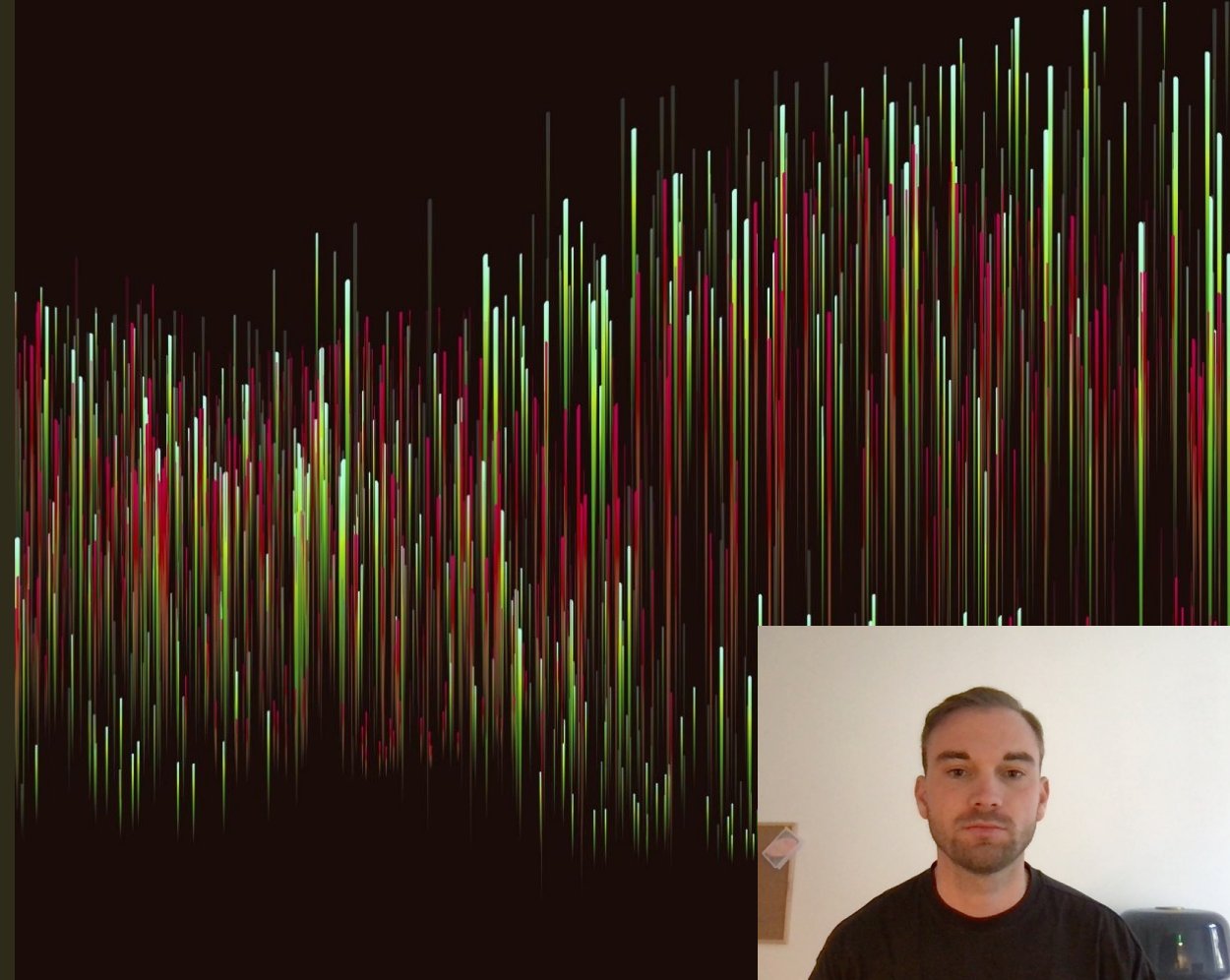
















Applications of Datascience



Preprocessing

< > Messung 180323 Fahrrad 2023-03-18...      

 Synchronisierung mit iCloud angehalten

Name	^	Änderungsdatum	Größe	Art
 Accelerometer.csv		18.03.23, 16:41	33 KB	Comm...et (.csv)
 Barometer.csv		18.03.23, 16:41	16 KB	Comm...et (.csv)
 Gyroscope.csv		18.03.23, 16:41	24 KB	Comm...et (.csv)
 Linear Accelerometer.csv		18.03.23, 16:41	24 KB	Comm...et (.csv)
 Location.csv		18.03.23, 16:41	127 KB	Comm...et (.csv)
 Magnetometer.csv		18.03.23, 16:41	23 KB	Comm...et (.csv)
>  meta		18.03.23, 16:42	--	Ordner
 Proximity.csv		18.03.23, 16:41	16 KB	Comm...et (.csv)



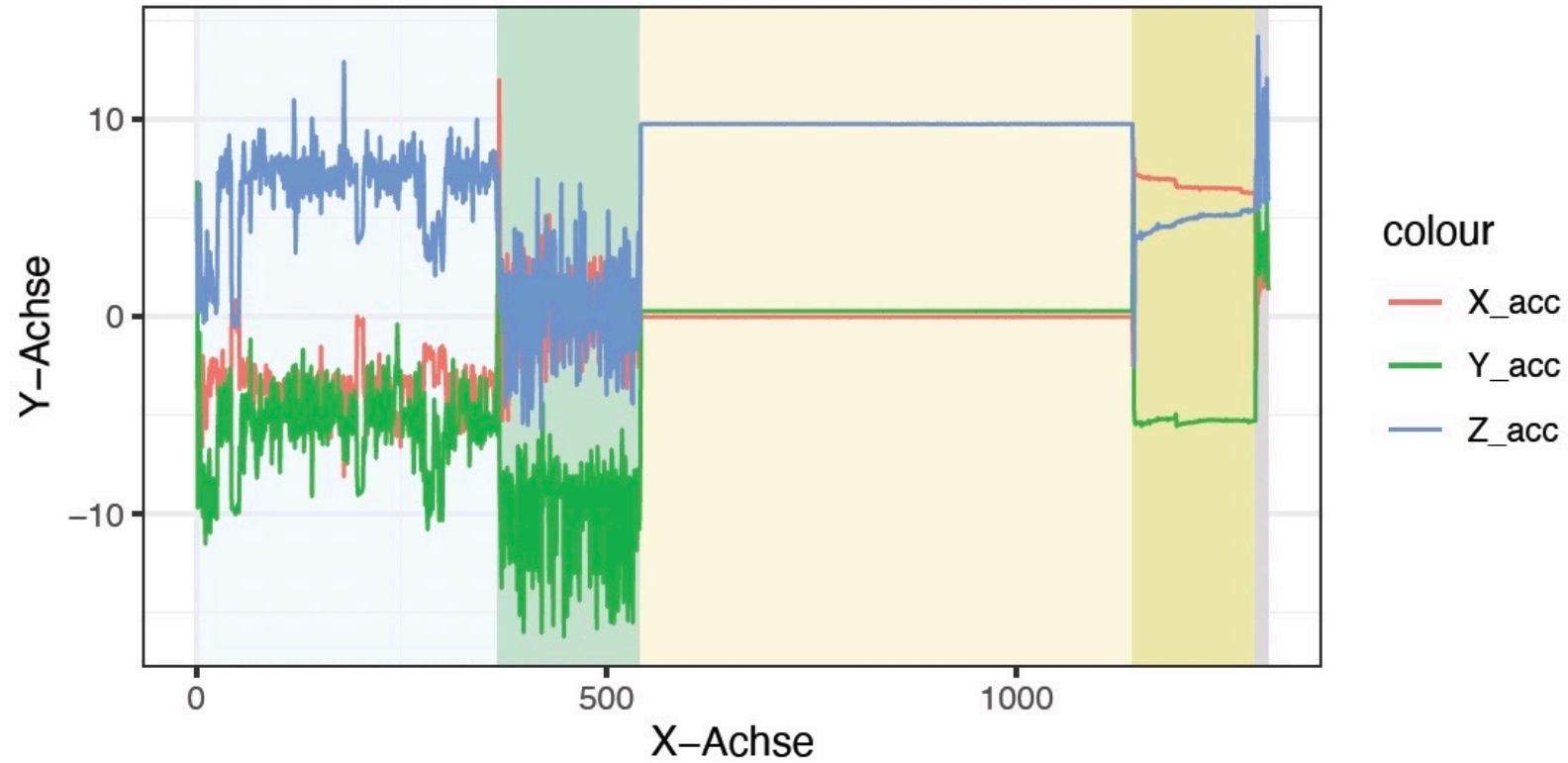
Dataframe I worked with

Head of the Dataframe with 1 column for every sensor

X_acc	X_hPa	X_rad	X_linacc	X_μT	Activity
-3.656746	993.4352	5.40708780	-3.9807026	-28.163292	fahrrad
-1.297652	993.4512	0.19785871	0.4741257	6.101456	fahrrad
-2.451302	993.4512	-0.09114313	-0.1024862	4.411934	fahrrad
-6.856941	993.4512	-0.69596463	2.1904757	3.494606	fahrrad
-2.876269	993.4512	0.22384980	-0.7286204	-4.406799	fahrrad

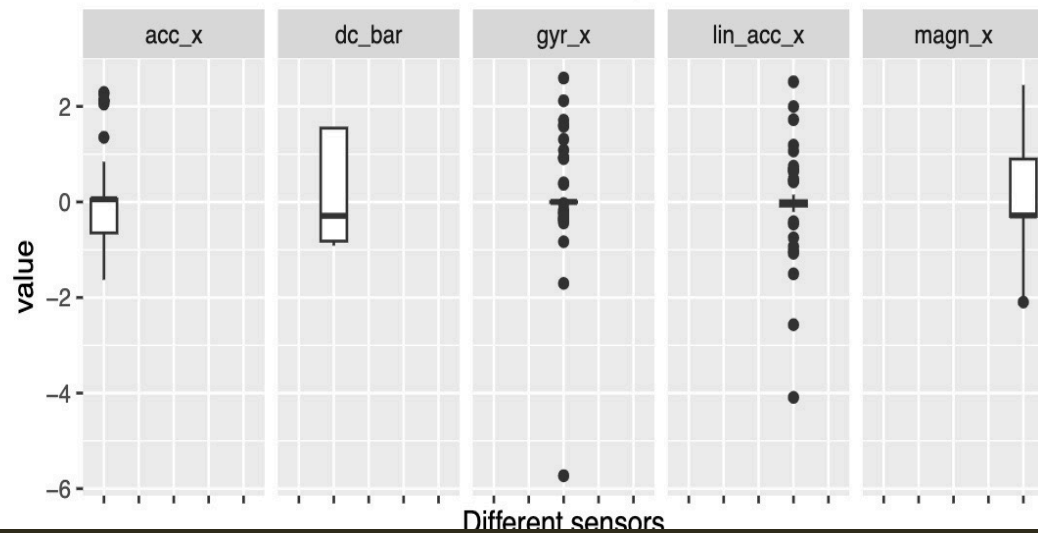


Plot of acceleration data

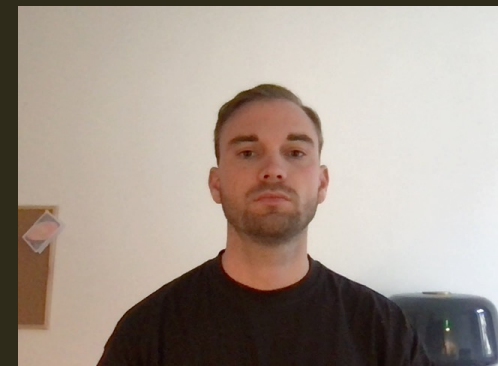
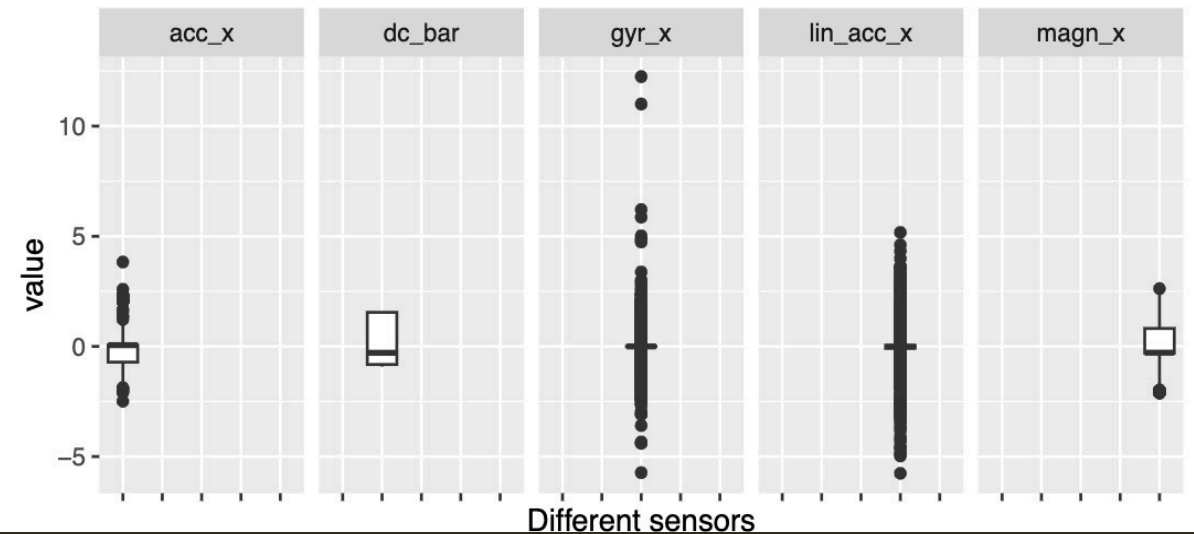


Difference in delta t

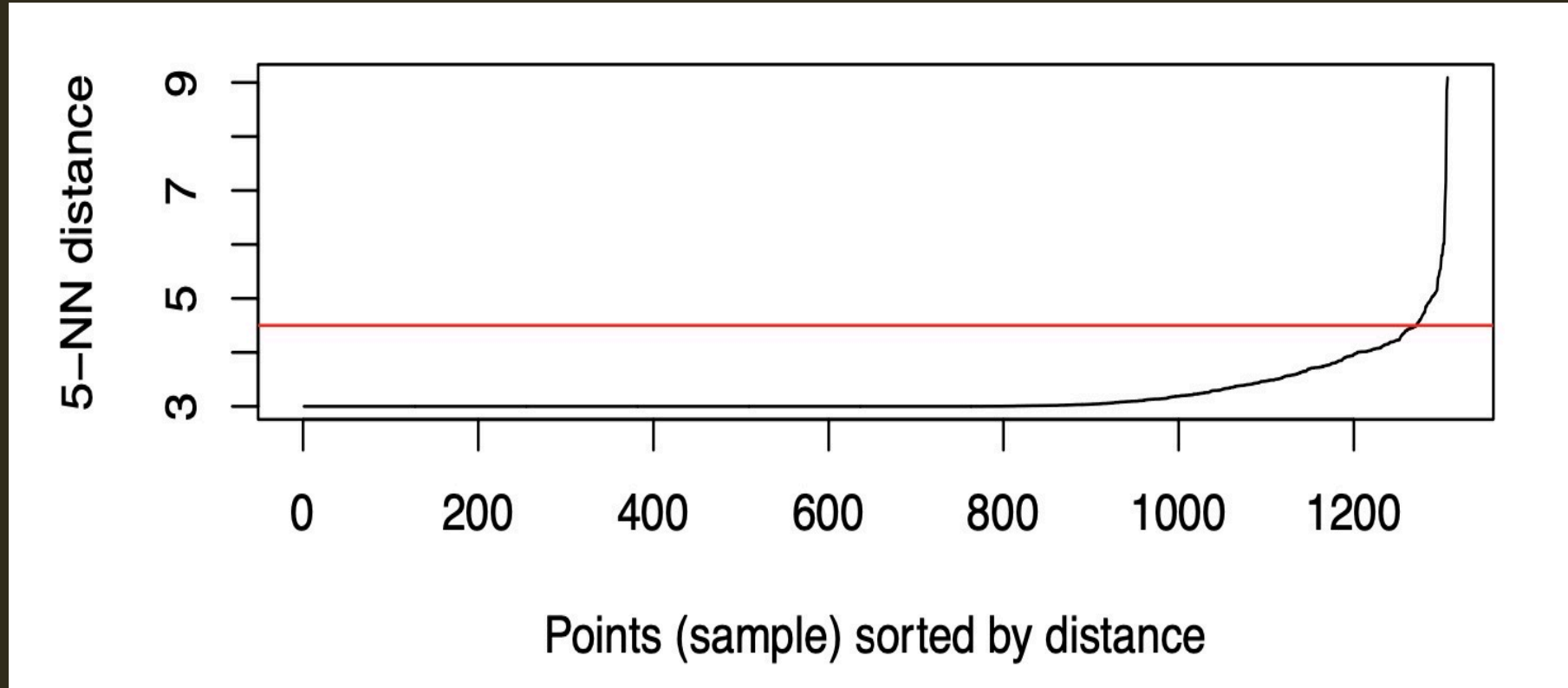
Box Plots of different sensory data – big delta t



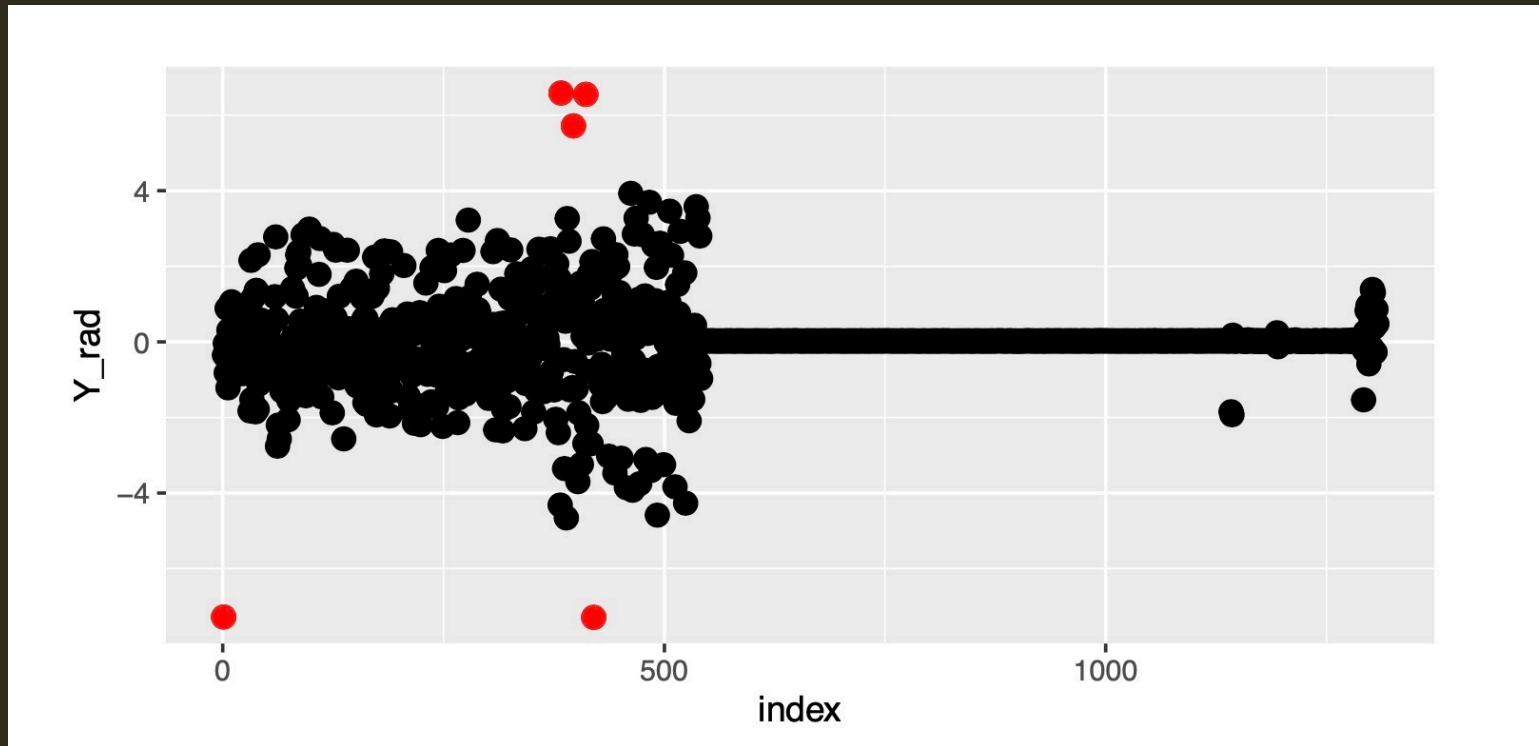
Box Plots of different sensory data – small delta t



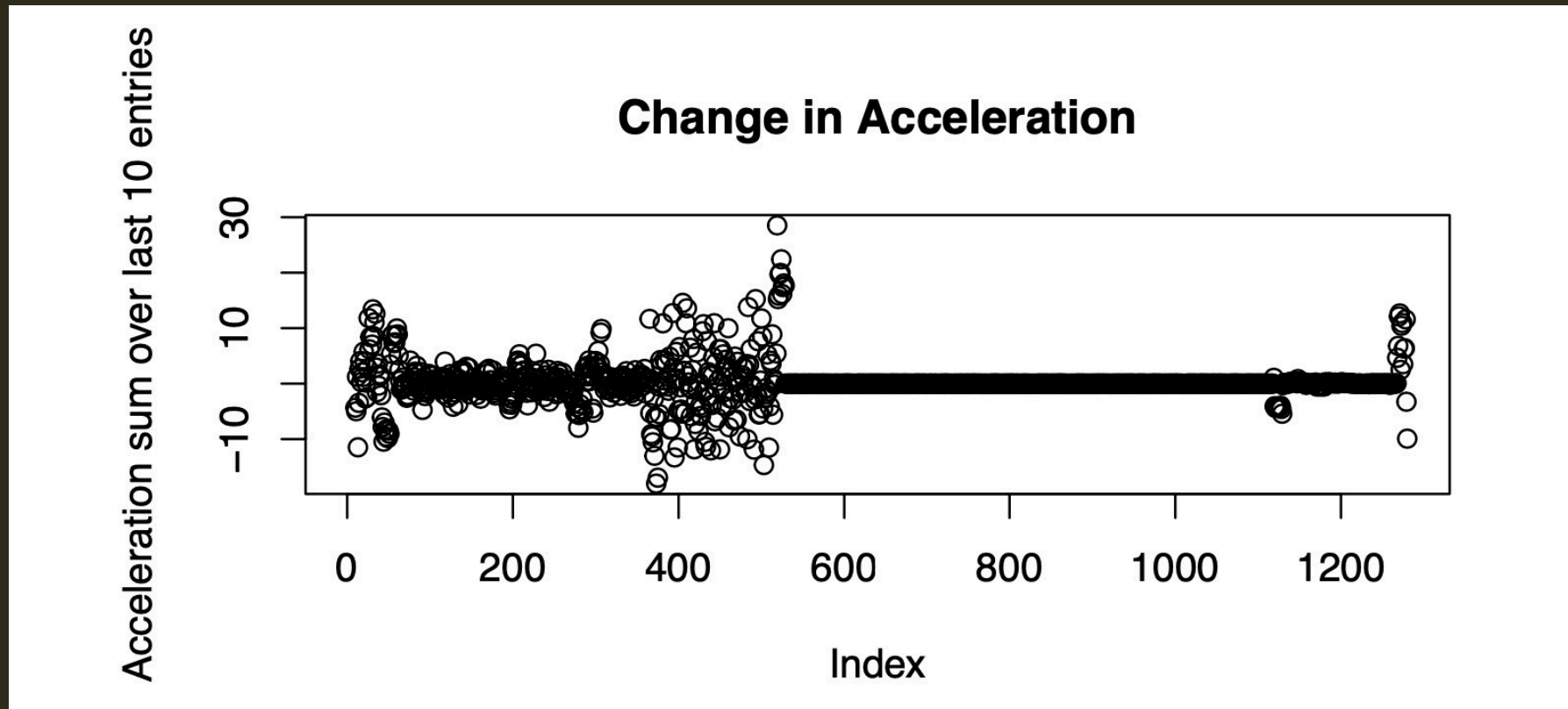
Outlier detection with dbscan - 1



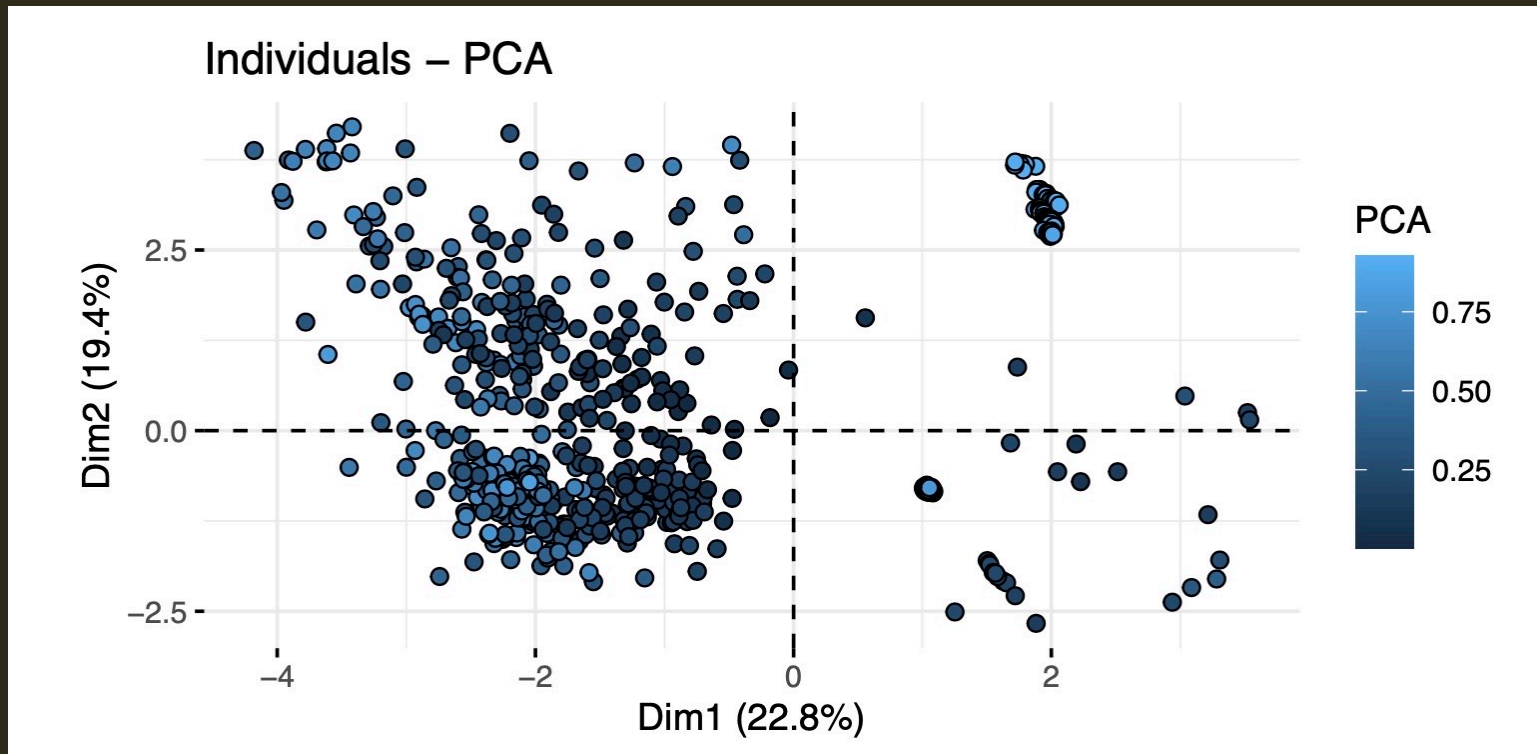
Outlier detection with dbscan - 2



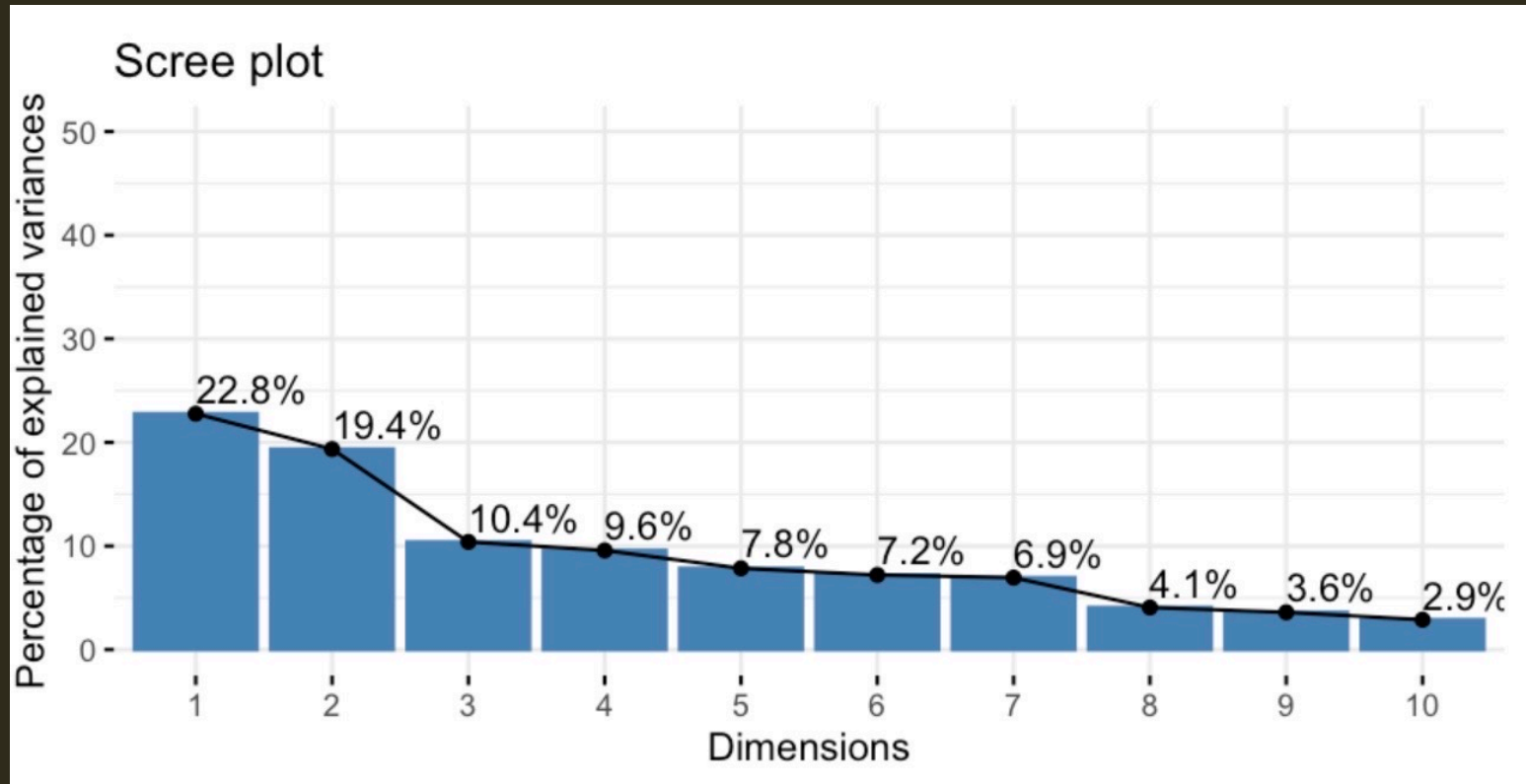
Feature engineering – Temporal Aggregation of acceleration data



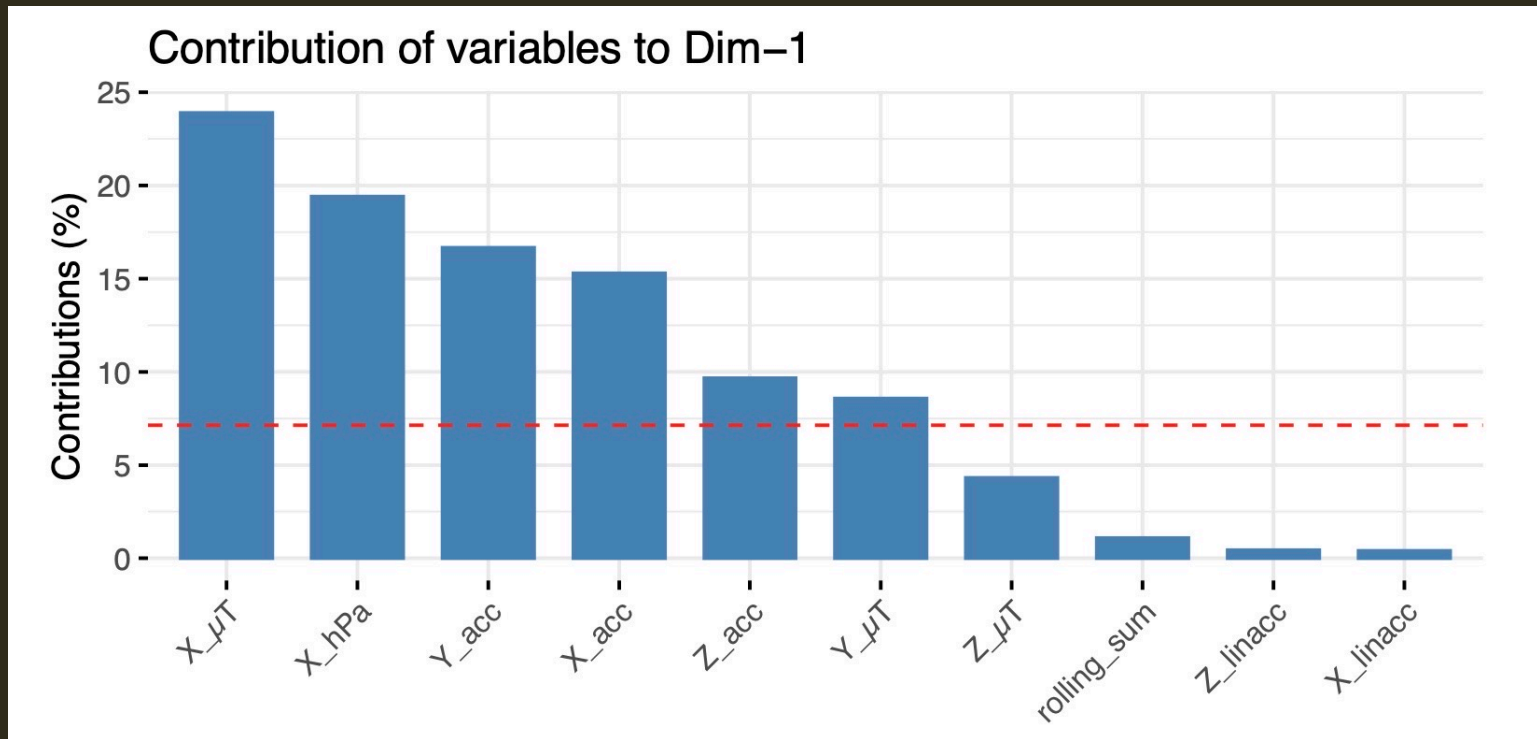
PCA to get a better understanding of the data



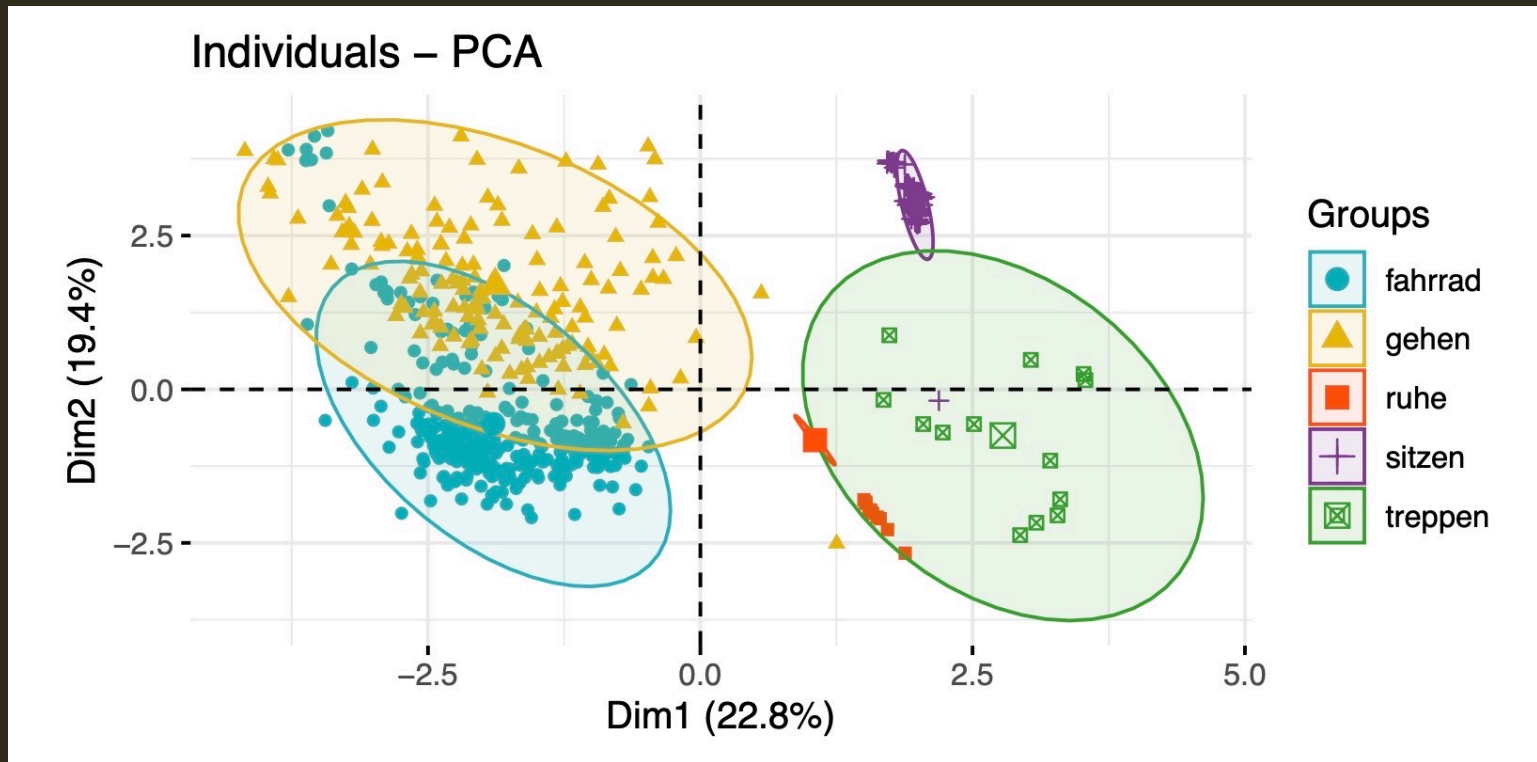
Scree plot



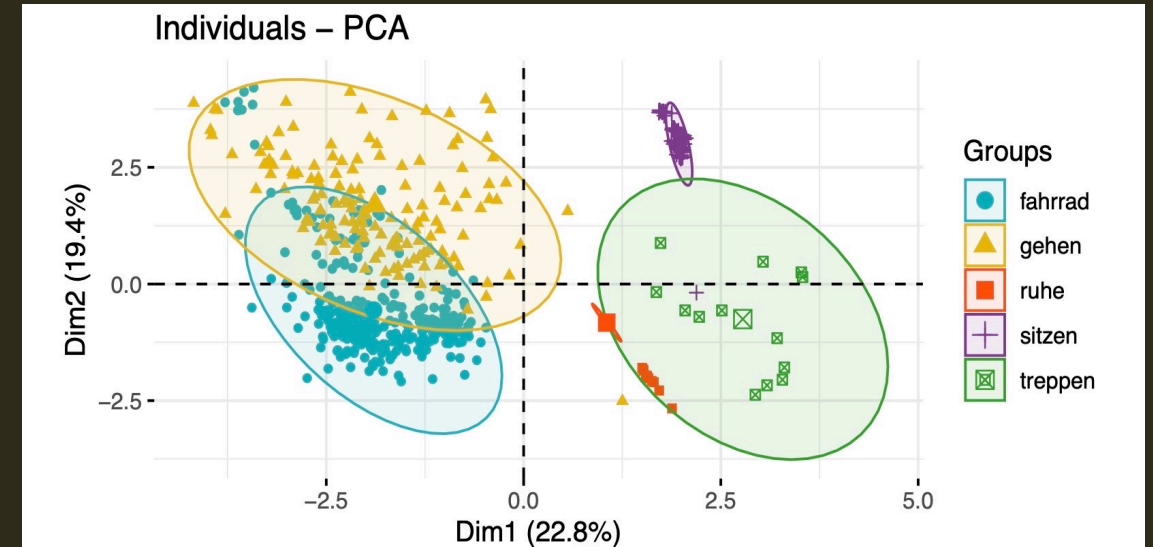
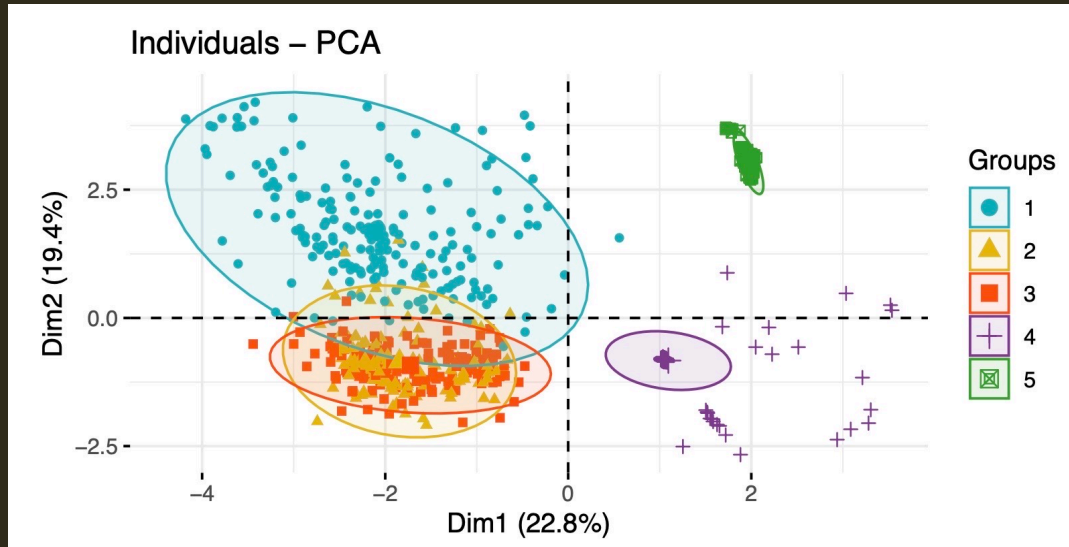
Contribution to principal component



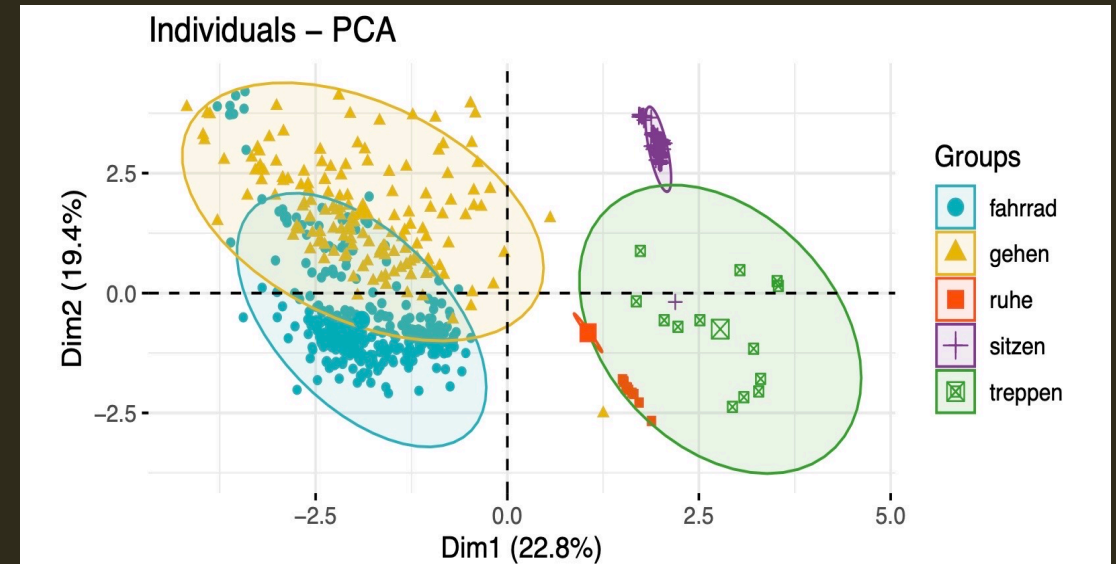
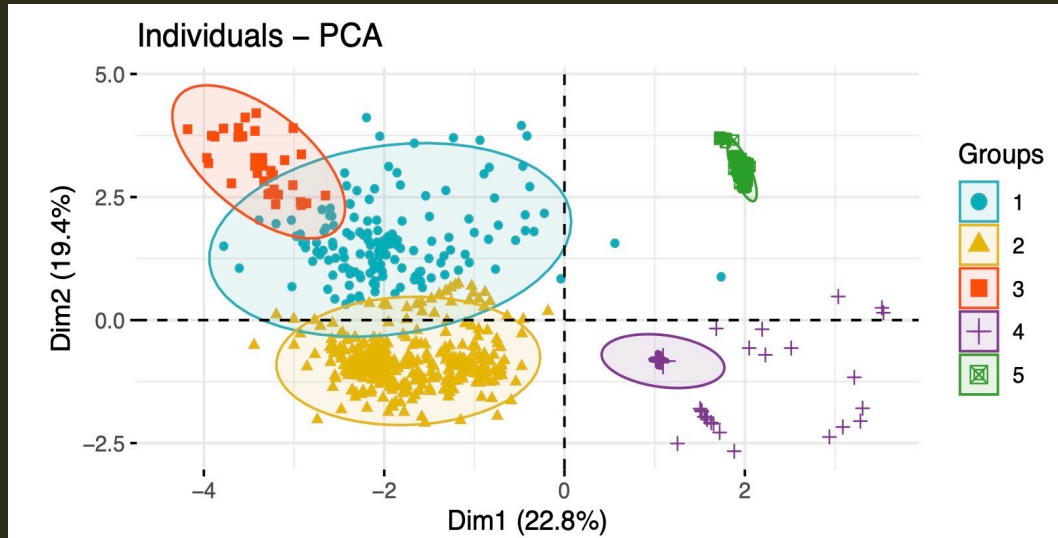
Real Cluster



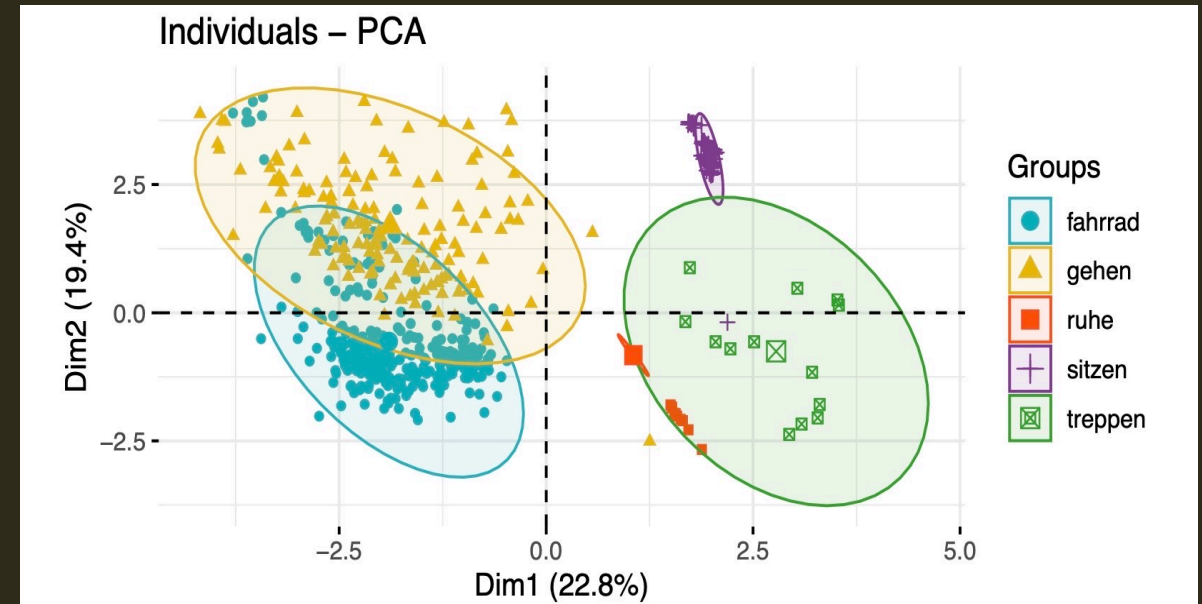
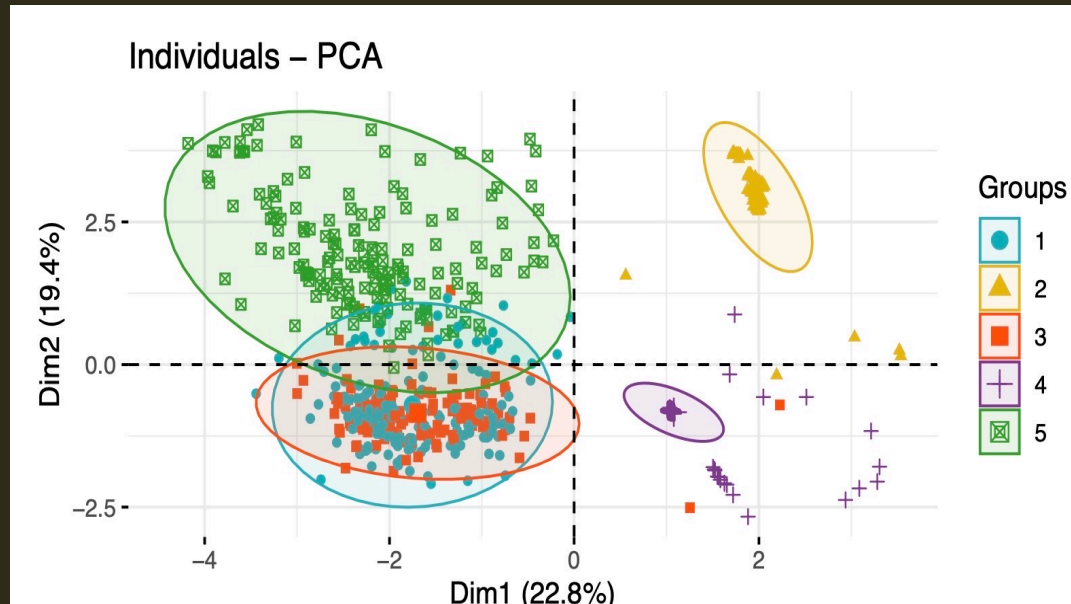
Hierarchical clustering - 1



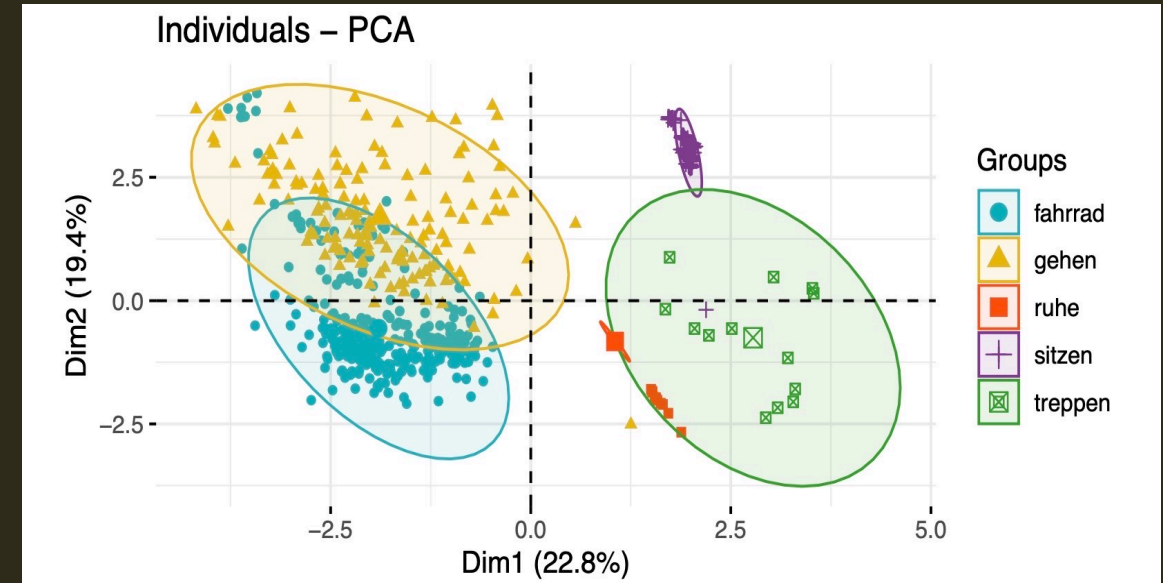
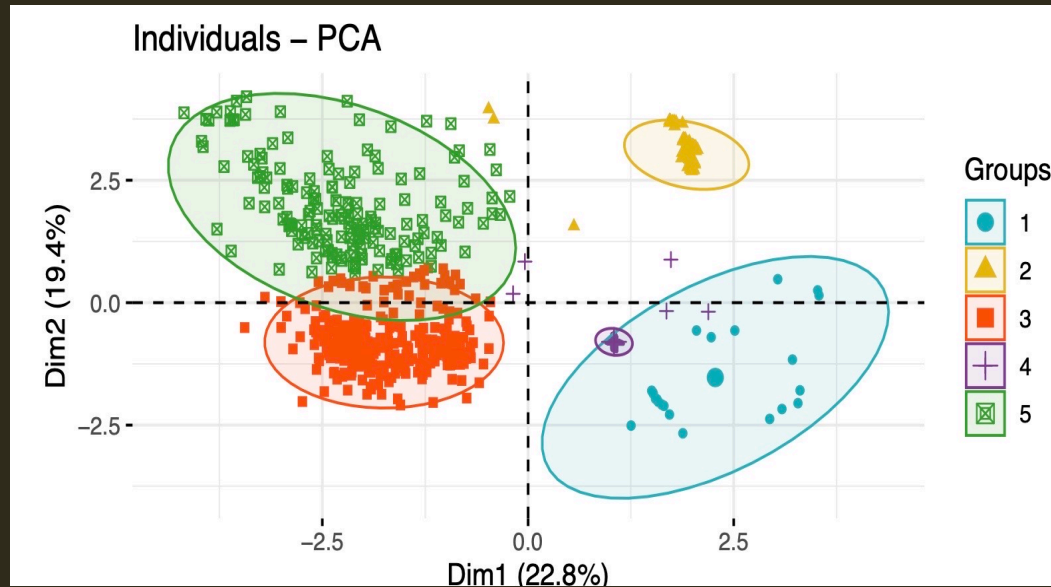
Hierarchical clustering - 2



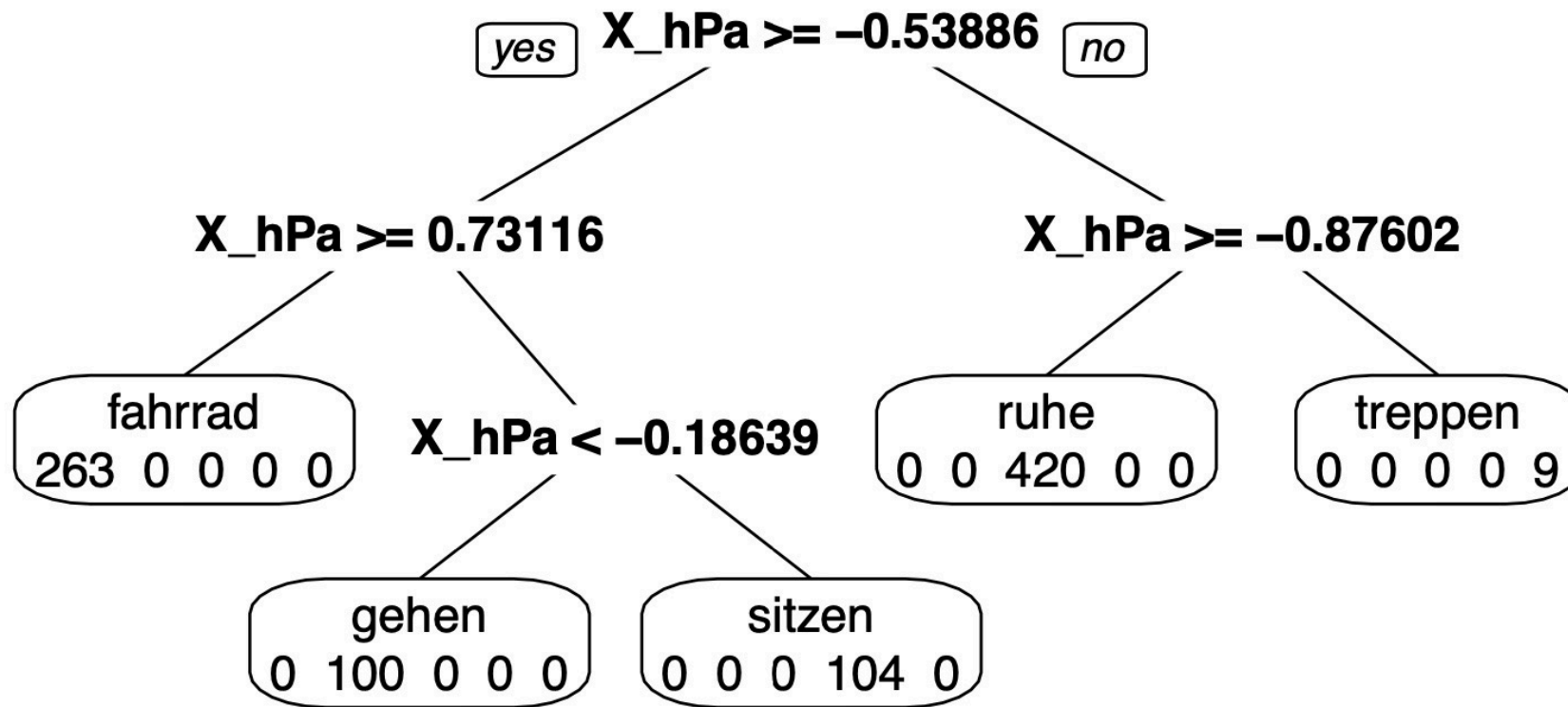
K-means clustering - 1



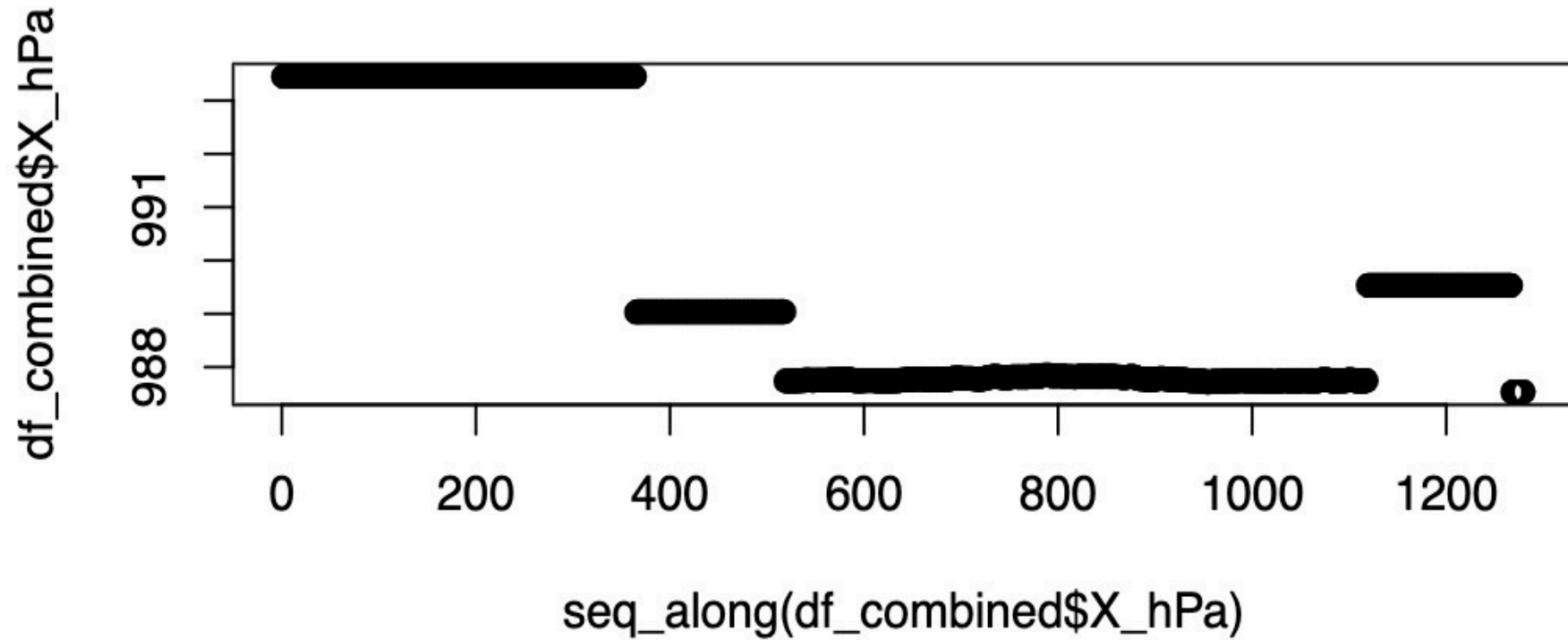
K-means clustering - 2



Classification tree – original data



Exploring reasons



Classification tree – confusion matrix

	fahrrad	gehen	ruhe	sitzen	treppen
fahrrad	101	0	0	0	0
gehen	0	54	0	0	0
ruhe	0	0	180	0	0
sitzen	0	0	0	45	0
treppen	0	0	0	0	4

