

MISSING DATA ANALYSIS



Identify missing data in the CHOLESTROL case study and provide statistical measures with

MISSING DATA ANALYSIS

APPROACH



- Identified missing data on each parameter
- Measure of Central Tendency and Descriptive statistics are calculated for each parameter.
- Mapped significant data with analyzed statistical values for each parameter.
- Core/raw statistical data should not impacted after mapping data.

MISSING DATA ANALYSIS

TOOLS USAGE

- Used Xcel, R and SPSS tools.
- Data verified on before and after data maps



MISSING DATA ANALYSIS

MISSING PARAMETER - AGE



Before missing data		After adding MCT data	
Age		Age	
Mean	47.08368	Mean	46.9986
Standard Error	0.338052	Standard Error	0.337939
Median	47	Median	47
Mode	47	Mode	47
Standard Deviation	12.74773	Standard Deviation	12.77482
Sample Variance	162.5046	Sample Variance	163.1961
Kurtosis	-0.45916	Kurtosis	-0.46966
Skewness	0.211519	Skewness	0.216408
Range	61	Range	61
Minimum	20	Minimum	20
Maximum	81	Maximum	81
Sum	66953	Sum	67161
Count	1422	Count	1429

MISSING DATA ANALYSIS

MISSING PARAMETER - WEIGHT



Before missing data		After MCT valued updated	
Weight		Weight	
Mean	82.65992	Mean	82.655
Standard Error	0.30995	Standard Error	0.308452
Median	81.65	Median	81.7
Mode	83.5	Mode	83.5
Standard Deviation	11.68804	Standard Deviation	11.66014
Sample Variance	136.6103	Sample Variance	135.9588
Kurtosis	-0.43083	Kurtosis	-0.41844
Skewness	0.239534	Skewness	0.241308
Range	60.6	Range	60.6
Minimum	54.8	Minimum	54.8
Maximum	115.4	Maximum	115.4
Sum	117542.4	Sum	118114
Count	1422	Count	1429

MISSING DATA ANALYSIS

MISSING PARAMETER - HEIGHT



Before – Missing data		After – Adding MCT data	
Height		Height	
Mean	180.8177	Mean	180.8055
Standard Error	0.187727	Standard Error	0.186881
Median	180.9	Median	180.8
Mode	174.5	Mode	174.5
Standard Deviation	7.076598	Standard Deviation	7.064484
Sample Variance	50.07823	Sample Variance	49.90693
Kurtosis	-0.42943	Kurtosis	-0.42098
Skewness	0.040588	Skewness	0.042669
Range	39.9	Range	39.9
Minimum	160.6	Minimum	160.6
Maximum	200.5	Maximum	200.5
Sum	256942	Sum	258371
Count	1421	Count	1429

MISSING DATA ANALYSIS



INSIGHTS

- Missing data is filled with Central tendency values like MEAN, MEDIAN and MODE.
 - Compared MCT & DS values after data mapping and observe there is no impact of raw data measurements.
-