Assessing Bridge Conditions in New York State: The Impact of Age, Municipality, and Structural Deficiencies*

My subtitle if needed

Ariel Xing

November 28, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

---

*Code and data are available at: https://github.com/Jiaqi-Xing/NYS_Bridge_Condition_Analysis.

1 Introduction

Overview paragraph

Estimand paragraph

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2….

2 Data

2.1 Overview

The dataset used in this analysis, titled "New_York_Bridges_2016", was accessed from the Data And Story Library (DASL) (Data And Story Library (DASL) 2024), which archives historical data originally compiled by the New York State Open Data platform (New York State Open Data 2020). The original dataset includes contributions from county-level data provided by the New York State Department of Transportation (NYSDOT) (New York State Department of Transportation (NYSDOT) 2016).

This dataset was chosen over the updated 2020 version from New York State Open Data due to the change in how bridge condition data is recorded. In the 2016 dataset, the condition variable is a continuous variable measured on a seven-point scale (1 = "Deficient Condition" to 7 = "Good Condition"). However, in the 2020 dataset, NYSDOT replaced this scale with a binary Poor/OK variable, recorded as Poor01 (1 = Poor, 0 = OK). This change limits the analytical potential of the dataset, as the binary variable lacks the granularity required for

linear regression that relies on continuous variables to detect nuanced trends or incremental differences. By using the 2016 dataset, this study retains the detailed numerical condition ratings, enabling to analysis impact of factors such as age, location, and owners on bridge conditions.

The raw dataset contains 17,502 rows and 13 columns, providing structural, geographical, and condition-related data on bridges across New York State. After data cleaning, four key variables were selected for analysis: the response variable "Condition" and three predictor variables, "AgeAtInspection", "Located_Municipality", and "Owner_Group". The variables "Condition" and "AgeAtInspection" are numeric and were directly retained from the raw data. The categorical variables "Located_Municipality" and "Owner_Group" were newly constructed by categorizing the original variables "Municipality" and "Owner", respectively.

To conduct the analysis for this paper, we utilized the statistical programming language R (R Core Team 2023). The workflow was supported by several packages, including tidyverse (Wickham et al. 2019) for data cleaning and manipulation, arrow (Richardson et al. 2023) for efficient data storage and retrieval, validate (van der Loo and de Jonge 2021) for data testing and validation, knitr(Xie 2023) for creating tables and ggplot2 (Wickham 2016) for creating visualizations.

2.2 Measurement

In New York State, bridge inspectors from the State Department of Transportation (NYSDOT) conduct evaluations of non-toll bridges every two years to ensure their safety and structural integrity. Their assessments are performed on a span-by-span basis, meaning that each section of the bridge between its supporting points is inspected individually. For every bridge, inspectors examine up to 47 structural elements, including 25 span-specific components such as the deck, beams, and supports. Each component is assigned a condition score, ranging from 1 (severely deficient) to 7 (new or excellent condition), which reflects its structural health at the time of inspection. Inspectors also note any defects, such as cracks, corrosion, or wear, and document associated quantities, such as the length, area, or number of affected elements, to capture the extent of the damage.

In addition to these detailed evaluations, inspectors assign federal ratings based on the overall average condition of a bridge's major components. These federal ratings correspond to the values of the **Condition** variable used in our paper and analysis. Bridges with condition scores below 5 are classified as being in "poor" condition and are further categorized as either Structurally Deficient (SD) or Functionally Obsolete (FO). All findings from these assessments are recorded in inspection reports, providing essential data to guide maintenance, rehabilitation, or replacement efforts.(New York State Department of Transportation (NYSDOT) 2024)

2.3 Outcome Variable

2.3.1 Condition

The **Condition** variable reflects the overall state of a bridge, encompassing both its structural integrity and functional adequacy as determined during inspections. It is measured on a continuous scale from 1 to 7, with higher scores indicating better condition. Bridges with condition scores above 5 are considered to be in good standing, while scores below 5 indicate that a bridge may be classified as either Structurally Deficient (SD) or Functionally Obsolete (FO).

Table 1: Summary Statistics of NYS Bridge Condition Ratings

Min	Max	Mean	Median	Standard Deviation	Count
1.93	7	5.377584	5.297	0.8390056	15855

Table 1 presents for the condition ratings of bridges in New York State, based on a dataset of 15,855 observations. The mean condition rating of 5.38 indicates that, on average, bridges are in fair to good condition. The median rating of 5.30 suggests that half of the bridges have condition ratings below or equal to 5.30. The close alignment of the mean and median reflects a relatively symmetrical distribution. While the minimum condition rating is 1.93, and the maximum is 7, the standard deviation of 0.83 suggests that most bridges cluster around the mean, with only a few outliers at the extreme ends of the scale. These statistics indicate that the majority of bridges are maintained at a stable and adequate level, with relatively few in very poor or excellent condition.
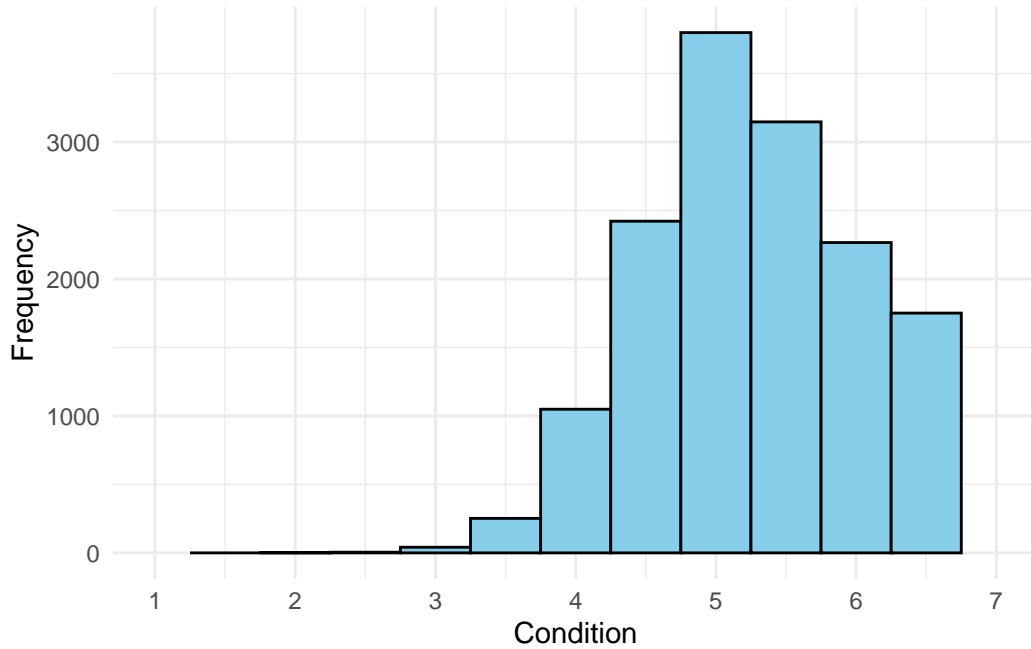
Figure 1: Distribution of Bridge Condition Ratings

Figure 1 The histogram displays the distribution of bridge condition ratings in New York State. The majority of bridges are clustered around the middle to upper end of the condition scale, with relatively few bridges at either the lower or higher ends. The left tail, representing bridges in severely poor condition, is thin, indicating that such bridges are uncommon. Similarly, the right tail, corresponding to bridges in excellent condition, is sparse, suggesting that high-quality, newly constructed bridges are not predominant. The distribution highlights a concentration of bridges in moderate condition, with a gradual decrease in frequency toward both ends, reflecting the natural aging process and variability in maintenance practices across the state.

2.4 Predictor variables

2.4.1 AgeAtInspection

The **AgeAtInspection** variable is a continuous measure that represents the age of a bridge in years. It is calculated as the time elapsed from the year the bridge was originally built or entirely replaced to its most recent inspection date.
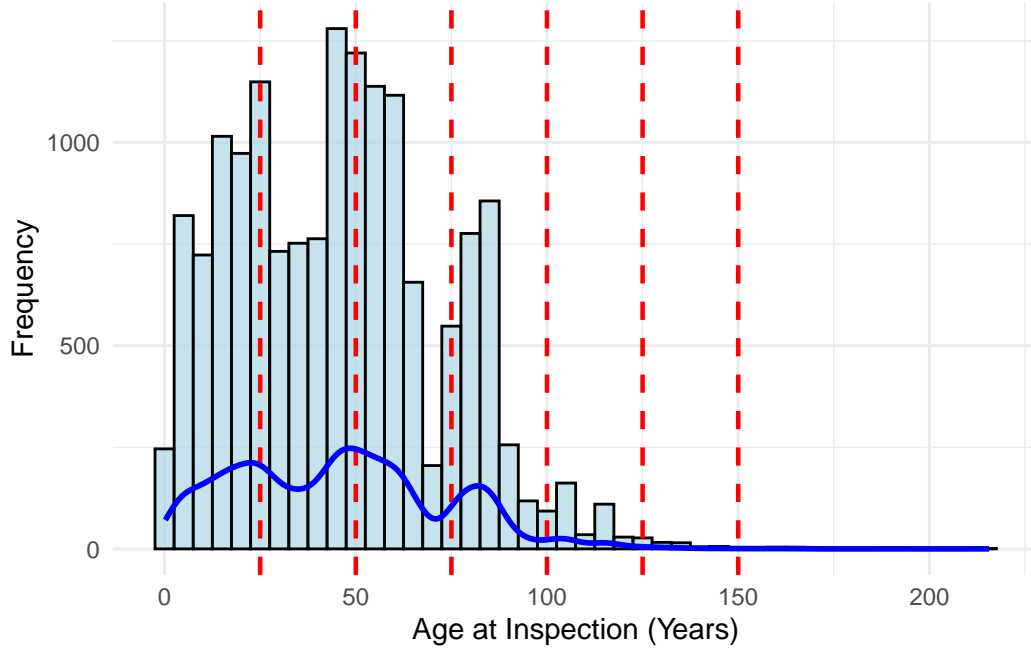
Figure 2: Distribution of Bridge Ages at the Time of Inspection. The red dashed lines mark intervals of 25 years and the blue density curve provides a smooth visualization of the distribution

Figure 2 shows distribution of bridge ages at the time of inspection, overlaid with red dashed lines and a blue density curve. The red dashed lines mark intervals of 25 years and align with clear peaks in the histogram, suggesting periodic patterns in bridge construction or major renovations. Prominent clusters are observed around 25, 50, and 75 years of age, indicating that many bridges were likely built or reconstructed during these timeframes. The blue density curve provides a smooth visualization of the distribution, confirming the periodic modes while highlighting the overall trends. Beyond 100 years, the frequency drops significantly, with very few bridges older than this still in operation, reflecting the typical lifecycle of infrastructure. The concentration of bridges between 20 and 80 years old suggests that a significant portion of bridges are mid-life and may soon require maintenance or replacement.@fig-age-histogram highlights both the historical patterns of bridge construction and the current age distribution of bridges in the dataset.

Figure 3: Relationship Between AgeAtInspection and Condition. Each blue dot represents a bridge, showing its age at the time of inspection and the corresponding condition rating obtained from that inspection.

Figure 3 illustrates the relationship between bridge age at inspection and their condition ratings. Most points are concentrated in the upper-left quadrant, showing that younger bridges tend to have higher condition ratings (close to 7). As bridge age increases, condition ratings generally decline, reflecting the expected deterioration over time. However, beyond approximately 80 to 100 years, the relationship becomes less defined, with condition ratings displaying a wide range. This variability may be influenced by differences in maintenance practices, varying levels of transportation load, or disparities in construction quality among older bridges. Overall, the pattern indicates that bridge age is a key factor in determining condition, particularly for younger and midlife bridges.

2.4.2 Located_Municipality

The **Located_Municipality** variable categorizes bridges based on the type of municipality where they are located. The three categories are City, Town, and Village.

Table 2: Proportion of Bridges by Municipality. The first column shows the type of municipality (Located_Municipality), and the second column represents the proportion of bridges in each category (Percentage Account).

Located_Municipality	Percentage Account
City	9.15%
Town	82.11%
Village	8.75%

Table 2 shows the distribution of bridges across different municipality types. Towns account for the majority of the bridges, representing 82.11% of the dataset, while cities and villages make up much smaller proportions, at 9.15% and 8.75%, respectively. This indicates that most bridges are located in towns, likely due to their larger geographic coverage.

Table 3: Average Bridge Condition by Municipality. The first column shows the type of municipality (Located_Municipality), and the second column provides the average condition rating (Average Condition) for bridges in each category.

Located_Municipality	Average Condition
City	5.16
Town	5.41
Village	5.33

Table 3 presents the average condition ratings of bridges by municipality type. Bridges in towns have the highest average condition rating of 5.41, followed by villages at 5.33 and cities at 5.16. The variation in average condition ratings may reflect differences in maintenance practices, traffic loads, or resource allocation among the municipality types.

2.4.3 Owner_Group

The **Owner_Group** variable categorizes bridges based on their ownership. This variable is designed to estimate potential differences in bridge conditions that may arise from varying ownership responsibilities. The categories include Municipalities, NYSDOT, and Other. The Other category combines bridges owned by private entities, authorities, commissions, and similar groups, as each accounts for a small portion of the data, making it practical to group them together

Table 4: Proportion of Bridges by Ownership. The first column shows the type of ownership group (Owner_Group), and the second column represents the proportion of bridges in each category (Percentage Account).

Owner_Group	Percentage Account
Municipalities	49.83%
NYSDOT	43.05%
Other	7.13%

Table 4 summarizes the ownership distribution of bridges. The two largest ownership groups are municipalities and NYSDOT, which account for 49.83% and 43.05% of the dataset, respectively. The remaining 7.13% of bridges fall under the "Other" category, which includes private, authority, and commission ownerships. The high percentage of bridges owned by municipalities and NYSDOT reflects their dominant role in managing public infrastructure.

Table 5: Average Bridge Condition by Ownership. The first column shows the type of ownership group (Owner_Group), and the second column provides the average condition rating (Average Condition) for bridges in each category.

Owner_Group	Average Condition
Municipalities	5.45
NYSDOT	5.33
Other	5.10

Table 5 shows the average condition ratings of bridges grouped by ownership. Bridges owned by municipalities have the highest average condition rating of 5.45, while NYSDOT-owned bridges have a slightly lower rating of 5.33. Bridges in the "Other" ownership category have the lowest average condition rating of 5.10, which may reflect less consistent maintenance or differing priorities among private or non-government entities.

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i | \mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of (**rstanarm?**). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table **??**.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows...

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows...

Examining how the model fits, and is affected
by, the data

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows... This suggests...

?@fig-stanareyouokay-2 is a Rhat plot. It shows... This suggests...

Checking the convergence of the MCMC algo-
rithm

References

Data And Story Library (DASL). 2024. "New York Bridges 2016 Dataset." https: //dasl.datadescription.com/datafile/new-york-bridges-2016/?_sf_s=Bridge&_sfm_ cases=4+59943.

New York State Department of Transportation (NYSDOT). 2016. "New York State Highway Bridge Data." https://www.dot.ny.gov/main/bridgedata.

———. 2024. "New York State Highway Bridge Data." https://www.dot.ny.gov/main/ bridgedata.

New York State Open Data. 2020. "Bridge Conditions: NYS Department of Transportation." https://data.ny.gov/Transportation/Bridge-Conditions-NYS-Department-of-Transportation/wpyb-cjy8/about_data.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, and Apache Arrow. 2023. *Arrow: Integration to 'Apache' 'Arrow'.* https://CRAN.R-project.org/package=arrow.

van der Loo, Mark P. J., and Edwin de Jonge. 2021. "Data Validation Infrastructure for R." *Journal of Statistical Software* 97 (10): 1–31. https://doi.org/10.18637/jss.v097.i10.

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 Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.