Analysing the frontline of diplomacy: the case for unleashing GIS

A proof of concept algorithm supplemented with i| indicative clustering analysis of diplomatic post locations and ii| conflict events within cities.

Associated code available here https://github.com/antoniosfiala/UCL_CASA_GIS

A | Introduction

The discipline of International Relations (IR) has, much like its subject of study, undergone significant change. This includes the gradual adoption of Geographic Information Systems (GIS) as part of its methodological toolkit. This process has not been without its methodological challenges and lessons. (Branch, 2016)

With this reality in mind, this paper draws on academic literature and projects in the field such as the *Global Diplomacy Index* (GDI) and the *Armed Conflict Location & Event Data* (ACLED) project in order to build a favourable case for the use of GIS tools in IR by focusing on the discipline's work in the, until recently neglected, diplomatic field and open a new level of analysis in the process.(Sharp, 1999)

Building on the lessons of Branch (2016) and other theoretical and practical observations across IR and GIS in section B, the scope of this paper is as much methodological as it is enquiry driven.

The process outlined in section C has two aims. Firstly, to outline the means through which datasets such as GDI can be 'enhanced' and address Branch's measurement validity critique. Secondly, have a publicly accessible, reusable and improvable 'proof of concept' data processing algorithm in R to serve as a base from which to explore this new dimension. Thus, lowering some barriers to entry in the pursuit of new lines of enquiry in the process. Embedded within the process are two case studies, explored in section D, that serve to illustrate academic and practical lines of enquiry that could be opened up, namely:

i | Testing concertation of diplomatic outposts within cities using DBSCAN.
What analysis can be done within geo-spatially enhanced datasets.

ii | Visualising potential exposure of diplomatic outposts to ACLED events.
How easy is it to link multiple datasets together?

Concluding remarks in section E, acknowledge the trade-off between breadth and depth in this work, while at the same time reflecting on the value of an open source workflow that has been opened up possibilities for intricately nuanced enquiries for academic.

B | Literature review

This paper first addresses three lines of enquiry in order to build an appropriate foundation.

i | B.1 | What role has GIS had in IR?
A new set of tools for a new world order.

ii | B.2 | What is the role of diplomatic outposts and those serving within them?
Why what diplomats do matters as much as exactly where they do it.

iii | B.3 | What role can GIS tools play in the field of diplomacy?
Why and how GIS can be applied.

B.1 | What role has GIS had in IR?

In order to understand the role of GIS in IR, it is useful to reflect on the subject matter covered by IR scholars over the last 30 years. This period covers systemic change covering dramatic shifts in the global balance of power (cold war onwards)(Kennedy, 1993), the displacement of the state as one of *the* central units of analysis with the proliferation of non-state actors and the blurred lines between domestic and foreign ties. (Brenner, 2017; Lacher, 2006; Marshall, 2016; Thompson, 2018) The centrality of individuals and local actors opened the black box and required a more precise toolkit to understand their relationships through network analysis and geographical positioning through GIS.

Although progress is noted in academia (Branch, 2016) and policy circles (Christian, 2005; de Groot, 2019), deployment is often fragmented. From the work so far, two methodological lessons are highlighted by Branch. It is important for the discipline to begin addressing though before they are embedded in a critical mass of work.

i |Selection bias: *"the methodological focus on spatial features built into GIS removes nonspatial units from the analysis"* giving the appearance of completeness when in fact there may be missing, non-spatial explanatory variables (Branch, 2016, p.852)

ii |Measurement validity: *"Measurement may not be valid when the political phenomena being studied are not being captured accurately by the data collected in a GIS data set."* (Branch, 2016, p.852)

The works that successfully mitigate these problems focus on point-data capture (countering measurement validity), coding each observation with coordinate data points and moving beyond the so-called 'territorial gap' in the process. (Agnew, 1994) Further, academics can then proceed to combine their analyses with theoretical and political layers to get a more complete picture when testing their hypotheses, addressing their selection bias. Ultimately, the key factor is an appropriate theoretical foundation and level of analysis for the question on hand. To facilitate that, easier access to relevant data is paramount. Cederman et al. (2009), Raleigh et al. (2010), Tollefsen et al. (2012) all stand as examples that *get it right* when it comes to ethnic conflict, aid and widely usable platforms demonstrating worthwhile avenues exist, the question remains whether our understanding of diplomacy stands to benefit.

B.2 | What is the role of diplomatic outposts and those serving within them?

Like the system it inhabits, diplomacy, as the study of how actors in international relations engage has undergone notable change from being "almost disconnected" from the wider IR world (Sharp, 1999) to "diplomacy 2.0", characterised by intricate networks across numerous spatial levels and dimensions. (Constantinou et al., 2016).

The role a diplomatic outpost, which can include embassies, consulates and permanent representations, is also not without debate. Moving beyond a pure state-to-state communication role, the works of Ferguson, Forslid and Rose support the argument that the renewed raison d'être of diplomatic outposts is to boosts exports. (Ferguson and Forslid, 2019; Rose, 2007)

Adding, as opposed to negating the above, Cooper and Cornut (2019) highlight the (re)widening of the scope of responsibility for frontline diplomacy. For instance, faster decision-making due to the pace of events indicating the return of some power from centralised ministries to the ground. Furthermore, public diplomacy and image management as well as intelligence gathering in a new 'polylateral' environment where state, non-state actors, enterprises and individuals can often have equal weight. (Cooper and Cornut, 2019; Ellen Huijgh and Caitlin Byrne, 2019; Murray et al., 2011) This suggests that the right placement of an outpost as well as digital means can play a significant role in reducing barriers in pursuing any one of these duties.

Nevertheless, most of work to date concerned with the *where* focus on a state-to-state level of analysis and where connections are made, not the micro-level interactions on the ground. (Webster, 2001; Webster and Ivanov, 2015) It is unclear whether this is a question of lack of demand for this kind of research or lack of supply of data to operationalise it.

The Lowy Institute took the first step in addressing the supply gap in 2016 with the launch of the GDI, a portal and dataset that captures the placement of diplomatic outposts for 61 countries (by 2019) down to city level.(Lowy Institute, 2020) While it's coverage is incomplete, the 61 countries for which is gathered account for ~92% of the global economy and ~75% of the world's population.[1] (World Bank, 2020) Crucially, it is curated that makes it a more credible source compared to others such as Open Street Map. Although the GDI is a useful utility, it stops short of the final spatial frontier which in light of the above is potential fertile ground for further enquiry and mitigating of Branch's measurement validity critique.

[1] Calculation performed based on World Bank figures for the globe and the sum of metadata for each country in the GDI dataset the capture GDP and population. See workflow for more detail.

B.3 | What role can GIS tools play in the field of diplomacy?

The intertwined nature of the global and the local trickling down from IR to the study of diplomatic relations is the principle argument for GIS having a greater role in expanding our understanding and improve policy decision-making. Enquiries that could benefit could be those of planning, security, resource consolidation. Few examples could be:

Academia:

i | Do diplomatic outpost clusters form in cities? What drives their location?
ii | Are there frequent groups of outposts found in the same cluster across locations? Do shared interest help explain this?
iii | Do cities (or areas within) benefit from diplomatic clusters similarly to industrial clusters?

Policy:

iv | Given certain policy preferences, where best to set up a diplomatic outpost?
v | How exposed are diplomatic missions to incidents of conflict? (de Groot, 2019)
vi | … and others

The answer to any of these questions depends on the availability of exact data points for the locations of diplomatic posts. The GDI dataset as it is does not provide this, but it is the means through which it can be achieved. This is the first output of this paper, an algorithm which adds this missing feature and can be used and improved on demand for subsequent data refreshes.

Secondly, to illustrate its utility, case studies touching on aspects of B.3.i clustering within cities and B.3.v exposure to conflict events.

C | Methodology

The content of section C is supplemented by code stored and available to download/view in a GitHub repository. The entire process is outlined in figures C.2.fig.2 and C.3.fig.3 in parts or in appendix G.2 as a whole. All code has been written in R with commentary beyond this work within to outline the role of each block of code. This section will describe key features of the GDI dataset, explain the role of Google's geocoding API in adding a coordinate dimension, and lay the foundation for the two case studies to be explored.

C.1 | Understanding the GDI dataset

The GDI dataset from the Lowy Institute is available for the years 2016, 2017 and 2019 where all follow a similar structure. As the most comprehensive, the 2019 instance is used to construct and explain the process through which an additional spatial dimension is added.

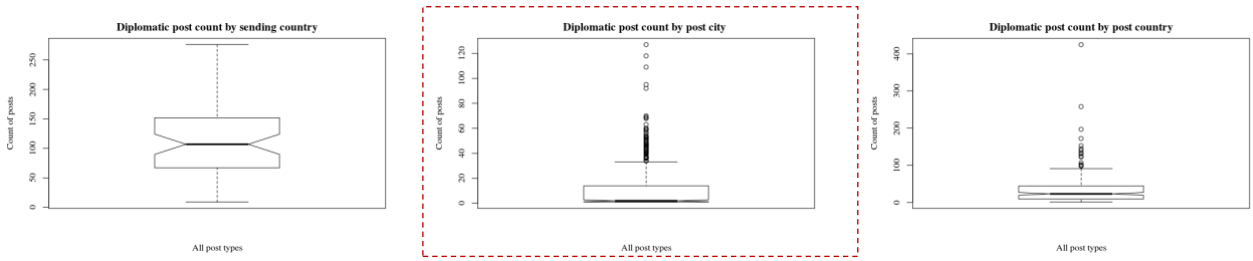
C.1 | Table 1 | Lowy Institute's Global Diplomacy Index (GDI) dataset

Type of data	Columns	Comments
Ranking data	[1] Overall rank; [2] G20 rank; [3] OECD rank; [4] Asia rank;	Rank data computed based on number of posts, more details in (Lowy Institute, 2019)
GDP and population	[5] GDP – USD billions; [6] Population in millions;	Sourced from World Bank
Diplomatic data	[7] Country sending the diplomatic mission [distinct count: 61]; [8] City mission is posted [d.c.: 686] [9] Country mission is posted [d.c: 207] [10] Type of post [distinct count: 4] [11] Title of post [d.c.: 141]	Locations of diplomatic posts, sourced from Ministries of Foreign Affairs and includes: [i] Embassies/High commissions; [ii] Consulates/Consulates-General [iii] Permanent missions to multilateral organisations [iv] Representations/delegations to states with no other formal diplomatic relations

***Note**: 'sending country' concept, refers to countries who dispatched diplomatic missions abroad and refers to values in column 7; 'post' city/country refers to the geographical unit of posting where the diplomatic mission is based.*
For example: the German Embassy in London, UK. Sending country == Germany; Post city == London; Post country == UK;

Referred to previously, the curation process behind the GDI is its unique selling point as it reduces likelihood of human error and increases reliability. However, its breadth is also of note. With 7,317 records it is almost double the number of datapoints in Open Street Map which stands at 3,754.(Open Street Map, 2020) See appendix G.1 for more detail. Scope for summary statistics is limited beyond unique values in C.1.Table 1. C.1.fig.1 however paints a highly skewed picture with the centres of diplomacy (i.e. cities of powerful states) showing up as outliers. Of interest is the fact that 50% of cities hosting outposts have 2 or less, the 75th Percentile has 14 and anything above the plots 33 count maximum is considered an outlier.

C.1 | Figure 1 | boxplots for summary of posts by sending country, post city and post country

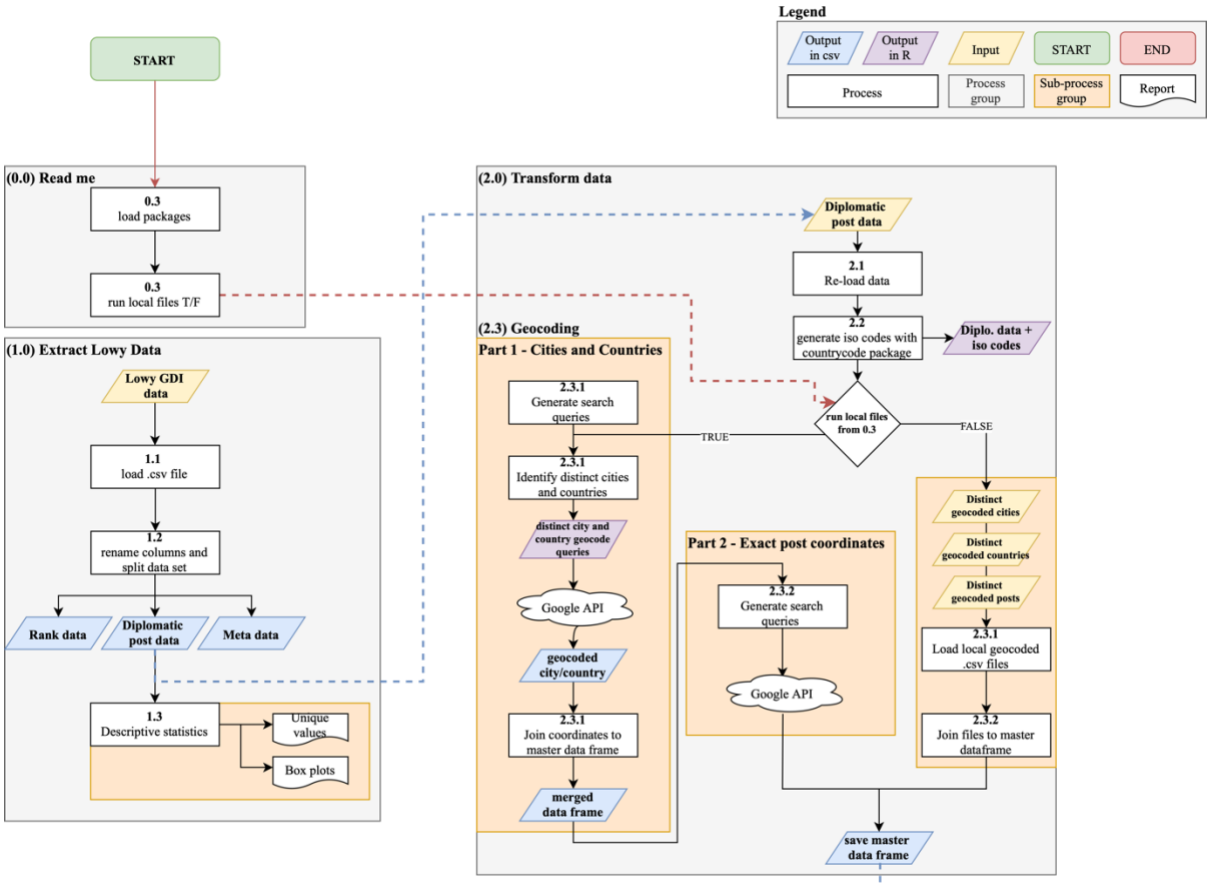


C.2 | Processing the GDI dataset

The central part of the algorithm is the enhancement of the GDI dataset with additional features that will make it easier to link with other datasets and add the coordinate dimension through geocoding. In this instance, Google's geocoding API is the chosen solution due to suggested higher degree of accuracy and previously discussed issues with Open Street Maps, if the features are not captured in the first place, one cannot hope to geocode them.(Lemke et al., 2015) The API is not free, however there is sufficient allowance of credits that allows for an infrequent request such as this to be arguably open to use, although not strictly open-source.

Figure C.2.fig.2 outlines the process for steps 0 through to 2, covering workflow set up through the geocoding.

C.2 | Figure 2 | Process map, steps 0 to 2



Parts of the workflow that warrant elaboration here are:

| Part 1.2 | Renaming of columns for easier legibility and separation of raw table to the three sub-tables described in C.1.Table 1

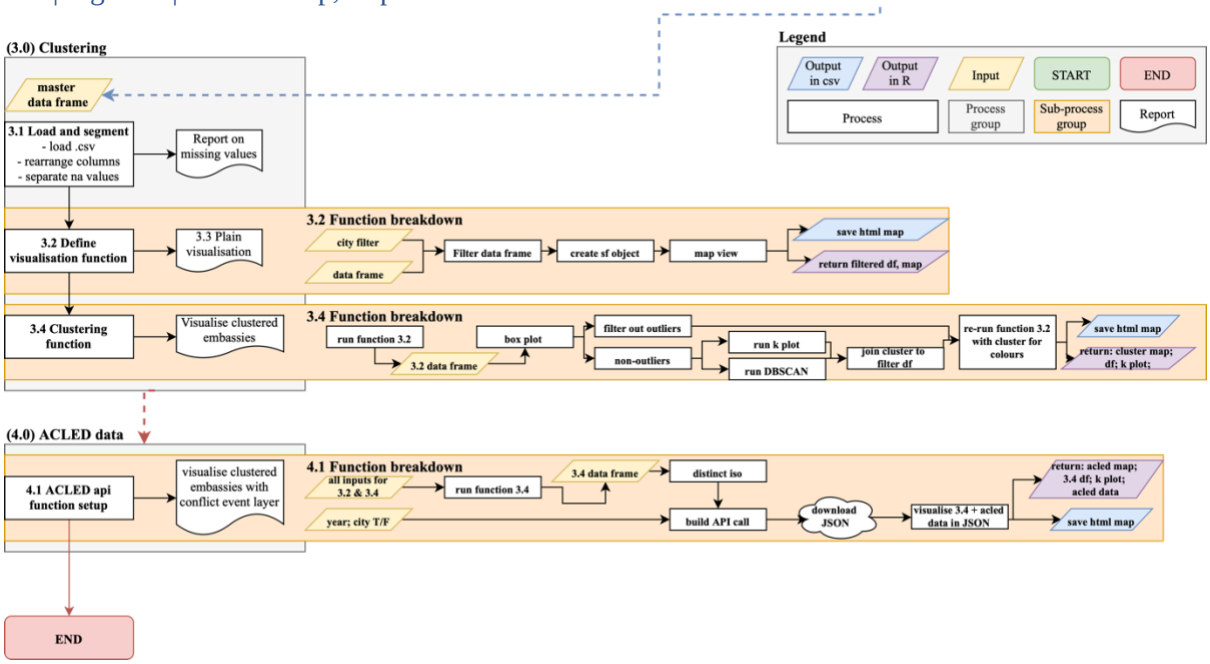
| Part 2.2 | Assign ISO codes against each country, important for Part 4.1 and a case of easier integration with other datasets

| Part 2.3 | API key and calls to geocode cities, countries and exact outpost locations:
| countries are based on country name, i.e. 'Australia'
| cities with diplomatic outposts are geocoded through combination of city and country, i.e. 'Alberta, Canada'
| exact outpost locations are coded based a string in the following format 'send_country' + "Embassy" + 'post_city' +'post_country',
i.e. Argentina, Embassy, Algiers, Algeria

C.3 | Clustering and visualisation

In the final steps, values that were not successfully geocoded are excluded in step 3.1.[2] Furthermore, an important part of the process is automatically identifying geographical outliers in step 3.4 using boxplot analysis ahead of using the DBSCAN algorithm to assess the diplomatic posts of the city.

C.3 | Figure 3 | Process map, steps 3 & 4



DBSCAN was chosen for its property for finding clusters within noisy data, compared to hierarchical or k-means methods. It requires the selection of an ε value (eps) and minimum points (minpts).(Schubert et al., 2017) To this end a k-nearest neighbour plot is used and added to all function outputs to allow adjustments.(Kurumalla and Rao, 2016) Finally, the results are visualised using the mapview package.

C.4 | Merging with external data: ACLED methodology

Finally, addressing point A.ii, step 4, carries out an API call to ACLED filtering down the data based on the characteristics of the clustered dataset. Taking into account the country (using iso codes from step 2.2) and city. As previous point data, the ACLED points are added as a layer to the mapview visualisation.

Outlining the data processing algorithm illustrates the extent of completeness (99.4%) and the geocoding means through which the GDI dataset can have a spatial dimension added. Subsequent functions outlined in steps 3.2 and 3.4 are subsequently ready to be used to carry out enquires into clustering of diplomatic posts within cities and overlay ACLED data respectively.

[2] In the Lowy GDI 2019 dataset, there are 7,317 entries with ~99.4% geocoded (7,274)

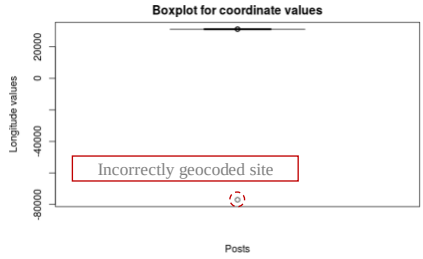
D | Proof of concept analysis

At this stage, the dataset contains 684 cities, selection of places to assess will firstly focus on areas common across both GDI and ACLED and subsequently choose cities of contemporary geo-political significance.

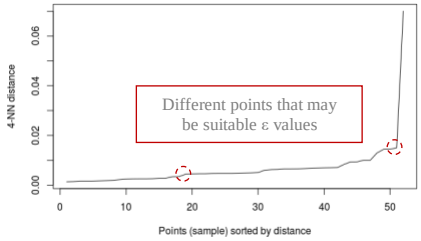
D.1 | Cluster analysis in Cairo

Egypt enjoys a unique geographical position with the role the Suez Canal plays in global trade. Cairo as its capital is a means to understand any policy change as well as assess how one the world's biggest cities is operating. Figures D.1.4-7 illustrate the process of removing clear outliers, selecting an appropriate eps value (0.003 in this case) and minimal points value of 4. There are five clusters of diplomatic outposts overall and the majority are within Gezira island or its vicinity.

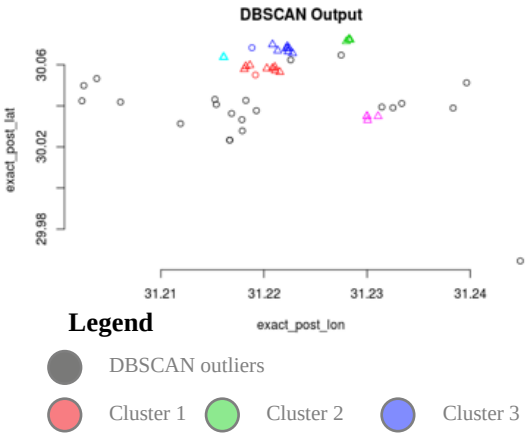
D.1 | Figure 4 | Cairo scaled longitude boxplot



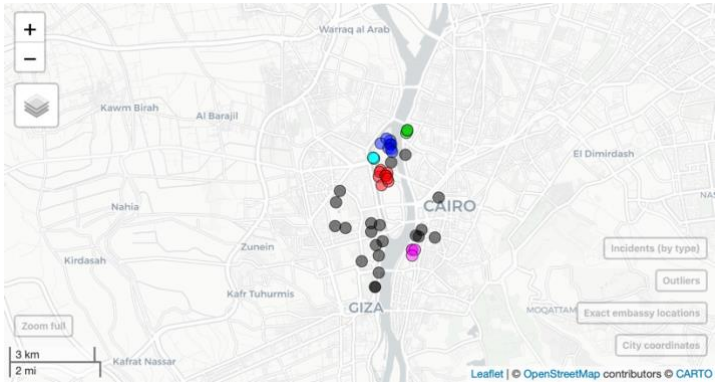
D.1 | Figure 5 | Cairo k nearest neighbour plot



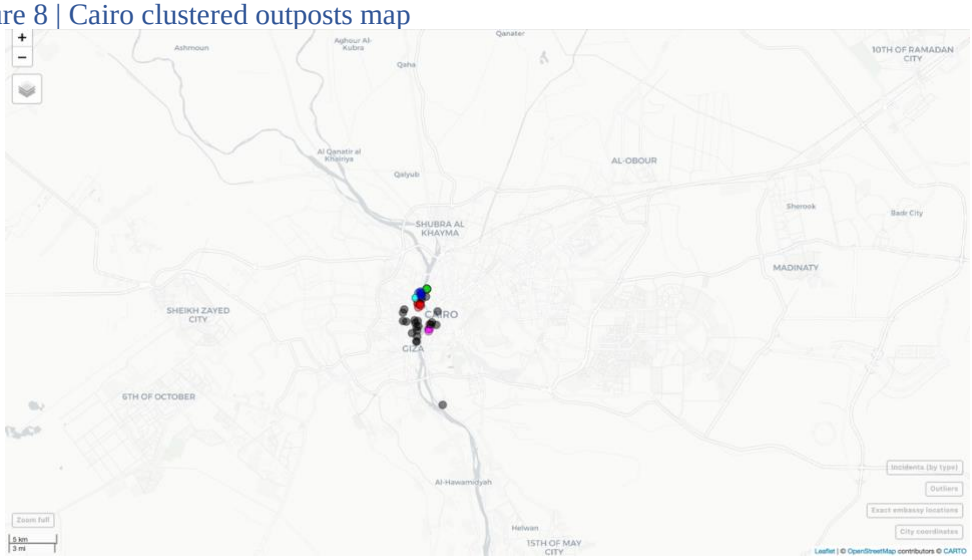
D.1 | Figure 6 | Cairo cluster scatter plot



D.1 | Figure 7 | Cairo clustered outposts map



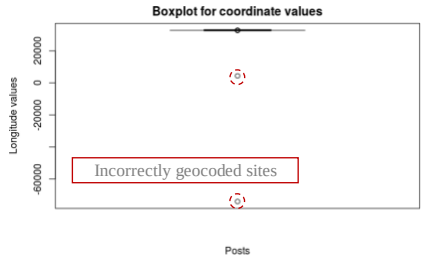
D.1 | Figure 8 | Cairo clustered outposts map



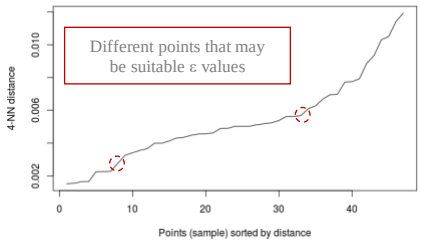
D.2 | Cluster analysis in Ankara

Turkey is another critical strategic actor. Figures D.2.8-11 illustrate the process of removing clear outliers, selecting an appropriate eps value (0.003 in this case) and minimal points value of 4. It suggests there are 3 areas of dense clusters within the city. In the context of the broader city, all appear to be concentrated in the south-eastern part of the city.

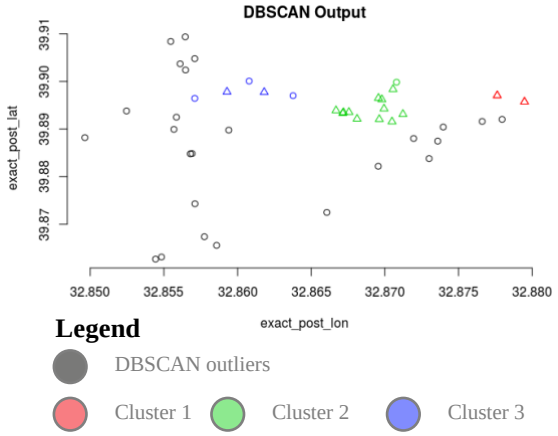
D.2 | Figure 9 | Ankara scaled longitude boxplot



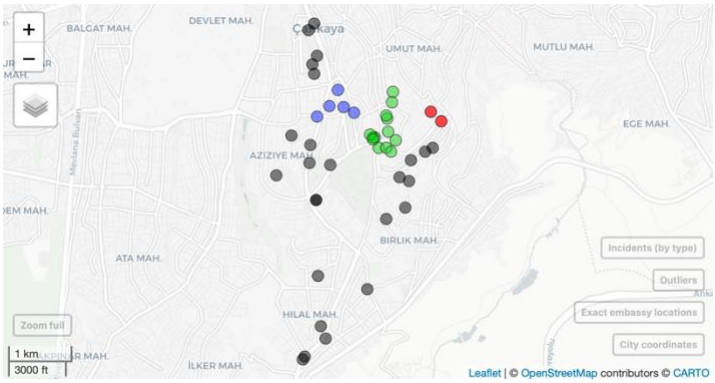
D.2 | Figure 10 | Ankara k nearest neighbour plot



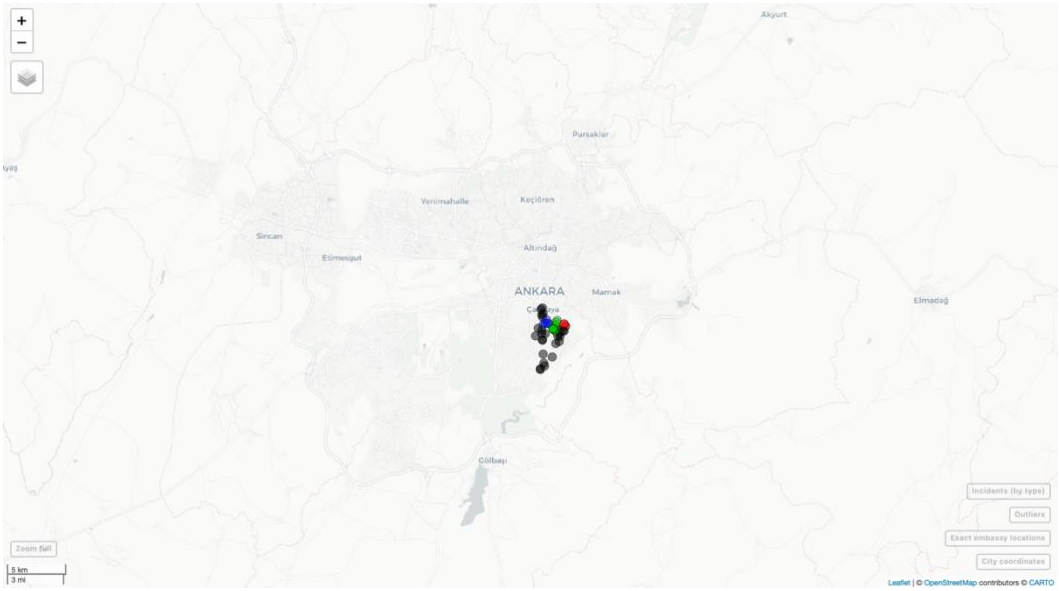
D.2 | Figure 11 | Ankara cluster scatter plot



D.2 | Figure 12 | Ankara clustered map



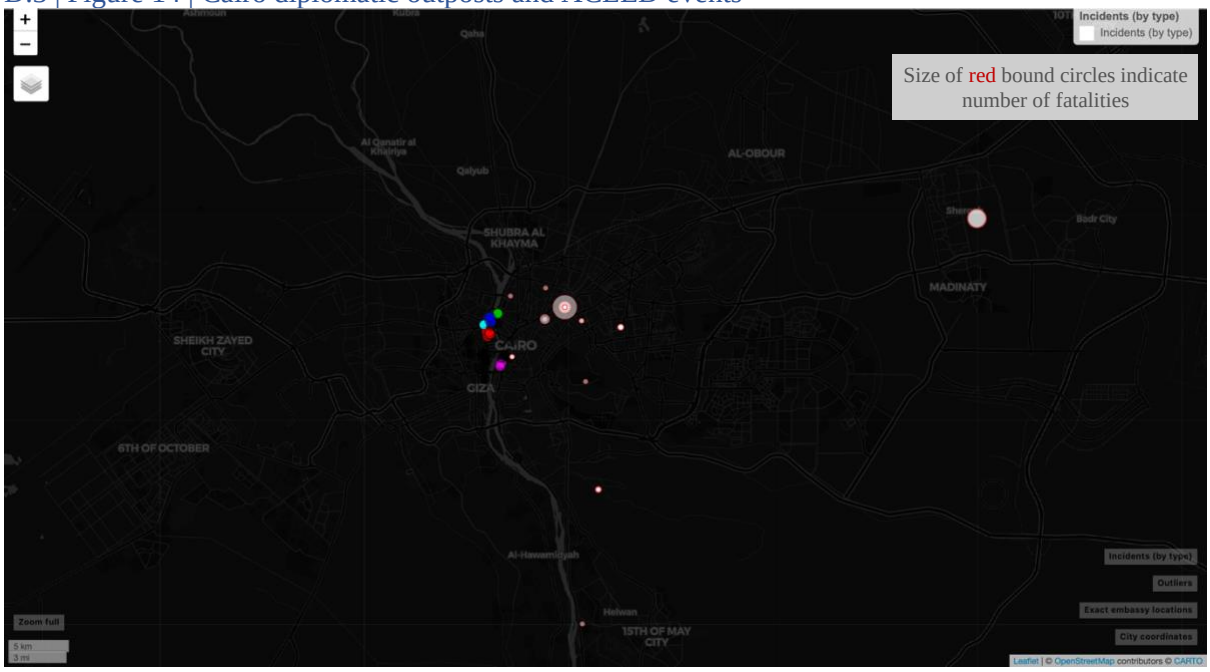
D.2 | Figure 13 | Overview of Ankara city and clusters within



D.3 | ACLED data overlay

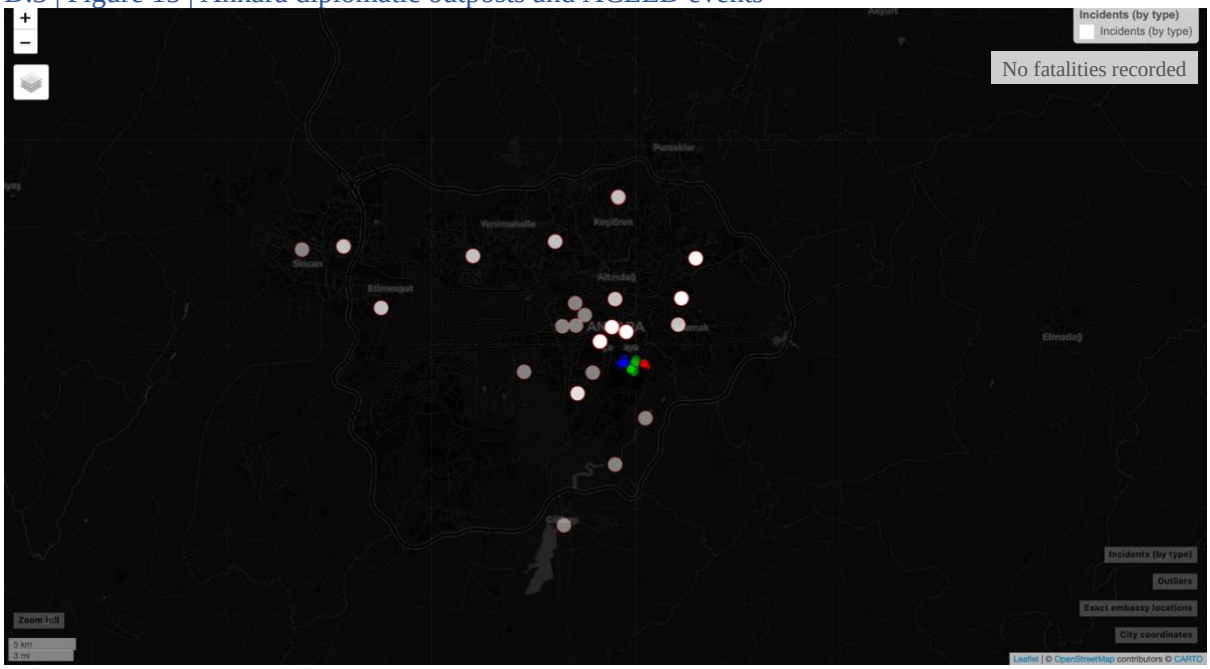
Within the same maps, ACLED events can be visualised to assess the location and proximity of protests, unrest and conflict events. D.3.fig.14 suggest that the concentration of ACLED events is largely to the east of where diplomatic outposts are located. Given the overlay of events more ought to be done whether it is an encoding issue or whether that is a particular hotspot in Egypt's capital.

D.3 | Figure 14 | Cairo diplomatic outposts and ACLED events



In contrast, while Ankara in D.3.fig.15 shows greater dispersal of ACLED events (mixture of riots, protests, violence against civilians), none take place directly among the clusters of diplomatic posts and generally in more central parts of the city, north and to the west of the diplomatic cluster.

D.3 | Figure 15 | Ankara diplomatic outposts and ACLED events



E | Conclusion

This work set out by reflecting on the prevalence of GIS within IR and set out to plug a gap in the availability of process, data and illustrate potential utility of filling this gap through case studies on clustering of diplomatic outposts within cities and their position within the city itself, as well as the overlay of additional datasets with the necessary coordinates.

While far from conclusive given the limited sample, the data presented supports the formation of a hypothesis that diplomatic outposts are formed in clusters and have groupings within. The significance or consistency of internal clusters between cities is an avenue that can be pursued to test the effect of alliances or rivalries on decision of where to locate diplomatic centres of operation.

The case of overlaying ACLED data served to illustrate the easy with which the coming together of disparate IR-relevant datasets when they both have common features, in this case coordinates and appropriate country identifiers such as ISO codes. From a research perspective, it serves to prompt the hypothesis that the concertation of conflict events take place around but rarely within diplomatic clusters. Methodologically, it is a call for more datasets to include a spatial dimension to allow far more nuanced work to take place with lower barriers to entry or preparation time.

Lowy's GDI dataset is an excellent example of a much-needed step in the right direction, however the scope for improvement is evident.

Finally, circling back to Branch's critiques, while additional or better coded spatial features can alleviate the second of his two critiques of GIS in IR, measurement validity (B.1.ii), selection bias (B.1.i) can only be tackled through the theoretical grounding of IR research and an understanding of what is needed to supplement and appropriately contextualise a line of enquiry.

All code, as well as outputs are available online, here. Additions, improvements are welcomed to what is a proof of concept algorithm (workflow)

Word count: 3,011

Page left intentionally blank

F | Bibliography

Agnew, J., 1994. The territorial trap: the geographical assumptions of international relations theory. Rev. Int. Polit. Econ. 1, 53–80.

Branch, J., 2016. Geographic information systems (GIS) in international relations. Int. Organ. 70, 845–869.

Brenner, N., 2017. Beyond state-centrism? Space, territoriality, and geographical scale in globalization studies, in: Theory and Methods. Routledge, pp. 313–336.

Cederman, L.-E., Buhaug, H., Rød, J.K., 2009. Ethno-nationalist dyads and civil war: A GIS-based analysis. J. Confl. Resolut. 53, 496–525.

Christian, C., 2005. Challenges: adopting GIS for diplomacy and foreign policy. Presented at the ESRI Federal User Conference Proceedings, 2005, id. 1.

Constantinou, C.M., Kerr, P., Sharp, P., 2016. The SAGE handbook of diplomacy. Sage.

Cooper, A.F., Cornut, J., 2019. The changing practices of frontline diplomacy: New directions for inquiry. Rev. Int. Stud. 45, 300–319. https://doi.org/10.1017/S0260210518000505

de Groot, J., 2019. Informal conversation on GIS projects at Ministry of Foreign Affairs in Netherlands.

Ellen Huijgh, Caitlin Byrne, 2019. Opening the Windows on Diplomacy: a Comparison of the Domestic Dimension of Public Diplomacy in Canada and Australia. Brill | Nijhoff, Leiden, The Netherlands, pp. 83–111. https://doi.org/10.1163/9789004394254_006

Ferguson, S., Forslid, R., 2019. Sizing Up the Impact of Embassies on Exports. Scand. J. Econ. 121, 278–297.

Kennedy, P., 1993. Preparing for the 21st Century. Harper Collins.

Kurumalla, S., Rao, P.S., 2016. K-nearest neighbor based DBSCAN clustering algorithm for image segmentation. J. Theor. Appl. Inf. Technol. 92, 395.

Lacher, H., 2006. Beyond globalization: capitalism, territoriality and the international relations of modernity. Routledge.

Lemke, D., Mattauch, V., Heidinger, O., Hense, H., 2015. [Who Hits the Mark? A Comparative Study of the Free Geocoding Services of Google and OpenStreetMap]. Gesundheitswesen Bundesverb. Arzte Offentlichen Gesundheitsdienstes Ger. 77, e160-5. https://doi.org/10.1055/s-0035-1549939

Lowy Institute, 2020. Lowy Global Diplomacy Index [WWW Document]. Lowy Glob. Dipl. Index. URL https://globaldiplomacyindex.lowyinstitute.org/about (accessed 1.5.20).

Lowy Institute, 2019. Lowy Global Diplomacy Index Methodology.

Marshall, T., 2016. Prisoners of geography: ten maps that explain everything about the world. Simon and Schuster.

Murray, S., Sharp, P., Wiseman, G., Criekemans, D., Melissen, J., 2011. The present and future of diplomacy and diplomatic studies. Int. Stud. Rev. 13, 709–728.

Open Street Map, 2020. Tag info, Open Street Map, office = diplomatic [WWW Document]. TagInfo. URL https://taginfo.openstreetmap.org/tags/?key=office&value=diplomatic#overview (accessed 10.10.20).

Raleigh, C., Linke, A., Hegre, H., Karlsen, J., 2010. Introducing ACLED: an armed conflict location and event dataset: special data feature. J. Peace Res. 47, 651–660.

Rose, A.K., 2007. The foreign service and foreign trade: embassies as export promotion. World Econ. 30, 22–38.

Schubert, E., Sander, J., Ester, M., Kriegel, H.P., Xu, X., 2017. DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. ACM Trans. Database Syst. TODS 42, 19.

Sharp, P., 1999. For diplomacy: representation and the study of international relations. Int. Stud. Rev. 1, 33–57.

Thompson, W.R., 2018. The Problem with Unipolarity. Return Geopolit. 6, 9.

Tollefsen, A.F., Strand, H., Buhaug, H., 2012. PRIO-GRID: A unified spatial data structure. J. Peace Res. 49, 363–374.

Webster, C., 2001. The Placement of Cyprus' Embassies: The Influence of Realpolitik and Culture. Cyprus Rev. 13, 103–118.

Webster, C., Ivanov, S., 2015. Placement of Romanian and Bulgarian embassies: EU aspirations and communist legacy. South-East. Eur. J. Econ. 5.

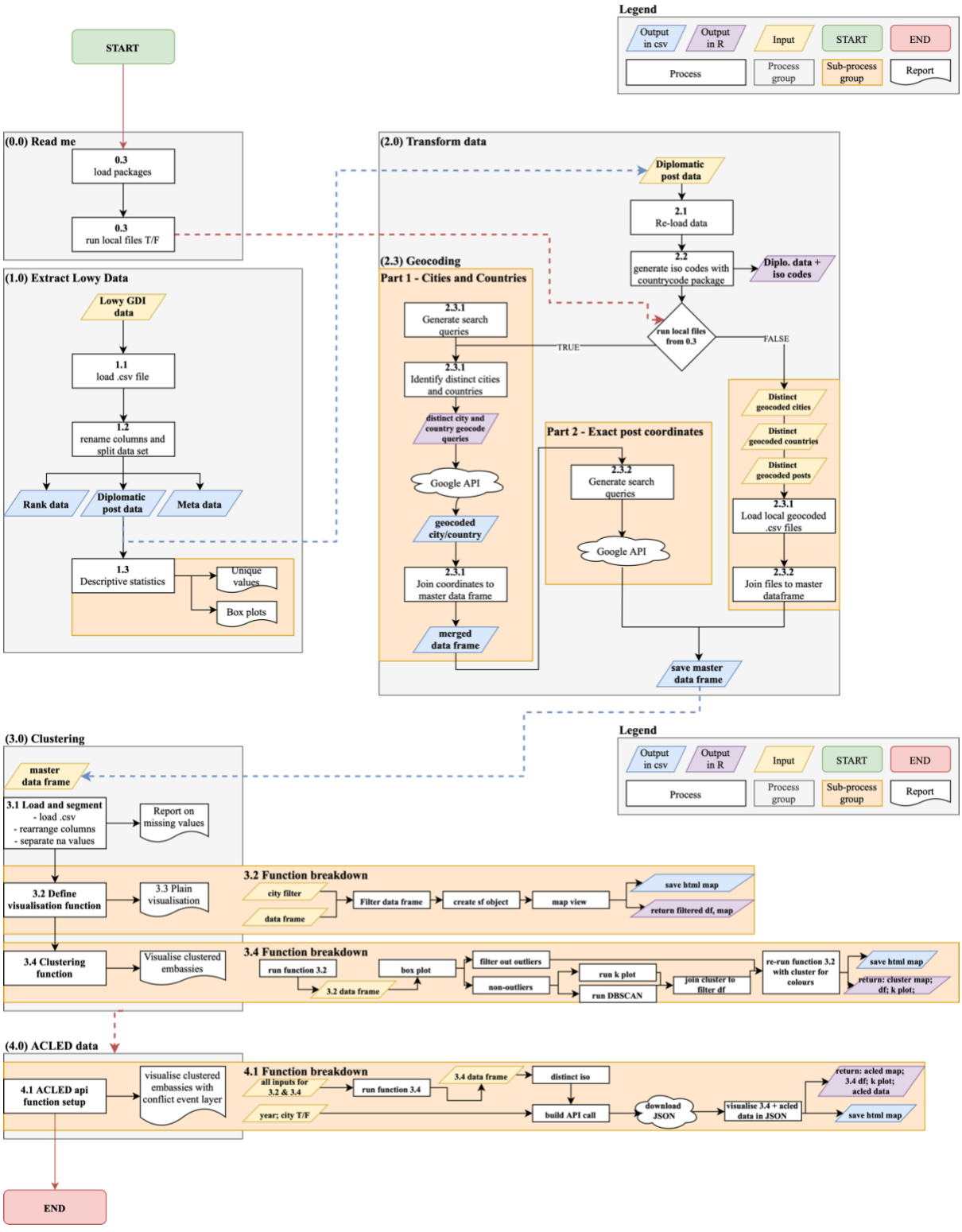
World Bank, 2020. Indicators [WWW Document]. URL https://data.worldbank.org/indicator (accessed 1.11.20).

G | Appendices

G.1 | Open Street Map vs. Lowy's GDI

OSM has a specific tag – 'office=diplomatic' – extractable via an Application Programming Interface (API). However, the number of 'node' datapoints (individual points) stands at 3,754 (Open Street Map, 2020) substantially less than GDI's 7,314. Furthermore, test extractions (not captured in example code) show a significant degree of inconsistency regarding the interpretation of what are columns [7] through to [11] in the GDI dataset due to variability in descriptions and sometimes a language barrier. Therefore, although easily accessible, the all-too-crucial metadata lacks sufficient credibility.

G.2 | Complete process map



G.3 | Code and r notebook links:

Code available here: https://github.com/antoniosfiala/UCL_CASA_GIS

R notebook published here: http://rpubs.com/antoniosfiala/casa_gis_diplomacy