

# STA 440 Final Project: Code Script

Eunice Lee

## Package Import & Data Cleaning

```
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(tidyr)
library(readr)
library(ggplot2)
```

```
load("REE.RData")
```

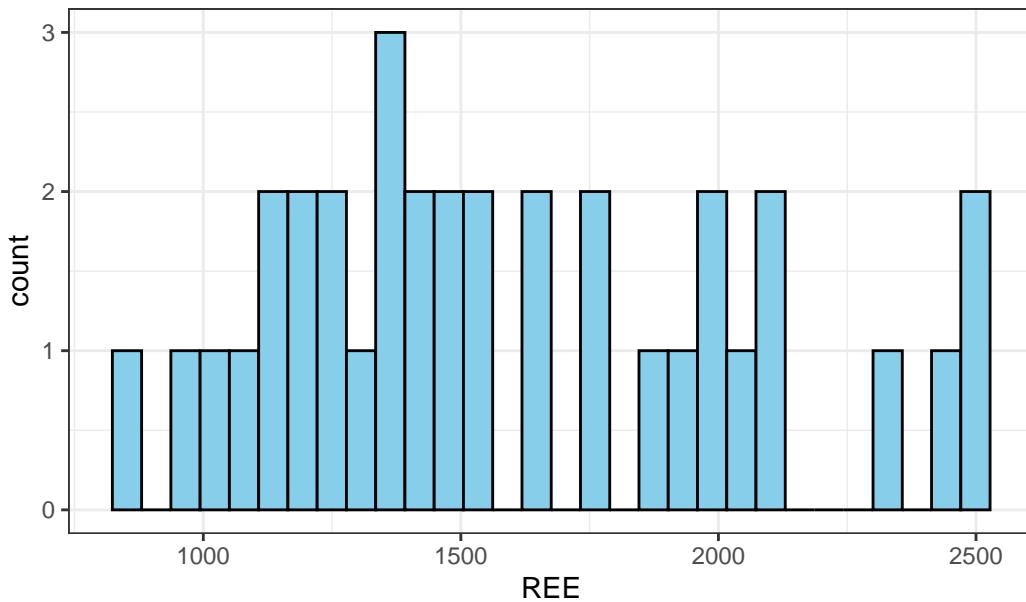
```
REE <- REE %>%
  mutate(
    ID = factor(ID),
    Obese = factor(Obese, levels = c(0, 1)),
    Ventilator = factor(Ventilator, levels = c(0, 1)),
    ICUorSDU = factor(ICUorSDU, levels = c("ICU", "SDU")),
    Measurement = as.numeric(Measurement)
  )
```

## EDA

```
# Histogram of REE
ggplot(REE, aes(x = REE)) +
  geom_histogram(fill = "skyblue", color = "black") +
  labs(title = "Distribution of REE") +
  theme_bw()
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

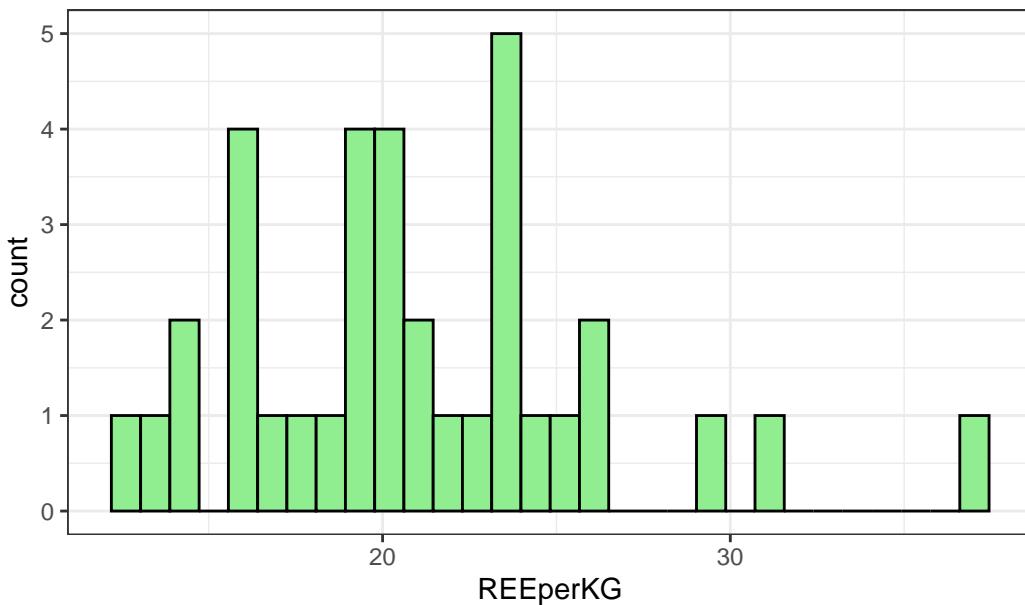
Distribution of REE



```
# Histogram of REE/kg
ggplot(REE, aes(x = REEperKG)) +
  geom_histogram(fill = "lightgreen", color = "black") +
  labs(title = "Distribution of REE per kg") +
  theme_bw()
```

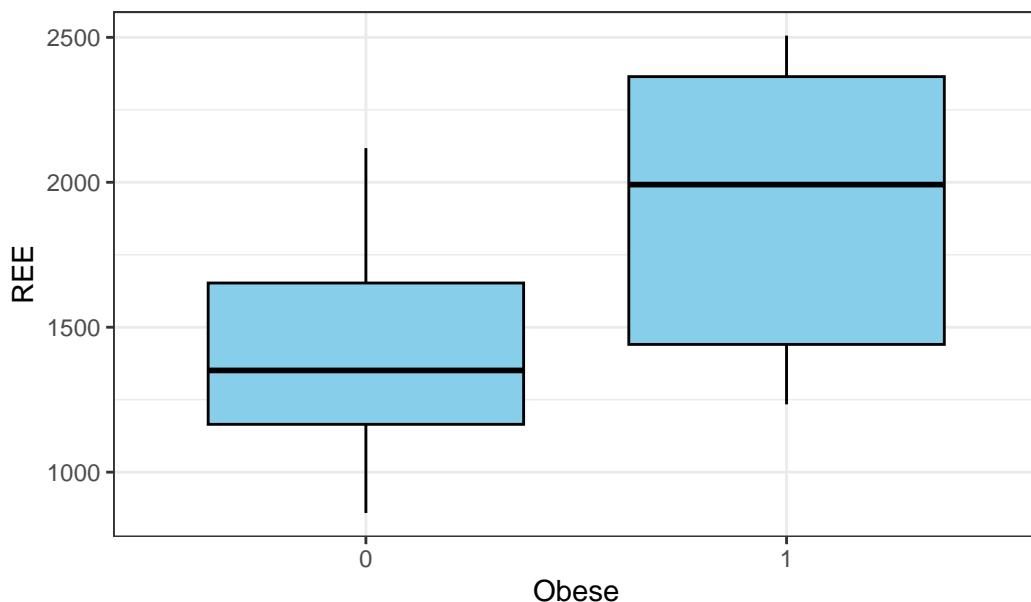
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

Distribution of REE per kg



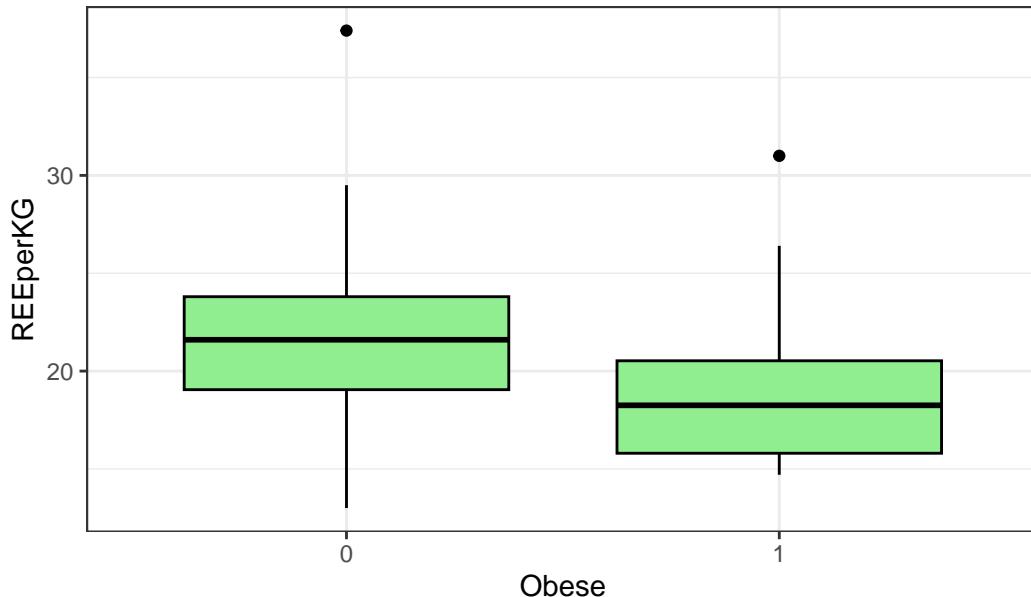
```
ggplot(REE, aes(x = Obese, y = REE)) +  
  geom_boxplot(fill = "skyblue", color = "black") +  
  labs(title = "REE by Obesity") +  
  theme_bw()
```

## REE by Obesity



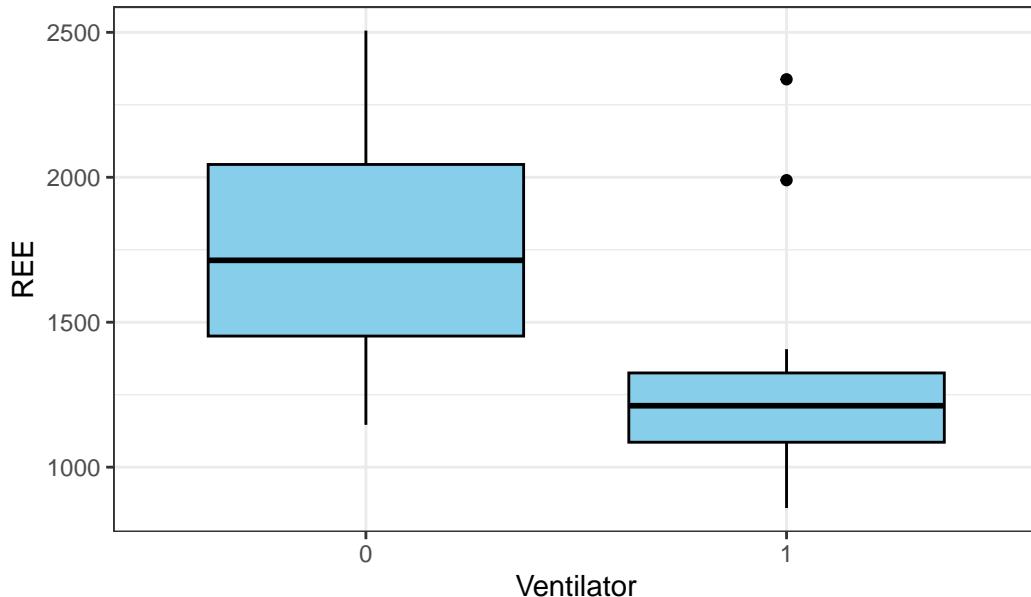
```
ggplot(REE, aes(x = Obese, y = REEperKG)) +  
  geom_boxplot(fill = "lightgreen", color = "black") +  
  labs(title = "REE/kg by Obesity") +  
  theme_bw()
```

REE/kg by Obesity



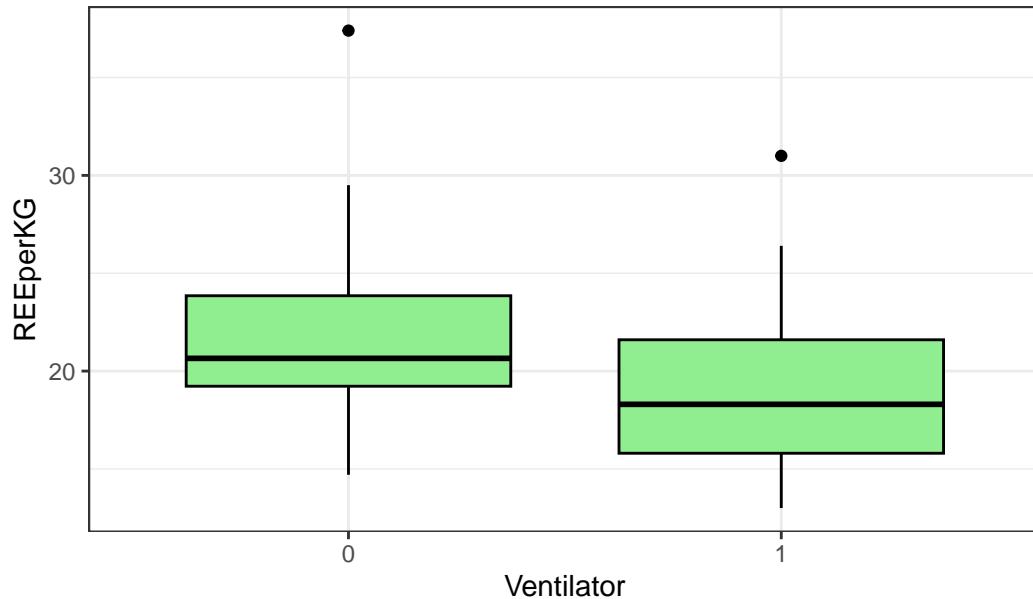
```
ggplot(REE, aes(x = Ventilator, y = REE)) +  
  geom_boxplot(fill = "skyblue", color = "black") +  
  labs(title = "REE by Ventilator Use") +  
  theme_bw()
```

### REE by Ventilator Use



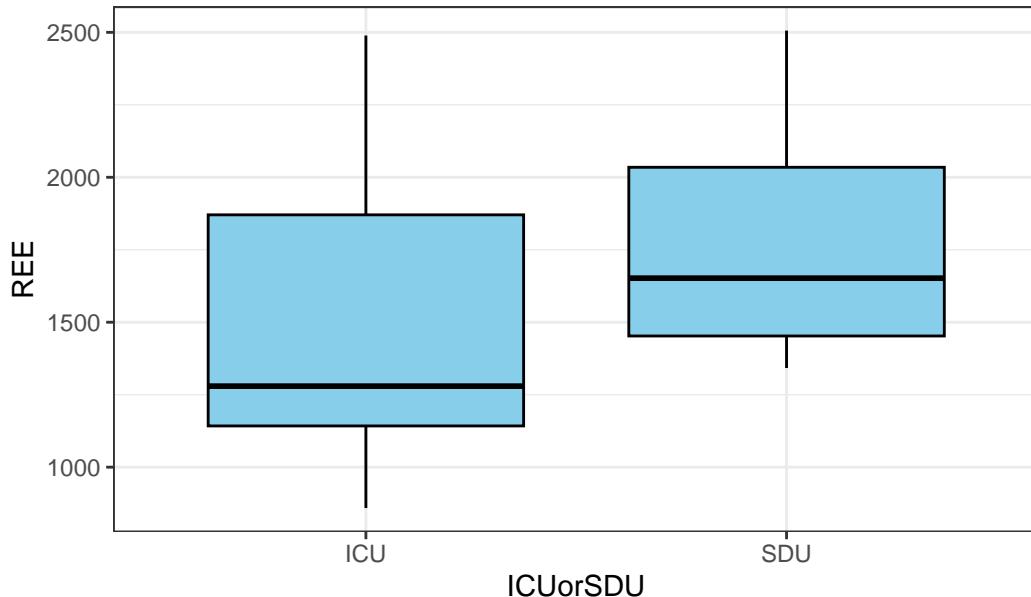
```
ggplot(REE, aes(x = Ventilator, y = REEperKG)) +  
  geom_boxplot(fill = "lightgreen", color = "black") +  
  labs(title = "REE/kg by Ventilator Use") +  
  theme_bw()
```

### REE/kg by Ventilator Use



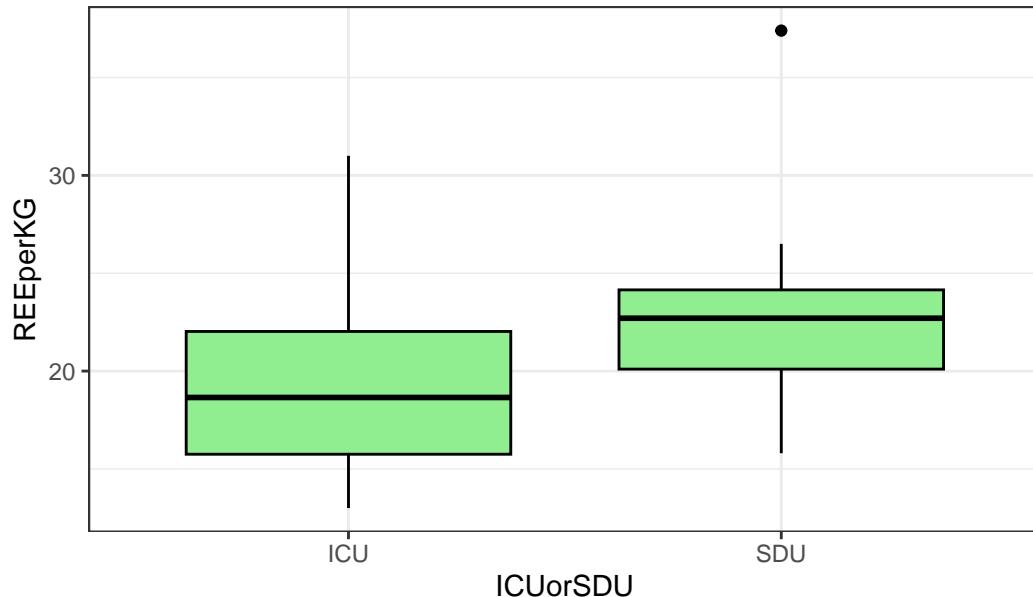
```
ggplot(REE, aes(x = ICUorSDU, y = REE)) +  
  geom_boxplot(fill = "skyblue", color = "black") +  
  labs(title = "REE by Location (ICU vs SDU)") +  
  theme_bw()
```

REE by Location (ICU vs SDU)



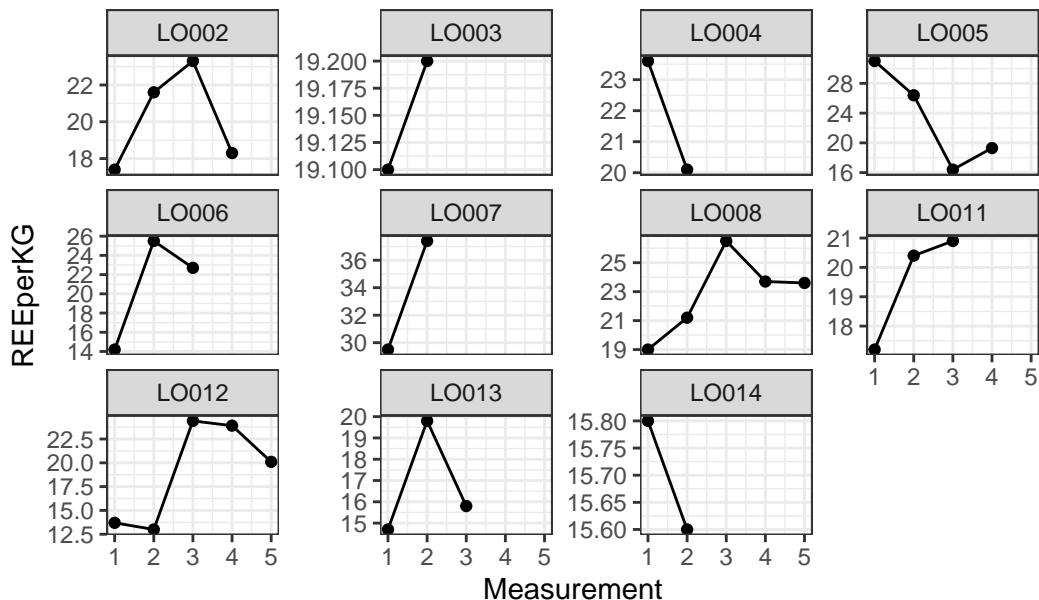
```
ggplot(REE, aes(x = ICUorSDU, y = REEperKG)) +  
  geom_boxplot(fill = "lightgreen", color = "black") +  
  labs(title = "REE/kg by Location (ICU vs SDU)") +  
  theme_bw()
```

REE/kg by Location (ICU vs SDU)



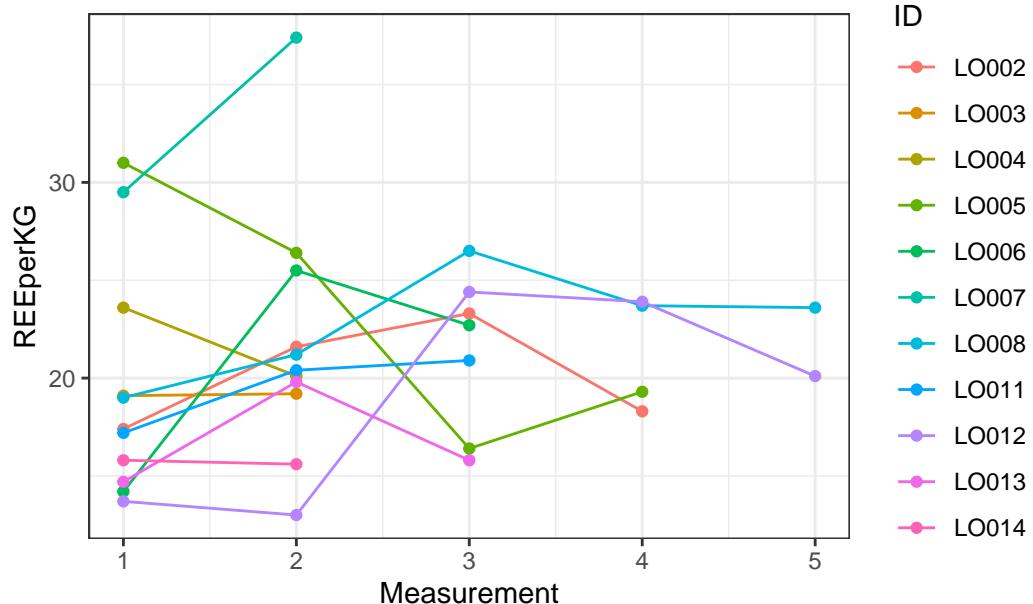
```
# Facet by patient: slope patterns
ggplot(REE, aes(x = Measurement, y = REEperKG)) +
  geom_line() +
  geom_point() +
  facet_wrap(~ ID, scales = "free_y") +
  labs(title = "REE per kg over Time for Each Patient") +
  theme_bw()
```

### REE per kg over Time for Each Patient



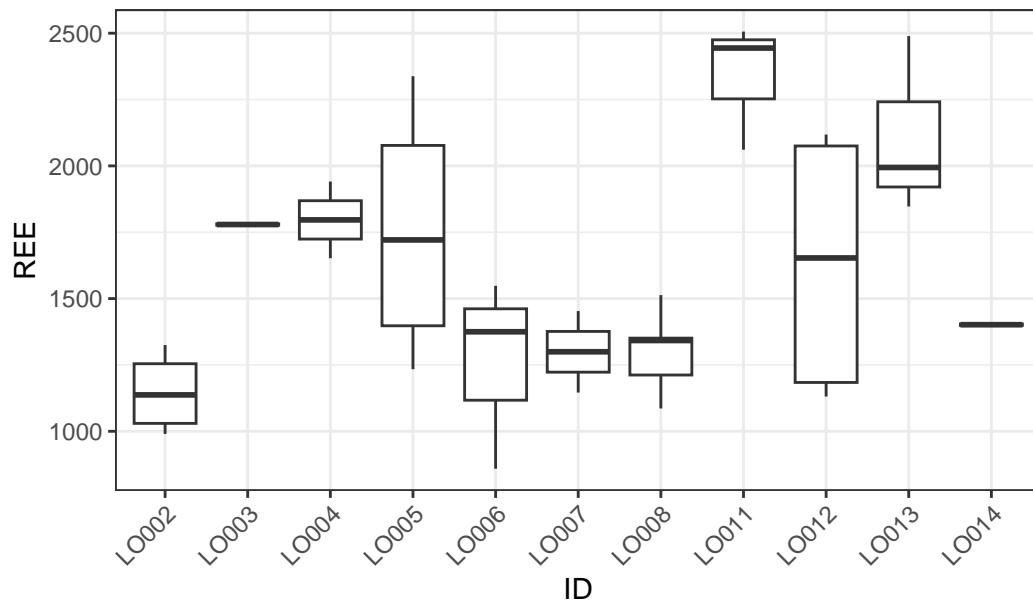
```
# Patient trajectories (spaghetti plot)
ggplot(REE, aes(x = Measurement, y = REEperKG, group = ID, color = ID)) +
  geom_line() +
  geom_point() +
  labs(title = "Longitudinal REE/kg Trajectories by Patient") +
  theme_bw() +
  theme(legend.position = "right")
```

## Longitudinal REE/kg Trajectories by Patient



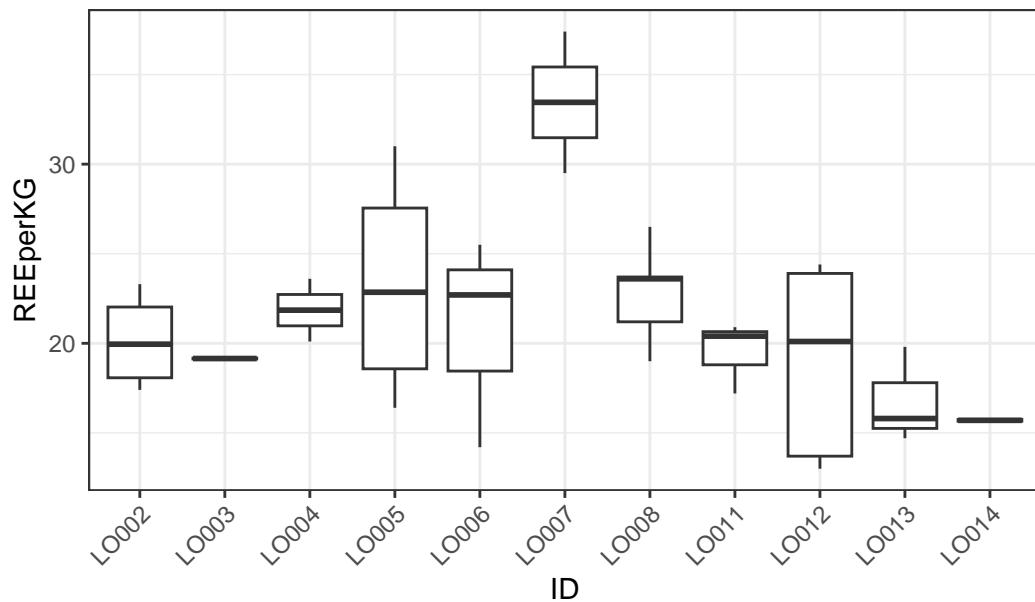
```
# Boxplot by patient
ggplot(REE, aes(x = ID, y = REE)) +
  geom_boxplot(outlier.colour = "red") +
  labs(title = "REE by Patient") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

REE by Patient

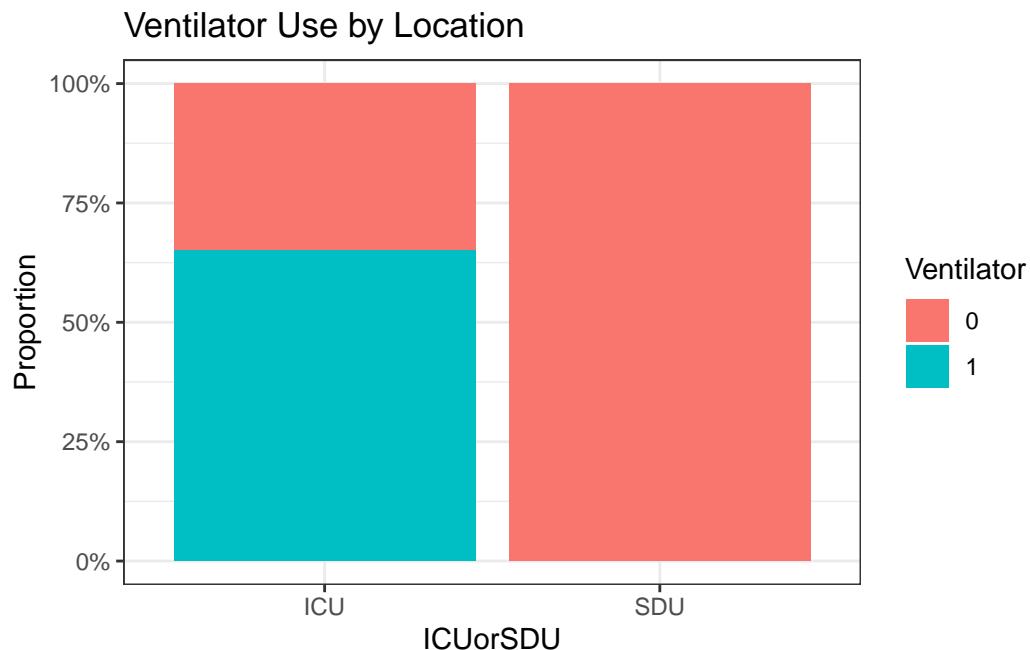


```
# Boxplot by patient
ggplot(REE, aes(x = ID, y = REEperKG)) +
  geom_boxplot(outlier.colour = "red") +
  labs(title = "REE/kg by Patient") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

## REE/kg by Patient



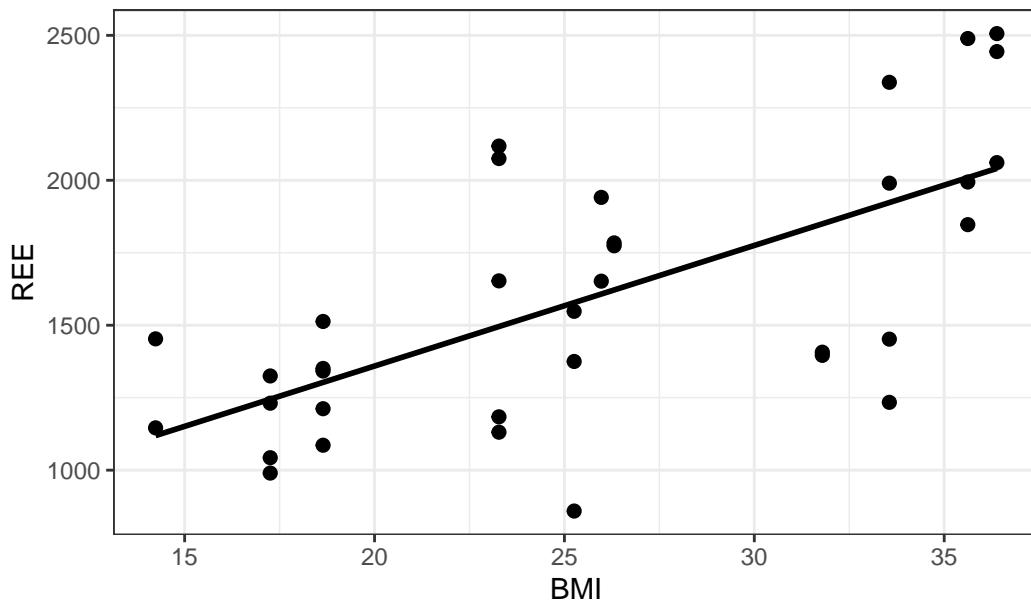
```
# Proportion of Ventilation by Location
ggplot(REE, aes(x = ICUorSDU, fill = Ventilator)) +
  geom_bar(position = "fill") +
  scale_y_continuous(labels = scales::percent) +
  labs(y = "Proportion", title = "Ventilator Use by Location") +
  theme_bw()
```



```
# REE vs BMI
ggplot(REE, aes(x = BMI, y = REE)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE, colour = "black") +
  labs(title = "REE vs BMI") +
  theme_bw()

`geom_smooth()` using formula = 'y ~ x'
```

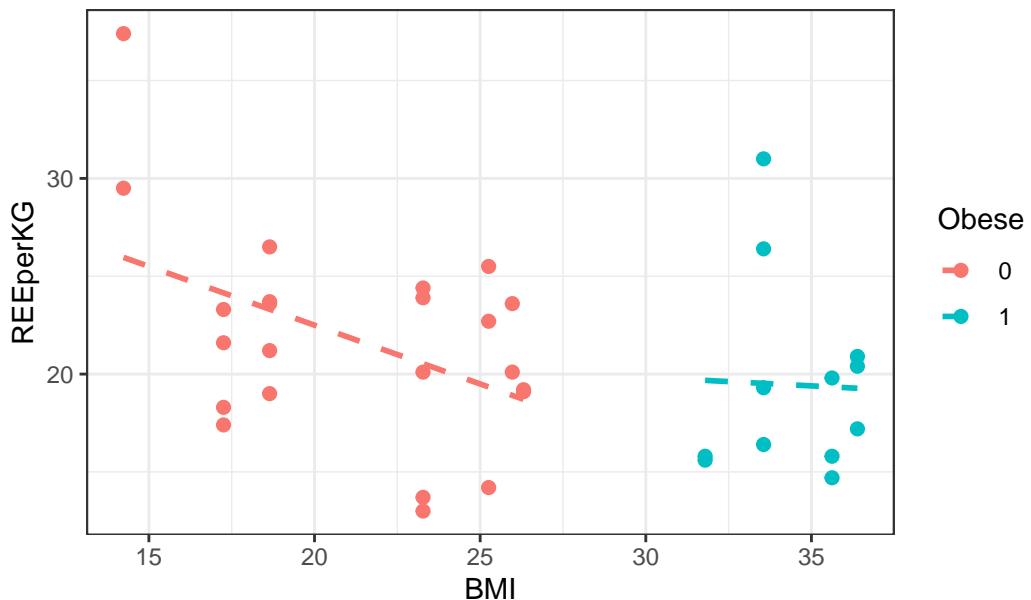
REE vs BMI



```
# REE/kg vs BMI
ggplot(REE, aes(x = BMI, y = REEperKG, colour = Obese)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE, linetype = "dashed") +
  labs(title = "REE/kg vs BMI") +
  theme_bw()
```

`geom\_smooth()` using formula = 'y ~ x'

REE/kg vs BMI



## Final Model

```
library(lme4)
```

```
Loading required package: Matrix
```

```
Attaching package: 'Matrix'
```

```
The following objects are masked from 'package:tidyR':
```

```
expand, pack, unpack
```

```
Warning in check_dep_version(): ABI version mismatch:
```

```
lme4 was built with Matrix ABI version 2
```

```
Current Matrix ABI version is 1
```

```
Please re-install lme4 from source or restore original 'Matrix' package
```

```

model_REEperKG <- lmer(
  REEperKG ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator + (1 | ID),
  data = REE
)

summary(model_REEperKG)

```

Linear mixed model fit by REML ['lmerMod']  
 Formula:  
 $\text{REEperKG} \sim \text{Obese} * \text{Measurement} + \text{ICUorSDU} * \text{Measurement} + \text{Ventilator} + (1 | \text{ID})$   
 Data: REE

REML criterion at convergence: 178.9

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.4128	-0.6014	-0.1273	0.5621	1.5307

Random effects:

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	20.83	4.564
	Residual	11.14	3.338

Number of obs: 35, groups: ID, 11

Fixed effects:

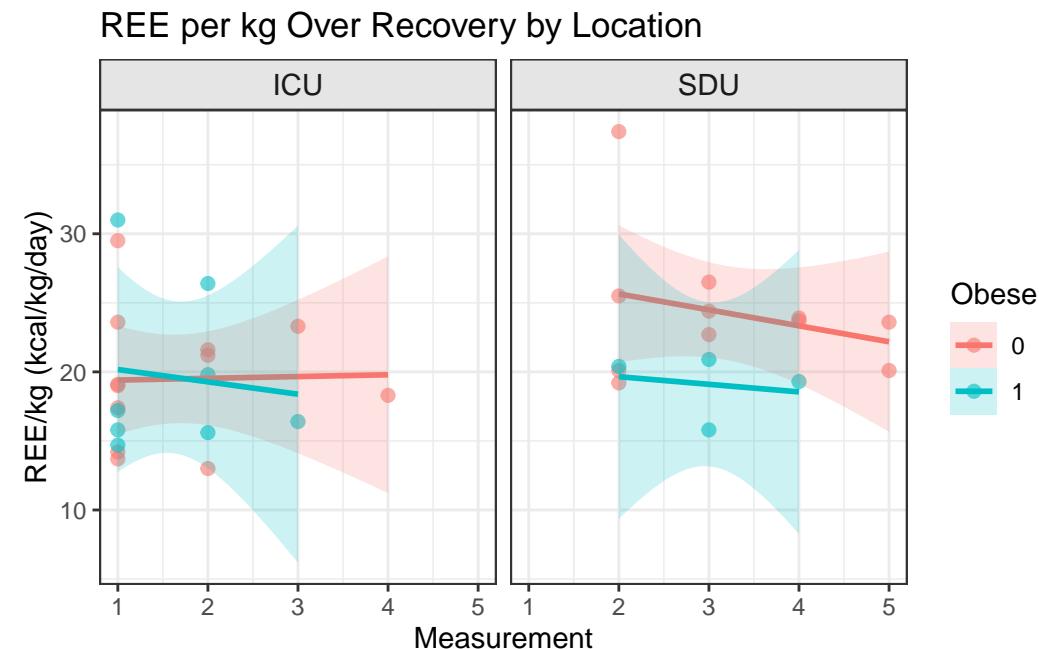
	Estimate	Std. Error	t value
(Intercept)	21.0144	2.9073	7.228
Obese1	4.2462	4.0805	1.041
Measurement	0.3738	1.0875	0.344
ICUorSDUSDU	6.5742	3.6543	1.799
Ventilator1	-3.0581	2.2974	-1.331
Obese1:Measurement	-3.5279	1.2526	-2.816
Measurement:ICUorSDUSDU	-1.2072	1.4533	-0.831

Correlation of Fixed Effects:

	(Intr)	Obese1	Msrmnt	ICUSDU	Vntlt1	Obs1:M
Obese1	-0.508					
Measurement	-0.557	0.166				
ICUorSDUSDU	-0.307	0.128	0.318			
Ventilator1	-0.445	0.097	-0.036	0.010		
Obs1:Msrmnt	0.289	-0.641	-0.285	-0.155	-0.072	

```
M:ICUSDUSDU 0.247 -0.038 -0.642 -0.812 0.314 0.080
```

```
ggplot(REE, aes(x = Measurement, y = REEperKG, color = Obese)) +  
  geom_point(size = 2, alpha = 0.6) +  
  geom_smooth(method = "lm", se = TRUE,  
              aes(fill = Obese), alpha = 0.2) +  
  facet_wrap(~ ICUorSDU) +  
  labs(title = "REE per kg Over Recovery by Location",  
       y = "REE/kg (kcal/kg/day)") +  
  theme_bw() +  
  theme(strip.background = element_rect(fill="grey90"),  
        strip.text = element_text(size=11))  
  
`geom_smooth()` using formula = 'y ~ x'
```



```
ggplot(REE, aes(x = Measurement, y = REEperKG, color = ICUorSDU)) +  
  geom_point(size = 2, alpha = 0.6) +  
  geom_smooth(method = "lm", se = TRUE,  
              aes(fill = ICUorSDU), alpha = 0.2) +  
  facet_wrap(~ Obese) +  
  labs(title = "REE per kg Over Recovery by Obesity Status",
```

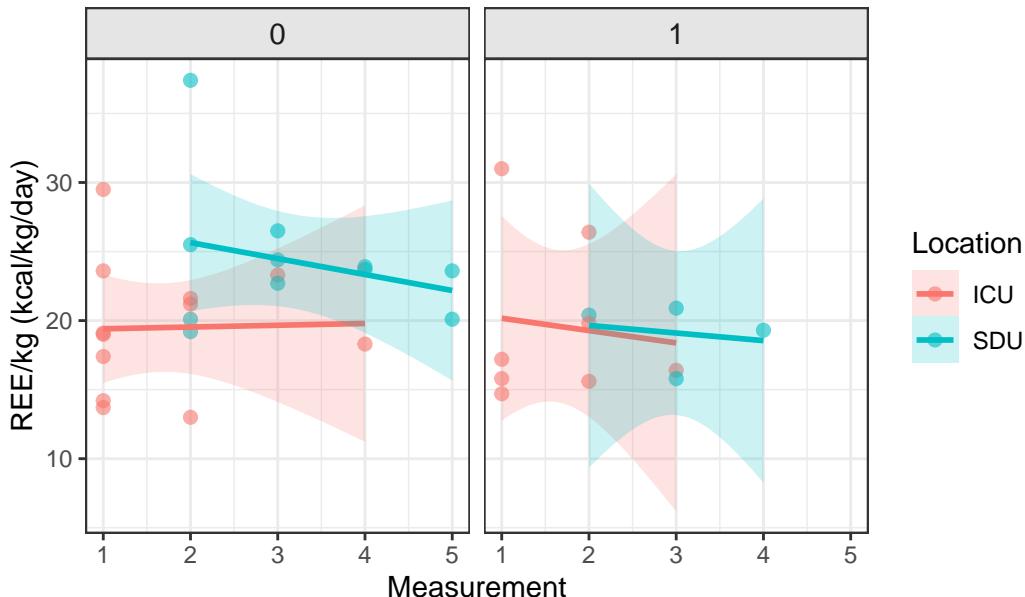
```

y = "REE/kg (kcal/kg/day)",
color = "Location",
fill = "Location") +
theme_bw() +
theme(strip.background = element_rect(fill="grey90"),
      strip.text = element_text(size=11))

`geom_smooth()` using formula = 'y ~ x'

```

REE per kg Over Recovery by Obesity Status



```

model_REE <- lmer(
  REE ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator + (1 | ID),
  data = REE
)

summary(model_REE)

```

```

Linear mixed model fit by REML ['lmerMod']
Formula: REE ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator +
  (1 | ID)
Data: REE

```

REML criterion at convergence: 418.9

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.3155	-0.5735	-0.2568	0.5738	1.6933

Random effects:

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	36139	190.1
Residual		79976	282.8

Number of obs: 35, groups: ID, 11

Fixed effects:

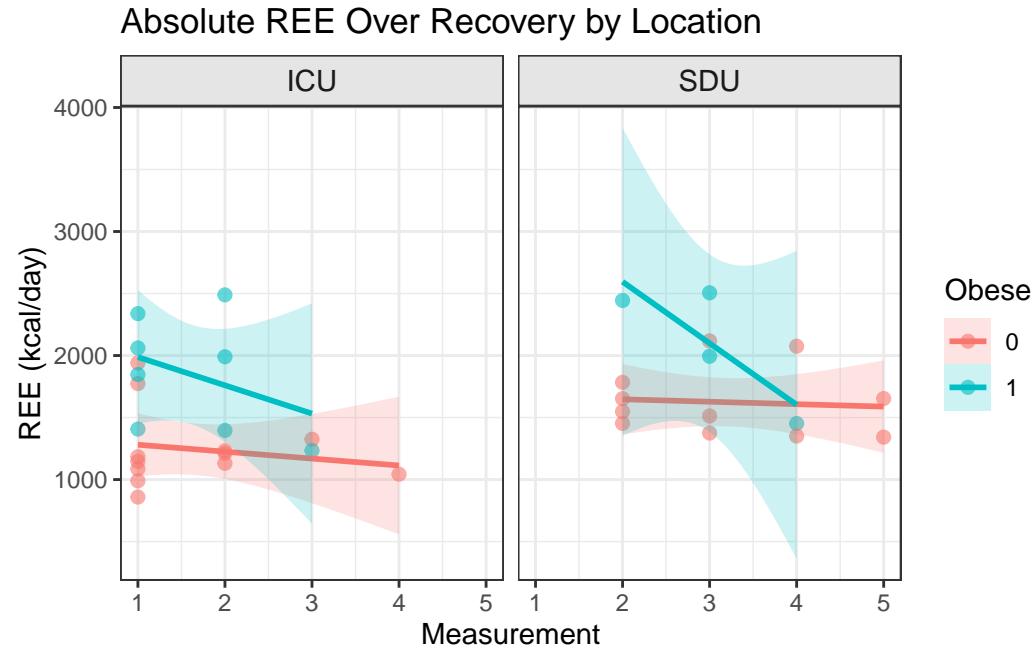
	Estimate	Std. Error	t value
(Intercept)	1448.707	197.797	7.324
Obese1	932.125	272.028	3.427
Measurement	3.954	87.954	0.045
ICUorSDUSDU	432.998	299.913	1.444
Ventilator1	-319.944	167.986	-1.905
Obese1:Measurement	-212.845	103.909	-2.048
Measurement:ICUorSDUSDU	-79.188	117.350	-0.675

Correlation of Fixed Effects:

	(Intr)	Obese1	Msrmnt	ICUSDU	Vnltl1	Obs1:M
Obese1	-0.494					
Measurement	-0.631	0.216				
ICUorSDUSDU	-0.439	0.176	0.365			
Ventilator1	-0.437	0.127	-0.157	0.102		
Obs1:Msrmnt	0.352	-0.809	-0.291	-0.180	-0.065	
M:ICUSDUSDU	0.365	-0.064	-0.698	-0.809	0.270	0.104

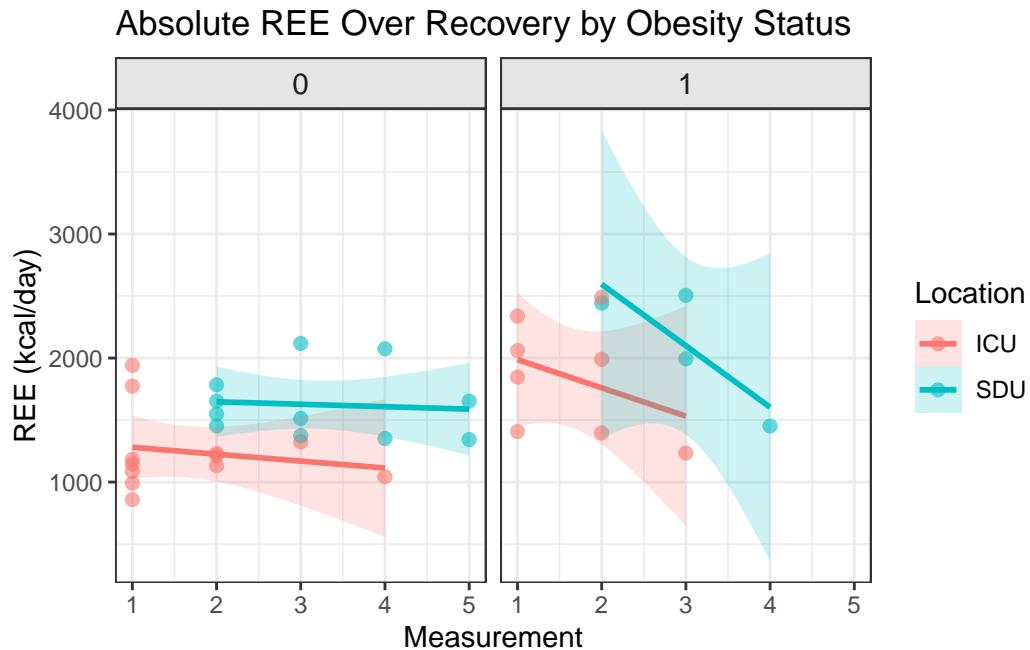
```
ggplot(REE, aes(x = Measurement, y = REE, color = Obese)) +
  geom_point(size = 2, alpha = 0.6) +
  geom_smooth(method = "lm", se = TRUE,
              aes(fill = Obese), alpha = 0.2) +
  facet_wrap(~ ICUorSDU) +
  labs(title = "Absolute REE Over Recovery by Location",
       y = "REE (kcal/day)") +
  theme_bw() +
  theme(strip.background = element_rect(fill="grey90"),
        strip.text = element_text(size=11))
```

```
`geom_smooth()` using formula = 'y ~ x'
```



```
ggplot(REE, aes(x = Measurement, y = REE, color = ICUorSDU)) +  
  geom_point(size = 2, alpha = 0.6) +  
  geom_smooth(method = "lm", se = TRUE,  
              aes(fill = ICUorSDU), alpha = 0.2) +  
  facet_wrap(~ Obese) +  
  labs(title = "Absolute REE Over Recovery by Obesity Status",  
       y = "REE (kcal/day)",  
       color = "Location",  
       fill = "Location") +  
  theme_bw() +  
  theme(strip.background = element_rect(fill="grey90"),  
        strip.text = element_text(size=11))
```

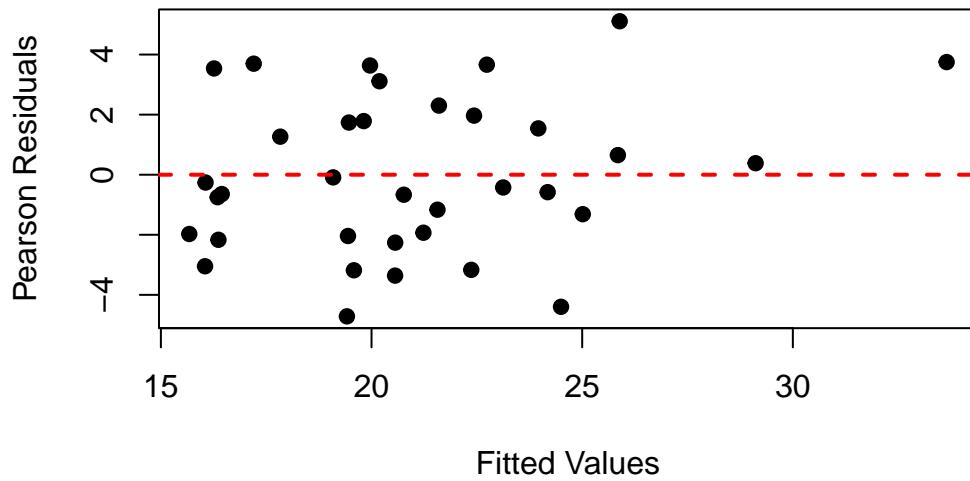
```
`geom_smooth()` using formula = 'y ~ x'
```



## Model Diagnostics

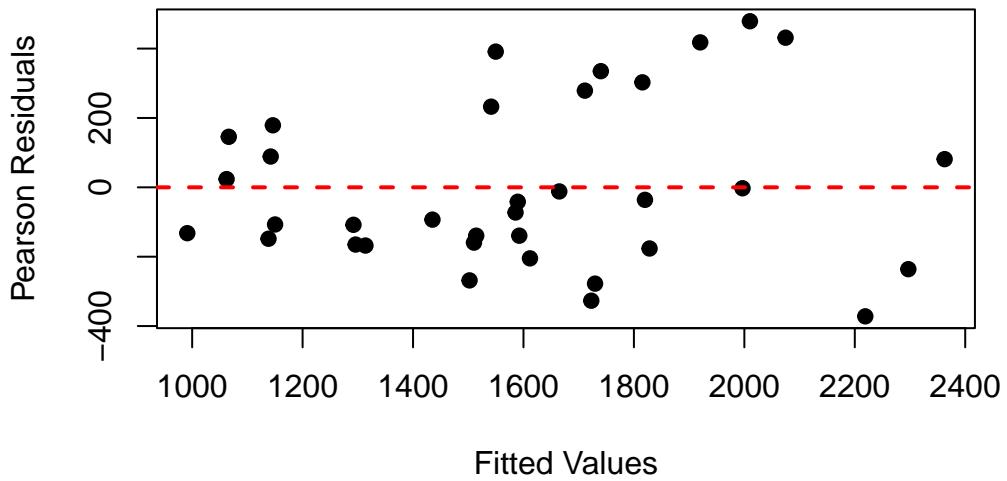
```
# Residual plots
plot(
  fitted(model_REEperKG),
  resid(model_REEperKG, type = "pearson"),
  pch = 19, col = "black",
  xlab = "Fitted Values",
  ylab = "Pearson Residuals",
  main = "Residuals vs Fitted: REEperKG"
)
abline(h = 0, col = "red", lty = 2, lwd = 2)
```

### Residuals vs Fitted: REEperKG



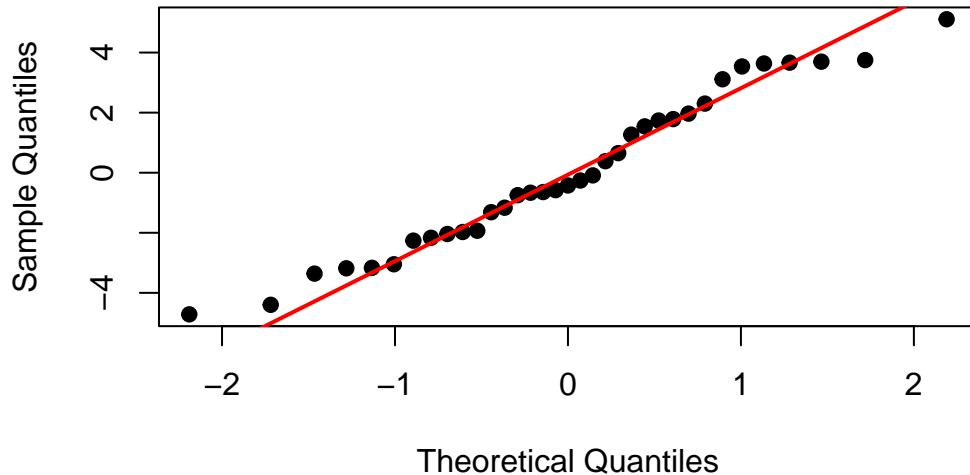
```
plot(
  fitted(model_REE),
  resid(model_REE, type = "pearson"),
  pch = 19, col = "black",
  xlab = "Fitted Values",
  ylab = "Pearson Residuals",
  main = "Residuals vs Fitted: REE"
)
abline(h = 0, col = "red", lty = 2, lwd = 2)
```

### Residuals vs Fitted: REE



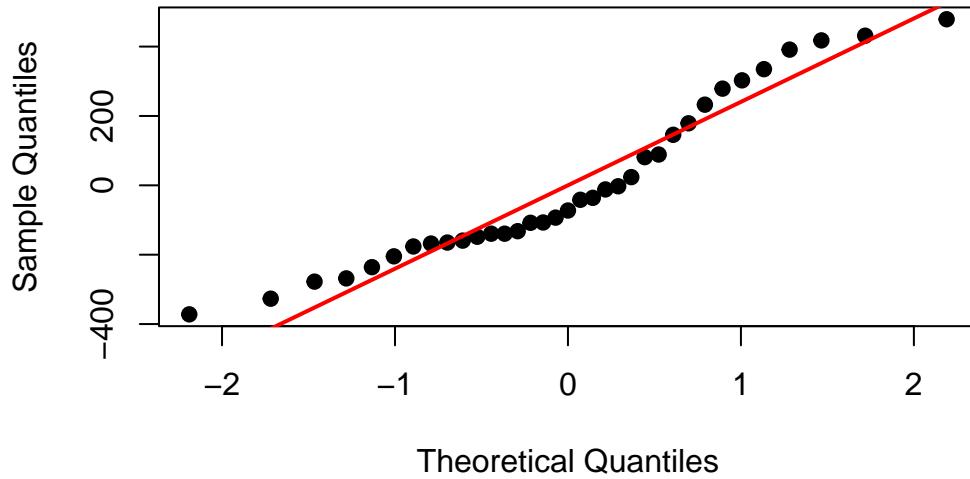
```
# QQ plots: residuals
qqnorm(resid(model_REEperKG),
      main = "Q-Q Plot of Residuals: REEperKG",
      pch = 19, col = "black")
qqline(resid(model_REEperKG), col = "red", lwd = 2)
```

### Q-Q Plot of Residuals: REEperKG



```
qqnorm(resid(model_REE),  
       main = "Q-Q Plot of Residuals: REE",  
       pch = 19, col = "black")  
qqline(resid(model_REE), col = "red", lwd = 2)
```

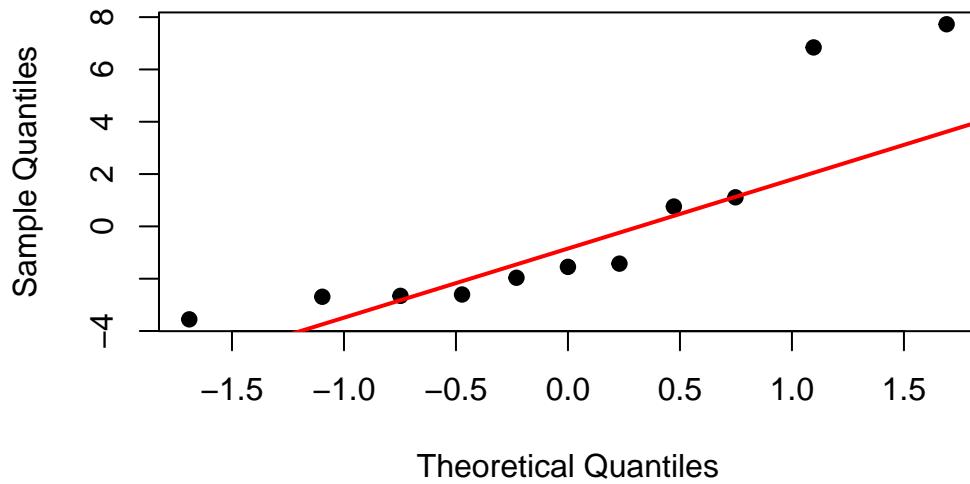
## Q-Q Plot of Residuals: REE



```
## QQ plots: random effects
re_REEperKG <- ranef(model_REEperKG)$ID[[1]]
re_REE      <- ranef(model_REE)$ID[[1]]

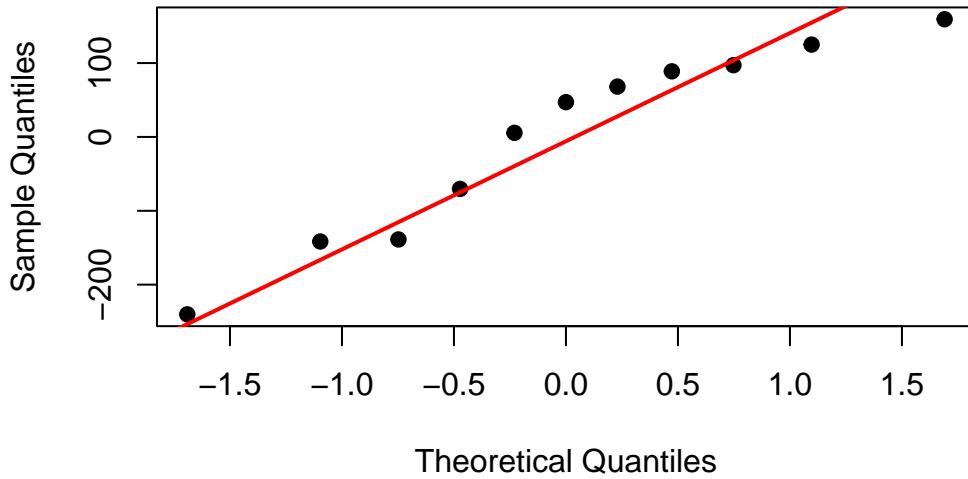
qqnorm(re_REEperKG,
       main = "Q-Q Plot of Random Intercepts: REEperKG",
       pch = 19, col = "black")
qqline(re_REEperKG, col = "red", lwd = 2)
```

### Q-Q Plot of Random Intercepts: REEperKG



```
qqnorm(re_REE,
       main = "Q-Q Plot of Random Intercepts: REE",
       pch = 19, col = "black")
qqline(re_REE, col = "red", lwd = 2)
```

## Q–Q Plot of Random Intercepts: REE



```
# REEperKG: Ratio test
model_REEperKG_full_ML <- update(model_REEperKG, REML = FALSE)

## Drop ICU x Measurement
model_REEperKG_no_locInt <- lmer(
  REEperKG ~ Obese * Measurement + ICUorSDU + Measurement + Ventilator + (1 | ID),
  data = REE,
  REML = FALSE
)

## Drop Obese x Measurement
model_REEperKG_no_obInt <- lmer(
  REEperKG ~ Obese + Measurement + ICUorSDU * Measurement + Ventilator + (1 | ID),
  data = REE,
  REML = FALSE
)

## Drop both interactions (main-effects only)
model_REEperKG_main <- lmer(
  REEperKG ~ Obese + Measurement + ICUorSDU + Ventilator + (1 | ID),
  data = REE,
  REML = FALSE
)
```

```

## AIC & BIC
AIC(model_REEperKG_full_ML, model_REEperKG_no_locInt, model_REEperKG_no_obInt, model_REEperKG_main)

      df      AIC
model_REEperKG_full_ML    9 214.7266
model_REEperKG_no_locInt  8 213.5626
model_REEperKG_no_obInt   8 220.1186
model_REEperKG_main        7 218.4164

BIC(model_REEperKG_full_ML, model_REEperKG_no_locInt, model_REEperKG_no_obInt, model_REEperKG_main)

      df      BIC
model_REEperKG_full_ML    9 228.7248
model_REEperKG_no_locInt  8 226.0054
model_REEperKG_no_obInt   8 232.5613
model_REEperKG_main        7 229.3039

## Likelihood ratio tests
anova(model_REEperKG_full_ML, model_REEperKG_no_locInt)

Data: REE
Models:
model_REEperKG_no_locInt: REEperKG ~ Obese * Measurement + ICUorSDU + Measurement + Ventilator
model_REEperKG_full_ML: REEperKG ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator
      npar      AIC      BIC  logLik deviance Chisq Df
model_REEperKG_no_locInt     8 213.56 226.00 -98.781   197.56
model_REEperKG_full_ML       9 214.73 228.72 -98.363   196.73 0.836  1
      Pr(>Chisq)
model_REEperKG_no_locInt
model_REEperKG_full_ML       0.3606

anova(model_REEperKG_full_ML, model_REEperKG_no_obInt)

Data: REE
Models:
model_REEperKG_no_obInt: REEperKG ~ Obese + Measurement + ICUorSDU * Measurement + Ventilator
model_REEperKG_full_ML: REEperKG ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator
      npar      AIC      BIC  logLik deviance Chisq Df
model_REEperKG_no_obInt     8 220.12 232.56 -102.059   204.12

```

```

model_REEperKG_full_ML      9 214.73 228.72 -98.363   196.73 7.3919  1
                           Pr(>Chisq)
model_REEperKG_no_obInt
model_REEperKG_full_ML     0.006552 **

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(model_REEperKG_full_ML, model_REEperKG_main)

Data: REE
Models:
model_REEperKG_main: REEperKG ~ Obese + Measurement + ICUorSDU + Ventilator + (1 | ID)
model_REEperKG_full_ML: REEperKG ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator
npar      AIC      BIC      logLik deviance Chisq Df
model_REEperKG_main      7 218.42 229.30 -102.208   204.42
model_REEperKG_full_ML    9 214.73 228.72 -98.363   196.73 7.6898  2
                           Pr(>Chisq)
model_REEperKG_main
model_REEperKG_full_ML    0.02139 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# REE: Ratio test
model_REE_full_ML <- update(model_REE, REML = FALSE)

# Drop ICU x Measurement
model_REE_no_locInt <- lmer(
  REE ~ Obese * Measurement + ICUorSDU + Measurement + Ventilator + (1 | ID),
  data = REE,
  REML = FALSE
)

# Drop Obese x Measurement
model_REE_no_obInt <- lmer(
  REE ~ Obese + Measurement + ICUorSDU * Measurement + Ventilator + (1 | ID),
  data = REE,
  REML = FALSE
)

# Drop both interactions (main-effects only)
model_REE_main <- lmer(
  REE ~ Obese + Measurement + ICUorSDU + Ventilator + (1 | ID),

```

```

    data = REE,
    REML = FALSE
)

# AIC & BIC
AIC(model_REE_full_ML, model_REE_no_locInt, model_REE_no_obInt, model_REE_main)

```

	df	AIC
model_REE_full_ML	9	514.1467
model_REE_no_locInt	8	512.6707
model_REE_no_obInt	8	516.2957
model_REE_main	7	514.4486

```
BIC(model_REE_full_ML, model_REE_no_locInt, model_REE_no_obInt, model_REE_main)
```

	df	BIC
model_REE_full_ML	9	528.1449
model_REE_no_locInt	8	525.1134
model_REE_no_obInt	8	528.7385
model_REE_main	7	525.3360

```
# Likelihood ratio tests
anova(model_REE_full_ML, model_REE_no_locInt)
```

Data: REE  
 Models:

model_REE_no_locInt: REE ~ Obese * Measurement + ICUorSDU + Measurement + Ventilator + (1   ID)	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
model_REE_full_ML: REE ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator + (1   ID)	8	512.67	525.11	-248.34	496.67			
model_REE_no_locInt	8	512.67	525.11	-248.34	496.67			
model_REE_full_ML	9	514.15	528.14	-248.07	496.15	0.5239	1	0.4692

```
anova(model_REE_full_ML, model_REE_no_obInt)
```

Data: REE  
 Models:

model_REE_no_obInt: REE ~ Obese + Measurement + ICUorSDU * Measurement + Ventilator + (1   ID)	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
model_REE_full_ML: REE ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator + (1   ID)	8	512.67	525.11	-248.34	496.67			
model_REE_no_obInt	8	512.67	525.11	-248.34	496.67			
model_REE_full_ML	9	514.15	528.14	-248.07	496.15	0.5239	1	0.4692

```
model_REE_no_obInt     8 516.30 528.74 -250.15   500.30
model_REE_full_ML      9 514.15 528.14 -248.07   496.15 4.1489  1     0.04166 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(model_REE_full_ML, model_REE_main)
```

```
Data: REE
Models:
model_REE_main: REE ~ Obese + Measurement + ICUorSDU + Ventilator + (1 | ID)
model_REE_full_ML: REE ~ Obese * Measurement + ICUorSDU * Measurement + Ventilator + (1 | ID)
npar      AIC      BIC  logLik deviance Chisq Df Pr(>Chisq)
model_REE_main    7 514.45 525.34 -250.22   500.45
model_REE_full_ML 9 514.15 528.14 -248.07   496.15 4.3019  2     0.1164
```