

Exploring User Behaviors in iNaturalist City Nature Challenge: Analysis of Crowd Participation

1. INTRODUCTION AND MOTIVATION

- Citizen science communities evolve over time. Using iNaturalist platform, City Nature Challenge as a showcase (CNC) we aim to identify common patterns and drivers that shape communities
- The research aims to understand and compare citizen science communities in different years based on the City Nature Challenge
- Research project analyzes patterns of engagement of users in London, San Francisco (SF), and Los Angeles (LA), with a focus on how user behaviors evolve over time (2018-2020)
- We look into the interaction network representing each community using the network analysis, combining this with statistical data analysis

2 HYPOTHESES

Participants' activity types in cities differs from year to year with main variations in the frequency and types of contributions made by users during the City Nature Challenges.

3. DATA

Data: <https://www.inaturalist.org/>
Format: CSV and JSON files | 9 files
Information: 200,000 + rows , 34 features: id, observation, identification counts, year, location etc.

4. METHODS & TOOLS

- Literature review
- Data Gathering
- Data preprocessing
- Data exploration
- Statistical analysis
- Clustering
- Data visualization
- Data interpretation
- Open Code Development



5. PRELIMINARY RESULTS /DISCUSSION

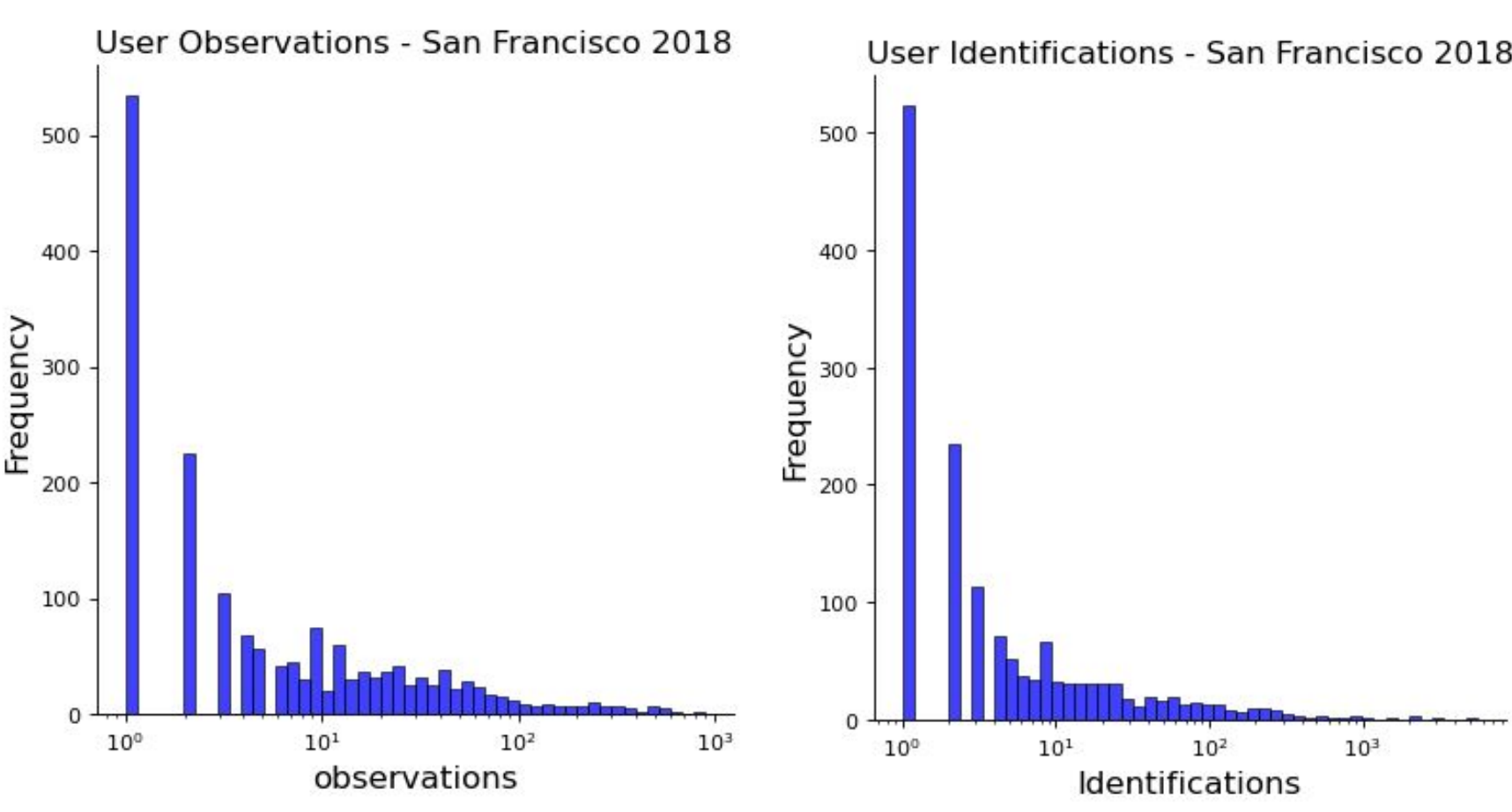


Fig 1: General overview of observations and identification made in year 2018. We can observe majority of the data is located on the right side and the tail of the distribution is skewed to the right. Yet here we do not assess the temporal aspects of contribution changes over time.

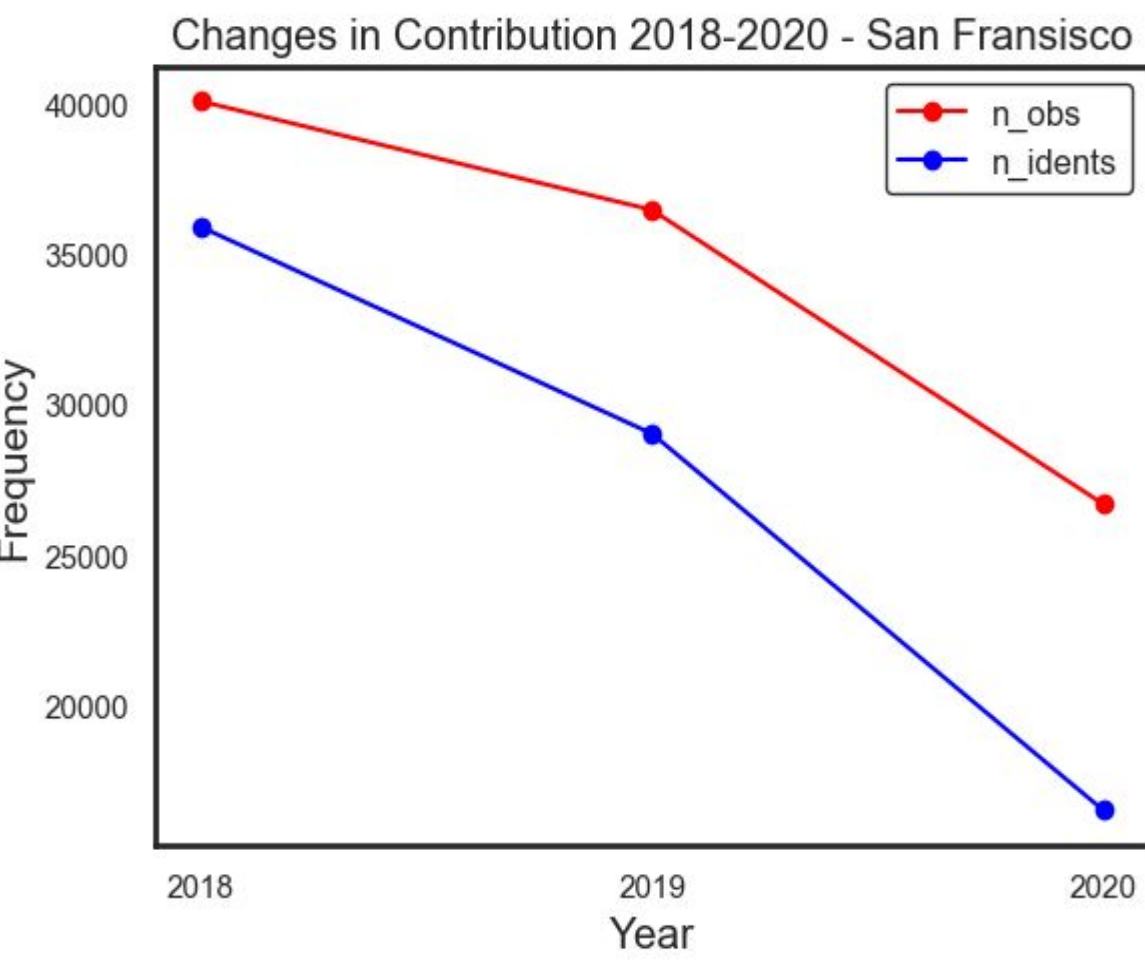


Fig 2 : Cumulative contributions made by users have decreased over 2018-2020.

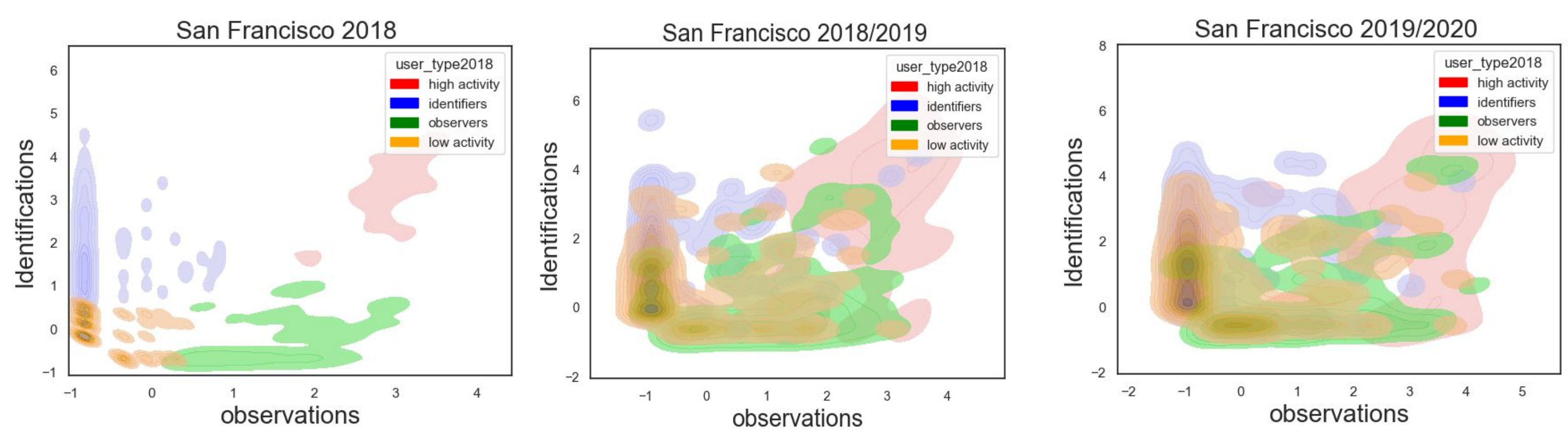


Fig 3 : Kernel Density Plot for K-Means clustering applied to users data in San Francisco. Each dot represents a user with identification and observations count during 2018. We analysed the temporal evolution of participants' activity types in cities. We see a change in user activity types. We observe a dispersion and changes of activity types throughout the years: from year 2018 to 2019 low activity type users are becoming observers and identifiers, same applies to high activity users in 2018, they no longer contribute extensively and are becoming identifiers or observers.

Users' interactions in a social network

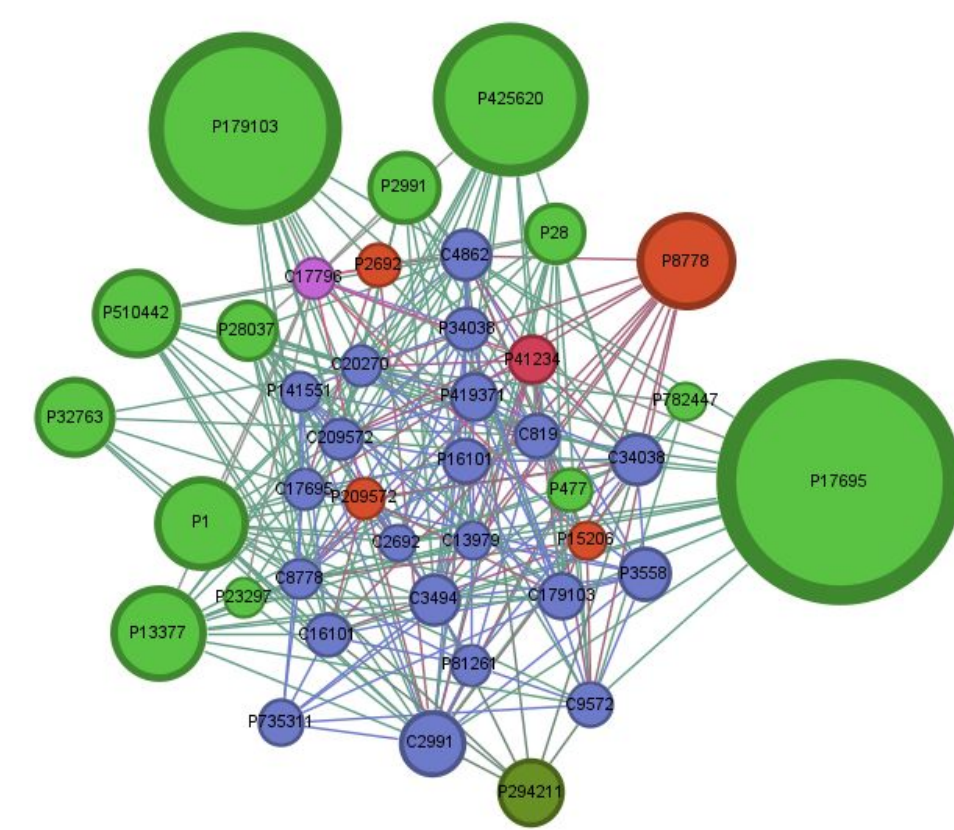


Fig 4 : The interaction among the users on the platform. Nodes represent users on the platform and Edges correspond to the identification-interaction of users' observations. The modularity classes [4] correspond to the communities of users who are likely to make identifications of observations of each other.

6.CHALLENGES

- Data Preparation and selection of appropriate analysis techniques
- Handling large data
- Interpreting results, discussions with iNaturalist community
- Theory vs Practice

Acknowledgement: We would like to thank Muki Haklay for discussions, LPI - for the support of the current project.

REFERENCES

1. Tupikina, L., Schlosser, F., Voskresenskii, V., Kloppenborg, K., Lopez, F., Mariz, A., Mogilevskaia, A., Haklay, M., & Tzovaras, B. G. (2021). iNaturalist citizen science community during City Nature Challenge: New computational approach for analysis of user activity (arXiv:2112.02693). arXiv. <https://doi.org/10.48550/arXiv.2112.02693>
2. Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for Biodiversity Science. BioScience, 71(11), 1179–1188. <https://doi.org/10.1093/biosci/biab093>
3. Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. Scientific Reports, 11(1), 3037. <https://doi.org/10.1038/s41598-021-82404-z>
4. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>

7. LEARNING

- Industry-specific knowledge: Citizen Science
- Understanding data analysis workflows
- Hands-on experience with data analysis tools
- Critical thinkings & Continuous learning

8. FUTURE WORK

- Analyze the social network of other cities and years
- Geo Spatial Analysis
- Identify more feasible analysis to get insights of citizen science data
- OCD : https://github.com/Liyubov/inaturalist_internship