding Distorting What the Data Have to Say

The hue (pigment), saturation (intensity), and value (darkness) chosen to portray data can dramatically change their perception.[40] Color is an important consideration for any graphic--however, it is especially crucial to consider when making maps due to the trust awarded to cartographers by the public. As Judith Tyner noticed in 1982, "map users, especially those who examine maps infrequently, tend to place inordinate faith in maps and accept them as true and complete representations."[41]

Research on the color choices that underpin readable maps and graphics has provided nuanced insight towards what kind of color palettes are most effective for representing numeric or sequential data. Broadly, three main categories of color scale exist: sequential, diverging, and categorical. Sequential palettes move through at least two colors, often going from light-to-dark (or vice versa). A diverging scheme passes through at least three colors, as two colors "diverge" from a shared middle ground (Figure 2-9). Categorical color schemes have no trajectory-- and are often chosen to be dissimilar from one another such that the reader can easily tell the difference between two categories. Broadly, sequential color schemes are used when the entire range of the data is important, but there is no central value with special significance that serves as a frame of reference. In contrast, diverging color schemes are recommended when the data have a meaningful central value that can be used as a frame of reference. This can be either a specific cutoff based on knowledge of values of relevance (such as values above and below 0, or a certain threshold), or based some parameter of the data (for example, mapping Z-scores that diverge from 0). One special case of the diverging color scheme is the "spectral scheme," alternatively known as the "rainbow scheme," that passed through the full (or modified) color space of red-orange-yellow-green-blue-indigo-violet. Although some researchers argue against this rainbow color scheme[42], others have advocated its use in specific scenarios, such as the mapping of mortality rates, where map readers could easily differentiate between the low values colored in blue, and the higher values coded in red.[43]

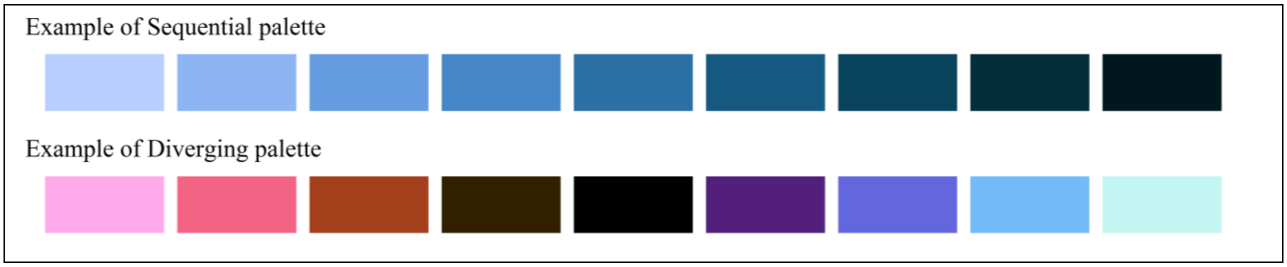


Figure 2-9. Sequential and diverging palettes.

The "ColorBrewer" project of Cynthia Brewer, available in R through the package **RColorBrewer**[44] contains a variety of color schemes used regularly within the mapping community. However, the sequential palettes presented in RColorBrewer often do not contain more than two or three colors, limiting the differentiation possible between data values. A member of the astronomy community, Dave Green, developed the "cubehelix" algorithm to differentiate star-brightness values from observations of the night sky. The "cubehelix" is a virtual corkscrew moving through a cube of red-green-blue color space on a trajectory from dark to light[20]. This function has been adapted for R in the **rje** package,[45] where users can define the following variables in order to produce a variety of diferent color palettes moving from dark to light: (i) the number of rotations, (ii) how large the radius of the corkscrew is, and (iii) the saturation. The permutations for color palettes that can be created with this function is nearly endless. Some examples of both a Color Brewer palette and two cubehelix palettes can be found in Figure 2-10.

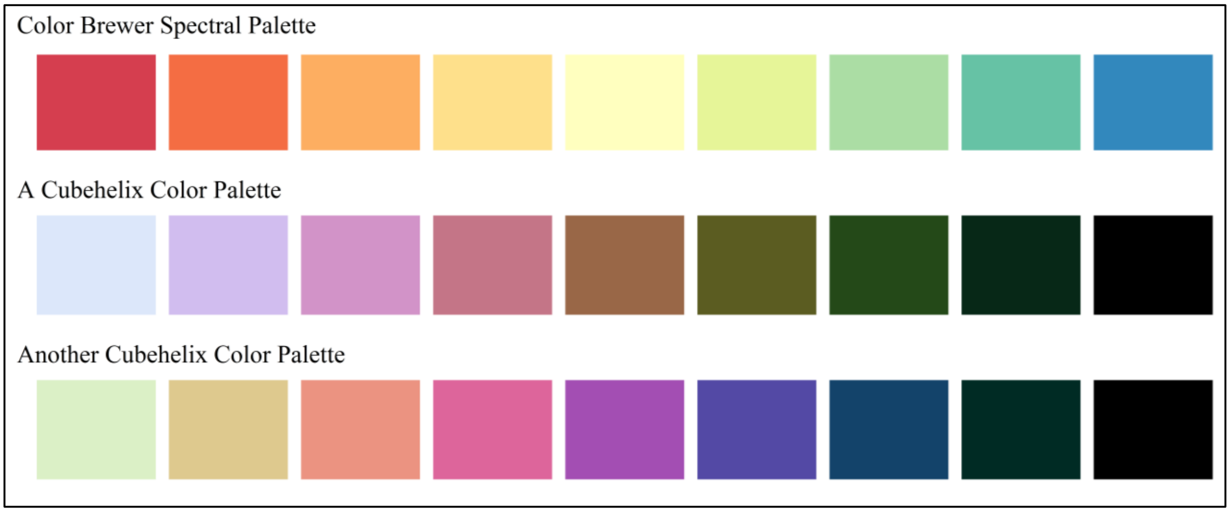


Figure 2-10. Cubehelix color palettes as compared to a Color Brewer palette.

Within the **MapSuite** library, a series of curated sequential and diverging palettes derived from the cubehelix algorithm are contained within the wpal function. The **MapSuite**'s "Woodson Palettes" are lists of color values that can be called upon by both the main mapping functions and any other function that accepts a list of colors in R. Using the MapSuite's wpal function, users can specify a number of colors that will be interpolated, as well as whether black will be included (if it exists in the color palette). If you wish to explore or view a specific Woodson Palette's default colors, the ViewPal function will plot the desired color ramp (Figure 2-11). The PlotColors function allows the plotting of any list of colors, including a modified Woodson Palette.

```
ViewPal("seaglass",base_size = 12)
## [1] "Plotting wpal color scheme seaglass"
```

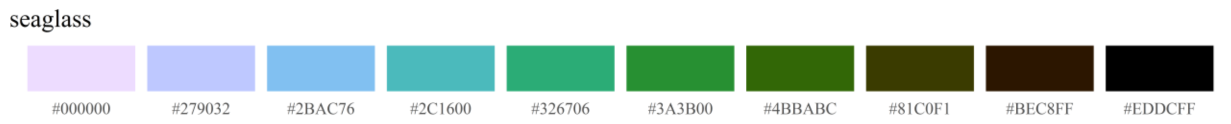


Figure 2-11. Plotting a Woodson Palette.

To use a Woodson Palette in reverse (re-ordered dark-to-light), the function rev() can be used sort the color pallette in reverse order, as seen in Figure (3-12).

```
more_colors<-rev(wpal("seaglass",n=12,noblack=T))
PlotColors(more_colors, color_list_name="Reversed Ramp with More Colors",base_size = 12)
```

Figure 2-12. Plotting a color ramp that has been reversed.

Our instinct to assign value judgments ("red is bad, blue is good") to color scales is simultaneously what makes them effective and conversely problematic. While there are some variables that can be clearly discerned as favorable and non-favorable (like mortality rates), visualizing quantities like the proportion of people belonging to a minority racial group on a similar color scale is far from appropriate. When visualizing a large quantity of data, where subtle spatial differences should be clear to the viewer, a diverging color palette that moves through three colors tends to add more clarity and visual precision than a sequential palette moving through only two. As such, there is the temptation to use diverging colors such as the spectral color scheme to represent data that may not have a meaningful central value, and may not have a clear "good" and "bad" connotation. For this reason, making color palettes based on a cubehelix color scheme that move from light to dark, rather than from a "good" color to a "bad" color, could serve an important purpose in visualizing data in social sciences. The default color scheme (called "earth"), moves in a sequence (yellow, blue, green, brown, purple, then black) designed to provide value differentiation without a moral or emotional value judgment. These scales, moving from light to dark, will also preserve the relationships between data values and pigment value, even when printed in black and white, and when viewed by the colorblind.

It is worth noting that the same color scale can look quite different based on the number of geometries or data points present-- while a color ramp that goes through many colors looks appropriate with large numbers of data points or geometries, too few can create a "chunky" or categorical look. In these circumstances, choosing a color palette with fewer colors may provide a more pleasing and readable map. To showcase the extent to which color and binning (converting a continuous variable to a categorical one) impact the visual appearance of data, the same data will be used in the fours figures, Figure 2-13 through Figure 2-16.

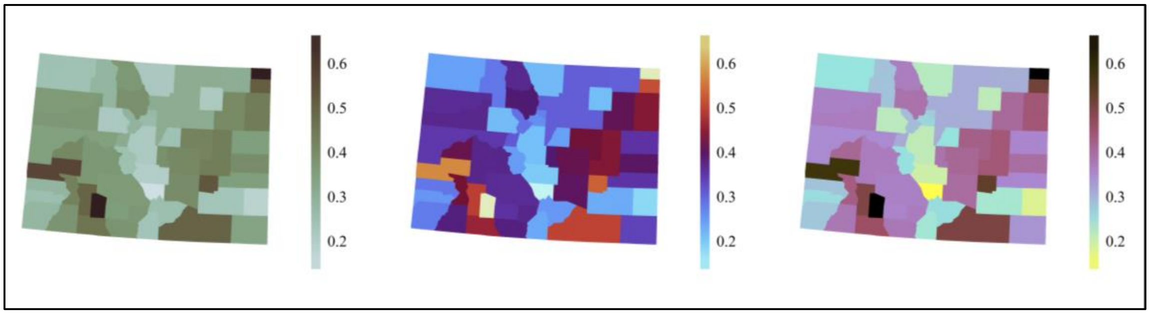


Figure 2-13. The same data, represented using three different color schemes.

The default use of continuous color scales for numeric data is designed to minimize the visual bias that occurs from data classification into discrete groups, which can "reveal meaningful spatial trends or promote misleading interpretations" of the data[46]. Although different color schemes and ramps can emphasize or subdue patterns in the data, a continuous color scale that ranges through the minimum and maximum values is subject to one fewer source of bias-- the choice of what classification scheme to use. Classifying or binning data has the danger of making values that are similar appear distant from one another due to proximity to the breaks values, and the size and divisions used to create data categories can dramatically change the visual appearance of the data.

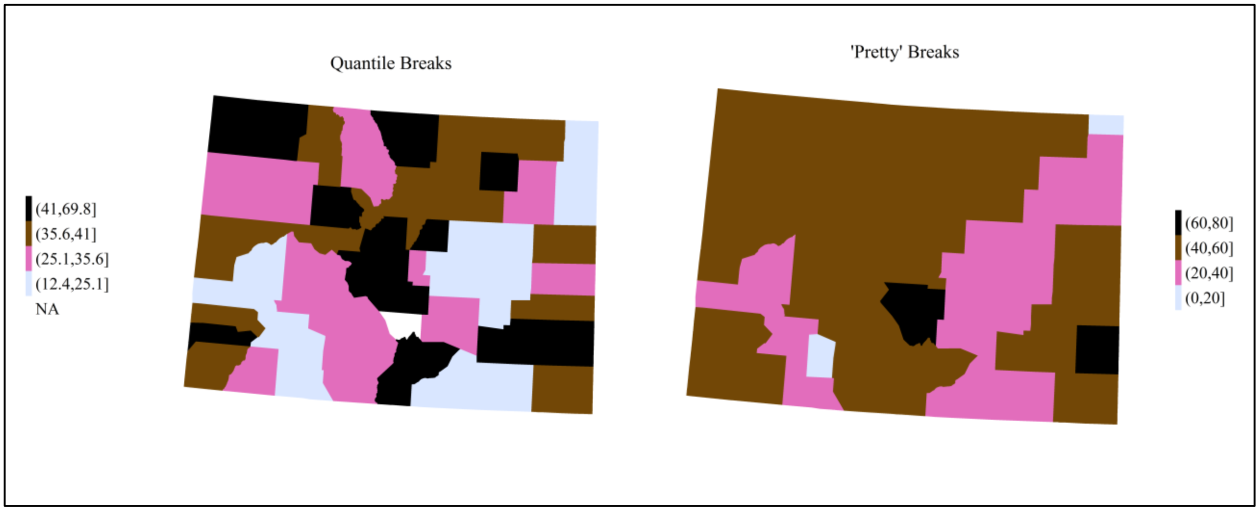


Figure 2-14. Two different binning methods.

To define the minimum and maximum of the color scale separately from the minimum and maximum of the data variable you are mapping, a two-item vector (with a minimum and maximum) can be passed to the map_colors_limits parameter. For those who wish to highlight or emphasize some portion of the data without using binned categories, the parameter map_color_breaks allows users to change how the color ramp is applied to the range of the data. Note that this approach does not change the values, or the "truth" presented by the maps-- it simply changes how the colors are stretched across the minimum and maximum of the scale. However, these changes can dramatically change the appearance of the data (Figure 2-15).

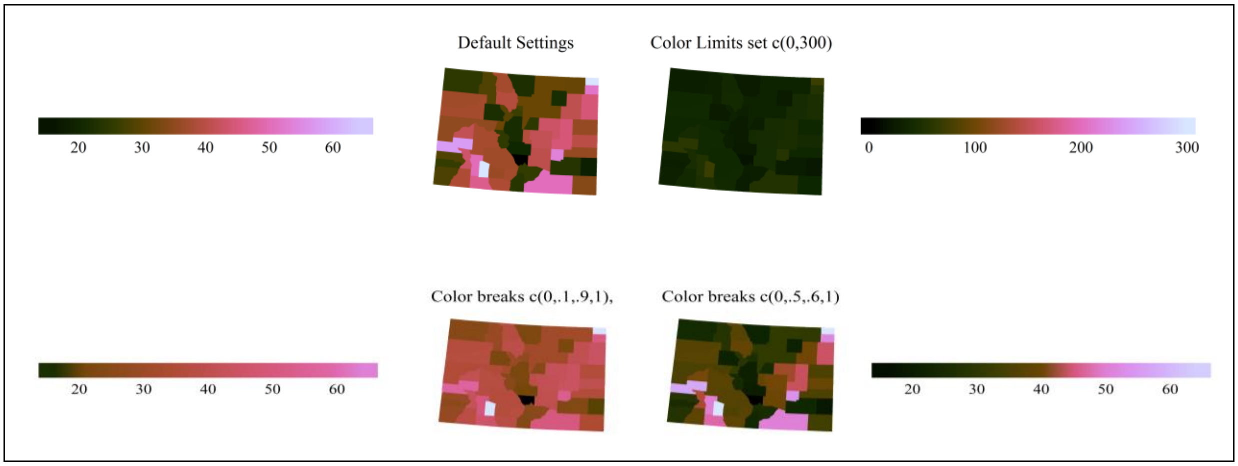


Figure 2-15. Changing how the colors are stretched across the data (i) Evenly from the minimum to the maximum (top left), (ii) Artificially extending the range of the colors to 0-300, which is larger than the range of the data (top right), and (iii) Changing whether light or dark colors are emphasized in the map by shifting how the colors are stretched across the data values (bottom).

If there is a meaningful central point or cutoff in your data, it is possible to set the numeric value that will serve as the central color in a diverging scale (Figure 2-16).

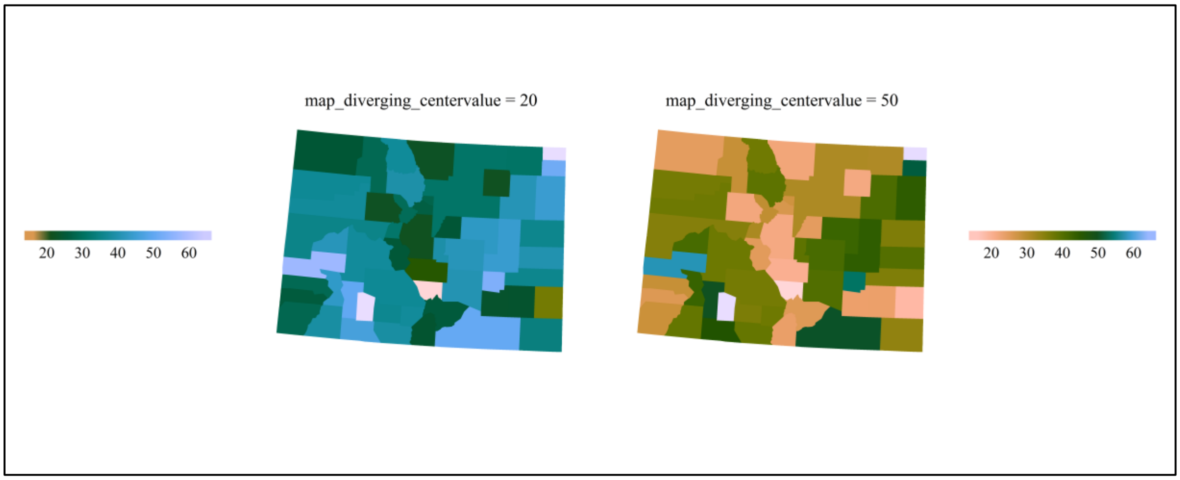


Figure 2-16. Using two different numeric values (20 and 50) to center the color scheme on a meaningful value.

2.5 PRESENT MANY NUMBERS IN A SMALL SPACE: COLOR-CODED HISTOGRAMS

While a good map can highlight spatial patterns, a histogram of the data being represented on the map illuminates the non-spatial trends in the distribution. As noted by Mark Monmonier, author of *How to Lie With Maps*, a secondary graphic that shows the data distribution is an important aspect of revealing the data's truth. He notes that "if the map author is at all concerned with full disclosure, a ...histogram... is a must.”[46]

However, visually linking a histogram to a 2-dimensional spatial distribution can be difficult. To make the comparison of spatial distributions and densities, the function HistogramColorStats provides extra context by linking histogram color to choropleth color. In addition, references to properties of the distribution can be referenced in the form of colored lines that highlight data parameters. Users can pass a vector of statistical terms (currently supported are the mean, median, standard deviation, and any number of quantiles) to the function, which will provide even more visual cues to help understand the patterns in the data. An example of the code used to generate a color-coded histogram is seen in Figure 2-17.

```
HistogramColorstats(datavector=county_simulated[["var2"]], color_ramp=wpal("seaglass"),
                title="Simulated variable 2",dist_stats=c("mean",.25,.75))
```

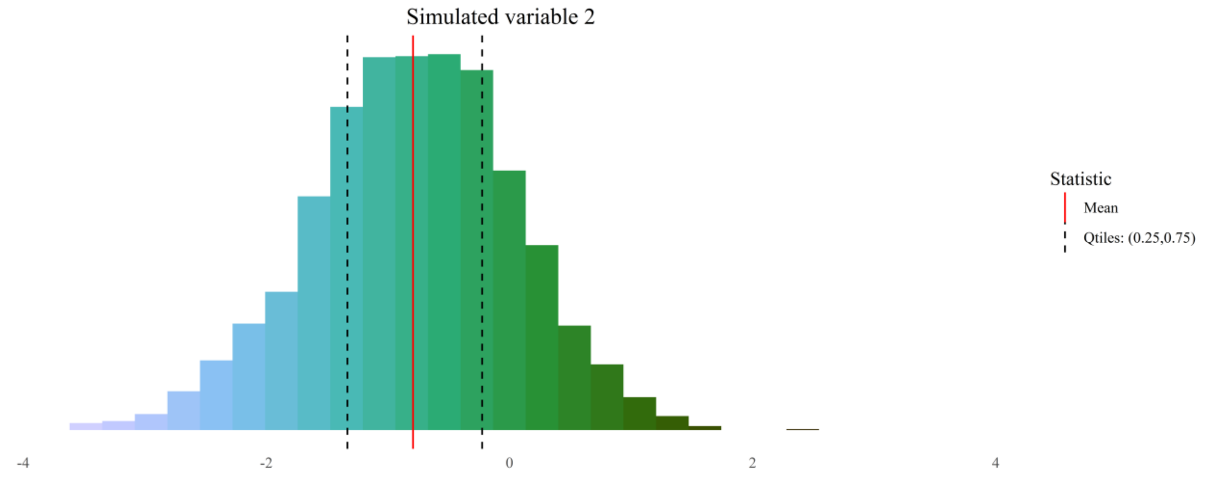


Figure 2-17. Generating a color-coded histogram.

This functionality is integrated into the main mapping functions with the use of a 'histogram' parameter. When set to TRUE, the histogram will be placed at the bottom of the map, with the same limits and scaling as the map (Figure 2-18). To modify the histogram or its position, users can generate the histogram separately, using the function HistogramColorstats(), and modify the placements and formatting of the histogram as desired.

```
RasterMap(coords=county_pixels[state_name=="Alaska"], id="cnty", xcol="x_albers", ycol="y_albers",
        data=county_simulated, variable="var3", map_colors=wpal("seaglass"),
        legend_position="left", legend_orientation="vertical", font_size=12,
        map_title="Simulated Data Distribution (Alaska)", legend_bar_length=unit(.5,"snpc"),
        histogram=T)
```

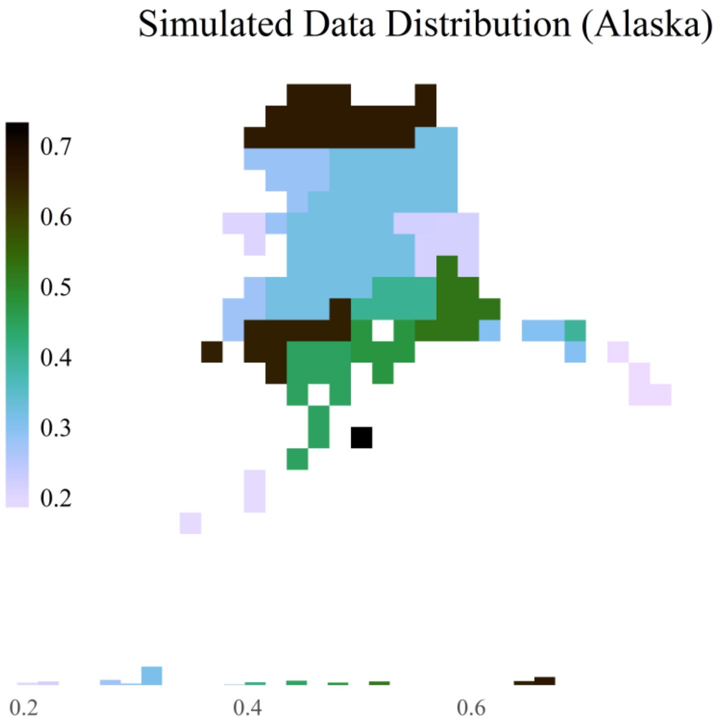


Figure 2-18. A map with a histogram of the data distribution.

2.6 ENCOURAGE THE EYE TO COMPARE DIFFERENT PIECES OF DATA: VISUALIZING BY SUBGROUP OR TIME

Even vast quantities of data (often referred to as "big data") can often be parsed into smaller, more illuminating subgroups. To map the same variable with repeated measurements in different time points, mapping software programs often require that these data are formatted "wide," such that each version of that variable is in its own separate column. Big, multi-dimensional data structures (such as a mortality rate by age, sex, and year) do not fit well in this paradigm. A more sensible structure (often used in data analysis) is the "long" format, where (i) the data identifiers or primary keys (such as the geometry's ID variable, age, sex, and time) each are represented as columns in a data set, and (ii) there is only one column for each variable of interest (available in different combinations of age, sex, and time).

Maintaining a consistent color classification scheme across these subgroups is a crucial part of comparing and contrasting the data. As Monmonier notes, "just as cut-points (data classification schemes) can be manipulated...pairs of choropleth maps can purposely heighten or suppress perceptions of bivariate association.”[46] MapSuite's mapping functions make this easy. By defining a column in the external data set as the 'series dimension' of the data, or the dimension of the data (for example, time or age) that you wish to iterate over, the functions will produce a series of maps (Figure 2-19). These maps will have the same ranges by default (consistent colors applied to each subset of the data based on the minimum and maximum of the observed data across all dimensions). If desired, the ranges used to produce the color scheme can be generated for each subset, by setting the map_colors_limits parameter to "each_dimension". One can restrict the quantity of maps made (if you are only interested in certain sub-groups of the data represented in the external data frame), providing a 'series sequence' that will restrict the loop of maps made to only the specified levels. By default, a subtitle will be generated that is the specific level or dimension of the data that is being mapped. For example, if the variable as the series dimension is 'year', the subtitle for each map would change to represent the specific time period represented in the graphic. When a series dimension is provided, text entered into the 'subtitle' parameter will serve as the prefix for the automatically generated subtitle based on the levels of the data that are being mapped.

```
PolygonMap(map=colorado, id="cnty", data=county_simulated_time,
        variable="var3", series_dimension="year",series_sequence=c(2000,2005,2009),
        map_colors = wpal("berries"), legend_position = "right",
        legend_orientation= "vertical", map_subtitle="Year: ", map_title="Var 3")
```

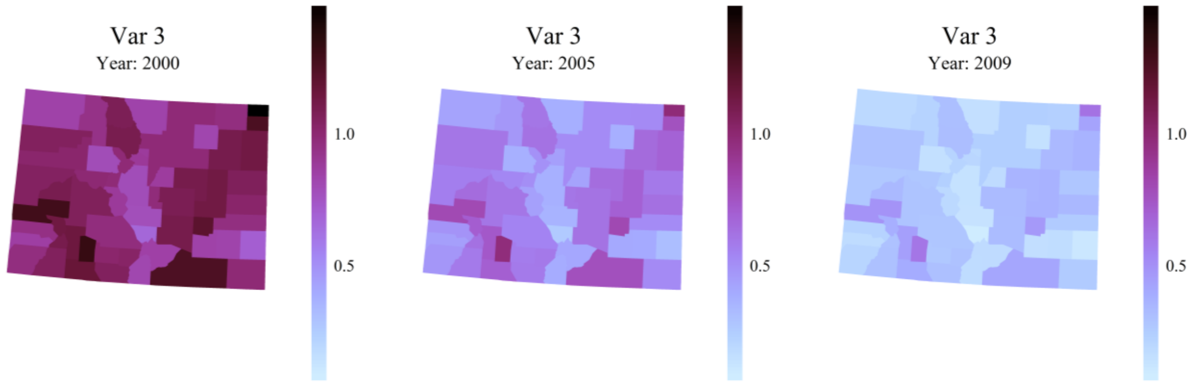


Figure 2-19. Three plots with standardized color schemes output from visualizing a time series. For the purposes of this illustration, these plots have been aligned horizontally, but would usually appear in separate plots.

If the data contains more than one dimension (such as age, sex, and time), a series of loops can be structured such that the data is subset, and plots can be created (Figure 2-20). To quickly create a PDF of maps of the variable, designate a PDF path, and a .pdf of the maps will be generated rather than printing the plot results to the screen or other viewport.

```
for(a in ages)@!?!
 for(s in sexes)@!?!
  PolygonMap(map=county_polygons, id="cnty",
      data=county_simulated_time[age==a&sex==s],
        variable="var3", series_dimension="year",
        map_subtitle="Year: ", pdf_path = paste0("var3",a,s,".pdf"))
 !?!@
!?!@
```

Figure 2-20. Example of code required to loop over multiple data dimensions.

To generate a series of complex plots in which some or all of the variables change over time or another dimension, a loop can be constructed to generate plots based on each subset of the data, combined with other map elements, such as a constant background map (Figure 2-21).

```
background<-PolygonMap(map=county_polygons, id="cnty",return_objects = T)$map

time_series<-PointMap(coords=county_centroids, data=county_simulated_time,
                      id="cnty", xcol="x_albers", ycol="y_albers",
                      variable="var2",series_dimension = "year",
                      return_objects=T)

for(y in c("2000","2005"))@!?!
  map<-StackMaps(maps=list(background,time_series$map[[y]]),
                 map_subtitle = time_series$subtitle[[y]])
  legends<-StackLegends(legends=list(time_series$map[[y]]))

  grid.newpage()
  pushViewport(viewport(layout = grid.layout(1, 3)))
  print(map, vp = vplayout(1, 1:2)); print(legends, vp = vplayout(1, 3))
!?!@
```

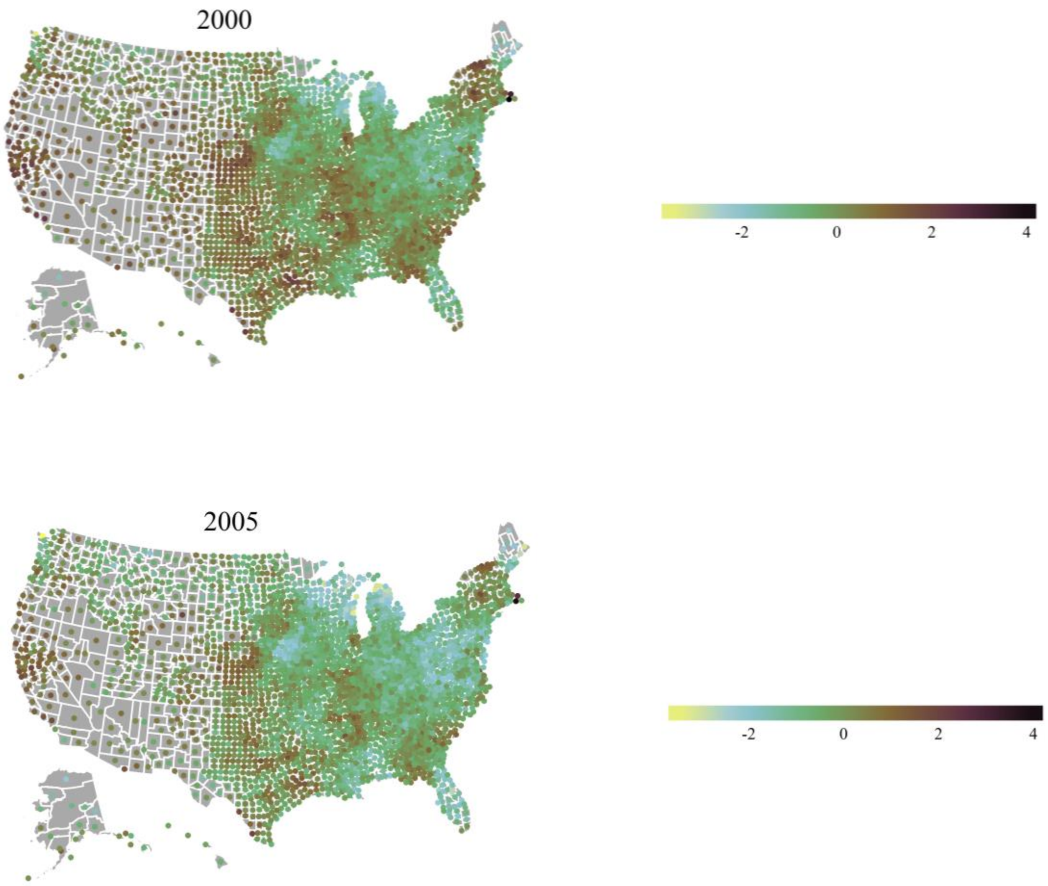


Figure 2-21. Generating multi-layered maps over time.

2.7 SUMMARY AND FURTHER WORK

A number of packages exist to visualize spatial data in R. The MapSuite package attempts to fill a niche that has the advantages of simplicity of use and graphic design, while remaining highly customizable for the more advanced user. Possible extensions and opportunities for future work on this package are the inclusion of line geometries, and the further development of functions and sensible default settings for legends that allow the stacking of map layers with even lower overhead. Interested users are invited to contribute to this effort on GitHub, where the full source code can be found at https://github.com/RebeccaStubbs/MapSuite.

REFERENCES

1. Woolhandler S, Himmelstein DU. Life or Debt: Underinsurance in America. *Journal of General Internal Medicine.* 2013;28(9):1122-1124.
2. Dusetzina SB, Basch E, Keating NL. For Uninsured Cancer Patients, Outpatient Charges Can Be Costly, Putting Treatments Out Of Reach. *Health Affairs.* 2015;34(4):584-591.
3. Martin AB, Hartman M, Washington B, Catlin A, Team tNHEA. National Health Spending: Faster Growth In 2015 As Coverage Expands And Utilization Increases. *Health Affairs.* 2016;36(1):166-176.
4. Himmelstein DU, Thorne D, Warren E, Woolhandler S. Medical Bankruptcy in the United States, 2007: Results of a National Study. *American Journal of Medicine.* 2009;122(8):741-746.
5. Ramsey S, Blough D, Anne Kirchhoff A, et al. Washington State Cancer Patients Found To Be At Greater Risk For Bankruptcy Than People Without A Cancer Diagnosis. *Health Affairs.* 2013;32(6):1143-1152.
6. Reinhardt UE. The Pricing Of U.S. Hospital Services: Chaos Behind A Veil Of Secrecy. *Health Affairs.* 2006;25(1):57-69.
7. National Center for Health Statistics USCB. Comparison of the Prevalence of Uninsured Persons from the National Health Interview Survey and the Current Population Survey,

2015 and 2016. 2016;
https://www.cdc.gov/nchs/data/nhis/earlyrelease/nchs_cps_comparison092016.pdf
Available from:
http://www.cdc.gov/nchs/data/nhis/health_insurance/NCHS_CPS_Comparison092016.pdf.

8. Services CgCfMM. NHE Fact Sheet: National Health Expenditure Data. 2017; https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html.
9. Anderson G. From 'Soak the Rich' To 'Soak the Poor': Recent Trends in Hospital Pricing. *Health Affairs (Millwood).* 2007;26(3):780-789.
10. Bai G, Anderson GF. Extreme Markup: The Fifty US Hospitals With The Highest Charge-To-Cost Ratios. *Health Aff (Millwood).* 2015 June;34(6):922-928.
11. Bai G, Anderson GF. US Hospitals Are Still Using Chargemaster Markups To Maximize Revenues. *Health Affairs.* 2016;35(9):1658-1664.
12. Ginsburg PB. Wide variation in hospital and physician payment rates evidence of provider market power. *Research Brief.* 2010;November(16):1-11.
13. P. WW, E. MM, G. WH, S. FE, E. WJ. Geographic variation in expenditures for physicians' services in the United States N Engl J Med 1993 328621-7. *M Engl J Med.* 1993;328(9):621-627.
14. Services CfMaM. Medicare Provider Charge Data: Inpatient Public Use Data Files FY 2011-2014. 2016; https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Inpatient.html.
15. Park JD, Kim E, Werner RM. Inpatient Hospital Charge Variability of U.S. Hospitals. *Journal of General Internal Medicine.* 2015;30(11):1627-1632.

16. Google. Google API Console. *Google* 2017; https://console.developers.google.com/apis/library.

17. Center: MP. Minnesota Population Center. National Historical Geographic Information System: Version 11.0 [Database]. 2016; http://doi.org/10.18128/D050.V11.0.

18. Bowers L, Gann C, Elser S. *Small Area Health Insurance Estimates: 2015.* Washington, D. C.2017.

19. Bureau of the Census USDoC. *U.S. Census Bureau, Small Area Income and Poverty Estimates (SAIPE): 2015 Highlights.* Washington, D.C.: U.S. Government Publishing Office;2016.

20. Statistics BoL. Local Area Unemployment Statistics. 2016; https://www.bls.gov/lau/lauov.htm.

21. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, al. e. US county-level trends in mortality rates for major causes of death, 1980-2014. *JAMA.* 2016;316(22):2385-2401.

22. R Development Core Team (2011) RALaEfSCV, Austria : the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/. R: A Language and Environment for Statistical Computing. 2011. http://www.R-project.org/.

23. Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software; Vol 1, Issue 5 (2016).* 2016.

24. Services CgCfMM. Hospital-Acquired Condition Reduction Program (HACRP). 2017; https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/HAC-Reduction-Program.html.

25. Burke LA, Ryan AM. The Complex Relationship between Cost and Quality in US Health Care. *Virtual Mentor.* 2014;16(2):124-130.

26. Patient Protection and Affordable Care Act USC. Patient Protection and Affordable Care Act, 42 U.S. Code § 18001 - Immediate access to insurance for uninsured individuals with a preexisting condition. 2010; https://www.law.cornell.edu/uscode/text/42/18001 https://archive.org/stream/392145-the-patient-protection-and-affordable-care-act/392145-the-patient-protection-and-affordable-care-act_djvu.txt.

27. Sieber R. Public Participation Geographic Information Systems: A Literature Review and Framework. *Annals of the Association of American Geographers.* 2006;96(3):491-507.

28. Steiniger S. An overview on current free and open source desktop GIS developments. *International Journal of Geographical Information Science.* 2009;23(10):1345-1370.

29. Pebesma EJ, Bivand RS. Classes and methods for spatial data in R. 2005;5(2):9-13.

30. Wickham H. *ggplot2: Elegant Graphics for Data Analysis* New York: Springer-Verlag; 2009.

31. Kahle D, Wickham H. ggmap: Spatial Visualization with ggplot2. 2013;5(1):144-161.

32. Dunnington D. ggspatial: Spatial Data Framework for ggplot2. 2017. https://cran.r-project.org/web/packages/ggspatial/index.html.

33. Giraud T. rCarto: This package builds maps with a full cartographic layout. 2013. https://CRAN.R-project.org/package=rCarto.

34. Lamstein A, Johnson BP. choroplethr: Simplify the Creation of Choropleth Maps in R 2017. https://CRAN.R-project.org/package=choroplethr.

35. Cheng J, Karambelkar B, Xie Y. leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. 2017. URL https://CRAN.R-project.org/package=leaflet.

36. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. Shiny: web application framework for R. *R package version 011*. 2015;1.

37. Wickham H. A Layered Grammar of Graphics. *Journal of Computational and Graphical Statistics*. 2010;19(1):3-28.

38. Tufte ER. *The Visual Display of Quantitative Information.* Graphics Press; 1983.

39. Bureau USC. TIGER/Line Shapefile, 2015 Cartographic Boundary File, State-County for United States, 1:20,000,000. 2017; https://catalog.data.gov/dataset/2015-cartographic-boundary-file-state-county-for-united-states-1-20000000.

40. Tyner J. *Principles of map design.* 2014.

41. Tyner JA. Persuasive cartography. *Journal of Geography.* 1982;81(4):140-144.

42. Light A, Bartlein PJ. The end of the rainbow? Color schemes for improved data graphics. 2004;85(40).

43. Brewer CA, MacEachren AM, Pickle LW, Herrmann D. Mapping mortality: Evaluating color schemes for choropleth maps. *Annals of the Association of American Geographers.* 1997;87(3):411-438.

44. Neuwirth E. RColorBrewer: ColorBrewer Palettes. 2014. https://cran.r-project.org/web/packages/RColorBrewer/index.html.

45. Evans R. *rje: Miscellaneous useful functions*. 2014. https://CRAN.R-project.org/package=rje.

46. Monmonier M. Lying with Maps. *Statist Sci.* 2005;20:215-222.

APPENDIX A

This appendix contains supplementary tables and figures pertaining to the hospital charge markup analysis detailed in Chapter 1.

Table A-1. Summary Statistics of the top 15 DRG categories in 2014, all states combined.

Disease reporting group	**Mean**	**Std.**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**N**
All Combined	5.00	2.60	0.15	3.24	4.44	6.18	33.00	7216853
Infectious & parasitic diseases, systemic or unspecified sites	4.46	2.28	0.56	2.91	3.99	5.51	22.60	680529
Injuries, poisonings & toxic effects of drugs	4.87	2.55	0.56	3.17	4.29	5.98	20.25	58953
Factors influencing health status & other contacts with health services	5.25	2.79	0.66	3.39	4.71	6.62	24.75	46081
Alcohol/drug use & alcohol/drug induced organic mental disorders	4.41	2.91	0.15	2.39	3.80	5.59	20.25	45217
Diseases # Disorders:								
circulatory system	5.22	2.67	0.30	3.40	4.66	6.44	30.97	1666746
respiratory system	4.87	2.57	0.51	3.08	4.29	6.08	26.50	1086318
musculoskeletal system & connective tissue	4.87	2.32	0.49	3.28	4.41	5.96	21.67	900426
digestive system	5.00	2.55	0.58	3.28	4.44	6.13	33.00	750694
kidney & urinary tract	4.79	2.56	0.37	3.07	4.20	5.94	31.63	592717
nervous system	5.42	2.82	0.65	3.52	4.79	6.71	29.78	521966
endocrine, nutritional & metabolic	4.86	2.55	0.63	3.17	4.31	6.02	32.97	249920
skin, subcutaneous tissue & breast	4.67	2.52	0.49	2.94	4.09	5.80	24.84	156402

hepatobiliary system & pancreas	5.22	2.65	0.74	3.47	4.67	6.42	23.58	125272
mental	4.01	2.36	0.68	2.40	3.43	5.10	17.44	104030
blood, blood forming organs, immunologic disorders	4.94	2.50	0.65	3.24	4.43	6.12	26.44	103864

Table A-2: Summary Statistics of Markup by State in 2014 for all DRGs in states with a mean markup greater than 4.00.

STATE	MEAN	STD.	MIN	Q1	MEDIAN	Q3	MAX	N
New Jersey	8.38	3.39	0.78	6.05	7.73	10.02	33.00	258611
Nevada	7.73	3.07	1.88	4.90	7.91	9.88	20.37	46252
Florida	7.19	3.01	1.03	5.06	6.57	8.78	29.78	594009
California	6.54	2.53	0.36	4.79	6.18	7.94	22.10	494022
Texas	6.23	2.57	0.53	4.35	5.82	7.68	22.73	495345
Colorado	6.19	2.24	1.31	4.61	5.86	7.48	17.06	66216
Pennsylvania	5.77	2.82	0.15	3.57	5.32	7.47	24.18	316731
Alabama	5.75	3.29	0.65	3.65	4.90	6.65	26.03	152111
South Carolina	5.68	2.37	1.25	4.01	5.07	6.87	24.28	113939
Arizona	5.44	2.27	0.37	3.98	5.09	6.41	24.75	111269
Illinois	5.18	1.87	0.56	3.93	4.91	6.20	19.49	351597
Kansas	5.16	1.89	0.65	3.94	5.02	6.20	14.79	71273
Mississippi	5.03	2.61	0.83	3.27	4.24	6.17	20.25	91670
Louisiana	4.87	2.07	0.30	3.51	4.60	5.91	15.16	109571
Tennessee	4.77	2.05	0.56	3.37	4.41	5.77	24.01	200056
Virginia	4.75	2.40	1.28	3.18	4.07	5.43	23.20	214416
Oklahoma	4.72	2.29	0.49	3.31	4.23	5.59	18.69	99552
Missouri	4.70	1.95	0.80	3.40	4.26	5.57	17.41	175765
Georgia	4.63	1.71	0.79	3.41	4.31	5.49	14.97	195162
Washington	4.49	1.54	1.04	3.38	4.24	5.37	13.37	117004
Indiana	4.44	1.42	0.88	3.50	4.23	5.14	15.17	182826
Nebraska	4.40	1.31	1.12	3.51	4.22	5.08	12.21	44448
Ohio	4.40	1.60	1.12	3.27	4.19	5.29	14.64	284040
Arkansas	4.38	1.96	0.95	3.15	4.14	5.25	16.27	86213
District of Columbia	4.36	1.17	1.21	3.60	4.41	5.05	9.56	26278
Kentucky	4.28	1.46	1.36	3.30	4.04	4.99	17.12	148772
Alaska	4.27	1.32	0.47	3.49	4.32	5.12	8.03	6515
New Mexico	4.16	1.98	0.72	2.76	3.67	5.47	12.13	27867
South Dakota	4.03	1.19	0.67	3.20	3.85	4.79	10.21	23575

Table A-3: Summary Statistics of Markup by State in 2014 for all DRGs in states with a mean markup less that 4.00.

STATE	MEAN	STD.	MIN	Q1	MEDIAN	Q3	MAX	N
Iowa	3.94	1.19	0.62	3.06	3.81	4.66	10.53	69970
North Carolina	3.91	1.30	1.15	2.99	3.69	4.59	13.70	235267
Wisconsin	3.88	1.27	0.86	3.03	3.68	4.54	11.60	106722
New Hampshire	3.87	1.36	1.52	2.89	3.54	4.58	10.12	38381
New York	3.87	1.72	0.47	2.59	3.60	4.84	14.64	430086
Connecticut	3.60	1.08	0.83	2.87	3.53	4.25	9.87	100251
Rhode Island	3.59	1.01	0.89	2.94	3.49	4.19	7.58	24572
Idaho	3.59	0.99	1.62	2.92	3.38	4.03	7.77	20765
Utah	3.52	1.11	1.22	2.71	3.38	4.06	9.89	28512
West Virginia	3.42	1.00	0.71	2.73	3.35	4.00	8.51	56323
Minnesota	3.41	1.01	0.30	2.69	3.30	4.01	11.02	117759
Delaware	3.36	1.11	1.15	2.58	3.11	3.95	8.11	35510
Hawaii	3.33	1.06	0.76	2.56	3.17	3.86	7.21	12211
Oregon	3.30	0.97	1.04	2.61	3.17	3.84	8.45	47557
Wyoming	3.29	1.21	1.10	2.36	3.18	3.97	9.18	7100
Montana	3.27	0.78	1.17	2.72	3.21	3.68	6.66	18205
Michigan	3.17	0.95	0.68	2.49	3.07	3.72	13.65	310619
North Dakota	3.05	0.81	1.26	2.52	2.96	3.42	10.20	19220
Maine	3.02	0.73	1.03	2.55	2.96	3.41	6.71	33873
Vermont	2.56	0.63	1.12	2.12	2.54	2.94	5.68	10393
Massachusetts	2.39	1.10	0.45	1.65	2.07	2.76	8.86	199370
Maryland	1.22	0.09	1.09	1.16	1.21	1.26	2.29	189082

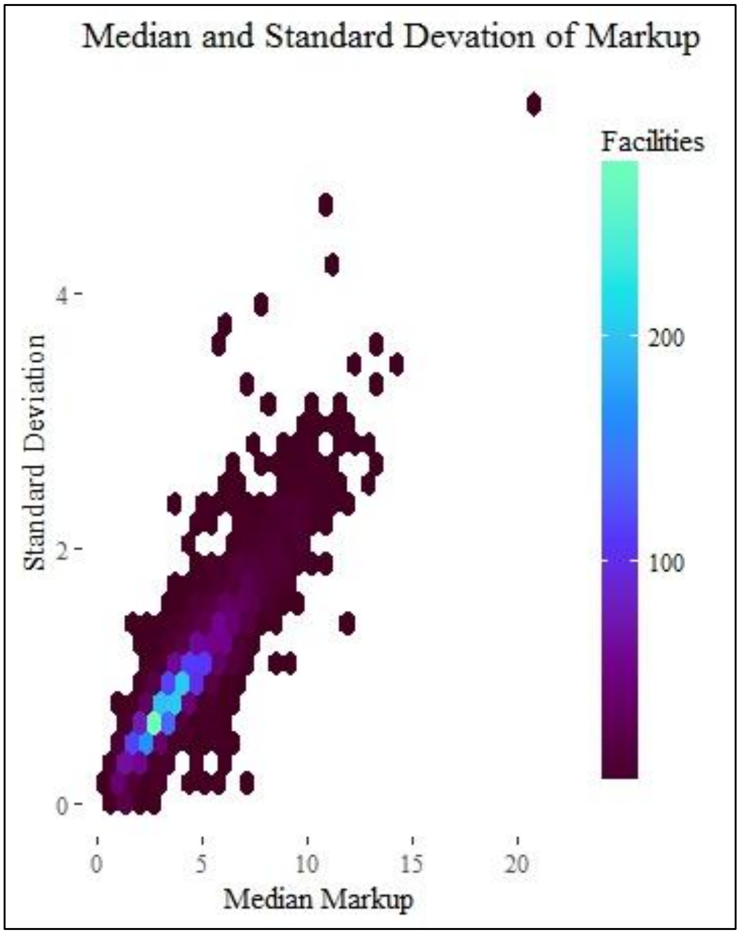


Figure A-1. Correlation between the median and standard deviation of markup.

APPENDIX B

This appendix contains the named color palettes available in the MapSuite's wpal() function.

salmon_run

sky_to_sea

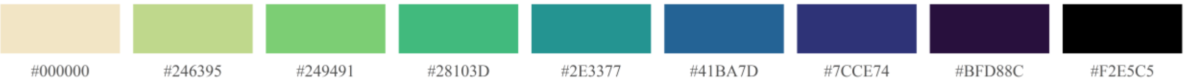
tropical_sunrise

parrotfish

seaglass

betafish

seaside



skyforest

bright_greens

bright_fire

bright_cool

bright_roygbiv

cool_blue_deepindigo

#070211 #110730 #1C1055 #261D83 #3133B5 #3D4DE1 #4C70F6 #6796FF #89C0FF

cool_blue_aqua

#002032 #003E53 #005E71 #007F8A #009F9C #04BEAB #1AD7B7 #44EBC4 #7CFBD3

cool_blue_bright

#0B1042 #102474 #143AA2 #1955CC #2173EA #2C92FF #42B0FF #64CCFF #8FE6FF

cool_blue_steel

#001C0D #00351F #015034 #046B4F #12866F #269F8F #45B7B0 #6DCDCE #9CE1E9

cool_blue_jeans

#00181D #022D3B #09435C #155981 #2B70A5 #4486C6 #669DE2 #8DB5F4 #B7CEFF

cool_green_grassy

#002800 #004700 #076500 #198100 #359A00 #55B003 #7DC313 #A4D43B #CCE474

cool_green_happy

#002A13 #004C1C #006E24 #00912A #00AF32 #07CA3C #2DDF4F #5EEE6D #95FA94

cool_green_deepforest

#141300 #232800 #334000 #425B00 #53780E #649522 #78B143 #93CB6C #B3E49D

cool_green_deeplake

#051A00 #093303 #0F4E0C #166C1B #238932 #33A64E #4CBF71 #70D597 #9BEAC0

cool_earthtones

#321E21 #4A362D #5E503B #6B6D4B #778965 #83A383 #97B9A6 #B1CCC6 #CFDFE4

cool_stormy

#2A2416 #394026 #465B3C #52765A #658D7C #7CA19F #9AB3BE #BAC4D6 #DBD7EA

warm_fire

#440022 #7C002A #B1042D #E00F27 #F4291E #FF4819 #FF7022 #FF9B40 #FFC86D warm_darkfire

warm_purple1

warm_purple2

warm_adultpink

warm_kidpink

warm_redpink

warm_mauve

warm_darkpeach

warm_brown

ld_reg1

ld_reg2

ld_reg3

ld_bright1

ld_bright2

ld_bright3

ld_muted1

ld_muted2

ld_muted3

diverging_intensity_purple_green

diverging_intensity_blue_red

diverging_pink_black_blue

diverging_orange_black_blue

diverging_green_black_purple

diverging_tan_black_green_multi

diverging_pink_white_blue_mutli

#4F196F #562D00 #6254D0 #67A1F8 #95E1F4 #BB452D #F96A92 #FFAAEA #FFFFFF

diverging_purple_white_blue

#103971 #11879F #39C8A0 #731332 #95EEAE #B73497 #C771F1 #CCBFFF #FFFFFF

diverging_tan_white_green_multi

#174D00 #4F196F #6254D0 #67A1F8 #6A7800 #95E1F4 #CB913A #FFB7A4 #FFFFFF

diverging_green_purple_pink

#0F6E95 #24B2A1 #2E2971 #68E0A2 #932B95 #CCF7CC #E351A1 #FF8FA2 #FFDBCC

diverging_orange_purple_blue

#5F3FB5 #611551 #6483EE #7AC7F7 #AA3036 #C3F5F3 #D06E38 #D5B560 #E2EEBE

diverging_tan_green_blue

#057B75 #0B5019 #3B9DD3 #4C6C00 #96B6FF #A7881D #E9DDFF #F0A373 #FFD6D8

diverging_green_blue_purple

#000F12 #00320E #131A49 #146B08 #562F99 #6BA01A #BE54D1 #DED38C #FFAEE6

diverging_pink_brown_green

#383C00 #387503 #50B12F #7BDE7D #815003 #C6F8D5 #C8742F #F89E7D #FFD8D5

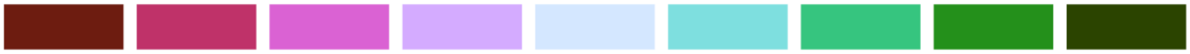
diverging_blue_green_pink

#21082B #331E6B #504EBE #5D0F4C #7392F9 #AF2B76 #BBDDFF #F26299 #FFBFCD

diverging_blue_lightpurple_pink

#1C0A30 #270627 #3B2882 #585ED2 #721358 #82A6FC #C43680 #DDCEE6 #F978A4

diverging_pink_light_green



#24901B #2B4400 #36C57F #6D1C10 #7EDFDF #BF3269 #D4ABFF #D4E7FF #DA62D3

intensity_tan_red

#540028 #804138 #8C0026 #8D1C2E #92573E #9C7046 #A08961 #A5A182 #B2B898 intensity_deepgreen_lavender

intensity_pink_purple

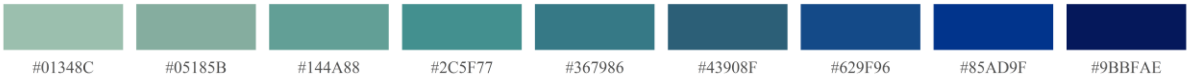
intensity_lightgreen_darkbrown

intensity_lavender_darkgreen

intensity_seagreen_purple

intensity_pastel_purple

intensity_seagreen_blue



diverging_succulents