**Lab 2: Node Centrality / Remittance Flows**

NAME \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

DUE \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part 1: Setting up and data wrangling**

**Step 1: Download** a matrix from the **BILATERAL REMITTANCE MATRICES** for the year 2017.

<https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data>

1. What is a remittance? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Step 2**: Open the Bilateral Remittances Matrix in Excel.

1. What is a matrix called when it has many 0s? Is this one of them? What type of matrix format is this? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. What would you have to delete / change to make this data more usable? Describe what field this data is used for—and the (a) challenges and (b) benefits to having the data in this format with these colors, tabs, etc.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Step 3:** A cleaned network has been provided. Open **remittances2017edgelist.csv** in Gephi or any of your software platforms. Compute node characteristics (degree, weighted degree, in-degree, out-degree, in-weighted, out-weighted and betweenness centrality (maybe weighted betweenness centrality, if possible).

**GRAPHIC 1: Display your network with interesting size and color symbology that shows a pattern. Include a caption and labels.**

1. Which type of centrality makes more sense here, betweenness centrality or closeness centrality? Why? What are the drawbacks for using betweenness centrality?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Part 2: Roles in the Global Economy**

**Step 1:** Export a table with *degree, weighted degree, in-degree, out-degree, in-weighted degree, out-weighted degree and betweenness centrality*. Call it ***Remittance2017netvar.csv*** **(or txt).**

**Step 2:** Open ArcMap or whatever mapping software you want to use (R, Python are find) and add the ***WorldCountries[[1]](#footnote-1)****.* Create points from the countries. These points are called CENTROIDS. In ArcMap you can do this by going to *Data Management 🡪 Features 🡪 Feature To Point*. Call these points **C*ountryCentroid*s**. It is probably a good idea to check that you want centroids to be inside the polygons.

**Step 3:** *Add in* ***Remittance2017netvar.csv****.* Join *Remittance2017netvar.csv* to *CountryCentroids on* Country.Put in a basemap as well. You can leave this alone for now—but you will make graphics 2 and 3 on your own.

**GRAPHICS 2 and 3: Make two maps. One of incoming and outgoing degree (same map, different circle sizes) and one of incoming and outgoing weighted degree (same map, different circle sizes). Hint: layer as incoming, then another layer as outgoing, then incoming again on top so no circle gets buried. ☺ Or use transparency.**

Optional: In your attribute table, create a new column and calculate the ratio of incoming to outgoing degree (InToOut); and incoming to outgoing weighted degree (InToOutWeight). Choose a few countries with the biggest disparity in *either* variable---you can choose. Who contributes way more than they take and who does the opposite? This will help you in the next part.

**Part 3: Data analysis on the topic of *Economics, Population and the Global Economy.* Choose which analysis you would like to do:**

**Option 1) Analysing global trends.**

**Are countries behaving how we would “think”, given population and economic stature?**

**Step 1:** Use your joined table of ***Remittance2017netvar.csv*** and ***CountryCentroids,*** to compare network variables to situational (gross domestic product and population) variables.

**Step 2:** Regress each network variable: degree, weighted degree, in-degree, out-degree, in-weighted, out-weighted and betweenness centrality, one at a time, with GDP, and then again with POPULATION (called POP) in the table. You will have up to 14 plots.

**Step 3:** For any plots that show correlation, paste the scatterplots, regression lines, and R squared correlations here. You can also use other methods besides R squared and scatter plots (multi-level regression, decision tree, etc). Your method must tell us whether GDP or population, both, or neither, tends to drive this network system.

**Step 4:** Write about ~100 words on the correlations in general (point to your plots), and ~150 words on interesting anomalies (countries that do not follow the trend) and why you think they do not follow the trend. Make sure you name specific countries. You can write more if you’d like. **TURN IN YOUR PLOT(S) AND TEXT.**

**Option 2) Zooming in on a case study.**

**What real-world case are you interested in?**

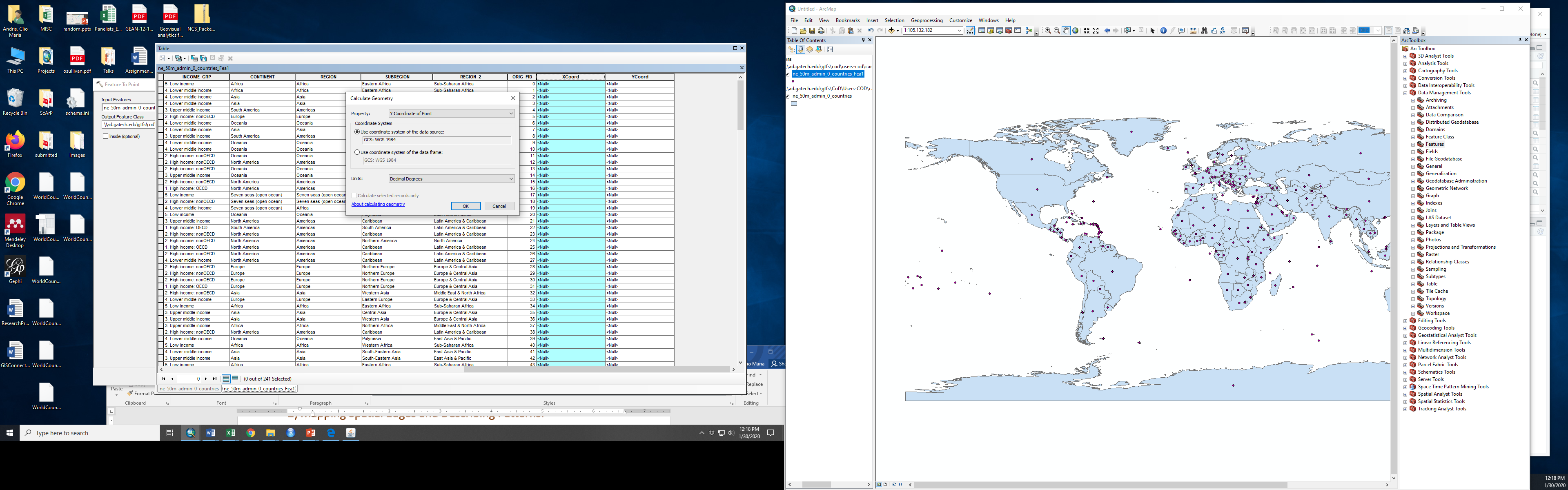
**Step 1:** Map your edges. To map edges, you can do this in many ways. Here is ArcMap help:

1. Add coordinate information to *CountryCentroids* bygoing to Data Management 🡪 Features 🡪 Add XY Coordinates. (Please note that you can also do this directly in the attribute table with Calculate Geometry—see figure below).
2. Export this table TWICE as a DBF. Call one *CountryCentroids\_Origin.dbf* and then do it again and call it *CountryCentroids\_Destination.dbf* and bring both versions back in.
3. Add *remittanceedgelist2017.csv*into ArcMap. Do a double join by joining both *CountryCentroids\_Origin.dbf* and *CountryCentroids\_Destination.dbf*to this *remittanceedgelist2017.csv* based on Source (matches Origin) and the Target (matches Destination), respectively. You’re now giving a coordinates to the edgelist.
4. I would export *remittanceedgelist2017.csv as a DBF here to make it more stable. Now call it RemittanceEdgelist.dbf.* Bring*RemittanceEdgelist.dbf* back in and go to Data Management 🡪 Features 🡪 XY to line. Use the origin and destination as your origin lon lat and your destination lon lats. You should now have edges! Export this and call it *RemittancesEdges.shp*. Now there are lots of edges! You can give the edges arrows to show directionality.

**Step 2:** Subset your data by a country or countries that interest you. *(Countries with changing economies? Countries with crises, refugees, war? Muslim countries? Tourist Countries? Caribbean? Middle East? East Asian countries? Scandinavian Countries?)* Remember to select incoming AND outgoing flows, if those are both part of your description.

**Step 3:** Make one or more maps of the flows.

**Step 4:** Tell a story about the data in ~500 words. Answer the following questions. A) Why you chose that country (or set of countries)? B) How many countries do they send to and how many do they receive from? C) To/from where are they contributing or taking from the economy? *Is that surprising or what you thought might occur?* (In your interpretation, are there few countries they are dependent on/depends on them, or a variety?) D) How much are they sending/receiving? E) Do they receive and send from far, or close? (Hint: like in the last lab, you can calculate the distances of all the edges!)—many continents or mostly neighbors? *Is this surprising or what you thought might occur?* F) **For** **5 Bonus Extra Points:** What is the total distance that each country is sending/receiving to/from? For A-F, try to tie your answers into why you chose the country/countries. **TURN IN YOUR MAP(S) AND TEXT.**



**EXTRA: FIGURE THAT SHOWS COMPUTING COORDINATES.**

1. Country data from: NAICS. <https://naciscdn.org/naturalearth/50m/cultural/ne_50m_admin_0_countries.zip> [↑](#footnote-ref-1)