Comprehensive EDA Report (Compact A4)
Dataset: C:\Users\DELL\Documents\DS\IPE Global\ML\lbw\data\AI_Cohort_LabeledD
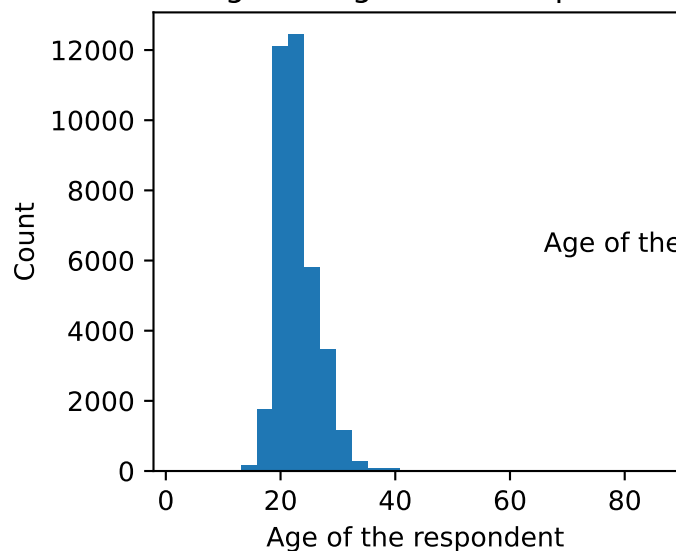Target: LBW
Observations: 37353
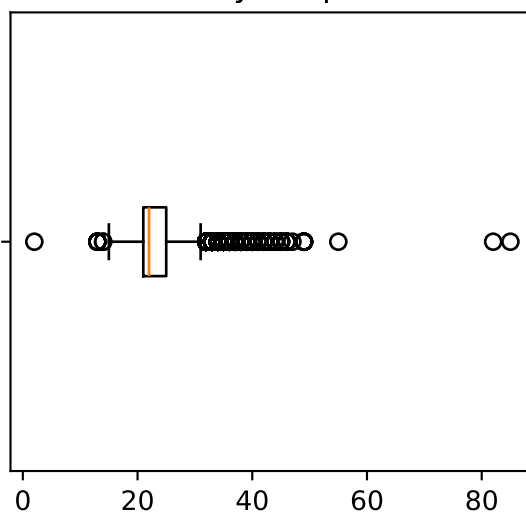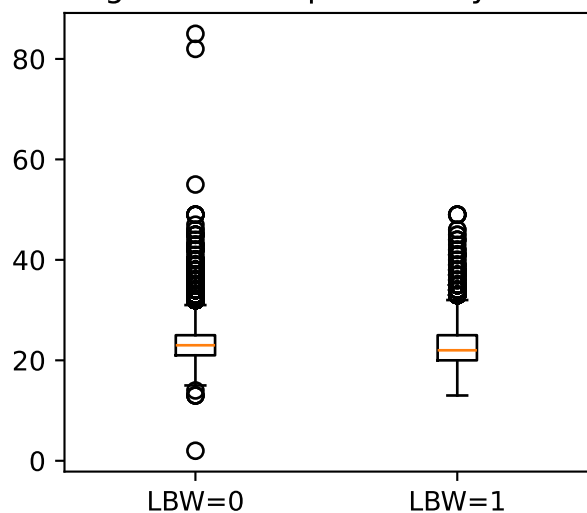Variables analyzed: 59

EDA – Age of the respondent

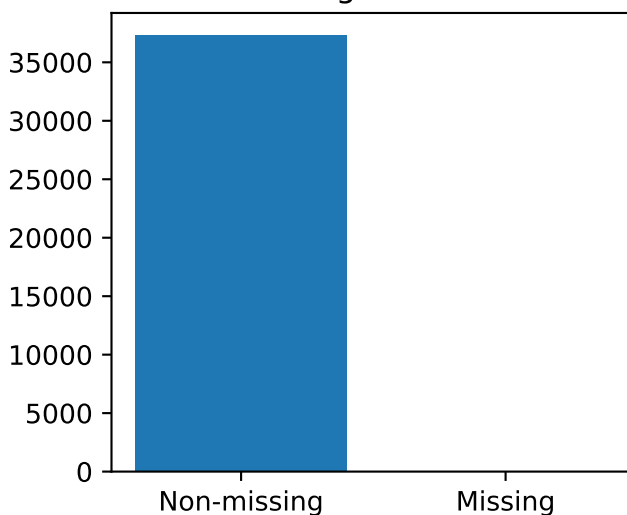## Histogram: Age of the respondent

Tukey Boxplot

Age of the respondent by LBW

Missing: 0/37353

Hypothesis Tests & Metrics for: Age of the respondent (numeric)
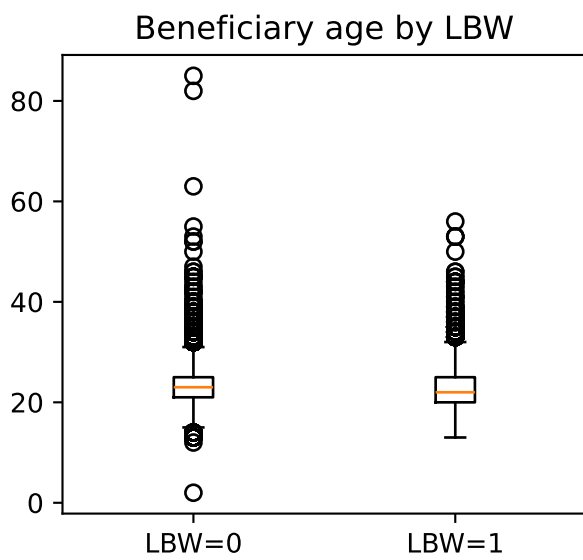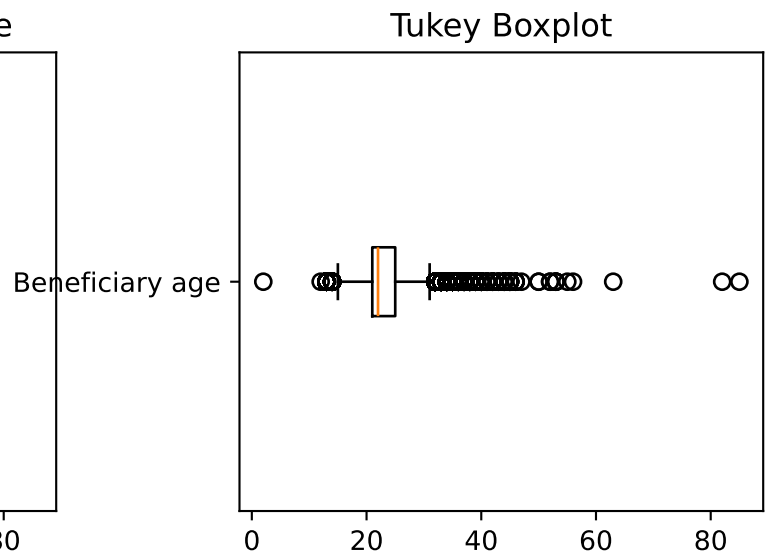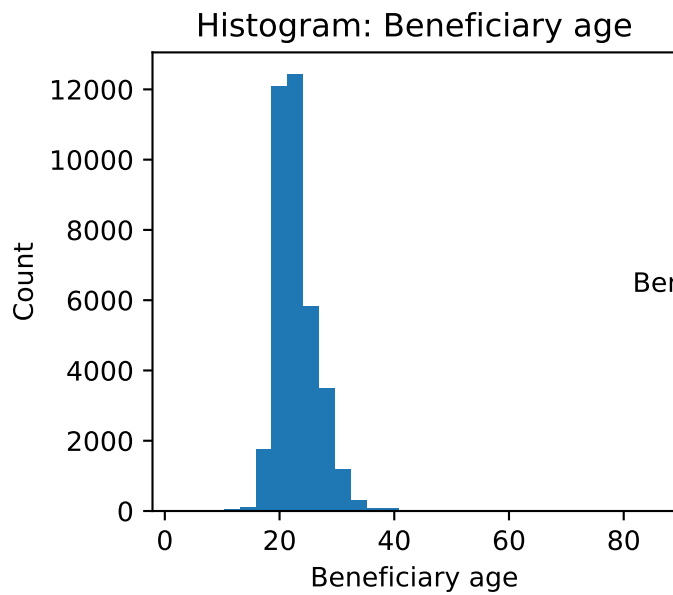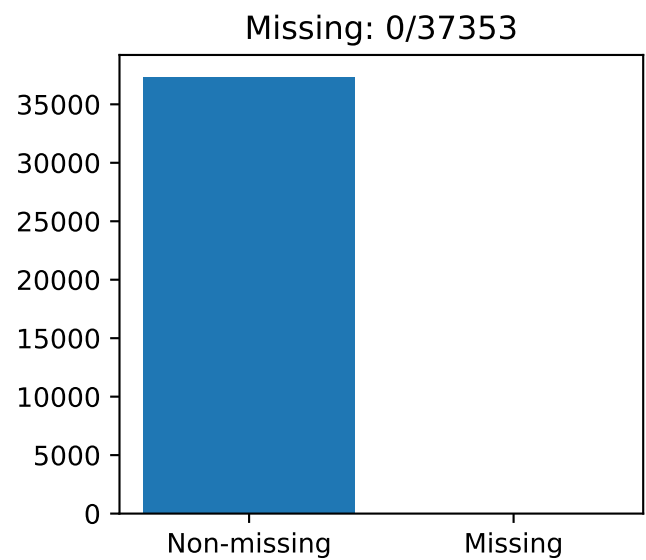Missing: 0.00% | Unique: 40
Welch t-test: t=13.1, p=4.1e-39
Point–Biserial r: r=-0.067, p=1.82e-38
Cohen's d: -0.146 | KS: D=0.0697, p=1.04e-33
AUROC (using Age of the respondent as score): 0.454

EDA – Beneficiary age

Histogram: Beneficiary age
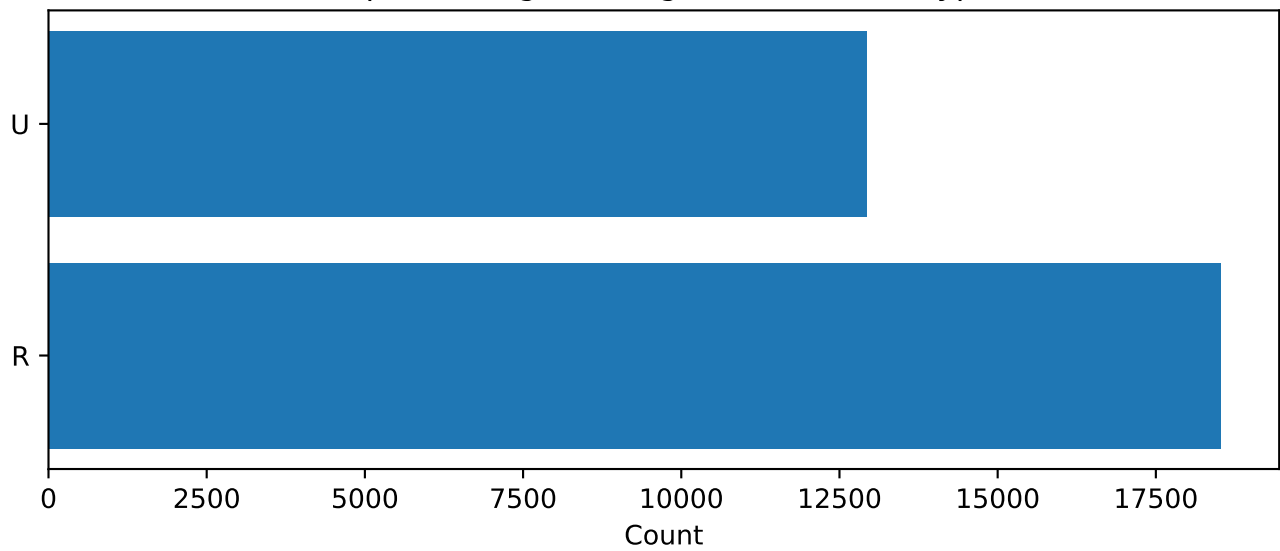

Tukey Boxplot
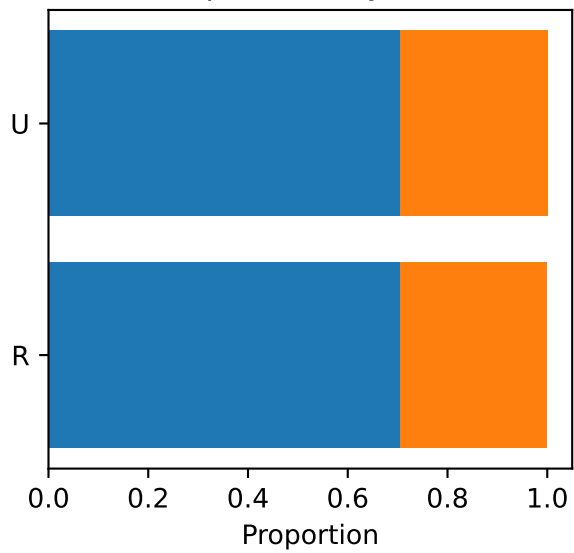

Beneficiary age by LBW

Missing: 0/37353

Hypothesis Tests & Metrics for: Beneficiary age (numeric)
Missing: 0.00% | Unique: 45
Welch t-test: t=13, p=1.86e-38
Point–Biserial r: r=-0.0666, p=5.93e-38
Cohen's d: -0.145 | KS: D=0.0695, p=1.58e-33
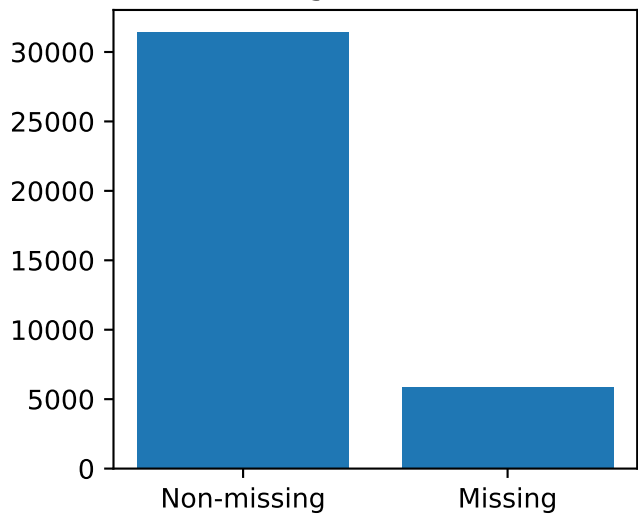AUROC (using Beneficiary age as score): 0.455

Top-15 categories: Aganwadi Center Type

Proportions by LBW

Missing: 5899/37353

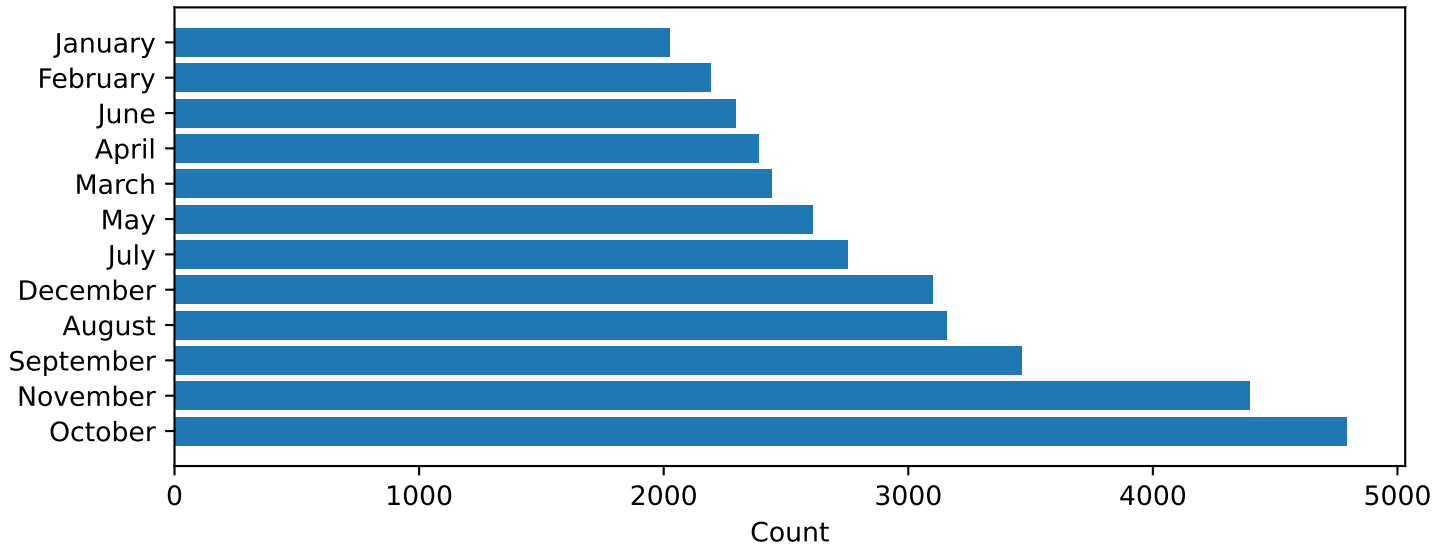Hypothesis Tests & Metrics for: Aganwadi Center Type (categorical)
Missing: 15.79% | Unique: 2
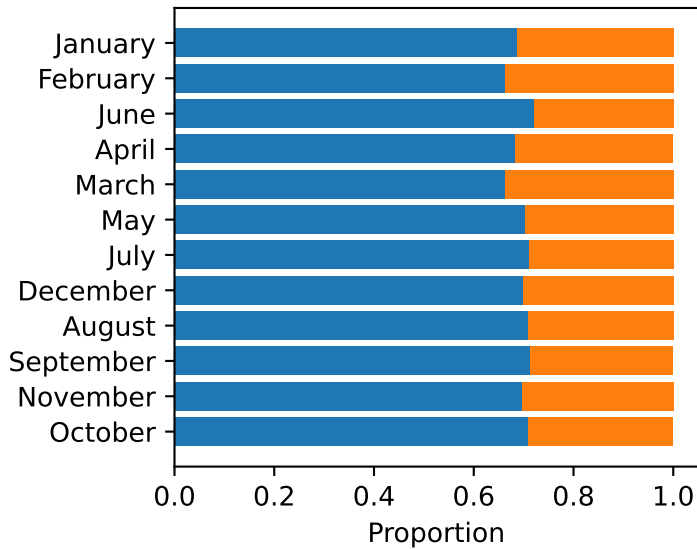Chi-square: chi2=0.032, dof=1, p=0.858
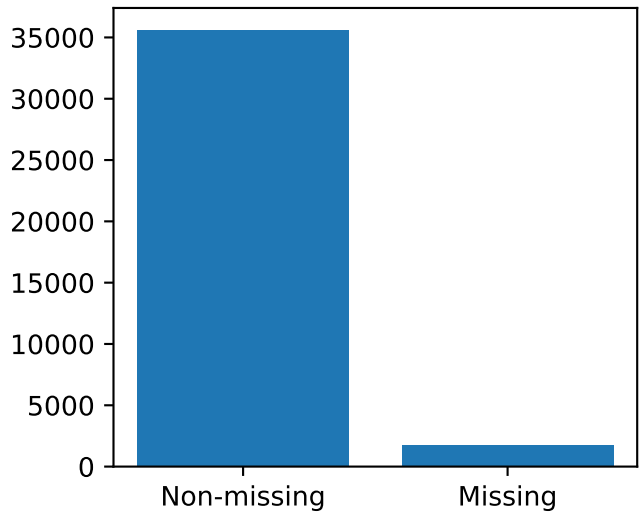Cramér's V (bias-corrected): 0.000

EDA – MonthConception

Top-15 categories: MonthConception

Proportions by LBW

Missing: 1731/37353

Hypothesis Tests & Metrics for: MonthConception (categorical)
Missing: 4.63% | Unique: 12
Chi-square: chi2=48.1, dof=11, p=1.37e-06
Cramér's V (bias-corrected): 0.032

EDA – education

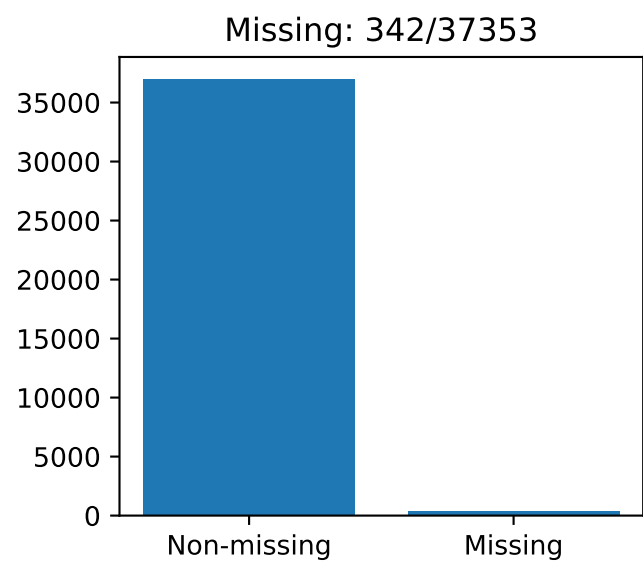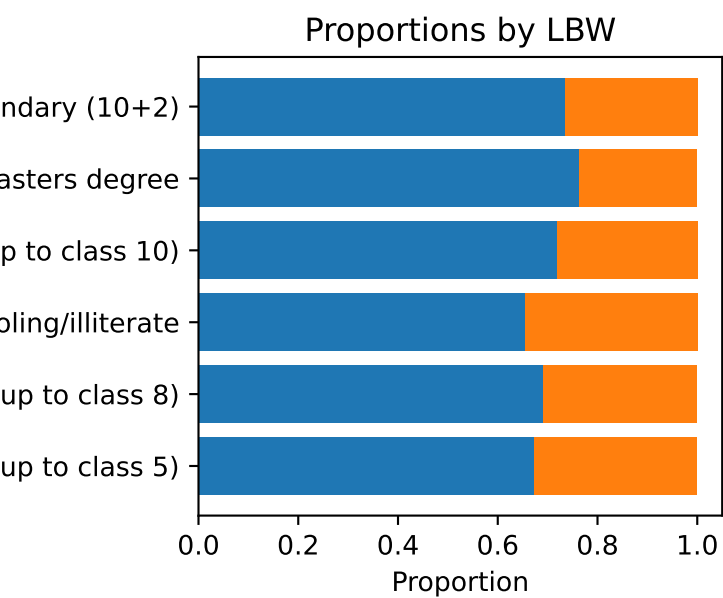## Top-15 categories: education

Proportions by LBW

Missing: 342/37353
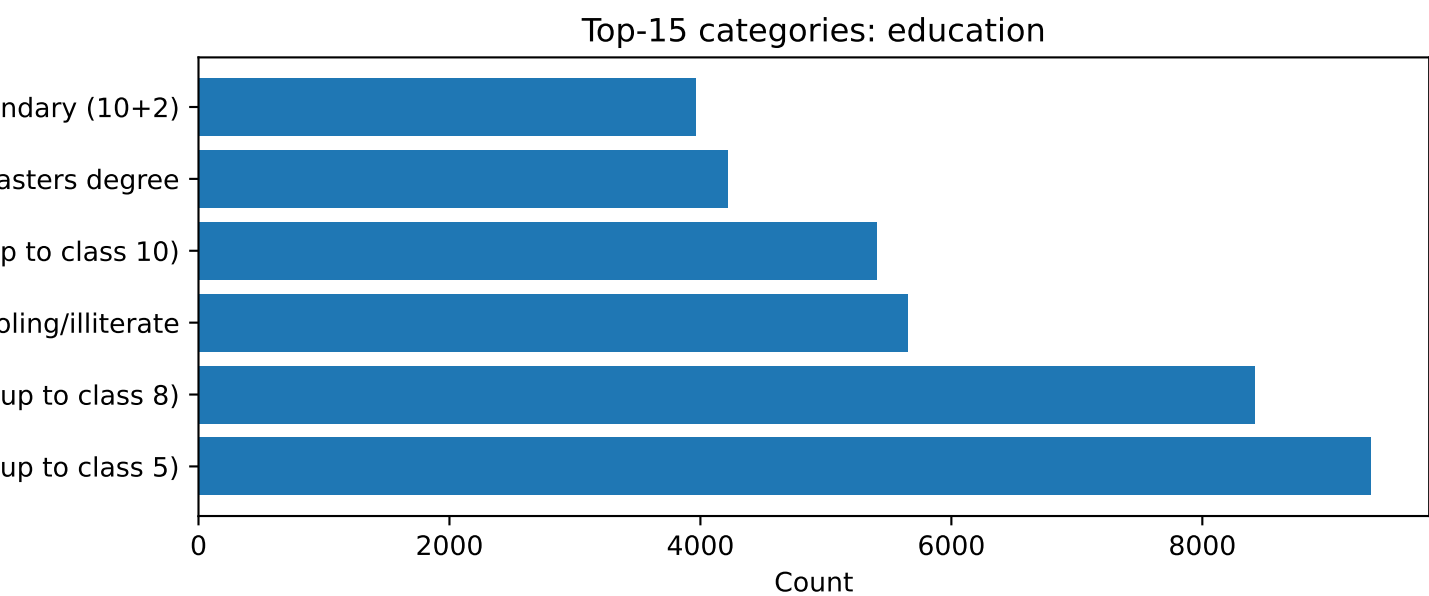

Hypothesis Tests & Metrics for: education (categorical)
Missing: 0.92% | Unique: 6
Chi-square: chi2=203, dof=5, p=6.59e-42
Cramér's V (bias-corrected): 0.073

EDA – Completed level of education

Top-15 categories: Completed level of education

Proportions by LBW

Missing: 0/37353
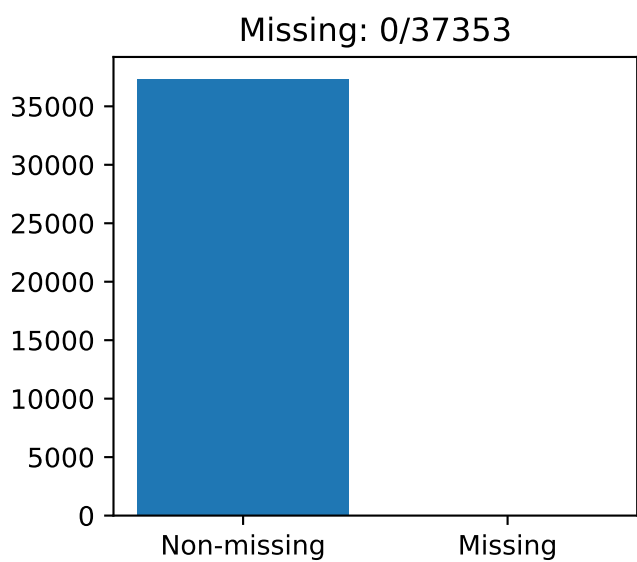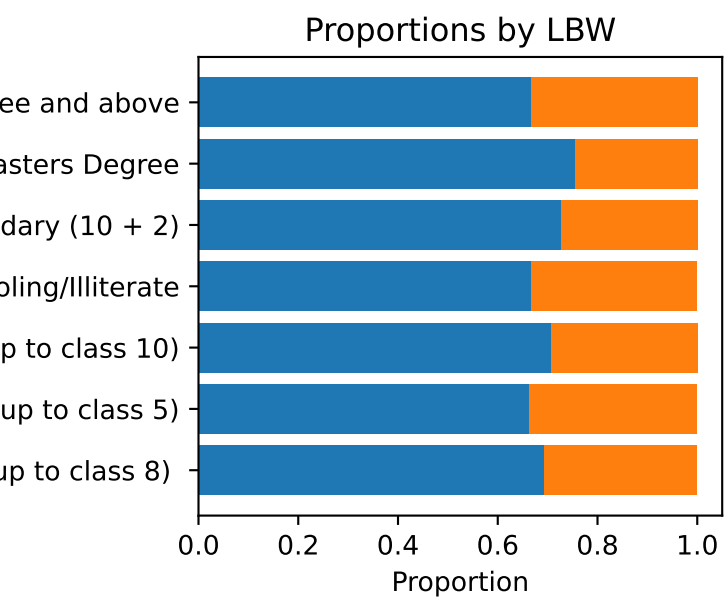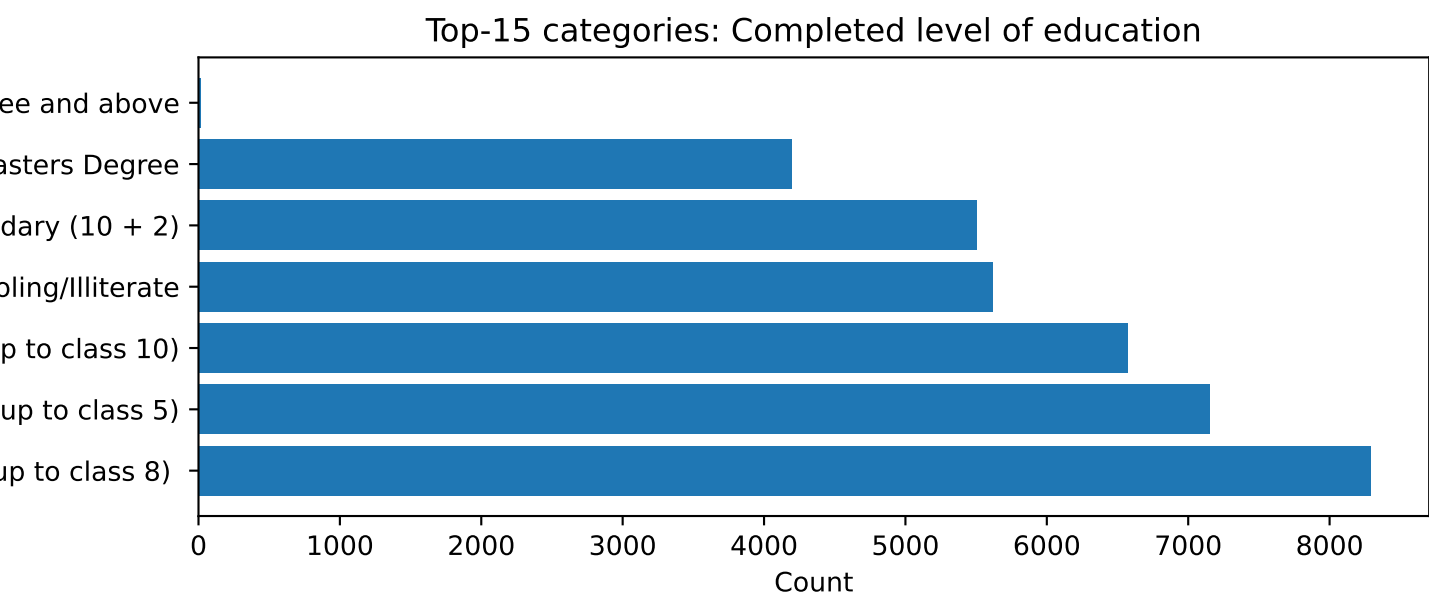
Hypothesis Tests & Metrics for: Completed level of education (categorical)
Missing: 0.00% | Unique: 7
Chi-square: chi2=152, dof=6, p=2.61e-30
Cramér's V (bias-corrected): 0.063

EDA – Number of members residing at respondent's house

Histogram: Number of members residing at respondent's house
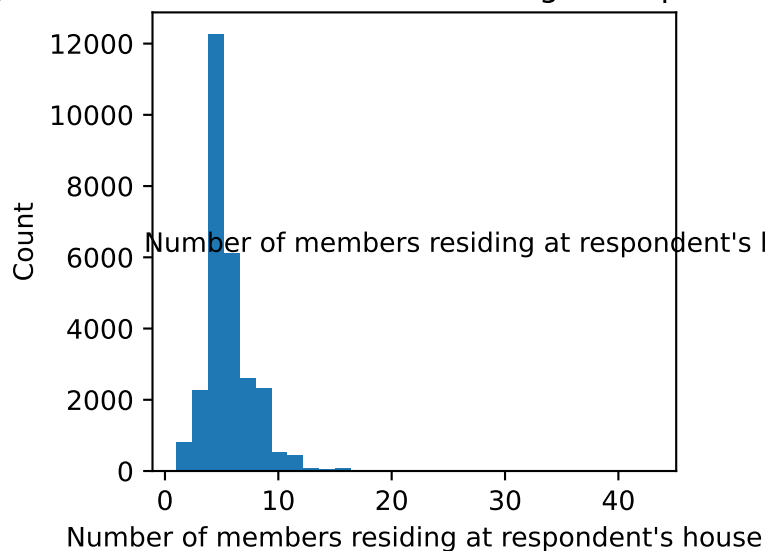
Tukey Boxplot

Number of members residing at respondent's house
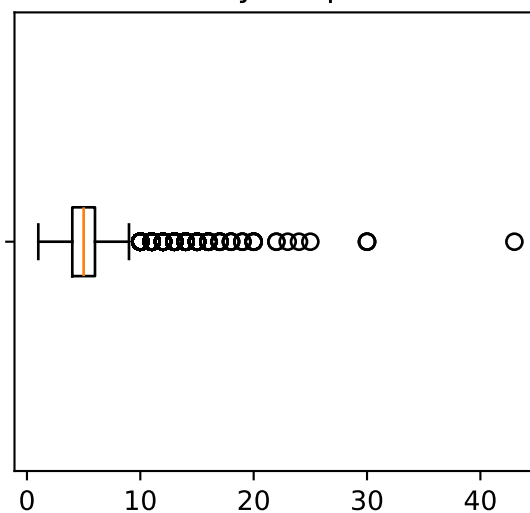
Number of members residing at respondent's house by LBW

Missing: 9865/37353
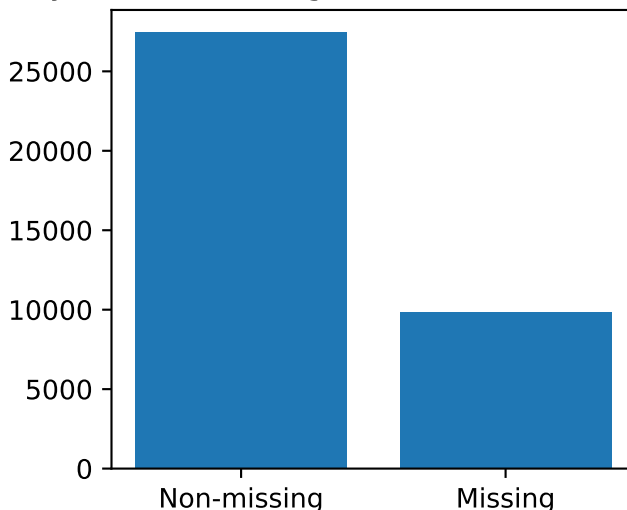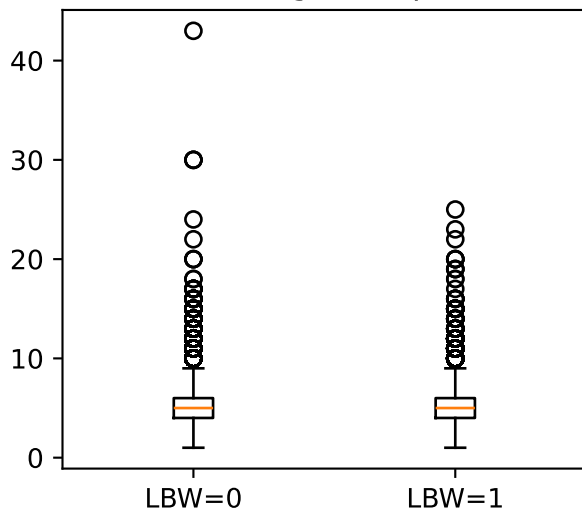
Hypothesis Tests & Metrics for: Number of members residing at respondent's house (numeric)
Missing: 26.41% | Unique: 26
Welch t-test: t=0.65, p=0.515
Point–Biserial r: r=-0.00393, p=0.515
Cohen's d: -0.00857 | KS: D=0.00723, p=0.921
AUROC (using Number of members residing at respondent's house as score): 0.497

EDA – Main Source of drinking water

Top-15 categories: Main Source of drinking water
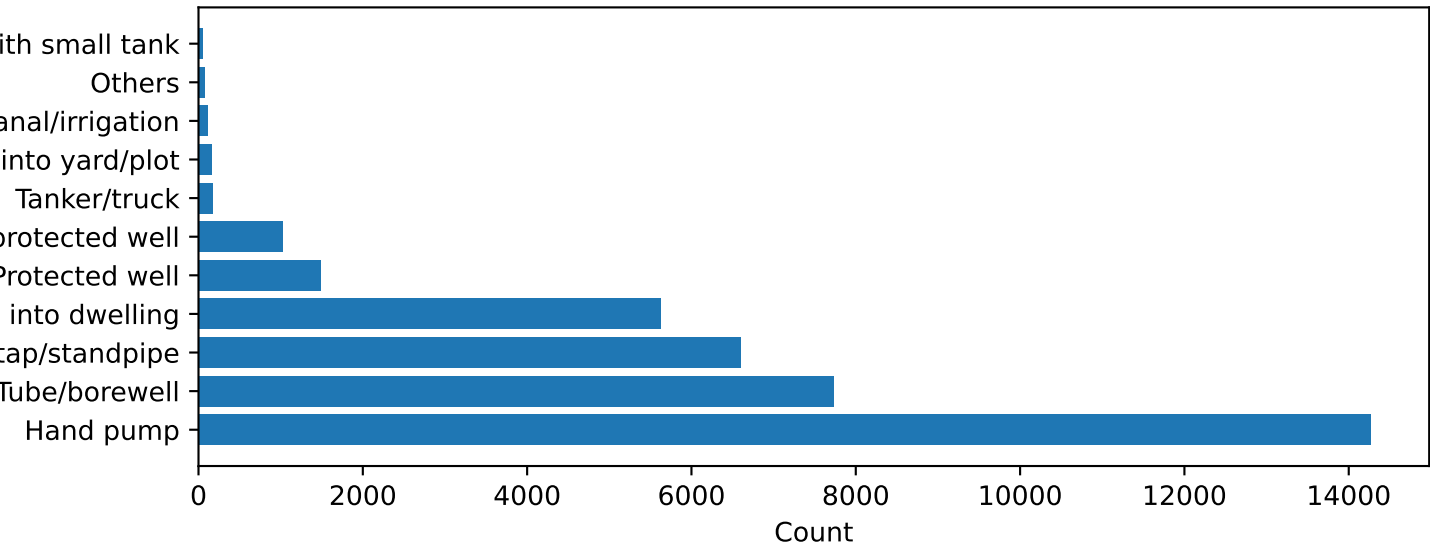
Proportions by LBW
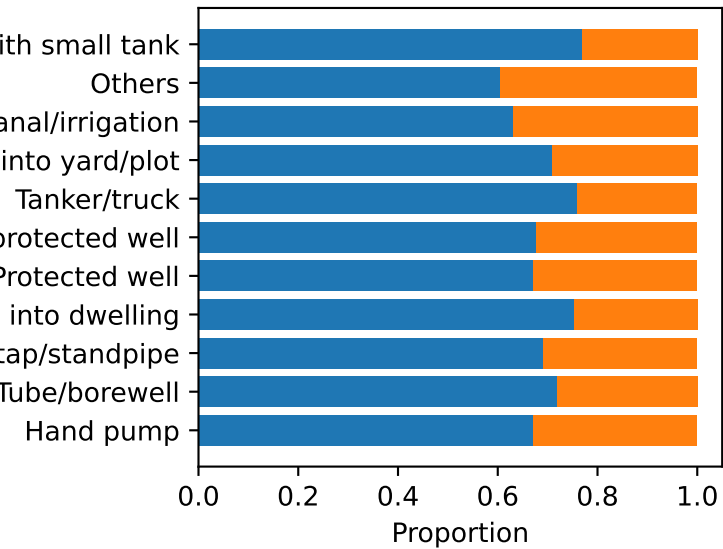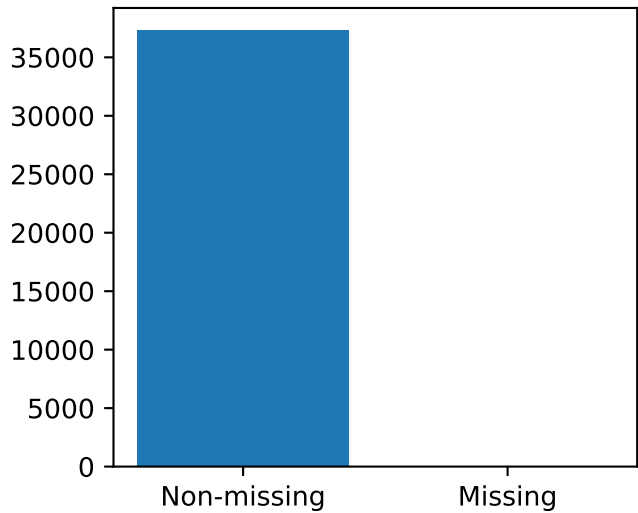
Missing: 0/37353

Hypothesis Tests & Metrics for: Main Source of drinking water (categorical)
Missing: 0.00% | Unique: 11
Chi-square: chi2=163, dof=10, p=7.54e-30
Cramér's V (bias-corrected): 0.064

EDA – toilet_type

Top-15 categories: toilet_type
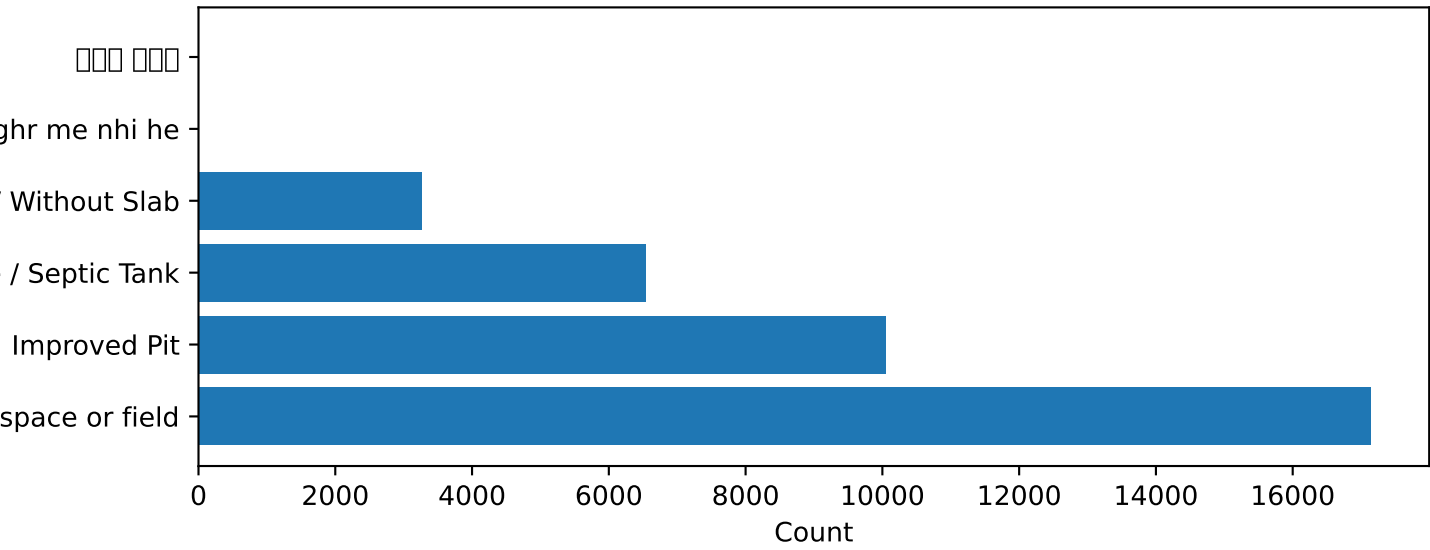
Proportions by LBW
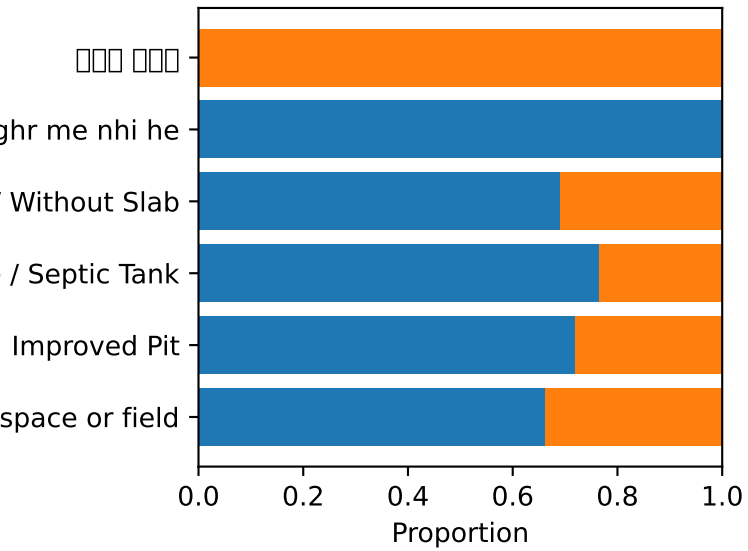
Missing: 342/37353
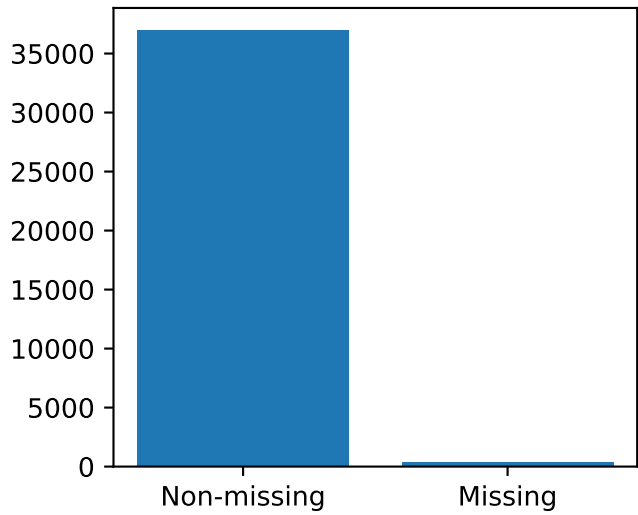
Hypothesis Tests & Metrics for: toilet_type (categorical)
Missing: 0.92% | Unique: 6
Chi-square: chi2=271, dof=5, p=1.4e-56
Cramér's V (bias-corrected): 0.085

EDA – food_group_n

## Histogram: food_group_n

Tukey Boxplot

food_group_n by LBW

Missing: 0/37353
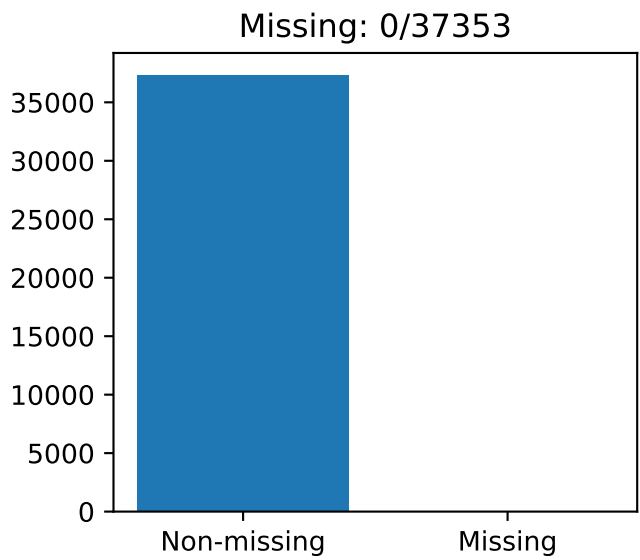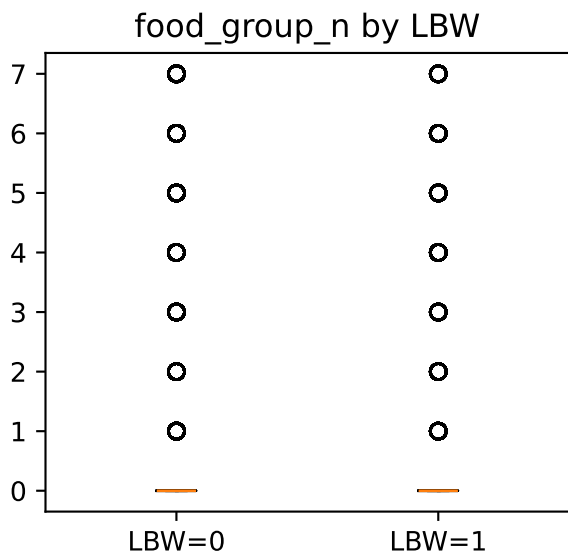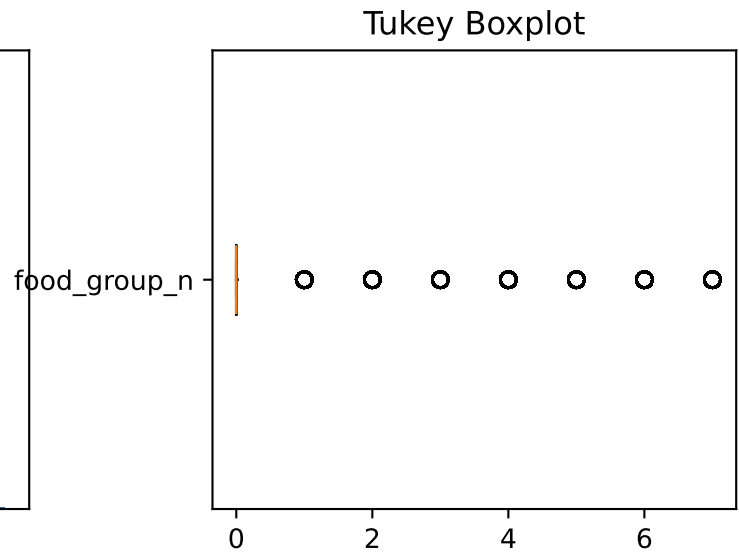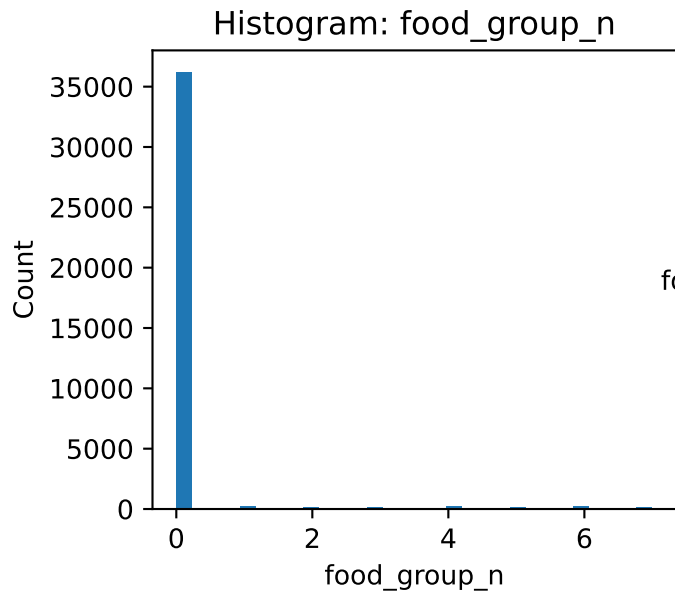
Hypothesis Tests & Metrics for: food_group_n (numeric)
Missing: 0.00% | Unique: 8
Welch t-test: t=1.89, p=0.0586
Point–Biserial r: r=-0.00949, p=0.0665
Cohen's d: -0.0207 | KS: D=0.0036, p=1
AUROC (using food_group_n as score): 0.499

Top-15 categories: Whether vegetables/fruits are grown by the household for own consumption

Proportions by LBW

Missing: 0/37353
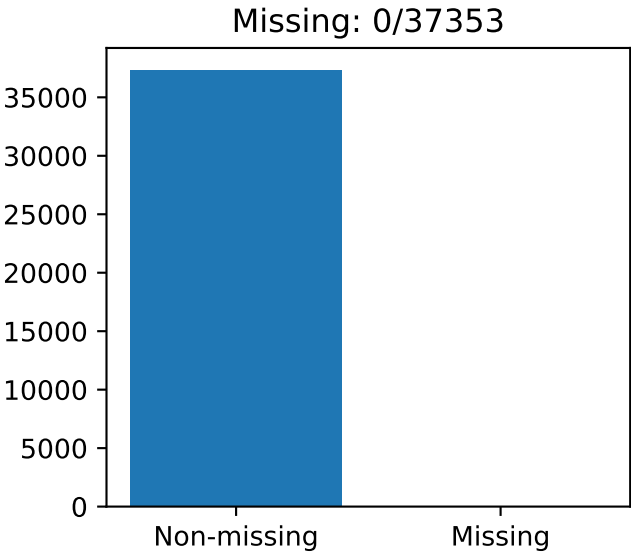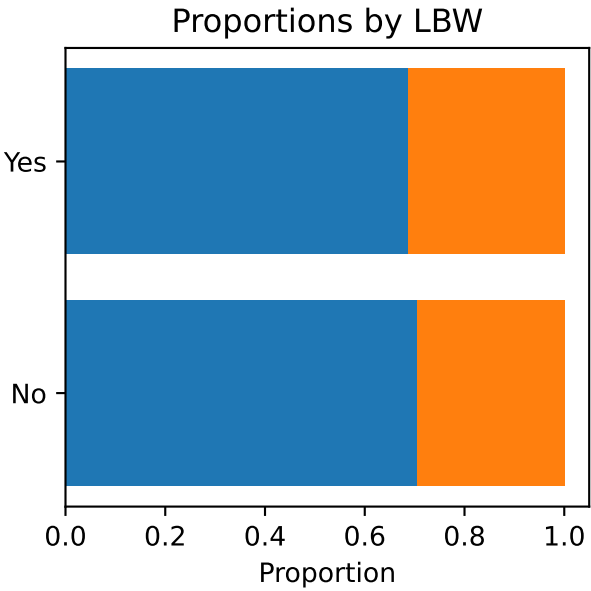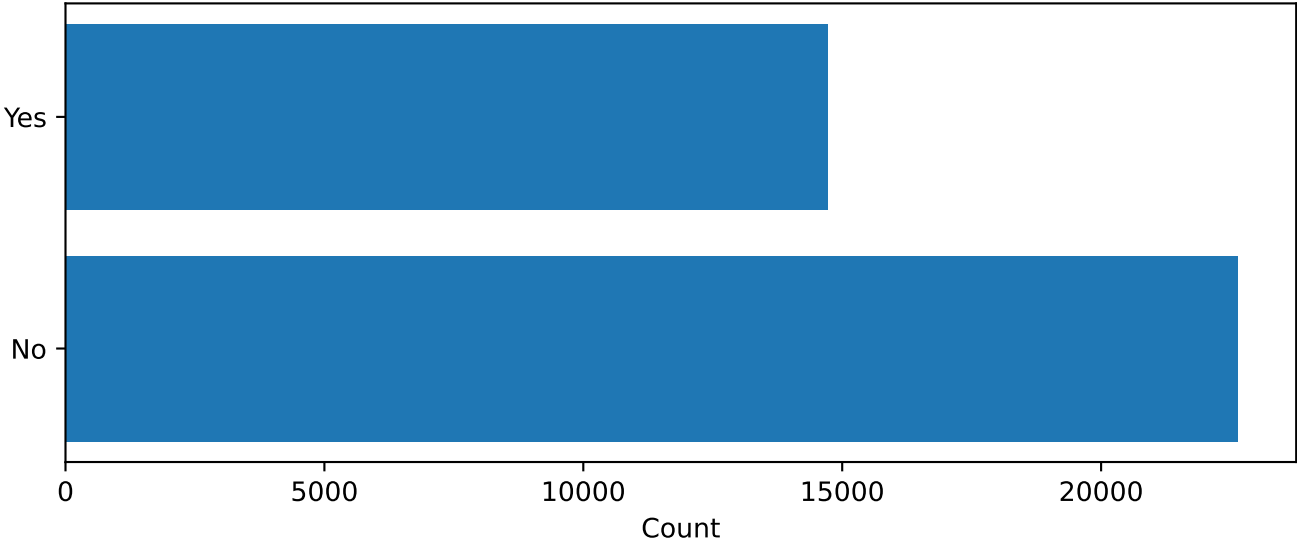
Hypothesis Tests & Metrics for: Whether vegetables/fruits are grown by the household for own consumption (cate
Missing: 0.00% | Unique: 2
Chi-square: chi2=13.6, dof=1, p=0.000228
Cramér's V (bias-corrected): 0.018

EDA – Household asset: Air Conditioner/ cooler

Top-15 categories: Household asset: Air Conditioner/ cooler
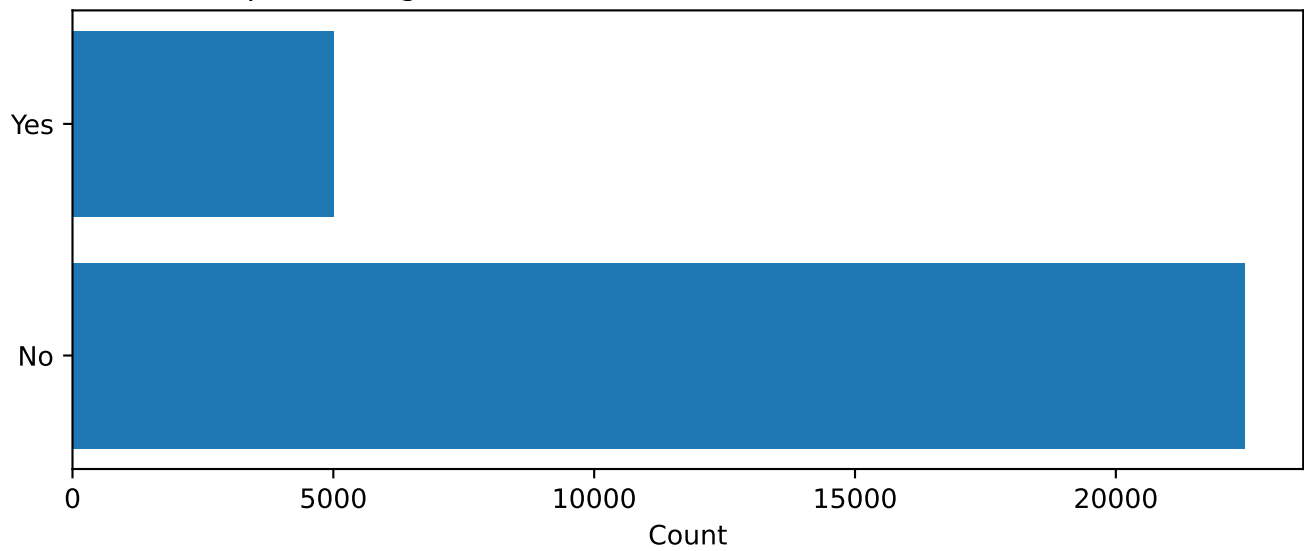
Proportions by LBW
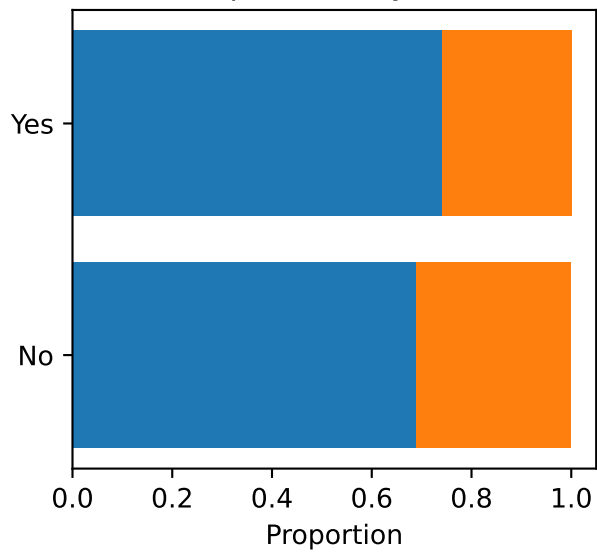
Missing: 9865/37353
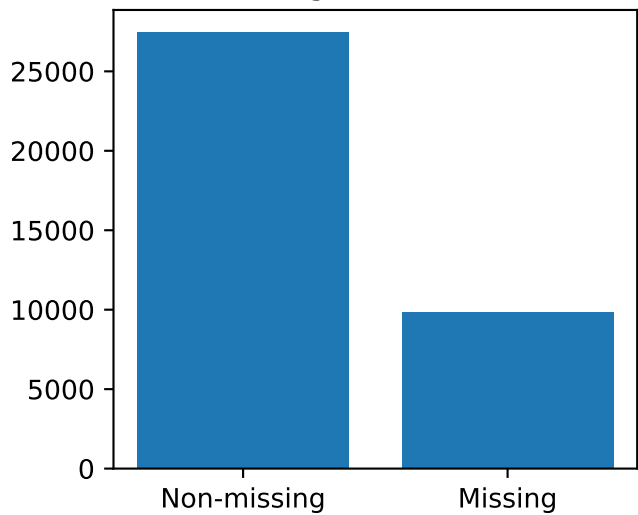
Hypothesis Tests & Metrics for: Household asset: Air Conditioner/ cooler (categorical)
Missing: 26.41% | Unique: 2
Chi-square: chi2=50.3, dof=1, p=1.33e-12
Cramér's V (bias-corrected): 0.042

EDA – Household asset: Washing Machine

Top-15 categories: Household asset: Washing Machine
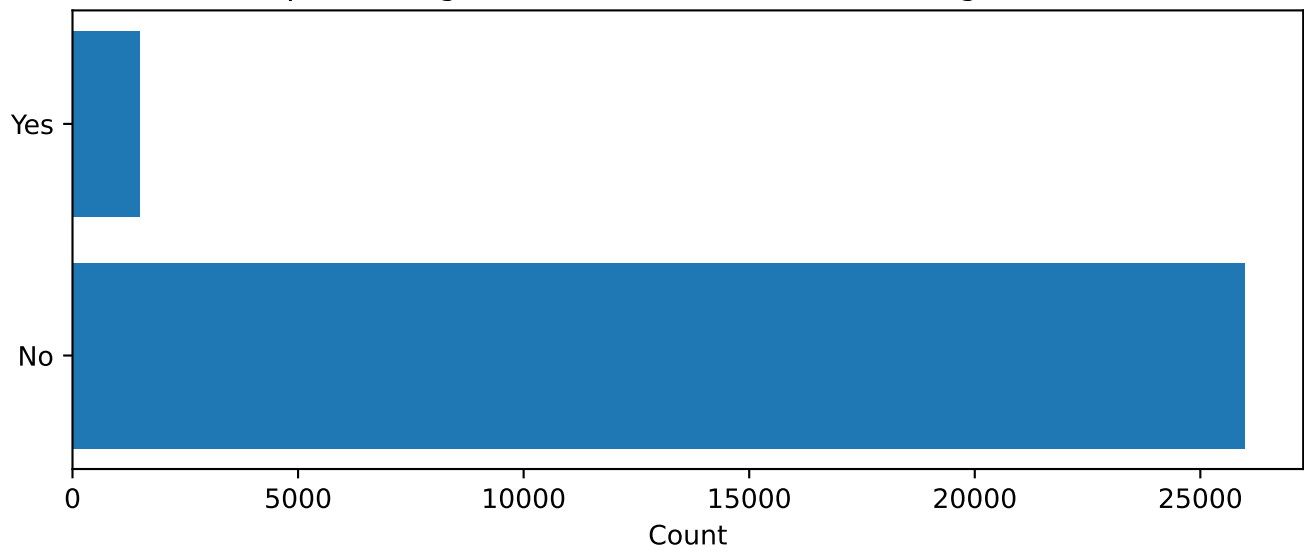
Proportions by LBW
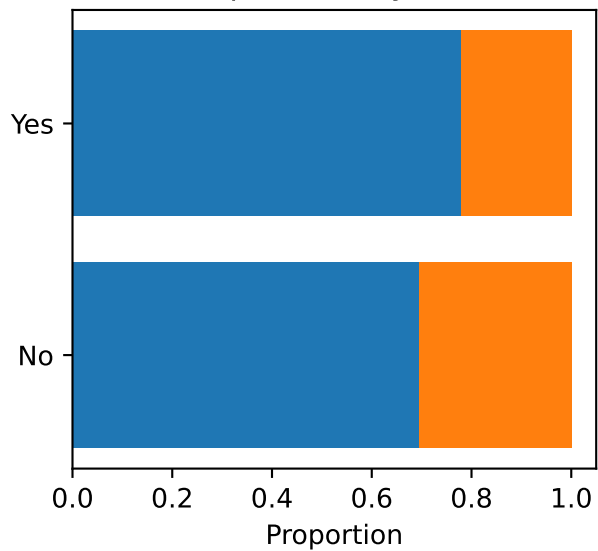
Missing: 9865/37353
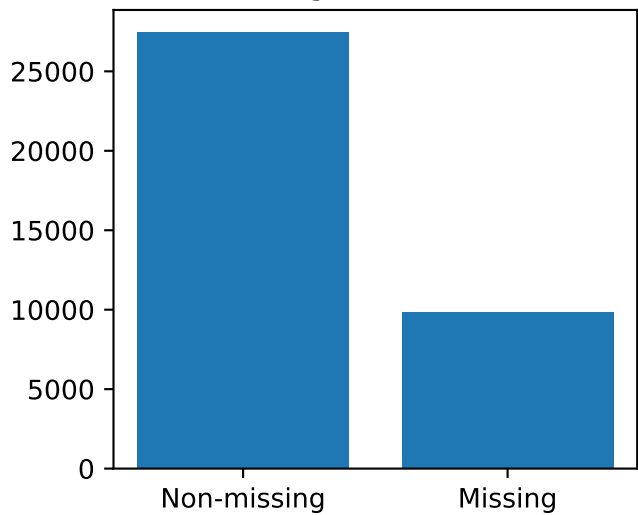
Hypothesis Tests & Metrics for: Household asset: Washing Machine (categorical)
Missing: 26.41% | Unique: 2
Chi-square: chi2=48.1, dof=1, p=4.05e-12
Cramér's V (bias-corrected): 0.041

EDA – Types of fuel used

Top-15 categories: Types of fuel used
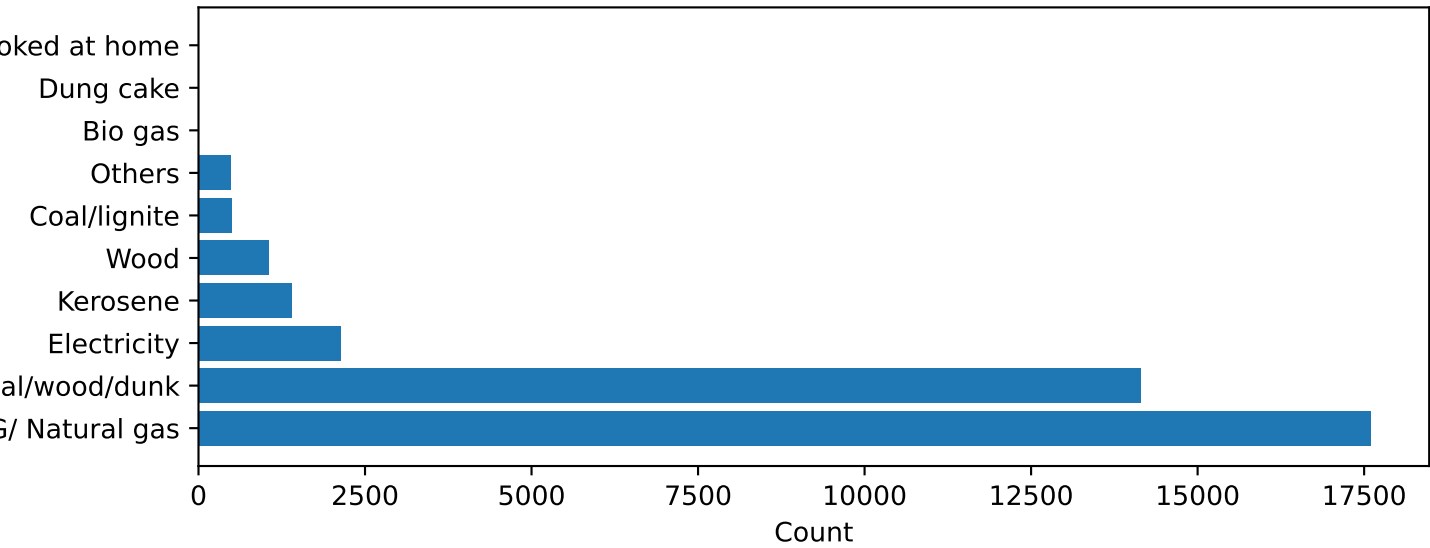
Proportions by LBW
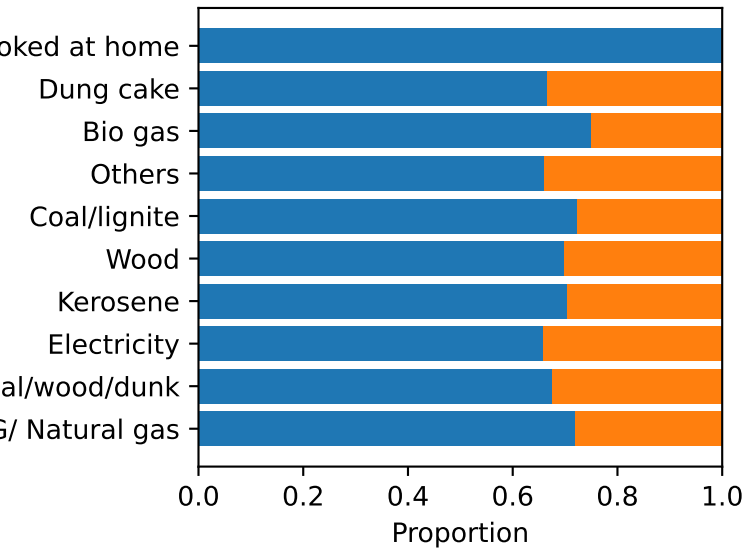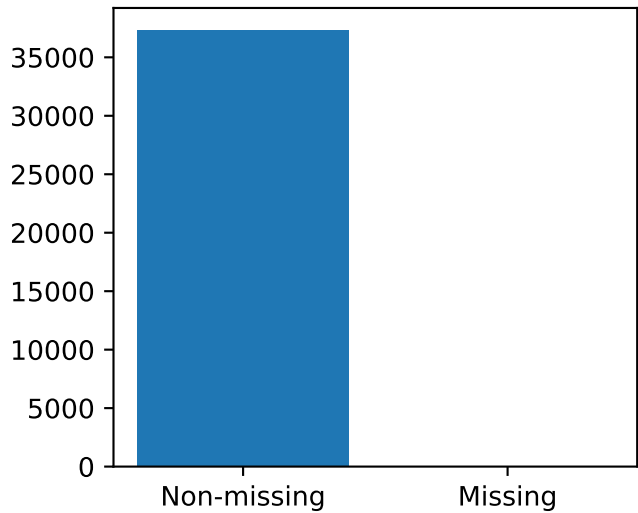
Missing: 1/37353

Hypothesis Tests & Metrics for: Types of fuel used (categorical)
Missing: 0.00% | Unique: 10
Chi-square: chi2=94.4, dof=9, p=2.11e-16
Cramér's V (bias-corrected): 0.048

EDA – Food_Groups_Category

Top-15 categories: Food_Groups_Category
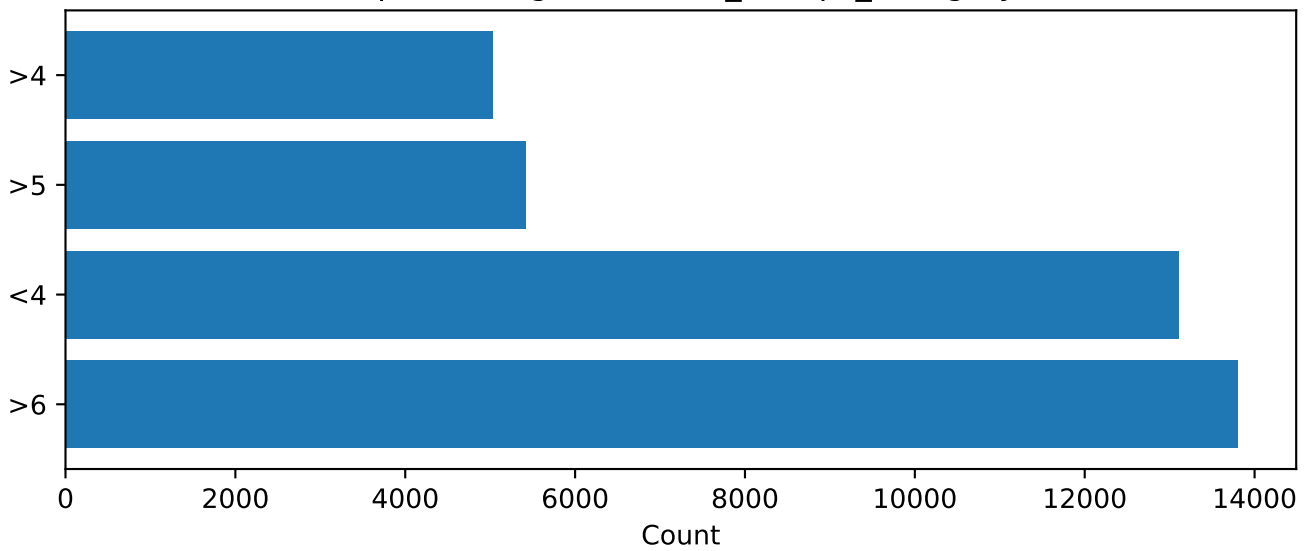
Proportions by LBW
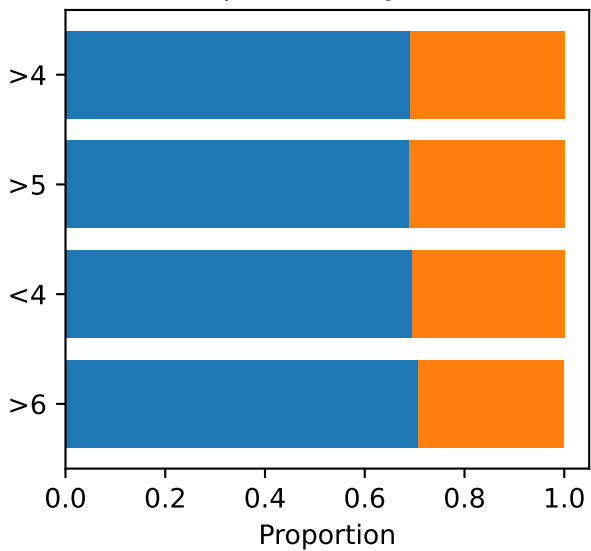
Missing: 0/37353
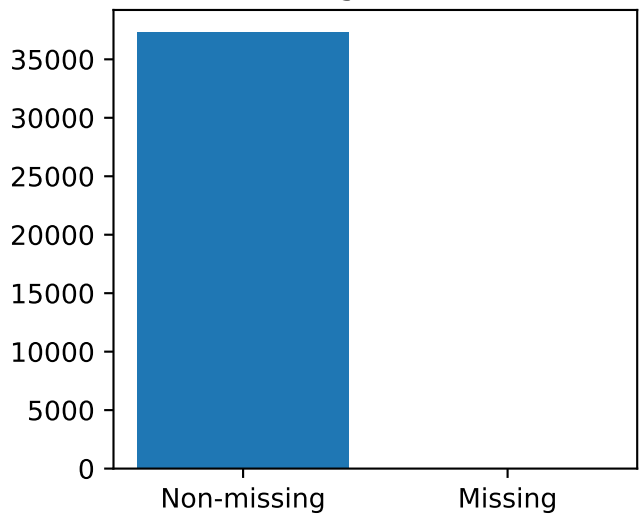
Hypothesis Tests & Metrics for: Food_Groups_Category (categorical)
Missing: 0.00% | Unique: 4
Chi-square: chi2=10.5, dof=3, p=0.0146
Cramér's V (bias-corrected): 0.014

EDA – Food_Intake_Quality_Awareness_Score

Histogram: Food_Intake_Quality_Awareness_Score
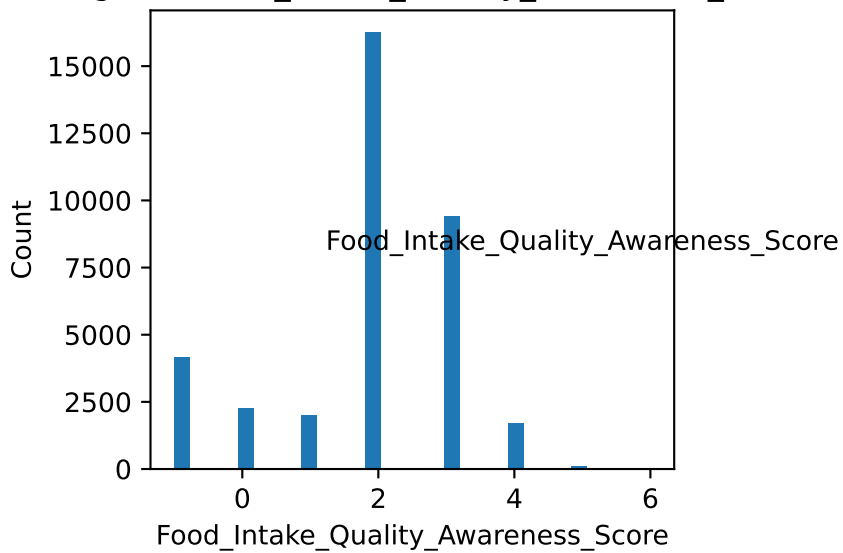
Tukey Boxplot
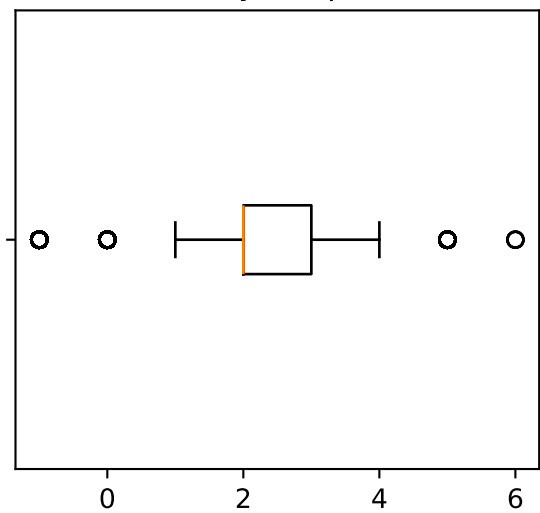
Food_Intake_Quality_Awareness_Score by LBW
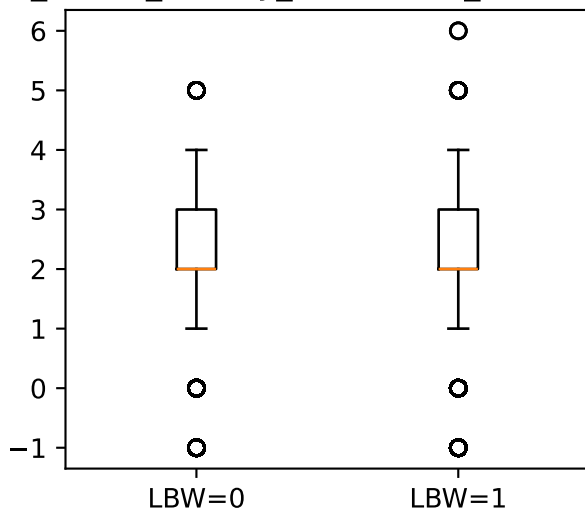
Missing: 1470/37353
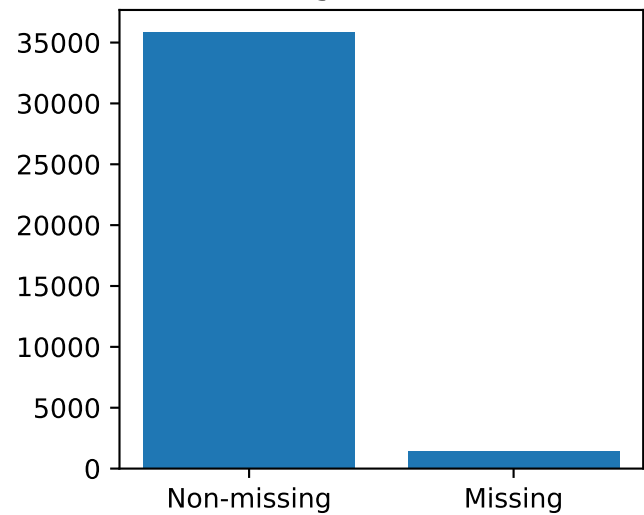
Hypothesis Tests & Metrics for: Food_Intake_Quality_Awareness_Score (numeric)
Missing: 3.94% | Unique: 8
Welch t-test: t=2.71, p=0.00668
Point–Biserial r: r=-0.0144, p=0.00646
Cohen's d: -0.0313 | KS: D=0.0136, p=0.119
AUROC (using Food_Intake_Quality_Awareness_Score as score): 0.491

Top-15 categories: Currently having disease- diabetes
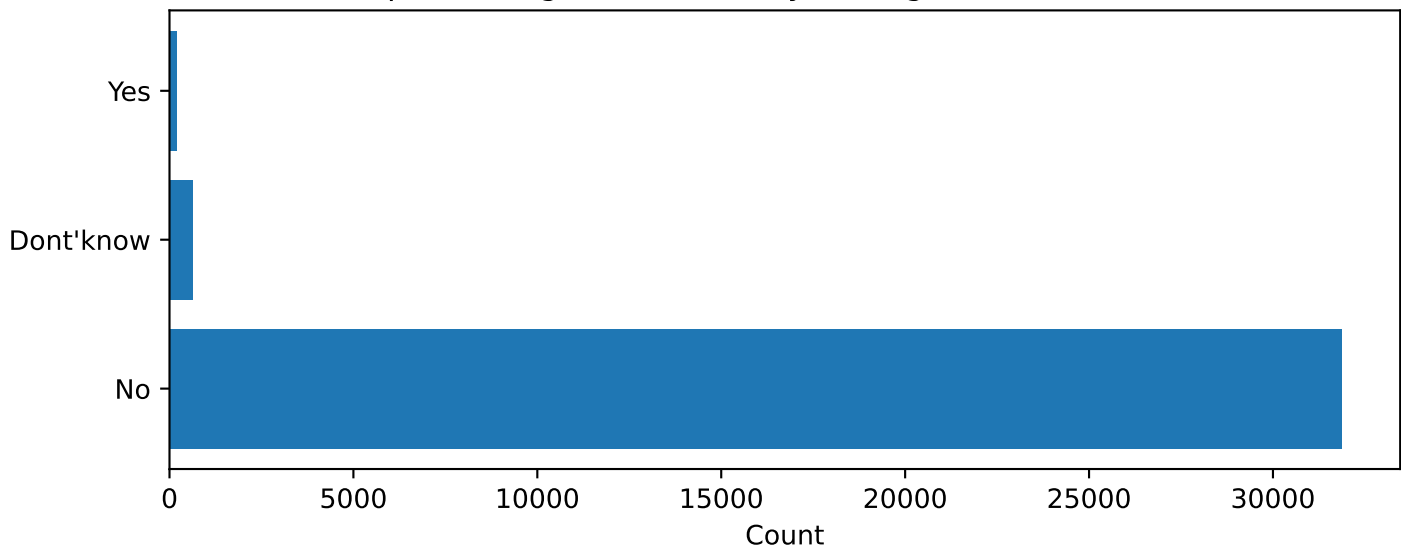
Proportions by LBW
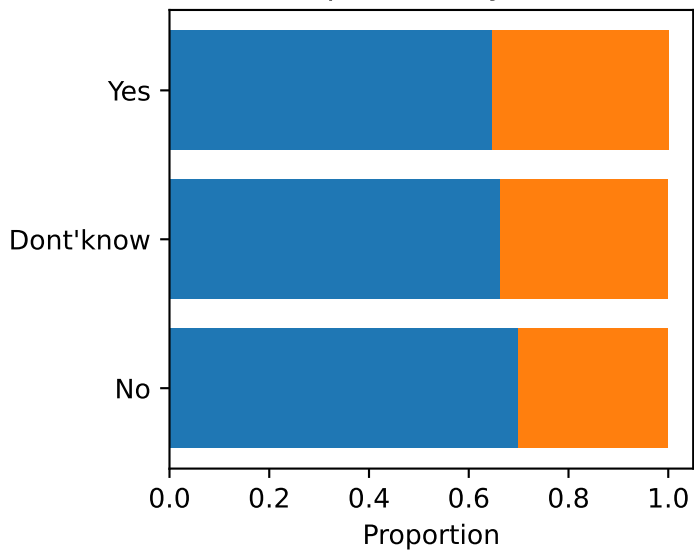
Missing: 4637/37353
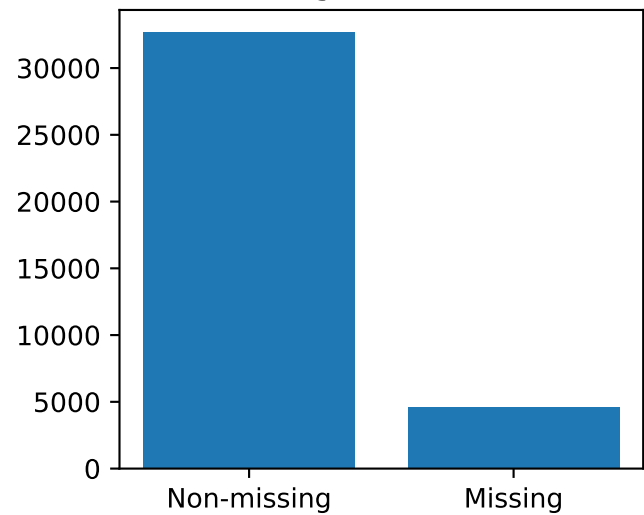
Hypothesis Tests & Metrics for: Currently having disease- diabetes (categorical)
Missing: 12.41% | Unique: 3
Chi-square: chi2=6.58, dof=2, p=0.0373
Cramér's V (bias-corrected): 0.012

EDA – Status of current smoking of bidi

Top-15 categories: Status of current smoking of bidi
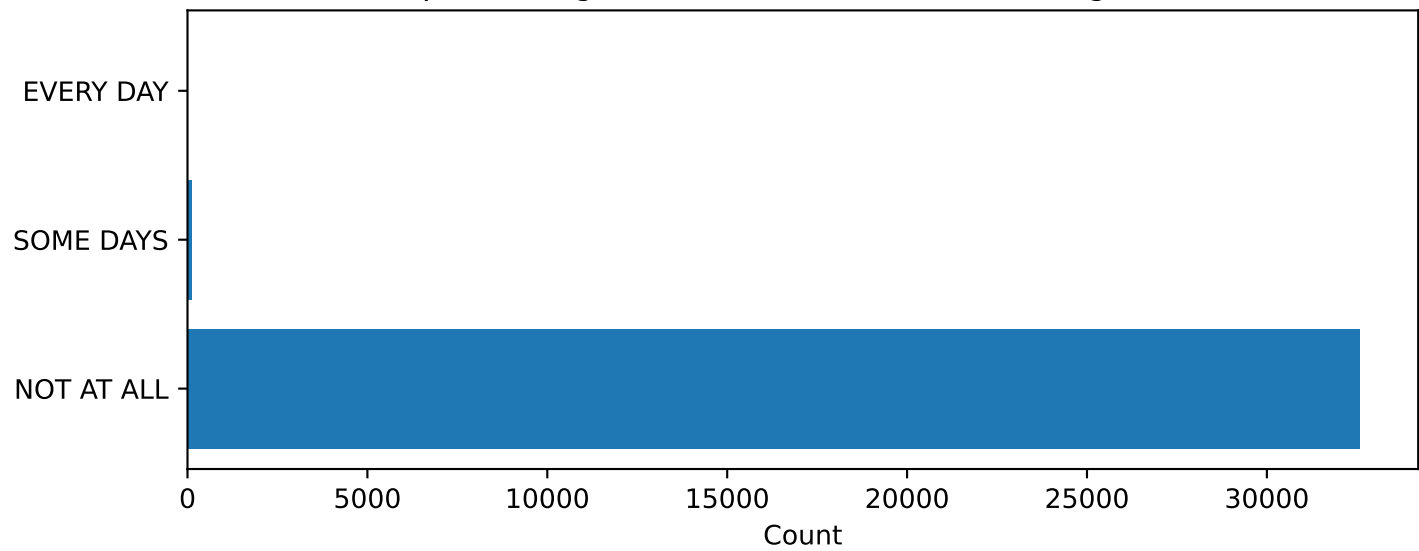
Proportions by LBW
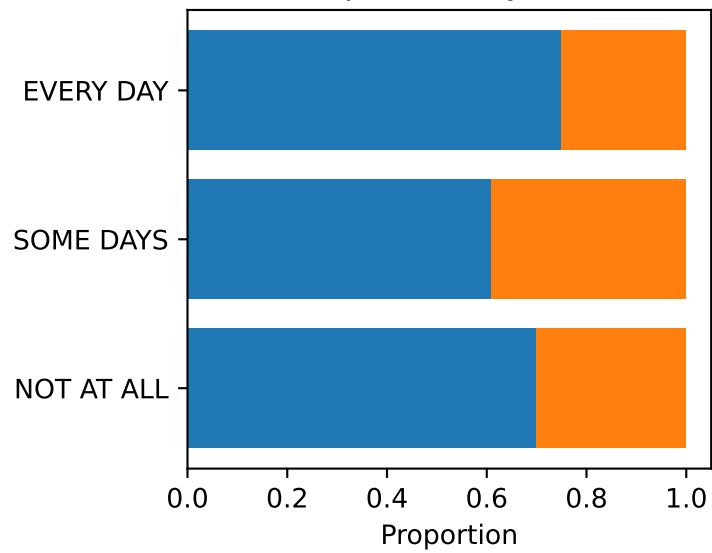
Missing: 4637/37353

Hypothesis Tests & Metrics for: Status of current smoking of bidi (categorical)
Missing: 12.41% | Unique: 3
Chi-square: chi2=4.88, dof=2, p=0.087
Cramér's V (bias-corrected): 0.009

EDA – consume_tobacco

Top-15 categories: consume_tobacco
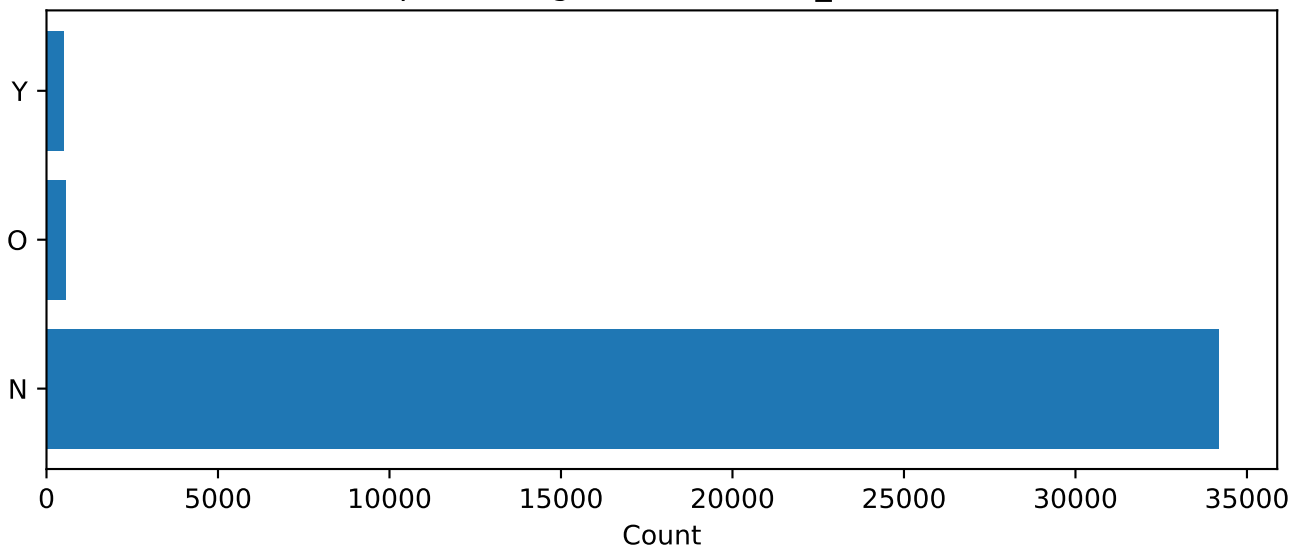
Proportions by LBW
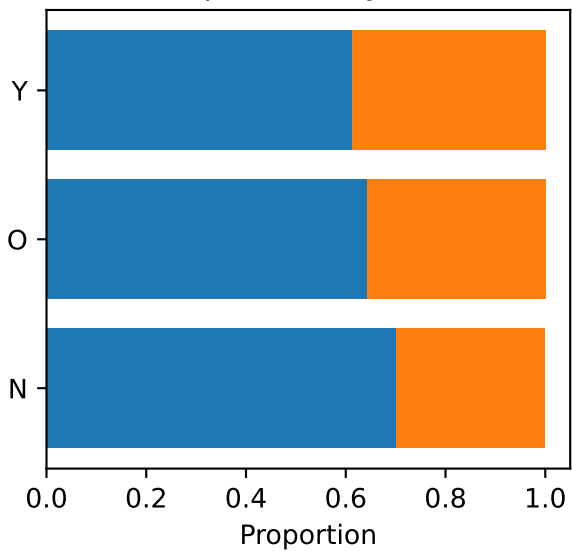
Missing: 2123/37353
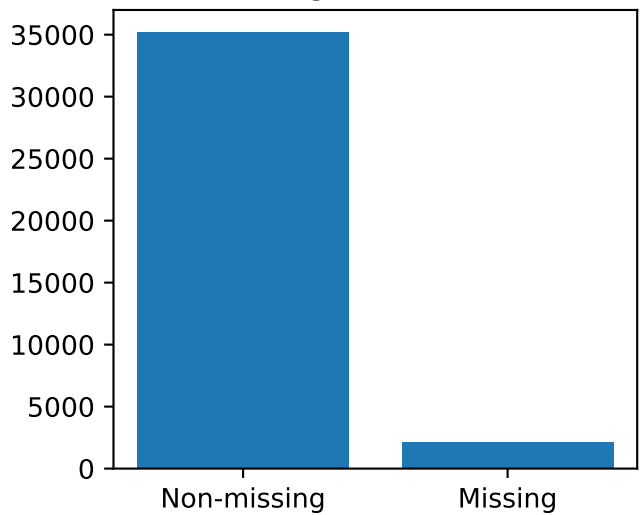
Hypothesis Tests & Metrics for: consume_tobacco (categorical)
Missing: 5.68% | Unique: 3
Chi-square: chi2=27.4, dof=2, p=1.1e-06
Cramér's V (bias-corrected): 0.027

EDA – Status of current chewing of tobacco

Top-15 categories: Status of current chewing of tobacco
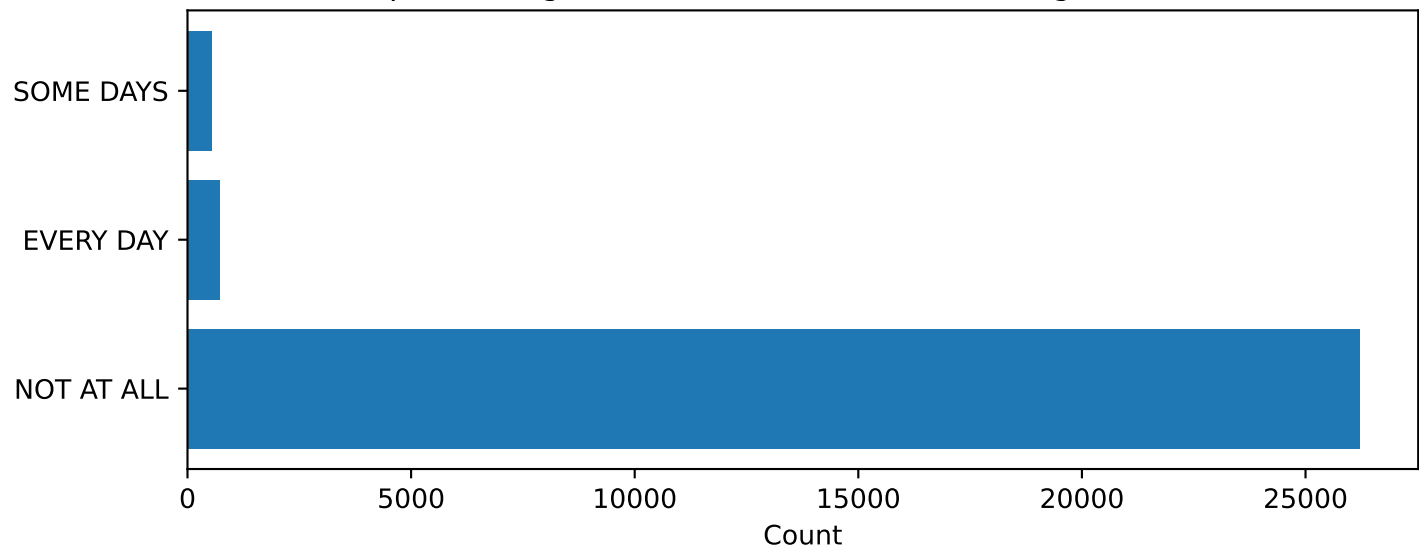
Proportions by LBW
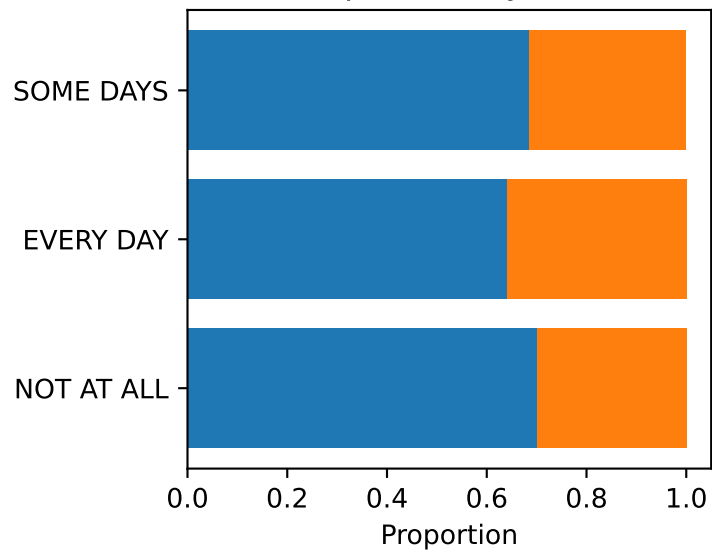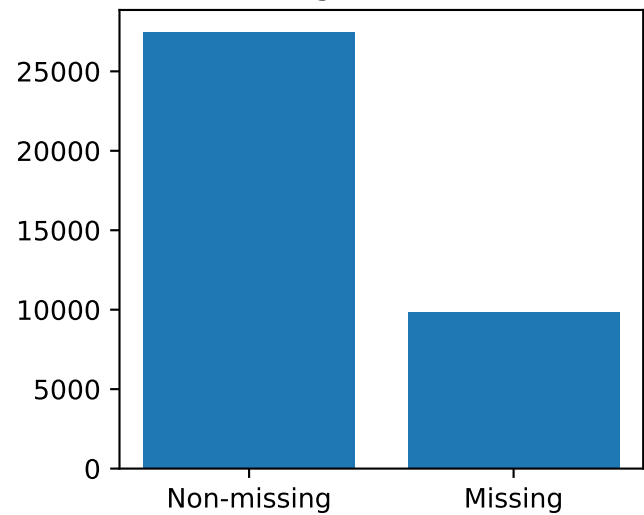
Missing: 9865/37353

Hypothesis Tests & Metrics for: Status of current chewing of tobacco (categorical)
Missing: 26.41% | Unique: 3
Chi-square: chi2=12.8, dof=2, p=0.00169
Cramér's V (bias-corrected): 0.020

EDA – Alcoholic beverages consumed yesterday

Top-15 categories: Alcoholic beverages consumed yesterday
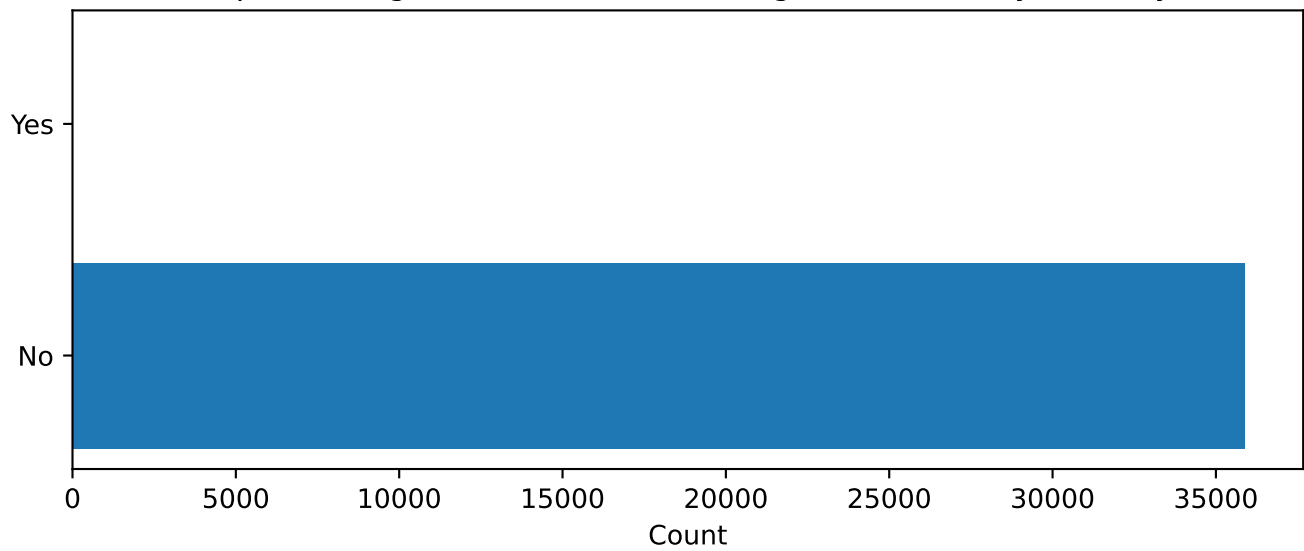
Proportions by LBW
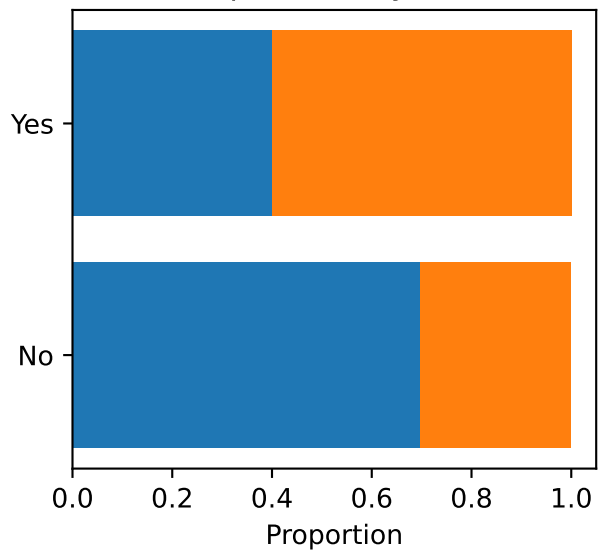
Missing: 1470/37353
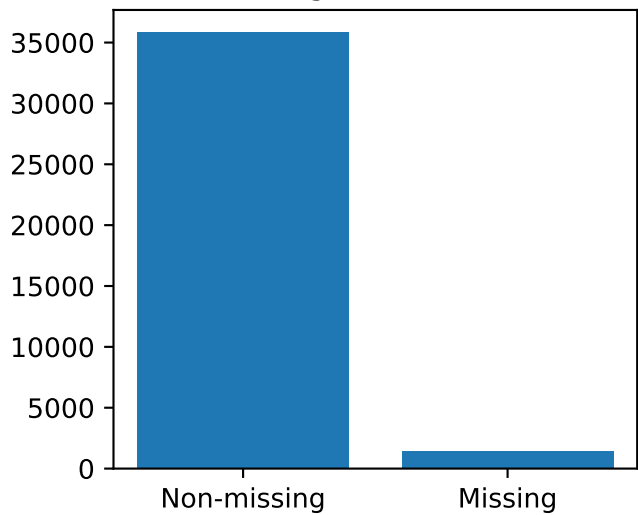
Hypothesis Tests & Metrics for: Alcoholic beverages consumed yesterday (categorical)
Missing: 3.94% | Unique: 2
Chi-square: chi2=0.923, dof=1, p=0.337
Cramér's V (bias-corrected): 0.000

EDA – consume_alcohol

Top-15 categories: consume_alcohol
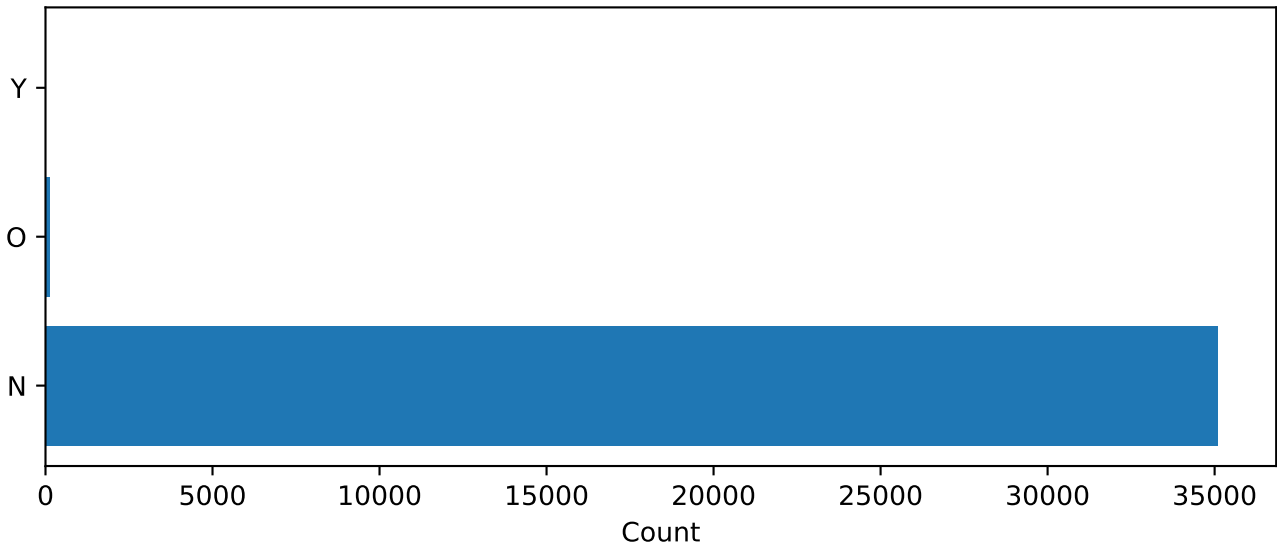
Proportions by LBW
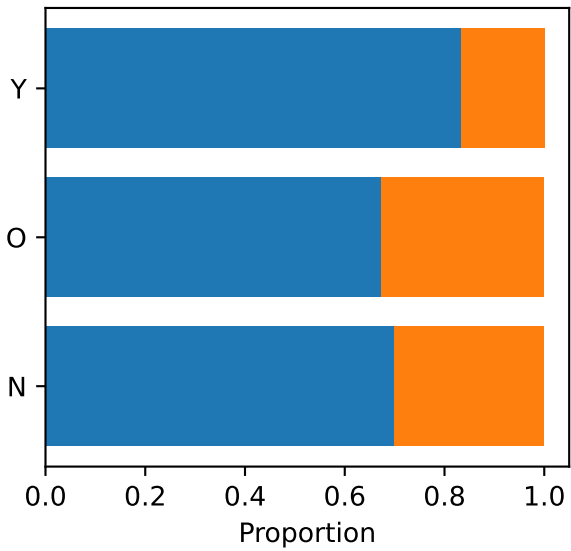
Missing: 2123/37353
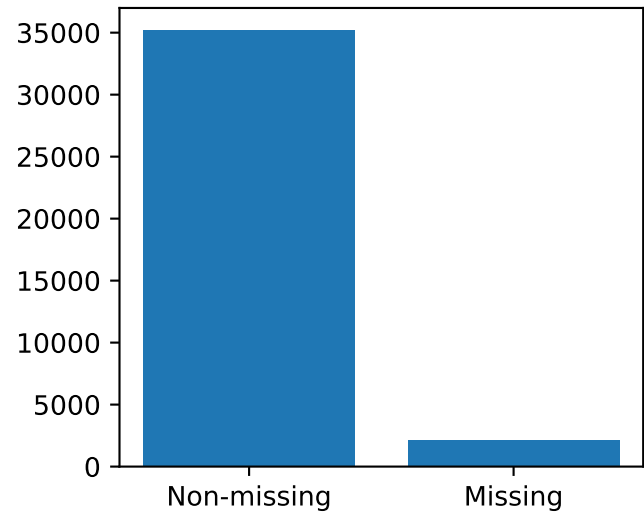
Hypothesis Tests & Metrics for: consume_alcohol (categorical)
Missing: 5.68% | Unique: 3
Chi-square: chi2=0.926, dof=2, p=0.629
Cramér's V (bias-corrected): 0.000

EDA – Status of drinking of alcohol

Top-15 categories: Status of drinking of alcohol
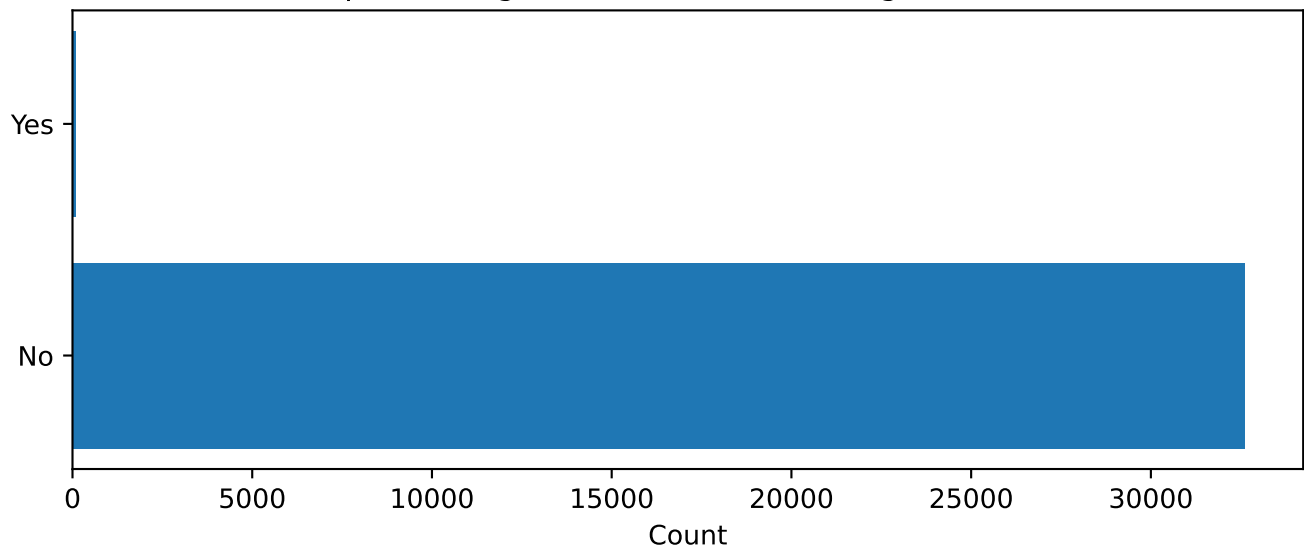
Proportions by LBW
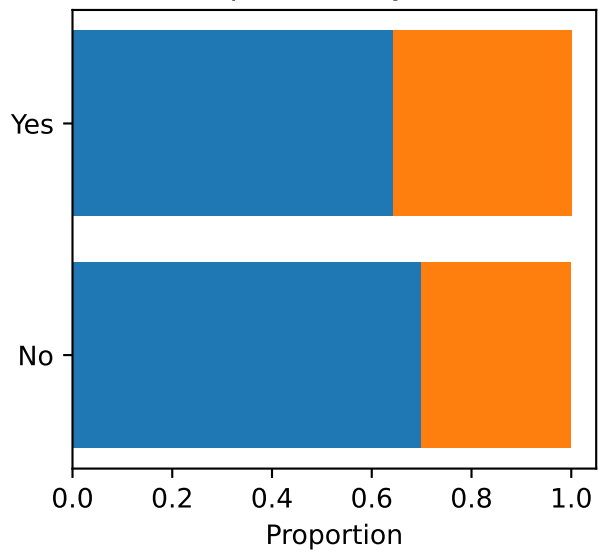
Missing: 4637/37353
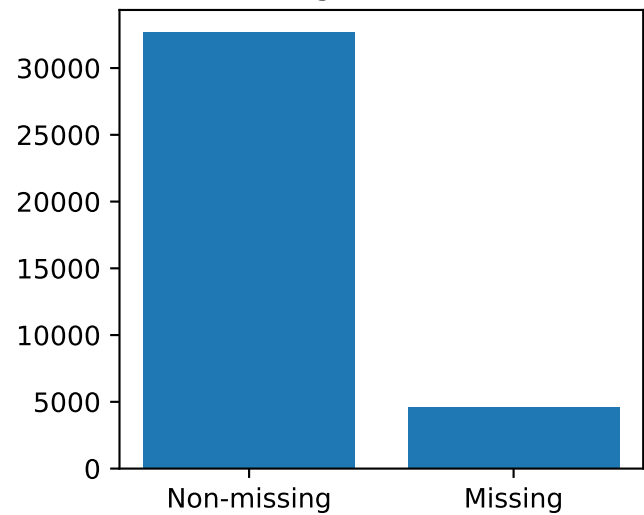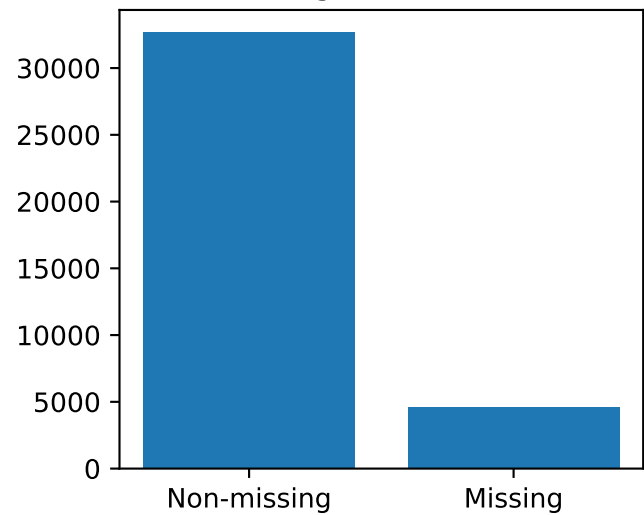

Hypothesis Tests & Metrics for: Status of drinking of alcohol (categorical)
Missing: 12.41% | Unique: 2
Chi-square: chi2=1.42, dof=1, p=0.233
Cramér's V (bias-corrected): 0.004
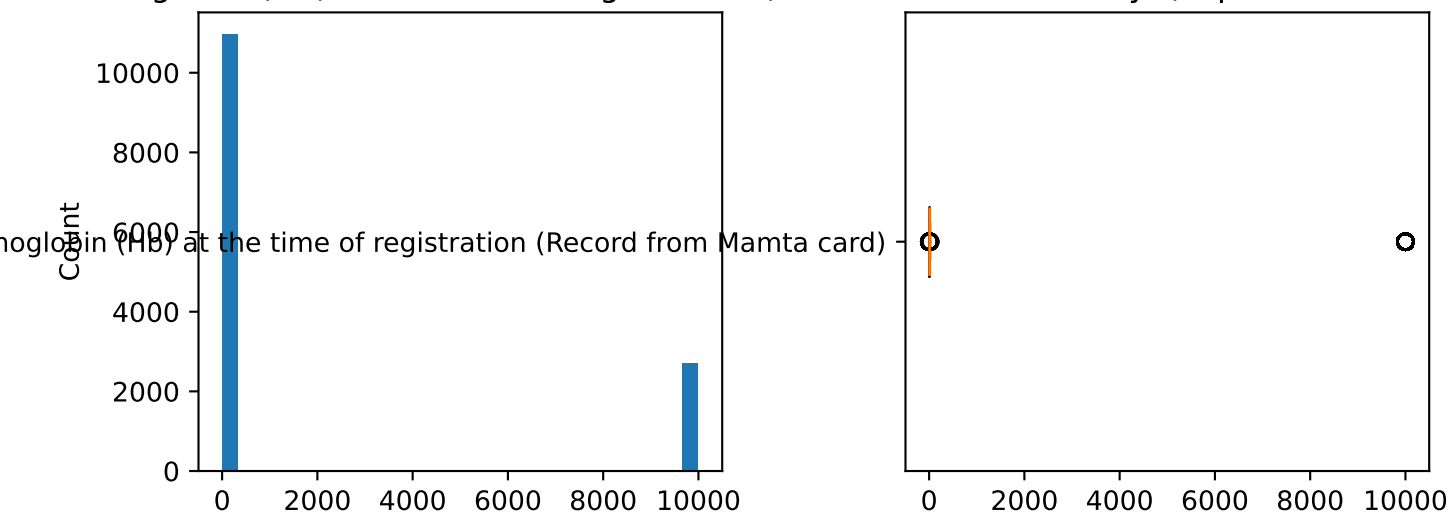
Level of hemoglobin (Hb) at the time of registration (Record from Mamta card) — Tukey Boxplot
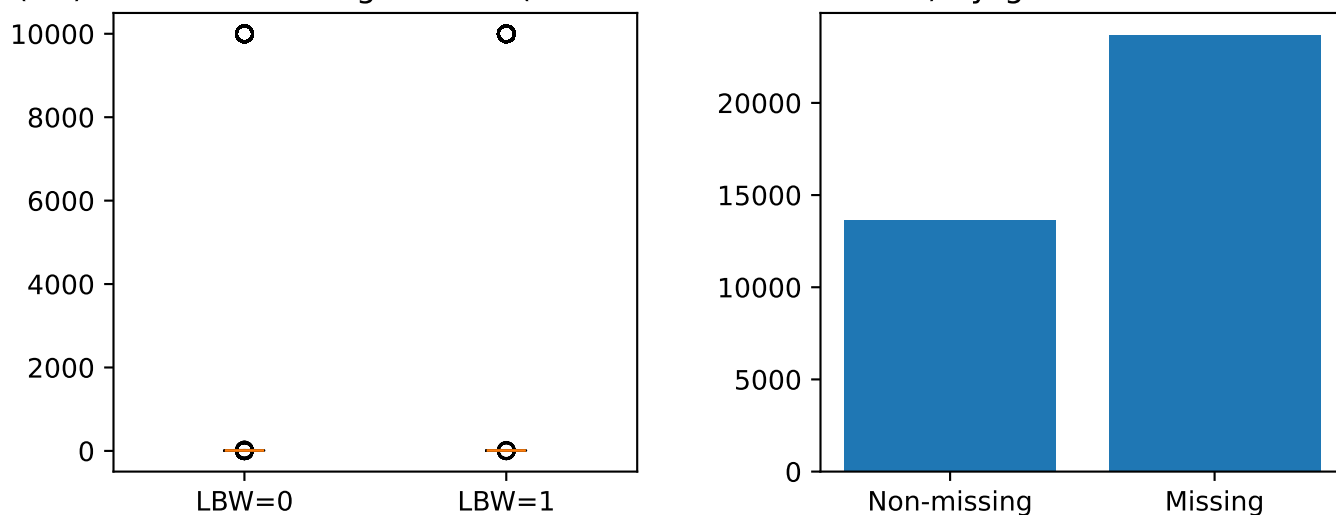
Hypothesis Tests & Metrics for: Level of hemoglobin (Hb) at the time of registration (Record from Mamta card) (nu...
Missing: 63.43% | Unique: 17
Welch t-test: t=1.27, p=0.205
Point–Biserial r: r=-0.0108, p=0.209
Cohen's d: -0.0234 | KS: D=0.0326, p=0.00421
AUROC (using Level of hemoglobin (Hb) at the time of registration (Record from Mamta card) as score): 0.480

EDA – Service received during last ANC: Hb test done

Top-15 categories: Service received during last ANC: Hb test done

Proportions by LBW

Missing: 23754/37353
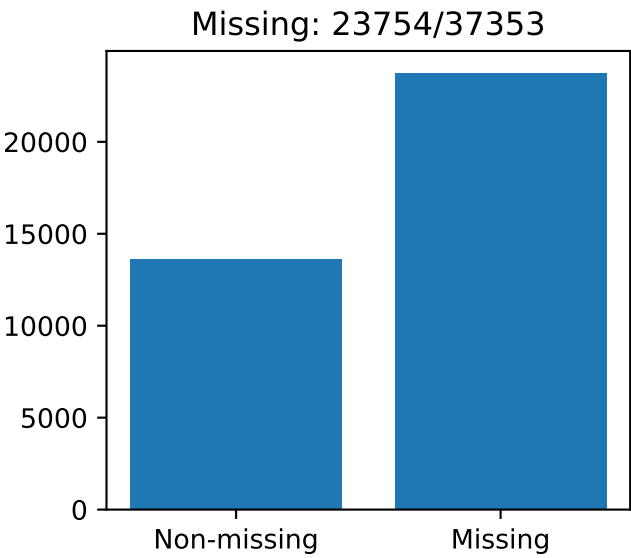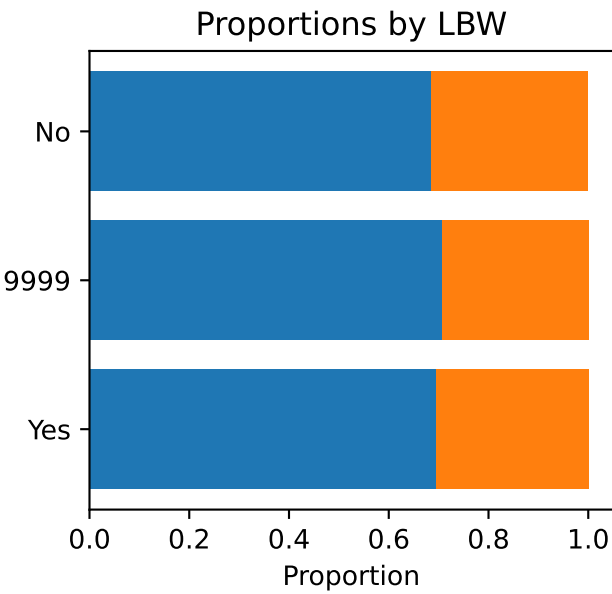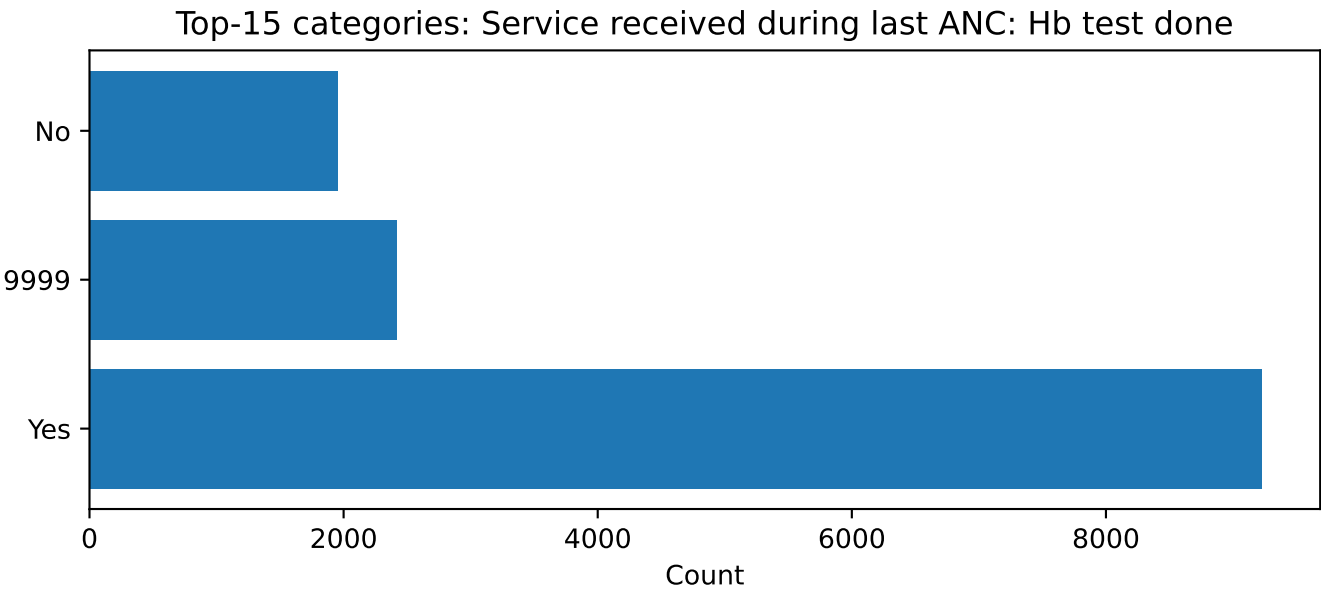
Hypothesis Tests & Metrics for: Service received during last ANC: Hb test done (categorical)
Missing: 63.59% | Unique: 3
Chi-square: chi2=2.18, dof=2, p=0.337
Cramér's V (bias-corrected): 0.004

EDA – measured_HB

Histogram: measured_HB

Tukey Boxplot

measured_HB by LBW

Missing: 27951/37353
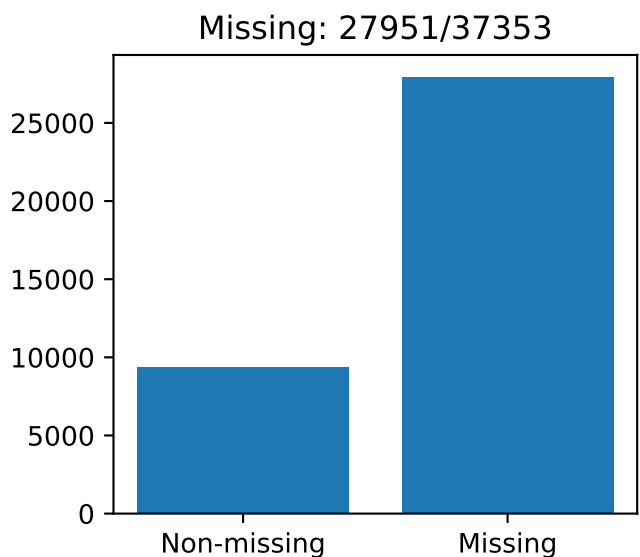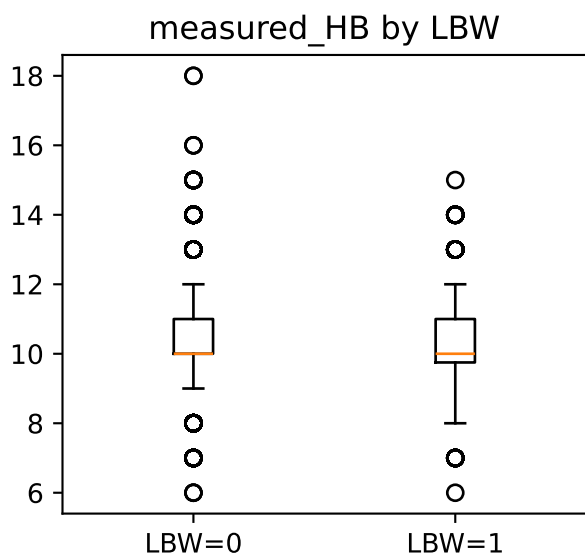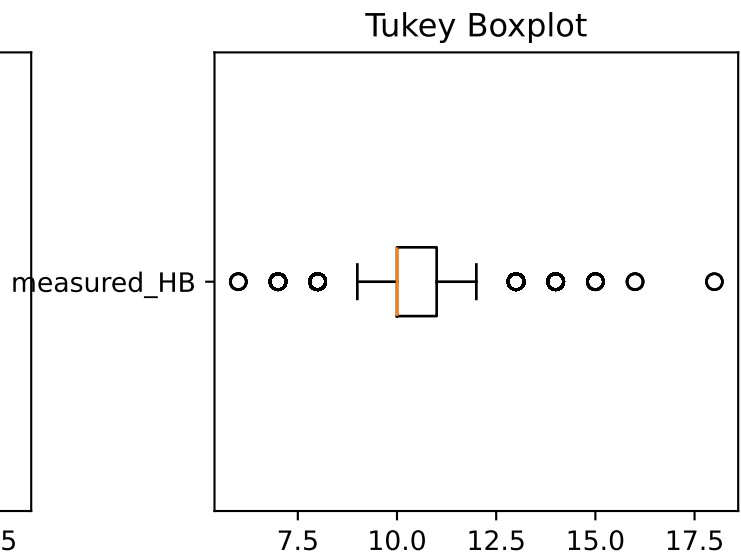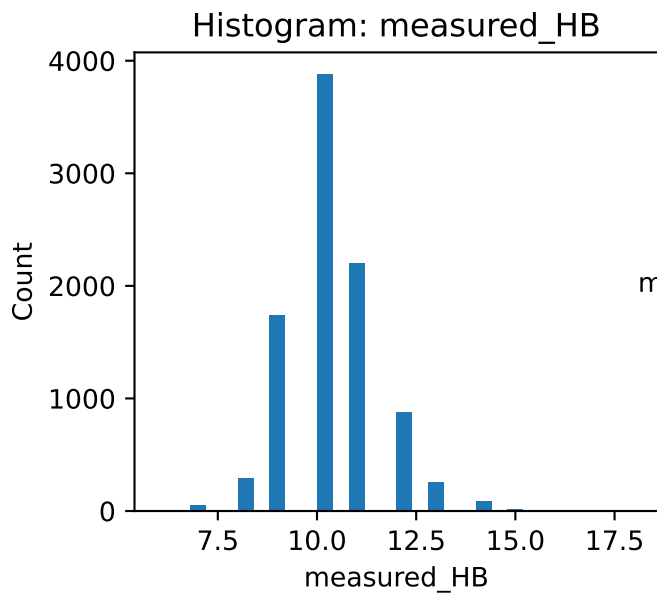
Hypothesis Tests & Metrics for: measured_HB (numeric)
Missing: 74.83% | Unique: 12
Welch t-test: t=6.16, p=7.71e-10
Point–Biserial r: r=-0.0618, p=1.95e-09
Cohen's d: -0.136 | KS: D=0.0512, p=7.15e-05
AUROC (using measured_HB as score): 0.464

Histogram: Child order/parity

Tukey Boxplot

Child order/parity by LBW

Missing: 4/37353
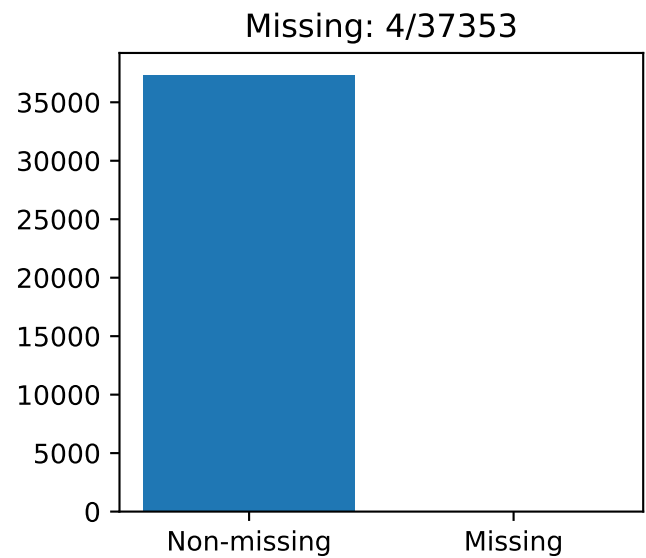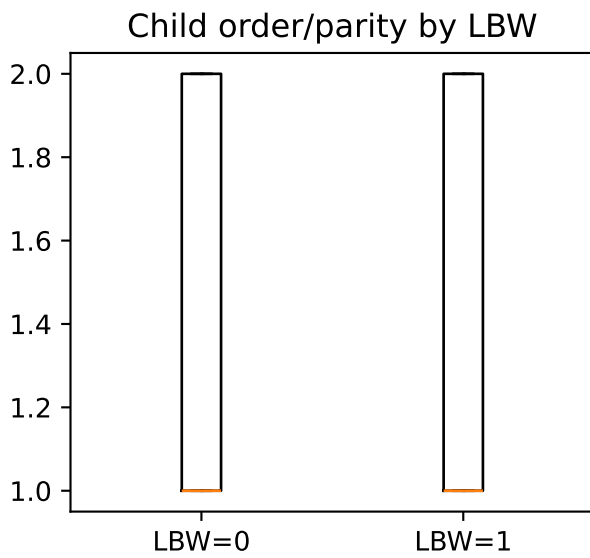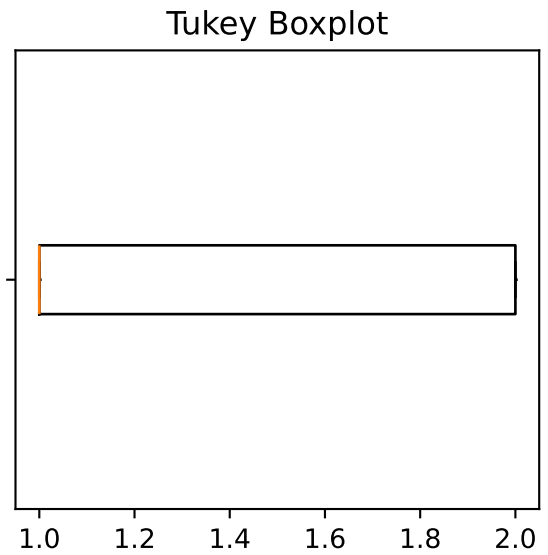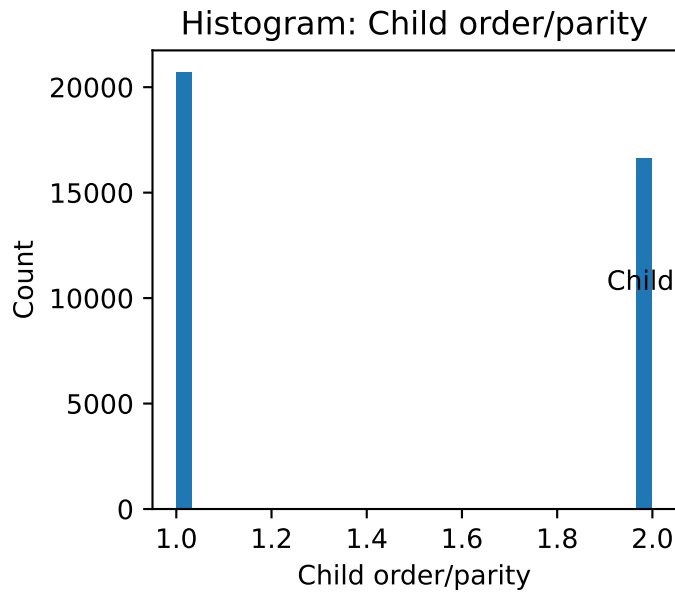
Hypothesis Tests & Metrics for: Child order/parity (numeric)
Missing: 0.01% | Unique: 2
Welch t-test: t=4.92, p=8.77e-07
Point–Biserial r: r=-0.0254, p=9.34e-07
Cohen's d: -0.0553 | KS: D=0.0275, p=1.35e-05
AUROC (using Child order/parity as score): 0.486

EDA – WeightGain_PW2

Histogram: WeightGain_PW2

Tukey Boxplot

WeightGain_PW2 by LBW

Missing: 28643/37353
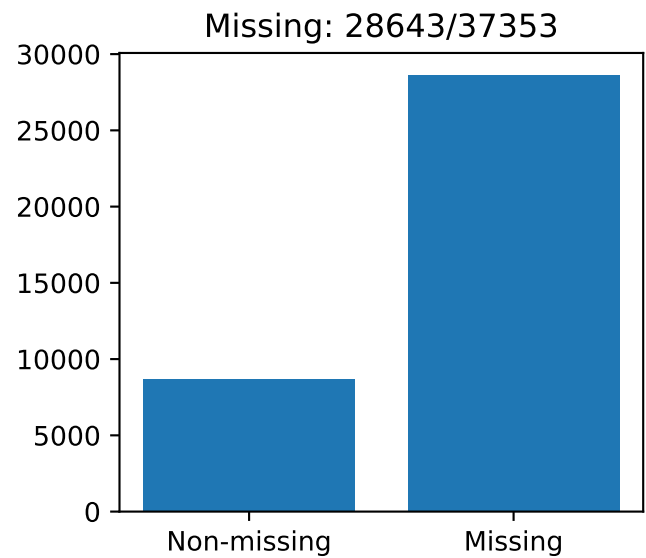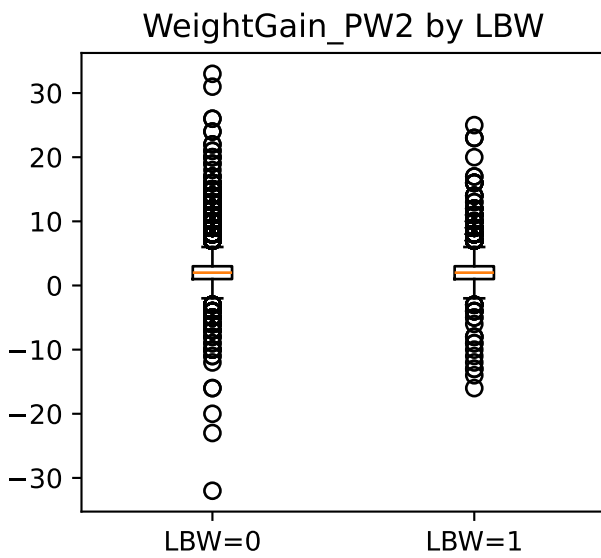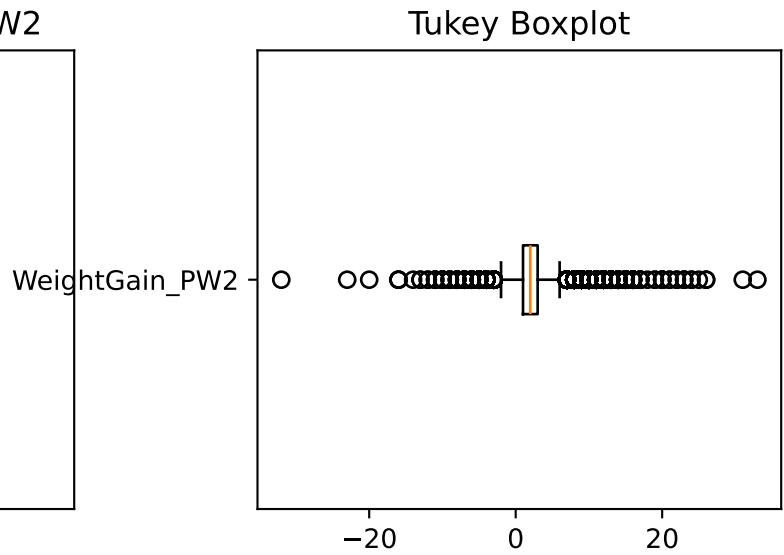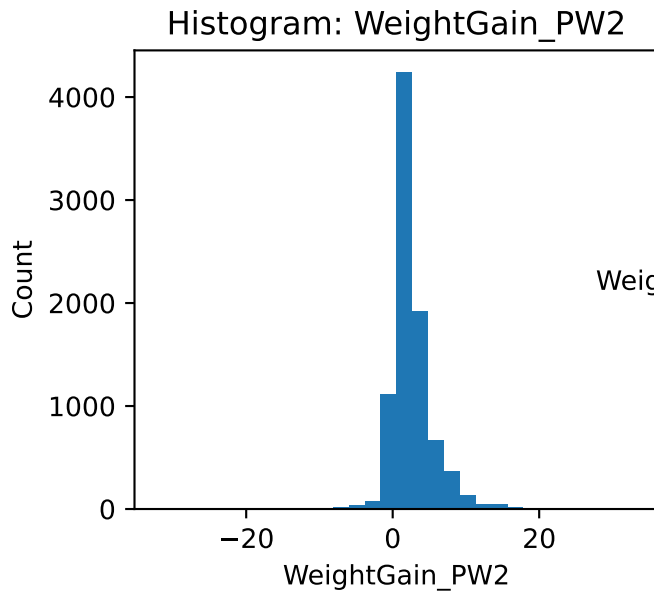
Hypothesis Tests & Metrics for: WeightGain_PW2 (numeric)
Missing: 76.68% | Unique: 47
Welch t-test: t=3.33, p=0.000859
Point–Biserial r: r=-0.0345, p=0.00127
Cohen's d: -0.0764 | KS: D=0.0295, p=0.0876
AUROC (using WeightGain_PW2 as score): 0.477

EDA – WeightGain_PW3

Histogram: WeightGain_PW3

Tukey Boxplot

WeightGain_PW3 by LBW

Missing: 17785/37353
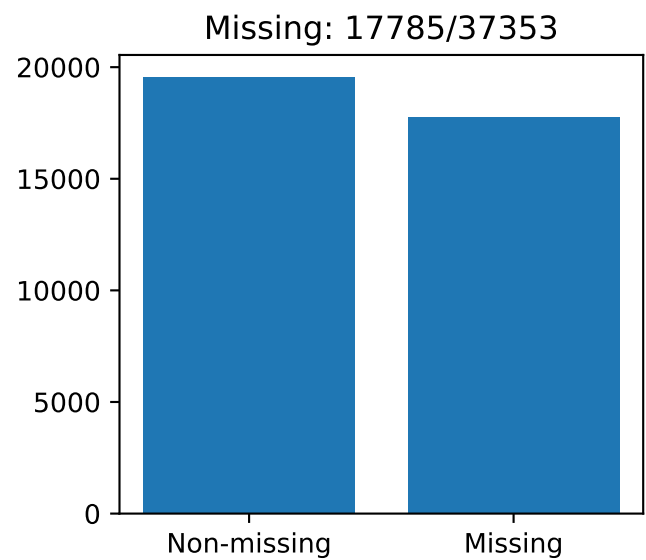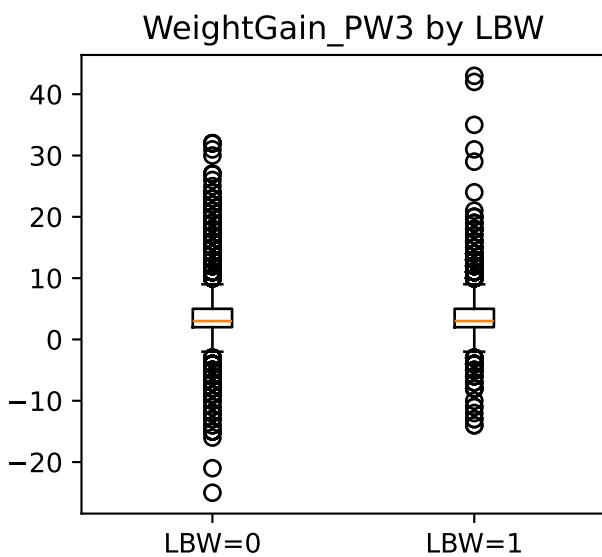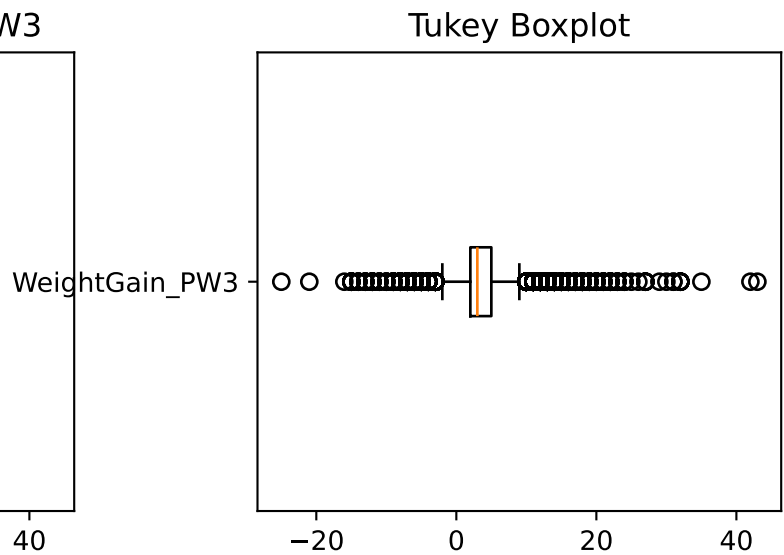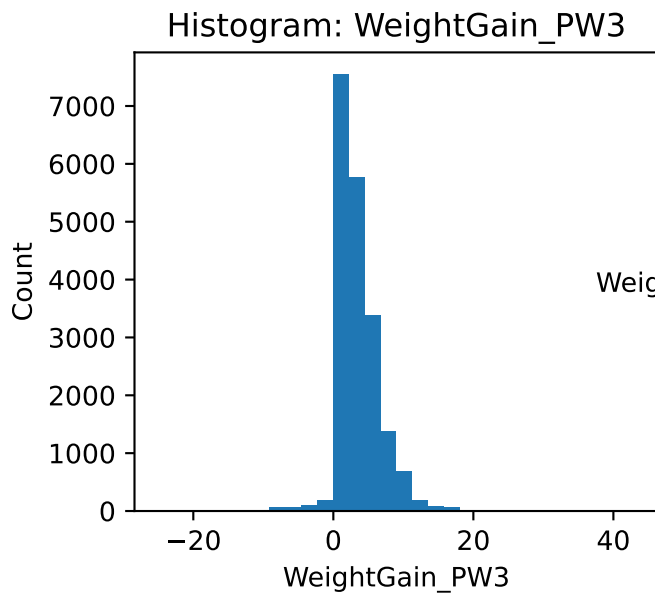
Hypothesis Tests & Metrics for: WeightGain_PW3 (numeric)
Missing: 47.61% | Unique: 53
Welch t-test: t=0.342, p=0.733
Point–Biserial r: r=-0.00238, p=0.739
Cohen's d: -0.00522 | KS: D=0.00989, p=0.819
AUROC (using WeightGain_PW3 as score): 0.499

EDA – WeightGain_PW4

Histogram: WeightGain_PW4

Tukey Boxplot

WeightGain_PW4 by LBW

Missing: 25510/37353
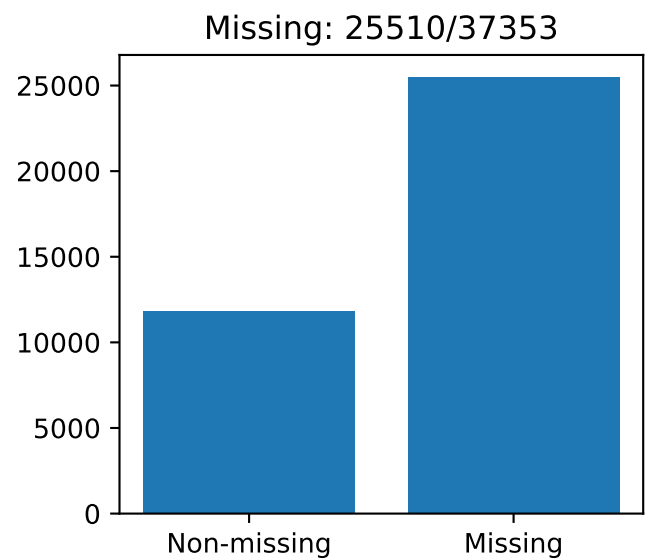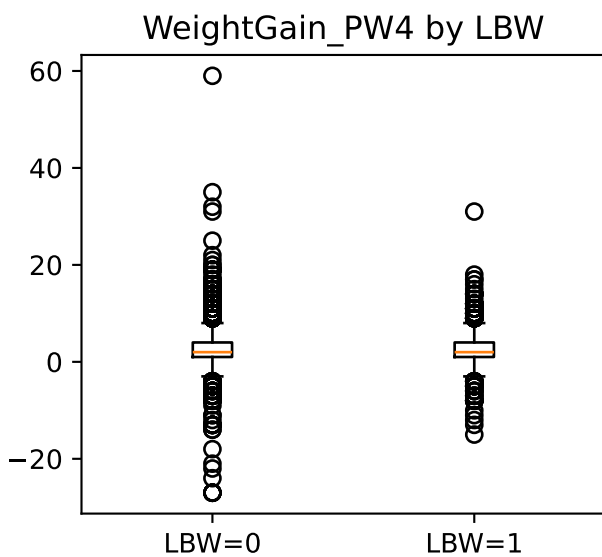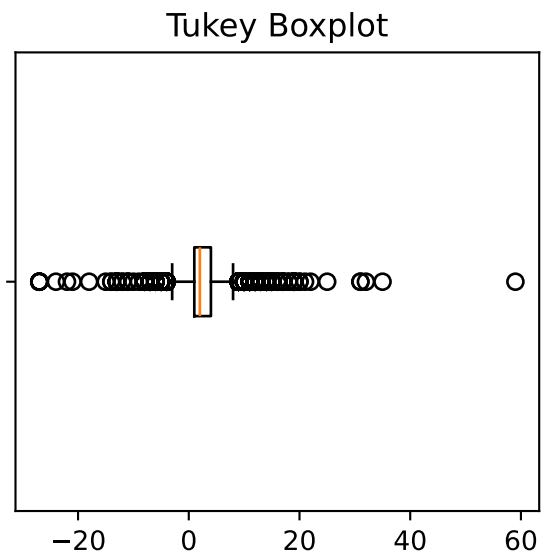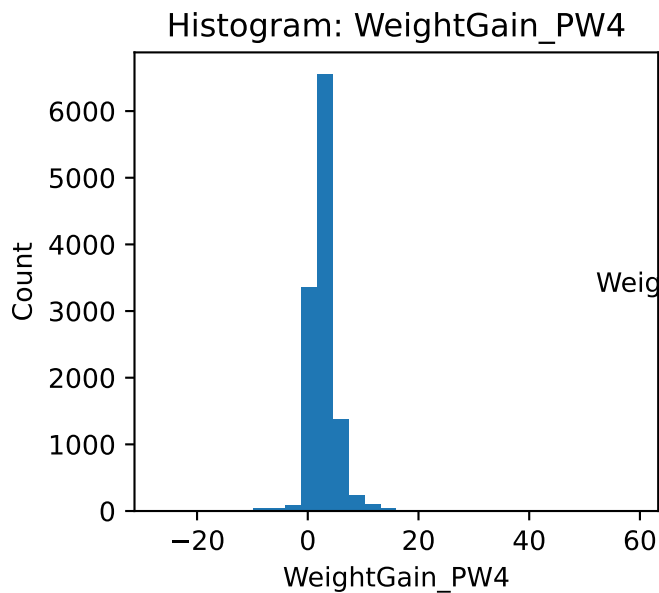
Hypothesis Tests & Metrics for: WeightGain_PW4 (numeric)
Missing: 68.29% | Unique: 48
Welch t-test: t=1.82, p=0.0682
Point–Biserial r: r=-0.0158, p=0.0858
Cohen's d: -0.0353 | KS: D=0.0144, p=0.699
AUROC (using WeightGain_PW4 as score): 0.491

EDA – Household_Assets_Score

Histogram: Household_Assets_Score

Tukey Boxplot

Household_Assets_Score by LBW

Missing: 0/37353
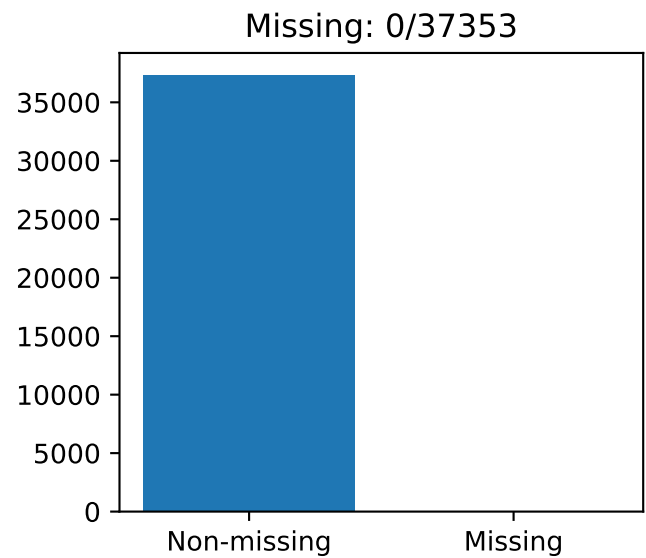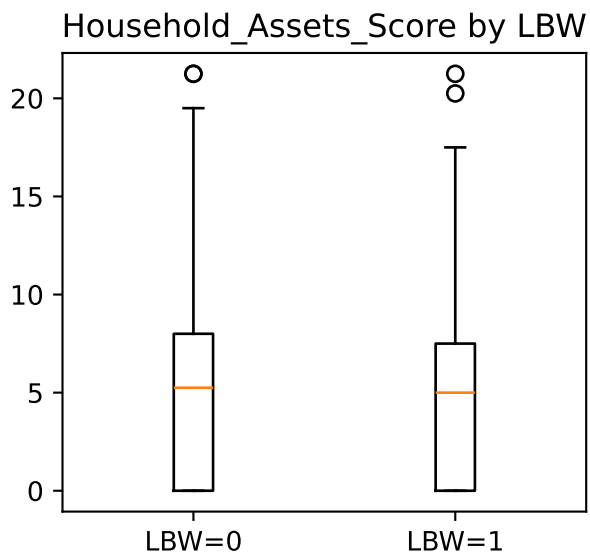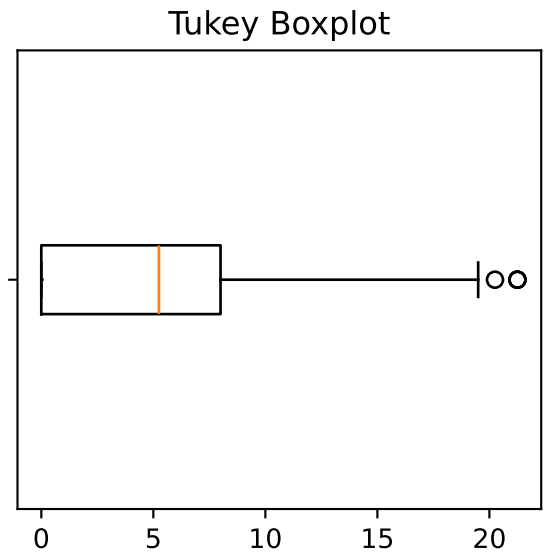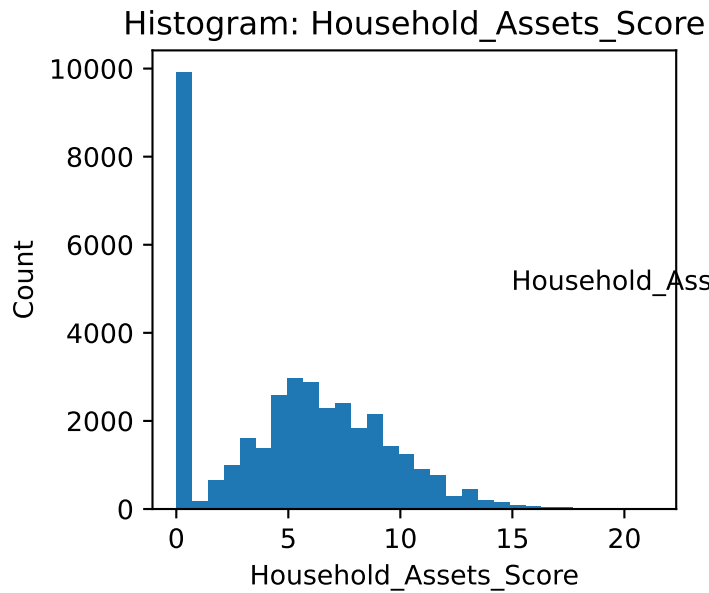
Hypothesis Tests & Metrics for: Household_Assets_Score (numeric)
Missing: 0.00% | Unique: 80
Welch t-test: t=8.17, p=3.34e-16
Point–Biserial r: r=-0.0412, p=1.72e-15
Cohen's d: -0.0897 | KS: D=0.0472, p=1.05e-15
AUROC (using Household_Assets_Score as score): 0.476

EDA – Household_Assets_Count

Histogram: Household_Assets_Count

Tukey Boxplot

Household_Assets_Count by LBW
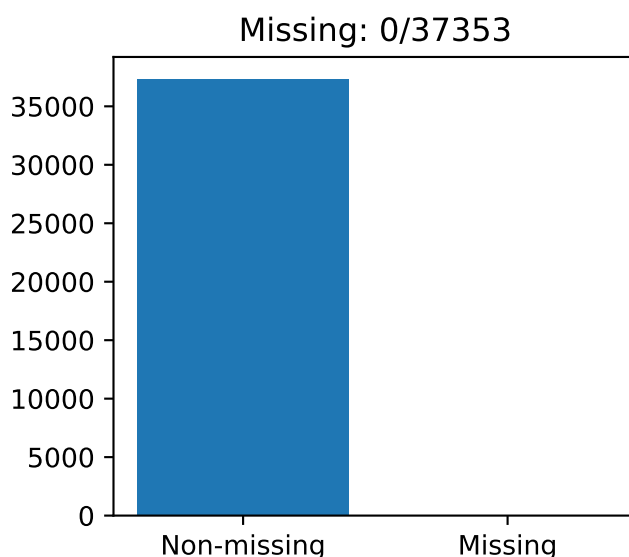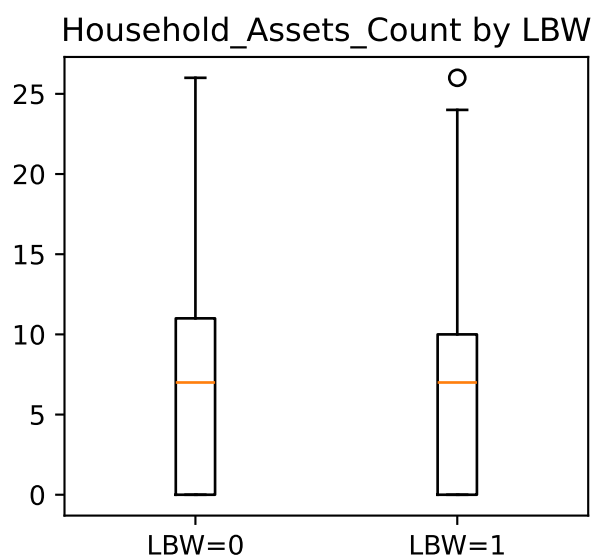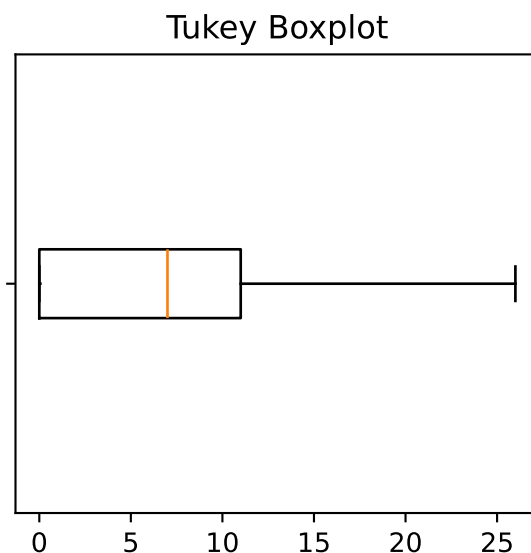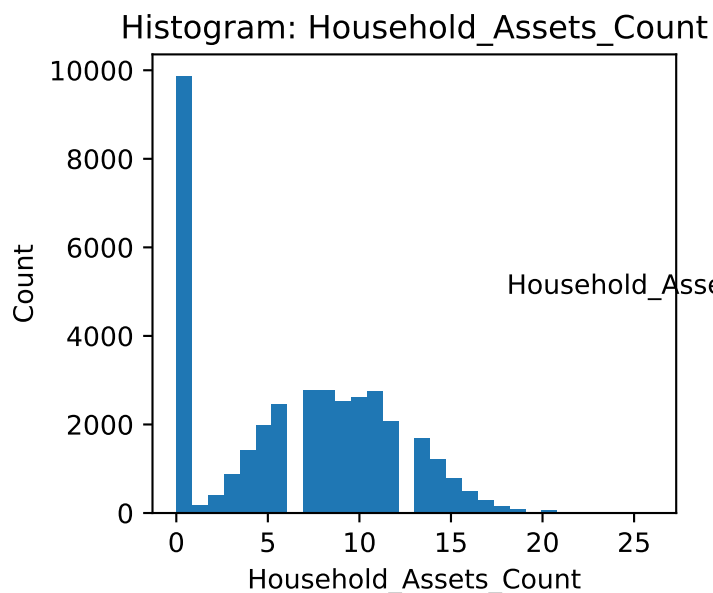
Missing: 0/37353

Hypothesis Tests & Metrics for: Household_Assets_Count (numeric)
Missing: 0.00% | Unique: 26
Welch t-test: t=7.35, p=2.05e-13
Point–Biserial r: r=-0.0373, p=5.81e-13
Cohen's d: -0.0812 | KS: D=0.0442, p=7.8e-14
AUROC (using Household_Assets_Count as score): 0.477

Top-15 categories: Household_Assets_Category
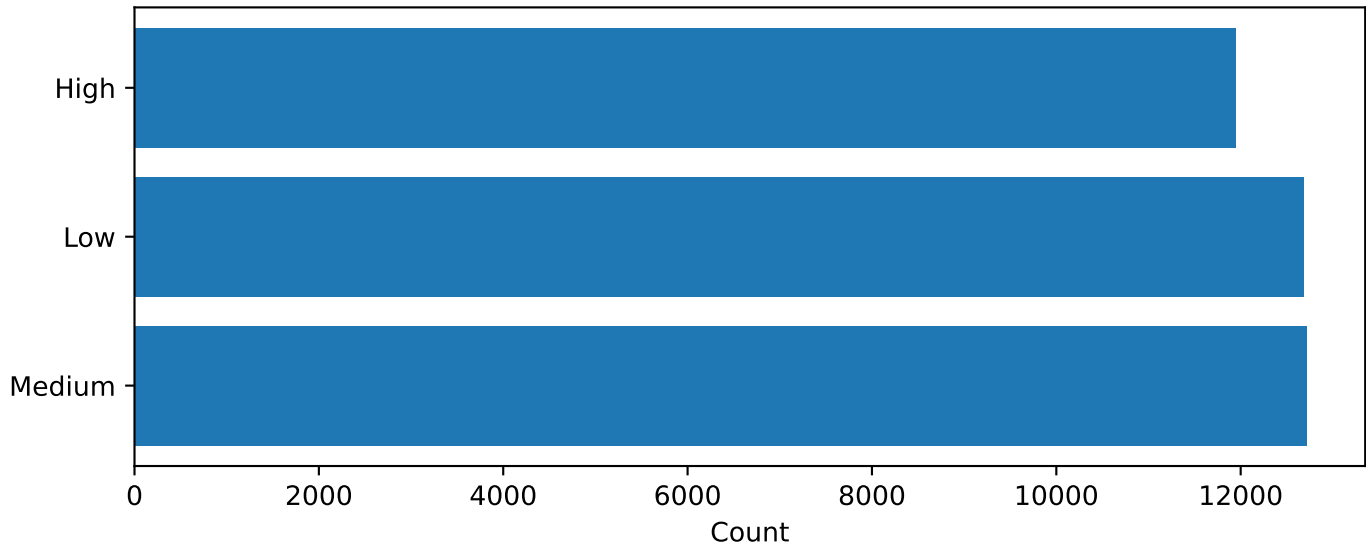
Proportions by LBW
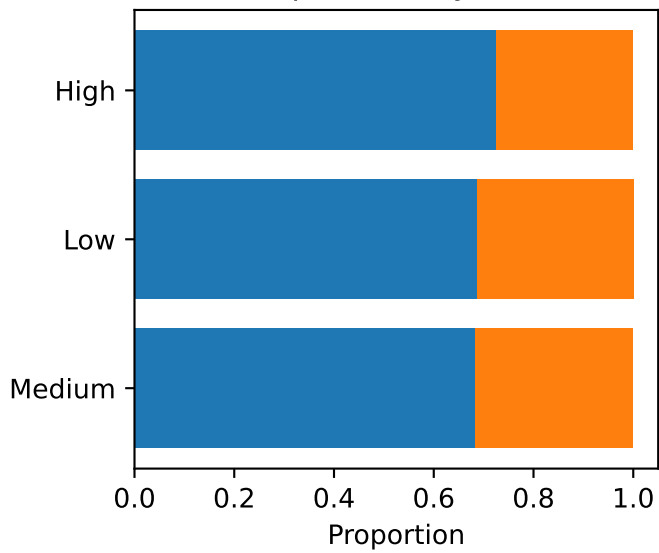
Missing: 0/37353
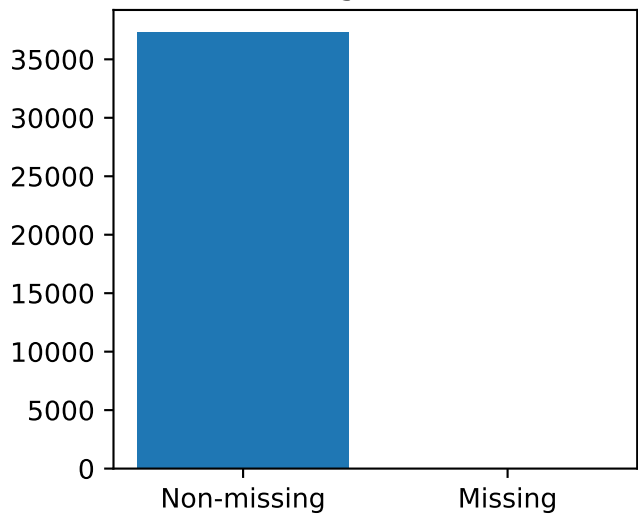
Hypothesis Tests & Metrics for: Household_Assets_Category (categorical)
Missing: 0.00% | Unique: 3
Chi-square: chi2=62.8, dof=2, p=2.36e-14
Cramér's V (bias-corrected): 0.040

EDA – No of ANCs completed

Histogram: No of ANCs completed

Tukey Boxplot

No of ANCs completed

No of ANCs completed by LBW

Missing: 0/37353
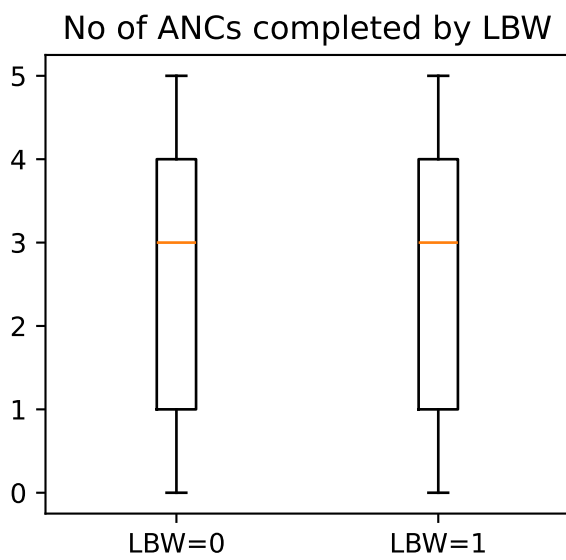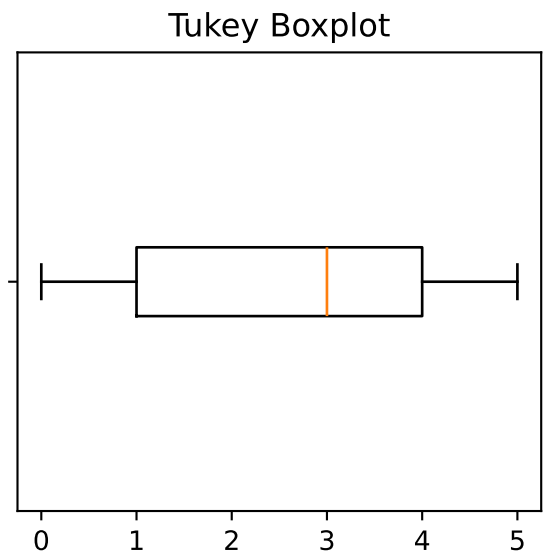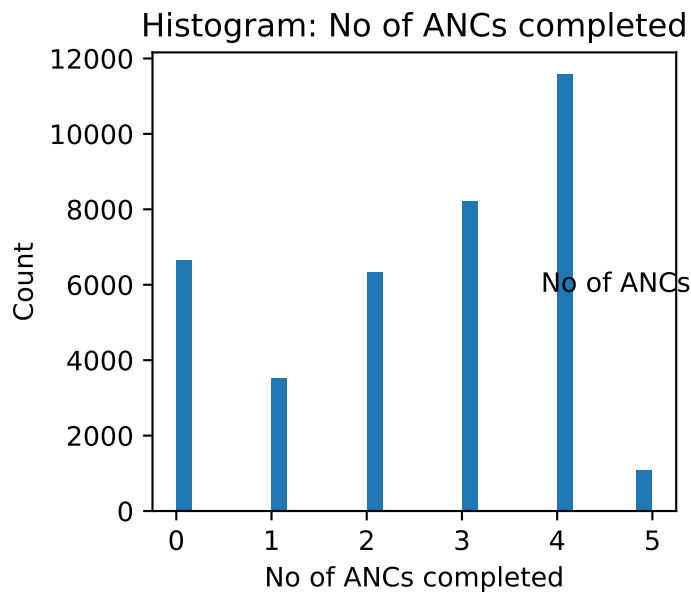
Hypothesis Tests & Metrics for: No of ANCs completed (numeric)
Missing: 0.00% | Unique: 6
Welch t-test: t=9.95, p=2.97e-23
Point–Biserial r: r=-0.0511, p=4.8e-23
Cohen's d: -0.111 | KS: D=0.06, p=4.74e-25
AUROC (using No of ANCs completed as score): 0.466

EDA – No of PC's total visit to PW

Histogram: No of PC's total visit to PW

Tukey Boxplot

No of PC's total visit to PW

No of PC's total visit to PW by LBW

Missing: 0/37353
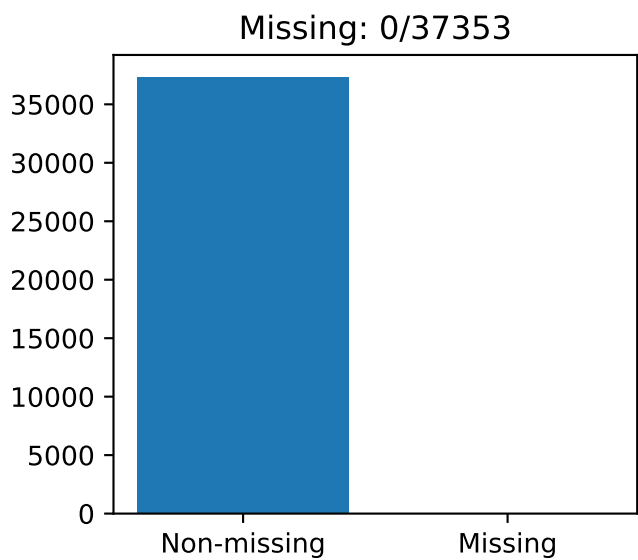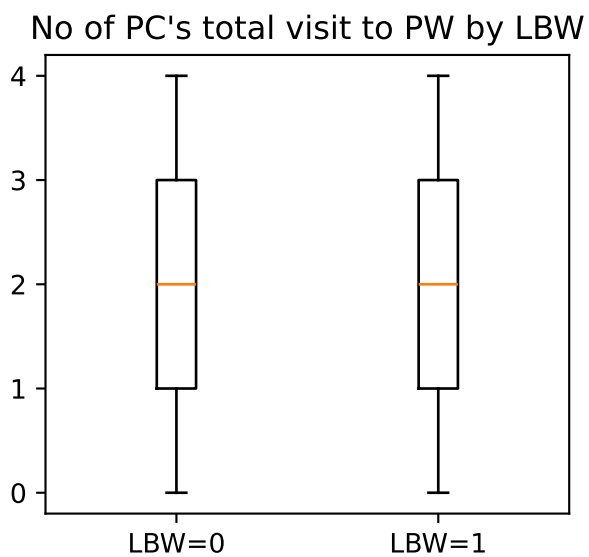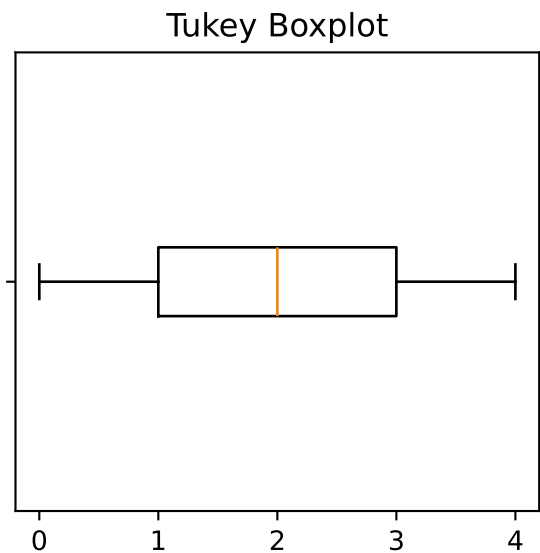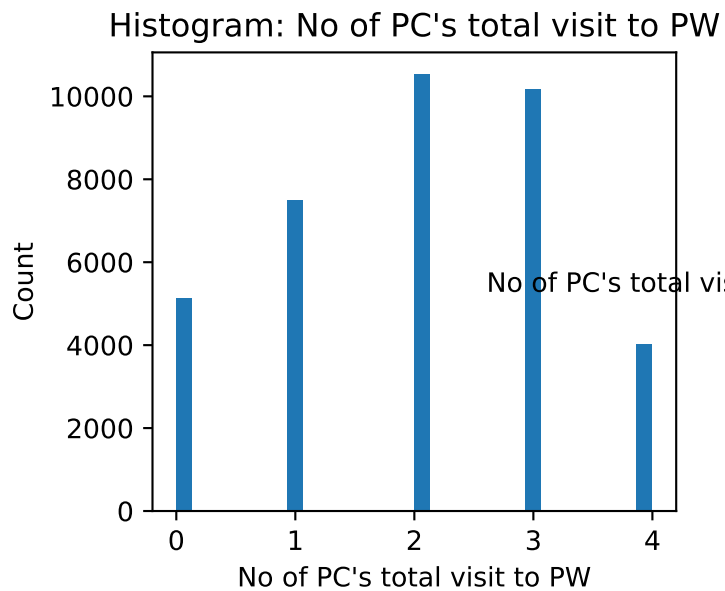
Hypothesis Tests & Metrics for: No of PC's total visit to PW (numeric)
Missing: 0.00% | Unique: 5
Welch t-test: t=6.51, p=7.88e-11
Point–Biserial r: r=-0.0334, p=1.11e-10
Cohen's d: -0.0727 | KS: D=0.0371, p=7.93e-10
AUROC (using No of PC's total visit to PW as score): 0.479
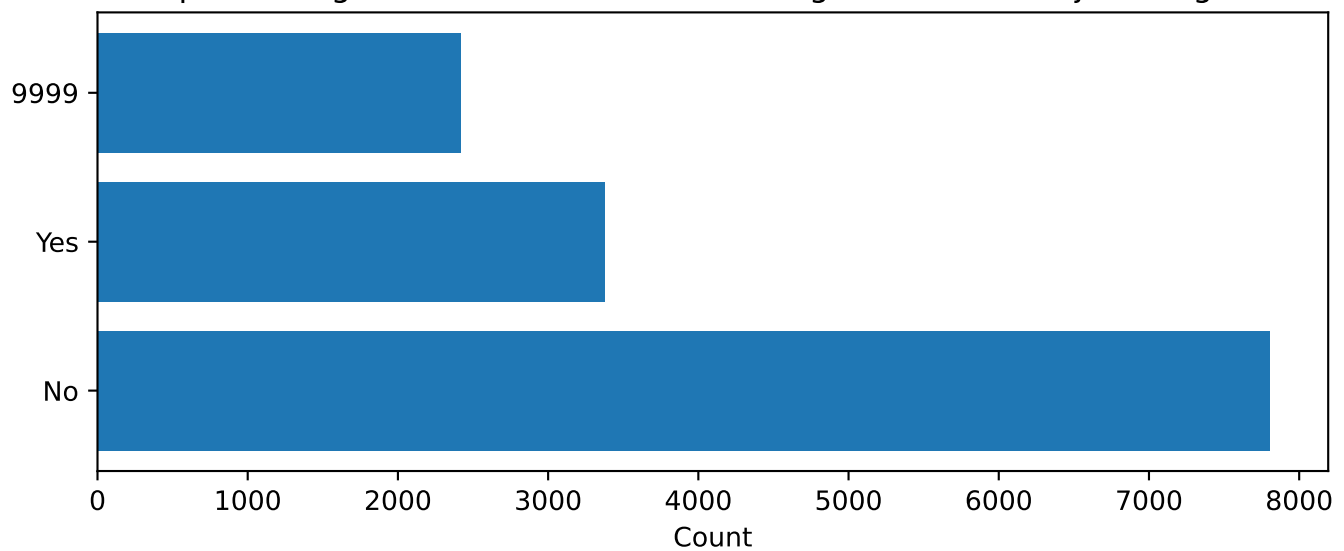
EDA – Service received during last ANC: TT Injection given

Top-15 categories: Service received during last ANC: TT Injection given

Proportions by LBW
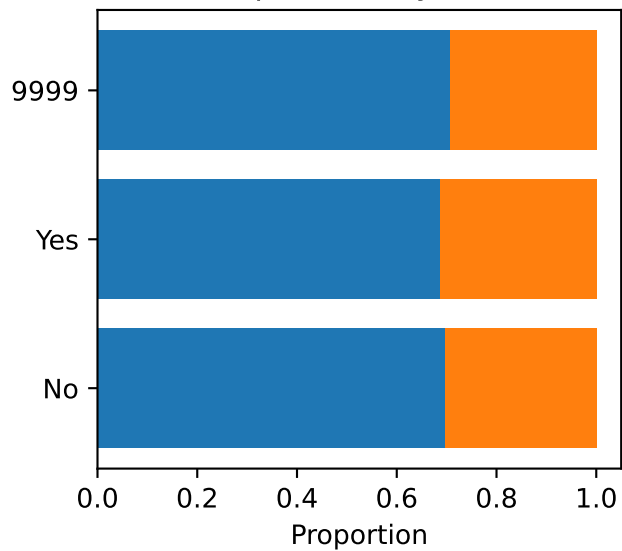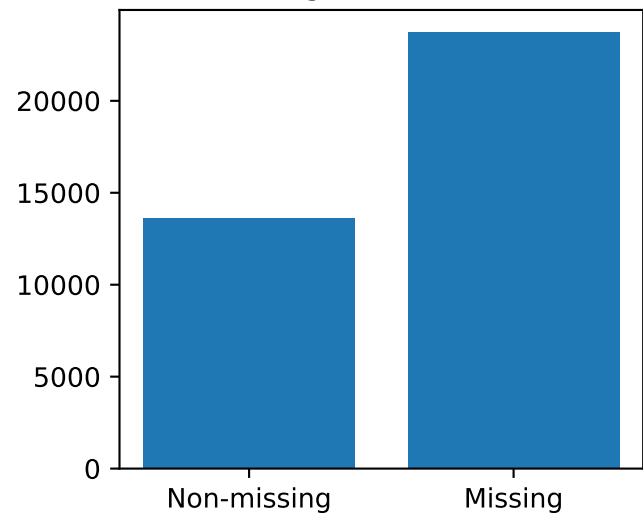
Missing: 23754/37353

Hypothesis Tests & Metrics for: Service received during last ANC: TT Injection given (categorical)
Missing: 63.59% | Unique: 3
Chi-square: chi2=2.67, dof=2, p=0.263
Cramér's V (bias-corrected): 0.007

EDA – No. of IFA tablets received/procured in last one month

ogram: No. of IFA tablets received/procured in last one month
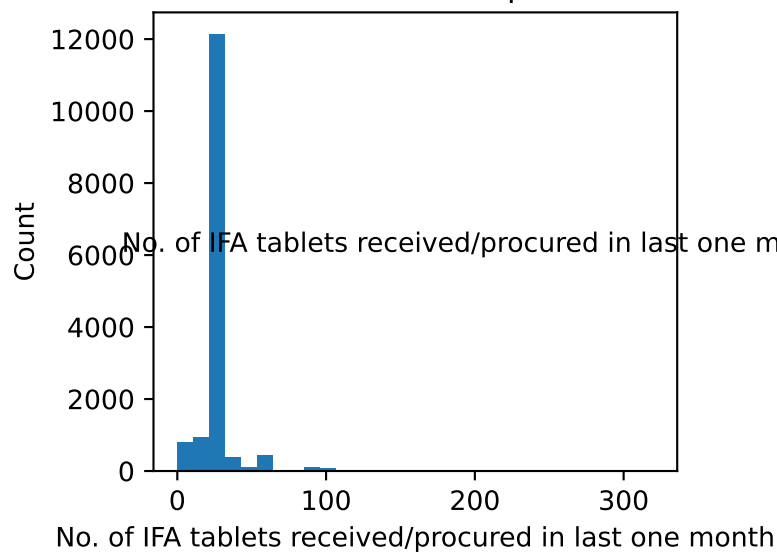
Tukey Boxplot
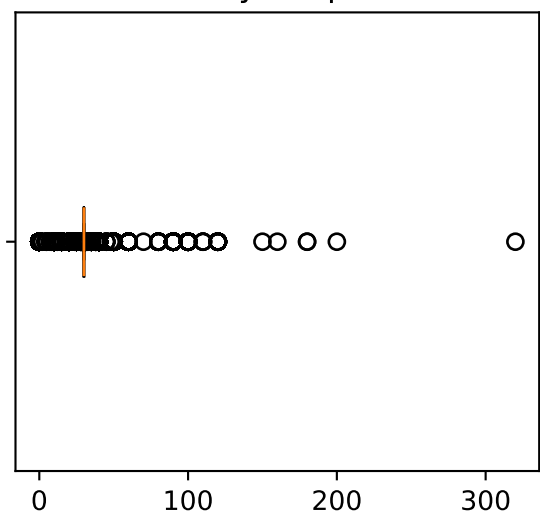
o. of IFA tablets received/procured in last one month by LBW   Missing: 22345/37353
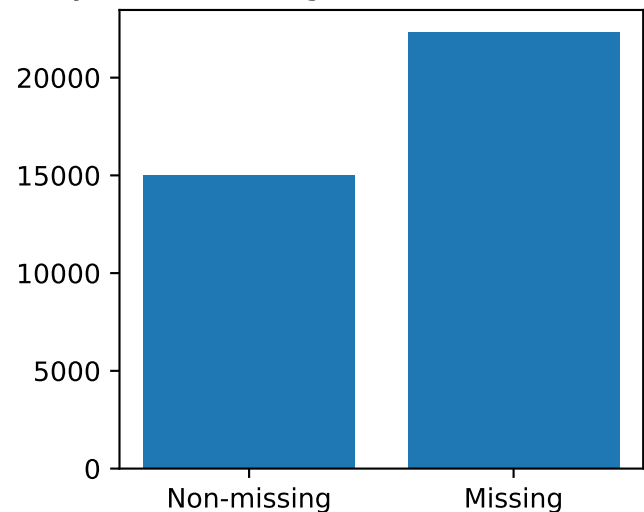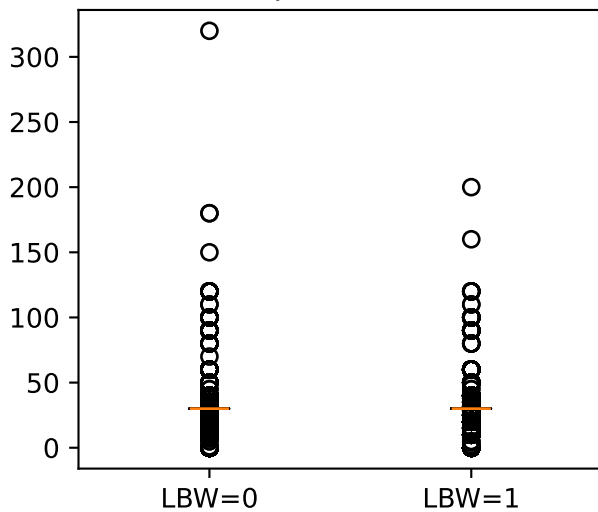
Hypothesis Tests & Metrics for: No. of IFA tablets received/procured in last one month (numeric)
Missing: 59.82% | Unique: 40
Welch t-test: t=-1.66, p=0.0964
Point–Biserial r: r=0.0136, p=0.096
Cohen's d: 0.0295 | KS: D=0.00932, p=0.943
AUROC (using No. of IFA tablets received/procured in last one month as score): 0.503

EDA – No. of calcium tablets consumed in last one month

Histogram: No. of calcium tablets consumed in last one month

No. of calcium tablets consumed in last one month
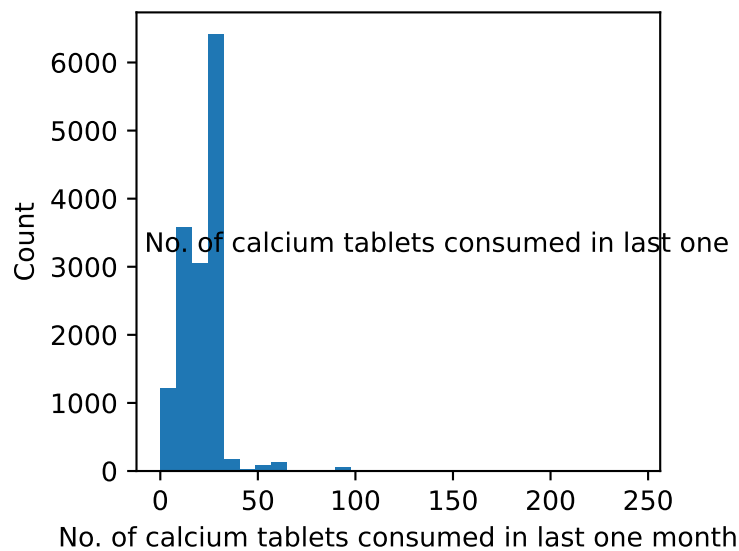
Tukey Boxplot
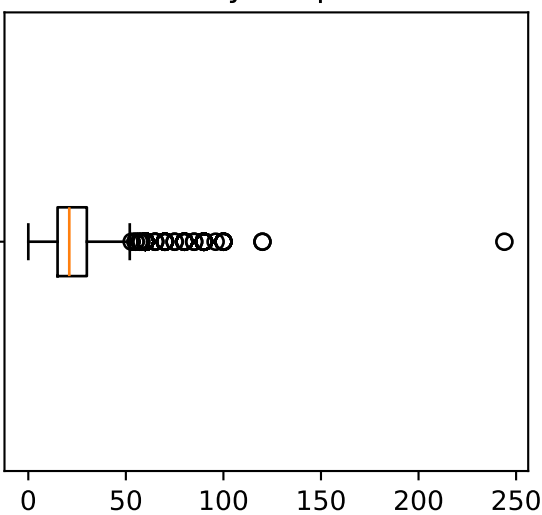
No. of calcium tablets consumed in last one month by LBW

Missing: 22618/37353
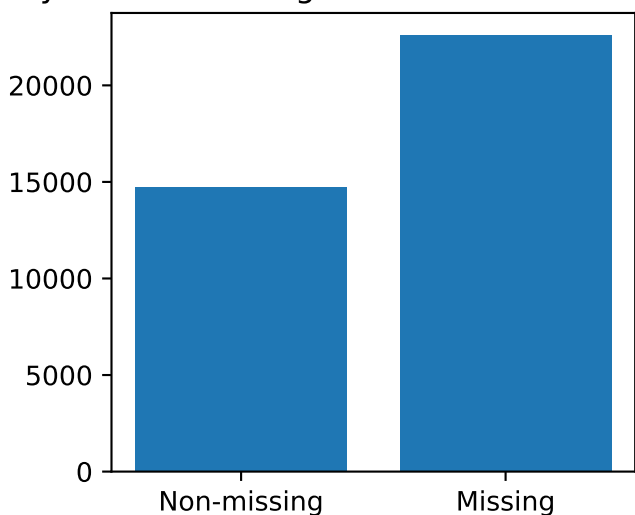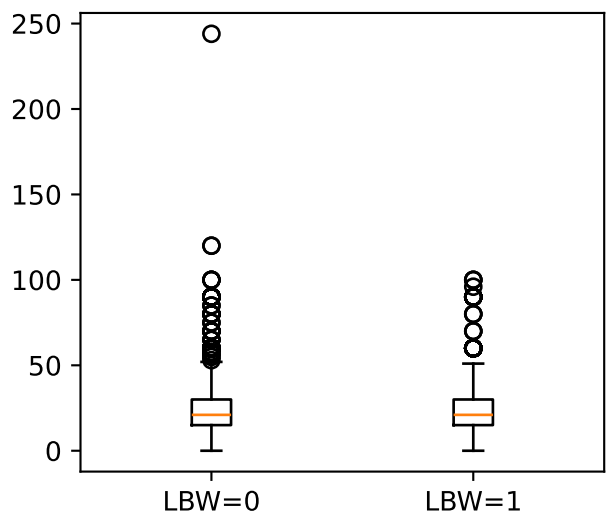
Hypothesis Tests & Metrics for: No. of calcium tablets consumed in last one month (numeric)
Missing: 60.55% | Unique: 63
Welch t-test: t=-0.238, p=0.812
Point–Biserial r: r=0.00192, p=0.816
Cohen's d: 0.00418 | KS: D=0.0124, p=0.723
AUROC (using No. of calcium tablets consumed in last one month as score): 0.501

Top-15 categories: Registered for cash transfer scheme: JSY
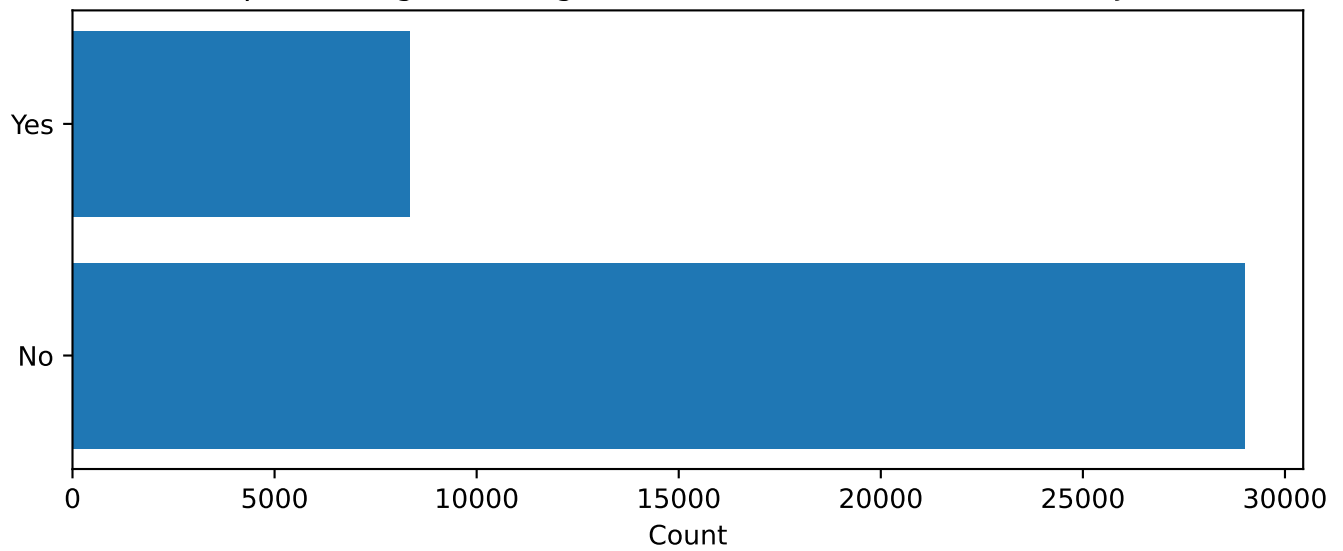
Proportions by LBW
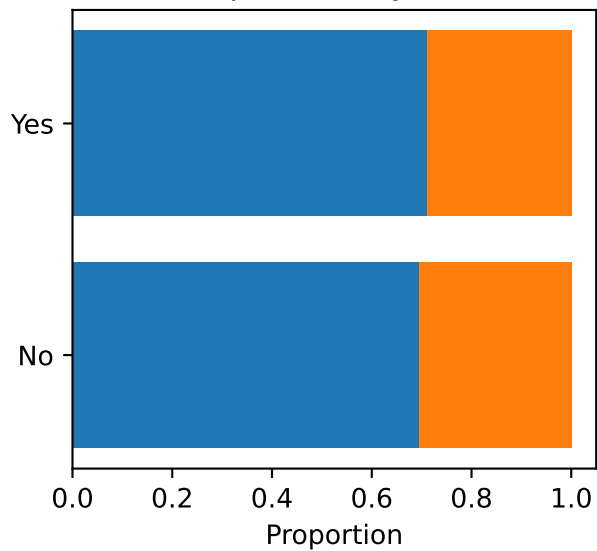
Missing: 1/37353
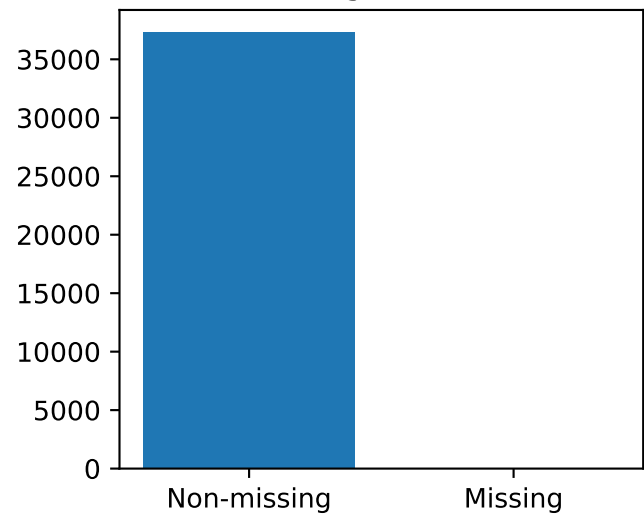
Hypothesis Tests & Metrics for: Registered for cash transfer scheme: JSY (categorical)
Missing: 0.00% | Unique: 2
Chi-square: chi2=8.6, dof=1, p=0.00336
Cramér's V (bias-corrected): 0.014

Top-15 categories: Registered for cash transfer scheme: RAJHSRI
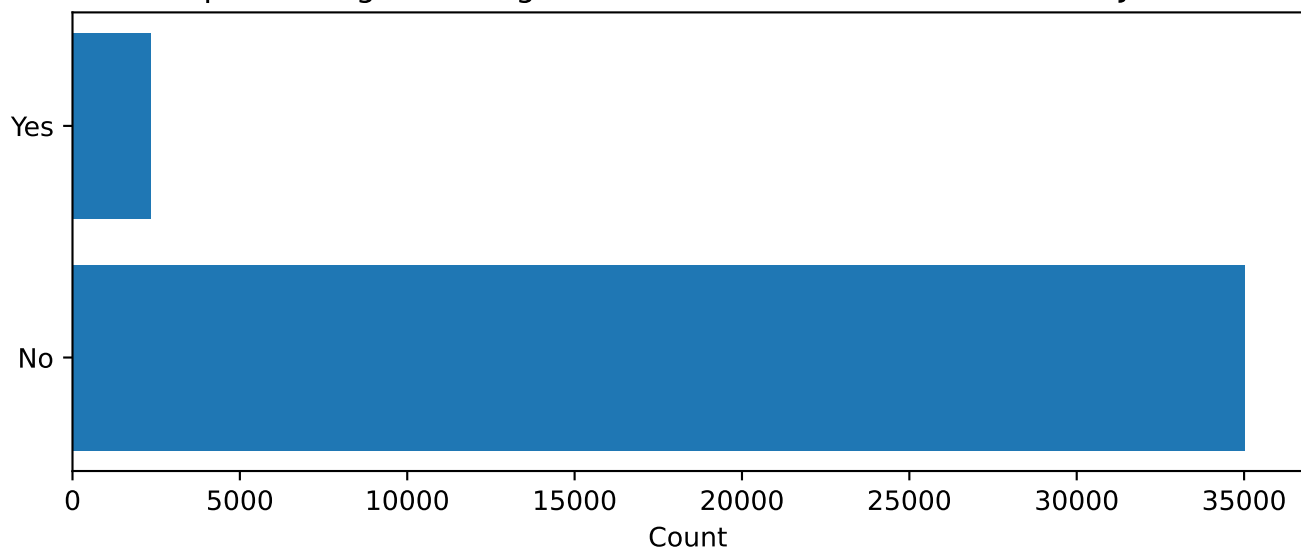
Proportions by LBW
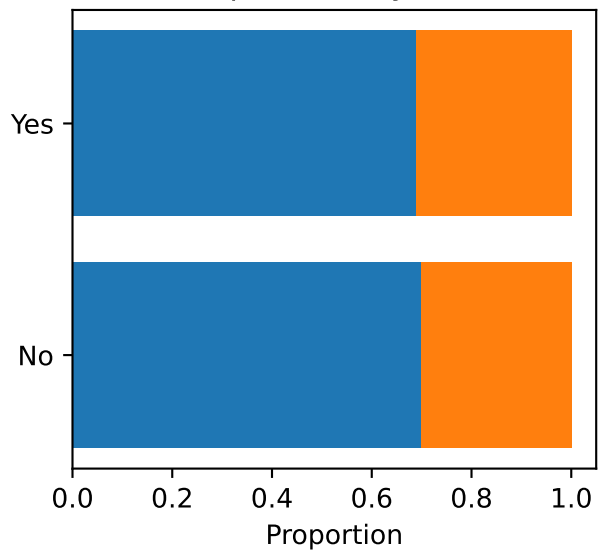
Missing: 1/37353
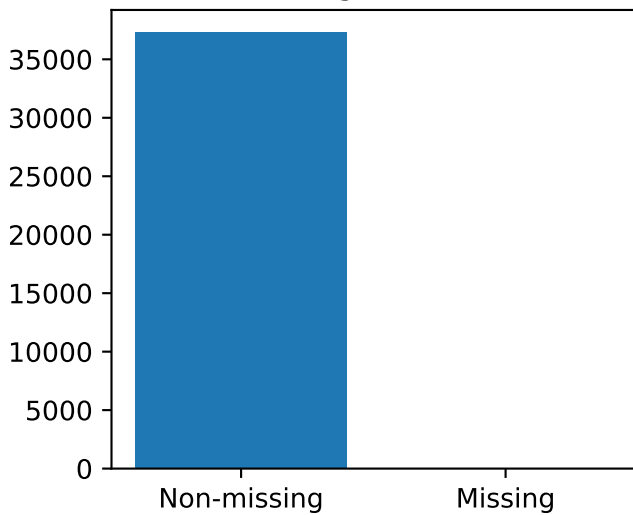
Hypothesis Tests & Metrics for: Registered for cash transfer scheme: RAJHSRI (categorical)
Missing: 0.00% | Unique: 2
Chi-square: chi2=1.02, dof=1, p=0.312
Cramér's V (bias-corrected): 0.001

EDA – JSY-Number of installment received

Histogram: JSY-Number of installment received
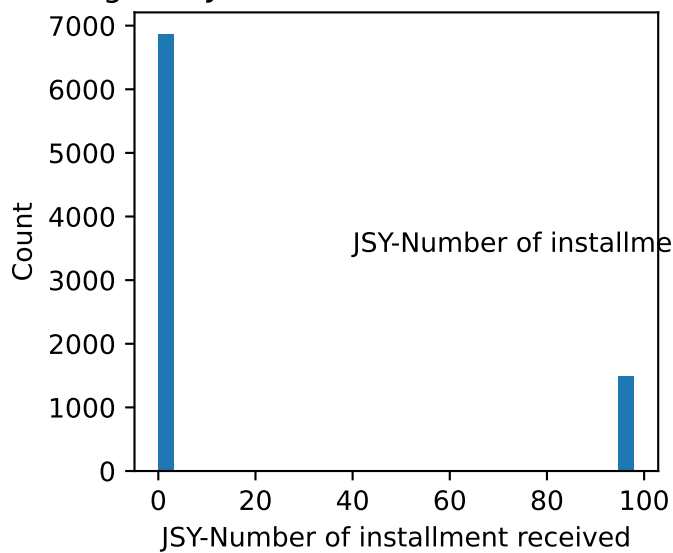
Tukey Boxplot
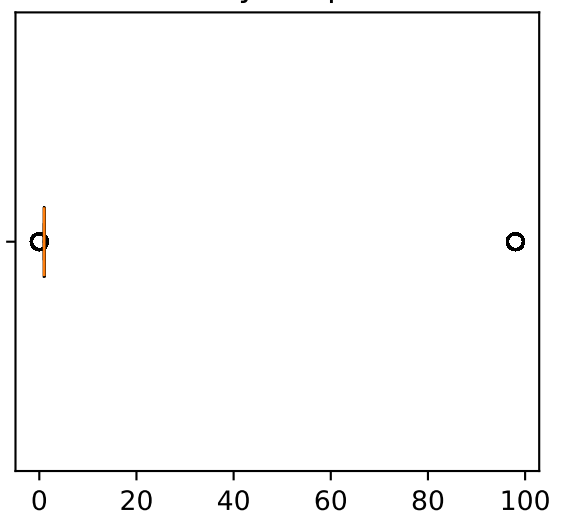
JSY-Number of installment received by LBW
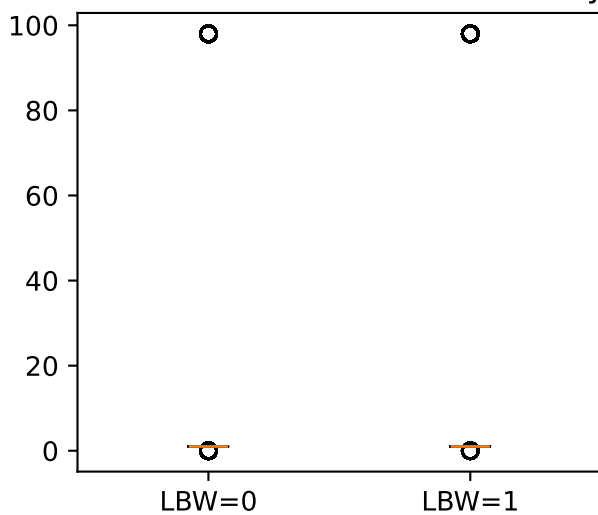
Missing: 29002/37353
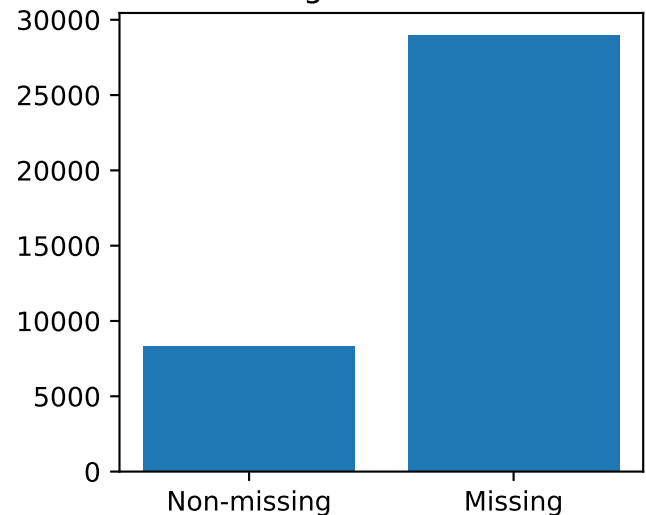
Hypothesis Tests & Metrics for: JSY-Number of installment received (numeric)
Missing: 77.64% | Unique: 3
Welch t-test: t=-0.482, p=0.63
Point–Biserial r: r=0.0053, p=0.628
Cohen's d: 0.0117 | KS: D=0.00444, p=1
AUROC (using JSY-Number of installment received as score): 0.503

Top-15 categories: PMMVY-Number of installment received
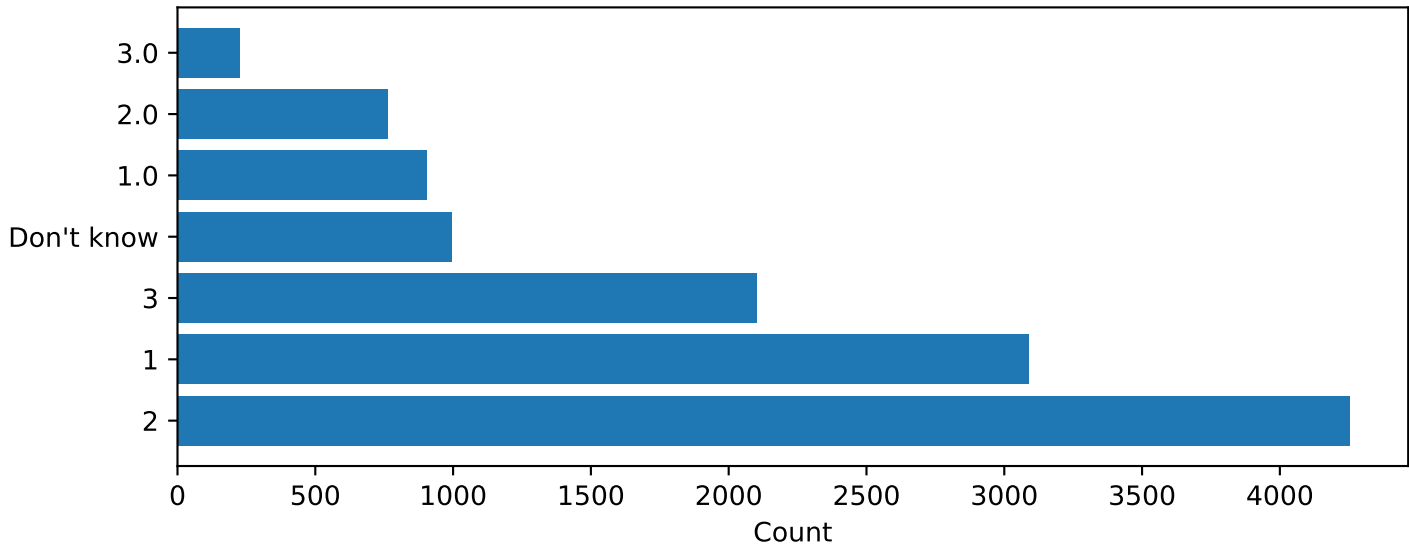
Proportions by LBW
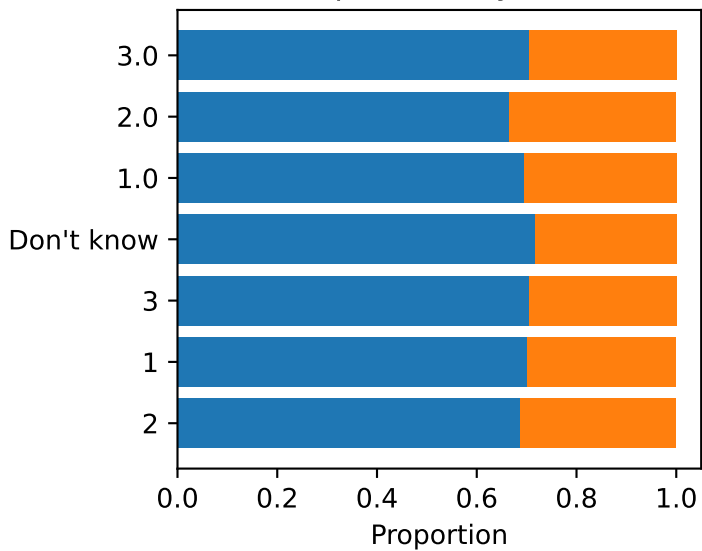
Missing: 25010/37353
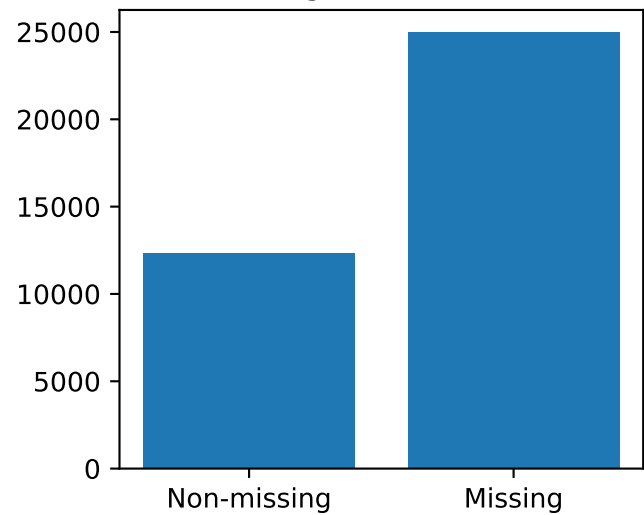
Hypothesis Tests & Metrics for: PMMVY-Number of installment received (categorical)
Missing: 66.96% | Unique: 7
Chi-square: chi2=8.14, dof=6, p=0.228
Cramér's V (bias-corrected): 0.013

EDA – LMPtoRegistration

Histogram: LMPtoRegistration
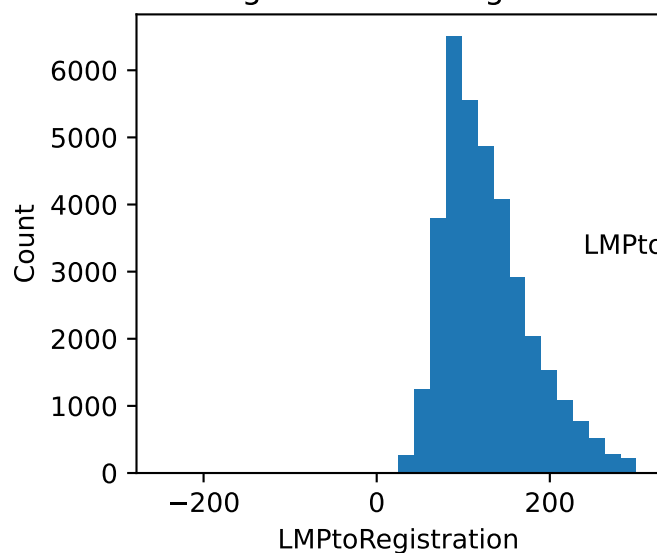
Tukey Boxplot
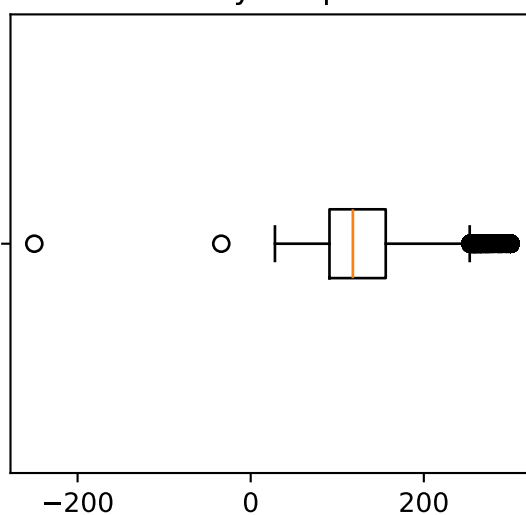
LMPtoRegistration by LBW
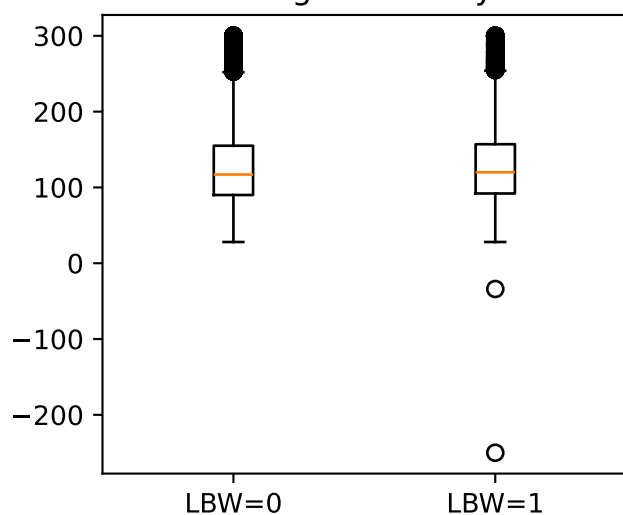
Missing: 1731/37353
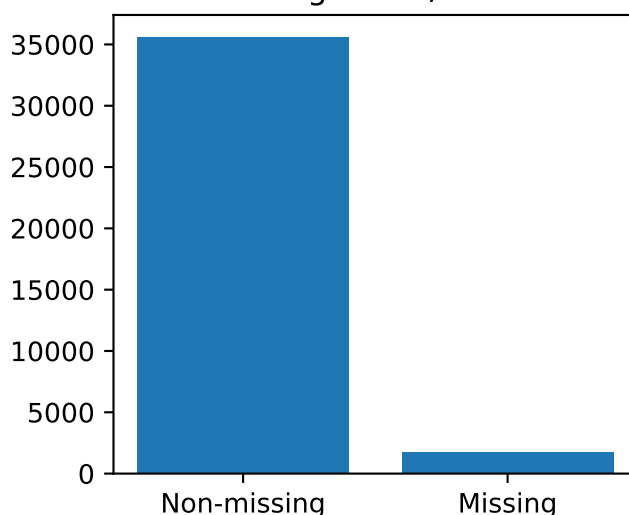
Hypothesis Tests & Metrics for: LMPtoRegistration (numeric)
Missing: 4.63% | Unique: 275
Welch t-test: t=-2.97, p=0.00299
Point–Biserial r: r=0.0157, p=0.00302
Cohen's d: 0.0342 | KS: D=0.0274, p=2.57e-05
AUROC (using LMPtoRegistration as score): 0.513

EDA – LMPtoPW1

Histogram: LMPtoPW1
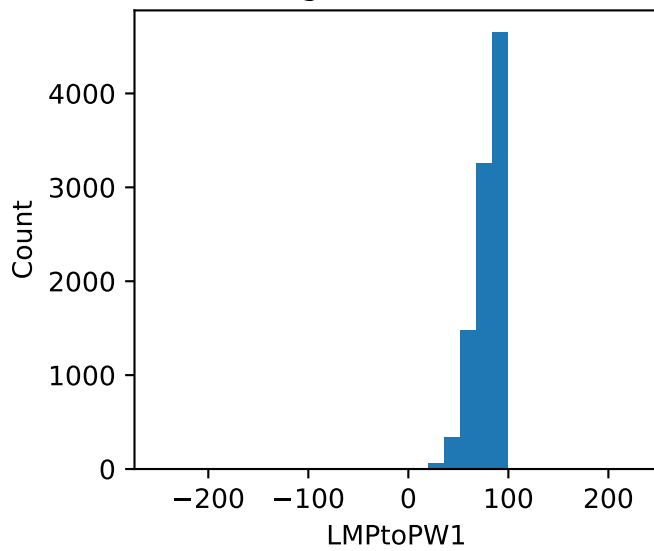
Tukey Boxplot
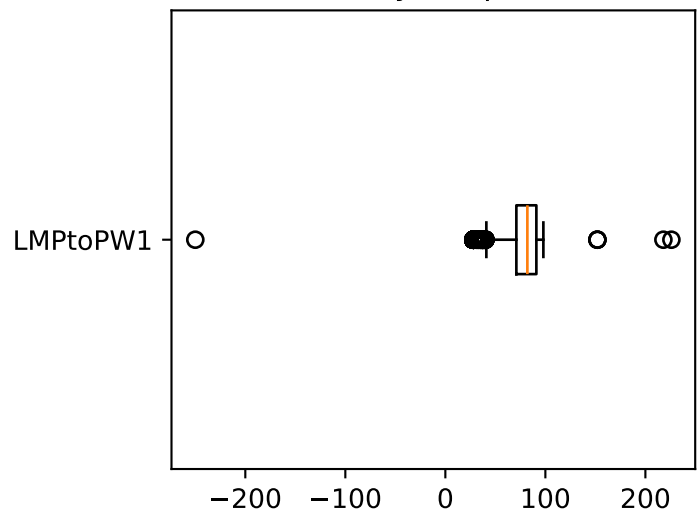
LMPtoPW1 by LBW
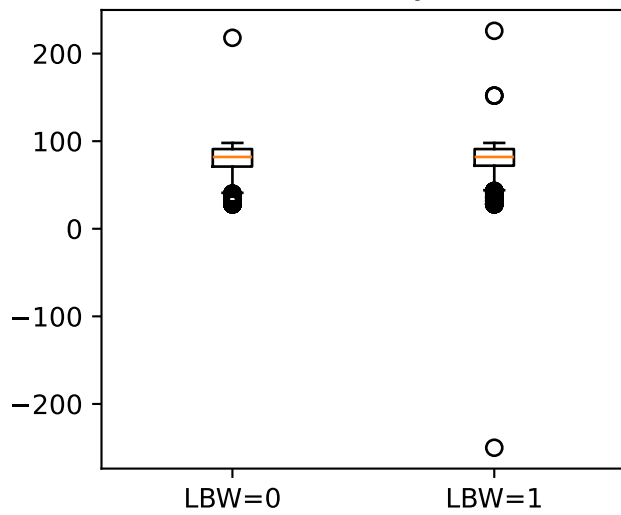
Missing: 27559/37353
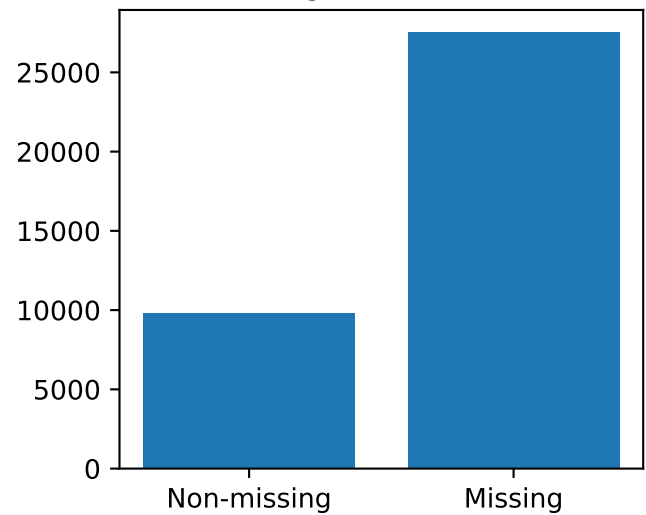
Hypothesis Tests & Metrics for: LMPtoPW1 (numeric)
Missing: 73.78% | Unique: 75
Welch t-test: t=-0.833, p=0.405
Point–Biserial r: r=0.00877, p=0.385
Cohen's d: 0.0193 | KS: D=0.0261, p=0.125
AUROC (using LMPtoPW1 as score): 0.505

EDA – LMPtoPW2

Histogram: LMPtoPW2

Tukey Boxplot

LMPtoPW2 by LBW

Missing: 10882/37353
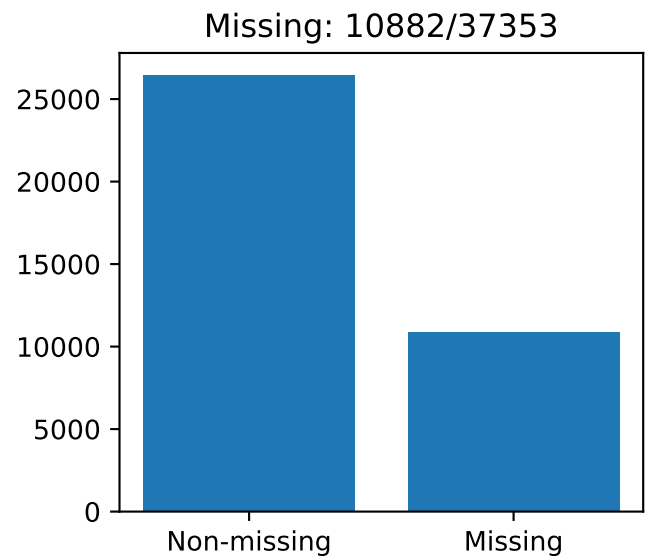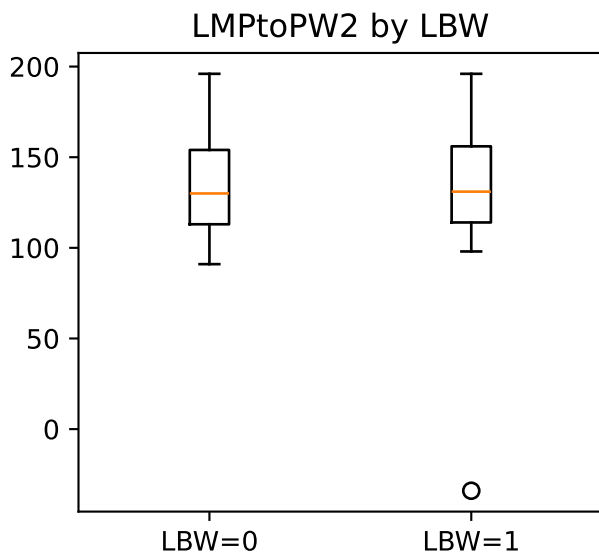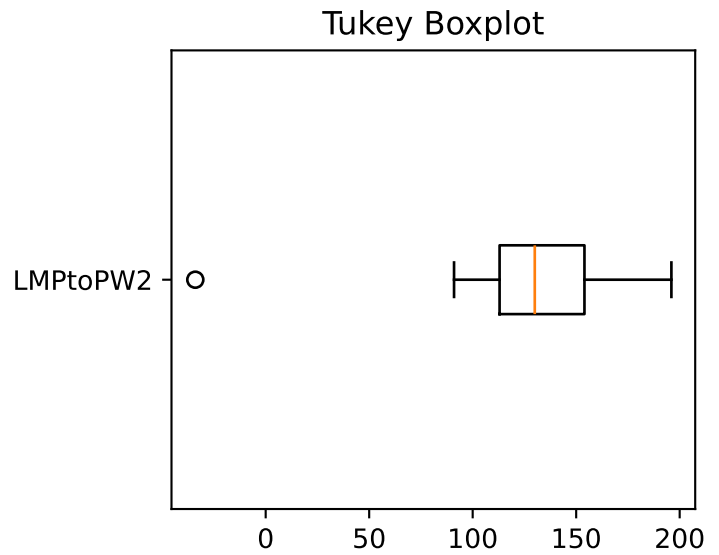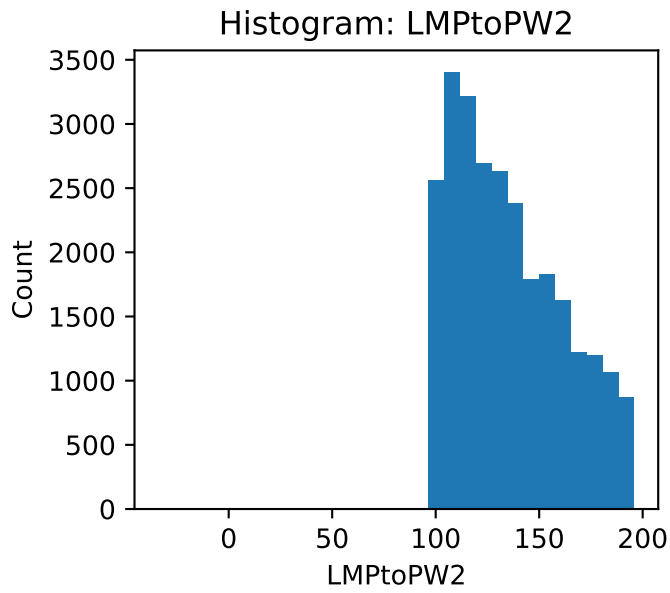
Hypothesis Tests & Metrics for: LMPtoPW2 (numeric)
Missing: 29.13% | Unique: 101
Welch t-test: t=-3.07, p=0.00213
Point–Biserial r: r=0.019, p=0.00198
Cohen's d: 0.0415 | KS: D=0.02, p=0.0231
AUROC (using LMPtoPW2 as score): 0.512

EDA – LMPtoPW3

Histogram: LMPtoPW3

Tukey Boxplot

LMPtoPW3 by LBW

Missing: 13382/37353
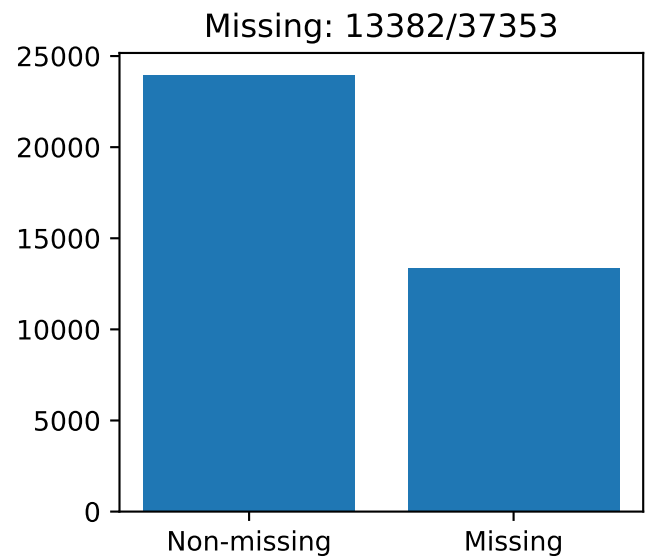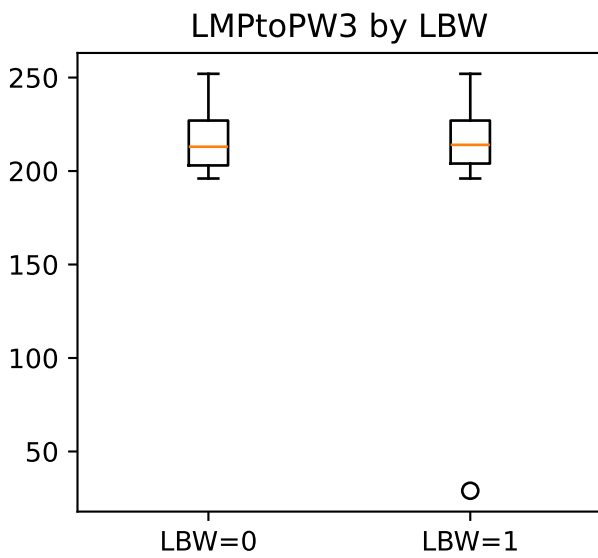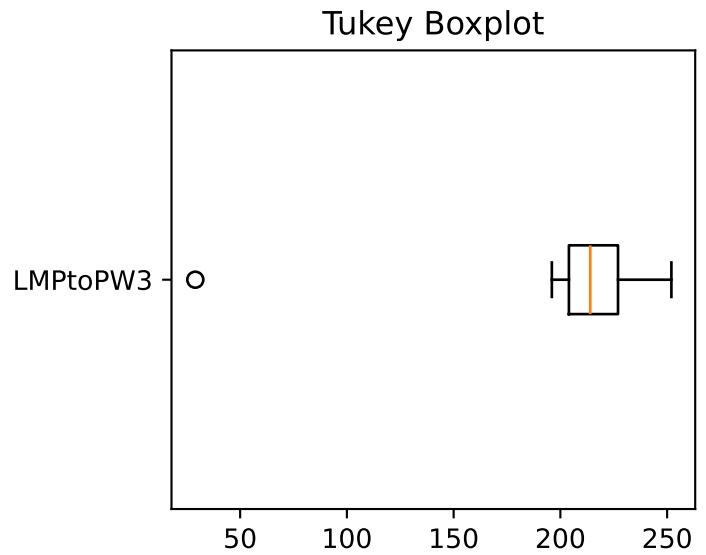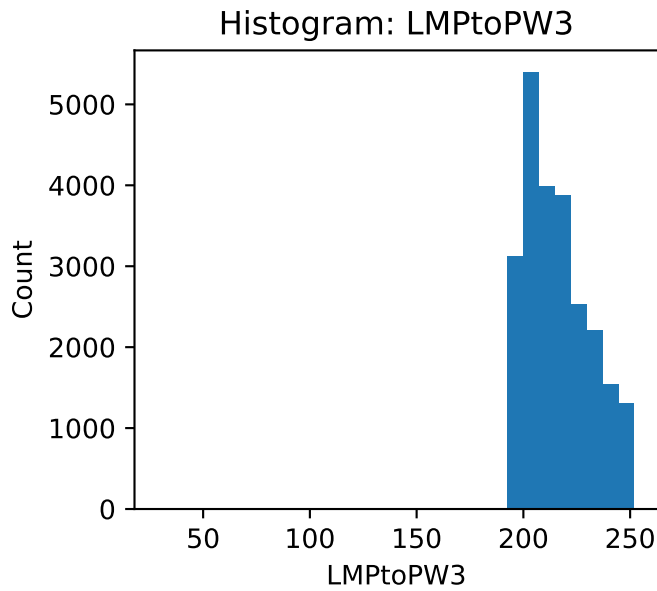
Hypothesis Tests & Metrics for: LMPtoPW3 (numeric)
Missing: 35.83% | Unique: 58
Welch t-test: t=-2.28, p=0.0227
Point–Biserial r: r=0.0148, p=0.0216
Cohen's d: 0.0325 | KS: D=0.0219, p=0.0159
AUROC (using LMPtoPW3 as score): 0.510

EDA – LMPtoPW4

Histogram: LMPtoPW4

Tukey Boxplot

LMPtoPW4 by LBW

Missing: 22423/37353
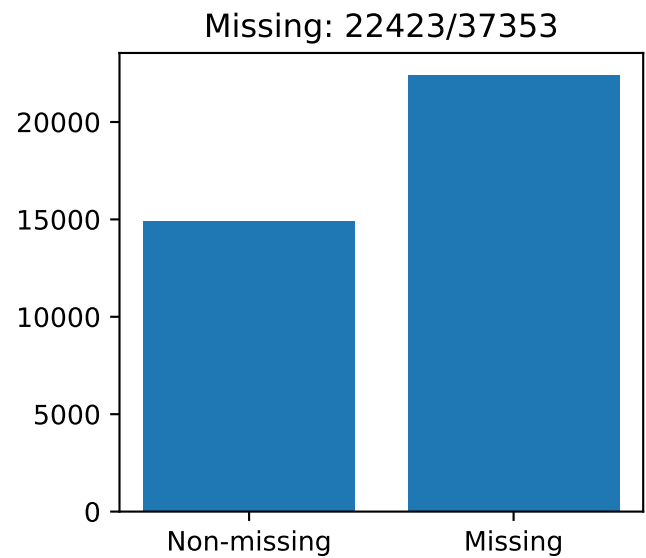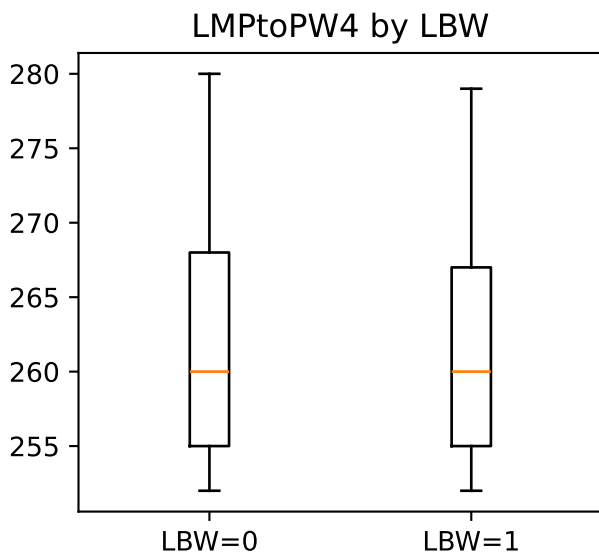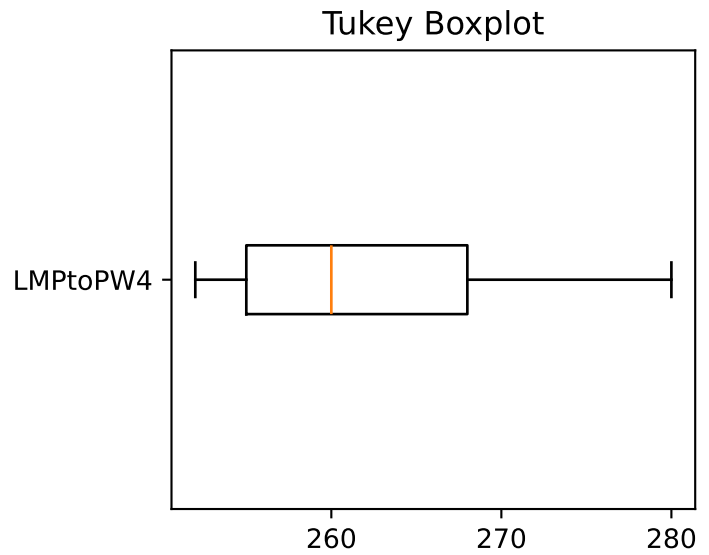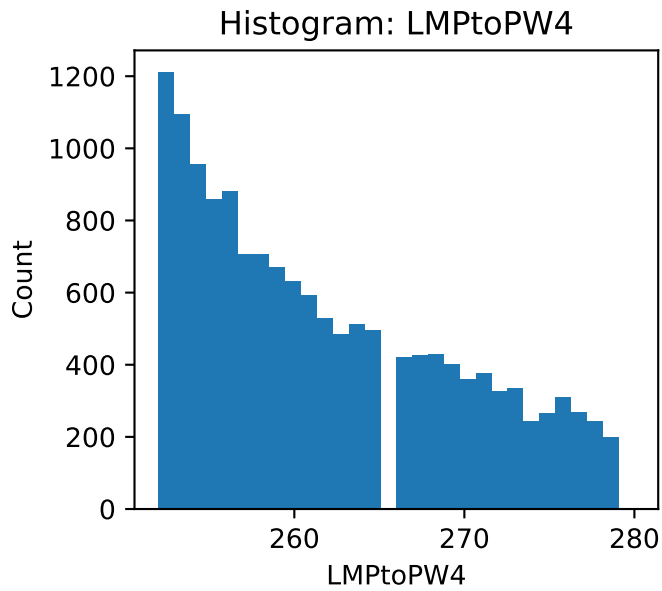
Hypothesis Tests & Metrics for: LMPtoPW4 (numeric)
Missing: 60.03% | Unique: 29
Welch t-test: t=1.82, p=0.0686
Point–Biserial r: r=-0.0148, p=0.0703
Cohen's d: -0.0331 | KS: D=0.0353, p=0.00114
AUROC (using LMPtoPW4 as score): 0.492

EDA – LMPtoINST1

Histogram: LMPtoINST1
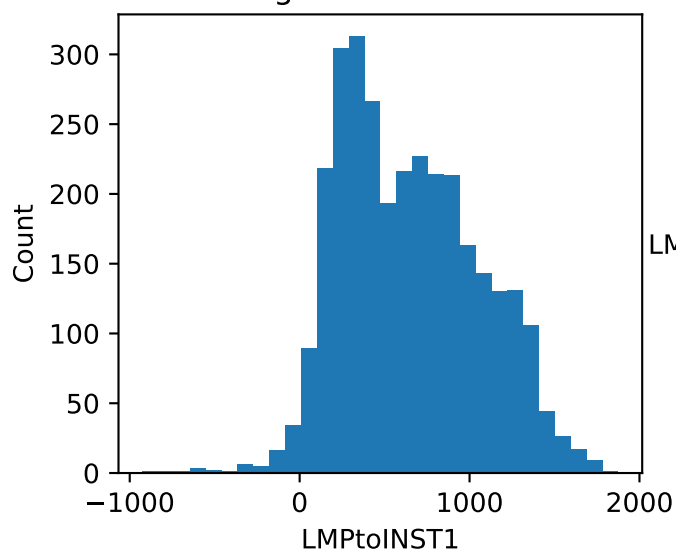
Tukey Boxplot
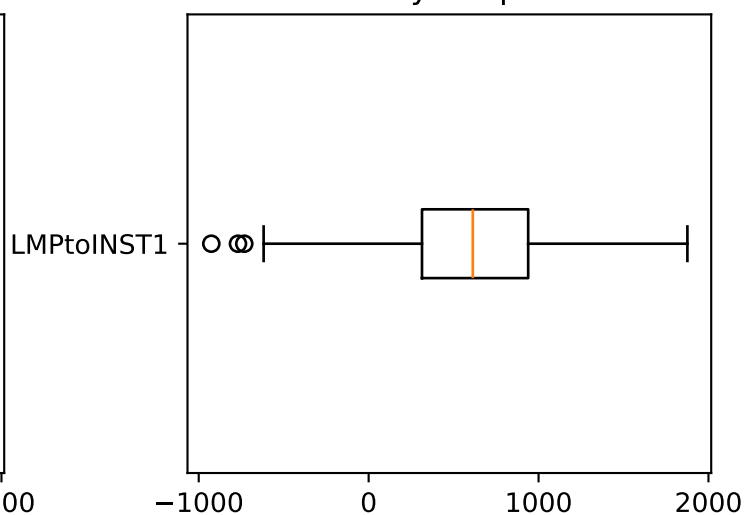
LMPtoINST1 by LBW
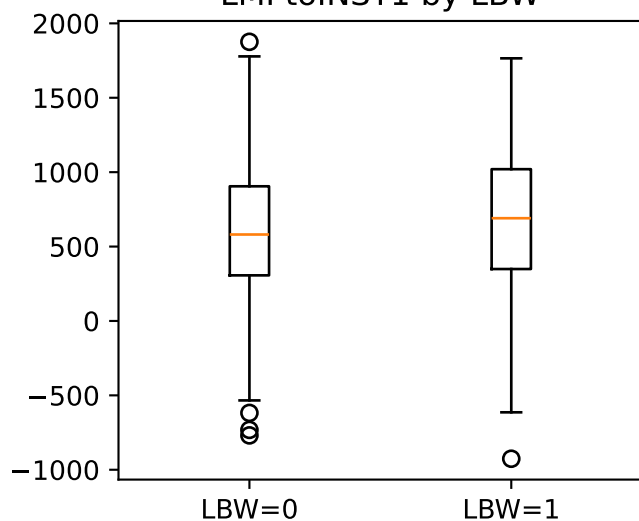
Missing: 34260/37353
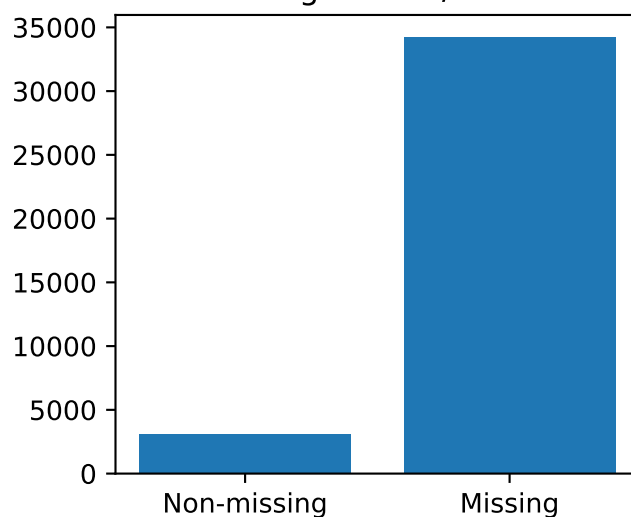
Hypothesis Tests & Metrics for: LMPtoINST1 (numeric)
Missing: 91.72% | Unique: 1194
Welch t-test: t=-4.61, p=4.44e-06
Point–Biserial r: r=0.0852, p=2.08e-06
Cohen's d: 0.186 | KS: D=0.11, p=2.66e-07
AUROC (using LMPtoINST1 as score): 0.551

EDA – LMPtoINST2

Histogram: LMPtoINST2

Tukey Boxplot

LMPtoINST2 by LBW

Missing: 34566/37353
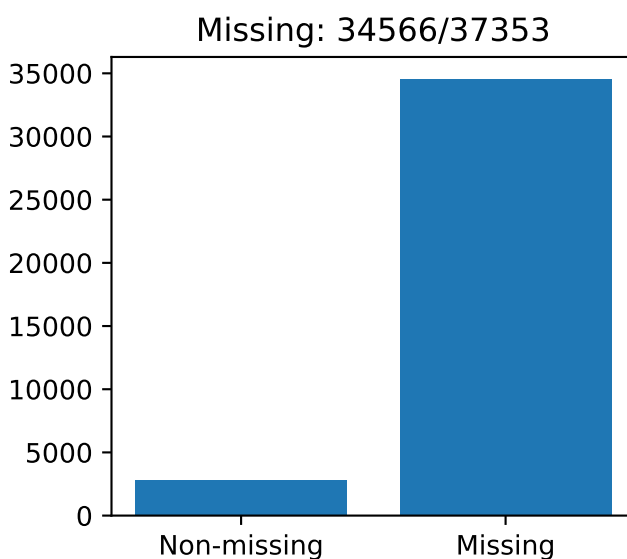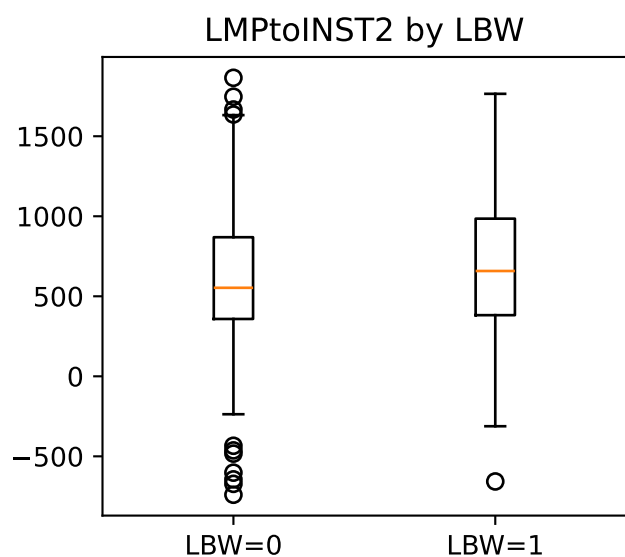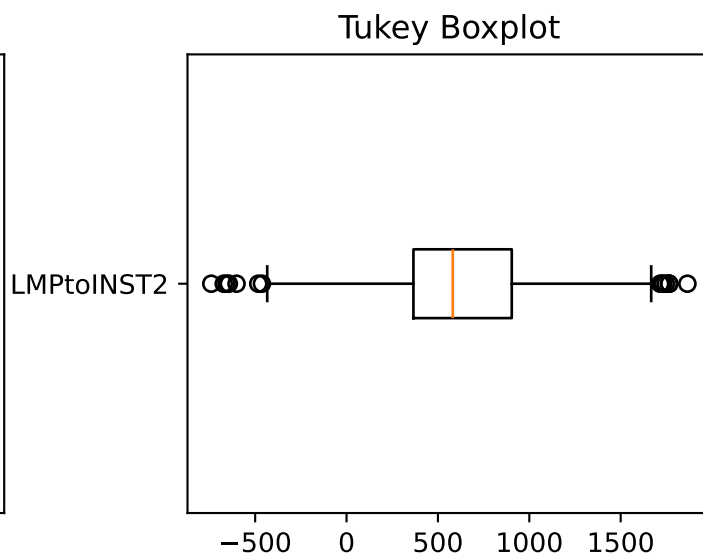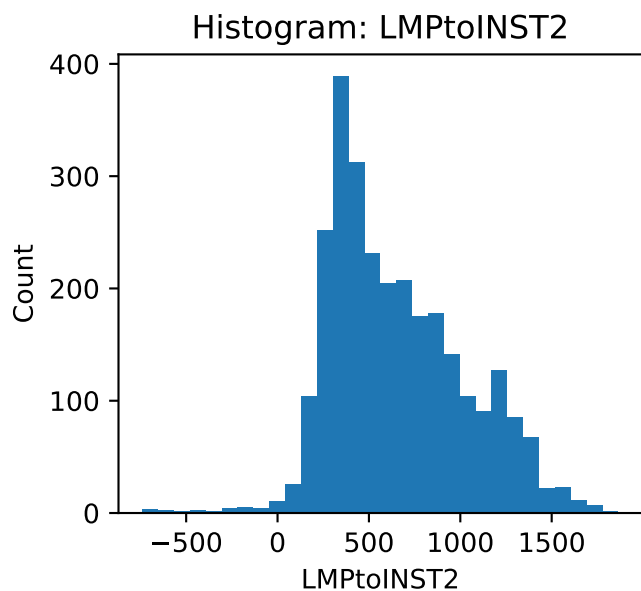
Hypothesis Tests & Metrics for: LMPtoINST2 (numeric)
Missing: 92.54% | Unique: 1033
Welch t-test: t=-4.64, p=3.87e-06
Point–Biserial r: r=0.0905, p=1.69e-06
Cohen's d: 0.199 | KS: D=0.125, p=2.22e-08
AUROC (using LMPtoINST2 as score): 0.553

Histogram: LMPtoINST3

Tukey Boxplot

LMPtoINST3 by LBW

Missing: 35419/37353
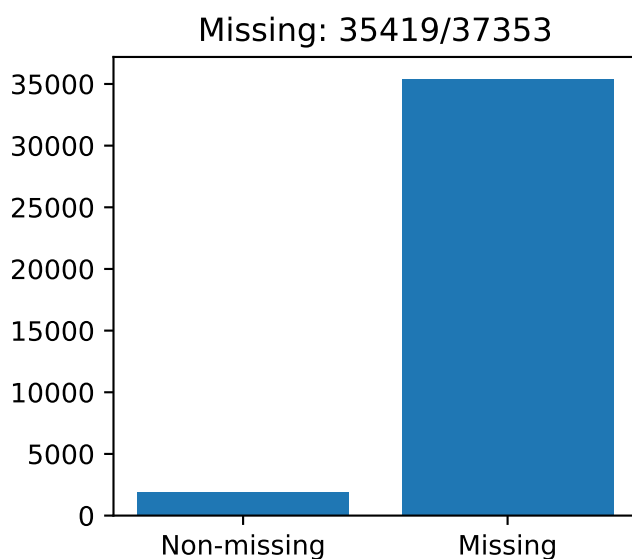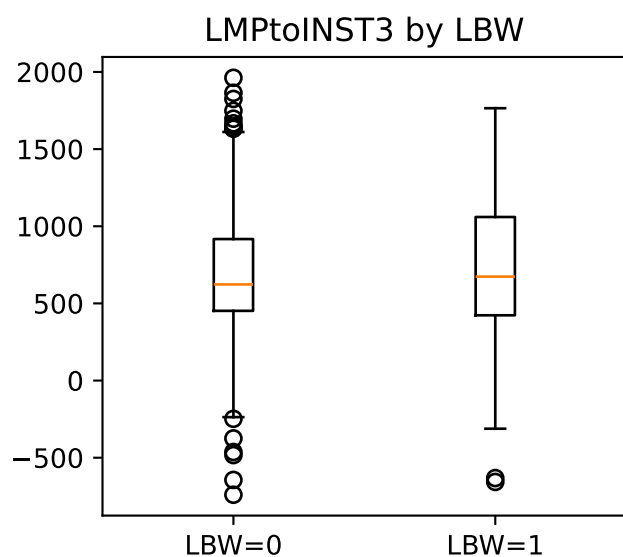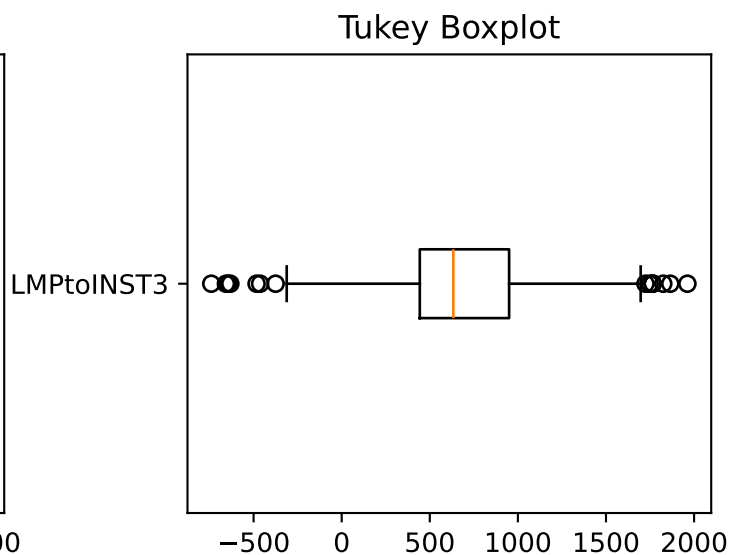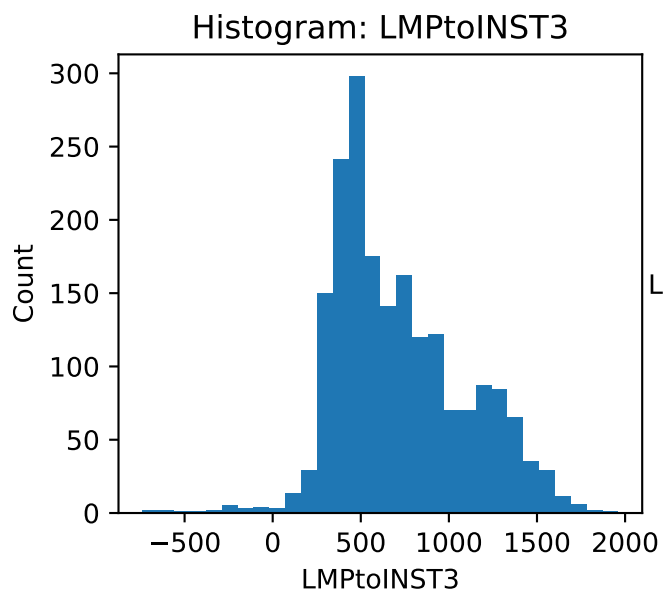
Hypothesis Tests & Metrics for: LMPtoINST3 (numeric)
Missing: 94.82% | Unique: 871
Welch t-test: t=-1.73, p=0.0846
Point–Biserial r: r=0.041, p=0.0717
Cohen's d: 0.0912 | KS: D=0.0987, p=0.000913
AUROC (using LMPtoINST3 as score): 0.518

EDA – LMPtoINST4

Histogram: LMPtoINST4

Tukey Boxplot

LMPtoINST4 by LBW

Missing: 37351/37353
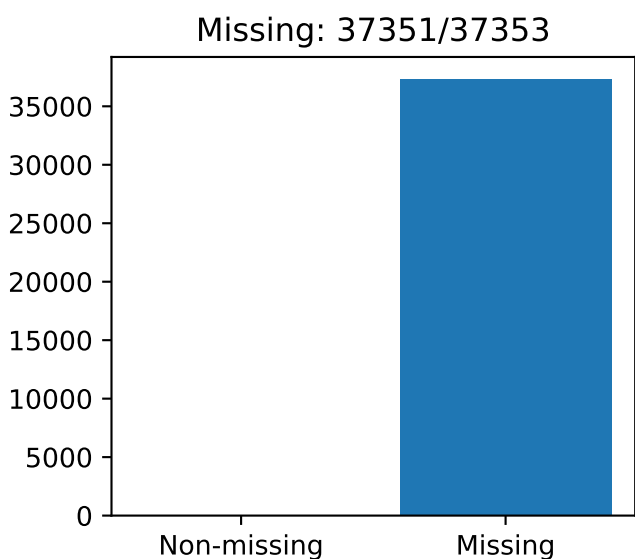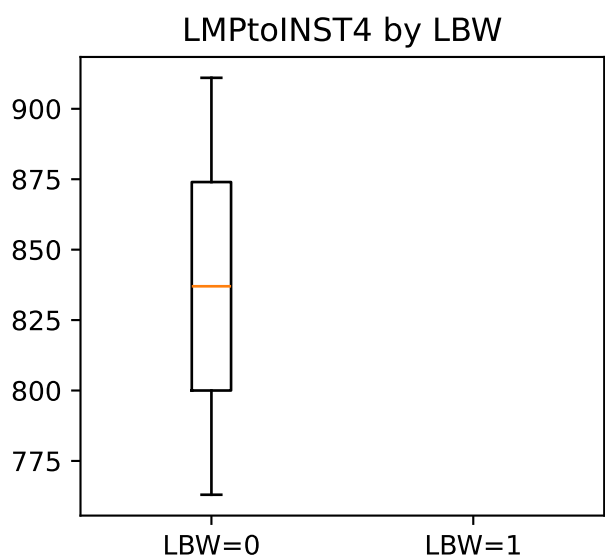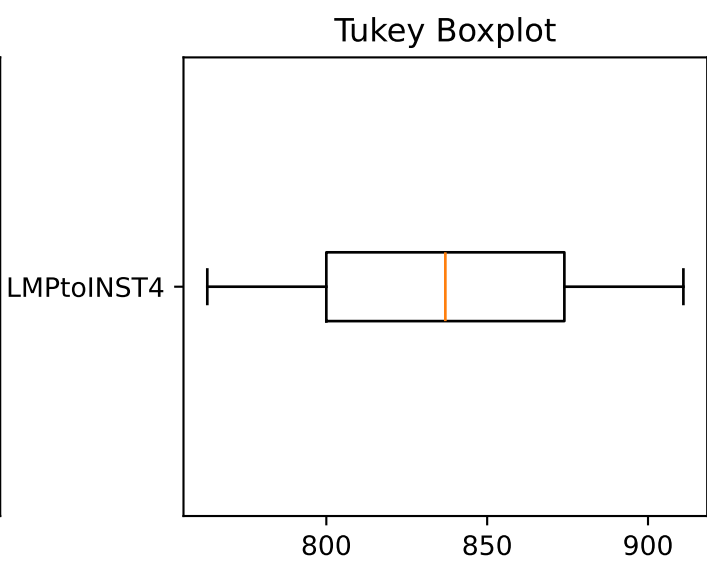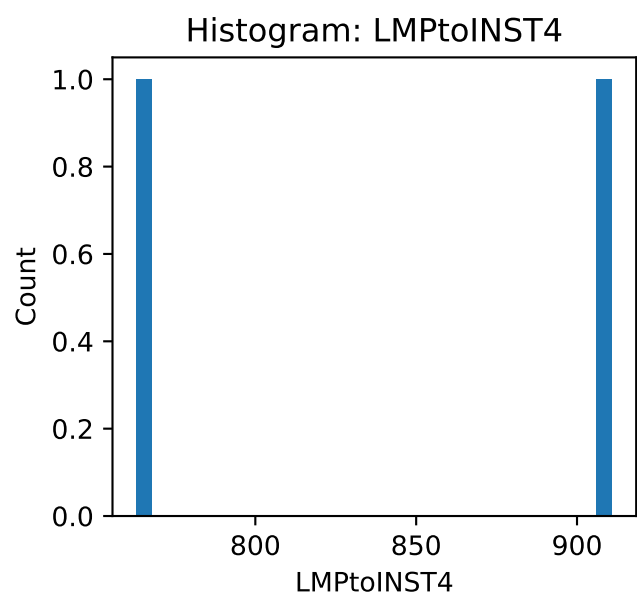
Hypothesis Tests & Metrics for: LMPtoINST4 (numeric)
Missing: 99.99% | Unique: 2
Welch t-test: t=nan, p=nan
Point–Biserial r: r=nan, p=nan
Cohen's d: nan | KS: D=nan, p=nan

Histogram: LMPtoINST5

Tukey Boxplot

LMPtoINST5 by LBW

Missing: 37256/37353
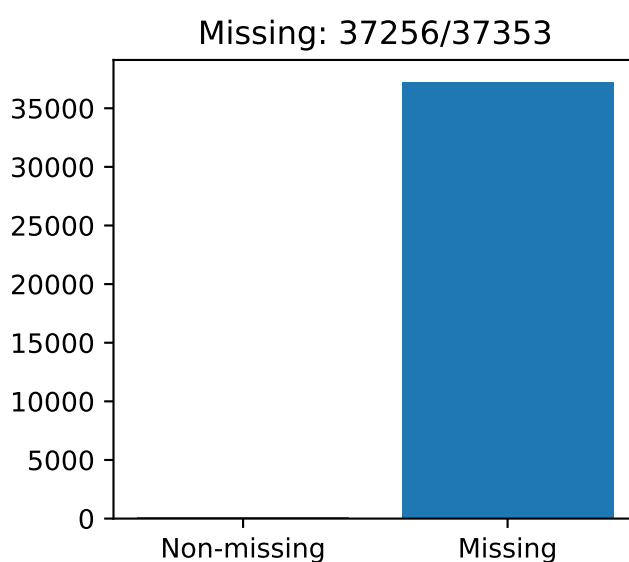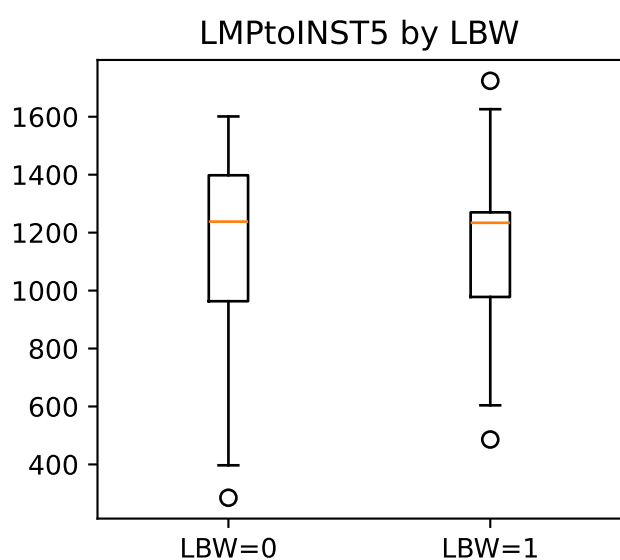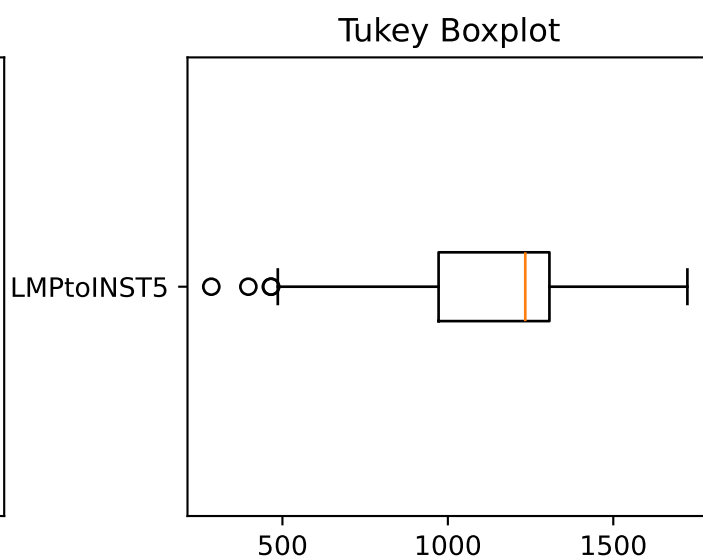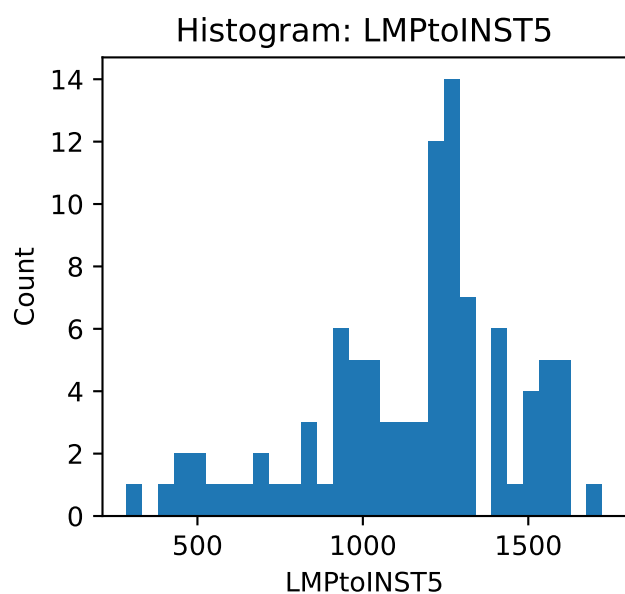
Hypothesis Tests & Metrics for: LMPtoINST5 (numeric)
Missing: 99.74% | Unique: 74
Welch t-test: t=0.395, p=0.694
Point–Biserial r: r=-0.0376, p=0.714
Cohen's d: -0.0852 | KS: D=0.212, p=0.32
AUROC (using LMPtoINST5 as score): 0.462

Top-15 categories: RegistrationBucket
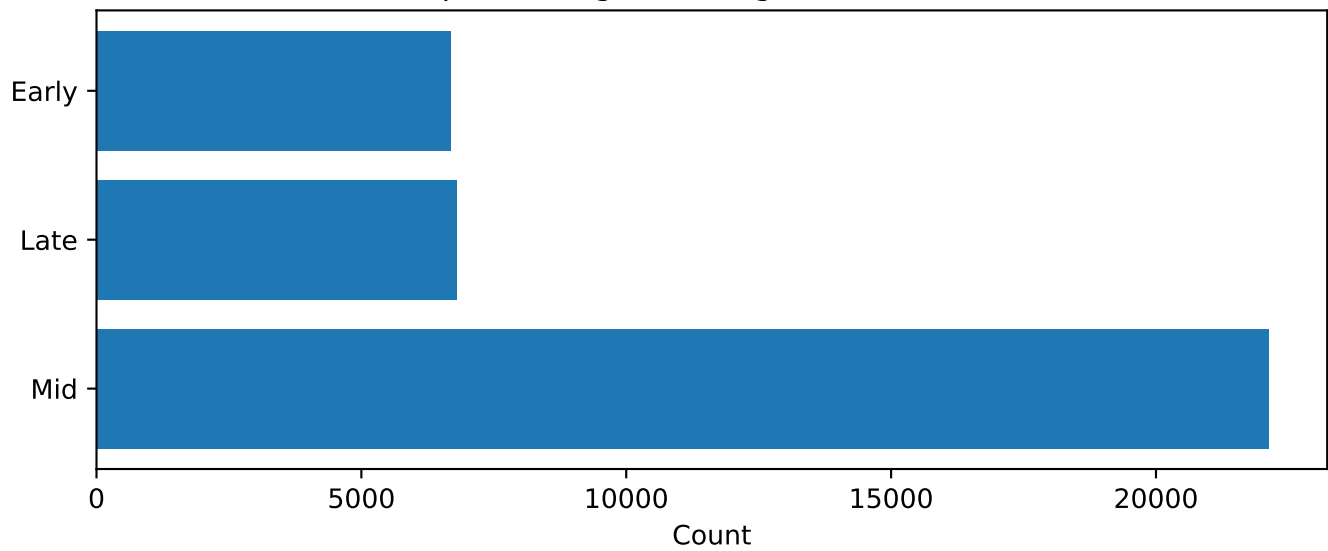
Proportions by LBW
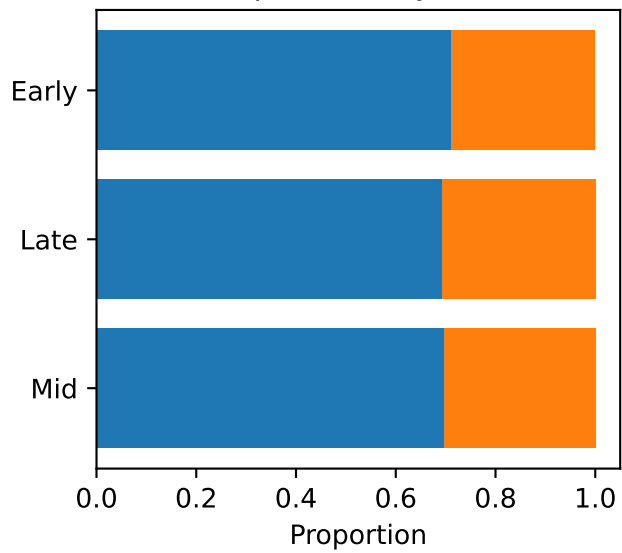
Missing: 1733/37353
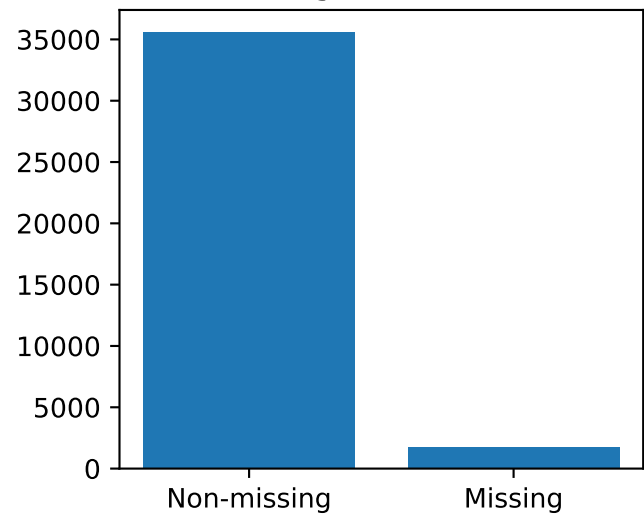
Hypothesis Tests & Metrics for: RegistrationBucket (categorical)
Missing: 4.64% | Unique: 3
Chi-square: chi2=6.89, dof=2, p=0.0319
Cramér's V (bias-corrected): 0.012

EDA – LMPtoEarliestANC

Histogram: LMPtoEarliestANC

Tukey Boxplot

LMPtoEarliestANC by LBW

Missing: 5132/37353
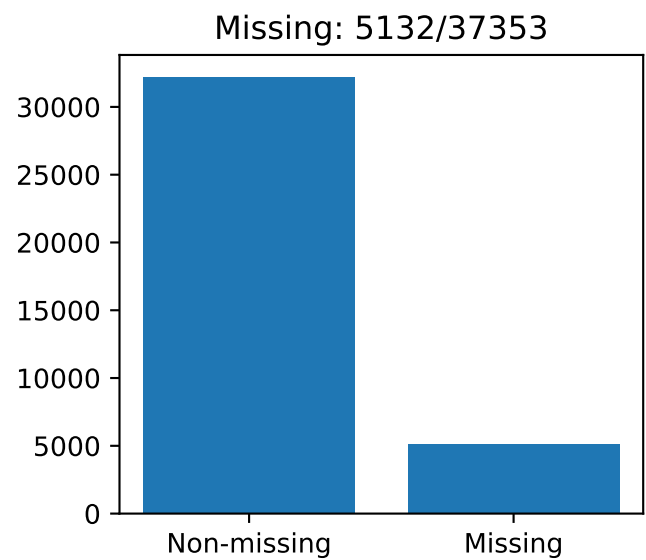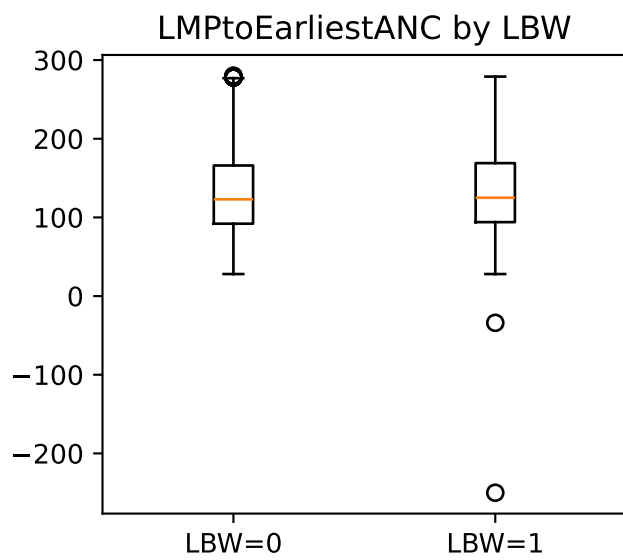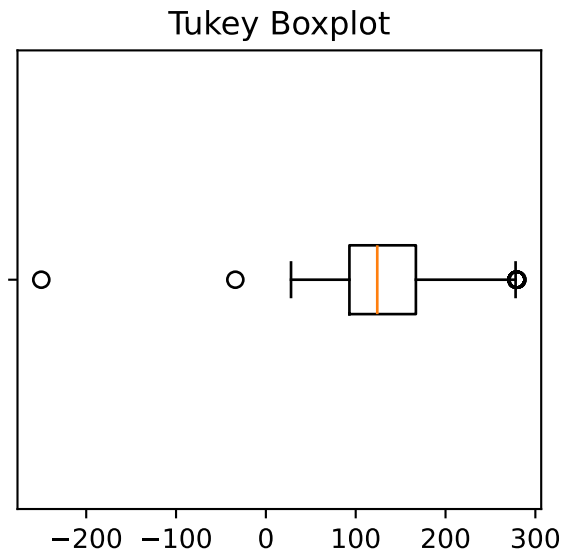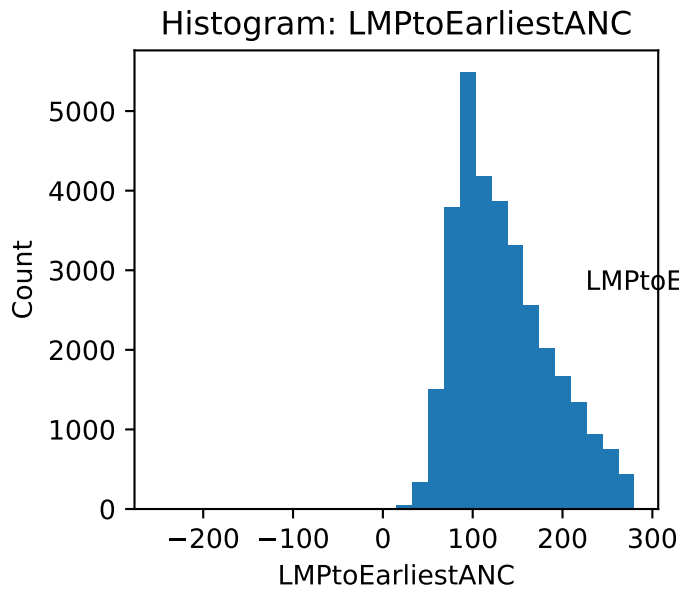
Hypothesis Tests & Metrics for: LMPtoEarliestANC (numeric)
Missing: 13.74% | Unique: 255
Welch t-test: t=-2.75, p=0.00591
Point–Biserial r: r=0.0154, p=0.00586
Cohen's d: 0.0335 | KS: D=0.0211, p=0.00478
AUROC (using LMPtoEarliestANC as score): 0.511

Top-15 categories: ANCBucket
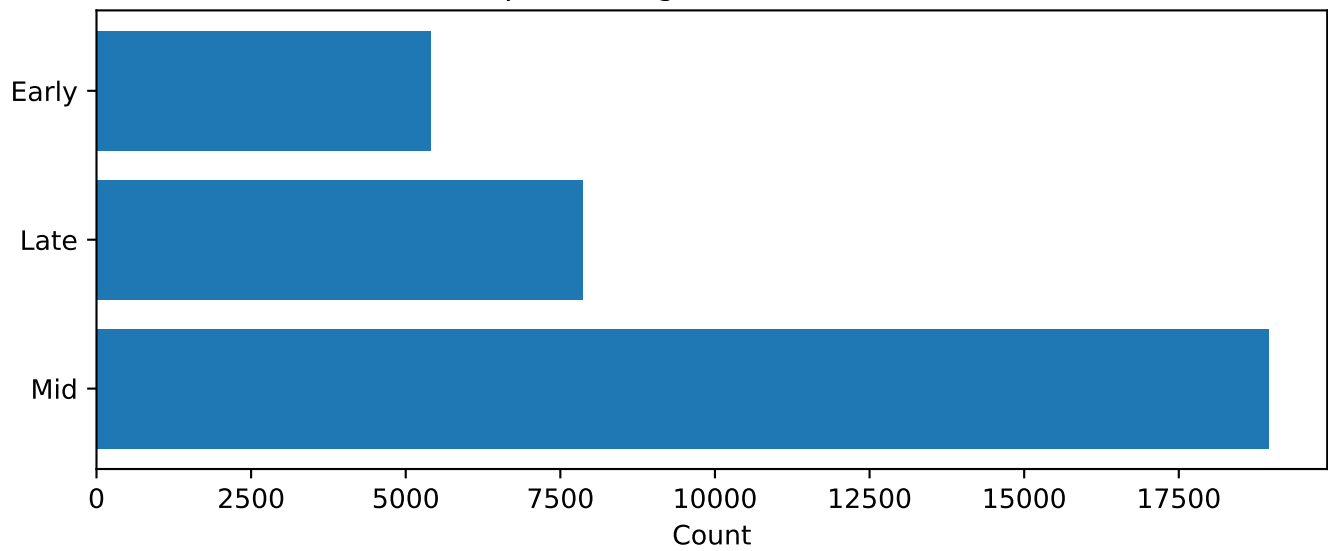
Proportions by LBW
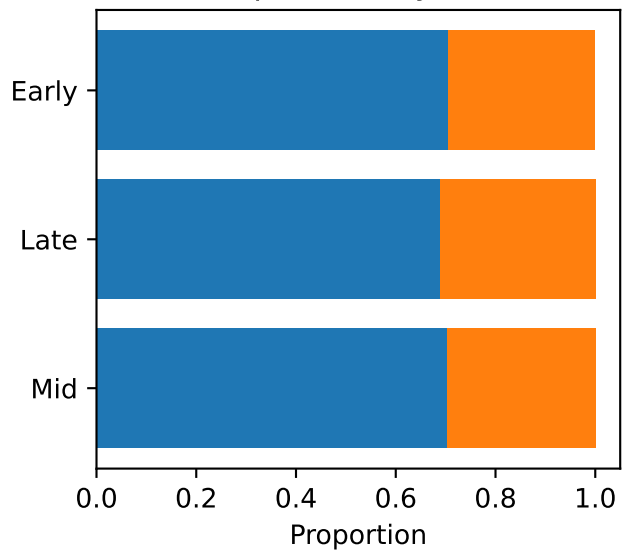
Missing: 5134/37353
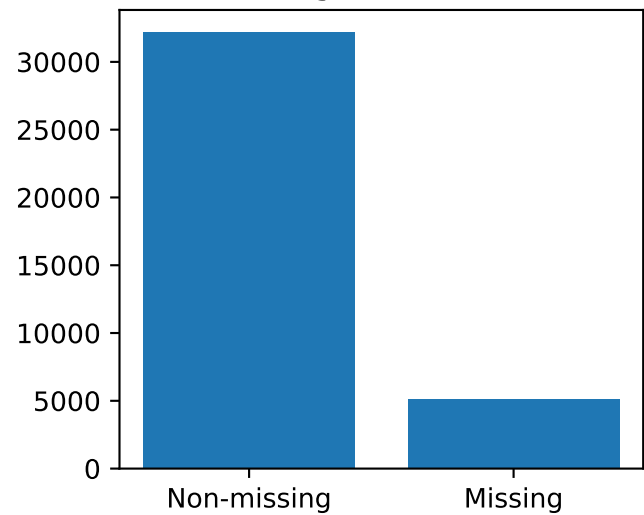
Hypothesis Tests & Metrics for: ANCBucket (categorical)
Missing: 13.74% | Unique: 3
Chi-square: chi2=6.14, dof=2, p=0.0464
Cramér's V (bias-corrected): 0.011

Top-15 categories: Knowledge of respondent whether IFA & Calcium tablets taken together

Proportions by LBW

Missing: 1155/37353
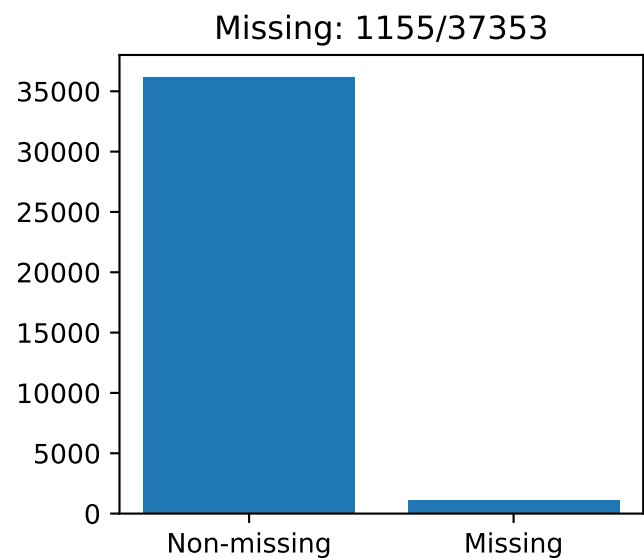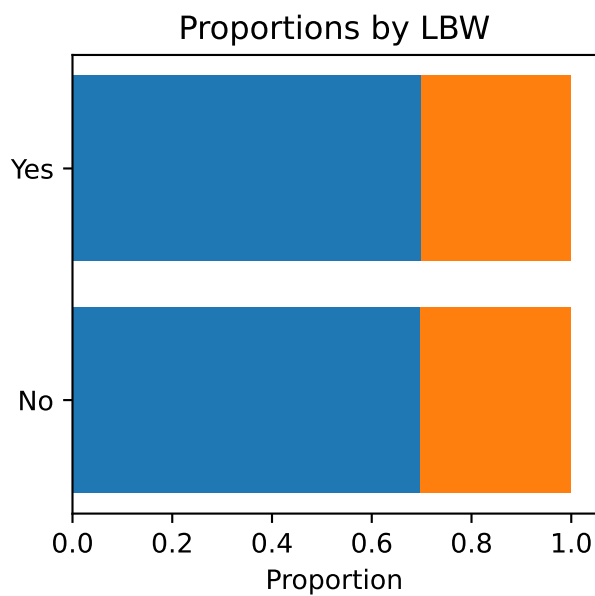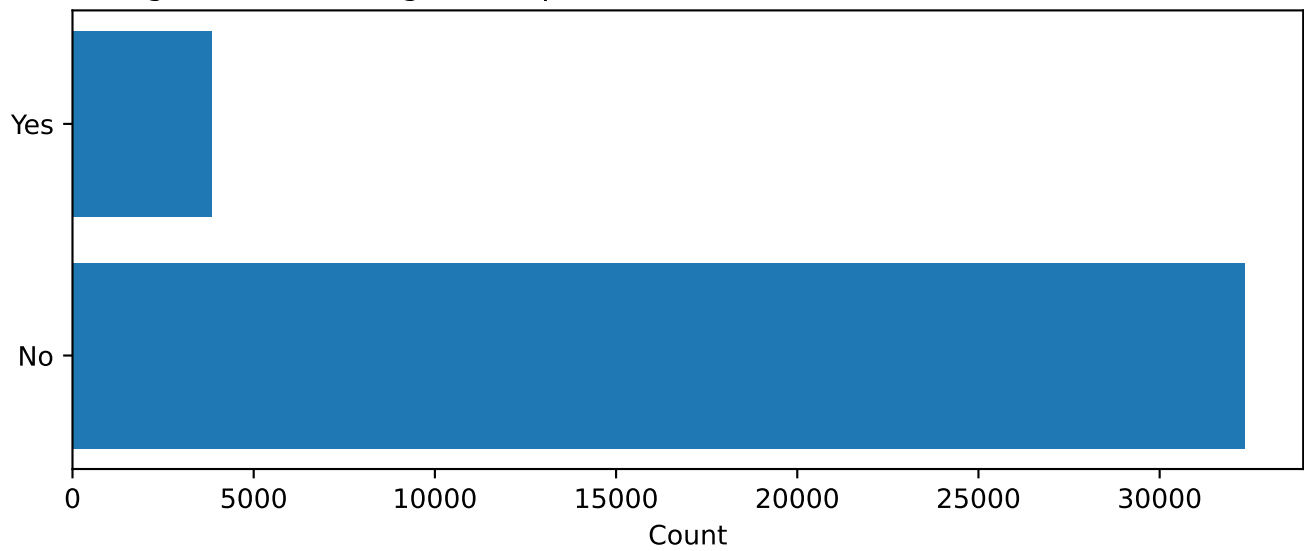
Hypothesis Tests & Metrics for: Knowledge of respondent whether IFA & Calcium tablets taken together (categoric
Missing: 3.09% | Unique: 2
Chi-square: chi2=0.0354, dof=1, p=0.851
Cramér's V (bias-corrected): 0.000

EDA – No. of times a pregnant woman should take her meals

Histogram: No. of times a pregnant woman should take her meals

Tukey Boxplot
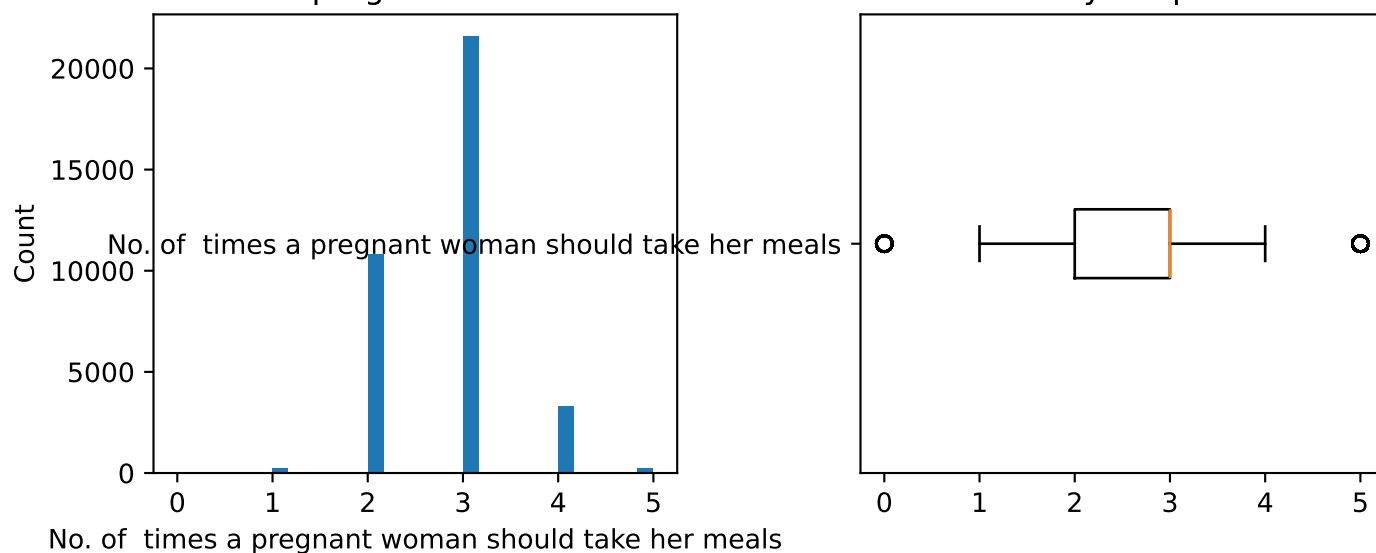
No. of times a pregnant woman should take her meals by LBW   Missing: 1155/37353
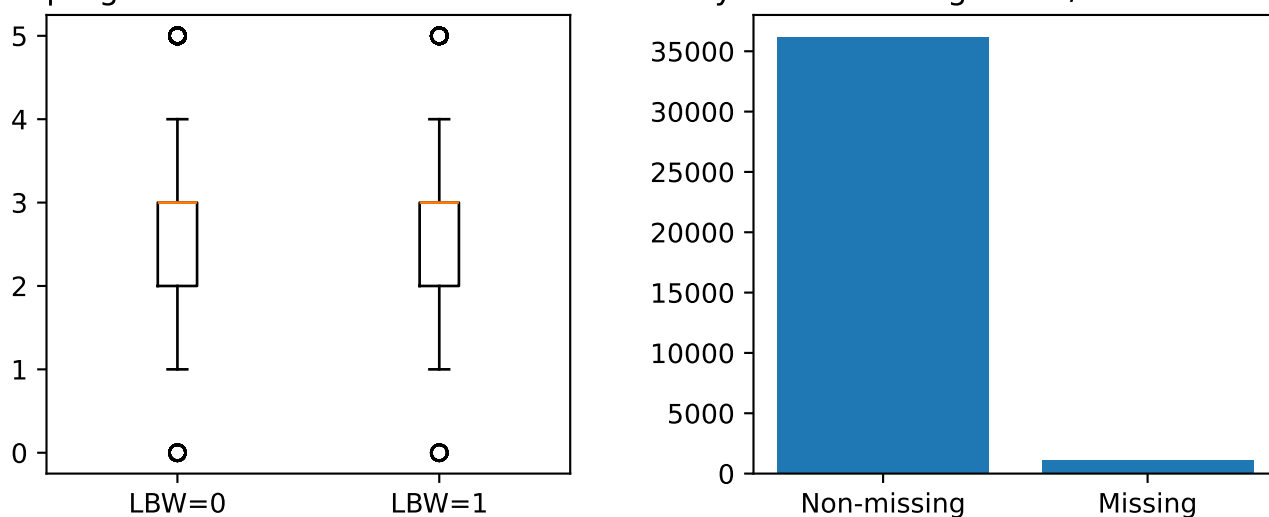
Hypothesis Tests & Metrics for: No. of times a pregnant woman should take her meals (numeric)
Missing: 3.09% | Unique: 6
Welch t-test: t=2.17, p=0.03
Point–Biserial r: r=-0.0115, p=0.0292
Cohen's d: -0.025 | KS: D=0.0118, p=0.234
AUROC (using No. of times a pregnant woman should take her meals as score): 0.494

EDA – Social_Media_Count

Histogram: Social_Media_Count
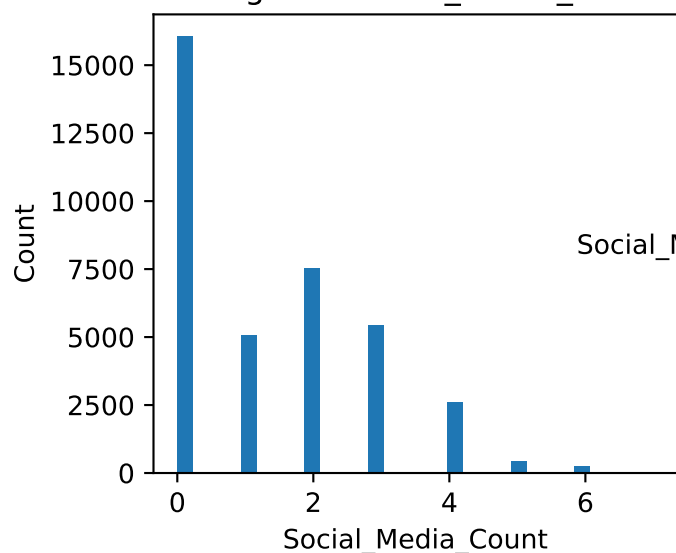
Tukey Boxplot
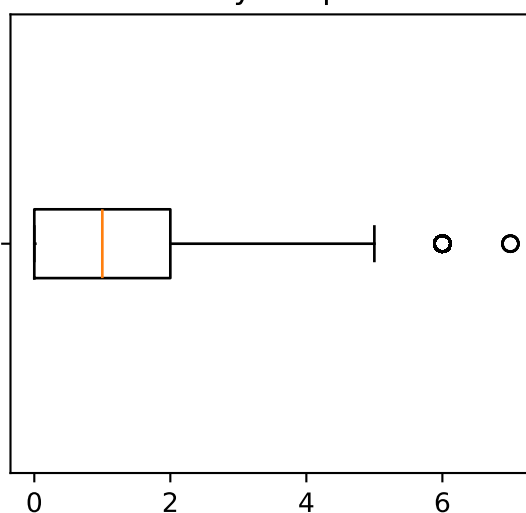
Social_Media_Count by LBW
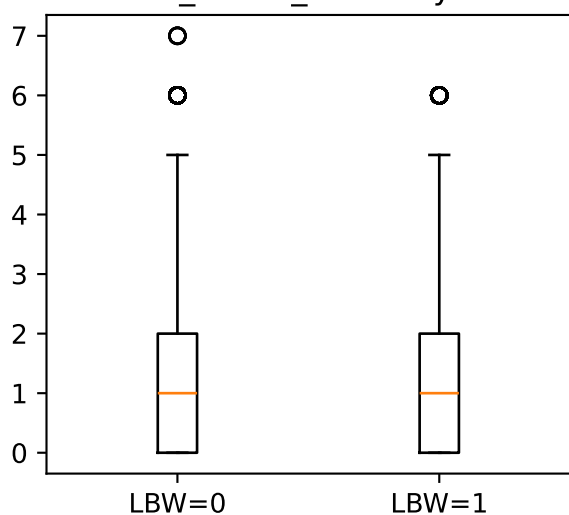
Missing: 0/37353
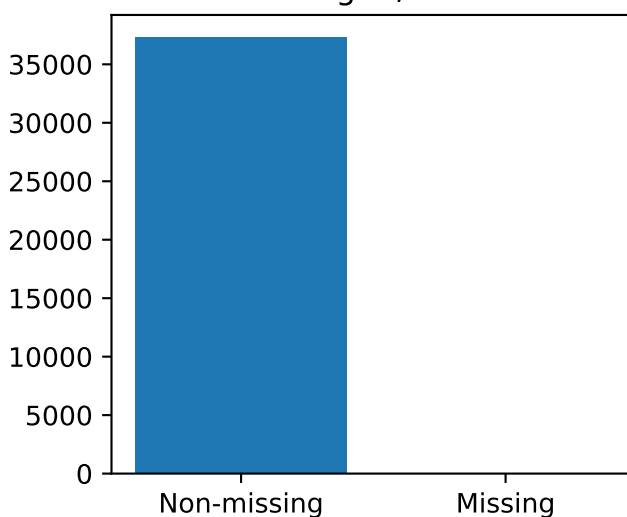
Hypothesis Tests & Metrics for: Social_Media_Count (numeric)
Missing: 0.00% | Unique: 8
Welch t-test: t=7.18, p=7.45e-13
Point–Biserial r: r=-0.0368, p=1.1e-12
Cohen's d: -0.0802 | KS: D=0.0399, p=2.37e-11
AUROC (using Social_Media_Count as score): 0.477

EDA – Social_Media_Category

Top-15 categories: Social_Media_Category
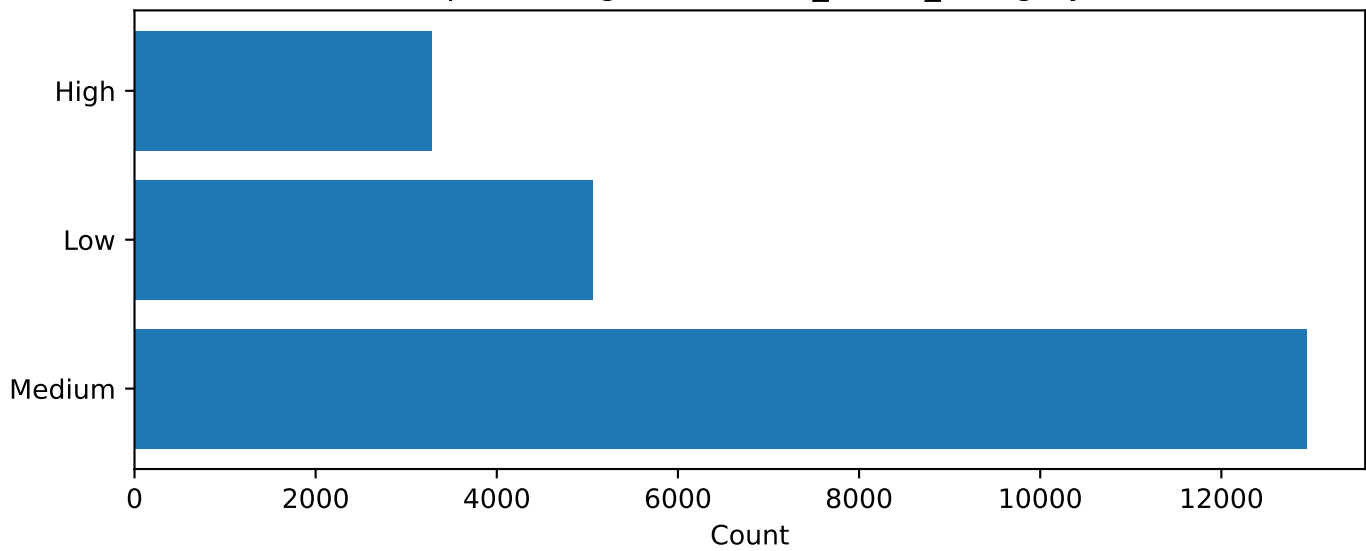
Proportions by LBW
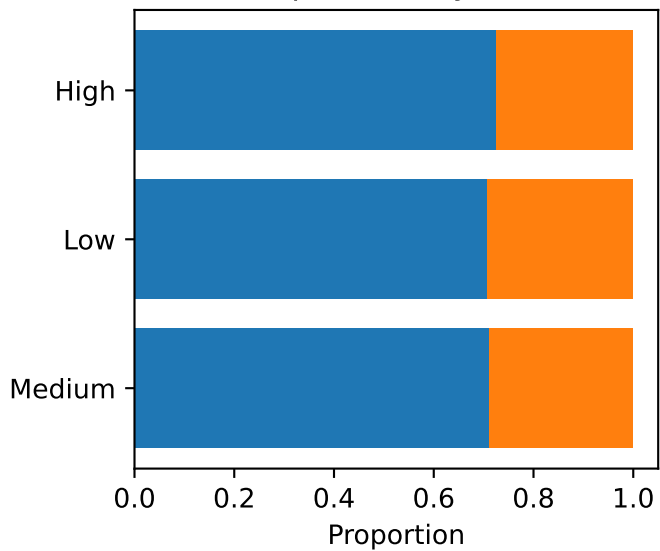
Missing: 16064/37353
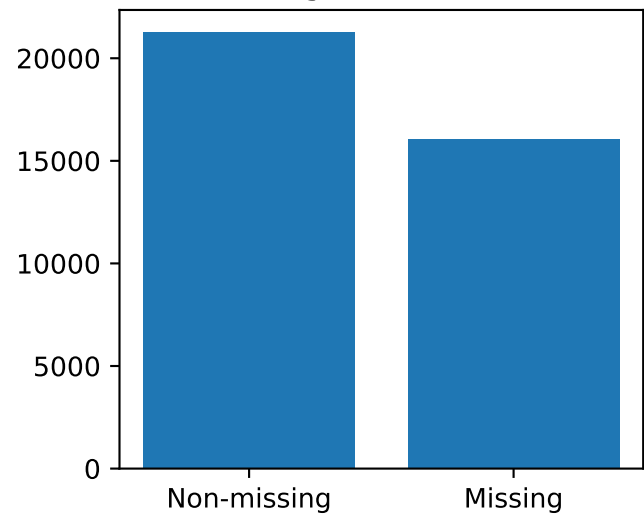
Hypothesis Tests & Metrics for: Social_Media_Category (categorical)
Missing: 43.01% | Unique: 3
Chi-square: chi2=3.55, dof=2, p=0.17
Cramér's V (bias-corrected): 0.009

Summary Table (page 1/3)

variable	type	pct_missing	n_unique	t_stat	t_p	point_biserial_r	pb_p	auroc	cohens_d	ks_stat	ks_p	chi2
Age of the respondent	numeric	0.0	40	13.109	0.0	-0.067	0.0	0.4544	-0.1463	0.0697	0.0	nan
Beneficiary age	numeric	0.0	45	12.9932	0.0	-0.0666	0.0	0.4546	-0.1453	0.0695	0.0	nan
Aganwadi Center Type	categorical	15.7926	2	nan	nan	nan	nan	nan	nan	nan	nan	0.032
MonthConception	categorical	4.6342	12	nan	nan	nan	nan	nan	nan	nan	nan	48.1016
education	categorical	0.9156	6	nan	nan	nan	nan	nan	nan	nan	nan	202.9656
Completed level of education	categorical	0.0	7	nan	nan	nan	nan	nan	nan	nan	nan	152.2352
Number of members residing at respondent's house	numeric	26.4102	26	0.6504	0.5155	-0.0039	0.5148	0.4966	-0.0086	0.0072	0.9211	nan
Main Source of drinking water	categorical	0.0	11	nan	nan	nan	nan	nan	nan	nan	nan	163.0645
toilet_type	categorical	0.9156	6	nan	nan	nan	nan	nan	nan	nan	nan	271.4023
food_group_n	numeric	0.0	8	1.8916	0.0586	-0.0095	0.0665	0.4986	-0.0207	0.0036	0.9999	nan
es/fruits are grown by the household for own consumption	categorical	0.0	2	nan	nan	nan	nan	nan	nan	nan	nan	13.5846
Household asset: Air Conditioner/ cooler	categorical	26.4102	2	nan	nan	nan	nan	nan	nan	nan	nan	50.2895
Household asset: Washing Machine	categorical	26.4102	2	nan	nan	nan	nan	nan	nan	nan	nan	48.0996
Types of fuel used	categorical	0.0027	10	nan	nan	nan	nan	nan	nan	nan	nan	94.4153
Food_Groups_Category	categorical	0.0	4	nan	nan	nan	nan	nan	nan	nan	nan	10.5171
Food_Intake_Quality_Awareness_Score	numeric	3.9354	8	2.7129	0.0067	-0.0144	0.0065	0.4908	-0.0313	0.0136	0.1192	nan
Currently having disease- diabetes	categorical	12.414	3	nan	nan	nan	nan	nan	nan	nan	nan	6.5799
Status of current smoking of bidi	categorical	12.414	3	nan	nan	nan	nan	nan	nan	nan	nan	4.8836
consume_tobacco	categorical	5.6836	3	nan	nan	nan	nan	nan	nan	nan	nan	27.4387
Status of current chewing of tobacco	categorical	26.4102	3	nan	nan	nan	nan	nan	nan	nan	nan	12.7659
Alcoholic beverages consumed yesterday	categorical	3.9354	2	nan	nan	nan	nan	nan	nan	nan	nan	0.9231
consume_alcohol	categorical	5.6836	3	nan	nan	nan	nan	nan	nan	nan	nan	0.9263
Status of drinking of alcohol	categorical	12.414	2	nan	nan	nan	nan	nan	nan	nan	nan	1.4248
(Hb) at the time of registration (Record from Mamta card)	numeric	63.4327	17	1.2661	0.2055	-0.0108	0.2087	0.4803	-0.0234	0.0326	0.0042	nan
Service received during last ANC: Hb test done	categorical	63.5933	3	nan	nan	nan	nan	nan	nan	nan	nan	2.1766
measured_HB	numeric	74.8293	12	6.1618	0.0	-0.0618	0.0	0.4638	-0.1362	0.0512	0.0001	nan
Child order/parity	numeric	0.0107	2	4.9187	0.0	-0.0254	0.0	0.4863	-0.0553	0.0275	0.0	nan
WeightGain_PW2	numeric	76.6819	47	3.3349	0.0009	-0.0345	0.0013	0.4771	-0.0764	0.0295	0.0876	nan

variable	type	pct_missing	n_unique	t_stat	t_p	point_biserial_r	pb_p	auroc	cohens_d	ks_stat	ks_p	chi2	chi2
WeightGain_PW3	numeric	47.6133	53	0.3417	0.7326	-0.0024	0.7394	0.499	-0.0052	0.0099	0.8187	nan	n
WeightGain_PW4	numeric	68.2944	48	1.8242	0.0682	-0.0158	0.0858	0.4914	-0.0353	0.0144	0.6989	nan	n
Household_Assets_Score	numeric	0.0	80	8.1665	0.0	-0.0412	0.0	0.4758	-0.0897	0.0472	0.0	nan	n
Household_Assets_Count	numeric	0.0	26	7.3501	0.0	-0.0373	0.0	0.4768	-0.0812	0.0442	0.0	nan	n
Household_Assets_Category	categorical	0.0	3	nan	nan	nan	nan	nan	nan	nan	nan	62.7569	
No of ANCs completed	numeric	0.0	6	9.945	0.0	-0.0511	0.0	0.4657	-0.1115	0.06	0.0	nan	n
No of PC's total visit to PW	numeric	0.0	5	6.5062	0.0	-0.0334	0.0	0.4791	-0.0727	0.0371	0.0	nan	n
Service received during last ANC: TT Injection given	categorical	63.5933	3	nan	nan	nan	nan	nan	nan	nan	nan	2.6684	0.26
No. of IFA tablets received/procured in last one month	numeric	59.8212	40	-1.6625	0.0964	0.0136	0.096	0.503	0.0295	0.0093	0.943	nan	n
No. of calcium tablets consumed in last one month	numeric	60.552	63	-0.2375	0.8122	0.0019	0.8158	0.501	0.0042	0.0124	0.7226	nan	n
Registered for cash transfer scheme: JSY	categorical	0.0027	2	nan	nan	nan	nan	nan	nan	nan	nan	8.6028	0.00
Registered for cash transfer scheme: RAJHSRI	categorical	0.0027	2	nan	nan	nan	nan	nan	nan	nan	nan	1.0239	0.31
JSY-Number of installment received	numeric	77.643	3	-0.4819	0.6299	0.0053	0.6285	0.5032	0.0117	0.0044	1.0	nan	n
PMMVY-Number of installment received	categorical	66.9558	7	nan	nan	nan	nan	nan	nan	nan	nan	8.1386	0.22
LMPtoRegistration	numeric	4.6342	275	-2.9696	0.003	0.0157	0.003	0.5126	0.0342	0.0274	0.0	nan	n
LMPtoPW1	numeric	73.7799	75	-0.8329	0.4049	0.0088	0.3855	0.5054	0.0193	0.0261	0.125	nan	n
LMPtoPW2	numeric	29.1329	101	-3.0725	0.0021	0.019	0.002	0.5115	0.0415	0.02	0.0231	nan	n
LMPtoPW3	numeric	35.8258	58	-2.2781	0.0227	0.0148	0.0216	0.5097	0.0325	0.0219	0.0159	nan	n
LMPtoPW4	numeric	60.03	29	1.8213	0.0686	-0.0148	0.0703	0.4918	-0.0331	0.0353	0.0011	nan	n
LMPtoINST1	numeric	91.7195	1194	-4.6053	0.0	0.0852	0.0	0.5512	0.1864	0.1099	0.0	nan	n
LMPtoINST2	numeric	92.5388	1033	-4.6362	0.0	0.0905	0.0	0.5531	0.1991	0.1251	0.0	nan	n
LMPtoINST3	numeric	94.8224	871	-1.7264	0.0846	0.041	0.0717	0.5181	0.0912	0.0987	0.0009	nan	n
LMPtoINST4	numeric	99.9946	2	nan	nan	nan	nan	nan	nan	nan	nan	nan	n
LMPtoINST5	numeric	99.7403	74	0.3953	0.6944	-0.0376	0.7143	0.4617	-0.0852	0.2117	0.3201	nan	n
RegistrationBucket	categorical	4.6395	3	nan	nan	nan	nan	nan	nan	nan	nan	6.8899	0.03
LMPtoEarliestANC	numeric	13.7392	255	-2.7529	0.0059	0.0154	0.0059	0.5108	0.0335	0.0211	0.0048	nan	n
ANCBucket	categorical	13.7445	3	nan	nan	nan	nan	nan	nan	nan	nan	6.1391	0.04
espondent whether IFA & Calcium tablets taken together	categorical	3.0921	2	nan	nan	nan	nan	nan	nan	nan	nan	0.0354	0.85

variable	type	pct_missing	n_unique	t_stat	t_p	point_biserial_r	pb_p	auroc	cohens_d	ks_stat	ks_p	chi2	chi2_p	chi2_dof
of times a pregnant woman should take her meals	numeric	3.0921	6	2.1708	0.03	-0.0115	0.0292	0.4936	-0.025	0.0118	0.2339	nan	nan	nan
Social_Media_Count	numeric	0.0	8	7.1752	0.0	-0.0368	0.0	0.477	-0.0802	0.0399	0.0	nan	nan	nan
Social_Media_Category	categorical	43.0059	3	nan	nan	nan	nan	nan	nan	nan	nan	3.5469	0.1697	2.0