**Part II: Visualizations and Descriptive Statistics**

1. A table with summary statistics for all of your variables

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Column1** | **sample size** | **household income** | **% nonhispanic whites** | **c/s delivery rate** | **early and adequate ANC** | **low birthweight** |
| low income communities | 11 | 28397 | 3.29 | 15.76 | 54.91 | 15.2 |
| not-low income communities | 66 | 703387 | 21.74 | 15.8 | 70.55 | 8.15 |

This table summarizes the data sets under two categories; low income communities (communities with median household income below $35000) and not-low income communities (communities with median household income equal to or above $35000). The mean for the various variables were used instead of the median since the data was evenly distributed**.**

1. One data visualization for each of your 5 main variables

According to the graph above, about 70.55% births were recorded among mothers living in not low-income compared to 54.91% mothers from low-income households. The population of non-hispanic whites were more (21.74%) in the not-low income households as opposed to the 3.29% in the low income households.

Contrary to the initial hypothesized above, early and adequate prenatal care did not yield any significant difference in the rate of cesarean session among the two population. It is worth mentioning that the rate of low birth weight was higher (15.2%) among births by mothers from low-income households than the 8.15% from the not low-income households.

**Part III: Basic Data Analysis/Bivariate Results**

A statistical test of the relationship between your dependent variable (outcome variable) and each of your 4 independent variables

A t-test of low-income and Early and adequate prenatal care rate

|  |  |
| --- | --- |
| Two-sample t test with equal variances |  |
|  |  |
| Variable Obs Mean Std. err. Std. dev. | [95% conf. interval] |
|  |  |
| lowincom 11 54.90909 1.385277 4.594443 | 51.8225 57.99568 |
| above\_lo 66 70.54697 1.18233 9.605291 | 68.18569 72.90824 |
|  |  |
| Combined 77 68.31299 1.206172 10.58412 | 65.91069 70.71529 |
|  |  |
| diff -15.63788 2.962958 | -31.275753 |
|  |  |
| diff = mean(lowincom) - mean(above\_lo) | t = -5.2778 |
| H0: diff = 0 Degrees | of freedom = 75 |
|  |  |
| Ha: diff < 0 Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.0000 Pr(T > t) = 0.0000 | Pr(T > t) = 1.0000 |

A t-test of low-income and low birth weight

|  |  |
| --- | --- |
| Two-sample t test with equal variances | |
|  |  |
| Variable Obs Mean Std. err. Std. dev. | [95% conf. interval] |
|  |  |
| lowincom 11 15.46364 .4910606 1.628664 | 14.36949 16.55779 |
| above\_lo 66 9.087879 .3352522 2.723601 | 8.418334 9.757424 |
|  |  |
| Combined 77 9.998701 .3904118 3.42585 | 9.221129 10.77627 |
|  |  |
| diff 6.375758 .8481555 | 4.686145 8.06537 |
|  |  |
| diff = mean(lowincom) – mean(above\_lo) | t = 7.5172 |
| H0: diff = 0 Degrees | of freedom = 75 |
|  |  |
| Ha: diff < 0 Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 1.0000 Pr(T > t) = 0.0000 | Pr(T > t) = 0.0000 |
|  |  |

A t-test of low-income and non-hispanic whites

|  |  |
| --- | --- |
| Two-sample t test with equal variances | |
|  |  |
| Variable Obs Mean Std. err. Std. dev. | [95% conf. interval] |
|  |  |
| lowincom 11 3.285885 .848542 2.814296 | 1.395216 5.176554 |
| above\_lo 66 31.74268 3.204386 26.03255 | 25.34308 38.14228 |
|  |  |
| Combined 77 27.67742 2.974155 26.09811 | 21.75388 33.60097 |
|  |  |
| diff -28.4568 7.899687 | -56.91359 |
|  |  |
| diff = mean(lowincom) - mean(above\_lo) | t = -3.6023 |
| H0: diff = 0 Degrees | of freedom = 75 |
|  |  |
| Ha: diff < 0 Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.0003 Pr(T > t) = 0.0006 | Pr(T > t) = 0.9997 |

A t-test of low-income and rate of caesarean session delivery

|  |  |
| --- | --- |
| Two-sample t test with equal variances | |
|  |  |
| Variable Obs Mean Std. err. Std. dev. | [95% conf. interval] |
|  |  |
| lowincom 11 15.76364 .871391 2.890077 | 13.82206 17.70522 |
| above\_lo 66 15.8 .2763397 2.244994 | 15.24811 16.35189 |
|  |  |
| Combined 77 15.79481 .2650584 2.325878 | 15.2669 16.32271 |
|  |  |
| diff -.0363636 .762489 | -1.55532 1.482592 |
|  |  |
| diff = mean(lowincom) - mean(above\_lo) | t = -0.0477 |
| H0: diff = 0 Degrees | of freedom = 75 |
|  |  |
| Ha: diff < 0 Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.4810 Pr(T > t) = 0.9621 | Pr(T > t) = 0.5190 |

The difference in means observed between the two populations (low and not low-income households) in relation to the mean births where mothers received early and adequate prenatal care is significant with all p-values close to zero (Pr(T < t) = 0.0000, Pr(T > t) = 0.0000, Pr(T > t) = 1.0000).

With relatively low p-values (Pr(T < t) = 1.0000, Pr(T > t) = 0.0000, Pr(T > t) = 0.0000) and a positive t-test (7.5172), there is significant difference in mean birth weight among the two populations.

There is a statistically significant difference in means of Hispanic whites between the two groups, and the "low income" group tends to have a lower mean.

The p-values are relatively high for the t-test of rate of caesarean session delivery in the two populations. The t-value is close to zero, and the confidence interval for the difference in means includes zero, indicating that the means of the two groups are not significantly different.

Therefore, there is no statistically significant difference in means.

**Part IV: Regression Analysis**

Logistic regression: fully adjusted model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Logistic regression |  |  | Number of obs | = | 77 |  |
|  |  |  | LR chi2(4) | = | 44.72 |  |
|  |  |  | Prob > chi2 | = | 0 |  |
| Log likelihood = -9.2178461 | |  | Pseudo R2 | = | 0.7081 |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| newlowincome | Odds ratio | Std. err. | z P>z |  | [95% conf. | interval] |
|  |  |  |  |  |  |  |
| earlyadequateANC | 1.07082 | 0.1474148 | 0.50 0.619 |  | 0.8175897 | 1.402482 |
| lowbirthweightrate | 0.185351 | 0.1100915 | -2.84 0.005 |  | 0.0578647 | 0.5937127 |
| percentageofnonhispanicwhite | 0.9086346 | 0.091001 | -0.96 0.339 |  | 0.7466898 | 1.105703 |
| csdeliveryrate | 2.408774 | 1.056481 | 2.00 0.045 |  | 1.019675 | 5.690239 |
| \_cons | 1096.287 | 12156.17 | 0.63 0.528 |  | 3.99E-07 | 3.01E+12 |
|  |  |  |  |  |  |  |
| Note: \_cons estimates baseline | odds. |  |  |  |  |  |

earlyadequateANC: The odds ratio is 1.07082 with a p-value of 0.619, indicating that birth rate among mothers receiving early and adequate prenatal care is not statistically significant in predicting the odds of being in the "low income" category. This suggests that, holding other variables constant, the odds of being in the "low income" category do not significantly change with earlyadequateANC.

lowbirthweightrate: The odds ratio is 0.185351 with a p-value of 0.005, suggesting that rate of low birthweight rate is statistically significant. A one-unit increase in lowbirthweightrate is associated with a decrease in the odds of being in the "low income" category. This variable seems to have a practical and statistically significant impact on the likelihood of being in the "low income" category.

percentageofnonhispanicwhite: The odds ratio is 0.9086346 with a p-value of 0.339, indicating that percentage of non-hispanic whites is not statistically significant in predicting the odds of being in the "low income" category. This variable does not have a significant impact on the odds of being in the "low income" category.

csdeliveryrate: The odds ratio is 2.408774 with a p-value of 0.045, suggesting that rate of caesarean delivery is statistically significant. A one-unit increase in csdeliveryrate is associated with an increase in the odds of being in the "low income" category. This variable appears to have a practical and statistically significant impact on the likelihood of being in the "low income" category.

1. Conclusion paragraph

It can be concluded that birth rate was higher among not-low income households. This observation could be affected by other factors such as socio-economic status and not necessarily access to early and adequate prenatal care. It is not surprising therefore, that rate of caesarean delivery was not different among the two populations. Rate of low birth weight was significantly higher in the low-income households. The t-test indicated a significant impact of low birth weight and cesarean delivery rate on the likelihood of being in the "low income" category. The logistic regression analysis provided a comprehensive examination of the variables' combined effects. While early and adequate prenatal care and percentage of non-hispanic whites did not show significant associations with the odds of being in the "low income" category, low birth weight and cesarean delivery rate emerged as statistically significant predictors. Relating this to the initial hypothesis, it is clear that rate of low birth weight and caesarean session significantly occur among mothers of the low-income households.