Automated data extraction from historical city directories: the rise and fall of mid-century gas stations in Providence, RI

Samuel Bell,[1]* Thomas Marlow,[2] Kai Wombacher,[1] Anina Hitt,[3] Neev Parikh,[3] Andras Zsom,[1] Scott Frickel[2]

[1]Advanced Research Computing, Center for Computation and Visualization, Brown University, Providence, Rhode Island, United States of America

[2]Institute at Brown for Environment and Society, Brown University, Providence, Rhode Island, United States of America

[3]Brown University, Providence, Rhode Island, United States of America

*Corresponding author

Email: samuel_bell@brown.edu

Abstract:

The location of defunct environmentally hazardous businesses like gas stations has many implications for modern American cities. To track down these locations, we present the *directoreadr* code (github.com/brown-ccv/directoreadr). Using scans of Polk city directories from Providence, RI, *directoreadr* extracts and parses business location data with a high degree of accuracy. The image processing pipeline ran without any human input for 94.4% of the pages we examined. For the remaining 5.6%, we processed them with some human input. Through hand-checking a sample of three years, we estimate that ~94.6% of historical gas stations are correctly identified and located, with historical street changes and non-standard address formats being the main drivers of errors. As an example use, we look at gas stations, finding that gas stations were most common early in the study period in 1936, beginning a sharp and steady decline around 1950. We are making the dataset produced by *directoreadr* publicly available. We hope it will be used to explore a range of important questions about socioeconomic patterns in Providence and cities like it during the transformations of the mid-1900s.

1. Background

Until the passage of the Resource Conservation and Recovery Act (RCRA) of 1976, waste produced during commercial and industrial activities in the United States was largely unregulated (1). It took another decade still until programs such as the Environmental Protection Agency's (EPA) Toxic Release Inventory (2), were established for keeping track of emissions from the largest and most hazardous facilities. These regulatory dynamics, combined with businesses tendencies to constantly churn in and out of operation over time, has created an urban environment covered with the relic sites and toxic legacies of past economic activity (3). This is a serious concern for both community members worried about their health (4), but also regulators and environmental professionals interested in locating and remediating contaminated sites. Thus, developing tools for the collection of data on the historical locations of businesses prone to producing waste represents an important contribution and is the goal of the current work.

Previous work focused on developing a software pipeline for processing historical directories specific to industrial manufacturing (5). The result was a software named *georeg* (https://github.com/brown-ccv/georeg), which was able to process digitized industrial directories to produce a near comprehensive dataset of industrial site locations and activities in Rhode Island for the years 1953-2012. This has been used productively for a range of scientific and community activities (6). However, while industrial production is a major source of urban pollution - it represents only a selection of economic activities that leave behind on-site contaminants. Gas stations are another such commercial activity of concern. According to the EPA, underground gas and oil storage tanks at these sites are a leading source of groundwater contamination (7). And while the federal government has been monitoring underground storage tanks (USTs) since the mid 1980s through RCRA, older USTs are added to lists only as they are discovered.

Therefore, in the current paper we develop an approach to collecting the historical location of commercial sites from city directories. Since the 1930s, the Polk Corporation has maintained detailed city directories for most American cities. Compiled annually, these books contain a comprehensive list of area businesses in the yellow pages.

Because the structure of the data in the city directories was considerably different from the industrial registries, we developed a new code, *directoreadr*, instead of adapting the *georeg* code of (5). We did use the same custom geocoder as (5), although the geocoding processing code was quite different.

To develop and test *directoreadr*, we have focused on city directories from Providence, Rhode Island. Using the scanned images of these directories, *directoreadr* is able to extract a company's name, address, and business type. The data are then geocoded to provide latitude and longitude. To show an example use of these data to examine environmentally hazardous sites, we focus on gas stations. However, the applications are not limited to tracking environmentally hazardous sites. These data can answer many important research questions across a range of topics that are of interest to many social science and environmental disciplines, from economics to ecology.

2. Data

For the purposes of this project, we have focused on the yellow pages business directory within the city directories. We examined 27 city directories from the city of Providence, RI with dates from 1936 to 1990. Beginning in 1940, the city directories were produced by the Polk corporation, but the three city directories from before 1940 were produced by the Sampson and Murdock corporation. By extracting this detailed spatio-temporal business data, we allow for socioenvironmental analysis of changes in the land use of industrial sites, manufacturing zones, or other potentially hazardous areas, such as current and former gas station sites.

Digitization was performed by the Internet Archive's office at the Boston Public Library, and the physical books were supplied by both the Boston Public Library and the Providence Public Library. The Internet Archive uses a standardized digitization process, delivering 300 dpi 8-bit color images with a lossy compression in the wavelet-based JPEG 2000 format. We convert these files to grayscale and do not use color information.

3 Methods

The *directoreadr* pipeline consists of a series of discrete processing steps that convert a series of page images into a database of businesses and locations. These steps are: grayscale thresholding, ad removal, margin cropping, column chopping, line chopping, Optical Character Recognition (OCR), header identification, entry concatenation, text cleaning, address parsing, street matching, and geocoding.

3.1 Image preprocessing

3.1.1 Grayscale thresholding

The original color images are read into *directoreadr* as 8-bit grayscale images with an integer pixel value ranging from 0 to 255, and the first step of the pipeline is to convert these images to a binary format, where each pixel value is either 0 or 1. This binarization step enables us to detect connected areas of black pixels, a core component of many of the computer vision algorithms we use. To do this, we use a fixed threshold across the entire page. Initially, *directoreadr* will attempt to estimate the threshold from the distribution of pixels at different pixel values. Incorrect grayscale thresholds are one of the largest sources of error in *directoreadr*, and its improvement would provide more accurate results. For this reason, we have chosen to allow a manual override of the grayscale threshold, and we recommend adjusting the grayscale threshold of pages with high error rates by hand.

3.1.2 Ad removal

After producing the binary images, we remove the advertisements along the border of the pages, as well as lines and decorations within and between the columns of text. For the sake of simplicity, we refer to all page features to be removed as "ads." To identify and separate out the ads from the text, we leverage two different geometric characteristics: First, the ads tend to be outlined by simple shapes like direct rectangles. Secondly, the ads tend to be much larger in extent than the characters of the text. Figure 1 shows an example of ad removal.

Figure 1: An example city directory page showing the ad removal process. The first panel shows the city directory page as a binary image. The second shows the contours identified as ads by the ad removal algorithm. The final panel shows the image after ad removal.

Using the OpenCV *contours* method, we identify regions of connected pixels. For each pixel contour, we calculate both the perimeter of the contour and the perimeter of the bounding box, the smallest possible horizontal rectangle circumscribing the contour. In most cases, ads can be separated from text simply by looking at the perimeter of the bounding box. However, in a few cases, when the grayscale threshold has been incorrectly set, many characters of text blur into each other, and the perimeter of the bounding box can be as large as it is for the smallest ads. To address this, we multiply the perimeter of the bounding box by the ratio between the perimeter of the bounding box and the perimeter:

$$p_{adjusted} = \frac{p_{bounding}^{2}}{p}.$$

Because text has a more complex shape than the ads, the ratio of bounded perimeter to contour perimeter is much lower for text than for ads, and it helps separate the text from the ads. Once we have identified the contours around the ads, we remove any black pixels within the bounding box around those contours.

In most cases, the ads around the edge of the page are surrounded by horizontal rectangles. We can address most cases that are not by identifying where the columns of ads are and removing all black pixels there. Even then, in a few cases ad removal fails, and the image has to be cropped by hand.

3.1.3 Margin cropping

Once the ads have been removed and replaced by whitespace, the columns of text in the center of the page are still surrounded by whitespace. To focus in on the text, the next step in the *directoreadr* pipeline is to remove the whitespace. To allow for specks, lines, and other noise on the page, we set a pixel threshold. Margin cropping is fairly straightforward and rarely creates problems.

3.2 Image segmentation

3.2.1 Column chopping

Each page is set up with columns of text (usually three columns), and in order to preserve information about text location, we separate the text into the columns. To identify the column breaks, we sum up the number of black pixels in each vertical line of pixels in the image. Around the column breaks, there are dips in the number of black pixels. To identify the location of the dips, we set a pixel threshold and identify the vertical lines with fewer black pixels than the threshold value. We cluster those vertical lines using the mean-shift machine learning algorithm. Unlike traditional clustering algorithms, like k-means, which take a number of clusters as an input, mean-shift figures out the optimal number of clusters. As the cut point for column separation, we pick the right-most vertical line in each cluster.

One of the key features of this algorithm is to err on the side of failure, throwing an error when the ad removal has performed poorly. The goal of this design is to allow for hand-chopping when it will meaningfully improve the results, and for all of the failure cases, we generated the columns through hand-chopping.

3.2.2 Line chopping

Once we have the columns of text, we then chop the columns into individual lines of text. To identify the lines of text, we use a similar process to identifying the columns. We calculate the number of pixels in each horizontal line of pixels in the column. Then, we cluster the horizontal lines of pixels that fall below the pixel threshold, using mean-shift to identify the entry breaks. If there are large blocks of entries that don't separate, we then run the algorithm on them with a higher black pixel threshold. This higher threshold is typically necessary when the page image is warped or tilted. This process is highly robust, and it rarely produces errors unless there are more serious problems with the image.

3.3 Address processing

3.3.1 OCR

Entering this part of the pipeline, we have a directory of images where each image represents a single line from one of the columns on the page. To convert these images of text to a string of text, we use the Tesseract OCR package developed by the Google corporation (8). OCR is not perfect, and it does produce some errors, and downstream text parsing parts of *directoreadr* must account for these errors.

3.3.2 Header determination

The data in the city directories are grouped under headers that describe the type of business, and these headers must be identified. Depending on the year, the city directories identify headers using a number of different characteristics. Headers are typically indented, and they sometimes contain all caps. Often, headers have asterisks before them. Depending on the year, *directoreadr* selects from five different header determination algorithms, most of which center around how many pixels each line is indented by. Because some columns are tilted, we calculate a relative indentation compared to nearby lines. Our header detection algorithms relied on indentation and capitalization as the primary detection features, and we did not build a robust header algorithm for 1964, the one year in which headers were not indented, and both the headers and the text were in all caps. As a result, most header identifications for 1964 are incorrect.

3.3.3 Entry concatenation

In many cases, entries in the columns of text are too long for one line and continue onto the next line. In all of these cases, the next line is indented, but not by as much as a header is. Using the indentation data, we concatenate the multi-line entries into single strings of text.

3.3.4 Text cleaning

Most of the raw entries just contain a business name and an address, but some of them contain additional information that must be removed, like a telephone number or a floor or room number. To clean these data, we used a complex series of pattern matching operations. In some cases in older books, there were multiple addresses for a business in a single entry, and we split these lists based on the positions of commas and the word “and.”

3.3.5 Address parsing

We start by using a regex to search the entry’s string of text for abbreviations like:St., Av., Ct., Dr., Rd., Ave., and Ln. in either upper or lower case. If one of those abbreviations is detected in the string, the algorithm searches for a group of digits before the abbreviation. It then classifies the string of text between the number and abbreviation the address and the text before the address number is classified as the company name. If the abbreviation is not detected, the algorithm will still try to parse out the address by searching for the address number and classifying the string of text after the number as the address, and string proceeding the number as the company name.

This parsing algorithm is not perfect – i.e. it requires digits (not spelled out numbers) for the address. However, it is generalized enough to work well across many different formats because it is built on simple components of the address that are consistently present.

3.3.6 Street matching

Because a number of the streets contained OCR errors, we used fuzzy matching to produce true street names. We developed two lists of streets, a list of current streets and a list of historical streets. The historical street list was developed through hand examination of historical maps and is not fully comprehensive. Because we only had a database of Providence streets, we removed the addresses we could identify as belonging to another Rhode Island municipality.

Using the *fuzzywuzzy* package in Python, we created a scoring algorithm to quantify how close an OCR reading of a street name is to a street in the true street name list. This scoring algorithm is based off the Levenshtein distance ratio:

$$ratio(s1,s2) = 100 * \left(1 - \frac{D(s1,s2)}{L(s1) + L(s2)}\right),$$

where *s1* and *s2* are the two strings being compared, *L* is a function giving the length of a string and *D* is a function giving the Levenshtein distance between two strings. The Levenshtein distance is the minimum number of edit operations (substitutions, deletions, or additions) required to convert one string into another. For instance, the

Levenshtein distance between “park” and “barks” is 2, one substitution and one addition. The ratio would be 100*(1-2/9) = 77.8%.

$$score = \frac{ratio(s1,s2) + ratio(LongestWord(s1),LongestWord(s2))}{2},$$

where *LongestWord* is a function that gives the longest word of a string. The reason for adding additional emphasis on the longest word was to emphasize the core street name. For instance, we wanted “BROADWAY” to match with “BROADWAY ST.”

For each street in the true streets list, we calculated the matching score between the OCR result and the street in the true streets list, selecting the true street with the highest score. To guard against false positives, we removed any matches with a score below 80.

To improve the efficiency of the street searching, we only searched for unique OCR results and built up a dictionary of OCR results and their cached street search results.

3.3.7 Geocoding

The last component of the *directoreadr* pipeline is geocoding the cleaned and parsed addresses to obtain the latitude and longitude coordinates of the businesses. After researching several different geocoding options, from paid services (SmartyStreets) to free APIs (Google Maps), we decided to implement the geocoder built for (5) using ArcGIS software, with data from Rhode Island's E911 database. This geocoder was free for our use, given that Brown University had an in-house ArcGIS server, but because geocoders are proprietary, in our published version of the code, we do not include the api key necessary to run the geocoder.

To improve the speed of the geocoding, we ran 50 concurrent searches and searched only for unique addresses, building up a dictionary of geocoder results to reference in future runs of the program. Because many addresses were repeated across many years, this drastically sped up the process.

The geocoder only contained data on current street layouts. Providence, however, like many American cities, has seen considerable change in its street pattern over the course of the study period. Many streets have been wholly or partially demolished, and others have been renumbered. To address this problem, we utilized the geocoder confidence score, and we removed any addresses with a confidence score under a perfect score of 100 from our final results. In a case where the entire street was no longer remaining, the geocoder returned an error. To address the most common of these addresses, we allowed for hardcoding of hand-identified historical geocodes, entering hard-coded locations for four large buildings with many businesses at those addresses.

4. Results and discussion

The image processing portion of the pipeline had a success rate of 94.4%. In our dataset, we ran the algorithm on 2,582 individual pages. For these pages, 144 or 5.6% required hand-chopping in order to process. We designed the column-chopping algorithm to deliberately fail when there were likely errors with the ad removal algorithm. The goal was to require hand-chopping whenever it would meaningfully improve the end result. Because of the hand-chopping, we were able to pass all of the pages through to the OCR and text parsing algorithms.

In the text parsing algorithm, 6.7% of all entries were dropped as not a successfully identified and matched address. These include both entries that should be dropped and entries that were dropped because of an error. In 38.2% of these cases, the algorithm failed to parse an address at all. In 10.3% of these cases, the algorithm parsed an address but returned an empty string for the street. In 4.2% of these cases, the algorithm parsed an address but threw an error in street matching. And in 47.2% of these cases, the algorithm successfully parsed an address and matched a street, but the confidence score was too low for us to be sure the address was correct. In some of the address drop cases, the addresses were outside of Providence, sometimes outside of Rhode Island. Others reflected an idiosyncratic address form the algorithm wasn't set up to parse. For instance, some addresses were named buildings without an address (e.g. "Arcade Bldg" or "Industrial Trust Bldg"). Others were street corners instead of numerical addresses. Of course, many of the address drop cases represented failures of the OCR or failures of the header identification, concatenation, and entry chopping algorithms. Address drop rates were not strongly correlated with time (Figure 2).

Figure 2: Address drop and geocoder error rates by year.

The geocoder algorithm produced errors in 4.7% of cases. Unsurprisingly, these errors were higher in earlier years when the Providence street pattern was considerably different (Figure 2). Towards the end of the study period, the percentage of addresses outside of Providence increased sharply. While we were able to capture most of these, we were not successful in all cases. Many addresses from a different city were not recognized as belonging to a different city, and when they were processed, they led to dropped addresses or geocoder errors.

These statistics only capture the places where the code generated errors or flags. In order to fully assess the ultimate accuracy of the code, we hand-examined the error rate for gas stations in three years: 1936, 1962, and 1990. In 1936, 220 out of 242 gas stations were correctly identified, for an accuracy rate of 90.9%. Of the 22 missing gas stations, eleven were missing due to geocoder errors cause by historical street changes, seven were missing because of non-standard address formats that *directoreadr* could not parse correctly, one had the wrong address read, and only three were entirely missing. In 1962, 219 out of 224 gas stations were correctly identified, for an accuracy rate of 97.8%. Of the five missing gas stations, one was dropped, one had a geocoder error, and three had their addresses read incorrectly. In 1990, 71 out of 73 gas stations were correctly identified, for an accuracy rate or 97.3%. Both of the missing gas stations were dropped. These statistics do not include errors in correctly reading and parsing the business names.

The total accuracy rate from the gas stations in the three years we examined by hand was 94.6%. (This number is somewhat skewed because 1936 had the most gas stations. The average of the three accuracy rates was 95.3%.)

We are making the data available for download from the Brown Digital Repository (https://doi.org/10.26300/typ4-nj27), and we are making the *directoreadr* code available at github.com/brown-ccv/directoreadr.

4.1 Efficiency

Once the geocoding, street matching, and OCR results have been cached, the parsing algorithm runs in roughly 20s per book on a standard laptop, enabling faster debugging and development. With no cached results, the full *directoreadr* pipeline still runs in under 30 minutes per book. Before our efficiency improvements, *directoreadr* would take many hours to process a single book.

4.3 Example hazardous site: gas stations

Because of their environmental importance, we selected gas stations as an example hazardous site type. Gas stations in Providence typically developed along main roads, avoiding wealthier neighborhoods like Providence's East Side (Figure 3).

Figure 3: Gas stations mapped in Providence from 1940 to 1990

Figure 4: Total number of gas stations recorded in Providence by directoreadr

Starting in the 1950s, gas stations began a precipitous decline in Providence (Figures 3 and 4). By 1990, there were only 75 gas stations in the city, a decline of 71% since 1950, when city directories list 257. This drop corresponds with a decline in the city's population, which dropped by a third between 1950 and 1980, the combined result of job loss from deindustrialization and displacement of minority residents whose neighborhoods were cleared for several ambitious "urban renewal" projects (9–13). These changes were part of broader national trends of suburbanization and economic decline in the urban core (14–16). Other factors specific to the service station and automobile industries also may have played a role. Broader regulatory changes likely also affected gas station counts, with zoning having a particularly important effect (17–19). Because the rate of geocoder errors was higher in the earlier years, these figures probably underestimate the dramatic drop in the number of gas stations. Overall, we identified 526 unique gas station addresses in Providence over the study period, compared to just 114 gas station addresses recorded in the Rhode Island Department of Environmental Management Underground Storage Tank (UST) database (http://www.dem.ri.gov/programs/wastemanagement/inventories.php).

5. Conclusions

We have successfully built a pipeline for the digitization, extraction, and processing of city directory data. While this approach was developed on directories from Providence, RI, these directory formats are fairly similar in different cities, and this approach should be adaptable to cities all across the country. There are many potential uses of these data, and we have demonstrated mapping of environmentally hazardous historical gas station sites as an example.

Acknowledgements

We acknowledge support from the Institute at Brown for Environment and Society, which funds a Research Assistantship for T.M., and from the Superfund Research Program of the NIEHS grant 2P42 ES013660. This work has also benefited from seed grants from the Brown University Office of the Vice President for Research (grant GR300065) and the Brown Social Sciences Research Institute. We would like to thank the Providence Public Library and the Boston Public Library for providing many of the physical directories for scanning, as well as the Internet Archive for doing the scanning. We would like to specially thank Kate Wells of the Providence Public Library for her invaluable advice and assistance in arranging the scanning.

Bibliography

1. Colton CE, Skinner PN. The road to love canal: managing industrial waste before EPA. Austin: U of Texas P. 1996;

2. Environmental Protection Agency US. U.S. EPA 1988 Toxics Release Inventory (TRI) Program. US EPA; 1988.

3. Frickel S, Elliott JR. Sites unseen: uncovering hidden hazards in american cities. Russell Sage Foundation; 2018.

4. Brown P. Toxic exposures: contested illnesses and the environmental health movement. New York Chichester, West Sussex: Columbia University Press; 2007.

5. Berenbaum D, Deighan D, Marlow T, Lee A, Frickel S, Howison M. Mining Spatio-temporal Data on Industrialization from Historical Registries. J ENV INFORM. 2018;

6. Guelfo JL, Marlow T, Klein DM, Savitz DA, Frickel S, Crimi M, et al. Evaluation and Management Strategies for Per- and Polyfluoroalkyl Substances (PFASs) in Drinking Water Aquifers: Perspectives from Impacted U.S. Northeast Communities. Environ Health Perspect. 2018 Jun 15;126(6):065001.

7. Environmental Protection Agency US. National Water Quality Inventory: 2000 Report. US EPA; 2002.

8. Smith R. An overview of the tesseract OCR engine. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2. IEEE; 2007. p. 629–33.

9. Antonucci C. Machine Politics and Urban Renewal in Providence, Rhode Island: The Era of Mayor Joseph A. Doorley, Jr., 1965-74. 2012;

10. Goldstein S, Mayer KB. Metropolitanization and population change in Rhode Island.

Metropolitanization and population change in Rhode Island. 1961;(3).

11. Goldstein S, Mayer K. Population decline and the social and demographic structure of an american city. Am Sociol Rev. 1964 Feb;29(1):48.
12. Zimmer BG, Hawley AH. Suburbanization and some of its consequences. Land Econ. 1961 Feb;37(1):88.
13. Zimmer BG. Rebuilding Cities: The Effects of Displacement and Relocation on Small Business. Providence, RI: Brown University; 1964.
14. Jackson KT. Crabgrass frontier: The suburbanization of the United States. Oxford University Press; 1987.
15. Sugrue TJ. The Origins of the Urban Crisis: Race and Inequality in Postwar Detroit-Updated Edition. Princeton University Press; 2014.
16. Wilson WJ. When work disappears: The world of the new urban poor. Vintage; 2011.
17. Campbell HE, Kim Y, Eckerd A. Local zoning and environmental justice. Urban Affairs Review. 2014 Jul;50(4):521–52.
18. Campbell HE. Rethinking Environmental Justice in Sustainable Cities: Insights from Agent-Based Modeling. Routledge; 2015.
19. Shertzer A, Twinam T, Walsh RP. Zoning and the economic geography of cities. J Urban Econ. 2018 May;105:20–39.

Daughters of St George—Perseverance Lodge 246 Weybosset
Daughters of the American Revolution 209 Williams
Gaspee Chapter Mrs T Frederick Chase regent 209 Point
Foresters of America—Grand Court
Fraternal Order of Eagles—Prov Aerie No 22, 631 Westminster
German Society—Harugari Lodge No 4, O D H S 24 Hartford av
Grand Army of the Republic—
Prescott Post No 1, 176 Benefit
Ladies Auxiliary Logan Circle 59 Chestnut
Improved Order of Red Men—
Wauola Council No 1, 246 Weybosset
Independent Order of Foresters—
Court Rochambeau 1029 Westminster
Companion Court Harriet No 726, 1029 Westminster
Irish National Foresters—
General Phil Sheridan No 451, 767 Westminster
Lady Sheridan No 1023, 767 Westminster
Italian Society—Basilicata Society 2036 Atwells av
Junior Order United American Mechanics—
Eagle No 5, 48 Snow
Rhode Island No 7, 380 Elmwood av
Washington Council No 2, 729 Westminster
Knights of Columbus 24 Greene—
Bishop Hendricken Assembly 4th Degree
Elmhurst Council No 31
Hope Council No 209
Tyler Council No 62
Providence Council No 95
Pleasant J A Council 220 Veazie
Knights of Pythias—
El Karnak No 45, 729 Westminster
Palestine Lodge No 2, 246 Weybosset
Elmwood Lodge No 38, 380 Elmwood av
Mystic Lodge No 3, 59 Chestnut
Red Cross Lodge No 35, 729 Westminster
Grand Lodge 48 Westminster rm 808 [1012]
What Cheer Lodge No 24, 59 Chestnut
Loyal Orange Institution Lodge No 109 and Lodge No 223, 59 Chestnut
Maccabees The 729 Westminster
Masonic Bodies—Freemasons' Hall 127 Dorrance Grand Lodge—
St Johns Lodge No 1
Mount Vernon Lodge No 4
What Cheer Lodge No 21
Corinthian Lodge No 27
Adelphoi Lodge No 33
Redwood Lodge No 35
Orpheus Lodge No 36
Nestell Lodge No 37
Overseas Lodge No 40
Roosevelt Lodge No 42
Thomas Smith Webb Lodge No 42
Thomas Smith Webb Chapter No 16
Grand Royal Arch Chapter of R I
Providence Royal Arch Chapter No 1
Providence Council R & S M No 1
Grand Chapter of R I, Order of the Eastern Star
Grand Council R & S M of R I
Calvary Commandery No 13
St Johns Commandery

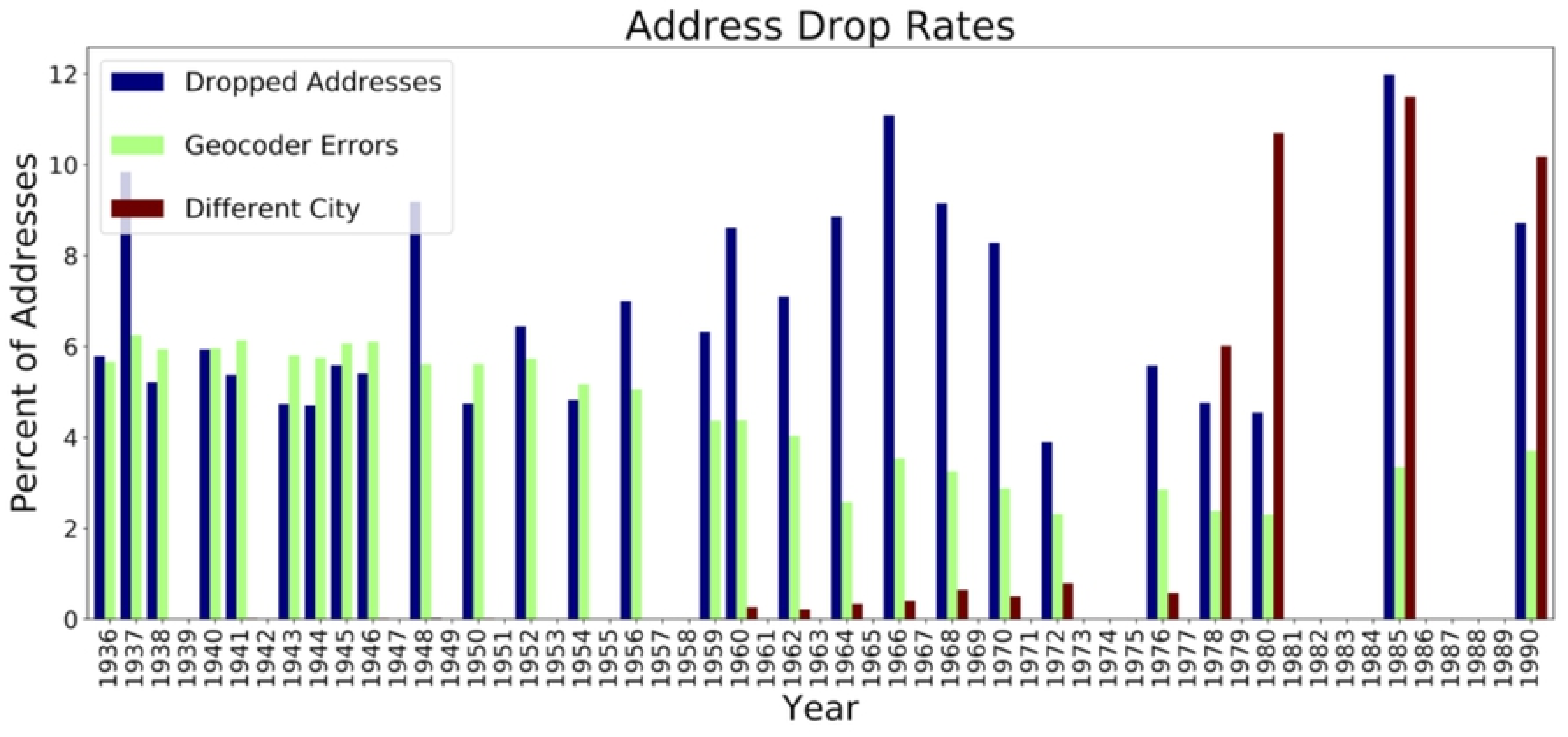
Thomas Smith Webb Commandery K T No 12
Rhode Island Consistory
Palestine Temple A A O N M S 40 Abore
Veterans Association of R I
Providence Chapter Order of the Eastern Star
Providence Chapter Order of De Molay
Ancient Accepted Scottish Rite
Modern Woodmen of America—
Royal Oak Camp No 7656, 830 Westminster
New England Order of Protection—
Vanguard Lodge No 27, 268 Weybosset
Westminster Lodge No 76, 268 Weybosset
Odd Fellows—
Grand Encampment of R I office 86 Weybosset rm 7
Canton Providence 276 Benefit
Eagle Lodge No 2, 48 Snow
Unity Lodge No 25, 59 Chestnut
Roger Williams Lodge No 3, 48 Snow
Narragansett Encampment No 1, 48 Snow
Massasoit Encampment No 12, 169 Potter's av
Canonicus Lodge No 3, 48 Snow
Past Grand M U 59 Chestnut
Manufacturers Lodge No 15, 279 Plainfield
Swarts Lodge No 18, 519 Westminster
North Star Lodge No 25, 36 Chalkstone av
Massasoit Encampment No 2, 48 Snow
Narragansett Encampment No 1, 48 Snow
James Wood Lodge No 30, 48 Snow
Westminster Lodge No 27, 48 Snow
What Cheer Lodge No 48, 558 Westminster
Rebekah Lodges—
Dorcas Lodge No 7, 269 Potter's av
Naomi Lodge No 1, 246 Weybosset
Arbor Lodge No 25, 519 Westminster
Priscilla Lodge No 30, 279 Plainfield
Rose Standish Lodge No 5, 48 Snow
Ruth Lodge No 8, 48 Snow
Manchester Unity—
Loyal Lily Lodge No 6075, 220 Veazie
Loyal Bud of Hope Lodge No 6517, 279 Plainfield
Odd Ladies Loyal Hope Lodge 246 Weybosset
Odd Fellows (Colored)—
Hope Lodge 315 Cranston
Westerner Lodge No 1934, 310 Cranston
Westminster Lodge No 2408, 310 Cranston
Providence Patriarchie, 310 Cranston
Rebekah Lodge—
Household of Ruth No 42, 310 Cranston
Household of Ruth No 1028, 310 Cranston
Order of Owls Nest 106, 71 Richmond
Order of Scottish Clans—
Clan Cameron No 7, 246 Weybosset
Daughters of Argyle No 1, 246 Weybosset
Providence Caledonian Society, 246
Order Sons of St George Pacheck Lodge No 194, 279 Plainfield

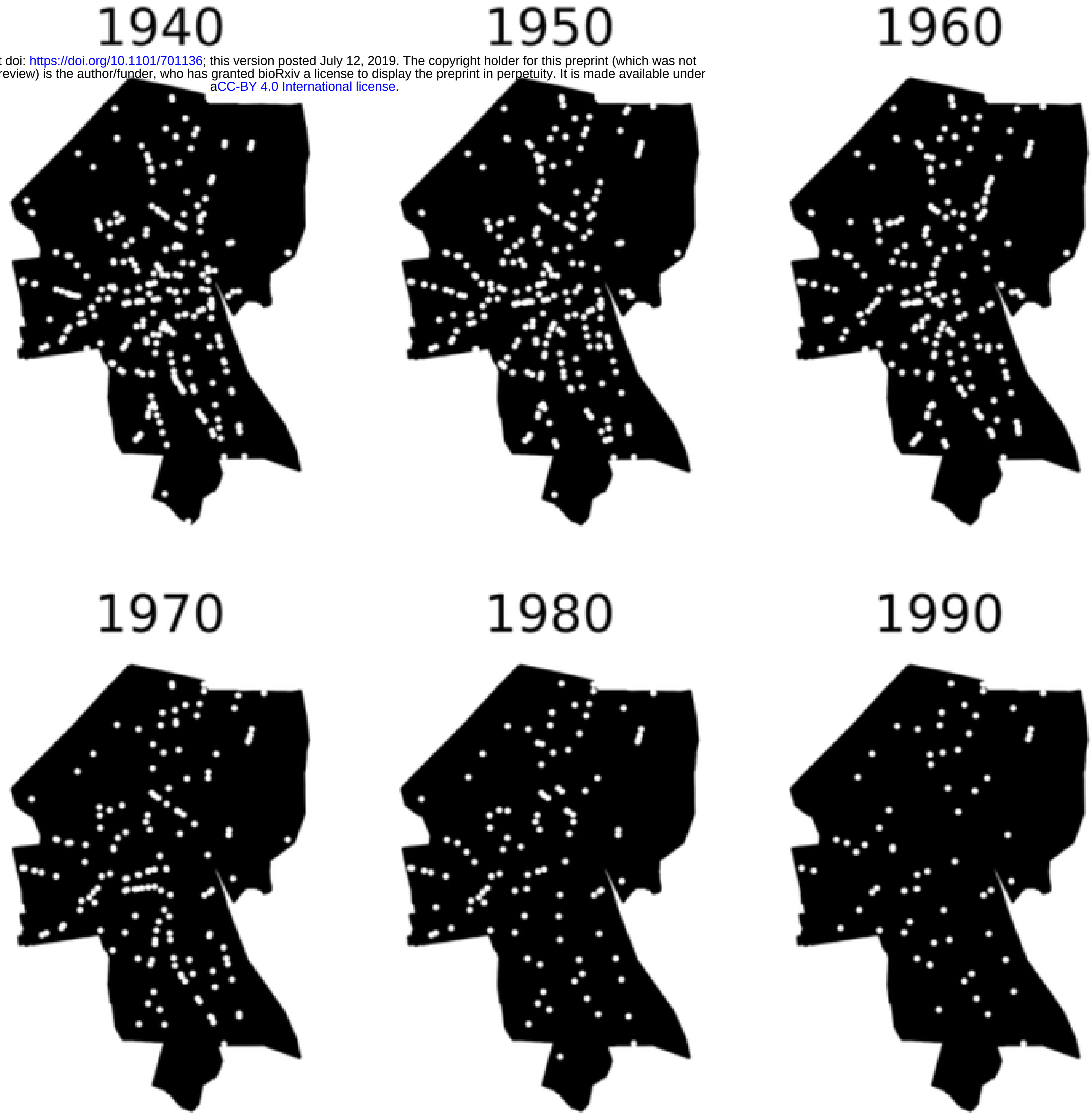
Order United American Mechanics—
Columbia Council No 21, 48 Snow
Providence Hebrew Sheltering Society 98 Jefferson
Provident Fraternity—Supreme Lodge 141 Weybosset rm 322
Pythian Sisters—
Auburn Temple 729 Westminster
Jennie Lind Temple 59 Chestnut
Royal Arcanum—
Chappoquaset Council No 1019, 380 Elmwood av
Mohamock Council 1172, 248 Weybosset
Royal Neighbors of America—
Evergreen Camp 729 Westminster
Free Camp 729 Westminster
Sons of Norway 59 Chestnut
Sons of Union Veterans of Civil War—
R I Department 47 Wash rm 4
Elisha H Rhodes No 11, 59 Chestnut
Sons & Daughters of Liberty—
Betsey Williams Council No 1, 729 Westminster
Betsey Ross Council No 25, 59 Chestnut
Lady Lincoln Council No 5, 246 Weybosset
United Spanish War Veterans—
Lieut Commander Gardiner C Sims Camp No 10, Benefit cor Meeting
Allyn K Capron Camp No 1, Benefit cor Meeting
Sidney F Starr Camp, Benefit cor Meeting
Auxiliary of Allyn K Capron Camp 176 Benefit
Providence Auxiliary No 4 of Sidney F Starr Camp 176 Benefit
Volunteer Service Club 45 Arcade Bldg
Woman's Relief Corps—Prescott No 24, 176 Benefit
Woodmen of the World—Royal Oak Camp No 1, 48 Snow

Miscellaneous

Alumni Associated of B U 75 Waterman
Alumni Assn of R I College of Pharmacy and Allied Science 235 Benefit
American Legion 100 Fountain rm 208
Antonio Mangione Post 54, 335 Dean
Melvin Hobson Post 626 Plainfield
Stafford Post No 3, 1 Burnside
Roger Williams Post No 21, 292 Smith
Thos L Ryan Post No 47, 138 Academy av
Y D 26th Div 176 Benefit
American Red Cross 100 N Main rm 20
Anti-Saloon League of R I 29 Weybosset rm 43
Barbers' Union Local No 224, 309 Wash rm 207
Bartenders Union Local 285, 281 Weybosset
Benito Mussolini Society 332 Dean
Bethany Home of R I 111 S Angell
Bible Readers Band of America 258 Hanover
Boy Scouts of America, Narragansett Council 100 N Main rm 24
British Empire Club 33 Westminster rm 808
Brotherhood of Railway Clerks, Free Lodge 47 Wash rm 12
Brotherhood of Utility Workers of NE 42 Weybosset rm 26
Carpenters Union Local 195, 281 Weybosset

PELLETT MONUMENTS


Address Drop Rates
Dropped Addresses
Geocoder Errors
Different City
Percent of Addresses
Year
0
2
4
6
8
10
12
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990


1940
1950
1960
1970
1980
1990