README

This README serves as a general instruction of how the codes are organised and how to replicate our results.

Important Remarks

1. The total run time is extremely long due to the air routing and online data gathering. The total run time is approximately 4/5 days to execute all codes. Only the air routing itself takes about 3 days. The code has some cached results such that you are able to generate quickly a look-a-like data set. However, we haven't cached the air routing result. If you want to check this, rather use the sample option.
2. We initially thought the data preparation part is not essential due to misinformation from the slides. Despite our best effort to reorganise all files, given the short time period before submission (~ 1 week before it), we do not have time to ensure total replication because of the remark #1. However, the code contains every step that we did.
3. We comment the code as best as possible to make it is understandable. However, there are still more than 2,000 lines of code in R and Python.
4. Please adjust the working.dir in the beginning of the Stata (Analysis.do) and R files (Index.R, Map Distribution Lenders.R, Visualise Flight Routing.R)

Run Order / Structure

1) Data Preparation:

{MAIN_FOLDER}/Index.R → See comments

2) Data Analysis:

{MAIN_FOLDER}/Analysis.do

3) Non-Essential Graphs:

{MAIN_FOLDER}/Map Distribution Lenders.do → Figure "Lender Distribution"

{MAIN_FOLDER}/Visualise Flight Routing.do → Figure "Flight Connection Routing"

4) Data Set:

{MAIN_FOLDER}/Data Set.dta

Python

All the Python codes can be ran with a normal compiler. However, we strongly suggest using Jupyter Notebook, which is an extension that comes with Anaconda to run the notebook file, namely ipynb.

The reason is because Jupyter Notebook has an interactive interface which allows users to interact with certain cell of code, plus one can add additional explanation in Mark-ups between cells. Within each notebook, there are already examples of how the function is called and also the time for running the function.