

Reflective Essay: A Digital Workflow for Analyzing The Scholars

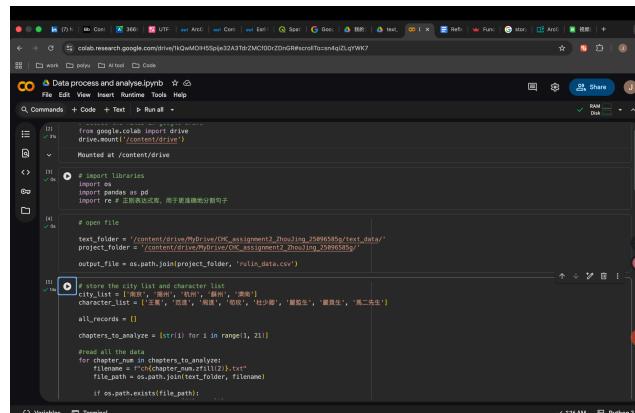
ZhouJing 25096585g

This project provided a hands-on opportunity to explore how a suite of digital tools can be orchestrated to analyze the spatial dimensions of a classic literary text. My goal was to investigate the geographical world of *The Scholars* (Rulin Waishi) by transforming its narrative into a structured dataset and, ultimately, an interactive web presentation. The process was a journey through three distinct stages: data processing, GIS visualization, and web deployment, each presenting unique challenges and learning opportunities.

Stage 1: Data Processing with Python and Human Insight

The project began with the foundational task of data extraction. My primary tool for this stage was Python, running in a Google Colab notebook. I wrote a script to parse the first twenty chapters of *The Scholars*, sourced from Ctext.org. The script's main function was to automatically scan the text and extract every sentence that contained one of five key city names: Nanjing, Yangzhou, Hangzhou, Suzhou, and Beijing. These sentences, along with their chapter and city context, were then organized into a clean CSV file, creating the initial dataset for my analysis.

However, relying solely on automation proved to be insufficient. The script operated on a simple sentence-level logic, which led to inaccuracies. For instance, my program identified a sentence in Chapter 1 mentioning "Suzhou" but failed to connect it to the character Wang Mian, whose name appeared in an adjacent sentence. This limitation taught me a critical lesson in digital humanities: a "human-in-the-loop" approach is essential. Code is powerful for the initial, large-scale processing, but human-led close reading is necessary for contextual accuracy. I therefore manually reviewed the entire output, correcting character attributions and refining the raw data into a trustworthy foundation for the next stage. After cleaning the data, I used a pivot table in Excel to quickly calculate the frequency of each city's appearance and sourced their historical coordinates from the China Historical GIS (CHGIS) project.



```
# Data process and analyze.ipynb
# File Edit View Insert Runtime Tools Help
# Commands + Code + Text > RunCell
# Mount at /content/drive
# Import Libraries
# Import os
# Import pandas as pd
# Import re # 正则表达式，用于更准确地匹配句子

# open file
text_folder = '/content/drive/MyDrive/CH_assignment2_ZhongJing_25096585g/text_data/'
project_folder = '/content/drive/MyDrive/CH_assignment2_ZhongJing_25096585g/'

output_file = os.path.join(project_folder, 'rulin_data.csv')

# use the city list and character list
city_list = ['南京', '扬州', '杭州', '苏州', '北京']
character_list = ['王冕', '周进', '范进', '包公', '杜少卿', '施耐庵', '鲁智深', '洪二先生']

all_records = []

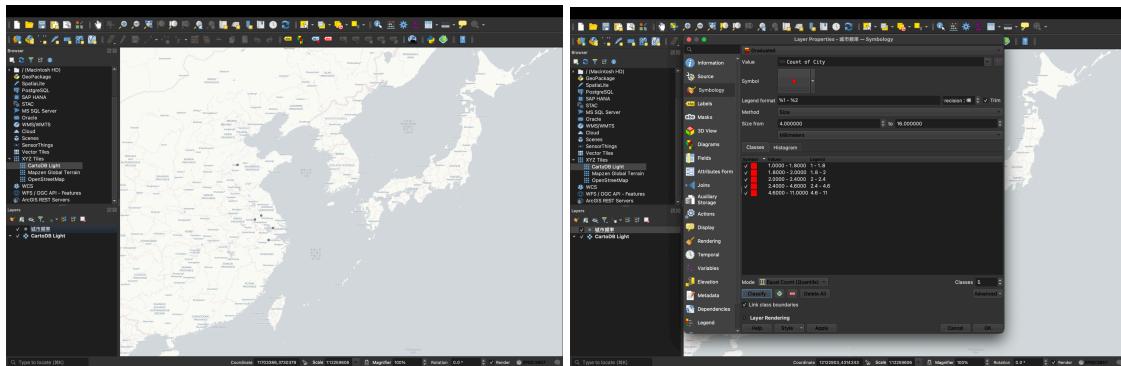
chapters_to_analyze = [str(i) for i in range(1, 21)]

# read all the data
for chapter in chapters_to_analyze:
    for filename in chapters_to_analyze:
        file_path = os.path.join(text_folder, filename)
        if os.path.exists(file_path):
            with open(file_path, 'r', encoding='utf-8') as f:
                lines = f.readlines()
                for line in lines:
                    if any(city in line for city in city_list) or any(character in line for character in character_list):
                        all_records.append(line)
```

Stage 2: GIS Visualization with QGIS

With a structured dataset ready, I moved to the visualization stage using QGIS. My objective was to create a map that clearly communicated the frequency of city mentions. This stage involved its own process of trial and error. My first map used the default OpenStreetMap basemap, but I found it was too visually "noisy," drawing attention away from my data points. I resolved this by switching to a minimalist basemap, CartoDB Light, which provided a clean, neutral background.

Using the graduated symbology feature in QGIS, I then linked the size of each city's marker to its mention frequency, making the data pattern instantly visible. To transform this static map into an interactive experience, I used the qgis2web plugin. This powerful tool seamlessly converted my QGIS project into a standalone web map (HTML, CSS, JavaScript), complete with pop-up windows that displayed city information when a user clicked on a point. This step was a bridge between data analysis and public-facing presentation.



Stage 3: Web Deployment with GitHub Pages and Streamlit

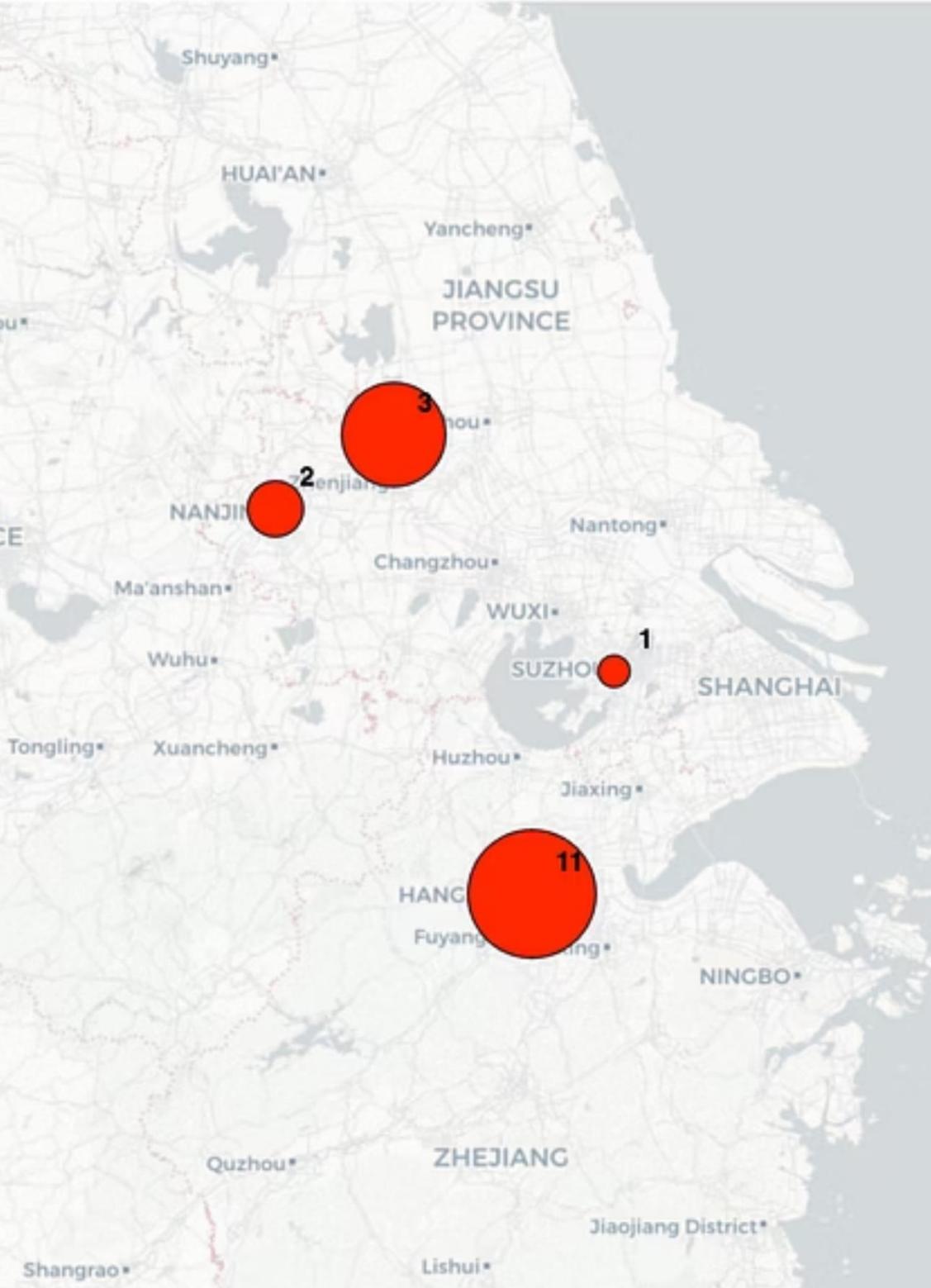
The final and most challenging stage was deploying the interactive map online. My initial attempt to embed the local map files directly into a Streamlit application resulted in a blank screen. Through troubleshooting, I discovered this was due to how web servers handle file paths. A local HTML file, when embedded, cannot find its associated CSS and JavaScript files in a web environment.

The solution was a professional two-step deployment workflow. First, I uploaded the entire qgis2web folder to a public GitHub repository. I then activated GitHub Pages, a feature that hosts static websites. However, I encountered another common issue: my map data, stored in a file named `_1.js`, was being ignored by GitHub's building process because files starting with an underscore are treated as special files. I resolved this by adding an empty `.nojekyll` file to my repository, which instructed GitHub to publish all my files without modification. This successfully generated a public, working URL for my interactive map.

With a stable online map, the role of Streamlit became clear. It was not just for showing the map, but for building a complete project dashboard. I used Streamlit to create a web page that included a project title, an introduction to my research, the embedded interactive map via its new URL, and a display of the raw data table. This transformed a single visualization

into a comprehensive and contextualized digital narrative, effectively communicating the entire scope of my research project.

Github: <https://github.com/JosicZhou/chc5904-assignment2>
Map: chc5904.streamlit.app/



A Digital Spatial Analysis of *The Scholars* (儒林外史)

Unveiling the Narrative Geography in the Opening Chapters (chapter1-20)

Josic Zhou | 25096585gCHC5904 Digital Studies of Chinese Culture

of *The Scholars* (儒林外史)

he First 20 Chapters

An analysis of key cities mentioned in the classic Chinese novel *Rulin Waishi* (*The Scholars*). This text.

nd its mention frequency.

ext.

graphical data from CHGIS.

creation, and Streamlit for web deployment.

□ chc5904.streamlit.app

A Spatial Analysis of The Scholars (儒林外史)

Visualizing Key Locations in the First 20 Chapters This interactive web application presents a spatial analysis

<https://chc5904.streamlit.app/>

**cZhou/chc5904-
ignment2**

Rulin survey

Contributor 0 Issues Stars 0 Forks 0

 GitHub

GitHub - JosicZhou/chc5904-assignment2: Rulin

Rulin survey. Contribute to JosicZhou/chc5904-assignment2 development by creating an account on

[GitHub - JosicZhou/chc5904-assignment2: Rulin survey](https://github.com/JosicZhou/chc5904-assignment2)

Central Research Question

This project seeks to answer two fundamental questions regarding the early chapters of *The Scholars*:



Cities as Narrative Space

How do different urban centers contribute to constructing the novel's unique narrative geography?



Functions & Socio-Economic Realities

What do the distinct functions of these cities reveal about the geographical focus and characters' societal positions?

Character Journeys and Urban Activities(Chapters 1-20)

A detailed overview of character movements and their primary activities within various cities in the opening chapters of *The Scholars*.

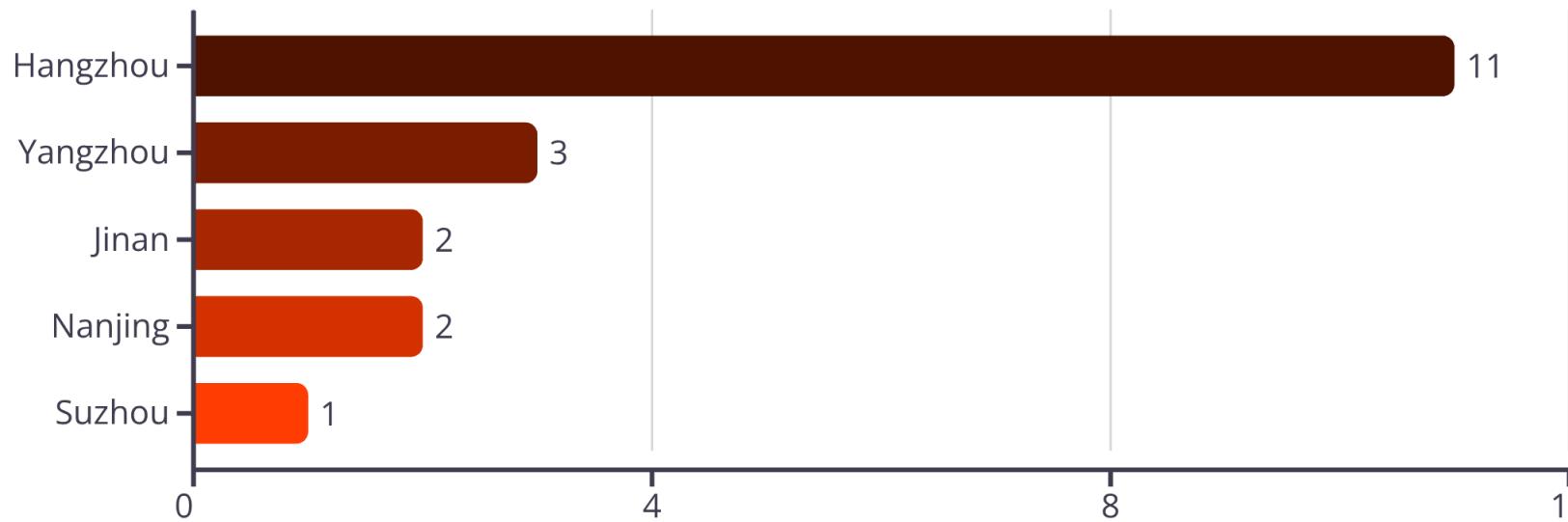
Chapter	City	Character	Activity_Summary
1	Jinan	王冕	跋涉至濟南府
1	Jinan	王冕	財主遣僕購畫擾之
8	Hangzhou	蘧公孫	贈銀作路費勸安身
12	Hangzhou	宦成	奉命乘船赴杭州
14	Hangzhou	馬二先生	籌盤費準備赴杭
14	Hangzhou	馬二先生	辭別動身赴杭州
14	Hangzhou	馬二先生	在杭州選書營務
14	Hangzhou	馬二先生	赴杭為書店選卷
16	Hangzhou	匡超人	得贈銀擬做小生意
17	Hangzhou	匡超人	熟悉杭州少人脈
17	Hangzhou	潘保正	寫信託帶往杭州
20	Hangzhou	匡超人	入城先投鄭丈人
20	Yangzhou	匡超人	包船至揚州登艙
20	Nanjing	牛布衣	至揚州轉船赴南京
20	Yangzhou	牛布衣	在揚州告別轉船
20	Hangzhou	匡超人	在杭積藏九十五卷
20	Yangzhou	匡超人	行至揚州抵達
20	Nanjing	牛布衣	經南京轉往蕪湖

This table illustrates the frequent movement and various activities undertaken by characters, particularly highlighting Hangzhou as a hub for cultural and economic pursuits, and Yangzhou/Nanjing as significant transit points.

The Primacy of the Jiangnan Region

An Unbalanced Geographical Focus

A comprehensive frequency analysis of five major cities across the first twenty chapters reveals a highly unbalanced geographical distribution. The narrative is not spread evenly across China but is intensely focused on the Jiangnan region—the prosperous area south of the Yangtze River.



As the data clearly indicates, **Hangzhou** is mentioned 11 times, establishing it as the undisputed narrative center. Yangzhou (3 mentions) and Nanjing (2 mentions) form a secondary tier of importance. In stark contrast, the northern city of Jinan (2 mentions) and the historically significant Suzhou (1 mention) remain distinctly peripheral.

This spatial imbalance suggests a deliberate authorial choice to ground the story in the cultural and economic heartland of Qing Dynasty China, reflecting where the opportunities, anxieties, and daily struggles of the scholar class were most intensely concentrated.

Analysis: Hangzhou as the "Center of Opportunity"

A Market for Knowledge:

Hangzhou is depicted as the commercial heart of intellectual life. **Ma Er** (馬二先生) explicitly states his purpose for being there: "我原在杭州選書" and explains his urgent need to return because "杭州各書店裏等著我選考卷." This portrays the city as a marketplace where scholarly skills are commodified.

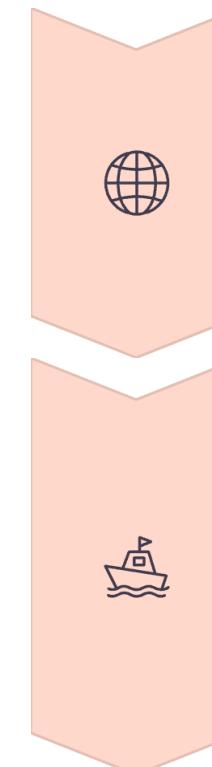
A Site for Social Mobility:

The city is also a crucial stage for personal transformation. The journey of **Kuang Chaoren** (匡超人) exemplifies this. He states, "我在杭州，虧遇著一個先生，他送了我十兩銀子，我明日做起個小生意." This initial support allows him to later accumulate significant intellectual capital: "家裏有個帳，共是九十五本." For Kuang, Hangzhou is the ladder of opportunity.

- The city is consistently portrayed as a **destination**—a place one actively *goes to* in order to advance one's career, seek patronage, or fundamentally change one's fortune. It is the gravitational center of ambition in the novel's world.



Analysis: Yangzhou as the "Transit Hub"



Yangzhou's role is purely logistical and transportation-focused. Every single mention of this city is directly linked to travel and transit. Kuang Chaoren "包了一只淌板船的頭艙, 包到揚州," while Niu Buyi "只到揚州, 弟就告別, 另上南京船."

The city functions as a critical junction in the Jiangnan transportation network, strategically positioned along the Grand Canal. It is a place of arrival and departure, of changing vessels and redirecting journeys, rather than a location for extended stays or significant events.

Analysis: Nanjing as the "Symbolic Passageway"

Despite its immense historical and cultural significance as the former capital, Nanjing's function in these chapters is surprisingly transient.

Narrative Function

Characters like **Niu Buyi** merely "獨自搭江船過了南京，來到蕪湖." The city is mentioned as a direction or a point on a larger journey, but no significant events take place there.

Symbolic Weight

In the early narrative, Nanjing serves more as a **symbolic geographical anchor** than as an active stage. Its historical weight exists in the background rather than driving the plot forward.





Analysis: Jinan as the "Moral Prologue"



Limited Narrative Scope

Jinan's role is exclusively confined to the opening story of Wang Mian (王冕), marking it as a distinct, self-contained narrative segment rather than an ongoing setting.



Integrity vs. Vulgarity

The city serves as the primary stage for a critical conflict between scholarly integrity and worldly vulgarity. This is vividly illustrated when "俗財主...遣幾個粗夯小廝, 動不動大呼小叫," directly prompting Wang Mian's retreat from urban life.



A Foundational Overture

This episode functions as a self-contained moral overture, establishing many of the novel's core values and ethical dilemmas. Once this foundational theme is set, the narrative abandons Jinan, underscoring its function as a powerful framing device rather than a continuous setting for literati interactions.

Analysis: Suzhou as the "Historical Backdrop"

Suzhou appears only once in the initial chapters, with the brief, non-narrative statement: "張士誠據了蘇州." This mention does not involve any of the novel's main characters or their direct actions. Instead, it serves a crucial temporal function, anchoring the beginning of the narrative in the turbulent transition from the Yuan to the Ming dynasty.

Consequently, Suzhou is not presented as a "place" where events unfold for the literati but rather as a "time"—a historical marker that sets the socio-political context for the world the characters inhabit.



A Jiangnan-Centric Narrative Network

Answering the Research Question

This digital spatial analysis of *The Scholars'* opening twenty chapters provides a clear and data-driven answer to our central research question by revealing a deliberately constructed and highly structured narrative geography.

Specialized Urban Functions

The novel utilizes different cities not as generic or interchangeable settings, but as specialized nodes within a carefully orchestrated network dominated by the Jiangnan region.

Hangzhou's Centrality

Hangzhou functions as the "Central Stage"—a dynamic hub of scholarly commerce and social mobility where the most important plotlines develop and character transformations occur.

Supporting Cast of Cities

Other cities serve essential but auxiliary roles: **Yangzhou** as the "Transit Hub," **Nanjing** as a "Symbolic Landmark," and **Jinan** as the "Prologue Space."

This pronounced geographical focus reveals a fictional world that is deeply rooted in the socio-economic realities of the Qing Dynasty. The intense concentration of narrative activity in the prosperous Jiangnan area—particularly in a major commercial center like Hangzhou—directly reflects the real-world environment where the scholar class navigated the complex pressures of a rapidly commercializing Confucian society.

- ❑ The narrative map is not merely a literary device but a mirror of historical reality, showing us where opportunities existed, where social mobility was possible, and where the anxieties of the literati were most acutely felt.