

# Mission Hospital Case Study

---

IDS-506

Health Information Management and Analytics

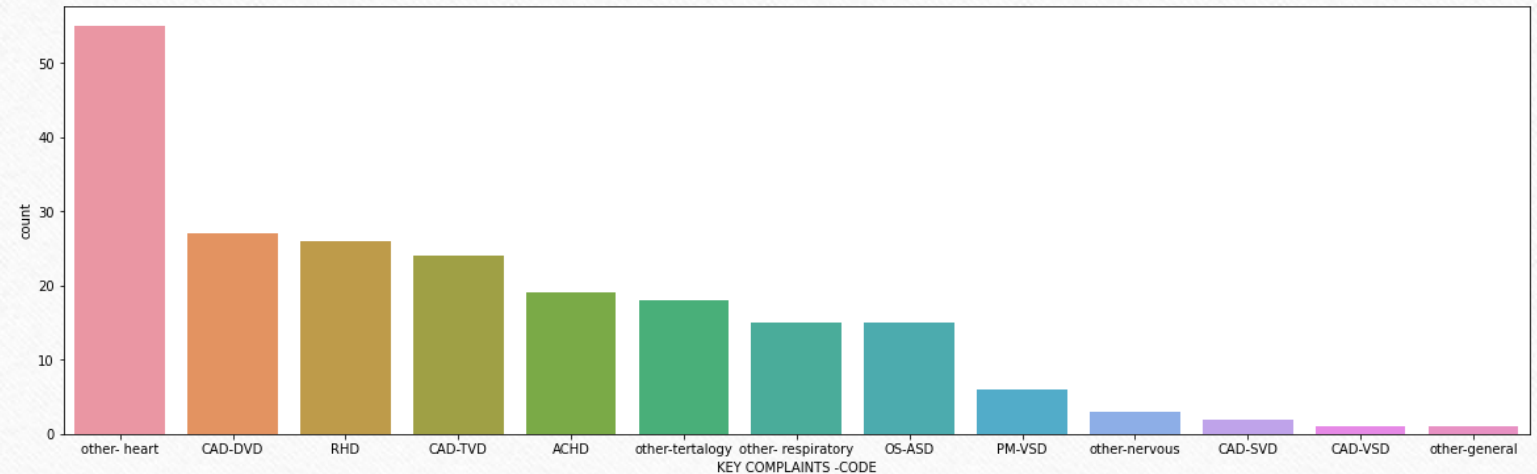
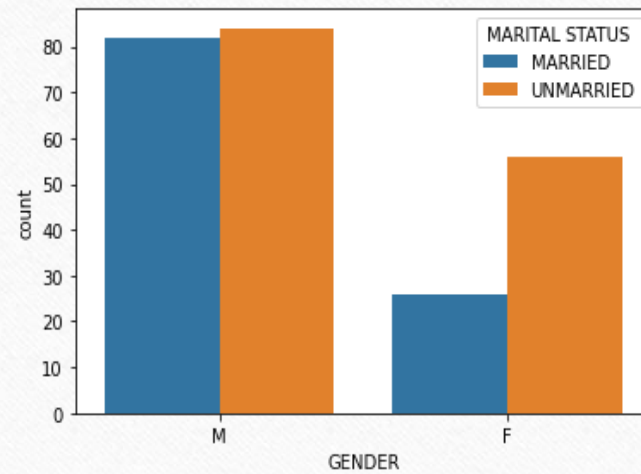
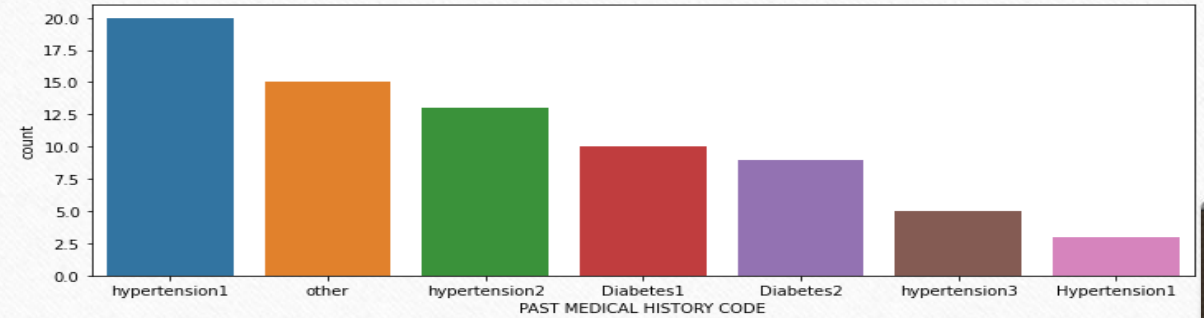
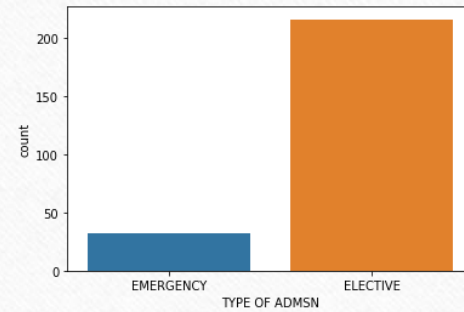
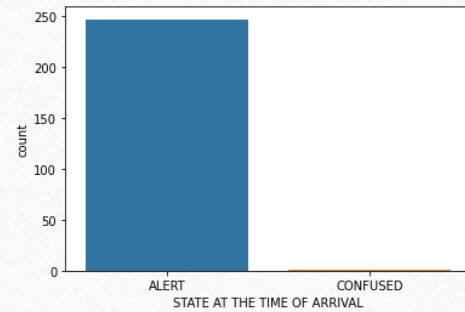
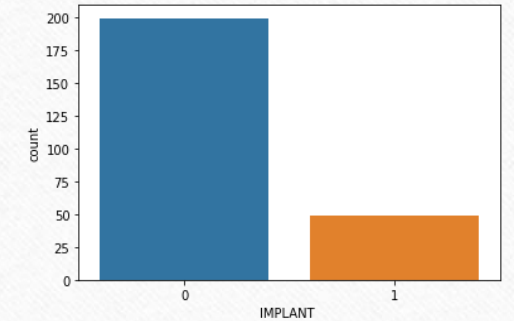
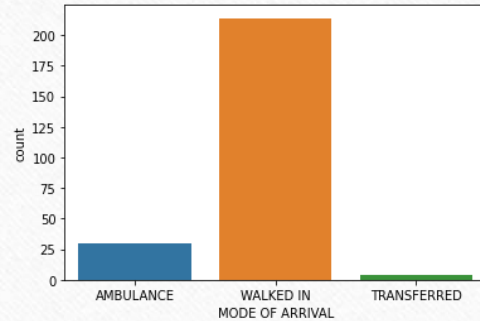
Vivek Kumar

# Data Transformation

---

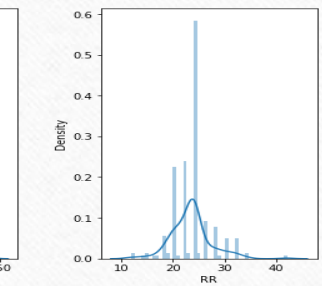
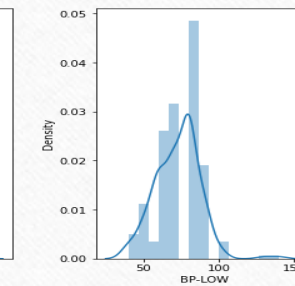
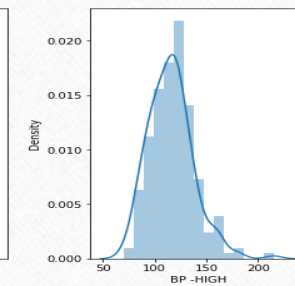
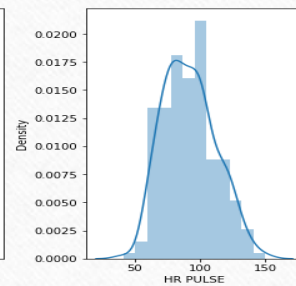
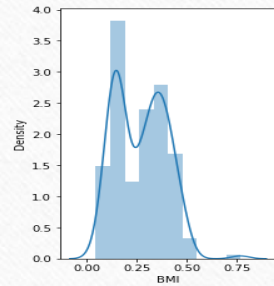
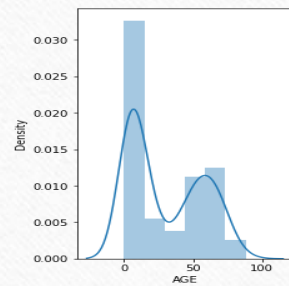
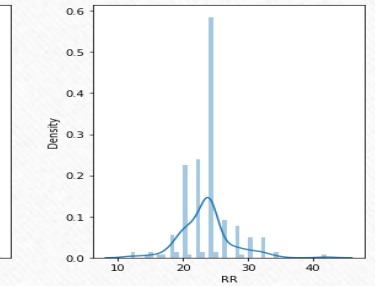
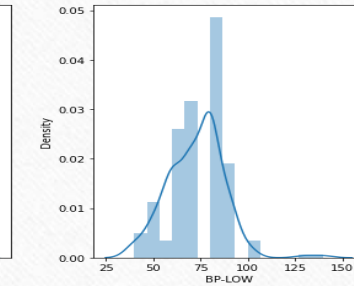
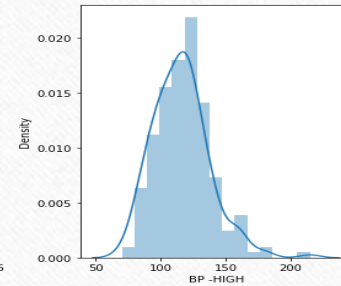
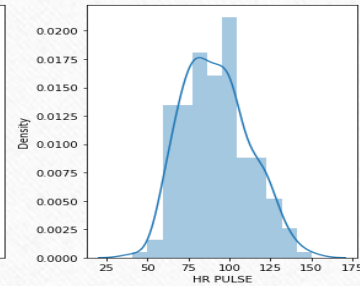
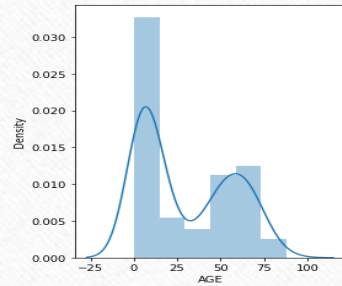
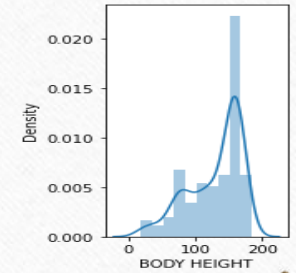
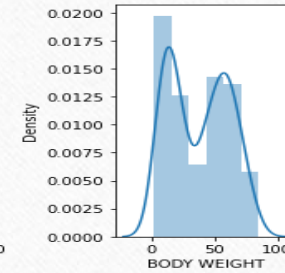
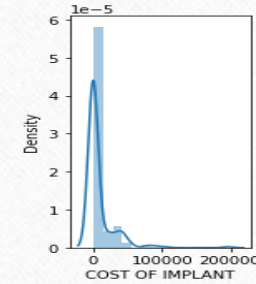
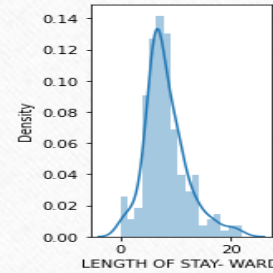
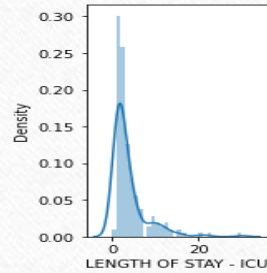
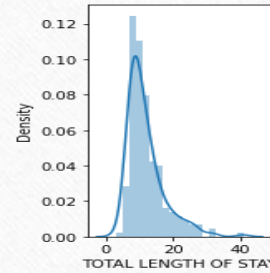
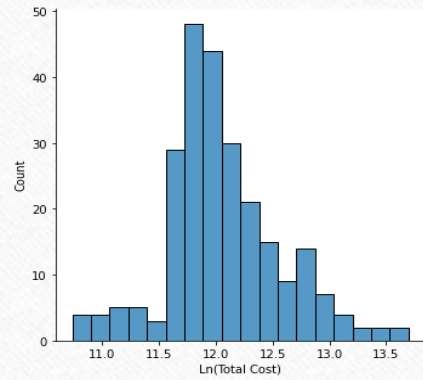
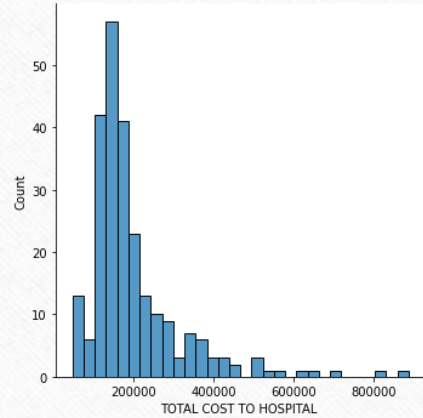
- The modified dataset is encoded for 'GENDER', 'MARITAL STATUS', 'KEY COMPLAINTS -CODE', 'PAST MEDICAL HISTORY CODE', 'MODE OF ARRIVAL', 'STATE AT THE TIME OF ARRIVAL', 'TYPE OF ADMSN', 'IMPLANT USED (Y/N)'
- Replaced missing values for 'BP -HIGH' with 120 and for 'BP-LOW' with 80 as all these missing values were for patients aged less than 16 and probably that's the reason their blood pressure was not measured
- Dropped records with missing values in 'HB', 'UREA', 'CREATININE'
- The final dataset is of shape (213, 41)

# Data Visualization





# Data Visualization



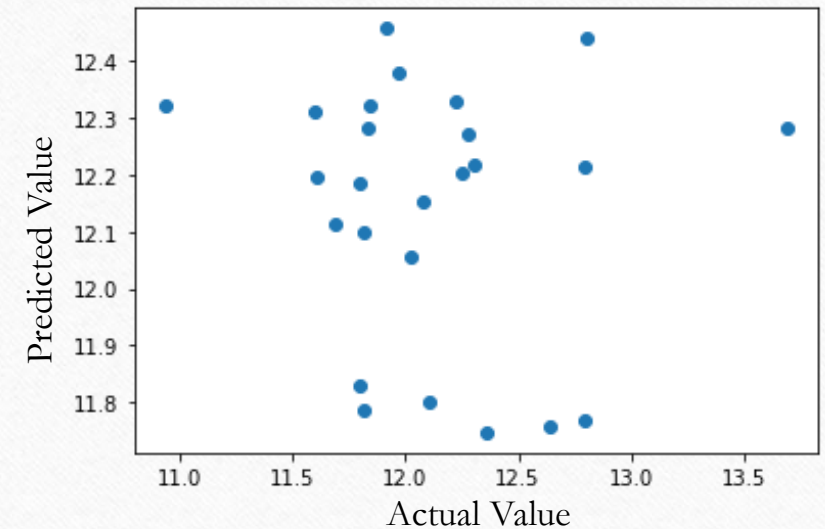
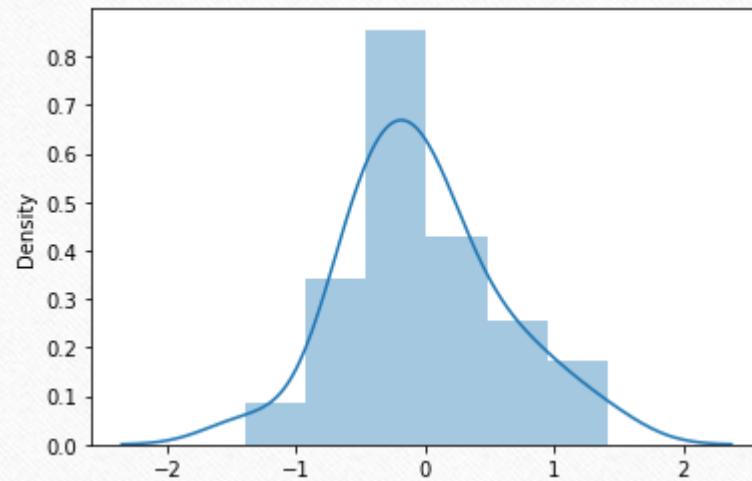
# Linear Regression

## Body Weight and Ln(Total Cost)

Intercept	Coefficient
11.70	0.0097

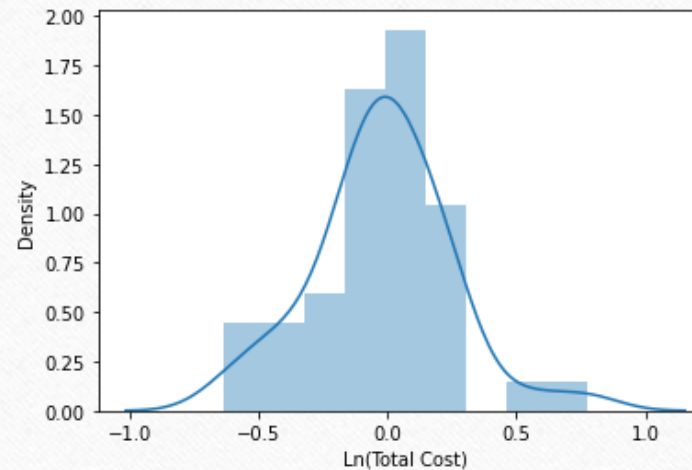
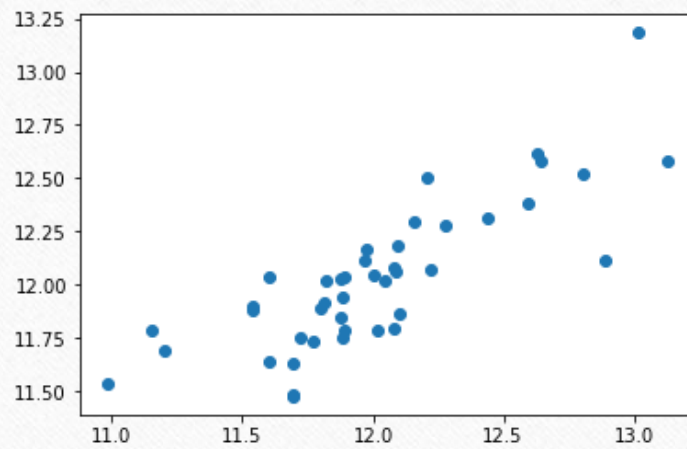
MAE	MSE	RMSE
0.45	0.36	0.596

<b>a</b>	11.70145898	
<b>b</b>	0.00967142	
<b>Weights</b>	<b>Ln(Total Cost)</b>	<b>Total Cost</b>
51	12.1947	197739
50	12.1850	195835
	<b>Difference</b>	<b>1903</b>



# Linear Regression (Baseline Model)

MAE	MSE	RMSE
0.196	0.071	0.267

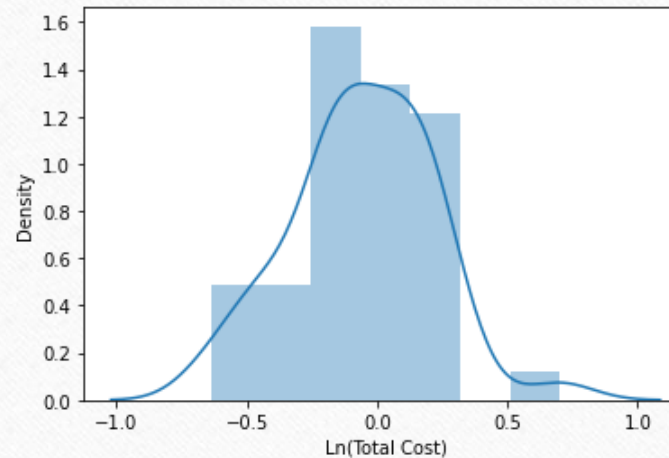
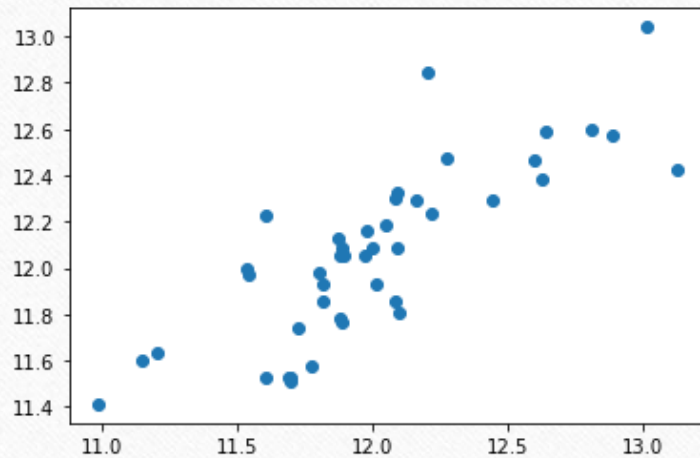


Attributes	Coefficients
other-teratology	0.2717150
PM-VSD	0.2707200
Diabetes2	0.2218983
IMPLANT	0.1991038
CAD-TVD	0.1654446
CAD-VSD	0.1380456
OS-ASD	0.1275094
UNMARRIED	0.0994622
CREATININE	0.0974861
other- heart	0.0974272
CAD-DVD	0.0904248
hypertension3	0.0783337
ELECTIVE	0.0695154
TOTAL LENGTH OF STAY	0.0691565
other	0.0587460
RHD	0.0540907
hypertension1	0.0467927
other-nervous	0.0162006
LENGTH OF STAY - ICU	0.0044479
AGE	0.0021740
BP -HIGH	0.0016353
HR PULSE	0.0016311
BODY WEIGHT	0.0016183
HB	0.0011759
BODY HEIGHT	0.0010661
COST OF IMPLANT	0.0000047
ALERT	(0.0000000)
UREA	(0.0004249)
BP-LOW	(0.0008112)
RR	(0.0062035)
hypertension2	(0.0093177)
AMBULANCE	(0.0137206)
LENGTH OF STAY- WARD	(0.0299812)
MALE	(0.0314364)
Diabetes1	(0.0624179)
other- respiratory	(0.0926849)
ACHD	(0.0998178)
TRANSFERRED	(0.1779651)
CAD-SVD	(0.4453927)
other-general	(0.9948150)



## Linear Regression (Reduced Variables)

	MAE	MSE	RMSE
Baseline	0.196	0.071	0.267
Reduced Variables	0.217	0.075	0.274



Attributes	Coefficients
other-teratology	0.352031
PM-VSD	0.274478
Diabetes2	0.302664
IMPLANT	0.490152
CAD-TVD	0.323761
CAD-VSD	0.098553
OS-ASD	0.17382
UNMARRIED	-0.078599
CREATININE	0.132017
other-heart	0.147751
CAD-DVD	0.287633
hypertension3	0.265867
ELECTIVE	0.110784
TOTAL LENGTH OF STAY	0.053637
other	-0.115763
RHD	0.087062