

# Split vs Stratification/Intersection

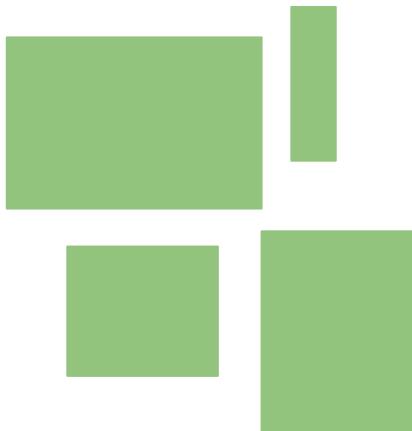
# A dataset can have many attributes (columns)

ECO_ID_U	ECO_CODE	ECO_NAME	WWF_MHT_NAM	RealmMHT	geometry
10184	AT1313	Masai Xeric Grasslands And Shrublands	Deserts and Xeric Shrublands	AT13	
10122	AT0705	East Sudanian Savanna	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	
10147	AT0901	East African Halophytics	Flooded Grasslands and Savannas	AT9	
10132	AT0715	Somali Acacia-Commiphora Bushlands And Thickets	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	

# **Split:** Divide the dataset by one of the attributes (columns)

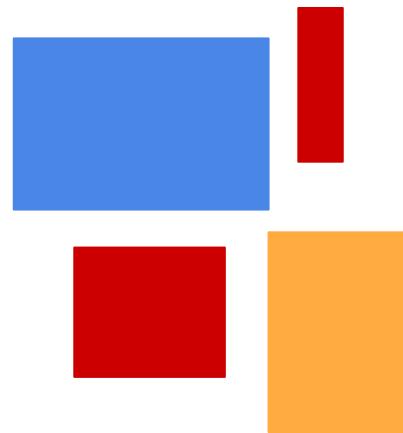
ECO_ID_U	ECO_CODE	ECO_NAME	WWF_MHT_NAM	RealmMHT	geometry
10184	AT1313	Masai Xeric Grasslands And Shrublands	Deserts and Xeric Shrublands	AT13	
10122	AT0705	East Sudanian Savanna	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	
10147	AT0901	East African Halophytics	Flooded Grasslands and Savannas	AT9	
10132	AT0715	Somali Acacia-Commiphora Bushlands And Thickets	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	

# The result of split is a new feature for each class



## WITHOUT SPLIT:

**Feature 1:** Terrestrial habitats



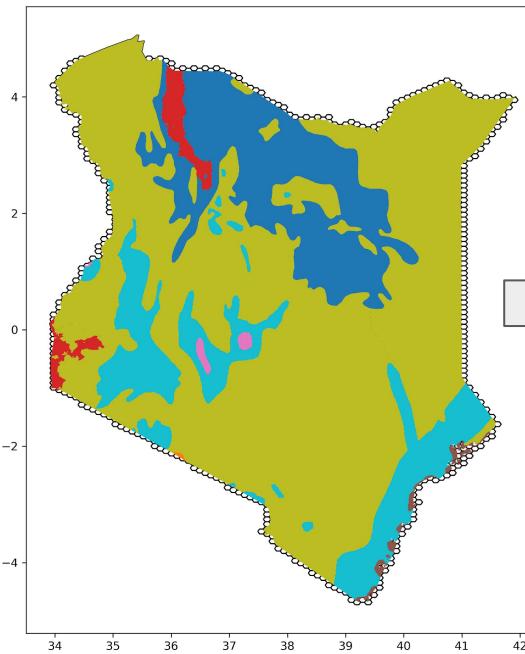
## WITH SPLIT:

**Feature 1:** Deserts and Xeric Shrublands

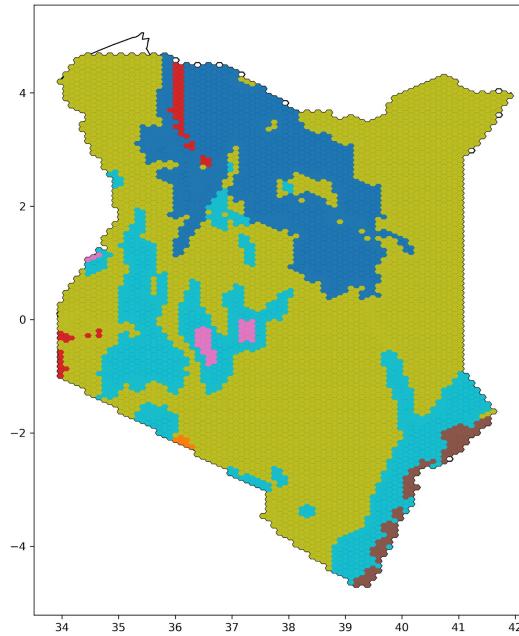
**Feature 2:** Tropical and Subtropical Grasslands,  
Savannas and Shrublands

**Feature 3:** Flooded Grasslands and Savannas

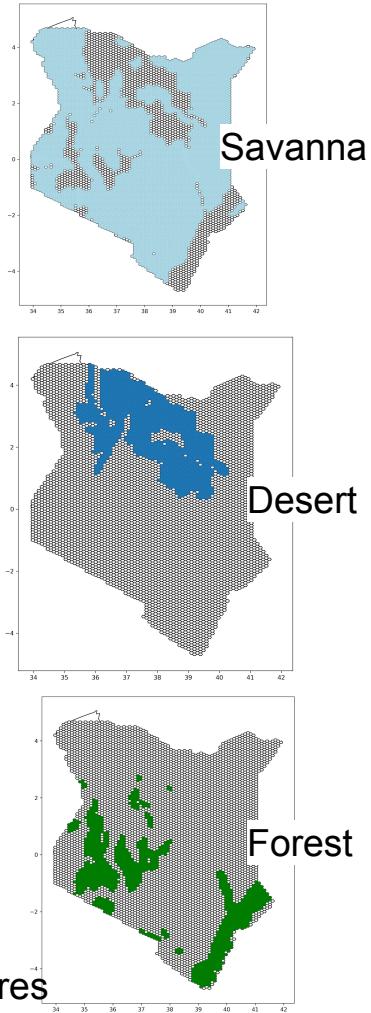
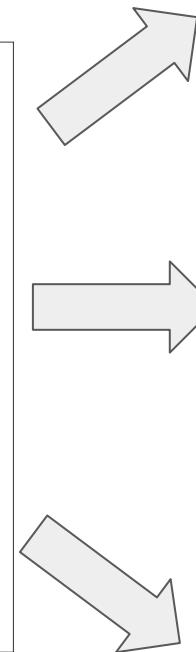
# Example of split with real data



Landscape/  
Bioregional layer



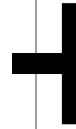
Landscape layer in grid  
(threshold applied)



Split layer:  
New features

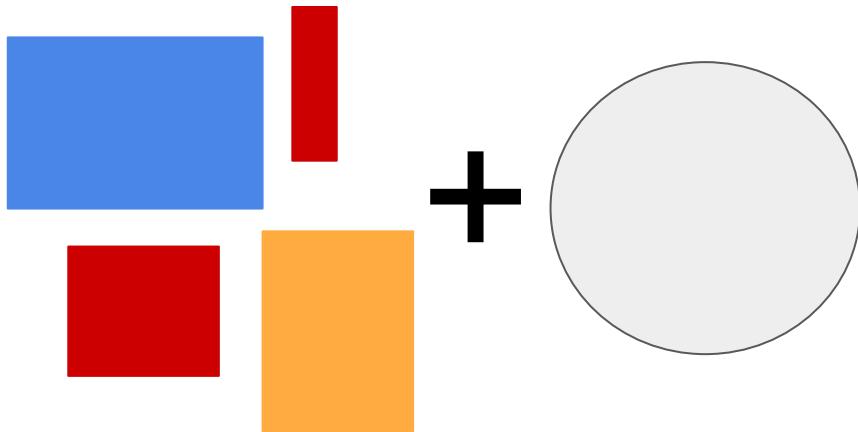
**Stratification:** Split a dataset by attributes and intersect with another feature

ECO_ID_U	ECO_CODE	ECO_NAME	WWF_MHTNAM	RealmMHT	geometry
10184	AT1313	Masai Xeric Grasslands And Shrublands	Deserts and Xeric Shrublands	AT13	
10122	AT0705	East Sudanian Savanna	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	
10147	AT0901	East African Halophytics	Flooded Grasslands and Savannas	AT9	
10132	AT0715	Somali Acacia-Commiphora Bushlands And Thickets	Tropical and Subtropical Grasslands, Savannas and Shrublands	AT7	



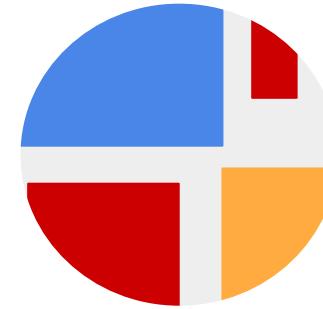
ASSESSMENT	ID_NO	BINOMIAL	PRESENCE	YEAR	geometry
115130419.0	15951.0	Panthera leo	4	2016	
██████████5130419.0	15951.0	Panthera leo	1	2016	
115130419.0	15951.0	Panthera leo	4	2016	
115130419.0	15951.0	Panthera leo	4	2016	

# The result of stratification is one feature for each combination of class + input feature



## WITHOUT STRATIFICATION:

- Feature 1: Deserts and Xeric Shrublands
- Feature 2: Tropical and Subtropical Grasslands, Savannas and Shrublands
- Feature 3: Flooded Grasslands and Savannas
- Feature 4: Lion

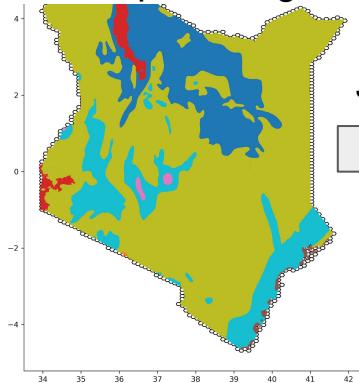


## WITH STRATIFICATION:

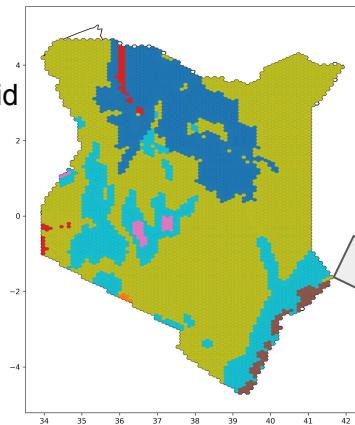
- Feature 1: Deserts and Xeric Shrublands + LION
- Feature 2: Tropical and Subtropical Grasslands, Savannas and Shrublands + LION
- Feature 3: Flooded Grasslands and Savannas + LION

# Example of stratification with real data

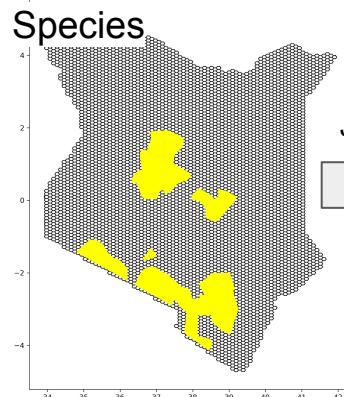
Landscape/Bioregional



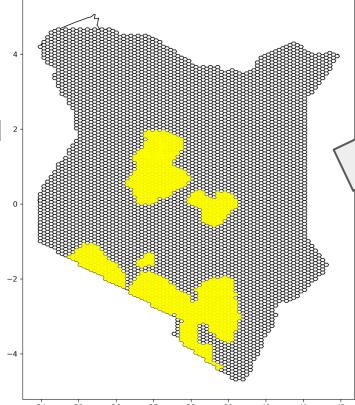
Join with PU grid



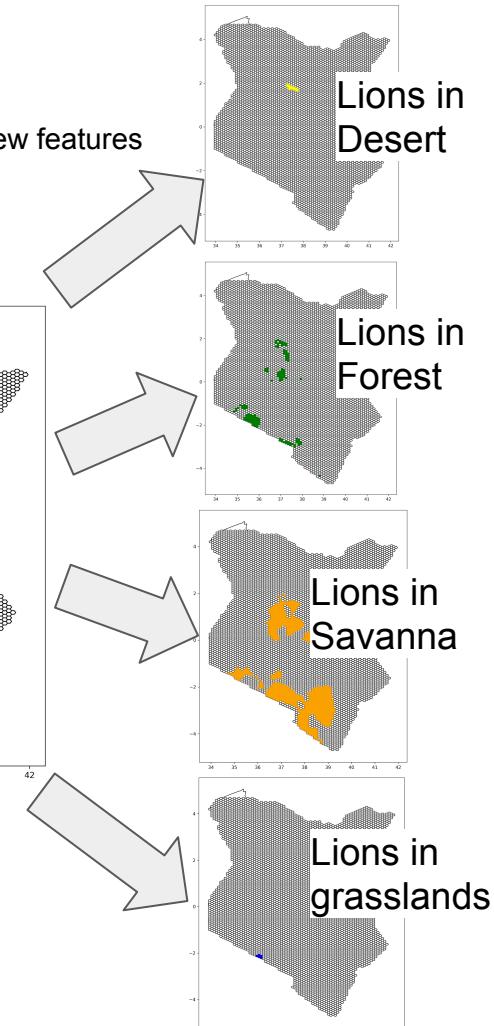
Species



Join with PU grid



New features



Lions intersected  
with habitats

# Steps

## Split:

- Landscape:
  - Read in terrestrial/marine landscape dataset
  - Select column to split by
  - Spatial join with PU grid and keep only one value per PU (majority of area)
  - Separate the dataset into different features

## Stratification

- Landscape:
  - Read in terrestrial/marine landscape dataset
  - Select column to split by
  - Spatial join with PU grid and keep only one value per PU (majority of area)
- Species
  - Read species dataset
  - Spatial join with PU grid
  -
- Join by PUid the landscape and species layer
- Separate the joined layer into different features, each with landscape + species