A Spatial-Temporal Analysis of Railroad Development, Abandonment, and Historical Population Density

Krystal M. Saverse

University of Oklahoma

Abstract

The U.S. has a rich history in railroad development and operation. The American rail system contributed to the growth and transportation of resources during the American industrial revolution. However, many abandoned railroads exist today. This project aims to explore local indicators which might explain railway abandonment. The study area is narrowed to Boone County, Missouri. The data mostly consisted of historic census data, historic railroad data, and abandoned railroad data. A major limitation of this study were the lack of recent census data, not acquired in the early stages of this analysis. The majority of this project was completed using python scripting.

Keywords: Abandoned railroads, Missouri historical census, ArcGIS, python, Spatial-temporal analysis

Introduction

In order to install or operate within right-of-way, a company or public entity must apply for a permit to the appropriate authority. Right-of-ways can be thought of as legal land access for the purpose of improving movement. Many transportation modes have ROW, such as highways, county roads, railroads. ROW is typically held by the railroad owner or a local jurisdiction. But what happens when a railroad becomes abandoned? This presents a murky legal issue. In the 1830s, the U.S. government granted thousands of miles of right-of-way to railroad operators (Roberts, 2008). Subsequently, millions of acres of land were also granted for the construction of railroads (Roberts, 2008). However, many railroads have been abandoned since then. In recent years, the federal law has allowed abandoned right-of-way to be repurposed for recreational bike and hiking trails (Roberts, 2008).

The U.S. has a rich and complicated history in railroad development and operations. Railroads played a major part in the industrial revolution and the early and present transportation network. This project aims to answer if there are any local indicators which might explain railway abandonment using historical census data. The study area is narrowed to counties within Missouri.

Materials (GIS Data)

The search for available abandoned railroad data proved to be difficult. There are two known sites where abandoned rail lines can be viewed, www.abandonedrails.com, and www.frrandp.com/p/the-map.html. Both sites contain maps of abandoned railroads in the U.S. and much information about the railroad using google analytics or submitted information. They are mostly “crowd-sourced” meaning, individuals may submit the locations and information about the abandoned rails to the map owner to be updated. However, neither site contain datasets available to the public. Two datasets were found on ArcGIS Online from a contributor from the U.S. Surface Transportation Board's office (STB). The STB is an independent agency of the federal government who regulate surface transportation, mostly rail freight (Surface Transportation Board, n.d.). The datasets are titled "Railbanked rail lines" and "Abandoned rail lines". The "Railbanked rail lines" dataset includes out-of-service railroads to be used by a trail agency through a Notice of Interim Trail Use. This status must be applied for by the rail owner through STB. The "Railbanked rail lines" dataset includes railroads which have filed through the STB but not specifically designated for trail use. Both datasets include polyline layers of rail lines, the name of the railroad, and STB information such as the application number, the date the application was filed, and approved. It should be noted that these datasets do not contain all abandoned railroads, just ones that were filed through STB. Addition railroad data was acquired from the ArcGIS Hub. This dataset, "Historical Railroads Vanderbilt", is comprised of historical railroads in the U.S. as early as 1826. The dataset includes polyline features of rail lines, the name of the railroad, and its operational date.

In order to examine local variables which might explain railway abandonment, a thorough search for historical census data was done. This was also a very difficult task. The U.S. Census Bureau does not have GIS data for years prior to 1992. Additionally, the American Community Survey which compiles "long-form" census data containing detailed information about population and housing characteristics was established after 2000 (U.S. Census Bureau, 2017). And so, neither sites would contain historical census data. Census records prior to 1950 are maintained by the National Archives and Records Administration (Census History Staff, n.d.). However, searching through the National Archives and Records Administration is not intuitive and nor fruitful. Ultimately, Missouri census data was found online, provided by the Missouri Spatial Data Information Service (MSDIS). The datasets included shapefiles for the decades1810-1990, and included multiple demographic fields. Due to various recording practices over the years and the cultural implications in early American history, the attribute tables for the datasets are not uniform. Because of this, it was decided that only population values would be used for this study.

Finally, *tigerline* polygons were downloaded from the U.S. Census Bureau website. This included all 2020 U.S. state boundaries, and all 2020 U.S. County boundaries. These datasets would be used to narrow the scope of the study area.

Methods

In total, our datasets included historical census polygons for each decade, historical rail polylines, abandoned rail polylines, and state and county polygons. The first step, after downloading all of the applicable datasets, was to create a file geodatabase to store our project information. Two feature datasets were created, *CensusData* and *RailData*. This was performed in order to keep the spatial reference consistent between all datasets. The projected geographic coordinate system chosen for the feature datasets was *NAD 1983 StatePlane Missouri Central FIPS 2402 (US Feet)*. Once the feature datasets were created with the selected spatial reference, the data were imported to the applicable feature datasets. These steps were performed using Python script. Figure 1 shows the overall flow of the methods performed for this project using ArcGIS and python. Step 1 being the creation of the geospatial workspace.

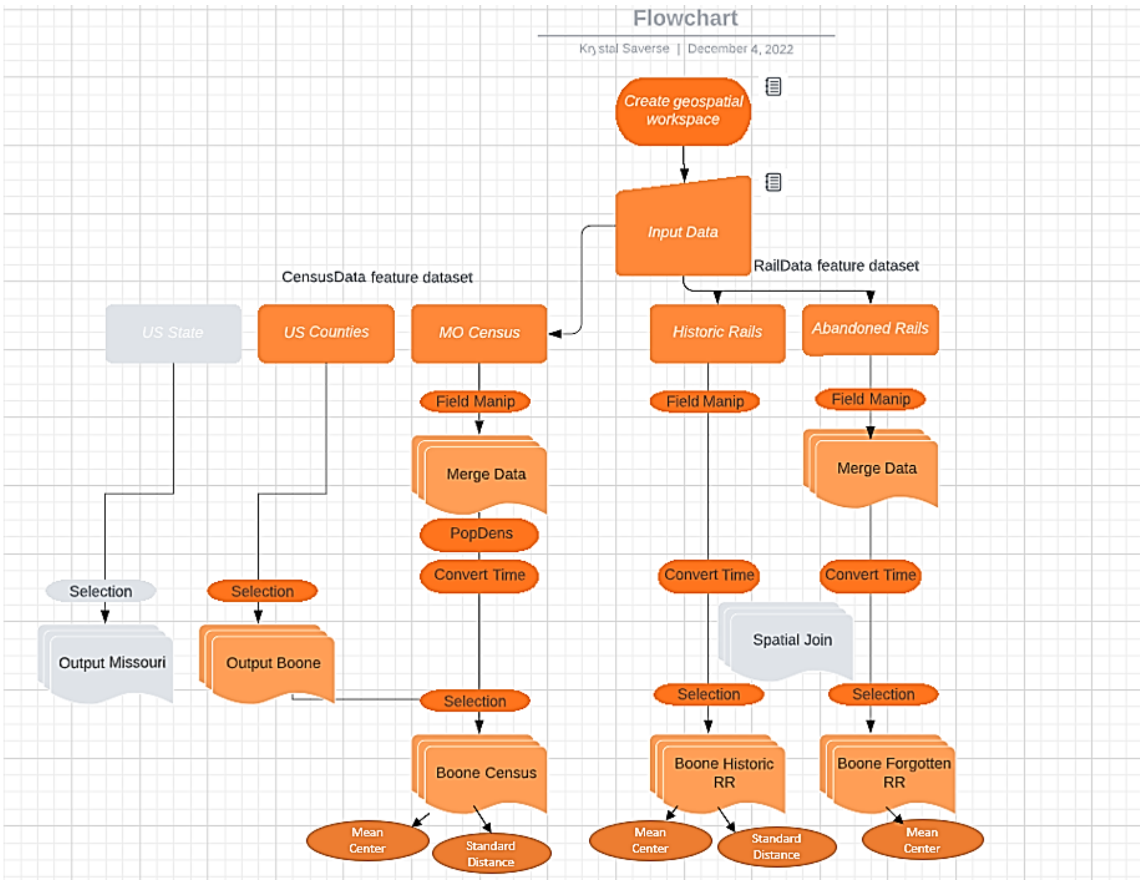

Figure 1. Flow chart of methods described. Items in grey were not necessary to calculate mean center and standard distance.

Next, fields were created and calculated for the datasets. There were 19 Missouri census polygons for the decades 1810-1990. However, there was no attribute in the tables to delineate the year the feature belonged to. A field, *YEAR* (TEXT datatype) was added to these datasets and calculated with the year information included in the file name. This step was performed using *os.walk()* and the *CalculateField* Data Management tool script. Similarly, there were 2 datasets representing railbanked/abandoned railroads. A field, *TYPE_ABAND* (TEXT datatype) was added to these datasets to delineate whether the railroad was railbanked or abandoned. This step also used the *CalculateField* script.

In order to use the datasets more efficiently, the Missouri census datasets were merged into a single output layer using arcpy. The *Merge* Data Management tool script was the best method for this step because the table fields were not uniform. The output for this step contained 1,826 features. Similarly, a merge was performed for the railbanked/abandoned datasets using arcpy. The output for this step contained 544 features.

Next, population density was calculated for the merged census data. The fields *AreaSqmi* and *POPDens* were added to the census dataset as FLOAT datatypes using the *AddField* Data Management tool script. *AreaSqMi* was calculated using the *CalculateGeometryAttribute* Data Management tool script. The script included optional parameters {*AREA}*, and {*SQUARE_MILES_US}*. Next, the *POPDens* field was calculated using fields indicating total population (*TOTPOP, TOTPOP30, TOTPOP40*) and the *AreaSQMi* field. This step's methods were intended to be completed using python script, however, due to limited time, the calculations for population density (total population/area square miles) was performed in ArcPro. This is because there were three different fields containing total population information. For future performance, a search cursor or code block expression should be used to calculate *POPDens* with information from multiple fields.

In order to view the map time series and perform spatial-temporal analysis, year and date information needed to be converted to a DATE datatype. First, the field *text_time* (TEXT data type) was added to the merged census dataset, the historic rail dataset, and the forgotten rail (abandoned/railbanked) dataset. Additionally, the field *YEAR* (TEXT data type) was added to the forgotten rail dataset. It was not necessary to add this field to the other datasets because the census dataset already contained the year per a previous step, and the historic rail dataset contained a field *InOpBy* containing a year value. These steps were performed using *AddField* python script. Next, the *text_time* field was calculated. This field needed to be a *YYYYMMDD* format. The year values were used to populate these datasets. The month and day values used were *0101* (January 01) since specific month and day values were not needed for this analysis. Subsequently, the *text_time* field was converted to a DATE datatype using the *ConvertTimeField* Data Management tool script.

A one-to-one spatial join was performed for the historic rail dataset and forgotten rail dataset using the *SpatialJoin* Analysis tool script. Forgotten rails were joined to historic rails (target feature) using a largest-overlap match option. This was performed in order to compare railroad operational date and abandonment dates.

Missouri census data and all railroad data were filtered using the *SelectLayerByLocation* and *SelectLayerByAttribute* Data Management tool scripts. First the study area Boone County and Missouri were extracted from the *tigerline* dataset using selection by attribute methods. Missouri census data were narrowed to features intersecting the Boone County boundary. Next, all railroad data were narrowed to features intersecting the selected Missouri Census Data. Selections were copied to feature classes using the *CopyFeatures* Data Management tool script.

Finally spatial-temporal analysis was completed by calculating the mean centers and standard distance for the datasets. These steps were performed using the *MeanCenter* and *Standard Distance* (Spatial Statistics) tool scripts. The inputs for the *MeanCenter* script included census, historic rail, and forgotten rail datasets (outputs from previous steps) for Boone County. Weighted fields used were *POPDens* and *Shape_Length*. Case fields used were *YEAR* and *InOpBy* (year). Similarly, these same variables were used for the *StandardDistance* script with the additional variable, Circle Size, set to *1_STANDARD_DEVIATION*. However, because of the vary small sample size for forgotten rails, the standard distance could not be performed for this dataset.

Results

The results from the spatial-temporal distributions are mean centers as point features, and spatial distributions as polygon features. Missouri Historical Census data were displayed using graduated colors (Natural Breaks) for population density. These features were grouped using year data. Figure 2 shows the final output for all datasets and the legend to be used for subsequent figures. The standard distribution has a very small grouping, which makes sense when considering its weighted values is population density. Population density is largest in Boon County and Cole County just South of Boone. This is because both counties consist of the cities Columbia which houses the University of Missouri, and Jefferson City which is the capital of Missouri. Surrounding counties also have a much less dense population density. From the census mean centers, we can theorize that the population density has shifted from northern areas of Boone County slightly to the South with the 1810 census being the exception. This is because Missouri’s census boundaries changed dramatically over the early 1800s. The rail operation mean centers and standard distribution appears to be extremely scattered. However, there seems to be a pattern for rail development to the states East and West boundary. This could be due to its’ proximity to the major cities Kansas City and St. Louis. Additionally, rail development appears to be centrally located, just Northwest of the present Boone County Boundary. This could be due to its proximity to Missouri River. It’s difficult to make any interpretation of the standard distance of rail operations from Figure 2. Rail abandonment mean centers also appear to be extremely disperse. However, there appears to be more abandoned rails to the state’s Western boundary, near Kansas City.

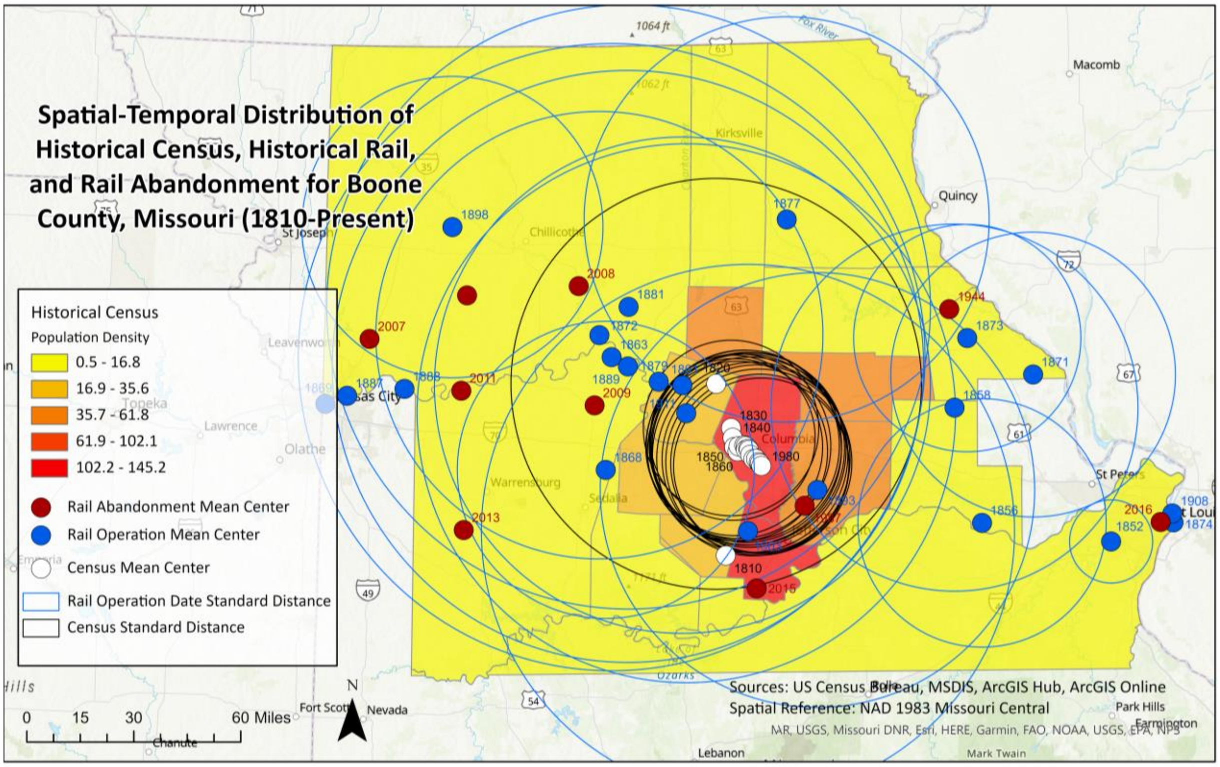

Figure 2. Spatial distribution, mean centers, and population density (1990) of all datasets.

The results from the timeseries analysis are easier to visualize. Once the Date datatype was calculated for the datasets, a timeseries (animated map) was able to be produced in ArcPro. This is not a geoprocessing tool, but a layer property setting. Time enabling allows the viewers to play back data at specified time intervals, or use the slider to go backward and forward thought the time series. Figure 3 shows the results from the enabled time series map in 10-year intervals (top left to right: 1850, 1870, 1880) (bottom left to right: 1900, 1940, 1980). The earliest developed railroads in the Boone census area were in operation around 1852. Railroads were developed a few years later in adjacent areas. The standard distance for developed rails in 1858 is much larger due to multiple rails being constructed during that time. Additionally, the standard distance for rails developed in 1856 is likely due to the large length of the rail. Over time, we see the distribution for rail development shift from East of Boone County to the West. This could be due to the expansion of the rail network over time. Additionally, there appears to be a pattern between rail development and population density. Rail development mean centers appear to be closer to areas with larger population densities. In regards to rail abandonment, a pattern to population density cannot be established. This is because rail abandonment and the acquired census data have no temporal overlap. This is because the census data is comprised of the years 1810-1990, and many railroads were classified as abandoned or railbanked occurred much later. This is the major downfall with this research. Future research should include census data after 1990.

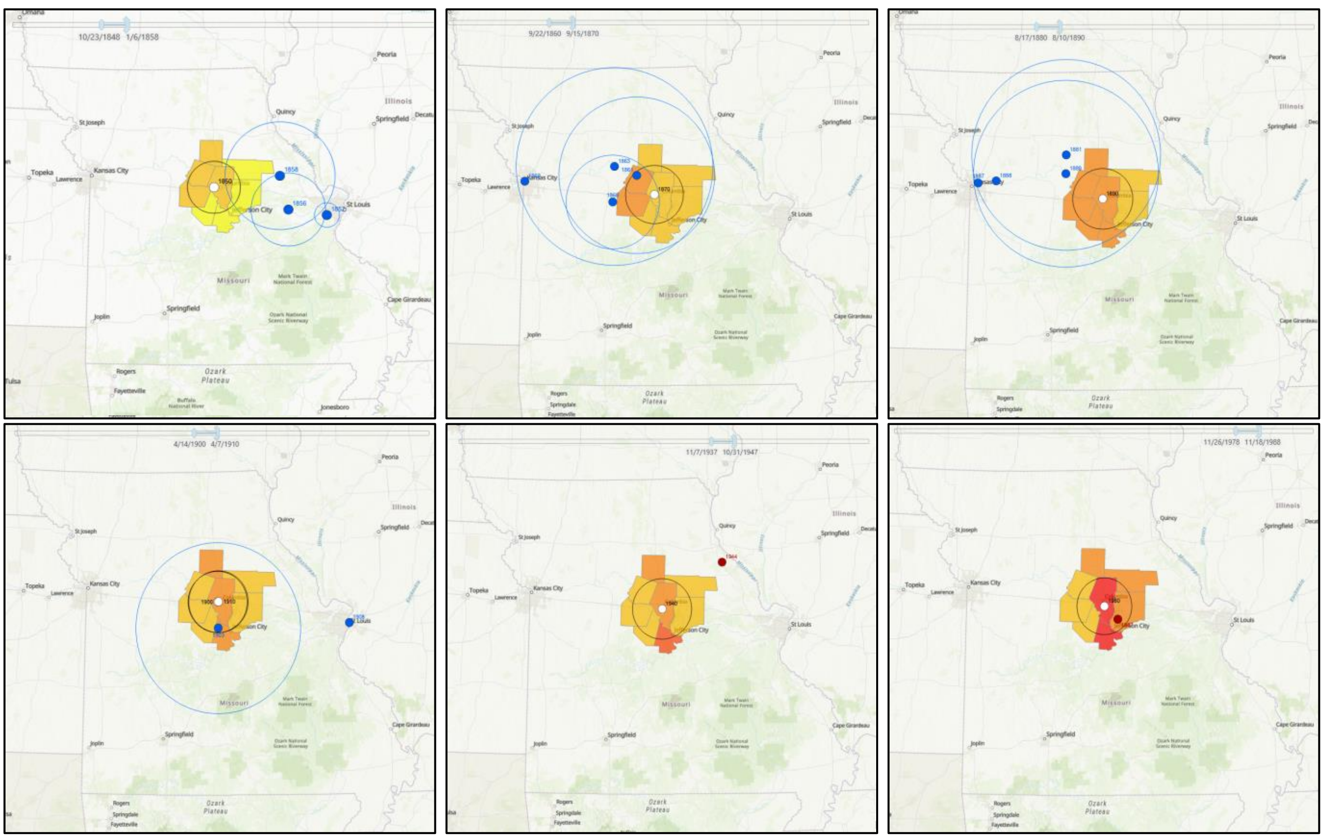

Figure 3. Time series of spatial distribution, mean center, and population density (top left to right: 1850, 1870, 1880) (bottom left to right: 1900, 1940, 1980). Features in blue represent

railroad operational date. Figures in dark red represent railroad abandonment date. Census data displayed are categorized population density.

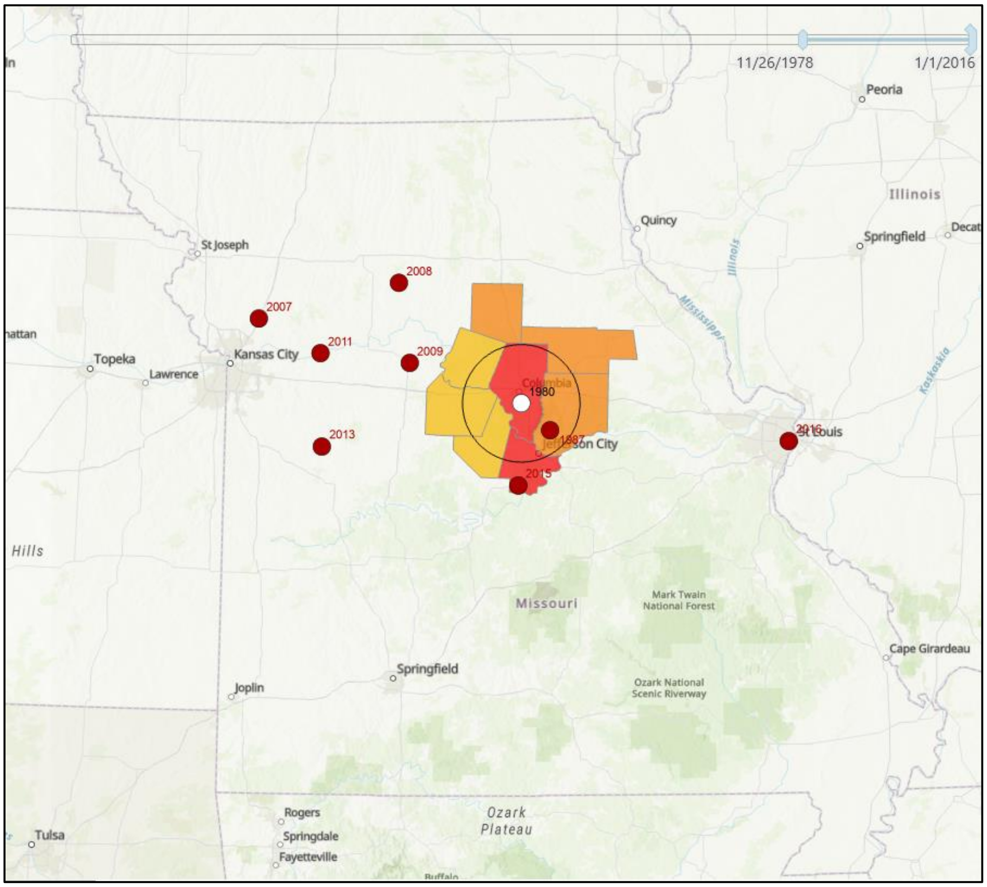

Figure 4. Mean centers of abandoned and railbanked rails (1978-2016) and 1990 population density.

Conclusion

This project ultimately set out to explore local indicators which might explain rail abandonment. Due to the limited census data, not much can be done to analyze local indicators, such as population density, to suggest any relationship. Additionally, abandoned rails were mostly recorded after the Surface Transportation Board (STB) was established. Also, not all abandoned railroads are recorded and consolidated, which is why crowd-sourced maps such as the ones described earlier are important. This project did fight a spatial pattern between railroad development and population density. This may seem intuitive, but actually the chronological accounts of railroad development and operation have very detailed historical accounts with many contributing factors ((Lamb, 2012). Mean centers and spatial distribution is a great way to display spatial change over time. It may be beneficial for future research to use these methods to explore other variables, such as economic variables or population migration. In order to further the results from this analysis, present day census data should be acquired (1990 and later). Additionally, if more detailed information on railroad maintenance decline becomes available, perhaps that data could be better suited for this analysis.

References

Census History Staff. (n.d.). *U.S. Census Bureau*. Retrieved from History: https://www.census.gov/history/www/genealogy/decennial_census_records/census_records_2.html

Lamb, P. J. (2012). A Macro-scale Look at Railroad History. *Railway and Locomotive Historical Society (R&LHS)*, 78-89.

Roberts, D. P. (2008). The legal history of federally granted railroad rights-of-way and the myth of congress's "1871 shift". *University of Colorado Law Review*, 113.

Surface Transportation Board. (n.d.). *About STB*. Retrieved from Surface Transportation Board: https://www.stb.gov/about-stb/

U.S. Census Bureau. (2017, December). *American Community Survey Information Guide.* Retrieved from U.S. Census Bureau: https://www.census.gov/programs-surveys/acs/library/information-guide.html