Team Data Wiz

Sponsored by

Washington State Library

Aayushi Gandhi, Manthan Mehta, Harshi Thaker

Washington Digital Newspaper Program

Digitizing WSL historic newspaper collections

Current State of WDN

– Unstructured data

– Chance of overlooking information

Newspapers available from King County

The Catholic Northwest Progress (Seattle — 21 December 1900 - 27 June 2013)
Cayton's Monthly (Seattle — 1 February 1921 - 1 March 1921)
Cayton's Weekly (Seattle — 14 July 1917 - 29 January 1921)
The Daily Intelligencer (Seattle — 5 June 1876 - 1 October 1881)
The Daily Republican (Seattle — 26 February 1896 - 2 March 1896)
Đất Mới = New Land (Seattle — 1 August 1975 - 1 January 1987)
The Enterprise (Seattle — 14 May 1921 - 4 April 1952)
Filipino Forum (Seattle — 15 October 1928 - 10 December 1963)
Monthly News Of The Department Of Washington (Seattle — 1 March 1943 - 1 June 1946)
The Northwest Times (Seattle — 1 January 1947 - 23 March 1955)
People's Telegram (Seattle — 3 November 1864 - 21 November 1864)
Puget Sound Dispatch (Seattle — 4 December 1871 - 4 October 1880)
Puget Sound Semi-weekly (Seattle — 5 April 1866 - 19 April 1866)
Puget Sound Weekly (Seattle — 30 April 1866 - 18 March 1867)
Puget Sound Weekly Gazette (Seattle — 25 March 1867 - 17 June 1867)
The Ranch (Seattle — 1 September 1902 - 1 June 1914)
The Republican (Seattle — 22 August 1896 - 17 June 1898)
Scandinavian American (Seattle — 1 January 1945 - 1 December 1956)
Seattle Daily Post-intelligencer (Seattle — 5 April 1882 - 10 May 1888)
Seattle Gay News (Seattle — 1 March 1977 - 28 December 2018)
The Seattle Gazette (Seattle — 10 December 1863 - 4 June 1864)
The Seattle Post-intelligencer (Seattle — 11 May 1888 - 31 December 1900)
The Seattle Republican (Seattle — 23 February 1900 - 2 May 1913)
The Seattle Star (Seattle — 27 February 1899 - 2 July 1930)
Seattle Weekly Gazette (Seattle — 6 August 1864 - 3 March 1866)
Vashon Island News-Record (Vashon — 24 May 1907 - 31 August 1933)
Veterans News (Seattle — 1 July 1946 - 1 December 1947)
The Veterans' Review (Seattle — 1 January 1936 - 1 March 1937)
Victory Worker (Seattle — 1 June 1942 - 2 October 1943)
Voice Of Action (Seattle — 24 April 1934 - 28 June 1935)
The Washington Gazette (Seattle — 15 August 1863)

Where Team Data Wiz comes in

1. **Research**
2. **Collaborate**
3. **Ideate**
4. **Deliver solution**
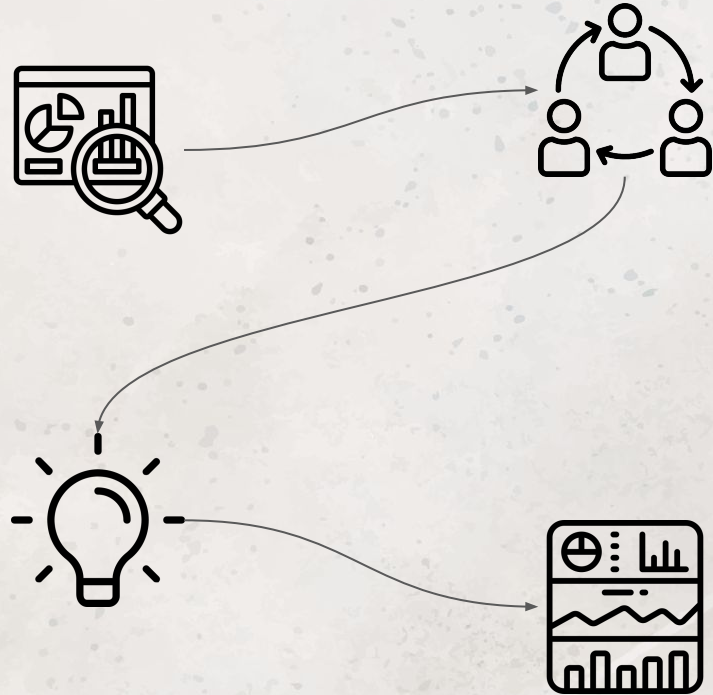
Project Schedule

	FALL' 22	WINTER' 23	SPRING' 23
Stakeholder Interactions	██████████		
Solution Ideation	████		
Requirements Gathering	█████		
Dataset development & transformation		████	
Design & Implementation			██████
Handover			███
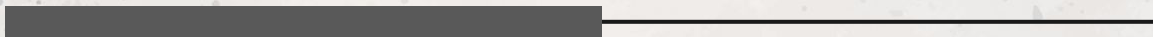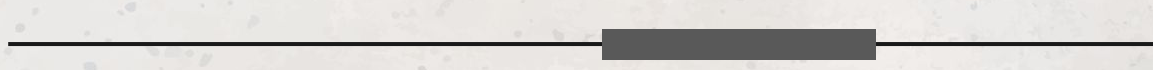
Data Extraction

Python script for extracting data

```python
93 lines (78 sloc)   3.41 KB

1   # -*- coding: utf-8 -*-
2   #!/usr/bin/env python
3   import json
4   import requests
5   import csv
6
7   # US Newspaper Directory results URL
8   results_json = 'https://chroniclingamerica.loc.gov/newspapers.json'
9
10  # List for storing digitzed titles JSON urls and title information
11  title_url_list = list()
12  all_digitized_titles = list()
13  digitized_title = []
14  count = 0
15  max_places_of_publication_count = 0
16
17  # Returns JSON results
18  def get_json(url):
19      data = requests.get(url)
20      return(json.loads(data.content))
21
22  data = get_json(results_json)
23
24  # Cycle through newspapers.json to get title url and state information
25  for record in data['newspapers']:
26      title_json_url = record['url']
27      state = record['state']
28      url_string = str(title_json_url)
29      all_digitized_info = [state, url_string]
30      title_url_list.append(all_digitized_info)
31
```
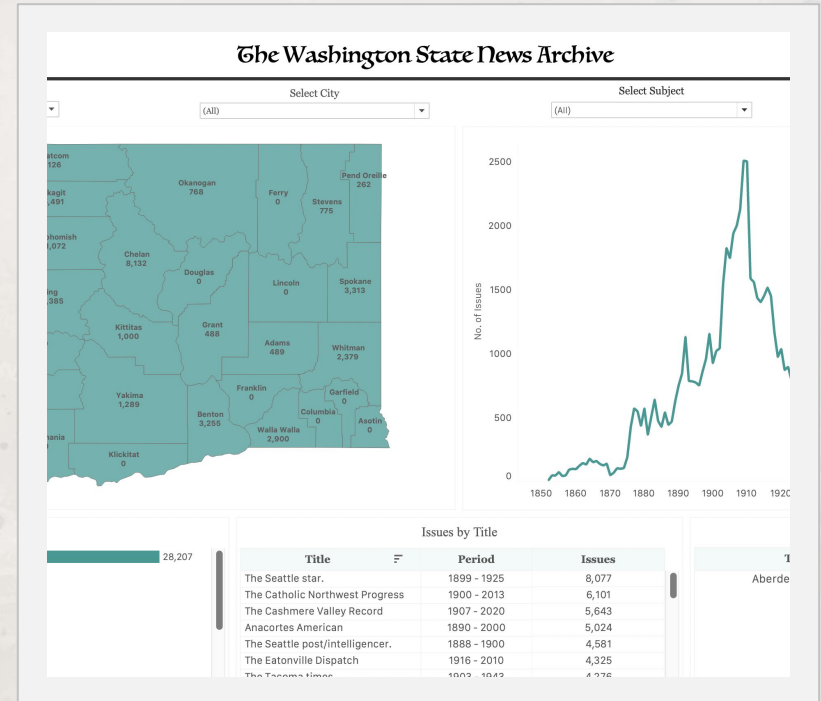
Data Transformation

Excel for housing the extracted data as well as performing data transformation

Project Solution

Tableau for Data Visualization and Insight Generation

Sample questions that users might have

- How many counties do we have represented in terms of digitized newspapers?

- Are there more titles and issues published between 1916 and 1919 in King County than in Spokane County?

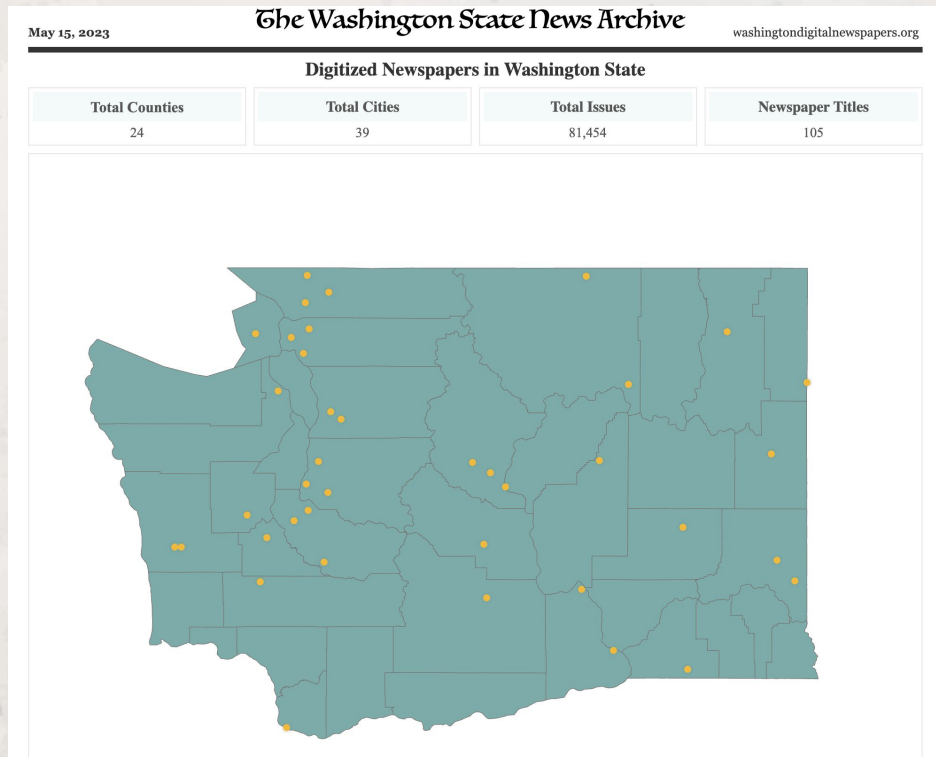- Which years are better represented in counties across the state?

Benefits

01

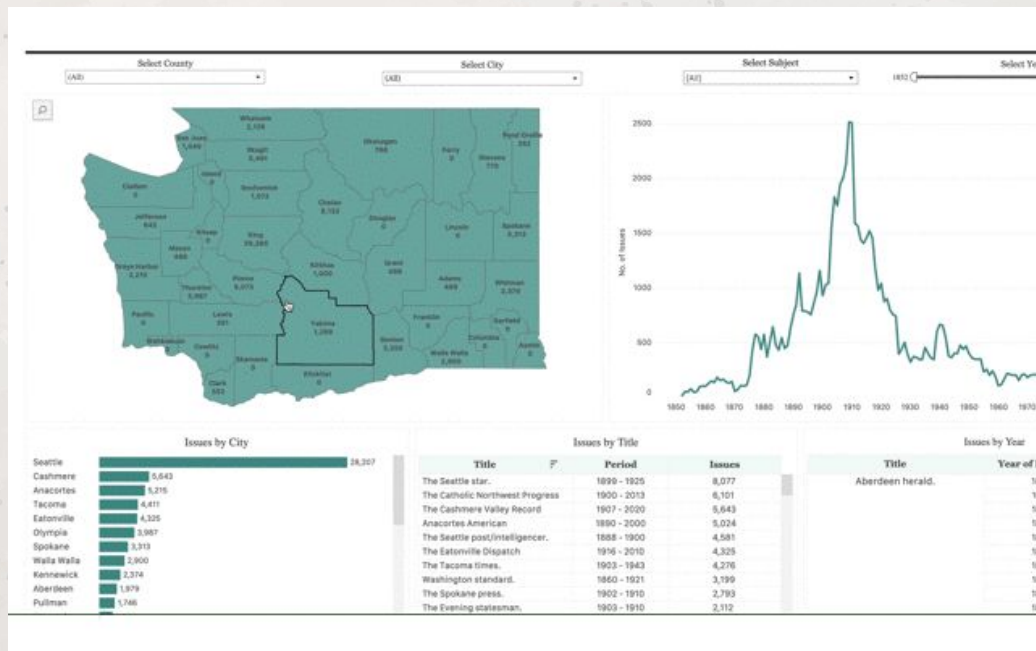## THE ULTIMATE PLATFORM

02

BRIDGING GAPS IN DATA

The Washington State News Archive

washingtondigitalnewspapers.org

Digitized Newspapers in Washington State

Total Counties	Total Cities	Total Issues	Newspaper Titles
24	39	81,454	105

Benefits

WIDEN THE REACH
03

INTERACTIVE INSIGHTS
04

Thank you!