

Structural Insights 2023: A Linear Analysis of Apartment Age and Safety Scores in Toronto*

Terry Tu, Jingyi Shen, Yaning Jin

March 14, 2024

This study harnesses a dataset from the Toronto Open Data portal, comprising 3,435 records of apartment evaluations, to analyze the correlation between building age and safety scores. Employing a linear regression model, we discovered a significant inverse relationship, highlighting that older buildings tend to have lower safety evaluations. The findings suggest a need for policy reform, particularly in the maintenance and inspection of Toronto’s aging residential infrastructure. Despite the study’s focus on building age, further research considering additional variables could offer a more comprehensive understanding of factors influencing building safety.

Table of contents

1	Introduction	2
2	Data	2
3	Model	6
4	Result	8
5	Discussion	11
6	Appendix	12
	References	12

*Code and data supporting this analysis are available at <https://github.com/TEJMaster/Toronto-Building-Safety-Analysis-2023.git>

1 Introduction

In the wake of the catastrophic collapse of an aged condominium in South Florida in June 2021, which resulted in 98 casualties, a subsequent lawsuit by the bereaved families underscored the notion that routine inspections and maintenance could have forestalled the disaster ([Anderson 2021](#)). This incident precipitates our investigation into forecasting the building evaluation scores for Toronto residences in 2023 via a simple linear regression model. An analysis was conducted on the apartment building evaluation dataset obtained from the City of Toronto’s open data portal to assess the relationship between the age of a property and its evaluation score assigned by the city. Our research is singularly focused on elucidating the effect of one independent variable—building age—on the dependent variable, which in this context is the building evaluation score.

Using simple linear regression analysis enables a detailed exploration of the linear relationship between the age of a building and its evaluation score. By quantifying this relationship through regression coefficients, we aim to reveal the expected change in evaluation scores for each incremental year in building age. The statistical significance of the age variable on evaluation scores will be ascertained using p-values. Provided the model satisfies the underlying assumptions, t-tests will be applied to test the hypotheses, while confidence and prediction intervals will be leveraged to quantify the effect size and enhance the accuracy of our forecasts, respectively.

2 Data

2.1 Raw Data

In this research, we examine apartment data sourced from the Toronto Open Data portal. The dataset encompasses 3,435 records, focusing on specific variables: the ward of each apartment, its evaluation score, and the construction year. The data covers evaluations across Toronto’s 25 wards, with scores varying between a minimum of 0 and some achieving the maximum score of 100. The study includes apartments that are over three stories in height and contain at least 10 units. These apartments span a wide range of construction years, from as early as 1805 to as recent as 2023. All the apartments in this dataset adhere to the City’s Apartment Building Standards bylaw, which is designed to safeguard the interests and safety of both landlords and tenants by mandating these evaluations ([Toronto 2023](#)).

2.2 Data Analysis Tools

The data analysis was performed using R ([R Core Team 2022](#)), a powerful open-source statistical programming language. Key packages from the tidyverse collection ([Wickham et al. 2019](#)) were employed to streamline data manipulation, visualization, and analysis processes. These

packages include ggplot2 (Wickham 2016) for creating advanced graphics, dplyr (Wickham et al. 2022) for data manipulation, readr (Wickham, Hester, and Bryan 2022) for its robust data reading functionality, here (Müller 2020) is used to avoid file path issue, and knitr (Xie 2014) for dynamic report generation.

2.3 Variable Description

CURRENT BUILDING EVAL SCORE: The Current Building Evaluation Score represents a measure of a building’s adherence to property standards. This score is derived by summing two components: the current reactive score, which reflects the compliance based on any outstanding Orders and Notices of Violations, and the proactive building score, which is based on the most recent comprehensive evaluation of the building.

YEAR BUILT: This indicates the year when the building was initially constructed, sourced from Building Owners/Managers.

YEAR EVALUATED: This denotes the year in which the building underwent evaluation, reflecting its condition and performance.

AGE: In this research, neither YEAR BUILT nor YEAR EVALUATED was used directly, we have generated a new variable called AGE which use YEAR EVALUATED minus YEAR BUILT.

2.4 Sample of Cleaned Building Evaluation Data

Table 1: Sample of Building Evaluation Data

X_id	AGE	CURRENT.BUILDING.EVAL.SCORE
1	62	86
2	44	94
3	54	78
4	73	89
5	54	99
6	53	99

Table 1 represents a subset of the larger building evaluation data. Each row corresponds to a unique building identified by X_id. The AGE column indicates the age of the building in years, while the CURRENT.BUILDING.EVAL.SCORE column shows the building’s current evaluation score on a scale from 0 to 100, with higher scores indicating better compliance with safety standards.

2.5 Data Exploration:

2.5.1 Data Summary

Table 2: Summary Statistics for the Cleaned Dataset

Count	Mean Age	SD of Age	Mean Score	SD of Score
1748	60.644	19.939	87.504	8.693

Table 2 presents a concise overview of the key descriptive metrics for the age of buildings and their evaluation scores within the dataset:

- **Count:** The dataset encompasses a total of 1,748 buildings, indicating a robust sample size for statistical analysis.
- **Mean Age:** The average age of the buildings is approximately 60.6 years, suggesting that the dataset primarily includes buildings that have been standing for over half a century.
- **Standard Deviation of Age:** With a standard deviation of about 19.9 years, there is considerable variability in the ages of the buildings. This range implies a diverse sample, encompassing both relatively new constructions and much older structures.
- **Mean Score:** The average building evaluation score is approximately 87.5 on a scale from 0 to 100. A higher score corresponds to better adherence to safety and maintenance standards, suggesting that the buildings in this dataset are generally well-maintained.
- **Standard Deviation of Score:** The standard deviation of approximately 8.7 for the evaluation scores indicates that while scores are clustered around a high mean, there is still notable variability, reflecting differences in the condition or compliance of the buildings.

These statistics offer a snapshot of the dataset’s characteristics, providing foundational insights for further analysis regarding the state of building safety and integrity in the population studied.

2.5.2 Age Distribution

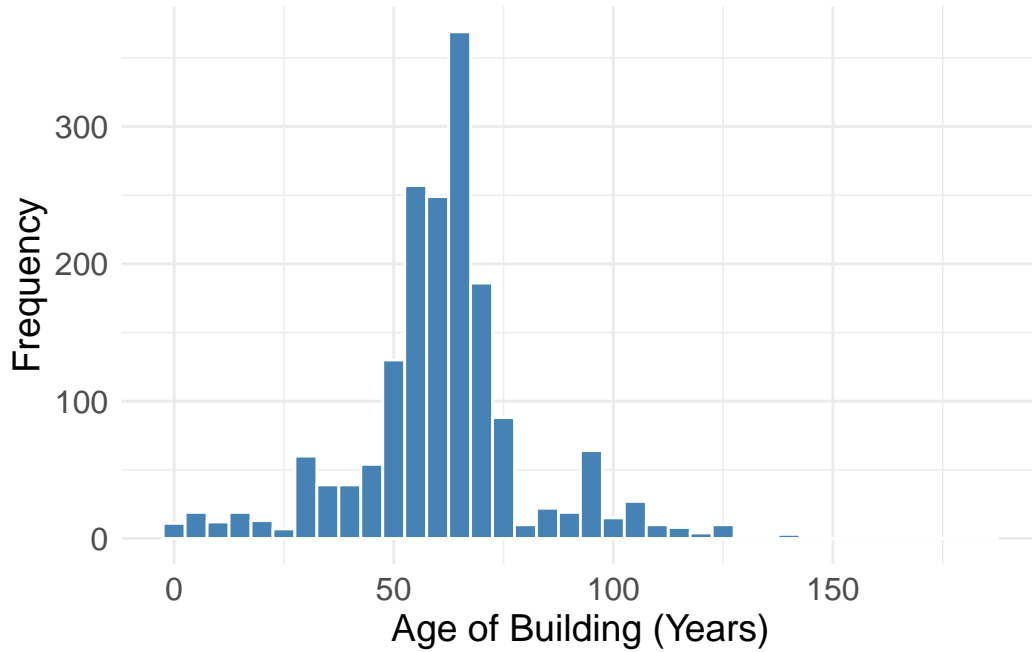


Figure 1: Distribution of building's age for the cleaned data set

Figure 1 illustrates the frequency distribution of the ages of buildings within the dataset. The x-axis represents the age of the buildings in years, while the y-axis shows the frequency of buildings within each age range. The distribution appears to be unimodal with a peak around the 50-year mark, indicating that a substantial number of the buildings within this dataset were constructed around half a century ago. Following this peak, the frequency gradually decreases, showing that fewer buildings have been in existence beyond this age. The distribution suggests a relatively young building population, with ages concentrated below 100 years, and very few buildings reaching beyond 120 years of age. This visualization aids in understanding the age profile of buildings under consideration for safety evaluation scoring.

2.5.3 Score Distribution

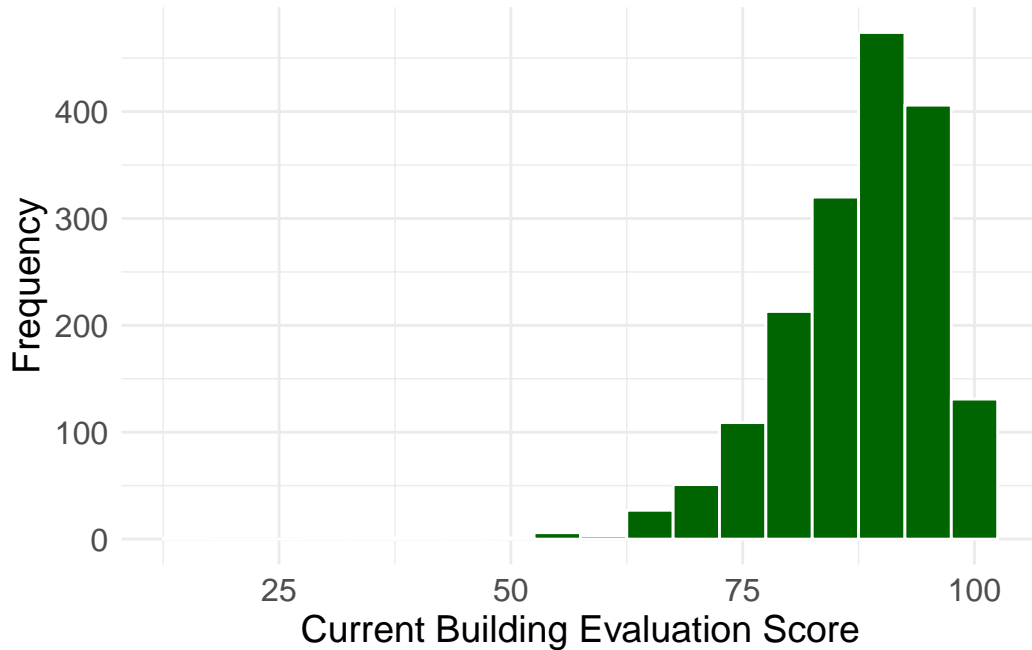


Figure 2: Distribution of building’s evaluation score for the cleaned data set

Figure 2 showcases the distribution of building evaluation scores. The x-axis represents the scores assigned to buildings, which can range from 0 to 100, with the y-axis displaying the frequency of buildings for each score interval. The chart reveals a concentration of high evaluation scores, suggesting that many of the buildings in the dataset are rated as being in good or excellent condition according to the evaluation criteria. The tallest bars are clustered towards the higher end of the score range, indicating that buildings with scores close to 100 are prevalent. This positive skew in the distribution might reflect effective building standards and maintenance practices in place. The spread of the scores and their frequency can offer insights into the overall health of the evaluated buildings, influencing urban planning and policy development.

3 Model

The objective of our model is dual: to discern the relationship between building age and safety evaluation scores and to quantify this relationship’s strength and direction. The study utilizes a simple linear regression model to analyze the dataset obtained from the Toronto Open Data portal.

3.1 Model set-up

Let y_i represent the safety evaluation score for each building. The building's age is denoted by x_i . The model is defined as follows:

$$y_i \sim \text{Normal}(\mu_i, \sigma) \quad (1)$$

$$\mu_i = \beta_0 + \beta_1 \times x_i \quad (2)$$

$$\beta_0 \sim \text{Normal}(90, 10) \quad (3)$$

$$\beta_1 \sim \text{Normal}(-0.1, 0.05) \quad (4)$$

The model is implemented in R using the `rstanarm` package, which allows us to incorporate prior beliefs about the parameters and to estimate the posterior distributions. For the intercept β_0 and slope β_1 , we choose normal prior distributions centered around our initial estimates based on exploratory data analysis. The `stan_glm` function from `rstanarm` is utilized to carry out the Bayesian regression, providing a framework for inference that incorporates both the data and our prior knowledge.

The priors are chosen to reflect reasonable beliefs about the parameters before observing the data. The intercept prior reflects an expectation that a new building will have a high safety score around 90, with a wide spread to accommodate uncertainty. Similarly, the slope prior indicates an expected small negative impact of age on the safety score, acknowledging some uncertainty in the precise effect. The `stan_glm` function combines these priors with the data to estimate the posterior distributions of the model parameters.

3.2 Model justification

We hypothesize that the coefficient for building age (β_1) will be negative, reflecting the tendency for older buildings to exhibit lower safety scores. This expectation is rooted in the general understanding that buildings can deteriorate over time due to wear, aging materials, and possible obsolescence in older construction methods. The intercept (β_0), representing the baseline evaluation score for a new building, is presumed to be relatively high, indicating the score a building would receive if it were built in the current period with modern standards. The slope (β_1) is therefore of particular interest, as it encapsulates the average annual decline in evaluation score attributable to aging. The Bayesian framework allows us to incorporate these hypotheses as priors in our model, thereby aligning the analysis with our substantive expectations and providing a more nuanced interpretation of the data.

4 Result

4.1 Model Coefficients Interpretation

(Intercept)	AGE
93.5598517	-0.1005911

The regression model provides two key coefficients to understand the relationship between building age and the evaluation score:

Intercept (β_0):

The model's intercept, β_0 , with a mean posterior estimate of 93.55985, represents the baseline evaluation score for a hypothetical newly constructed building. In Bayesian terms, this value is not a single point estimate but rather the mean of the posterior distribution for the intercept. This distribution reflects our belief about the probable values of the evaluation score for a new building, informed by the data and our prior assumptions.

Slope (β_1):

Similarly, the slope parameter β_1 has a mean posterior estimate of -0.10059. It describes the average decrease in the evaluation score associated with each additional year of a building's age. Again, this is the mean of the posterior distribution for the slope, encompassing both our prior beliefs and the observed data. A mean estimate that is negative suggests a trend where older buildings have lower evaluation scores.

Interpretation: The Bayesian analysis suggests that there is a likely negative relationship between building age and evaluation score. The mean estimates of the intercept and slope, along with their posterior distributions, indicate that newer buildings tend to have higher evaluation scores, which decrease as buildings age. This trend highlights the importance of age as a factor in building evaluations and underscores the need for ongoing maintenance and updates to preserve building safety and integrity over time.

4.2 Model Equation

$$y_i \sim \text{Normal}(\mu_i, \sigma) \quad (5)$$

$$\mu_i = \beta_0 + \beta_1 \times x_i \quad (6)$$

$$\beta_0 = 93.55985 \quad (7)$$

$$\beta_1 = -0.10059 \quad (8)$$

In this model, y_i represents the predicted evaluation score for the i -th building, x_i represents the age for the i -th building, μ_i is the mean of the normal distribution for the i -th observation, β_0 is the intercept, and β_1 is the slope representing the change in evaluation score per year of building age.

4.3 Analysis of Residuals

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-69.61906	-4.32025	1.75670	0.04447	5.87798	15.89571

1. Minimum Residual: The minimum residual is approximately -69.619, indicating that there is at least one building for which the model's predicted evaluation score was lower than the actual score by approximately 69.662 points. This could be an outlier or a building with exceptional characteristics not captured by the model.

2. First Quartile Residual: The first quartile (Q1) of the residuals is approximately -4.320, meaning that 25% of the buildings have evaluation scores that are less than 4.364 points below the model's prediction for their age.

3. Median Residual: The median residual is approximately 1.757, suggesting a slight positive bias in the model's predictions, as the median of the error distribution is above zero. This indicates that, on average, the model tends to slightly overestimate the evaluation scores.

4. Third Quartile Residual: The third quartile (Q3) is approximately 5.878, indicating that 75% of the buildings have evaluation scores within 5.833 points of the model's predictions or better. This shows that the model provides reasonably accurate predictions for the majority of buildings.

5. Maximum Residual: The maximum residual is approximately 15.896, suggesting that there is at least one building for which the model's predicted evaluation score was higher than the actual score by approximately 15.896 points. This could be another outlier or a building that performed exceptionally well.

Interpretation: The residuals provide insights into the model's fit and the distribution of errors. The presence of outliers and the distribution's slight asymmetry warrant further investigation. It's crucial to assess the model's assumptions, such as linearity, homoscedasticity, and normality of residuals, to ensure the validity of the statistical inferences made from the model. Understanding the reasons behind significant deviations can help improve the model and provide more accurate predictions.

4.4 Plot for the linear model

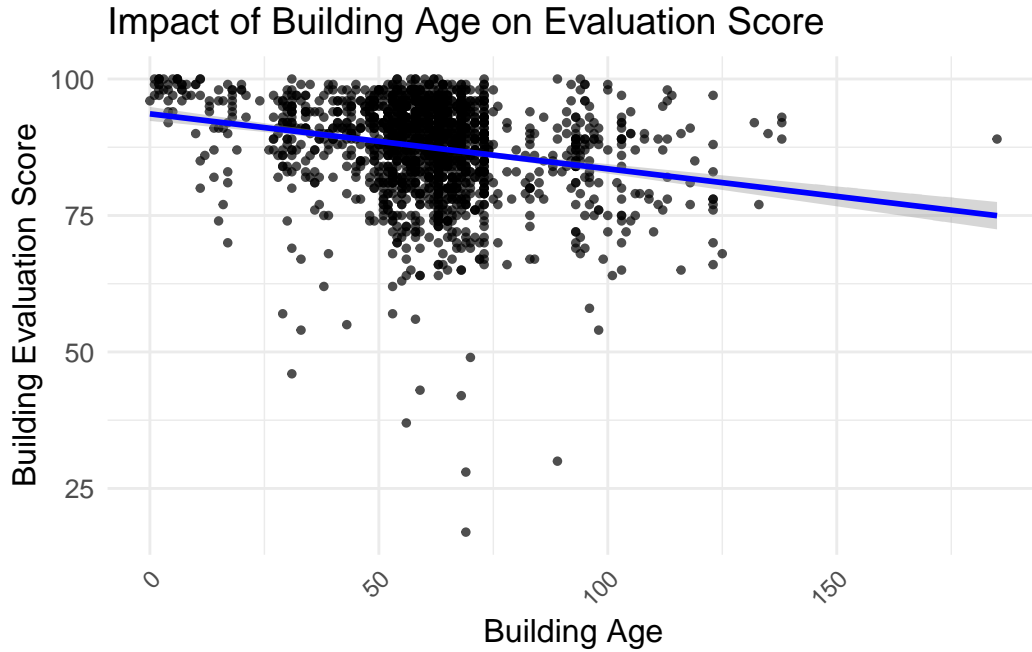


Figure 3: plot for the linear relationship between building age and evaluation score

Figure 3 provides a clear visual relationship between the age of buildings and their evaluation scores. Displayed with building age on the x-axis and evaluation score on the y-axis, the data points represent individual apartment buildings. The blue line running through the cluster of points represents the linear regression line, which captures the average effect of age on evaluation score. Noticeably, the slope of the line is downward, suggesting a negative correlation; as the buildings get older, their evaluation scores tend to decrease. This trend is consistent across the dataset, indicating a general decline in scores as building age increases, a reflection of the typical effects of wear and time on building integrity.

In statistical analysis, the significance of this downward trend is supported by a p-value much lower than the conventional threshold of 0.05, confirming the relationship between building age and evaluation score is not due to random chance. However, the plot also reveals that while younger buildings cluster more densely around higher evaluation scores, older buildings exhibit a wider spread of scores. This variation suggests that other factors may influence the condition and safety of older buildings, which are not captured solely by age. The plot does not show any clear pattern or funnel shape in the residuals, indicating that the variance of the evaluation scores is fairly consistent across the range of building ages, and the assumption of homoscedasticity is not visibly violated. Nonetheless, the model's findings imply the importance of consistent maintenance and updates to older buildings to ensure safety standards are upheld over time.

5 Discussion

5.1 Overview of the Study

This paper presented a detailed analysis using simple linear regression to investigate the impact of building age on the evaluation scores assigned by the City of Toronto. By analyzing data from the Toronto Open Data portal, the study aimed to forecast evaluation scores for Toronto residences in 2023. The analysis revealed a statistically significant inverse relationship between building age and evaluation score, suggesting that older buildings tend to have lower evaluation scores.

5.2 Insights into Urban Development and Maintenance

One key insight from this study is the importance of building age in the overall health and safety of urban residential environments. The findings underscore the potential risks associated with older buildings, which may not meet modern standards or may have deteriorated over time. This knowledge emphasizes the need for regular inspections, maintenance, and upgrades to older buildings to ensure they remain safe and habitable. Additionally, the study highlights the value of continuous monitoring and evaluation by urban authorities to maintain and improve housing standards across the city.

5.3 Policy Implications and Housing Quality

Another significant insight relates to the policy implications for housing quality and safety standards. The correlation between building age and lower evaluation scores may prompt policymakers to reconsider regulations surrounding building inspections, maintenance schedules, and renovation requirements. It suggests an opportunity for targeted interventions in older buildings to prevent potential hazards and improve living conditions for residents. Moreover, these findings could guide urban planning strategies, prioritizing the revitalization of aging housing stock and ensuring equitable access to safe and quality housing.

5.4 Limitations of the Study

Despite its contributions, this study has several limitations. Firstly, the model accounts for only a small fraction of the variability in evaluation scores, indicating that other factors besides building age significantly influence these scores. The exclusion of variables such as building materials, design, maintenance history, and location may limit the comprehensiveness of the findings. Additionally, the model's assumptions (e.g., linearity and homoscedasticity) and the presence of outliers suggest caution in generalizing the results without further investigation.

5.5 Future Directions

Given the limitations and the insights gained, future research should aim to incorporate additional variables that could affect building evaluation scores. A more comprehensive model including factors like renovation history, compliance with current building codes, and socio-economic characteristics of the neighborhood could provide a deeper understanding of what influences building evaluation scores. Longitudinal studies could also shed light on the trends over time, offering insights into the effectiveness of policy interventions and maintenance practices. Moreover, comparative analyses across different cities or countries could highlight unique challenges and best practices in building maintenance and safety standards.

In conclusion, while this study has provided valuable insights into the impact of building age on evaluation scores, it also opens the door for further research to explore the multifaceted nature of building quality and safety. Addressing these questions is crucial for developing effective strategies to improve urban living conditions and ensure the longevity and safety of the housing stock.

6 Appendix

References

- Anderson, Curt. 2021. “Lawsuit Claims Surfside Condo Collapse Triggered by Building Work.” *NBC 6 South Florida*. NBC 6 South Florida. <http://www.nbcmiami.com/news/local/lawsuit-claims-surfside-condo-collapse-triggered-by-building-work/2622574/>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Toronto, City of. 2023. “Open Data.” *City of Toronto*. <https://www.toronto.ca/city-government/data-research-maps/open-data/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data*. <https://CRAN.R-project.org/package=readr>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich

Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.