



# EDA REPORT

## MONK



### Report Overview

This report was created for the EDA of **monk** data. It helps explore data to **understand the data and find scenarios for performing the analysis.**

# Contents

<b>Overview</b>	<b>2</b>
Data Structures	2
Job Informations	2
<b>Univariate Analysis</b>	<b>3</b>
Descriptive Statistics	3
Numerical Variables	3
Categorical Variables	6
Normality Test	7
<b>Bivariate Analysis</b>	<b>15</b>
Compare Numerical Variables	15
Compare Categorical Variables	37
<b>Multivariate Analysis</b>	<b>38</b>
Correlation Analysis	38
Correlation Coefficient Matrix	38
Correlation Plot	39

# Overview

## Data Structures

division	metrics	value	division	metrics	value
size	observations	111	data type	numerics	6
size	variables	7	data type	integers	1
size	values	777	data type	factors/ordered	0
size	memory size (KB)	0	data type	characters	0
duplicated	duplicate observation	0	data type	Dates	0
missing	complete observation	111	data type	POSIXcts	0
missing	missing observation	0	data type	others	0
missing	missing variables	0			
missing	missing values	0			

Table 1: Data structures and types

## Job Informations

division	metrics	value
dataset	dataset	monk
dataset	dataset type	data.frame
dataset	target	not defied
job	samples	111 / 111 (100%)
job	created	2022-07-08 12:35:11
job	created by	dlookr

Table 2: Job informations

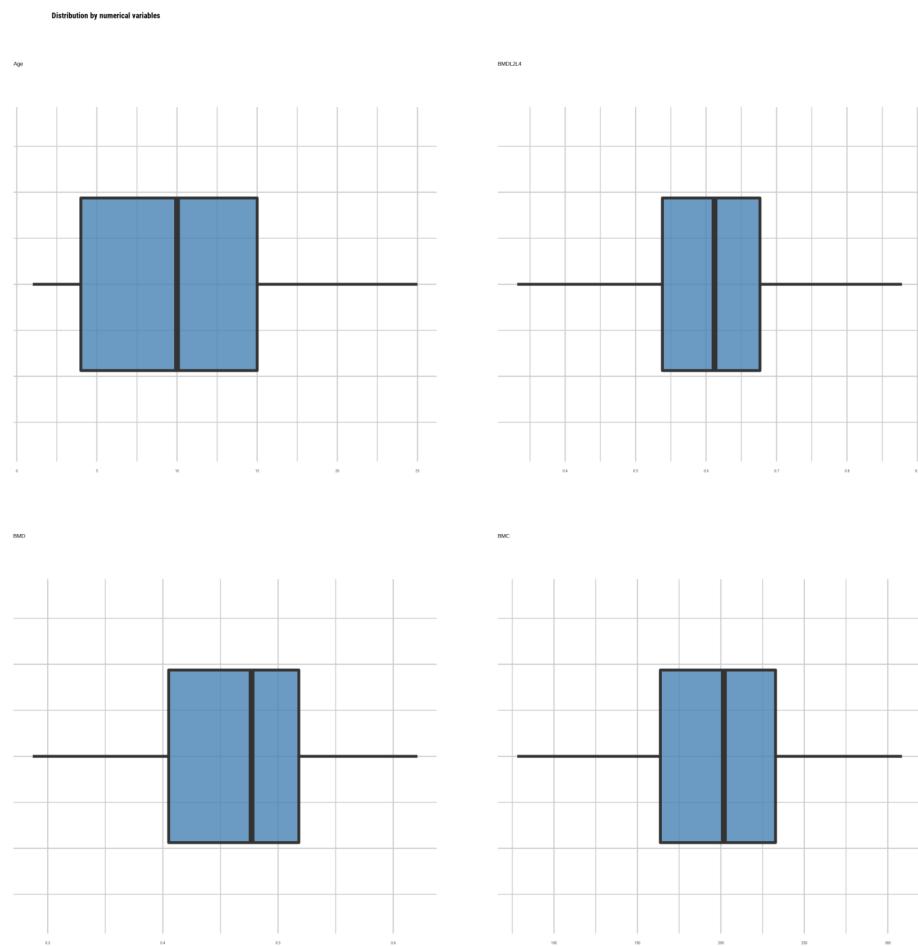
# Univariate Analysis

## Descriptive Statistics

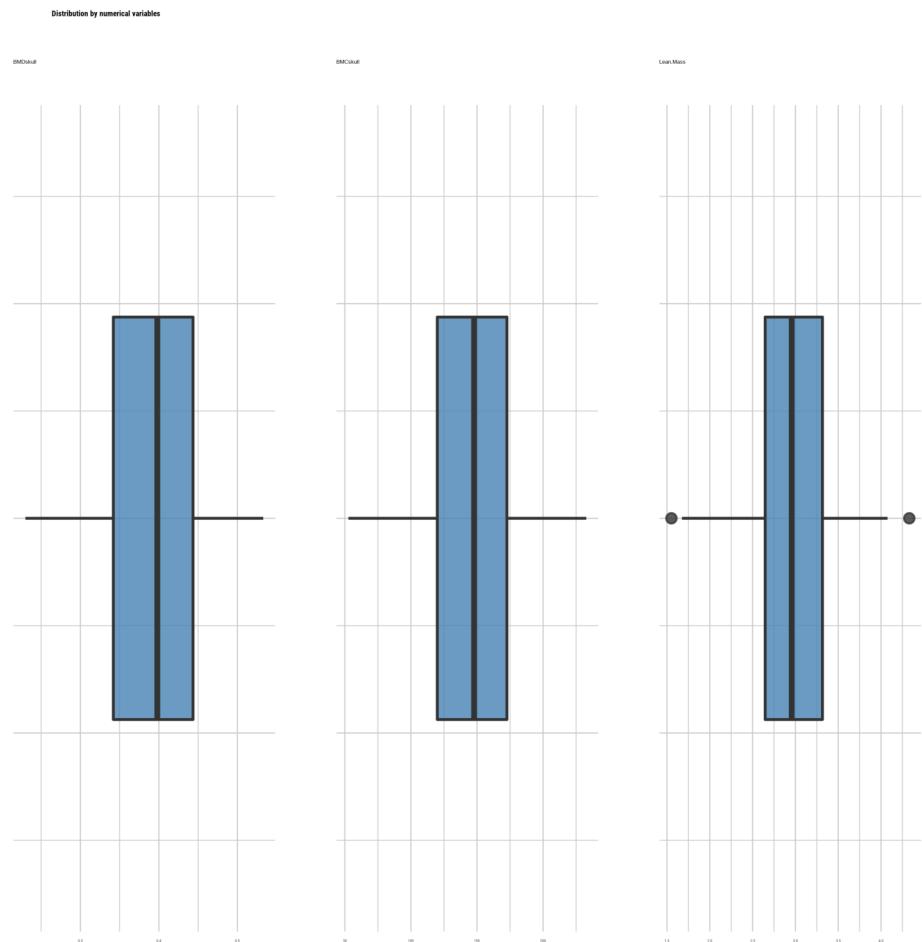
### Numerical Variables

variables	missing	mean	sd	min	Q1	median	Q3	max
Age	0	11.00	7.55	1.00	4.00	10.00	15.00	25.00
BMDL2L4	0	0.60	0.12	0.33	0.54	0.61	0.68	0.88
BMD	0	0.46	0.08	0.29	0.41	0.48	0.52	0.62
BMC	0	192.16	54.73	78.00	163.75	201.80	232.75	308.50
BMDskull	0	0.39	0.07	0.23	0.34	0.40	0.44	0.53
BMCskull	0	141.73	42.01	52.70	119.95	147.80	172.65	232.70
Lean.Mass	0	2.95	0.58	1.55	2.65	2.96	3.32	4.33

Table 3: Descriptive statistics of numerical variables



variables	data types	distinct	skewness	kurtosis	zero	negative	outlier
Age	integer	23	0.37	-1.15	0	0	0
BMDL2L4	numeric	96	-0.15	-0.22	0	0	0
BMD	numeric	90	-0.56	-0.48	0	0	0
BMC	numeric	108	-0.55	-0.38	0	0	0



variables	data types	distinct	skewness	kurtosis	zero	negative	outlier
BMDskull	numeric	92	-0.57	-0.44	0	0	0
BMCskull	numeric	106	-0.52	-0.36	0	0	0
Lean.Mass	numeric	104	-0.37	-0.13	0	0	2

## Categorical Variables

There are no categorical variables.

The number of categorical(factor/ordered) variables is 0.

## Normality Test

described_variables	min	Q1	median	Q3	max	skewness	kurtosis	balance
Age	1.0	4.0	10.0	15.0	25.0	0.4	-1.1	Balanced
BMDL2L4	0.3	0.5	0.6	0.7	0.9	-0.2	-0.2	Balanced
BMD	0.3	0.4	0.5	0.5	0.6	-0.6	-0.5	Balanced
BMC	78.0	163.8	201.8	232.8	308.5	-0.5	-0.4	Balanced
BMDskull	0.2	0.3	0.4	0.4	0.5	-0.6	-0.4	Balanced
BMCskull	52.7	120.0	147.8	172.7	232.7	-0.5	-0.4	Balanced
Lean.Mass	1.6	2.6	3.0	3.3	4.3	-0.4	-0.1	Balanced

Table 4: Descriptive statistics of numerical variables

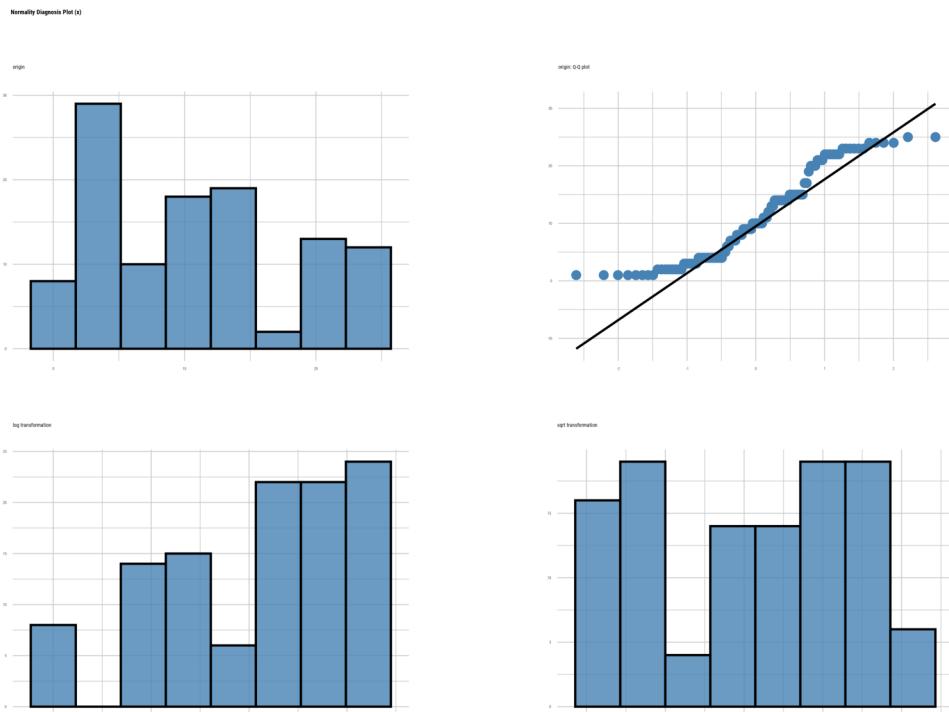
# Age

statistic	p_value	remark
0.91683	3.5024e-06	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	0.3615	1.8478
log transformation	-0.7076	2.4817
sqrt transformation	-0.0983	1.8060

Table 5: skewness and kurtosis



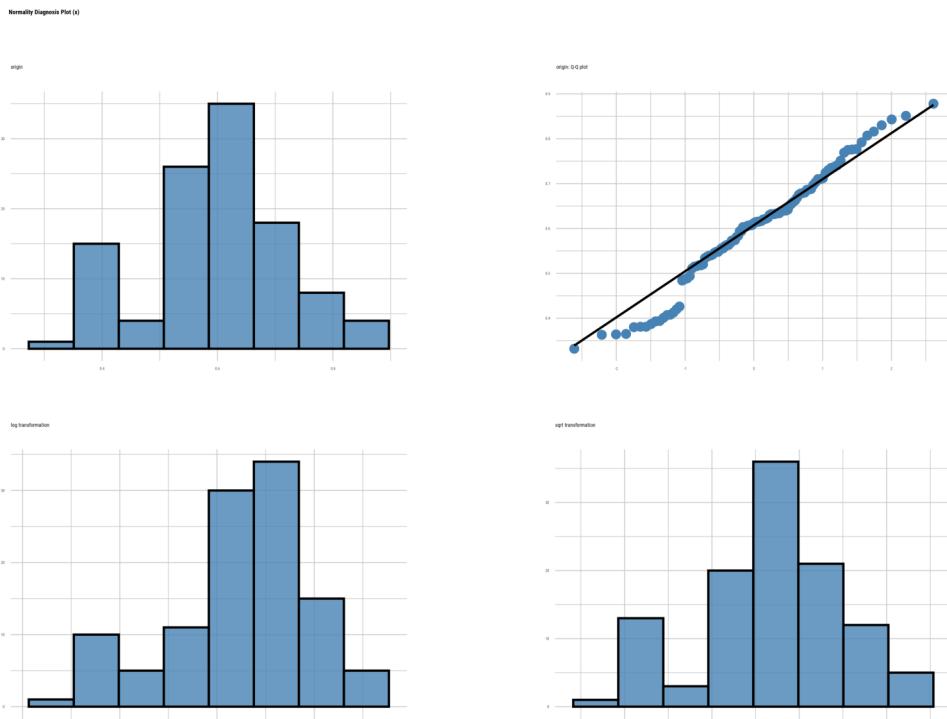
## BMDL2L4

statistic	p_value	remark
0.97746	0.056941	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.1497	2.7407
log transformation	-0.6708	3.0243
sqrt transformation	-0.4134	2.8201

Table 5: skewness and kurtosis



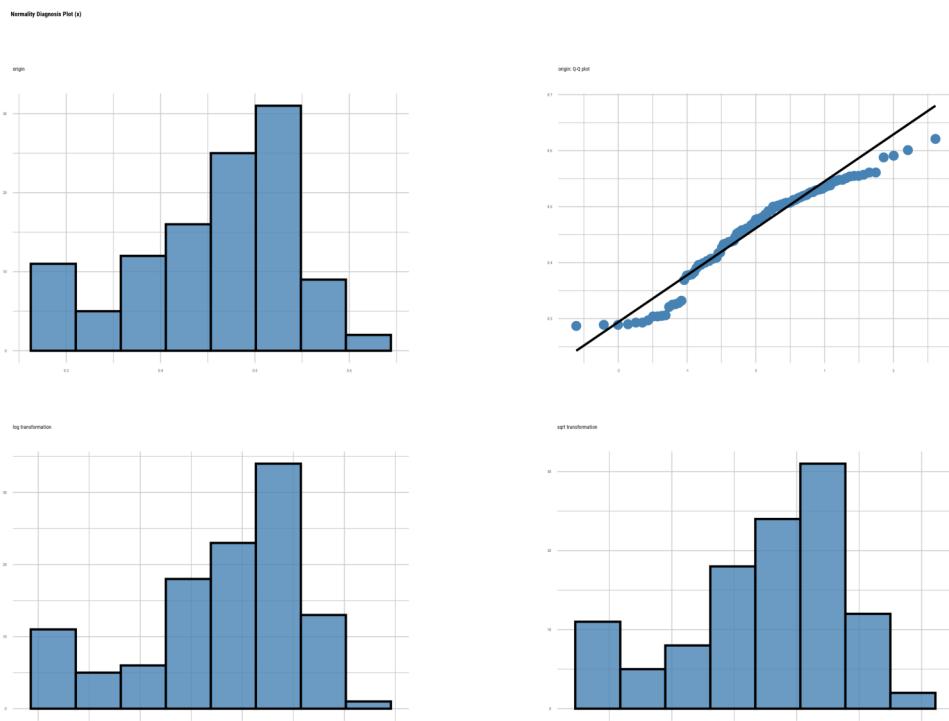
# BMD

statistic	p_value	remark
0.94577	0.00020057	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.5543	2.4881
log transformation	-0.8834	2.8795
sqrt transformation	-0.7206	2.6572

Table 5: skewness and kurtosis



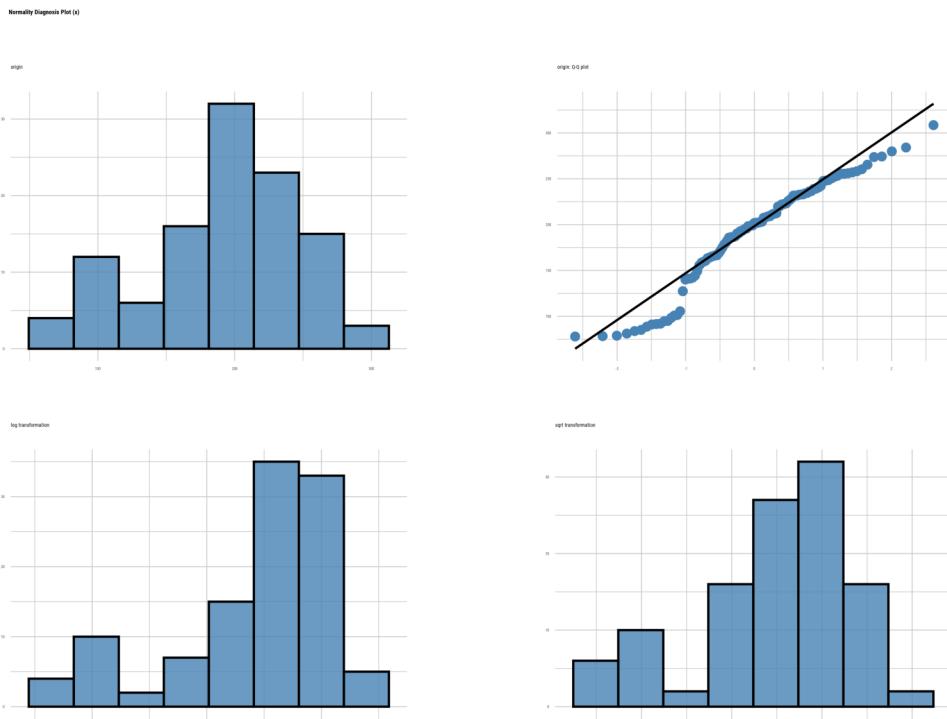
# BMC

statistic	p_value	remark
0.94905	0.00033675	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.5422	2.5820
log transformation	-1.1274	3.3565
sqrt transformation	-0.8414	2.8957

Table 5: skewness and kurtosis



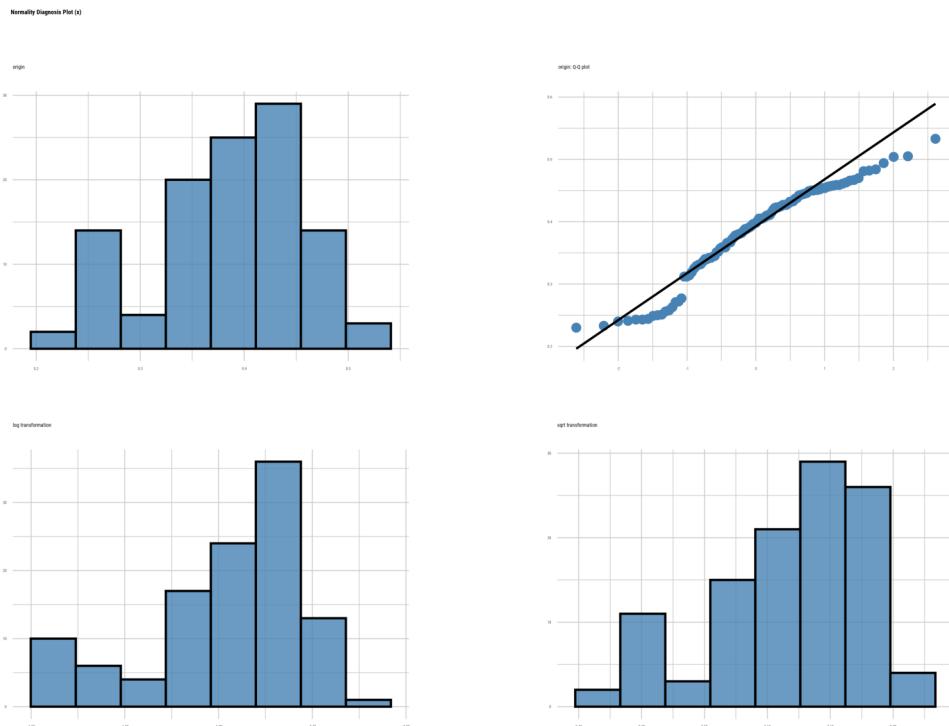
## BMDskull

statistic	p_value	remark
0.94678	0.00023495	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.5657	2.5240
log transformation	-0.9178	2.9606
sqrt transformation	-0.7436	2.7129

Table 5: skewness and kurtosis



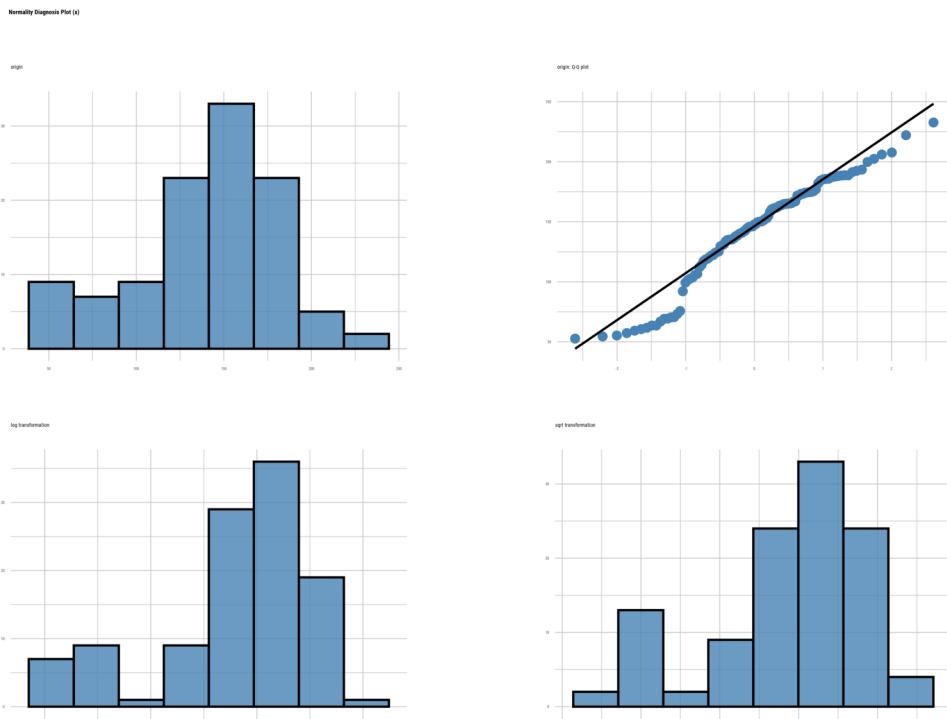
## BMCskull

statistic	p_value	remark
0.95327	0.00066979	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.5091	2.6062
log transformation	-1.1427	3.4206
sqrt transformation	-0.8336	2.9220

Table 5: skewness and kurtosis



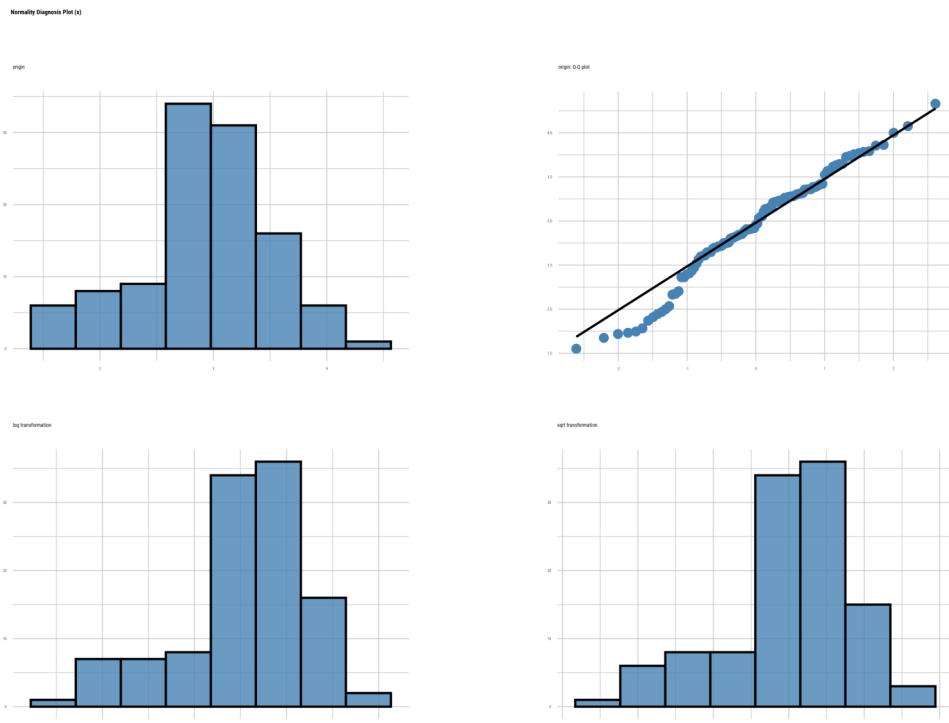
## Lean.Mass

statistic	p_value	remark
0.97792	0.062419	No sample

Table 5: Shapiro-Wilk normality test

type	skewness	kurtosis
original	-0.3606	2.8180
log transformation	-0.8895	3.4532
sqrt transformation	-0.6239	3.0568

Table 5: skewness and kurtosis



# Bivariate Analysis

## Compare Numerical Variables

first variable	second variable	correlation coefficient
Age	BMDL2L4	0.63868
Age	BMD	0.69010
Age	BMC	0.61467
Age	BMDskull	0.70983
Age	BMCskull	0.61501
Age	Lean.Mass	0.33518
BMDL2L4	BMD	0.91142
BMDL2L4	BMC	0.86853
BMDL2L4	BMDskull	0.92100
BMDL2L4	BMCskull	0.86995
BMDL2L4	Lean.Mass	0.62007
BMD	BMC	0.95877
BMD	BMDskull	0.99258
BMD	BMCskull	0.94974
BMD	Lean.Mass	0.69381
BMC	BMDskull	0.95637
BMC	BMCskull	0.99577
BMC	Lean.Mass	0.77948
BMDskull	BMCskull	0.95586
BMDskull	Lean.Mass	0.70400
BMCskull	Lean.Mass	0.78195

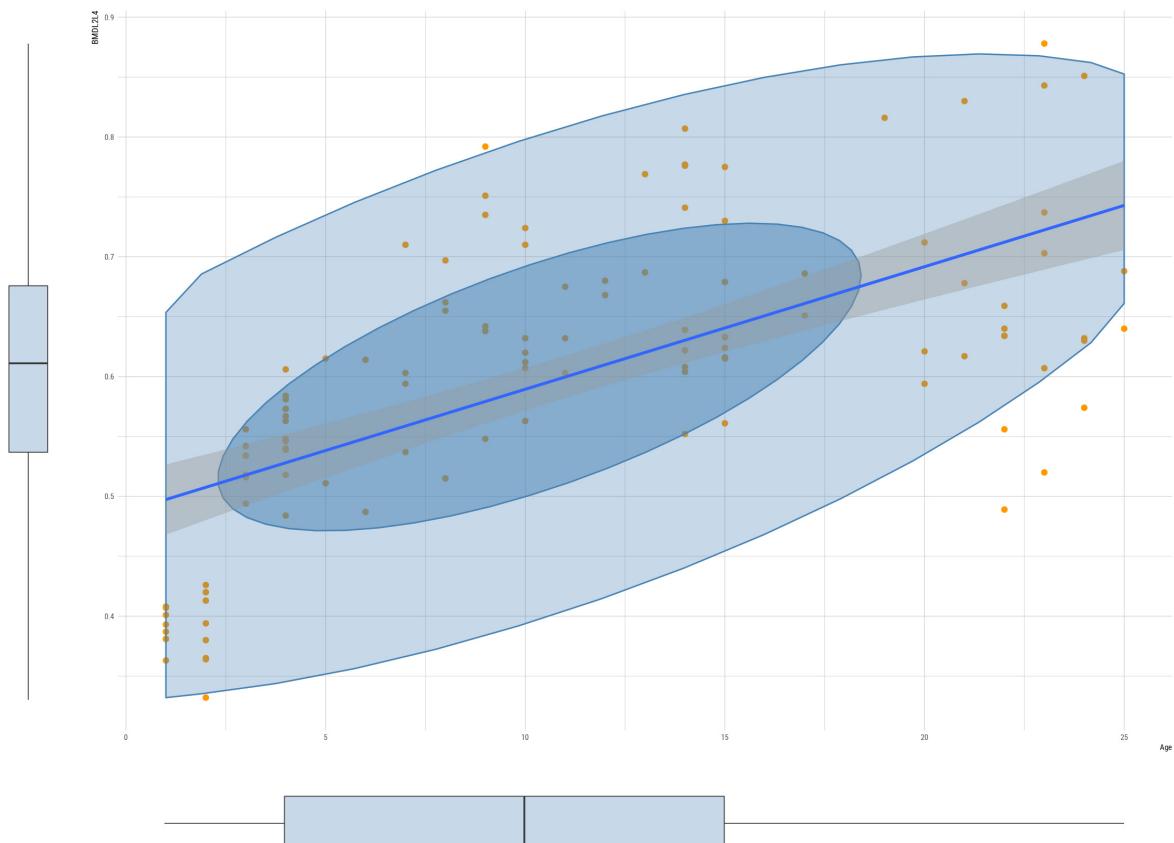
Table 6: Correlation coefficient

## 'Age' vs 'BMDL2L4'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	BMDL2L4	0.4079145	0.4024825	5.839688	75.09504	0	1

Table 6: Summary of linear model

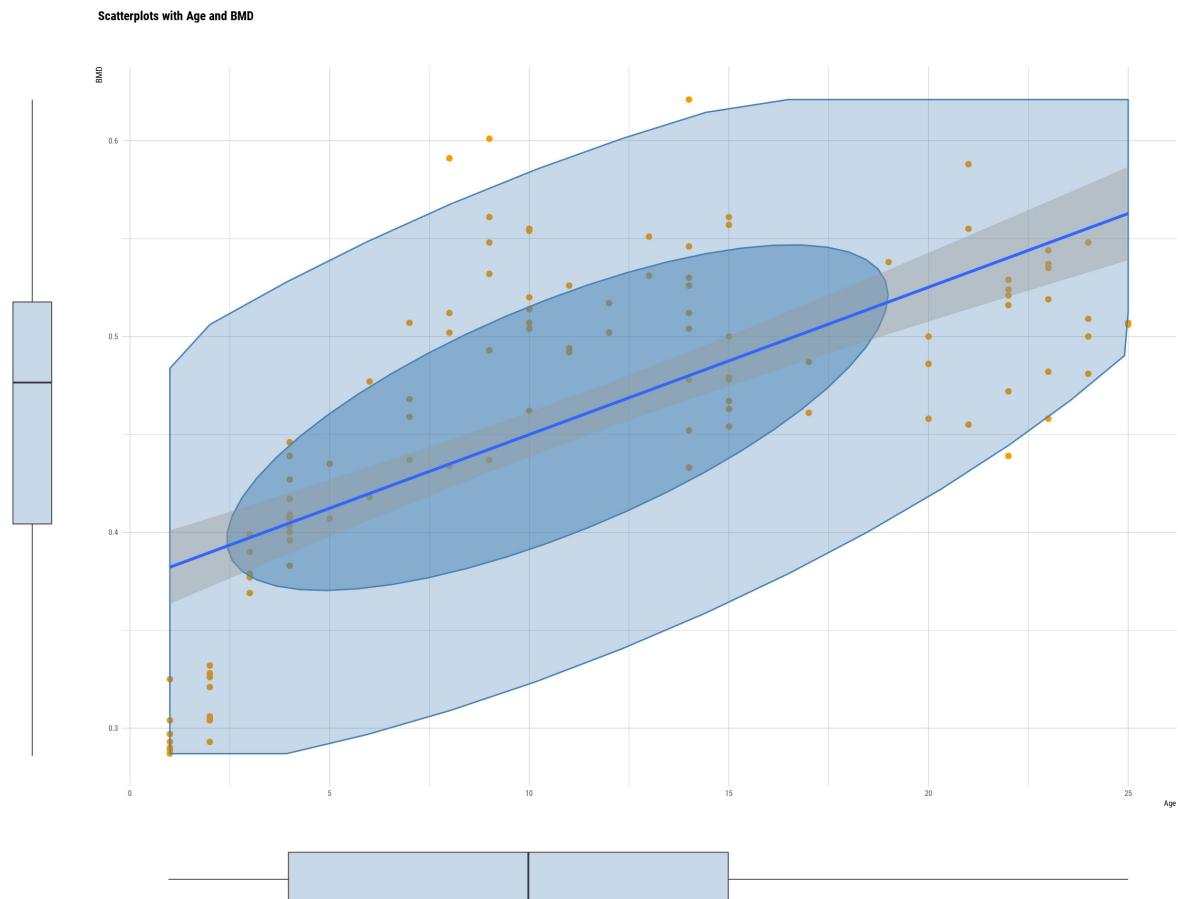
Scatterplots with Age and BMDL2L4



## 'Age' vs 'BMD'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	BMD	0.4762316	0.4714264	5.492462	99.10725	0	1

Table 6: Summary of linear model

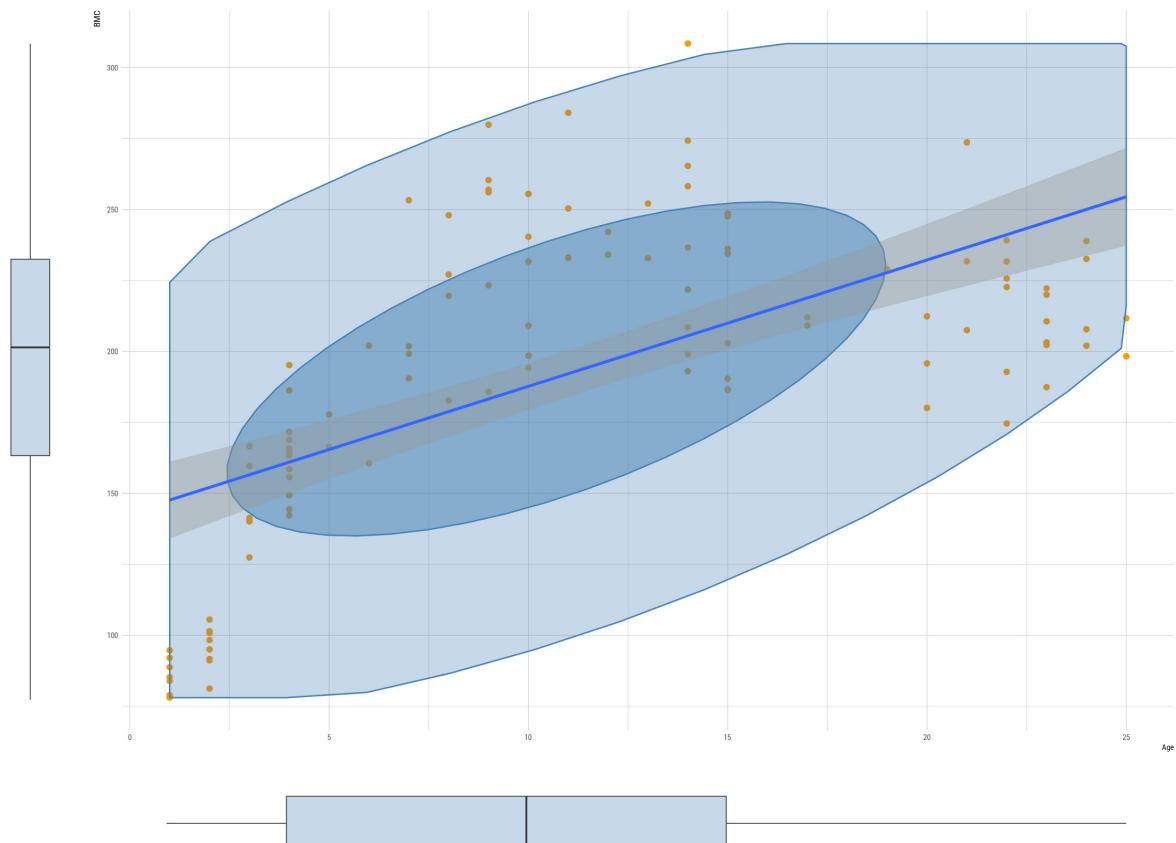


## 'Age' vs 'BMC'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	BMC	0.3778235	0.3721155	5.986241	66.19145	0	1

Table 6: Summary of linear model

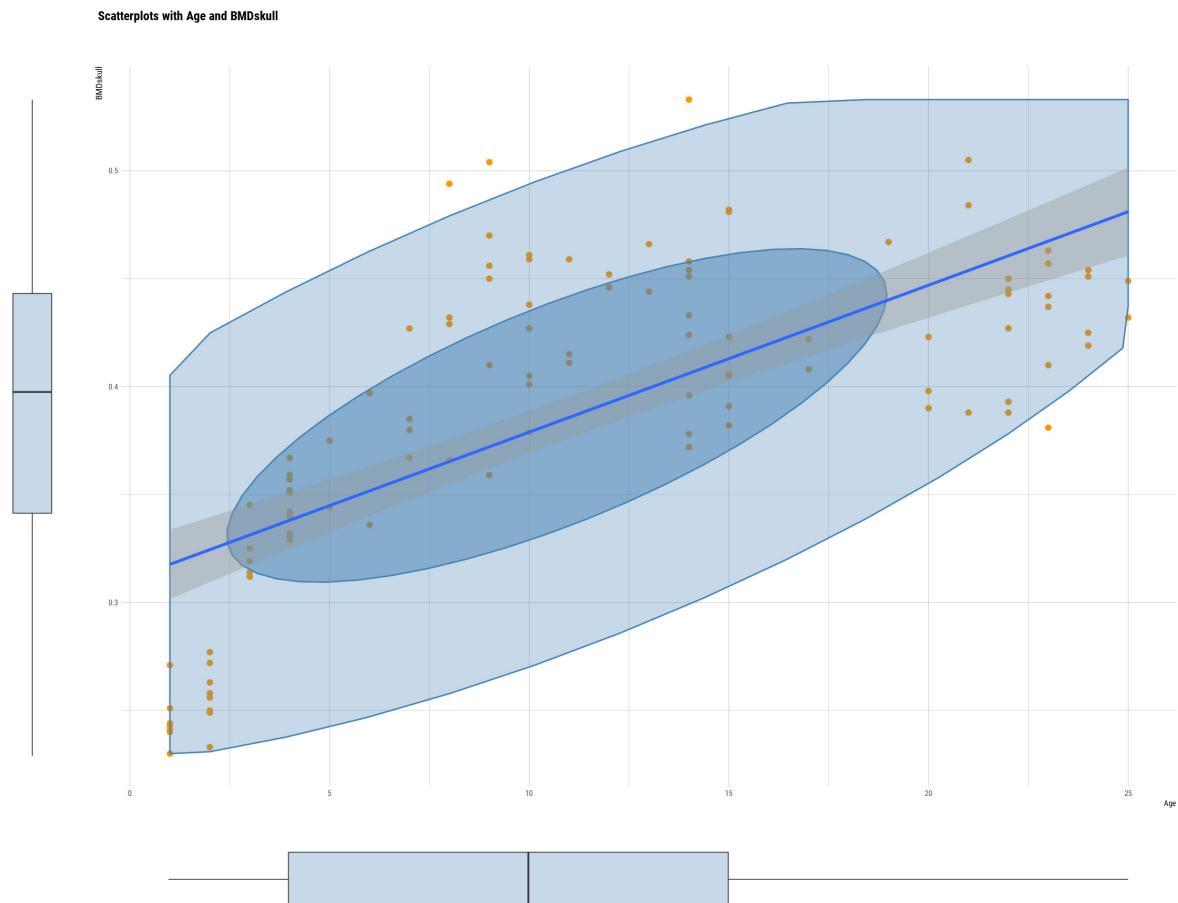
Scatterplots with Age and BMC



## 'Age' vs 'BMDskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	BMDskull	0.5038619	0.4993102	5.345628	110.6969	0	1

Table 6: Summary of linear model

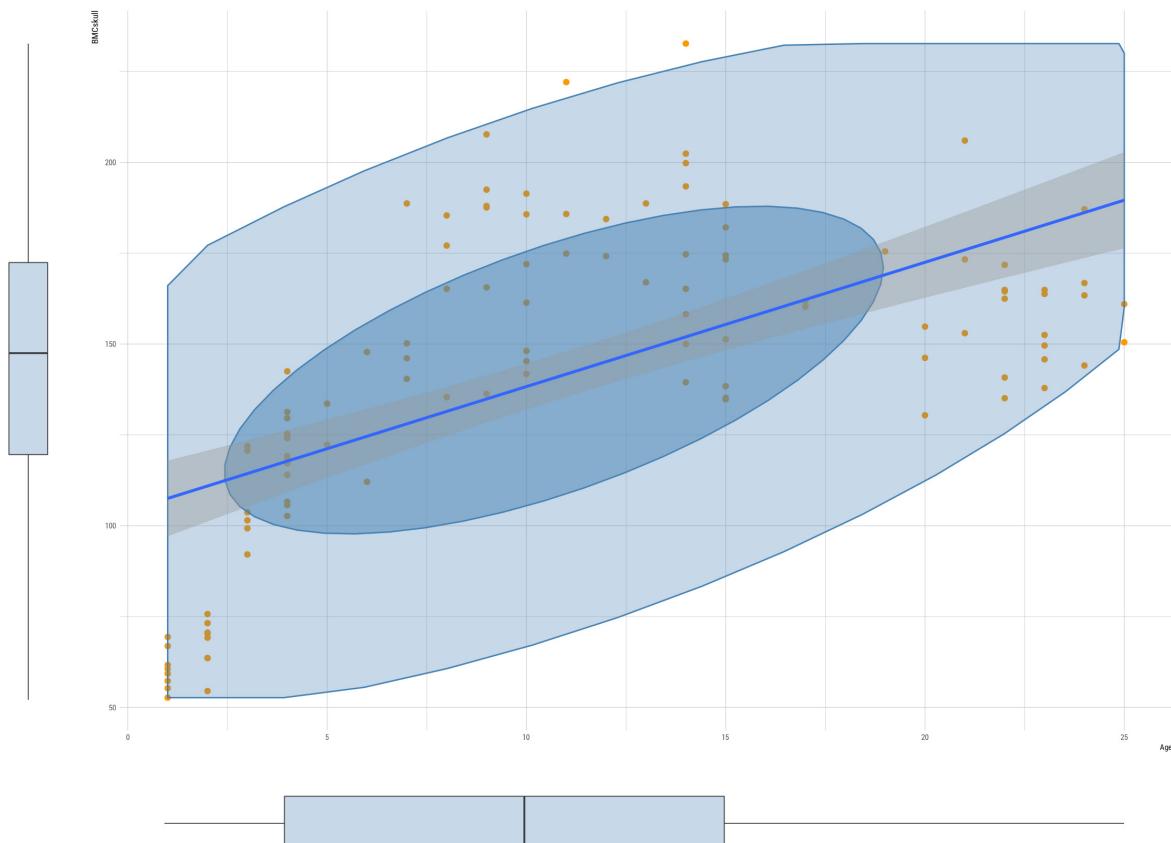


## 'Age' vs 'BMCskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	BMCskull	0.3782343	0.3725301	5.984264	66.30721	0	1

Table 6: Summary of linear model

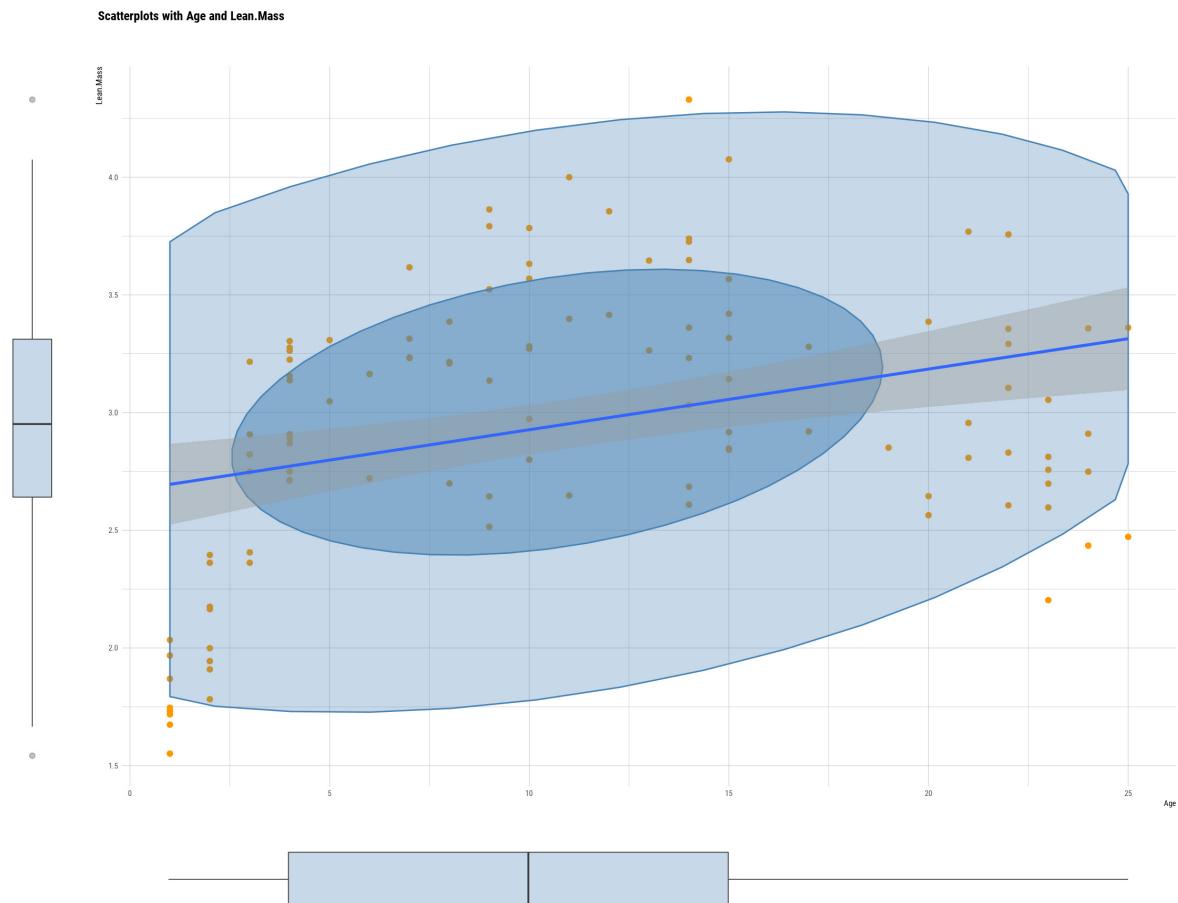
Scatterplots with Age and BMCskull



## 'Age' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
Age	Lean.Mass	0.1123475	0.1042039	7.150212	13.7958	0.0003231	1

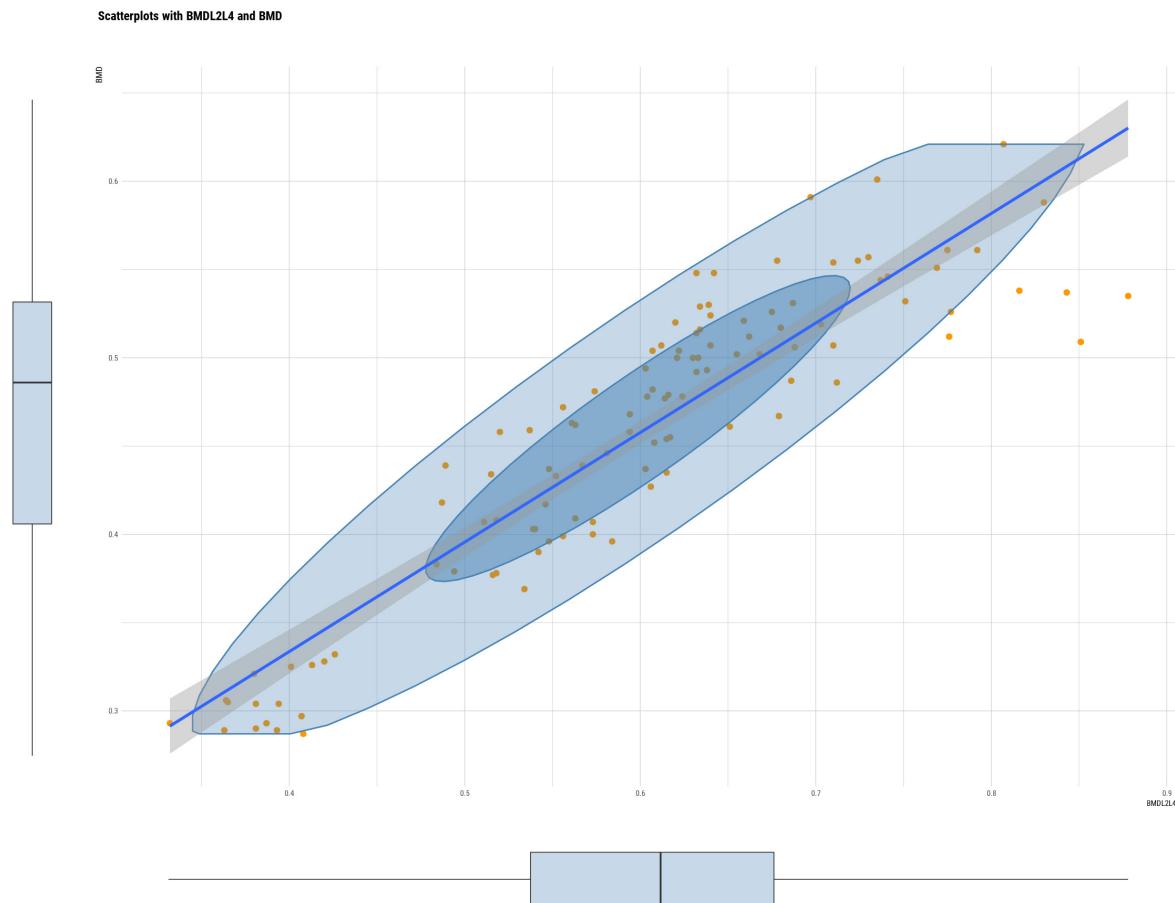
Table 6: Summary of linear model



## 'BMDL2L4' vs 'BMD'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDL2L4	BMD	0.8306909	0.8291376	0.050023	534.793	0	1

Table 6: Summary of linear model

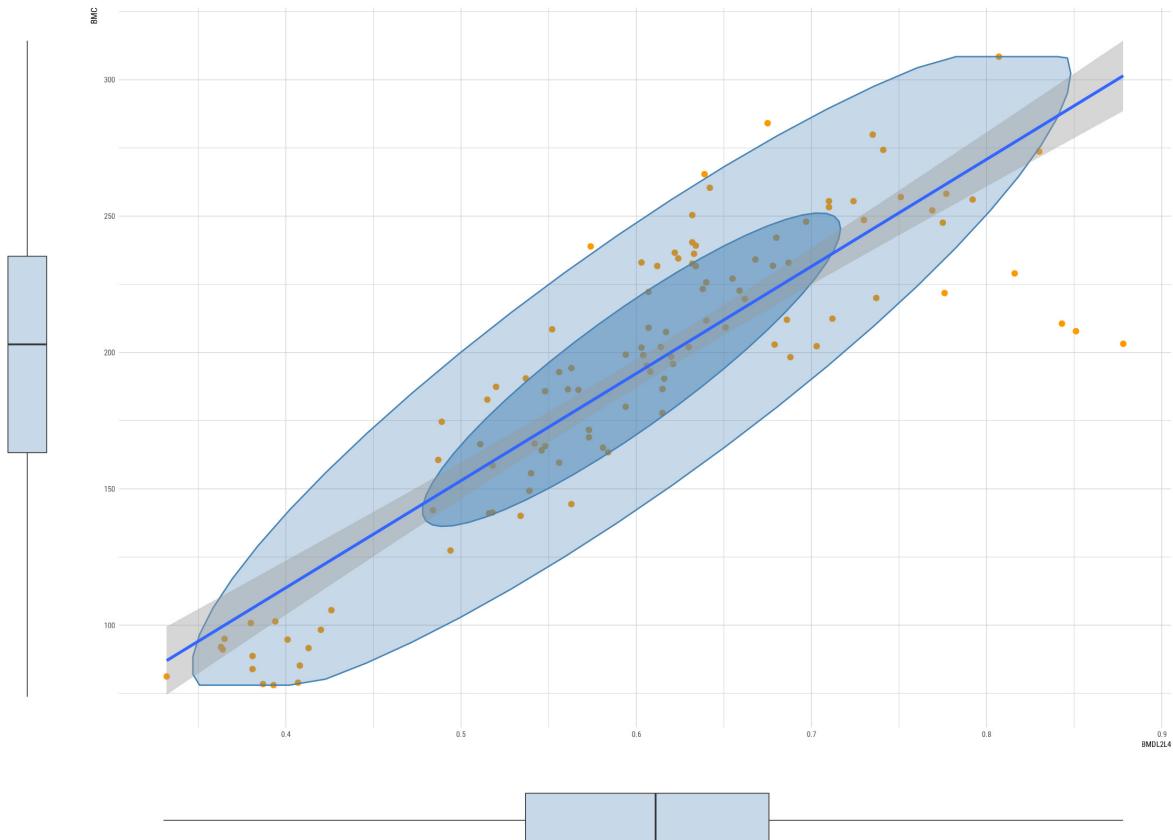


## 'BMDL2L4' vs 'BMC'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDL2L4	BMC	0.7543478	0.7520942	0.0602545	334.7169	0	1

Table 6: Summary of linear model

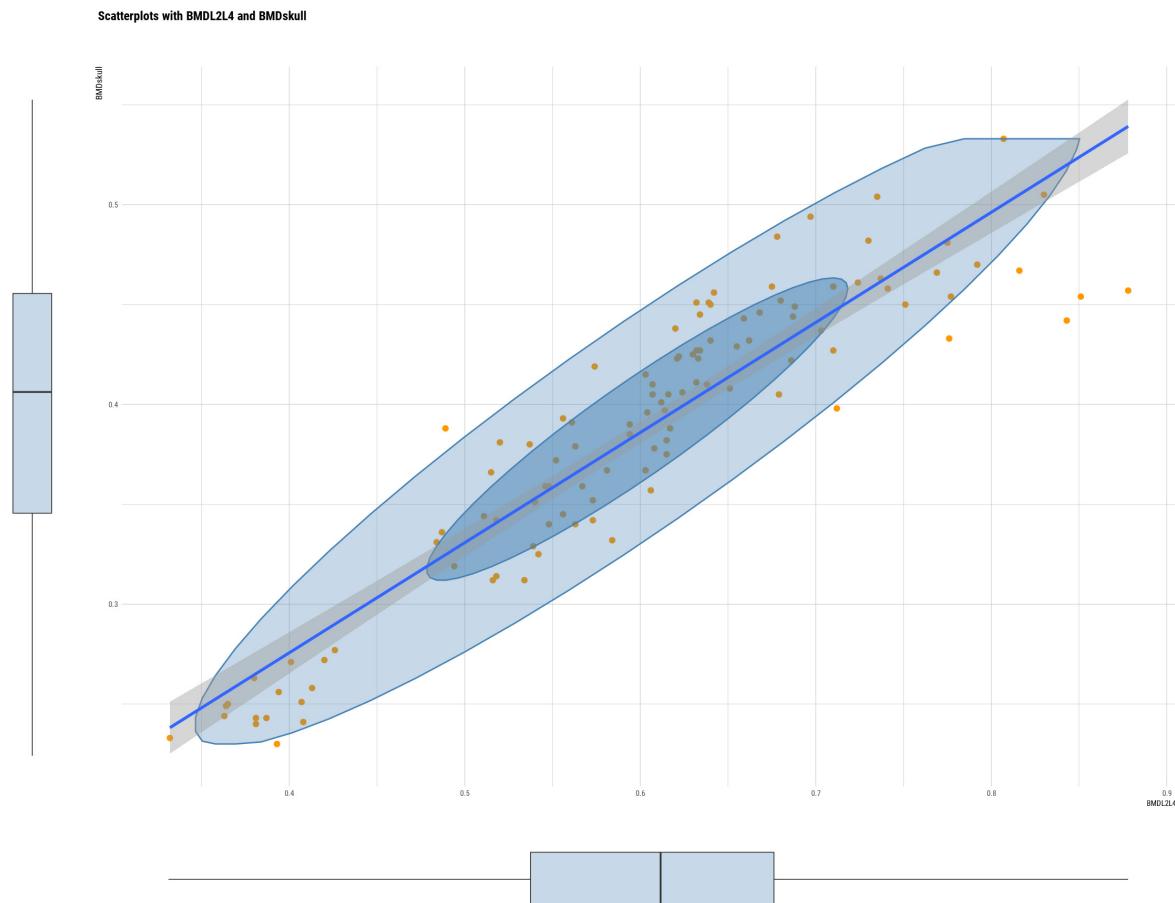
Scatterplots with BMDL2L4 and BMC



## 'BMDL2L4' vs 'BMDskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDL2L4	BMDskull	0.8482449	0.8468526	0.0473588	609.2624	0	1

Table 6: Summary of linear model

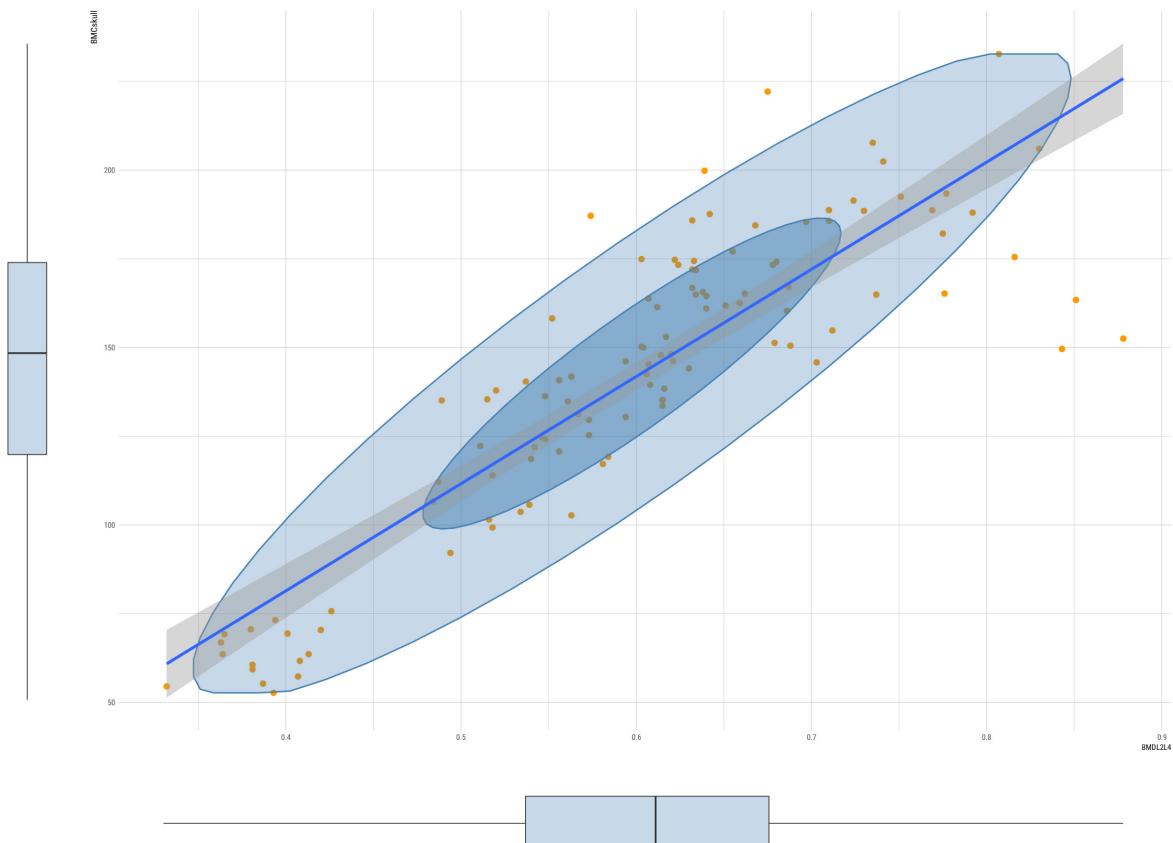


## 'BMDL2L4' vs 'BMCskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDL2L4	BMCskull	0.7568192	0.7545882	0.0599507	339.2262	0	1

Table 6: Summary of linear model

Scatterplots with BMDL2L4 and BMCskull

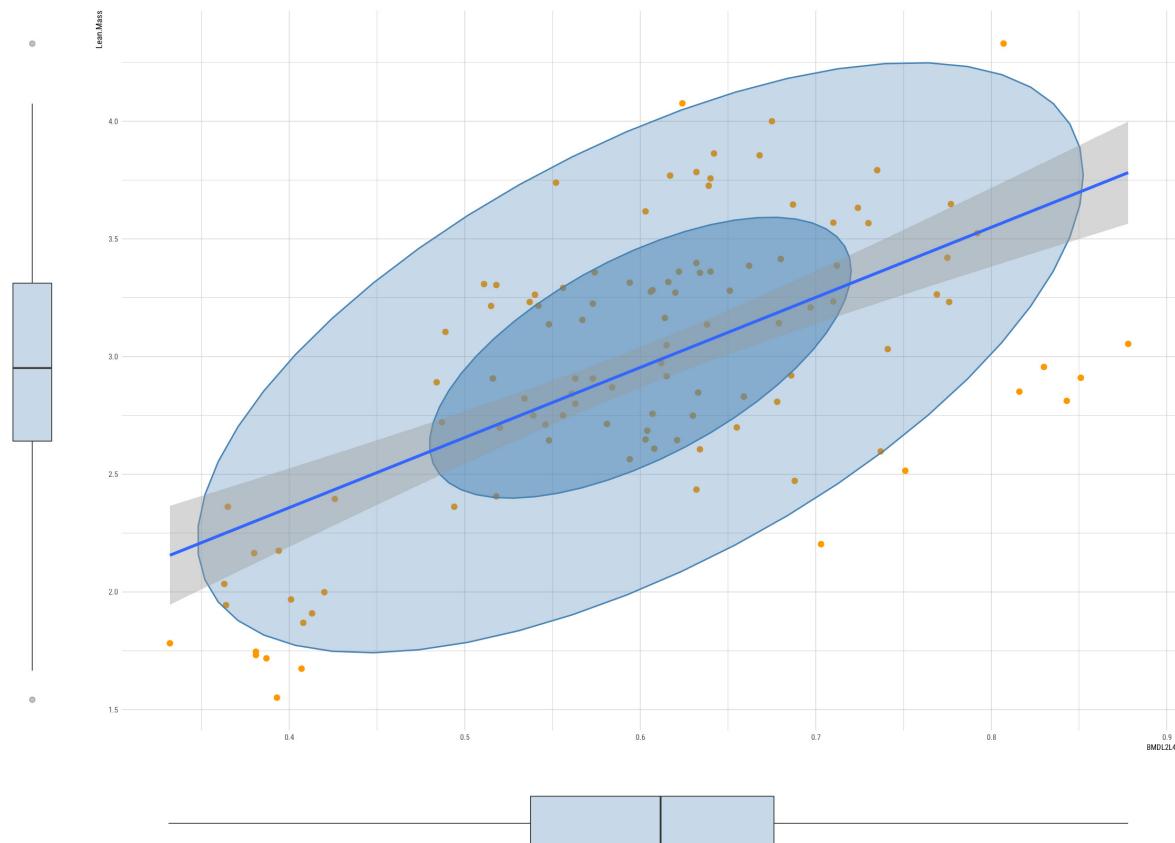


## 'BMDL2L4' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDL2L4	Lean.Mass	0.3844816	0.3788347	0.0953784	68.08651	0	1

Table 6: Summary of linear model

Scatterplots with BMDL2L4 and Lean.Mass

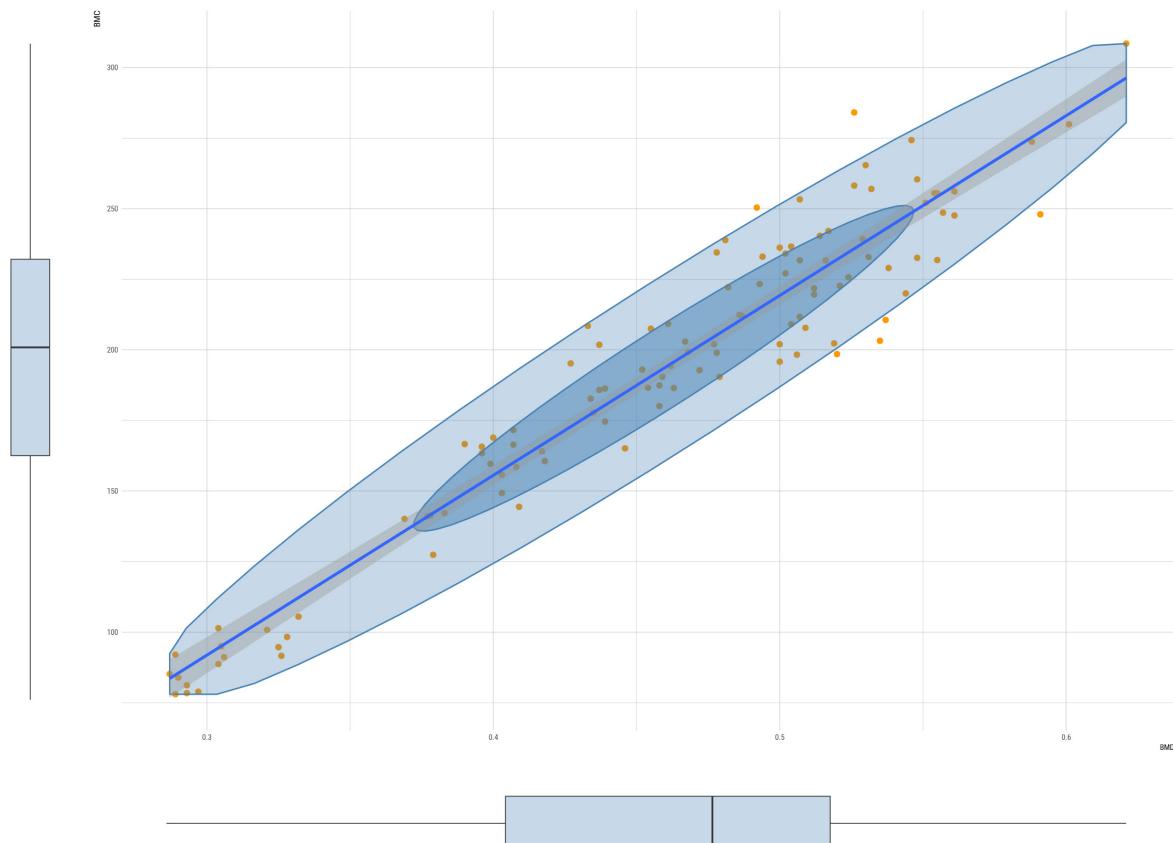


## 'BMD' vs 'BMC'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMD	BMC	0.9192402	0.9184993	0.0235181	1240.682	0	1

Table 6: Summary of linear model

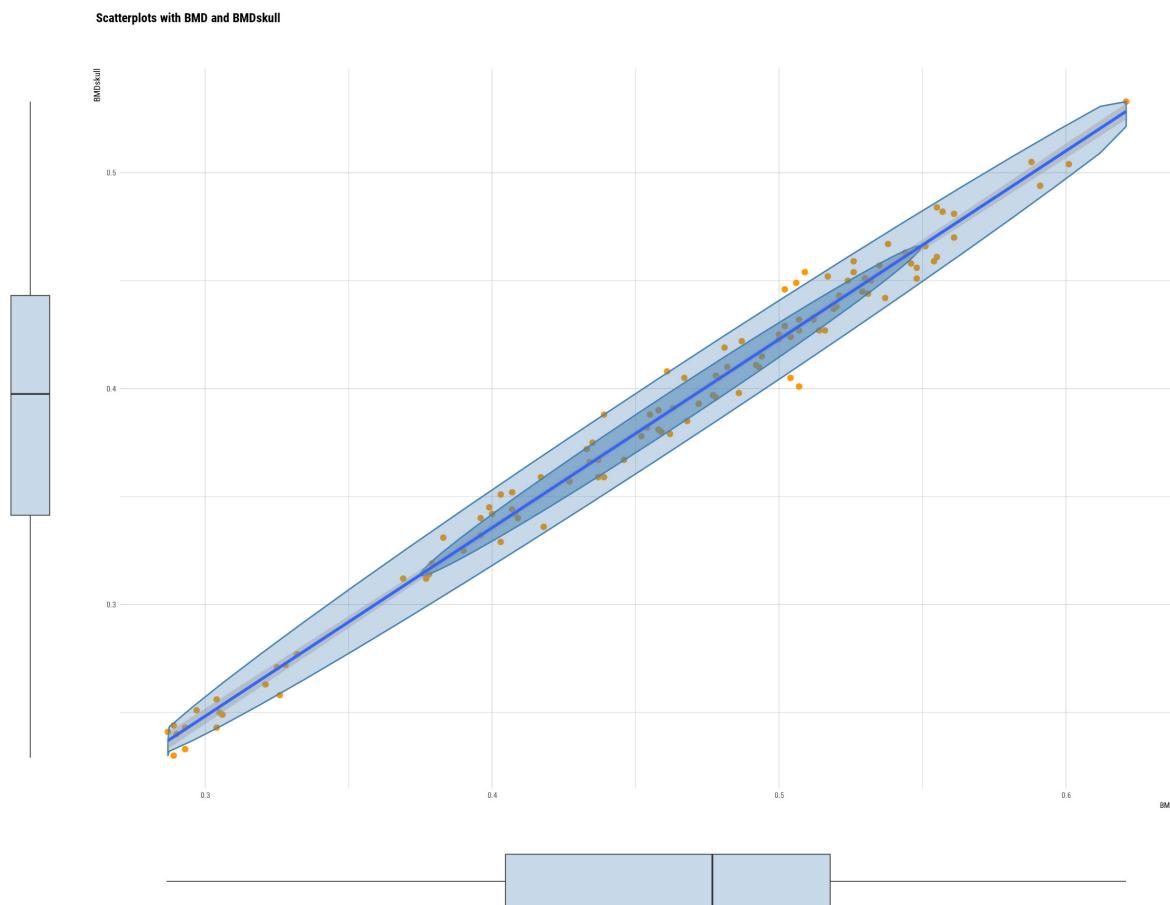
Scatterplots with BMD and BMC



## 'BMD' vs 'BMDskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMD	BMDskull	0.9852201	0.9850846	0.010061	7265.904	0	1

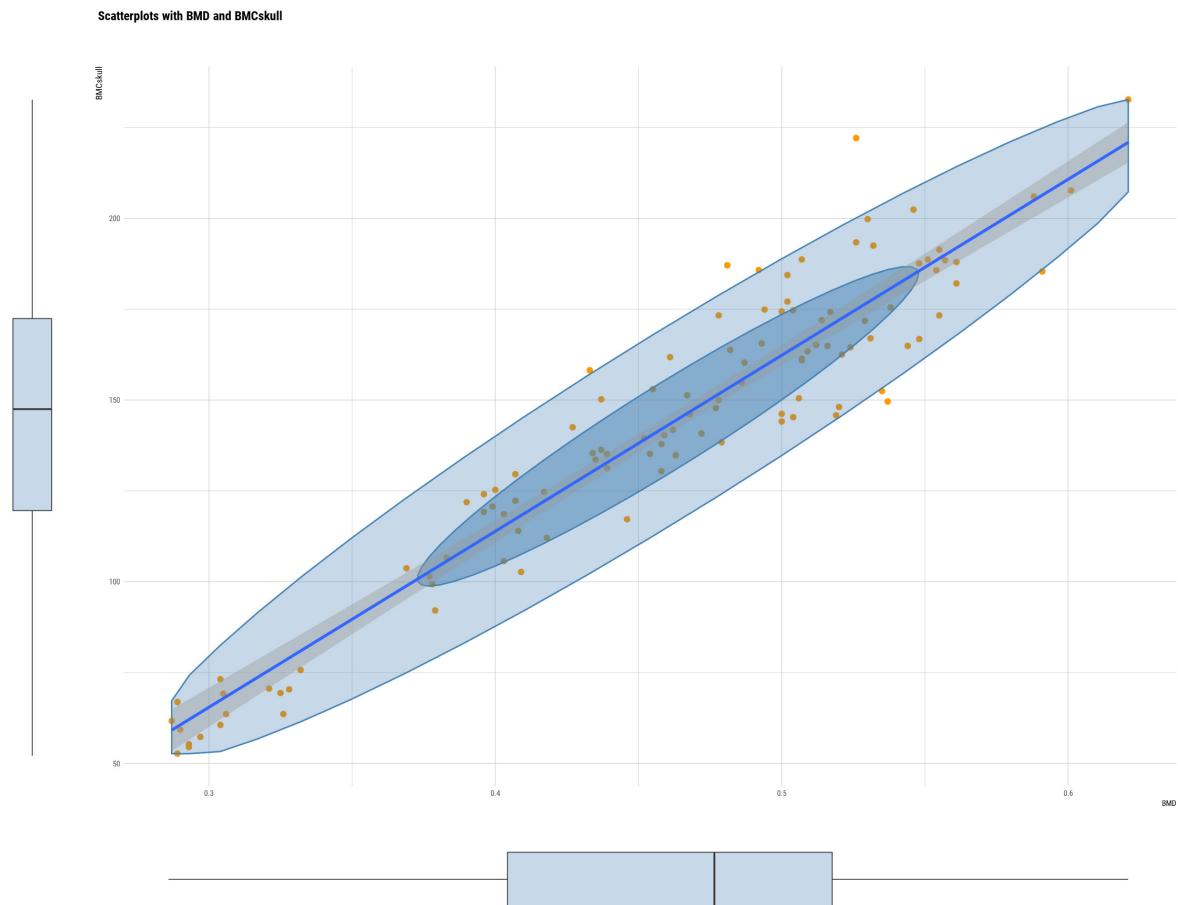
Table 6: Summary of linear model



## 'BMD' vs 'BMCskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMD	BMCskull	0.9020113	0.9011123	0.0259055	1003.373	0	1

Table 6: Summary of linear model

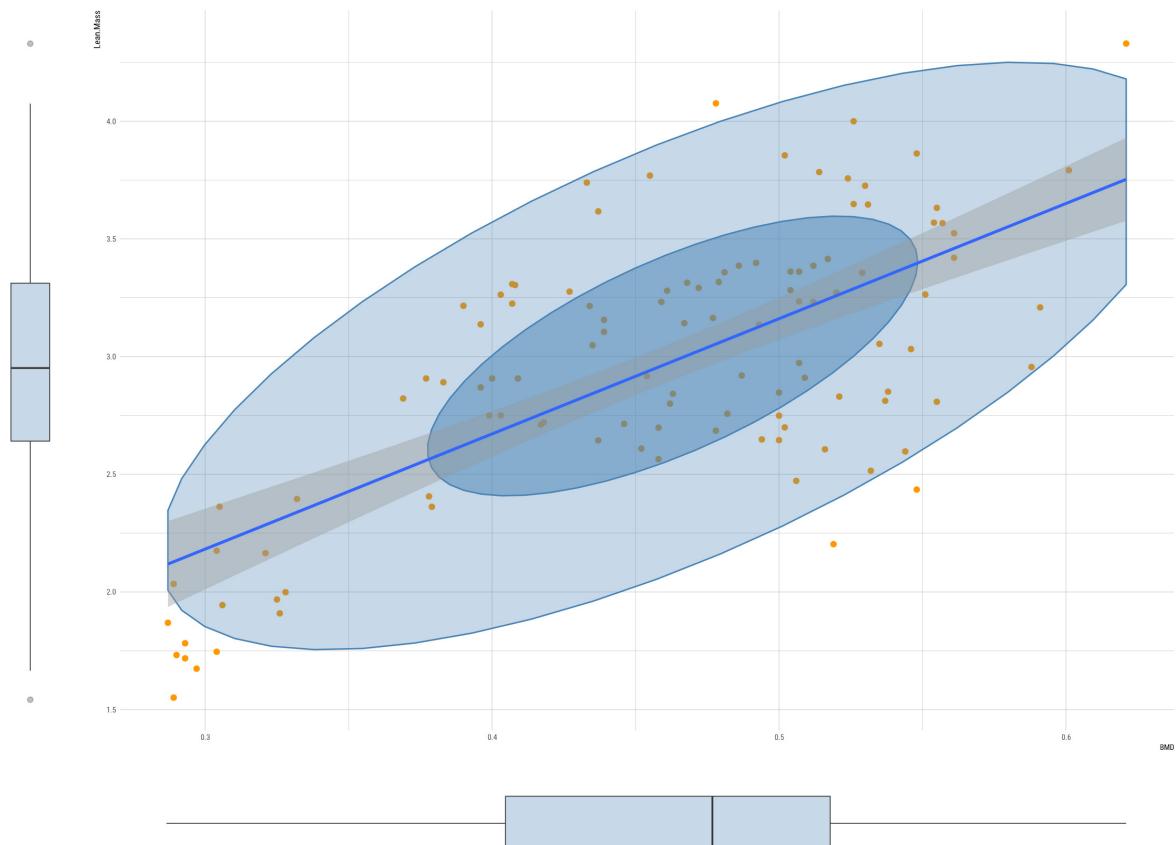


## 'BMD' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMD	Lean.Mass	0.481369	0.4766109	0.0595983	101.1687	0	1

Table 6: Summary of linear model

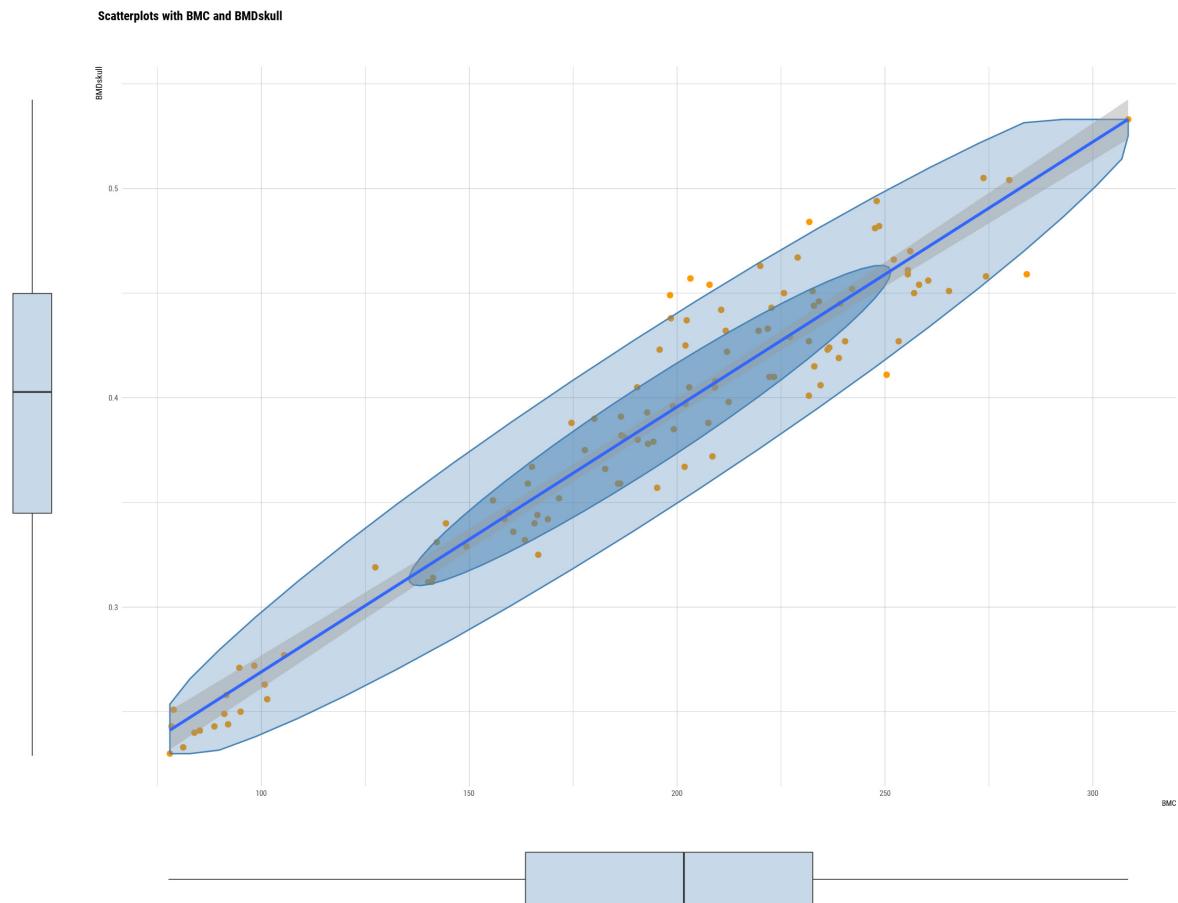
Scatterplots with BMD and Lean.Mass



## 'BMC' vs 'BMDskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMC	BMDskull	0.9146531	0.9138701	16.06186	1168.14	0	1

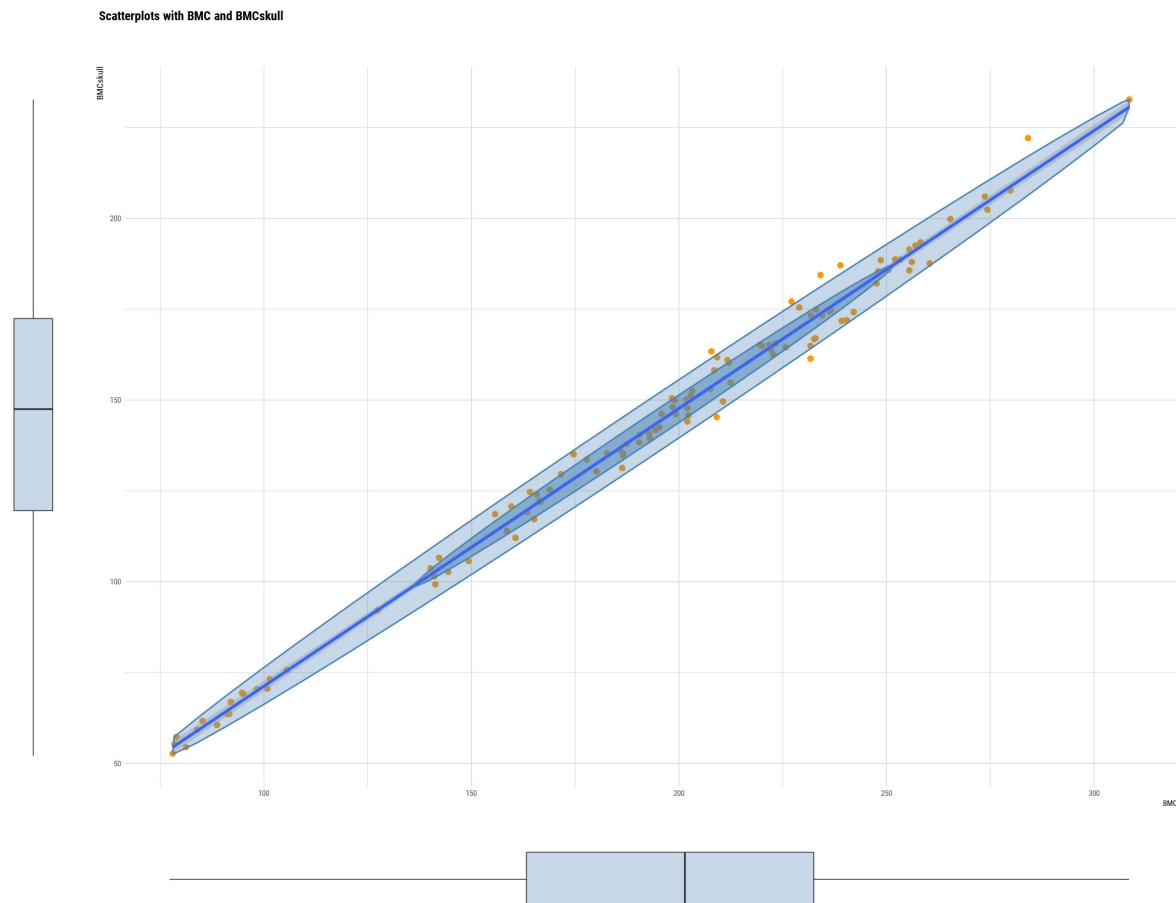
Table 6: Summary of linear model



## 'BMC' vs 'BMCskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMC	BMCskull	0.9915676	0.9914902	5.048675	12817.34	0	1

Table 6: Summary of linear model

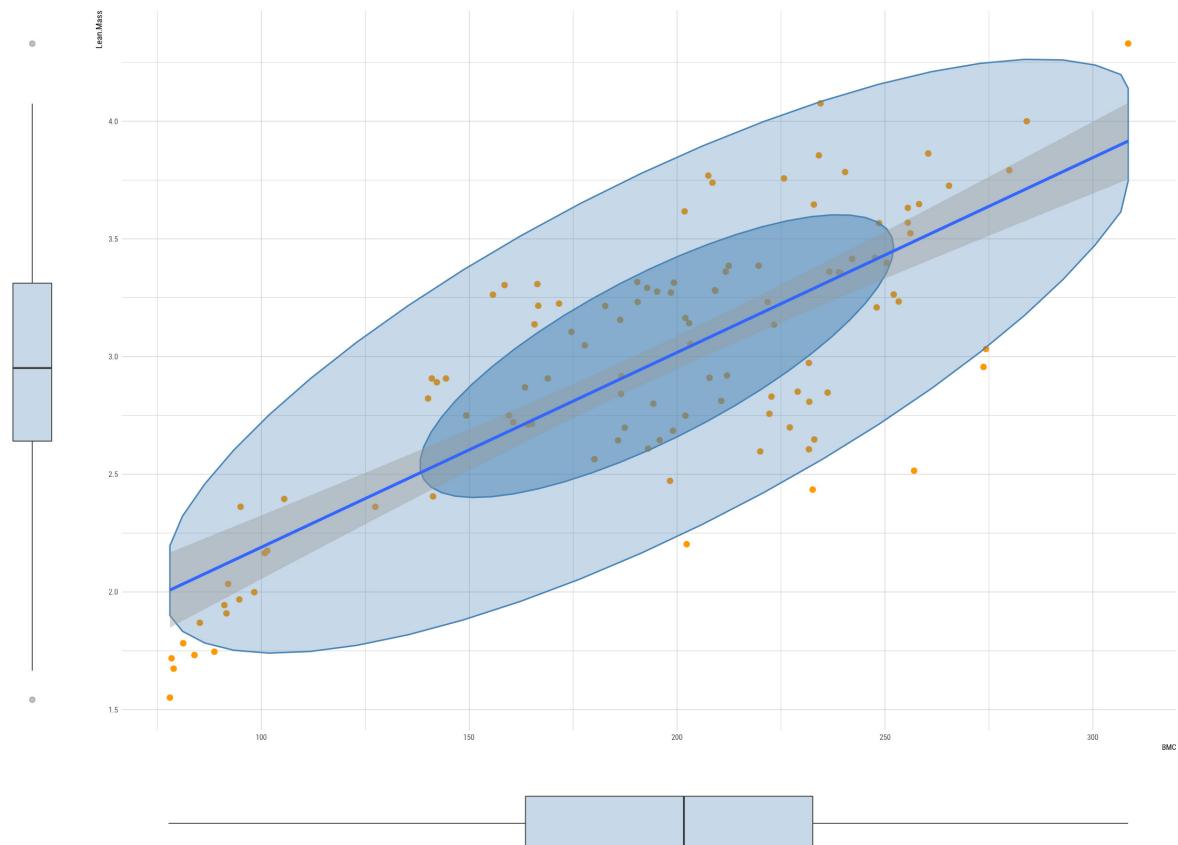


## 'BMC' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMC	Lean.Mass	0.6075864	0.6039863	34.44086	168.7682	0	1

Table 6: Summary of linear model

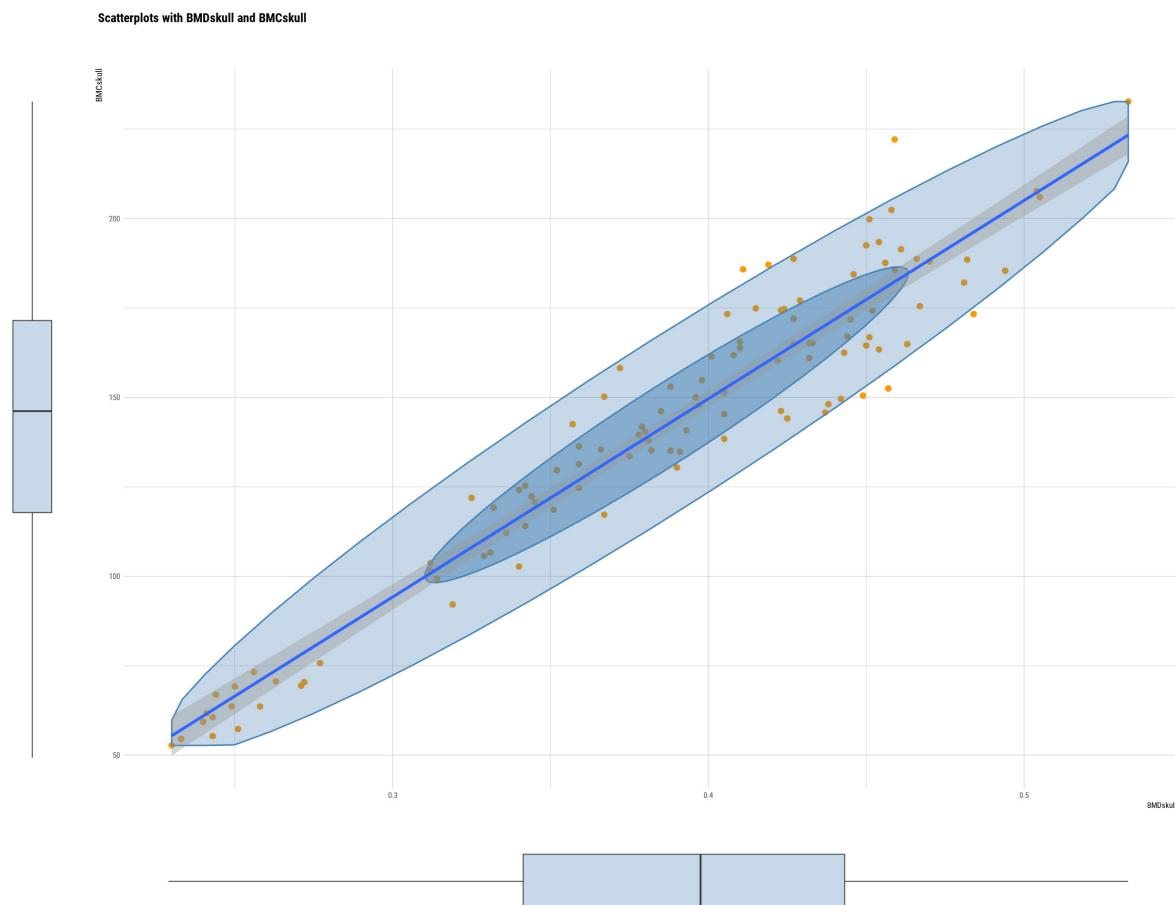
Scatterplots with BMC and Lean.Mass



## 'BMDskull' vs 'BMCskull'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDskull	BMCskull	0.913675	0.912883	0.0213837	1153.67	0	1

Table 6: Summary of linear model

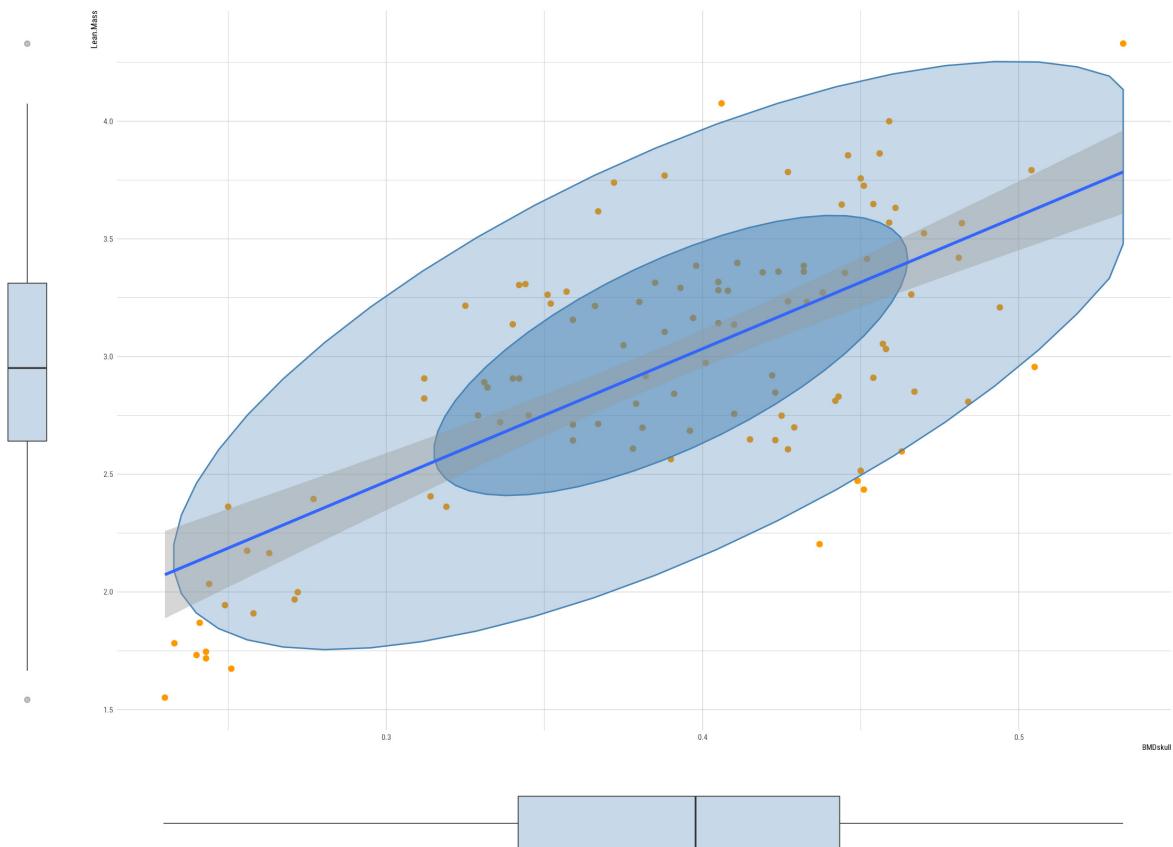


## 'BMDskull' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMDskull	Lean.Mass	0.4956227	0.4909954	0.0516883	107.1081	0	1

Table 6: Summary of linear model

Scatterplots with BMDskull and Lean.Mass

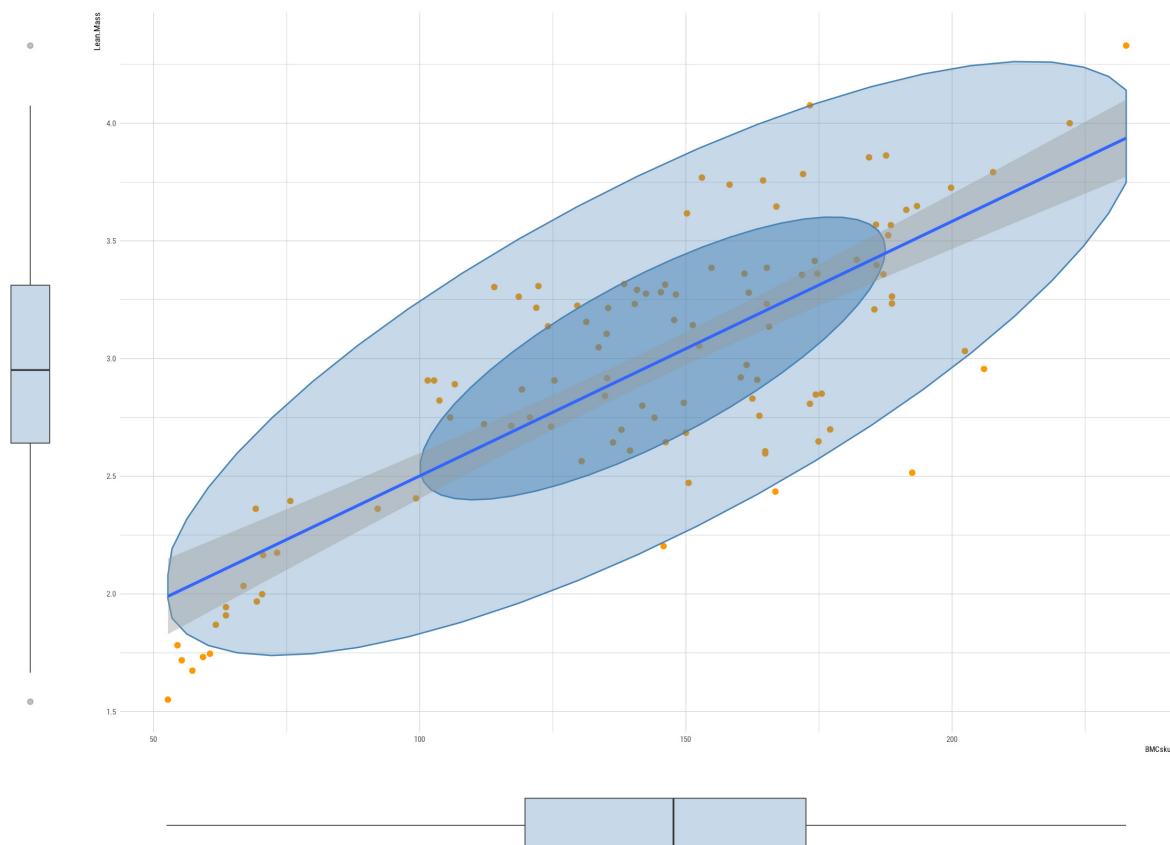


## 'BMCskull' vs 'Lean.Mass'

first variable	second variable	r.squared	adj.r.squared	sigma	statistic	p.value	df
BMCskull	Lean.Mass	0.6114464	0.6078816	26.30651	171.5275	0	1

Table 6: Summary of linear model

Scatterplots with BMCskull and Lean.Mass



## Compare Categorical Variables

The number of categorical variables is less than 2.

# Multivariate Analysis

## Correlation Analysis

### Correlation Coefficient Matrix

first variable	second variable						
	Age	BMDL2L4	BMD	BMC	BMDskull	BMCskull	Lean.Mass
Age	NA	0.639	0.690	0.615	0.710	0.615	0.335
BMDL2L4	0.639	NA	0.911	0.869	0.921	0.870	0.620
BMD	0.690	0.911	NA	0.959	0.993	0.950	0.694
BMC	0.615	0.869	0.959	NA	0.956	0.996	0.779
BMDskull	0.710	0.921	0.993	0.956	NA	0.956	0.704
BMCskull	0.615	0.870	0.950	0.996	0.956	NA	0.782
Lean.Mass	0.335	0.620	0.694	0.779	0.704	0.782	NA

Table 7: Matrix table of correlation coefficient

## Correlation Plot

