CFARG Weekly Report – Dennis & Barry

Date: Friday, 08 December 2023

# What I did this week

This work’s objective was to better understand the data and create some software tools to read and visualise the data. As part of the understanding of the data, we will try to detect pathology from the 6 parameters provided in the PDF reports. We do not anticipate to be able to reliably classify pathology from just these 6 values, but this should act as a baseline for further data analysis. In particular, this week we spent time:

* Analysing the lung function features
  + We considered the 6 parameters (VDP, MSV, TV, VH, VHSS, VHLS) from the PDF reports, and created the 3D correlation graph, Principal Component Analysis (PCA), and basic classifiers.

à We found 3 features (VDP, MSV, VH) perform better in 3D correlation graph and PCA, as shown in the interesting results section.

* Analysing the histograms
  + We examined the histograms and calculated the Heterogenous Disease (HD) value for each histogram.

à We will have further investigations on HD values.

* 3D visualization
  + We created a 3D visualization of the lungs from the specific ventilation csv files.

à This might help the doctors to visualise the lungs, and locate the affected region easier.

# My administrative tasks

* + N/A

# Paper writing

* + N/A

# Things For Next Week

* Analysing histograms:
  + Extract more metrics from the histograms:
    - Mean, Median, Mode, Skewness, Kurtosis (Shape)
    - Quartiles, Interquartile Range, Variance and Standard Deviation (Spread)
    - Signal to Noise Ratio, Entropy (Randomness)
  + Try out some novel ideas like vectorization to analyse the data.

# Future Plans

* Analysing 3D data:
  + Use image feature extraction with traditional Machine Learning classification approach
  + Use Auto-encoder, features extraction, dimension reduction, etc. to find interesting results.
  + Use techniques for the time-series data (from Sheep Asthma Study).
* 3D visualization:
  + Create more diverse options to help doctors with the graphs (need more insights from doctors).

# Issues Limiting My Activity

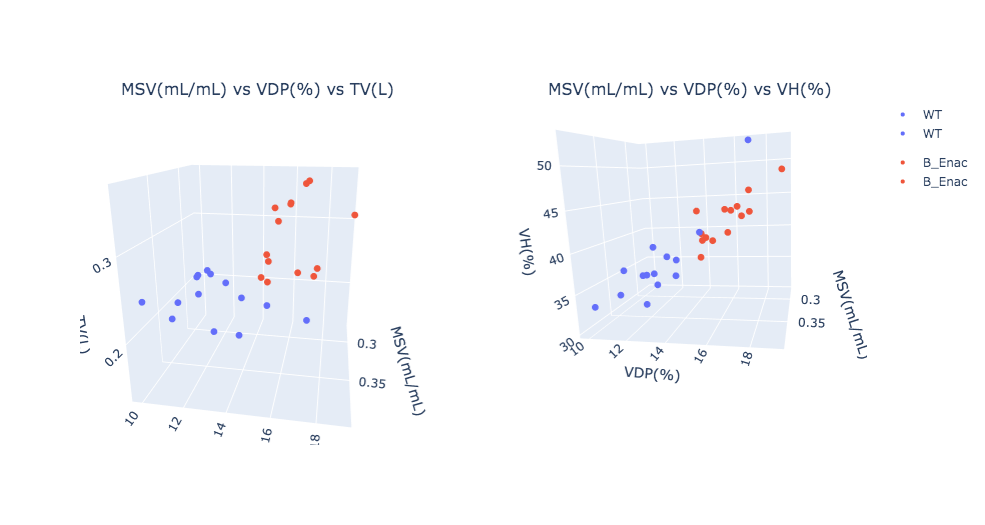
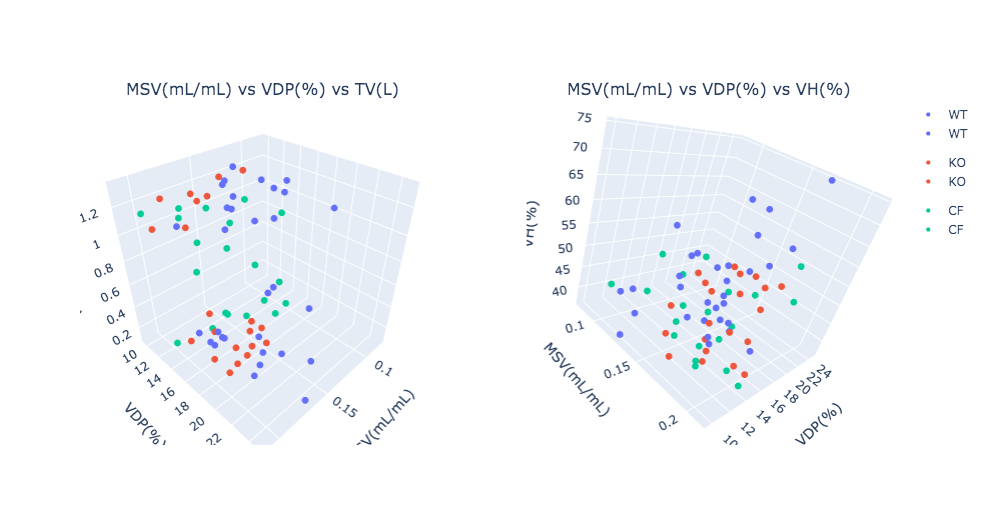
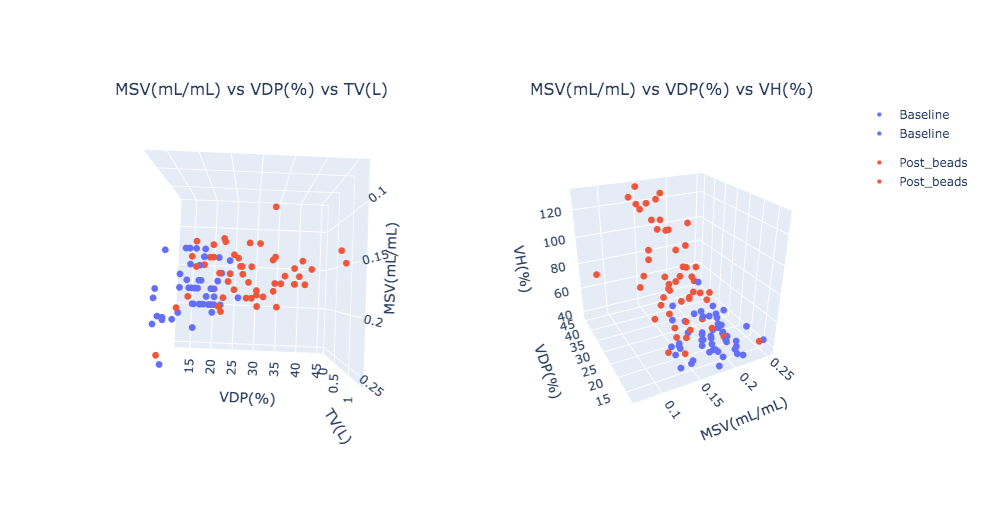
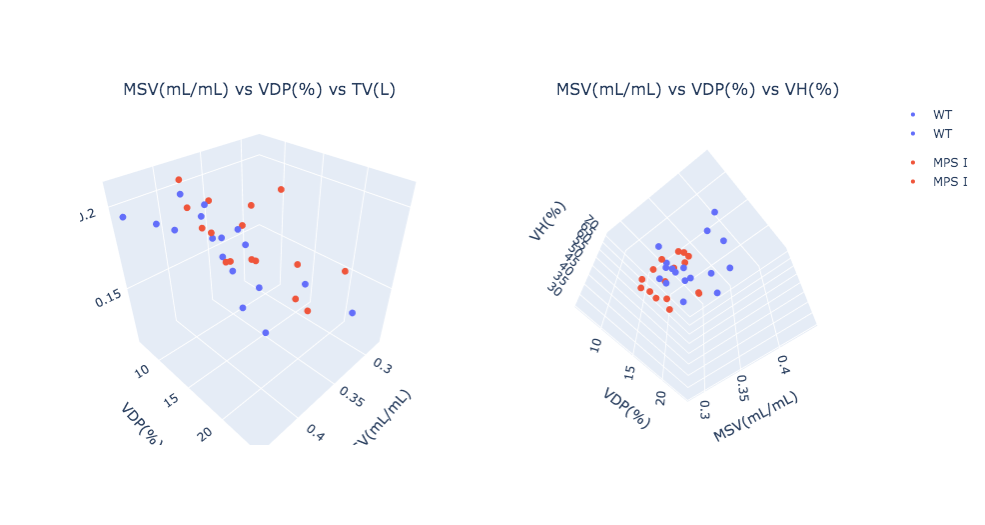
* Limited dataset

# Laboratory issues

* N/A

# Interesting results

3 Features Correlation Graph: Separation of clusters can be observed in ¾ studies: Mouse MPS, Rat Sterile Bead, and Mouse B-ENaC study. The result of Rat PA Study might be affected by other factors like weight, sex, age, etc.



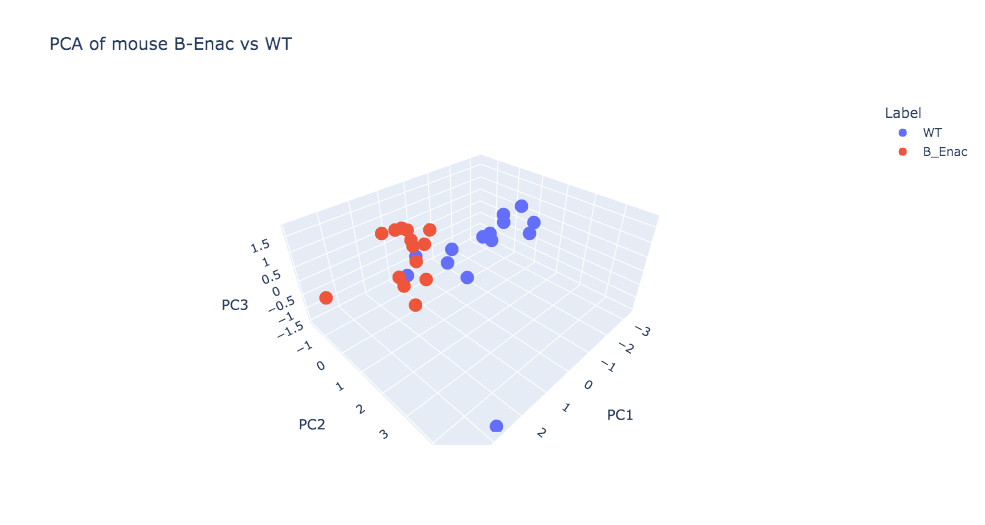
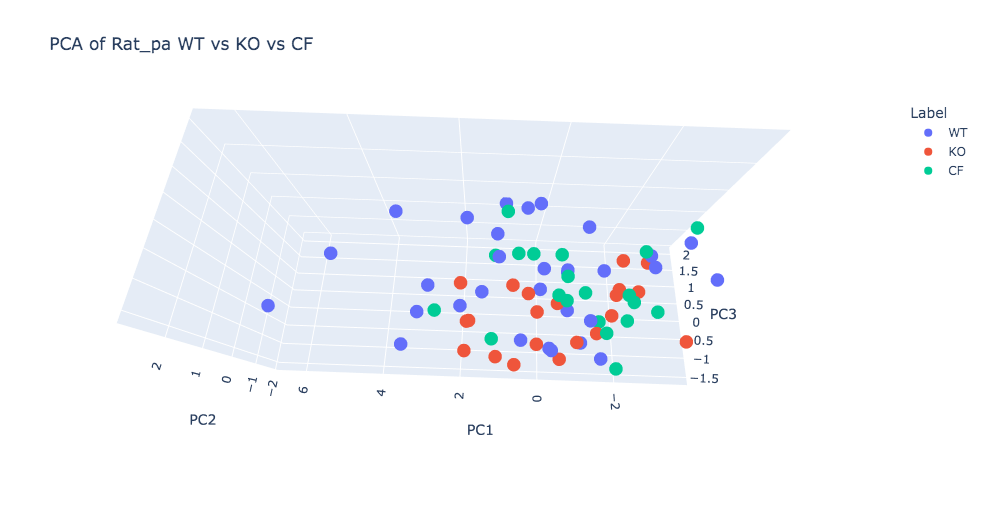
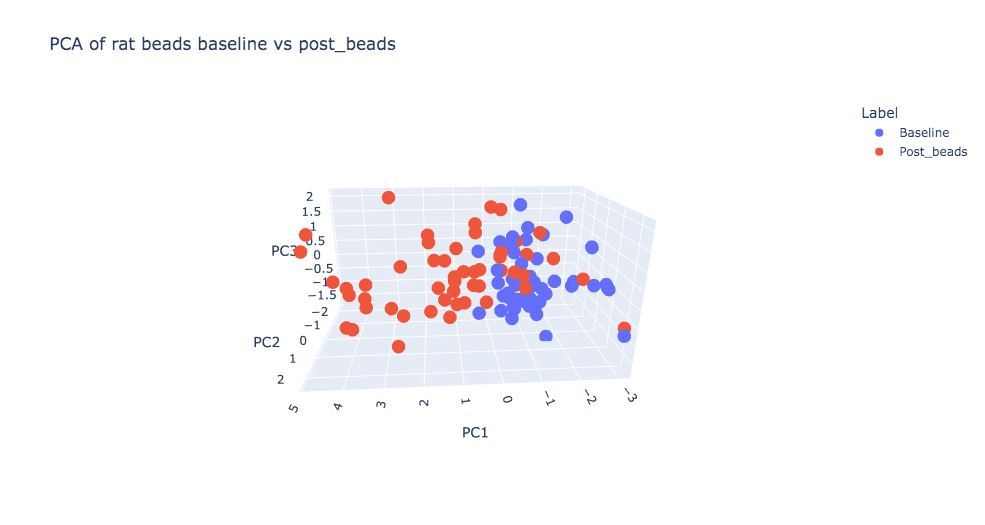
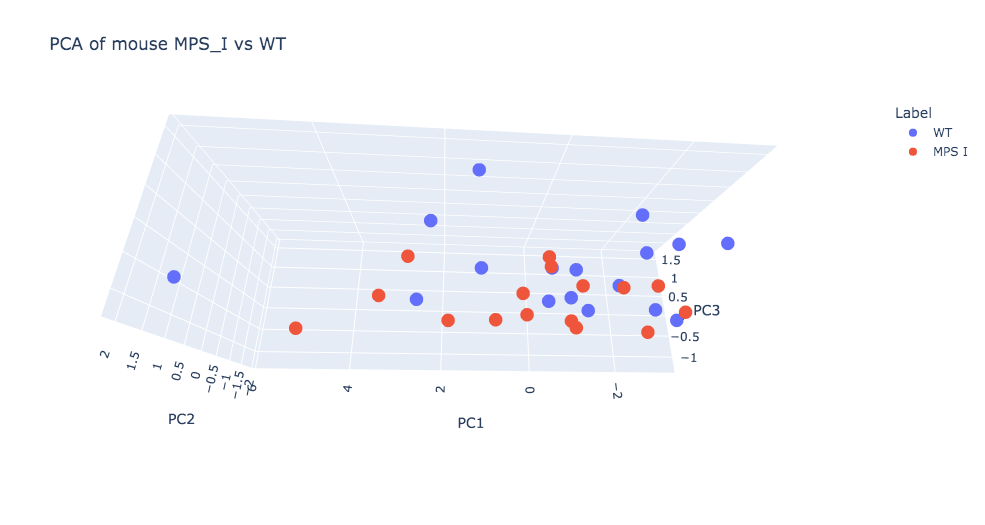
Rat PA Study

Mouse B-ENaC Study

Rat Sterile Bead Study

Mouse MPS Study

3D Principal Component Analysis (PCA): 6 features were used for PCA and we extract PC1, PC2 and PC3. Similar results can be observed.

****