Identifying Mountain Bike Categories from Frame Geometry

with PCA & Clustering

Presentation:
Saf Flatters 21827361

Identifying Mountain Bike Categories from Frame Geometry

with PCA & Clustering

Presentation:
Saf Flatters 21827361

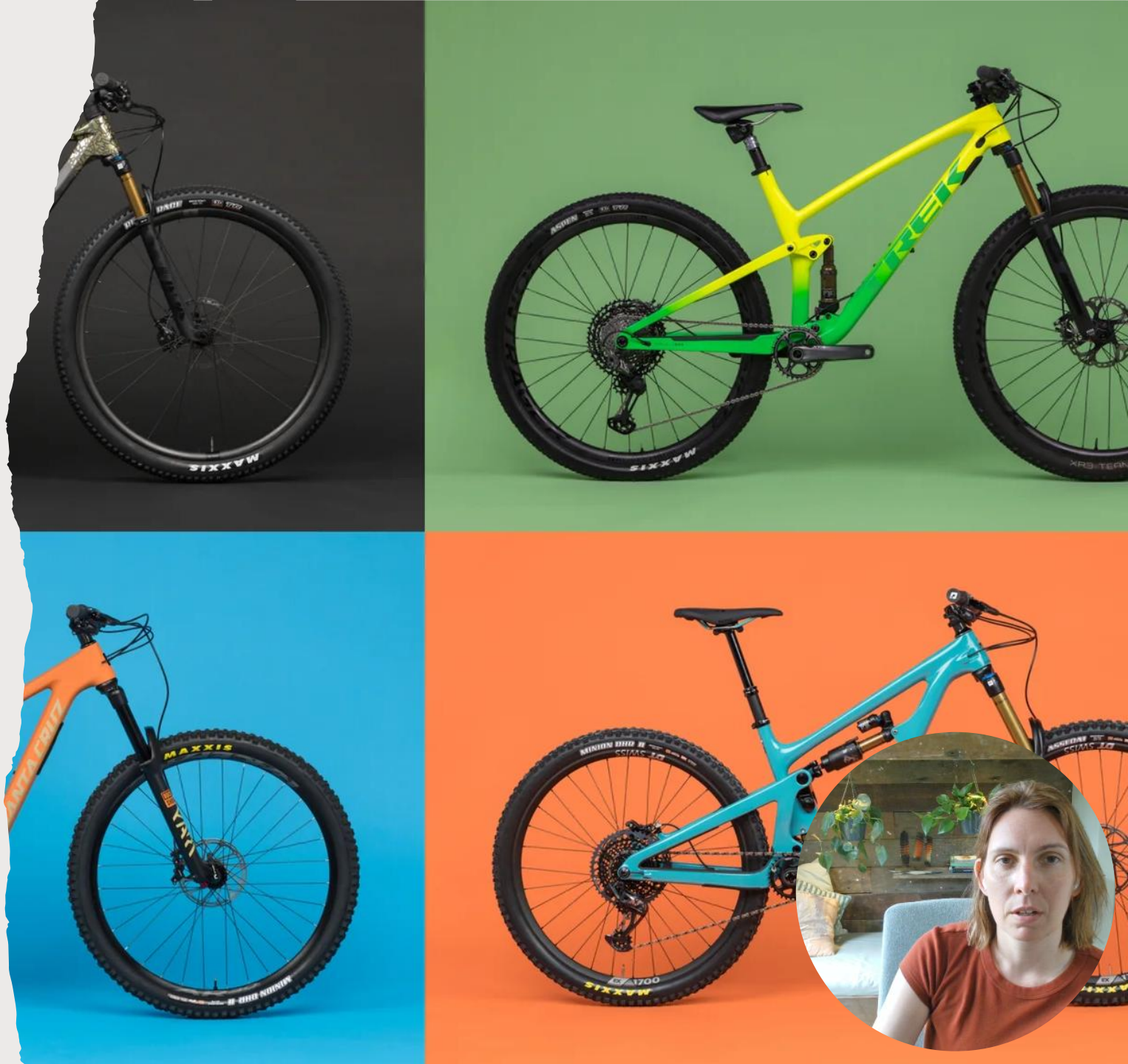
Project Aim

Can frame geometry alone uncover riding styles?

→ Data Wrangling
→ PCA
→ k-Means Clustering
→ k-Prototypes Clustering
→ SVM Classifications

Check against industry labels:

- Cross Country
- Trail
- Enduro
- Downhill

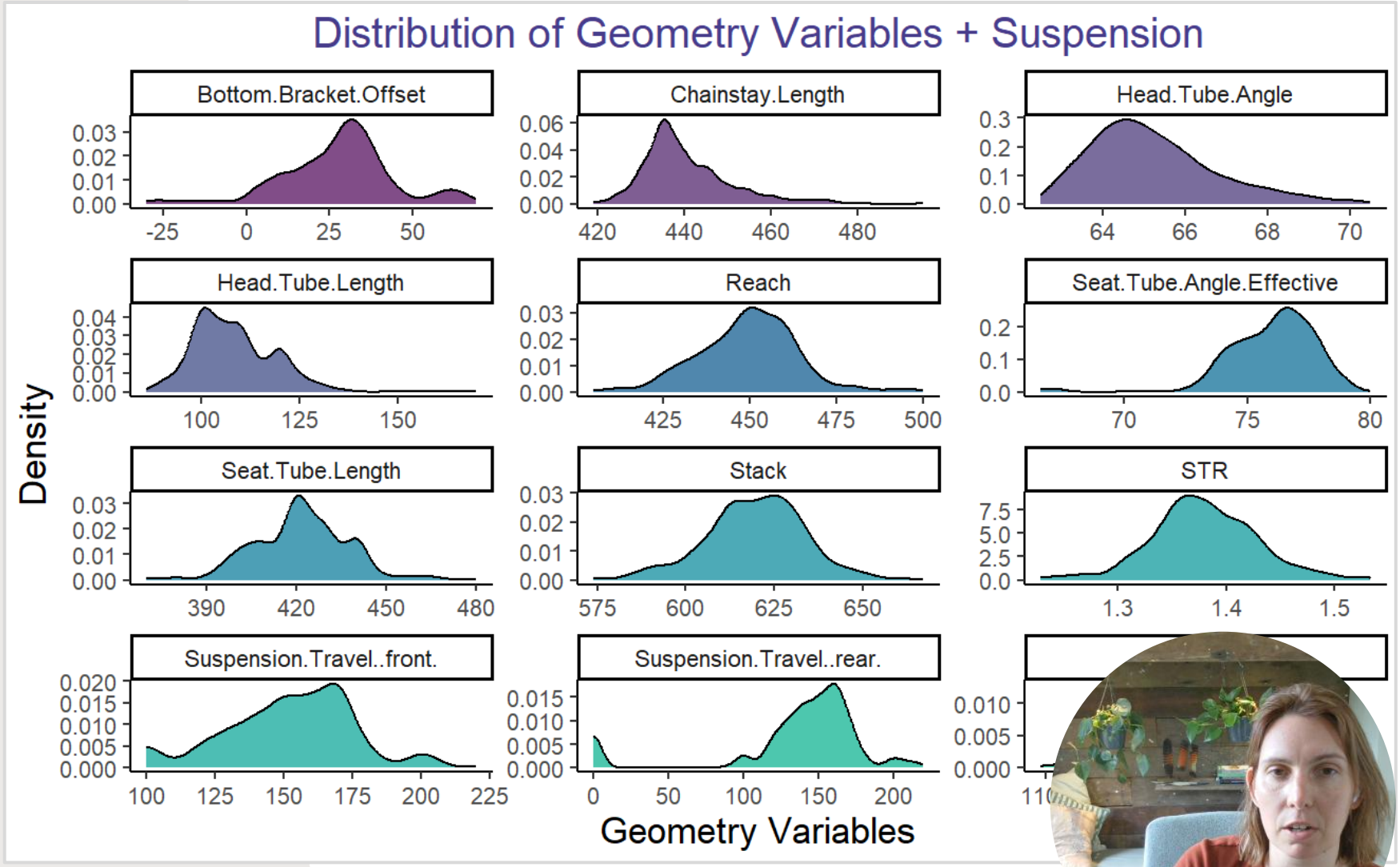


The Data

Data: https://github.com/dorianprill/dataset-bicycle-geometry

616 mountain bike frames that are:

- Medium Sized
- 27” or 29” tires
- 12 numeric variables (geo related)
- 6 other geo-related variables
- 2 suspension related variables reserved for post-hoc validation

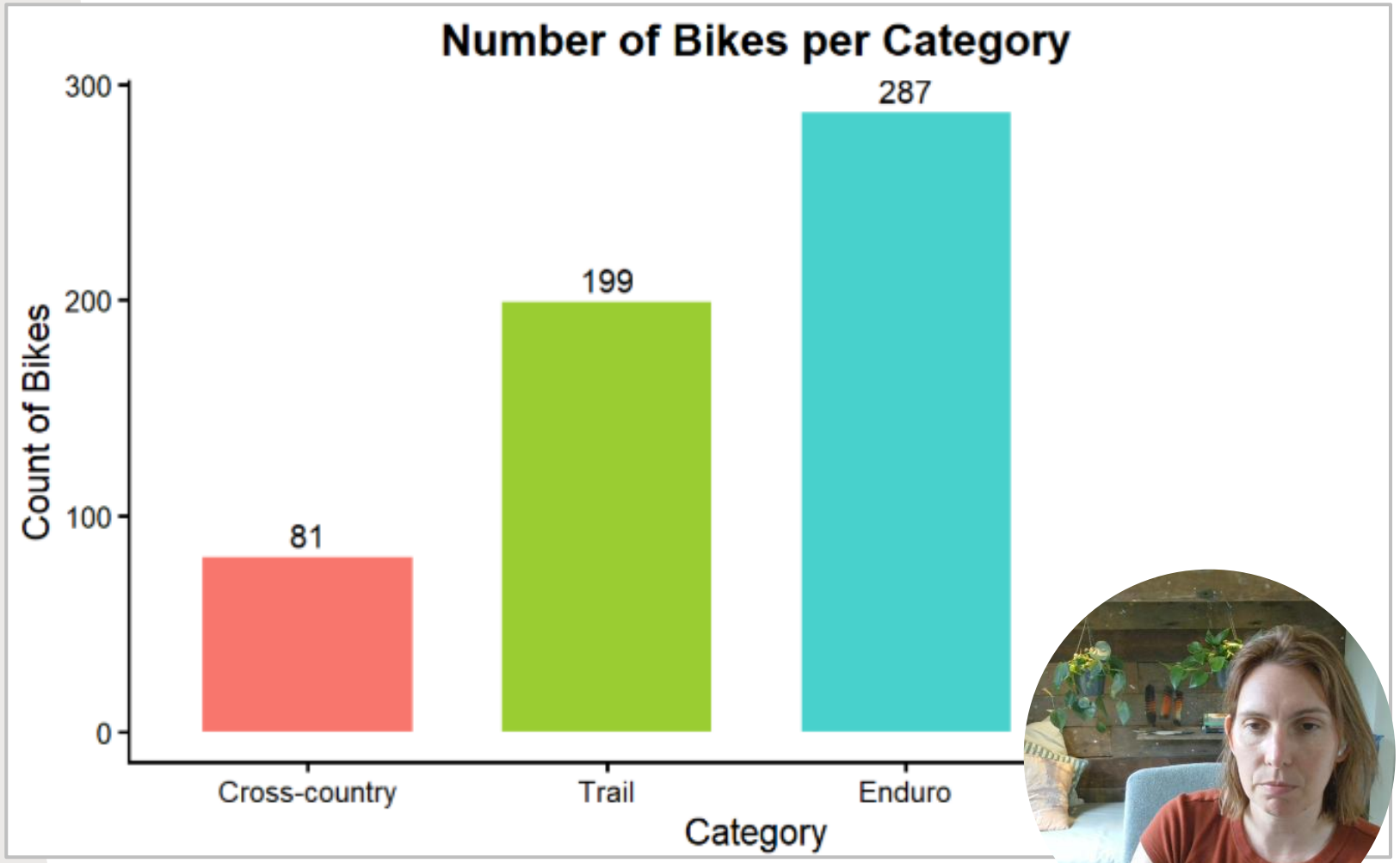


Pseudo-labels for validation

Use front and rear suspension variables to put MTBs into categories.

To be used to compare our clusters later.

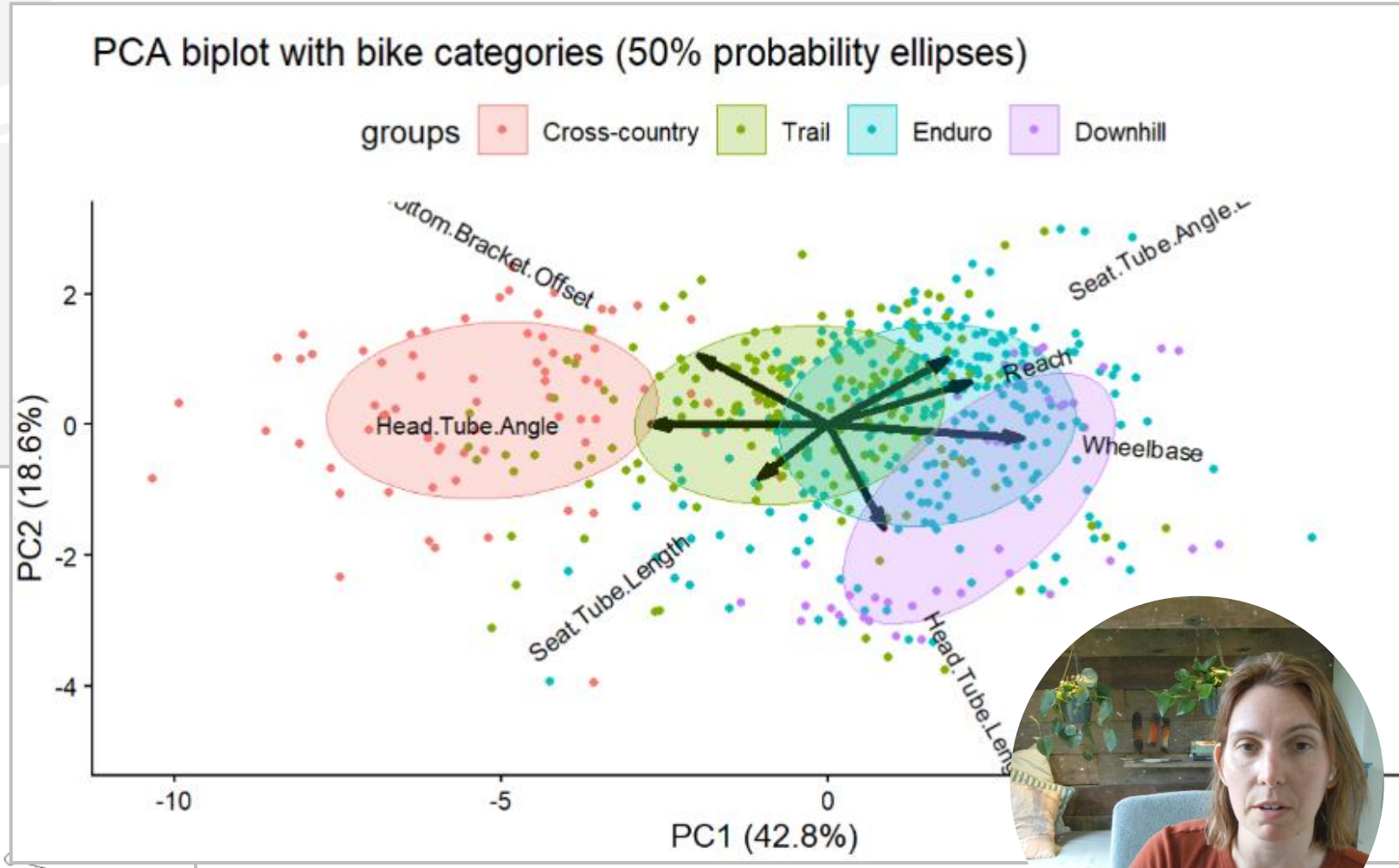
Shorter to longer suspension (left to right) →

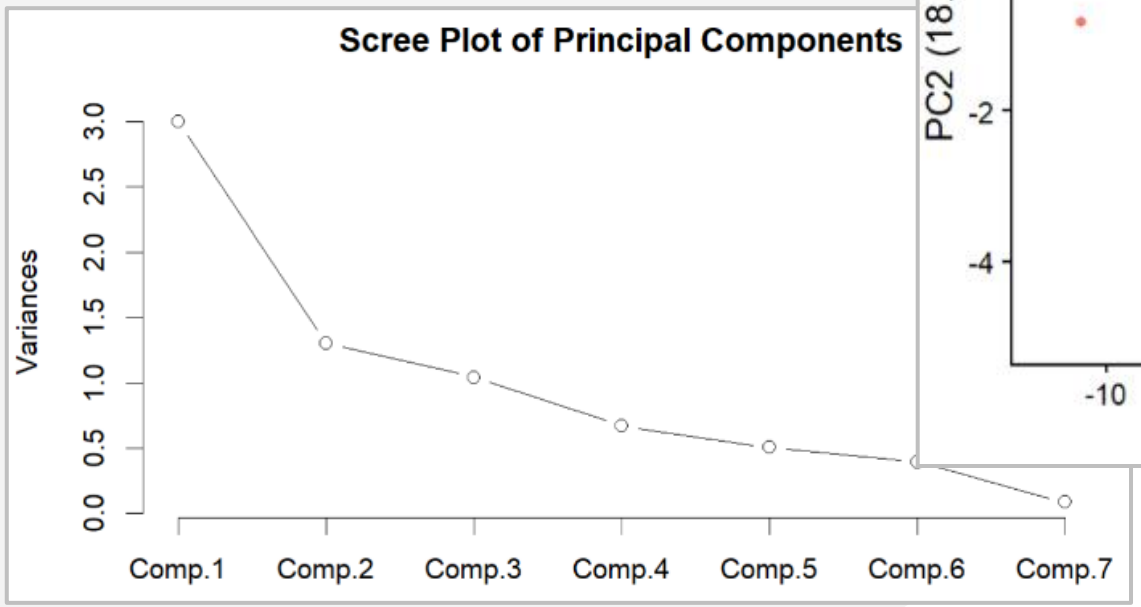


Principal Component Analysis (PCA)

PC1 → Slack: Mainly “reach” , “wheelbase” and “head tube angle”

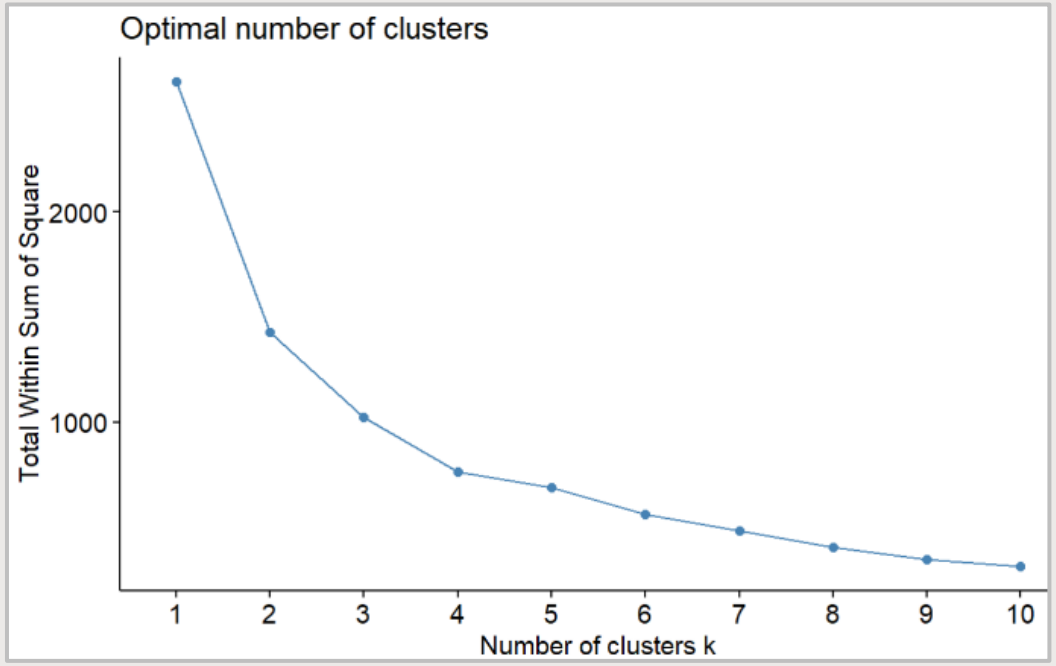
PC2 → Rider Position: Mainly “head tube length”, “seat tube angle” and “bottom bracket offset”

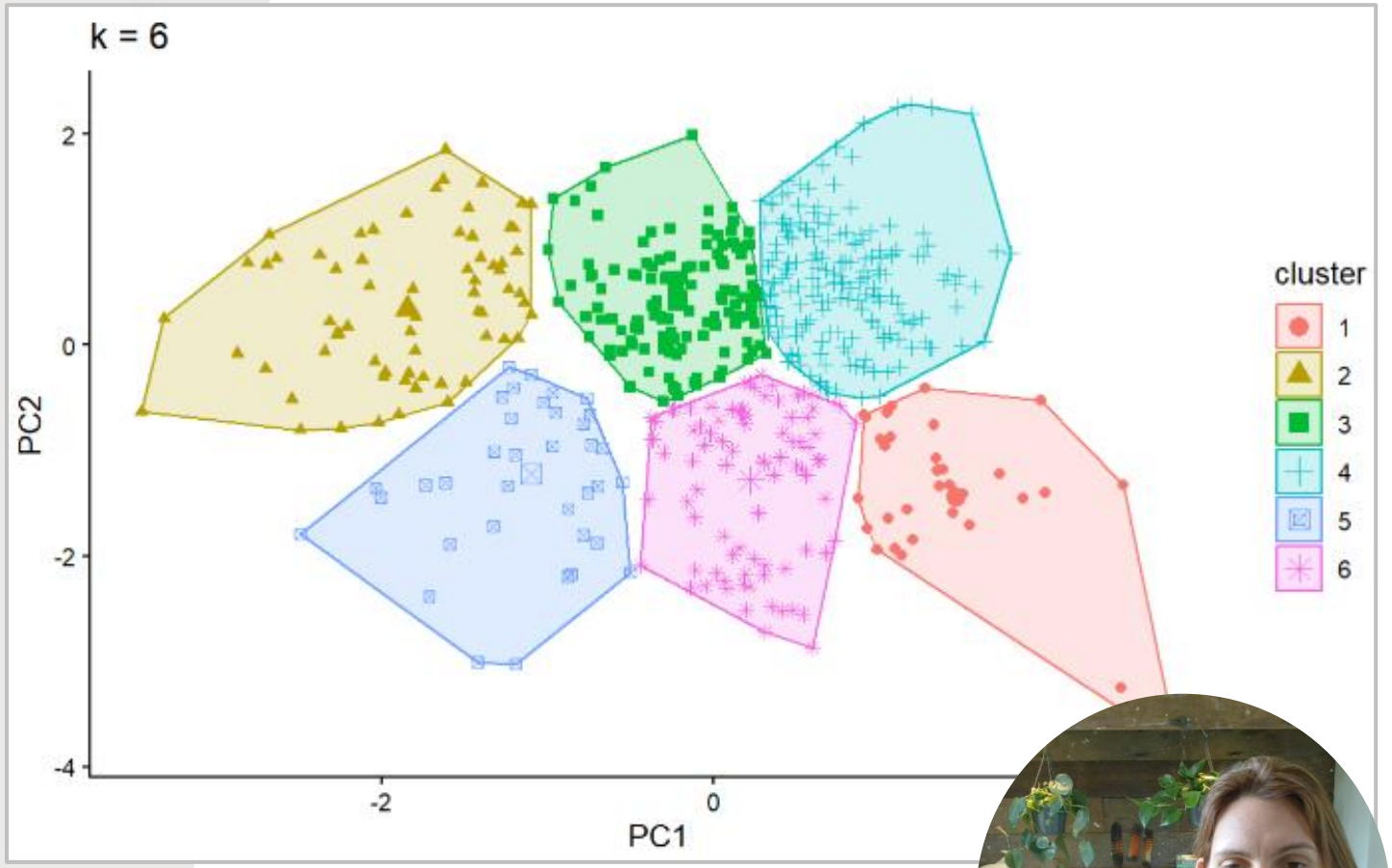




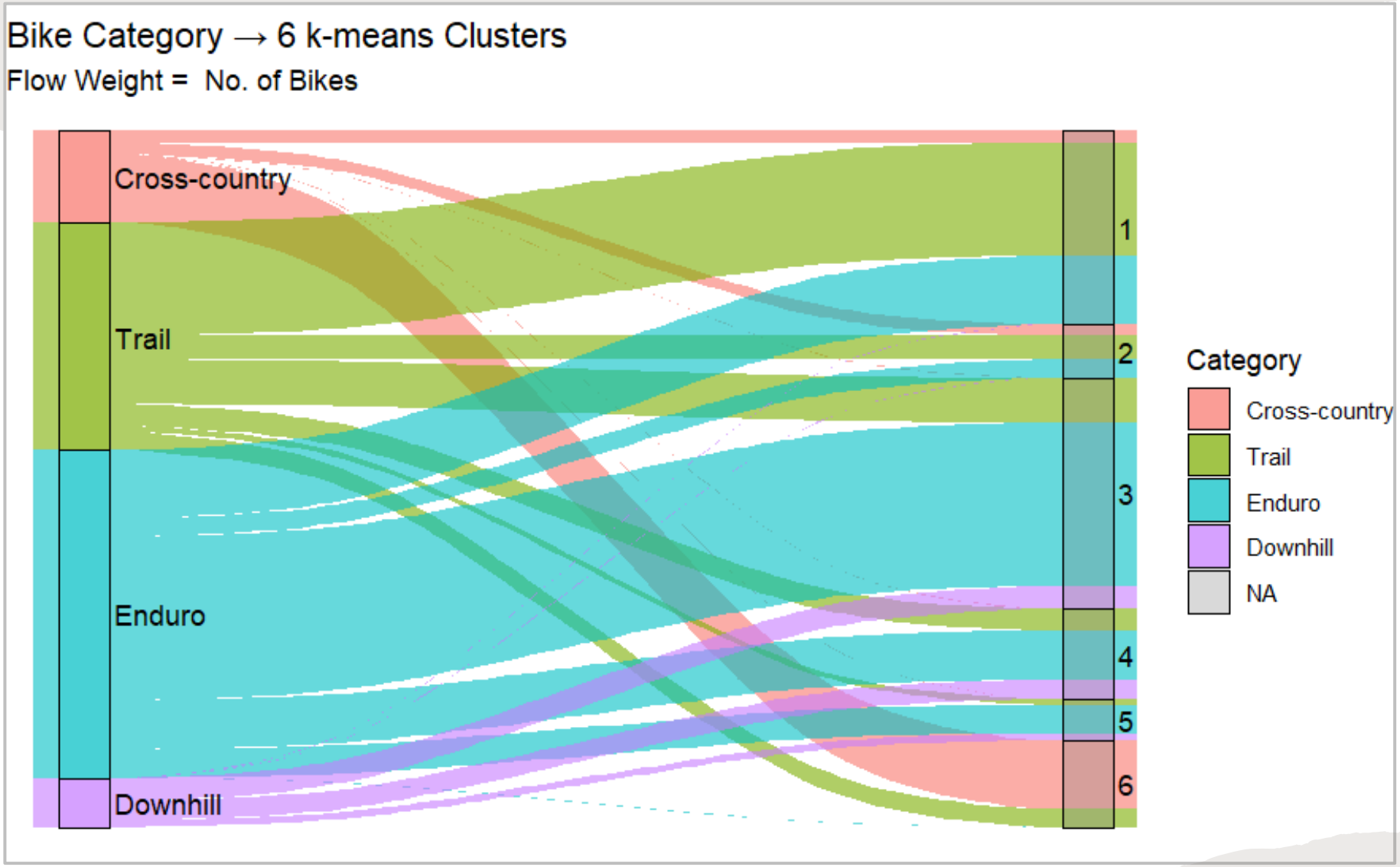
k-Means Clustering

- Performed k-Means Clustering on 2 PCs
- Visual check chose k=6





Validation of k=6 Clustering



Cluster 1: Trail / light Enduro

Cluster 2: Mixed bag (Enduro or park bikes)

Cluster 3: Enduro

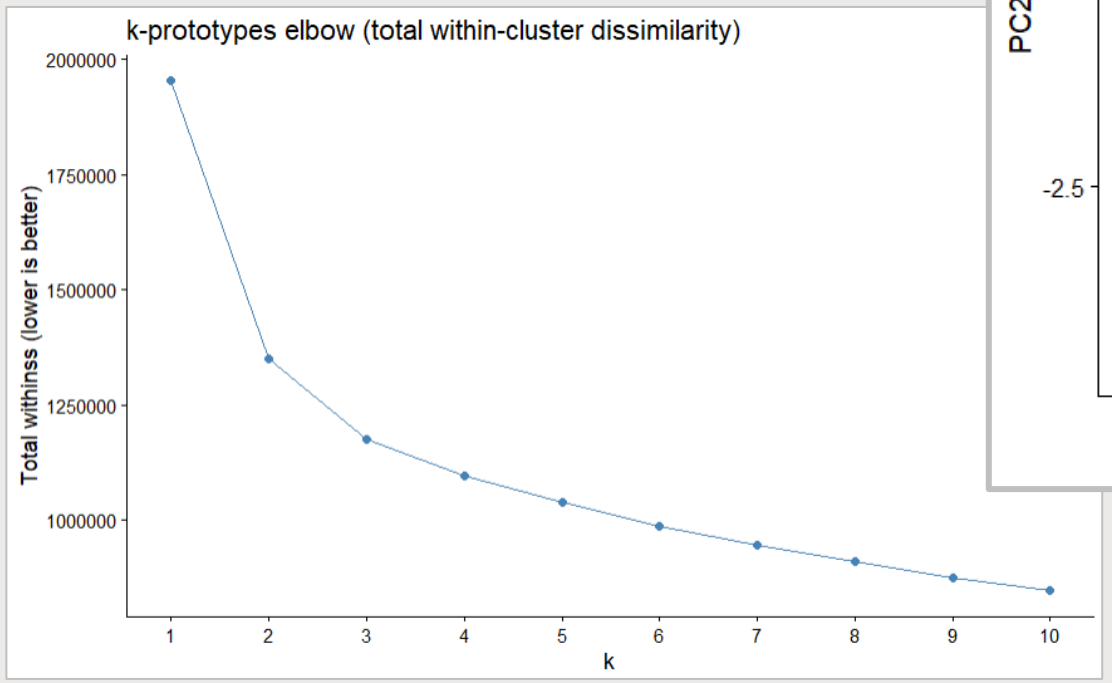
Cluster 4: All Mountain (light Enduro / heavy Trail + Downhill)

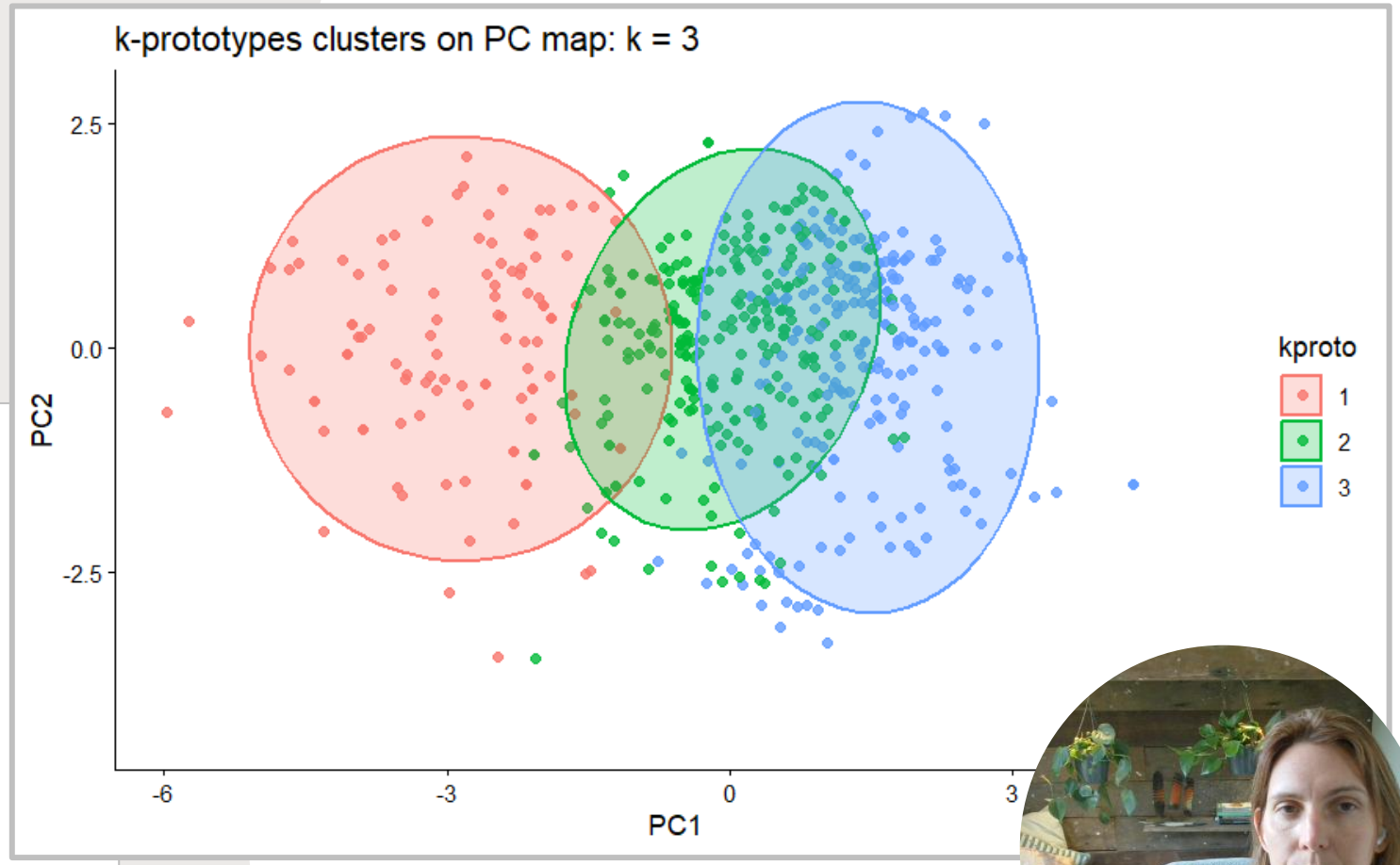
Cluster 5: M
(specialise

Cluster 6:

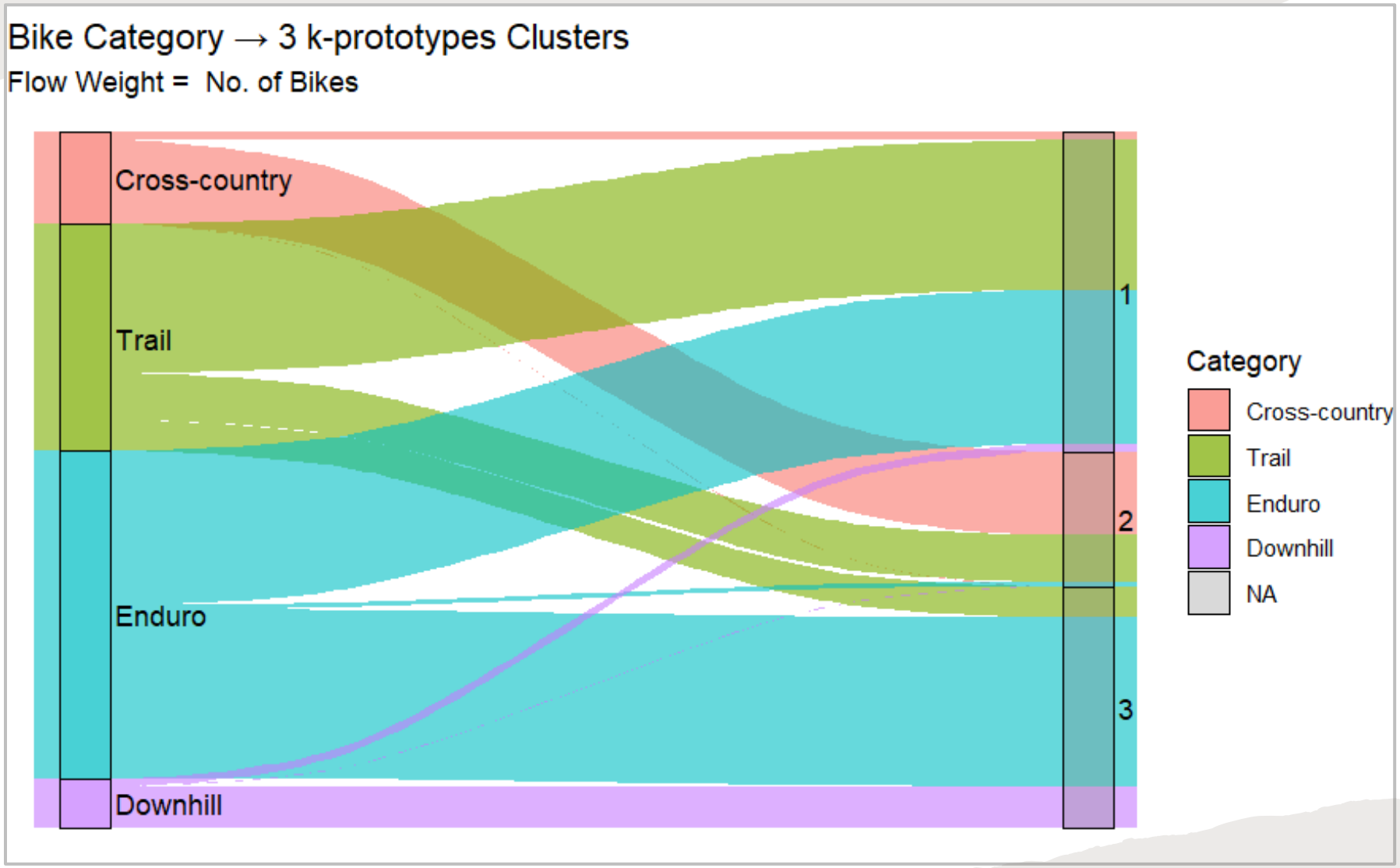
ClustMixType R Package

+ Brand
+ Year
+ Ebike (binary)
+ Wheel Size





Validation of ClustMixType Clustering

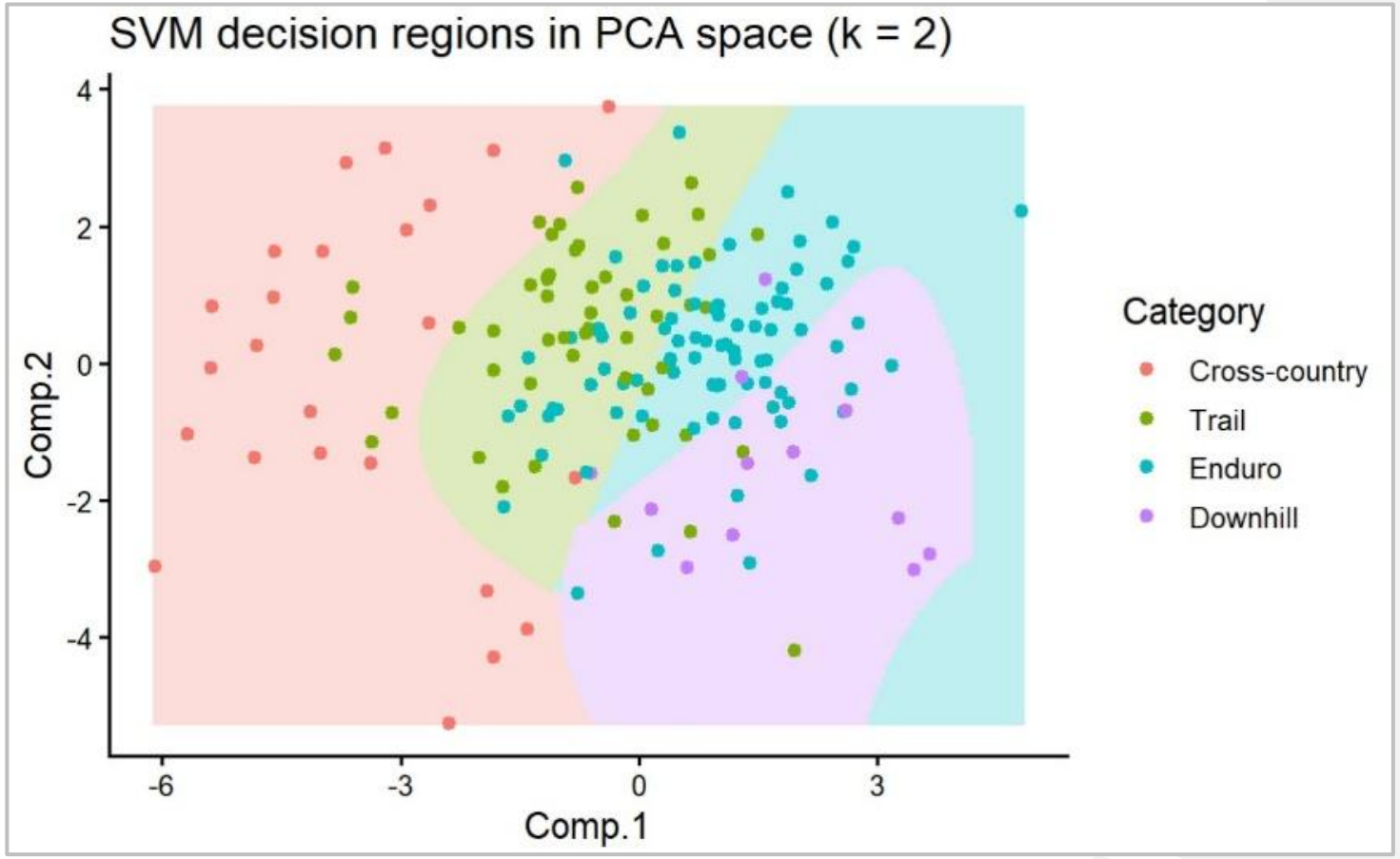


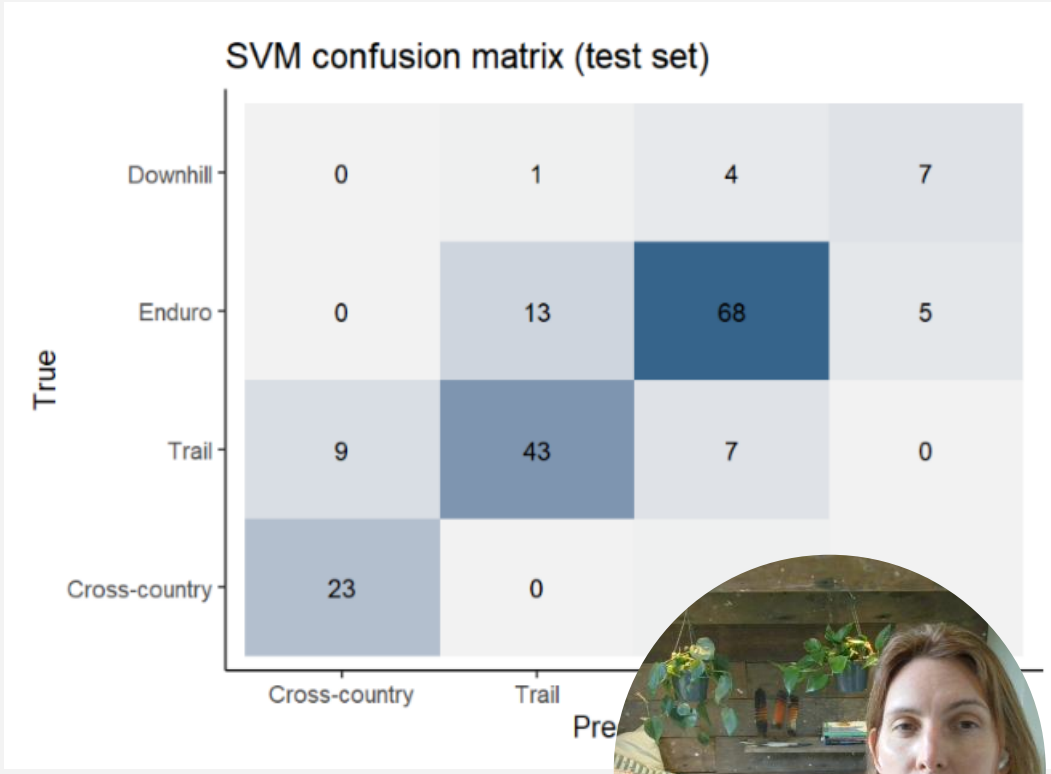
Cluster 1: Trail / light Enduro

Cluster 2: Cross Country

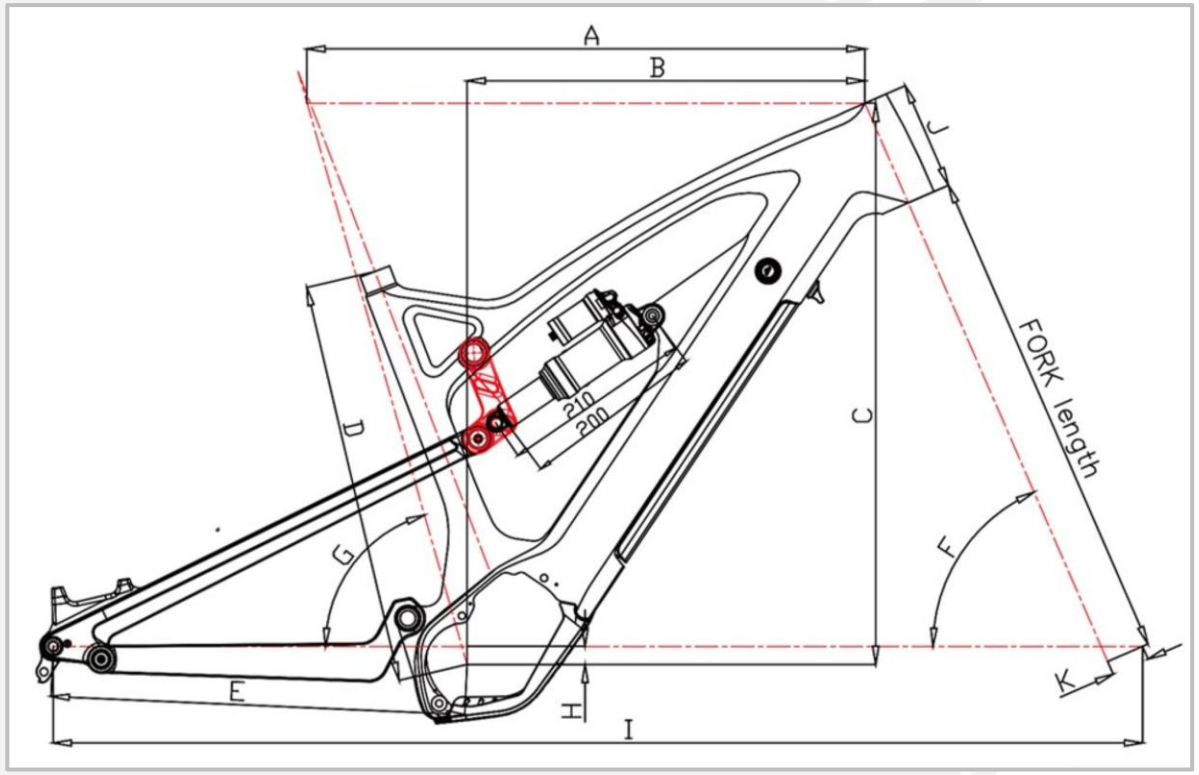
Cluster 3: Enduro / Downhill

Multi-class Support Vector Machine Classification (Extension)





Future Work



The End