**Atlanta SeeClickFix Project**

**CS 109 Final Project**

**Olalekan Sodeinde, Jon Keen, Matt Lutze, and Adele Faure**

Dear reader:

We divided our **process notebook** into five parts representing the different aspects of our analysis:

1. Introduction and exploratory data analysis: “1. SeeClickFix.ipynb”
2. Network analysis: “2. SeeClickFix\_Network.ipynb”
3. Classification analysis (status): “3. SeeClickFix\_Classification \_Status.ipynb”
4. Classification analysis (type): “4. SeeClickFix\_Classification \_Type.ipynb”
5. Regression analysis: “5. SeeClickFix\_Regression.ipynb”

Please start with Part 1, which introduces the project and the data we are using.

We use four **datasets**:

1. “SeeClickFixData\_Primary”: our primary dataset, which lists each report submission by users of SeeClickFix, as well as the incident location, user id, and description.
2. “SeeClickFixData\_Issue\_History\_ATL”: all the comments made by users for each report submitted.
3. “SeeClickFixData\_Users\_ATL”: for each SeeClickFix user, the number of issues they have commented on, chosen to follow, and closed, their number of “civic points,” and whether they are authorized to acknowledge other reports.
4. “SeeClickFixData\_GISExport”: this spreadsheet connects the report submissions to their geographic location, namely city and neighborhood (if in Atlanta).

To visit our **website** and view our screencast, please go to: <http://seeclickfix-atlanta.weebly.com/>

If you have any questions, please email us at [cs109-atlanta-project@googlegroups.com](mailto:cs109-atlanta-project@googlegroups.com).

Thank you,

Lekan, Adele, Matt, and Jon