AI Documentation

conversation with AI for help regarding section to Create a gazetteer for the NER places, mapping the regex-extracted placenames and mapping the NER-extracted placenames

**Conversation 1:**

# Merge on the shared column: 'name' chatgpt solution-1

merged\_df = pd.merge(ner\_df, gazetteer\_df, left\_on="name", right\_on="Name", how="inner")

hey chatgpt i am facing a problem with merging place place frequency data with the coordinates from a gazetteer file but it wasn’t working. Don't give me the exact code just tell me what to do

Hey! No problem — here's what you can do step by step:

1. \*\*Check the column names\*\* in both your place frequency data and the gazetteer file. Make sure the place names are stored under the same column name.

2. \*\*Clean the column names\*\* by removing any extra spaces or invisible characters that might cause mismatches. Sometimes just a small typo or space can break the merge.

3. \*\*Standardize the data\*\* — make sure all place names are in the same case (like all lowercase) to avoid case-sensitive mismatches.

4. \*\*Merge the datasets\*\* using the shared place name column once everything matches.

5. \*\*Check the result\*\* by printing a few rows to confirm that the merge worked as expected.

Let me know how it goes or if you run into another issue!

**Conversation 2:**

# Better map style with labels , chatgpt solution -2

fig.update\_layout(

mapbox\_style="carto-positron",

margin={"r":0,"t":40,"l":0,"b":0}

Hey chatgpt I made a map showing the places using the data ang it looks really distorted and kind of messy the labels are hard to read. How can I make it look better

Hey! It sounds like the default map style might be cluttered or not well-suited for your data. Here's what you can do to improve it:

1. \*\*Switch to a cleaner map style\*\* — something like `"carto-positron"` or `"open-street-map"` works well for readability. These styles have light backgrounds and clear labels that don’t overpower your data.

2. \*\*Adjust zoom levels and marker sizes\*\* so the map doesn’t look overcrowded.

3. \*\*Add hover tooltips or popups\*\* only if necessary — too much text on the map can make it look messy.

4. \*\*Check your data points\*\* to make sure they aren’t overlapping too much in one area. If they are, you can cluster or group them.

Changing the map style alone usually makes a big difference. Want help choosing a specific style based on how your data looks?

**Conversation 3:**

# Save both formats

fig.write\_html("ner\_map.html")

fig.write\_image("ner\_map.png", width=800, height=600, scale=2) #chatgpt-sol 3

I’ve finished creating my map, but now I want to save it. I want to save it in html and png format. how should i do it. dont give me the code just give me instructions

Great! Here are the steps to save your map in both \*\*HTML\*\* and \*\*PNG\*\* formats:

1. \*\*Save as HTML\*\*:

Use your plotting library's built-in function to save the interactive version of the map as an HTML file. This version will let you zoom, pan, and hover over points. Choose a meaningful file name like `map\_output.html`.

2. \*\*Save as PNG\*\*:

To save a static image version, you'll usually need to use an export function that supports static rendering (some libraries may need additional setup like installing \*kaleido\* or \*orca\*). Make sure your map is properly zoomed and styled before exporting.

3. \*\*Check your output folder\*\*:

Make sure both files were saved in your desired directory, and give them clear, related names (e.g., `map\_output.html` and `map\_output.png`).

4. \*\*Test both files\*\*:

Open the HTML file in a browser to check interactivity, and open the PNG in an image viewer to confirm the appearance is correct.

Let me know what plotting library you're using (like Plotly or Folium), and I can tailor the steps if needed.

2.can i also use ner\_map.png for map\_output.png

Yes, absolutely! You can name the file \*\*`ner\_map.png`\*\* instead of `map\_output.png` — the filename is completely up to you. Just make sure the name is:

\* \*\*Descriptive\*\* (like `ner\_map.png` if it's a Named Entity Recognition map),

\* \*\*Consistent\*\* (use a similar name for the HTML version, like `ner\_map.html`),

\* And has the correct \*\*file extension\*\* (`.png` for the image, `.html` for the interactive version).

This makes your files easier to organize and understand later. Want help coming up with a good naming system for your project files?

**Conversation 4:**   
#print(merged\_df)

# Cleaning data (remove rows with missing lat, lon, or frequency)

merged\_df = merged\_df.dropna(subset=["latitude", "longitude", "count"])#chatgpt sol-4

i am facing an issue after merging the place frequency data with coordinates I tried to print the merged dataframe but some rows look incomplete and the map still isn’t displaying all the places. What should I do?

give me instructions

No problem! Here's what you should do step by step to fix this issue:

Check for missing values: After merging, look at your dataframe to see if any rows have missing values in key columns like latitude, longitude, or frequency.

Drop incomplete rows: Remove any rows where any of those important columns are empty (i.e., NaN). These incomplete rows can cause your map to leave out locations or crash during plotting.

Double-check the cleaned data: After dropping the missing rows, print the dataframe again (head() or sample few rows) to make sure the remaining data looks complete.

Proceed with mapping: Now that your dataset only includes complete, valid rows, recreate or refresh your map — it should work as expected.

Let me know if you'd like help checking which rows are missing data before you clean them!