

Water infrastructure in Uganda

Capstone Progress Stand up – Brainstation Data Science
July 2022 – Thomas Adler

Situation

- World not on track for SDG 6: Clean Water and Sanitation
- ~400m lack basic drinking water in sub-Saharan Africa
- ½ healthcare facilities in LIC don't have basic water services

Problem

- Life threatening
- Cycle of poverty
- Basic economic input

Solution

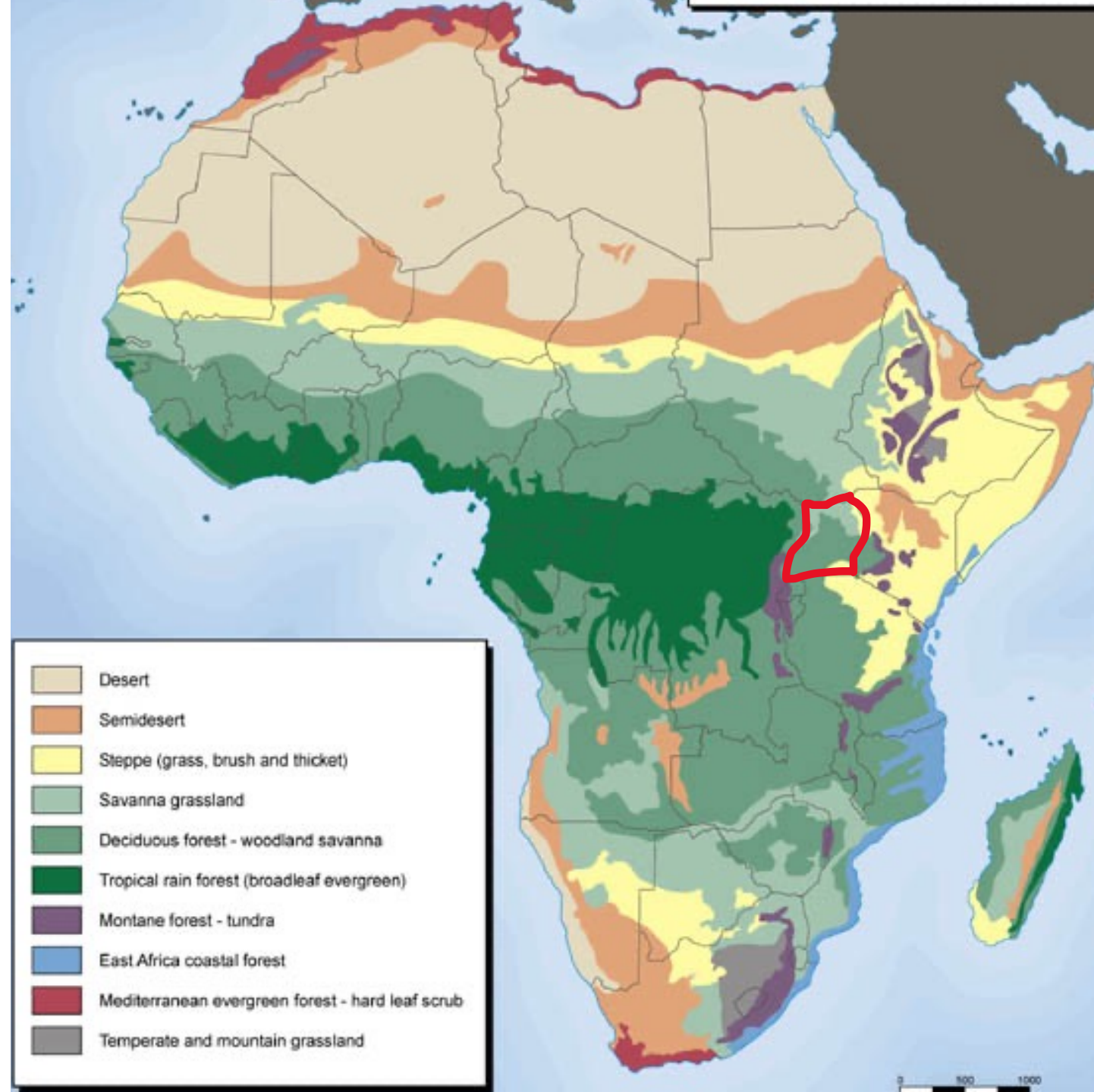
- Better water infrastructure management:
- Monitoring
- Repairing
- Improving

Can we use AI to predict the functionality of water infrastructure to repair them quicker and enable citizens to access water so that they can prosper?

Combined 3 data sources...

1. Water Point Data Exchange (WPDx - API)
2. Uganda Regional Demographics (Bureau of Statistics - manual)
3. Armed Conflict Location & Event Data (ACLED - API)

Why Uganda?



WPDx API
n=110k - 50 features

ACLED API
n=50k - 20 features

Uganda BoS
n=2k - 40 features

Cleaning

Water Point Data
n=105k - 30 features

Conflicts Data
n=2k - 6 features

Demographic Data
n=2k - 30 features

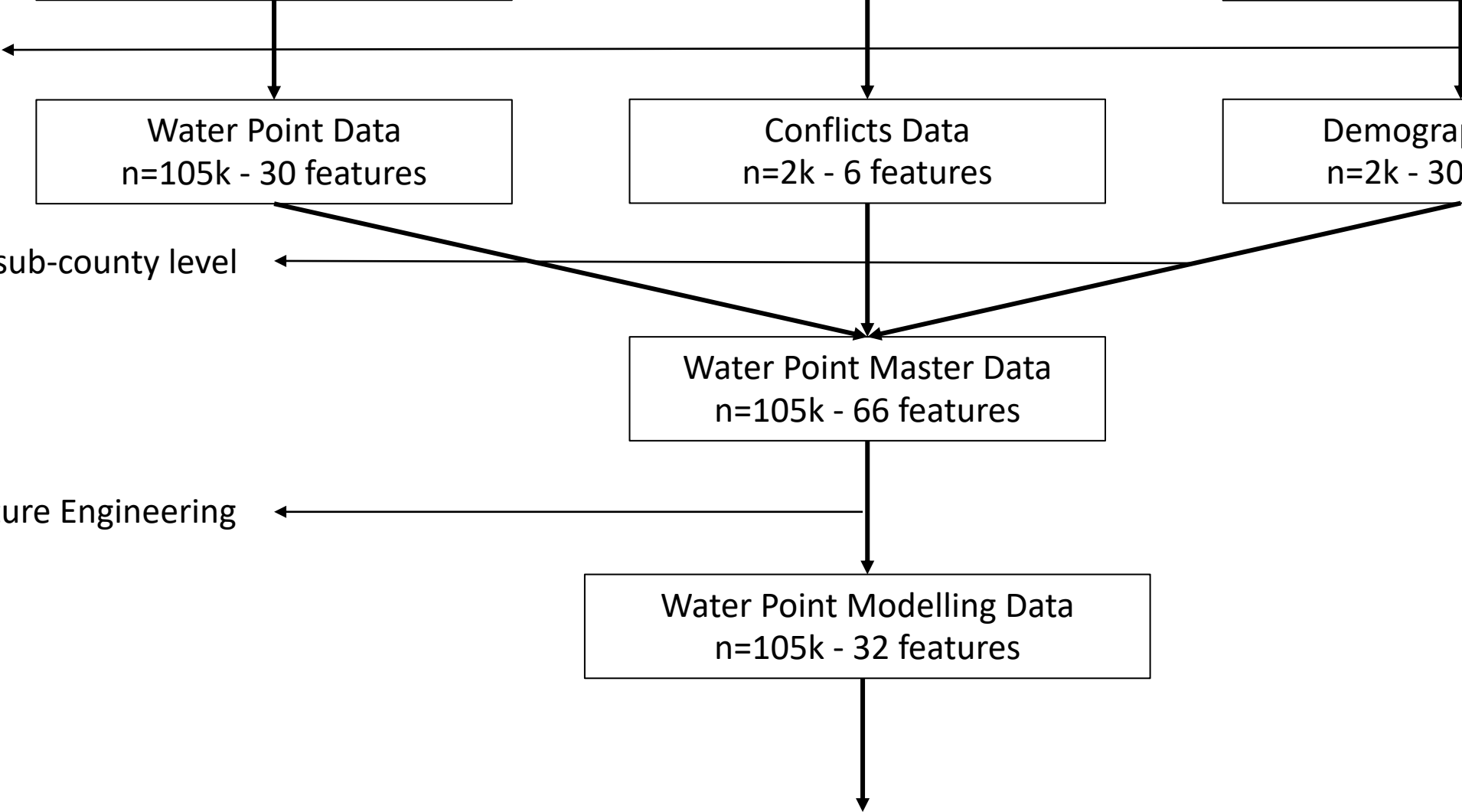
Joining on sub-county level

Water Point Master Data
n=105k - 66 features

EDA & Feature Engineering

Water Point Modelling Data
n=105k - 32 features

...



Feature Engineering resulted in these features...

Water Point

| Location |
|-----------------------|
| lat_deg |
| lon_deg |
| distance_to_primary |
| distance_to_secondary |
| distance_to_tertiary |
| distance_to_city |
| distance_to_town |
| is_central |
| is_eastern |
| is_western |

| Infrastructure |
|-------------------------|
| is_complex_tech |
| is_installed_after_2006 |
| is_public_management |
| crucialness |
| usage_cap |
| perc_local_served |

Regional

| Economic |
|-----------------------|
| perc_pop1464_working |
| perc_hh_temp_dwelling |
| perc_hh_own_house |
| perc_hh_bank_acc |
| perc_hh_electricity |

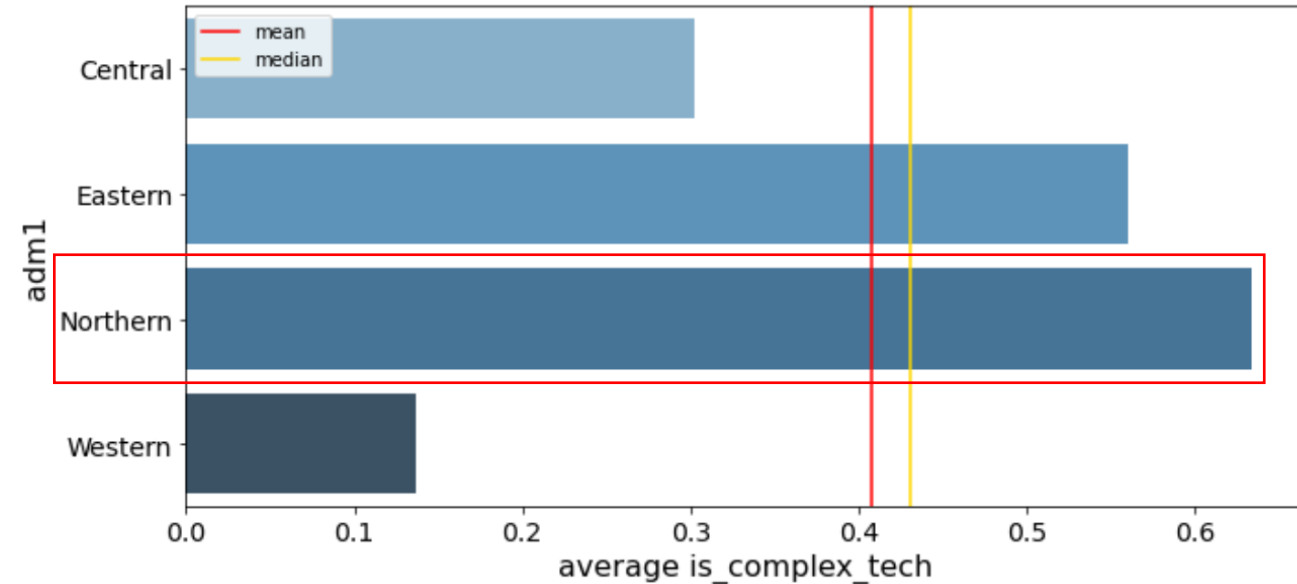
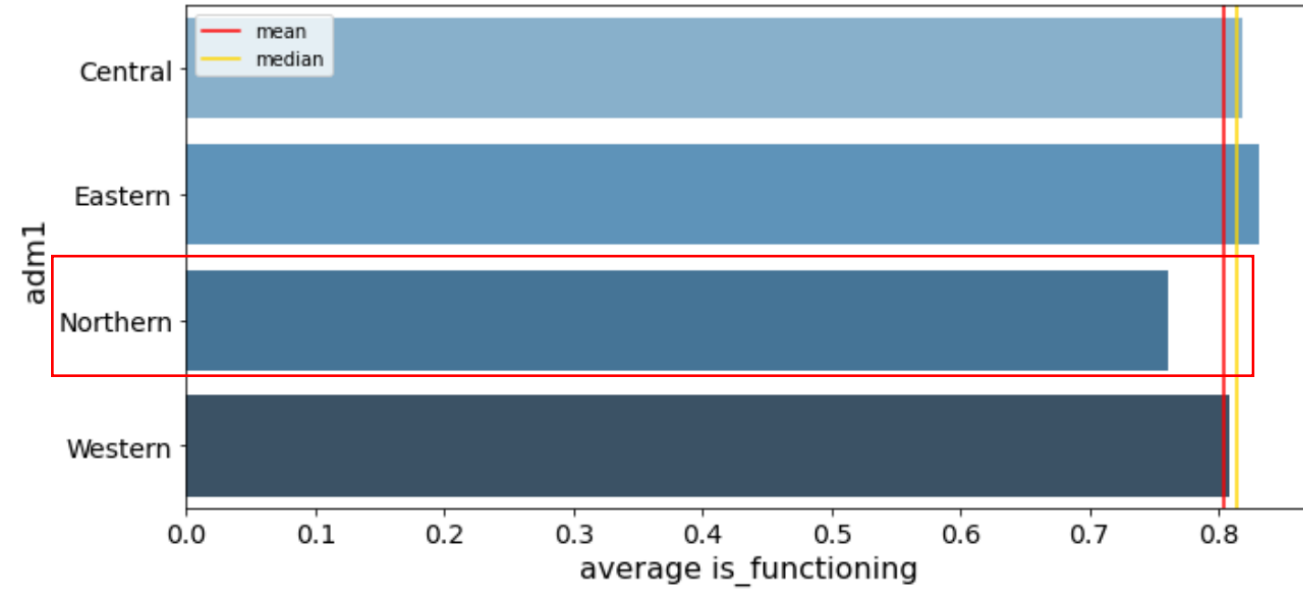
| Social |
|--------------------------|
| perc_hh_head_male |
| perc_pop017_both_parents |
| perc_pop2p_disability |
| perc_pop1017_married |
| perc_pop1217_birth |

| Public Service |
|-------------------------|
| perc_pop1318_secondary |
| perc_pop017_certificate |

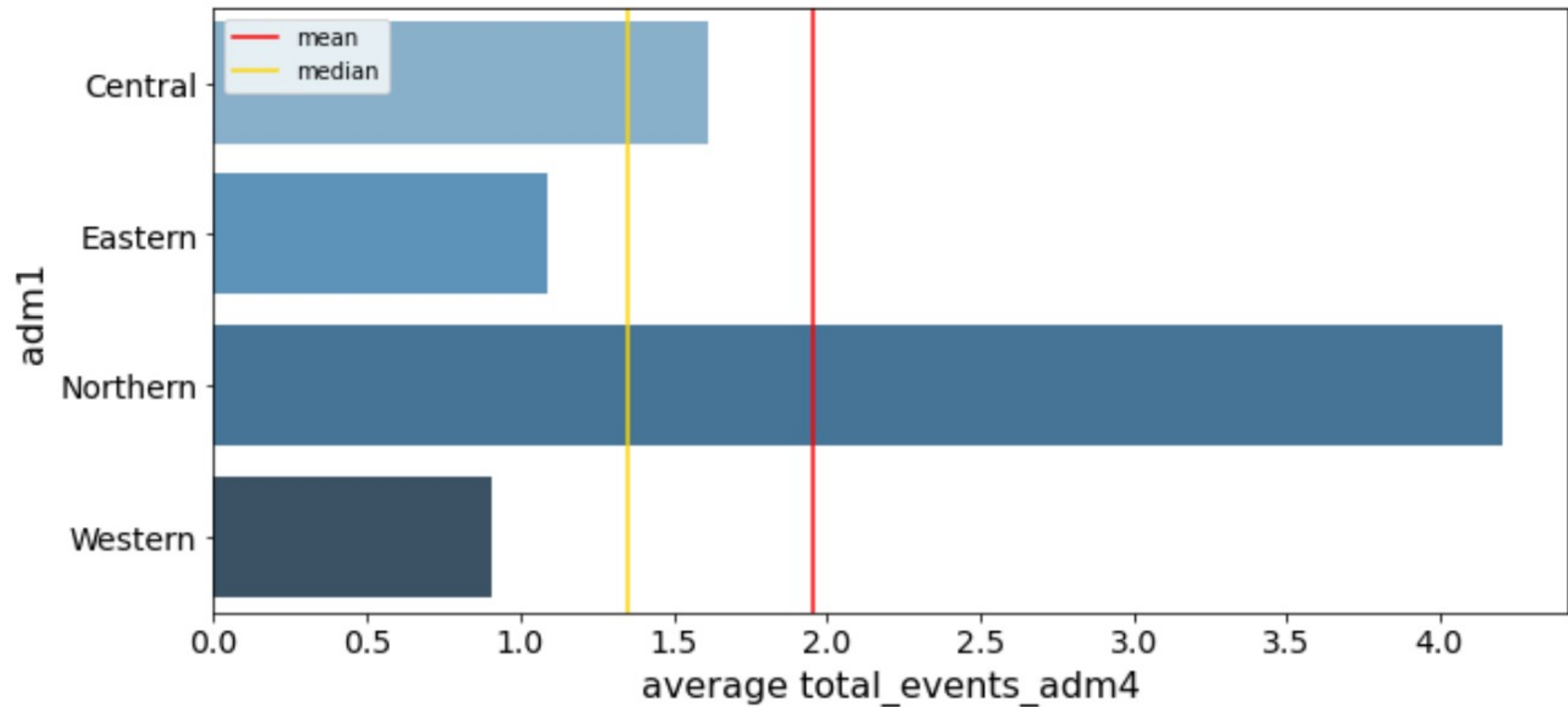
| Health |
|----------------------|
| perc_hh_mosquito_net |
| perc_hh_toilet |

| Conflict |
|-------------------|
| total_events_adm4 |

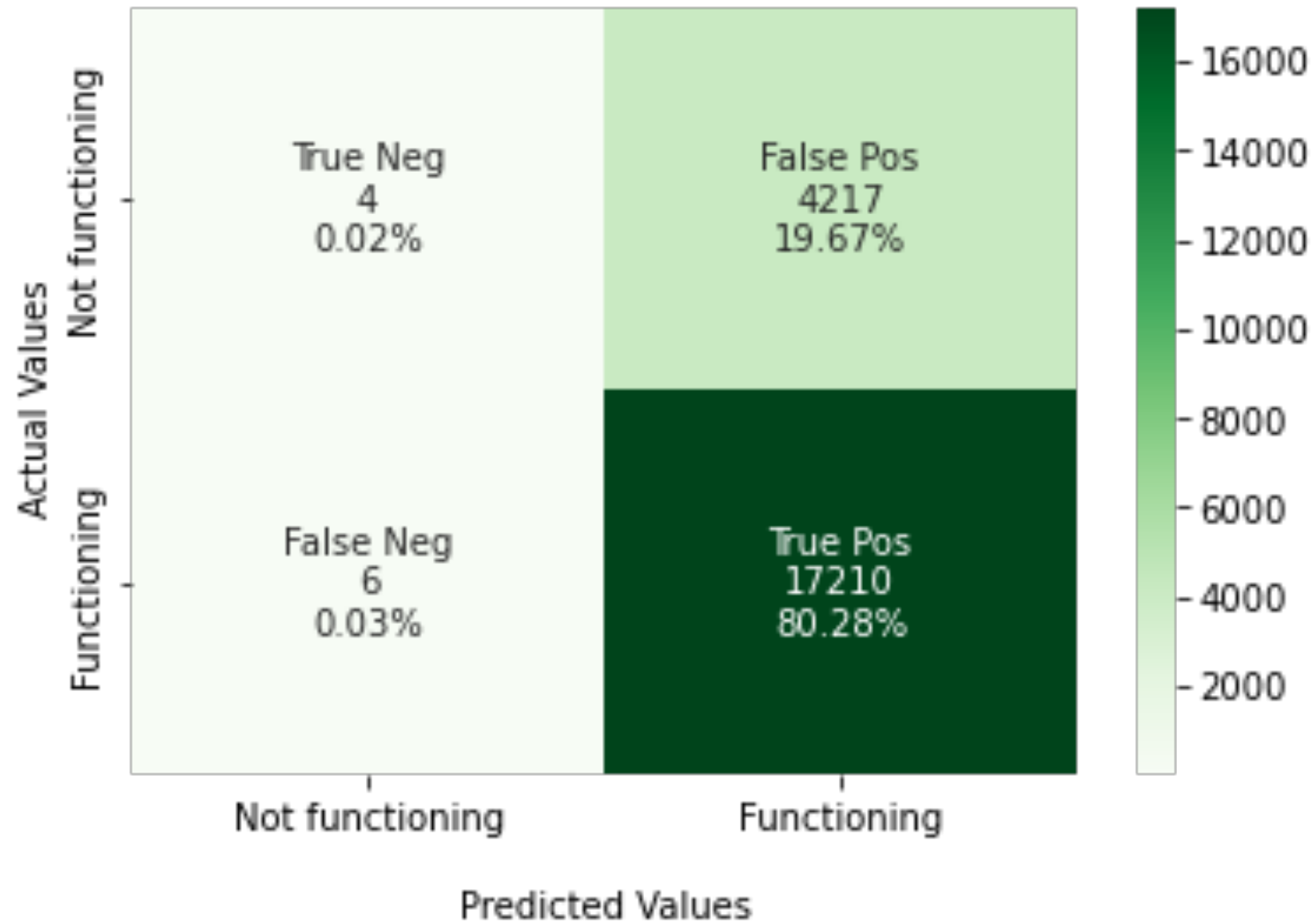
EDA - Do modern technologies break down more easily...



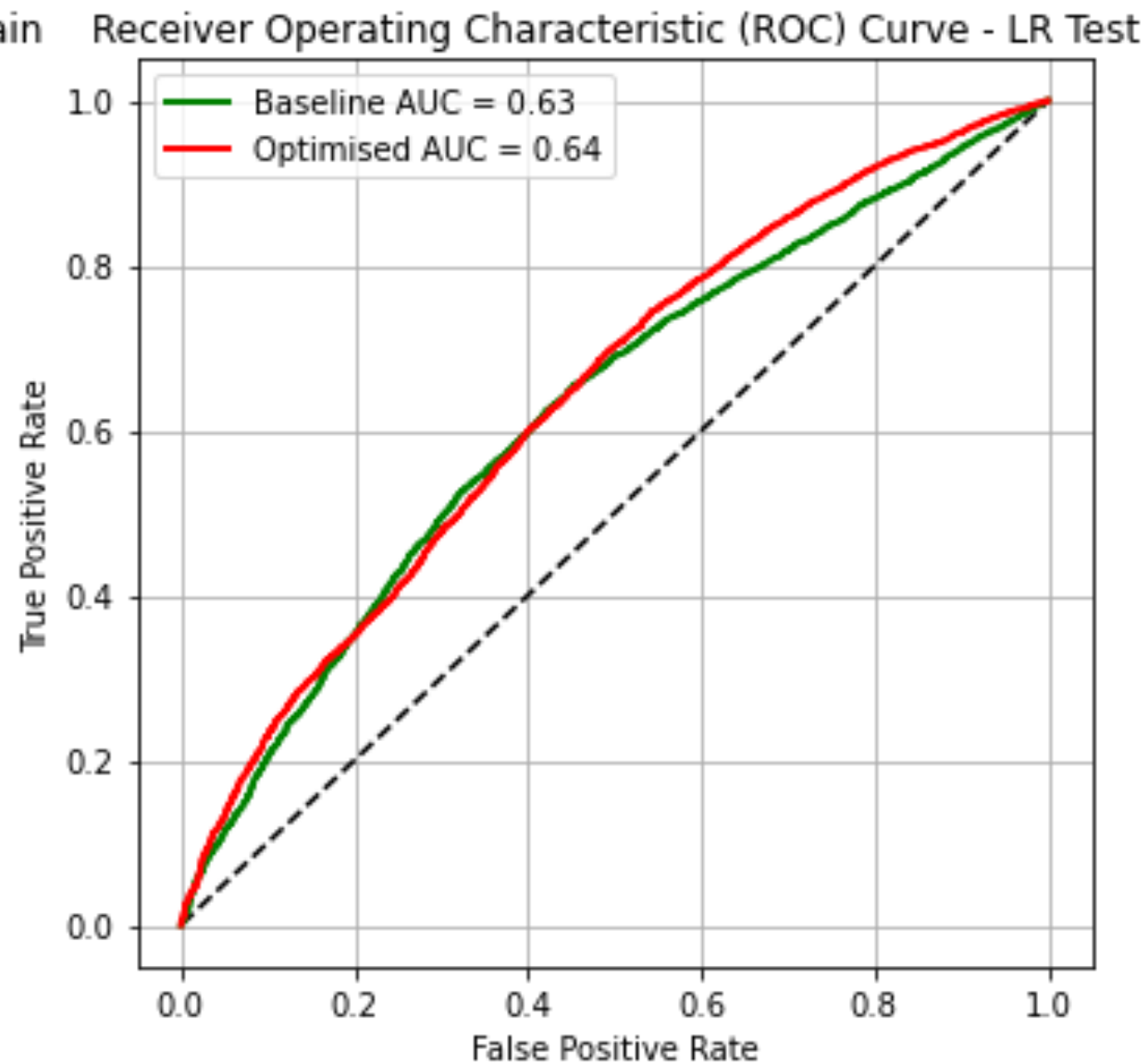
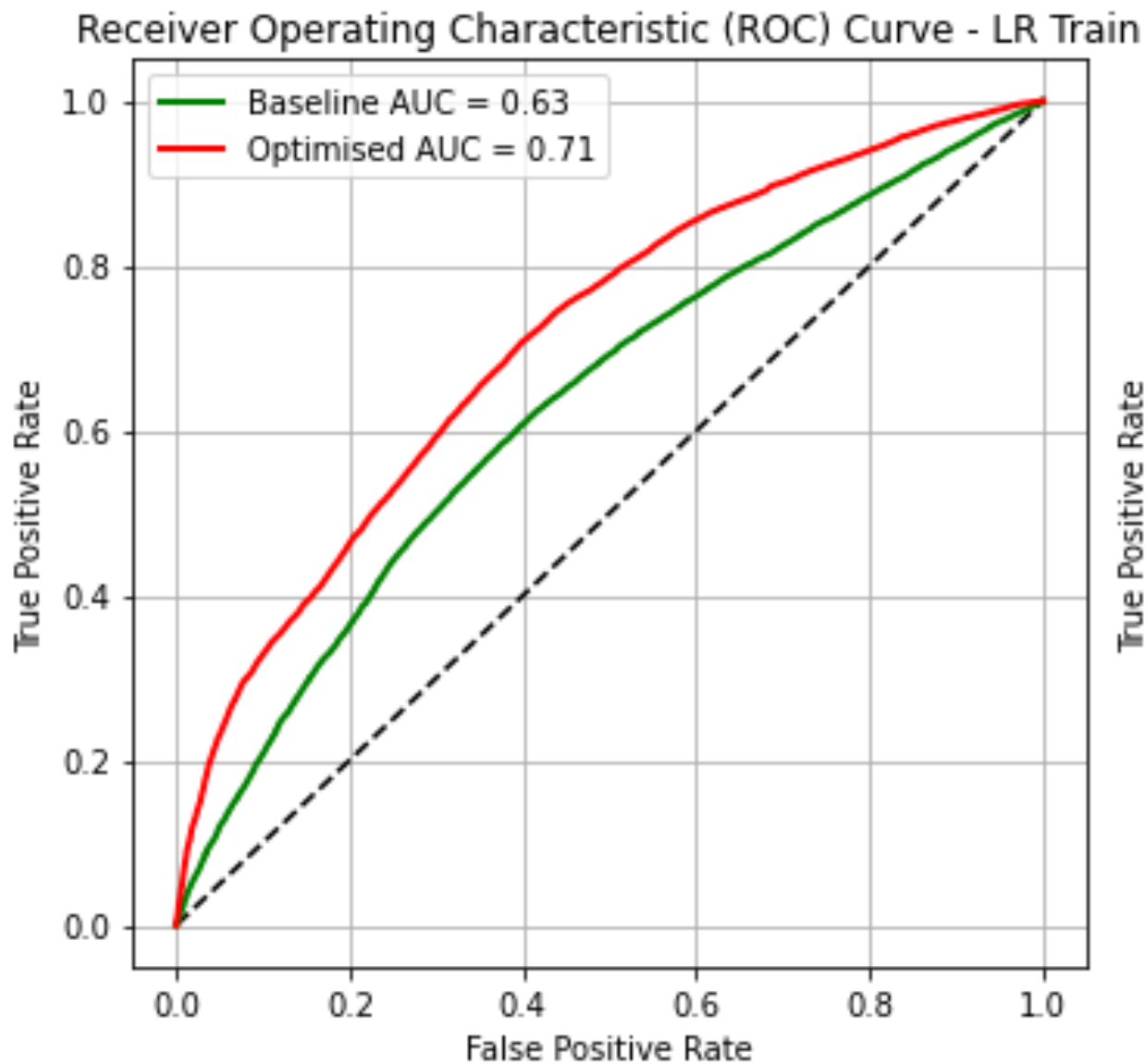
EDA – ... or is it because of violence?



Modelling – Sampling was a big issue...



Modelling – Still work to be done on model performance...

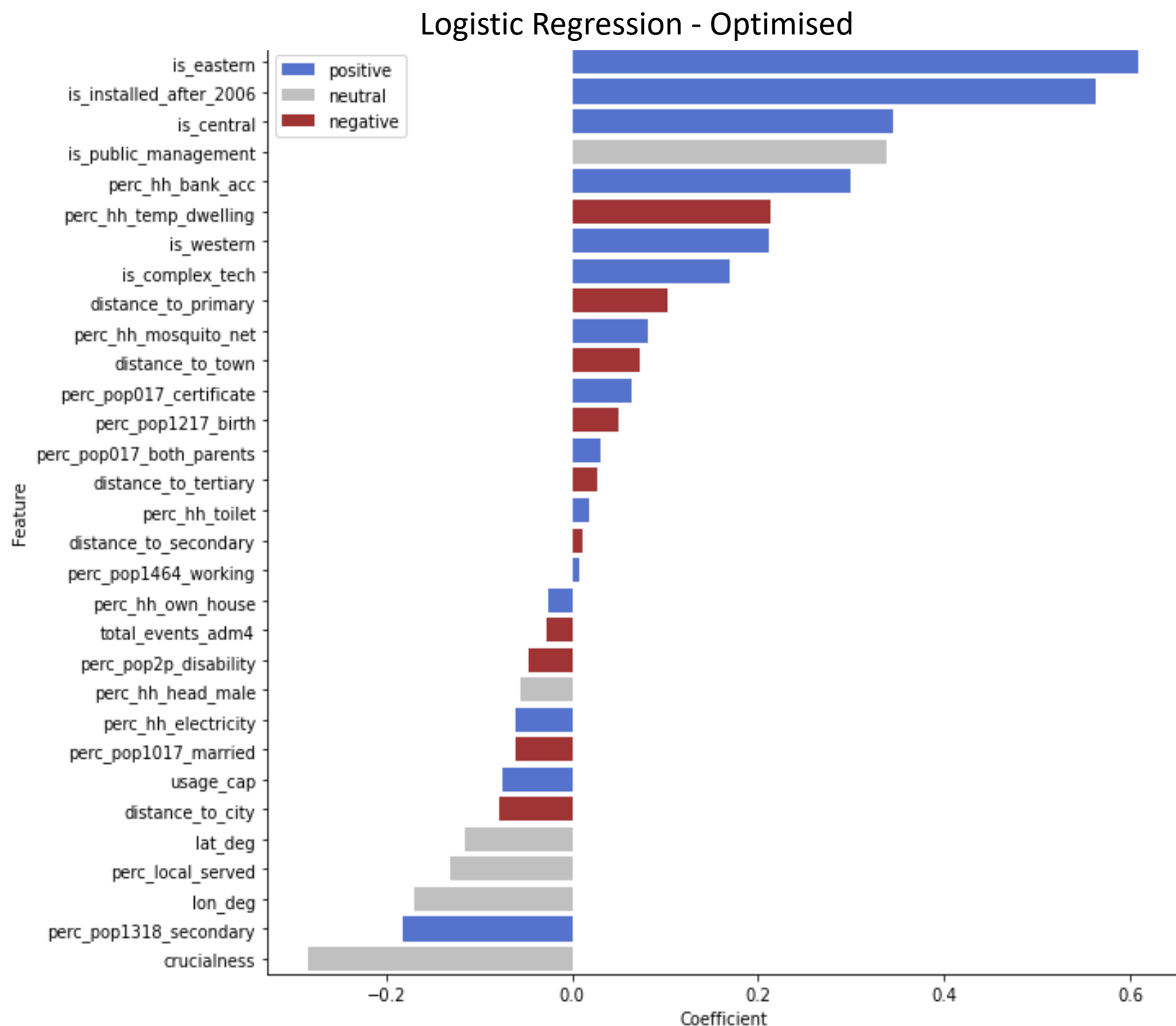


Modelling – ..but initial findings are interesting

+ installed after 2006
+ not in Northern region
+ publicly managed

- school enrollment
- crucialness

~ conflicts



Next steps...

- Modelling, modelling, modelling...but will it work?
- Neural Network?
- Visualisations

