

# An R workflow for lipidomics data

Lisa Schneider

June 14, 2020

# Data creation

## Samples

- ▶ Grass fed cattle
- ▶ Stall fed cattle
- ▶ Deer
- ▶ Roe deer
- ▶ Salmon (farmed)

## Standards

- ▶ Equi Lipido SplashMix
  - ▶ deuterium labelled standards with lipid ratios similar to human plasma

## LC-MS/MS

- ▶ positive ionization mode

## Feature selection and lipid identification

Using Thermo Scientific TraceFinder \* identification via exact mass  
\* positive mode

What the data looked like

```
## knitr
```

```
## dplyr
```

```
## stringr
```

```
## data.table
```

```
## impute
```

```
## imputeLCMD
```

```
## ggplot2
```

```
## tibble
```

## Slide with R Output

```
## knitr
```

```
## dplyr
```

```
## stringr
```

```
## data.table
```

```
## impute
```

```
## imputeLCMD
```

```
## ggplot2
```

```
## tibble
```

```
## viridis
```

```
## GGally
```

## Slide with Plot

```
## knitr
```

```
## dplyr
```

```
## stringr
```

```
## data.table
```

```
## impute
```

```
## imputeLCMD
```

```
## ggplot2
```

```
## tibble
```

```
## viridis
```

```
## GGally
```

## Next Slide

- ▶ Theory
- ▶ My Workflow
  - ▶ Create data matrix
  - ▶ Log-Transformation
  - ▶ Normalization
  - ▶ Imputation
  - ▶ Exploratory data analysis
  - ▶ Univariate analysis
  - ▶ PCA
  - ▶ Hierarchical clustering
- ▶ What next?
  - ▶ Supervised learning
  - ▶ Lipid set enrichment analysis