

Quality by Design Template Investigate Data Quality by Profiling

Table of Contents

Data Set Provenance and Scope.....	2
Listing of Data Set Attributes	2
Listing of Variables and Their Attributes	2
Statistical Profile of Data Set	3
Visual Profile of Data Set.....	6
Heatmap Display of Data Frame Data For Numeric variables	6
Histograms and Frequency Displays of Data Frame Data	7

Data Set Provenance and Scope

Listing of Data Set Attributes

Company	Title	Record Create Date
Discovery Pharmaceuticals, Inc.	Drug Product Manufacturing Data Set	6/23/2007 12:00:00 AM
Drug Name	Drug CAS	Investigator
Famotidine	76824-35-6	John Smith
ATC Code	Dosage	Formulation Code
A02BA03	10 mg	ZX2345

Listing of Variables and Their Attributes

	Variables	Labels	Units	QbDTypes
1	LotNo	API Lot No.		material
2	ParticleSize	API Particle Size		material
3	MillTime	Mill Time	min	process
4	ScreenSize	Screen Size		process
5	MgStSup	MgStearate Supplier		material
6	LactoseSup	Lactose Supplier		material
7	SugarSup	Sugar Supplier		material
8	TalcSup	Talc Supplier		material
9	BlendTime	Blend Time	min	process
10	BlendSpeed	Blend Speed	rpm	process
11	CompUnit	Compaction Unit		equipment
12	CompForce	Compaction Force	kN	process
13	CoatingSup	Coating Supplier		material
14	CoatingVis	Coating Viscosity	poise	material
15	InletTemp	Inlet Temp	deg C	process
16	ExhaustTemp	Exhaust Temp	deg C	process
17	SprayRate	Spray Rate	ml/min	process
18	AtomPres	Atomizer Pressure	atm	process
19	Dissolution	Dissolution at 60 min	%	performance
20	LotAccept	Lot Acceptance		quality

Statistical Profile of Data Set

ManufacturingData

20 Variables 90 Observations

LotNo

n	missing	unique
90	0	90

lowest : API0001 API0004 API00017 API00036 API00055
highest: API000857 API000865 API000888 API000895 API000897

ParticleSize

n	missing	unique
90	0	3

Small (28, 31%), Medium (33, 37%), Large (29, 32%)

MillTime

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	26	17.01	5.0	6.9	10.0	17.5	24.0	28.0	29.0

lowest : 5 6 7 8 9, highest: 26 27 28 29 30

ScreenSize

n	missing	unique
90	0	3

4 (24, 27%), 5 (34, 38%), 3 (32, 36%)

MgStSup

n	missing	unique
90	0	2

Smith Ind (41, 46%), Jones Inc (49, 54%)

LactoseSup

n	missing	unique
90	0	2

James Ind (40, 44%), Bond Inc (50, 56%)

SugarSup

n	missing	unique
90	0	2

Sour (36, 40%), Sweet (54, 60%)

TalcSup

n	missing	unique
90	0	2

Rough (45, 50%), Smooth (45, 50%)

BlendTime

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	14.92	12.09	13.18	13.98	15.05	16.09	16.81	17.19

lowest : 10.99 11.14 11.71 11.90 11.99, highest: 17.23 17.38 17.41 17.57 17.61

BlendSpeed

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	60.04	58.75	59.08	59.41	59.99	60.63	61.15	61.33

lowest : 57.93 57.96 58.65 58.65 58.70, highest: 61.34 61.46 61.48 62.11 62.27

CompUnit

n	missing	unique
90	0	2

COMPRESS2 (51, 57%), COMPRESS1 (39, 43%)

CompForce

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	25.01	24.48	24.62	24.77	24.99	25.24	25.46	25.51

lowest : 24.07 24.24 24.29 24.43 24.44, highest: 25.52 25.56 25.60 25.64 25.81

CoatingSup

n	missing	unique
90	0	3

Mac (29, 32%), Down (33, 37%), Coat (28, 31%)

CoatingVis

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	99.32	92.83	94.00	96.41	98.79	102.16	105.55	108.04

lowest : 86.02 87.42 88.57 90.70 92.52, highest: 108.29 108.94 109.21 109.36 110.13

InletTemp

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	106.9	105.4	105.6	106.1	106.8	107.7	108.3	108.5

lowest : 104.6 105.0 105.1 105.2 105.3, highest: 108.6 108.6 108.6 108.8 108.8

ExhaustTemp

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	69.94	68.30	68.66	69.27	69.94	70.59	71.31	71.55

lowest : 67.61 67.77 68.19 68.25 68.26, highest: 71.58 71.87 71.99 72.02 72.95

SprayRate

n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	69.94	68.30	68.66	69.27	69.94	70.59	71.31	71.55

90	0	90	399.7	391.0	392.6	396.0	399.3	403.6	407.1	407.6
lowest : 388.4 388.6 389.3 389.4 390.9, highest: 407.7 407.7 408.5 411.8 418.5										

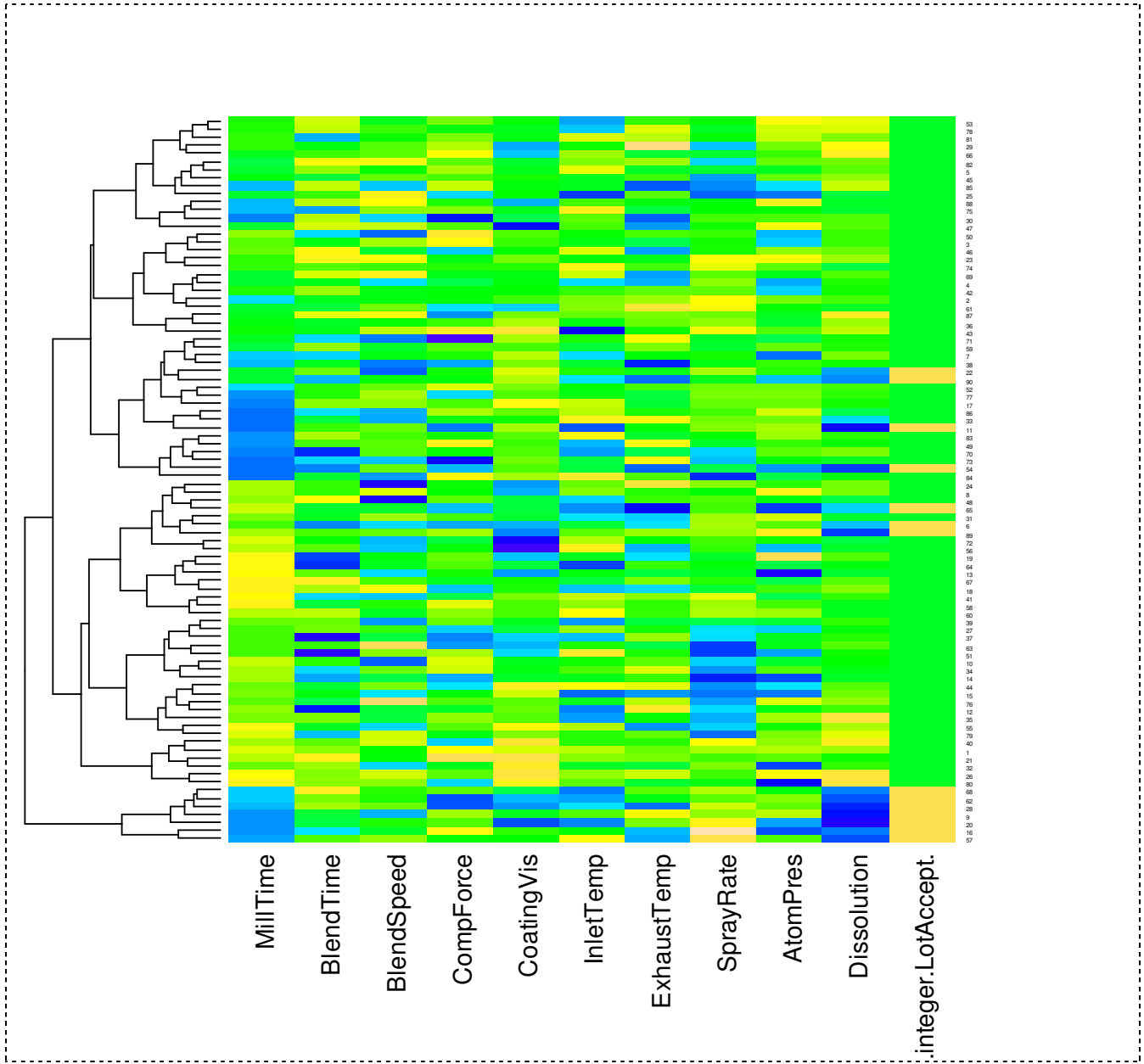
AtomPres										
n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	90	59.98	58.19	58.68	59.26	60.13	60.62	61.23	61.48
lowest : 57.46 57.68 58.04 58.12 58.18, highest: 61.51 61.65 61.72 61.74 62.35										

Dissolution										
n	missing	unique	Mean	.05	.10	.25	.50	.75	.90	.95
90	0	89	72.86	66.25	67.88	71.16	73.05	74.97	77.39	78.99
lowest : 64.15 64.55 65.11 65.67 66.17, highest: 79.07 79.26 80.03 80.13 80.23										

LotAccept										
n	missing	unique								
90	0	2								
Accept (76, 84%), Reject (14, 16%)										

Visual Profile of Data Set

Heatmap Display of Data Frame Data For Numeric variables



Histograms and Frequency Displays of Data Frame Data

