**Codes/Python**

1. *code\_download\_oss\_selenium.py* This is the code to extract and download html files from openstack gitweb webpage. The code is commented with explanations for each definition.
2. *create\_analyze\_network.py* This code creates network of developers. The code is commented at various stages explaining the steps. In our first idea, we created the network of developers if the work on the same problem in Nova. In the 2nd idea, we took the same approach but this time for every time bin of 2 months. In the 3rd idea we constructed a weighted network of developers where two developers are linked if they committed on the same code. Now, two developers could be separated in time (time of commit) or separation in line number of codes during commit. We introduce a restriction on the link formation – two developers are linked if they commit on same code and also the commit lines overlap. The temporal weight is proportional to the difference in time of commit and spatial weight is proportional to difference in separation of lines.
3. *code\_data\_analysis.py* This code is for analyzing and processing data for regression purpose.

**Codes/Stata**

1. *imputation\_code\_DV\_ratio.do* Stata code for running imputation code using Gaussian regression
2. *code\_diversity\_interactions.do* Stata code for running time fixed effect regressions with explanatory variables being diversity indices – Herfindahl index and insularity index
3. *Idea03\_code\_mediation.do* Stata code implementing the Sobel-goodman test for mediation in spatial and temporal weighed networks (Idea 3)

**Data**

1. **/Stata\_files/** Contains the stata files used in the regression analysis
2. **/Idea03CSV/** contains csv files for mediation analysis using the temporal and spatial networks for years 2012, 2013, ..., 2016.
3. **/cleaned\_data\_from\_html/** Here we store the text formatted “|” separated data obtained from gitweb pages and gitlog of Nova package.