Hike with a view

Helping hikers select trails

Tatiana Konstantinova
Insight Data Science

Hiking in the USA

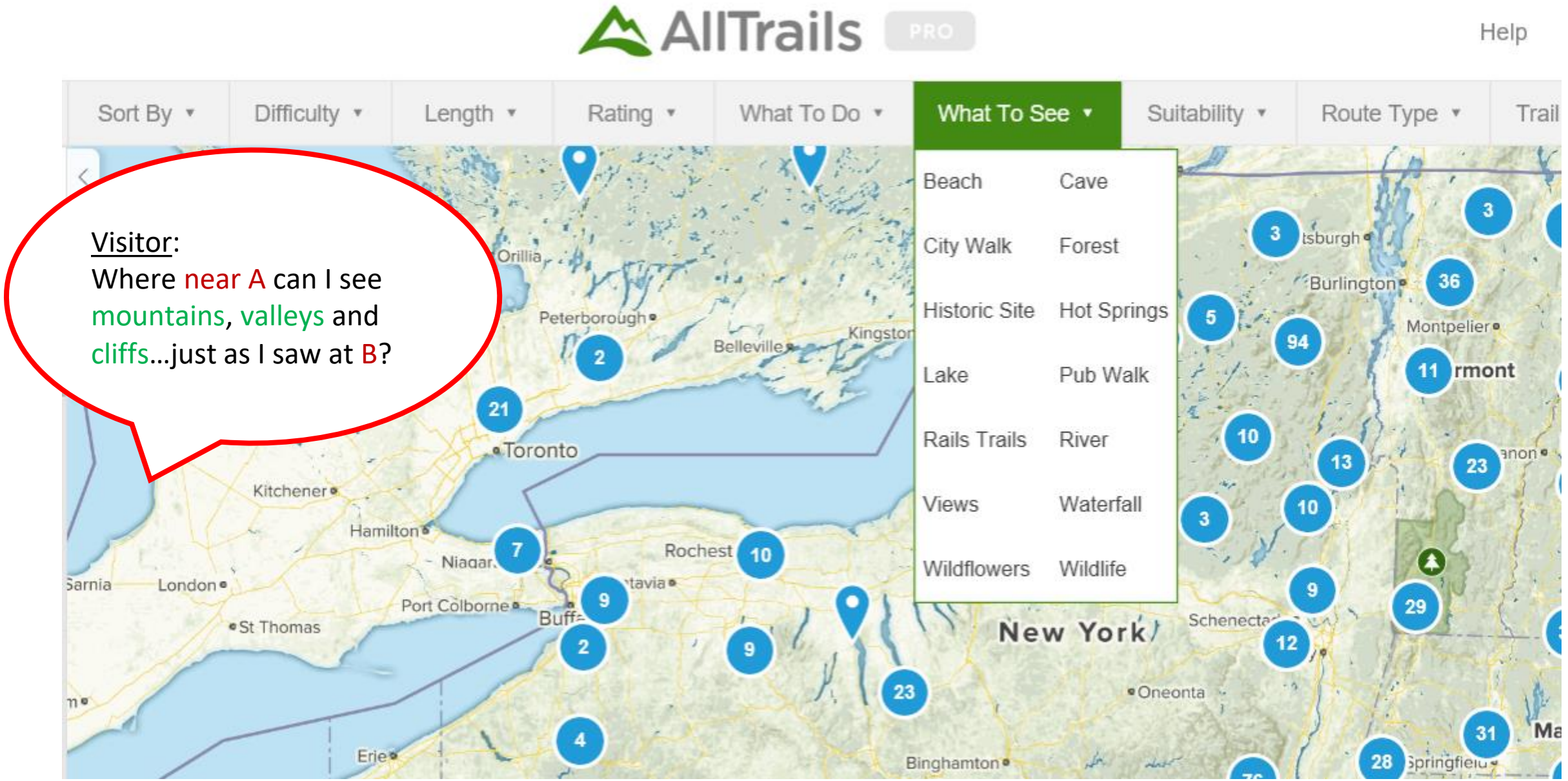
45M hike each year

"Observing scenic beauty"
in top 5 reasons
for getting outdoors

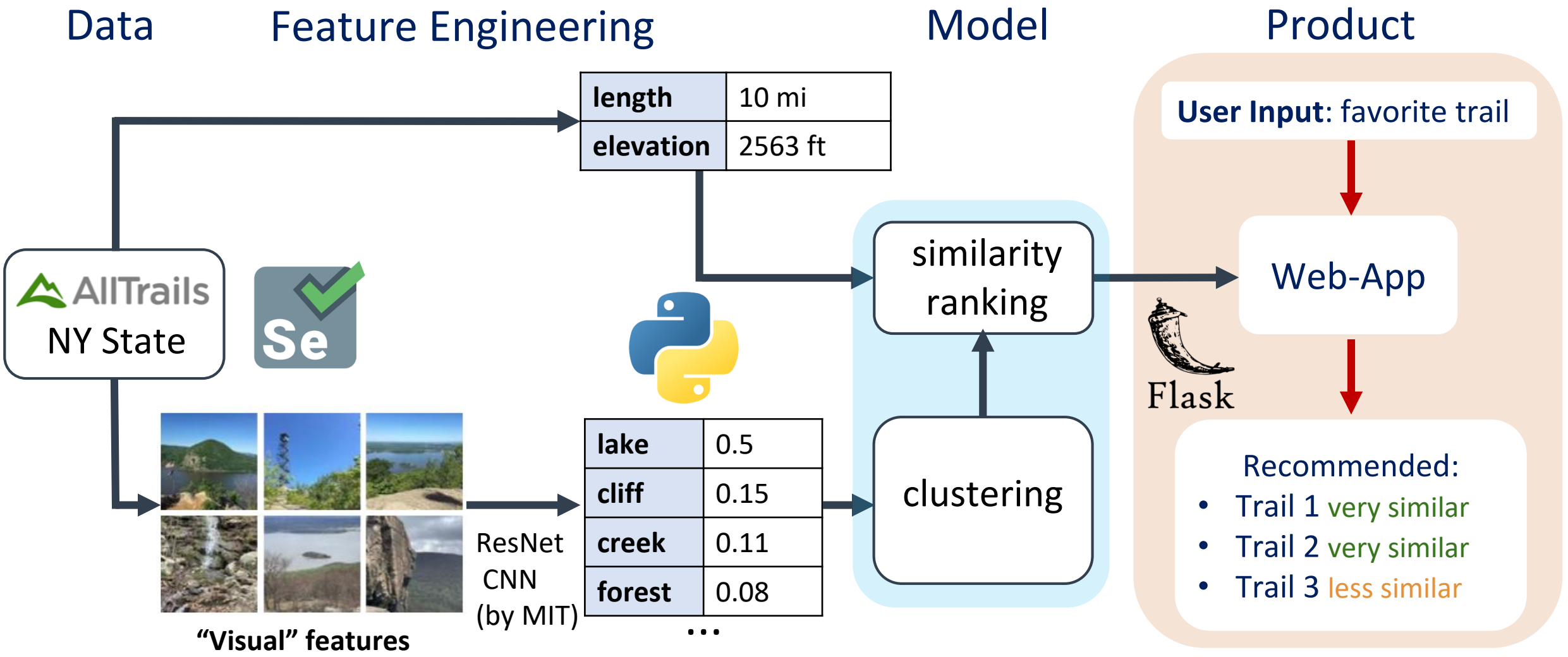
~ 2500 trails in NY
Which one will please your eye?

data source: https://www.statista.com/

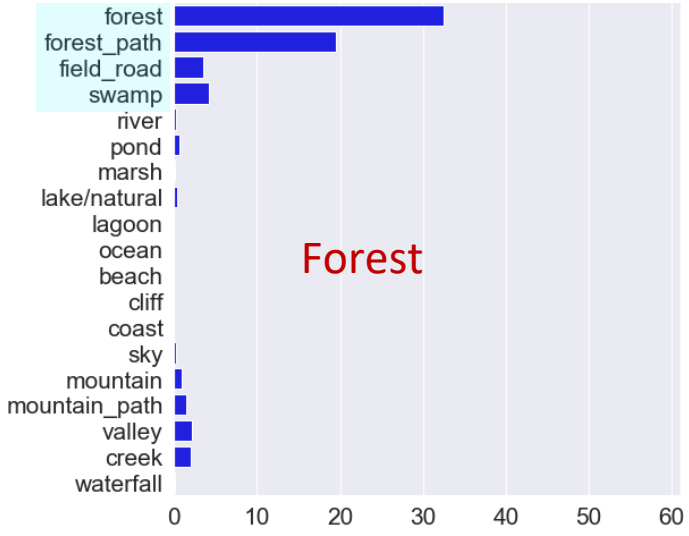
Current solution for trail selection

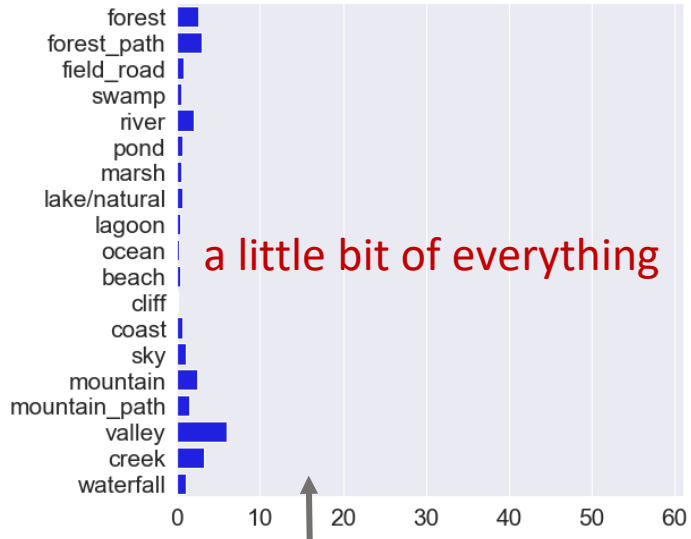


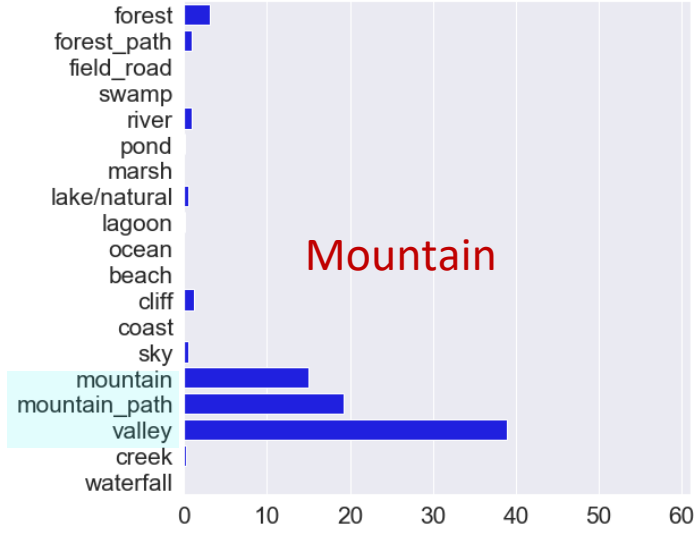
My approach

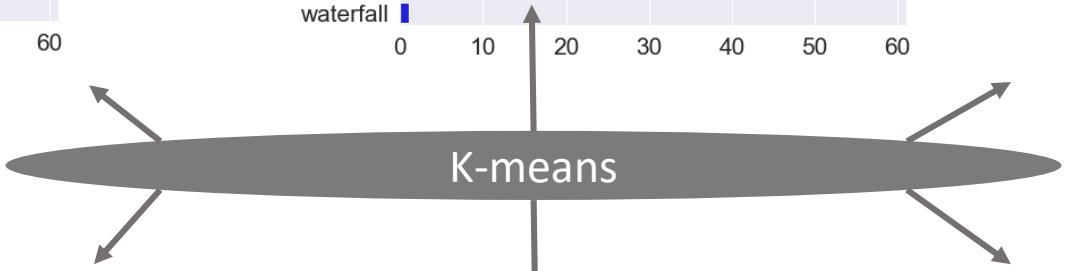

Data
Feature Engineering
Model
Product
AllTrails
NY State
Se
length
10 mi
elevation
2563 ft
similarity ranking
clustering
User Input: favorite trail
Web-App
Flask
Recommended:
Trail 1 very similar
Trail 2 very similar
Trail 3 less similar
lake
0.5
cliff
0.15
creek
0.11
forest
0.08
...
ResNet CNN (by MIT)
"Visual" features

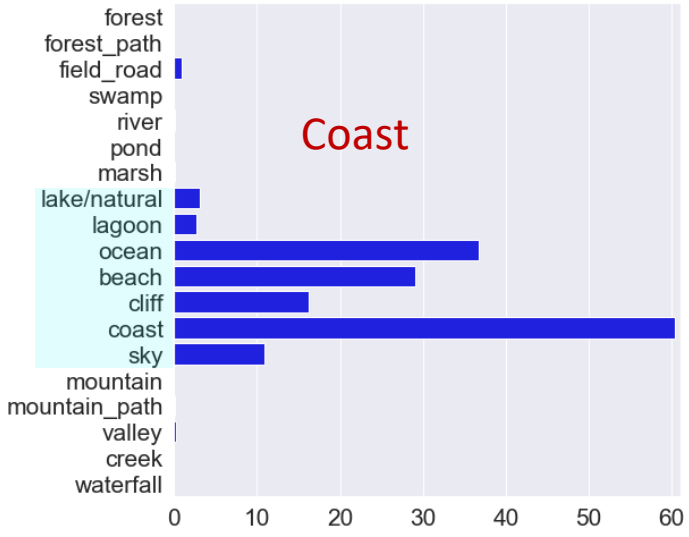
Trails Views Groups

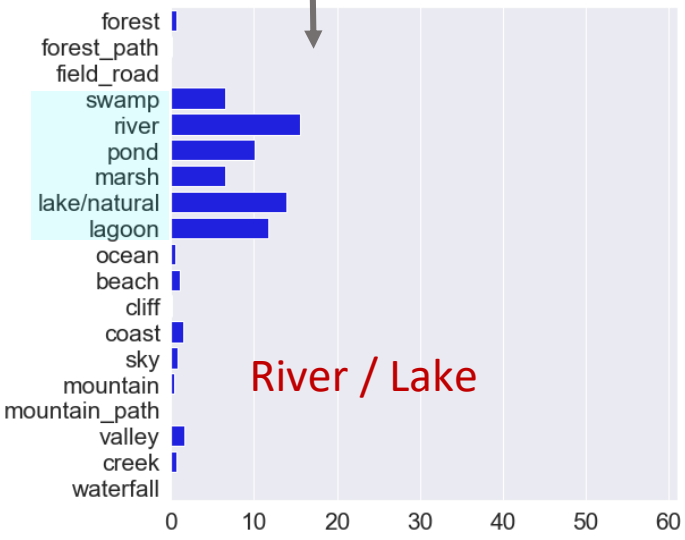


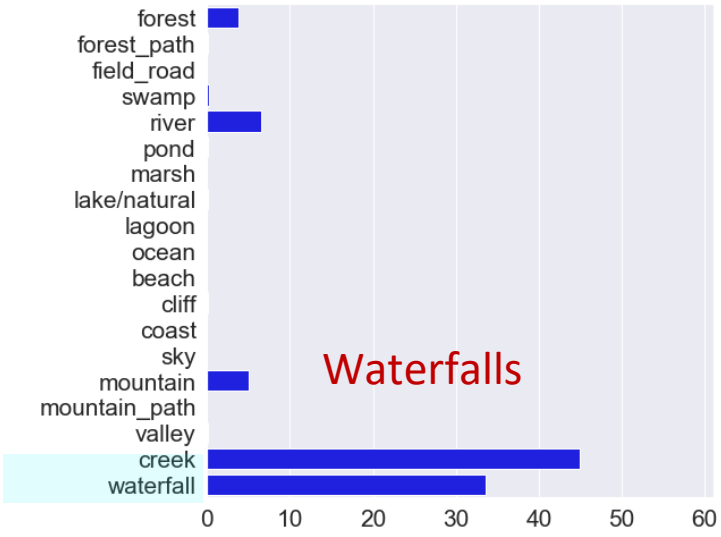












Similarity ranking

User's Trail

Length	Elevation	lake	river	forest	mountain	filed	waterfall	beach
5.0	250	0.86	0.5	0.91	0	0.1	0.02	0.5

Other trails in the cluster

Length	Elevation	lake	river	forest	mountain	filed	waterfall	beach
3.2	0	0.28	0.8	0.69	0	0.3	0	0.1
0.25	0	0.7	0.6	0.46	0	0.7	0	0
…								
15.2	3650	0.56	0.1	0.6	0.25	0.21	0.4	0

Cosine Similarity

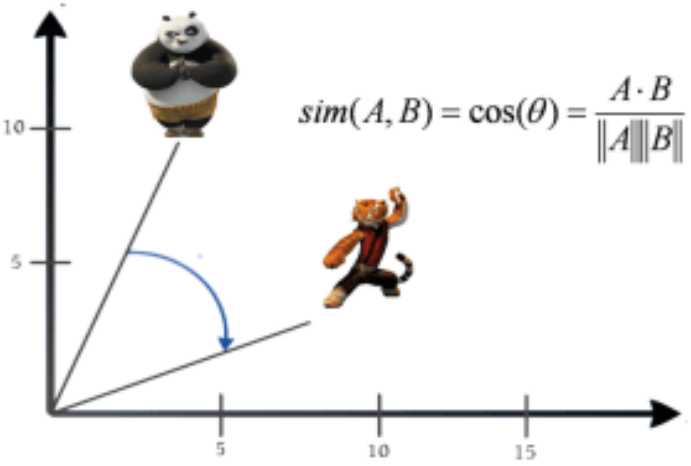
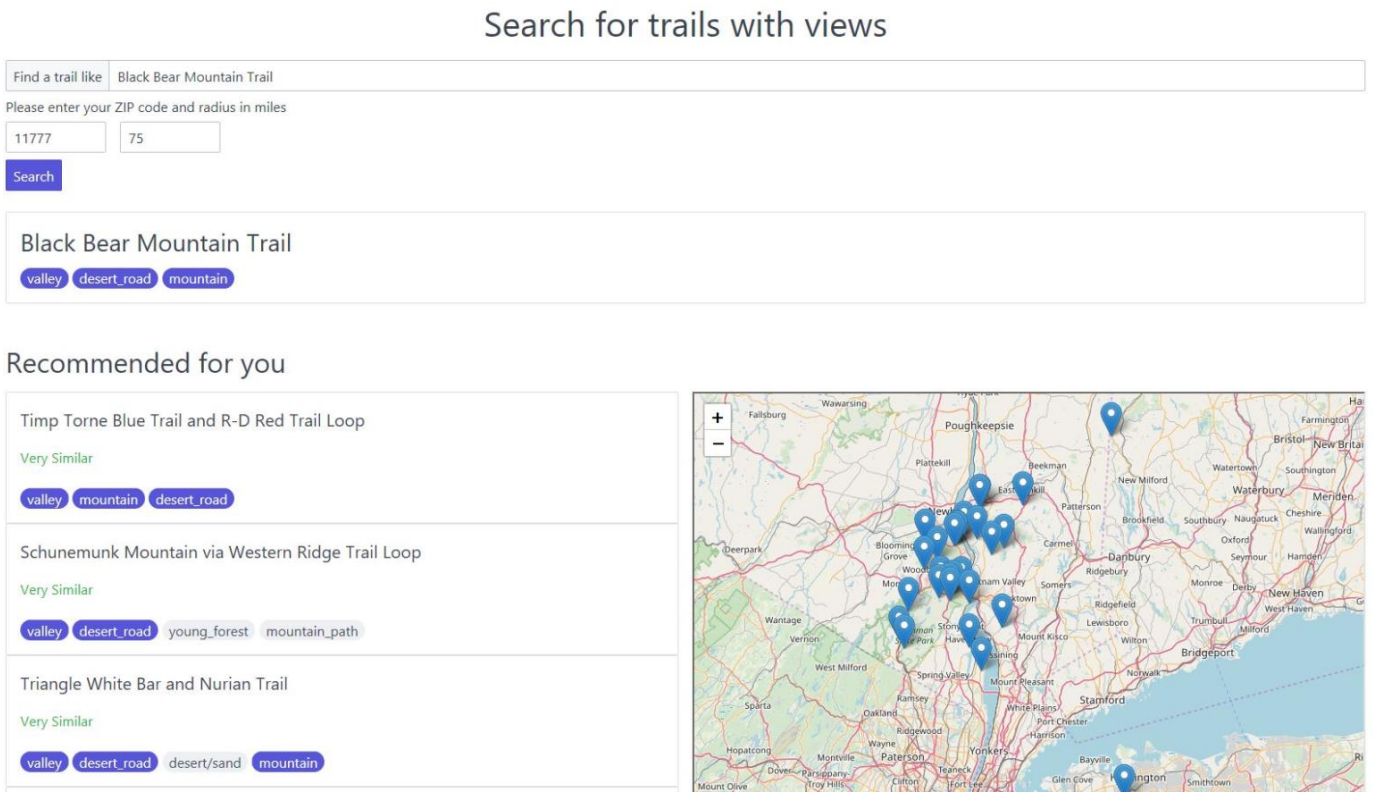


Image: techinpink.com

Web App

Hike-with-view.club



Search for trails with views

Find a trail like | Buttermilk Falls Gorge and Rim Trail Loop

Please enter your ZIP code and radius in miles

10001 | 50

Search

Buttermilk Falls Gorge and Rim Trail Loop

creek mountain forest swimming_hole

Recommended for you

Diamond Mountain Loop via Stony Brook and Pine Meadows

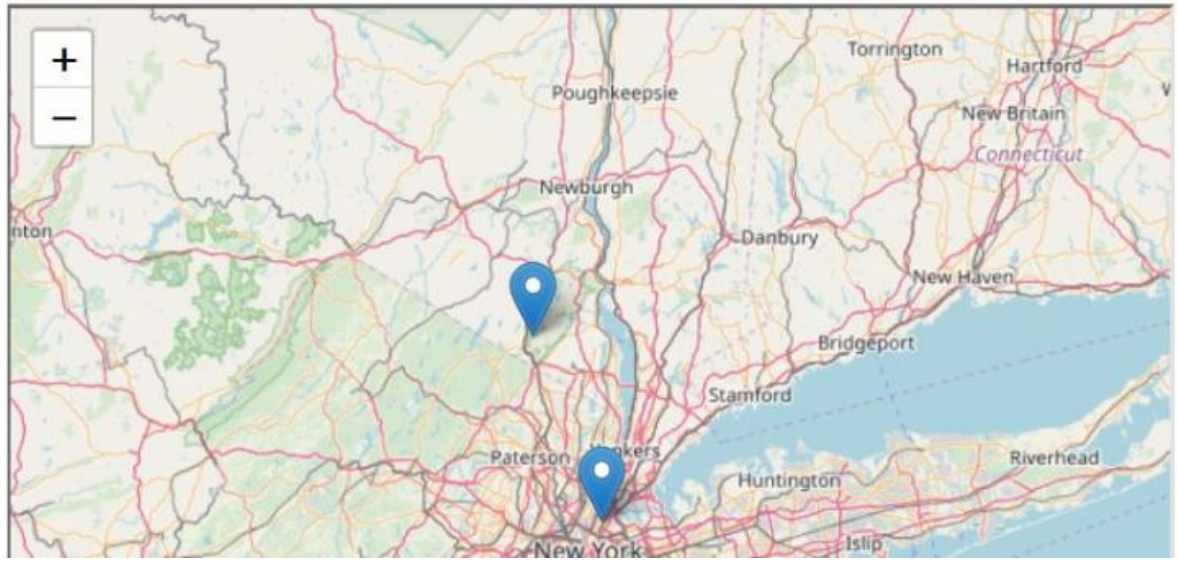
Very Similar

creek swamp forest river

East Drive 102nd Street Cutoff and West Drive Loop

Very Similar

creek waterfall swimming_hole hot_spring



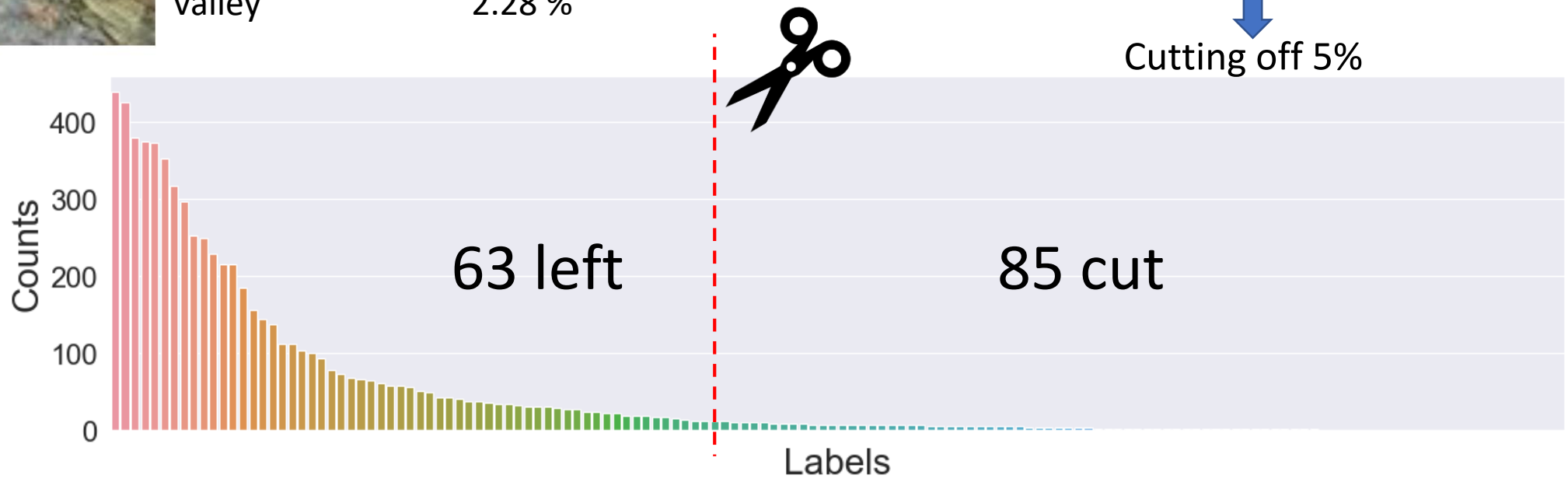
Visual Features

Labels	Probability
creek	41.09 %
river	15.60 %
swimming hole	8.77 %
hot spring	7.37 %
dam	6.09 %
raft	4.25 %
valley	2.28 %

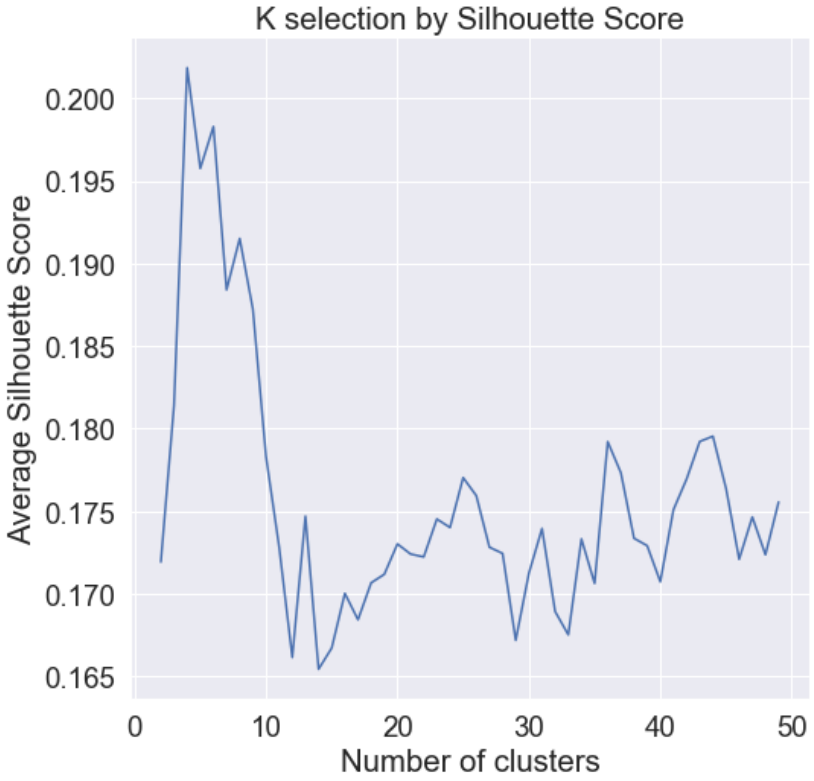
Average within the trail

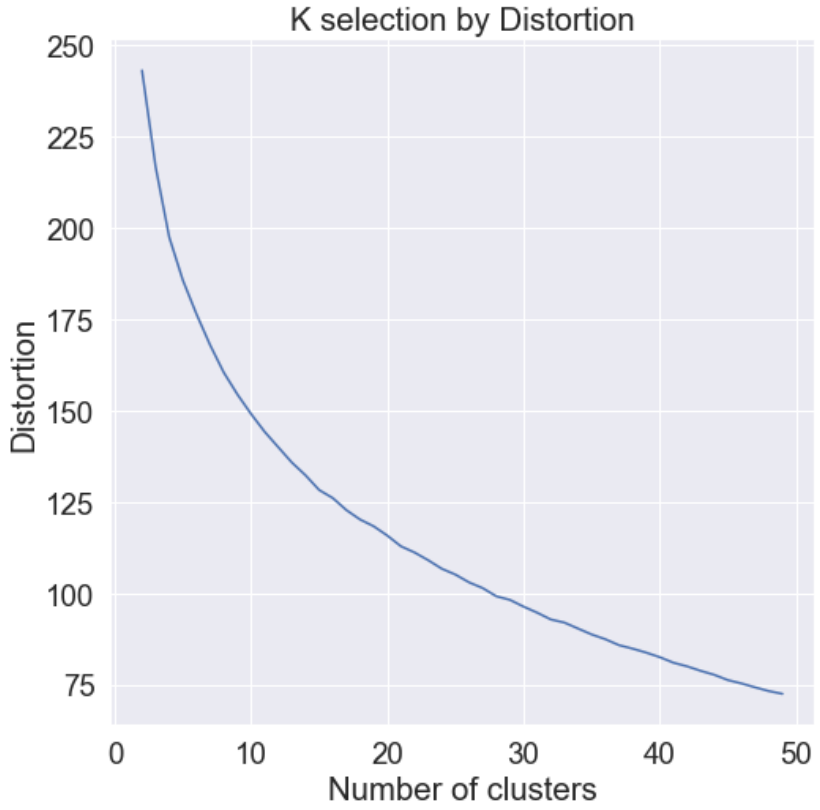
Select top 5 for every trail

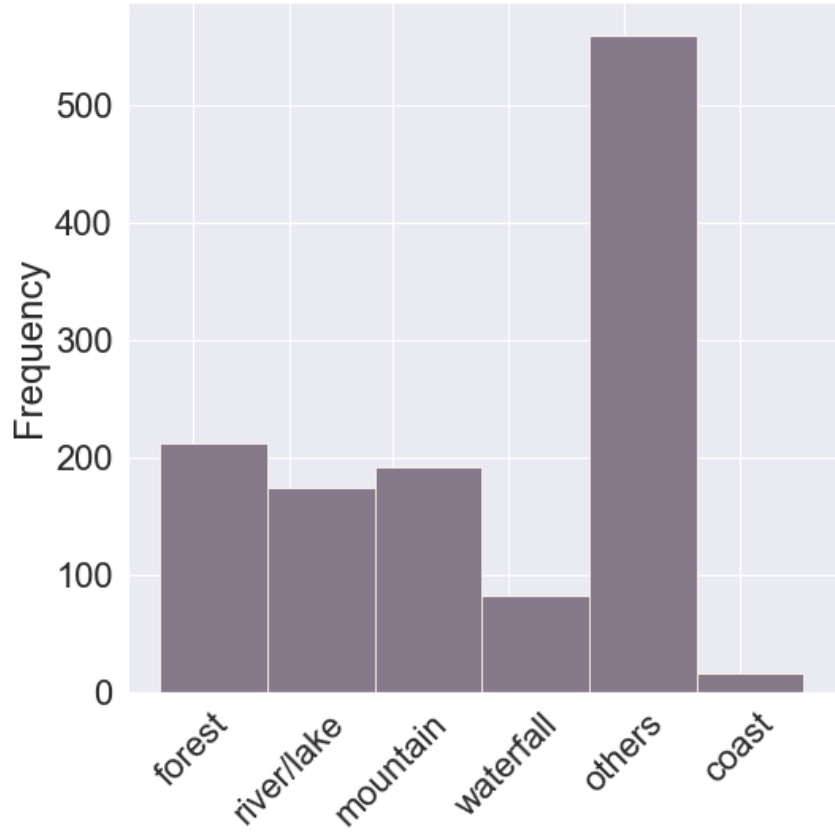
Cutting off 5%

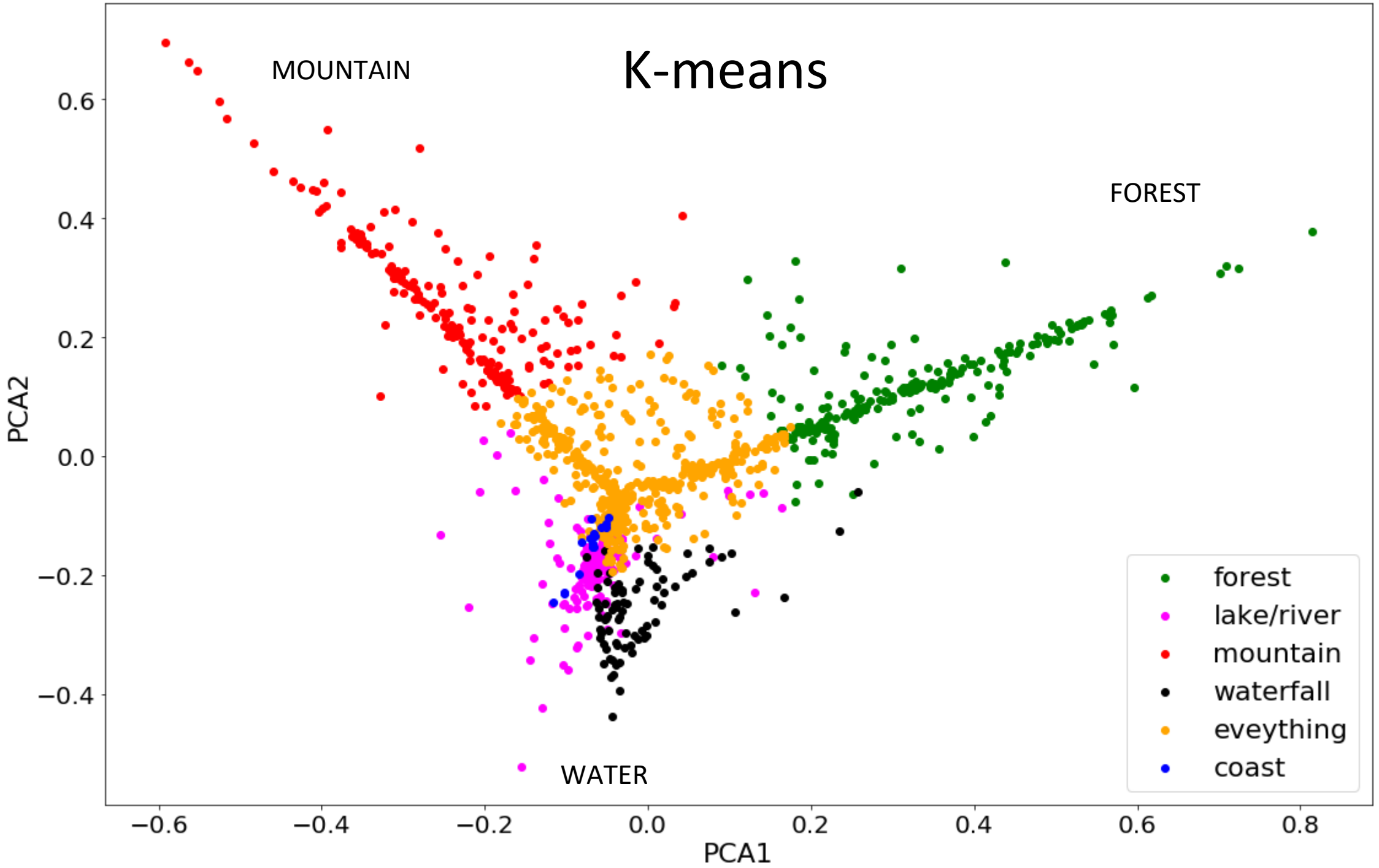


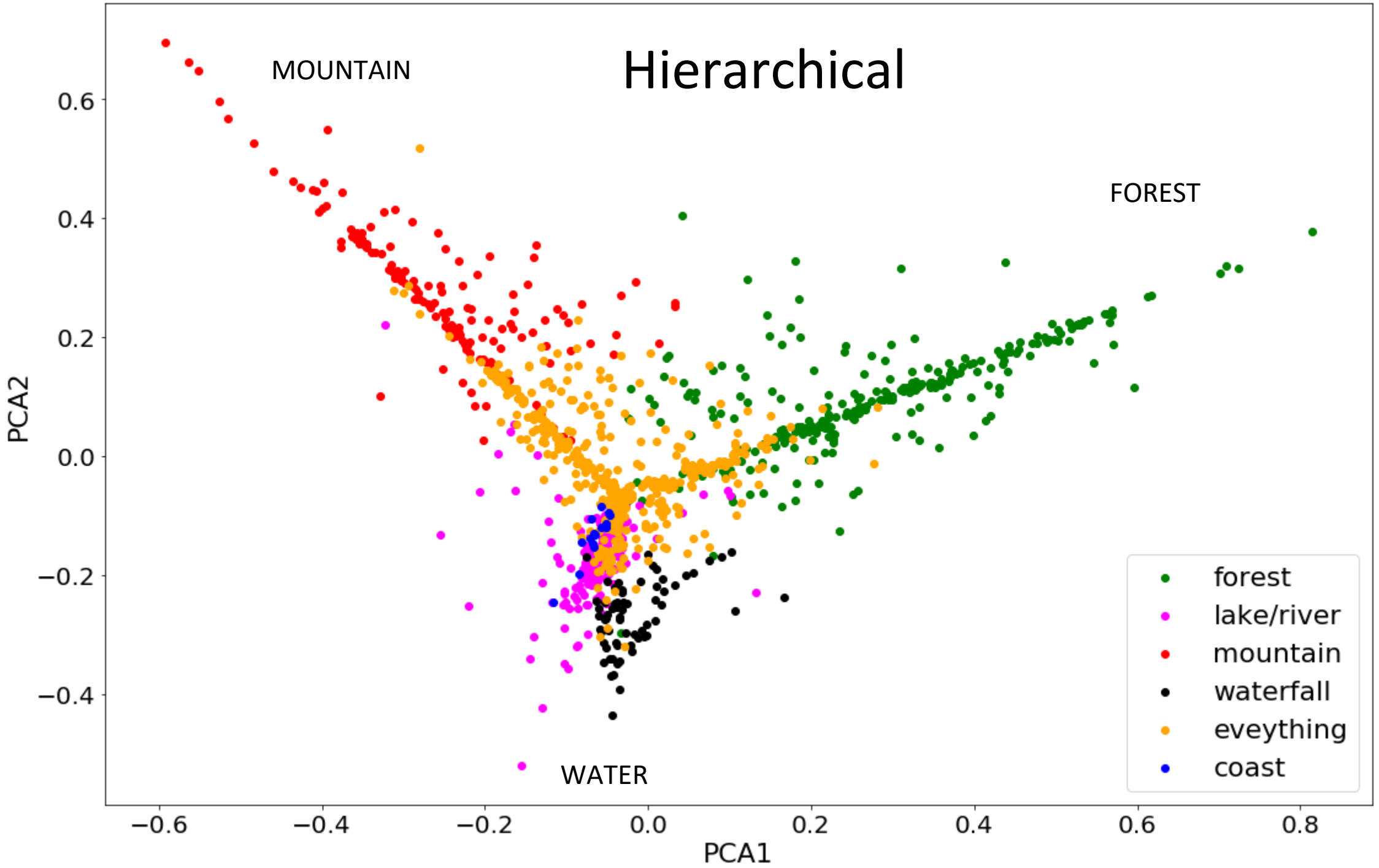
K-means clustering








K-means
MOUNTAIN
FOREST
WATER
PCA1
PCA2
forest
lake/river
mountain
waterfall
eveything
coast

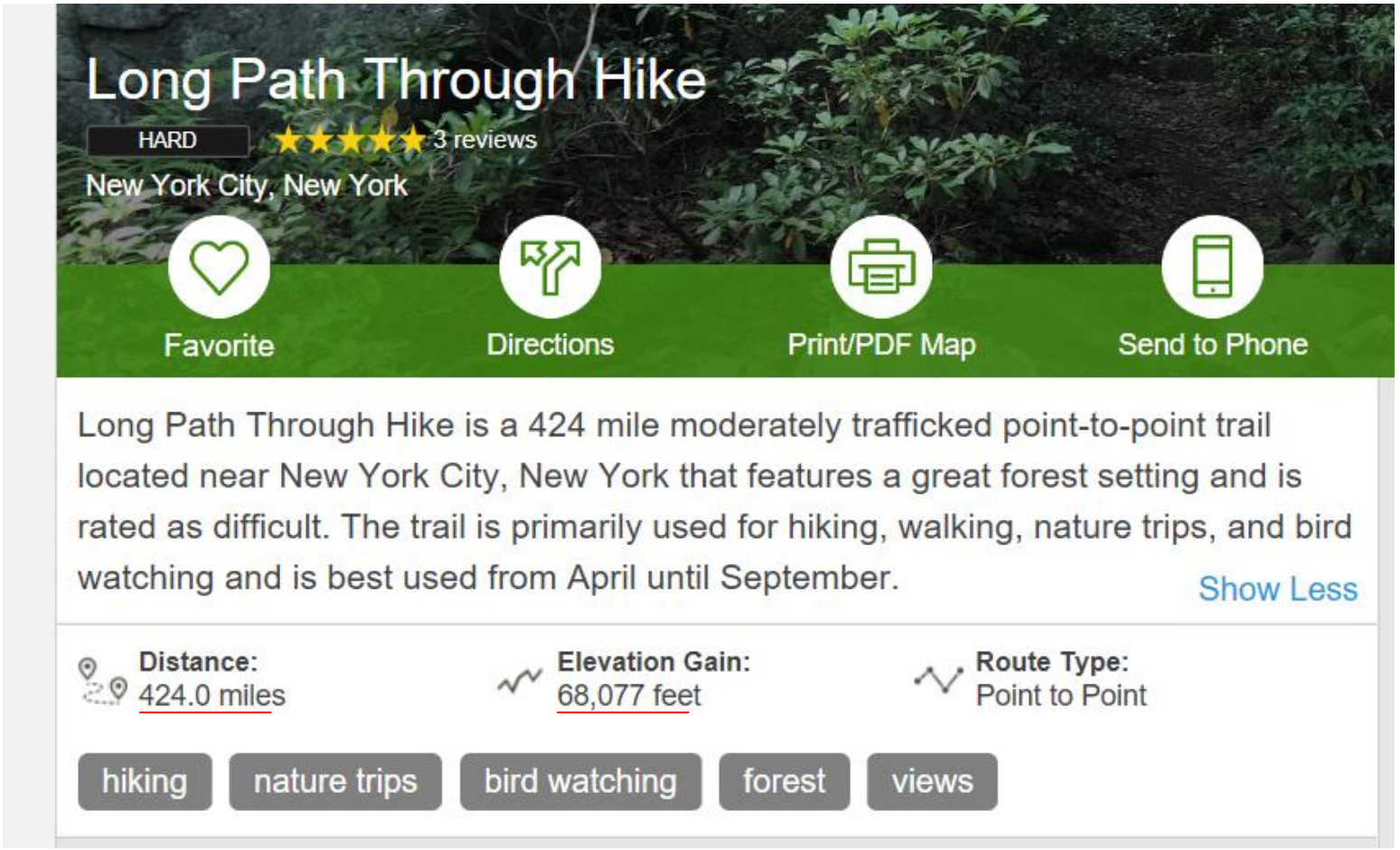

Hierarchical
MOUNTAIN
FOREST
WATER
PCA1
PCA2
forest
lake/river
mountain
waterfall
eveything
coast

tagged by alltrails.com

Clusters from photo labels

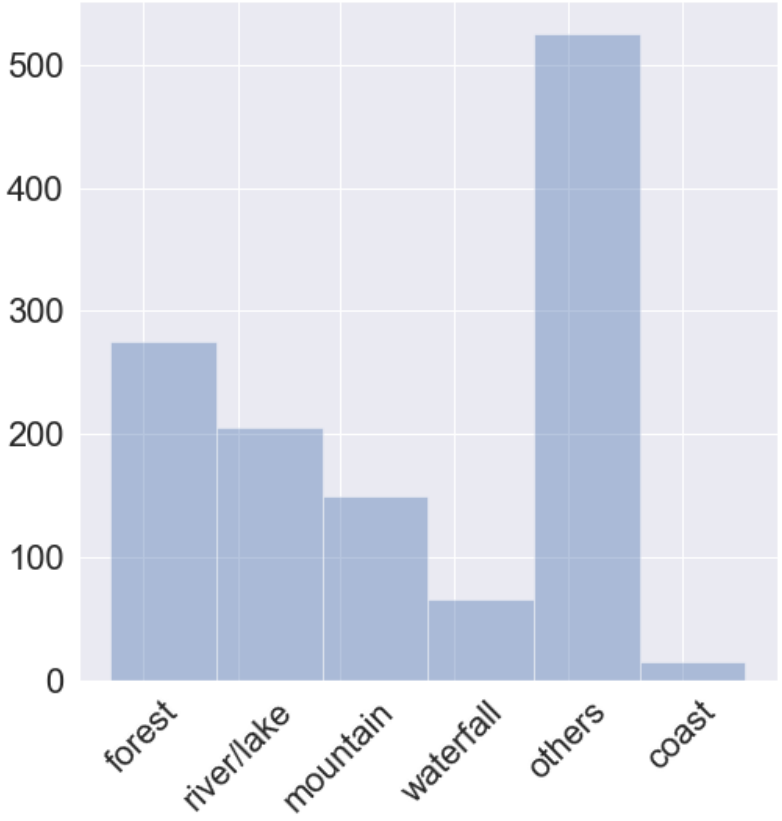
	forest	beach	river	lake	waterfall
Forest	81 (%)	2	18	19	11
Mountain	88	1	20	24	17
Coast	33	73	7	27	0
River/Lake	68	7	25	61	11
Waterfall	76	0	47	10	81
A bit of everything	76	6	26	28	18

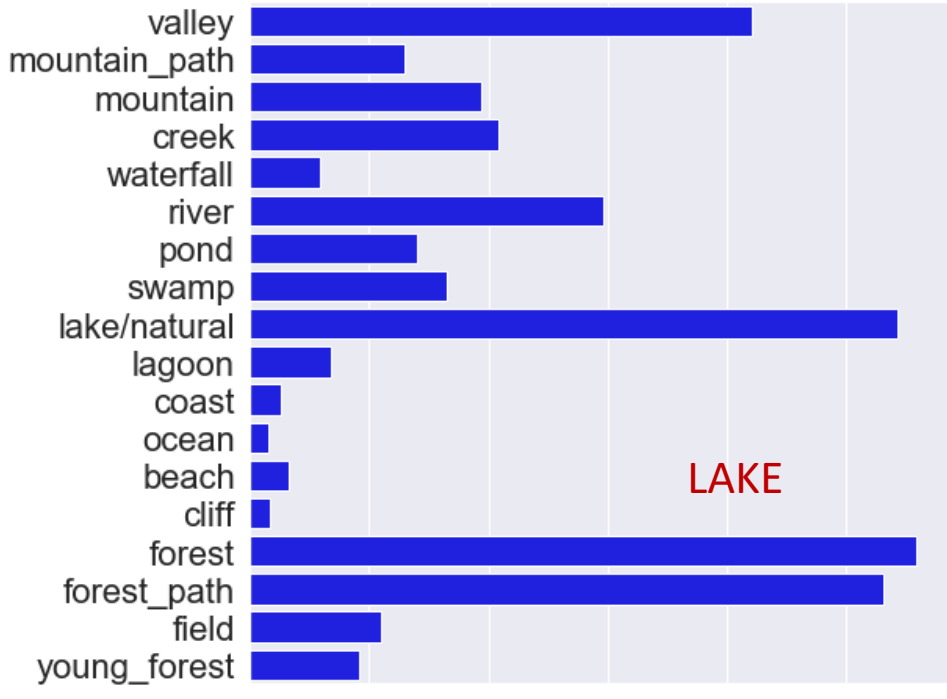
Outliers

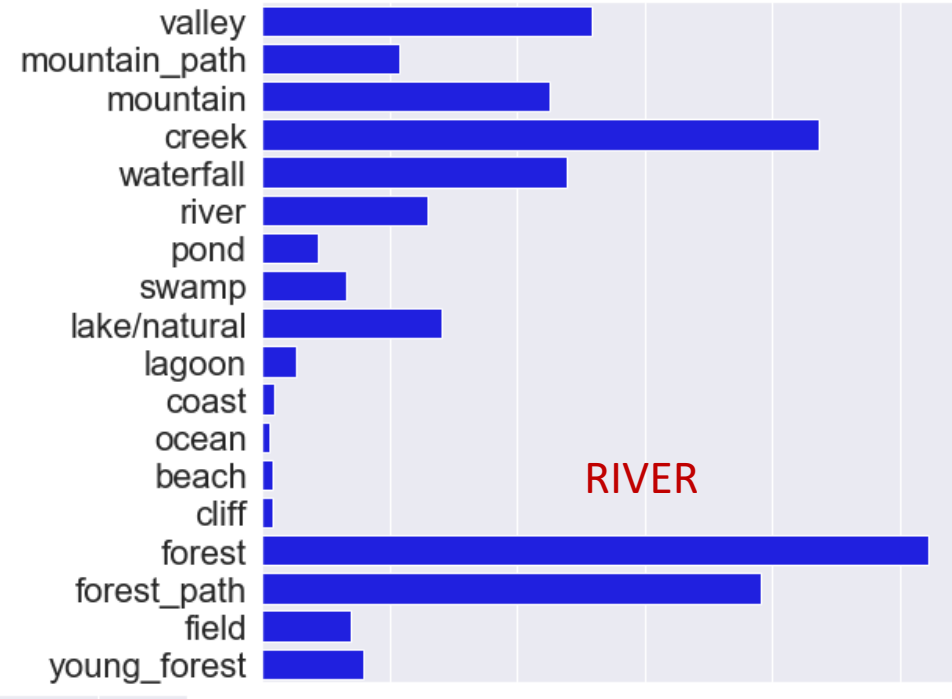

Long Path Through Hike
HARD
3 reviews
New York City, New York
Favorite
Directions
Print/PDF Map
Send to Phone
Long Path Through Hike is a 424 mile moderately trafficked point-to-point trail located near New York City, New York that features a great forest setting and is rated as difficult. The trail is primarily used for hiking, walking, nature trips, and bird watching and is best used from April until September.
Show Less
Distance:
424.0 miles
Elevation Gain:
68,077 feet
Route Type:
Point to Point
hiking
nature trips
bird watching
forest
views

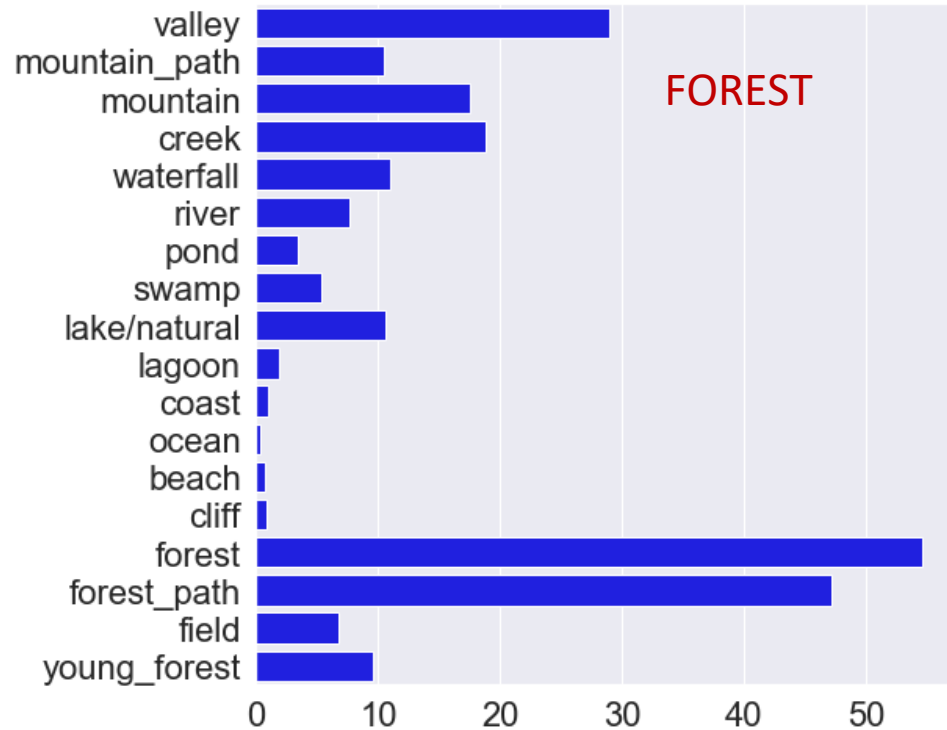
Hierarchical Clustering

- ❑ Centered around certain theme
- ❑ Include “ocean” theme (1%)
- ❑ Do not depend on random assignment




valley
mountain_path
mountain
creek
waterfall
river
pond
swamp
lake/natural
lagoon
coast
ocean
beach
cliff
forest
forest_path
field
young_forest
LAKE
0
2
4
6
8


valley
mountain_path
mountain
creek
waterfall
river
pond
swamp
lake/natural
lagoon
coast
ocean
beach
cliff
forest
forest_path
field
young_forest
RIVER
0.0
2.5
5.0
7.5
10.0
12.5


valley
mountain_path
mountain
creek
waterfall
river
pond
swamp
lake/natural
lagoon
coast
ocean
beach
cliff
forest
forest_path
field
young_forest
FOREST
0
10
20
30
40
50