Identifying Slum Areas in Indonesia

WI26

Group Results: Progress Data

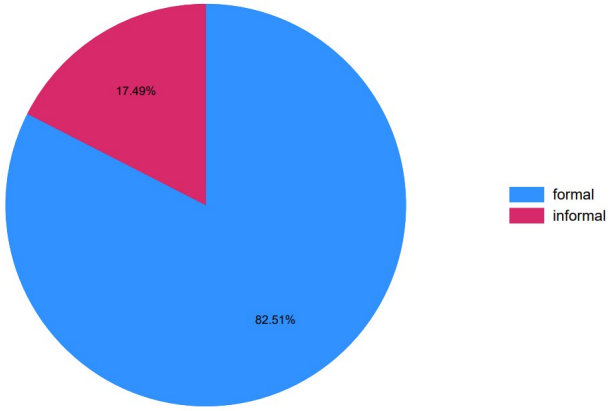
Name	City	Number of images
Santiago O.	Padang	52
Nikolai M.	Bandaar Lampung	36
Holly	Balikpapan	Google Satellite
Louis K.	Samarinda	13 S.C.C | 45 S.C.S 58 individual images 17 composite images
Fahim C.	Makassar	14
Ji M.	Palembang	14
Kamran S.	Gorontalo	17

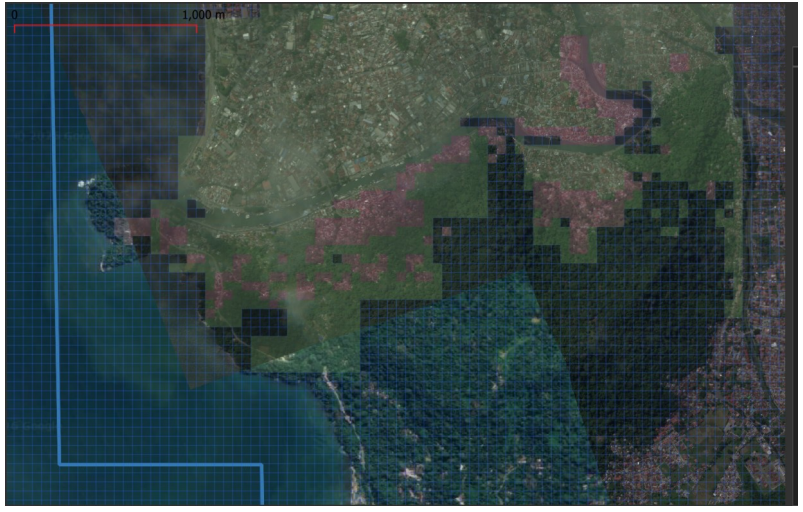
Group Results: Classification Summary Statistics

Name	Total number of 50x50m cells labeled	% formal	% informal	% Unidentified
Santiago O.	3,367	82.51%	17.49%	
Nikolai M.	3,254	71.94%	28.06%	
Louis K.	-	-	-	
Fahim C.	466			
Ji M.	1,298	0.39%	0.22%	0.39%
Kamran S.	N/A (labeled numerically)			

Screenshots of Output: Santiago O.

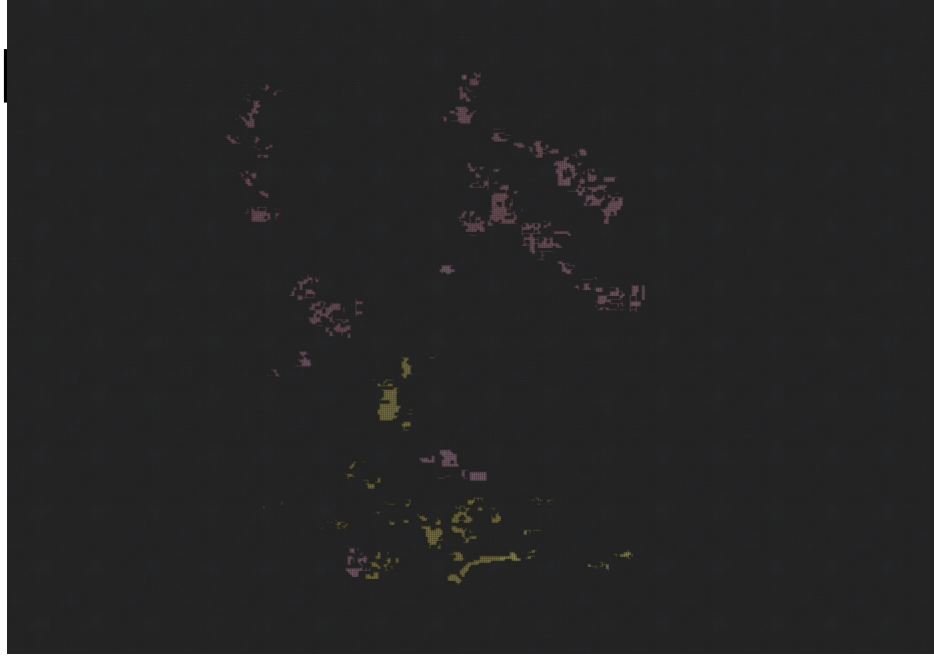
```
. tab class

      class |      Freq.     Percent        Cum.
------------+-----------------------------------
     formal |      2,778       82.51       82.51
   informal |        589       17.49      100.00
------------+-----------------------------------
      Total |      3,367      100.00

. count if class != ""
  3,367

. count if class == "informal"
  589

. count if class == "formal"
  2,778
```





Note: My sample was not truly randomly drawn, so the result should be interpreted as descriptive stats, not as an unbiased representation of the city's slum proportions.

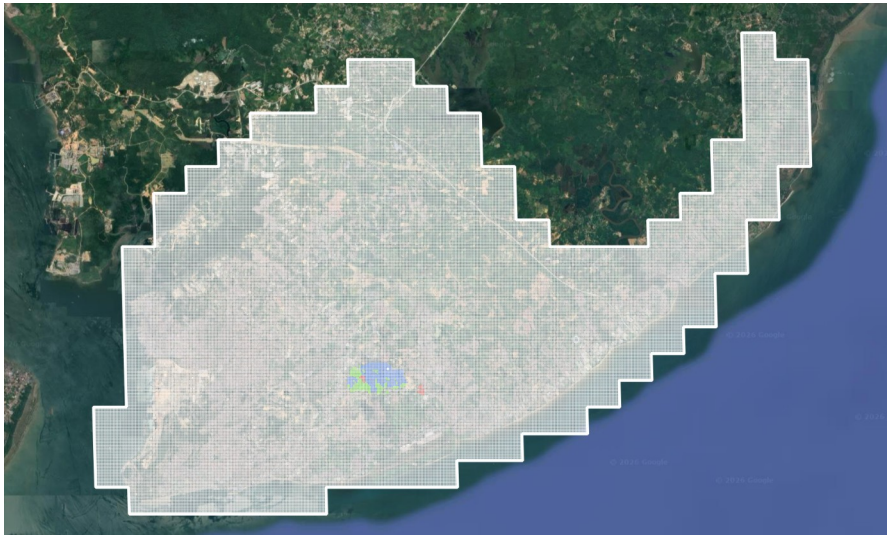
Bandaar Lampung: Nikolai

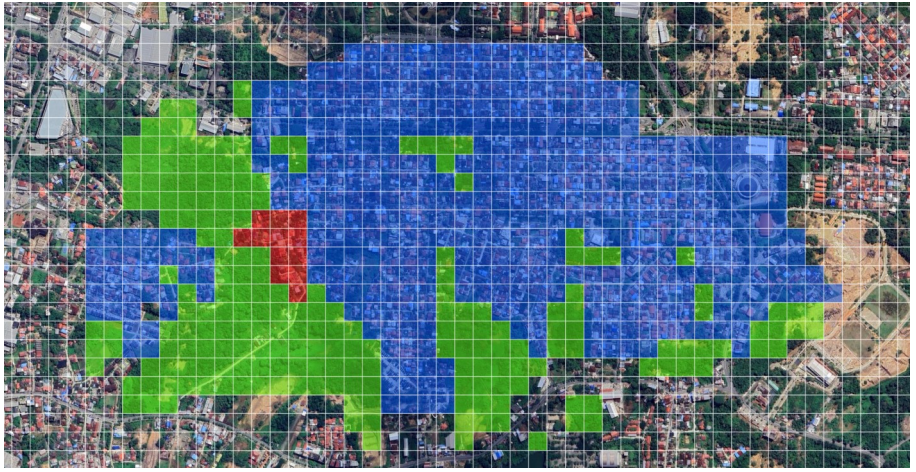


Formal	2341	71.94%
Informal	913	28.06%
Total	3254	100%

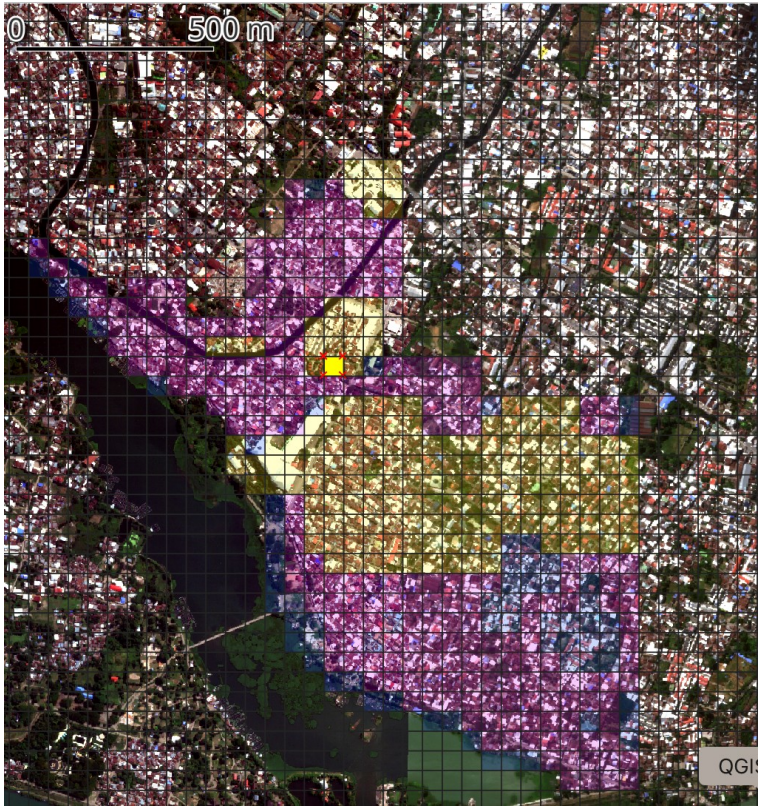
0.5% of the total city is mapped

Balikpapan: Holly Zhang



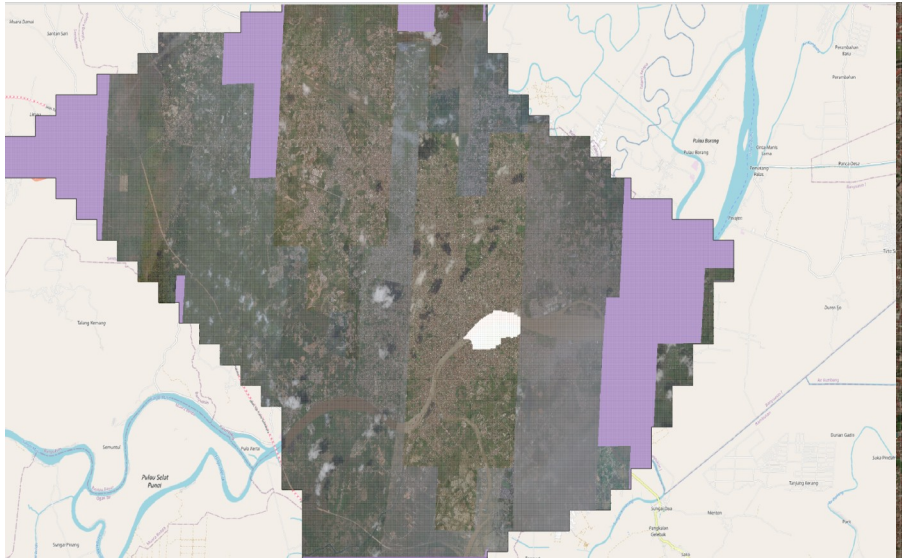


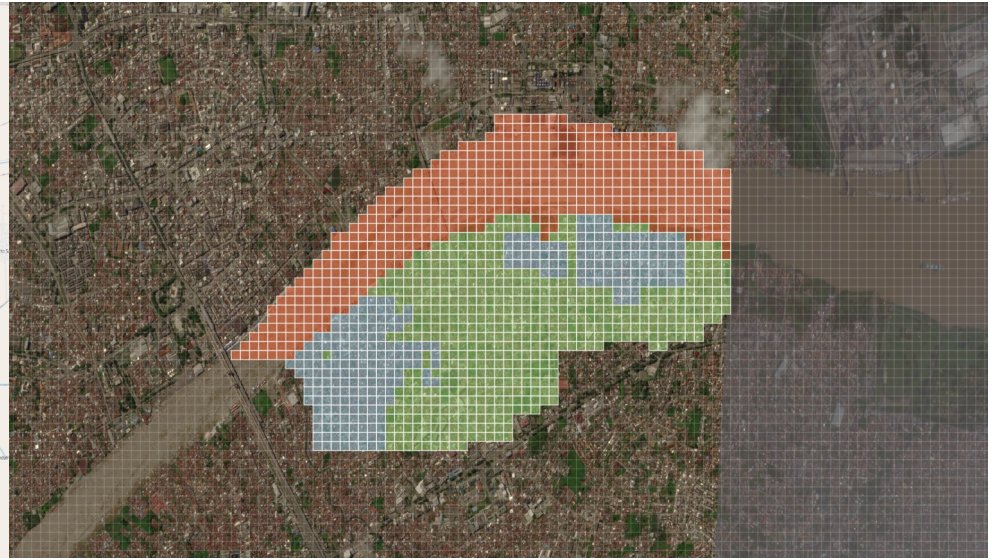
Makassar



Informal	Formal	Mixed	Total
236	164	66	466
51%	35%	14%	100%

Palembang

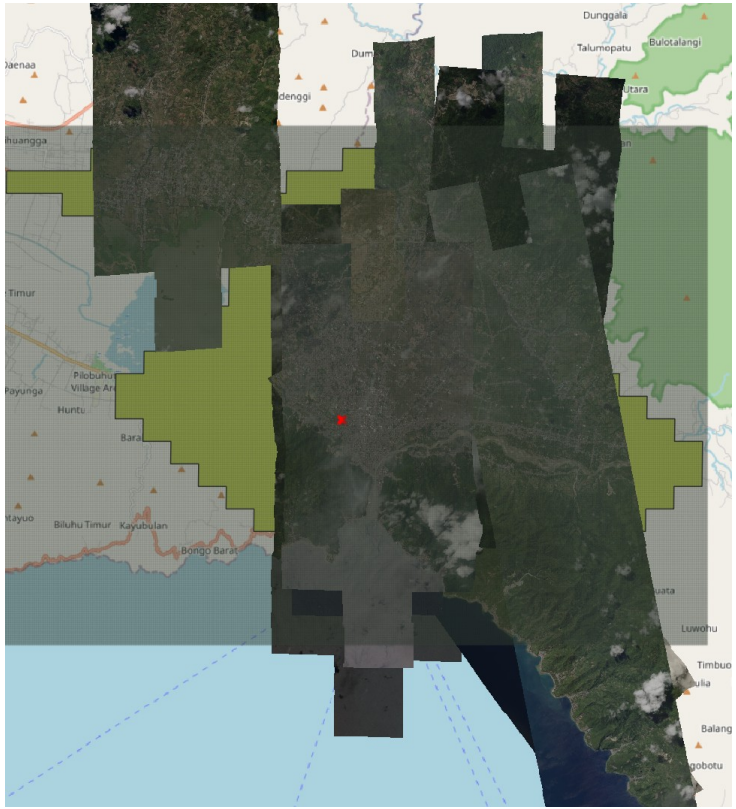




Out of 250,000 grids, 1,298 grids are mapped
Formal : 503 Unidentified : 515
Informal : 280

To estimate the proportion of population of a city, does it make sense to select the areas we map and then calculate the ratio?

Gorontalo



Trends of Slum Locations

- Located by sea/bodies of water
  WHY?
  - Land near water is often undesired, so homes are priced cheaper.
  - People living in slums often lack access to clean, prevalent water.
  - Work opportunities (ex: fishing, boat crafting, etc.)
- Steep hillsides and slopes
  - Land on slopes is undesirable due to landslides, so prices are low in such areas.

Gorontalo

Near body of water

Slum on a hill

Common Challenges

- Saving Issues in QGIS
- Labeling ambiguous areas as formal/informal
- Finding enough high resolution images to cover entire city

Link to guide - Louis

https://docs.google.com/presentation/d/1wlqHL3hwhZLUf_4iAr2Dor6iGCDF86pp4vrreePYurA/edit?usp=sharing

Suggestion? - Ji

- Why don’t we randomly select grids and then map them?
- Standardization is necessary to properly distinguish between slum, non-slum, and unidentified areas.

Ex)

- if an image contains more than 40% forest or roads, it should be classified as unidentified.
  > it happens because of the conflict between ml algorithm vs stata research. To prepare proper features for algorithm, I avoided image which contains more than 40% forest or roads. However, it might cause bias on stata research

- A slum is defined as an area where houses have little to no distance between them and are characterized by very high housing density.