IDENTIFYING PREDICTORS OF BRIDGE DETERIORATION IN THE UNITED STATES FROM A DATA SCIENCE PERSPECTIVE

A Thesis

Presented to the

Department of Computer Science

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment
of the Requirements for the Degree

Master of Science

University of Nebraska at Omaha

by

Akshay Kale

May, 2019

Supervisory Committee:

Dr. Robin Gandhi
Dr. Brian Ricks
Dr. Jong-Hoon Youn

IDENTIFYING PREDICTORS OF BRIDGE DETERIORATION IN THE UNITED STATES FROM A DATA SCIENCE PERSPECTIVE

Akshay Kale, M.S.

University of Nebraska, 2019

Advisor:

Dr. Robin Gandhi

Dr. Brian Ricks

Dr. Jong-Hoon Youn

US Bridges scored a C+ on the 2017 infrastructure report card. There is a need for substantial improvement in bridge conditions as many of them are structurally deficient and can become unsafe in the near future. The nation's most recent bridge rehabilitation estimate is $123 billion.[1]

Many state's department of transportation (DOT) have limited resources, leaving them with difficult decisions about where to invest and allocate limited resources. To make cost-effective decisions, these bridge stakeholders need clean data and studies to estimate the future bridge conditions. This will give them data-driven, accurate life-cycle models for bridges and improved inspections intervals.

Previous researchers have identified factors that may cause bridge deterioration. Unfortunately, these researchers limit their data to specific regions and bridge types. This severely limits their result's general applicability.

In this thesis, we approach bridge health-related decision making challenges using a novel data science perspective. This bridge health deterioration study provides new insights into making bridge rehabilitation and reconstruction decisions. In this research, we use all US inspection record data regulated by the Federal Highway Agency that is available in the National Bridge Inventory (NBI)

database and precipitation data from the Center for Disease Control and Prevention (CDC).

Our specific contributions are 1) providing a reference big data pipeline implementation for bridge health-related datasets; 2) demonstrating the feasibility of data science to study bridge deterioration; 3) developing repeatable methods for sharing large datasets with reproducible analysis driven by data science and making them available to other researchers. Further, our curated datasets and platforms are used to analyze the statistical significance of bridge deterioration factors as identified by the literature and subject matter experts at the Nebraska State DOT. From our results, we found that bridge material type has the highest association in comparison to other factors such as average daily traffic, average daily truck traffic, structure length, maintainer, region, and precipitation. This research used all NBI inspection records and precipitation rates from all US counties.

COPYRIGHT

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Highways and bridges are cornerstones of the US Transportation system. It is essential for commerce and economic activity and is the primary mode of transportation in America. The 2017 infrastructure report states that there are about 600,000 bridges in the US. The average age of four out of ten bridges is 50 years. Approximately, 39% of the bridges will soon require rehabilitation as most of them were designed for a lifespan of around 50 years. About 9.1% of the nation's bridges were structurally deficient by 2017.[1] These statistics highlight the urgent need for innovative solutions to better understand and manage bridge health.

Through this research, we provide our stakeholders tools and insights to make data driven decisions regarding bridges' maintenance, appropriate inspection intervals, rebuild and reconstruction, using bridge life cycle and bridge deterioration models.

We use survey records of over 600,000 bridges in the United States from National Bridge Inventory, also known as NBI, to guide the development of effective decision making tools. State Departments of Transportation (DOTs) nationwide survey every bridge, every two years. The results of this survey of bridges is submitted to the NBI every year.

Previous research efforts have attempted to address problems of estimating appropriate inspection intervals and future condition of bridges, by developing deterioration models and life-cycle models. In particular, the following limitations in prior research efforts are noteworthy:

1. Previous research has failed to provide a generalized strategy to clean bridge survey records across all states in the U.S.

2. Previous research has not provided clear specification of data sources, working data samples, data cleaning strategies employed, and development environment used to conduct their study.

3. The analysis presented by previous research has limited their data to only specific regions and type of bridges. Hence, the insights gained from these studies are not generalizable.

4. Previous research has not devised methods that systematically identify and eliminate factors that contribute to condition of the bridges.

In this thesis, we address these limitations. First, we contribute in providing a generalized strategy to clean the NBI dataset. Second, we provide a way of using bridge inspection records (NBI) through a online data analytics platform. Third, we demonstrate the use of data from all available inspection records in the US, for all years to support reproducible analysis and generalizable insights. Fourth, provide a way to identify or eliminate factors that affect bridge conditions over time.

More formally, this thesis addresses the questions listed in the following subsections:

1.2 Research Question 1

The NBI dataset is a collection of the data submitted by all the states in the United States, and the data from different states might have variability in the number of fields, missing data, and how data is recorded. These discrepancies can be a challenge for data processing, cleaning, and curation when working with the NBI dataset. Therefore, the research study is interested in investigating the following question:

1. How can data cleaning be generalized across NBI data submitted from different states?

1.3 Research Question 2

Many DOT stakeholders and researchers rely on the NBI dataset to develop and validate their claims about bridge health. Due to the lack of a shared scientific compute and storage service, efforts to analyze and build decision support tools are in silos. This disconnect among NBI data users leads to rework, inconsistent cleaning practices, in-compatible analytic platforms, non-reproducible research, and lack of strong evidence for asserted claims about bridge health deterioration.

2. How can data processing platforms for large bridge inspection datasets be shared among researchers and practitioners?

1.4 Research Question 3

Average daily traffic is an indicator of frequency of bridge usage. The amount of precipitation helps estimate the exposure of the bridge to rain and snow events. Both these factors are used in many decision making tools and research efforts as the leading factors of bridge health deterioration. While average daily traffic is part of the NBI, precipitation data is not. As a result, examining the efforts of precipitation on bridge health deterioration is a non-trivial endeavor. For large datasets like the NBI combined with other large datasets like precipitation, methods to quickly identify or eliminate factors hypothesized to contribute to bridge health deterioration, and understanding the strength of their association are needed.

3. How can bridge inspection related datasets be used to identify or eliminate factors that affect bridge conditions?

In summary, we identified three research question, and in the following chapters we will do the following:

- In Chapter 2, we will review literature in the field of modeling bridge health.
- In Chapter 3, we will introduce our methods for carrying a systematic analysis to identify influential factors that affect bridge conditions.
- In Chapter 4, we will test various factors that are in NBI dataset and external factors using methods described in chapter 3, and understand the results.
- In Chapter 5, Chapter 6, we will discuss the results from chapter 4 and provide conclusion and discuss possible improvements for future work respectively.

Chapter 2

Previous Work

2.1 Overview

To address the research questions introduced in Chapter 1 Introduction of this thesis, we reviewed the literature regarding bridges from the following perspectives:

1. Data sources used to perform analysis and data cleaning strategies employed.

2. Data cleaning and data analysis environment made available to other researchers, and practitioners through the research.

3. Techniques and methods employed to identify and eliminate factors that contribute to the condition and deterioration of bridges, evaluate the current state of the bridge, and predict future condition of the bridges.

In reviewing the literature from the first perspective, we found that the data used is often limited to a single state. Hence, provide limited information on other state's bridges.[3–5] Several of these studies also rely on augmenting the NBI dataset with additional data collected by the state Department of Transportation (DoT)'s.[3–7] As a result, the combined dataset used in one research effort for a particular state is different from another. Getting access to or creating a consistent and consolidated dataset for all states is a challenge. Some of these databases are also restricted for access, which makes related studies challenging to reproduce and used by other civil engineering practitioners.

With respect to the second perspective, most prior research does not make their working data and data processing environment available. The description of data cleaning procedures are often unclear, and the original source of the data is unavailable. Unobtainable working data and descriptions of the data processing environment used for data analysis pose challenges in reproducing the results.

Finally, the review of the literature from the third perspective has revealed significant efforts made in the evaluation of current bridge condition and predicting future condition. These efforts rely on statistical analysis, geographical information system and artificial intelligence. These studies have previously identified several influential factors in determining deterioration of bridges based on certain established techniques for conducting the predictive analysis. Several[8–10] researchers have used Geographical Information System (GIS) techniques to show deficient bridges. Bolukbasi et al.[3] 2004 conducted regression analysis on bridges in Illinois. Washer et al. 2014[4] performed statistical analysis to understand patterns in deterioration and estimate bridge inspection cycles, rather than doing it every two years for all bridges.

There has been a considerable effort in improving existing artificial intelligence techniques used in predicting future conditions of the bridge. Markov chain-based techniques are popular throughout the literature,[5,11,11–13] Artificial Neural Network (ANN) are also used in developing predictive models of deterioration.[14] Apart from looking at problems only from the statistical point of view, there has been an exploration of various new methods for predicting future condition of bridges. The visual patterns from GIS provide valuable information and make results understandable. Kim et al. (2009)[6] observe that there are critical findings from research predicting the future condition of constructed bridges. There are lack of insights into factors owners and maintainers, regions of the U.S having similar weather influence in deterioration and condition of the bridges. There is a need to have a standard set of tools and services to conduct large-scale research.

2.2 Summary of Prior State-of-the-Art

In Table 2.2, select papers from prior state-of-the-art is summarized. This compact format is intended to highlight the perspectives identified the opening discussion in this Chapter. Table 2.1 provides term used in the literature review table 2.2:

Term	Research Paper
PAPER 1	Estimating the Future Condition of Highway Bridge Components using National Bridge Inventory Data[3]
PAPER 2	Identifying Critical Sources of Bridge Deterioration in Cold Regions through the constructed Bridges in North Dakota[6]
PAPER 3	The Methodology for Probabilistic Modeling of Highway Bridge Infrastructure: Accounting for Improvement Effectiveness and Incorporating Random Effects[5]
PAPER 4	Comparative Analysis of Bridge Superstructure Deterioration[15]
PAPER 5	An In-Depth Analysis of the national bridge inventory database utilizing data mining, GIS, and advance statistical methods[8]
PAPER 6	Estimation of Infrastructure Transition Probabilities from Condition Rating Data Research Purpose[16]
PAPER 7	Estimating inspection intervals for bridges based on statistical analysis of national bridge inventory data[4]
PAPER 8	Modeling Bridge Deterioration Using Case-Based Deterioration[7]

Table 2.1: Terms used to represent research papers

PAPER: 1 **TIME SPAN:** 1976-98 **BRIDGES:** 2601 **DATA SOURCE 1:** NBI **DATA SOURCE 2:** ISIMS **PURPOSE:** Develop methods to estimate planning of future conditions of bridges.	**METHODS:** Deterioration curves and calculating expected duration of condition ratings **INDICATOR:** Average daily traffic, Type of Service. Type of Material. **PROS:** Elimination of unclean records. Takes reconstruction into account. Small sample size. Limited to one state. No source code available. **CONS:** Small sample size. Limited to one state. No source code available.

PAPER: 2 **TIME SPAN:** 2006-07 **BRIDGES:** 5289 **DATA SOURCE 1:** NBI **PURPOSE:** Performance of constructed bridges in cold regions by examining the bridges in North Dakota.	**METHODS:** Multiple linear regression and GIS techniques. **INDICATOR:** Interstate bridges. Large city bridges with high population, Concrete bridges are better in colder regions than steel, Year built, Volume of traffic, Structural System. Presence of water. **PROS:** Elimination of unclean records. Takes reconstruction into account. Small sample size. Limited to one state. No source code available. **CONS:** Small sample size. Limited to one state. No source code available.

PAPER: 3 **TIME SPAN:** 1992-2014 **BRIDGES:** 5600 **DATA SOURCE 1:** NBI **PURPOSE:** Develop methods to account bias and random effects in NBI dataset.	**METHODS :** Ordered Probit Models **INDICATOR :** Replacement, Repair. **PROS:** Considers random effects and bias. **CONS:** Small sample size. Limited to one state. No source code available.

Table 2.2: Literature Summary Table - A

PAPER: 4 **TIME SPAN:** 1990 **BRIDGES:** 57700 **DATA SOURCE 1:** NBI **PURPOSE:** Data analysis on the pre-stressed superstructure of the bridge.	**METHODS :** Regression analysis. **INDICATOR:** Age, Average daily traffic. **PROS:** Considers more than one state. **CONS:** Small sample size. No source code available,No environmental factors, Focuses only on the structural material of superstructure.

PAPER: 5 **TIME SPAN:** 1996 **BRIDGES:** 30000 **DATA SOURCE 1:** NBI **DATA SOURCE 2:** Weather data **DATA SOURCE 3:** Natural hazard **PURPOSE:** Purpose of this study is to look into deterioration of the bridges from the geo- spatial layer in the GIS system.	**METHODS:** Regression modelling. **INDICATOR:** Age, Average daily traffic, Predominant structural material, Annual precipitation, Frequency of salting, Temperature range, Freeze-thaw cycle **PROS:** Data cleaning of outlier and erroneous entries. Environmental and natural hazard factors considered. **CONS:** Small sample size. Limited to one state. No source code available.

PAPER: 6 **TIME SPAN:** 1978-1986 **BRIDGES:** 5700 **DATA SOURCE 1:** NBI **PURPOSE :**Introduces new method for bridge deterioration model.	**METHODS:** Markov decision modeling,Ordered Probit modeling. **INDICATOR:** Wearing surface type 1,2,6,9. Climate region. Age. Average daily traffic. **PROS:** Difference in environment, inspection procedures may explain this pattern. **CONS:** Small sample size. Limited to one state, No source code available, Study is focused on only concrete bridges.

Table 2.3: Literature Summary Table - B

PAPER: 7 **TIME SPAN:** 1992-2011 **BRIDGES:** 4270 **DATA SOURCE 1:** NBI **PURPOSE :** Statistical methods to estimate inspection intervals.	**METHODS:** Weibull distribution. Anderson Darling test. **INDICATOR:** Weibull scale parameter. **PROS:** Simplicity of model. **CONS:** Small sample size. Limited to one state.No source code available.Study is focused on only on superstructure.

PAPER: 8 **TIME SPAN:** 1993-99 **BRIDGES:** 512 **DATA SOURCE 1:** MTQ **PURPOSE:** Introduces a new method to predict future condition of bridges.	**INDICATOR:** Highway class, Region, Material, Structural System, Wearing surface. **PROS:** Method for predicting future condition of bridges. **CONS:** Small sample size. No environmental factors considered.

Table 2.4: Literature Summary Table - C

2.2.1 Abbreviations of summary literature

Table 2.2 is a summary of the literature review and few of the abbreviations used are as follows:

Abbreviation	Terms
NBI	National Bridge Inventory
MTQ	Ministry of Transportation,Quebec
GIS	Geographical Information Systems
AADT	Annual Average Daily Traffic
ISIM	llinois Structure Information Management System

Table 2.5: Abbreviations used in summary literature

In summary, we reviewed literature from perspective of data source, data cleaning and analysis environment made available for other researches, and methods employed to identify and eliminate factors that most influence deterioration, and condition of the bridges.

From our review, the following are the observation regarding the previous research:

- Previous research concluded their finding using a limited set of data.
- Previous research does not make their working data and data processing and analytic environment available.
- Previous research have not explored the effects of factors such as region, owners and maintainers.
- Previous research have not provided with a systematic approach to identify the influential factors that affects deterioration and condition of the bridges.

This research lays a solid groundwork in the area of modeling bridge condition. In the following chapters, we show how we improve on this work.

Chapter 3

Methods

3.1 Overview

In this chapter, we introduce our data cleaning and transformation strategies, we explain a new method to compute scores that indicate the condition of bridges and we these to explore bridges in the U.S.

3.2 Data Cleaning and Transformation

In our exploration of the NBI data, we observed that some of the data needed to cleaned and some data needed to be transformed.

3.2.1 Data Normalization

Since our work is focused on bridge deterioration, we are most interested in the condition of the bridge, which is indicated in the deck rating, substructure rating, and superstructure fields of a bridge inspection record, we also examined other fields in the NBI such as latitude, longitude, and structure type. We found that many bridges had inconsistencies such as such as:

- Condition ratings such as deck, substructure, and superstructure rating had missing values.
- Fields such as Longitude and latitude were not in usable format.
- Fields were inconsistent when checked using the given cross-checking guidelines.[17]
- Other fields had invalid values that are out of range as per described in the NBI recording guide,[18] and NDOT[17]
- Repeated inspection records of the same structure number early in the lifecyle of the database or records with changing values for the year of built.

3.2.1.1 Data Cleaning

In data cleaning stage of our pipeline we addressed the above-mentioned inconsistencies by:

- Discarding records with missing values for deck, substructure, and superstructure.
- Maintaining a log of invalid values and repeated records.
- Adding a new field for location: Longitude and Latitude, to have a usable format of degrees and minutes.
- Reintroducing bridges in database by identifying the change in year-built of the bridge.

Tables 3.1 and 3.2 show the number of culled inspection records by year and reason for removal. These stables show the substantial decrease in the repeated records and missing geo-coordinates of the bridges over the years in the state of Nebraska.

Year	**Repeated Records**
1992	454
1993	453
1994	468
1995	472
1996	477
1997	476
1998	477
1999	484
2000	486
2001	488
2002	493
2003	503
2004	504
2005	514
2006	517
2007	517
2008	532
2009	532
2010	522
2011	517
2012	515
2013	0
2014	0
2015	0
2016	0
2017	0

Table 3.1: Table showing the number of repeated NBI inspection records for the state of Nebraska, 1992 - 2017

Year	Valid Geo-Coordinates	Missing Geo-Coordinates	Total
1992	3983	12238	16221
1993	4026	12212	16238
1994	4109	12199	16308
1995	4142	12185	16327
1996	4189	12153	16342
1997	4193	12135	16228
1998	4205	12088	16293
1999	16181	108	16289
2000	16272	1	16273
2001	16260	2	16262
2002	16237	2	16240
2003	16241	1	16242
2004	16238	5	16243
2005	16254	1	16255
2006	16254	1	16255
2007	16300	1	16301
2008	16294	1	16295
2009	16243	1	16244
2010	16185	1	16186
2011	16195	1	16196
2012	15392	1	15393
2013	15369	1	15370
2014	15372	2	15374
2015	15341	0	15341
2016	15334	0	15334

Table 3.2: Table showing the number of NBI inspection records with invalid geo-coordinates in the state of Nebraska, 1992-2017

NBI Data Field	**Criteria**
Item 058 Deck	Field coded as 'N' or 'NA'
Item 059 Substructure	Field coded as 'N' or 'NA'
Item 060 Superstructure	Field coded as 'N' or 'NA'
Item 108A Type of Wearing Surface	Field coded as '6'
Item 43 Structure Type	Field is coded as '19'

State	**Total No. of Survey Records**	**Bridge Survey Records Considered In Study**	**% of Bridge Survey in Study**
CO	223645	55259	24.71
WY	81257	48403	59.57
MT	143904	86220	59.91
ID	110415	37871	34.30
WA	219318	111952	51.05
OR	198421	55259	33.01
UT	86707	21616	24.93
NV	45179	13013	28.80
CA	761313	336288	44.17
AK	334522	21204	63.39
HI	28868	10968	37.99
TX	1347902	473361	35.12
OK	623417	330803	53.06
NM	105730	38725	36.63
AZ	198595	53104	26.74
WV	194514	87703	45.09
VA	396744	164877	41.56
KY	373241	214966	57.59
TN	545968	138745	25.41
NC	517708	132552	25.60
SC	246494	93626	37.98
GA	405216	186850	46.11
AL	417539	212866	50.91

Table 3.3: Table showing the number of NBI inspection records available after data cleaning - A

State	Total No. of Survey Records	Bridge Survey Records Considered In Study	% of Bridge Survey in Study
MS	442724	278356	62.87
LA	350988	192878	54.95
FL	339786	168565	49.61
NE	400539	295655	73.81
IA	641429	500280	77.99
IL	714936	362921	50.76
IN	484585	260133	53.68
KS	663487	387056	58.34
MI	641698	369188	57.53
ND	116171	71287	61.36
MO	641698	369188	57.53
SD	158215	98721	62.40
OH	1570646	531684	33.85
WI	386700	225939	58.43
MN	480773	156378	32.53
MA	137664	20186	14.66
CT	126267	6118	4.85
ME	67423	22560	33.46
NH	82523	12203	14.79
RI	22591	2521	11.16
VT	73408	20385	27.77
NJ	216169	76105	35.21
NY	522369	179217	34.31
PA	665314	267464	40.20
DC	7808	2623	33.59
MD	144047	60340	41.89
DE	28935	11135	38.48
AR	339492	178566	52.60
PR	61350	25421	41.44

Table 3.4: Table showing the number of NBI inspection records available after data cleaning - B

Tables 3.3 and 3.4 show the results of our data cleaning on the number of inspection records available in all states of the U.S. We observed that a large part of the data is unavailable for further analysis of identifying influential factors–on average only 42% of the original dataset is available for data analysis after the data cleaning and filtration process.

NBI Dataset also has substantial number of bridges with constant condition ratings. About 48% of the bridges used in this analysis post data cleaning and data filtration criteria have constant condition ratings.

3.2.2 Data Transformation

In this work we study bridges over the course of two decades to determine which factors have the greatest effect on bridge health. However, as we transformed the database of inspection records into a database of bridges with timeseries of data, we noted that the year built for some bridges was not consistent. As a result, Bridges appeared as unnaturally aged or newly built. Based on our discussion with subject matter experts, we learned that when an inspection record shows a newer year built than previous inspection records, it often means the bridge was rebuild. To capture this fact into our data, we divided the time-series of a bridge data into several consistent segments with the same year-built, and then each consistent segment was treated as a different bridge in our working data. We maintain the traceability of every segment to the original bridge by adding a segment number as a suffix to the original structure number of the bridge.

In summary, by dividing bridges' timeseries data into consistent segments and reintroducing each consistent segment of bridge time series as a new bridge, we were able to create a consistent dataset of bridges' timeseries for our analysis of identifying influential factors of bridge conditions.

By applying data cleaning and transforming the NBI data, we could create a definition of bridge condition and develop method using our definition of bridge condition. In next section, we described the condition of the method for computing Baseline Difference method.

3.3 Measurements

Our goal is to provide data-driven tools to our bridge stakeholder collaborators. With the use of these data-driven tools, stakeholders can allocate their limited resources to bridges that require the most attention.

To effectively allocate resources, stakeholders require methods that evaluate bridges individually. As discussed in our literature review, the current state-of-the-art only provide collective measures, not the individual measures that stakeholders need (e.g. Markov Decision Process,[16] Ordered Probit Model,[5] and Regression Models[8]).

Using the cleaned NBI data just discussed, we propose a method that scores bridges individually, which we call the Baseline Difference Score (BDS) method. BDS determines the performance of a bridge relative to a baseline computed from the national average of BDS scores. With this method, we can examine how a bridge may differ from the established national baseline, giving each bridge a quantitative, individualized score.

To compute BDS, we consider only one of the three ratings (deck, substructure, superstructure) of the bridge. A review of literature in deterioration of the bridges indicate that deterioration of superstructure is considered crucial, as the function of the superstructure is of the backbone of the bridge, which plays an important role in the safety of the bridge.[4] In Figure 3.1, The condition ratings of deck, superstructure, and substructure are highly correlated. Therefore, we compute BDS for bridges using superstructure condition ratings of the bridge.

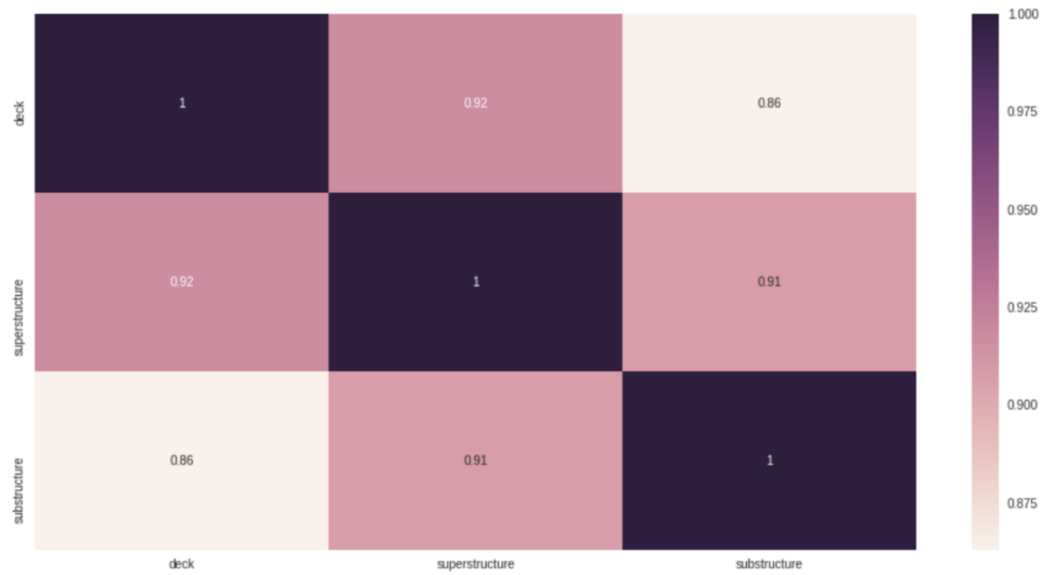

Figure 3.1: Pearson Correlation of bridge Deck, Substructure, and Superstructure

3.3.1 Baseline Difference Score

To calculate the BDS S_a for Bridge A.

Let i be the age of Bridge A, when bridge A was first inspected.

k be the age of Bridge A, when bridge A was last inspected.

$\vec{C}$ be the vector of condition ratings of the Bridge A from age i to age k.

$\vec{X}$ be the average condition rating of bridges of all bridges from age 1 to 100.

X_y be the vector of average condition rating from age i to age j,

such that $j \leq k$.Then, Deterioration Score S_a

$$S_a = \vec{C} - \vec{X_y}$$

3.3.1.1 Computing National Baseline

To provide a clear example of the computation, we provide a graphical example of the baseline difference computation of a bridge. In Figure 3.2, we compute a baseline from the condition rating of the Bridge A, Bridge B, and Bridge C. Here, a baseline is the average condition rating at each age.

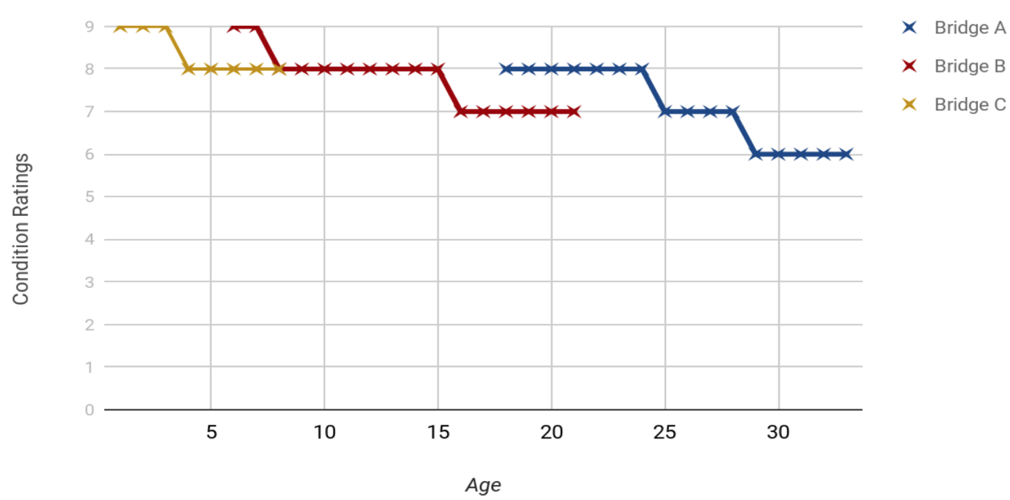

Figure 3.2: Condition Ratings of the Bridge A, Bridge B, and Bridge C.

In Figure 3.3, we compute the average of condition rating of the bridges for all age.

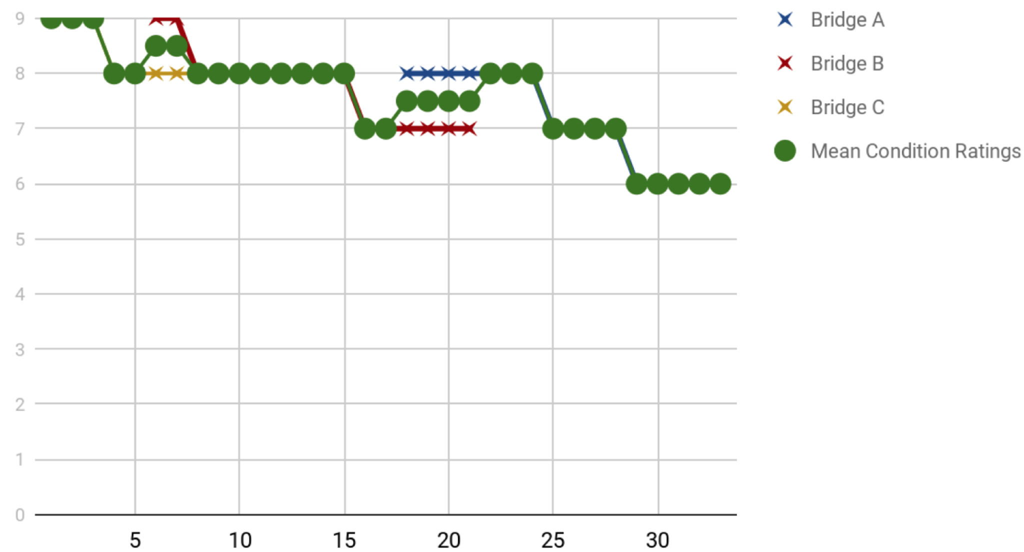

Figure 3.3: Computation of baseline from the condition rating of Bridge A, Bridge B, and Bridge C.

3.3.1.2 Computing BDS of an individual bridge

After computing the baseline, in Figure 3.4 and 3.5, we compute the difference in condition ratings and the baseline.

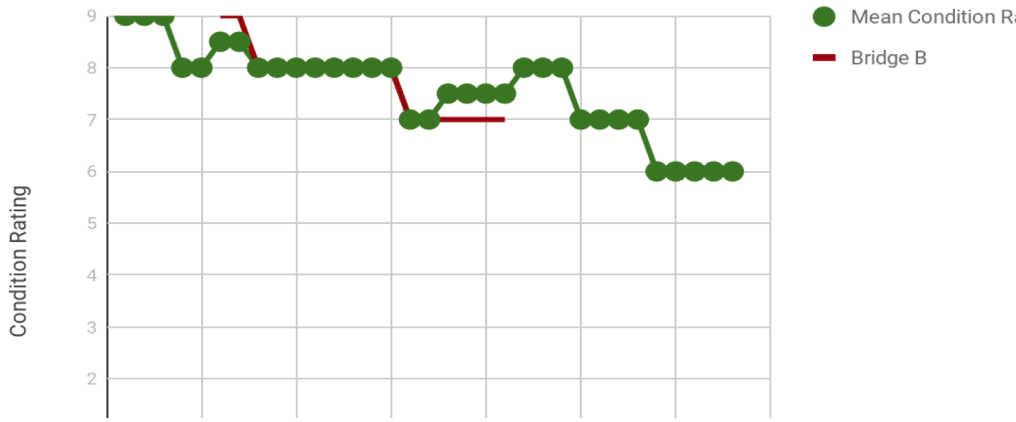

Figure 3.4: Computation of baseline difference score of Bridge B.

In figure 3.5, we compute the final BDS for Bridge 'B' by averaging the difference.

M - B = (8.5 - 9) + (8.5 - 9) + (8 - 8) + (8 - 8) + (8 - 8) + (8 - 8) + (8 - 8) + (8 - 8) + (8 - 8) + (8 - 8) + (7 - 7) + (7 - 7) + (7.5 - 7) + (7.5 - 7) + (7.5 - 7) + (7.5 - 7)

M - B = -0.5 + -0.5 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0.5 + 0.5 + 0.5 + 0.5

M - B / Length of the condition rating = 1 / 16

Score (B) = 0.0625

Figure 3.5: Computation of baseline difference score of Bridge B (Continued).

3.3.2 Psuedocode of Baseline Difference Score

The pseudocode for computing BSD score has a main method that encompass three main functions:

1. getCountAndSum
2. ComputeBaseline
3. ComputeBaselineDifferenceScore.

First, getCountAndSum function will compute sum and count of all the condition ratings for all ages in a given region. Second, ComputeBaseline function will compute a baseline of bridges by averaging condition rating for all years in a given region. Since, Bridges are designed for the service of 50 years, we only take into consideration of the baseline till the year 50; Finally, ComputeBaselineDifference function compare each bridge with the baseline and returns the difference in between the bridge and baseline and computes mean of the differences and returns the BDS score of the bridge.

Algorithm 1: mainFunction

Input: $Condition\_Rating$: A sequence of condition ratings (integers) of bridges
$\langle\langle c_{1_1}, c_{1_2}, \ldots, c_{1_n}\rangle, \langle c_{2_1}, c_{2_2}, \ldots, c_{2_n}\rangle, \ldots, \langle c_{2_n}, c_{2_n}, \ldots, c_{2_n}\rangle\rangle$,
$Age\_Of\_Bridges$: A sequence of condition ratings (integers) of bridges
$\langle\langle A_{1_1}, A_{1_2}, \ldots, A_{1_n}\rangle, \langle A_{2_1}, A_{2_2}, \ldots, A_{2_n}\rangle, \ldots, \langle A_{2_n}, A_{2_n}, \ldots, A_{2_n}\rangle\rangle$

Output: Baseline_Difference_Score: A List of scores for every segment in the List_Of_Segments
$\langle\langle B_{R_{1_1}}, B_{R_{2_2}}, \ldots, B_{R_{2n}}\rangle, \langle B_{R_{2_1}}, B_{R_{2_2}}, \ldots, S_{R_{2_n}}\rangle, \ldots, \langle B_{R_{n_1}}, B_{R_{n_2}}, \ldots, B_{R_{n_n}}\rangle\rangle$

Count_Of_Bridges_at_Age, Sum_Of_Condition_Rating_at_Age = getCountAndSum(Condition_Rating, Age_Of_Bridges)

Dict_Of_Age_Average_Condition_Rating = computeBaseline(Count_Of_Bridges_at_Age, Sum_Of_Condition_Rating_at_Age)

Baseline_Difference_Score = computeBaselineDifferenceScore(Condition_Ratings, Dict_Of_Age_And_Average_Condition_Rating_Ages)

return $Baseline\_Difference\_Score$

Algorithm 2: getCountAndSum

Input: $Condition\_Rating$: A sequence of condition ratings (integers) of bridges
$\langle\langle c_{1_1}, c_{1_2}, \ldots, c_{1_n}\rangle, \langle c_{2_1}, c_{2_2}, \ldots, c_{2_n}\rangle, \ldots, \langle c_{2_n}, c_{2_n}, \ldots, c_{2_n}\rangle\rangle$,
$Age\_Of\_Bridges$: A sequence of Age (integers) of bridges
$\langle 1, 2, \ldots, n\rangle$

Output: $Count\_Of\_Bridges$: A sequence of total count (integers) of bridges at all ages r age of the bridge $\langle T_1, T_2, \ldots, T_3\rangle$
$Sum\_Of\_Condition\_Rating\_at\_Age$: A sequence of Sums of the condition Rating (integers) of bridges at all ages. $\langle X_1, X_2, \ldots, X_3\rangle$

$i \leftarrow 0$, be the outer pointer
$j \leftarrow 0$, be the inner pointer
$Count\_Of\_Bridges\_at\_Age \leftarrow []$, be the list of deteriorating segments of condition ratings
$Len \leftarrow Length(Condition\_Rating)$, be the length of list of deteriorating segments of condition ratings
$Sum\_Of\_Condition\_Rating\_at\_Age \leftarrow []$

```
while i == Len do
    Condition_Rating_Bridge = Condition_Ratings_Of_All_Bridges[i]
    Ages = Age_Of_Bridges[i]
    j=0
    Len_Condition_Rating_of_A_Bridge = len(Condition_Rating_Bridge)
    while j == Len_Condition_Rating_of_A_Bridge do
        Age = Ages[j] Count_Of_Bridges_at_Age[Age] =
          Count_Of_Bridges_at_Age[Age] + 1
        j = j + 1
    i = i + 1
```

return *Count_Of_Bridges_at_Age, Sum_Of_Condition_Rating_at_Age*

Algorithm 3: computeBaseline

Input: $Condition\_Rating$: A sequence of condition ratings (integers) of bridges
$\langle\langle c_{1_1}, c_{1_2}, \ldots, c_{1_n}\rangle, \langle c_{2_1}, c_{2_2}, \ldots, c_{2_n}\rangle, \ldots, \langle c_{2_n}, c_{2_n}, \ldots, c_{2_n}\rangle\rangle$

Output: $Dict\_Of\_Age\_And\_Average\_Condition\_Rating$: A Key-Value pair of Age and Baseline_Condition_Rating

$Age\_Of\_Bridges$: A sequence of Age (integers) of bridges $\langle 1, 2, \ldots, n\rangle$

Counter $\leftarrow$ 0
Dict_Of_Age_And_Average_Condition_Ratings $\leftarrow$,
while *Counter == Len(Count_Of_Bridges_at_Age)* **do**
    Dict_of_Age_and_Average_Condition_Rating[Counter] = (Sum_Of_Condition_Rating_at_Age / Count_Of_Bridges_at_Age)

return $Dict\_of\_Age\_and\_Average\_Condition\_Rating$

Algorithm 4: computeBaselineDifferenceScore

Input: $Condition\_Rating$: A sequence of condition ratings (integers) of bridges
$\langle\langle c_{1_1}, c_{1_2}, \ldots, c_{1_n}\rangle, \langle c_{2_1}, c_{2_2}, \ldots, c_{2_n}\rangle, \ldots, \langle c_{2_n}, c_{2_n}, \ldots, c_{2_n}\rangle\rangle$,
$Age\_Of\_Bridges$: A sequence of Age (integers) of bridges
$\langle 1, 2, \ldots, n\rangle$,
$Dict\_Of\_Age\_And\_Average\_Condition\_Rating$: A Key-Value pair of Age and Baseline Condition Rating

Output: An Integer: BaselineScore; A Baseline Difference Score of a bridge

Let Mean(), be the function to calculate mean of a list.
Counter ← 0 Temp_List ← []
while *Counter == len(Condition_Ratings)* **do**
    Condition_Rating = Condition_Ratings[Counter]
    Age = Ages[Counter]
    Baseline_Condition_Rating =
      Dict_Of_Age_And_Average_Condition_Rating[Age]
    Temp_List.append(Condition_Rating - Baseline_Condition_Rating)
BaselineScore = Mean(Temp_List)

return *BaselineScore*

Overall, BDS evaluates bridges individually and provides a measure of comparison against the national baseline.

3.4 Distribution of Baseline Difference Scores in the U.S.

To select appropriate statistical techniques to guide us in identifying the factors that affect the condition of the bridges, we have to understand the distribution of the BSD score of the bridges in the U.S.

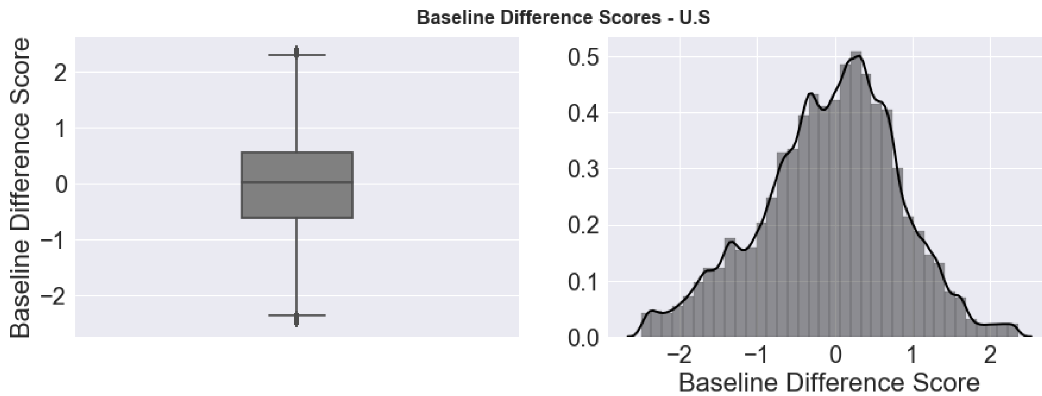

Figure 3.6: Baseline Difference Scores - U.S.

	U.S.
Count	1142331
Mean	-0.061
Std. dev	0.887
Min	-2.50
25%	-0.617
50%	0.014
75%	0.54
Max	2.368

Table 3.5: The Summary Statistics of Distribution of BDS of the bridges in the U.S.

We provide a box-plot and the density plot of the BDS in Figure 3.6 and summary statistics in the Table 3.7 . Note that the median (0.014) of the BDS score distribution in the box plot is positive and close to zero. The summary statistics of the distribution reveals that the mean (-0.061) is negative and close to zero and the standard deviation (0.887) close to one.

These results suggest that the variance, mean, and median of the distribution may satisfy the assumption of the normal distribution. To confirm our hypotheses about the distribution of the data, we performed a normality test of the distribution of the data.

3.4.1 Normality Test

We performed two normality test: D'Agostino's K^2 Test and Kolmogorov-Smirnov Test. In Table 3.6, with p-value of 0.0 for both test: D'Agostino's K^2 Test and Kolmogorov, we reject the null hypothesis. The results of the normality test suggest that the distribution of the baseline difference score (BDS) of all available bridges in the U.S is not normal; However, we performed a similar test on the random sample of 400 bridges, the normality test 3.7 results suggest the distribution of the random sample is normal.

Test	**Statistic**	**p-value**	**Null Hypothesis**
D'Agostino's	12917.120	0.0	reject
Kolmogorov-Smirnov	0.064	0.0	reject

Table 3.6: Normality Test on all Bridges in the U.S.

Test	**Statistic**	**p-value**	**Null Hypothesis**
D'Agostino's	4.204	0.122	fail to reject
Kolmogorov-Smirnov	0.090	0.370	fail to reject

Table 3.7: Normality Test on Random Sample of 400 Bridges

In the following chapter, we will present our findings on the effect of the following factors on bridge health as described by our BDS algorithm: Region, Average Daily Traffic, Average Daily Truck Traffic, Precipitation, Maintainer, Structure Length, and Material. We also classify bridges into appropriate groups with respect to every factor and perform ANOVA to understand the differences in the mean BDS in between the groups.

3.5 Summary

In summary, we employed data cleaning and transformation strategy on 17 million inspection records across 53 states and territories. We introduced the Baseline Difference Score and computed BDS for every bridge to enable comparison among bridges. The distribution of randomly selected 400 bridges' BDS in the U.S. is a normal distribution.

In next chapter, we will test factors such as Region, Maintainers, Precipitation, Average Daily Traffic, Average Daily Truck Traffic, Structure Material, and Structure Length, to see if these factors affect the bridge conditions.

Chapter 4

Results

4.1 Overview

In this chapter, we identify factors that have a significant effect on the condition of bridges based on Baseline Difference Score (BDS).

For every factor, we will categorize bridges into several groups as appropriate and then perform ANOVA to find the difference between the mean Baseline Difference (BDS) score of the groups.

We choose to do ANOVA because BDS is a continuous variable and factors have independent categories. We learned from chapter 3 that the normality test of the BDS score of a small sample of BDS is normally distributed and a large sample is similar to a normal distribution. Hence, all these mentioned criteria fulfill the assumptions of the ANOVA, and that makes ANOVA suitable for this analysis.

The degree of association between the factors and BDS score is measured using effect size. A measure of effect size can also be thought of as a correlation between factor and BDS score. In ANOVA, a commonly used measure of effect size is Eta Squared.

Cohen’s d is a standardized measure that is easy to interpret. For our analysis, we used Cohen's to measure the effect size between two means. Therefore, we converted eta squared value from our analysis of ANOVA to Cohen's d using a web resource.[19] The Table 4.1 provides a description for magnitude of Cohen's d.

Effect Size	**Cohen's d**
Very Small	0.01
Small	0.20
Medium	0.50
Large	0.80
Very Large	1.20
Huge	2.00

Table 4.1: Description for magnitudes of Cohen's d by Sawilowsky[2]

In the following section, we will explore the effect of Region, Precipitation, Average Daily Traffic, Average Daily Truck Traffic, Maintainers, Structure Material, and Structure Length on the bridge condition based on BDS. These factors are either suggested by the NDOT (Structure Length, Material, ADT, and ADTT), are commonly cited (Precipitation, ADT, and ADTT), or may serve as a proxy for the compound effect of various factors (Region).

4.2 Effect of Average Daily Traffic

Average Daily Traffic is one of the most commonly studied attribute of the bridge as the factor that affects the condition of the bridges.[3,6,8,15,16,20] To observe the effect of Average Daily Traffic (ADT) on condition of the bridges, Morcous et al.[20] provided criteria to classify ADT into four classes shown in the Table 4.2:

1. Very Light
2. Light
3. Moderate
4. Heavy

ADT Group	**Criteria**
Very Light	ADT $<$ 100
Light	$100 \leq$ ADT ≤ 1000
Moderate	$1000 \leq$ ADT ≤ 5000
Heavy	ADT ≥ 5000

Table 4.2: Grouping of Bridges by Average Daily Traffic

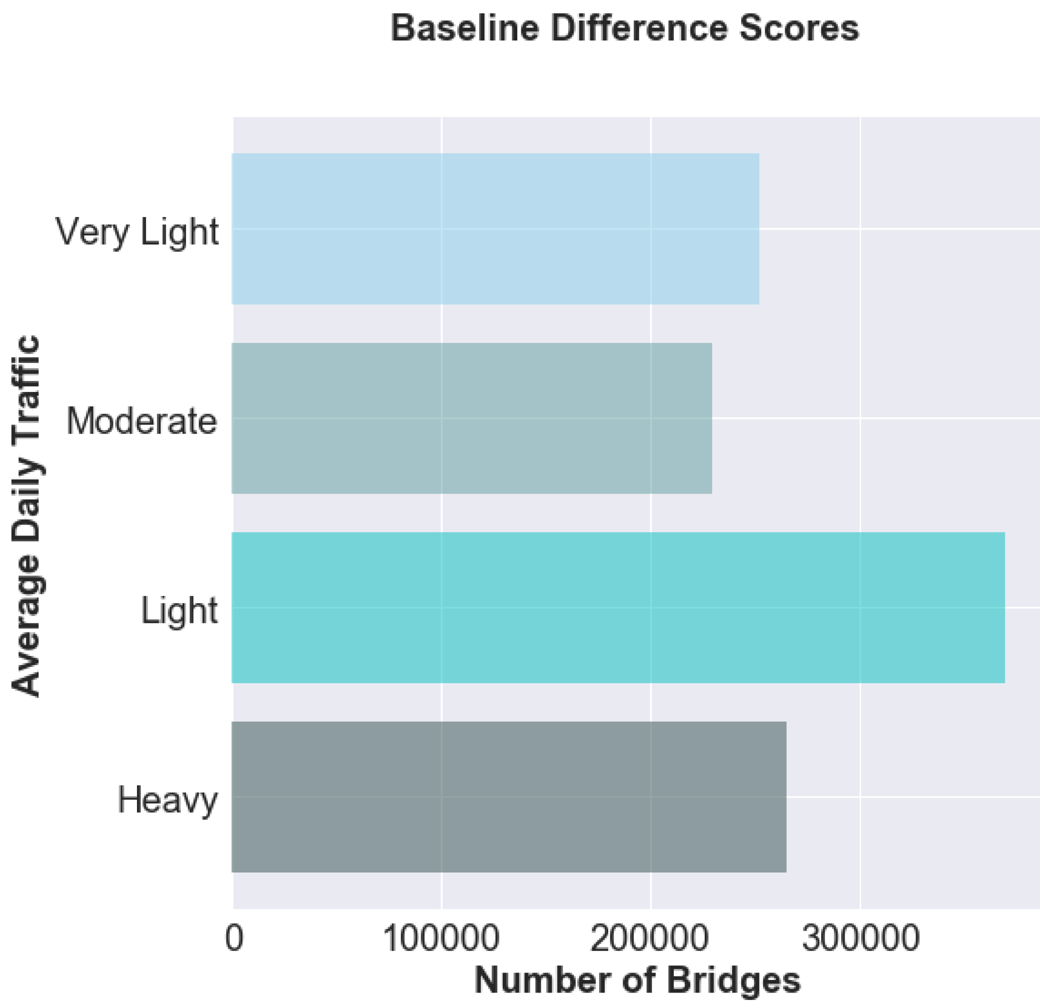

Figure 4.1: Number of Bridges in Very Light, Moderate, Light, and Heavy Average Daily Traffic

From Figure 4.1, a large number of bridges belong to Light ADT. Heavy and Very Light ADT have an similar number of bridges, and Moderate ADT has the least number of bridges.

Figure 4.2 compares the BDS distribution of the four different ADT classes of bridges and reveals that the mean and median are very similar within these groups and across other groups.

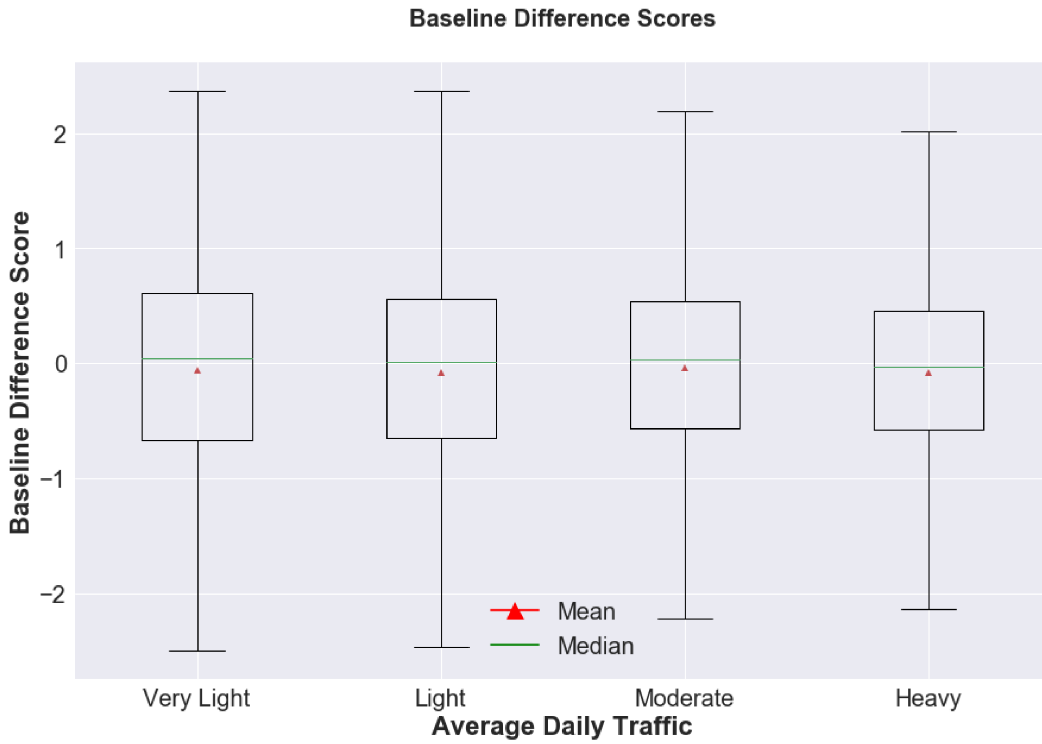

Figure 4.2: Distribution of Baseline Difference Scores of Very Light, Light, Moderate, and Heavy Average Daily Traffic

As seen in the Figure 4.2, the mean and the medians are similar within these distributions.

4.2.1 Results

We performed one-way ANOVA on a small sample size and a large sample size of data.

4.2.1.1 One-way ANOVA - Small Sample

We randomly selected 400 bridges (100 bridges from each ADT group) to perform ANOVA.

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
ADT	0.493	3.0	0.21	0.88	0.0016	0.08	very small
Residual	305.93	396.0	-	-	-	-	-

Table 4.3: One-way ANOVA of Average Daily Traffic on sample size of 400 bridges

With PR(>F) > 0.05 as seen in the Table 4.3, we fail to reject the null hypothesis based on ANOVA results. The effect size (Cohen's d = 0.08) is very small by convention.

However, the analysis is performed on small sample size (0.04% of the available data.) Therefore, it is likely that a small sample size might have not captured the variance in the data of a large sample. Hence, we did a similar analysis on a large sample.

4.2.1.2 One-way ANOVA - Large Sample

From Figure 4.1, we noticed that ADT group Moderate has the lowest number (229,442) of bridges. To have an equal sample size from all groups, we randomly selected 229,442 (least number of bridges belonging to Moderate Group) bridges from other groups (Very Light, Moderate, Light, and Heavy).

The analysis on 917,768 bridges (229,442 bridges in each group) reveals that the PR(>F) value < 0.05 as seen in the Table 4.4, so we reject the null hypothesis. We observed the change in PR(>F) of a small sample and a large sample analysis,

	sum sq	**df**	**F**	**PR(>F)**
ADT	289.31	3.0	123.70	0.0
Residual	715466.50	917764	-	-
	eta sq	**cohen's d**	**effect size**	**-**
ADT	0.0004	0.04	very small	-
Residual	-	-	-	-

Table 4.4: One-way ANOVA of Average Daily Truck Traffic on the sample size of 917,768 bridges

this change in PR(>F) value suggests that as the sample size increased, we observed a difference in mean BDS among the ADT groups. However, the effect size (Cohen's d) of 0.04 is very small that suggests a small association between ADT and condition of the bridges.

From the analysis of small samples and a large sample, our findings suggest that the mean BDS score among four groups of the bridges (Very Light, Moderate, Light, and Heavy) is statistically insignificant. These results seem to indicate that, Average Daily Traffic fails to explain the differences in the bridge condition in comparison to the national baseline.

4.3 Effect of Average Daily Truck Traffic

Heavy trucks can have an impact on bridge substructure that can lead to a progressive collapse of the bridge superstructure and disastrous accidents.[21] Several studies investigated the effect of ADTT on the condition of the bridge,[21,22] and cite ADTT as an indicator of bridge condition.

In the NBI dataset, Average Daily Truck Traffic is reported as a percentage of Average Daily Traffic. We grouped bridges into three Average Daily Truck Traffic groups: Light, Moderate, and Heavy in the U.S. respectively.[20]

1. Light
2. Moderate
3. Heavy

In Table 4.5, we show the criteria for grouping of bridges with respect to ADTT.

ADT Group	**Criteria**
Light	ADTT < 100
Moderate	$100 \leq$ ADTT ≤ 500
Heavy	ADTT ≥ 500

Table 4.5: Grouping of the bridges by Average Daily Truck Traffic

As seen in Figure 4.3 a large number of bridges belong to Light ADT Heavy ADTT and Moderate ADTT have a similar number of bridges.

Figure 4.4 compares the BDS distribution of Moderate, Light, and Heavy ADTT groups of bridges and reveals that the mean and median are similar within the groups and across other groups. The range of the distribution of bridges with Light ADT is the largest, which could be a result of a large number of bridges in this grouping.

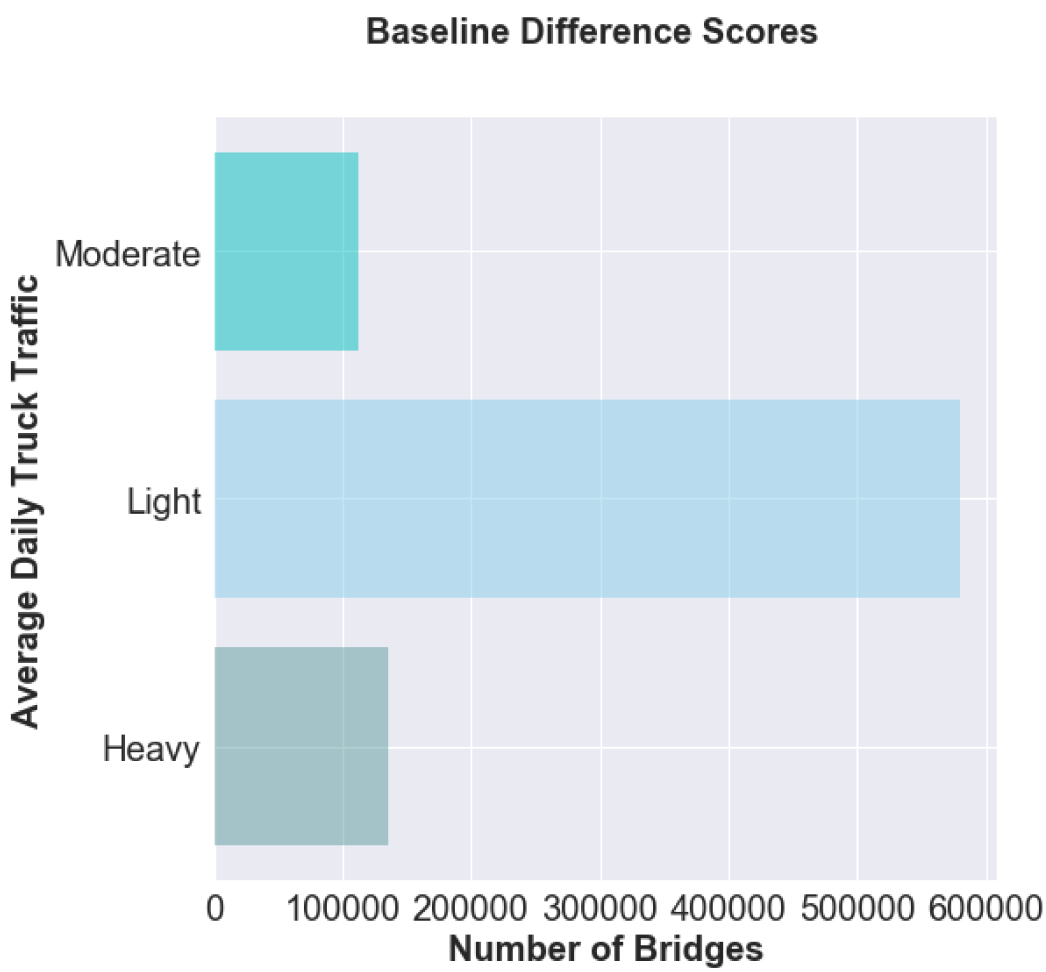

Figure 4.3: Number of Bridges in Light, Moderate, and Heavy Average Daily Truck Traffic group

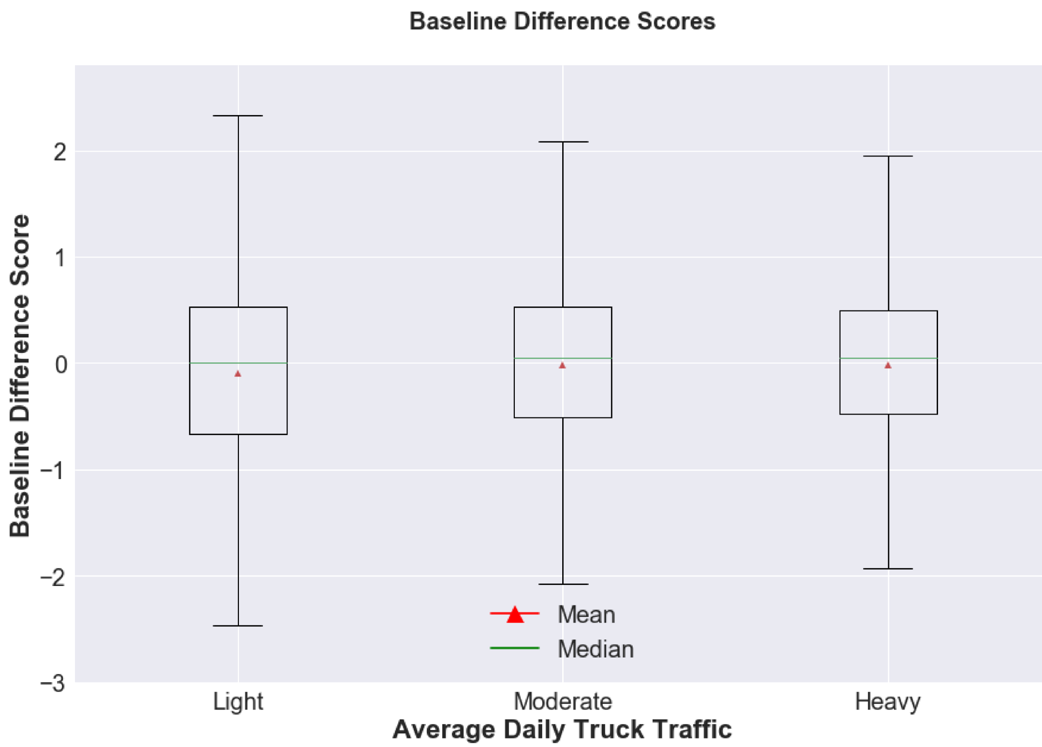

Figure 4.4: Distribution of Baseline Difference Scores of Light, Moderate, and Heavy ADTT group

4.3.1 Results

We performed one-way ANOVA on a small sample and a large sample of data.

4.3.1.1 One-way ANOVA - Small Sample

We randomly selected 300 bridges (100 bridges from each ADTT group) to perform ANOVA.

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
ADTT	2.84	2.0	2.12	0.12	0.01	0.20	small
Residual	199.02	297.0	-	-	-	-	-

Table 4.6: One-way ANOVA of Average Daily Truck Traffic on the sample size of 300 bridges

The results seen in Table 4.6 of the ANOVA revealed that the PR(>F) > 0.05, so we fail to reject the null hypothesis. The effect size (Cohen's d) of 0.20 calculated in the analysis reveals a small effect size by convention.

However, the analysis is performed on small sample size (0.08% of the available data). Therefore, it is likely that the small sample size might not have captured the variance in the data of a large sample. Hence, we did a similar analysis on a large sample.

4.3.1.2 One-way ANOVA - Large Sample

From Figure 4.1, we noticed that group Moderate has the lowest number (112,983) of bridges. To have an equal sample size from all groups, we randomly selected 112,983 (least number of bridges belonging to Moderate Group) bridges from other groups (Heavy and Light).

The analysis on 338,949 bridges (112,983 bridges in each group) reveals that the PR(>F) value < 0.05 as given in Table 4.7, so we reject the null hypothesis. We observed that there is a change in PR(>F) of a small sample and a large sample

	sum sq	df	F	PR(>F)
ADTT	469.33	2.0	33.44	0.0
Residual	238552.42	338946	-	-
	eta sq	**cohen's d**	**effect size**	**-**
ADTT	0.001	0.063	very small	-
Residual	-	-	-	-

Table 4.7: One-way ANOVA of Average Daily Truck Traffic on the sample size of 338,949 bridges

analysis, which suggests that as the sample size increased, we also see the difference in mean BDS among the ADTT groups. However, the effect size (Cohen's d = 0.063) is very small. A very small effect size suggest a very small association between ADTT and condition of the bridges. From our analysis on small sample size and a large sample size, our findings suggest that the mean BDS score among three groups of the bridges (Light, Moderate, and Heavy) is statistically insignificant. In other words, Average Daily Truck Traffic fails to explain the differences in the bridge condition in comparison to national baseline.

4.4 Effect of Bridge Maintainer

Maintenance practices related to bridges such as Rehabilitation, Replacement and Repair (RRR) of bridge components are essential for the durability of bridges.[5] These bridge maintenance practices may differ across various states and maintainer groups. In this research study we limited our analysis to only the top four agencies that maintain bridges:

1. State Highway Agency
2. County Highway Agency
3. Town or Township Agency
4. City or Municipal Highway Agency

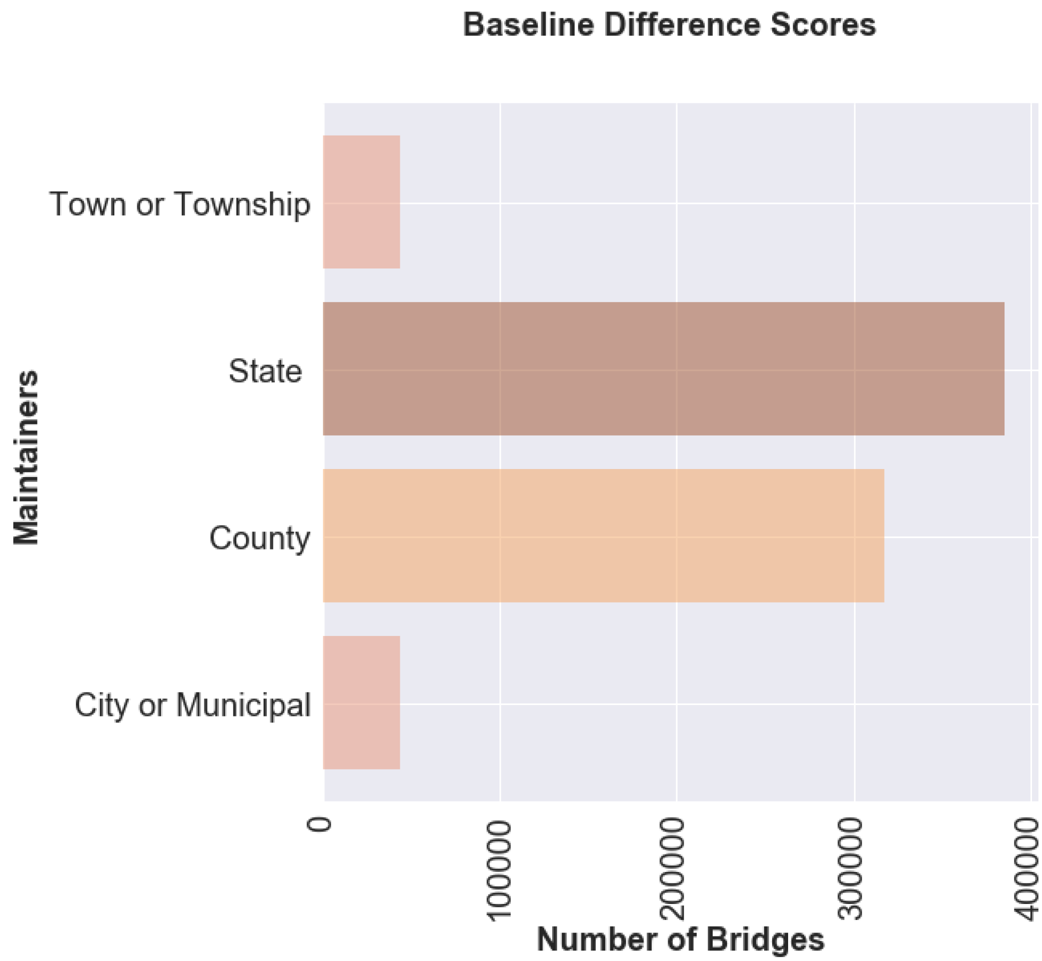

Figure 4.5: Number of Bridges in Town, State, County, and City Maintainers groups

From figure 4.5, a large number of bridges are maintained by the State and County. A smaller, similar number of bridges are maintained by either the Town or the City.

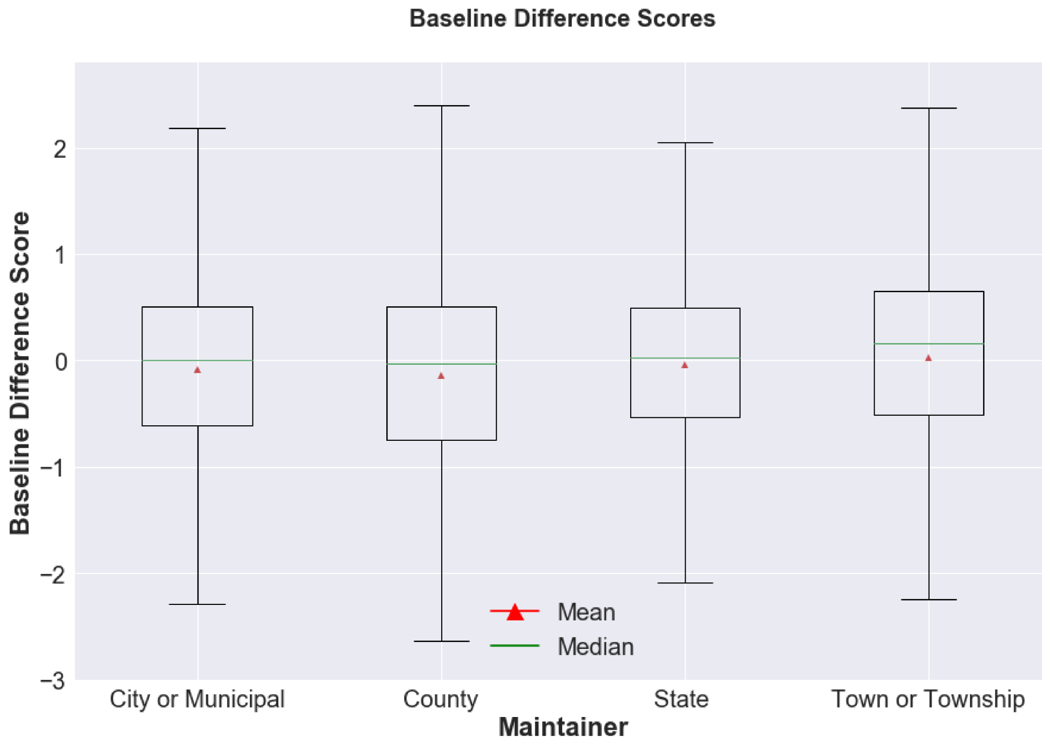

Figure 4.6: Distribution of Baseline Difference Scores of Town, State, County, and City Maintainer group

Figure 4.6, compares the BDS distribution of City, County, State, and Town maintained bridges. The groups of BDS reveals that the mean and the median are similar within the groups of City, County, and State groups. The Town maintainer group have a higher mean and the median. The range of distribution of bridges with State maintained bridges is the smallest. The mean and median are close to zero, the mean and the median suggest that state highway maintained bridges are in better condition.

4.4.1 Results

We performed one-way ANOVA on maintainers: City or Municipal Highway Agency, County Highway Agency, State Highway Agency, and Town or Township Highway Agency with a small sample and a large sample of data.

4.4.1.1 One-way ANOVA - Small Sample

We randomly selected 400 bridges (100 bridges from each maintainers group) to perform ANOVA.

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
Maintainer	4.96	3.0	1.87	0.13	0.013	0.15	very small
Residual	349.54	396.0	-	-	-	-	-

Table 4.8: One-way ANOVA of Maintainer on the sample size of 400 bridges

The results of the ANOVA as seen in the Table 4.8 revealed that the PR(>F) > 0.05, so we fail to reject the null hypothesis. The effect size (Cohen's d) of 0.15 calculated in the analysis reveals a small effect size by convention. This means that the association between maintainer and the condition of the bridge is very small in this sample.

However, the analysis is performed on small sample size (0.22% of the available data). Therefore, it is likely that a very small sample size might not have captured the variance in the data of a large sample. Hence, we did a similar analysis on a large sample.

4.4.1.2 One-way ANOVA - Large Sample

From Figure 4.1, we noticed that group Town or Township Highway Agency have the lowest number (43,530) of bridges. To have an equal sample size from all groups, we randomly selected 43,530 bridges from other groups (State, County, and City Highway Agency).

	sum sq	**df**	**F**	**PR(>F)**
Maintainer	577.07	3.0	245.50	0.0
Residual	136421.92	174116	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Maintainer	0.004	0.12	very small	-
Residual	-	-	-	-

Table 4.9: One-way ANOVA of Maintainer on the sample size of 174,120 bridges

The analysis on 174,120 bridges (43,530 bridges in each group) as seen in the Table 4.9 reveals that the PR(>F) value < 0.05, since the PR(>F) value is less than 0.05 we reject the null hypothesis. This change in PR(>F) value suggests as the sample size increased, we observed a difference in mean BDS among the maintainer groups. However, the effect size (Cohen's d) of 0.12 is very small that suggests a small association between maintainer and condition of the bridges.

From our analysis on small and large samples, our findings suggest that mean BDS score among four groups of the bridges (City, County, State, and Town) are statistically insignificant. In other words, the Maintainer factor alone fails to explain the differences in the bridge condition in comparison to the national baseline.

4.5 Effect of Material

Material of the bridge plays an important role in the durability of the bridge,[23] and are commonly analyzed as an indicator of the bridge condition.[3,6,8,20]

Bridges are often classified by the material of the bridge. In NBI dataset, there are nine Material Types for bridges. But for our analysis, we limited our analysis to four most commonly used Material Types:

1. Concrete
2. Prestressed Concrete
3. Steel
4. Wood and Timber

From Figure 4.7, a large number of bridges in the U.S. are made of Steel and Prestressed Concrete. A smaller number of bridges are made of the Wood or Timber.

From Figure 4.8, the range of the BDS distribution of Prestressed Concrete is lower than the other bridges. The mean and median of the Prestressed Concrete is positive, which suggest that Prestressed Concrete bridges are in better condition than another Material Type of bridges. This may also be due to the fact that Prestressed Concrete bridges are relatively newer technology in bridge construction material. Wood and Timber bridges have the highest range of BDS distribution, and the mean and median of Wood and Timber bridges are below zero.

Figure 4.7: Number of bridges in Wood, Steel, Prestressed Concrete, and Concrete Material Type

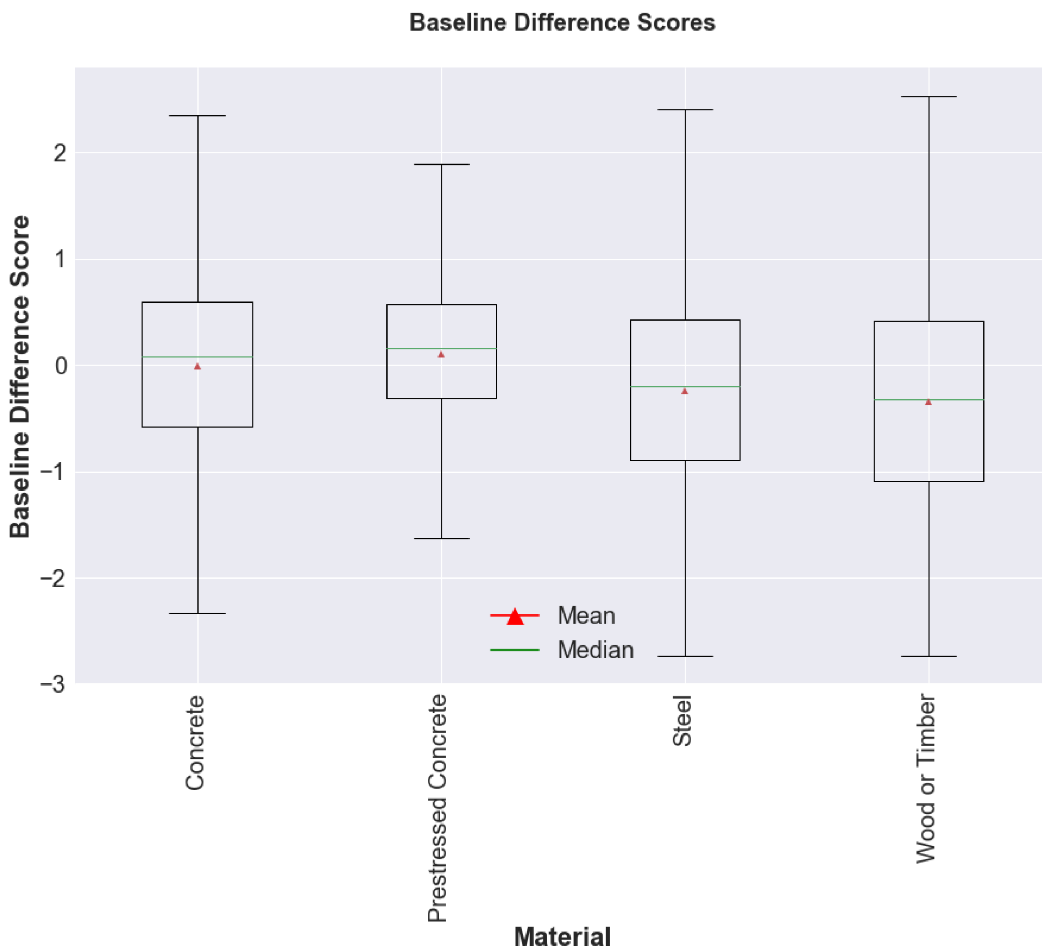

Figure 4.8: Distribution of Baseline Difference Score of Concrete, Prestressed Concrete, Steel, and Wood or Timber Material Type. Notice Prestressed Concrete bridges have the highest mean and median values

To understand boxplot of the BDS score from a different perspective, we provided a density plot 4.9 of the BDS scores with respect to the Material Type of the bridges, that Prestressed Concrete has the sharpest/tallest peak with most of its values skewed towards the positive side.

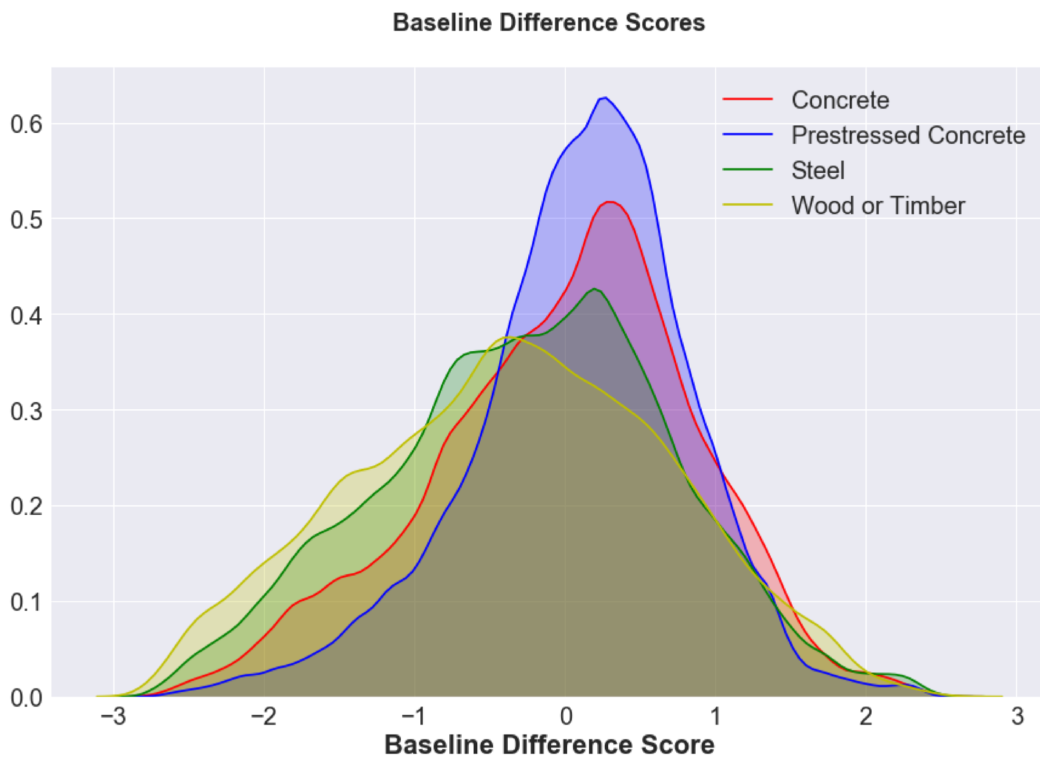

Figure 4.9: Density plot of Baseline Difference Scores of Concrete, Prestressed Concrete, and Steel the bridges in the U.S.

As seen in the Figure 4.9 Wood/Timber bridges has the lowest peak and skewed towards negative values. It also noteworthy that the steel bridges appear to have a bimodal distribution.

4.5.1 Results

In our analysis of the effect of the material on bridge condition, we performed ANOVA on a small sample of randomly selected 100 bridges from each of the four types of material: Concrete, Prestressed Concrete, Steel, and Wood or Timber bridges.

We further performed a similar analysis on a larger sample of bridges 421,792 bridges, that have equal size representation of 105,448 bridges from Concrete, Prestressed Concrete, Steel, and Wood or Timber bridges.

4.5.1.1 One-way ANOVA - Small Sample

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
Material	11.872406	3.0	2.80	0.010	0.02	0.30	small
Residual	277.4	396.0	-	-	-	-	-

Table 4.10: One-way ANOVA of Material Type on the sample size of 400 bridges

The results of the ANOVA as seen in the Table 4.10 revealed that the PR(>F) value < 0.05, so we reject the null hypothesis. This suggests that there is a difference between mean BDS across the four types of materials. The Cohen's d value is 0.30 in the analysis that suggests that there is a small effect size, or in other words a small association between Material Type and condition of the bridges compared to the national baseline.

4.5.1.2 One-way ANOVA - Large Sample

To capture the variance in data, we performed a similar analysis on a large sample of the data. Similar to the previous analysis as seen in the Table 4.11, the results of the ANOVA revealed that with the PR(>F) value < 0.05, so we reject the null hypothesis. The Cohen's d value is 0.35 in the analysis that suggests that there is small effect size.

	sum sq	df	F	PR(>F)
Material	5368.913	3.0	1259.8	0.0
Residual	159216	421788	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Material	0.03	0.35	small	-
Residual	-	-	-	-

Table 4.11: One-way ANOVA of Material Type on the sample size of 421,792 bridges

4.5.1.3 Post-hoc Test

The p-value of one-way ANOVA is not significant at the 99% confidence level, we know that the BDS of different Material Type differs. To understand how these Materials Types differ, we can perform a follow up 'post-hoc test'. One of the post-hoc tests to perform is a separate t-test for each pair of regions.

Group 1	Group 2	reject	Statistic	p-value
Wood or Timber	Steel	True	-12.47	0
Wood or Timber	Concrete	True	-44.44	0
Wood or Timber	Prestressed Concrete	True	-63.28	0
Steel	Concrete	True	-32.78	0
Steel	Prestressed Concrete	True	-51.72	0
Concrete	Prestressed Concrete	True	-17.13	0

Table 4.12: Pair-wise T-Test

In the Table 4.12, the p-values for each t-test suggest that the bridges of each Material Type are different from other Material Types. Since the p-values for each t-test is less than 0.0, using unadjusted pairwise t-tests can overestimate significance. However, we can perform Tukey's test comparison between Material Types.

Group 1	Group 2	meandiff	upper	lower	reject
Concrete	Prestressed Concrete	0.1101	0.0915	0.1287	True
Concrete	Steel	-0.2408	-0.2594	-0.2222	True
Concrete	Wood or Timber	-0.34	-0.3586	-0.3215	True
Prestressed Concrete	Steel	-0.3509	-0.3694	-0.3323	True
Prestressed Concrete	Wood or Timber	-0.4501	-0.4687	-0.4316	True
Steel	Wood or Timber	-0.0993	-0.1178	-0.0807	True

Table 4.13: Tukey's Test on Bridges in the U.S. by Construction Material Type reveals the difference in BDS of all Material Type of the bridges. Note that None of the Material Type pairs fail to reject null-hypothesis

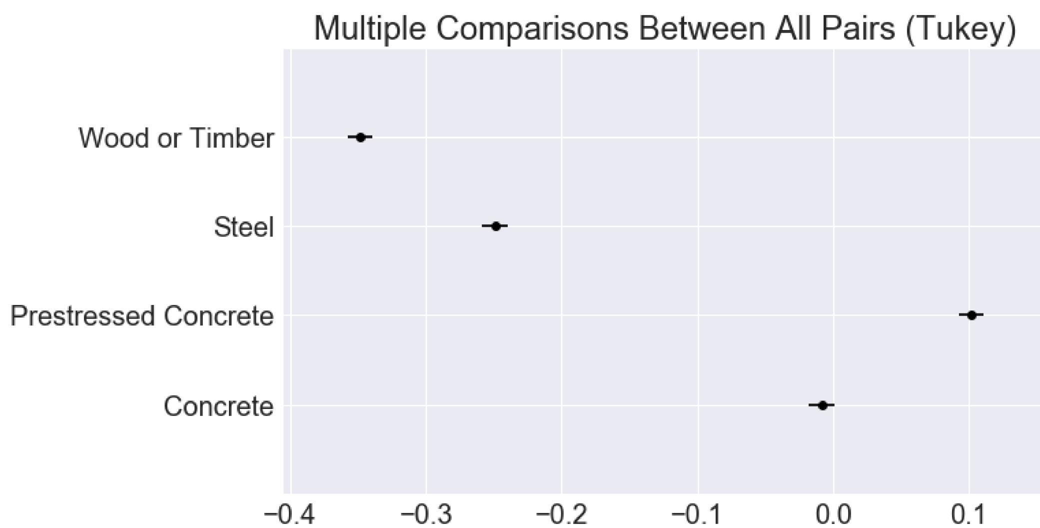

Figure 4.10: Confidence interval plot of all Material Types of the bridges based on Tukey's Test. The results include all available bridges from the U.S.

In Table 4.13, the output of the Tukey's test shows both the average difference between the Material Types and the confidence interval. It also shows whether to reject or fail to reject the null hypothesis for each pair of groups at the given significance level. In this case, the test suggests we reject the null hypothesis for all pairs of Material Type, since the confidence intervals of all pair do not overlap at all, this suggests that condition of bridges with each material is different from each other. Figure 4.10 shows the 95% confidence interval plot reinforces the results visually. As it can be seen in this figure that no Material type's confidence interval overlaps with any other material group's confidence interval.

4.5.2 Further Analysis

Given that prestressed concrete is a relatively newer material, it is appropriate to suggest that age might have an impact on the different Material Types. In other words, prestressed concrete bridges are newer and hence might be better performing. We accounted for age as a factor in our analysis to study the material's effect on the condition of the bridge by holding age constant.

Figure 4.11 shows bridges made of Prestressed Concrete are the youngest in comparison to other material type bridges in this analysis, while Steel and Concrete Bridges have the highest mean age.

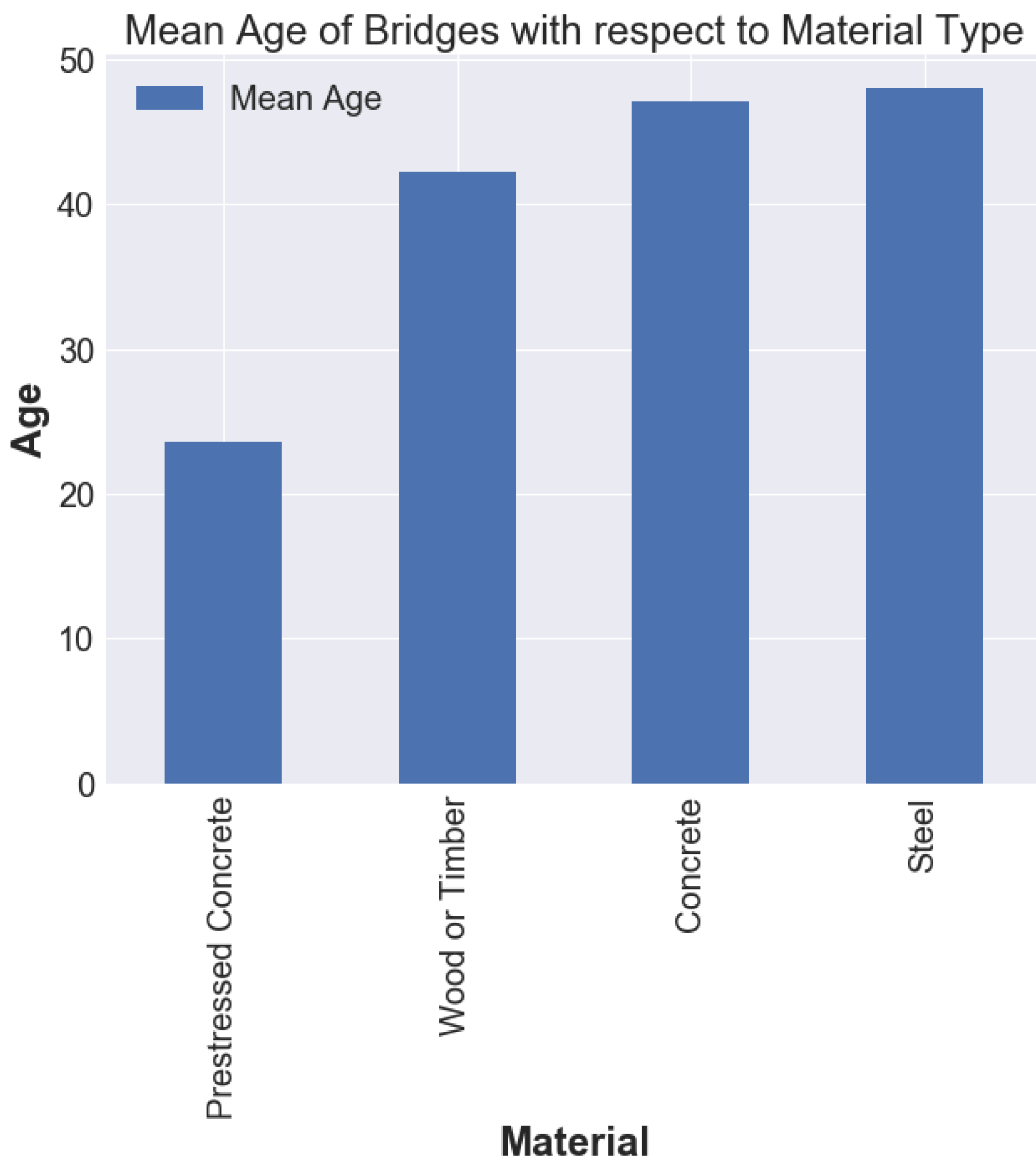

Figure 4.11: Mean Age of the bridges with respect to Materials Type

Based on Figure 4.11, we can further assume that bridges of a given Material Type are not equally represented at any given age. Using the four quartiles of age distribution of the bridges, we can categorize bridges into four age category: Very

Young, Young, Mid Age, and Old. Figure 4.12 shows that our initial assumption of unequal representation is seen clearly as there is a very small number of Wood or Timber bridges in the categories such as Very Young, Young, Mid Age and Old category of the bridge. Prestressed Concrete bridges are nearly absent in the Old category. Steel bridges have a similar proportion of the bridges in each category.

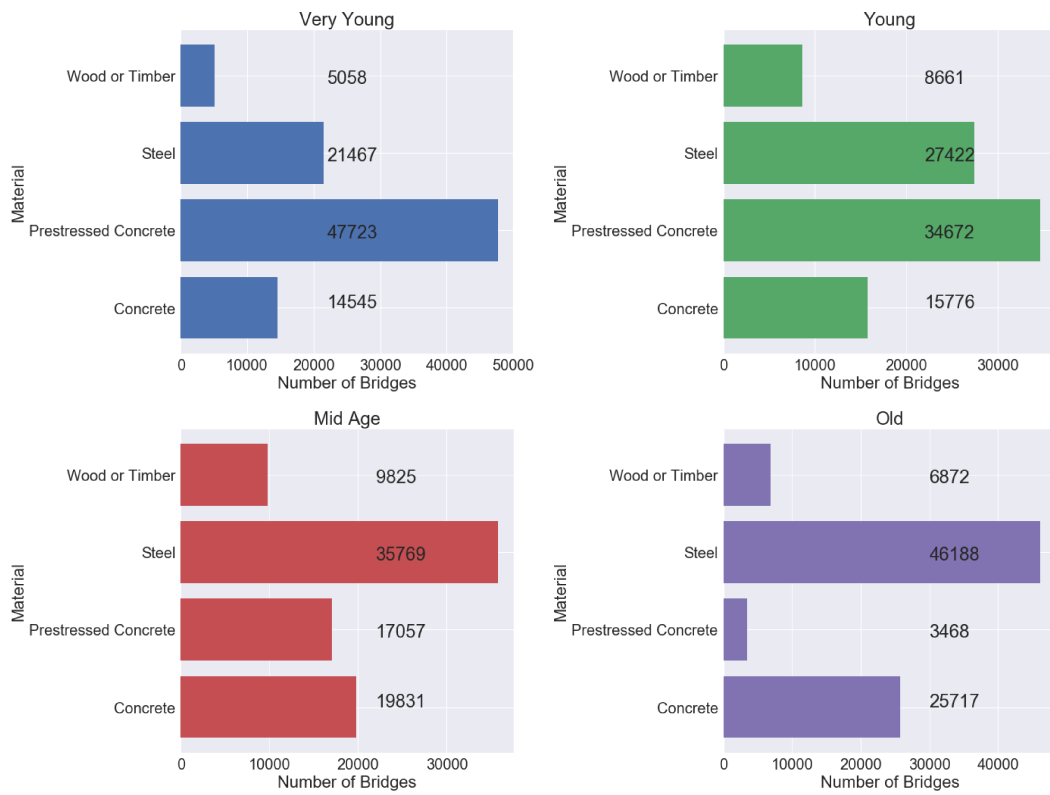

Figure 4.12: Number of Bridges in each categorization of bridges in Age group

To perform ANOVA by accounting age in this analysis, we selected a category of bridges with the most and equal data sample representing every Material Type and considered that Prestressed Concrete are likely to be more younger than other Material Types of bridge.

In the Young category of the bridges, Wood or Timber type of bridges are the most underrepresented. However, in comparison to other categories, Wood or Timber have the most number of bridges in the Young criteria. Therefore, an undersampling of the Young criteria of the bridges allowed us to have the highest number of data possible having equal representation from Material Type of the bridges.

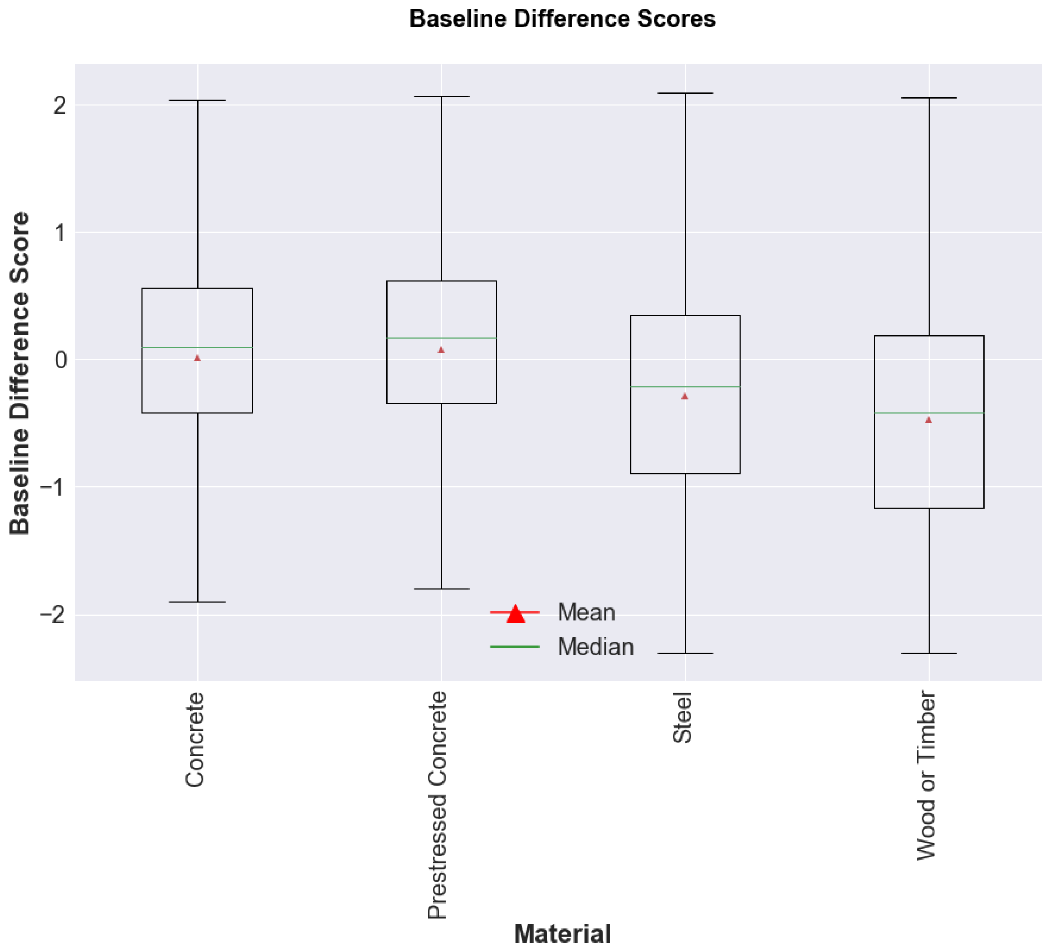

Figure 4.13: Distribution of Baseline Difference Score in Young Category of Concrete, Prestressed Concrete, Steel, and Wood or Timber. Notice Prestressed Concrete bridges have the highest mean and median values

In Figure 4.13, the distribution of BDS is normally distributed for each Material Type in the Young category. The mean and median are almost the equal within the category for all type of bridges, Concrete Bridges and Prestressed Concrete bridges have the same distribution. Steel bridges have a wider range of BDS scores and Wood or Timber bridges have the lowest mean and median compared to other Material Types.

4.6 Results of Young category of Bridges

We performed a similar ANOVA on a small sample of randomly selected 100 bridges and large sample analysis on selected 8,661 bridges (a total of 34,644 bridges) from Young category bridges.

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
Material	22.51	3.00	10.68	0.0	0.07	0.54	medium
Residual	278.25	396.00	-	-	-	-	

Table 4.14: One-way ANOVA of Material Type on the sample size of 400 bridges - Young

The results of the ANOVA as seen in the Table 4.14 revealed that the PR(>F) value < 0.05, so we reject the null hypothesis. By keeping the age constant for all bridges, there is still a difference between mean BDS across the four types of materials. The Cohen's d value is 0.54 in the analysis that suggests that there is a medium association between Material Type and condition of the bridges in comparison to the national baseline.

	sum sq	**df**	**F**	**PR(>F)**
Material	1745.02	3.0	888.50	0.0
Residual	22677.81	34640	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Material	0.07	0.54	medium	-
Residual	-	-	-	-

Table 4.15: One-way ANOVA of Material Type on the sample size of 34,644 bridges - Young

In a similar large sample analysis seen in the Table 4.15, the results of the ANOVA revealed that the PR(>F) value < 0.05, so we reject the null hypothesis. The Cohen's d value is 0.54 in the analysis that suggests that there is a medium effect size between the Material Type and condition of the bridges in comparison the national average within Young category.

4.6.1 Post-hoc Test of Young category of bridges

In our similar post-hoc test of Young category of bridges, the test result 4.16 each Material Type is different from other Material Types.

Group 1	Group 2	reject	Statistic	p-value
Wood or Timber	Steel	True	-11.73	0
Wood or Timber	Concrete	True	-44.09	0
Wood or Timber	Prestressed Concrete	True	-63.31	0
Steel	Concrete	True	-33.27	0
Steel	Prestressed Concrete	True	-51.71	0
Concrete	Prestressed Concrete	True	-16.69	0

Table 4.16: Pair-wise T-Test

In Table 4.17, the p-values for each t-test suggest that the bridges of each Material Type are different from other Material Types. Since the p-values for each t-test is less than 0.05, we also performed did Tukey's test to take into account overestimation of significance from the t-test.

Group 1	Group 2	meandiff	upper	lower	reject
Concrete	Prestressed Concrete	0.1071	0.0885	0.1256	True
Concrete	Steel	-0.2432	-0.2617	-0.2247	True
Concrete	Wood or Timber	-0.3362	-0.3547	-0.3177	True
Prestressed Concrete	Steel	-0.3503	-0.3688	-0.3317	True
Prestressed Concrete	Wood or Timber	-0.4433	-0.4618	-0.4248	True
Steel	Wood or Timber	-0.093	-0.1116	-0.0745	True

Table 4.17: Tukey's Test on Bridges in Young Category in the U.S. by Construction Material Type reveals the difference in BDS of all Material Type of the bridges. Note that none of the Material Type pairs fail to reject null-hypothesis

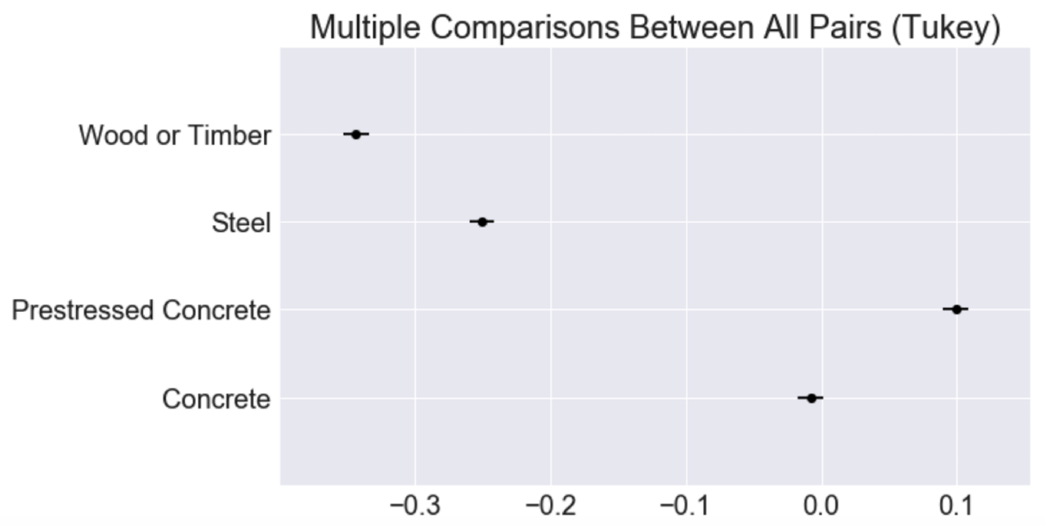

Figure 4.14: Confidence interval plot of all Material Types of the bridges based on Tukey's Test. The results include all available bridges from the U.S.

The test suggests that we reject the null hypothesis for all pairs of Material Type within the Young category of bridges. No overlapping of the confidence interval between the pairs of Material Type suggests that each Material Type is different from each other. Figure 4.14 suggests that the difference between Wood and Steel is the smallest followed by Concrete and Prestressed Concrete. Prestressed Concrete and Wood bridges have the biggest difference.

4.7 Effect of Structure Length

Structure Length of the bridge is also commonly studied as the factor that can determine the condition of the bridges.[15,16,20] To understand the effects of Structure Length on the condition of the bridges, we classify bridges into two Structure Length groups as seen in the Table 4.18, the Very Long (Top 5% of the length of bridges in the U.S.) and Very Short (Bottom 5% of the length of bridges in the U.S.) in the U.S. respectively, and performed ANOVA.

Group	**Range of Structure Length**
Very Short	6.1 - 8.5 (m)
Very Long	58.9 - 38421 (m)

Table 4.18: Grouping of Bridges by Structure Length in meters

In Figure 4.15, there is a larger number of bridges in the Very Long category in comparison to the Very Short Category of the bridges. However, Figure 4.16 reveals a wider range of BDS in the Very Short category of the bridges.

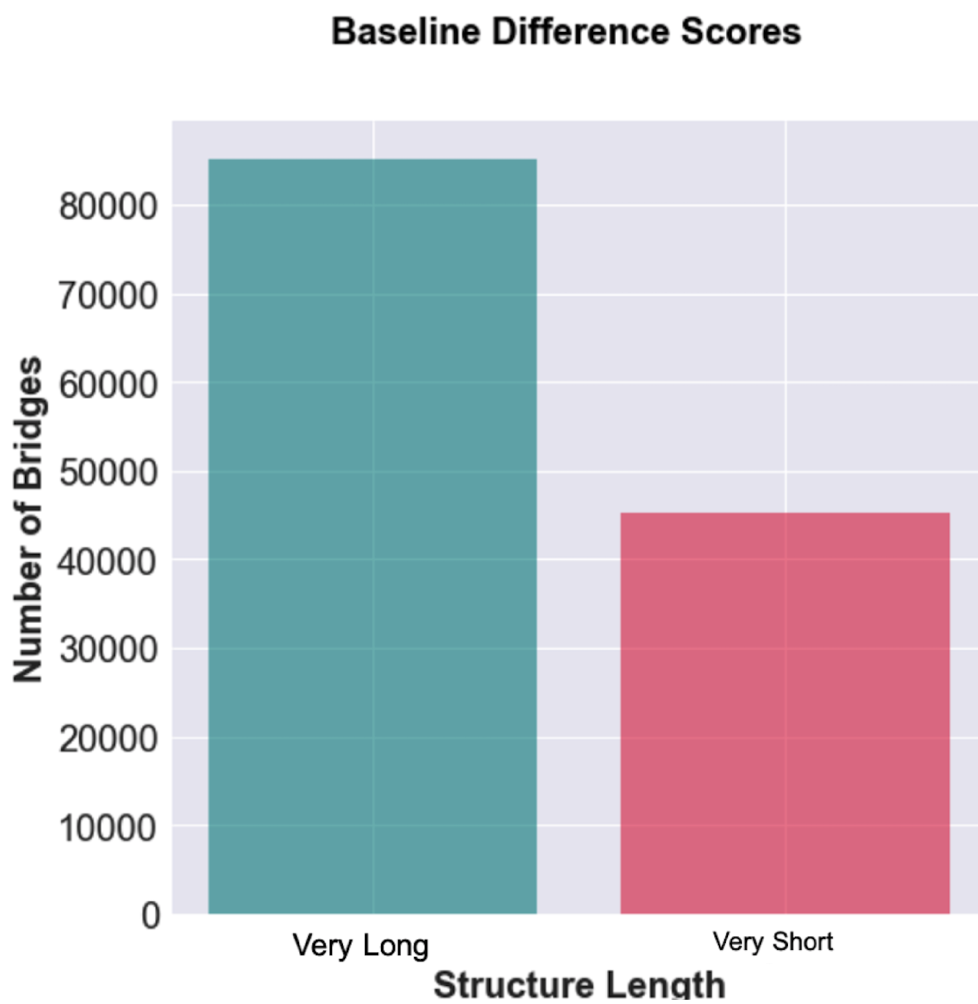

Figure 4.15: Number of Bridges with Very Long and Very Short Structure Length groups. Notice a high representation of Very High Structure Length group

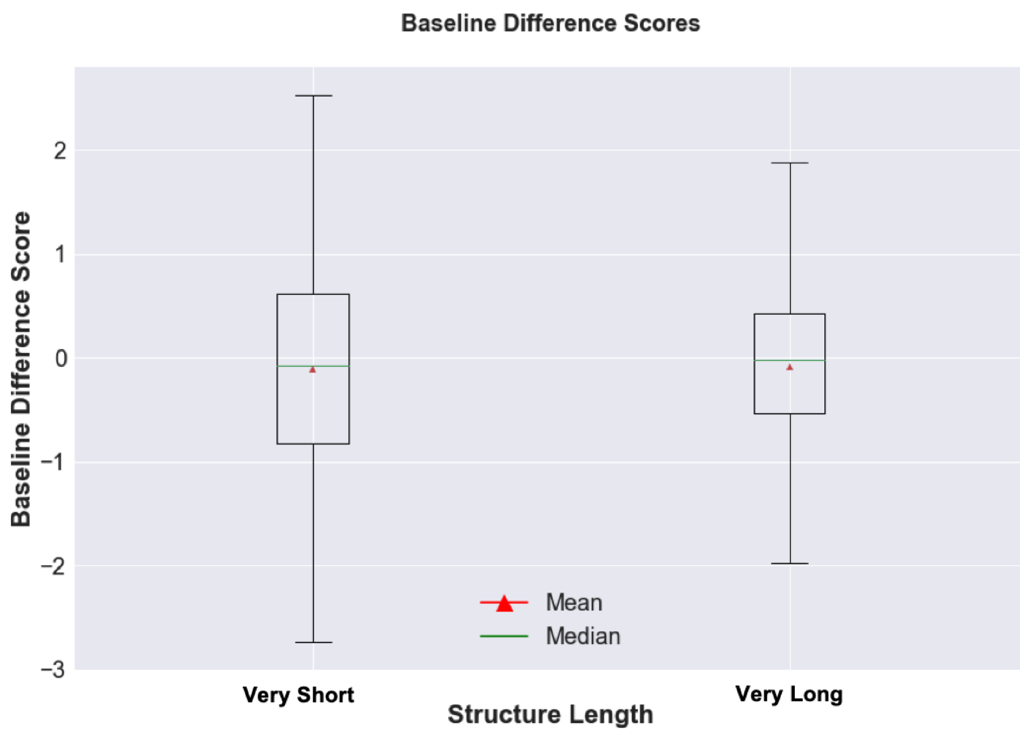

Figure 4.16: Distribution of Baseline Difference Scores of Very Long and Very Short Structure Length group

4.7.1 Results

For performing ANOVA on small sample, we selected a random sample of 200 bridges (100 from each group), and large sample data that includes randomly selected 90,242 bridges (45,121 from Very Long and Very Short groups).

4.7.1.1 One-way ANOVA - Small Sample

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
Structure Length	0.19	1.0	0.21	0.63	0.001	0.06	very small
Residual	176.32	198.0	-	-	-	-	-

Table 4.19: One-way ANOVA of Structure Length on the sample size of 200 bridges

The result of ANOVA on a data sample of 200 bridges (100 bridges each structure length group), has the PR(>F) value (0.63) > 0.05 as seen in Table 4.19, so we fail to reject the null hypothesis. In order words, there is a difference between mean BDS across Very Long and Very Short Structure Length. The Cohen's d value of 0.06 reveals a very small effect size between the structure length and BDS. This means there is a very small association between Structure Length and condition of bridges in comparison to the national average.

4.7.1.2 One-way ANOVA - Large Sample

	sum sq	**df**	**F**	**PR(>F)**
Structure Length	21.4	1.0	26.02	0.0
Residual	74295.28	90240.00	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Structure Length	0.0002	0.02	very small	-
Residual	-	-	-	-

Table 4.20: One-way ANOVA of Structure Length on the sample size of 90,242 bridges

Our results suggest that with a PR(>F) value (0.0) < 0.05 as seen in the Table 4.20, so we reject the null hypothesis in a large sample size of the bridges for each group of structure length. There are differences in mean between Very Long and Very Short Structure Length of the bridges.

However, with Cohen'd value of 0.02 suggest that association between structure length and BDS is very small. This means that Structure Length alone fails to explain the differences in bridge condition compared to the national baseline.

4.8 Effect of Precipitation

External factor such as precipitation are cited as an indicator for the condition of bridges.[8] The challenge of analyzing the effect of external factors on the condition of the bridges is that the data for these factors are not available in the NBI dataset.

We use the dataset from the Center for Disease Control and Prevention to analyze the effects of Precipitation on bridge conditions. In Figure 4.17, we see eight bands of precipitation, from a range of 0 to 7.64 average daily precipitation (mm) at county - level.

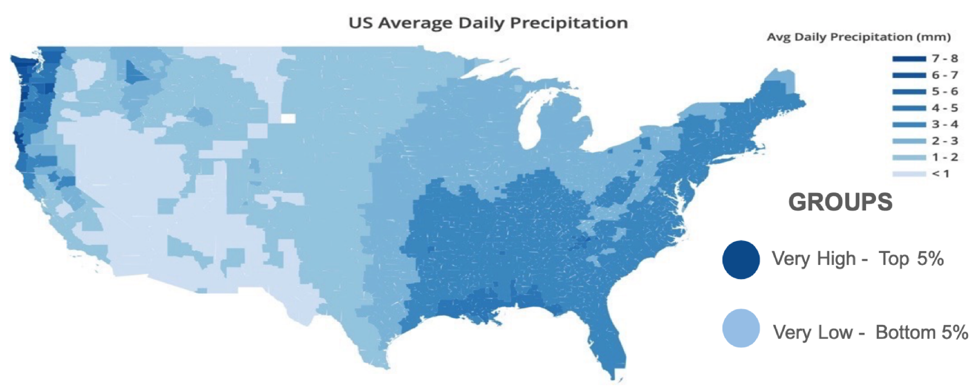

Figure 4.17: County-level Precipitation Data from CDC

To understand the effects of precipitation on the condition of the bridges, we classify bridges into two precipitation groups: Very High (Top 5% of the Precipitation in the U.S) and Very Low (Bottom 5% of the Precipitation in the U.S.) respectively, and performed ANOVA.

In Table 4.21, shows the range of precipitation in the Very Low and Very High group.

Since the distribution of the BDS is a normal distribution, The grouping of bridges in top 5% and bottom 5% precipitation as described in Figure 4.21, allowed to have an equal number of bridges.

Precipitation Group	Range of Precipitation
Very High	0.43 - 1.23 (mm)
Very Low	3.76 - 7.64 (mm)

Table 4.21: Classification of Groups by Precipitation

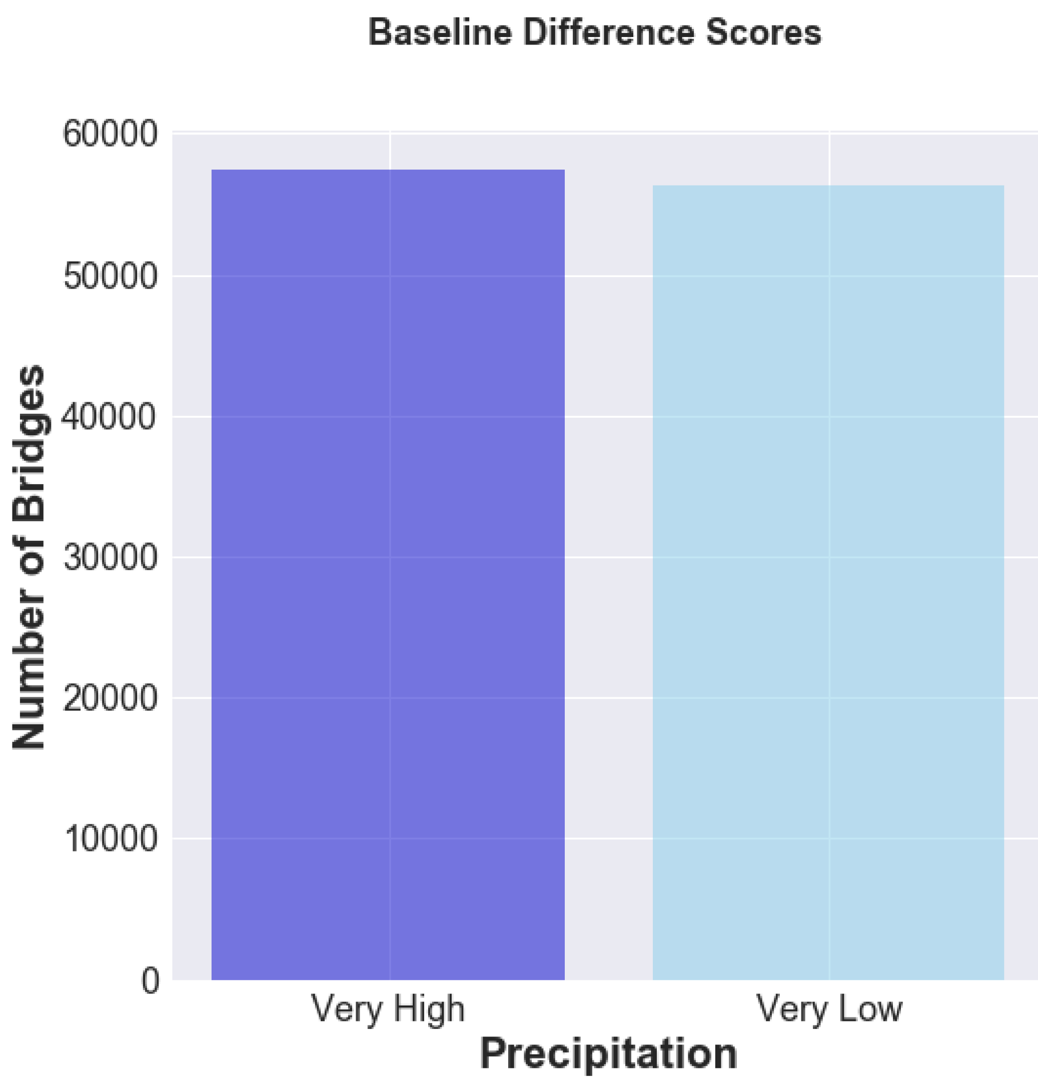

Figure 4.18: Number of Bridges in Very High and Very Low Precipitation regions

In Figure 4.19, we can see the distribution for BDS that Very High precipitation and Very Low precipitation are very similar. The mean and median score of Very Low precipitation is higher than Very High precipitation.

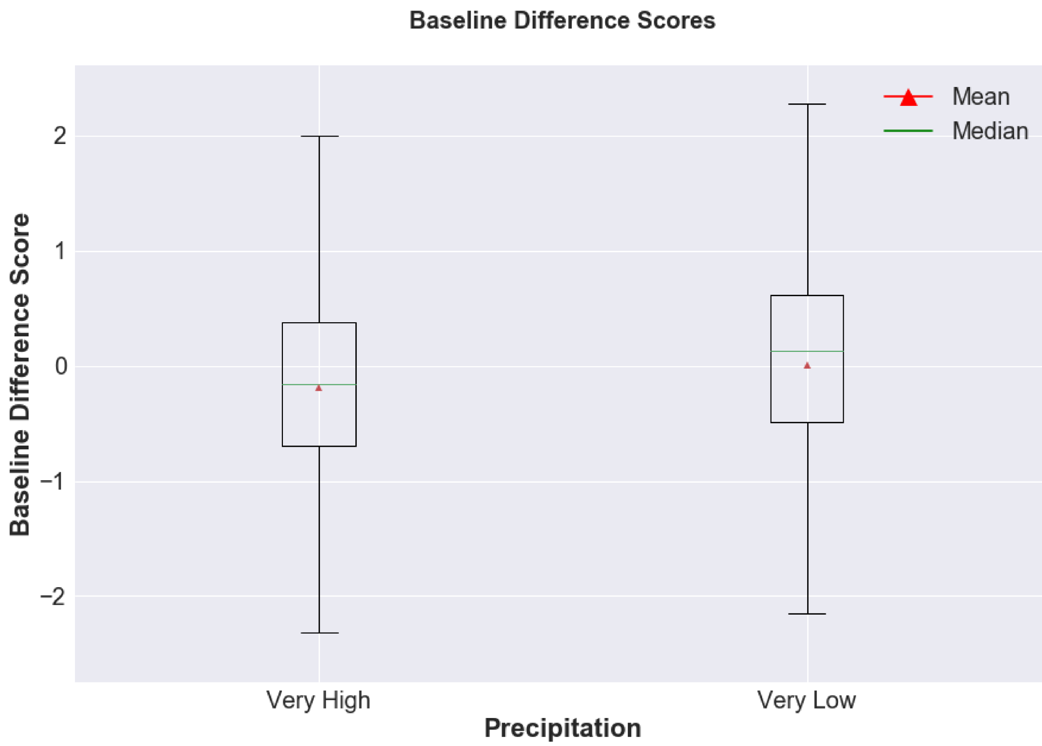

Figure 4.19: Distribution of Baseline Difference Scores of Very High and Very Low Precipitation regions of the U.S. Notice the mean and the median of Very Low is higher than the mean and median of Very Low Precipitation region

4.8.1 Results

For performing ANOVA on small sample, we selected a random sample of 200 bridges (100 from each group), and large sample data that includes randomly selected 112,000 bridges (56,000 from each group).

4.8.1.1 One-way ANOVA - Small Sample

	sum sq	**df**	**F**	**PR(>F)**	**eta sq**	**cohen's d**	**effect size**
Precipitation	1.73	1	2.54	0.11	0.01	0.20	small
Residual	134.88	198.0	-	-	-	-	-

Table 4.22: One-way ANOVA of Precipitation on the sample size of 200 bridges

From our results of ANOVA on a data sample of 200 bridges (100 bridges each precipitation region) as seen in 4.22, the PR(>F) value (0.11) > 0.05, so we fail to reject the null hypothesis. In order words, there is a difference between mean BDS across Very High and Very Low precipitation regions. The Cohen's d value of 0.20 reveals a small effect size between the precipitation and BDS. This means there is a small association between precipitation and condition of bridges.

4.8.1.2 One-way ANOVA - Large Sample

	sum sq	**df**	**F**	**PR(>F)**
Precipitation	1070.55	1.0	1520.14	0.0
Residual	78873.58	111998.0	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Precipitation	0.01	0.20	small	-
Residual	-	-	-	-

Table 4.23: One-way ANOVA of Precipitation on the sample size of 112,000 bridges

In Table 4.23, our results suggest that the PR(>F) value (0.0) < 0.05, so we reject the null hypothesis in a large sample size of the bridges for each group.

There are differences in mean between very high precipitation region and very low precipitation region.

However, with Cohen'd of 0.20 the effect size in both the large sample analysis and small sample analysis suggest that association between precipitation and BDS is small by convention. This means that precipitation alone fails to explain the differences in bridge condition compared to the national baseline.

4.9 Effect of Region

Effect of factors on the bridge condition may vary across different regions of the country. For instance: type of construction, type of material, development maintenance practices may vary depending on the owner and maintainer of the bridges. Externals factors such as temperature also vary across regions. Therefore, the region serves not only as a proxy for a particular effect but, also as a proxy for the compound effect of factors such as precipitation, maintenance that might affect the condition of the bridges.

For our analysis, we group bridges into four regions according to U.S. Census Bureau Regions and Divisions.

1. Northeast
2. Midwest
3. South
4. West

Table 4.24 provides the following grouping of states into regions is widely used for data collection and analysis.

Region	States
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania.
Midwest	Illinois, Indiana, Michigan, Ohio, and Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota
South	Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, District of Columbia, West Virginia Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas
West	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, California, Hawaii, and Oregon.

Table 4.24: Grouping of bridges into their Regions with respect to each bridge's State

As seen in Figure 4.20, there is an imbalance in the number of bridges in each region. Midwest and South have the largest number of bridges, they have an equal number of bridges. West region has the least number of bridges.

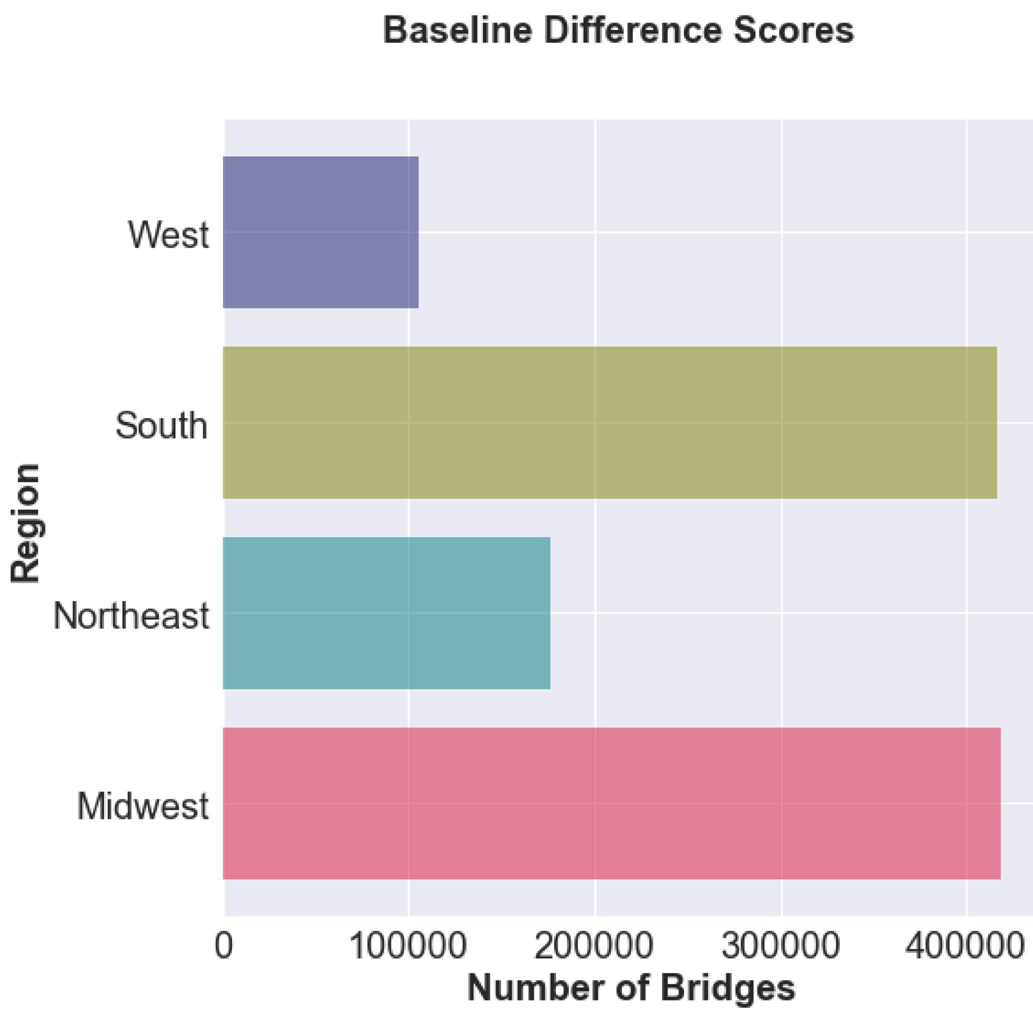

Figure 4.20: Number of Bridges in West, South, Northeast, and Midwest of the U.S, Notice the high concentration of bridges are in the regions of South and Midwest

Figure 4.21, reveals that the BDS across the region are similar. Note that the mean and median values of BSD in the West are equal to zero, that suggesting a perfectly normal distribution. The distribution of Northeast is similar to South. The condition of bridges in the Midwest are better than other regions, as seen in the figure, the mean and median value is higher than any other regions.

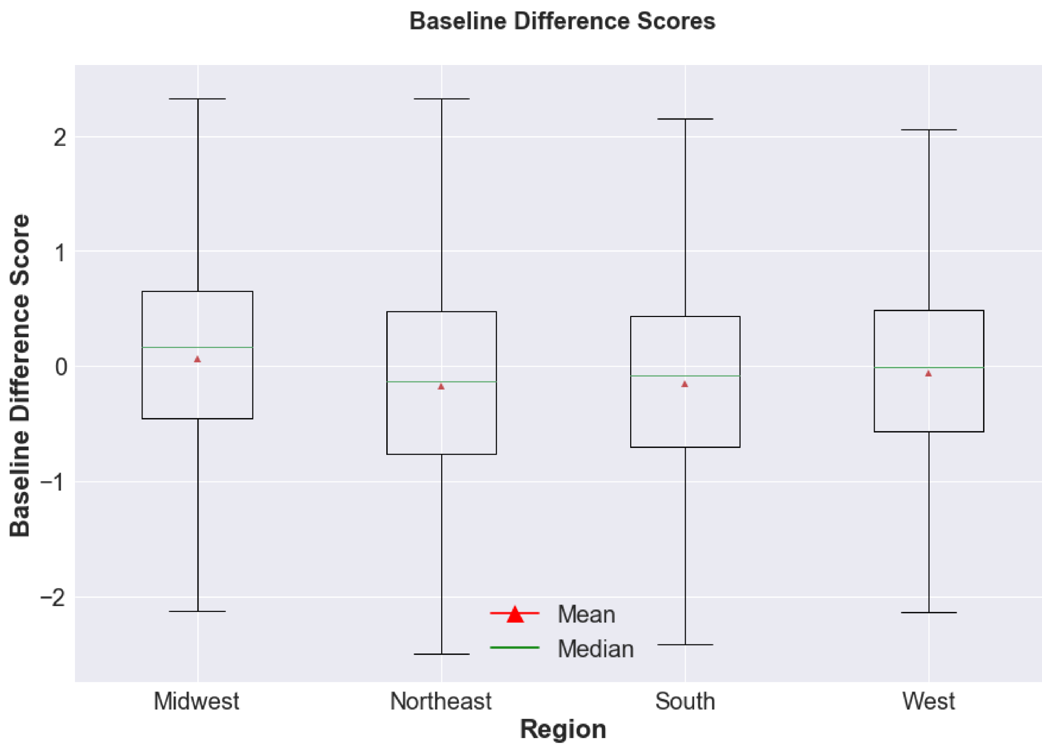

Figure 4.21: Distribution of Baseline Difference Scores of Midwest, Northeast, South, and West Regions of the U.S., Notice the mean and the median are almost equal and close to zero with the distribution of all groups

4.9.1 Results

Similar to previous analysis, we performed ANOVA on a small sample of 400 bridges (100 from each region: West, Northeast, Midwest, and South) and a large Sample of 421,792 (105,448 each region) bridges.

4.9.1.1 One-way ANOVA - Small Sample

	sum sq	df	F	PR(>F)	eta sq	cohen's d	effect size
Region	2.87	3.0	1.19	0.31	0.008	0.17	very small
Residual	317.5	396.0	-	-	-	-	-

Table 4.25: One-way ANOVA of Region on the sample size of 400 bridges

The results seen in Table 4.26 of the ANOVA revealed that with the PR(>F) value > 0.05. Hence, we do not reject the null hypothesis. Therefore, the difference in the mean BDS across all regions are not statistically significant. The Cohen's d value is 0.17 in the analysis reveals a very small association between region and condition of the bridges.

4.9.1.2 One-way ANOVA - Large Sample

	sum sq	df	F	PR(>F)
Region	3580.13	3.0	1558.97	0.0
Residual	322874.96	421788.0	-	-
	eta sq	**cohen's d**	**effect size**	**-**
Region	0.01	0.20	small	-
Residual	-	-	-	-

Table 4.26: One-way ANOVA of Region on the sample size of 421,792 bridges

With a large sample size of randomly selected 421,792 bridges, we are able to notice a difference between the mean BDS of different regions. The PR(>F) value (0.20) < 0.05 as seen in the Table 4.26. Hence, we reject the null hypothesis. However, there is a small difference between region and condition of bridges. In other words, the region of the bridge fails to explain the differences in the bridge condition in comparison to the national baseline.

4.10 Summary

In summary, we performed one-way ANOVA to test the effects of factors: Average Daily Traffic, Average Daily Truck Traffic, Maintainers, Material, Structure Length, Precipitation, and Region on the condition of the bridges. The Summary of ANOVA analysis is provide in the Table 4.27. The Table 4.28 shows the effect size observed in the analysis of a large sample.

Factors	**Fail to Reject Null Hypothesis - Small Sample**	**Fail to Reject of Null Hypothesis - Large Sample**
Average Daily Traffic	True (0.8)	False (0.0)
Average Daily Truck Traffic	True (0.12)	False (0.0)
Maintainer	True (0.13)	False (0.0)
Material	False(0.01)	False (0.0)
Material (Age Constant)	False(0.0)	False (0.0)
Structure Length	True (0.63)	False (0.0)
Precipitation	True (0.11)	False (0.0)
Region	True (0.3)	False (0.0)

Table 4.27: Summary of Analysis - ANOVA

Factors	**Effect Size of a Small Sample**	**Effect Size of a Large Sample**
Average Daily Traffic	Very Small (0.08)	Very Small (0.04)
Average Daily Truck Traffic	Small (0.20)	Very Small (0.06)
Maintainer	Small (0.15)	Very Small (0.12)
Material	Small (0.30)	Small (0.35)
Material (Age Constant)	Medium (0.54)	Medium (0.54)
Structure Length	Very Small (0.06)	Very Small (0.02)
Precipitation	Small (0.20)	Small (0.20)
Region	Very Small (0.17)	Small (0.20)

Table 4.28: Summary of Analysis - Effect Size

Our findings reveal that mean BDS of bridges across Material groups were significant in the analysis of the small sample and a large sample of the bridges. Further post-hoc analysis (Young Category) revealed that the bridges of different material type have a confidence interval that does not overlap with other material types. Our post-hoc of results suggest that there is a difference in the condition of bridges with respect to the material type of the bridge.

In our analysis of external factors, there is a small association (Cohen'd = 0.20) between precipitation and BDS. Similarly, region alone can not explain the variance in the condition of the bridges as the association (Cohen'd = 0.20) between region and condition of the bridges is small.

Bridge with BDS of zero is equal to the average condition of the bridge, so an alternative perspective in understanding the effect of factors on the condition of the bridges is to look at the percentages of bridges above the BDS of zero and below the BDS of zero.

In Figure 4.22, we see that Prestressed Concrete bridges have the highest number of bridges above the BDS of zero from the age 30 to age 45 (Young Category). Wood bridges have a lower percentage of bridges above the BDS score of zero. The Figure 4.22, reveals that Prestressed Concrete bridges perform better than other Material Types in this analysis.

A similar analysis of precipitation reveals a counter-intuitive result. In Figure 4.23, we observe that regions with very high precipitation region have a higher percentage of the bridge above the BDS of zero in comparison to the very low precipitation region.

However, The results from Figure 4.24 and Figure 4.25 suggest the choice of the material of the bridges in these precipitation regions might explain the counter-intuitive result.

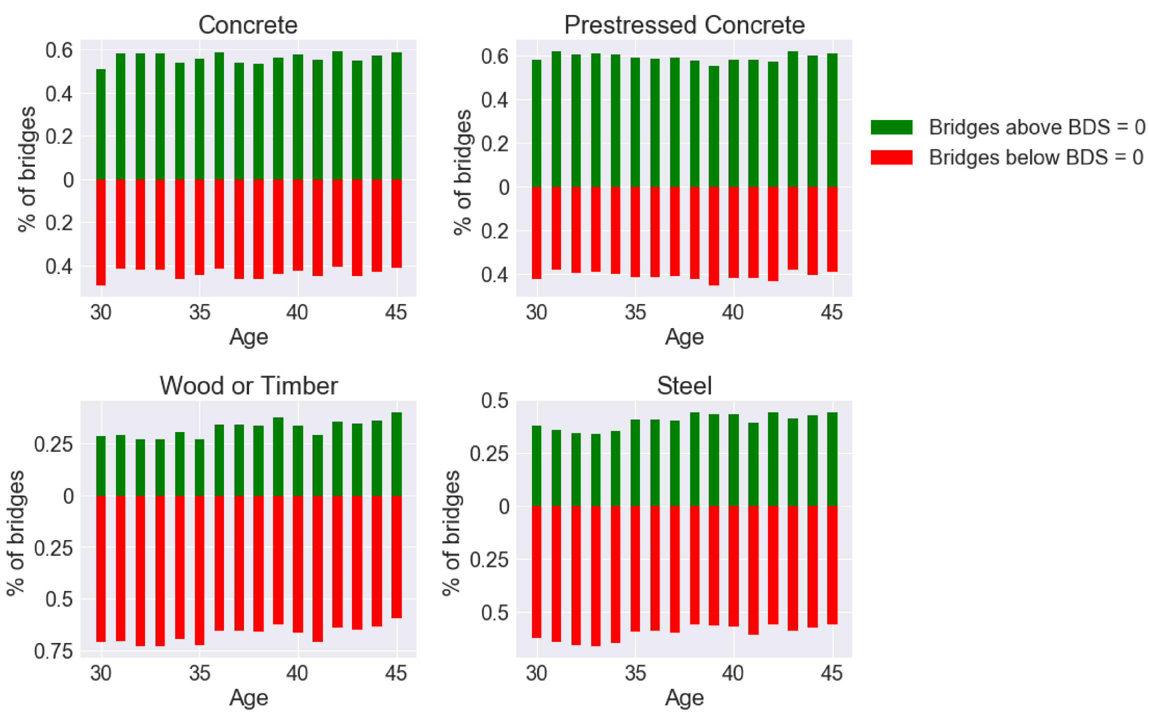

Figure 4.22: Percentage of bridges above and below the BDS of zero with respect to Material Type

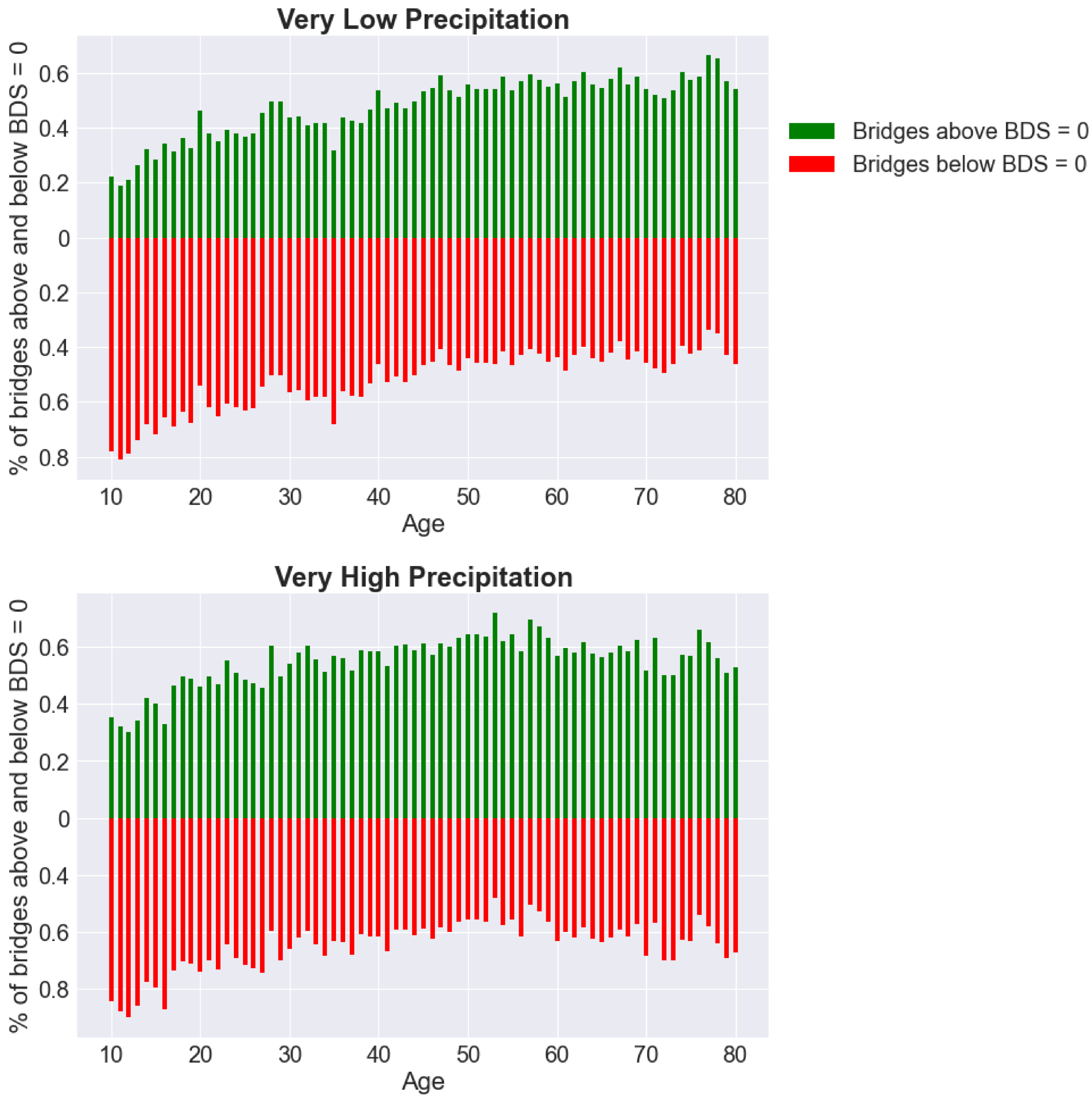

Figure 4.23: Percentage of bridges above and below the BDS of zero with respect to Precipitation regions

We know from the previous results 4.22 that Prestressed Concrete bridges perform better than Wood or Timber bridges. In the Figure 4.24 that the distribution of Prestressed Concrete bridges are low in very low precipitation region and the distribution of Prestressed Concrete bridges are high in very high precipitation region 4.25. Similarly, the Wood or Timber bridges are nearly absent in high precipitation regions. Hence, the choice of the material type of the bridges might explain the counter-intuitive results.

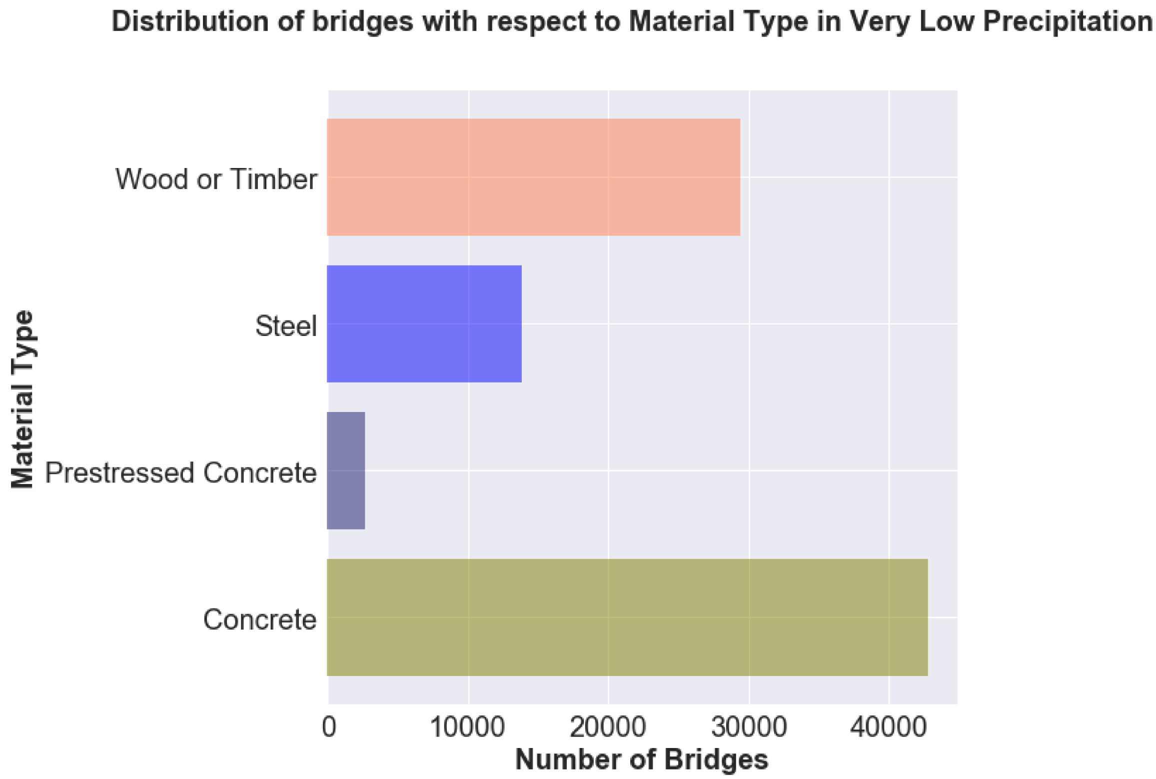

Figure 4.24: Distribution of bridges with respect to Material Type in Very Low Precipitation

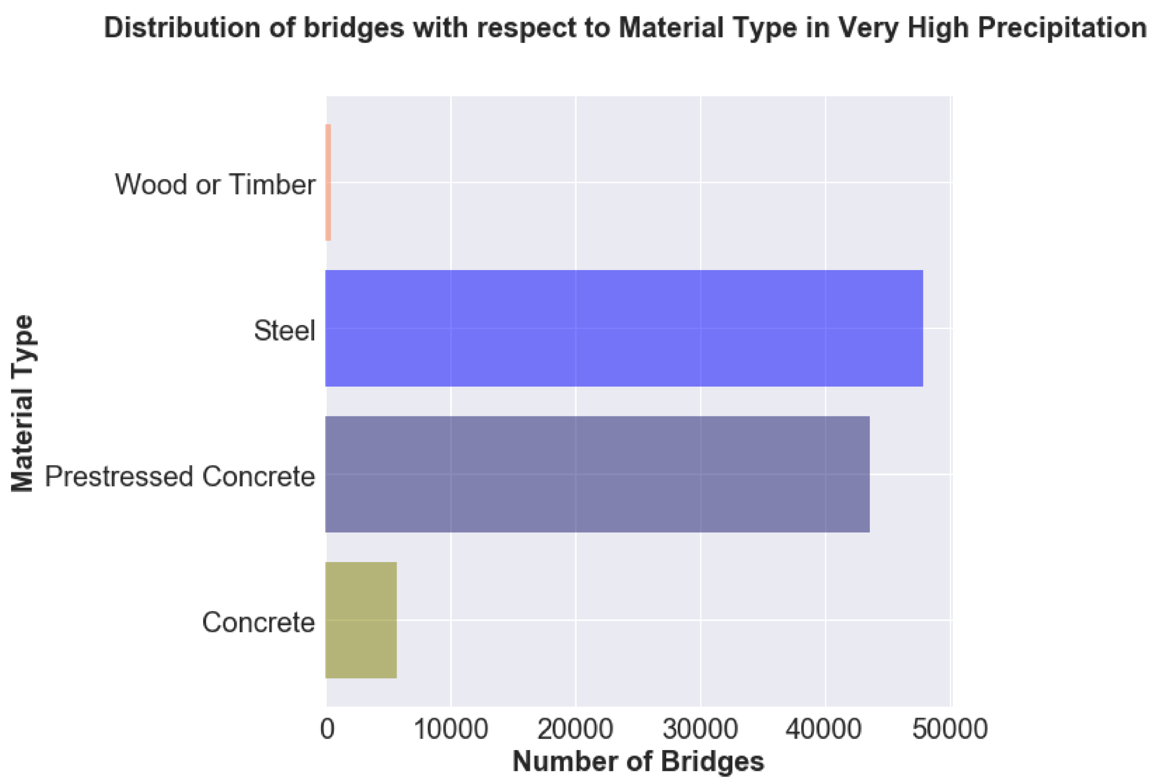

Figure 4.25: Distribution of bridges with respect to Material Type in Very High Precipitation

Overall, our findings suggest that we need to identify other relevant factors and develop a hypothesis that can guide in testing the compound effects of the factors on the condition of the bridges.

Chapter 5

Discussion

5.1 Overview

In this chapter, we will reflect on previous literature in the new light of our findings from Chapter 4. In the following sections, our discussion will relate to the quality of the dataset; development process and environment; and the assumptions, methods, factors, and limitations of our work.

5.2 Data

Each condition rating in the NBI dataset is based on subjective assessments by a bridge inspector. In addition to this subjectivity, we also observed that the NBI dataset has a lot of missing data, such as missing condition ratings in some years of a bridge and missing data related to repairs and reconstruction.[3] The NBI dataset also has a substantial number of bridges where the condition ratings do not change in any of the recorded inspections.

Researchers have proposed methods that address many of the previously mentioned inconsistencies in the NBI dataset,[3,4] but there are no methods available that can account for the subjectivity in the condition ratings. We attempt to address this issue by using a national baseline for estimating bridge performance.

The bridge inspection cycle are done biannually. Researchers have debated that this inspection cycle is an inefficient use of time and resources because newer bridges are in good condition and hence will take a long time before they will deteriorate, while bridges with a bad condition might require more frequent inspections.[4]

Washer et al. further concluded that their model of inspection cycle estimation will help provide efficient data collection with respect to the resources. However, the data collected will still be subjective. Therefore, there is also a requirement to improve data collection. A possible solution could be the use of Internet-of-Things (IoT) devices to collect data from the bridges.

In our study, we address issues related to NBI by applying cleaning and filtering criteria, that we identified in chapter 3. A cleaned NBI dataset is necessary to build a complete and robust model of bridge conditions in the U.S. Unfortunately, we were not able to find any cleaned dataset or working datasets from previous research. Therefore in order to make our work reproducible, we have shared our working data and cleaned NBI dataset available on both the National Data Service (NDS) and on GitHub. We hope this will help other researchers by saving them time, effort, and other computing resources by providing them with a clean dataset.

5.3 Development Environment

In this section, we address Research Question 2 related to exploring data sharing platforms for large datasets among researcher and practitioners. Large datasets like the NBI database are difficult to share among other researchers and practitioners. In our research, the Labs Workbench from National Data Service (NDS) provides and integrated environment for storing research datasets, data sharing, data exploration, and research collaboration using Jupyter notebooks.

The NDS Labs workbench not only reduces the work required in data cleaning and setting up a work environment, but also maintains reproducibility of this research. The Labs Workbench also offers features such as a sharable filesystem and security mechanisms for logging and monitoring. These features allows controlled sharing of datasets and workspace with other researchers and practitioners.

The central idea of the NDS is to discover available management tools from catalogs and services, compare and evaluate various technologies, deploy test instances and provide tools and technology to support cloud-based development and publish/share your tools for others to discover.[24]

In this study, the Labs Workbench environment consisted of Jupyter Notebooks and a centrally hosted MongoDB database with NBI records from all years and all states. The NBI dataset is populated from in CSV format files available from FHWA website. The database is curated, cleaned, and populated in a MongoDB Database using Python scripts. The Jupyter Notebook environment enables us to execute python scripts and display charts in a shareable and readable document. Python scripts are developed as part of this research study for data processing such as extraction, formatting and curating dataset.

Links to the Data Cleaning and Analysis Scripts can be found here:

- Data Cleaning and Filtration
- Data Analysis

5.4 Method

Baseline Difference Score method provides a score for a bridge, which indicates the condition of the bridge in comparison to the national average over its lifetime. We used the computed BDS of bridges to identify factors that had an effect the BDS by performing ANOVA. We performed both a small sample analysis and a large sample analysis. This allows us to watch for inflated p-values due to large sample sizes. Both small and large sample analysis provided a range for the effect size of the analysis. The necessary assumptions for ANOVA of a normally distributed population of scores is preserved in both small and large samples.

We can extend the use of BDS for other types of analysis. As discussed in Chapter 4: Results, the distribution of the BDS scores are near normal distribution. Using standard deviation, the set of all scores can be classified using Z-values as cut-offs. The new labels will allow bridge condition ratings to be further understood using classification and clustering.

The results of this study were computed based on all data available in the NBI database, i.e.17 million bridge inspections. We also demonstrated the use of integration of external open datasets, such as precipitation data in our analysis. In prior studies, the results were limited to data from a single state or a few states with a mix of open and restricted data. Their analysis pipeline is also not available for inspection.

5.5 Factors

From our literature review, we know that Precipitation, Average Daily Traffic (ADT), and Average Daily Truck Traffic (ADTT) are commonly believed to the determinants of the deterioration in bridges. ADTT is considered a stronger determinant than ADT in deterioration of the bridges, because the truck loading may have a serious impact on the condition of the deck,[22] and there is an observed positive correlation between structure length and average daily traffic.

But our results suggest that the material type of bridges has a higher association with the condition of the bridges, followed by precipitation and region. The maintainer has a higher association than average daily traffic and average daily truck traffic, while structure length has the least association with the condition of the bridges in comparison to all the factors analyzed in this study.

Upon analyzing the bridge material types further, our analysis revealed that Prestressed Concrete and Concrete bridges perform better than Wood or Timber, and Steel bridges. However, we suspect that data quality issues, such as a high number of bridges with no change in condition ratings in their life time, could introduce bias in our analysis. We hope to investigate this in our future work.

5.6 Other Limitations

The climate of any region is influenced by various factors such as latitude, elevation, topography, and prevailing winds. The cumulative effect of multiple variations in climate factors may also affect the condition of the bridges. We included geographic regions of the bridges as a factor in our test as a proxy for the cumulative effect of climate. However, we plan to perform additional tests at a more granular geographic level in our future work.

Our method includes many complex steps. Many of these steps are data-intensive and require a balanced representation of the data across several factors. For example, based on the discussion in Chapter 3, we know that after data cleaning and filtration, bridges from all states are not equally represented. We address the issue of imbalanced representation of the data by sub-sampling the over-represented states.

Finally, the baseline difference score method is new. We expect that it will take time before the method is widely adopted and further improved. Overall, the results of this study were validated and deemed useful by subject matter experts.

5.7 Summary

In summary, we discussed the limitations of our development environment, methods, and selected factors for analysis. In the next chapter, we outline our conclusions and explore potential future work.

Chapter 6

Conclusion and Future Work

In conclusion, this thesis:

- Provided a reference implementation of a big data pipeline for bridge health related datasets conforming to the standards described in the FHWA coding guide[18]
- Provided a method to compute and score bridge condition over a period of time using bridge inspection records.
- Provided methods based on sound statistical analysis to identify the level of association of certain factors such as construction material, precipitation and traffic with bridge conditions.
- Demonstrated the use of the Labs Workbench data analytics platform for sharing large datasets and a reproducible research environment.

6.1 Summary of Significant Findings

In this section, we provide a summary of the significant finding of our analysis of with respect to the research question.

1. How can data cleaning be generalized across NBI data submitted from different states?

- The cleaning and filtration strategy used in this research can be generalized to NBI records across all the states. The strategy is also traceable to the coding guide and cross-checks as provided by FHWA.
- After cleaning and filtration of the NBI dataset, we found that only 42% of the original dataset is suitable for analysis.
- A substantial number of bridges in the NBI dataset have condition ratings that do not change across all reported inspections.

2. How can data processing platforms for large bridge inspection datasets be shared among researchers and practitioners?

The Labs Workbench provide services for researchers and practitioners. They can share their working datasets and the environment. The Labs Workbench also provides a solution for long term preservation and archival of datasets and data analysis scripts. The Labs Workbench provides services for evaluating software, tools, and database management system.

3. How can bridge inspection related datasets be used to identify or eliminate factors that affect bridge conditions?

- The condition rating of substructure, superstructure, and deck are highly correlated. Hence, we selected the superstructure component of the bridge to measure the condition of bridges because superstructure is considered the backbone of the bridge.
- The normal distribution of the BDS of the bridges in the U.S. allows classifying bridges into the good, the bad, and the average with respect to the condition of the bridge.

- The association between all factors selected in this analysis and condition of the bridges varies from very small to small, which means that factors alone cannot explain the condition of the bridges.
- Material type has the highest association with the condition of bridges in comparison to the other factors selected in this study. The least associated factor to the condition of the bridges is average daily traffic.
- External factors such as Precipitation, have a small association with the condition of the bridges.

6.2 Future Work

In our attempt to address the research questions, we discovered several areas of this research for advancement that are related to data cleaning, data sharing platforms, and identifying factors that contribute to the deterioration of the bridges. The following are the suggested recommendation for future work.

Improvement in data quality will provide us with better bridge condition models of the bridges. Due to the missing values of condition rating of the bridges and other inconsistencies, only 42% of the bridge inspection records are used to conduct data analysis. Therefore, estimating the missing values will provide more data points for analysis.

In addition to missing data, inspections of bridges are biennial and the long intervals between inspections, prevent accurate characterization of the bridges. Further, inspectors conduct visual inspections of the bridges that are highly subjective. Data collection using IoT (Internet of Things) devices could provide reliable, frequent and objective data.

In our method, we compare bridges to national baseline of the U.S., which provides a perspective of the condition of a given bridge in comparison to the average National Bridge Condition. An alternative perspective to look at bridges might be from a bridge condition deterioration point of view. Understanding the patterns of decreasing condition ratings will guide in developing a better predictive model of the bridges. The two challenges in devising a method to compute deterioration would be:

- Accounting for intervention (Repair, Reconstruction, and Rehabilitation) of bridges that increase the condition ratings of the bridge.
- Considering only monotonically decreasing segments of the bridges, as the condition of the bridge tend to stay the same across many inspection records in the NBI.

Future work can extend the method of computing score by using a baseline of monotonically decreasing segments of the bridge to compute deterioration score of the bridges.

Another approach to consider for future work is to compute a national baseline that is specific for each material type. Then compare the performance of the bridges relative to their material-specific national baseline.

Rural bridges are in bad condition in several states.[25–27] A similar analysis of urban or rural bridges will help understand the reason behind their condition.

Future work can extend the analysis of external factors by integrating datasets such as heat and cold cycles, regional snowfall index and non-snowfall related regions. De-icing is also considered a factor that has the protective layers of the deck.

The following are links to various datasets that can be used for future works:

- Infobridge
- Weather Data Environment

- U.S. Climate Normals
- History of the U.S. Climate Divisional Dataset
- U.S. Climate Atlas
- Regional Snowfall Index

The classification of bridges according to BDS produces labels. These labels can be made use of with clustering and classification based machine learning algorithms in building a predictive model of bridge conditions.

In this study, we looked into three research question focused on data cleaning and transformation, analysis of commonly studied factors and unexplored factors as an indicator of bridge condition, and making the research reproducible.

We were able to clean and transform the NBI data extracted from FHWA and integrate external data sources such as CDC data. We tested commonly cited factors and unexplored factors as determinants of the bridge condition. We saw interesting and counter-intuitive results regarding Precipitation and Material Type. We also saw factors such as ADT and ADTT were not found highly associated with bridge condition in comparison to other factors.

Lastly, we made our research reproducible by adopting cloud-based work environment using the Labs Workbench. We developed our source code in python using Jupyter Notebook and MongoDB and we also made the source code available on GitHub.

[1]

Our vision for this research is to understand bridge performance using data-driven methods and identify factors that affect the condition of the bridges. In our future work we expect to collect real-time data of the bridge condition using

[1] Data Cleaning and Filtration, Data Analysis

IOT devices, and develop models that best predict future performance and needed maintenance to extend the bridge's service life.

References

[1] ASCE. 2017 infrastructure report card, 2017.

[2] Shlomo S Sawilowsky. "new effect size rules of thumb". *Journal of Modern Applied Statistical Methods*, 8, 2009.

[3] Melik Bolukbasi, Jamshid Mohammadi, and David Arditi. Estimating the future condition of highway bridge components using national bridge inventory data. *Practice Periodical on Structural Design and Construction*, 9(1):16–25, 2004.

[4] Massoud Nasrollahi and Glenn Washer. Estimating inspection intervals for bridges based on statistical analysis of national bridge inventory data. *Journal of Bridge Engineering*, 20(9):04014104, 2014.

[5] Tariq Usman Saeed, Yu Qiao, Sikai Chen, Konstantina Gkritza, and Samuel Labi. Methodology for probabilistic modeling of highway bridge infrastructure condition: Accounting for improvement effectiveness and incorporating random effects. *Journal of Infrastructure Systems*, 23(4):04017030, 2017.

[6] Yail J Kim and Dong Keun Yoon. Identifying critical sources of bridge deterioration in cold regions through the constructed bridges in north dakota. *Journal of Bridge Engineering*, 15(5):542–552, 2009.

[7] G Morcous, H Rivard, and AM Hanna. Modeling bridge deterioration using case-based reasoning. *Journal of Infrastructure Systems*, 8(3):86–95, 2002.

[8] STEVEN B Chase, EDGAR P Small, and CHRIS Nutakor. An in-depth analysis of the national bridge inventory database utilizing data mining, gis and advanced statistical methods. *Transportation Research Circular*, 498:1–17, 1999.

[9] Chang-Guk Sun, Sung-Ho Chun, Tae-Gyun Ha, Choong-Ki Chung, and Dong-Soo Kim. Development and application of a gis-based tool for earthquake-induced hazard prediction. *Computers and Geotechnics*, 35(3):436–449, 2008.

[10] Mucip Tapan and Riyad S. Aboutaha. Strength evaluation of deteriorated rc bridge columns. *Journal of Bridge Engineering*, 13(3):226–236, 2008.

[11] William T Scherer and Douglas M Glagola. Markovian models for bridge maintenance management. *Journal of Transportation Engineering*, 120(1):37–51, 1994.

[12] Mark A Cesare, Carlos Santamarina, Carl Turkstra, and Erik H Vanmarcke. Modeling bridge deterioration with markov chains. *Journal of Transportation Engineering*, 118(6):820–833, 1992.

[13] Paul D DeStefano and Dimitri A Grivas. Method for estimating transition probability in bridge deterioration models. *Journal of Infrastructure Systems*, 4(2):56–62, 1998.

[14] Ying-Hua Huang. Artificial neural network model of bridge deterioration. *Journal of Performance of Constructed Facilities*, 24(6):597–602, 2010.

[15] David Veshosky, Carl R Beidleman, Gerald W Buetow, and Muge Demir. Comparative analysis of bridge superstructure deterioration. *Journal of Structural Engineering*, 120(7):2123–2136, 1994.

[16] Samer Madanat, Rabi Mishalani, and Wan Hashim Wan Ibrahim. Estimation of infrastructure transition probabilities from condition rating data. *Journal of Infrastructure Systems*, 1(2):120–125, 1995.

[17] FHWA. Fhwa cross-check, 2017.

[18] Federal Highway Administration (FHWA). *FHWA Recording and Coding Guide for the Nation's Bridges, Office of Engineering , Bridge Division*, 1995.

[19] Jamie DeCoster. Converting effect sizes, July. 25, 2009 2009.

[20] George Morcous. *DEVELOPING DETERIORATION MODELS FOR NEBRASKA BRIDGES*. Nebraska Department of Roads, 2011.

[21] Pavel Jiříček and Marek Foglar. Numerical analysis of a bridge pier subjected to truck impact. *Structural Concrete*, 17(6):936–946, 2016.

[22] Hani Nassif Andrzej S. Nowak, Jeffery A. Laman. Effect of truck loading on bridges. Struct. Engrg. Rep. No. UMCE 94-22, University of Michigan at Ann Arbor, Ann Arbor, Michigan, 2009.

[23] GE Ramey and RL Wright. Assessing and enhancing the durability/longevity performances of highway bridges. *Final Report on Highway Research Record. Ctr. Proj*, 2:13506, 1994.

[24] Craig Willis, Mike Lambert, Kenton McHenry, and Christine Kirkpatrick. Container-based analysis environments for low-barrier access to research

data. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*, page 58. ACM, 2017.

[25] Associated Equipment Distributors. Ohio faces deficient roads and bridges, high rural fatality rates, and mounting congestion.

[26] Annette Dubas. The alarming condition of nebraska's rural bridges, December. 12 ,2014 2014.

[27] WHOTV. Research shows many rural iowa roads and bridges not in good condition, June. 27, 2017 2017.