



UBID for CEC

Mark Borkum
Computer Scientist



PNNL is operated by Battelle for the U.S. Department of Energy



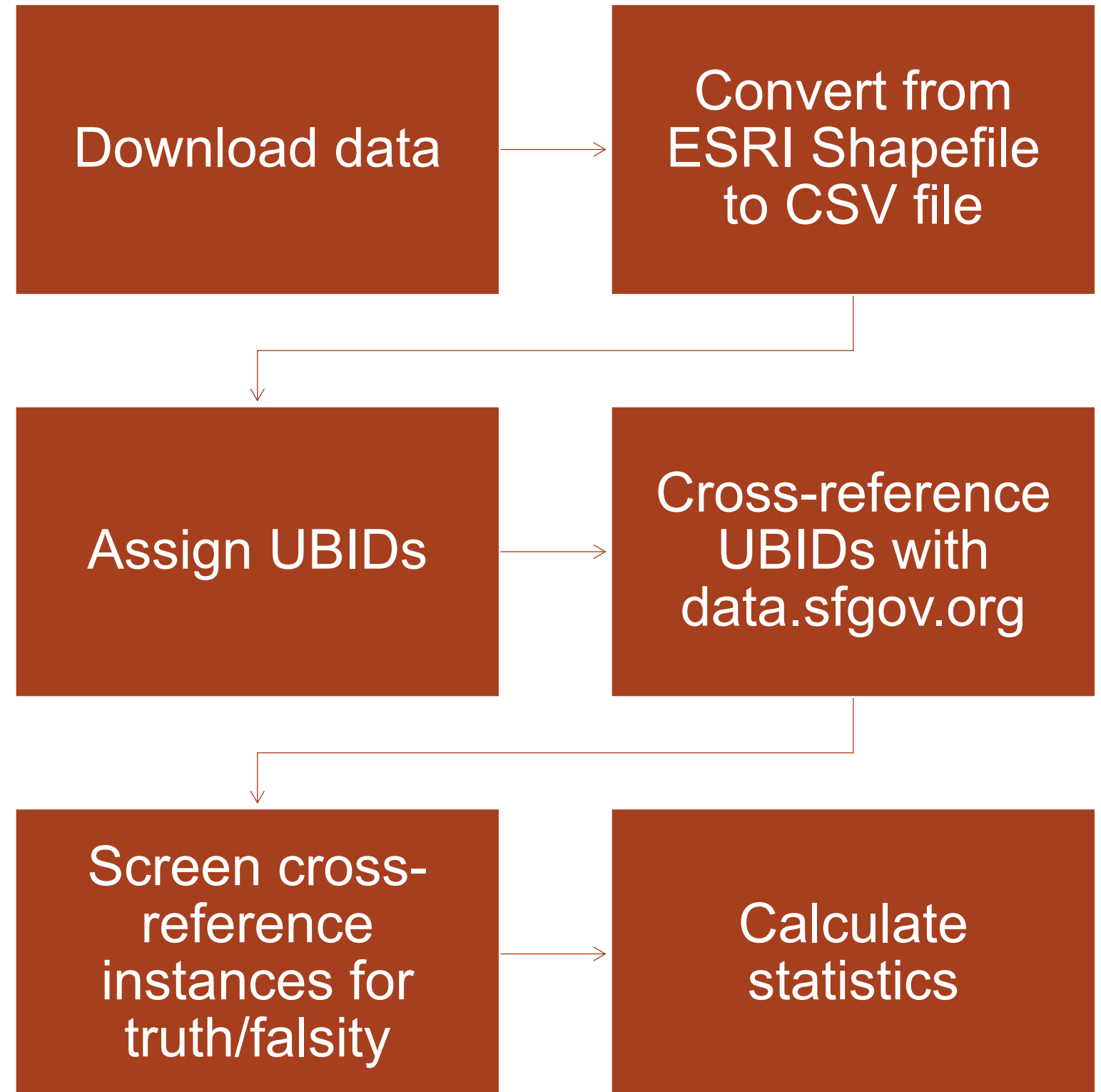
- CEC_Buildings_shapefile
 - UBID_Berkeley
 - UBID_LosAngeles
 - UBID_SanFrancisco ← Used for this study
 - UBID_Statewide1
 - UBID_Statewide2

Data Source

- ESRI Shapefile
 - Geographic coordinates
 - WGS-84

Methodology

- Open source:
 - UBID assignment
- Closed source:
 - Cross-reference UBIDs



Cross-reference

- Takes two files as input:
 - LHS = CEC
 - RHS = data.sfgov.org “2s2t-jwzp” dataset of building footprints
- Decide truth/falsity by comparing MBLRs
- By intersection of bounding boxes:
 - 407/435 records matched
 - 327 true positives
 - 80 false positives
 - Unknown number of true/false negatives
 - Precision = 80.3%
 - Recall = 100% (no false negatives), 92.1% (no true negatives)
 - F1 score = 89.1% (no false negatives), 85.8% (no true negatives)
- By Euclidean distance between centroids:
 - 407/435 records matched
 - 346 true positives
 - 61 false positives
 - Unknown number of true/false negatives
 - Precision = 85%
 - Recall = 100% (no false negatives), 92.5% (no true negatives)
 - F1 score = 91.9% (no false negatives), 88.6% (no true negatives)
- Combining the results of both approaches has 100% precision for the input datasets! 😊



- CEC = 3709/011
- False +ve = SF3709011
 - By intersection
- True +ve = SF3709012
 - By Euclidean distance

[Back to top](#)



- CEC = 0292/004
- False +ve = SF0292015
 - By Euclidean distance
- True +ve = SF0292004
 - By intersection

[Back to top](#)



Thank you

