

# Overview

## Dataset info

Number of variables	19
Number of observations	59757
Total Missing (%)	1.0%
Total size in memory	8.7 MiB
Average record size in memory	152.0 B

## Variables types

Numeric	6
Categorical	10
Boolean	0
Date	0
Text (Unique)	0
Rejected	3
Unsupported	0

## Warnings

Appointment Date has a high cardinality: 1086 distinct values	Warning
CHI has a high cardinality: 4914 distinct values	Warning
Drug Name has a high cardinality: 62 distinct values	Warning
Drug Type has constant value Cytotoxic	Rejected
Given Dose has 11078 / 18.5% zeros	Zeros
Height has 2227 / 3.7% zeros	Zeros
Last Toxicity Date has 10857 / 18.2% missing values	Missing
Last Toxicity Date has a high cardinality: 871 distinct values	Warning
Regime has a high cardinality: 223 distinct values	Warning
Required Dose is highly correlated with Drug Dose ( $\rho = 0.96787$ )	Rejected
Surface Area is highly correlated with Weight ( $\rho = 0.91716$ )	Rejected
Weight has 2227 / 3.7% zeros	Zeros
Dataset has 23 duplicate rows	Warning

# Variables

## Appointment Date

Categorical

**Distinct count** 1086

**Unique (%)** 1.8%  
**Missing (%)** 0.0%  
**Missing (n)** 0

05-01-2016	113	
02-03-2016	111	
30-12-2015	110	
Other values (1083)		59423

[Toggle details](#)

## CHI

Categorical

**Distinct count** 4914  
**Unique (%)** 8.2%  
**Missing (%)** 0.0%  
**Missing (n)** 0

0307541363	148	
2606701259	138	
1010952269	123	
Other values (4911)		59348

[Toggle details](#)

## Consultant Code

Categorical

**Distinct count** 43  
**Unique (%)** 0.1%  
**Missing (%)** 0.0%  
**Missing (n)** 0

LH	5420	
AB	4173	
MM	3858	
Other values (40)		46306

[Toggle details](#)

## Cycle

Numeric

**Distinct count** 85  
**Unique (%)** 0.1%  
**Missing (%)** 0.0%  
**Missing (n)** 0

Infinite (%)

0.0%

Infinite (n)

0

Mean

4.0003

Minimum

1

Maximum

85

Zeros (%)

0.0%



[Toggle details](#)

### Drug Dose

Numeric

Distinct count

102

Unique (%)

0.2%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

Infinite (n)

0

Mean

566.32

Minimum

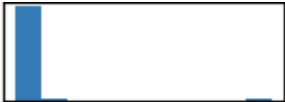
0

Maximum

30000

Zeros (%)

0.5%



[Toggle details](#)

### Drug Name

Categorical

Distinct count

62

Unique (%)

0.1%

Missing (%)

0.0%

Missing (n)

0

FLUOROURACIL	6994	
PACLITAXEL	6097	
CAPECITABINE	5823	
Other values (59)		40843

[Toggle details](#)

29/04/2019

Profile report

Drug Status

Categorical

Distinct count

5

Unique (%)

0.0%

Missing (%)

0.2%

Missing (n)

99

Given

48682

Authorised

6017

Planned

4597

Toggle details

Drug Type

Constant

This variable is constant and should be ignored for analysis

Constant value

Cytotoxic

Given Dose

Numeric

Distinct count

502

Unique (%)

0.8%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

Infinite (n)

0

Mean

728.33

Minimum

0

Maximum

30000

Zeros (%)

18.5%

Toggle details

Height

Numeric

Distinct count

63

Unique (%)

0.1%

Missing (%)

0.0%

Missing (n)

0

Infinite (%)

0.0%

file:///D:/Users/eblackle/AppData/Local/Temp/Temp1\_profiling (002).zip/profiling/Chemocare\_Treatment.html

4/10

29/04/2019

Profile report

Infinite (n)

0

Mean

1.6037

Minimum


0

Maximum

1.98

Zeros (%)

3.7%



Toggle details

Intention

Categorical

Distinct count

8

Unique (%)

0.0%

Missing (%)

0.0%

Missing (n)

0

Palliative

29628

Adjuvant

12651

Curative

7220

Other values (5)

10258

Toggle details

Last Toxicity Date

Categorical

Distinct count

871

Unique (%)

1.5%

Missing (%)

18.2%

Missing (n)

10857

29-12-2015

115

03-05-2016

109

08-07-2014

108

Other values (867)

48568

(Missing)

10857

Toggle details

Patient Type

Categorical

Distinct count

3

Unique (%)

0.0%

Missing (%)

0.0%

29/04/2019	Profile report			
Missing (n)	0			
		D		48248
		I	8045	
		O	3464	
Toggle details				
<div>Regime</div> <div>Categorical</div> <div>Distinct count223</div> <div>Unique (%)0.4%</div> <div>Missing (%)0.0%</div> <div>Missing (n)0</div> <div>CAPOX3501</div> <div>FEC 802931</div> <div>PACLITAX WKLY2880</div> <div>Other values (220)50445</div> <div>Toggle details</div>				
<div>Required Dose</div> <div>Highly correlated</div> <div>This variable is highly correlated with Drug Dose and should be ignored for analysis</div> <div>Correlation0.96787</div>				
<div>Surface Area</div> <div>Highly correlated</div> <div>This variable is highly correlated with weight and should be ignored for analysis</div> <div>Correlation0.91716</div>				
<div>Tumour Group</div> <div>Categorical</div> <div>Distinct count15</div> <div>Unique (%)0.0%</div> <div>Missing (%)0.0%</div> <div>Missing (n)0</div> <div>Breast14417</div> <div>GI Lower11131</div> <div>Lung and Chest7744</div> <div>Other values (12)26465</div>				

[Toggle details](#)

Weight

Numeric

Distinct count	858
Unique (%)	1.4%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0
Mean	72.172
Minimum	0
Maximum	167.5
Zeros (%)	3.7%



[Toggle details](#)

age\_at\_diagnosis

Numeric

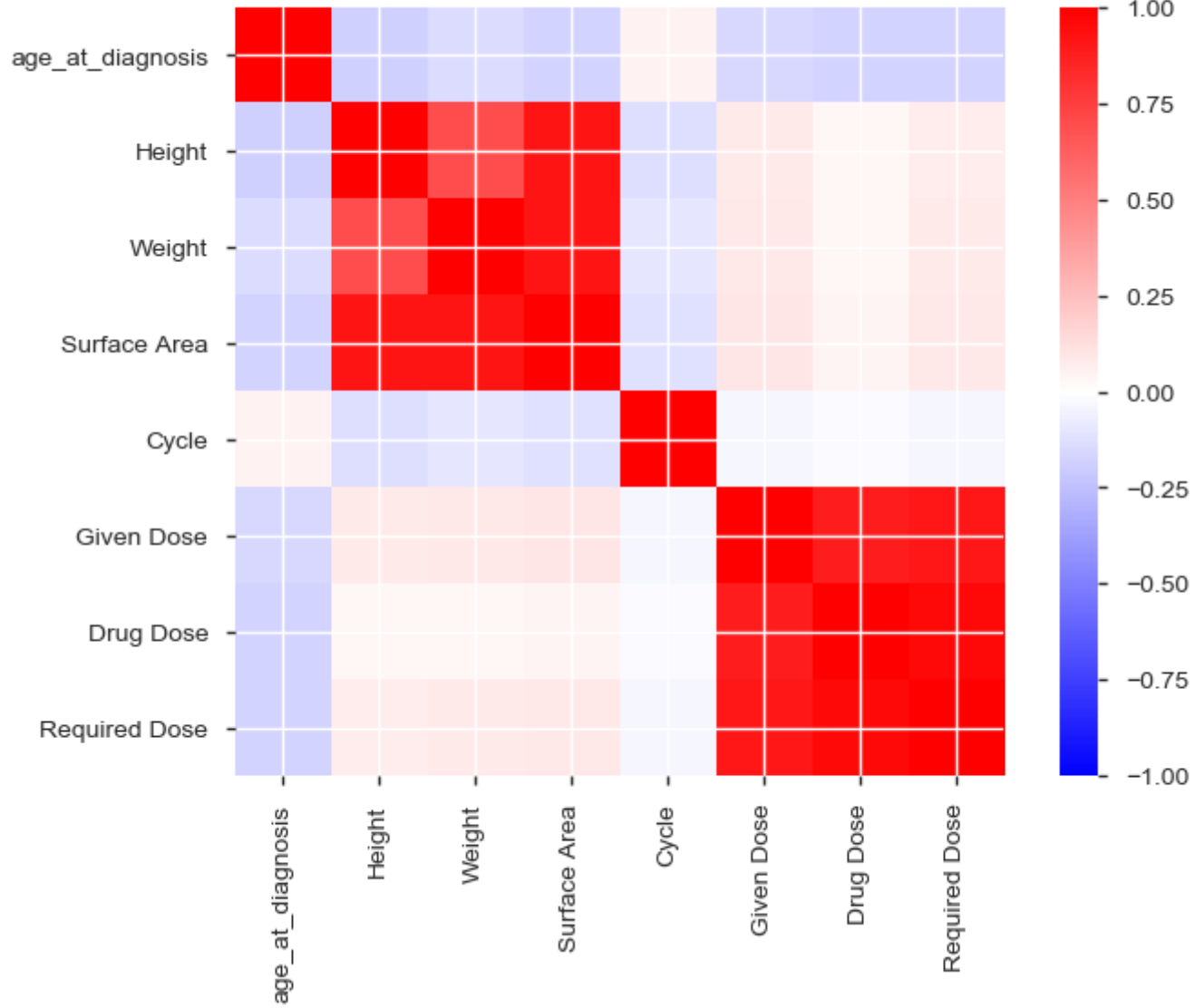
Distinct count	16140
Unique (%)	27.0%
Missing (%)	0.1%
Missing (n)	34
Infinite (%)	0.0%
Infinite (n)	0
Mean	59.307
Minimum	16.321
Maximum	94.827
Zeros (%)	0.0%



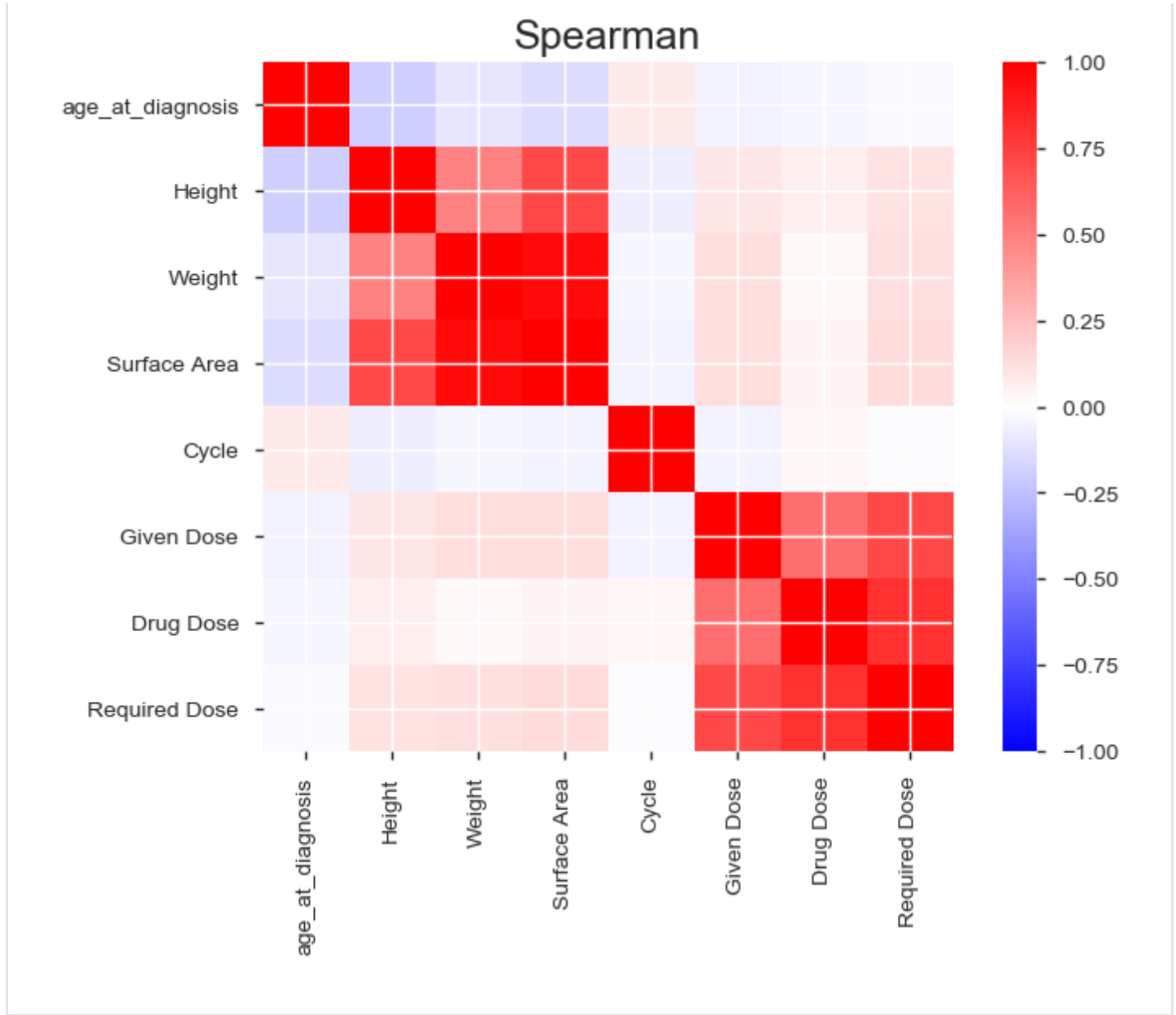
[Toggle details](#)

Correlations

Pearson







# Sample

	CHI	Appointment Date	Last Toxicity Date	Tumour Group	age_at_diagnosis	Height
0	2608891225	02-01-2014	31-12-2013	Bone Sarcoma	24.369863	1.00
1	2608891225	02-01-2014	31-12-2013	Bone Sarcoma	24.369863	1.00
2	2608891225	02-01-2014	31-12-2013	Bone Sarcoma	24.369863	1.00
3	2608891225	02-01-2014	31-12-2013	Bone Sarcoma	24.369863	1.00
4	2608891225	03-01-2014	31-12-2013	Bone Sarcoma	24.372603	1.00

